Homework 12: Orthogonality

This homework is due on Monday, February 26, respectively on Tuesday, February 27, 2018. Advise for Tue-Thu sections: try to do it earlier due to the exam.

1 a) There exists exactly one vector $\vec{v} = [a,b,c,d]^T$ in $\vec{R}^4$ with integer entries $1 < a < b < c < d < 10$ so that the length is an integer. Hunt it down!

b) Find the angle between $\vec{v} = \begin{bmatrix} 3 \\ 1 \\ -1 \\ -1 \\ 0 \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} 0 \\ 1 \\ 3 \\ 1 \\ 1 \end{bmatrix}$.

c) What is the length of the vector $\begin{bmatrix} 1 \\ 2 \\ \vdots \\ 24 \end{bmatrix}$? **Remark:** By a theorem of Watson, this is the only vector of the form $[1,\dots,n]$ with $n > 1$ and non-negative integer length. The idea here is to simply add up the numbers. If you cite the formula for the sum of the first n squares, please give a reference. If you should compute the sum using a computer program, add the source code.

2 A vector $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ encodes data $(x_1,\dots,x_n)$. The **expectation** $\mathrm{E}[x]$ of x is defined as the average $m = (x_1 + \cdots + x_n)/n$. The vector $X = \begin{bmatrix} x_1 - m \\ \vdots \\ x_n - m \end{bmatrix}$ is called the **centered form** of x. Its expectation is zero. If X, Y are the centered versions of x, y then $(X \cdot Y)/n$ is called the covariance of X and Y and $\mathrm{Var}[X] = (X \cdot X)/n$ the variance of X and $\sigma(x) = |X|/\sqrt{n}$ the standard deviation. The correlation coefficient is $Cov[X]/(\sigma[X]\sigma[Y])$ which

simplifies to $X \cdot Y/(|X||Y|)$, the cosine of the angle between X and Y. Working with $\vec{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}$ and $\vec{y} = \begin{bmatrix} 4 \\ 3 \\ 5 \\ 7 \\ 11 \end{bmatrix}$.

a) Find the expectations of $\vec{x}, \vec{y}$.

b) Find the variance and standard deviation of $\vec{x}$ and $\vec{y}$.

c) Find the correlation coefficient of $\vec{x}, \vec{y}$.

3 If $\vec{x}, \vec{y}$ are two vectors, we get data points $(x_1, y_1), (x_2, y_2), ...(x_n, y_n)$ in the plane. The line $y = ax+b$ is called the **best linear fit**. We have $b = \mathrm{E}[y] - a\mathrm{E}[x]$, where $a = \mathrm{Cov}[X, Y]/Var[X]$. Draw the 5 data points from problem 2 and find the best fit $y = ax + b$.

4 An orthogonal basis in $\mathbb{R}^n$ for which every vector has either entries -1 or 1 is called a Walsh basis. and the corresponding matrix a Walsh matrix. Check that the columns of $W = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} /2$ form an orthonormal basis. Use W to build a 8×8 matrix encoding an orthonormal basis in R^8 by scaling $A = \begin{bmatrix} W & W \\ W & -W \end{bmatrix}$ in the right way. (Joseph Walsh graduated from Harvard in 1916 and also taught here from 1935-1966. Here is an open problem: nobody knows whether there is an orthogonal basis of vectors with entries $-1, 1$ in R^{668}. The corresponding matrices are called Hadamard matrices. The Walsh matrices above allow to construct examples for $n = 2^m$.)

5 Use an orthonormal basis of the plane $x + y + z = 0$ to find the matrix of the projection onto that plane. The formula for that projection matrix is given below: it is QQ^T, where Q contains the orthogonal basis in the columns and Q^T is the transpose,

containing the orthogonal basis in the rows.

Orthogonality

Two vectors are **orthogonal** if $\vec{v} \cdot \vec{w} = \begin{bmatrix} v_1, v_2, \dots, v_n \end{bmatrix} \begin{bmatrix} w_1 \\ \cdots \\ \cdots \\ w_n \end{bmatrix} = v_1 w_1 + \cdots + v_n w_n = 0.$ The **length** of a vector is is $|\vec{v}| = \sqrt{\vec{v} \cdot \vec{v}}$. The vector $\vec{v}/|\vec{v}|$ is called a **unit vector**. The Cauchy-Schwarz inequality $|\vec{v} \cdot \vec{w}| \leq |\vec{v}||\vec{w}|$ allows to define the angle α by $\cos(\alpha) = (\vec{v} \cdot \vec{w})/(|\vec{v}| \cdot |\vec{w}|)$. The number $\cos(\alpha)$ is called the **correlation coefficient**. If it is positive, the vectors are **positively correlated**, if it is negative they are **negatively correlated**. Orthogonal vectors are **uncorrelated**. A basis is an orthonormal basis, if all vectors are perpendicular and have length 1. If they are just orthogonal, they form an orthogonal basis. If we have an orthonormal basis of V, and Q be the matrix containing the basis vectors as column vectors. then the projection onto the space V is given by the matrix $P = QQ^T$, where Q^T is the transpose matrix.

Here is a run through for some of the statistics terms. We work with the vectors $x = \begin{bmatrix} 4 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ $y = \begin{bmatrix} 0 \\ 4 \\ 8 \\ 0 \end{bmatrix}$

Expectation	$\mathrm{E}[x]$	1
Expectation	$\mathrm{E}[y]$	3
Centered	$x - \mathrm{E}[x]$	$\begin{bmatrix} 3 \\ -1 \\ -1 \\ -1 \end{bmatrix}$
Centered	$y - \mathrm{E}[y]$	$\begin{bmatrix} -3 \\ 1 \\ 5 \\ -3 \end{bmatrix}$
Variance	$\mathrm{Var}[X] = X \cdot X/4$	3
Variance	$\mathrm{Var}[Y] = Y \cdot Y/4$	11
Standard deviation	$\sigma[x]$	$\sqrt{3}$
Standard deviation	$\sigma[y]$	$\sqrt{11}$
Covariance	$\mathrm{Cov}[x, y] = X \cdot Y/4$	-3
Correlation	$\mathrm{Cor}[x, y] = X \cdot Y/(\lvert X\rvert\lvert Y\rvert)$	$-3/\sqrt{33}$

We just did geometry in four dimensions computed the cosine of the angle between two vectors. This has an interpretation as a correlation. If the angle between X and Y is obtuse, the data are negatively correlated. If the angle is acute, then the date are positively correlated. If the angle is a 90 degree angle, the data are uncorrelated. In our example, the data were negatively correlated. Yes, the original vectors x, y were orthogonal but we first center the random variables before doing geometry. You can check that the correlation is invariant if we scale or translate x, y together.