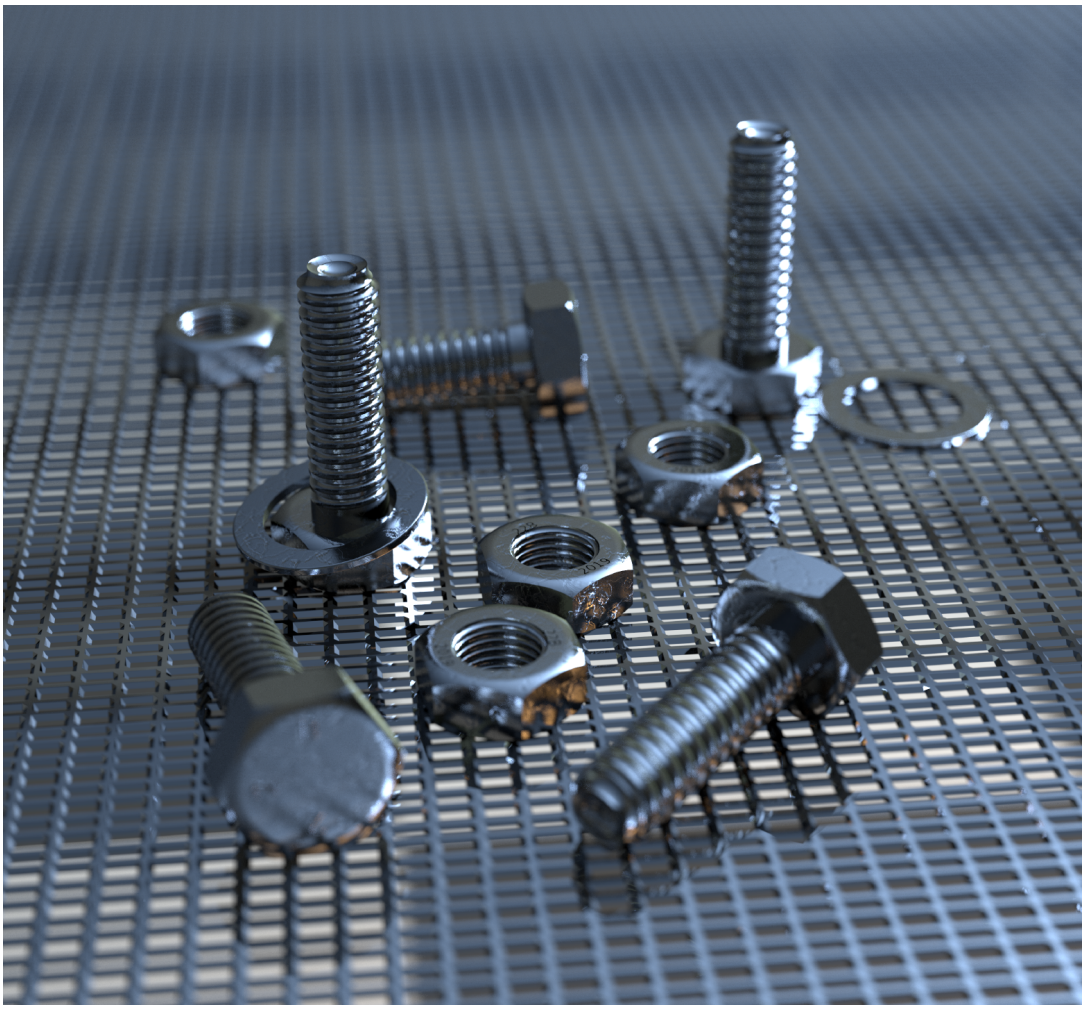
Linear Algebra And Vector Calculus II

Oliver Knill



Title image: Povray code by Jaime Vives

This document is in the public domain.

Oliver Knill, Harvard College, Spring 2019

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 1: Linear Spaces

Lecture

1.1. X is called a **linear space** over the real numbers $\mathbb{R}$ if there is an **addition** $+$ on X, a **zero element** in X and a **scalar multiplication** $x \to \lambda x$ with $1x = x$ in X. Additionally, we want that every x in X can be **negated**; this additive inverse element $-x$ satisfying $x + (-x) = 0$. The zero element 0 is required to satisfy $x + 0 = x$ for all x. With **addition**, we mean an operation which satisfies the **associativity** law $(x + y) + z = x + (y + z)$, the **commutativity** laws $x + y = y + x$, $\lambda x = x\lambda$ and the **distributivity** laws $\lambda(x + y) = \lambda x + \lambda y$, $\lambda\mu(x + y) = \lambda(\mu x + \mu y)$.

1.2. We are familiar with the real numbers $X = \mathbb{R}$. They form a linear space and we have learned to compute with these numbers early on like $7(3 + 5) = 56$. The rules of computation, like the associativity rule are not **results** which are proven but are considered **axioms** meaning that they are assumptions. There are simpler structures requiring less axioms: an example is the set of natural numbers $\mathbb{N} = \{0, 1, 2, 3 \dots\}$, where we have no additive inverse. To have an additive inverse, we need to extend the natural numbers to $\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$.

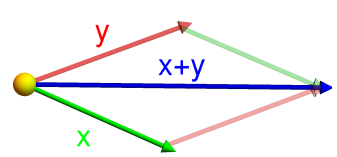
1.3. The set $M(n, m)$ is the space of all $n \times m$ matrices, arrays of numbers in which there are n rows and m columns. It is an example of a **linear space**: it contains a **zero element** in the form of the 0**-matrix**. We can **add** $(A + B)_{ij} = A_{ij} + B_{ij}$, **subtract** $(A - B)_{ij} = A_{ij} - B_{ij}$ and **multiply with scalars** λA. An important class of matrices is the set $\mathbb{R}^n = M(n, 1)$ of **column vectors**. It is the n-dimensional **Euclidean space**. We especially like **the plane** $\mathbb{R}^2$ which we use for writing and $\mathbb{R}^3$, the space we live in.

Theorem: $X = M(n, m)$ is a linear space.

Proof. The associativity, commutativity and distributivity properties are inherited from the reals because each component A_{ij} is a real number. □

1.4. If we want to check whether a subset X of $M(n, m)$ is a linear space, the associativity, commutativity and distributivity properties are inherited from the ambient space. We now only need to check the following three properties:

i: $0 \in X$.
ii: $x + y \in X$ if $x, y \in X$.
iii: $\lambda x \in X$ if $x \in X$.



Theorem: If $X \subset M(n,m)$ satisfies $(i)-(iii)$, then X is linear.

Proof. The addition, scalar multiplication and zero element which satisfy the associative, commutative and distributive properties are given in $M(n,m)$. They hold therefore also in X. We only need to make sure therefore that addition, and scalar multiplication keeps us in X and also that 0 is in X and this is what the three conditions tell. $\square$

1.5. Examples.
a) The set X of non-negative 3×4 matrices for example is not a linear space. Yes, it satisfies $i)$ and $ii)$ but property $iii)$ fails. For $\lambda = -1$, and a non-zero $A \in X$, the matrix λA is not in X.
b) The set X of all 3×4 matrices for which the sum of all matrix entries is zero is a linear space. Proof: i) check. The zero matrix entries add up to zero. ii) check. If $\sum_{i,j} A_{ij} = 0$ and $\sum_{i,j} B_{ij} = 0$, then $\sum_{i,j}(A+B)_{ij} = 0$. Finally, if A satisfying $\sum_{i,j} A_{ij} = 0$ and $\lambda \in \mathbb{R}$ are given then $\sum_{i,j}(\lambda A)_{ij} = \lambda \sum_{i,j} A_{ij} = 0$.

1.6. If there is also a multiplication $A \cdot B$ defined on a linear space X for which associativity, distributivity hold and a 1-element exists, one calls X an **algebra**. We do not require the multiplication to be commutative. Remember that the **matrix multiplication** $(AB)_{ij} = \sum_{k=1}^{p} A_{ik}B_{kj}$ was defined if $A \in M(n,p)$ and $B \in M(p,m)$. The result was a matrix in $M(n,m)$. If $m = n$, then the result is again in X.

Theorem: The space $M(n,n)$ is an algebra.

Proof. The identity matrix 1 satisfies $1 \cdot A = A$. We write 1 but note that 1 is an element in $M(n,n)$ and not an element in $\mathbb{R}$. We also write $\cdot$ for the multiplication if there can be some ambiguity like $2 \cdot 1$ is twice the identity matrix and not 21. The associativity property $A \cdot (B \cdot C) = (A \cdot B) \cdot C$ can be established formally as both sides are $(ABC)_{k,l} = \sum_{i,j} A_{ki}B_{ij}C_{jk}$. The distributivity property $A \cdot (B+C) = A \cdot B + A \cdot C$ also can be checked with $(A \cdot (B+C))_{k,l} = \sum_j A_{kj}(B+C)_{jl} = \sum_j A_{kj}(B_{jl} + C_{jl}) = \sum_j A_{kj}B_{jl} + \sum_j A_{kj}C_{jl} = AB + AC$. $\square$

1.7. We can compute with square matrices as with numbers. Here are three things which are different in the algebra $M(n,n)$.

The algebra $M(n,n)$ is not commutative if $n > 1$

1.8. For $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$, we have $AB = \begin{bmatrix} 2 & 0 \\ 1 & 0 \end{bmatrix}$, $BA = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

1.9.

There are infinitely many non-invertible elements if $n > 1$.

The 2×2 matrix with all entries being 1 is non-zero but not invertible.

1.10.

A^n can grow different than exponentially for $n > 1$.

In dimension 1, we have A^n either growing exponentially, decaying exponentially or staying bounded. For $n = 2$ already, we can have growth which is linear: for $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ we have $A^n = \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}$.

Examples

1.11. Let $A = \begin{bmatrix} 3 & 4 \\ -1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$. Then $A+B = \begin{bmatrix} 5 & 5 \\ -1 & 2 \end{bmatrix}$ and $5A-3B = \begin{bmatrix} 9 & 17 \\ -5 & 2 \end{bmatrix}$ and $A^3 = \begin{bmatrix} -1 & 36 \\ -9 & -19 \end{bmatrix}$ and $7A^{-1} = \begin{bmatrix} 1 & -4 \\ 1 & 3 \end{bmatrix}$.

1.12. Is the upper half plane $H = \{(x, y) \in \mathbb{R}^2 \mid y \geq 0\}$ a linear space? It contains a zero element, is stable under addition and if v is there, also any multiple λv is there. Yes, almost. But it fails with scalar multiplication. $x \to \lambda x$ does not preserve H for negative λ. It is not a linear space.

1.13. Is the set X of 2×2 matrices for which $A_{11} = 0$ a linear space? Yes, it is. We check the three properties. The zero matrix is in the space. The sum of two matrices of this form is a matrix of this form. And if we multiply a matrix in X with a constant, then also λx is in X.

1.14. Is the set X of 2×2 matrices for which all entries are rational numbers a linear space? We can call this space $M_{\mathbb{Q}}(2, 2)$. Check the properties. It is your turn.

1.15. Is the space $\{(x, y, z) \in \mathbb{R}^3 \mid xyz = 0\}$ a linear space?

Illustrations

1.16. Is the set X of all pictures with 800×600 pixels a linear space? We can add two pictures, multiply (make it brighter) and also have the zero picture, where red, green and blue entries are zero. While X is part of a linear space it is not a linear space itself. The pixel color range is an integer $[0, 255]$. Even if we allow real color values, the bounded range prevents X to become a linear space. But X is **part** of a linear space.

Figure 1. A is a flower from the garden of Emily Dickinson in Amherst. B is a portrait of Emily Dickinson herself. We then formed $0.3A + 0.7B$.

Homework

This homework is due on Tuesday, 2/6/2019.

Problem 1.1: Which of the following spaces are linear spaces?
a) the set of symmetric 2×2 matrices. ($A^T = A$).
b) the set of anti-symmetric 2×2 matrices. ($A^T = -A$).
c) the set of 2×2 diagonal matrices with zero trace.
d) the set of 2×2 matrices for which all entries are ≥ 0.
e) the set of 2×2 matrices with determinant 1.
f) the set of 2×2 matrices which are not invertible.
g) the set of 2×2 matrices which are in row reduced echelon form.
h) the set of 2×2 matrices with zero trace.

Problem 1.2: a) Take the matrix $A = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$, then compute A^2, A^3 etc. Can you find a pattern for A^n?
b) Do the same for $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$.

Problem 1.3: a) Find the inverse of the matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & -1 & 2 \\ 1 & 1 & 1 \end{bmatrix}$ using row reduction.
b) Assume $A^{10} = 1$, can you find an expression for the inverse of A which only involves addition and multiplication?
c) Write down the inverse $(AB)^{-1}$ as a product of two inverses. Is it $A^{-1}B^{-1}$?

Problem 1.4: a) Solve the equation $AXB = 3X$ for X in the case $A = \begin{bmatrix} 3 & 4 \\ -1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$.
b) Find a 3×3 matrix A such that A, A^2 are not zero but A^3 is the zero matrix.

Problem 1.5: For any function $f(x)$ with a Taylor expansion $f(x) = \sum_{n=0}^{\infty} a_n x^n$ and a matrix A we can also define $f(A) = \sum_{n=0}^{\infty} a_n A^n$.
a) Assume $f(x) = (1-x)^{-1}$, compute $f(\text{Diag}(1/2, 1/3))$ and the series.
b) Assume $f(x) = e^x$. Compute $f(\begin{bmatrix} 0 & -t \\ t & 0 \end{bmatrix})$ using the Taylor expansion.
c) Find a matrix $A \in M(2,2)$ which satisfies $\sin(A) = 0$ but which is not of the form $A = 0$ or $A = \pi = \pi I_2$.

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 2: Linear transformations

Lecture

2.1. A matrix $A \in M(n,m)$ defines a **linear map** $T(x) = Ax$ from $\mathbb{R}^m$ to $\mathbb{R}^n$. For example, if $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix}$ and $v = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$, then $Av = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. The matrix A defines a linear map from $\mathbb{R}^3$ to $\mathbb{R}^2$.

2.2. To verify that a map T is linear, we have to check three things:

$T(0) = 0$	compatibility with zero
$T(x+y) = T(x) + T(y)$	compatibility with addition
$T(\lambda x) = \lambda T(x)$	compatibility with scaling

2.3. The vector $e_k = [0, \dots, 0, 1, 0, \dots, 0]^T \in \mathbb{R}^n$ is called the k'th **basic basis vector**. We can also see the vector e_k is the k'th column of the identity matrix 1 in $M(n,n)$.

2.4. An important link between geometry and algebra is the following simple fact.

Theorem: The k'th column vector of A is the image of e_k.

Proof. Just compute the image of e_k which is Ae_k. Formally, we can just write $(Ae_k)_i = \sum_j A_{ij}(e_k)_j = A_{ik}$, where we have used that $(e_k)_j = 1$ if $k = j$ and 0 else. $\square$

2.5. This theorem is pivotal as it implies that if a transformation T satisfies the three properties above, then there is a matrix A which has the property that $T(v) = Av$. The proof is constructive in that the matrix A is explicitly given by its column vectors. The linearity then assures that if $v = [a_1, \dots, a_m]^T = a_1 e_1 + \cdots + a_m e_m$, then $Tv = a_1 T(e_1) + \cdots + a_m T(e_m)$ is the matrix multiplication Av. Since the $T(e_k)$ are columns, one calls this the **column picture**.

2.6. To see for example what the effect of $T(x) = Ax$ with $A = \begin{bmatrix} 1 & -2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ is, we look at the column vectors. They have the same length and are perpendicular to each other. The transformation is a rotation dilation about the z-axes as rotation.

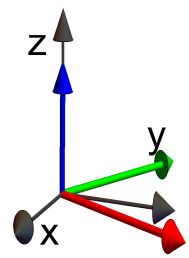


FIGURE 1. A rotation dilation in space. We link the transformation with the matrix by looking at the image of the basis vectors.

2.7. Here are the 4 most important types of linear transformations in the plane $\mathbb{R}^2$. Shear means "horizontal shear".

Rotation	Reflection	Projection	Shear
$\begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix}$	$\begin{bmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{bmatrix}$	$\begin{bmatrix} \cos^2(\alpha) & \cos(\alpha)\sin(\alpha) \\ \sin(\alpha)\cos(\alpha) & \sin^2(\alpha) \end{bmatrix}$	$\begin{bmatrix} 1 & c \\ 0 & 1 \end{bmatrix}$

2.8. When combined with a dilation, the structure of the matrices becomes simpler: allowing dilations is simpler.

Rotation Dilation	Reflection Dilation	Projection Dilation	Shear Dilation
$\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$	$\begin{bmatrix} a & b \\ b & -a \end{bmatrix}$	$\begin{bmatrix} a^2 & ab \\ ba & b^2 \end{bmatrix}$	$\begin{bmatrix} a & c \\ 0 & a \end{bmatrix}$

FIGURE 2. What kinds of transformations are these? In each of the two cases, we see the original square on the left and the transformed square to the right.

These four examples allow for building more complicated linear transformations.

2.9. A linear transformation $T : X \to X$ is called **invertible** if there exists another transformation $S : X \to X$ such that $TS(x) = x$ for all x.

Theorem: If T is linear and invertible, then T^{-1} is linear and invertible.

Proof. To invert $T(x) = Ax$, we have to be able to solve $Ax = b$ uniquely for every b. Let B be the matrix in which the k'th column b_k is the solution of $Ax = e_k$. Then $S(e_k) = b_k$ so that $ABx = x$. $\square$

Examples

2.10. Is the map $T(x) = x + 1$ a linear transformation from $\mathbb{R}$ to $\mathbb{R}$? By definition, we would have to find a matrix A such that $T(x) = Ax$. But a 1×1 matrix A is just a real number. We would have to find a real number a such that $x + 1 = ax$ for all x. For $x = 0$ this gives $1 = 0$ which is not possible.

Illustrations

2.11. The tesseract is a 4 dimensional cube.

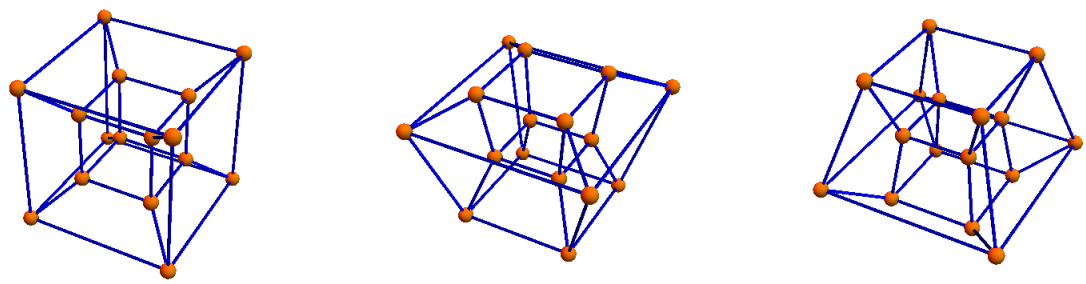


Figure 3. We see here the effect of a rotation in four dimensional space. The tesseract is first rotation by $\pi/3$, then again by $\pi/3$. The picture shows a projection into three dimensions.

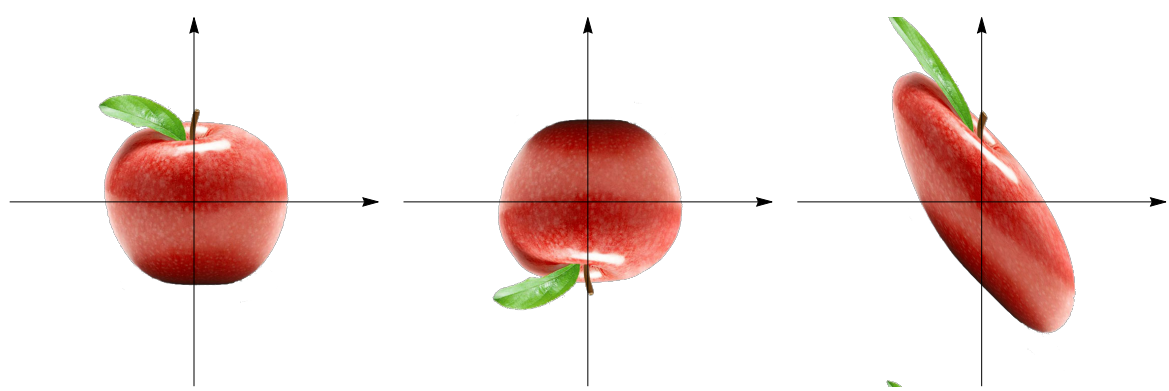


Figure 4. Can you find which of the matrices A-E implements the transformation from the left apple to the middle one. Which one implements the transformation from the left apple to the right one?

2.12. A=$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ B=$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ C=$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ D=$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ E=$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

Homework

This homework is due on Tuesday.

Problem 2.1: Which of the following transformations are linear?
a) $T : \mathbb{R}^4 \to \mathbb{R}^2$ given by $T([x, y, z, w]^T) = [y, z]^T$.
b) $T : \mathbb{R}^4 \to \mathbb{R}^2$ given by $T([x, y, z, w]^T) = [y + 1, w + x]^T$.
c) $T : M(3, 3) \to \mathbb{R}$ the map $T(A) = A_{22}$
d) $T : M(3, 4) \to M(4, 3)$ given by $T(A) = A^T$.
e) $T : M(2, 2) \to \mathbb{R}$ given by $T(A) = \det(A)$.
f) $T : M(2, 2) \to M(2, 2)$ given by $T(A) = A^2$.
g) $T : \mathbb{R} \to M(2, 2)$ by $T(\alpha) = R(\alpha)$, the rotation matrix by α.

Problem 2.2: a) What is the composition of two reflections at lines.
b) What is the composition of two rotation dilations.
c) What is the composition of two horizontal shears.
d) Verify that for a projection P, the identity $P^2 = P$ holds.
e) Verify that for a reflection T, the inverse is equal to T.
f) Find a composition of two projections A, B such that $AB = 0$.
g) What is the composition of a reflection and a rotation?

Problem 2.3: a) Find the matrix in $M(3, 3)$ which reflects about the x-axes.
b) Find the matrix in $M(3, 3)$ which maps $i = [1, 0, 0]$ to $j = [0, 1, 0]$ and $j = [0, 1, 0]$ to $k = [0, 0, 1]$ and maps $k = [0, 0, 1]$ to $i = [1, 0, 0]$.
c) Let A be the matrix in $M(2, 2)$ which reflects about the x-axes and B be the matrix which reflects about the y axes. Check that $AB = BA$.
d) Is it true in general that for two reflections A, B, the identity $AB = BA$ holds?
e) Find a 2×2 matrix which is neither a reflection, nor a rotation, nor a shear nor a projection.

Problem 2.4: a) Write down the matrix A of a rotation about the z-axes by an angle θ.
b) Write down the matrix B of a rotation about the y axes by an angle ϕ.
c) Compute AB. What is the vector $AB([1, 0, 0]^T)$? Do you recognize the coordinates of this vector?

Problem 2.5: Find the 4×4 matrix A which implements the multiplication of a quaternion z by $a + ib + jc + kd$.

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 3: Axioms

Seminar

3.1. An **axiom system** is a collection of statements which define a **mathematical structure** like a **linear space**. The statements of an axiom system are not proven; they are assertions which are **assumed** to be true. They need to be "interesting" in the sense that there should be realizations which satisfy these axioms. You should think of an axiom system as **the rules** of the game. There are games which are interesting like chess, there are games which are not interesting: a chess game in which the queen can jump anywhere on the board would not be interesting, as the player who starts wins the game. An axiom system stating that $x = 0$ for all x would not be interesting. Also similar to games, new interesting mathematical structures are constantly invented and played with. Playing the rules of an axiom system and finding new theorems in it is the **mathematician's game**.

3.2. In the first lecture we have seen axioms which define a **linear space**. Some linear spaces also feature a multiplicative structure and an additional set of axioms which define an **algebra**. These axioms for linear spaces are reasonable because $M(n, m)$ realizes it. The algebra structure is reasonable because $M(n, n)$ is a model for an algebra.

3.3. Here is a first example of an axiom system which is much simpler than the axiom system for a linear space. It defines the structure of a **monoid** which is an important structure in mathematics and computer science.

> **Definition:** $(X, +, 0)$ is a **monoid**, if $+$ defines an addition on X which is associative $(x + y) + z = x + (y + z)$ and which is compatible with **zero** 0 in the sense that $x + 0 = x$ for all x in X. The addition has to be defined between any two elements in X and give again an element in X.

3.4. Examples:
a) The real numbers form a monoid with 0 as the zero element.
b) The nonzero real numbers with multiplication form a monoid. The "zero" is 1.
c) If X is a linear space, then $(X, +, 0)$ is a monoid.
d) The set of even integers defines a monoid.
e) The set of functions from $\mathbb{R}$ to $\mathbb{R}$ form a monoid with composition as addition $f \circ g$.
f) Let us define an addition of graphs $G = (V, E)$ and $H = (W, F)$, where V, W are

the vertex sets and E, F are the edge sets. The sum of the two graphs is $G + H = (V \cup W, E \cup F \cup U(V, W))$, where $U(V, W)$ is the set of all connections from V to W.
g) The set of complex numbers $\mathbb{C}$ form a monoid under addition.
h) The set of functions from $\mathbb{R}^2$ to $\mathbb{R}$ form a monoid under addition.

3.5.

Problem A: a) Verify that the natural numbers $\mathbb{N} = \{0, 1, \dots\}$ form a monoid with addition.
b) Do the negative numbers $\{-1, -2, -3, \dots\}$ form a monoid?
c) What about $X = \{0, -1, -2, -3, \dots\}$.

Problem B: a) Verify that the natural numbers $\mathbb{N} = \{0, 1, \dots\}$ form a monoid with multiplication. What is the "zero element"?
b) Verify that $M(2, 2)$ forms a monoid under multiplication. What is the zero element?

Problem C: a) Why does the set of prime numbers $(P, +)$ not form a monoid?
b) Do the numbers $\{2^n, \{n = 0, 1, 2, 3, \dots\}$ with the usual multiplication as "monoid addition" form a monoid?

3.6. Monoids play an important role also in computer science. The reason is that many **languages** are monoids. Given a finite set A called **alphabet** with letters, then a finite sequence of such letters is called a **word**. The addition of two words x, y is obtained by just **concatenating** them $x + y = xy$. For example, $x =$ "*milk*" and $y =$ "*shake*" combine to $xy = milkshake$ and $yx = shakemilk$. The alphabet could also contains punctuation signs and space, so that every text is an element in a monoid. While we usually count as a word a list of letters, in this mathematical framework, any finite sequence of letters is a word.

Problem D: a) Is the monoid of a language commutative?
b) Illustrate with an example, why the monoid of a language is associative.

3.7. To illustrate a bit more on the last example, we can make a language more interesting by giving additional rules. These additional rules are usually called a **grammar**. For example, one can look at the monoid X of words over the alphabet $A = \{a, b\}$ with the rule that b always has a letter a right before and right after. Example of words in X are $abaaaabaabaaaaaaba$.

3.8. An axiom system is called **consistent** if there is a model for it, and if one can not prove something wrong like $1 = 0$ from it. It should also not be silly like the axiom system **Null**: there is only axiom in it: **all elements are** 0**.** This axiom system obviously has a realization as the space $X = \{0\}$. It is consistent, but it is of no interest at all.

3.9. Let us look at the space $M(n,n)$ and let us define the addition $A \oplus B = AB$.

Problem E: Verify that $(M(3,3), \oplus)$ defines a monoid. What is the zero element?

3.10. An important example of a monoid is a **group**. It is a monoid, in which every element x has an inverse y. The inverse of x is an element y satisfying $x + y = 0$.

Problem F: Verify that if X is a linear space, then $(X, +, 0)$ is a group.

Problem G: Verify that the $SL(2,R)$ the subset of $M(2,2)$ for which the determinant is 1 is a group with the multiplication as addition. (Remember a formula for the inverse of a matrix.)

FIGURE 1. Euclid of Alexandria (325 BC-265 BC) built the first axiom system for geometry and Giuseppe Peano (1858-1932) formulated the Peano axioms of arithmetic.

3.11.

Homework

Exercises A)-D) are done in the seminar. This homework is due on Tuesday:

Problem 3.1 Which of the following sets form monoids? If it is, identify the zero.
a) The set of rotation dilation matrices with multiplication as addition.
b) The set of rotation dilation matrices with matrix addition as addition.
c) The set of horizontal shear matrices with multiplication as addition.
d) The set of horizontal shear matrices with matrix addition as addition.
e) The set of all sets with addition $A+B = A\Delta B$ (symmetric difference).
f) The set of all sets with addition $A + B = A \cup B$.

Problem 3.2 Which of the following sets form groups?
a) The set of all diagonal matrices with matrix addition as addition.
b) The set of all diagonal matrices with matrix multiplication as addition.
c) The set of all 2×2 matrices satisfying $A^2 = 0$.
d) The set of all 2×2 matrices for which the trace is 0 and with matrix addition as addition.
e) The set of all invertible functions from $\mathbb{R} \to \mathbb{R}$ with composition.
f) The set of all sets with addition $A + B = A\Delta B = (A \setminus B) \cup (B \setminus A)$.

Problem 3.3 Most of the axiom systems which are used in mathematics have many rules. Here is an structure, which needs only one axiom to be defined: X is a set of non-empty sets which is closed under the operation of taking finite-nonempty subsets. This structure is called a **simplicial complex**. The sets are the simplices. Let us take the example where X is the set of all non-empty subsets of $\{1, 2, 3\}$. Visualize this structure by drawing the graph in which the sets are the nodes and where two are connected, if one is contained in the other.

Problem 3.4 Look up the nine **Peano axioms** and write them down.

Problem 3.5 Look up the notions of **Magma** and **Semigroup** and compare their axiom systems with the axiom system of a monoid and group. Maybe order the structures according to their generality.

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 4: Basis and dimension

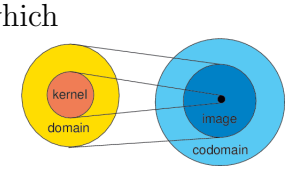
Lecture

4.1. Let X be a linear space. A collection $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$ of vectors in X **spans** X if every x in X can be written as a **linear combination** $x = a_1 v_1 + \cdots + a_n v_n$. The set $\mathcal{B}$ is called **linearly independent** if $a_1 v_1 + \cdots + a_n v_n = 0$ implies that all a_i are zero. The set $\mathcal{B}$ is a **basis** if it both spans X and if it is linearly independent.

4.2. For $X = \mathbb{R}^2$, three vectors $v_1 = [1,1]^T$, $v_2 = [2,3]^T$ and $v_3 = [0,1]^T$ span the plane but they are not linearly independent because $v_1 - v_2 + v_3 = 0$. Indeed, we only need two vectors to span the entire plane. Already $\mathcal{B} = \{v_1, v_2\}$ spans the plane. They are also linearly independent because $a_1 v_1 + a_2 v_2 = 0$ would mean the two vectors are parallel. So, a collection of vectors in $\mathbb{R}^2$ spans if and only if it is not contained in a line. Similarly, a collection of vectors in $X = \mathbb{R}^3$ spans X if and only if the vectors are not contained in a plane.

4.3. A natural basis of $\mathbb{R}^2$ is given by the vectors $[1,0]^T$ and $[0,1]^T$. The collection $\mathbb{B} = \{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\}$ is called the **standard basis** of the plane. The standard basis in $\mathbb{R}^3$ is $\mathbb{B} = \{i = e_1, j = e_2, k = e_3\}$. The standard basis in the quaternion space is $\mathbb{H} = \mathbb{R}^4$ is $e_1 = 1, e_2 = i, e_3 = j, e_4 = k$.

4.4. The **kernel** of a $n \times m$ matrix A is the set $\ker(A) = \{x \in \mathbb{R}^m \mid Ax = 0\}$. The **image** of A is the set $\mathrm{im}(A) = \{Ax \mid x \in \mathbb{R}^m\} \subset \mathbb{R}^n$. These are important constructs because every linear space can be written either as a kernel or as an image of some matrix. Remember that we wrote a two-dimensional plane containing the origin as $ax + by + cz = 0$. This is the kernel of a matrix $A = [a, b, c]$. We also learned how to write the plane as a linear combination of two vectors v, w. The matrix A which contains these two vectors as column vectors has this plane as the image.



Lemma: Both the kernel and image of a matrix are linear spaces.

Proof. a) 0 is the kernel because $A0 = 0$. If x and y are in the kernel, then $Ax = 0$ and $Ay = 0$ implying $A(x + y) = 0$. So, $x + y$ is in the kernel. Also, if $Ax = 0$, then $A(\lambda x) = \lambda(Ax) = 0$ so that λx is in the kernel.
b) Zero is in the image because $A0 = 0$. If $x = Au$ and $y = Av$, then $x + y = Au + Av = A(u + v)$ so that $x + y$ is in the image. If $x = Au$, then $\lambda x = A(\lambda u)$ so that also λx is in the image of A. $\square$

4.5. How do we check whether a set of vectors is a basis? The key is the **transformation matrix** S which contains the elements of $\mathcal{B}$ as columns. Given a matrix S, the **image** of S is the set of all vectors Sx, where x is in X. The **kernel** of S is the solution set of $Ax = 0$.

Lemma: $\mathcal{B}$ is linearly independent if and only if S has a trivial kernel.

Lemma: $\mathcal{B}$ spans X if and only if the image of S contains X.

4.6. We can find the kernel and image of a matrix by row reduction: the kernel is parametrized by free variables. The image is spanned by the columns of A. Here is something one at first does not appreciate. In principle, it would be possible that there is one basis of X with m vectors and a basis with n vectors. In principle, it would be possible that we can implement X as a space of vectors with m components and implement X differently as a space of vectors with n components. But this is not the case:

Theorem: Every basis of X has the same number of elements.

Proof. (i) We first show that if $\mathcal{A} = \{v_1, \dots, v_q\}$ span and $\mathcal{B} = \{w_1, \dots, w_p\}$ are linearly independent, then $q \geq p$.
Assume $q < p$. Because the v_j span, each vector w_i can be written as $w_i = \sum_{j=1}^{q} a_{ij} v_j$.
Now row reduce the augmented $(p \times (q+n))$-matrix $\left| \begin{array}{ccc|c} a_{11} & \dots & a_{1q} & w_1^T \\ \dots & \dots & \dots & \dots \\ a_{p1} & \dots & a_{pq} & w_p^T \end{array} \right|$, where v_i^T is the vector v_i written as a row vector. Each row of A of this matrix contains some nonzero entry. We end up with a matrix which contains a last row $\left| \begin{array}{ccc|c} 0 & \dots & 0 & b_1 w_1^T + \dots + b_q w_q^T \end{array} \right|$ showing that $b_1 w_1^T + \dots + b_q w_q^T = 0$. Not all b_j are zero because we had to eliminate some nonzero entries in the last row of A. This nontrivial relation of w_i contradicts linear independence. The assumption $q < p$ can not be true.
(ii) Because $\mathcal{A}$ spans X and $\mathcal{B}$ is linearly independent, we know that $q \leq p$. Because $\mathcal{B}$ spans X and $\mathcal{A}$ is linearly independent also $p \leq q$ holds. Together, $p \leq q$ and $q \leq p$ implies $p = q$. $\square$

4.7. It follows that every linear space X which is spanned by finitely many vectors has an integer attached to it. The **dimension** of a linear space is defined as the number of basis elements for a basis.

4.8. The dimension of the image of a matrix A is called the **rank** of A. The dimension of the kernel of a matrix A is called the **nullity** of A. The nullity of A is the number of columns without leading 1 in $\text{rref}(A)$; the rank is the number of columns with leading 1. The following theorem is called the **rank-nullity theorem** or **fundamental theorem of linear algebra**.

Theorem: $\text{rank}(A) + \text{nullity}(A) = n$

Proof. There are n columns. Either a column has no leading 1 after row reduction, leading to a free variable and so to a basis vector in the kernel, or then the column will lead to a leading one, meaning that the original column is in the image of A. $\square$

4.9. Can you write the basis for the kernel of the partitioned 4×4 matrix $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ in terms of a basis $\{v_1, v_2\}$ of the kernel of A and the basis $\{w_1, w_2\}$ of the kernel of the 2×2 matrix B. Answer: $\mathcal{B} = \{ \begin{bmatrix} v_1 \\ 0 \end{bmatrix}, \begin{bmatrix} v_2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ w_1 \end{bmatrix}, \begin{bmatrix} 0 \\ w_2 \end{bmatrix} \}$.

Examples

4.10. The kernel and image of the projection matrix A which projects onto the plane $\Sigma : x + y + z = 0$ can be seen geometrically: the kernel of the transformation consists of all vectors orthogonal to Σ it has a basis $\mathbb{B} = \{[1, 1, 1]^T\}$, the image is the plane Σ. It has a basis $\mathbb{B} = \{[1, -1, 0]^T, [1, 0, -1]^T\}$.

4.11. A basis of the kernel of $A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ is $\{e_1, e_2\}$. A basis for the image is $\{e_1\}$. This is an example where the image is part of the kernel.

4.12. Problem: Find a basis for the image and a basis for the kernel of $A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 6 & 8 & 10 \\ 3 & 6 & 9 & 12 & 16 \end{bmatrix}$.

Solution: We have $\text{rref}(A) = \begin{bmatrix} 1 & 2 & 3 & 4 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$. There are two leading 1's. The rank of the matrix is 2, the nullity is 3. The basis of the image is given by the first and last column of the original matrix A:

$$\mathcal{B} = \{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 5 \\ 10 \\ 16 \end{bmatrix} \} .$$

To get a basis for the kernel, introduce free variables for the columns which have no leading 1 and write down the equations $x_1 + 2u - 3v - 4w = 0, x_2 = u, x_3 = v, x_4 = w, x_5 = 0$ then write down the solution in vector form

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = u \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + v \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + w \begin{bmatrix} -4 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} .$$

A basis of the kernel is

$$\mathcal{B} = \{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -4 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \} .$$

Homework

This homework is due on Tuesday, 2/13/2019.

Problem 4.1: Check whether the given set of vectors is linearly independent. Use the S-matrix to find out.
a) $\{\begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}\}$. b) $\{\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 1 \\ -1 \end{bmatrix}\}$.
c) $\{\begin{bmatrix} 3 \\ 16 \end{bmatrix}, \begin{bmatrix} 4 \\ 18 \end{bmatrix}, \begin{bmatrix} 5 \\ 19 \end{bmatrix}\}$. d) $\{[\begin{bmatrix} 1 \end{bmatrix}, \begin{bmatrix} 2 \end{bmatrix}\}$.

Problem 4.2: Find a basis for the image as well as as a basis for the kernel of the following matrices
a) $\begin{bmatrix} 7 & 0 & 7 \\ 2 & 3 & 8 \\ 9 & 0 & 9 \\ 5 & 6 & 17 \end{bmatrix}$, b) $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$. c) $\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$. d) $\begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix}$.

Problem 4.3: a) Find a basis for the kernel and a basis for the image of the projection dilation matrix

$$A = \begin{bmatrix} a^2 & ab \\ ab & b^2 \end{bmatrix} .$$

b) Under which conditions do reflection dilations or rotation dilations or shear dilations all have a trivial kernel?

Problem 4.4: Look up unit 5 of Math 22a. We have seen there the theorem that the image of A^T is perpendicular to the kernel of A. Understand and write down the proof of this important result in your own words.

Problem 4.5: Write down statements which are equivalent to the fact that A is an invertible matrix:
a) A statement involving the rank of A.
b) A statement involving the nullity of A.
c) A statement involving the linear independence of the columns of A.
d) A statement involving the term basis and column vectors.

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 5: Change of Coordinates

Lecture

5.1. Given a basis $\mathcal{B}$ in a linear space X, we can write an element v in X in a unique way as a sum of basis elements. For example, if $v = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ is a vector in $X = \mathbb{R}^2$ and $\mathcal{B} = \{v_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, v_2 = \begin{bmatrix} 1 \\ 6 \end{bmatrix}\}$, then $v = 2v_1 + v_2$. We say that $\begin{bmatrix} 2 \\ 1 \end{bmatrix}_{\mathcal{B}}$ are the $\mathcal{B}$ **coordinates** of v. The **standard coordinates** are $v = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ are assumed if no other basis is specified. This means $v = 3e_1 + 4e_2$.

5.2. If $\mathcal{B} = \{v_1, v_2, \cdots, v_n\}$ is a basis of $\mathbb{R}^n$, then the matrix S which contains the vectors v_k as column vectors is called the **coordinate change matrix**.

Theorem: If S is the matrix of $\mathcal{B}$, then $S^{-1}v$ are the $\mathcal{B}$ coordinates of v.

5.3. In the above example, $S = \begin{bmatrix} 1 & 1 \\ -1 & 6 \end{bmatrix}$ has the inverse $S^{-1} = \begin{bmatrix} 6 & -1 \\ 1 & 1 \end{bmatrix}/7$. We compute $S^{-1}[3,4]^T = [2,1]^T$.

Proof. If $[v]_{\mathcal{B}} = [a_1, \ldots, a_n]$ are the new coordinates of v, this means $v = a_1v_1 + \cdots + a_nv_n$. But that means $v = S[v]_{\mathcal{B}}$. Since $\mathcal{B}$ is a basis, S is invertible and $[v]_{\mathcal{B}} = S^{-1}v$. $\square$

Theorem: If $T(x) = Ax$ is a linear map and S is the matrix from a basis change, then $B = S^{-1}AS$ is the matrix of T in the new basis $\mathcal{B}$.

Proof. Let $y = Ax$. The statement $[y]_{\mathcal{B}} = B[x]_{\mathcal{B}}$ can be written using the last theorem as $S^{-1}y = BS^{-1}x$ so that $y = SBS^{-1}x$. Combining with $y = Ax$, this gives $B = S^{-1}AS$. $\square$

5.4. If two matrices A, B satisfy $B = S^{-1}AS$ for some invertible S, they are called **similar**. The matrices A and B both implement the transformation T, but they do it from a different perspective. It makes sense to adapt the basis to the situation. For example, here on earth, at a specific location, we use a coordinate system, where v_1

points east, where v_2 points north and where v_3 points straight up. The natural basis here in Boston is different than the basis in Zürich. [1]

FIGURE 1. A good coordinate system is adapted to the situation. When talking about points on the globe, we can use a global coordinate system with e_3 in the earth axes. When working on earth say near Boston, we need another basis.

5.5. Using a suitable basis is one of the main reasons why linear algebra is so powerful. This idea will be a major one throughout the course. We will use "eigen basis" to diagonalize a matrix, we will use good coordinates to solve ordinary and partial differential equations.

5.6. For us, the change of coordinates now is a way to figure out the matrix of a transformation

To find the matrix A of a reflection, projection or rotation matrix, find a good basis for the situation, then look what happens to the new basis vectors. This gives B. Now write down the matrix S and get $A = SBS^{-1}$.

EXAMPLES

5.7. Problem. Find the matrix A which implements the reflection T at the plane $X = \{x + y + 2z = 0\}$. **Solution.** We take a basis adapted to the situation. Take $v_3 = [1, 1, 2]^T$ which is perpendicular to the plane, then choose $v_1 = [1, -1, 0]^T, v_2 = [2, 0, -1]^T$ which are in the plane. Now, since $T(v_1) = v_1, T(v_2) = v_2$ and $T(v_3) = -v_3$, the transformation is described in that basis with the matrix $B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$. The

[1] Figure (1) was rendered using creative commons Povray code by Gilles Tran authored 2000-2004.

basis change transformation S is $S = \begin{bmatrix} 1 & 1 & 2 \\ 1 & -1 & 0 \\ 2 & 0 & -1 \end{bmatrix}$. We can now get $A = SBS^{-1} = \begin{bmatrix} -1 & -2 & 2 \\ 1 & 2 & -1 \\ 1 & 1 & 0 \end{bmatrix}$.

5.8. Problem. Find the matrix which rotates about the vector $[3, 4, 0]^T$ by 90 degrees counter clockwise when looking from the tip $(3, 4, 0)$ of the vector to the origin $(0, 0, 0)$.
Solution. We build a basis adapted to the situation. Of course, we use $v_1 = [3, 4, 0]^T$. We need now two other vectors which are perpendicular to each other. The vectors $v_2 = [-4, 3, 0]^T$ and $v_3 = [0, 0, 5]$ present themselves. It is good to have the two vectors which are moving to have the same length because then the matrix B is particularly simple: since $v_2 \to -v_3, v_3 \to v_2, v_1 \to v_1$, we have $B = \begin{bmatrix} 1 & 0 & 0 \\ x0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$. With $S = \begin{bmatrix} 3 & -4 & 0 \\ 4 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$, we get $A = SBS^{-1}$. This is now just matrix multiplication and can be computed $\begin{bmatrix} 9 & 12 & -100 \\ 12 & 16 & 75 \\ 4 & -3 & 0 \end{bmatrix}/25$. It would have been quite hard to find the column vectors of this matrix by figuring out where each of the standard basis vectors e_k goes. Still, we have used that basic principle when figuring out what B is.

5.9. Find the matrix of the projection on the line perpendicular to the hyperplane $x + y + z + w = 0$ in $\mathbb{R}^4$. Solution: there is a nice basis adapted to that situation. It gives

$$\mathcal{B} = \{v_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, v_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} v_3 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} v_4 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}\}, S = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}.$$

S is invertible. In this case $S^{-1} = 4S$. Now, in the new basis, the transformation matrix is very simple. As v_1 goes to v_1 and v_2 and v_3 and v_4 all go to zero, we have

$$B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

and $A = SBS^{-1} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}/4$. In this case, we might also have been able to write down the matrix without going to a new coordinate system as the image of the first basis vector is the vector projection of $[1, 0, 0, 0]$ onto $[1, 1, 1, 1]$.

Homework

This homework is due on Tuesday, 2/13/2019.

Problem 5.1: What are the $\mathcal{B}$-coordinates of $\vec{v}$ in the basis $\mathcal{B}$.

$$\vec{v} = \begin{bmatrix} 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}, \mathcal{B} = \{ \begin{bmatrix} 3 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \} ?$$

Problem 5.2: What is the matrix B for the transformation $A = \begin{bmatrix} 3 & 1 \\ -1 & 4 \end{bmatrix}$ in the basis $\mathcal{B} = \{ \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \}$:

Problem 5.3: Chose a suitable basis to solve:
a) What matrix A implements the reflection at the plane $3x+3y+6z=0$?
b) What matrix A implements the reflection at the line spanned by $[2,2,4]^T$?

Problem 5.4: Find the matrix A corresponding to the orthogonal projection onto the plane spanned by the vectors $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$.

Problem 5.5: "Graphene" are hexagonal planar structures. We can work with them when using a good adapted basis. Assume the first is $v = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. a) Find w so that $\mathcal{B} = \{v, w\}$ is the basis as seen in the picture.
b) What are the standard coordinates of $\begin{bmatrix} 3 \\ -1 \end{bmatrix}_{\mathcal{B}}$?
c) Is $\begin{bmatrix} 23 \\ 72 \end{bmatrix}_{\mathcal{B}}$ a vertex of a hexagon or the center of one?

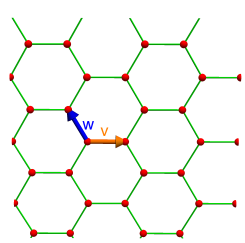


Figure 2. Graphene are single layer hexagonal lattice carbon structures.

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 6: Prototypes

Seminar

6.1. One of the powerful mind-hacks of **Richard Feynman** was his ability to **use examples** to make a point. He was known for acute sharpness to spot errors in arguments and once explained what was the secret behind that: rather than following the general abstract presentation, he would make up a basic example and project the general case to that. The principle is simple: if the example fails, then the general theory fails. While examples do not cover an entire theory, they serve as **prototypes** which capture the theory. We want to explore this today.

Figure 1. Richard Feynman playing bongo drums (1962).

[1]

[1]We colorized the picture and enlarged part to get the right picture. The book "Feynman Tips on Physics" uses the left picture as a cover and states that the photographer is unknown. The widespread use, both commercially as well as in picture collections, appears to put the photo into the public domain.

6.2. We explore today the power of examples. Examples not only provide **counter examples** to statements we believe to be true, examples also are **prototypes** for general results. In many cases, one can reduce the general case to a prototype. Feynman was good at listening to a general statement, then running it through a well chosen example. This immediately allowed him to spot errors simply because the structure failed in the example. This principle is in particular powerful in linear algebra.

6.3. Let's look at the example of a reflection at a line. We have derived in unit 2, why the matrix of a reflection at a line containing the vector $[\cos(\alpha), \sin(\alpha)]$ was the matrix $A = \begin{bmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{bmatrix}$. A special case is if we deal with the x-axis. In that case, the matrix is $B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.

6.4. Let's look at the statement

Theorem: A reflection at a line preserves length and angles.

Problem A: Look at a simple special case to verify that it works if the basis is chosen in an appropriate way.

6.5. Checking the result in a special coordinate system is simpler than verifying the theorem in full generality by comparing the lengths of

$$v = \begin{bmatrix} x \\ y \end{bmatrix}, Av = \begin{bmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix},$$

and then checking that the dot product between $[x, y]^T$ and $[u, v]^T$ is the same as the dot product between $A[x, y]^T$ and $A[u, v]^T$. For the later, there is an elegant reduction to length:

Problem B: Verify that if one can measure length $|x|$ in a linear space, then one can also get the dot product. Hint: compute $|x - y|^2, |x + y|^2$ and $|x|^2$ and $|y|^2$ and see how to get $x \cdot y$.

6.6. Why does the fact that a reflection preserves the dot product imply that angles are preserved?

Problem C: How again can one get the angle between two vectors from the dot product?

6.7. Here is a tougher question: is it possible that for a 2×2 matrix A and a vector x, the length of $A^n x$ grows quadratically?

Problem D: Explore the quadratic growth question with the eyes of Feynman. We don't have the tools yet to answer this definitely, but we want you to try finding a 2×2 matrix, where this happens. As you don't succeed finding a 2×2 matrix, can you find a 3×3 matrix where quadratic growth happens?

6.8. In any subject, whether it is science, art, law or humanities, it is good to have a collection of prototypes available. One should have examples which are **typical** and also have examples which are **unusual** and **special**. The latter serve as counter examples or examples of surprising things. Here are some examples in linear algebra:

- $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ is an example of a matrix which has the property that $A^2 = -1$. This can not be realized in the real numbers.
- $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ is an example of a matrix which has the property that $A^2 = 0$. This can not be realized in the real numbers without being zero. It is also an example, for which the image of A is the kernel of A.

6.9. Now we can ask ourselves, what is the collection of all the 2×2 matrices which have the property that $A^2 = -1$? This is a tougher question but we can easily give more examples by changing basis. If $B = S^{-1}AS$, then $B^2 = S^{-1}ASS^{-1}AS = S^{-1}A^2S = S^{-1}(-1)S = -S^{-1}S = -1$. We can also just by brute force, look for $A^2 = -1$ and see that $d = -a$ and $cb = -1 - a^2$. in the second case $A^2 = 0$, we get $d = -a$ and $cb^2 = -a^2$.

6.10. Let $r(t) = [x(t), y(t)]$ denote a closed curve. To play **billiards** inside this table, we reflect a ball with the **reflection rule** telling that the incoming angle is equal to the outgoing angle.

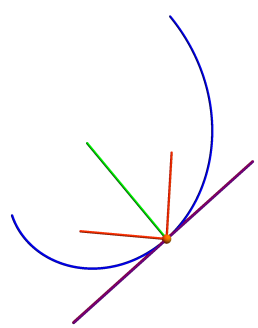


FIGURE 2. How do we compute the reflection at the table boundary?

Assume we have $r(t) = [\sin(t) + t^2, t^4 + t]$ and want to compute the matrix A which reflects in coming ball with velocity v at the point $r(0)$. The outgoing vector should be Av.

Problem E: What is a good basis for this problem? How does the reflection matrix B look like in this coordinate system?

Problem F: If you have time, find the matrix A.

HOMEWORK

Exercises A)-D) are done in the seminar. This homework is due on Tuesday:

Problem 6.1 a) Is it true that if A is an invertible 2×2 matrix with rational entries, then A^{-1} is a 2×2 matrix with rational entries? If yes, give a proof.
b) Is it true that if A is an invertible 2×2 matrix, where all entries are 1 or 0, then A^{-1} is a 2×2 matrix, where all entries are 0 or 1.
c) Is it true that if A is an invertible 2×2 matrix where all entries are 1 or 0, then A^{-1} is a 2×2 matrix, where all entries are 0 or 1 or -1.

Problem 6.2 a) Is it true that if A is an invertible $n \times n$ matrix with rational entries, then A^{-1} is a $n \times n$ matrix with rational entries. b) Is it true that if A is an invertible $n \times n$ matrix where all entries are 1 or 0, then A^{-1} is a $n \times n$ matrix where all entries are 0 or 1 or -1.

Problem 6.3 A matrix is called **unimodular**, if its determinant is 1.
a) Is it true that if A is a unimodular 2×2 matrix with integer entries, then its inverse is a 2×2 matrix with integer entries?
b) Is it true that if A is a unimodular 3×3 matrix with integer entries, then its inverse is a 3×3 matrix with integer entries?

Problem 6.4 Let's look at a plane $ax + by + cz = 0$, where a, b, c are integers in which not all are zero.
a) Is it true that the matrix of reflection at the plane has integer entries?
b) Is it true that the matrix of reflection at the plane has rational entries?
c) Is it true that the matrix of orthogonal projection onto the plane has integer entries?
d) Is it true that the matrix of orthogonal projection onto the plane has rational entries?

Problem 6.5 Experiment with a computer algebra system. Take a random matrix and take its inverse and then form its exponential. Make some pictures like $MatrixPlot[MatrixExp[Table[Random[], \{100\}, \{100\}]]]$ Explore using experiments, whether you can find A where $\exp(A)$ or A^{-1} look random.

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 7: Gram-Schmidt

Lecture

7.1. For vectors in the linear space $\mathbb{R}^n$, the **dot product** is defined as $v \cdot w = \sum_i v_i w_i$. More generally, in the linear space $M(n,m)$ there is a natural dot product $v \cdot w = \text{tr}(v^T w)$, where tr is the trace, the sum of the diagonal entries. It is the sum $\sum_{i,j} v_{ij} w_{ij}$. The dot product allows to compute **length** $|v| = \sqrt{v \cdot v}$ and **angles** α between two vectors defined by the equation $v \cdot w = |v||w| \cos(\alpha)$. If the relation $v \cdot w = 0$ holds, the vectors v and w are called **orthogonal**.

7.2. A collection of pairwise orthogonal vectors $\{v_1, v_2, \cdots, v_n\}$ in $\mathbb{R}^n$ is linearly independent because $a_1 v_1 + \cdots a_n v_n = 0$ implies that $v_k \cdot (a_1 v_1 + \cdots a_n v_n) = a_k v_k \cdot v_k = a_k |v_k|^2 = 0$ and so $a_k = 0$. A collection of n orthogonal vectors therefore automatically forms a basis.

7.3. Definition. A basis is called **orthonormal** if all vectors have length 1 and are orthogonal. Why do we like to have an orthogonal basis? One reason is that an orthogonal basis looks like the standard basis. Another reason is that rotations preserving a space V or orthogonal projections onto a space V are easier to describe if an orthogonal basis is known on V. Let's look at projections as we will need them to produce an orthonormal basis. Remember that the **projection** of a vector x onto a unit vector v is $(v \cdot x)v$. We can now give the matrix of a projection onto a space V if we know an orthonormal basis in V:

> **Lemma:** If $\mathcal{B} = \{v_1, v_2, \cdots, v_n\}$ is an orthonormal basis in V, then the projection P onto V satisfies $Px = (v_1 \cdot x)v_1 + \cdots + (v_n \cdot x)v_n$.

Proof. By Pythagoras, $(x - Px) \cdot x = |x|^2 - (v_1 \cdot x)^2 - \cdots - (v_n \cdot x)^2 = 0$, so that $x - Px$ is perpendicular to x.

Let Q be the matrix containing the basis v_k as columns. We can rewrite the result as $\boxed{P = QQ^T}$. We write Q because it is not a $n \times n$ matrix like S. The matrix Q contains the basis of the subspace V and not the basis of the entire space. We will see next week a more general formula for P which also holds if the vectors are not perpendicular.

7.4. Let $v_1, \dots, v_n$ be a basis in V. Let $w_1 = v_1$ and $u_1 = w_1/|w_1|$. The **Gram-Schmidt process** recursively constructs from the already constructed orthonormal set $u_1, \dots, u_{i-1}$ which spans a linear space V_{i-1} the new vector $w_i = (v_i - \text{proj}_{V_{i-1}}(v_i))$ which is orthogonal to V_{i-1}, and then normalizes w_i to get $u_i = w_i/|w_i|$. Each vector w_i is orthonormal to the linear space V_{i-1}. The vectors $\{u_1, \dots, u_n\}$ form then an orthonormal basis in V.

7.5. The formulas can be written as

$$v_1 = |w_1| u_1 = r_{11} u_1$$

...

$$v_i = (u_1 \cdot v_i) u_1 + \cdots + (u_{i-1} \cdot v_i) u_{i-1} + |w_i| u_i = r_{1i} u_1 + \cdots + r_{ii} u_i$$

...

$$v_n = (u_1 \cdot v_n) u_1 + \cdots + (u_{n-1} \cdot v_n) u_{n-1} + |w_n| u_n = r_{1n} u_1 + \cdots + r_{nn} u_n \ .$$

In matrix form this means

$$A = \begin{bmatrix} | & | & \cdot & | \\ v_1 & v_2 & \cdot & v_n \\ | & | & \cdot & | \end{bmatrix} = \begin{bmatrix} | & | & \cdot & | \\ u_1 & u_2 & \cdot & u_n \\ | & | & \cdot & | \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \cdot & r_{1n} \\ 0 & r_{22} & \cdot & r_{2n} \\ 0 & 0 & \cdot & r_{nn} \end{bmatrix} = QR \ ,$$

where A and Q are $m \times n$ matrices and R is a $n \times n$ matrix with

$$\boxed{r_{ij} = v_j \cdot u_i, \text{ for } i < j \text{ and } v_{ii} = |w_i|}$$

We have just seen:

Theorem: A matrix A with linearly independent columns v_i can be decomposed as $A = QR$, where Q has orthonormal column vectors and where R is an upper triangular square matrix with the same number of columns than A. The matrix Q has the orthonormal vectors u_i in the columns.

7.6. The recursive process was stated first by Erhard Schmidt (1876-1959) in 1907. The essence of the formula was already in a 1883 paper by J.P.Gram in 1883 which Schmidt mentions in a footnote. The process seems to already have been anticipated by Laplace (1749-1827) and was also used by Cauchy (1789-1857) in 1836.

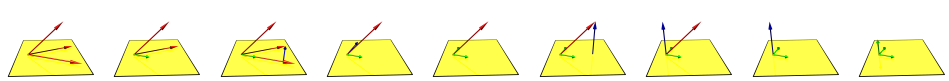


Figure 1.

Examples

7.7. Problem. Use Gram-Schmidt on $\{v_1 = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}, v_3 = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}\}$.

Solution. 1. $w_1 = \frac{v_1}{|v_1|} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $u_1 = w_1$.

2. $w_2 = (v_2 - \text{proj}_{V_1}(v_2)) = v_2 - (u_1 \cdot v_2)u_1 = \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix}$. $u_2 = \frac{w_2}{|w_2|} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$.

3. $w_3 = (v_3 - \text{proj}_{V_2}(v_3)) = v_3 - (u_1 \cdot v_3)u_1 - (u_2 \cdot v_3)u_2 = \begin{bmatrix} 0 \\ 0 \\ 5 \end{bmatrix}$, $u_3 = \frac{w_3}{|w_3|} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

7.8. From $v_1 = |v_1|u_1$, $v_2 = (u_1 \cdot v_2)u_1 + |w_2|u_2$, and $v_3 = (u_1 \cdot v_3)u_1 + (u_2 \cdot v_3)u_2 + |w_3|u_3$, we get the QR decomposition

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 0 & 3 & 2 \\ 0 & 0 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 0 & 3 & 2 \\ 0 & 0 & 5 \end{bmatrix} = QR\,.$$

7.9. One reason why we are interested in orthogonality is that in statistics, "orthogonal" means "uncorrelated". **Data** are often also arranged in matrices as relational databases. Let's take the matrices $v_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ and $v_2 = \begin{bmatrix} 0 & 3 \\ 3 & 0 \end{bmatrix}$. They span a two dimensional plane in the linear space $M(2,2)$ of 2×2 matrices. We want to have an orthogonal set of vectors in that plane. Now, how do we do that? We can use Gram-Schmidt in the same way as with vectors in $\mathbb{R}^n$. One possibility is to write the matrices as vectors like $v_1 = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}^T$ and $v_2 = \begin{bmatrix} 0 & 3 & 3 & 0 \end{bmatrix}^T$ and proceed with vectors. But we can also remain within matrices and do the Gram-Schmidt procedure in $M(2,2)$. Let us do that. The first step is to normalize the first vector. We get $u_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}/2$. The second step is to produce $w_2 = v_2 - (u_1 \cdot v_2)u_1 = \begin{bmatrix} 0 & 3 \\ 3 & 0 \end{bmatrix} - 3\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}/2 = \begin{bmatrix} -3/2 & 3/2 \\ 3/2 & -3/2 \end{bmatrix}$. Now, we have to normalize this to get $u_2 = w_2/|w_2| = w_2/3 = \begin{bmatrix} -1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix}$. Now, $\mathcal{B} = \{u_1, u_2\}$ is an orthonormal basis in the space X spanned by $\{v_1, v_2\}$.

7.10.

Theorem: If $S^T = S^{-1}$ the map $T : A \to S^{-1}AS$ is an orthogonal transformation from $M(n,n) \to M(n,n)$.

Proof. The dot product between $A_1 = S^{-1}AS$ and $B_1 = S^{-1}BS$ is equal to the one between A and B:

$$\text{tr}(A_1^T B_1) = \text{tr}((S^{-1}AS)^T S^{-1}BS) = \text{tr}((S^{-1}A^T SS^{-1}BS) = \text{tr}(S^{-1}A^T BS) = \text{tr}(A^T B)\,.$$

□

We have used in the last step that similar matrices always have the same trace. We prove this later. For 2×2 matrices we can check it by brute force:

```
A={{a,b},{c,d}};S={{p,q},{r,s}};
Simplify[Tr[Inverse[S].A.S]==Tr[A]]
```

Homework

This homework is due on Tuesday, 2/20/2019.

Problem 7.1: Perform the Gram-Schmidt process on the three vectors $\{ \begin{bmatrix} 5 \\ 5 \\ 5 \\ 5 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 4 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} \}$ and then write down the QR decomposition.

Problem 7.2: a) Find an orthonormal basis of the plane $x + y + z = 0$ and form the projection matrix $P = QQ^T$.
b) Find an orthonormal basis of the hyper plane $x_1 + x_2 + x_3 + x_4 + x_5 = 0$ in $\mathbf{R}^5$.

Problem 7.3: a) Produce an orthonormal basis of the kernel of

$$A = \begin{bmatrix} 1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & 1 & 1 \end{bmatrix}.$$

b) Write down an orthonormal basis for the image of A.

Problem 7.4: Find the QR factorization of the following matrices

$$A = \begin{bmatrix} 2 & 4 \\ 1 & 1 \end{bmatrix}, B = \begin{bmatrix} 0 & -3 & 0 \\ 0 & 0 & 0 \\ 7 & 0 & 0 \\ 0 & 0 & 4 \end{bmatrix}, C = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, D = \begin{bmatrix} 12 & 5 \\ -5 & 12 \end{bmatrix}.$$

Problem 7.5: Find the QR factorization of the following matrices (D was the quaternion matrix you have derived earlier)

$$A = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 5 \end{bmatrix},$$

$$B = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}, C = \begin{bmatrix} 5 & 12 \\ 12 & -5 \end{bmatrix}.$$

$$D = \begin{bmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{bmatrix}$$

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 8: The orthogonal group

Lecture

8.1. The **transpose** of a matrix A is $A^T_{ij} = A_{ji}$, the matrix A in which rows and columns are interchanged. The **transpose operation** $A \to A^T$ is a linear map from $M(n,m)$ to $M(m,n)$. Here are some properties of this operation:

a) $(AB)^T = B^T A^T$
b) $x \cdot Ay = A^T x \cdot y$.
c) $(A^T)^T = A$
d) $(A^T)^{-1} = (A^{-1})^T$
e) $(A+B)^T = A^T + B^T$

Proof: a) $(AB)^T_{kl} = (AB)_{lk} = \sum_i A_{li} B_{ik} = \sum_i B^T_{ki} A^T_{il} = (B^T A^T)_{kl}$.
b) $x \cdot Ay = x^T Ay = (A^T x)^T y = A^T x \cdot y$.
c) $((A^T)^T)_{ij} = (A^T)_{ji} = A_{ij}$.
d) $1_n = 1_n^T = (AA^{-1})^T = (A^{-1})^T A^T$ using a).
e) $(A+B)^T_{ij} = (A+B)_{ji} = A_{ji} + B_{ji} = A^T_{ij} + B^T_{ij}$.

8.2. An $n \times n$ matrix A is called **orthogonal** if $A^T A = 1$. The identity $A^T A = 1$ encodes the information that the columns of A are all perpendicular to each other and have length 1. In other words, the columns of A form an **orthonormal basis**. [1]

8.3. Examples of orthogonal matrices are rotation matrices and reflection matrices. These two types are the only 2×2 matrices which are orthogonal: the first column vector has as a unit vector have the form $[\cos(t), \sin(t)]^T$. The second one, being orthogonal has then two possible directions. One is a rotation, the other is a reflection. In three dimensions, a reflection at a plane, or a reflection at a line or a rotation about an axis are orthogonal transformations. For 4×4 matrices, there are already transformations which are neither rotations nor reflections.

8.4. Here are some properties of orthogonal matrices:

[1] Why not call it **orthonormal matrix**? It would make sense, but **orthogonal matrix** is already strongly entrenched terminology.

a) If A is orthogonal, $A^{-1} = A^T$.
b) If A is orthogonal, then not only $A^T A = 1$ but also $AA^T = 1$.

Theorem: A transformation is orthogonal if and only if it preserves length and angle.

Proof. Let us first show that an orthogonal transformation preserves length and angles. So, let us assume that $A^T A = 1$ first. Now, using the properties of the transpose as well as the definition $A^T A = 1$, we get $|Ax|^2 = Ax \cdot Ax = A^T Ax \cdot x = 1x \cdot x = x \cdot x = |x|^2$ for all vectors x. Let α be the angle between x and y and let β denote the angle between Ax and Ay and α the angle between x and y. Using $Ax \cdot Ay = x \cdot y$ again, we get $|Ax||Ay|\cos(\beta) = Ax \cdot Ay = x \cdot y = |x||y|\cos(\alpha)$. Because $|Ax| = |x|, |Ay| = |y|$, this means $\cos(\alpha) = \cos(\beta)$. As we have defined the angle between two vectors to be a number in $[0, \pi]$ and cos is monotone on this interval, it follows that $\alpha = \beta$.
To the converse: if A preserves angles and length, then $v_1 = Ae_1, \cdots v_n = Ae_n$ form an orthonormal basis. By looking at $B = A^T A$ this shows off diagonal entries of B are 0 and diagonal entries of B are 1. The matrix A is orthogonal. □

8.5. Orthogonal transformations form a group with multiplication:

Theorem: The composition and the inverse of two orthogonal transformations is orthogonal.

Proof. The properties of the transpose give $(AB)^T AB = B^T A^T AB = B^T B = 1$ so that AB is orthogonal if A and B are. The statement about the inverse follows from $(A^{-1})^T A^{-1} = (A^T)^{-1} A^{-1} = (AA^T)^{-1} = 1$. □

8.6. Orthogonal transformations are important because they are natural **symmetries**. Many coordinate transformations are orthogonal transformations. We will also see that the Fourier expansion is a type of orthogonal transformation.

Examples

8.7. Here is an orthogonal matrix, which is neither a rotation, nor a reflection. it is an example of a **partitioned matrix**, a matrix made of matrices. This is a nice way to generate larger matrices with desired properties. The matrix

$$A = \begin{bmatrix} \cos(1) & -\sin(1) & 0 & 0 \\ \sin(1) & \cos(1) & 0 & 0 \\ 0 & 0 & \cos(3) & \sin(3) \\ 0 & 0 & \sin(3) & -\sin(3) \end{bmatrix}$$

produces a rotation in the xy-plane and a reflection in the zw-plane. It is not a reflection because A^2 is not the identity. It is not a rotation either as the determinant is not 1 nor -1. We will look at determinants later.

8.8. What is the most general rotation in three dimensional space? How many parameters does it need? You can see this when making a photograph. You use two angles to place the camera direction, but then we can also turn the camera.

8.9. What is the most general rotation matrix in three dimensions? We can realize that as

$$A = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos(\phi) & 0 & -\sin(\phi) \\ 0 & 1 & 0 \\ \sin(\phi) & 0 & \cos(\phi) \end{bmatrix} \begin{bmatrix} \cos(\gamma) & -\sin(\gamma) & 0 \\ \sin(\gamma) & \cos(\gamma) & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

8.10. Problem: Find a rotation which rotates the earth (latitude,longitude)=(a_1, b_1) to a point with (latitude,longitude)=(a_2, b_2)? Solution: The matrix which rotates the point $(0, 0)$ to (a, b) is a composition of two rotations. The first rotation brings the point into the right latitude, the second brings the point into the right longitude. $R_{a,b} = \begin{bmatrix} \cos(b) & -\sin(b) & 0 \\ \sin(b) & \cos(b) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos(a) & 0 & -\sin(a) \\ 0 & 1 & 0 \\ \sin(a) & 0 & \cos(a) \end{bmatrix}$. To bring a point (a_1, b_1) to a point (a_2, b_2), we form $A = R_{a_2,b_2} R^{-1}_{a_1,b_1}$.
With Cambridge (USA): $(a_1, b_1) = (42.366944, 288.893889)\pi/180$ and Zürich (Switzerland): $(a_2, b_2) = (47.377778, 8.551111)\pi/180$, we get the matrix

$$A = \begin{bmatrix} 0.178313 & -0.980176 & -0.0863732 \\ 0.983567 & 0.180074 & -0.0129873 \\ 0.028284 & -0.082638 & 0.996178 \end{bmatrix}.$$

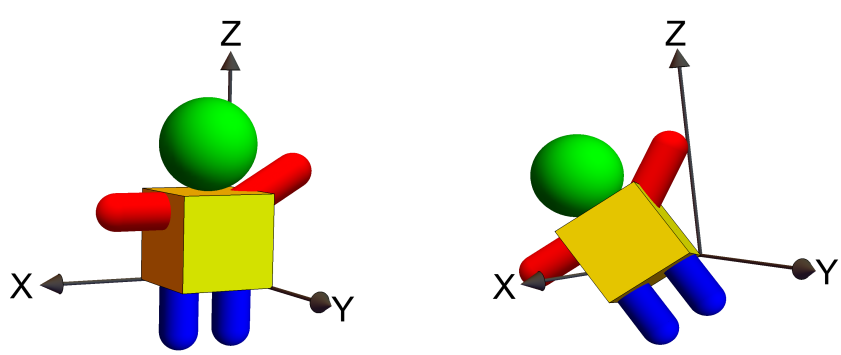


FIGURE 1. The rotation group $SO(3)$ is three dimensional. There are three angles which determine a general rotation matrix.

Homework

This homework is due on Tuesday, 2/20/2019.

Problem 8.1: Which matrices are orthogonal? a) $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$, b) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix}$, c) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$. d) $[-1]$, e) $\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, f) $\begin{bmatrix} \cos(7) & -\sin(7) & 0 & 0 \\ \sin(7) & \cos(7) & 0 & 0 \\ 0 & 0 & \cos(1) & \sin(1) \\ 0 & 0 & \sin(1) & -\cos(1) \end{bmatrix}$, g) $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

Problem 8.2: If A, B are orthogonal, then
a) Is A^T orthogonal? b) Is B^{-1} orthogonal? c) Is $A - B$ orthogonal? d) Is $A/2$ orthogonal? e) Is $B^{-1}AB$ orthogonal? f) Is BAB^T orthogonal?

Problem 8.3: Rotation-Dilation matrices $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ form an algebra in which the multiplication $(a + ib)(c + id)$ corresponds to the matrix multiplication $\begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} c & -d \\ d & c \end{bmatrix}$. Draw the vectors $z = \sqrt{3} + i$, $w = 1 - i$ in the complex plane, then draw its product. Find the angles α, β and lengths r,s so that $z = re^{i\alpha}, w = se^{i\beta}$ and verify that $zw = rse^{i(\alpha+\beta)}$.

Problem 8.4: Matrices of the form $A(p, q, r, s) = \begin{bmatrix} p & -q & -r & -s \\ q & p & s & -r \\ r & -s & p & q \\ s & r & -q & p \end{bmatrix}$ are called **quaternions**. a) Find a basis for the set of quaternion matrices. b) Under which condition is $A(p, q, r, s)$ orthogonal?

Problem 8.5: a) Explain why the identity matrix is the only $n \times n$ matrix that is orthogonal, upper triangular and has positive entries on the diagonal. b) Conclude, using a) that the QR factorization of an invertible $n \times n$ matrix A is unique. That is, if $A = Q_1R_1$ and $A = Q_2R_2$ are two factorizations, argue why $Q_1 = Q_2$ and $R_1 = R_2$.

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 9: SU(2)

Seminar

9.1. We have seen this week that the set of orthogonal $n \times n$ matrices forms a group. This means that one can multiply two such matrices AB and still have an orthogonal matrix. Also, the inverse of an orthogonal matrix is orthogonal. The group of orthogonal $n \times n$ matrices is called $O(n)$. We will talk about determinants later in this course but there is a subgroup $SO(n)$ consisting of all orthogonal matrices which have determinant 1. It is a subgroup because $\det(AB) = \det(A)\det(B)$ and $\det(A^T) = \det(A)$. If you have time, verify these identities in the case of 2×2 matrices or 3×3 matrices at the end of this seminar.

9.2. The class of 2×2 matrices with determinant 1 forms a group called $SL(2, R)$. It is called the **special linear group**. We can say that the intersection of $SL(2, R)$ with $O(2)$ is $SO(2)$. It is the special orthogonal group or rotation group in two dimensions.

Problem A: a) Write down the general matrix in the form $SO(2)$.
b) Write down the 2×2 matrix in $O(2)$ which is not in $SO(2)$?

Problem B: Can you find a continuous path $A(t)$ of matrices, where $A(0)$ is an orthogonal 2×2 matrix with determinant 1 and $A(1)$ is an orthogonal 2×2 matrix with determinant -1? If yes, give a path connecting $1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ with $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. If none should exist, prove why such a deformation can not exist. What about 3×3 matrices? Can one find a path $A(t)$ in $O(3)$ such that $A(0) = 1$ and $A(1)$ is a reflection at the xy-plane?

9.3. In this course, we again use complex numbers. Let us just recall some things: the set of complex numbers $z = a + ib$ is a linear space and called $\mathbb{C}$. It is also an **algebra** because we can not only add but also multiply. The algebra $\mathbb{C}$ is isomorphic to the space of **rotation dilation matrices**. Indeed, if $z = a + ib$, then we can write $1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $i = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ so that $a1 + bi$ is a rotation dilation matrix. We can say that the rotation-dilation group is a **representation** of the complex algebra.

9.4. Let us look now at complex 2×2 matrices. The same computation as above shows that $\det(AB) = \det(A)\det(B)$. The matrices with determinant 1 are called $SL(2,\mathbb{C})$. They form a group too.

> **Problem C:** Why is $SL(2,\mathbb{C})$ a group? How does one invert a general matrix in $SL(2,\mathbb{C})$?

9.5. As for real matrices, we can look at the transpose matrix A^T. In the complex, there is a more natural involution on matrices, which is called the **adjoint** $A^* = \overline{A}^T$. To get the adjoint of a matrix, one both transposes and conjugates it. One reason why one takes also the conjugate is that then the inner product $A \cdot B = \mathrm{tr}(A^*B)$ behaves like an inner product and especially has the property that $\mathrm{tr}(A^*A) = \sum_{i,j} |A_{ij}|^2$ is non-zero for nonzero matrices.

9.6. The analogue of orthogonal matrices are the **unitary matrices.** A matrix is called **unitary** if $U^*U = 1$. The class of unitary 2×2 matrices is denoted $U(2)$. It has a subgroup $SU(2)$ of unitary transformations with determinant 1. It is called the **special unitary group**.

9.7. So, $SU(2)$ consists of all matrices

$$A = \begin{bmatrix} z & -\overline{w} \\ w & \overline{z} \end{bmatrix}$$

with $\det(A) = |z|^2 + |w|^2 = 1$.

> **Problem C:** Write $z = a + ib$ and $w = c + id$ and verify that $SU(2)$ is the **three dimensional sphere** $a^2 + b^2 + c^2 + d^2 = 1$ in $\mathbb{R}^4$.

9.8. We have mentioned the three dimensional sphere a couple of times in Math 22a and seen that the volume is $2\pi^2$. Every compact simply connected three dimensional manifold is topologically a 3-sphere. Very interesting is that S^3 is besides the circle the only unit sphere in a Euclidean space which is also a Lie group.

9.9. We have seen quaternions a couple of times already. These are numbers of the form $a+bi+cj+dk$. The norm of a quaternion was $\sqrt{a^2 + b^2 + c^2 + d^2}$. If a quaternion has norm 1, then it is a **unit quaternion**.

> **Problem D:** Verify that if $a + ib + cj + dk$ is a unit quaternion, then
>
> $$\begin{bmatrix} a + ib & c + id \\ -c + id & a - ib \end{bmatrix}$$
>
> is a matrix in $SU(2)$.

9.10.

> **Problem E:** Verify the determinant identity $\det(AB) = \det(A)\det(B)$ for 2×2 matrices. Either do it by hand or run the following Mathematica code to verify it. While you are at it, verify it also for 3×3 matrices.

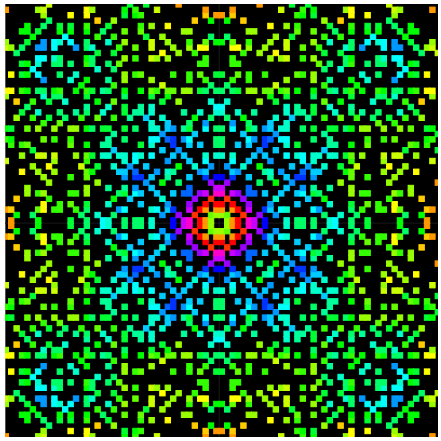


FIGURE 1. We see a few Hurwitz primes on the slice $a = 1/2, b = 1/2$ in the space of integer quaternions (a, b, c, d). For this illustration we take primes of the form $(a, b, c, d) = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}) + (0, 0, k, l)$. This is a prime if and only if $1 + k + k^2 + l + l^2$ is a **rational prime** (usual prime). In the picture, a Hurwitz prime is colored according how many primes there are close by.

```
A={{a,b},{c,d}};   B={{p,q},{r,s}};
Simplify[Det[A.B] == Det[A] Det[B]]
Det[Transpose[A]]==Det[A]
```

9.11. Finally, we should mention that the group $SU(2)$ is of great interest in physics as it is the gauge group of the **electroweak interaction**.

FIGURE 2. Sheldon Glashow, Abdus Salam and Steven Weinberg won the Nobel Prize for electroweak unification in 1979.

Homework

Exercises A)-E) are done in the seminar. This homework is due on Tuesday:

Problem 9.1 Verify that if $X = a + ib + jc + kd$ is a quaternion, then the corresponding matrix

$$Q(X) = \begin{bmatrix} a+ib & c+id \\ -c+id & a-ib \end{bmatrix}$$

has the property that $Q(XY) = Q(X)Q(Y)$. Either do it by hand or look up the Mathematica code presented in unit 40 of Math 22a.

Problem 9.2 Assuming the property $\det(AB) = \det(A)\det(B)$ for 2×2 matrices, use the previous problem to see that if $N(X) = N(a+ib+jc+kd) = a^2 + b^2 + c^2 + d^2$, then $N(XY) = N(X)N(Y)$.
P.S. This compatibility of the quaternions with multiplication makes $\mathbb{H}$ a **division algebra**. One can divide X/Y by forming $X\overline{Y}/|Y|^2$ where $\overline{a+ib+jc+kd} = a - ib - jc - kd$.

Problem 9.3 a) The **Lagrange four square theorem** assures that every positive integer can be written as a sum of four integer squares. How does the previous identity assure that the theorem is proven if one has verified it for primes?
b) Find the smallest integer which can not be written as a sum of 3 squares.
c) Find the smallest integer which can not be written as a sum of 4 non-negative cubes.

Problem 9.4 a) Find a quaternion prime of norm 17.
b) Use the Quaternion identity $Q(XY) = Q(X)Q(Y)$ to find a, b, c, d such that $a^2 + b^2 + c^2 + d^2 = 229 * 179 = 40991$.

Problem 9.5 a) The **Hurwitz quaternions** in $\mathbb{Q}$ are the analog of integers in $\mathbb{R}$ They are either quaternon numbers of the form $a + ib + jc + kd$ with integers a, b, c, d or then half integers $(a + ib + jc + kd) + (1/2, 1/2, 1/2, 1/2)$. A Hurwitz quaternion X is called **prime**, if it can not be written as a product of two Hurwitz quaternions with smaller norm. How can you find primes?
b) Write down the 24 Hurwitz quaternions X have norm $N(X) = 1$. Eight of them have integer entries. Sixteen have half integer entries. They are called the unit quaternions. They form a group. Why?

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 10: Data Fitting

Lecture

10.1. A $n \times m$ matrix A defines a linear space $V = \text{im}(A) \subset \mathbb{R}^n$. If there is no kernel of A, then the image is a m-dimensional subspace of the n-dimensional space. Given a vector $b \in \mathbb{R}^n$, we can solve the equation $Ax = b$ only if b is in V. If n is larger than m, we in general can not find a solution, but we can try to find a "solution which is best".

10.2. The best possible solution x^* to this equation is the vector x^* for which Ax^* is closest to b. This means that $b - Ax^*$ is perpendicular to V, meaning that x^* is in the kernel of A^T. In other words, $A^T(b - Ax^*) = 0$ or $A^T b = A^T A x^*$. We see:

$$x^* = (A^T A)^{-1} A^T b.$$

10.3. The vector x^* is called the **least square solution** of $Ax = b$ because the $|Ax - b|^2$ is minimized. This is a sum of squares. As a bonus, we get a formula for the orthogonal projection P of $\mathbb{R}^n$ to the m-dimensional space V.

$$P = A(A^T A)^{-1} A^T \text{ is the orthogonal projection onto } \text{im}(A).$$

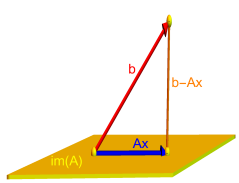


Figure 1. $Ax - b$ perpendicular to $\text{im}(A)$ means $Ax - b$ is in $\ker(A^T)$.

10.4. The least square principle is one of the most important paradigms in mathematics. It solves an optimization problem. The use is everywhere. If we write down a wish-list of things we want to have satisfied in the form of a system of linear equations $Ax = b$, we are in a situation where the least square solution is the best possible form. In virtually all problems of science, technology or engineering, we are exposed to a huge amount of data which we want to organize using few parameters. By projecting

the data onto the image of a matrix, we get the **best possible solution** for the model at hand.

10.5. In order for the projection formula to work, we needed the formula $(\mathrm{im}(A))^\perp = \ker(A^T)$. We also need to know when A^TA is invertible. This is the case if and only if A does not have a kernel:

Theorem: $\ker(A) = \ker(A^TA)$

Proof. We prove that in class. You should be able to reproduce this proof yourself. □

Examples

10.6. The maximum $f(x) = \max_A \log(\max_{ij} A^{-1}_{ij})$ over all Boolean $x \times x$ matrices A is measured with the computer and gives the data points $(x, y = f(x))$ given by $(3, \log(1))$, $(4, \log(2))$, $(5, \log(3))$, $(6, \log(5))$, $(7, \log(9))$, $(8, \log(18))$, $(9, \log(36))$. How do these data grow? We are not aware that this problem has been asked already so that we have to experiment first. What is the best linear fit? To find the best function $ax + c = y$, we plug in the values x_k, y_k in this equation, to get

$$\begin{array}{rcl} a3+c &=& \log(1) \\ a4+c &=& \log(2) \\ a5+c &=& \log(3) \\ a6+c &=& \log(5) \\ a7+c &=& \log(9) \\ a8+c &=& \log(18) \\ a9+c &=& \log(36) \end{array}, A = \begin{bmatrix} 3 & 1 \\ 4 & 1 \\ 5 & 1 \\ 6 & 1 \\ 7 & 1 \\ 8 & 1 \\ 9 & 1 \end{bmatrix}, b = \begin{bmatrix} \log(1) \\ \log(2) \\ \log(3) \\ \log(5) \\ \log(9) \\ \log(18) \\ \log(36) \end{bmatrix}.$$

We compute $(A^TA)^{-1} = \begin{bmatrix} \frac{1}{28} & -\frac{3}{14} \\ -\frac{3}{14} & \frac{10}{7} \end{bmatrix}$ and $A^Tb = [88.6775, 12.0723]^T$. The best parameters are $(a, c) = (-0.595901, 0.580129)$ which corresponds to the line $y = -0.595901x + 0.580129$.

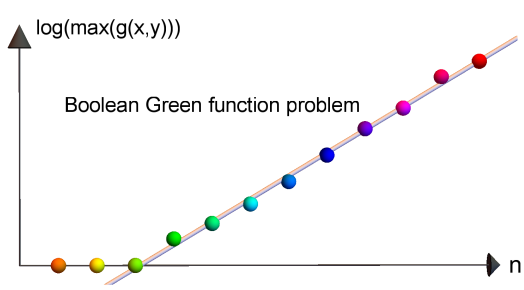


Figure 2. Linear data fitting of the Boolean Green function problem: $\max A^{-1}_{ij}$ among all invertible Boolean $n \times n$ matrices A.

10.7. Problem. Find a quadratic polynomial $p(x) = ux^2 + vx + w$ which best fits the four data points $(-1, 8), (0, 8), (1, 4), (2, 16)$. **Solution.** We write down the equations $ux^2 + vx + w = y$, for every data point (x, y). This gives us a system of four equations

$Ax = b$ with $A = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{bmatrix}$ $b = \begin{bmatrix} 8 \\ 8 \\ 4 \\ 16 \end{bmatrix}^T$. $A^T A = \begin{bmatrix} 18 & 8 & 6 \\ 8 & 6 & 2 \\ 6 & 2 & 4 \end{bmatrix}$. The solution is $x^* = (A^T A)^{-1} A^T b = \begin{bmatrix} 3, -1, 5 \end{bmatrix}^T$.

10.8. Problem. Prove that $\text{im}(A) = \text{im}(AA^T)$.
Solution. The image of AA^T is contained in the image of A because we can write $v = AA^T x$ as $v = Ay$ with $y = A^T x$. On the other hand, if v is in the image of A, then $v = Ax$. If $x = y + z$, where y in the kernel of A and z orthogonal to the kernel of A, then $Ax = Az$. Because z is orthogonal to the kernel of A, it is in the image of A^T. Therefore, $z = A^T u$ and $v = Az = AA^T u$ is in the image of AA^T.

10.9. Here is a solution with a higher dimensional fitting problem: **Problem.** Which paraboloid $ax^2 + by^2 = z$ best fits the data

x	y	z
0	1	2
-1	0	4
1	-1	3

Solution. Find the least square solution for the system of equations for the unknowns a, b.

$$\begin{aligned} a0 + b\,1 &= 2 \\ a1 + b\,0 &= 4 \\ a1 + b\,1 &= 3\,. \end{aligned}$$

In matrix form this can be written as $Ax = b$ with

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}, b = \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix}.$$

We have $A^T A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ and $A^T b = \begin{bmatrix} 7 \\ 5 \end{bmatrix}$. We get the least square solution with the formula

$$x = (A^T A)^{-1} A^T b = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

The best fit is the function $\boxed{f(x, y) = 3x^2 + y^2}$. It has as a graph an elliptic paraboloid.

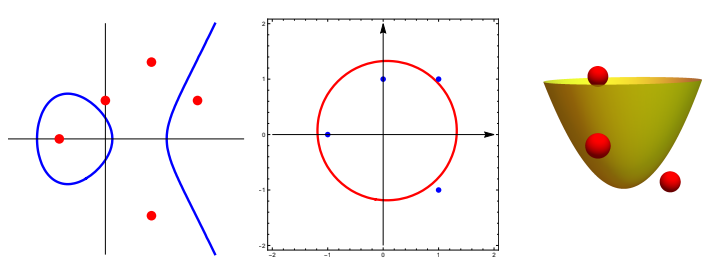


FIGURE 3. We can fit virtually anything. We see a fit with elliptic curves, circles or higher dimensional objects like quadrics.

Homework

This homework is due on Tuesday, 2/27/2019.

Problem 10.1: The first 7 prime numbers $2, 3, 5, 7, 11, 13$ define the data points $(1, 2), (2, 3), (3, 5), (5, 7), (6, 11), (7, 13)$ in the plane. Find the best line $y = ax + b$ which fits these data.

Problem 10.2: a) Find the least square solution x^* of the system $Ax = b$ with $A = \begin{bmatrix} 2 & -1 \\ 1 & 2 \\ 2 & -4 \end{bmatrix}$, and $b = \begin{bmatrix} 125 \\ 125 \\ 125 \end{bmatrix}$.
b) What is the matrix P which projects on the image of A?

Problem 10.3: A curve of the form

$$y^2 = x^3 + ax + b$$

is called an **elliptic curve** in Weierstrass form. Elliptic curves are important in cryptography. Use data fitting to find the best parameters (a, b) for an elliptic curve given the following points: $(x_1, y_1) = (1, 2), (x_2, y_2) = (-1, 0), (x_3, y_3) = (2, 1), (x_4, y_4) = (0, 1)$.

Problem 10.4: Find the circle $a(x^2 + y^2) + b(x + y) = 1$ which best fits the data

x	y
0	1
-1	0
1	-1
1	1

In other words, find the least square solution for the system of equations for the unknowns a, b which aims to have all 4 data points (x_i, y_i) on the circle.

Problem 10.5: Let us look at some extreme cases.
a) Analyze the best linear fit $f(x) = ax + b$ for the three data points $(1, 1), (1, 2), (1, 3)$.
b) To find the best linear $f(x) = a + bx$ for the four data points $(1, 1), (1, 1), (2, 2), (4, 7)$ with $f(x) = a + bx$, we try to take a short cut and pick the points $(1, 1), (2, 2), (4, 7)$ instead, as one data point obviously was redundant. Do we get the same solution?

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 11: Determinants

Lecture

11.1. We have already seen the determinants of 2×2 and 3×3 matrices:

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc, \quad \det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = aei + bfg + dhc - gec - hfa - dbi \ .$$

Our goal is to define the determinant for arbitrary matrices and understand the properties of the **determinant functional** det from $M(n,n)$ to $\mathbb{R}$.

11.2. A **permutation** of a set is an invertible map π on this set. It defines a rearrangement of the set. The point x goes to $\pi(x)$. Inductively, one can see that there are $n! = n \cdot (n-1) \cdots 1$ **permutations** of the set $\{1, 2, \ldots, n\}$: fixing the position of first element leaves $(n-1)!$ possibilities to permute the rest. For example, there are $6 = 3 \cdot 2 \cdot 1$ permutations of $\{1, 2, 3\}$. They are $(1,2,3), (1,3,2), (2,1,3), (2,3,1), (3,1,2), (3,2,1)$. A permutation can be visualized in the form of a **permutation matrix** A. It is a Boolean matrix which has zeros everywhere except at the positions $A_{k\pi(k)}$, where it is 1. An **up-crossing** is a pair $k < l$ such that $\pi(k) < \pi(l)$. When drawing out a permutation matrix, we also call it a **pattern**. The **sign** of a permutation π is defined as $\text{sign}(\pi) = (-1)^u$, where u is the **number of up-crossings** in the pattern of π.

11.3. The **determinant** of a $n \times n$ matrix A is defined by Leibniz as the sum

$$\sum_{\pi} \text{sign}(\pi) A_{1\pi(1)} A_{2\pi(2)} \cdots A_{n\pi(n)} \ ,$$

where π is a permutation of $\{1, 2, \ldots, n\}$. We see that for $n = 2$, we get two possible permutations, the identity permutation $\pi = (1,2)$ and the flip $\pi = (2,1)$. The determinant of a 2×2 matrix therefore is a sum of two numbers, the product of the diagonal entries minus the product of the side diagonal entries. For $n = 3$, we have 6 permutations and get the **Sarrus formula** stated initially above.

11.4. To organize the summation, one can first choose all the permutations for which $\pi(1) = 1$, then look at all permutations for which $\pi(1) = 2$ etc. This produces the **Laplace expansion**. Let $M(i,j)$ denote the matrix in which the i'th row and j'th column are deleted. Its determinant is called a **minor** of A. For every $1 \leq i \leq n$:

Theorem: $\det(A) = \sum_{j=1}^{n} (-1)^{i+j} A_{ij} \det(M(i,j))$

11.5. This expansion allows to compute the determinant a $n \times n$ matrix by reducing it to a sum of determinants of $(n-1) \times (n-1)$ matrices. It is still not suited to compute the determinant of a 20×20 matrix for example as we would need to sum up $20! = 2432902008176640000$ elements.

11.6. The fastest way to compute determinants for general matrices is by doing a **row reduction**. To understand this, we need the following properties:

Subtracting a row from another row does not change the determinant.
Swapping two rows changes the sign of the determinant.
Scaling a single row by a factor λ multiplies the determinant by λ.

11.7. Let s be the number of swaps and $\lambda_1, \ldots, \lambda_k$ the scaling factors which appear when bringing A into row reduced echelon form.

Theorem: $\det(A) = (-1)^s \lambda_1 \cdots \lambda_k \det(\mathrm{rref}(A))$

11.8. We see from this that the determinant "determines" whether a matrix is invertible or not:

Theorem: $\det(A)$ is non-zero if and only if A is invertible.

Here are more properties for $n \times n$ matrices which we prove in class:

$\det(AB) = \det(A)\det(B)$
$\det(A^{-1}) = \det(A)^{-1}$
$\det(SAS^{-1}) = \det(A)$
$\det(A^T) = \det(A)$
$\det(\lambda A) = \lambda^n \det(A)$
$\det(-A) = (-1)^n \det(A)$

11.9. An important thing to keep in mind is that the determinant of a **triangular** matrix is the product of its diagonal elements.

Example: $\det(\begin{bmatrix} 1 & 0 & 0 & 0 \\ 4 & 5 & 0 & 0 \\ 2 & 3 & 4 & 0 \\ 1 & 1 & 2 & 1 \end{bmatrix}) = 20.$

11.10. Another useful fact is that the determinant of a **partitioned matrix** $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ is the product $\det(A)\det(B)$. Example: $\det(\begin{bmatrix} 3 & 4 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 4 & -2 \\ 0 & 0 & 2 & 2 \end{bmatrix}) = 2 \cdot 12 = 24.$

Examples

11.11. The determinant of a rotation matrix is either $+1$ or -1: Proof: we know $A^T A = 1$. So, $1 = \det(1) = \det(A^T A) = \det(A^T)\det(A) = \det(A)\det(A) = \det(A)^2$ which forces $\det(A)$ to be either 1 or -1. For a rotation in $\mathbb{R}^2$ the determinant is 1 for a reflection, it is -1. In general, for any rotation the determinant is 1 as we can change the angle of rotation continuously to 0 forcing the determinant to be 1. The determinant depends continuously on the matrix. It can not jump from -1 to 1. Check the proof seminar in Unit 6.

11.12. Find the determinant of the partitioned matrix

$$A = \begin{bmatrix} 3 & 3 & 7 & 3 & 7 & 1 \\ 3 & 5 & 3 & 4 & 1 & 1 \\ 0 & 0 & 4 & 3 & 1 & 1 \\ 0 & 0 & 2 & 2 & 0 & 3 \\ 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}.$$

The determinant is $6 * 2 * 3 = 36$.

11.13. Use row reduction to compute the determinant of the following matrix:

$$A = \begin{bmatrix} 1 & 1 & 1 & 5 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 2 & 0 & 1 \\ 2 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

The answer is 8.

11.14. In this example, Laplace expansion is nice. Also row reduction works.

$$A = \begin{bmatrix} 0 & 0 & 0 & 5 & 8 & 0 \\ 3 & 1 & 3 & 4 & 0 & 0 \\ 0 & 5 & 1 & 3 & 2 & 7 \\ 0 & 0 & 7 & 1 & 3 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 9 & 0 \end{bmatrix}.$$

Homework

This homework is due on Thursday, 2/28/2019.

Problem 11.1: Find the determinants of A, B, C: $A = \begin{bmatrix} a^2 & ab \\ ba & b^2 \end{bmatrix}$,

$$B = \begin{bmatrix} 0 & 5 & 7 & 3 & 7 & 1 \\ 6 & 0 & 0 & 0 & 0 & 1 \\ 0 & 4 & 0 & 3 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 6 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \end{bmatrix}, C = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 3 \\ 3 & 3 & 0 & 0 & 6 & 0 \\ 4 & 2 & 0 & 4 & 0 & 0 \\ 5 & 3 & 2 & 0 & 0 & 0 \\ 6 & 3 & 0 & 0 & 4 & 0 \\ 7 & 0 & 5 & 0 & 0 & 0 \end{bmatrix}$$

Problem 11.2: Is the following determinant positive, zero or negative? (no technology!)

$$\begin{bmatrix} 22 & 100^9 & 7 & -6 & 3 & 1 \\ 100^9 & 22 & 2 & 2 & 2 & 2 \\ 6 & 4 & 22 & 1 & 100^9 & -1 \\ 2 & 2 & 100^9 & 22 & -5 & 9 \\ 9 & 1 & -1 & 100^9 & 22 & 2 \\ 7 & 4 & -1 & 2 & 4 & 100^9 \end{bmatrix}.$$

Problem 11.3: a) Use the Leibniz definition of determinants to show that the **partitioned matrix** satisfies $\det \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} = \det(A)\det(B)$.

b) Assume now that A, B are $n \times n$ matrices. Can you find a formula for $\det \begin{bmatrix} 0 & A \\ B & 0 \end{bmatrix}$? (It will depend on n.)

c) Show that number of up-crossings of a pattern is the same if the pattern is transposed and that therefore $\det(A^T) = \det(A)$.

Problem 11.4: Find the determinant of the matrix $A_{ij} = 2^{ij}$ for $i, j \leq 4$. It is $\begin{bmatrix} 2 & 4 & 8 & 16 \\ 4 & 16 & 64 & 256 \\ 8 & 64 & 512 & 4096 \\ 16 & 256 & 4096 & 65536 \end{bmatrix}$. First scale some rows to make the computation more manageable.

Problem 11.5: Find a formula for the determinant of the $n \times n$ matrix $L(n)$ which has 2 in the diagonal and 1 in the side diagonals and 0 everywhere else. Compute first $L(2), L(3), L(4)$, then $L(5) = \begin{bmatrix} 2 & 1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix}$.

Now, you see a pattern. Prove it by induction.

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 12: Complexity

Seminar

12.1. We are heading to the first midterm. In a timed exam, **speed** is always an issue. As an exam writer it is therefore important to know whether a particular problem can be solved in the allocated time. Now, when solving a problem, one could be lucky and get a particular solution fast, by pure luck. As an engineer, it is important however to know how quickly the task can be finished in general. Estimating the speed is a rigorous game and usually asks for finding out how long the worst case takes. To do so we are rather rough and do not care whether an algorithm takes $100n$ steps or n steps. We care however whether the algorithm takes n steps (linear complexity) or n^2 (quadratic complexity) or e^n (exponential complexity).

Figure 1. Cheetah at Ree Park, Denmark, 24 April 2010, Malene Thyssen

12.2. Let us look at the **Landau big O** notation. Given two functions $f(n)$ and $g(n)$, we say $f(n) = O(g(n))$ as $n \to \infty$ if there exists a constant C such that $|f(n)| \leq C|g(n)|$ if n is large enough. If we say $f = O(g)$, we mean the limit $n \to \infty$. If $f = O(x^n)$ for some n, we say f has **polynomial growth** and if $f = O(e^x)$ but not better, we say f has **exponential growth**. A growth rate like $O(e^{\sqrt{x}})$ is an example of what one calls **sub-exponential growth**. We can also have **super exponential** growth, like $O(e^{x^2})$ or $O(e^{e^x})$. Sometimes one also writes $e^{O(x)}$ in order to characterize exponential growth. Now $e^{O(5x)} = e^{O(x)}$ but it is **not true** that $O(e^{5x})$ is equal to $O(e^x)$. If you see a statement using the big O notation, when in doubt, rewrite it using the definition.

12.3. For example $5x^2 + \sin(x) = O(x^2)$ or $7e^{5x} - 1000e^{3x} + x^{22} = O(e^{5x})$ or $5 + 1/(1 + x^2) + \cos(x^7) = O(1)$.

12.4. In **complexity theory**, the functions under consideration are usually a function of an integer n and $f(n)$ gives the number of computation steps which are needed to do the task for an object of size n. Let us first make sure we understand the big O notation:

> **Problem A:** Determine from each of the following functions whether they show polynomial, exponential, sub-exponential (and super-polynomial), or super-exponential growth.
> a) $f(n) = n^n$
> b) $f(n) = e^{5\log(n)}$
> c) $f(n) = \log(e^{n^7}) + \log(n)$.
> d) $f(n) = n^3 \sin(e^n)$
> e) $f(n) = e^{\sin(n)+n^7}$
> f) $f(n) = 1/((1/n) + (1/\log(n))$
> g) $f(n) = \log(\log(n^{10n}))$

12.5. How fast can we compute determinants of $n \times n$ matrices? Following the definition summing over all possible permutations is too costly as it requires to compute $n!$ products of n numbers. This appears to cost $O(n!n)$.

> **Problem B:** Show that the Laplace expansion reduces the cost to $O(n!)$ multiplications. Is $O(n!)$ polynomial, sub-exponential, exponential or super-exponential?

12.6. In reality we can compute determinants much faster. Try out the following:

```
A=Table[Random[],{10000},{10000}]; Timing[Det[A]]
A=Table[Random[Integer,1],{1000},{1000}]; Timing[Det[A]]
```

Indeed, determinants can be computed in polynomial time. Let us explain:

> **Problem C:** Why can we compute the determinant of a $n \times n$ matrix as fast as row reducing a $n \times n$ matrix? Can you estimate in the big O notation how fast determinants can be computed?

12.7. Related to determinants are **permanents**. The permanent of a matrix A is defined as

$$\mathrm{per}(A) = \sum_{\pi} A_{1\pi(1)} A_{2\pi(2)} \cdots A_{n\pi(n)} \ .$$

We see that they are defined like the by Leibniz defined determinants but the $\mathrm{sign}(\sigma)$ is missing. It looks like a small matter to leave away the -1 factors, but it changes the completely of the problem. The computation of permanents is believed to be **very hard**. An algorithm which does it in polynomial time is not unknown. If such an algorithm would exist, it would be a huge surprise or even a sensation.

12.8. Question: is there a fast way to determine whether a given orthogonal matrix has determinant 1 or -1? If you should find a fast way to do that let us know. It must be faster than just row reducing!

12.9.

Problem D: Time for Speed training! How fast can you compute the following determinants:

$$A = \begin{bmatrix} 0 & 0 & 2 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 5 \\ 7 & 0 & 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 4 & 5 & 0 & 0 \\ 1 & 9 & 2 & 0 \\ 2 & 2 & 6 & 2 \end{bmatrix}, C = \begin{bmatrix} 7 & 2 & 5 & 1 \\ 1 & 1 & 2 & 1 \\ 0 & 0 & 3 & 2 \\ 0 & 0 & 2 & 2 \end{bmatrix}, D = \begin{bmatrix} 0 & 0 & 2 & 1 \\ 0 & 0 & 3 & 4 \\ 2 & 2 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$

12.10.

Problem E: Time for Speed training! How fast can you find the QR decomposition of the following matrices

$$A = \begin{bmatrix} 0 & 0 & 2 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 5 \\ 7 & 0 & 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 3 & 4 & 2 & 1 \\ 0 & 5 & 3 & 1 \\ 0 & 0 & 5 & 5 \\ 0 & 0 & 0 & 2 \end{bmatrix}, C = \begin{bmatrix} 3 & -2 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 4 & -2 \end{bmatrix}, D = \begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 0 & 1 & 4 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & 6 \end{bmatrix}.$$

12.11.

Problem F: Time for Speed training! How fast can you find a basis for the image of the following matrices?

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{bmatrix}, B = \begin{bmatrix} 2 & 4 & 6 & 8 \\ 4 & 8 & 12 & 16 \\ 8 & 16 & 24 & 32 \\ 16 & 32 & 48 & 64 \end{bmatrix}, C = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}, D = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 8 & 10 \\ 4 & 8 & 10 & 12 \end{bmatrix}.$$

12.12.

Problem G: Time for Speed training! How fast can you find a basis for the kernel of the following matrices?

$$A = \begin{bmatrix} 1 & 2 & 2 & 1 \\ 1 & 2 & 2 & 1 \\ 1 & 2 & 2 & 1 \\ 1 & 2 & 2 & 1 \end{bmatrix}, B = \begin{bmatrix} 2 & 4 & 2 & 4 \\ 4 & 8 & 4 & 8 \\ 8 & 16 & 8 & 16 \\ 16 & 32 & 16 & 32 \end{bmatrix}, C = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}, D = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 8 & 10 \\ 4 & 8 & 10 & 12 \end{bmatrix}.$$

Homework

Exercises A)-G) are done in the seminar. This homework is due on Tuesday:

Problem 12.1 There is a nice formula

$$A_{ij}^{-1} = (-1)^{i+j}\det(M(j,i))/\det(\mathrm{A}) \ ,$$

where $M(i,j)$ is the matrix in which the i'th row and jth column is deleted. The value $\det(M(i,j))$ is called the (i,j) **minor** of A.
a) First of all, this formula immediately follows from the Laplace expansion. Look at the formula at the bottom of the first page of lecture 11. Can you see it?
b) How fast can we compute the inverse of a $n \times n$ matrix using this formula?
c) Remember how we have computed the inverse of a matrix using row reduction. Does the determinant formula produce an advantage in terms of complexity?

Problem 12.2 How fast can we do the QR decomposition of a matrix? Is it polynomial or exponential?

Problem 12.3 In cryptology, it is important to be able to factor an integer x with n digits fast. Any significant progress there is monitored closely as somebody with a fast algorithm could bring down modern financial and communication systems.
a) How fast is the "Baby algorithm" which just tries out all factors up to $\sqrt{x}$. You have to give the answer in terms of $O(f(n))$ of the number n of digits of x.
b) You will find notions like $O(f(b))$, where b is the number of bits to represent an integer. Why can you replace the number of bits with the number of digits of a number?
c) Look up on the web how fast the fastest known integer factorization algorithm is.

Problem 12.4 The most important problem in computer science is the $P - NP$ problem. Describe this problem in one short paragraph.

Problem 12.5 What is "NP complete". List 3 examples of NP complete problems.

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 13: Checklist for First Hourly

Definitions

- **Linear space X**: $0 \in X$, $x + y \in X$, $\lambda x \in X$
- **Linear transformation**: $T : x \mapsto Ax$, $T(x + y) = T(x) + T(y), T(\lambda x) = \lambda T(x)$.
- **Column vectors of** A: images of standard basis vectors $e_1, \dots, e_n$.
- **Coefficient matrix of the system** $Ax = b$: is the matrix A.
- **Matrix multiplication**: $[AB]_{ij} = \sum_k A_{ik} B_{kj}$
- **Row reduction steps**: Swapping rows, Scaling row, Subtracting row.
- **Row reduced**: nonzero rows have **leading** 1, columns with leading 1 are clean, every row above row with leading 1 has leading 1 to left.
- **Pivot column**: column with leading 1 in rref(A).
- **Free variable**: a variable for which we have no leading 1 in rref(A).
- **Redundant column**: column with no leading 1 in rref(A).
- **Rank of matrix** A: number of leading 1 in rref(A). Is equal dim(im(A)).
- **Nullity of matrix** A: the number of free variables. It is dim(ker(A)).
- **Kernel of matrix**: $\{x \in \mathbf{R}^m, Ax = 0\}$.
- **Image of matrix**: $\{Ax, x \in \mathbf{R}^m\}$.
- **Inverse matrix of** A: matrix $B = A^{-1}$ satisfies $AB = BA = 1$
- **Orthogonality**: dot product $v \cdot w = v^T w$ is zero. (In stats: uncorrelated if centered)
- **Length of a vector**: $\sqrt{v \cdot v}$ (=standard deviation if centered)
- **Angle between vectors**: $\cos(\alpha) = v \cdot w/(|v||w|)$ is in $[0, \pi]$ (in stats: correlation)
- **Gram-Schmidt Orthogonalization** (QR decomposition)
- **Computing Determinants** (patterns,Laplace,row reduce,partition)
- **Least square**: Produce the least square solution of a linear equation
- **Solve data fitting problem** (linear system, least square problem)
- **Algebra of matrices** (multiplication, inverse, transpose)
- $\mathcal{B} = \{v_1, \dots, v_n\}$ **spans** V: every $x \in V$ can be written as $x = a_1 v_1 + \cdots + a_n v_n$.
- $\mathcal{B}$ **linear independent** V: $a_1 v_1 + \cdots a_n v_n = 0$ implies $a_1 = \cdots = a_n = 0$.
- $\mathcal{B}$ **basis in** V: linear independent in V and span V.
- **Dimension of linear space** V: number of basis elements of a basis in V.
- S**-matrix** : The matrix which contains the basis vectors as columns.
- $\mathcal{B}$ **coordinates**: $c = S^{-1} v$, where $S = [v_1, \dots, v_n]$ contains $\mathcal{B}$ as columns.
- $\mathcal{B}$ **matrix**: of T in basis $\mathcal{B}$. The matrix is $B = S^{-1} A S$.
- A **similar to** B: defined as $B = S^{-1} A S$. Write $A \sim B$.
- **Orthogonal vectors** $v \cdot w = 0$

- **Length** $||v|| = \sqrt{v} \cdot v$, **Unit vector** v with $||v|| = \sqrt{v \cdot v} = 1$
- **Orthogonal set** $v_1, \ldots, v_n$: pairwise orthogonal
- **Orthonormal set** orthogonal and length 1
- **Orthonormal basis** A basis which is orthonormal
- **Orthogonal to V** v is orthogonal to V if $v \cdot x = 0$ for all $x \in V$
- **Orthogonal complement** $V^\perp = \{v \in \mathbf{R}^n \mid v \text{ perpendicular to } V\}$
- **Complex numbers** $z = a + ib = re^{i\theta} = r\cos(\theta) + ir\sin(\theta)$, $\overline{z} = a - ib$, $|z|^2 = a^2 + b^2$.
- **Quaternions** $z = a + ib + jc + kd$, $\overline{z} = a - ib - jc - kd$. $N(z) = |z|^2 = a^2 + b^2 + c^2 + d^2$.

Theorems:

- **Matrix space** $M(n, m)$ is a linear space, $M(n, n)$ is an algebra.
- **Kernel and Image** $Ker(A)$ and $Im(A)$ are linear spaces.
- **Column picture** The k'th column of A is equal to Ae_k.
- **Invertibility** If T is linear invertible, then T^{-1} is linear.
- **linear independence** $\mathcal{B}$ linear independent $\Leftrightarrow$ S has trivial kernel
- **kernel and image** $\ker(A^T) = \mathrm{im}(A)^\perp$
- **Existence of dimension** Number of basis elements is the same for every basis of V.
- **kernel of** $A^T A$ $\ker(A^T A) = \ker(A)$
- **Unique QR** QR decomposition is unique
- **Rank-nullity theorem**: $\dim(\ker(A)) + \dim(\mathrm{im}(A)) = n$, where A is $m \times n$ matrix.
- **Angles and Length**: Linear trafo T is orthogonal $\Leftrightarrow$ T preserves length/angles.
- **Laplace expansion**: $\det(A) = \sum_{j=1}^{n} (-1)^{i+j} A_{ij} \det(M(i, j))$.

Properties:

- **Cute projection onto V** $P = QQ^T$ if Q has orthonormal columns
- **Gram-Schmidt** $w_i = v_i - \mathrm{proj}_{V_{i-1}} v_i$, $u_i = w_i / ||w_i||$
- **QR-factorization** $Q = [u_1 \cdots u_n]$, $R_{ii} = |w_i|$, $R_{ij} = u_i \cdot v_j, j > i$
- **Transpose matrix** $A^T_{ij} = A_{ji}$. Transposition switches rows and columns.
- **Symmetric matrix** $A^T = A$ and **skew-symmetric** $A^T = -A$
- **Orthogonal matrix** $A^T A = 1$ (examples are rotations or reflections)
- **Orthogonal projection** onto V is $P = A(A^T A)^{-1} A^T$
- **Least square solution** of $Ax = b$ is $x^* = (A^T A)^{-1} A^T b$
- **Determinant** $\det(A) = \sum_\pi \mathrm{sign}(\pi) A_{1\pi(1)} A_{2\pi(2)} \cdots A_{n\pi(n)}$
- **Trace** is $\mathrm{tr}(A) = \sum_i A_{ii}$, sum of diagonal elements.

- **Orthogonal projection** $P = A(A^T A)^{-1} A^T$ onto $\mathrm{im}(A)$
- **Orthogonal projections** Only 1 is both projection and orthogonal matrix
- **Kernel of** A **and** $A^T A$ are the same: $\ker(A) = \ker(A^T A)$
- **Determinant** $\det(\begin{bmatrix} a & b \\ c & d \end{bmatrix}) = ad - bc$.
- **Det** $\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + bfg + cdh - ceg - fha - bdi$. Sarrus rule.

- **Pattern:** Determinant is sum of patterns, sign depends on number of up-crossings.
- **Determinant of triangular matrix** product of diagonal entries
- **Det of partitioned matrix** $\det(\begin{bmatrix} A & C \\ 0 & B \end{bmatrix}) = \det(A)\det(B)$.
- **Product properties** $\det(AB) = \det(A)\det(B)$, $\det(A^{-1}) = 1/\det(A)$
- **Determinants** $\det(SAS^{-1}) = \det(A)$, $\det(A^T) = \det(A)$.
- **Adding rows or columns** $\det([AvB]) + \det([AwB]) = \det([A(v+w)B])$
- **Scaling rows or columns** $\det([A(\lambda v)B]) = \lambda\det([AvB])$
- **Swapping of two rows** $\det(B) = -\det(A)$
- **Adding row to given row** $\det(B) = \det(A)$
- **rref** $\det(A) = (-1)^s(\lambda_1\lambda_2\cdots\lambda_k)\det(\mathrm{rref}(A))$, λ_i scales and s swaps
- **Properties of transpose** $(A^T)^T = A$, $(AB)^T = B^TA^T$, $A^T + B^T = (A+B)^T$
- **Properties of transpose** $(A^{-1})^T = (A^T)^{-1}$, $\det(A^T) = \det(A)$
- **Orthogonal Matrices** A have $\det(A) = \pm 1$
- **Rotations** satisfy $\det(A) = 1$ in all dimensions
- **QR Decomposition** $A = QR$ orthogonal A, upper triangular R
- **QR Decomposition** $|\det(A)| = R_{11}R_{22}\cdots R_{nn}$ if $A = QR$

Algorithms:

- **Determinants** Laplace, row reduce, partition triangular, ..., patterns you can use.
- **Gram-Schmidt** Perform Gram Schmidt
- **Check linear space**: both for matrix space or function space
- **Check linear transformation**: is a given transformation linear or not?
- **Kernel - Image** : find a basis for the kernel and basis for image by row reduction.
- **Row reduction**: scale rows, swap rows, subtract row from other row.
- **Matrix algebra**: multiply, invert, manipulate matrices. Solve simple matrix equations
- **Find space orthogonal to** V: write basis as rows of a matrix and find kernel.
- **Find transformation**: use good basis, form S, B and then find $A = SBS^{-1}$.
- **Similarity**: If A is similar to B, then $B^n = S^{-1}A^nS^n$. If A is invertible, so is B.
- **Find angle**: $\cos(\alpha) = (v \cdot w)/(|v||w|)$. Is zero for orthogonal vectors.
- **Number of basis elements**: is independent of basis. Allows to define dimension.
- **Basis of image of** A: pivot columns of A form a basis of the image of A.
- **Basis of kernel of** A: introduce free variables for each redundant column of A.
- **Inverse of** 2×2 **matrix**: switch diagonal, sign the wings and divide by determinant.
- **Inverse of** $n \times n$ **matrix**: Row reduce $[A|1]$ to get $[1|A^{-1}]$.
- **Matrix algebra**: $(AB)^{-1} = B^{-1}A^{-1}$, $A(B+C) = AB + AC$, etc. no commutativity
- **A is invertible**: $\Leftrightarrow \mathrm{rref}(A) = 1 \Leftrightarrow$ columns are basis $\Leftrightarrow \mathrm{rank}(A) = n, \Leftrightarrow \mathrm{nullity}(A) = 0$.

Transformations:

- **Identity**: $x \mapsto x$, $1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
- **Rotation in plane**: $x \mapsto Ax$, $A = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix}$, counterclockwise.

- **Dilation in plane**: $x \mapsto \lambda x$, also called scaling. Given by diagonal $A = \lambda 1$
- **Rotation-Dilation**: $x \mapsto Ax$, $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$. Scale by $\sqrt{a^2+b^2}$, rotate by α
- **Reflection-Dilation**: $x \mapsto Ax$, $A = \begin{bmatrix} a & b \\ b & -a \end{bmatrix}$. Scaling factor $\sqrt{a^2+b^2}$.
- **Horizontal and vertical shear**: $x \mapsto Ax$, $A = \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}$, $x \mapsto Ax$, $A = \begin{bmatrix} 1 & 0 \\ b & 1 \end{bmatrix}$.
- **Reflection about line**: $x \mapsto Ax$, $A = \begin{bmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{bmatrix}$.
- **Projection**: $x \mapsto Ax$, $A = \begin{bmatrix} a^2 & ab \\ ba & b^2 \end{bmatrix}$.
- **Magic bullet**: $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ satisfies $A^2 = 0$. $\mathrm{Ker}(A) = \mathrm{Im}(A)$.

Proofs:

- **Axioms**: set of rules which are assumed and not proven.
- **Monoid**: an algebraic structure with associativity and neutral element.
- **Group**: monoid where each element has an inverse.
- **Examples**: an example can be a counter example or a prototype.
- $SL(n)$: special linear group (matrices with determinant 1).
- $O(n)$: orthogonal group (determinant 1 or -1). Defined by $A^T A = 1$
- $SO(n)$: special orthogonal group (determinant 1).
- $U(n)$: unitary group. Defined by $\overline{A}^T A = 1$
- $SU(n)$: special unitary group (determinant 1).
- $SU(2)$: unitary group $A = \begin{bmatrix} z & -\overline{w} \\ w & \overline{z} \end{bmatrix}$.
- **Inverse**: $A^{-1}_{ij} = (-1)^{i+j}\det(M(j,i))/\det(\mathrm{A})$.
- **Landau O**: $f(x) = O(g(x))$ if $|f(x)| \leq C|g(x)|$ for some constant C and all x large.

People:

- **Feynman** (examples)
- **Peano** (Peano axioms)
- **Euclid** (Axioms of geometry)
- **Gram** (QR)
- **Schmidt** (QR)
- **Gauss** (row reduction)
- **Jordan** (row reduction)
- **Hamilton** (quaternions)
- **Lagrange** (four square theorem)
- **Glashow, Salam, Weinberg** (electoweak unification)
- **Landau** (complexity)

1	
2	
3	
4	
5	
6	
7	
8	
9	
10	

Name:

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Total:

Unit 13: First Hourly Practice

Welcome to the first hourly. It will take place on February 26, 2019 at 9:00 AM sharp in Hall D. You can already fill out your name in the box above.

- You only need this booklet and something to write. Please stow away any other material and electronic devices. Remember the honor code.
- Please write neatly and give details. We want to see details, even if the answer should be obvious to you.
- Try to answer the question on the same page. There is also space on the back of each page.
- If you finish a problem somewhere else, please indicate on the problem page so that we find it.
- You have 75 minutes for this hourly.

Problems

Problem 13P.1 (10 points):
We prove $\text{im}(A) = \text{im}(AA^T)$ in two steps. Do them both.
a) Prove that $\text{im}(A)$ contains $\text{im}(AA^T)$.
b) Prove that $\text{im}(AA^T)$ contains $\text{im}(A)$.

Problem 13P.2 (10 points):
Decide in each case whether the set X is a linear space. If it is, prove it, if it is not, then give a reason why it is not.
a) The space of 4×4 matrices with zero trace.
b) The space of 4×4 quaternion matrices.
c) The space of 2×2 matrices λQ, where Q is an orthogonal matrix and λ is real.
d) The space of 3×3 matrices λQ, where Q is an orthogonal matrix and λ is real.
e) The space of 5×5 matrices with entries 0 or 1.
f) The image $\text{im}(1)$ of the identity 3×3 matrix 1.
g) The kernel of the matrix $A = [1, 2, 3, 4]$.
h) The set of all vectors in $\mathbb{R}^3$ for which $x^2 + y^2 + z^2 = 0$.
i) The set of all vectors in $\mathbb{R}^3$ for which $xyz = 0$.
j) The set of 3×3 matrices with orthogonal columns.

Problem 13P.3 (10 points):
a) (2 points) Is the transformation $T(x, y) = (x^2, y^2)$ from $\mathbb{R}^2$ to $\mathbb{R}^2$ linear?
b) (2 points) Is the map $T(A) = \mathrm{tr}(A)1$ as a map from $M(3, 3)$ to $M(3, 3)$ linear? Here 1 is the identity matrix.
c) (4 points) Match the transformation type:
A: **rotation dilation**, B: **reflection dilation**, C: **shear dilation**, D: **projection dilation**

Fill A-D				
Matrix	$\begin{bmatrix} 4 & 4 \\ 4 & 4 \end{bmatrix}$	$\begin{bmatrix} 4 & -4 \\ 4 & 4 \end{bmatrix}$	$\begin{bmatrix} 4 & 4 \\ 4 & -4 \end{bmatrix}$	$\begin{bmatrix} 4 & 4 \\ 0 & 4 \end{bmatrix}$

d) (4 points) SO, SU or NO (no SO nor SU)? Yo, you have to decide so!

SO,SU,NO				
Matrix	$\begin{bmatrix} i & -i \\ i & i \end{bmatrix}$	$\begin{bmatrix} i & 0 \\ 0 & 0 \end{bmatrix}$	$\begin{bmatrix} i & i \\ i & i \end{bmatrix}$	$\begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$

Problem 13P.4 (10 points, each sub problem is 2 points):
a) Which physicist promoted the use of examples to understand a theory?
b) We describe a general linear map T from $M(2, 2)$ to $M(2, 2)$. In a suitable basis, is this map given by a 2×2 matrix or a 4×4 matrix?
c) Is it true that $f(x) = e^{3x}$ satisfies $f(x) = O(e^x)$?
d) Give an example of a 4×4 matrix A for which $\ker(A) = \mathrm{im}(A)$.
e) If $B = S^{-1}AS$, what can you say about the determinants of A and B?

Problem 13P.5 (10 points):
In the **checkers matrix**, the entry 1 means that the checkers initial condition has a checker piece there and 0 means that that field is empty:

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}.$$

a) (6 points) Find a basis for the kernel of A.
b) (4 points) Find a basis for the image of A.

Problem 13P.6 (10 points):

a) (5 points) The projection-dilation matrix $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ in the basis

$$\mathcal{B} = \{v_1, v_2, v_3\} = \{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \}$$

is given by a matrix B. Find this 3×3 matrix B.

b) (5 points) A linear transformation T satisfies

$$T(v_1) = v_2, T(v_2) = v_3, T(v_3) = v_1$$

where v_1, v_2, v_3 are given in a). Find the matrix R implementing this transformation in the standard basis.

Problem 13P.7 (10 points):

Find the QR decomposition of the following matrices

a)(2 points) $A = \begin{bmatrix} 1 & -2 & 0 \\ 2 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$.

b) (2 points) $B = \begin{bmatrix} 2 & -1 \\ 1 & 2 \\ 1 & 1 \\ 3 & 1 \\ 1 & -3 \end{bmatrix}$.

c) (2 points)$C = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$.

d) (2 points)$D = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$.

e) (2 points) $E = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

Problem 13P.8 (10 points):

a) (2 points) Find the determinant of the "prime" matrix

$$A = \begin{bmatrix} 0 & 2 & 0 & 3 \\ 13 & 0 & 0 & 11 \\ 0 & 7 & 0 & 0 \\ 0 & 0 & 5 & 0 \end{bmatrix}.$$

b) (2 points) Find the determinant of the "count to 12" matrix

$$B = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 0 & 0 & 9 & 10 \\ 0 & 0 & 11 & 12 \end{bmatrix}.$$

c) (2 points) Find the determinant of the "2-1" matrix

$$C = \begin{bmatrix} 2 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 1 & 2 \end{bmatrix}.$$

d) (2 points) Find the determinant of the "Pascal triangle" matrix

$$D = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 1 & 3 & 3 & 1 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}.$$

e) (2 points) Find the determinant of the "mystery" matrix:

$$E = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 3 & 3 & 3 & 3 \\ 1 & 1 & 4 & 4 & 4 \\ 1 & 1 & 1 & 5 & 5 \\ 1 & 1 & 1 & 1 & 6 \end{bmatrix}.$$

Problem 13P.9 (10 points):
Find the function

$$a|x| - b|x-1| = y$$

which is the best fit for the data

x	y
1	2
3	2
-2	1
2	0

Problem 13P.10 (10 points):
The matrices

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}$$

$$C = \begin{bmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{bmatrix} \quad D = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

are called the **Gamma matrices** or **Dirac matrices**.
a) (4 points) Are the columns orthonormal? That is, is it true that $A^*A = 1$, where $A^* = \overline{A}^T$.

	A	B	C	D
Orthonormal columns?				

b) (3 points) Compute A^2 and D^2 and $AD + DA$.
c) (3 points) Write down the inverse of A,B and D.

1	
2	
3	
4	
5	
6	
7	
8	
9	
10	

Name:

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Total:

Unit 13: First Hourly

Welcome to the first hourly. Please don't get started yet. We start all together at 9:00 AM. You can already fill out your name in the box above. Then relax at the beautiful pond (a Povray scene using code of Gilles Tran who wrote this in 2004).

- You only need this booklet and something to write. Please stow away any other material and electronic devices. Remember the honor code.
- Please write neatly and give details. Except when stated otherwise, we want to see details, even if the answer should be obvious to you.
- Try to answer the question on the same page. There is also space on the back of each page and at the end.
- If you finish a problem somewhere else, please indicate on the problem page so that we find it. Make sure we find additional work.
- You have 75 minutes for this hourly.

Problems

Problem 13.1 (10 points):
Prove that $\ker(A) = \ker(A^T A)$ for any matrix A. We have seen this to be useful in the data fitting part.

Problem 13.2 (10 points) Each problem is 1 point:
Decide in each case whether the set X is a linear space. If it is, prove it, if it is not, then give a reason why it is not.

a) The space of upper triangular 4×4 matrices with zero trace.

b) The space of 2×2 rotation dilation matrices.

c) The space of 2×2 diagonal matrices.

d) The space of 4×4 matrices which have all zeros in the diagonal.

e) The space of all 3×3 matrices with non-trivial kernel.

f) The kernel of $A^T A$, where A is a 10×3 matrix.

g) The 3×3 matrices which row reduce to the 1 matrix.

h) The 3×3 matrices A which have the property that $e^A = 1$.

i) The 2×2 rotation dilation matrices zero trace.

j) The set of 4×4 quaternion matrices satisfying $A^2 = -1$.

Problem 13.3 (10 points):
Decide in each case, whether the transformation T is linear. If it is, prove it, if it is not, then give a reason why it is not.

a) (1 point) $T(A) = A \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$, from $M(2,2)$ to $M(2,2)$.

b) (1 point) $T(A) = A^2$ from $M(2,2)$ to $M(2,2)$.

c) (8 points) Match each of matrices with one of the geometric descriptions below. You don't have to give explanations in this part c)

Matrix	Enter A-H here.	Matrix	Enter A-H here.
$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$		$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$	
$\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$		$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	
$\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$		$\begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	
$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$		$\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$	

A) Shear along a plane.
B) Projection onto a plane.
C) Rotation around an axes.
D) Reflection about a point.
E) Projection onto a line.
F) Reflection about a plane.
G) Reflection about a line.
H) Identity transformation.

Problem 13.4 (10 points, each sub problem is 2 points):
a) Who was the inventor of quaternions?

b) The SU in $SU(2)$ stands for Super dUper. No seriously, what does SU abbreviate?

c) What can you say about the determinant of an orthogonal matrix?

d) What are the properties of a monoid?

e) Is it possible to have a 3×5 matrix with orthonormal columns? If yes, give one.

Problem 13.5 (10 points):
a) (6 points) Find a basis for the kernel of the matrix

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 1 & 0 \\ 2 & 2 & 2 & 2 & 2 \end{bmatrix}$$

b) (4 points) Find a basis for the image of A.

Problem 13.6 (10 points):

a) (5 points) What are the coordinates of the vector $\begin{bmatrix} 4 \\ 3 \\ 3 \end{bmatrix}$ in the basis

$$\mathcal{B} = \{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \} \ ?$$

b) (5 points) A transformation T is described in the standard basis by

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} .$$

What is the matrix B of the transformation T in the basis $\mathcal{B}$?

Problem 13.7 (10 points):

a) (2 points) Find the QR Decomposition of

$$A = \begin{bmatrix} 1 & -2 \\ 2 & 1 \\ 3 & -4 \\ 4 & 3 \end{bmatrix}$$

b) (2 points) Find the QR decomposition of the product

$$A = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 2 & 1 \\ 0 & 4 & 2 \\ 0 & 0 & 5 \end{bmatrix} .$$

c) (2 points) Find the QR decomposition of

$$A = \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix} .$$

d) (2 points) Find the QR decomposition of

$$A = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

e) (2 points) Find the QR decomposition of the 1×1 matrix

$$A = [5] .$$

Problem 13.8 (10 points):
Make sure to indicate which method you use to compute the determinant!

a) (2 points) Find the determinant $A = \begin{bmatrix} 1 & 2 & 4 & 8 & 2 \\ 0 & 0 & 3 & 4 & 3 \\ 7 & 0 & 0 & 0 & 4 \\ 9 & 0 & 0 & 7 & 3 \\ 6 & 0 & 0 & 0 & 0 \end{bmatrix}$

b) (2 points) Find the determinant of $A = \begin{bmatrix} 2 & 1 & 0 & 0 & 0 \\ 1 & 7 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

c) (2 points) Find the determinant $A = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 \\ 1 & 2 & 2 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 2 & 2 & 0 & 1 & 3 \\ 2 & 2 & 0 & 4 & 2 \end{bmatrix}$

d) (2 points) Find the determinant $A = \begin{bmatrix} 1 & 6 & 10 & 1 & 15 \\ 2 & 8 & 17 & 1 & 29 \\ 0 & 0 & 3 & 8 & 12 \\ 0 & 0 & 0 & 4 & 9 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix}$

e) (2 points) Find the determinant $A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 6 & 8 & 10 \\ 1 & 2 & 6 & 8 & 13 \\ 1 & 2 & 3 & 8 & 10 \\ 1 & 2 & 3 & 4 & 9 \end{bmatrix}$

Problem 13.9 (10 points):
Find the equation of the form

$$x^2 + axy + by^2 = 1$$

that best fits the data points:

x	y
2	1
-1	1
1	0
0	1

Problem 13.10 (10 points) every entry is 1 point:
Match the following matrices with the correct label. No justifications are needed. Fill in the attribute acronym into the middle boxes. One attribute might apply to more than one. But there is a unique match which works for all. To the right, indicate whether the transformation is invertible or not.

	Transformation matrix	attribute	invertible (yes or no)?
A=	$\begin{bmatrix} 2 & 3 & 4 \\ 3 & 5 & 7 \\ 4 & 7 & 9 \end{bmatrix}$		
B=	$\begin{bmatrix} 0 & 3 & 4 \\ -3 & 0 & 7 \\ -4 & -7 & 0 \end{bmatrix}$		
C=	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$		
D=	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$		
E=	$\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$		

Here are the attribute acronyms.

ANTI) anti-symmetric matrix

REFL) reflection matrix

PROJ) orthogonal projection

SYMM) symmetric matrix

ORTH) orthogonal matrix

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 14: Characteristic Polynomial

Lecture

14.1. We have seen the definition of determinants as well as methods to compute them. In this lecture, we will learn another method which is based on the **fundamental theorem of algebra**.

14.2. If A is a $n \times n$ matrix and v is a non-zero vector such that $Av = \lambda v$, then v is called an **eigenvector** of A and λ is called an **eigenvalue**. We see that v is an eigenvector if it is in the kernel of the matrix $A - \lambda 1$. We know that this matrix has a non-trivial kernel if and only if $p(\lambda) = \det(A - \lambda 1)$ is zero. By the definition of determinants the function $p(\lambda)$ is a polynomial of degree n.

14.3. In order to study the **characteristic polynomial**

$$p_A(\lambda) = \det(A - \lambda 1)$$

we first of all need to know the **fundamental theorem of algebra**:

Theorem: A polynomial $f(x)$ of degree n has exactly n roots in $\mathbb{C}$.

The roots are counted with multiplicity. $f(x) = x^2 + 2x + 1$ for example has two roots $-1, -1$ so that it factors as $f(x) = (x + 1)(x + 1)$. The roots can be complex like $f(x) = x^2 + 1$, which factors as $(x - i)(x + i)$. The theorem implies

$$p_A(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \cdots (\lambda_n - \lambda) \ .$$

It is enough to prove that one root λ_1 exists. By induction we can then factor the degree $n - 1$ polynomial $p(\lambda)/(\lambda - \lambda_1)$.

Proof. The Bolzano extremal value theorem assures that a continuous function f on a closed disk has a minimum and maximum. This implies that the function $|f(z)|$ has a global minimum in $\mathbb{C}$. (There exists r_0 such that the minimum of $|f(z)|$ with $|z| = r$ is larger than say $f(0)$ for $r \geq r_0$ so that the minimum has to be in the disk $\{|z| \leq r_0\}$.) We use the method of contradiction to show that the minimal value $|f(z_0)|$ is zero. Assume $f(z_0) \neq 0$ allows to introduce $g(z) = f(z_0 + z)/f(z_0)$ which is also a polynomial of degree n with minimum 1 at z_0. We have $g(z) = 1 + b_k z^k + \cdots + b_n z^n$ with $b_k = |b_k| e^{i\theta} \neq 0$ for some k. Define the line $z(t) = t|b_k|^{-1/k} e^{i(\pi - \theta)/k}$. Then $g(z(t)) = 1 - t^k + t^{k+1} h(t)$ with some continuous function $h(t)$. The value of $|g(z(t))| \leq 1 - t^k + t^{k+1}|h(t)|$ is less than 1 for small positive t contradicting that g had a minimal value 1 at z_0. $\square$

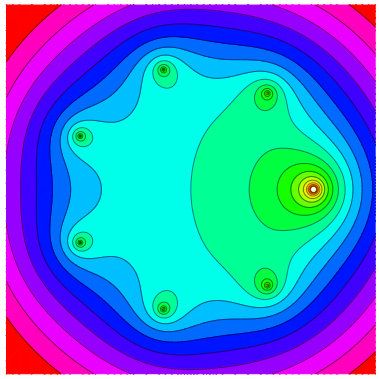


FIGURE 1. Contour map $(x+iy) \to |p(x+iy)|$ of the characteristic polynomial of the 8-queen matrix. Here, from the 8 roots, $\lambda = 1$ appear with multiplicity 2. (The algebraic multiplicity of 1 is 2.)

14.4. This implies

The determinant of A is the product of the eigenvalues of A.

The trace of A is the sum of the eigenvalues of A.

14.5. This gives us a new way to compute determinants. Find the eigenvalues and take the product of the eigenvalues. Example: to find the determinant of

$$A = \begin{bmatrix} 11 & 1 & 1 & 1 & 1 \\ 1 & 11 & 1 & 1 & 1 \\ 1 & 1 & 11 & 1 & 1 \\ 1 & 1 & 1 & 11 & 1 \\ 1 & 1 & 1 & 1 & 11 \end{bmatrix}$$

define $B = A - 10I$ which has a 4 dimensional kernel and so 4 eigenvalues 0. The last eigenvalue of B is 5 as the sum of the eigenvalues is $tr(B) = 5$. Since B has the eigenvalues $0, 0, 0, 0, 5$, the matrix A has the eigenvalues $10, 10, 10, 10, 15$. The determinant is 150000. We can even write down the characteristic polynomial

$$p_A(\lambda) = (\lambda - 10)^4(\lambda - 15) \ .$$

14.6. We are interested in the coefficients of the characteristic polynomial.

The polynomial starts with $(-\lambda)^n$ so that $a_n = (-1)^n$.

The coefficient $(-1)^{n-1}a_{n-1}$ is the trace of A.

The coefficient a_0 is the determinant of A.

Examples

14.7. Problem: Find $p_Q(\lambda$ for the **magic square**: $A = \begin{bmatrix} 4 & 9 & 2 \\ 3 & 5 & 7 \\ 8 & 1 & 6 \end{bmatrix}$ and factor it.

Solution. Just compute the determinant

$$\det(A-\lambda) = \det(\begin{bmatrix} 4-\lambda & 9 & 2 \\ 3 & 5-\lambda & 7 \\ 8 & 1 & 6-\lambda \end{bmatrix}) = -\lambda^3 + 15\lambda^2 - 24\lambda + 360 \ .$$

We double check that a_0 is the determinant of A. The start is always $(-\lambda)^n$. The term in from of λ^2 is the trace of A. The only coefficient which was not so easy to get is the -24λ. How do we factor the polynomial? We know that $\lambda = 15$ is a root because $[1,1,1]^T$ is an eigenvector as it is the sum of the row elements. So, we can divide $(-\lambda^3 + 15\lambda^2 - 24\lambda + 360)/(15-\lambda)$ which is $24 + x^2$. So, we have $p_A(\lambda) = (15-\lambda)(i\sqrt{24}-\lambda)(-i\sqrt{24}-\lambda)$.

Problem: Find the characteristic polynomial of the **Boolean matrix** encoding one of the solutions of the **8 queen problems**

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} .$$

Solution. The matrix A has only one nonzero pattern. The determinant is 1. The trace is 1. So, we know the polynomial looks like $\lambda^8 - \lambda^7 + ... + 1$. The matrix $A - \lambda 1$ is partitioned with a 1×1 and 7×7 matrix. The characteristic polynomial of the 7×7 matrix is $(-\lambda^7 + 1)$. The characteristic polynomial of the 1×1 matrix is $1 - \lambda$. We get (see homework 14.3) $(1-\lambda)(1-\lambda^7) = \lambda^8 - \lambda^7 - \lambda + 1$.

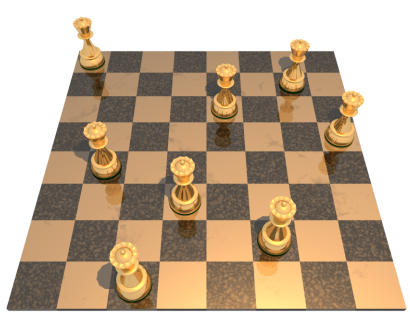


Figure 2. A solution to the 8 queen problem

Homework

This homework is due on Tuesday, 3/5/2019.

Problem 14.1: Find $p(\lambda)$ and factor: a) $A = 4$, b) $B = \begin{bmatrix} 3 & 4 \\ 5 & 2 \end{bmatrix}$, c) $C = \begin{bmatrix} 2 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 1 \end{bmatrix}$, d) $D = \begin{bmatrix} 11 & 1 & 1 & 1 & 1 & 1 \\ 11 & 1 & 1 & 1 & 1 & 1 \\ 11 & 1 & 1 & 1 & 1 & 1 \\ 11 & 1 & 1 & 1 & 1 & 1 \\ 11 & 1 & 1 & 1 & 1 & 1 \\ 11 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$, e) $E = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix}$.

Problem 14.2: Find the determinant and characteristic polynomial:

$$A = \begin{bmatrix} 101 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 102 & 3 & 4 & 5 & 6 & 7 \\ 1 & 2 & 103 & 4 & 5 & 6 & 7 \\ 1 & 2 & 3 & 104 & 5 & 6 & 7 \\ 1 & 2 & 3 & 4 & 105 & 6 & 7 \\ 1 & 2 & 3 & 4 & 5 & 106 & 7 \\ 1 & 2 & 3 & 4 & 5 & 6 & 107 \end{bmatrix}$$

Problem 14.3: a) Express the characteristic polynomial of the **partitioned matrix** $\det \begin{bmatrix} A & C \\ 0 & B \end{bmatrix}$ in the form of the characteristic polynomials of A and B.
b) What is the relation between the $p_{A^T}(\lambda)$ and $p_A(\lambda)$?

Problem 14.4: a) Given the eigenvalues $\lambda_1 = 3, \lambda_2 = 6, \lambda_3 = 4, \lambda_4 = 9$, find a non-triangular matrix which has these eigenvalues.
b) Given the eigenvalues $\lambda_1 = 1 + i$, $\lambda_2 = 3 + 4i$, $\lambda_3 = 2 - 2i$ is there a real matrix 3×3 matrix which has these eigenvalues? If no, why not?
c) Is it true that if λ is an eigenvalue of A and μ is an eigenvalue of B, then $\lambda\mu$ is an eigenvalue of AB?
d) Is it true that if λ is an eigenvalue of A, then λ^2 is an eigenvalue of A^2.
e) True or False: if λ is a non-zero eigenvalue of A, then $1/\lambda$ is an eigenvalue of A^{-1}.
f) True or False: if λ is an eigenvalue of A, then λ is an eigenvalue of A^T.

Problem 14.5: There are sometimes fast ways to compute the characteristic polynomials: An example

$$A = \begin{bmatrix} 4 & 4 & 4 \\ 2 & 2 & 8 \\ 3 & 3 & 6 \end{bmatrix} .$$

You can virtually "see" the characteristic polynomial. How?

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 15: Rising sea

Seminar

15.1. Alexander Grothendieck was a special mathematician. Both extremely creative, charismatic as well as eccentric, he not only revolutionized huge parts of mathematics, he also chose in the later part of his life to shut himself off from the world and live alone as a hermit. Today we learn about a picture which Grothendieck drew about methods to solve mathematical problems. It is the **hammer and chisel principle** versus the **rising sea approach** to solve mathematical problems. We look at this here in the context of determinants.

15.2. The **Hammer and Chisel principle** is to put the cutting edge of the chisel against the shell and to strike hard. If needed, begin again at many different points until the shell cracks-and you are satisfied. Grothendieck puts this poetically by comparing the theorem to a nut to be opened. The mathematician uses the Hammer and Chisel to reach "the nourishing flesh protected by the shell". It can also be called the Sledgehammer method. You just hit the problem with all you have and grind it through until the problem is solved.

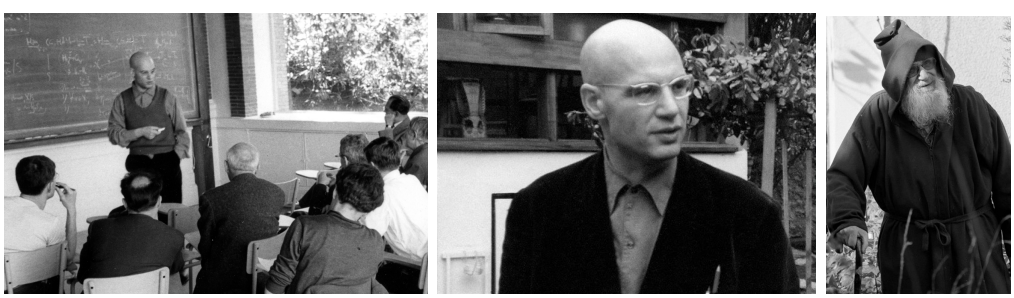


Figure 1. Three pictures of Alexander Grothendieck.

15.3. The **rising sea approach** submerges the problem first into a theory, going well beyond the results originally to be established. Grothendieck imagined to immerse the nut in some softening liquid until the shell becomes flexible. After weeks and months, by mere hand pressure, the shell opens like a perfectly ripened avocado.

15.4. We will illustrate the two principles when looking at some properties of determinants and permutation matrices. One can learn a lot about proofs in this context. Here is a statement you have proven in a homework already.

Theorem: $\det(A^T) = \det(A)$.

15.5. The proof goes by noticing that every permutation π has a dual permutation π^T which has the same sign. We just noticed that the number of up-crossings of π and π^T are the same. This might be a bit hard to see.

15.6. The "rising sea" approach is to put the problem into the larger context of **permutation groups**. The set of all permutations become a group if one realizes the elements as permutation matrices as matrices can be multiplied.

Problem A: Given a permutation matrix A belonging to a permutation π. What is the permutation matrix of the inverse permutation?

15.7. A permutation or pattern can be written as a 0-1 matrix where $P_{ij} = 1$ if $\pi(j) = i$. Now, building a theory takes time and effort. In the case of permutations, one first shows that every permutation can be written as a product of **transpositions**. A transposition is a matrix obtained from the identity matrix by swapping two rows. As we are not doing an abstract algebra course, let us assume the following fact to be for granted. The number of up-crossings of π is even if and only if one can generate π with an even number of transpositions.

15.8.

Problem B: Is it true that the number of up-crossings is equal to the minimal number of transpositions needed to realize the permutation? If yes, prove it. If not, give a counter example.

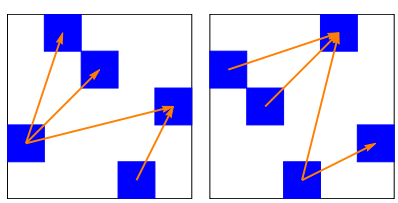


FIGURE 2. A pattern π and the inverse (transpose) pattern π^T We also draw the up-crossings. There are the same number of up-crossings for π and π^T.

15.9. This rising sea approach is cool, but one needs time to build a theory. Let us therefore play a bit more with the hammer and chisel. It is actually quite a nice task if one has the help of a computer.

15.10. Permutation matrices are matrices in which each of the columns is a basis vector e_k and no columns are the same. They solve the 8-**rook problem**:

> **Problem C:** Here is a combinatorial problem in chess. Assume you have a chess game with 8 rooks (rooks can move freely along rows or columns). How many chess rook configurations are there on a 8×8 board in which no rook threatens any other rook?

> **Problem D:** Find a chess game configuration with 4 queens on a 4×4 board (queens can move also diagonally) such that none hits the other. How many are there with 4 queens?

15.11. The Mathematica code to find the number of non-attacking Queen positions for a general n is in the homework part. In the case $n = 4$ you can do it by hand.

15.12. Here is another identity we have mentioned. You have checked it for 2×2 and 3×3 cases, but we did not prove it yet in full generality. We do that now.

> **Theorem:** $\det(AB) = \det(A)\det(B)$.

15.13. Proof: Let us see whether we can understand the proof. Let $C = AB$. Now write

$$A|B = C$$

Now, if we scale a row in A, this corresponds to scaling a row in C. Switching two rows of A corresponds to switching two rows in C. Subtracting a multiple of a row to another row also produces a subtraction of a multiple of a row to another row. While row reducing A, we do the same operations also on C. The sign changes on the left and right hand side are the same. Every division by λ on the left also reduces the determinant of C by λ. After finishing up, we have on the left hand side the determinant of $1 \cdot B$ which is the determinant of B and on the right hand side a matrix with determinant $\det(C)/\det(A)$. We see that $\det(B) = \det(C)/\det(A)$ meaning $\det(A)\det(B) = \det(C)$. QED.

> **Problem E:** The above proof does not quite work if A is not invertible. What happens in that case? Why is it that if A is not invertible, still, the identity holds?

15.14. In the proof that A^T and A have the same determinant, we need the fact that for a pattern, the transposed pattern has the same number of up-crossings. A pattern P defines an orthogonal matrix P. Now, the transpose pattern is the inverse matrix. If we can show that P and P^T have the same determinant, then we are done.

Homework

> **Problem 15.1** Find the characteristic polynomial of the permutation matrix where e_{k-1} is in the k'th column and e_n is in the first column. What are the roots of this polynomial in the case $n = 4$?

Problem 15.2 The Matrix Tree theorem assures that if L is the Laplacian of a graph with n nodes, then the coefficient a_{n-1} of the characteristic polynomial counts the number of rooted spanning trees in the graph. For the complete graph with 5 elements, the Laplacian is

$$L = \begin{bmatrix} 4 & -1 & -1 & -1 & -1 \\ -1 & 4 & -1 & -1 & -1 \\ -1 & -1 & 4 & -1 & -1 \\ -1 & -1 & -1 & 4 & -1 \\ -1 & -1 & -1 & -1 & 4 \end{bmatrix}.$$

What are the eigenvalues of L. What is the characteristic polynomial of L? What is a_4, the number of rooted spanning trees?

Problem 15.3 Here is the Laplacian of the circular graph with 5 nodes. Write this matrix in the form $L = 2 - Q - Q^{-1}$.

$$L = \begin{bmatrix} 2 & -1 & 0 & 0 & -1 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ -1 & 0 & 0 & -1 & 2 \end{bmatrix}.$$

Find the eigenvalues of Q and express the eigenvalues of L using the eigenvalues of Q.

Problem 15.4 Hammer and chisel: Experiment with the following code to get the number of n queen problems for $n = 3, ..., 8$. Alternatively, outline the life of Alexander Grothendieck.

```
PermutationMatrix[p_]:=Module[{n=Length[p],A},A=Table[0,{n},{n}];
   Do[A[[p[[k]],k]]=1,{k,Length[p]}]; A];
QueenConflicts[p_]:=Sum[Sum[If[Abs[p[[i]]-p[[j]]]==Abs[i-j],1,0],
   {i,j+1,Length[p]}],{j,Length[p]}];
F[n_]:=Module[{P=Permutations[Range[n]]},
    U=Flatten[Position[Map[QueenConflicts,P],0]];
    Table[PermutationMatrix[P[[U[[k]]]]],{k,Length[U]}]];
MatrixForm[F[6]]
```

Problem 15.5 Hammer and chisel: Experiment with the following code to get the number of n super queen solutions for $n = 10$. If you have the energy while waiting, write a mathematical essay of the length SGA (a document written by Alexander Grothendieck).

```
PermutationMatrix[p_]:=Module[{n=Length[p],A},A=Table[0,{n},{n}];
   Do[A[[p[[k]],k]]=1,{k,Length[p]}]; A];
SuperQueenConflicts[p_]:=Sum[Sum[If[Abs[p[[i]]-p[[j]]]==Abs[i-j] ||
   (Abs[i-j]==2 && Abs[p[[i]]-p[[j]]]==1) ||
   (Abs[i-j]==1 && Abs[p[[i]]-p[[j]]]==2),1,0],
   {i,j+1,Length[p]}],{j,Length[p]}];
F[n_]:=Module[{P=Permutations[Range[n]]},
    U=Flatten[Position[Map[SuperQueenConflicts,P],0]];
    Table[PermutationMatrix[P[[U[[k]]]]],{k,Length[U]}]];
MatrixForm[F[10]]   (* PATIENCE, SOAKING FOR AN HOUR!!!! *)
```

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 16: Diagonalization

Lecture

16.1. We say that $\mathcal{B} = \{v_1, v_2, \cdots, v_n\}$ is an **eigenbasis** of a $n \times n$ matrix A if it is a basis of $\mathbb{R}^n$ and every vector $v_1, \ldots, v_n$ is an eigenvector of A. The matrix $A = \begin{bmatrix} 2 & 4 \\ 3 & 3 \end{bmatrix}$ for example has the eigenbasis $\mathcal{B} = \{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -4 \\ 3 \end{bmatrix}\}$. The basis might not be unique. The identity matrix for example has **every basis** of $\mathbb{R}^n$ as eigenbasis.

16.2. Does every matrix have an eigenbasis? One could conjecture it and try to prove it, but one would fail. The answer is no! How do we find a counter example? Remember the **magic matrix** $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$? Its characteristic polynomial is $p_A(\lambda) = \lambda^2$ so that $\lambda_1 = 0, \lambda_2 = 0$ are the eigenvalues of A. The eigenvectors are in the kernel of A which is one-dimensional only as A has only one free variable. For a basis, we would need two linearly independent eigenvectors to the eigenvalue 0.

16.3. We say a matrix A is **diagonalizable** if it is similar to a diagonal matrix. This means that there exists an invertible matrix S such that $B = S^{-1}AS$ is diagonal. Remember that we often have created transformations like a reflection or projection at a subspace by choosing a suitable basis and diagonal matrix B, then get the similar matrix A.

Theorem: A is diagonalizable if and only if A has an eigenbasis.

Proof. Assume first that A has an eigenbasis $\{v_1, \cdots v_n\}$. Let S be the matrix which contains these vectors as column vectors. Define $B = S^{-1}AS$. Since

$$Be_k = S^{-1}ASe_k = S^{-1}Av_k = S^{-1}\lambda_k v_k = \lambda_k S^{-1}v_k = \lambda_k e_k$$

for every k and Be_k is the k'th column vector of B, the matrix B is diagonal with entries λ_k in the diagonal.
Assume now that A is diagonalizable. There exists an invertible matrix S such that $S^{-1}AS = B$ is a diagonal matrix with diagonal entries λ_k. The equation $Be_k = \lambda_k e_k$ means $S^{-1}ASe_k = \lambda_k e_k$ which means after multiplying with S from the left $ASe_k = S\lambda_k e_k = \lambda_k Se_k$. So, $v_k = Se_k$ are eigenvectors with eigenvalues λ_k. Because $\{v_k\}$ is the set of column vectors of S and S is invertible, $\{v_1, \cdots, v_n\}$ is a basis. $\square$

16.4. We don't yet have the answer to the question when an eigenbasis exists, but the theorem shows that the question is important. Here is a sufficient condition:

Theorem: If all eigenvalues of A are different, then an eigenbasis exists.

Proof. If λ_k is an eigenvalue of A, then $A - \lambda_k 1$ has a non-trivial kernel. Let v_k be a non-zero vector in that kernel. Then v_k is an eigenvector of λ_k. To show that $\mathcal{B} = \{v_1, \dots, v_n\}$ are linearly independent, assume that $a_1 v_1 + \cdots + a_n v_n = 0$. Applying $A - \lambda_1 1$ to this from the left gives gives $a_2(\lambda_2 - \lambda_1)v_2 + ... + a_n(\lambda_n - \lambda_1)v_n = 0$. Multiplying this with $A - \lambda_2$ gives $a_3(\lambda_3 - \lambda_2)(\lambda_3 - \lambda_1)v_3 + ... + a_n(\lambda_n - \lambda_2)(\lambda_n - \lambda_1)v_n = 0$. Multiplying with all $A - \lambda_j$ except λ_k gives $\prod_{j \neq k}(\lambda_j - \lambda_k)a_k v_k = 0$. Since v_k is not zero and all eigenvalues are distinct, $a_k = 0$. This finishes the proof of linear independence. A set of n linear independent vectors in $\mathbb{R}^n$ automatically spans and therefore is a basis. $\square$

16.5. The condition is not necessary: the identity matrix for example is a matrix which is diagonalizable (as it is already diagonal) but which has all eigenvalues 1. The eigenvalues are not distinct. Let us refine the question a bit.

16.6. The **algebraic multiplicity** of an eigenvalue λ of A is the number of times the eigenvalue appears in the list of eigenvalues. More precisely, we can write $p_A(\lambda) = (\lambda_1 - \lambda)^{a_1} \cdots (\lambda_k - \lambda)^{a_k}$, where $\lambda_1, \cdots, \lambda_k$ are all distinct and the multiplicities $a_1, \cdots a_k$ are all positive integers. The integer a_j is called the algebraic multiplicity of the eigenvalue λ_j. For example, the algebraic multiplicity of $\lambda = 1$ in the identity $n \times n$ matrix is n. The statement that all eigenvalues of A are different means that all algebraic multiplicities are 1.

16.7. The **geometric multiplicity** of an eigenvalue λ of A is the dimension of the eigenspace $\ker(A - \lambda 1)$. By definition, both the algebraic and geometric multiplies are integers larger than or equal to 1.

Theorem: geometric multiplicity of λ_k is $\leq$ algebraic multiplicity of λ_k.

Proof. If $v_1, \cdots v_m$ is a basis of $V = \ker(A - \lambda_k)$, we can complement this with a basis $w_1 \cdots, w_{n-m}$ of $V^\perp$ to get a basis of $\mathbb{R}^n$. Let S be the matrix with column vectors $\{v_1, \dots, v_m, w_1, \dots, w_{n-m}\}$. Then $B = S^{-1}AS$ is a partitioned matrix of the form

$$B = \begin{bmatrix} \lambda_k 1 & C \\ 0 & D \end{bmatrix}$$

which has the same characteristic polynomial as A. You have seen in the last homework that the characteristic polynomial of B is the product of the characteristic polynomial of $\lambda_k 1$ and the characteristic polynomial of D. This is $(\lambda_k - \lambda)^m p_D(\lambda)$ showing that $m_{alg}(A) \geq m$. $\square$

16.8. One can do this also by deformation: if $\{v_1, \dots, v_m\}$ is an orthonormal basis of $\ker(A - \lambda_k)$, define $A(t) = A + \sum_{j=1}^m (v_j v_j^T) t^j$ whose eigenvalues to the eigenvector v_j are $\lambda_k + t^j$. If t is positive but small enough, all these eigenvalues are different and also different from any other eigenvalue of A. The characteristic polynomial of $A(t)$ contains therefore the factor $(\lambda - \lambda_k - t)(\lambda - \lambda_k - t^2) \cdots (\lambda - \lambda_k - t^m)$ which in the

limit $t \to 0$ gives the factor $(\lambda - \lambda_k)^m$ showing $m_{alg}(\lambda_k) \geq m$. We come back to this perturbation method.

16.9. How do we decide whether a matrix is diagonalizable?

Theorem: A is diagonalizable $\Leftrightarrow m_{\text{alg}}(\lambda) = m_{\text{geom}}(\lambda)$ for all eigenvalues.

Proof. In that case, we have an eigenbasis for A. It is the union of the bases of the individual eigenspaces: $\mathcal{B} = \bigcup_{j=1}^{k} \mathcal{B}(\ker(A - \lambda_j))$. $\square$

If A is diagonalizable, then any polynomial of A is diagonalizable.

Proof. $S^{-1}f(A)S = f(S^{-1}AS)$ is first shown for polynomials, then by approximation, it follows for any continuous function f.

If A is diagonalizable, then A^T is diagonalizable.

Proof. Assume $S^{-1}AS = B$ is diagonal. Take the transpose: $(S^T)^{-1}A^TS^T = B$.

Examples

16.10. The matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$ is similar to $C = \begin{bmatrix} 6 & 0 & 0 \\ 7 & 4 & 0 \\ 8 & 9 & 1 \end{bmatrix}$. Proof. Both are similar to the same diagonal matrix because both are diagonalizable.

16.11. A rotation dilation matrix $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ has the eigenbasis $v_1 = [1, -i]^T$ and $v_2 = [1, i]^T$. The corresponding eigenvalues are $a - ib, a + ib$, where $\alpha = \arg(a + ib)$ is the polar angle to the vector $[a, b]^T$. There is no real eigenbasis if $b \neq 0$. Here is the calculation which shows all this

$$\begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix}^{-1} \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix} = \begin{bmatrix} a - ib & 0 \\ 0 & a + ib \end{bmatrix}.$$

16.12. The magic matrix $A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ can not be diagonalized because there is no eigenbasis. The rank of A is 1 so that the kernel, the eigenspace to the eigenvalue 0 is only one-dimensional. Similarly, any shear dilation $A = \begin{bmatrix} a & b \\ 0 & a \end{bmatrix}$ for $b \neq 0$ can not be diagonalized. Note however that $A = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$ can be diagonalized if $a \neq c$.

Homework

This homework is due on Tuesday, 3/12/2019.

Problem 16.1: Assume, A is an invertible matrix with eigenbasis $\mathcal{B}$. Decide whether this is also an eigenbasis for B. If yes, give an argument why (like stating what the eigenvalues are), if no, find a counter example.

a) $B = A^{-1}$. d) $B = e^A$. g) $B_{ij} = A_{ij}^2$.
b) $B = A^T$. e) $B = 1 + A$. h) $B = A^T A$
c) $B = A^3$. f) $B = \text{rref}(A)$. i) $B = A + A^T$

Problem 16.2: Decide from the following matrices, whether there is an eigenbasis. Decide also whether the eigenbasis is real or complex.

a) $\begin{bmatrix} 2 & -3 \\ 3 & 2 \end{bmatrix}$. b) $\begin{bmatrix} 2 & 6 \\ 6 & 9 \end{bmatrix}$. c) $\begin{bmatrix} 2 & 3 \\ 3 & -2 \end{bmatrix}$. d) $\begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix}$. e) $\begin{bmatrix} -2 & 3 \\ 0 & 2 \end{bmatrix}$.

Problem 16.3: Compute an eigenbasis. If there is an orthonormal eigenbasis, find one. If the matrix is diagonalizable, find B and S such that $S^{-1}AS = B$ is diagonal.

a) $A = \begin{bmatrix} 3 & 17 \\ 10 & 10 \end{bmatrix}$, b) $A = \begin{bmatrix} 4 & 1 & 1 \\ 1 & 2 & 3 \\ 4 & 1 & 1 \end{bmatrix}$, c) $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$.

Problem 16.4: Group the matrices which are similar!

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad D = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

$$E = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad F = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad G = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad H = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Problem 16.5: You should see the eigenvalues and eigenvectors of $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ (Make sure you get this, by discussing it.) If you write down the eigensystem or type $Eigensystem[\{\{1,1,1\},\{1,1,1\},\{1,1,1\}\}]$ in mathematica (heaven forbid you should do that!) you obtain an eigenbasis which is not orthonormal. In the next lecture, we will prove that symmetric matrices have an orthonormal eigenbasis.

a) Find an orthonormal eigenbasis to A.
b) Change one 1 to 0 so that there is an eigenbasis but no orthogonal one.
c) Change three entries 1 to 0 in A so that there is no eigenbasis.

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 17: Spectral theorem

Lecture

17.1. A real or complex matrix A is called **symmetric** or **self-adjoint** if $A^* = A$, where $A^* = \overline{A}^T$. For a real matrix A, this is equivalent to $A^T = A$. A real or complex matrix is called **normal** if $A^*A = AA^*$. Examples of normal matrices are symmetric or anti-symmetric matrices. Normal matrices appear often in applications. Correlation matrices in statistics or operators belonging to observables in quantum mechanics, adjacency matrices of networks are all self-adjoint. Orthogonal and unitary matrices are all normal.

17.2.

Theorem: Symmetric matrices have only real eigenvalues.

Proof. We extend the dot product to complex vectors as $(v, w) = v \cdot w = \sum_i \overline{v}_i w_i$ which extends the usual dot product $(v, w) = \overline{v} \cdot w$ for real vectors. This dot product has the property $(A^*v, w) = (v, Aw)$ and $(\lambda v, w) = \overline{\lambda}(v, w)$ as well as $(v, \lambda w) = \lambda(v, w)$. Now $\overline{\lambda}(v, v) = (\lambda v, v) = (Av, v) = (A^*v, v) = (v, Av) = (v, \lambda v) = \lambda(v, v)$ shows that $\overline{\lambda} = \lambda$ because $(v, v) = \overline{v} \cdot v = |v_1|^2 + \cdots + |v_n|^2$ is non-zero for non-zero vectors v. $\square$

17.3.

Theorem: If A is symmetric, then eigenvectors to different eigenvalues are perpendicular.

Proof. Assume $Av = \lambda v$ and $Aw = \mu w$. If $\lambda \neq \mu$, then the relation $\lambda(v, w) = (\lambda v, w) = (Av, w) = (v, A^T w) = (v, Aw) = (v, \mu w) = \mu(v, w)$ is only possible if $(v, w) = 0$. $\square$

17.4. If A is a $n \times n$ matrix for which all eigenvalues are different, we say such a matrix has **simple spectrum**. The "wiggle-theorem" tells that we can approximate a given matrix with matrices having simple spectrum:

Theorem: A symmetric matrix can be approximated by symmetric matrices with simple spectrum.

Proof. We show that there exists a curve $A(t) = A(t)^T$ of symmetric matrices with $A(0) = A$ such that $A(t)$ has simple for small positive t.
Use induction with respect to n. For $n = 1$, this is clear. Assume it is true for n, let A be a $(n+1) \times (n+1)$ matrix. It has an eigenvalue λ_1 with eigenvector v_1 which we assume to have length 1. The still symmetric matrix $A + tv_1 \cdot v_1^T$ has the same eigenvector v_1 with eigenvalue $\lambda_1 + t$. Let $v_2, \dots, v_n$ be an orthonormal basis of $V^\perp$ the space perpendicular to $V = \text{span}(v_1)$. Then $A(t)v = Av$ for any v in $V^\perp$. In that basis, the matrix $A(t)$ becomes $B(t) = \begin{bmatrix} \lambda_1 + t & C \\ 0 & D \end{bmatrix}$. Let S be the orthogonal matrix which contains the orthonormal basis $\{v_1, v_2, \dots, v_n\}$ of $\mathbb{R}^n$. Because $B(t) = S^{-1}A(t)S$ with orthogonal S, also $B(t)$ is symmetric implying that $C = 0$. So, $B(t)$ preserves D and $B(t)$ restricted to D does not depend on t. In particular, all the eigenvalues are different from $\lambda_1 + t$. By induction we find a curve $D(t)$ with $D(0) = D$ such that all the eigenvalues of $D(t)$ are different and also different from $\lambda_1 + t$. □

17.5. This immediately implies the **spectral theorem**

Theorem: Every symmetric matrix A has an orthonormal eigenbasis.

Proof. Wiggle A so that all eigenvalues of $A(t)$ are different. There is now an orthonormal basis $\mathcal{B}(t)$ for $A(t)$ leading to an orthogonal matrix $S(t)$ such that $S(t)^{-1}A(t)S(t) = B(t)$ is diagonal for every small positive t. Now, the limit $S(t) = \lim_{t\to 0} S(t)$ and also the limit $S^{-1}(t) = S^T(t)$ exists and is orthogonal. This gives a diagonalization $S^{-1}AS = B$. The ability to diagonalize is equivalent to finding an eigenbasis. As S is orthogonal, the eigenbasis is orthonormal. □

17.6. What goes wrong if A is not symmetric? Why can we not wiggle then? The proof applied to the magic matrix $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ gives $A(t) = A + te_1 \cdot e_1^T = \begin{bmatrix} t & 1 \\ 0 & 0 \end{bmatrix}$ which has the eigenvalues $0, t$. For every $t > 0$, there is an eigenbasis with eigenvectors $[1, 0]^T, [1, -t]$. We see that for $t \to 0$, these two vectors collapse. This can not happen in the symmetric case because eigenvectors to different eigenvalues are orthogonal there. We see also that the matrix $S(t)$ converges to a singular matrix in the limit $t \to 0$.

17.7. First note that if A is normal, then A has the same eigenspaces as the symmetric matrix $A^*A = AA^*$: if $A^*Av = \lambda v$, then $(A^*A)Av = AA^*Av = A\lambda v = \lambda Av$, so that also Av is an eigenvector of A^*A. This implies that if A^*A has simple spectrum, (leading to an orthonormal eigenbasis as it is symmetric), than A also has an orthonormal eigenbasis, namely the same one. The following result follows from a Wiggling theorem for normal matrices:

17.8.

Theorem: Any normal matrix can be diagonalized using a unitary S.

Examples

17.9. A matrix A is called doubly stochastic if the sum of each row is 1 and the sum of each column is 1. Doubly stochastic matrices in general are not normal, but they

are in the case $n = 2$. Find its eigenvalues and eigenvectors. The matrix must have the form

$$A = \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}$$

It is symmetric and therefore normal. Since the rows sum up to 1, the eigenvalue 1 appears to the eigenvector $[1,1]^T$. The trace is $2a$ so that the second eigenvalue is $2a-1$. Since the matrix is symmetric and for $a \neq 0$ the two eigenvalues are distinct, by the theorem, the two eigenvectors are perpendicular. The second eigenvector is therefore $[-1,1]^T$.

17.10. We have seen the quaternion matrix belonging to $z = p + iq + jr + ks$: $\begin{bmatrix} p & -q & -r & -s \\ q & p & s & -r \\ r & -s & p & q \\ s & r & -q & p \end{bmatrix}$. As an orthogonal matrix, it is normal. Let $v = [q, r, s]$ be the space vector defined by the quatenion. Then the eigenvalues of A are $p \pm i|v|$, both with algebraic multiplicity 2. The characteristic polynomial is $p_A(\lambda) = (\lambda^2 - 2p\lambda + |z|^2)^2$.

17.11. Every normal 2×2 matrix is either symmetric or a rotation-dilation matrix. Proof: just write down $AA^T = A^TA$. This gives a system of quadratic equations for four variables a, b, c, d. This gives $c = b$ or $c = -b, d = a$.

Illustrations

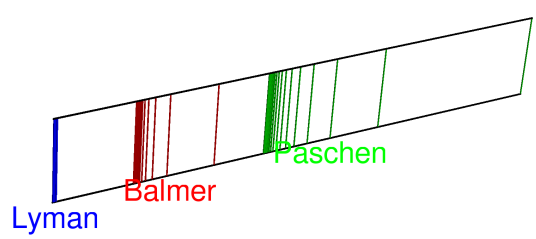


Figure 1. The atomic hydrogen emission spectrum is given by eigenvalue differences $1/\lambda = R(1/n^2 - 1/m^2)$, where R is the **Rydberg constant**. The **Lyman series** is in the ultraviolet range. The **Balmer series** is is visible in the solar spectrum. The **Paschen Series** finally is in the infrared band. By Niels Bohr, the n'th eigenvalue of the self-adjoint Hydrogen operator A is $\lambda_n = -Rhc/n^2$, where h is the **Planck's constant** and c is the **speed of light**. The spectra we see are differences of such eigenvalues.

Homework

This homework is due on Piday Tuesday, 3/12/2019.

Problem 17.1: Give a reason why its true or provide a counterexample.
a) The product of two symmetric matrices is symmetric.
b) The sum of two symmetric matrices is symmetric.
c) The sum of two anti-symmetric matrices is anti-symmetric.
d) The inverse of an invertible symmetric matrix is symmetric.
e) If B is an arbitrary $n \times m$ matrix, then $A = B^T B$ is symmetric.
f) If A is similar to B and A is symmetric, then B is symmetric.
g) $A = SBS^{-1}$ with $S^T S = I_n$, A symmetric $\Rightarrow$ B is symmetric.
h) Every symmetric matrix is diagonalizable.
i) Only the zero matrix is both anti-symmetric and symmetric.
j) The set of normal matrices forms a linear space.

Problem 17.2: Find all the eigenvalues and eigenvectors of the matrix

$$A = \begin{bmatrix} 2222 & 2 & 3 & 4 & 5 \\ 2 & 2225 & 6 & 8 & 10 \\ 3 & 6 & 2230 & 12 & 15 \\ 4 & 8 & 12 & 2237 & 20 \\ 5 & 10 & 15 & 20 & 2246 \end{bmatrix}.$$

Problem 17.3: a) Find the eigenvalues and orthonormal eigenbasis of $A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$. b) Find $\det(\begin{bmatrix} 7 & 2 & 2 & 2 & 2 \\ 2 & 7 & 2 & 2 & 2 \\ 2 & 2 & 7 & 2 & 2 \\ 2 & 2 & 2 & 7 & 2 \\ 2 & 2 & 2 & 2 & 7 \end{bmatrix})$ using eigenvalues

Problem 17.4: a) Group the matrices which are similar.

$$A = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \quad C = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \quad D = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

b) Which of the above matrices are normal?

Problem 17.5: Find the eigenvalues and eigenvectors of the Laplacian of the Bunny graph. The Laplacian of a graph with n nodes is the $n \times n$ matrix A which for $i \neq j$ has $A_{ij} = -1$ if i, j are connected and 0 if not. The diagonal entries A_{ii} are chosen so that each row adds up to 0.

$$A = \begin{bmatrix} 2 & -1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ -1 & -1 & 4 & -1 & -1 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 0 & 1 \end{bmatrix}$$

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 18: Spectra

Seminar

18.1. If you hit a drum, you hear eigenvalues of its Laplacian. An actual drum is a finite network of fibres. The sound we hear when hitting the drum is made of spectra of the eigenvalues of a **Laplacian matrix** defined by the network. You remember the Laplacian matrix from last semester, when we defined the discrete gradient d, the discrete divergence d^* and formed the Laplacian $L = d^*d$. We even computed some spectra already then.

18.2. Mark Kac asked in 1966 whether one can hear the shape of a drum. [1] [2] This problem is still not solved for convex drums. A drum is a planar region with piecewise smooth boundary for which the line between two arbitrary points is contained in the drum.

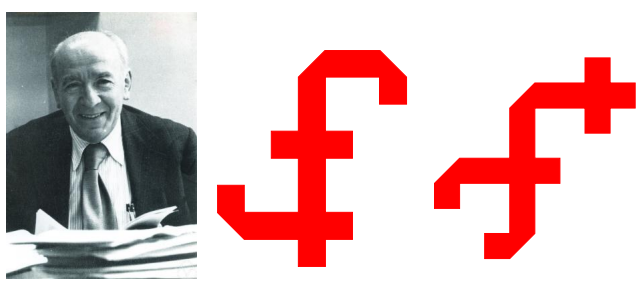


Figure 1. Mark Kac asked in 1966 whether one can hear the shape of a drum. Carolyn Gordon, David Webb and Scott Wolpert found in 1992 the first non-isometric isospectral domains. They are shown to the right.

18.3. Instead of looking at actual drums, we can look at a small finite network and compute the eigenvalues of the Laplacian matrix L associated to this network. It is defined as follows. L is a $n \times n$ matrix if there are n nodes. If node i is connected to node j, let $L_{ij} = L_{ji} = -1$, otherwise put a zero. In the diagonal,place $L_{ii} = d(i)$, where $d(i)$ is the number of neighbors of the node i.

[1] M. Kac, Can One Hear the Shape of a Drum?, American Mathematical Monthly, 73, 1966, page 1-23

[2] C. Gordon, D.L. Webb and S. Wolpert, One can not hear the shape of a drum, Bull. Amer. Math. Soc. 27 (1992), page 134-138

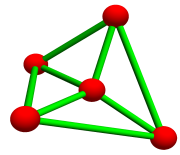


FIGURE 2. A small drum with 5 nodes.

18.4. For a wheel graph with 5 nodes, we have

$$\begin{bmatrix} 4 & -1 & -1 & -1 & -1 \\ -1 & 3 & -1 & 0 & -1 \\ -1 & -1 & 3 & -1 & 0 \\ -1 & 0 & -1 & 3 & -1 \\ -1 & -1 & 0 & -1 & 3 \end{bmatrix}$$

There are two eigenvalues 5 with eigenvectors $[-2, 0, 1, 0, 1]$ and $[-2, 1, 0, 1, 0]$ and two eigenvalues 3 with eigenvectors $[0, -1, 0, 1, 0]$ and $[0, 0, -1, 0, 1]$. Finally, there is the eigenvalue 0 with eigenvector $[1, 1, 1, 1, 1]$.

18.5. Here are two isospectral networks.

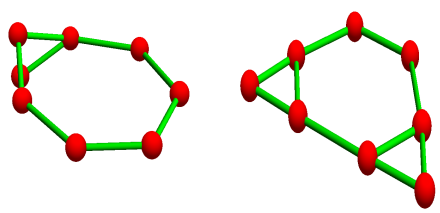


FIGURE 3. An example of a cospectral pair for the Laplace operator L_0 given by W. Hamers and E. Spence in 2004. Both have 8 nodes.

18.6. What are the properties of the Laplacian L?

Problem A: The Laplacian L of a network always has an eigenvalue 0.

Hint. Remember the lecture last week.

18.7. We can write $L = d^*d$, where d is a $m \times n$ matrix, where m is the number of edges and n the number of nodes.

Problem B: Any matrix L of the form $L = A^*A$ has real eigenvalues.

Hint. Use a theorem.

Problem C: Show that $L = A^*A$ has non-negative eigenvalues.

Hint: Here is a start: assume v is an eigenvector of L to the eigenvalue λ then $(v, Lv) = (v, A^*Av) = (Av, Av) = |Av|^2$.

18.8. A matrix is doubly stochastic if all matrix entries are non-negative and each row and each column adds up to 1. Each column vector is then a discrete probability distribution.

Theorem: The inverse g of $1 + L$ is doubly stochastic.

Problem D: Either prove this theorem or look whether you find the theorem.

18.9. Hint: This result follows from the matrix forest theorem which tells that $\det(L+1)$ is the number of rooted forests in a graph. The entries of g_{ij} can be interpreted as probabilities of forests.

18.10. In the case of the wheel graph before

$$24g = \begin{bmatrix} 8 & 4 & 4 & 4 & 4 \\ 4 & 9 & 4 & 3 & 4 \\ 4 & 4 & 9 & 4 & 3 \\ 4 & 3 & 4 & 9 & 4 \\ 4 & 4 & 3 & 4 & 9 \end{bmatrix} .$$

All rows and columns add up to 24 so that g is stochastic.

18.11. Why are spectra interesting. One reason is Hückel theory.

FIGURE 4. Erich Hueckel devised a method to compute approximate molecular orbital electron systems. Pavel Chebotarev and Elena Shamis proved the matrix forest theorem and also were the first to see that the inverse of $1 + L$ is doubly stochastic.

18.12. The Wikipedia article about Hueckel mentions that Hueckel had a lack of communication skills. But a poem he wrote about Schrödinger shows his humor (it was freely translated by Felix Bloch)

Erwin with his psi can do
Calculations quite a few.
But one thing has not been seen:
Just what does psi really mean?

Homework

This homework is due on 3/12/2019

Problem 18.1 Compute the eigenvalues of the Laplacian of K_n, the complete graph with n nodes.

Problem 18.2 Below is Mathematica code encoding two networks. The **Halbeisen-Hungerbuehler isospectral graphs**. Verify that the Laplacians have the same characteristic Polynomial. What is $\det(L+1)$, the number of rooted forests of the Halbeisen-Hungerbuehler graphs? Check that $(L+1)^{-1}$ is a doubly stochastic matrix.

```
s1=UndirectedGraph[Graph[{71->72,72->73,73->74,74->75,75->76,76->77,77->70,70->71,
  71->73,75->77,77->1,1->2,2->73,75->3,3->4,4->71,72->76,74->5,76->6,70->7,72->8}]];
s2=UndirectedGraph[Graph[{71->72,72->73,73->74,74->75,75->76,76->77,77->70,70->71,
  71->73,75->77,74->1,1->2,2->70,72->3,3->4,4->76,72->76, 71->5,73->6,75->7,77->8}]];
L1=Normal[KirchhoffMatrix[s1]]; L2=Normal[KirchhoffMatrix[s2]];
K1=IdentityMatrix[Length[L1]]+L1;
```

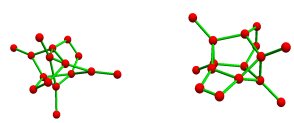


Figure 5. The Halbeisen Hungerbuehler isospectral graphs

Problem 18.3 According to **Hueckel theory**, the eigenvectors of L can help to understand the information about the distribution of electrons in a molecule. Below is an example, where we compute the eigenvalues and eigenvectors of the water molecule. What is the largest eigenvalue of the caffeine molecule? What is the corresponding eigenvector?

```
A=Normal[ChemicalData["Water", "AdjacencyMatrix"]];
L=DiagonalMatrix[Total[A]]-A; Eigensystem[L]
```

Problem 18.4 **Euler's handshake formula** tells that the trace of L is twice the number of edges in the graph of L. Prove this formula.

Problem 18.5 The Pseudo determinant $\mathrm{Det}(L)$ of a matrix L is the product of the non-zero eigenvalues of L. In comparison, the determinant is the product of all the eigenvalues. The Matrix tree theorem tells that $\mathrm{Det}(L)$ with Laplacian L is the number of rooted trees in a graph. Find a formula for the number of rooted trees in a complete graph K_n.

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 19: Discrete dynamical systems

Lecture

19.1. A $n \times n$ matrix A defines a linear transformation $T(x) = Ax$. Iterating gives a sequence of vectors $x, Ax, A^2x, \dots A^nx, \dots$. It is called the **orbit** of x of the **discrete dynamical system** defined by A. We write $x(t) = A^t x$ so that $x(0)$ is the **initial condition**. We think of t as **time** and $x(t)$ as the situation at time t.

19.2. For example, for $A = \begin{bmatrix} 5 & 1 \\ 2 & 4 \end{bmatrix}$ we have $A^{10} = \begin{bmatrix} 40330467 & 20135709 \\ 40271418 & 20194758 \end{bmatrix}$. For an initial condition like $\begin{bmatrix} 3 \\ 4 \end{bmatrix}$, we get $A^{10}x = \begin{bmatrix} 201534237 \\ 201593286 \end{bmatrix}$. Of course, we want a better way to compute the orbit than just multiplying the matrix again and again.

19.3. Eigenvectors provide relief: if v is an eigenvector, then $Av = \lambda v$ and $A^t v = \lambda^t v$. Using linearity of $Tx = Ax$, we see that if x is a sum of eigenvectors like $x = c_1 v_1 + \cdots + c_n v_n$, then $Ax = A(c_1 v_1 + \cdots + c_n v_n) = c_1 A v_1 + \cdots + c_n A v_n = c_1 \lambda_1 v_1 + \cdots + c_n \lambda_n v_n$.

> If $\mathcal{B} = \{v_1, \cdots, v_n\}$ is an eigenbasis of A and $x(0) = c_1 v_1 + \cdots + c_n v_n$ then $x(t) = c_1 \lambda_1^t v_1 + \cdots + c_n \lambda_n^t v_n$ solves $x(t+1) = Ax(t)$.

This solution is called the **closed-form solution**.

19.4. In the above example, we have the eigenvectors $v_1 = [1, 1]^T, v_2 = [-1, 2]^T$ to the eigenvalues $\lambda_1 = 6, \lambda_2 = 3$ and $x = [3, 4]^T$ is $c_1 v_1 + c_2 v_2$ with $c_1 = 10/3$ and $c_2 = 1/3$. The closed-form solution is

$$x(t) = \frac{10}{3} 6^t \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{1}{3} 3^t \begin{bmatrix} -1 \\ 2 \end{bmatrix} .$$

It is no problem to evaluate this at any t, like $t = 10$ which gives the above result for $A^{10}x$. We can even evaluate this for $t = 10^{100}$ but can not produce A^t.

Examples

19.5. The sequence of numbers $u(0), u(1), u(2), \dots, u(t), \dots$ defined by the **recursion**

$$u(t+1) = u(t) + u(t-1)$$

with initial condition $u(0) = 0$, $u(1) = 1$ is called the **Fibonacci sequence**. We will solve this system in class. It leads to a recursion

$$\begin{bmatrix} u(t+1) \\ u(t) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} u(t) \\ u(t-1) \end{bmatrix}$$

so that we want to compute the eigenvalues and eigenvectors of the matrix $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$. This leads to the formula of Binet

$$u(t) = \frac{\left(\frac{1}{2}\left(1+\sqrt{5}\right)\right)^t - \left(\frac{1}{2}\left(1-\sqrt{5}\right)\right)^t}{\sqrt{5}} .$$

FIGURE 1. Leonardo Pisano 1170-1250 was later called Fibonacci by historians. He published his book "Liber Abaci" in 1202, in which the Hindo-Arabic numeral system and place value was introduced to the West. Also discussed was the Fibonacci sequence. The picture of Fibonacci shown is of unknown origin and almost certainly a work of fiction. The original edition of the text is lost, but there is a version from 1228. Mathematics related to the Fibonacci numbers already appeared in a combinatorial context between 450 and 200 BC in "Chanda-Sutra" = "The art of prosody" written by the Indian mathematician Acharya Pingala.

[1]

19.6. For a **discrete recursion** equation like $u(t+1) = 2u(t) + u(t-1)$ and initial conditions like $u(0) = 1$ and $u(1) = 1$ and get all the other values fixed. We have $u(2) = 3, u(3) = 10$, etc. A discrete recursion can always be written as a discrete dynamical system. Just use the vector $x(t) = [u(t), u(t-1)]^T$ and write

$$x(t+1) = \begin{bmatrix} u(t+1) \\ u(t) \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} u(t) \\ u(t-1) \end{bmatrix} = AX(t) .$$

Now we can compute the closed-form solution. The eigenvalues are $1+\sqrt{2}$ and $1-\sqrt{2}$ to the eigenvectors $v_1 = [1+\sqrt{2}, 1]^T, v_2 = [1-\sqrt{2}, 1]^T$. The initial condition $[1, 1]$ is $x(0) = (1/2)v_1+(1/2)v_2$. The closed form solution is $x(t) = (1/2)(1+\sqrt{2})^t v_1+(1/2)(1-\sqrt{2})^t v_2$.

[1]K. Plofker, Mathematics in India, Princeton University press, 2009

19.7. If $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ is a rotation dilation then we can compute A^n quickly. As it is a rotation dilation matrix with angle $\theta = \arg(a+ib)$ and scaling factor $r = \sqrt{a^2+b^2}$, we have $A^n = r^n \begin{bmatrix} \cos(n\theta) & -\sin(n\theta) \\ \sin(n\theta) & \cos(n\theta) \end{bmatrix}$.

19.8. The recursion $u(t+1) = u(t) - u(t-1)$ with $u(0) = 0, u(1) = 1$ leads to the discrete dynamical system with the matrix $A = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$. We see that A^6 is the identity. Every initial vector is mapped after 6 iterations back to its original starting point. The eigenvalues are complex.

19.9. A matrix with non-negative entries for which the sum of the columns entries add up to 1 is called **stochastic matrix** or **Markov matrix**.

Theorem: Markov Matrices have an eigenvalue 1

Proof. The matrix A^T has rows which add up to 1. The vector $[1, 1, \dots, 1]^T$ therefore is an eigenvector of A^T to the eigenvalue 1. As A has the same eigenvalues as A^T, also A has an eigenvalue 1. □

19.10. The next result is part of the **Perron-Frobenius theorem**.

Theorem: If all entries of the Markov matrix A are positive, then the eigenvalue 1 is maximal and has algebraic multiplicity 1.

Proof. Given $Av = v$, we show that all v_k are the same. We first prove by contradiction that $|v_k|$ is constant. Otherwise, let m be the index where $|v_k|$ is largest and let i be an index where $|v_k|$ is smallest. $|v_i| < |v_m|$ and

$$|v_m| = |\sum_{k=1}^n A_{mk} v_k| \leq \sum_{k=1}^n A_{mk}|v_k| \leq \sum_{k=1, k\neq i}^n A_{mk}|v_m| + A_{mi}|v_i| < \sum_{k=1}^n A_{mk}|v_m| = |v_m| \ .$$

Contradiction. Now, we show that v_k everywhere has the same sign. By scaling, we can assume $v_k = \pm 1$. If there should be different signs let m be an index where $v_m = 1$ and i be an index where $v_i = -1$. Repeat the same computation:

$$1 = v_m = \sum_{k=1}^n A_{mk} v_k \leq \sum_{k=1}^n A_{mk}|v_k| \leq \sum_{k=1, k\neq i}^n A_{mk} 1 + A_{mi}(-1) < \sum_{k=1}^n A_{mk} 1 = 1$$

which is again a contradiction. □

19.11. Intuitively, applying the matrix A averages the coordinate values of A and smooths out the distribution making it constant. In the homework you prove

Theorem: The product of Markov matrices is a Markov matrix.

Homework

This homework is due on Tuesday, 3/26/2019. [2]

Problem 19.1: The vector $A^t x$ gives pollution levels in the Silvaplana, Sils and St Moritz lake t weeks after an oil spill. The matrix is $A = \begin{bmatrix} 0.7 & 0 & 0 \\ 0.1 & 0.6 & 0 \\ 0 & 0.2 & 0.8 \end{bmatrix}$ and $x(0) = \begin{bmatrix} 100 \\ 0 \\ 0 \end{bmatrix}$ is the initial pollution level. Find a closed-form solution for $x(t)$.

Problem 19.2: A **Lilac bush** has $n(t)$ new branches and $o(t)$ old branches at the beginning of each year t. During the year, each old branch will grow two new branches and remain old and every new branch will becomes an old branch. Write down the matrix A such that $\begin{bmatrix} n(t+1) \\ o(t+1) \end{bmatrix} = A \begin{bmatrix} n(t) \\ o(t) \end{bmatrix}$ and find closed-formulas for $n(t), o(t)$ if $\begin{bmatrix} n(0) \\ o(0) \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$.

Problem 19.3: Given a polyhedron with vertex,edge and face data $x = [v, e, f]^T$, one can do a **Barycentric refinement**. This produces a new polyhedron with data $x(1) = Ax(0)$ written out as $\begin{bmatrix} v(1) \\ e(1) \\ f(1) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 6 \\ 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} v(0) \\ e(0) \\ f(0) \end{bmatrix}$. How many vertices, edges and faces does the icosahedron with $[v(0), e(0), f(0)]^T = [12, 30, 20]^T$ have after 10 Barycentric refinements?

Problem 19.4: The **Lucas numbers** are defined by the same recursion as the Fibonacci numbers. It is only that we don't start with $(u(0), u(1)) = (0, 1)$ but with $(u(0), u(1)) = (2, 1)$. Find a close formula for the t'th Lucas numbers $2, 1, 3, 4, 7, 11, 18, ...$

Problem 19.5: a) Prove that the product of two Markov matrices A, B is a Markov matrix.
b) What happens if we would define Markov as the row sum adding up to 1. Why does the result still hold? Hint for a) You have to show that if $\sum_{j=1}^n A_{jk} = 1$ for all k and $\sum_{j=1}^n B_{jk} = 1$ for all k then $\sum_{j=1}^n \sum_{l=1}^n A_{jl} B_{lk} = 1$ for all k.

Oliver Knill, knill@math.harvard.edu, Math 22b, Harvard College, Spring 2019

[2] 19.1 and 19.2 are slightly modified problems from the book of Otto Bretscher.

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 20: Differential equations

Lecture

20.1. A **differential equation** is an equation for an **unknown function** or vector valued function $x(t)$ which involves at least one derivative $\frac{dx}{dt} = x'$ for x. An example is $x'(t) = x(t)^2 + t^3$ or $x''(t) + tx'(t)x(t) = \sin(x(t))$. The **order** of a differential equation is the highest derivative which appears. The first example was a **first order differential equation**, the later was a **second order differential equation**.

20.2. Many differential equations can be solved using **separation of variables**. Let us look at the equation

$$\frac{dx}{dt} = 3x \ .$$

To solve this, put all x on one side and all t on the other side:

$$\frac{dx}{x} = 3dt, \qquad \Rightarrow \qquad \int \frac{dx}{x} = \int 3dt + c$$

which gives $\log(x) = 3t + c$ and $x(t) = e^{3t+c} = Ce^{3t}$ for some constant C.

20.3. It is custom to write the constant C as $x(0)$ because this is the value we get for $t = 0$. Also used is the notation $x'(t)$ or $\dot{x}(t)$ for the derivative. The former has been used by Leibniz, the later by Newton.

$x' = \lambda x$ has the solution $x(0)e^{\lambda t}$.

20.4. For positive λ we get **exponential growth**. This is the most primitive model for population growth. For negative λ we get **exponential decay**. This is used to model **radioactive decay** and is used in **carbon-14 dating**. In both cases the rate of change is proportional to the sample size. The λ is the fertility in population dynamics or **reciprocal mean life** which is 1/8.267 years for C^{14}. This means that in 8.2 years 1/e=1/2.718. Using $\log_2$ this means that in 5.72 years half of the C^{14} isotopes have decayed to C^{12}.

20.5. Given a $n \times n$ matrix A, we can look at the **system of linear differential equations** $x'(t) = Ax(t)$. If $x(0)$ is given, one can write $x(t) = e^{At}x(0)$. But as before in the discrete case, where the power A^t was hard to compute, also the exponential e^{At} is not convenient. We get more insight if we use an eigenbasis.

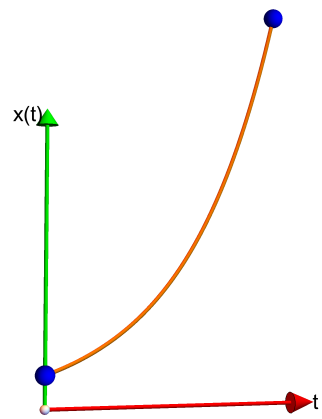


FIGURE 1. The exponential function.

20.6. For example, for $A = \begin{bmatrix} 5 & 1 \\ 2 & 4 \end{bmatrix}$ and an initial condition like $\begin{bmatrix} 3 \\ 4 \end{bmatrix}$, we proceed exactly in the same way. We have already the eigenvectors $v_1 = [1,1]^T, v_2 = [-1,2]^T$ to the eigenvalues $\lambda_1 = 6, \lambda_2 = 3$ and $x = [3,4]^T$ is $c_1 v_1 + c_2 v_2$ with $c_1 = 10/3$ and $c_2 = 1/3$. The closed-form solution is

$$x(t) = \frac{10}{3} e^{6t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{1}{3} e^{3t} \begin{bmatrix} -1 \\ 2 \end{bmatrix} .$$

Also here, we have no problem evaluating this at any time t.

If $\mathcal{B} = \{v_1, \cdots, v_n\}$ is an eigenbasis of A and $x(0) = c_1 v_1 + \cdots + c_n v_n$ then $x(t) = c_1 e^{\lambda_1 t} v_1 + \cdots + c_n e^{\lambda_n t} v_n$ solves $x'(t) = Ax(t)$.

20.7. Both for differential equations $\dot{x}(t) = Ax(t)$ in two dimensions as well as for discrete systems $x(t+1) = Ax(t)$, one can see the right hand side as a vector field and the solution curve $x(t)$ as a **flow line** of $x' = F(x(t))$. These **phase portraits** together with some solution curves can give more qualitative information.

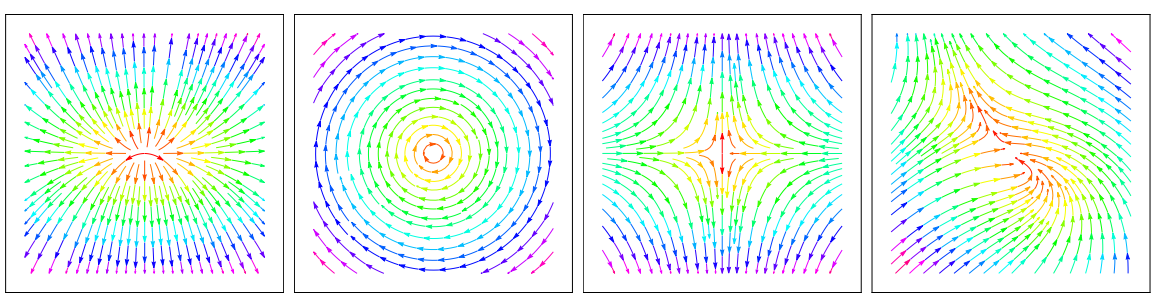


FIGURE 2. Stream plots of $F(x,y) = (x, 2y)$, $F(x,y) = (y, -x)$, $F(x,y) = (-x, 2y)$, $F(x,y) = (-y - x)$, $x^2 + y^2)$. The first three are linear systems, the last one is not linear.

Examples

20.8. The linear equation $x' = ax$ produces exponential growth $x(t) = x(0)e^{at}$. This is not sustainable in the long term. For a population, the food might run out for example. The **logistic equation** $x' = ax - x^2$ takes care of this. It is a model, where the growth is stopped when $x = a$. Let us solve a concrete problem. Assume $x' = x(1-x)$ and $x(0) = 10$. What is $x(t)$? Using **separation of variables** we get $dx/(x(1-x)) = dt$ so that $\log(x) - \log(1-x) = t + c$. This Leads to $x(t) = 10e^t/(10e^t - 9)$.

20.9. The equation $x' = x^2$ with $x(0) = 1$ is solved by $1/(1-t)$. The solution blows up in finite time. This example shows the need for some conditions so that solutions exist for all times. We will not go into this here. For linear differential equations, we always have solutions for all times.

20.10. The equation $x' = 2\sqrt{x}$ with $x(0) = 0$ is solved by $x(t) = t^2$ as well as with $x(t) = 0$. We see that there is no unique solution. Also this is a phenomenon which only happens for non-linear systems.

20.11. Find a closed-formula for the solution of the system

$$\begin{aligned} \dot{x}_1 &= x_1 + 2x_2 \\ \dot{x}_2 &= 4x_1 + 3x_2 \end{aligned}$$

with $x(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. The system can be written as $\dot{x} = Ax$ with $A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$. The matrix A has the eigenvector $v_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ to the eigenvalue -1 and the eigenvector $v_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ to the eigenvalue 5.
Because $Av_1 = -v_1$, we have $v_1(t) = e^{-t}v$. Because $Av_2 = 5v_1$, we have $v_2(t) = e^{5t}v$. The vector v can be written as a linear-combination of v_1 and v_2: $v = \frac{1}{3}v_2 + \frac{2}{3}v_1$. Therefore, $x(t) = \frac{1}{3}e^{5t}v_2 + \frac{2}{3}e^{-t}v_1$.

Remarks

20.12. Finding solutions of nonlinear differential equations can be hard. Sometimes, we can not find closed-form solutions. Like for $x'(t) = e^{-t^2}$, where $x(t)$ as an anti-derivative of e^{-t^2} has no elementary solution. But if you write down a random equation, you will probably not be able to give a concrete solution. Try Mathematica's "Dsolve" procedure on $x' = x + \sin(x)$ for example.
The closed-form solution like $x(t) = e^{At}x(0)$ for $\dot{x} = Ax$ does not give us much insight what happens. One wants to understand the solution quantitatively and answer questions like: what happens in the long term? Is the origin stable? Are there periodic solutions? Can one decompose the system into simpler subsystems?

Homework

This homework is due on Tuesday, 3/27/2019.

Problem 20.1: Solve the differential equations.
a) $x' = \sin(t)x$ with $x(0) = 2$, b) $x' = t/(5x^4)$,with $x(0) = 2$.
c) $x' = 1 + x^2$,with $x(0) = 0$, d) $x' = 1/\cos(x)$,with $x(0) = 9$.

Problem 20.2: Solve the differential equation $x' = 6x^{3/2}$. This equation does not have a unique solution with $x(0) = 0$. Find two.

Problem 20.3: Solve the system

$$\frac{dx}{dt} = Ax, \quad A = \begin{bmatrix} 4 & 9 \\ 7 & 6 \end{bmatrix}$$

with initial condition $x(0) = \begin{bmatrix} 10 \\ -6 \end{bmatrix}$. Draw the phase portrait.

Problem 20.4: A population model is given by

$$\begin{aligned} \frac{dx}{dt} &= 15x - 13y \\ \frac{dy}{dt} &= 8x - 11y \end{aligned}$$

First decide whether it is a symbiosis, competition or predator-prey model. then sketch the phase portrait in the first quadrant and decide for which initial conditions the populations die out.

Problem 20.5: A door opens on one side only. A spring mechanism closes the door which forms an angle $\theta(t)$ with the frame. The angular velocity is $\omega(t) = \frac{d\theta}{dt}(t)$. The differential equations are

$$\begin{aligned} \frac{d\theta}{dt} &= \omega \\ \frac{d\omega}{dt} &= -2\theta - 3\omega \end{aligned}$$

The first equation is the definition, the second incorporates the force -2θ of the spring and the friction -3ω.
Sketch a phase portrait for the system and use this to answer the question, for which initial conditions, the door reaches $\theta = 0$ with negative ω.

[1]

[1] 20.5 is the slamming door problem by Otto Bretscher.

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 21: The golden mean

SEMINAR

21.1. The **Fibonacci recursion** $F_{n+1} = F_n + F_{n-1}$ leads to the **Fibonacci sequence**

$$1, 1, 2, 3, 5, 8, \ldots .$$

There is an explicit formula of Binet involving the **golden mean**

$$\phi = \frac{(1+\sqrt{5})}{2} .$$

The golden mean has also been called the **"divine proportion."** The number appeared first in Euclid's elements around 350 BC.

Problem A: Verify from the **formula of Binet** that $F_{n+1}/F_n \to \phi$.

21.2. Both the Fibonacci numbers and the golden ratio are popular subjects. A few books dedicated to the subject are listed at the end. If we look at the number ϕ, except maybe for π, there is no other number about which so much has been written about. Fibonacci numbers connect with nature: Devlin writes in "Fibonacci's arithmetic revolution" *"an iris has 3 petals; primroses, buttercups, wild roses, larkspur, and columbine have 5; delphiniums have 8; ragwort, corn marigold, and cineria 13; asters, black-eyed Susan, and chicory 21; daisies 13, 21, or 34; and Michaelmas daisies 55 or 89. Sunflower heads, and the bases of pine cones, exhibit spirals going in opposite directions. The sunflower has 21, 34, 55, 89, or 144 clockwise, paired respectively with 34, 55, 89, 144, or 233 counterclockwise; a pine-cone has 8 clockwise spirals and 13 counterclockwise. All Fibonacci numbers.*

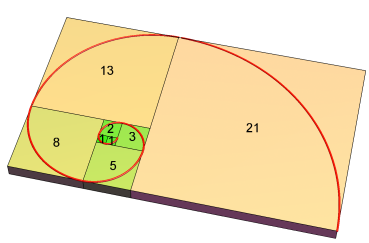


FIGURE 1. The Fibonacci spiral.

21.3.

Problem B: Which number x has the property that if you subtract one of it, then it it is its reciprocal?

Problem C: Which rectangle has the property that if you cut away a square with the length of the smaller side, you get a similar rectangle.

Problem D: Which number is a fixed point of the map $T(x) = 1 + 1/x$?

Problem E: Which number is given by the limit $1/(1+1/(1+1/1+\dots))$?

Problem F: Which number is given by $x = \sqrt{1 + \sqrt{1 + \sqrt{1 + \sqrt{1 + \dots}}}}$

21.4. A regular **pentagon** can be obtained conveniently in the complex plane as the solution set of the equation $z^5 = 1$. The solutions are $e^{2\pi ik/5}$, with $k = 1, 2, 3, 4, 5$.

Problem G: Use this to find the ratio between the diagonal of the pentagon and the side length.

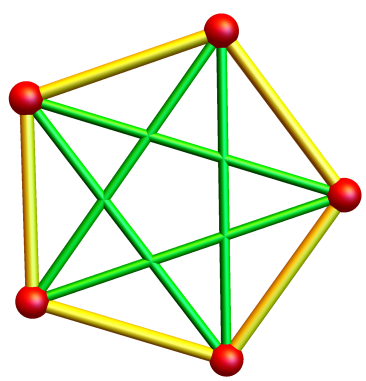


FIGURE 2. The regular Pentagon, when equipped with diagonals becomes the Pentagram. Don't draw it, the figure is known to be magic.

21.5. The **Fibonacci spiral** is obtained by drawing quarter-circle arcs into squares of size F_n. Putting these arcs together produces an approximation of the **golden spiral**.

Problem H: What is $1 + 2\cos(2\pi/5)$? Compute $1 + 2\sin(\pi/10)$.

Problem I: The **triacontahedron** is a thirty-faced convex Catalan solid Each of the 30 congruent rhombic faces is a **golden rhombus** The coordinates of one of these rhombi is $A = (-\phi, 0, 0), B = (0, -1, 0), C = (\phi, 0, 0), D = (0, 1, 0)$. What is the surface area of the triacontahedron?

21.6. Some literature:

M. Livio, The golden ratio, Broadway books, 2002

G.B. Meisner, The golden ratio, Race Point, 2018.

N.N. Vorobev, Fibonacci numbers

R.A. Dunlap The Golden Ratio and Fibonacci Numbers

A.S. Posamentier, I. Lehmann The Fabulous Fibonacci Numbers

R. Knot, Fibonacci Numbers and the Golden Section, 2001

L.E. Sigler, Fibonacci's Liber Abaci, 2003

T. Koshy, Fibonacci and Lucas Numbers with Applications, Wiley, 2001

K. Devlin, The Man of Numbers, Fibonacci's Arithmetic Revolution, Bloomsbury, 2011

G.E.Bergum, A.N.Philippou, A.F.Horadam, Applications of Fibonacci Numbers, Springer, 1998

R.C. Johnson, Fibonacci Numbers and Matrices, 2016

R.Herz-Fischler, A mathematical History of the golden number, 1998

H. Walser, J. Pedersen, P. Hilton, The golden section, 2001

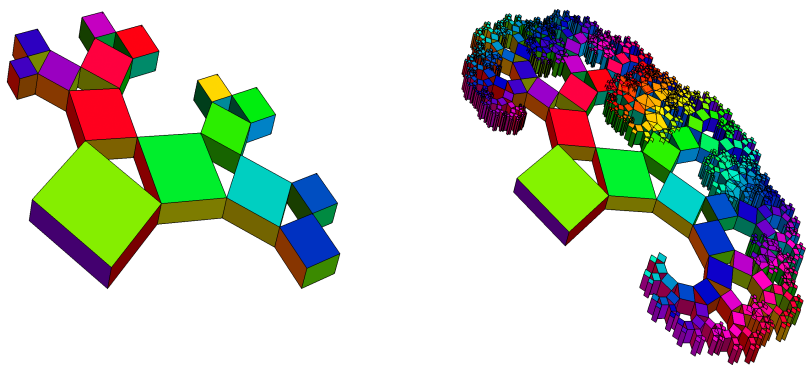


FIGURE 3. Stage 2 and 8 of the tree of Pythagoras with golden ratio.

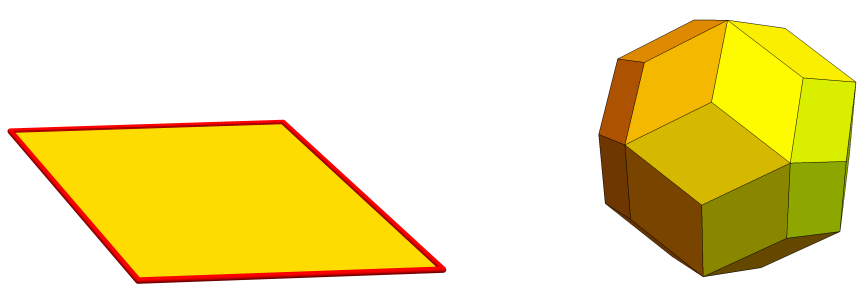


FIGURE 4. The golden rhombus and the triacontahedron.

Homework

Problem 21.1 a) The **sublime triangle** is an isosceles triangles with 36,72,72 degrees. Prove that the ratio between the long and short side is the golden mean. It is the reason why the triangle is also called the **golden triangle**.
b) The **golden Gnomon** is an isosceles triangle with angles 36,36,108 degrees. What is the ratio between the long and short side?

Problem 21.2 Let us experiment with determinants again and define $L(n)$ as the $n \times n$ matrix with -1 in the upper side diagonal and 1 in the diagonal and 1 in the lower side diagonal. For example,

$$L(7) = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} .$$

Compute $\det(L(n))$ using the Laplace expansion and verify that $\det(L(n))/\det(L(n-1)) \to \phi$.

Problem 21.3 Verify that the regular octagon has width which compares to the base as the **silver ratio** $1 + \sqrt{2}$. Why is this number called the silver ratio? What is the third **metallic mean** satisfying $x^2 - 3x = 1$ called?

Problem 21.4 Find a picture of a plant in which some Fibonacci numbers are present. Either sketch it or print out that picture and write in the Fibonacci numbers.

Problem 21.5
a) How is the **lute of Pythagoras** related to the golden mean?
b) How is the **tree of Pythagoras** related to the golden mean?
Look up and draw both objects.

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 22: Stability

Lecture

22.1. A **linear dynamical system** is either a **discrete time dynamical system** $x(t+1) = Ax(t)$ or a **continuous time dynamical systems** $x'(t) = Ax(t)$. It is called **asymptotically stable** if for all initial conditions $x(0)$, the orbit $x(t)$ converges to the origin 0 as $t \to \infty$. The one-dimensional case is clear: the discrete time system $x(t+1) = \lambda x(t)$ has the solution $x(t) = \lambda^t x(0)$ and is asymptotically stable if and only if $|\lambda| < 1$. The continuous time system $x'(t) = \lambda x(t)$ has the solution $x(t) = e^{\lambda t} x(0)$. This is asymptotically stable if and only if $\lambda < 0$. If $\lambda = a + ib$, then $e^{(a+ib)t} = e^{at} e^{ibt}$ shows that asymptotic stability happens if $\mathrm{Re}(\lambda) = a < 0$.

22.2. Let us first discuss the stability for discrete dynamical systems $x(t+1) = Ax(t)$, where A is a $n \times n$ matrix.

22.3.

Theorem: A discrete dynamical system $x(t+1) = Ax(t)$ is asymptotically stable if and only if all eigenvalues of A satisfy $|\lambda_j| < 1$.

Proof. (i) If A has an eigenbasis, this follows from the closed-form solution: assume $||\lambda_k| \leq \lambda < 1$. From $|x(t)| \leq |c_1||\lambda_1|^t + \cdots + |c_n||\lambda_n|^t \leq (\sum_i |c_i|)|\lambda|^t$, we see that the solution approaches 0 exponentially fast.
(ii) The general case needs the **Jordan normal form theorem** proven below which tells that every matrix A can be conjugated to $B + N$, where B is the diagonal matrix containing the eigenvalues and $N^n = 0$. We have now $(B+N)^t = B^t + B(n,1)B^{t-1}N + \cdots + B(n,n)B^{t-n}N^{n-1}$, where $B(n,k)$ are the Binomial coefficients. The eigenvalues of A are the same as the eigenvalues of B. By (i), we have $B^t \to 0$. So, also $A^t \to 0$. $\square$

22.4. In the case of **continuous time dynamical system** $x'(t) = Ax(t)$. the complex eigenvalues will later play an important role but they are also important for discrete dynamical systems.

22.5.

Theorem: A continuous dynamical system is asymptotically stable if and only if all eigenvalues satisfy $\mathrm{Re}(\lambda_j) < 0$.

Proof. We can see this as a discrete time dynamical system with time step $U = e^A$ because the solution e^{At} can be written as U^t. We need therefore that $e^{\lambda_j(U)} < 1$. Which is equivalent to $\mathrm{Re}(\lambda_j) < 0$. □

22.6. A $m \times m$ matrix J is a **Jordan block**, if $Je_1 = \lambda e_1$, and $Je_k = \lambda e_k + e_{k+1}$ for $k = 2, \dots, m$. A matrix is A in **Jordan normal form** if it is block diagonal, where each block is a **Jordan block**. The shear matrix $J = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ is an example of a 2×2 Jordan block.

Theorem: Every $A \in M(n,n)$ is similar to $B \in M(n,n)$ in Jordan normal form.

Proof. A vector v is called a **generalized eigenvector** of the eigenvalue λ of A if $(A - \lambda)^m v = 0$ for some $m > 0$. The smallest integer m for which the relation holds is called the **eigenvector rank** of v. If the eigenvector rank is 1, then v is an actual eigenvector.

22.7. Take an eigenvalue λ of A and pick the maximal m for which there is a generalized eigenvector of rank m. This means that there is a vector v_m such that $(A-\lambda)^m v_m = 0$ but $(A-\lambda)^{m-1} v_m \neq 0$. Now define $v_k = (A-\lambda)^{m-k} v_m$. The vector $v_1 = (A-\lambda)^{m-1} v_m$ is an eigenvector of A because $(A-\lambda)v_1 = 0$. Because $(A-\lambda)v_{k+1} = v_k$ and $(A-\lambda)v_1 = 0$, on the space V spanned by $\mathcal{B} = \{v_1, v_2, \cdots, v_m\}$, the matrix is given by a Jordan block with 0 in the diagonal. This space V is called a **generalized eigenspace** of A. It is left invariant by A.

22.8. Take a generalized eigenspace V defined by an eigenvector v_1 and a generalized eigenspace W defined by another eigenvector w_1 not parallel to v_1. Claim: V and W have no common vector. Proof: Assume $x = \sum_{i=1}^m a_i v_i = \sum_{j=1}^l b_j w_j$. Then $(A-\lambda)x = \sum_{i=1}^m a_i v_{i-1} = \sum_{j=1}^l b_j w_{j-1}$. If $m \neq l$, and say $m < l$ then applying $(A - \lambda)$ $m - 1$ times gives $a_m v_1 = b_{1+l-m} w_{1+l-m}$ but since v_1 is an eigenvector and v_{1-l-m} is not, this does not work. If $m = l$, then we end up with $a_m v_1 = b_m w_1$ which is not possible as v_1 and w_1 are not parallel unless $a_m = b_m = 0$. Now repeat the same argument but only apply $m - 2$ times. This gives $a_{m-1} v_{m-1} = b_{m-1} w_{m-1}$ but this implies v_{m-1}, w_{m-1} are parallel showing again v_m is parallel to w_m We get $a_{m-1} = b_{m-1} = 0$ and eventually see that all $a_k = b_k = 0$.

22.9. The proof of the theorem for $n \times n$ matrices uses induction with respect to n. The case $n = 1$ is clear as a 1×1 matrix is already in Jordan normal form. Now assume that for every $k < n$, every $k \times k$ matrix is similar to a matrix in Jordan normal form. Take a generalized eigenvector v and build the Jordan normal block acting on the generalized eigenspace V. By the previous paragraph, we can find a basis such that $V^\perp$ is invariant. Using induction, there is a Jordan decomposition for A acting on the $V^\perp$. The matrix A has in this bases now a Jordan decomposition with an additional block. □

22.10. This implies the **Cayley-Hamilton theorem**:

> **Theorem:** If p is the characteristic polynomial of A, then $p(A) = 0$.

Proof. It is enough to show this for a matrix in Jordan normal form for which the characteristic polynomial is λ^m. But $A^m = 0$. $\square$

Examples

22.11. You have a **checking account** of one thousand dollars and a **savings account** with one thousand dollars. Every day, you pay 0.001 percent to the bank for both accounts. But the deal is sweet: your checking account will get every day 1000 times the amount in the savings account. Will you get rich?

$$\begin{aligned} c(t+1) &= 0.999c_n + 1000s_n \\ s(t+1) &= 0.999s_n \ . \end{aligned}$$

We will discuss this in class.

22.12. For which constants a is the system $x(t+1) = Ax(t)$ stable?
a) $A = \begin{bmatrix} 0 & a \\ a & 0 \end{bmatrix}$, b) $A = \begin{bmatrix} a & 1 \\ 1 & 0 \end{bmatrix}$, c) $A = \begin{bmatrix} a & a \\ a & a \end{bmatrix}$.
Solution.
a) The trace is zero, the determinant is $-a^2$. We have stability if $|a| < 1$. You can also see this from the eigenvalues, $a, -a$.
b) Look at the trace-determinant plane. The trace is a, the determinant -1. This is nowhere inside the stability triangle so that the system is always unstable.
c) The eigenvalues are $0, 2a$. The system is stable if and only if $|2a| < 1$ which means $|a| < 1/2$.

22.13. In two dimensions, we can see asymptotic stability from the trace and determinant. The reason is that the characteristic polynomial and so the eigenvalues only need the trace and determinant.

> A two dimensional discrete dynamical system has asymptotic stability if and only if $(\mathrm{tr}(A), \det(A))$ is contained in the interior of the **stability triangle** bounded by the lines $\det(A) = 1, \det(A) = \mathrm{tr}(A) - 1$ and $\det(A) = -\mathrm{tr}(A) - 1$.

Proof: Write $T = \mathrm{tr}(A)/2, D = \det(A)$. If $|D| \geq 1$, there is no asymptotic stability. If $\lambda = T + \sqrt{T^2 - D} = \pm 1$, then $T^2 - D = (\pm 1 - T)^2$ and $D = 1 \pm 2T$. For $D \leq -1 + |2T|$ we have a real eigenvalue ≥ 1. The conditions for stability is therefore $D > |2T| - 1$. It implies automatically $D > -1$ so that the triangle can be described shortly as $\boxed{|\mathrm{tr}(A)| - 1 < \det(A) < 1}$.

> For a two-dimensional continuous dynamical system we have asymptotic stability if and only if $\mathrm{tr}(A) < 0$ and $\det(A) > 0$.

Homework

This homework is due on Tuesday, 4/02/2019.

Problem 22.1: Determine whether the matrix is stable for the discrete dynamical system or for the continuous dynamical system or for both: a) $A = \begin{bmatrix} -2 & -3 \\ 3 & -2 \end{bmatrix}$, b) $B = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$, c) $C = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$.

Problem 22.2: True or false? We say A is stable if the origin $\vec{0}$ is asymptotically stable for $x(t+1) = A(x(t))$. Give short explanations:
a) 1 is stable.
b) 0 matrix is stable.
c) a horizontal shear is stable
d) a reflection matrix is stable.
e) A is stable if and only if A^T is stable.
f) A is stable if and only if A^{-1} is stable.
g) A is stable if and only if $A+1$ is stable.
h) A is stable if and only if A^2 is stable.
i) A is stable if $A^2 = 0$.
j) A is unstable if $A^2 = A$.
k) A is stable if A is diagonalizable.

Problem 22.3: a) Check the Cayley-Hamilton theorem for the matrix $A = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$. b) Ditto for a rotation dilation matrix defined by a, b.

Problem 22.4: For which real values k does the drawing rule

$$\begin{aligned} x(t+1) &= x(t) - ky(t) \\ y(t+1) &= y(t) + kx(t+1) \end{aligned}$$

produce trajectories which are ellipses? Write first the system as a discrete dynamical system $v(t+1) = Av(t)$, then look for k values for which the eigenvalues satisfy $|\lambda_k| = 1$.

Problem 22.5: Find the eigenvalues of

$$A = \begin{bmatrix} 0 & a & b & c & 0 & 0 \\ 0 & 0 & a & b & c & 0 \\ 0 & 0 & 0 & a & b & c \\ c & 0 & 0 & 0 & a & b \\ b & c & 0 & 0 & 0 & a \\ a & b & c & 0 & 0 & 0 \end{bmatrix}$$

Where a, b and c are arbitrary constants. Verify that the discrete dynamical system is stable for $|a| + |b| + |c| < 1$.

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 23: Nonlinear systems

Lecture

23.1. We look at **nonlinear differential equations** for differentiable functions f, g:

$$\begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix} = \begin{bmatrix} f(x,y) \\ g(x,y) \end{bmatrix} .$$

In vector form, with $r(t) = [x(t), y(t)]^T$ and vector field $F(x,y) = [f(x,y), g(x,y)]^T$ the differential equation can be written as $r'(t) = F(r(t))$. To say it in the language of vector calculus, we aim to find the **flow lines** of the vector field F. Even having left the linear context, we can still use **linear algebra** to analyze such systems.

23.2. A nonlinear system in population dynamics is the **Murray system**

$$\begin{aligned} x' &= x(6-2x) - xy \\ y' &= y(4-y) - xy . \end{aligned}$$

It is a coupled pair of **logistic systems** which without the xy interaction term would evolve independently of each other. With the interaction, which implements a **competition** situation, we cannot write down a closed-form solutions. Even a computer algebra system is unable to do that for that above Murray system.

23.3. A point (x, y) is called an **equilibrium point** if $F(x, y) = 0$. It is helpful to look for **x-nullclines**, points where $f(x, y) = 0$ and also for y-nullclines, where $g(x, y) = 0$. On x-nullclines, the vector field is vertical, while on y-nullclines, the vector field is horizontal. We are already familiar with the **Jacobian matrix** $A = dF(x, y) = \begin{bmatrix} f_x & f_y \\ g_x & g_y \end{bmatrix}$ which is the **linearization**. Given an equilibrium point (x_0, y_0), we get the linearized system $r'(t) = Ar(t)$. We call the equilibrium point **asymptotically stable** if its linearization is asymptotically stable. Here is a simple principle which helps to analyze the flow phase space.

Theorem: If two solution curves cross, we have an equilibrium point.

Proof. This follows from the fact that the system has a unique solution. If two solutions intersect, then there are two solution curves through this point. □

23.4. To analyze a non-linear system, we find the nullclines, the equilibrium points, linearize the system near each equilibrium point, then draw the phase portraits near the equilibrium points and finally connect the dots to see the global phase portrait. Let us do that in the case of the Murray system. Since $f(x,y) = x(6-2x-y)$, the x-nullclines consist of two lines $x = 0$ and $y = 6-2x$. Since $g(x,y) = y(4-x-y)$ the y-nullclines are $y = 0$ and $y = 4-x$. The equilibrium points are $(0,0), (3,0), (0,4), (2,2)$.

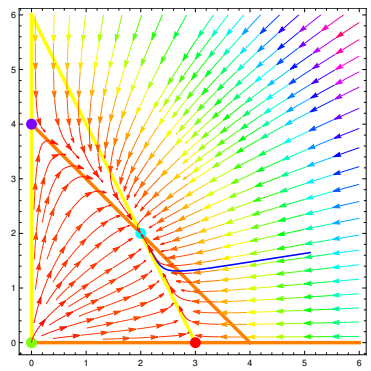


FIGURE 1. The Murray system is a coupled logistic system which describes a **competition model**.

Equilibrium	Jacobian	Eigenvalues	Type
(0,0)	$\begin{bmatrix} 6 & 0 \\ 0 & 4 \end{bmatrix}$	$\lambda_1 = 6,\ \lambda_2 = 4$	Unstable source
(3,0)	$\begin{bmatrix} -6 & -3 \\ 0 & 1 \end{bmatrix}$	$\lambda_1 = -6,\ \lambda_2 = 1$	Hyperbolic saddle
(0,4)	$\begin{bmatrix} 2 & 0 \\ -4 & -4 \end{bmatrix}$	$\lambda_1 = 2,\ \lambda_2 = -4$	Hyperbolic saddle
(2,2)	$\begin{bmatrix} -4 & -2 \\ -2 & -2 \end{bmatrix}$	$\lambda_i = -3 \pm \sqrt{5}$	Stable sink

EXAMPLES

23.5. In the 1920's, the **Volterra-Lodka systems** appeared:

$$\begin{aligned} \dot{x} &= 0.4x - 0.4xy \\ \dot{y} &= -0.1y + 0.2xy \end{aligned}$$

has equilibrium points $(0,0)$ and $(1/2,1)$. It describes a predator-prey situation like for example a shrimp-shark population. The shrimp population $x(t)$ becomes smaller with more sharks. The shark population grows with more shrimp. Volterra explained so first the oscillation of fish populations in the mediterranean sea

23.6. Given a function $H(x,y)$ of two variables it defines a system

$$\begin{aligned} \dot{x} &= \partial_y H(x,y) \\ \dot{y} &= -\partial_x H(x,y) \end{aligned}$$

called Hamiltonian systems. An example is the **pendulum** $H(x,y) = y^2/2 - \cos(x)$ appearing in the homework, in which x denotes the angle between the pendulum and y-axes, and y is the angular velocity, the value $\sin(x)$ is the potential energy. Hamiltonian

systems preserve energy $H(x,y)$ because $\frac{d}{dt}H(x(t),y(t)) = \partial_x H(x,y)\dot{x} + \partial_y H(x,y)\dot{y} = \partial_x H(x,y)\partial_y H(x,y) - \partial_y H(x,y)\partial_x H(x,y) = 0$. Orbits stay on level curves of H.

23.7. Lienhard systems are differential equations of the form $\ddot{x}+\dot{x}F'(x)+G'(x) = 0$. With $y = \dot{x} + F(x), G'(x) = g(x)$, this gives

$$\begin{aligned} \dot{x} &= y - F(x) \\ \dot{y} &= -g(x) \end{aligned}$$

An example is the **Van der Pol equation** $\ddot{x} + (x^2 - 1)\dot{x} + x = 0$ appears in electrical engineering, biology or biochemistry. Since $F(x) = x^3/3 - x, g(x) = x$.

$$\begin{aligned} \dot{x} &= y - (x^3/3 - x) \\ \dot{y} &= -x \ . \end{aligned}$$

Lienhard systems have **limit cycles**. A trajectory always ends up on that limit cycle. This is useful for engineers, who need oscillators which are stable under changes of parameters. One knows: if $g(x) > 0$ for $x > 0$ and F has exactly three zeros $0, a, -a$, $F'(0) < 0$ and $F'(x) \geq 0$ for $x > a$ and $F(x) \to \infty$ for $x \to \infty$, then the corresponding Lienhard system has exactly one stable limit cycle.

23.8. Chaos can occur for systems $\dot{x} = f(x)$ in three dimensions. Here are three examples. They lead already to **strange attractors**.

23.9. The **Roessler system**

$$\begin{aligned} \dot{x} &= -(y + z) \\ \dot{y} &= x + y/5 \\ \dot{z} &= 1/5 + xz - 5.7z \end{aligned}$$

23.10. The **Lorentz system**

$$\begin{aligned} \dot{x} &= 10(y - x) \\ \dot{y} &= -xz + 28x - y \\ \dot{z} &= xy - \frac{8z}{3} \end{aligned}$$

23.11. The **Duffing system** models moving plate: $\ddot{x} + \frac{\dot{x}}{10} - x + x^3 - 12\cos(t) = 0$

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= -y/10 - x + x^3 - 12\cos(z) \\ \dot{z} &= 1 \end{aligned}$$

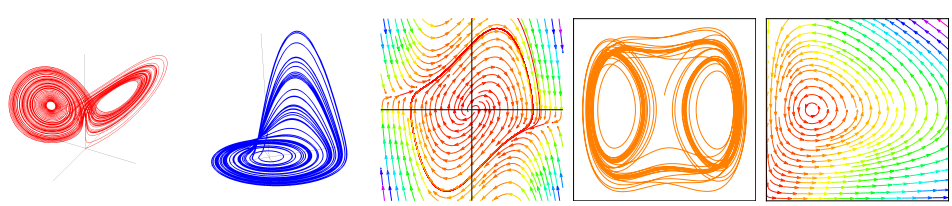


FIGURE 2. Lorentz, Roessler, Vanderpool, Duffing, Volterra

Homework

Due on Tuesday, 4/02/2019. **Analysis** means 1) find nullclines and equilibria 2) determine stability 3) draw phase space 4) list typical trajectories. Some problems adapted from unpublished notes by Otto Bretscher.

Problem 23.1: Analyze the **Volterra-Lodka system** on $x \geq 0, y \geq 0$:

$$\begin{aligned} \frac{dx}{dt} &= 2x + xy - x^2 \\ \frac{dy}{dt} &= 4y - xy - y^2 \end{aligned}$$

Problem 23.2: Analyze the population model on $x \geq 0, y \geq 0$.

$$\begin{aligned} \frac{dx}{dt} &= x(1 - x + 2y - 2) \\ \frac{dy}{dt} &= y(1 - y + 2x - 2) \end{aligned}$$

Problem 23.3: Analyze the **pendulum Hamiltonian system**

$$\begin{aligned} \frac{dx}{dt} &= y \\ \frac{dy}{dt} &= -2\sin(x) \ , \end{aligned}$$

Problem 23.4: Analyze the pendulum with friction

$$\begin{aligned} \frac{dx}{dt} &= y \\ \frac{dy}{dt} &= -2\sin(x) - y \ . \end{aligned}$$

Problem 23.5: Analyze the system

$$\begin{aligned} \frac{dx}{dt} &= x^2 + y^2 - 1 \\ \frac{dy}{dt} &= xy \end{aligned}$$

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 24: Chaos

Seminar

24.1. Simple transformations can produce **chaotic orbits**. Let us experiment! Make sure your calculator is in the "Rad" mode, in which 2π radians mean 360 degrees. You can check the mode by computing $\cos(\pi)$. You should see the result -1.

Figure 1. Rotate the phone by 90 degrees to get the scientific calculator. Be in Rad mode!

Problem A: Push repetitively $\boxed{\cos}$. What do you observe?

Problem B: Don't clear and repeat pushing $\boxed{x^2}$. Observe.

Problem C: Repeat pushing $\boxed{\sqrt{x}}$. What limit do you get?

Problem D: Repeat pushing $\boxed{\sin}$ then $\boxed{1/x}$ button. Observe.

Problem E: Repeat pushing $\boxed{\tan}$ then $\boxed{1/x}$ button. Observe.

24.2. Experiments like what you just did have led **Mitchell Feigenbaum** to the discovery of **universality**. It is now part of **chaos theory**. Feigenbaum played with a calculator and looked for bifurcation points in the sequence of systems like $T(x) = c\sin(\pi x)$ on $[0,1]$. For small c, all orbits converge to a fixed point, then there will be an interval in which there is a periodic point of periodic 2, then 4 etc. This period doubling bifurcations will accumulate at a parameter point in a way which is independent of the map. It would be the same for the map $T(x) = cx(1-x)$ on $[0,1]$.

24.3. Given a **discrete time dynamical system** $x(t+1) = F(x(t))$ or a **continuous time dynamical system** $x'(t) = F(x(t))$, we say that the orbit $x(t)$ shows **sensitive dependence of initial conditions** if the map $U(t)$ which maps $x(0)$ to $x(t)$ has the property that $|dU(t)|$ grows exponentially in the sense that the **Lyapunov exponent** $\gamma(x) = \liminf_{t\to\infty}(1/t)\log|dU(t)|$ is positive.

Remember that for a matrix A, we defined $|A| = \mathrm{tr}(A^T A)$ which is the length of A if we look at the matrix as a vector. The liminf γ means that $(1/t)\log|dU(t)| \geq \gamma$ for arbitrarily large t and that we have taken the largest γ with that property. For example, $\liminf \sin(t) = -1$. It is a basic fact that if a function $f(t)$ is bounded, then $\liminf_t f(t)$ exists. It follows that if F is differentiable and $|dF|$ is globally bounded, then the Lyaponuov exponent exists. This does not mean that the limit exists. It is a liminf only.

24.4. Look at the system on the square $X = \mathbb{T}^2 = [0,1) \times [0,1)$ with $T(x,y) = (2x+y, x+y) \bmod 1$. The notation $u \bmod 1$ is a number in $[0,1)$ defined by removing the integer part. for example $1.73 \bmod 1 = 0.73$. Mathematically X is the **two-dimensional torus** which is the square on which the left and right side are identified and where the bottom side and top side are identified too. The map is defined by a linear relation but we also wrap things around the torus. We have $dT = A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$. The map T is called the **Arnold cat map**.

24.5. We say that a system shows **chaos** on a bounded subset X of the phase space if X is left invariant and the **metric entropy** $\mu(T) = \int_X \gamma(x)\, dV(x) > 0$. The subset X can be a bounded region of Euclidean space or then a bounded surface $S = r(R)$ which is left invariant by the system and where R is a parameter domain. In that case, $dV(x) = |dr(u)|du$ is the usual measure we take for example to compute length, area or volume.

24.6. A popular choice is the **torus** $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$. It can be visualized as the unit cube $[0,1]\times[0,1]\times\cdots\times[0,1]$, where in each interval the left and right side are identified. The cat map $(x,y) \to (2x+y, x+y)$ is a transformation on the 2-torus $\mathbb{T}^2$. When iterating the map, we just discard any integer parts. For example $T(0.3, 0.5) = (0.1, 0.8)$.

Theorem: If $T(x) = Ax$ is given by a $n \times n$ matrix A with integer coefficients and the eigenvalues of A are all positive, then the corresponding map on the torus $\mathbb{T}^n$ shows chaos. The Lyapunov exponent is $\log|\lambda_1|$ where λ_1 is the largest eigenvalue of A.

24.7. In the case of the cat map, the eigenvalues of A are $(3 \pm \sqrt{5})/2$. The Lyapunov exponent is constant $\log((3+\sqrt{5})/2) = 0.962424\ldots$

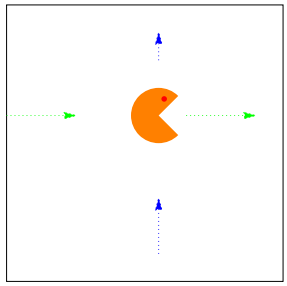


FIGURE 2. Pac-Man is played on a torus. When reaching the right wall, it comes in on the other side. When identifying the left and right side we get a finite piece of cylinder. Now identifying the two ends produces a torus.

FIGURE 3. The cat map $T(x, y) = (2x + y, x + y)$ on $\mathbb{T}^2$. Vladimir Arnold visualized the map with a cat subjected to the dynamics. In the middle we see a set evolve under the Standard map $T(x, y) = (2x - y + 4\sin(x), x)$. To the right, some orbits of the standard map.

24.8. There are many open problems in chaos theory. Also, when dealing with differential equations. For example, nobody has shown that a system like the ABC system shows chaos in the above sense. It is a differential equation on the three dimensional torus.

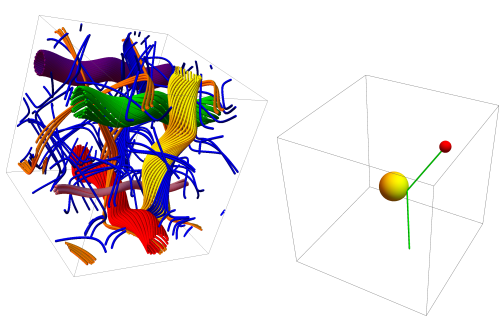


FIGURE 4. The ABC system $x' = A\sin(z) + C\cos(y)$, $y' = B\sin(x) + A\cos(z)$, $z' = C\sin(y) + B\cos(x)$, with $A = 1, B = 1.5, C = 1.5$ is a differential equation on the three dimensional torus. The Sinai billiard to the right too. A ball is a fixed obstacle at which a gas particle reflects with the billiard law. This system is known to be chaotic. It is a model for chaos of a Boltzman gas in the kinetic theory of gases.

Homework

Problem 24.1 a) If $T(x) = 10x \bmod 1$, find $T^5(\pi)$.
b) What is the entropy of the map $T(x) = 10x \bmod 1$?
c) What is the entropy of the Fibonacci map $T(x, y) = (x + y, x) \bmod 1$?

Problem 24.2 Experiment with the map $f(x) = c\cos(e^x))$. There is a threshold c_0 such that for $c > c_0$ is chaotic. Find it approximately.

```
T[x_] := 1.9 Cos[Exp[x]];
ListPlot[NestList[T, 0.3, 100000],PlotRange -> {-1, 1}]
```

Problem 24.3 Find your own two key combination on the calculator which produces chaos. We have seen that the combination $\boxed{1/x}$ and $\boxed{\tan}$ works. There are more. If you have found a good one, plot the orbit using the above mathematica code and submit that.

Problem 24.4 Use the page http://www.dynamical-systems.org/twist/Entropy/Entropy.html from 20 years ago, to compute numerically the entropy of the standard map $T(x, y) = (2x - y + ksin(x), x)$ on the torus with $k = 5.5$. Or use the following Mathematica code to compute an estimate for $k = 5.5$.

```
(* Open Standard map entropy conjecture Entropy(g) >= log(g/2)      *)
pi=N[Pi]; R:=N[2*pi*Random[]];
T[{x_,y_},g_]:=Mod[{2*x-y+g*Sin[x],x},2*pi];   (* Chirikov map       *)
A[{x_,y_},g_]:={{2+g*Cos[x],-1},{1,0}};        (* Jacobian matrix    *)
Lya[{x_,y_},g_,n_]:=Module[{B=B={{1,0},{0,1}},p={x,y},t=0,a},
 Do[B=A[p,g].B;p=T[p,g];a=Abs[B[[1,1]]];If[a>1,B=B/a;t+=Log[a]],{n}];
 (t+Log[Sqrt[Tr[Transpose[B].B]]])/n];          (* Lyapunov exponent *)
entropy[g_,n_,m_]:={Sum[Lya[{R,R},g,n],{m}]/m,N[Log[g/2]]};
entropy[4.0,1000,100] (* g=4, got 100 orbits of length 1000         *)
```

Problem 24.5 A piecewise smooth convex closed curve in the plane defines a **billiard dynamical system** in which the billiard ball reflects at the boundary. An example is a rectangle or a circle. These are examples of billiards which do not show chaos. Look whether you find something about chaotic billiards. There are tables which produce chaos. Find one, then use ruler and compass (or a good eye) to plot some orbits in that table.

Oliver Knill, knill@math.harvard.edu, Math 22b, Harvard College, Spring 2019

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 25: Function Spaces

Lecture

25.1. We have worked so far with $M(n,m)$, the linear space of all $n \times m$ matrices and especially with the Euclidean space $\mathbb{R}^n = M(n,1)$. When working with differential equations, it is necessary to work also with **spaces of functions**. Like vectors, functions can be added, scaled and contain a zero element, the function which is constant 0. From now on, when we speak about a linear space, we mean an **abstract linear space**, a set X which we can add, scale and have a zero element.

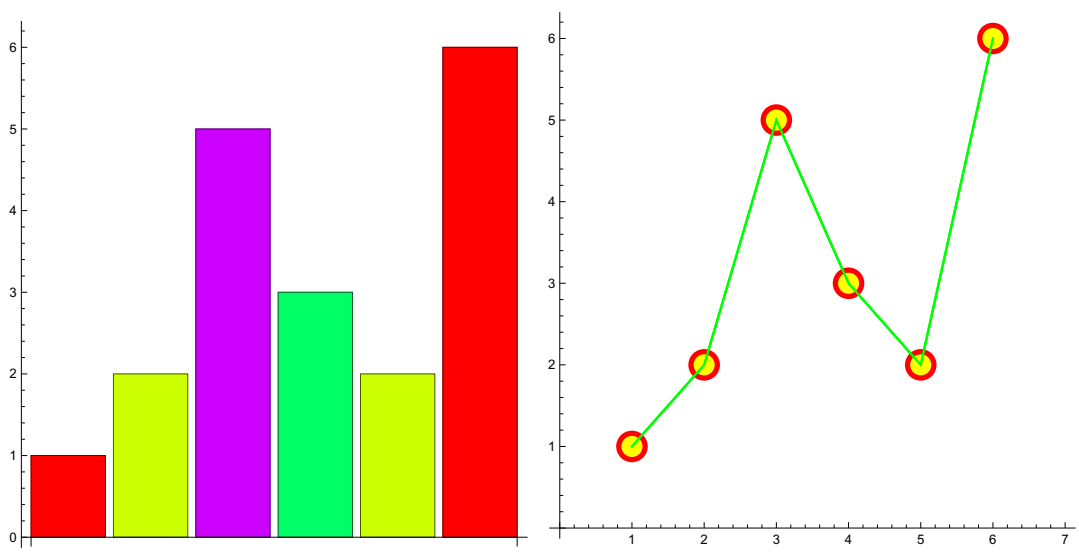


Figure 1. The vector $[1,2,5,3,2,6]$ can be interpreted as a function: just define $f(k) = v_k$. The index k is the input and the output is the value of the function is v_k. By looking at the function, we can visualize and see this 6-dimensional vector.

25.2. The space $C(\mathbb{R})$ of all continuous functions is a linear space. It contains vectors like $f(x) = \sin(x), g(x) = x^3 + 1$ or $h(x) = \exp(x)$. Functions can be added like $(f+h)(x) = \sin(x) + e^x$, they can be scaled like $(7f)(x) = 7\sin(x)$. Any function space also needs to contain the **zero function** $0(x) = 0$ satisfies $(f+0)(x) = f(x)$.

25.3. Why do we want to look at spaces of functions? One of the main reasons for us here is that **solutions spaces of linear systems of differential equations are function spaces.** Another reason is that in probability theory, **random variables** are elements in function spaces. In physics, fields are function spaces. This includes scalar fields, vector fields, wave functions or parametrizations for describing geometric objects like surfaces or curves. Finally, **functions** are a universal language to describe **data**. In figure 1 we see a data set with 6 points. We can not draw the vector in $\mathbb{R}^6$ but we can draw the bar-chart. With many data points, a bar chart can look like the graph of a function. It is a function on the set $\{1,2,3,4,5,6\}$ anf $v_k = f(k)$.

25.4. Here is a general principle to generate linear spaces:

Principle: If X is a set, all maps from X to $\mathbb{R}^m$ form a linear space.

25.5. For example, if $X = \{1,2,3\}$ then the set of all maps from X to $\mathbb{R}$ is equivalent to $\mathbb{R}^n$. With $[f(1), f(2), f(3)]^T$ we get a vector. If $X = \{(1,1),(1,2),(2,1),(2,2)\}$ then the set of all maps from X to $\mathbb{R}$ is equivalent to $M(2,2)$. With $\begin{bmatrix} f(1,1) & f(1,2) \\ f(2,1) & f(2,2) \end{bmatrix}$ we get a matrix. If $X = \mathbb{R}$, we get the set of all maps from X to $\mathbb{R}$. It is a large infinite dimensional space. If $X = \mathbb{R}^2$, we get the set of all functions $f(x,y)$ of two variables. The space of all maps from $\mathbb{R}$ to $\mathbb{R}^2$ is the space of all **parametrized planar curves**. The space of all maps from $\mathbb{R}^2$ to $\mathbb{R}^3$ is the space of all **parametrized surfaces**.

25.6. We can select subspaces of function spaces. For example, the space $C(\mathbb{R})$ of **continuous functions** contains the space $C^1(\mathbb{R})$ of all **differentiable functions** or the space $C^\infty(\mathbb{R})$ of all **smooth functions** or the space $P(\mathbb{R})$ of polynomials. It is convenient to look at $P_n(\mathbb{R})$, the space of all polynomials of degree $\leq n$. Also the space $C^\infty(\mathbb{R},\mathbb{R}^3)$ of all smooth parametrized curves in space is a linear space. Another important space is $C^\infty(\mathbb{T})$ of 2π-periodic smooth functions. They can be seen as functions on the circle $\mathbb{T} = \mathbb{R}/(2\pi\mathbb{Z})$, which is the line in which all points in distance 2π are identified.

25.7. Let us look at the space $P_n = \{a_0 + a_1x + a_2x^2 + \cdots + a_nx^n\}$ of polynomials of degree $\leq n$.

Principle: The space P_n is a linear space of dimension $n+1$.

Proof. It is a linear space because we can add such functions, scale them and there is the zero function $f(x) = 0$. The functions $\mathcal{B} = \{1, x, x^2, x^3, \ldots, x^n\}$ form a basis. First of all, the set $\mathcal{B}$ spans the space P_n. To see that the set is linearly independent assume that $f(x) = a_0 1 + a_1x + a_2x^2 + \cdots + a_nx^n = 0$. By evaluating at $x = 0$, we see $a_0 = 0$. By looking at $f'(0) = 0$, we see that $a_1 = 0$, by looking at $f''(0) = 0$ we see $a_2 = 0$. Continue in the same way and compute the n'th derivative to see $a_n = 0$. $\square$

25.8. As in the space of Euclidean spaces, we can find new linear spaces by looking at the kernel or the image of some transformation T. The most important transformation for us is the **derivative map** $T(f) = f'$. We call it D. So, $D\sin = \cos$ and $Dx^5 = 5x^4$.

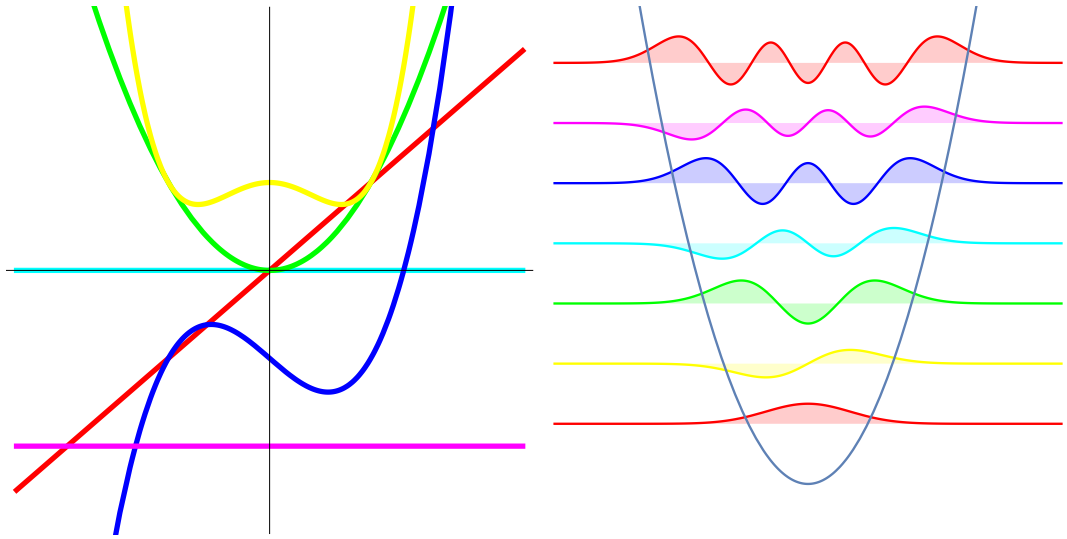


FIGURE 2. Left: graphs of 5 polynomials $f_0(x) = 0$, $f_1(x) = -2$, $f_2(x) = x$, $f_3(x) := x^2$, $f_4(x) := x^3 - x - 1$, $f_5(x) = x^4 - x^2 + 1$. The function $f_0(x)$ is the zero function. The functions $\{f_1, f_2, f_3, f_4, f_5\}$ form a basis of P_4. Right: graphs of eigenfunctions $f_n(x)$ of the harmonic oscillator operator $T = -D^2 + x^2$. The graphs are lifted to have average $(2n+1)$, the n'th energy level.

Theorem: Kernel and image of a linear transformation are linear spaces.

Proof. Let $X = \ker(T)$. To verify that X is a linear space, we check three things: (i) if x, y are in X, then $x + y$ is in X. Proof: If $T(x) = 0, T(y) = 0$, then $T(x + y) = T(x) + T(y) = 0 + 0 = 0$. (ii) if x is in X, then λx in X. Proof: If $T(x) = 0$, then $T(\lambda x) = \lambda T(x) = \lambda 0 = 0$. (iii) We have 0 in X. Proof: $T(0) = 0$. $\square$

25.9. What is the kernel and image of the transformation $Df = f'$ on $C^\infty(\mathbb{R})$? To find the kernel, we look at all functions f which satisfy $Df = f' = 0$. By integration, we see $f = c$ is a constant. So, the nullity of D is 1:

Principle: The kernel $\ker(D) = \{c \,|c \text{ is real}\}$ is one-dimensional.

25.10. To find the image, we want to see which functions f can be reached as $f = Dg$. Given f, we can form $g(x) = \int_0^x f(t)\, dt$. By the fundamental theorem of calculus, we see $Dg = g' = f(x)$.

Principle: The image of D is the entire space $\text{im}(D) = C^\infty$.

25.11. In the next lecture we will learn how to find solutions to differential equations like $f''(x) + 3f'(x) + 2f(x) = 0$. We will write this as an equation $(D^2 + 3D + 2)f = 0$ which means that the solution is the kernel of a transformation $T = D^2 + 3D + 2$. Now, because this is $(D + 2)(D + 1)f = 0$. Solutions can now be obtained by looking at $(D + 1)f = 0$ and $(D + 2)f = 0$, which has solutions $C_1 e^{-x}$ and $C_2 e^{-2x}$. So, the general solution is $f(x) = C_1 e^{-x} + C_2 e^{-2x}$.

Homework

This homework is due on Tuesday, 4/09/2019.

Problem 25.1: Which spaces X are linear spaces?
a) All polynomials of degree 2 or 3.
b) All smooth functions with $f'(1) = 0$.
c) All continuous periodic functions $f(x+1) = f(x)$ with $f(0) = 1$.
d) All functions satisfying $f''(x) - f(x) = 0$.
e) All smooth functions with $\lim_{|x|\to\infty} f'(x) = 0$.
f) All continuous real valued function $f(x, y, z)$ of three variables.
g) All continuous vector fields $F(x, y) = [P(x, y), Q(x, y)]$.
h) All parametrizations $r(t) = r(t + 2\pi) = [x(t), y(t), z(t)]$.
i) All curves $r(t) = [x(t), y(t)]$ in the plane which pass through $(1, 1)$.
j) All 4K movies, maps from $[0, 1]$ to $M(3200, 2400)$.

Problem 25.2: A polynomial $p(x, y)$ is of degree n, if the largest term $a_{kl}x^k x^l$ satisfies $k + l = n$. For example, $f(x, y) = 3x^4y^5 + xy + 3$ has degree 9. a) What is the dimension of the set of polynomials of degree less than 3? b) write down a basis. c) find a formula for the dimension of the space of all polynomials of degree n?

Problem 25.3: The linear map $Df(x) = f'(x)$ is an example of a **differential operator**. As it has a kernel, there is no unique inverse. One inverse is $Sf(x) = D^{-1}f(x) = \int_0^x f(t)\, dt$.
a) Evaluate $D\sin, D\cos, D\tan, S1/(1 + x^2), S\tan$.
b) Find an eigenfunction f of D to the eigenvalue -22.
c) Verify that if f is an eigenfunction of D to the eigenvalue 2, then f is also an eigenfunction of $D^4 - 2D + 22$. What is the eigenvalue?

Problem 25.4: a) Find a basis for the kernel of D^3 on the linear space P of polynomials.
b) Find the image $D^3 + D + 1$ on the linear space P. c) Find the kernel of $Af = (D - \sin(t))f(t)$ on $C^\infty(\mathbb{T})$.

Problem 25.5: a) Solve $D^3 f = 0$ with the additional condition $f(0) = 3, f'(0) = 1, f''(0) = 2$. b) Solve $D^3 f = \cos(x)$ with $f(0) = 3, f'(0) = 1, f''(0) = 2$.

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 26: Operators

Lecture

26.1. A linear map on a function space X is called an **operator**. An example is $Df(x) = f'(x)$ on $X = C^\infty$. More generally, if $p(x)$ is a polynomial, we can look at $p(D)$ on X and look at differential equations $p(D)f = g$ for an unknown function. This is similar as we solved equations $Ax = b$. Now, the operator $p(D)$ plays the role of the matrix A, and f, g replace vectors x and b.

26.2. With the polynomial $p(x) = x^3 + x$ and $g(x) = x^2 + \sin(x)$ we get the problem $p(D)f = g$ which is $f'''(x) + f' = x^2 + \sin(x)$. The equation $p(D)f = g$ is the analogue of an equation $Ax = b$. For example, if $Df = g$, then f is the **anti-derivative** of g. There is a one-dimensional solution space.

Theorem: $p(D)f = g$ has a $\deg(p)$ dimensional solution space.

26.3. If $g = 0$, we get $p(D)f = 0$ and deal with the kernel of a linear transformation and so a linear space. In general, as in the case $Ax = b$, the solution space is **affine**, it is a translated linear space. For example, the harmonic oscillator $(D^2 + 9)f = 0$ has the solution $C_1 \cos(3t) + C_2 \sin(3t)$. To prove the theorem, we need a lemma:

Lemma: $f(t) = e^{\lambda t}(C + \int_0^t e^{-\lambda s} g(s)\, ds)$ solves $f' - \lambda f = g$.

Proof. We just check that it solves the equation. □

26.4. In other words, we have an explicit formula for the inverse $(D - \lambda)^{-1}$. In the case $\lambda = 0$, the formula is

$$D^{-1}g = \int_0^t g(s)\, ds + C \ .$$

So, we can see the formula as a generalized integration formula. To the proof:

Proof. The fundamental theorem of algebra implies that $p(x) = c(x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_n)$. In order to solve $p(D)f = g$, we have

$$f = p(D)^{-1}g = c^{-1}(D - \lambda_1)^{-1} \cdots (D - \lambda_n)^{-1} g \ .$$

Now we can invert one of the operators after the other and each integration introduces a new constant. □

26.5. We have compared $Tf = g$ with solving the system of linear equations $Ax = b$. There are some things which go over, there are other things which don't. For example, we don't have a row reduction process to solve the equation $Tf = g$ in infinite dimensions. We also do in general not have a notion of a determinant of D, at least not in an elementary way.

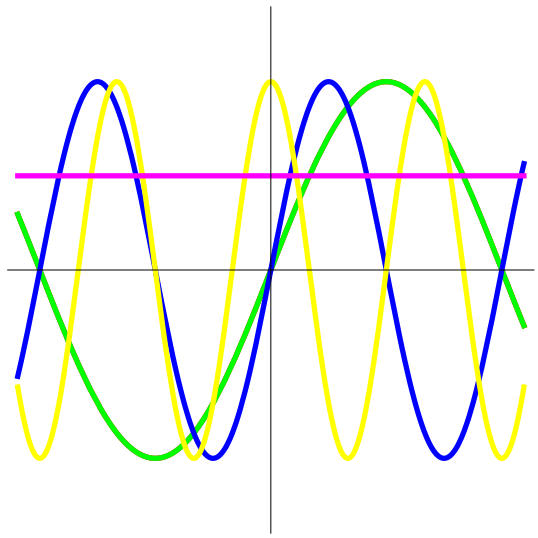


FIGURE 1. Trigonometric polynomials $\cos(nx), \sin(nx)$ and the constant function 1 all are functions in $C^\infty(\mathbb{T})$ that solve the eigenvalue problem $D^2 f = \lambda f$ for some λ. In the homework you show that these are the only ones and that the eigenvalues of D^2 therefore are in a discrete quantized set.

26.6. The method of solving a differential equation by inverting the operator T is called the **operator method**. It is quite powerful. The advantage is that one does not have to think much: just factor T into linear parts, then apply the inversion formulas. There is a bit more convenient way to do that without computers. We will learn an **engineer method** in Unit 27.

EXAMPLES

26.7. To check that an operator is linear, we have to check three things: $T(f+g) = T(f)+T(g)$, $T(\lambda f) = \lambda T(f)$ and $T(0) = 0$ as we did for linear transformations earlier.
a) $T(f)(x) = x^2 f(x-4)$ is linear.
b) $T(f)(x) = f'(x)^2$ is not linear.
c) $T = D^2 + D + 1$ is linear.
d) $T(f)(x) = e^x \int_0^x e^{-t} f(t)\, dt$ is linear.
e) $T(f)(x) = f(\sin(x))$ is linear
f) $T(f)(x) = x^5 f(x^3)$ is linear
g) $T(f)(x) = f(f(x))$ is not linear.
Here is a formal verification for a):
i) $T(0)(x) = x^2 0(x-4) = 0$. ii) $T(f)(x) + T(g)(x) = x^2 f(x-4) + x^2 g(x-4) = x^2[f(x-4)+g(x-4)] = x^2(f+g)(x-4) = T(f+g)(x)$. iii) $T(\lambda f)(x) = x^2 \lambda f(x-4) =$

$\lambda x^2 f(x-4) = \lambda T(f)(x)$.
For a refutation, we need a counter example. For g) for example, take $f(x) = x^3$. Then $T(f)(x) = (x^3)^3 = x^9$. Now, $T(2x^3) = 2(2x^3)^3 = 16x^9$ which is different from $2T(x^3)$.

26.8. The equation $D^3 f = f''' = t^5$ is solved by integrating g three times. This gives $f(x) = x^8/(6*7*8) + C_1 t^2 + C_2 t + C_3$.

26.9. What is the kernel and image of the linear operators $T = D+3$ and $D-2$? Use this to find the kernel of $p(D)$ for $p(x) = x^2 + x - 6$?
The kernel of $T = D+3$ is $\{Ce^{-3x}\}$. The kernel of $D-2$ is $\{Ce^{2x}\}$. The kernel of $(D+3)(D-2)$ contains these two kernels and is $\{Ae^{-3x} + Be^{2x}\}$.

26.10. Verify whether the function $f(x) = e^{-x^2/2}$ is in the kernel of the differential operator $T = D + x$: Solution: We compute $Dg = -xg(x)$. So, $f' + xf(x) = 0$.

26.11. The differential equation $f' - 3f = \sin(x)$ can be written as $Tf = g$ with $T = D-3$ and $g = \sin$. We need to invert the operator T. But we have an inversion formula. We get $f = Ce^{3x} + e^{3x}\int_0^x e^{-3t}\sin(t)\ dt$ which is $Ce^{3x} + (3/10)\sin(x) - \cos(x)/10$. In the proof seminar, we will learn another method to solve this.

Applications

26.12. In quantum mechanics, the operator $P = iD$ is called the **momentum operator** and the operator $Qf(x) = xf(x)$ is called the **position operator**.
Every λ is an eigenvalue of P. The eigenvalue is $e^{i\lambda x}$. What operator is

$$[Q, P] = QP - PQ\ ?$$

You check this in the homework.

26.13. The operator

$$Tf(x) = -f''(x) + x^2 f(x)$$

is called the **energy operator** of the **quantum harmonic oscillator**.
Task: Check that $f(x) = e^{-x^2/2}$ is an eigenfunction of T. What is the eigenvalue?
Solution. Differentiate $f(x) = e^{-x^2/2}$ twice with respect to x. This gives $x^2 f(x) - f(x)$. Since $Tf = f$, the eigenvalue is 1.

26.14. In statistics, one looks at real valued functions f on a probability space Ω. There is a natural linear map from functions to $\mathbb{R}$, which is the **expectation** $\mathrm{E}[f] = \int_\Omega f(x)\ dP(x)$, where P is a probability measure. One usually assumes that these functions are integrable, meaning that $E[|f|] = \int_\Omega |f(x)|\ dP(x) < \infty$ and also that f has the property that $\mathrm{E}[f^2] < \infty$ as this assures that one has finite variance $\mathrm{Var}[f] = \mathrm{E}[(f-m)^2]$. Which of the maps "expectation" or "variance" is linear? Answer: the expectation is a linear map, the variance is not. While $\mathrm{Var}[0] = 0$, it is not true in general that $\mathrm{Var}[f+g] = \mathrm{Var}[f] + \mathrm{Var}[g]$ (an example is $\Omega = [-1,1]$, $f(x) = x, g(x) = x^3$ then $\mathrm{Var}[f] = \int_{-1}^1 x^2\ dx = 2/3$ and $\mathrm{Var}[g] = \int_{-1}^1 x^6\ dx = 2/7$ and $\mathrm{Var}[f+g] = \int_{-1}^1 (x+x^3)^2\ dx = 184/105$. Even easier to check is that $\mathrm{Var}[2f] = 4\mathrm{Var}[f]$ which verifies that the variance does not scale linearly.

Homework

This homework is due on Tuesday, 4/09/2019.

Problem 26.1: Which of the following are linear transformations?
a) $T(f)(x) = f'''(x)$ on $X = C^\infty(\mathbb{R})$.
b) $T(f)(x) = f'''(x) + 1$ on $X = C^\infty(\mathbb{R})$.
c) $T(f)(x) = x^3 f(x)$ on $X = C^\infty(\mathbb{R})$.
d) $T(f)(x) = f(x)^2$ on $X = C^\infty(\mathbb{R})$.
e) $T(f)(x) = f(0) + \int_0^x f(x)\, dx$ on $X = C(\mathbb{R})$.
f) $T(f)(x) = f(x+2)$ on $X = C^\infty(\mathbb{R})$.
g) $T(f)(x) = \int_{-1}^1 f(x-s)s^2\, ds$ on $C(\mathbb{R})$.
h) $T(f)(x) = f'(x^3) + f(3)$ on $C^\infty(\mathbb{R})$.
i) $T(f)(x) = f'(x^2) - f(5) + 1$ on $C^\infty(\mathbb{R})$.
j) $T(f)(x) = f^2(x)$ on $C(\mathbb{R})$.

Problem 26.2: a) Solve the differential equation $f'' - 5f' + 6f = x^2$ by applying the inversion formula.
b) Solve the differential equation $f'' + f = 0$ using the inversion formula (it will be complex).

Problem 26.3: a) Solve the differential equation $f' - f = e^x$ using the inversion formula.
b) Now find the solution with $f(0) = 4$.

Problem 26.4: The operator $Tf = D^2$ will play an important role throughout the rest of the course. It is the negative of the energy operator P^2. Since $Te^{\lambda x} = \lambda^2 e^{\lambda x}$ we see that every real λ is an eigenvalue of T on $C^\infty(\mathbb{R})$. This completely changes when we look at the same operator on the space $C^\infty(\mathbb{T})$ of 2π periodic functions. Instead of having a continuum of eigenvalues, we will get a discrete set of values. Find all the eigenvalues and eigenvectors of D^2!

Problem 26.5: What is the commutator of the momentum and position operator in quantum mechanics? Remember that $Pf(x) = if'(x)$ and $Qf(x) = xf(x)$. What is $[Q, P] = QP - PQ$? Your result will give the **Heisenberg uncertainty relations**. It has important consequences like that we can not measure both position as well as momentum at the same time.

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 27: Cookbook

Seminar

27.1. There is something nice about **recipes**. They relax, produce **robust** and **reliable** results, boost **confidence** and build a fertile ground for creativity. An important principle at the heart of the learning process is now also appreciated in artificial intelligence:

> Used once, it is a trick. Used twice, it has become method.

27.2. When learning something it is good to follow rules at first. **Understanding** happens in different levels: **knowing** the terms, being able to **use** it, **understanding** its mechanism, be able to **teach** it, and finally to **extend** it. When speaking a new language which can be a programming language too, we first "parrot" to get a feel for the language. It does not make much sense to start reading the grammar book or to read the language specification before one has tried a few examples.

27.3. The same applies when solving mathematical problems or even when building theories. There are rules which help to build theory. There are rules which help to find proofs. There are methods to be creative. In this seminar, we want to illustrate in the context of differential equations, how powerful a "solution manual" can be. This **cookbook** not only allows us to understand the operator method better; it also enables us to solve differential equations in a fraction of the time.

27.4. If you have never been taught to solve a differential equation like $f''(x)+f(x) = x$ you have to be ingenious. You might try to push the **method of separation of variables** further, you might try Taylor expansions and find relations of the Taylor coefficients, you might to find a solution by trial and error. Why not just try $f(x) = x$ for example? Wow, bingo we were lucky. This guess is actually a solution because $f'' = 0$. But this "trial and error" needs in general quite a bit of luck. For $f''(x)+f(x) = x^3$, the function $f(x) = x^3$ no more solves the problem and we have to find another way.

27.5. We have seen in the last lecture a general method to solve differential equations like $f'' - 2f' + f = x^2$. We sometimes use also the variable t instead of x. This is on purpose. If we use t, we indicate that t is **time**. If we use x, we indicate that x is **position**. It does not matter. You need to be aware that we can deal with any variable. Let us review first the operator method.

A) Write the equation in operator form $p(D)f = g$.
B) Factor the polynomial p to get $(D-\lambda_1)\cdots(D-\lambda_n)f = g$.
C) Invert each factor using $(D-\lambda)^{-1}(g)(x) = e^{\lambda x}(\int_0^x e^{-\lambda s} g(s)\, ds + C)$.

27.6. A) In the example, the operator rewrite is $(D^2 - 2D + 1)f = x^2$.
b) Factored, it is $(D-1)^2 f = x^2$. B) Invert once to get $(D-1)f = e^x(\int_0^x e^{-s} s^2\, ds + C_1)$ which is (integration by parts) $= C_1 e^x + (2e^x - x^2 - 2x - 2)$. Now invert again to get $f = e^x((C_1+2)x + C_2 - 6) + x^2 + 4x + 6$. Here is the cookbook method

A) Solve the **homogeneous solution** f_h where the right hand side is 0.
B) Find a **particular solution** f_p using the cookbook.
C) The **general solution** is the sum $f_h + f_p$.

27.7. Example: A) Since the eigenvalue 1 appears twice, we have $f_h(x) = C_1 e^x + C_2 x e^x$. B) For the particular solution, try $f = Ax^2 + Bx + C$. It gives $f' = 2Ax + B$, $f'' = 2A$. Plug this into the equation to get $f'' - 2f' + f = 2A - 2(2Ax + B) + Ax^2 + Bx + C = Ax^2 + (B - 4A)x + C + 2A - 2B = x^2$. Comparing coefficients gives $A = 1, B = 4, C = 6$ so that $f_p = x^2 + 4x + 6$. C) The general solution is $x^2 + 4x + 6 + C_1 e^x + C_2 x e^x$.

Problem A: Do as many examples as possible from the list below. Try first without looking at the solution, then compare.

27.8. The "recipe" for finding solutions is to **feed in the same class of functions which you see on the right** and **multiply with x when stuck**.

$\lambda_1 \neq \lambda_2$ real	$f_h = C_1 e^{\lambda_1 x} + C_2 e^{\lambda_2 x}$
$\lambda_1 = \lambda_2$ real	$f_h = C_1 e^{\lambda_1 x} + C_2 x e^{\lambda_1 x}$
$\lambda_1 = ik, \lambda_2 = -ik$	$f_h = C_1 \cos(kx) + C_2 \sin(kx)$
$\lambda_1 = a + ik, \lambda_2 = a - ik$	$f_h = C_1 e^{ax} \cos(kx) + C_2 e^{ax} \sin(kx)$

27.9. Here are examples how to get the particular solution Ansatz:

$g(x) = a$	$f_p(x) = A$ constant
$g(x) = ax + b$	$f_p(x) = Ax + B$
$g(x) = ax^2 + bx + c$	$f_p(x) = Ax^2 + Bx + C$
$g(x) = a\cos(bx)$	$f_p(x) = A\cos(bx) + B\sin(bx)$
$g(x) = a\sin(bx)$	$f_p(x) = A\cos(bx) + B\sin(bx)$

$g(x) = a\cos(bx)$ with $p(D)g = 0$	$f(x) = Ax\cos(bx) + Bx\sin(bx)$
$g(x) = a\sin(bx)$ with $p(D)g = 0$	$f(x) = Ax\cos(bx) + Bx\sin(bx)$
$g(x) = ae^{bx}$	$f(x) = Ae^{bx}$
$g(x) = ae^{bx}$ with $p(D)g = 0$	$f(x) = Axe^{bx}$

EXAMPLE 1: $f'' = \cos(5x)$

This is of the form $D^2 f = g$ and can be solved by inverting D which is integration: integrate a first time to get $Df = C_1 + \sin(5x)/5$. Integrate a second time to get $f = C_2 + C_1 x - \cos(5x)/25$ This is the operator method in the case $\lambda = 0$.

EXAMPLE 2: $f' - 2f = 2x^2 - 1$

This homogeneous differential equation $f' - 2f = 0$ is hardwired to our brain. We know its solution is Ce^{2x}. To get a particular solution, try $f(t) = Ax^2 + Bx + C$. We have to compare coefficients of $f' - 2f = -2Ax^2 + (2A - 2B)x + B - 2C = 2x^2 - 1$. We see that $A = -1, B = -1, C = 0$. The special solution is $-x^2 - x$. The complete solution is $f = -x^2 - x + Ce^{2x}$

EXAMPLE 3: $f' - 2f = e^{2x}$

In this case, the right hand side is in the kernel of the operator $T = D - 2$ in equation $T(f) = g$. The homogeneous solution is the same as in example 2, to find the inhomogeneous solution, try $f(x) = Axe^{2x}$. We get $f' - 2f = Ae^{2x}$ so that $A = 1$. The complete solution is $f = xe^{2x} + Ce^{2x}$

EXAMPLE 4: $f'' - 4f = e^t$

To find the solution of the homogeneous equation $(D^2 - 4)f = 0$, we factor $(D-2)(D+2)f = 0$ and add solutions of $(D - 2)f = 0$ and $(D + 2)f = 0$ which gives $C_1 e^{2t} + C_2 e^{-2t}$. To get a special solution, we try Ae^t and get from $f'' - 4f = e^t$ that $A = -1/3$. The complete solution is $f = -e^t/3 + C_1 e^{2t} + C_2 e^{-2t}$

EXAMPLE 5: $f'' - 4f = e^{2t}$

The homogeneous solution $C_1 e^{2t} + C_2 e^{-2t}$ is the same as before. To get a special solution, we can not use Ae^{2t} because it is in the kernel of $D^2 - 4$. We try Ate^{2t}, compare coefficients and get $f = te^{2t}/4 + C_1 e^{2t} + C_2 e^{-2t}$

EXAMPLE 6: $f'' + 4f = e^t$

The homogeneous equation is a harmonic oscillator with solution $C_1 \cos(2t) + C_2 \sin(2t)$. To get a special solution, we try Ae^t compare coefficients and get $f = e^t/5 + C_1 \cos(2t) + C_2 \sin(2t)$

EXAMPLE 7: $f'' + 4f = \sin(t)$

The homogeneous solution $C_1 \cos(2t) + C_2 \sin(2t)$ is the same as in the last example. To get a special solution, we try $A \sin(t) + B \cos(t)$ compare coefficients to get $f = \sin(t)/3 + C_1 \cos(2t) + C_2 \sin(2t)$

EXAMPLE 8: $f'' + 4f = \sin(2t)$

The solution $C_1 \cos(2t) + C_2 \sin(2t)$ is as before. To get a special solution, we can not try $A \sin(t)$ because it is in the kernel of the operator. We try $At \sin(2t) + Bt \cos(2t)$ instead and compare coefficients $f = t \sin(2t)/16 - t \cos(2t)/4 + C_1 \cos(2t) + C_2 \sin(2t)$

EXAMPLE 9: $f'' + 8f' + 16f = \sin(5t)$

The homogeneous solution is $C_1 e^{-4t} + C_2 t e^{-4t}$. To get a special solution, we try $A \sin(5t) + B \cos(5t)$ compare coefficients and get $f = -40 \cos(5t)/41^2 + -9 \sin(t)/41^2 + C_1 e^{-4t} + C_2 t e^{-4t}$

EXAMPLE 10: $f'' + 8f' + 16f = e^{-4t}$

The homogeneous solution is still $C_1 e^{-4t} + C_2 t e^{-4t}$. To get a special solution, we can not try e^{-4t} nor te^{-4t} because both are in the kernel. Add another t and try with $At^2 e^{-4t}$. $f = t^2 e^{-4t}/2 + C_1 e^{-4t} + C_2 t e^{-4t}$

EXAMPLE 11: $f'' + f' + f = e^{-4t}$

By factoring $D^2 + D + 1 = (D - (1 + \sqrt{3}i)/2)(D - (1 - \sqrt{3}i)/2)$ we get the homogeneous solution $C_1 e^{-t/2} \cos(\sqrt{3}t/2) + C_2 e^{-t/2} \sin(\sqrt{3}t/2)$. For a special solution, try Ae^{-4t}. Comparing coefficients gives $A = 1/13$. $f = e^{-4t}/13 + C_1 e^{-t/2} \cos(\sqrt{3}t/2) + C_2 e^{-t/2} \sin(\sqrt{3}t/2)$

Homework

Problem 27.1 Find the general solution of $f'' + 225f = e^{44t} + t$.

Problem 27.2 Find the general solution of $f'' + 225f = \cos(15t)$.

Problem 27.3 Find the general solution of $f'' + 225f = \cos(10t)$.

Problem 27.4 Find the general solution to $f'' - 10f' + 25f = 4e^{5t}$.

Problem 27.5 Find the general solution to $f''' - 2f'' - f' + 2f = e^t + e^{-t}$.

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 28: Checklist for Second Hourly

Definitions

- ☐ **Characteristic polynomial of A** $p(\lambda) = \det(A - \lambda)$
- ☐ **Orthonormal eigenbasis** An eigenbasis which is orthonormal
- ☐ **Orthogonal complement in $\mathbf{R}^n$** space $V^\perp = \{v \in \mathbf{R}^n \mid v \text{ perpendicular to } V\}$
- ☐ **Transpose matrix** $A^T_{ij} = A_{ji}$. Transposition switches rows and columns.
- ☐ **Symmetric matrix** $A^T = A$ and **skew-symmetric** $A^T = -A$, are both normal
- ☐ **Structure of p:** $p(\lambda) = (-\lambda)^n + \mathrm{tr}(A)(-\lambda)^{n-1} + \cdots + \det(A)$
- ☐ **Eigenvalues and eigenvectors** $Av = \lambda v, v \neq 0$.
- ☐ **Factorization. Have** $p(\lambda) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda)$ **with roots** λ_i**.**
- ☐ **Algebraic multiplicity** is k, if eigenvalue appears exactly k times.
- ☐ **Geometric multiplicity** the dimension of the kernel of $A - \lambda I_n$
- ☐ **Simple spectrum** The property of having all eigenvalues to be different.
- ☐ **Kernel and eigenspace** $\ker(A - \lambda)$ is eigenspace.
- ☐ **Eigenbasis** Basis which consists of eigenvectors of A
- ☐ **Complex numbers** $z = x + iy = |z| \exp(i\,\theta) = r\cos(\theta) + ir\sin(\theta)$
- ☐ **n-Root of 1** $e^{2\pi ik/n}$, $k = 0, \ldots, n-1$ solve $\lambda^n = 1$ (are on regular n-gon)
- ☐ **Discrete dynamical system** $\vec{v}(t+1) = A\vec{v}(t)$ with initial condition $v(0)$ known
- ☐ **Continuous dynamical system** $\vec{v}'(t) = A\vec{v}(t)$ with initial condition $v(0)$ known
- ☐ **Asymptotic stability** $A^n\vec{x} \to 0$ for all $\vec{x}$ (equivalent $|\lambda_i| < 0$)
- ☐ **Markov matrix** Non-negative entries and the sum in each column is 1.
- ☐ **Mother of all ODEs** $f' = \lambda f$ has the solution $f(t) = f(0)e^{\lambda t}$.
- ☐ **Father of all ODEs** $f'' + c^2 f = 0$ has the solution $f(t) = f(0)\cos(ct) + f'(0)\sin(ct)/c$.
- ☐ **Linear operator** $T : X \to X$, where X is a linear space like a function space.

Theorems:

- ☐ **Wiggle theorem** Perturbations { symmetric matrices} can give simple spectrum.
- ☐ **Diagonalization** Matrices with simple spectrum can be diagonalized.
- ☐ **Spectral theorem** A symmetric matrix can be diagonalized with orthogonal S.
- ☐ **Spectral theorem** A normal matrix can be diagonalized with unitary S.
- ☐ **Jordan normal form theorem** Any matrix can be brought into Jordan normal form
- ☐ **Fundamental Theorem of Algebra** $p(\lambda)$ has exactly n roots
- ☐ **Asymptotic stability for discrete** $\Leftrightarrow |\lambda_i| < 1$ for all i.
- ☐ **Asymptotic stability for continuous** $\Leftrightarrow \mathrm{Re}(\lambda_k) < 0$ for all k.

Number of eigenvalues A $n \times n$ matrix has exactly n eigenvalues.
Inequality The geometric multiplicity of an eigenvalue is $\leq$ algebraic multiplicity.
Perron-Frobenius Markov matrices have a simple eigenvalue 1.
Markov processes The product of Markov matrices is Markov.
Solution space The ODE $p(D)f = g$ has a $\deg(p)$ dimensional solution space.

Properties:

Circular matrix $Q = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$ has eigenvalues $\lambda_k = e^{2\pi i k/5}$ with $k = 0, 1, 2, \ldots, 4$ and eigenvectors $v_k = [1, \lambda_k, \lambda_k^2, \lambda_k^3, \lambda_k^4]^T$.
Diagonalization S has eigenvectors of A in columns, $B = S^{-1}AS$ is diagonal
Nontrivial kernel $\Leftrightarrow \det(A) = 0 \Leftrightarrow$ have eigenvalue 0
Orthogonal Matrices A have eigenvalues of length 1.
Determinant is Product of eigenvalues. $\det(A) = \lambda_1 \cdot \cdots \cdot \lambda_n$
Trace is Sum of eigenvalues. $\mathrm{tr}(A) = \lambda_1 + \cdots + \lambda_n$
Geometric Multiplicity of $\lambda \leq$ Algebraic Multiplicity of λ
All different Eigenvalues $\Rightarrow$ can diagonalize.
Symmetric matrix Have real eigenvalues, can diagonalize with orthogonal S!
Eigenvalues of A^T agree with eigenvalues of A (same $p(\lambda)$)
Rank of A^T is equal to the rank of A
Reflection at k-dimensional subspace of $\mathbf{R^n}$ has eigenvalues 1 or -1.
Euler $e^{i\theta} = \cos(\theta) + i\sin(\theta)$. Is real for $\theta = k\pi$.
De Moivre $z^n = \exp(in\theta) = \cos(n\theta) + i\sin(n\theta) = (\cos(\theta) + i\sin(\theta))^n$
Number of eigenvalues A $n \times n$ matrix has exactly n eigenvalues
Power A^k has eigenvalue λ^k if A has eigenvalue λ. Same eigenvector!
Eigenvalues of $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ are $\lambda_\pm = \mathrm{tr}(A)/2 \pm \sqrt{(\mathrm{tr}(A)/2)^2 - \det(A)}$
Eigenvectors of $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ with $c \neq 0$ are $v_\pm = \begin{bmatrix} \lambda_\pm - d, c \end{bmatrix}^T$.
Rotation-Dilation Matrix $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, eigenvalues $a \pm ib$
Rotation-Dilation Matrix has eigenvectors $[\pm i, 1]^T$
Permutation Matrix $Ae_i = e_{i+1}$, roots are on unit circle on regular polygon
Rotation-Dilation Matrix linear stable origin if and only if $|\det(A)| < 1$.
Similarity A, B with simple eigenvalues are similar if eigenvalues are the same.
Similarity A, B both symmetric are similar if eigenvalues are the same
Similarity If A, B are similar, then eigenvalues, trace and det are the same
Similarity If some λ_j of A^k has different geometric multiplicity, then not similar
Spectral Theorem Symmetric and normal matrices are diagonalizable.
Simple spectrum Matrices with simple spectrum are diagonalizable.
Cyclic matrices $Ae_i = e_{i+1}$ and $Ae_n = e_1$ has eigenvalues λ satisfying $\lambda^n = 1$.

- [ ] **Sum of cyclic matrices** Can determine all eigenvalues and eigenvectors.
- [ ] **Row sum** If every row sum is constant λ, then λ is an eigenvalue.

Algorithms:

- [ ] **Computing Eigenvalues of A** (factor $p(\lambda)$, trace, non-invertibility)
- [ ] **Use eigenvalues** to compute determinant of a matrix.
- [ ] **Computing Eigenvectors of** A (determining kernel of $A - \lambda$)
- [ ] **Calculate with complex numbers** (add, multiply, divide, take $n-$'th roots)
- [ ] **Computing algebraic and geometric multiplicities**
- [ ] **Diagonalize Matrix** (find the eigen system to produce S)
- [ ] **Decide about similarity** (e.g. by diagonalization, geometric multiplicities)
- [ ] **Solve discrete systems** (use eigenbasis to get closed-form solution)
- [ ] **Solve continuous systems** (use eigen basis to get closed-form solution)
- [ ] **Find orthonormal eigenbasis** (always possible if $A = A^T$)
- [ ] **Important ODE's** Solve $(D - \lambda)f = 0$ or $(D^2 + c^2)f = 0$. Ask Ma or Pa.

Proof seminar:

- [ ] **Rising sea**: A picture of Grothendieck to crack a nut.
- [ ] **Chaos** $T : X \to X$ has entropy limsup of $\int_X (1/n) \log(|dT^n(x)|)dx$.
- [ ] **Matrix forest theorem**. $\det(1 + L)$ is the number of rooted forests.
- [ ] **Formula of Binet** $F_{n+1}/F_n \to \phi = (\sqrt{5} + 1)/2$.
- [ ] **Cookbook** Solve homogeneous system, find particular solution, then add.
- [ ] **Cat map** on torus $T(x, y) = (2x + y, x + y)$.
- [ ] **Standard map** on torus $T(x, y) = (2x - y + c \sin(x), y)$.

People:

- [ ] **Grothendieck** (rising sea)
- [ ] **von Neumann** (wiggle)
- [ ] **Wigner** (wiggle)
- [ ] **Jordan** (Jordan normal form)
- [ ] **Kac** (Can one hear a drum?)
- [ ] **Gordon-Webb-Wolpert** (No one can't!)
- [ ] **Leonardo Pisano** (Fibonacci Rabbits)
- [ ] **Lyapunov** (Lyapunov exponent)
- [ ] **Arnold** (cat map)
- [ ] **Cayley and Hamilton** (theorem)
- [ ] **Lorentz** (Lorentz system)

1	
2	
3	
4	
5	
6	
7	
8	
9	
10	

Name:

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Total:

Unit 28: Second Hourly Practice

Welcome to the second hourly. It will take place on April 9, 2019 at 9:00 AM sharp in Hall D. Please fill out your name in the box above.

- You only need this booklet and something to write. Please stow away any other material and electronic devices. Remember the honor code.
- Please write neatly and give details. We want to see details, even if the answer should be obvious to you.
- Try to answer the question on the same page. There is also space on the back of each page.
- If you finish a problem somewhere else, please indicate on the problem page so that we find it.
- You have 75 minutes for this hourly.

Problems

Problem 28P.1 (10 points):
a) Prove, using one of the theorems we have seen in the course that any matrix of the form $A - A^T$ is diagonalizable if A is an arbitrary real $n \times n$ matrix.
b) Prove, using one of the theorems we have seen in the course that any matrix of the form AA^T is diagonalizable if A is an arbitrary real $n \times n$ matrix.
c) The matrix

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

has the characteristic polynomial $-\lambda^5 + 1 = 0$. What does the theorem of Cayley-Hamilton state?

Problem 28P.2 (10 points):
Find the characteristic polynomial and the eigenvalues of the following matrices:

a) $\begin{bmatrix} -1 & 3 \\ 4 & 0 \end{bmatrix}$

b) $\begin{bmatrix} 1 & 4 & 3 \\ 1 & 4 & 3 \\ 3 & 4 & 1 \end{bmatrix}$

c) $\begin{bmatrix} 2 & 3 & 1 & 4 \\ 0 & 5 & 2 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 6 \end{bmatrix}$.

Problem 28P.3 (10 points):
Check the boxes which apply for all matrices of the type:

	invertible	diagonalizable	symmetric	real eigenvalues
Projection matrix				
Shear matrix				
Rotation matrix				
Reflection matrix				
$A^2 = 0, A \neq 0$				
Diagonal matrix				

Problem 28P.4 (10 points, each sub problem is 2 points):
a) (5 points) Find the determinant of the following matrix. You have to give reasoning!

$$A = \begin{bmatrix} 22 & 2 & 2 & 2 & 2 \\ 2 & 22 & 2 & 2 & 2 \\ 2 & 2 & 22 & 2 & 2 \\ 2 & 2 & 2 & 22 & 2 \\ 2 & 2 & 2 & 2 & 22 \end{bmatrix}$$

b) (5 points) Find an eigenbasis of A. It does not have to be orthonormal.

Problem 28P.5 (10 points):
Find the possibly complex eigenvalues for the following matrices:
a) (2 points)

$$A = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$$

b) (2 points)

$$B = \begin{bmatrix} 3 & 4 \\ 4 & -3 \end{bmatrix}$$

c) (2 points)

$$C = \begin{bmatrix} 3 & -4 \\ 4 & 3 \end{bmatrix}$$

d) (2 points)

$$D = \begin{bmatrix} 3 & 4 \\ 4 & 3 \end{bmatrix}$$

e) (2 points)

$$E = \begin{bmatrix} 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \\ 2 & 0 & 0 & 0 \end{bmatrix}.$$

Problem 28P.6 (10 points):
Find a closed-form solution of the recursion

$$x_{n+1} = 3x_n - 2x_{n-1}$$

with $x(1) = 1, x(0) = 4$ and determine with the system is stable. Make sure to write this first as a discrete dynamical system $(x(t+1), x(t)) = A(x(t), x(t-1))$ then use the initial condition $[1, 4]$.

Problem 28P.7 (10 points):
When we try to find a closed-form solution of the system

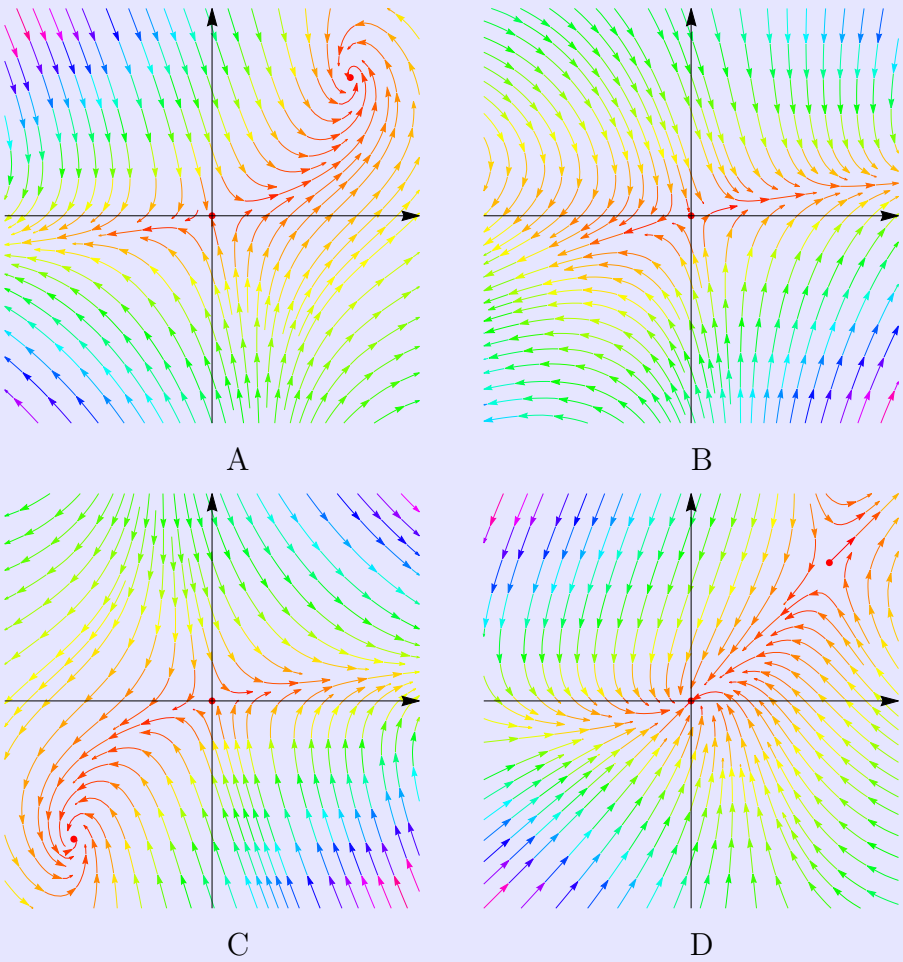
$$\begin{aligned} x' &= x - 2y \\ y' &= 2x - 3y \end{aligned}$$

with $x(0) = 2, y(0) = 1$ we run into trouble. Outline why, then tell how we still can find a closed-form solution. You don't have to do that explicitly. Just show to which matrix A of the system is similar to. What we want to know is whether the system is stable.

Problem 28P.8 (10 points):
We consider the nonlinear system of differential equations

$$\begin{aligned} \frac{d}{dt}x &= x + y - xy \\ \frac{d}{dt}y &= x - 3y + xy \ . \end{aligned}$$

a) (2 points) Find the equilibrium points.
b) (3 points) Find the Jacobian matrix at each equilibrium point.
c) (3 points) Use the Jacobean matrix at an equilibrium to determine for each equilibrium point whether it is stable or not.
d) (2 points) Which of the diagrams A-D is the phase portrait of the system above?



Problem 28P.9 (10 points):
(4 points) a) (5 points) Assume T is a transformation on $C^\infty(\mathbb{T})$, the linear space of 2π-periodic functions on the real line. Which transformations are linear?

Transformation	Check if linear
$Tf(x) = f(x+1)$	
$Tf(x) = f(\cos(x))$	
$Tf(x) = f'(x + \cos(x))$	
$Tf(x) = f(f(x)\cos(x))$	
$Tf(x) = \cos(f(x))$	
$Tf(x) = \cos(x) + f(x)$	

(5 points) The rest are knowledge questions which do not need any reasoning.
a) Which mathematician was a hermite in the later part of his life?
b) What is the entropy of the map $T(x) = 22x$ on $\mathbb{R}^1/\mathbb{Z}^1$.
c) Assume a 5×5 matrix has 2 Jordan blocks. Is it diagonalizable?
d) Find the eigenvector of the eigenvalue problem $Df = 3f$.
e) How big is the dimension of the solution space $(D^5 + D^3 + D)f = 0$?

Problem 28P.10 (10 points):
Find the general solution to the following differential equations:
a) (1 point)

$$f'(t) = 1/(t+1)$$

b) (1 point)

$$f''(t) = e^t + t$$

c) (2 points)

$$f''(t) + f(t) = t + 2$$

d) (2 points)

$$f''(t) - 2f'(t) + f(t) = e^t$$

e) (2 points)

$$f''(t) - f(t) = e^t + \sin(t)$$

f) (2 points)

$$f''(t) - f(t) = e^{-3t}$$

1	
2	
3	
4	
5	
6	
7	
8	
9	
10	

Name:

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Total:

Unit 28: Second Hourly

Welcome to the second hourly. Please don't get started yet. We start all together at 9:00 AM. You can already fill out your name in the box above. Then grab some cereal in the beautiful kitchen (a Povray scene using code of Jaime Vives Piqueres from 2004).

- You only need this booklet and something to write. Please stow away any other material and electronic devices. Remember the honor code.
- Please write neatly and give details. Except when stated otherwise, we want to see details, even if the answer should be obvious to you.
- Try to answer the question on the same page. There is also space on the back of each page and at the end.
- If you finish a problem somewhere else, please indicate on the problem page so that we find it. Make sure we find additional work.
- You have 75 minutes for this hourly.

Problems

Problem 28.1 (10 points):
a) (3 points) What basic fundamental theorem in mathematics is involved to prove that the sum of the algebraic multiplicities of a $n \times n$ matrix is equal to n?
Name of the theorem: (1 point)

State the theorem (1 point) and tell why it implies the statement (1 point) .

b) (3 points) What theorem in linear algebra implies that the sum of the geometric multiplicities of an **orthogonal** $n \times n$ matrix is n so that A is diagonalizable over the complex numbers?
Name of the theorem: (1 point)

State the theorem (1 point) and why does the theorem imply the statement? (1 point)

c) (4 points) What theorem mentioned in this course assures that **any matrix** (not only diagonalizable ones) with eigenvalues 0 or 1 is similar to a matrix in which every entry is 0 or 1.
Name of the theorem: (1 point)

State the theorem (2 points) and why does the theorem imply the statement? (1 point)

Problem 28.2 (10 points):
Match the following matrices with the sets of eigenvalues. You are told that there is a unique match. It is not always necessary to compute all the eigenvalues to do so. You have to give a reason although for each choice (one reason could be that it is the last possible match as you are told there is an exact match). Two points for each sub problem.

Enter 1-5	The matrix
	$A = \begin{bmatrix} -1 & -2 & 8 \\ -7 & -3 & 19 \\ -3 & -2 & 10 \end{bmatrix}$
	$A = \begin{bmatrix} 5 & -9 & -7 \\ 0 & 5 & 2 \\ 0 & 0 & 6 \end{bmatrix}$
	$A = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$
	$A = \begin{bmatrix} 5 & -6 & 0 \\ 6 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix}$
	$A = \begin{bmatrix} 13 & 11 & 13 \\ -2 & -1 & -2 \\ -8 & -7 & -8 \end{bmatrix}$

1) $\{3, 2, 1\}$.

2) $\{1, 0, 3\}$.

3) $\{6, 5, 5\}$.

4) $\{1, i, -i\}$.

5) $\{5 + 6i, 5 - 6i, 5\}$.

Problem 28.3 (10 points):
Which of the following matrices are diagonalizable?
If it is not diagonalizable, tell why. If it is, write down the diagonal matrix B it is conjugated to.

a) (2 points) $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

b) (2 points) $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$.

c) (2 points) $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 1 & 1 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 4 \end{bmatrix}$.

d) (2 points) $\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

e) (2 points) $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 \end{bmatrix}$.

Problem 28.4 (10 points, each sub problem is 2 points):
a) (4 points) Fill in $\leq, =, \geq$ so that the statement is true for an arbitrary real $n \times n$ matrix A. The "number of eigenvalues" is the sum of all algebraic multiplicities of all eigenvalues. No justifications are needed in this problem.

The algebraic multiplicity of an eigenvalue of A is		its geometric multiplicity.
The number of $\mathbb{C}$ eigenvalues of A is		n.
The number of $\mathbb{R}$ eigenvalues of A is		n.
The rank of A is		the number of nonzero eigenvalues of A

b) (2 points) Give an example of a normal 3×3 matrix A, which is not symmetric.

A =

c) (2 points) Give an example of a real 2×2 matrix B which has eigenvalues $6 + 7i$ and $6 - 7i$.

B =

d) (2 points) Give an example of a real 3×3 matrix C which is not diagonalizable.

C =

Problem 28.5 (10 points):
The following 6 matrices can be grouped into 3 pairs of similar transformations. Find these three pairs and justify in each case why the matrices are similar.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 2 \\ 0 & 0 & 3 \end{bmatrix}, B = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 1 & 2 & 2 \end{bmatrix}, C = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 3 & 2 \\ 0 & 0 & 0 \end{bmatrix},$$

$$D = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ 2 & 2 & 2 \end{bmatrix}, E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}, F = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}.$$

Group 1:

Why are they similar?

Group 2:

Why are they similar?

Group 3:

Why are they similar?

Problem 28.6 (10 points):
a) (4 points) Find the eigenvalues λ_1, λ_2 and eigenvectors v_1, v_2 of the matrix

$$A = \begin{bmatrix} 9 & 1 \\ 2 & 8 \end{bmatrix} .$$

b) (6 points) Write down a closed-form solution for the discrete dynamical system

$$\begin{aligned} x(t+1) &= 9x(t) + y(t) \\ y(t+1) &= 2x(t) + 8y(t) \end{aligned}$$

for which $x(0) = 2$, $y(0) = -1$.

Problem 28.7 (10 points):
a) (6 points) Find a closed-form solution to the system

$$\begin{aligned} x'(t) &= 9x(t) + y(t) \\ y'(t) &= 2x(t) + 8y(t) , \end{aligned}$$

for which $x(0) = 2$, $y(0) = -1$.

b) (4 points) Determine the stability of the linear system of differential equations:

$$\begin{aligned} x'(t) &= x(t) - 7y(t) - 9z(t) \\ y'(t) &= x(t) - 2y(t) + z(t) \\ z'(t) &= x(t) + 5y(t) + z(t) . \end{aligned}$$

Problem 28.8 (10 points):
We consider the nonlinear system of differential equations

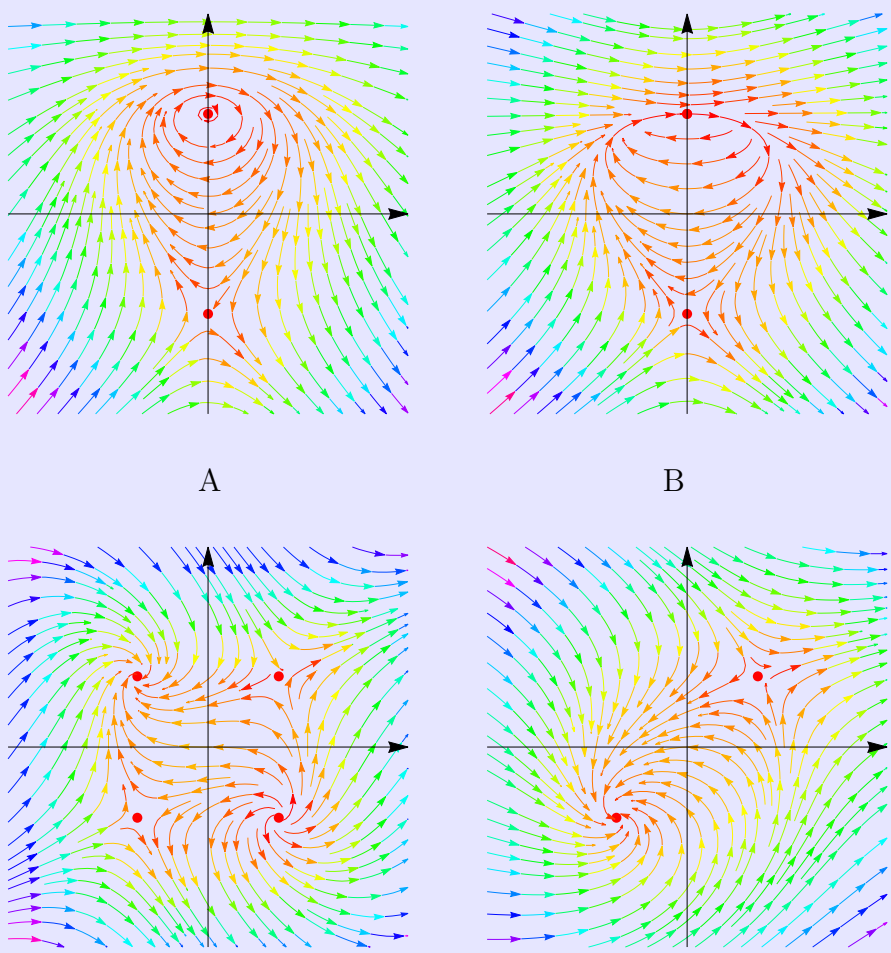
$$\begin{aligned} \frac{d}{dt}x &= x^2 + y^2 - 1 \\ \frac{d}{dt}y &= xy - 2x \ . \end{aligned}$$

a) (2 points) Find the nullclines and equilibrium points.

b) (3 points) Find the Jacobian matrix at each equilibrium point.

c) (3 points) Use the Jacobean matrix at an equilibrium to determine for each equilibrium point whether it is stable or not.

d) (2 points) Which of the diagrams A-D is the phase portrait of the system above? **Draw the nullclines and equilibrium points into the portrait!**



Problem 28.9 (10 points):
Which of the following are linear spaces? If the space is a linear space we need to see you checked its properties. If it is not, we want to see why it is not linear.

a) (2 points) The space of smooth functions f satisfying $f(x) = x + \sin(f(x))$.

b) (2 points) The space of smooth functions f satisfying $f(x) \geq -100$.

c) (2 points) The space of smooth functions f satisfying $f(x) + f(x)\sin(x) = 0$.

Which of the following are linear operators? If it is a linear operator, we need to see you have checked its properties. If it is not, we want to see a reason why it fails to be linear.

d) (2 points) The operator $T(f)(x) = f'(\cos(x))\sin(x)$.

e) (2 points) The operator $T(f)(x) = e^{f(x)} - 1$.

Problem 28.10 (10 points):
Cookbook or operator method. It is your choice! But we want to see protocol and steps, not just the answer!

a) (2 points) Find the general solution of the system $f''' = 24t$.

b) (2 points) Find the general solution of the system $f'' + 9f = 1$.

c) (3 points) Find the general solution of the system $f'' - 4f = 2t$.

d) (3 points) Find the general solution of the system $f'' + 10f' + 16f = 2t$.

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 29: Fourier series

Lecture

29.1. It is convenient for applications to extend the linear space $C^\infty(\mathbb{T})$ of all smooth 2π periodic functions and consider the larger linear space $\mathcal{X}$ of **piecewise smooth periodic functions**. We can draw them functions on the interval $[-\pi, \pi]$. It contains functions as drawn in figure (1). We always draw functions in $\mathcal{X}$ as functions on $[-\pi, \pi]$ and do not insist that the left and right value agree as this just produces another jump when seeing as a function on the circle $\mathbb{T}$. In particular, we just write $f(x) = x$ for example, and draw it as a function on $[-\pi, \pi]$ then think of it 2π-periodically continued.

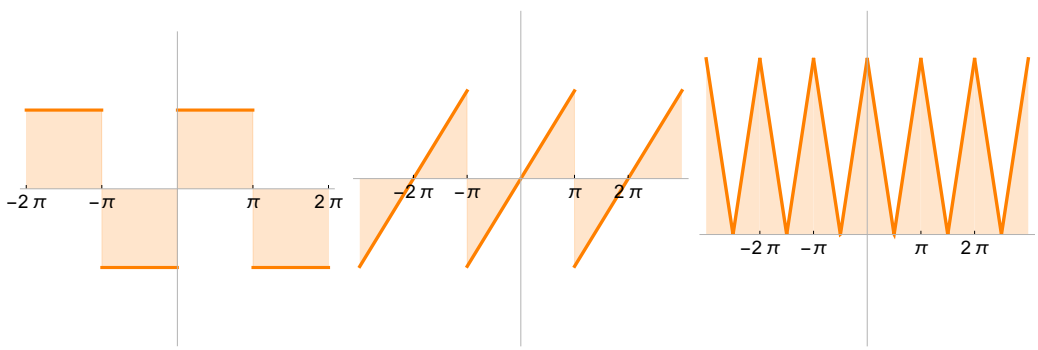


Figure 1. Piecewise smooth 2π-periodic functions in the linear space $\mathcal{X}$.

29.2. On the space $\mathcal{X}$ of piecewise smooth functions $f(x)$ on $[-\pi, \pi]$ there is an **inner product** defined by

$$\langle f, g \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) g(x) \, dx \ .$$

It plays the role of the **dot product** $v \cdot w$ in $\mathbb{R}^n$ or $\mathrm{tr}(A^T B)$ in $M(n, m)$.

29.3. This product allows to define angles, **length** $|f|$, **distance** $|f - g|$ or **projections** in X as we did in finite dimensions: the length is $\sqrt{\langle f, f \rangle}$. The **angle** α is defined by by $\cos(\alpha) = \langle f, g \rangle / (|f||g|)$ and the projection of f onto g is $\frac{\langle f, g \rangle}{|g|^2} g$.

29.4. The function $f(x) = x^2$ for example has length $|f| = \sqrt{(1/\pi)\int_{-\pi}^{\pi} x^4\,dx} = \sqrt{2\pi^5/(5\pi)} = \sqrt{2/5}\pi^2$. It is perpendicular to the function $g(x) = x^3$. It illustrates the general principle that even and odd functions are perpendicular to each other.

Lemma: $\{\cos(nx), \sin(nx), 1/\sqrt{2}\,\}$ form an orthonormal set in $\mathcal{X}$.

The addition formulas

$$\begin{aligned} 2\cos(nx)\cos(mx) &= \cos(nx-mx)+\cos(nx+mx) \\ 2\sin(nx)\sin(mx) &= \cos(nx-mx)-\cos(nx+mx) \\ 2\sin(nx)\cos(mx) &= \sin(nx+mx)+\sin(nx-mx) \quad . \end{aligned}$$

allow to verify these things. What helps is that integral of an odd function over $[-\pi,\pi]$ is zero:

Proof. $\langle 1/\sqrt{2}, 1/\sqrt{2}\,\rangle = 1$
$\langle \cos(nx), \cos(nx)\rangle = 1, \langle \cos(nx), \cos(mx)\rangle = 0$
$\langle \sin(nx), \sin(nx)\rangle = 1, \langle \sin(nx), \sin(mx)\rangle = 0$
$\langle \sin(nx), \cos(mx)\rangle = 0$
$\langle \sin(nx), 1/\sqrt{2}\rangle = 0$
$\langle \cos(nx), 1/\sqrt{2}\rangle = 0$ □

29.5. The **Fourier coefficients** of a function f in X are defined as

$$a_0 = \langle f, 1/\sqrt{2}\rangle = \tfrac{1}{\pi}\int_{-\pi}^{\pi} f(x)/\sqrt{2}\,dx$$

$$a_n = \langle f, \cos(nt)\rangle = \tfrac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(nx)\,dx$$

$$b_n = \langle f, \sin(nt)\rangle = \tfrac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(nx)\,dx$$

29.6. The **Fourier representation** of a piecewise smooth function f is the identity

$$f(x) = \frac{a_0}{\sqrt{2}} + \sum_{k=1}^{\infty} a_k\cos(kx) + \sum_{k=1}^{\infty} b_k\sin(kx)$$

We will see later that the series converges and that the identity holds at all points x where f is continuous. We first want to learn how to compute the Fourier expansion.

29.7. Remember that f is **odd** if $f(x) = -f(-x)$ for all x. A function f is **even** if $f(x) = f(-x)$ for all x.

If f is odd then f has a sin-series.
If f is even then f has a cos-series.

This follows from the fact that if the definite integral of an odd function over $[-\pi,\pi]$ is always 0, and that the product between an even and an odd function is always odd.

29.8. Find the Fourier series of $f(x) = x$ on $[-\pi, \pi]$. This is an odd function $f(-x) = -f(x)$ so that it has a sin series: with

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} x\sin(nx)\,dx = \frac{-1}{\pi}(x\cos(nx)/n + \sin(nx)/n^2|_{-\pi}^{\pi}) = \frac{2(-1)^{n+1}}{n}\ ,$$

we get

$$x = \sum_{n=1}^{\infty} 2\frac{(-1)^{n+1}}{n}\sin(nx)\ .$$

29.9. If we evaluate both sides at a point x, we obtain identities. For $x = \pi/2$ for example, we get

$$\frac{\pi}{2} = 2(\frac{1}{1} - \frac{1}{3} + \frac{1}{5} - \frac{1}{7}...)\ .$$

This is a **formula of Leibniz**.

29.10. Let $f(x) = 1$ on $[-\pi/2, \pi/2]$ and $f(x) = 0$ else. This is an even function $f(-x) - f(x) = 0$. It has a cos-series: with $a_0 = 1/(\sqrt{2}), a_n = \frac{1}{\pi}\int_{-\pi/2}^{\pi/2} 1\cos(nx)\,dx = \frac{\sin(nx)}{\pi n}|_{-\pi/2}^{\pi/2} = \frac{2(-1)^m}{\pi(2m+1)}$ if $n = 2m+1$ is odd and 0 else. So, the series is

$$f(x) = \frac{1}{2} + \frac{2}{\pi}(\frac{\cos(x)}{1} - \frac{\cos(3x)}{3} + \frac{\cos(5x)}{5} - ...)$$

What happens at the discontinuity? The Fourier series converges to 1/2. Diplomatically it has chosen the point in the middle of the limits from the right and the limit from the left. By the way, also in this example with $x = 0$, we get 1 on the left and $1/2 + 2/\pi(1 - 1/3 + 1/5 - 1/7 + \cdots)$ confirming again the Leibniz formula.

29.11. The function $f_n(x) = \frac{a_0}{\sqrt{2}} + \sum_{k=1}^{n} a_k\cos(kx) + \sum_{k=1}^{n} b_k\sin(kx)$ is called a **Fourier approximation** of f. The picture below plots a few approximations in the case of a piecewise continuous even function given in the above example.

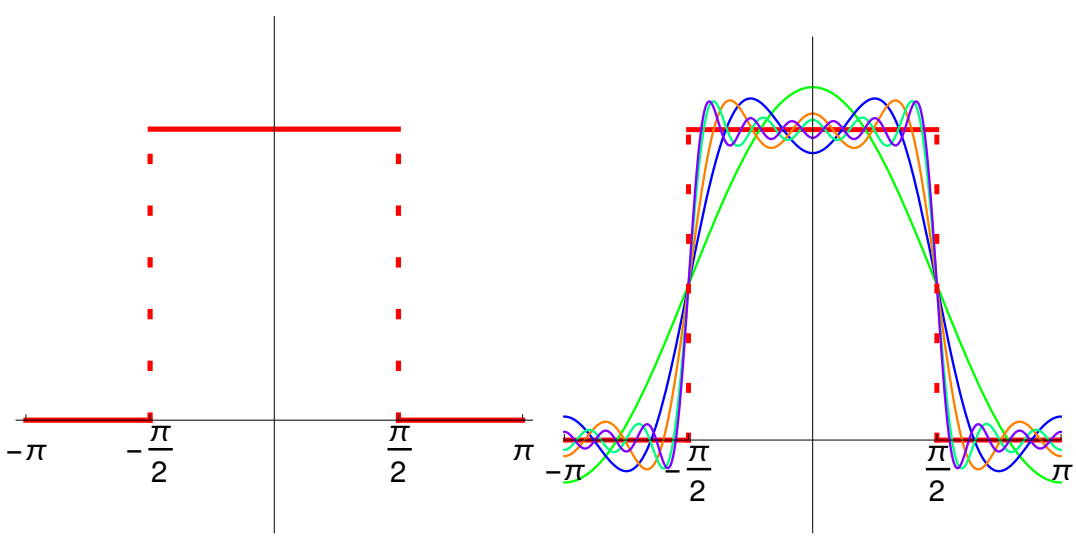


FIGURE 2. A piecewise continuous function and some Fourier approximations.

Homework

This homework is due on Tuesday, 4/16/2019.

Problem 29.1: Find the Fourier series of the function $f(x) = 22 + |6x|$.

Problem 29.2: a) Find the Fourier series of the function $4\cos(3x) + \sin^2(22x) + 22$. (Note, there is almost nothing to do here).
b) What is the Fourier series of the function which is x for $x \geq 0$ and 0 else?

Problem 29.3: a) Find the angle between the functions $f(x) = x^2$ and $g(x) = x^3$.
b) Project $f(x) = \sin^2(x)$ onto the plane spanned by $\sin(2x), \cos(2x)$.
c) Find the length of the function $f(x) = x^3$ in C^∞_{per}.

Problem 29.4: Find the Fourier series of the function $f(x) = |\sin(x)|$.

Problem 29.5: The **inner product** of two complex functions f, g in $C(\mathbb{T})$ is defined as

$$\langle f, g \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} \overline{f(x)} g(x) \, dx \ .$$

a) Check that $\mathcal{B} = \{e^{inx}\}$ with $n \in \mathbb{Z}$ is an orthonormal family in $C^\infty(\mathbb{T})$.
b) Write $\cos(nx)$ and $\sin(nx)$ as a linear combination of functions in $\mathcal{B}$.
c) Write e^{inx} as a linear combination of real trig functions.
d) Now write $f(x) = x$, the example done in this text again, but as a Fourier series using the complex basis $\mathcal{B}$.

P.S. Fourier theory in the complex are a bit more natural. We do not have to treat three different cases like constant function, even or odd trig functions. The Fourier basis is $\{e^{inx}, n \in \mathbb{Z}\}$ which is the eigenbasis of D on the circle. The complex Fourier coefficients are

$$c_n = \langle e^{inx}, f \rangle \ .$$

The **complex Fourier series** is

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx} \ .$$

Oliver Knill, knill@math.harvard.edu, Math 22b, Harvard College, Spring 2019

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 30: Dirichlet's Proof

Seminar

30.1. The historical development of a mathematical topic not only gives background, it also illustrates the struggle in the search of a theory and often mirrors the difficulties which a modern student experiences, when learning the subject. We will spend some time in this proof seminar with the battle for proving that the Fourier series converges. Fourier claimed this to be true without justification. The proof given here came only later with Dirichlet. Be advised that this is not an easy proof as there are several bits and pieces which come together. It can be rewarding however to battle it and appreciate its difficulty.

30.2. Fourier theory is an important topic also because it was a turning point in understanding the concept of a function. The origin however came from a concrete problem. How can one describe the diffusion of heat? Fourier started to think about this problem while in the service of Napoleon during the campaign from 1798-1801 in Egypt. But the book on heat appeared only in 1822.

> **Problem A:** Find the part in Fourier's book, where the Fourier series is introduced. The document which Fourier wrote in 1822 can be found on the website.

30.3. Before Fourier, one has seen functions only tied to their analytic expressions. Indeed one can see a Taylor series in the complex related to a Fourier series. Look at the series $\sum_{k=1}^n a_k z^k$. If we evaluate this on $|z| = 1$, we can plug in $z = e^{ix}$ and get $\sum_{k=1}^n a_k e^{ikx}$ which as a function of x has real and imaginary parts which are now given as Fourier series.

> **Problem B:** What is the real part and imaginary part of $\sum_{k=1}^n a_k e^{ikx}$?

30.4. Because of this relation of Taylor series and Fourier series, one might think that Fourier series work only in analytic situations and fail for a function which has discontinuities. The surprise is that Fourier series can handle also discontinuous functions!

30.5. A major theorem about Fourier series deals with functions in $\mathcal{X}$, the space of piece-wise smooth functions on $[-\pi, \pi]$. It is a theorem due to **Peter Gustav Dirichlet** from 1829.

Theorem: The Fourier series of $f \in \mathcal{X}$ converges at every point of continuity. At discontinuities, it takes the middle value.

30.6.

Problem C: Try to understand as much as possible from the following proof of the theorem. During the proof seminar, you go through the main line of the proof. In the homework you flesh out some details.

$\boxed{\text{Part I}}$ of the proof is a computation:
Let $S_n(x) = \frac{1}{2}a_0 + \sum_{k=1}^n a_k \cos(kx) + b_k \sin(kx)$ denote the n'th partial sum of f. By plugging in the formulas

$$\begin{aligned} a_0 &= \tfrac{1}{\pi}\textstyle\int_{-\pi}^{\pi} \tfrac{1}{2} f(y)\, dy \\ a_k &= \tfrac{1}{\pi}\textstyle\int_{-\pi}^{\pi} \cos(ky) f(y)\, dy \\ b_k &= \tfrac{1}{\pi}\textstyle\int_{-\pi}^{\pi} \sin(ky) f(y)\, dy, \end{aligned}$$

we get

$$S_n(x) = \int_{-\pi}^{\pi} D_n(x-y) f(y)\, dy$$

where

$$D_n(x-y) = \frac{1}{\pi}(\frac{1}{2} + \sum_{k=1}^n \cos(kx)\cos(ky) + \sin(kx)\sin(ky)) = \frac{1}{\pi}(\frac{1}{2} + \sum_{k=1}^n \cos(k(x-y)))\ .$$

30.7. $\boxed{\text{Part II}}$: the function $D_n(z)$ is called the **Dirichlet kernel**. The next lemma gives a simple closed-form expression for it, which does not involve a sum:

Lemma: $D_n(x) = \frac{1}{2\pi}\frac{\sin((n+\frac{1}{2})x)}{\sin(\frac{x}{2})}$.

Proof. There are three identities which tie things together. You verify them in the homework.
a) $\frac{1}{\pi}(\frac{1}{2} + \sum_{k=1}^n \cos(kx)) = \frac{1}{2\pi}\sum_{k=-n}^n e^{ikx}$.
b) $\frac{1}{2\pi}\sum_{k=-n}^n e^{ikx} = \frac{1}{2\pi}\frac{e^{i(n+1)x} - e^{-inx}}{e^{ix}-1}$.
c) $\frac{1}{2\pi}\frac{e^{i(n+1)x} - e^{-inx}}{e^{ix}-1} = \frac{1}{2\pi}\frac{\sin((n+\frac{1}{2})x)}{\sin(x/2)}$.
This expression is understood by l'Hopital as $\frac{1}{\pi}(2n+1)$ for $x = 0$. □

30.8. $\boxed{\text{Part III}}$ of the proof gives a formula for the difference between $S_n(x)$ and $f(x)$. We see from the definition of $D_n(x) = \frac{1}{\pi}(\frac{1}{2} + \sum_{k=1}^n \cos(kx))$ that $\int_{-\pi}^{\pi} D_n(y)\, dy = 1$. We can therefore write

$$f(x) = \int_{-\pi}^{\pi} D_n(y) f(x)\, dy\ .$$

By a change of variables,

$$S_n(x) = \int_{-\pi}^{\pi} D_n(x-y)f(y)\ dy = \int_{-\pi}^{\pi} D_n(y)f(x+y)\ dy\ .$$

Therefore,

$$S_n(x) - f(x) = \int_{-\pi}^{\pi} D_n(y)(f(x+y) - f(x))\ dy\ .$$

30.9. $\boxed{\text{Part IV}}$ introduces a function

$$F_x(y) = \frac{f(x+y) - f(x)}{2\sin(y/2)}\ .$$

with the understanding that $F_x(0) = f'(x)$. Even so f is only piecewise continuous, the function $y \to F_x(y)$ is continuous and 2π-periodic. With this function, we can write

$$S_n(x) - f(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} F_x(y)\sin((n+\frac{1}{2})y)\ dy\ .$$

30.10. $\boxed{\text{Part V}}$ first uses a trig identity $\sin(n+\frac{1}{2}) = \cos(y/2)\sin(ny) + \sin(ny/2)\cos(ny)$, then introduces two more continuous periodic functions

$$G_x(y) = F_x(y)\cos(y/2), H_x(y) = F_x(y)\sin(y/2)\ .$$

The n'th Fourier coefficients of these functions are

$$B_n = \frac{1}{\pi}\int_{-\pi}^{\pi} G_x(y)\sin(ny)\ dy$$

$$A_n = \frac{1}{\pi}\int_{-\pi}^{\pi} H_x(y)\cos(ny)\ dy$$

Putting things together, we see

$$S_n(x) - f(x) = B_n + A_n\ .$$

30.11. $\boxed{\text{Part VI}}$ is called the **Riemann-Lebesgue lemma**. It tells that for continuous functions g, h, we have in the limit $n \to \infty$

$$A_n = \frac{1}{\pi}\int_{-\pi}^{\pi} g(x)\cos(nx)\ dx \to 0$$

$$B_n = \frac{1}{\pi}\int_{-\pi}^{\pi} h(x)\sin(nx)\ dx \to 0\ .$$

we can then apply this in the above situation for $g(y) = G_x(y)$ or $h(y) = H_x(y)$ and see that $A_n \to 0$ and $B_n \to 0$. The Riemann-Lebesgue lemma follows from the **Bessel inequality**

$$a_0^2 + \sum_{k=1}^{n} a_k^2 + b_k^2 \leq \langle f, f\rangle = ||f||^2\ .$$

By the Pythagoras theorem, as $f(x) - S_n(x)$ and $S_n(x)$ are perpendicular, we have

$$||f - S_n||^2 + ||S_n||^2 = ||f||^2\ .$$

This implies that $||S_n||^2 \leq ||f||^2$ for all n as $||S_n||^2 = a_0^2 + \sum_{k=1}^{n} a_k^2 + b_k^2$ is a sum of non-negative terms, the infinite sum has to converge and a_n, b_n have to converge to zero.

Homework

This homework is due on Tuesday, 4/16/2019.

Problem 30.1: Verify

$$\frac{1}{\pi}(\frac{1}{2}+\sum_{k=1}^{n}\cos(kx)) = \frac{1}{2\pi}\sum_{k=-n}^{n} e^{ikx} \ .$$

Problem 30.2: Verify

$$\frac{1}{2\pi}\sum_{k=-n}^{n} e^{ikx} = \frac{1}{2\pi}\frac{e^{i(n+1)x}-e^{-inx}}{e^{ix}-1} \ .$$

Problem 30.3: Verify

$$\frac{1}{2\pi}\frac{e^{i(n+1)x}-e^{-inx}}{e^{ix}-1} = \frac{1}{2\pi}\frac{\sin((n+\frac{1}{2})x)}{\sin(x/2)} \ .$$

Problem 30.4: Verify that in the limit $x \to 0$, the expression

$$\frac{1}{\pi}\frac{\sin((n+\frac{1}{2})x)}{2\sin(x/2)}$$

becomes $\frac{1}{2\pi}(2n+1)$.

Problem 30.5: We have not yet seen what happens, with the convergence if x is a point of discontinuity. How do you settle the general case, in which several jump discontinuities can occur?
Here is a start: let us assume to have an odd function $g(x)$ which has a jump discontinuity at 0 with $g(x+0) = a$, $g(x-0) = -a$. The proof above shows that the Fourier series $\sum_n b_n \sin(nx)$ converges to $g(x)$ at every point $x \neq 0$. At $x = 0$, the series converges to 0, the middle point.
Now, if $f(x)$ has a single discontinuity at 0 jumping by $2a$, look at $f(x) - g(x)$. This function is continuous and has a convergent Fourier series. The Fourier series of f therefore converges to the middle of the discontinuity.

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 31: Parseval's theorem

Lecture

31.1. We have seen that every $f \in \mathcal{X}$ can be represented as a series

$$f(x) = \frac{a_0}{\sqrt{2}} + \sum_{n=1}^{\infty} a_n \cos(nx) + \sum_{n=1}^{\infty} b_n \sin(nx) \ ,$$

where the Fourier coefficients a_n, b_n are were given by integrals

$$\begin{aligned} a_0 &= \langle f, 1/\sqrt{2} \rangle = \tfrac{1}{\pi} \int_{-\pi}^{\pi} f(x)/\sqrt{2}\ dx, \\ a_n &= \langle f, \cos(nt) \rangle = \tfrac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx)\ dx, \\ b_n &= \langle f, \sin(nt) \rangle = \tfrac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx)\ dx. \end{aligned}$$

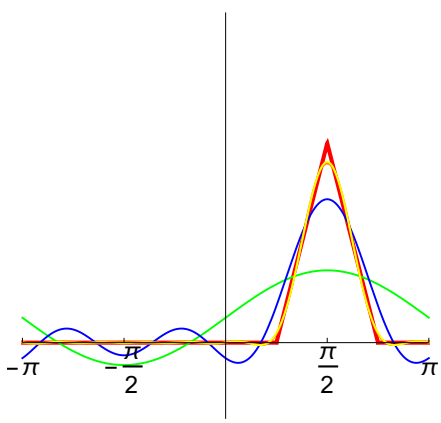


Figure 1. A Fourier approximation of a function $f \in \mathcal{X}$ which is neither even nor odd. The function f is piecewise linear and continuous.

31.2. The inner product allowed us to define the length $||f||^2 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)^2\ dx$ of a vector. The following theorem is called the **Parseval's identity**. It is the **Pythagoras theorem** for Fourier series.

Theorem:

$$||f||^2 = a_0^2 + \sum_{n=1}^{\infty} a_n^2 + b_n^2 \ .$$

Proof. The function $g(x) = \frac{a_0}{\sqrt{2}} + \sum_{n=1}^{\infty} a_n \cos(nx) + \sum_{n=1}^{\infty} b_n \sin(nx)$ agrees with $f(x)$ except at finitely many points. This implies $||f||^2 = ||g||^2$. Let us compute the later using that $\cos(nx), \sin(nx)$ and $1/\sqrt{2}$ form an orthonormal basis:

$$\frac{1}{\pi}\int_{-\pi}^{\pi}[\frac{a_0}{\sqrt{2}} + \sum_{n=1}^{\infty} a_n \cos(nx) + \sum_{n=1}^{\infty} b_n \sin(nx)][\frac{a_0}{\sqrt{2}} + \sum_{n=1}^{\infty} a_n \cos(nx) + \sum_{n=1}^{\infty} b_n \sin(nx)]\, dx$$

which is after foiling out $a_0^2 + \sum_{n=1}^{\infty} a_n^2 + b_n^2$. □

31.3. Here is an example: We have seen the Fourier series for $f(x) = x$ as

$$f(x) = 2(\sin(x) - \frac{\sin(2x)}{2} + \frac{\sin(3x)}{3} - \frac{\sin(4x)}{4} + \dots).$$

The coefficients $b_k = 2(-1)^{k+1}/k$ and so

$$4(\frac{1}{1} + \frac{1}{4} + \frac{1}{9} + \cdots) = \frac{1}{\pi}\int_{-\pi}^{\pi} x^2\, dx = \frac{2\pi^2}{3}\ .$$

This can be written as

$$1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \frac{1}{25} + \cdots = \frac{\pi^2}{6}\ .$$

This is another solution of the Basel problem. (See Problem 18.3 in Math 22a).

31.4. There are different types of convergence for functions. From the Parseval's identity we have

$$||f - f_n|| \to 0\ .$$

We call this L^2 convergence. If f is smooth, then from the pointwise convergence we have a stronger **uniform convergence** using $||f - g||_\infty = \max|f(x) - g(x)|$.

Corollary: If f is smooth, then $||f_n - f||_\infty \to 0$.

Proof. This follows from the Dirichlet proof on Fourier series and the Cantor-Heine Theorem (see Unit 8 in Math 22a). □

31.5. What about the case with discontinuities? Here uniform convergence can fail. It is called the **Gibbs phenomenon**. Remember from the proof seminar, where we had

$$S_n(x) - f(x) = \int_{-\pi}^{\pi} D_n(y)(f(x+y) - f(x))\, dy\ .$$

Now, in the case when $f(x) = \text{sign}(x)$ with $f(x) = 1$ for $x > 0$ and $f(x) = -1$ for $x < 0$, we can evaluate the error $S_n(x) - f(x)$ at $x = \pi/n$ and get

$$S_n(x) - f(x) = \int_{\pi/n}^{\pi} D_n(y) 2\, dy\ .$$

This can be shown to have a non-zero limit for $n \to \infty$.

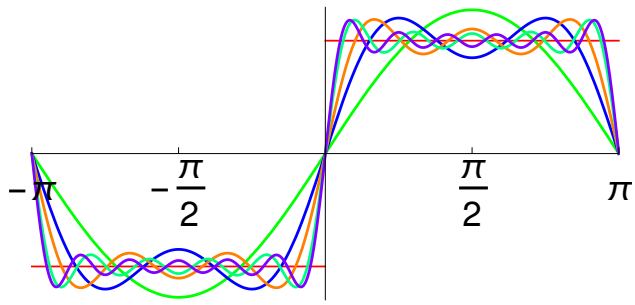


FIGURE 2. The Gibbs phenomenon: the Fourier approximation does not converge uniformly to f if f is not continuous. The series converges pointwise.

EXAMPLES

31.6. Find the Fourier series of the function $f(x) = \begin{cases} 1 & , |x| < \pi/4 \\ 0 & , |x| \geq \pi/4 \end{cases}$. What is the sum of the squares of the Fourier coefficients? Answer: The function is even. It has a cos series. We compute $a_0 = (2/\pi)\pi/(4\sqrt{2}) = \frac{\sqrt{2}}{4}$ and $a_n = \frac{2}{n\pi}\sin(n\pi/4)$. The Fourier series is

$$f(x) = \frac{\sqrt{2}}{4}\frac{1}{\sqrt{2}} + \sum_{n=1}^{\infty} \frac{2}{n\pi}\sin(n\pi/4)\cos(nx) \ .$$

By Parseval, we have $a_0^2 + \sum_n a_n^2 = ||f||^2$ which is $(\pi/2)/\pi = 1/2$. We did not have to sum up the sum after all! Let us also illustrate how an AI computes things:

```
FourierCosSeries[If[x^2 < Pi^2/4, 1, 0], x, 10]
```

31.7. The function $g(x) = x^3 - \pi^2 x$ has the Fourier series

$$g(x) = \sum_{n=1}^{\infty} \frac{12(-1)^n}{n^3}\sin(nx) \ .$$

What is

$$\sum_{n=1}^{\infty} \frac{1}{n^6} \ ?$$

This number is called $\zeta(6)$, the value of the **Riemann Zeta function** at 6. Parseval's theorem $|g|^2 = \sum_{n=1}^{\infty} b_n^2$ shows that the result $144\zeta(6) = \sum_n b_n^2$ is

$$\frac{2}{\pi}\int_0^{\pi} (x^3 - \pi^2 x)^2 \, dx = \frac{16\pi^6}{105} \ .$$

From the Parseval identity we get

$$\zeta(6) = \frac{1}{144}\sum_n b_n^2 = \frac{1}{144}\frac{16\pi^6}{105} = \frac{\pi^6}{945} \ .$$

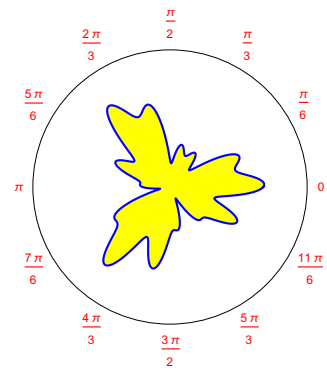
31.8. It is possible like this to get explicit expressions for $\zeta(2n)$ for even positive n. For odd n, such prepresentations are not known. The smallest, where one does not know it is the Apéry's constant $\zeta(3)$. A Parseval theorem type approach seems not to help in computing this constant.

Homework

This homework is due on Tuesday, 4/23/2019.

Problem 31.1: Find the Fourier series of the function which is 22 on $[0, \pi/4]$ and zero everywhere else. What is the sum $a_0^2 + \sum_{n=1}^{\infty} a_n^2 + b_n^2$?

Problem 31.2: a) Use Parseval to find the length $||f||$ for the **fairy godmother curve**
$f(x) = 10 + 2\cos(14x) + \cos(11x) + 2\sin(7x) - \cos(4x) + 5\cos(3x)$. b) What does $||f||$ mean geometrically for the polar curve $r(t) = f(t)$?



Problem 31.3: Compute both sides of the Parseval identity for $f(x) = x + |x|$.

Problem 31.4: Find $\sum_{n=1}^{\infty} \frac{1}{(2n)^2} = 1/4 + 1/16 + 1/36 + \dots$ from the known Basel problem formula of $\sum_n \frac{1}{n^2}$ and use this to compute the sum $\sum_{n=0}^{\infty} \frac{1}{(2n+1)^2}$ over the odd numbers. Written out, this is

$$\frac{1}{1} + \frac{1}{9} + \frac{1}{25} + \cdots .$$

Problem 31.5: We have seen the Parseval identity

$$\langle f, f \rangle = a_0^2 + \sum_{n=1}^{\infty} a_n^2 + b_n^2 .$$

Why does this imply more generally that if f is a function with Fourier coefficients a_n, b_n and g is a function with Fourier coefficients c_n, d_n, then

$$\langle f, g \rangle = a_0 c_0 + \sum_{n=1}^{\infty} a_n c_n + b_n d_n \ ?$$

Either prove this directly analogously to what we did when proving the Parseval identity or then reduce it to the Parseval identity.

P.S. Here is a historical challenge: we know very little about **Marc-Antoine Parseval des Chenes**. The result is named after Parseval as there was a note written in 1799 which contains a statement looking similar. In the St-Andrews article of J.J. O'Connor and E.F. Robertson about Parseval, it is stated that *it would not be unfair to say that Parseval has fared well in having a well known result, which is quite far removed from his contribution, named after him. However he remains a somewhat shadowy figure and it is hoped that research will one day provide a better understanding of his life and achievements..*

Oliver Knill, knill@math.harvard.edu, Math 22b, Harvard College, Spring 2019

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 32: Fourier Applications

Lecture

32.1. Fourier theory has many applications. There are mathematical applications in **number theory**, **arithmetic**, **ergodic theory**, **probability theory** as well as applications in applied sciences like **signal processing**, **quantum dynamics**, **data compression** or **tomography**. In this lecture, we mention a few applications, sometimes a bit informally as subjects like probability theory, ergodic theory, number theory or inverse problems are subjects which each would fill courses by themselves.

32.2. In **probability theory**, a non-negative function f which has the property that $\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\ dx = 1$ is called a **probability density function**. This is abbreviated often as PDF. The complex Fourier coefficients $c_n = \mathrm{E}[fe^{inx}]$ of f form what one calls the **characteristic function** of the distribution. Why is this useful? If we have two distributions f and g representing **independent data**, then the **convolution**

$$f \star g(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-y)g(y)\ dy$$

represents the distribution of the sum of the data. Here is the math:

Lemma: If c_n and d_n are the Fourier coefficients of f and g, then $c_n d_n$ is the characteristic function of $f \star g$.

Proof. The n'th Fourier coefficient of $f \star g$ is

$$\frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f(x-y)g(y)\ dy e^{-inx}\ dx\ .$$

A change of variables $z = x - y$ gives

$$\frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f(z)g(y)\ dy e^{-in(z+y)}\ dx\ .$$

This can be written as

$$[\frac{1}{2\pi}\int_{-\pi}^{\pi} f(z)e^{-inz}\ dz][\frac{1}{2\pi}\int_{-\pi}^{\pi} g(y)e^{-iny}\ dy] = c_n d_n\ .$$

□

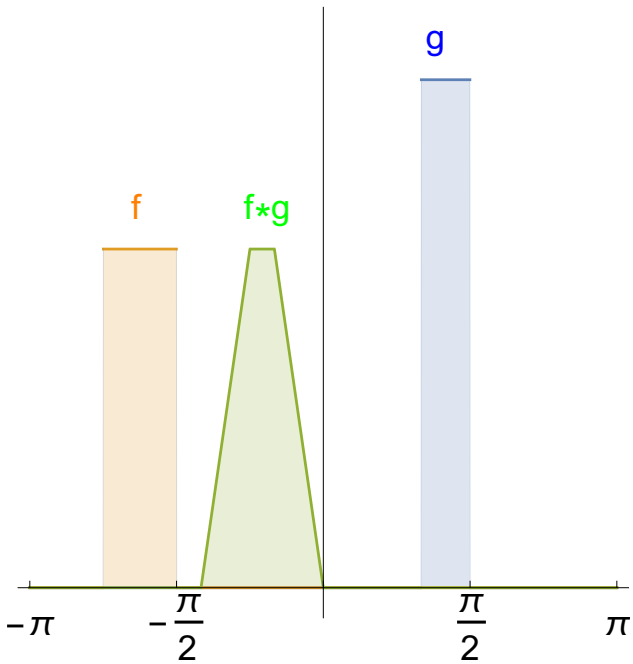


FIGURE 1. We see f, g and its convolution $f \star g$. The Fourier coefficients of $f \star g$ is the product of the Fourier coefficients of f.

32.3. All this can be done also for distributions on $\mathbb{R}$ rather than $[-\pi, \pi]$. The characteristic function is then a **Fourier transform**. When dealing with **circular data**, then the **Fourier series** are important. Now, if f is a constant distribution, then its Fourier data are $c_n = 0$ except $c_0 = 1$. For a probability distribution in general all $|c_n| < 1$ for $n \neq 0$ and $c_0 = 1$. We get now immediately the **central limit theorem for circular data** with data which have identical distribution:

Theorem: Given a sequence of independent circular data, then the distribution of the sum converges to the constant distribution.

Proof. Each density function f_k has Fourier coefficients c_n and the sum of m independent data has the Fourier coefficients c_n^m. Since for $n \neq 0$, we have $|c_n| < 1$ we have $c_n^m \to 0$. In the limit we have a distribution which has only one non-zero Fourier coefficient $c_0 = 1$. This is the constant distribution. □

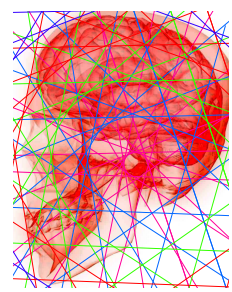
32.4. A similar analysis works also in the continuum case. There, the limiting distribution is the **standard normal distribution** $f(x) = 1/\sqrt{2\pi}e^{-x^2/2}$. It has the property that the convolution of $f(x)$ with itself is again a normal distribution but with variance 2. The **central limit theorem** now uses the **Fourier transform**.

32.5. In **ergodic theory**, which is also the mathematical frame work for "chaos", one studies the long term behavior of dynamical systems. Let us look at the transformation $T(x) = x + \alpha$ where α is some irrational multiple of 2π. Given a function $f(x)$, what happens with **time average** $S_n/n = \frac{1}{n}[f(x) + f(x+\alpha) + \cdots + f(x + (n-1)\alpha)]$ in the limit $n \to \infty$. The expectation $\mathrm{E}[f]$ of f is the **space average** $\int_{-\pi}^{\pi} f(x)\, dx/(2\pi)$. There is the following **ergodic theorem**

Theorem: Time average is space average $\frac{S_n}{n} \to \mathrm{E}[f]$ for $n \to \infty$

Proof. The Fourier coefficients of the sum is $c_n(1 + e^{i\alpha} + \cdots + e^{i(n-1)\alpha})$ which is $c_n(1 - e^{in\alpha})/(1 - e^{i\alpha})$. Because we divide by n, each Fourier coefficient converges to 0 except for the zero'th coefficient c_0 which is always $\mathrm{E}[f]$. $\square$

32.6. In **magnetic resonance imaging** one has the problem of finding the density function $g(x, y, z)$ of a three dimensional body from measuring the **absorption rate** along lines. One can reduce it to two dimensions by looking at **slice** $f(x, y) = g(x, y, c)$, where $z = c$ is kept constant. The **Radon transform**, introduced by Johann Radon in 1917, produces from f another function



$$R(f)(p, \theta) = \int_{\{x\cos(\theta)+y\sin(\theta)=p\}} f(r(t))\ |r'(t)|dt$$

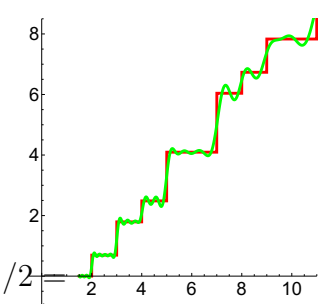
measuring the absorption along the line L of polar angle α in distance p from the center and where the line is parametrized by a curve $r(t)$. Reconstructing $f(x, y) = g(x, y, c)$ for different c allows to recover the **tissue density** g and so "see inside the body".

Theorem: The Radon transform can be diagonalized using Fourier theory

To do so, we need some regularity: we need that $\phi \to f(r, \phi)$ is piecewise smooth which then assures that the Fourier series $f(r, \phi) = \sum_n f_n(r)e^{in\phi}$ converges. We also need that $r \to f_n(r)$ has a Taylor series. The expansion $f(r, \phi) = \sum_{n\in\mathbb{Z}} \sum_{k=1}^{\infty} f_{n,k}\psi_{n,k}$ with $\psi_{n,k}(r, \phi) = r^{-k}e^{in\phi}$ is an eigenfunction expansion with explicitly known eigenvalues $\lambda_{n,k}$. The **inverse problem** is subtle due to the existence of a **kernel** spanned by $\{\psi_{n,k} \mid (n+k) \text{ odd }, |n| > k\}$. In applied situations, one calls it an **ill posed problem**.

32.7. Fourier theory also helps to understand **primes**. For an integer n, let Λ denote the **Mangoldt function** defined by $\Lambda(n) = \log(p)$ if $n = p^k$ is a power of some prime p and $\Lambda(n) = 0$ else. Its sum $\psi(x) = \sum_{n\le x} \Lambda(n)$ is called the **Chebyshev function**. Riemann indicated and Mangoldt proved first that it satisfies the **Riemann-Mangoldt formula**

$$\psi(x) = x - \sum_w \frac{x^w}{w} - \log(2\pi) - \frac{1}{2}\log(1 - x^{-2}) \ ,$$



where w runs over the non-trivial roots of the **Riemann zeta function**

$$\zeta(s) = \sum_{n\ge 1} n^{-s} \ ,$$

and where $\log(2\pi) = \zeta'(0)/\zeta(0)$ comes from the simple pole at 1 and $\log(1 - x^{-2})/2 = \sum_{k=1}^{\infty} -x^{-2k}/(2k)$ is the contribution of the **trivial roots** $-2, -4, -6, \ldots$ of the zeta function and $f_j(x) = x^{w_j}/w_j = e^{\log(x)a_j+\log(x)ib_j}/w_j$ is the contribution from the nontrivial zeros w_j (which are believed on the line $\mathrm{Im}(z) = 1/2$). Pairing complex conjugated roots $w_j = a_k + ib_j = |w_j|e^{i\alpha_j}, \overline{w}$ gives a sum of functions

$$f_j(x) = e^{\log(x)a_j} 2\cos(\log(x)b_j - \alpha_j)/|a_j + ib_j| \ .$$

The functions f_j are the tunes of the **music of the primes**. There is a book and movie with this title by Marcus du Sautoy.

Homework

This homework is due on Tuesday, 4/23/2019.

Problem 32.1: The piecewise linear continuous function g has a graph connecting $(-\pi,0),(-\pi/2,\pi/2-1),(0,-1/2),(\pi/2,\pi/2-1),(\pi,0)$. It satisfies $g=f\star f$, where f be the function which is -1 on $[-\pi/2,0]$ and 1 on $[0,\pi/2]$. Use HW 32.5 to compute the Fourier coefficients of g.

Problem 32.2: a) In a magnetic resonance problem, we measure the density function $f(r,\phi)=\sum_n r^n\cos(n\phi)$. Find a closed-form for $f(r,\phi)$. **Hint:** the series is the real part of $\sum_{n=1}^\infty r^n e^{in\phi}$. You can assume $|r|<1$.
b) Give an explicit expression for $\sum_{n=1}^\infty \frac{1}{2^n}\cos(nx)$.

Problem 32.3: There is a general principle which tells that the smoother a function is as faster the Fourier series decays.
Given a Fourier series $f(x)=\sum_n b_n\sin(nx)$ of a smooth function, can you give the Fourier series of derivative $f'(x)$? Conclude that for an odd $f\in C^\infty$ and any k, like $k=22$, one has $b_n n^{22}\to 0$ as $n\to\infty$.

Problem 32.4: Something in number theory: Define the Fourier series $f(x)=\sum_p e^{ipx}/2^p$, where p runs over all primes. Define the function $g(x)=f(x)^2$ and compute its Fourier coefficients c_n. Why is the Goldbach conjecture equivalent to the fact that all $c_{2n}\neq 0$ for $n>1$?

Problem 32.5: Prove that if f is an odd function with Fourier coefficients b_n, then $g=f\star f$ is even with Fourier coefficients $a_0=0$ and $a_n=-b_n^2/2$. Hint: use the theorem and hat for an odd function $b_n=-2\mathrm{Im}(c_n)$ and $a_n=2\mathrm{Re}(c_n)$, where c_n is the n'th complex Fourier coefficient of f. A key relation therefore is $2c_n=a_n-ib_n$.

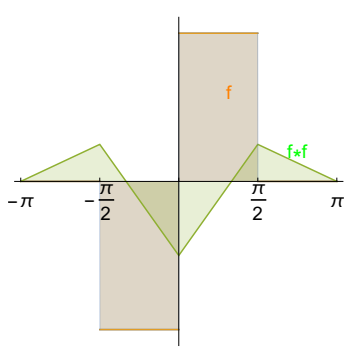


Figure 2. The convolution $g=f\star f$ in HW 32.1. We see both the odd step function f and the even convolution g.

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 33: DFT

Lecture

33.1. Remember the circular matrices like

$$Q = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} ?$$

You should. We have seen that they are orthogonal matrices for which we can find the eigenvalues explicitly. Indeed the characteristic polynomial $-\lambda^5 + 1$ has as roots the numbers $\lambda_k = e^{2\pi k i/5}$ with $k = 0, \cdots, 4$. What are the eigenvectors?

> **Problem A:** Verify that if λ is an eigenvalue of Q, then $[1, \lambda, \lambda^2, \cdots \lambda^4]^T$ is an eigenvector of Q to the eigenvalue λ.

33.2. Remember that orthogonal matrices are normal. By the spectral theorem for normal matrices we know that Q can be diagonalized $B = S^* Q S$, where S is unitary $S^* S = 1$ and B is diagonal with eigenvalues in the diagonal.

> **Problem B:** Write down the unitary matrix S for the matrix A above.

33.3. Given $n = 5$ numbers a, b, c, d, e, we can look at the matrix $A = a + bQ + cQ^2 + dQ^3 + eQ^4$.

> **Problem C:** How do you diagonalize the matrix A? What are the eigenvalues, what are the eigenvectors?

33.4. You have now already covered the discrete Fourier transform. The map, which assigns to the list $(a_1, a_2, a_3, a_4, a_5)$ the eigenvalues $\hat{a}(k) = \sum_j a_j e^{2\pi i j k/5}$ is called the **Discrete Fourier transform** abbreviated DFT. We can write

$$\hat{a} = Sa$$

where S is the unitary coordinate change matrix containing the eigenvectors as column vectors.

33.5. How can we use this? One application is the multiplication of numbers.

Problem D: Compute $32 * 45$ using school arithmetic.

33.6. Now form the matrices $A = 2 + 3Q$ and $B = 5 + 4Q$ and form $A * B$ to verify that $AB = 10 + 23Q + 12Q^2$.

$$AB = \begin{bmatrix} 2 & 3 & 0 & 0 & 0 \\ 0 & 2 & 3 & 0 & 0 \\ 0 & 0 & 2 & 3 & 0 \\ 0 & 0 & 0 & 2 & 3 \\ 3 & 0 & 0 & 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} 5 & 4 & 0 & 0 & 0 \\ 0 & 5 & 4 & 0 & 0 \\ 0 & 0 & 5 & 4 & 0 \\ 0 & 0 & 0 & 5 & 4 \\ 4 & 0 & 0 & 0 & 5 \end{bmatrix} = \begin{bmatrix} 10 & 23 & 12 & 0 & 0 \\ 0 & 10 & 23 & 12 & 0 \\ 0 & 0 & 10 & 23 & 12 \\ 12 & 0 & 0 & 10 & 23 \\ 23 & 12 & 0 & 0 & 10 \end{bmatrix}.$$

Problem D: What is $10 + 23 \cdot 10 + 12 \cdot 10^2$?

33.7. Well, this seems to be an awfully complicated way to compute the product of two numbers. Let us see what happened. We encoded the first number as a matrix A and then encoded the second matrix as a matrix B. After diagonalizing the two matrices, we just can compute their diagonal entries. This can be done fast. It turns out that the map from the number $a = (a_1, \dots, a_n)$ to its discrete Fourier transform $\hat{a} = (\hat{a}_1, \dots, \hat{a}n)$ can be computed fast.

Theorem: One can multiply two integers of length n in time $n \log(n)$

Sound

33.8. We can listen to a function by replacing Plot with Play:

```
Play[Sin[4 x] Abs[Sin[1000 x]], {x, 0, 3}]
```

Here is the function $f(x) = x$ periodically continued

```
Plot[SawtoothWave[x], {x, -Pi, Pi}]
```

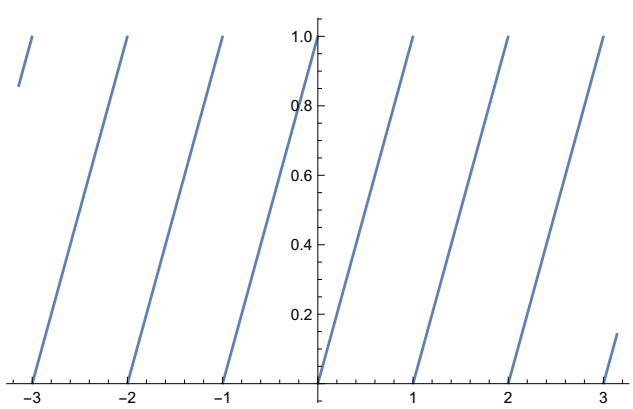


Figure 1. The Sawtooth Wave function.

```
Play[SawtoothWave[1000 x]], {x, 0, 1}]
```

It does not sound very nice. The reason is that the Fourier series of $f(x) = x$ does not decay very fast.

```
Play[Sin[Pi SawtoothWave[1000 x]], {x, -Pi, Pi}]
```

Problem F: Experiment with various functions and plot and play at least one.

33.9. If you have a sound file lying around, you can import it and look at the wave.

```
A=Import["https://www.quantumcalculus.org/sound/wiggle.wav"];
B=AudioData[A];
Audio[SampledSoundList[B, 48000]]
ListPlot[Table[B[[1, k]], {k, 1000}], Joined -> True]
```

33.10. Mathematica can play instruments. We can take the sound, represent the sound as a list of numbers and plot it:

```
A = Sound[SoundNote[1, 1, "Violin"]];
B = AudioData[A]; B1 = B[[1]]; B2 = B[[2]];
ListPlot[Table[B2[[k]], {k,500}], Joined -> True]
```

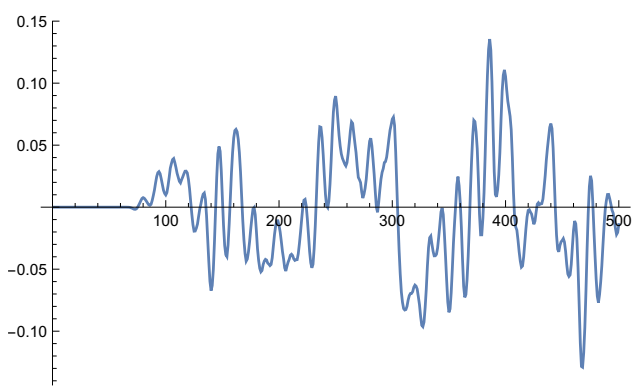


FIGURE 2. The first 500 entries in the sound wave representing a piano sample.

Homework

This homework is due on Tuesday, 4/23/2019.

Problem 33.1: The procedures Fourier and InverseFourier are implemented already in Mathematica Here are emulations showing what they do. Why does the $\sqrt{n}$ term appear in the matrix S?

```
X={5,9,7,4,2,3,2,1};
S=N[(Table[Exp[-I 2Pi k l/n],{k,0,n-1},{l,0,n-1}]/Sqrt[n])];
S.X
InverseFourier[X]
```

Now, let us get back X.

```
X={5,9,7,4,2,3,2,1};
S=N[(Table[Exp[I 2Pi k l/n],{k,0,n-1},{l,0,n-1}]/Sqrt[n])];
S.X
Fourier[X]
```

Problem 33.2: Make your own sound file (from a music piece or recording) and display part of it in Mathematica.

Problem 33.3: Mathematica has quite many MIDI instruments implemented. Use this to make your own little mini song.

Problem 33.4: Autotune is a technique which allows you to sing and never sing wrong. Explain how to implement this using Fourier theory.

Problem 33.5: Our ear can in the Cochlea do a Fourier analysis of sound. How does it work? Look it up and write a short paragraph explaining the principle.
Alternatively, if you should freak out thinking about your own ears, explain what the Mathematica code below does. Especially, why is "Reverse" used. Why is "PadRight" used?

```
FastMultiplication[x_,y_]:=Module[{X,Y,U,V,n,Q},
  X = Reverse[IntegerDigits[x]]; Y = Reverse[IntegerDigits[y]];
  n =Length[X]+Length[Y]+1; X=PadRight[X,n];    Y=PadRight[Y,n];
  U=InverseFourier[X]; V=InverseFourier[Y];
  Q=Round[Re[Fourier[U*V]*Sqrt[n]]];
  Sum[Q[[k]] 10^(k-1), {k,n}]];
x0 = 11234; y0 = 52342; FastMultiplication[x0,y0]==x0*y0
```

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 34: Heat equation

Lecture

34.1. The partial differential equation

$$f_t = f_{xx}$$

is called the **heat equation**. It is an equation for an unknown function $f(t,x)$ of two variables t and x. The interpretation is that $f(t,x)$ is the temperature at **time** t and **position** x. In order to use Fourier theory, we assume that f is a function on the interval $[-\pi,\pi]$. The problem is: given an initial heat distribution $f(0,x)$, what is the situation $f(t,x)$ at a later time? The process does what one expects from heat. It produces **diffusion**.

34.2. What does the equation tell? We have a temperature distribution $x \to f(t,x)$. The rate of change in time of this temperature is the second space derivative of f. If x is a location, where $f(t,x)$ is concave down as a function of x, this means that f_t is negative and that the function will decrease there in the near future. If $f(t,x)$ is concave up, then this means that f_t is positive, meaning that f increases there. While the partial differential equation describes a motion of a function f the set-up is as before, where we looked at the motion $v(t)$ of a vector v.

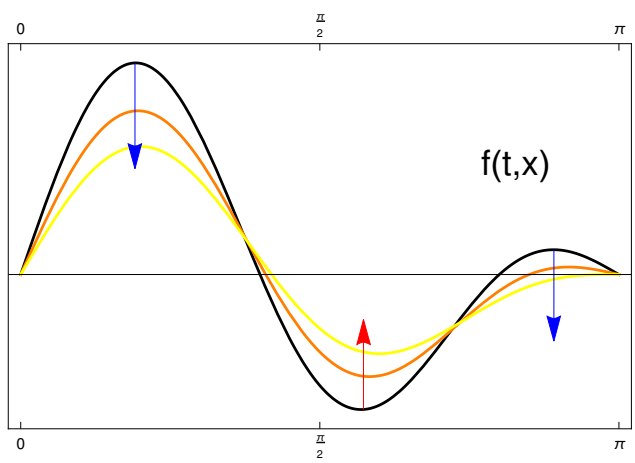


Figure 1. The heat equation describes the motion of a function. The evolution has a smoothing effect. The height of the function diffuses and settles to the average.

34.3. In order to use the **closed-form solution method** from the earlier part of the course, we write the heat equation as

$$f_t = D^2 f$$

and think of $A = D^2$ as a transformation or a matrix. Now remember what we did in the case of differential equations $x' = Ax$; we found the eigenvalues and eigenvectors of A. This is what we do here too. But we know already that $\cos(nx), \sin(nx)$ are eigenfunctions of D^2 to the eigenvalue $-n^2$. In the ordinary differential equation case, we also expressed $v(0) = c_1 v_1 + \cdots + c_n v_n$ as a sum of eigenfunctions, then wrote down the **closed-form solution** $v(t) = c_1 e^{\lambda_1 t} v_1 + \cdots + c_n e^{\lambda_n t} v_n$.

34.4. We do the same thing here and can use that the Fourier basis is an eigenbasis of D^2. This is the great discovery of Fourier:

Theorem: The solution to the heat equation is $f(t,x) = a_0 \frac{1}{\sqrt{2}} + \sum_{n=1}^{\infty} a_n e^{-n^2 t} \cos(nx) + \sum_{n=1}^{\infty} b_n e^{-n^2 t} \sin(nx)$.

Proof. For $t = 0$, we get the Fourier series of $f(0,x)$. A direct differentiation shows that $f_t = f_{xx}$. □

34.5. Fourier himself used the sin series. We could continue the function as an odd function on $[-\pi, \pi]$ and use only the sin-series. We could also continue the function as an even function on $[-\pi, \pi]$ and use the cos-series. The later is done in applications like JPG encoding of pictures as there, the average is important as it represents brightness. Let us look at an example of $f(0,x) = \sin(x)$ on $[0, \pi]$. If we continue that as an odd function, then this is just $\sin(x)$ and $f(t,x) = e^{-t^2} \sin(x)$ solves that equation. If we continue the function as an even function, then we deal with $f(x) = |\sin(x)|$.

34.6. A function of two variables $f(x,y)$ can be expanded into a Fourier series too. The Fourier basis is

$$\mathcal{B} = \{\cos(nx)\cos(ny), \sin(nx)\sin(ny), \cos(nx)\sin(my), \frac{1}{\sqrt{2}}\cos(nx), \frac{1}{\sqrt{2}}\sin(nx), 1/2\} .$$

It is here much more convenient to use the **complex basis** $\mathcal{B} = \{e^{i(nx+my)}\}$ and write

$$f(x,y) = \sum_{n,m} c_{n,m} e^{i(nx+my)}$$

with $c_{n,m} = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} f(x,y) e^{-inx-imy}\, dxdy$. The real Fourier coefficients would include terms like $a_{nm} = \frac{1}{\pi^2} \int_{-\pi}^{\pi} f(x,y) \cos(nx)\cos(ny)\, dxdy$. Let $\mathcal{X}$ be the set of functions $f(x,y)$ with the property that for every x the function $g(y) = f(x,y)$ is piecewise smooth, then the Dirichlet theorem implies that the Fourier series converge. And we have

Theorem: The heat equation $f_t = f_{xx} + f_{yy}$ with initial condition $f(0,x,y)$ is solved by $f(t,x,y) = \sum_{n,m} c_{n,m} e^{-(n^2+m^2)t} e^{i(nx+my)}$.

Proof. For $t = 0$, is the Fourier series of $f(0,x,y)$. When differentiating each part $f_{nm} = e^{-(n^2+m^2)t} e^{i(nx+my)}$ with respect to t, this gives the same as when applying the Laplacian $\Delta f = f_{xx} + f_{yy}$ on f_{nm}. □

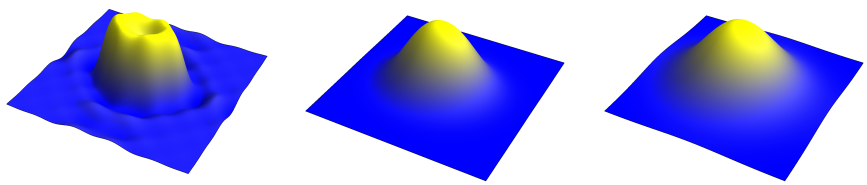


FIGURE 2. The evolution of a two dimensional heat equation is solved also explicitly using Fourier series.

EXAMPLES

34.7. Problem: Find the general solution of the modified heat equation $f_t = 3f_{xx} + f$, where $f(0)$ is 1 for $x \in [\pi/3, 2\pi/3]$ and 0 else. Let us assume that f stays zero at the boundary of $[0, \pi]$ and is continued in an odd way so that it has a sin-series. This equation is a driven heat equation which can model a **fire** which in which heat produces more fuel. **Solution:** f has a sin series with

$$b_n = \frac{2}{\pi} \int_{\pi/3}^{2\pi/3} \sin(nx)\, dx = \frac{2}{n\pi}[-\cos(2n\pi/3) + \cos(n\pi/3)] \ .$$

Now look the operator $A = 3D^2 + 1$ on the right hand side so that $f_t = Af$. What are the eigenvalues of A? Since the eigenvalues of D^2 are $-n^2$, the operator A has the eigenvalues $\lambda_n = -3n^2 + 1$. The closed-form solution is $f(t, x) = \sum_{n=1}^{\infty} b_n e^{(-3n^2+1)t} \sin(nx)$.

34.8. Problem: Let us take the same problem as before but increase the fuel strength feeding the fire $f_t = 3f_{xx} + 5f$. What happens now is that

$$f(t, x) = \sum_{n=1}^{\infty} b_n e^{(-3n^2+5)t} \sin(nx) \ .$$

The high frequency parts still die out but there is one mode which explodes now exponentially because $\lambda_n = -3n^2 + 5$. The fire takes over. Let us write down the solution in a bit more intelligible way. The coefficient $b_n = -\cos(2n\pi/3) + \cos(n\pi/3)$ is 0 for even n, It is alternating 1 and -2 for odd n. We have the initial condition

$$f(0, x) = \frac{1}{\pi}(\frac{\sin(x)}{1} - \frac{2\sin(3x)}{3} + \frac{\sin(5x)}{5} - \cdots) \ .$$

The solution of the heat equation is now

$$f(t, x) = \frac{1}{\pi}(\frac{e^{2t}\sin(x)}{1} - \frac{e^{-7t}2\sin(3x)}{3} + \frac{e^{-22t}\sin(5x)}{5} - \ldots) \ .$$

It is the e^{2t} part which renders the fire out of control.

Homework

This homework is due on Tuesday, 4/30/2019.

Problem 34.1: Solve the heat equation $f_t = 2019 f_{xx}$ on $[-\pi, \pi]$ with the initial condition $f(0,x) = 0$ if $x \in [-\pi/2, \pi/2]$ and $f(x) = \sin(2x)$ if $|x| \in [\pi/2, \pi]$.

Problem 34.2: Solve the partial differential equation $f_t = 3 f_{xxxxxx} + 5 f_{xx}$ with initial condition $f(0,x) = 22x$.

Problem 34.3: Solve the partial differential equation $f_t = -f_{xxxx} - f_{yyyy}$ with initial condition $f(0,x) = 2\cos(22x)\sin(33y) + 17\sin(12x)\sin(11y)$.

Problem 34.4: Solve the partial differential equation $f_t = f_{xx} + \sin(3t)$ with initial condition $f(0,x) = 3x + \cos(5x) + \sin(7x)$. **Hint:** First solve the homogeneous equation $f_t = f_{xx}$, then add a special solution which only depends on t.

The next problem deals with a very important partial differential equation. It looks like the heat equation and can be treated like the heat equation but its solutions behaves in a different way. It is that i which changes everything. Instead of $e^{-n^2 t}$ which goes to zero very fast, we have $e^{in^2 t}$ which is a wave.

Problem 34.5: a) The equation $i f_t = f_{xx}$ is the **Schrödinger equation**. Assume $f(0,x) = \sin(17x) + \cos(12x)$, find the solution $f(t,x)$.
b) Write down the solution to the Schrödinger equation $i f_t = f_{xx} + f_{yy} + f_{zz}$ with $f(0,x,y,z) = \sin(13x)\cos(15y)\sin(17z)$.

Figure 3. Firepit near the Harvard Science center. Contemplating heat.

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 35: Wave equation

Lecture

35.1. The partial differential equation

$$f_{tt} = f_{xx}$$

is called the **wave equation**. It is an equation for an unknown function $f(t,x)$ of two variables t and x. The interpretation is that $f(t,x)$ is the string amplitude at **time** t and position x. Again, we assume that f is a function on the interval $[-\pi,\pi]$. One problem is: given an initial string position $f(0,x)$, what is $f(t,x)$ at a later time? Another problem is to give the initial velocity $f_t(0,x)$ and determine from this the string position at time t. We can also give both the initial position and velocity in which case we just add up the solution for the initial position and the solution for the initial velocity.

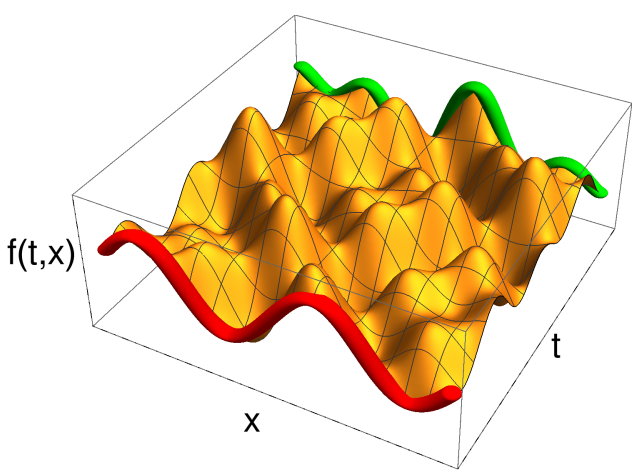


Figure 1. We see a solution $f(t,x)$ of the wave equation $f_{tt} = f_{xx}$. The initial wave front is a sin function, but there is also an initial velocity given. We outlined $f(0,x)$ and $f(1,x)$.

35.2. What is the meaning of the wave equation? We can interpret the acceleration f_{tt} as a **force** acting on the string. By Newton's law acceleration is proportional to force. That force acts in such a way that the string wants to be flattened out. But as the system not only has position but also momentum, it does not just flatten out as in

the heat case. It is a **conservation of energy** which prevents the wave from going to zero without friction:

Lemma: The **energy** $H(f) = \frac{1}{\pi}\int_{-\pi}^{\pi} f_t^2/2 + f_x^2/2\ dx$ is constant.

Proof. Using integration by parts (using that we are on a circle and so have no boundary) we have the general formula for differentiable functions

$$\int_{-\pi}^{\pi} f'(x)g'(x)\ dx = \int_{-\pi}^{\pi} -f(x)g''(x)\ dx\ .$$

We therefore have $\frac{d}{dt}H(f) = \frac{1}{\pi}\int_{-\pi}^{\pi} f_t f_{tt} + f_x f_{tx}\ dx$. Using the wave equation $f_{tt} = f_{xx}$ and the integration by parts formula $\int_{-\pi}^{\pi} f_x f_{tx}\ dx = \int_{-\pi}^{\pi} -f_{xx} f_t\ dx$, we get $\frac{1}{\pi}\int_{-\pi}^{\pi} f_t f_{xx} - f_{xx} f_t\ dx = 0$. □

35.3. For the heat equation, the solution of $x'(t) = \lambda x$ was important. It was $x(t) = e^{\lambda t}x(0)$. For the wave equation, the solution of $x''(t) = -c^2 x$ is important. It is $x(t) = x(0)\cos(ct) + x'(0)\sin(ct)/c$. We usually have given things in the form $x''(t) = \lambda x$, so that $c = \sqrt{-\lambda}$. Now, if we have an initial condition which is a Fourier basis vector like $f = \cos(nx)$, where $D^2 f = (-n^2)f$, where $\lambda = -n^2$ and $c = n$, we have the solution $f(t,x) = \cos(ct)\cos(nx)$.

Theorem: If $f_t(0,x) = 0$ and $f(0,x) = \frac{a_0}{\sqrt{2}} + \sum_{n=1}^{\infty} a_n \cos(nx) + b_n \sin(nx)$, then $f(t,x) = \frac{a_0}{\sqrt{2}} + \sum_{n=1}^{\infty} a_n \cos(nt)\cos(nx) + b_n \cos(nt)\sin(nx)$ solves the wave equation $f_{tt} = f_{xx}$.

Proof. We check that $f_t(0,x) = 0$ and that $f(0,x)$ agrees with the Fourier expansion of $f(0,x)$. □

35.4. The solution to the **harmonic oscillator** $x''(t) = -c^2 x$ also has a contribution $x'(0)\sin(ct)/c$ which takes care of the **initial velocity**. This allows us to write down the **closed-form solution** if the initial velocity is given.

Theorem: If $f(0,x) = 0$ and $f_t(0,x) = \frac{a_0}{\sqrt{2}} + \sum_{n=1}^{\infty} a_n \cos(nx) + b_n \sin(nx)$, then $f(t,x) = t\frac{a_0}{\sqrt{2}} + \sum_{n=1}^{\infty} \frac{a_n}{n}\sin(nt)\cos(nx) + \frac{b_n}{n}\sin(nt)\sin(nx)$ solves the wave equation $f_{tt} = f_{xx}$.

Proof. We see that for $t = 0$, $f(0,x) = 0$ and that $f_t(0,x)$ agrees with the Fourier expansion of $f_t(0,x)$. □

35.5. The energy can be written as $||f_x||^2/2 + ||f_t||^2/2$ which is a sum of a **potential energy** and **kinetic energy** of the string. It is custom to use the factor $1/2$ as we have in physics energy = mass times velocity squared divided by 2.

35.6. More generally, if we want to solve a partial differential equation $f_{tt} = Af$, where A is a function of D^2 we do:

A) find the eigenvalues λ_n of A and form $c_n = \sqrt{-\lambda_n}$.
B) Decompose the initial position $f(0,x)$ as a Fourier series and write down $f(t,x) = \frac{a_0}{\sqrt{2}} + \sum_{n=1}^{\infty} a_n \cos(c_n t)\cos(nx) + b_n \cos(c_n t)\sin(nx)$
C) Decompose the initial velocity $f_t(0,x)$ as a Fourier series and write down $f(t,x) = t\frac{a_0}{\sqrt{2}} + \sum_{n=1}^{\infty} \frac{a_n}{c_n} \sin(c_n t)\cos(nx) + \frac{b_n}{c_n} \sin(c_n t)\sin(nx)$.
D) If both initial position and velocity are given, we use both B) and C) and add the two solutions together.

35.7. For example, for the **driven wave equation** $f_{tt} = f_{xx} - f$ we first compute the eigenvalues of the operator $A = D^2 - 1$. Since the eigenvalues of D^2 are $-n^2$, the eigenvalues of A are $\lambda_n = -n^2 - 1$. This gives $c_n = \sqrt{n^2+1}$. If the initial wave $f(0,x) = 4\sin(222x)$ and initial velocity $f_t(0,x) = 7\sin(365x)$, then we don't have to compute the Fourier series as the functions are already Fourier series. The closed-form solution for the initial position is

$$f(t,x) = 4\cos(\sqrt{222^2+1}t)\sin(222x)$$

The closed-form solution of the initial velocity is

$$f(t,x) = 7\sin(\sqrt{365^2+1}t)\sin(365x)/\sqrt{365^2+1}\ .$$

The solution of the problem is the sum

$$f(t,x) = 4\cos(\sqrt{222^2+1}t)\sin(222x) + 7\sin(\sqrt{365^2+1}t)\sin(365x)/\sqrt{365^2+1}\ .$$

35.8. Let us assume that $f(0,x,y)$ is a given function like $f(0,x,y) = \sum_{n,m} \frac{1}{n+m}\cos(nx)\cos(my)$. and assume $f(0,t,x,y) = 0$. What is the solution of the wave equation

$$f_{tt} = f_{xx} + f_{yy}\ ?$$

The functions $\cos(nx)\cos(my)$ are eigenfunctions of $Af(x,y) = f_{xx} + f_{yy}$ with eigenvalues $\lambda_n = -n^2 - m^2$. The solution is

$$f(t,x,y) = \sum_{n,m} \frac{\cos(\sqrt{n^2+m^2})}{n+m}\cos(nx)\cos(my)\ .$$

Examples

35.9. What is the solution of the driven wave equation $f_{tt} = f_{xx} - f + 6t$ if $f(0,x) = \sum_n \frac{1}{n^3}\sin(nx)$. We first solve the **homogeneous problem**

$$f_{tt} = f_{xx} - f\ .$$

Since the right hand side is Af with $A = D^2 - 1$, which has the eigenvalues $-n^2 - 1$, the solution is

$$f(t,x) = \sum_n \frac{\cos(\sqrt{n^2+1})}{n^3}\sin(x)\ .$$

A special solution which does not depend on x satisfies $f_{tt} = 6t$ which has the solution t^3. The final solution is

$$f(t,x) = t^3 + \sum_n \frac{\cos(\sqrt{n^2+1})}{n^3}\sin(x)\ .$$

Homework

This homework is due on Tuesday, 4/30/2019.

Problem 35.1: A piano string is fixed at the ends $x = 0$ and $x = \pi$ and is initially undisturbed $f(0, x) = 0$. The piano hammer induces an initial velocity $f_t(x, 0) = g(x)$ onto the string, where $g(x) = \sin(3x)$ on the interval $[-\pi/2, \pi/2]$ and $g(x) = 0$ on $[\pi/2, \pi]$ or $[-\pi, -\pi/2]$. a) How does the string amplitude $f(t, x)$ move, if it follows the wave equation $f_{tt} = f_{xx}$? b) Now we replace the piano by a **harpsichord**, where the string is plucked. In that case $f(0, x) = g(x)$ is given and $f_t(0, x) = 0$.

Problem 35.2: A laundry line is excited by the wind. It satisfies the differential equation $f_{tt} = f_{xx} + \cos(t) + \cos(3t)$. Assume that the amplitude u satisfies initial position $f(0, x) = x$ and $f_t(0, x) = 4\sin(5x) + 10\sin(6x)$. Find the function $f(t, x)$ which satisfies the differential equation.

Problem 35.3: Solve the partial differential equation $f_{tt} = -f_{xxxx} + f_{xx}$ with initial condition $f_t(0, x) = x^3$ and $f(0, x) = x^3$.

Problem 35.4: a) Solve the wave equation $f_{tt} = 9f_{xx}$ on $[0, \pi]$ with the initial condition $f(0, x) = \max(\cos(x), 0)$.
b) Solve the wave equation $f_{tt} = f_{xx} + f_{yy}$ with $f(0, x, y) = 9\sin(3x)\cos(5y)$.

Problem 35.5: Verify that the energy $\frac{1}{\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f_t^2/2 + f_x^2/2 + f_y^2/2 \, dxdy$ is invariant under the evolution of the wave equation $f_{tt} = f_{xx} + f_{yy}$.

Figure 2. Oliver playing a harpsichord in the Swiss mountains. The wood stove to the right is from 1883, the house built in 1864. The 360 degree camera was placed in a tray. We see its boundary as a piecewise smooth function of the direction angle.

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 36: Discrete PDE

Lecture

36.1. We have seen the Fourier theory allowed to solve the **heat equation**

$$f_t = -Lf \ ,$$

where $L = -D^2$ is the second derivative operator. The negative sign is added so that $-D^2$ has non-negative eigenvalues. The reason why things worked out so nicely was that the Fourier basis was an eigenbasis of D^2. Indeed, $L\sin(nx) = (-n^2)\sin(nx)$ and $L\cos(nx) = (-n^2)\cos(nx)$ and $L\frac{1}{\sqrt{2}} = 0$. We got a **closed-form solution** of the heat equation by writing the initial heat as a Fourier series then evolving each eigen function to get $f(t,x) = a_0\frac{1}{\sqrt{2}} + \sum_{n=1}^{\infty} a_n e^{-n^2 t}\cos(nx) + b_n e^{-n^2 t}\sin(nx)$.

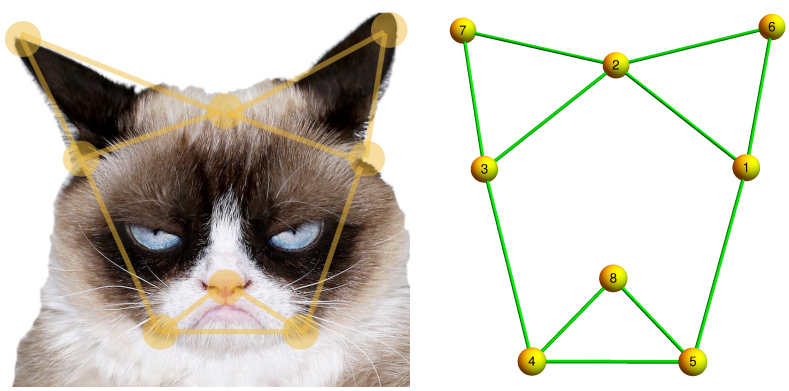


Figure 1. The "grumpy cat graph". The eigenvalues of the Kirchhoff Laplacian L are $\{3+\sqrt{5}, 3+\sqrt{3}, 4, 3, 3, 3-\sqrt{3}, 3-\sqrt{5}, 0\}$.

36.2. The same idea works also in a discrete framework when space is a graph. The analogue of the Laplacian is now the Kirchhoff matrix $L = A - B$, where A is the **adjacency matrix** and B is the diagonal matrix containing the vertex degrees. You have proven last semester that the eigenvalues are non-negative. The reason was that L could be written as d^*d for the gradient matrix d so that $\lambda(v,v) = (Lv,v) = (d^*dv,v) = (dv,dv)$ implying that $\lambda = (v,v)/(dv,dv) = |v|^2/|dv|^2 \geq 0$. The **discrete heat equation**

$$x' = -Lx$$

is now a discrete dynamical system we have seen before.

36.3. We can run a partial differential equation on any graph. Let's take the "Grumpy Cat Graph". It is especially fun to run the Schrödinger equation

$$if_t = Lf$$

on that graph. It is **Schrödinger's cat**. Grumpy cat has $v_0 = 8$ vertices and $v_1 = 11$ edges and $v_2 = 3$ triangles as ears and the snout. Its Euler characteristic $v_0 - v_1 + v_2 = 0$ is zero, one reason why the cat is so grumpy. It is also unhappy not knowing whether it is dead or alive and because his friend, "Arnold the cat" can live on a doughnut.

36.4. The Laplacian of the Grumpy cat graph encodes the graph because the entries -1 tell which vertices are connected.

$$L = \begin{bmatrix} 3 & -1 & 0 & 0 & -1 & -1 & 0 & 0 \\ -1 & 4 & -1 & 0 & 0 & -1 & -1 & 0 \\ 0 & -1 & 3 & -1 & 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & 3 & -1 & 0 & 0 & -1 \\ -1 & 0 & 0 & -1 & 3 & 0 & 0 & -1 \\ -1 & -1 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & -1 & -1 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -1 & -1 & 0 & 0 & 2 \end{bmatrix} .$$

The eigenvalues of L are $\{\lambda_1 = 3 + \sqrt{5}, \lambda_2 = 3 + \sqrt{3}, \lambda_3 = 4, \lambda_4 = 3, \lambda_5 = 3, \lambda_6 = 3 - \sqrt{3}, \lambda_7 = 3 - \sqrt{5}, \lambda_8 = 0\}$. We give the eigenvectors $v_3 = [1, 1, 1, -1, -1, -1, -1, 1], v_4 = [-1, 0, 0, 0, -1, 1, 0, 1], v_5 = [0, 0, -1, -1, 0, 0, 1, 1]$ and $v_8 = [1, 1, 1, 1, 1, 1, 1, 1]$.

Theorem: For a connected graph, the solution $x(t)$ to the heat equation converges to a constant function which is the average value of $x(0)$.

Problem A: Prove this theorem. You can use that all eigenvalues of L are positive except one which is 0.

Problem A': Solve the heat equation for the grumpy cat with initial condition $f(0) = v_3 + 5v_4 + 2v_5$.

36.5. Let us now look at the **discrete wave equation**

$$f_{tt} = -Lf \ ,$$

where again L is the discrete Laplacian of a connected graph. Assume λ_k are the eigenvalues of L and v_k the eigenvectors.

Theorem: The function $f(t) = \sum_k c_k \cos(\sqrt{\lambda_k} t) v_k$ solves the discrete wave equation with initial condition $f(0) = \sum_k c_k v_k$.

Problem B: Verify this theorem by verifying that each part in the sum satisfies the wave equation.

Problem B': Solve the wave equation for the grumpy cat with initial condition $f(0) = v_3 + 5v_4 + 2v_5$.

Theorem: The function $f(t) = \sum_k c_k \sin(\sqrt{\lambda_k} t)/\sqrt{\lambda_k} v_k$ solves the wave equation with initial velocity $f'(0) = \sum_k c_k v_k$.

Problem C: Prove this theorem.

Problem C': Solve the wave equation for the grumpy cat with initial condition $f_t(0) = v_3 + 5v_4 + 2v_5$.

Problem D: Formulate the theorem for the discrete Schrödinger equation $if_t = Lf$.

Problem D': Solve the Schrödinger equation for the grumpy cat with initial condition $f(0) = v_3 + 5v_4 + 2v_5$.

36.6. Partial differential equations which are not linear are hard. An example is the **sine-Gordon equation**

$$f_{tt} = -Lf - c\sin(f) \ ,$$

where c is a constant. This can also be considered in the discrete, where L is the Kirchhoff matrix. One of the simplest examples is when the graph is the complete graph with 2 vertices. In that case

$$L = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

If $f = (x, y)$, then

$$\begin{aligned} x'' &= -x + y - c\sin(x) \\ y'' &= x - y - c\sin(y) \end{aligned}$$

This is a non-linear system if c is different from zero.

Problem E: Solve this system for $c = 0$ in the case when $(x(0), y(0)) = (2, 1)$ and $(x'(0), y'(0)) = (0, 0)$.

Problem F: Verify that the energy of the sine-Gordon equation, $H(x, y, x', y') = (x'^2 + y'^2)/2 + (x - y)^2/2 + c\cos(x) + c\cos(y)$, is constant.

Homework

This homework is due on Tuesday, 4/30/2019.

In the next three problems, we take G be the complete graph with 3 vertices.

Problem 36.1: Write down the discrete heat equation $f_t = -Lf$ and find the closed-form solution $f(t)$ with $f(0) = (0, 2, 1)$.

Problem 36.2: Write down the discrete wave equation $f_{tt} = -Lf$ and find the closed-form solution $f(t)$ with $f(0) = (0, 2, 1)$ and $f_t(0) = (0, 0, 0)$.

Problem 36.3: Write down the discrete Schrödinger equation $if_t = -Lf$ and find the closed-form solution $f(t)$ with $f(0) = (0, 2, 1)$.

Problem 36.4: Remember that if f is a function on vertices of a graph, then df is a function on the edge by $df((a, b)) = f(b) - f(a)$. Verify that the energy $H = \sum_e df(e)^2/2 + \sum_v f_t(v)^2/2$ is time invariant under the wave equation $f_{tt} = -Lf$. Hint: You can use that $L = d^*d$, where d is a $m \times n$ matrix, where n is the number of vertices and m the number of edges. Use that $(df, df) = (d^*df, f) = (Lf, f)$.

Problem 36.5: Pick a graph of your choice, write down the matrix L and write down a closed-form solution for
a) The discrete heat equation.
b) The discrete wave equation
c) The discrete Schrödinger equation.

Oliver Knill, knill@math.harvard.edu, Math 22b, Harvard College, Spring 2019

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 37: Last unit

37.1. In this last lecture, we look at some key points and connections. Of course, this is just an attempt. It is your task to do this yourself when reviewing the material. It can be helpful to see things from far also.

Objects and Arrows

37.2. It is important that you know what are **objects** of a theory and what are relations called **morphisms** between the objects here given by **transformations**. It is a fancy point of view to see objects and arrows between objects. The subject is called **category theory**. It is a field which builds bridges between different subjects. It is a bit abstract at first like set theory but it starts to get hold also in computer science. [1]

37.3. In linear algebra, we deal with the **category of vector spaces**. The objects are the **linear spaces**. Examples are $\mathbb{R}^n$ or $M(n,m)$ or **linear space of functions** like $C^\infty([-\pi,\pi])$. And then we looked at linear transformations which are implemented by matrices or operators.

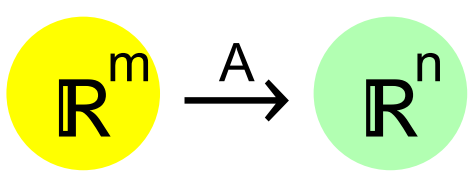


Figure 1. A matrix $A \in M(n,m)$ defines a linear map from $\mathbb{R}^m$ to $\mathbb{R}^n$. It is a morphism meaning that it respects the structure addition and scalar multiplication of the vector space.

Linear Spaces	**Linear Transformations**
Vector spaces	Matrices
Function spaces	Operators

37.4. Important examples of spaces are image and kernel of a transformation. Later in the course, we looked at **solution spaces of differential equations**. Also these spaces can be described in the form of kernels of transformations

[1] B. Milewski: Category Theory for Programmers, 2018

Time

to not only be able to survive in the short term but also to plan for the long term

37.5. It is pivotal for us to be able to predict the future. We would like to know the development of the weather or seasons. We need to plan ahead to avoid problems. It is a sign of intelligence to not only be able to survive in the short term but to also to worry about the long term prosperity. We might, in the future, change the trajectory of an asteroid targeting the earth, change the energy consumption habits to avoid a climate disaster, or to adopt a long term financial planning strategy for times when betting on exponential growth is no option any more. In order to make informed decisions, we have to be able to gather data, fit these data so that we can make a model or theory, then use this knowledge to predict the situation at a later time.

37.6. The science of time is called **dynamical systems theory**.

Discrete dynamical systems	$x(t+1) = Ax(t)$
Ordinary differential equations	$x'(t) = Ax(t)$
Partial differential equations	$f_t = Af.$

Solving equations

37.7. A major theme in this course was the problem to **solve equations**:

Systems of linear equations	$Ax = b.$
Find roots of a polynomial	$f_A(x) = 0$
Find least square solution	$(A^T A)^{-1} A^T b$
Find equilibrium points	$f(x, y) = 0, g(x, y) = 0$
Solve ordinary differential equations	$f'' = -c^2 f.$
Solve partial differential equations	$f_{tt} = f_{xx}$

Invariants

37.8. Invariants are quantities which do not change under coordinate change or do not change under time evolution. Here are examples:

Trace of a matrix
Determinant of a matrix
Eigenvalues of a matrix
Geometric multiplicity of eigenvalues
algebraic multiplicity of eigenvalues
energy
Markov property

37.9. A nice big picture had been painted by **Emmy Noether**. She related **invariant quantities** with **symmetries**. In linear algebra, symmetries pop up at various places. Examples:

Symmetry	Invariant
Similarities like $A \to S^{-1}AS$	Eigenvalue
even function	$b_n = 0$
odd function	$a_n = 0$

Building blocks

37.10. A nice idea is to split things up into smaller parts, solve the smaller parts, then put things together. **Linearity** makes this possible. If we have two solutions to a linear equation, we can add them and get a solution again.

37.11. In linear algebra, the situation is particularly nice if we have spaces which are invariant. If a transformation on $\mathbb{R}^4$ for example which preserves the xy and the zw plane is written in **block form**. If we can find a **basis** for which all invariant blocks are one dimensional, then we have **diagonalized the situation**. In that case we can completely solve the system using **closed-form solutions.**

37.12. We use that in one dimensions

equation	solution
$x(t+1) = \lambda x(t)$	$x(t) = \lambda^t x(0)$.
$x'(t) = \lambda x(t)$	$x(t) = e^{\lambda t} x(0)$.
$x''(t) = -c^2 x(t)$	$x(t) = \cos(ct)x(0) + \sin(ct)\frac{x'(0)}{c}$.

Examples

37.13. Good examples are pivotal for understanding and ideas and developments. Fourier theory has lots of open ends. Let us look at stranger series:

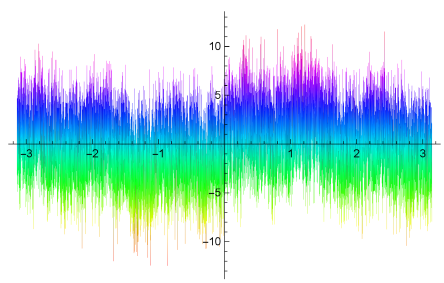


Figure 2. The Fourier series $\sum_{n=1}^{\infty} \sin(n!x)$ converges for every $x = \pi p/q$. We see its graph. It was first studied by Riemann.

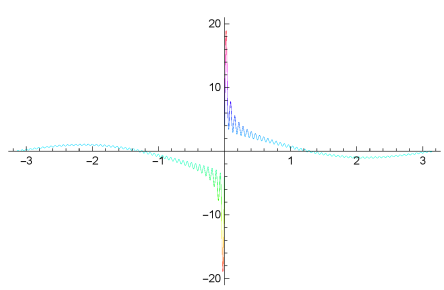


Figure 3. The function $\sum_{n=1}^{\infty} \sin(nx)/\log(n)$, an example due to Fatou. It converges everywhere, but is not even Lebesgue integrable.

37.14. Here is a challenge. Can you see what $\sum_{n=1}^{\infty} \frac{\sin(n)}{n}$ is? Hint. Try to find a formula for the Fourier sin series with $b_n = 1/n$.

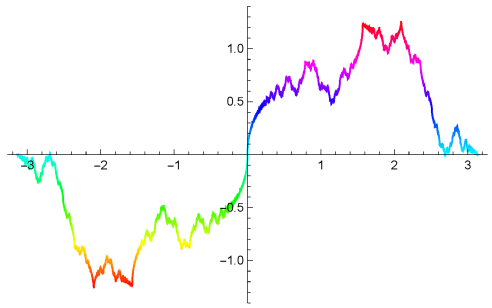


FIGURE 4. The function $\sum_{n=1}^{\infty} \sin(n^2 x)/n^2$ is due to Riemann who told his students that this might be a function with no derivative at any point. Hardy proved later that it is indeed nowhere differentiable except perhaps at points $\pi p/q$ with odd integers p and q. Only in 1970, it was shown that it is differentiable at those points.

History

37.15. The story of linear algebra, differential equations and Fourier theory cover an interesting time in mathematics of the 19th century. Between the time of Euler and Riemann, the understanding of functions completely changed.

37.16. The Dirichlet convergence theorem for example is not only an important mathematical result; it is also an important turning point in the history of mathematics as it led to a different level of rigor. To cite from the book of Kahane and Lemarié-Rieusset on Fourier series, "the article of Dirichlet on Fourier series is a turning point in the theory and also in the way mathematical analysis is approached and written."

37.17. Even Dirichlet, who worked with piece wise monotone functions had the impression that "these are all the functions encountered in nature". But more generalizations were necessary to deal more fractal type structures or have completeness as needed in quantum mechanics.

Beyond

37.18. After multivariable calculus and linear algebra a couple of interesting fields to explore like functional analysis, complex analysis, algebraic geometry or measure theory. One of the objects studied in functional analysis are Hilbert spaces and Banach spaces, which are all linear spaces but which also feature an inner product or norm. Fourier theory can then be formulated in a Hilbert space like $L^2(\mathbb{T})$ where the inner product what we have considered but where much for functions are allowed.

37.19. Mathematical subjects like statistics, discrete math, graph theory, theory of computation, dynamical systems, and numerical analysis all benefit from a solid knowledge of calculus and linear algebra. There are also immediate applications in other sciences like biology, chemistry, computer science, physics or astronomy. Of course, the engineering applications in computer graphics, artificial intelligence, economics, finance and even good old rocket science are enormous.

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 38: Checklist for Final

Glossary mostly since unit 28

- **linear space** X If f, g are in X, then $f+g, \lambda g$ are in X. Especially, 0 is in X.
- **linear map** $T(f+g) = T(f) + T(g), T(\lambda f) = \lambda T(f)$ and $T(0) = 0$.
- **diagonalization** possible if A is symmetric or normal or if spectrum is simple.
- **trace** $\text{tr}(A) =$ sum of diagonal entries $= \lambda_1 + \lambda_2 + \cdots + \lambda_n$.
- **determinant** $\det(A) =$ product of diagonal entries $= \lambda_1 \cdot \lambda_2 \cdots \lambda_n$.
- **diagonalization** $B = S^{-1}AS$, with diagonal B, where S contains eigenbasis of A.
- **Jordan normal form** $\lambda 1 + B$, where B has only super diagonal entries 1.
- **differential operator** like $p(D) = D^2 + 3D$ then $p(D)f = g$ reads $f'' + 3f' = g$.
- **homogeneous ODE** $p(D)f = 0$. Example: $f'' + 3f' = 0$.
- **inhomogeneous** ODE $p(D)f = g$. Example: $f'' + 3f' = \sin(t)$.
- **first order linear** $f' = \lambda f$, $f(t) = e^{\lambda t} f(0)$.
- **operator method** $((D-\lambda)^{-1}g)(t) = Ce^{\lambda t} + e^{\lambda x} \int_0^t e^{-\lambda s} g(s)\, ds$.
- **inner product** $\langle f, g \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) g(x)\, dx$, **length** $\sqrt{\langle f, f \rangle} = ||f||$.
- **Fourier series** $f(x) = a_0/\sqrt{2} + \sum_{n=1}^{\infty} a_n \cos(nx) + b_n \sin(nx)$.
- **Fourier basis** $1/\sqrt{2}, \cos(nx), \sin(nx)$ for piecewise smooth functions on $[-\pi, \pi]$.
- **Fourier coefficients** $a_0 = \langle f, 1/\sqrt{2} \rangle$, $a_n = \langle f, \cos(nx) \rangle$, $b_n = \langle f, \sin(nx) \rangle$.
- **odd case** use $b_n = \frac{2}{\pi} \int_0^{\pi} f(x) \sin(nx)\, dx$.
- **even case** use $a_n = \frac{2}{\pi} \int_0^{\pi} f(x) \cos(nx)\, dx$
- **Parseval identity** $a_0^2 + \sum_{n=1}^{\infty} a_n^2 + b_n^2 = ||f||^2$
- **stability** $|\lambda_i| < 1$ for discrete systems and $\text{Re}(\lambda_i) < 0$ for continuous systems.
- **nonlinear differential equation** $x' = f(x, y), y' = g(x, y)$.
- **equilibrium points** points, where $f(x, y) = g(x, y) = 0$.
- **X-nullclines** $f(x, y) = 0$. **Y-nullclines** are curves, where $g(x, y) = 0$.
- **Jacobian** $\begin{bmatrix} f_x(x_0, y_0) & f_y(x_0, y_0) \\ g_x(x_0, y_0) & g_y(x_0, y_0) \end{bmatrix}$.
- **heat equation** $f_t = \mu D^2 f$ solution example $\sum_{n=1}^{\infty} b_n e^{-n^2 \mu t} \sin(nx)$
- **wave** $f_{tt} = c^2 D^2 f$ solution $\sum_{n=1}^{\infty} b_n \cos(nct) \sin(nx)$ or $\sum_n \frac{b_n}{nc} \sin(nct) \sin(nx)$
- **generalized heat** $f_t = p(D)$ has solution $f(x, t) = \sum_{n=1}^{\infty} b_n e^{\lambda_n t} \sin(nx)$.
- **generalized wave** $f_{tt} = p(D)$ has solution $f(x, t) = \sum_{n=1}^{\infty} b_n \cos(\sqrt{-\lambda_n} t) \sin(nx)$ (position) or $\sum_n \frac{b_n}{\sqrt{-\lambda_n}} \sin(\sqrt{-\lambda_n} t) \sin(nx)$, with λ_n eigenvalue of $p(D)$ (velocity). Add both, if both position and velocity are given.
- **Parseval identity** The Pythagoras identity for Fourier $||f||^2 = a_0^2 + \sum_n a_n^2 + b_n^2$.

Key Points

- **Transformations** Relate a geometric transformation with matrix: look at columns!.
- **Determinants** Laplace, Row reduce, patterns, partitioned, triangular, eigenvalues.
- **Determinants** are compatible with matrix multiplication $\det(AB) = \det(A)\det(B)$.
- **Data fitting** First write down system, then use least square solution formula.
- **Diagonalization** is possible for normal matrices OR if all eigenvalues are different.
- **Similarity** see whether two matrices are similar.
- **Jordan normal form** is defined for all matrices.
- **QR decomposition** is defined for all matrices A with independent columns.
- **stability** Eigenvalues determine the asymptotic stability. $|\lambda| < 1$ or $\mathrm{Re}(\lambda) < 0$.
- **closed form solution** use eigenbasis both in the discrete and continuous case.
- **basic differential equations** $\boxed{f' = \lambda f}$, $\boxed{f'' + c^2 f = 0}$.
- **Fourier basis** diagonalizes D^2. Watch even and odd case. For PDE's, use sin-series.
- **inhomogeneous cases** are solved by cookbook.
- **nonlinear systems** can be understood by analyzing equilibrium points.
- **Markov matrices** have a Perron-Frobenius eigenvector 1. Can be multiplied.

Skills

- Solve systems of linear equations. Find kernel and image.
- Understand linear spaces, linear maps. Distinguish whether linear or not.
- Find eigenvalues/eigenvectors. Use tricks of large kernels, sums of rows are constant.
- When are matrices similar? Criteria for similarity and contradicting it.
- Fit data with functions. Use the least square solution $(A^T A)^{-1} A^T b$ of $Ax = b$.
- Make QR decomposition of a matrix.
- Algebra of complex numbers: Add, subtract, multiply, take roots.
- Solve discrete dynamical systems $x(n+1) = Ax(n)$. Closed form.
- Solve continuous dynamical systems $x' = Ax$. Closed form.
- Solve differential equations $p(D)f = g$ by factoring p or using "cookbook".
- Decide stability for continuous and discrete dynamical systems.
- Analyze nonlinear systems: equilibrium points, null-clines, stability.
- Match phase space with system. Both linear and nonlinear.
- make Fourier synthesis of function $f(x)$ on $[-\pi, \pi]$.
- Know that Fourier basis diagonalizes D^2 or $p(D^2)$ like $D^2 - D^4 + 1$.
- Apply Parseval to relate Fourier coefficients with length $||f||$ of f.
- Solve heat type equations $f_t = p(D)f + g(t)$ with closed form solution.
- Solve wave type equations $f_{tt} = p(D)f + g(t)$ with closed form solution.
- Find projection onto the image of a matrix.

Closed form solutions

Partial differential equations like the **heat equation** $f_t = D^2 f$ or modifications like $f_t = (D^2 - D^4)f$ or $if_t = D^2 f$ or the **wave equation** $f_{tt} = D^2 f$ or modifications like $f_{tt} = (D^2 - D^4)f$ are solved with the Fourier basis $\sin(nx), \cos(nx), 1/\sqrt{2}$ which is an eigenbasis for D^2 or modifications like $p(D) = D^4 + 2D^6 + 4$. To do so, we use $\boxed{f'(t) = \lambda f(t)}$ with solution $f(t) = f(0)e^{\lambda t}$ or the **harmonic oscillator** $\boxed{f''(t) = -c^2 f(t)}$ with solution $f(0)\cos(ct) + f'(0)\sin(ct)/c$, where $c = \sqrt{-\lambda}$.

For discrete dynamical systems, write the initial condition as a linear combination of eigenvectors, then write down the solution $\sum_{k=1}^{n} c_k \lambda_k^t v_k$.

For continuous dynamical systems, write the initial condition as a linear combination of eigenvectors, then write down the solution $\sum_{k=1}^{n} c_k e^{\lambda t} v_k$.

Words of wisdom

- "Columns are the basis of an image"
- "Round and round you go with circular matrices."
- "Laplace row reduce ..."
- "If you feel tears, think shears".
- "PDE's are solved easily with Fourier!"
- "Odd functions provoke sins."
- "Even functions have a cause."
- "$f'(t) = \lambda f(t)$ is the mother of ODE's."
- "$f''(t) = -c^2 f(t)$ is the father of ODE's."
- "Oh, PDE, oh PDE, solved easily with Fourier."

Type of matrices

- projection dilations
- reflection dilations
- rotation dilations
- dilations
- shear dilations
- SU(2) matrices
- The magic matrix
- Circular matrices
- normal matrices $A^T A = AA^T$.

People

Since second midterm:

- **Fourier** (series)
- **Dirichlet (kernel)**
- **Parseval** (identity)
- **Pythagoras (tree)**
- **Schroedinger (cat)**
- **Arnold** (cat)
- **Grumpy** (cat)
- **Perron and Frobenius (maximal eigenvalue)**

- ☐ **Markov (matrix)**
- ☐ **Lorenz (attractor)**
- ☐ **Lyapunov (exponent)**
- ☐ **Goldbach (conjecture)**
- ☐ **Feigenbaum (universality)**

From Second midterm:

- ☐ **Murray** (system)
- ☐ **Euler** (Euler identity)
- ☐ **Grothendieck** (rising sea)
- ☐ **von Neumann** (wiggle)
- ☐ **Wigner** (wiggle)
- ☐ **Jordan** (Jordan normal form)
- ☐ **Laplace** (Laplace expansion)
- ☐ **Kac** (Can one hear a drum?)
- ☐ **Gordon-Webb-Wolpert** (No one can't!)
- ☐ **Leonardo Pisano** (Fibonacci Rabbits)
- ☐ **Lyapunov** (Lyapunov exponent)
- ☐ **Arnold** (cat map)
- ☐ **Cayley and Hamilton** (theorem)
- ☐ **Lorentz** (Lorentz system)

From First midterm:

- ☐ **Feynman** (examples)
- ☐ **Peano** (Peano axioms)
- ☐ **Euclid** (Axioms of geometry)
- ☐ **Gram** (QR)
- ☐ **Schmidt** (QR)
- ☐ **Gauss** (row reduction)
- ☐ **Jordan** (row reduction)
- ☐ **Hamilton** (quaternions)
- ☐ **Lagrange** (four square theorem)
- ☐ **Glashow, Salam, Weinberg** (electoweak unification)
- ☐ **Landau** (complexity)

Proof seminar

- ☐ **Axiom systems** What is a monoid. What is a group. What is a linear space.
- ☐ **Unitary matrices** What is $SO(3)$, $SU(2)$?
- ☐ **Complexity** What does $O(x^3)$ mean?
- ☐ **Raising sea** Why is it good to have a theory?
- ☐ **Spectra** Isospectral drums or graphs.
- ☐ **Golden mean** What is it? Where does it appear?
- ☐ **Chaos** What is the Lyapunov exponent?
- ☐ **Circular matrices** For Discrete Fourier transform.

- **Cookbook** Why does it work?
- **Dirichlet's proof** Know some ingredients like what done in HW
- **Applications of Fourier:** Sound, tomography, fast multiplication, number theory.

Oliver Knill, knill@math.harvard.edu, Math 22b, Harvard College, Spring 2019

1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	

Name:

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Total:

Unit 38: Final Exam Practice

Problems

Problem 38P.1) (10 points):
a) Write down the Laplacian L of the complete graph with four vertices and find its eigenvalues and eigenvectors.
b) Solve the wave equation $f_{tt} = -Lf$ on the graph with initial position $f(0) = [2, 4, 3, 3]$.

Problem 38P.2) (10 points):
a) The set of all 800×600 gray scale pictures are a subspace of a large vector space $M(m, n)$ What is the dimension of this space?
b) The map $(x, y, z, w) \to (y, z, w, x)$ defines a linear transformation on $\mathbb{R}^4$. What is its determinant?
c) The set of all words in the alphabet $A - Z$ feature an addition $+$ given by concatenation. The zero element 0 is the empty word. What is the name of this algebraic structure?
d) Which of the three following people is not associated to an axiom system *Euclid*, *Peano*, *Jordan*.
e) What does the rank-nullity theorem tell for a $n \times m$ matrix?
f) If S is the eigenbasis of a $n \times n$ matrix A, what can you say about $S^{-1}AS$?
g) If A is a $n \times 1$ matrix and $A = QR$ is the QR-decomposition of A, what shape does the matrix R have?
h) What is the name of a matrix A with complex entries for which $\overline{A}^T A = 1$.
i) What is the determinant of a matrix $A \in SU(2)$?
j) Which of the fundamental forces are associated to $SU(2)$. The electromagnetic, the weak, the strong or the gravitational force?

Problem 38P.3) (10 points):
We look for 4 numbers x, y, z, w. We know their sum is 20 and that their "super sum" $x - y + z - w$ is 10. As a matter of fact these two equations form a system $Ax = b$ which defines a 2-dimensional plane V in 4-dimensional space.
a) (6 points) Find the solution space of all these numbers by row reducing its augmented matrix $B = [A|b]$ carefully.

$$B = \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 20 \\ 1 & -1 & 1 & -1 & 10 \end{array}\right].$$

b) (4 points) Find two linearly independent vectors which are perpendicular to the kernel of A.

Problem 38P.4) (10 points):
People on social media have been in war about expressions like $\boxed{2x/3y - 1 \text{ if } x = 9 \text{ and } y = 2}$. Computers and humans disagree: most humans get 2, while most machines return 11. A psychologist investigates whether the size of the numbers influences the answer and asks people. This needs data fitting: using the least square method, find those a and b such that

$$\frac{ax}{3y} - b = 2$$

best fits the data points in the following table:

x	y
9	3
6	1
-3	1
0	1

Problem 38P.5) (10 points):

Let $\vec{x} = \begin{bmatrix} v \\ e \\ f \end{bmatrix}$ denote the number of vertices, edges and faces of a polyhedron. During a **Barycentric refinement**, this vector transforms as

$$A\vec{x} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 6 \\ 0 & 0 & 6 \end{bmatrix} \vec{x} \, .$$

a) (5 points) Verify that $\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, and $\vec{v}_3 = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$ are eigenvectors of A and find their eigenvalues.

b) (5 points) Write down a closed form solution of the discrete dynamical system $\vec{x}(t+1) = A\vec{x}(t)$ with the initial condition $\begin{bmatrix} v \\ e \\ f \end{bmatrix} = \begin{bmatrix} 7 \\ 12 \\ 6 \end{bmatrix}$.

Problem 38P.6) (10 points):

The **Arnold cat map** is $T\vec{v} = A\vec{v}$ where

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \, .$$

It is an icon of chaos theory.

a) (2 points) What is the characteristic polynomial of A?

b) (2 points) Find the eigenvalues of A.

c) (2 points) Find the eigenvectors of A.

d) (2 points) Is the discrete dynamical system defined by A asymptotically stable or not?

e) (2 points) Write down an orthogonal matrix S and a diagonal matrix B such that $B = S^{-1}AS$.

Problem 38P.7) (10 points):
The following configuration is called the "Beacon Oscillator" in the **Game of Life**.

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

a) (2 points) What is the rank and the nullity of A?
b) (4 points) Find a basis for the kernel and a basis for the image of A.
c) (4 points) The following matrix is called the "glider configuration" in the **Game of life**.

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

Find the inverse of A using row reduction.

Problem 38P.8) (10 points):
Remember to give computation details. Answers alone can not be given credit.
a) (2 points) The following matrix displays the solution of the Cellular automaton 10. Find its determinant

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

b) (2 points) Find the determinant of

$$B = \begin{bmatrix} 0 & 0 & 3 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 0 & 5 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

c) (2 points) Find the determinant of

$$C = \begin{bmatrix} 1 & 2 & 3 & 8 & 8 \\ 4 & 5 & 0 & 8 & 8 \\ 6 & 0 & 0 & 8 & 8 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 & 3 \end{bmatrix}.$$

d) (2 points) Find the determinant of

$$D = \begin{bmatrix} 11 & 2 & 3 & 2 & 1 \\ 1 & 12 & 3 & 2 & 1 \\ 1 & 2 & 13 & 2 & 1 \\ 1 & 2 & 3 & 12 & 1 \\ 1 & 2 & 3 & 2 & 11 \end{bmatrix}.$$

e) (2 points) Find the determinant of $E = 2Q + 5Q^{-1} + 7I$: (you can leave it in terms of eigenvalues of the basic circulant matrix Q you have seen. No simplifications are required):

$$E = \begin{bmatrix} 7 & 2 & 0 & 0 & 5 \\ 5 & 7 & 2 & 0 & 0 \\ 0 & 5 & 7 & 2 & 0 \\ 0 & 0 & 5 & 7 & 2 \\ 2 & 0 & 0 & 5 & 7 \end{bmatrix}.$$

Problem 38P.9) (10 points):
Find the general solution to the following differential equations:
a) (1 point)

$$f'(t) = 1/(t+1)$$

b) (1 point)

$$f''(t) = e^t + t$$

c) (2 points)

$$f''(t) + f(t) = t + 2$$

d) (2 points)

$$f''(t) - 2f'(t) + f(t) = e^t$$

e) (2 points)

$$f''(t) - f(t) = e^t + \sin(t)$$

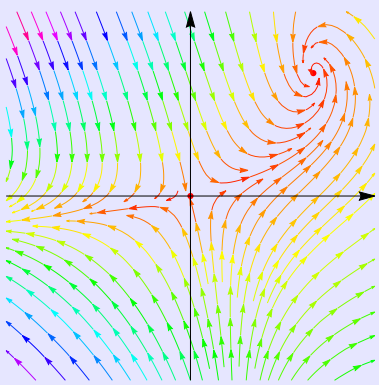
f) (2 points)

$$f''(t) - f(t) = e^{-3t}$$

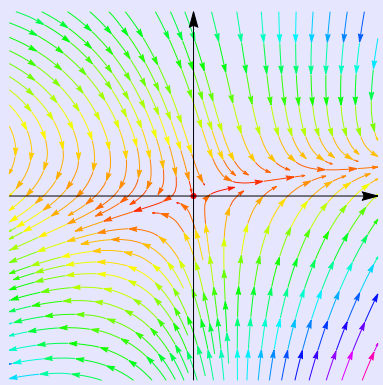
Problem 38P.10) (10 points):
We consider the nonlinear system of differential equations

$$\begin{aligned} \frac{d}{dt}x &= x + y - xy \\ \frac{d}{dt}y &= x - 3y + xy \ . \end{aligned}$$

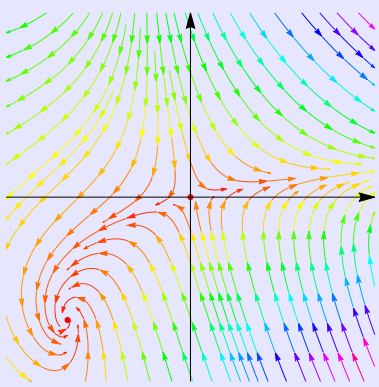
a) (2 points) Find the equilibrium points.
b) (3 points) Find the Jacobian matrix at each equilibrium point.
c) (3 points) Use the Jacobian matrix at an equilibrium to determine for each equilibrium point whether it is stable or not.
d) (2 points) Which of the diagrams A-D is the phase portrait of the system above?



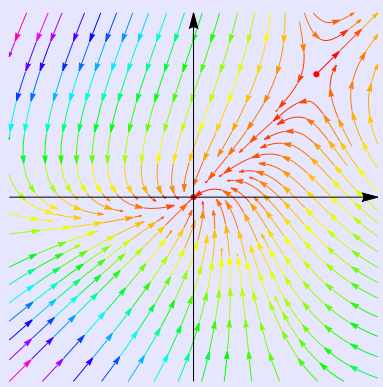
A



B



C



D

Problem 38P.11) (10 points):
a) (6 points) Find the **Fourier series** of the function which is 1 if $|x| > 1$ and -1 else. We call it the **Pacific rim** function.

$$f(x) = \begin{cases} 1 & |x| > 1 \\ -1 & |x| \leq 1 \end{cases} .$$

b) (4 points) Find the value of the sum of the squares of all the Fourier coefficients of f.

Problem 38P.12) (10 points):
a) Solve the system $f_t = 3f_{xx} - f + t$ with $f(0,x) = x$ on $[-\pi, \pi]$
b) Solve the system $f_t = 3f_{xx} - 9f_{xxxx}$ with $f(0,x) = x$.

Problem 38P.13) (10 points):
a) Solve the system $f_{tt} = 9f_{xx}$ with $f(0,x) = x$ and $f_t(0,x) = 11\sin(88x)$.
b) Solve $f_{tt} = 9f_{xx} - f_{xxxx}$ with $f_t(0,x) = x$.

1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	

Name:

LINEAR ALGEBRA AND VECTOR ANALYSIS

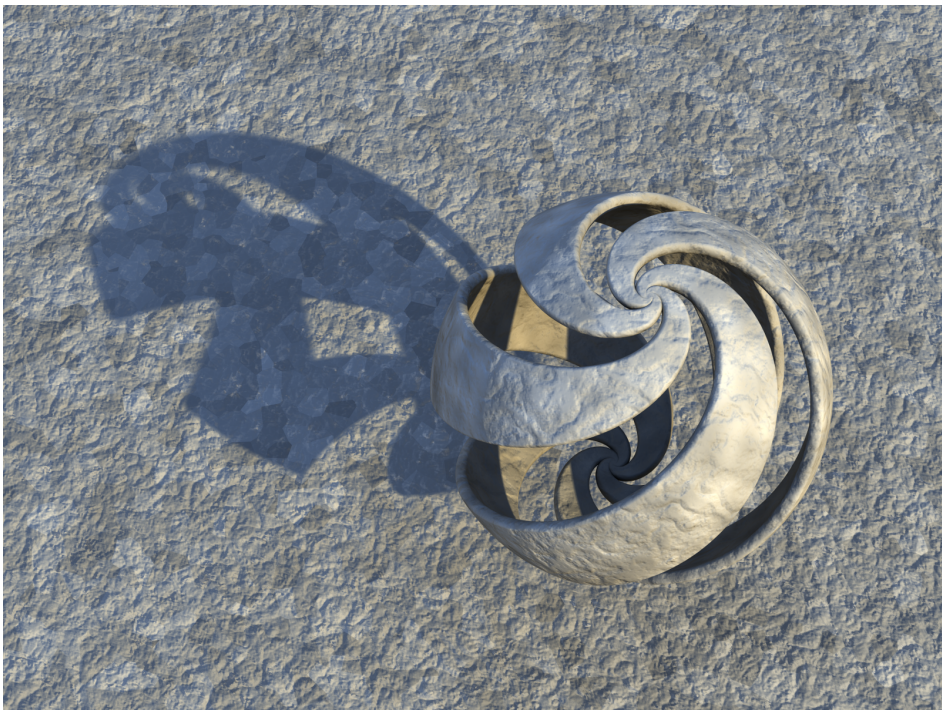
MATH 22B

Total:

Unit 38: Final Exam, 5/15/2019

Welcome to the final exam. Please don't get started yet. We start all together at 2:00 PM after reviewing some exam formalities. You can fill out the attendance slip already. Also, you can already enter your name into the larger box above.

- You only need this booklet and something to write. Please stow away any other material and any electronic devices. Remember the honor code.
- Please write neatly and give details. Except for problem 2 we want to see details, even if the answer should be obvious to you.
- Try to answer the question on the same page. There is additional space on the back of each page. If you must, use additional scratch paper at the end.
- If you finish a problem somewhere else, please indicate on the problem page where we can find it.
- You have 180 minutes for this 3-hourly.



Picture credit: rendered with Povray code of Tor Olav Kristensen.

Problems

Problem 38.1) (10 points):
a) (3 points) Write down the Laplacian $L \in M(2,2)$ of the complete graph with two vertices and find its two eigenvalues and eigenvectors.

b) (3 points) Write down a closed form solution of the **discrete Schrödinger equation** $if_t = Lf$ on this graph if the initial condition is $f(0) = [3, 5]$.

c) (2 points) Prove that the set $X = \{f \in C^\infty(\mathbb{R}) \mid f(x) = -f(-x-1)\}$ is a linear space of functions.

d) (2 points) Prove that the map $T(f) = -f(-1-x)$ is a linear transformation on $C^\infty(\mathbb{R})$.

Problem 38.2) (10 points). Each question is one point:

a) Marc Kac asked a famous question in 1966. We have demonstrated the topic in class. What was the question?
b) What general solution method can be used to solve the differential equation $dx/dt = x^3t^2$?
c) What is the value of the golden mean ϕ? If you should have forgotten, write down a matrix for which it is the eigenvalue and compute the eigenvalue.
d) Which condition has to be satisfied so that a continuous dynamical system $x'(t) = Ax(t)$ is stable?
e) Write down the 5×5 Jordan block to the eigenvalue $\lambda = 4$.
f) Fill in second column of the 2×2 matrix $A = \begin{bmatrix} (1+i)/2 & \text{.....} \\ (1-i)/2 & \text{.....} \end{bmatrix}$ so that the matrix is in the group $SU(2)$.
g) Who published the book Liber Abaci in 1202?
h) You have proven the identity $QP - PQ = i$ for some operators Q and P acting on smooth functions. What is the name of this identity?
i) What is the name of the mathematician who gave a first rigorous proof that the Fourier series of a differentiable function converges?
j) What does the Parseval identity say?

Problem 38.3) (10 points):
In order to solve the system of equations

$$\left|\begin{array}{ccccccccc} x & & & & & & + & v & = 10 \\ & & y & + & z & & & & = 6 \\ x & & & & & + & u & + & v & = 6 \end{array}\right|$$

we have to row reduce the following matrix:

$$B = \left|\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 1 & 10 \\ 0 & 1 & 1 & 0 & 0 & 6 \\ 1 & 0 & 0 & 1 & 1 & 6 \end{array}\right| .$$

Carefully do that one step at a time. Even if you know how to do it faster, we want you to use one row reduction step at a time. Then, we want you to write down the general solution (x, y, z, u, v) of the system.

Problem 38.4) (10 points):
a) (8 points) Find the best function

$$f(x, y) = ax^4 + by^5 = z$$

which fits the data points $(0, 1, 1), (1, 1, 2), (1, 0, 4)$ using the least square method.

b) (2 points) Somewhere in the computation you had to invert a 2×2 matrix. What is the Jordan normal form of that matrix $A^T A$?

Problem 38.5) (10 points):
a) (4 points) Find all the eigenvalues and eigenvectors of the matrix

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} .$$

If you should use some ingenuity to find the eigenvalues, give your reasoning. Only then, your genious will be appreciated! Without reasoning, we assume you were lucky.

b) (2 points) Find a matrix S such that $S^{-1}AS$ is a diagonal matrix B.

c) (1 point) What is the matrix B?

d) (1 point) Give the name of a theorem which assures without computation that A is diagonalizable.

e) (1 point) Is the system $x(t + 1) = Ax(t)$ stable or not?

f) (1 point) Is the system $x'(t) = Ax(t)$ stable or not?

Problem 38.6) (10 points):
The **Barycentric refinement process** for one-dimensional discrete geometries takes a graph with x vertices and y edges and produces a new graph with $x+y$ vertices and $2y$ edges. This defines a discrete dynamical system $T(x,y) = (x+y, 2y)$.

a) (3 points) Write down the 2×2 matrix A which implements this transformation T.

b) (3 points) We start with a graph having $x = 10$ vertices and $y = 11$ edges. How many vertices and edges are there after applying T two times?

c) (4 points) Find a closed form solution which gives $[x(t), y(t)]$ at time t with initial condition from b).‘

Problem 38.7) (10 points):
a) (2 points) Find the QR decomposition of the matrix

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} .$$

b) (2 points) Find the QR decomposition of the matrix

$$B = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} .$$

c) (1 point) Find the QR decomposition of the matrix AB, where A, B are given in a) and b).

d) (1 point) What are the eigenvalues of A?

e) (1 point) What are the eigenvalues of B?

f) (1 point) What is the determinant of AB?

g) (1 point) Why are the matrices AB and BA similar?

h) (1 point) What is the kernel of AB?

Problem 38.8) (10 points):

a) (2 points) What is the determinant of

$$A = \begin{bmatrix} 3 & 4 \\ 4 & 3 \end{bmatrix} ?$$

b) (2 points) A 22×22 matrix B unknown to you has determinant 22. What is the determinant of the matrix

$$(B^{-22})^T \ ?$$

c) (6 points) Find the determinant of the following 22×22 matrix: (make sure to document what you are doing! You are allowed to write into the matrix. There is more to the matrix than you might think.)

$$\begin{bmatrix}
0 & 2 & 0 & 0 & 0 & 0 & 0 & 8 & 0 & 0 & 8 & 0 & 0 & 8 & 8 & 8 & 8 & 0 & 8 & 8 & 8 & 8 \\
0 & 0 & 2 & 0 & 0 & 0 & 0 & 8 & 0 & 0 & 8 & 0 & 0 & 8 & 0 & 8 & 8 & 0 & 8 & 8 & 0 & 0 \\
0 & 0 & 0 & 2 & 0 & 0 & 0 & 8 & 0 & 0 & 8 & 0 & 0 & 8 & 0 & 8 & 0 & 8 & 0 & 8 & 8 & 8 \\
0 & 0 & 0 & 0 & 2 & 0 & 0 & 8 & 0 & 0 & 8 & 0 & 0 & 8 & 0 & 8 & 0 & 0 & 0 & 8 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 2 & 0 & 8 & 0 & 0 & 8 & 8 & 0 & 8 & 8 & 8 & 0 & 0 & 0 & 8 & 8 & 8 \\
2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 2 & 22 & 22 & 22 & 22 & 0 & 0 & 8 & 8 & 8 & 0 & 8 & 8 & 8 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 22 & 2 & 22 & 22 & 22 & 0 & 0 & 0 & 0 & 8 & 0 & 0 & 0 & 8 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 22 & 22 & 2 & 22 & 22 & 0 & 0 & 8 & 8 & 8 & 0 & 8 & 8 & 8 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 22 & 22 & 22 & 2 & 22 & 0 & 0 & 8 & 0 & 0 & 0 & 8 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 22 & 22 & 22 & 22 & 2 & 0 & 0 & 8 & 8 & 8 & 0 & 8 & 8 & 8 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 22 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 22 & 22 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 22 & 22 & 22 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 22 & 22 & 22 & 22 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 & 2 & 2 & 2 & 2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 & 2 & 2 & 2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 & 2 & 2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 & 2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2
\end{bmatrix}$$

.

Problem 38.9) (10 points):
Solve the following differential equations. You can of course use any method you know but you have to document what you do.

a) (2 points) $f'(t) + 3f(t) = e^{-2t}$.

b) (2 points) $f''(t) - 2f'(t) + f(t) = t$

c) (2 points) $f''(t) + 4f(t) = \cos(t)$

d) (2 points) $f''(t) + 4f(t) = \cos(2t)$

e) (2 points) $f'''(t) - f''(t) - f'(t) + f(t) = 1$

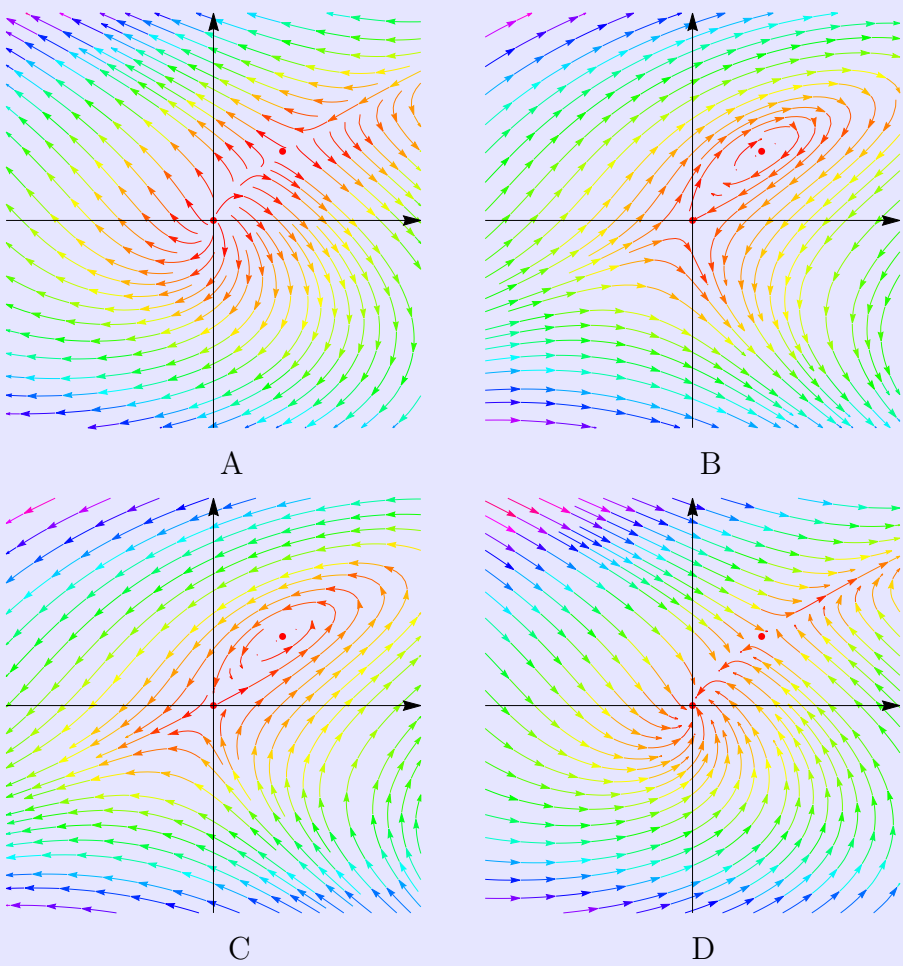
Problem 38.10) (10 points):
Analyze the solutions $(x(t), y(t))$ for the following nonlinear dynamical system

$$\begin{aligned} \frac{d}{dt}x &= x - y^2 \\ \frac{d}{dt}y &= y - x \end{aligned}$$

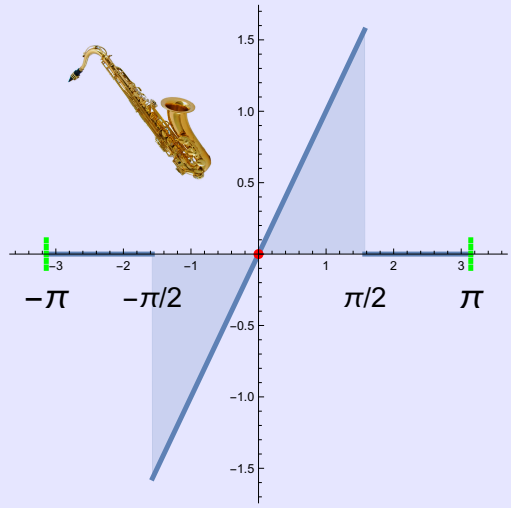
a) (3 points) Find the equations of the null-clines and find all the equilibrium points.

b) (4 points) Analyze the stability of all the equilibrium points.

c) (3 points) Which of the phase portraits A,B,C,D below belongs to the above system?



Problem 38.11) (10 points):
We try to model the sound of a **saxophone**. Its hull function is roughly given as a 2π periodic function which is on $[-\pi, \pi]$ defined as follows: the function is equal to x on $[-\pi/2, \pi/2]$ and $f(x)$ is 0 on $(\pi/2, \pi]$ and on $[-\pi, -\pi/2)$. You see the graph below. Your task is to find the Fourier series of this function.



Problem 38.12) (10 points):
In order to find out whether **grumpy cat** is dead or alive, you solve the **non-linear Schrödinger equation**

$$f_t = 3i f_{xxxx}$$

with initial condition $f(0, x) = \sin^2(x) + 7\sin(10x) + 9\cos(11x)$.

Problem 38.13) (10 points):
You are a scientist investigating **"monster waves"**. Also called **"rogue waves"**. The subject is also of interest to surfers. [This problem was written while watching again the classic "Point break" (the 1991 one, the 2015 remake has great stunts too):]
Find the solution of the nonlinear wave equation

$$f_{tt} = 2f_{xx} + 3f_{xxxxxx} + 12t^2$$

for which

$$f(0,x) = 9\sin(4x) + 8\cos(6x)$$

and

$$f_t(0,x) = 7\sin(13x) + \sum_{n=1}^{\infty} \frac{1}{n^4}\cos(nx) \ .$$

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Unit 39: Some literature

39.1. Here are some books in linear algebra.

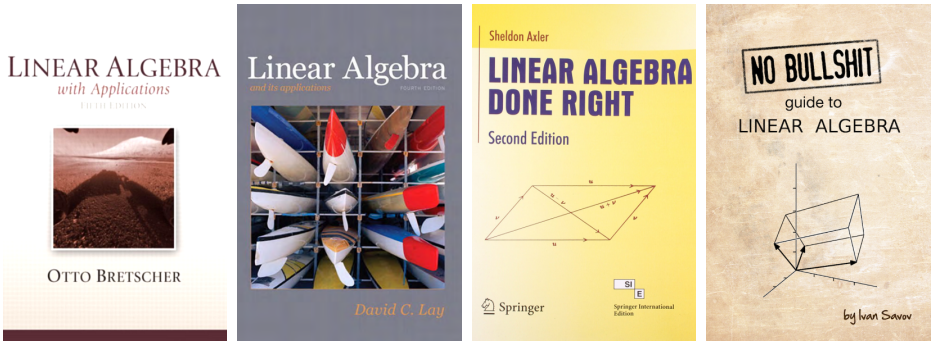


FIGURE 1. Bretscher, Lays, Axler, Savov.

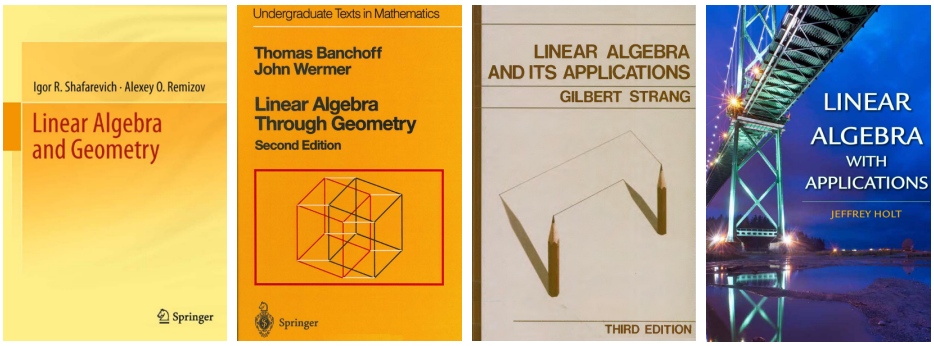


FIGURE 2. Shafarevich, Banchoff, Strang, Holt.

39.2. Some books in Fourier:

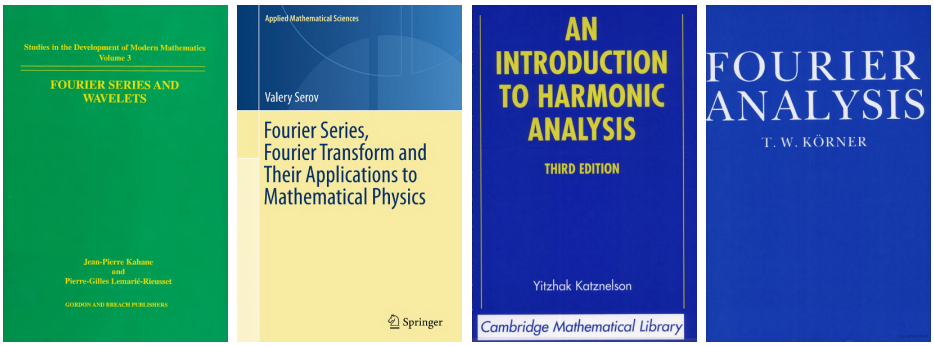


FIGURE 3. Kahane, Serov, Katznelson, Koerner.

39.3. Some books in PDE's:

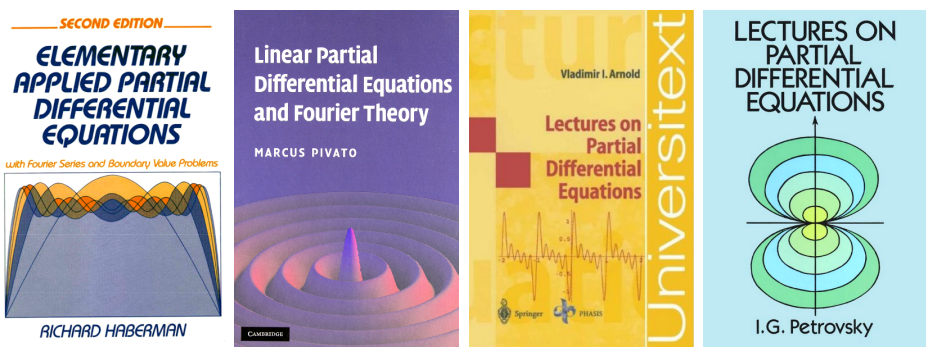


FIGURE 4. Haberman, Pivato, Arnold, Petrovsky.

OLIVER KNILL, KNILL@MATH.HARVARD.EDU, MATH 22B, HARVARD COLLEGE, SPRING 2019

LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22B

Acknowledgements:

43.1. Thanks to the Math 22b class Jordan Barkin, Vlad Batagui, Silvia Casacuberta, Ian Chan, Candice Chen, Michael Chen, Isabel Diersen, Shelby Elder, William Elder, Eric Hansen, Emily He, Jerry Huang, Spencer Hurt, Ruth Jaensubhakij, Drew Kelner, Madeleine Klebanoff, Mary Kolesar, Jordan Lawanson, Julian Lee, Jake Lim, Robert Malate, Samantha OSullivan, Christopher Ong, Moni Radev, Sanjana Ramrajvel, Isabella Rhyu, Isaac Robinson, Julian Schmitt, Daniel Shin, Ross Simmons, Sophia Sun, Varun Tekur, Grace Tian, Eddie Tu, Carissa Wu, Rebecca Xi, Mark Xu, Jenny Yao, Nicole Zhang, Jen Zhu, Maxwell Zhu, and Richard Zhu for listening to the lectures trying out the problems and proofreading.

43.2. Thanks to Robin Gottlieb and Cliff Taubes for initiating the course and getting it off the ground.

43.3. Thanks to Jameel Al-Aidroos, Wes Cain, Janet Chen and Dusty Grundmeyer for valuable discussions and feedback for early versions during spring and summer of 2018 while planning the course.

43.4. And to the TF's Yongquan Zhang, Elliot Parlin and Forrest Flesher for helping running the course smoothly.

43.5. The cover picture uses Povray code by Jaime Vives.

Oliver Knill, knill@math.harvard.edu, Math 22b, Harvard College, Spring 2019