Lecture Notes Math 371: Algebra (Fall 2006)

by Nathanael Leedom Ackerman

September 19, 2006

0.1 TALK SLOWLY AND WRITE NEATLY!!

0.2 Introduce Bilinear Forms

0.2.1 Definitions

Now that we have dealt with the structure theorem for finitely generated abelian groups it is time to go back to vector spaces an study another element of them, bilinear forms Bilinear forms are meant to be a generalization of the dot product on $\mathbb{R}^n$. So before we continue recall

Dot Product Definition

Definition 0.2.1.1. Let $X, Y \in \mathbb{R}^n$ then we define $(X \cdot Y) = X^t Y = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$.

The important features of the dot product are:

Bilinearity

$$((X_1 + X_2) \cdot Y) = (X_1 \cdot Y) + (X_2 \cdot Y)$$

$$(X \cdot (Y_1 + Y_2)) = (X \cdot Y_1) + (X \cdot Y_2)$$

$$(cX \cdot Y) = c(X \cdot Y) = (X \cdot cY)$$

Symmetry $(X \cdot Y) = (Y \cdot X)$

Positivity $X \neq 0 \Rightarrow (X \cdot X) > 0$

Notice that bilinearity says that if we fix one element of the dot product then $(- \cdot Y) : \mathbb{R}^n \to \mathbb{R}$ is a linear transformation. And it is this property which we will focus on first.

Bilinear Form Definition

Definition 0.2.1.2. Let V be a vector space over F. We define a bilinear form to be a function $f : V \times V \to F$ such that

$$(\forall v_1, v_2, w \in V) f(v_1 + v_2, w) = f(v_1, w) + f(v_2, w)$$

$$(\forall v, w_1, w_2 \in V) f(v, w_1 + w_2) = f(v, w_1) + f(v, w_2)$$

$$(\forall v, w \in V, c \in F) f(cv, w) = cf(v, w) = f(v, cw)$$

We will often use the notation $\langle v, w \rangle$ for $f(v, w)$.

Symmetric Bilinear Form Definition

Definition 0.2.1.3. We say a bilinear form $\langle , \rangle$ is Symmetric if

$$(\forall v, w)\langle v, w\rangle = \langle w, v\rangle$$

Skew-Symmetric Bilinear Form Definition

Definition 0.2.1.4. We say a bilinear form $\langle , \rangle$ is Skew-Symmetric if

$$(\forall v)\langle v, v\rangle = 0$$

0.2.2 Theorems

Explanation of Skew-Symmetric Definition

Lemma 0.2.2.1. *Let V be a vector space over a field F of characteristic $\neq 2$. Let $\langle , \rangle$ be a bilinear form on*

V. Then $\langle , \rangle$ is skew-symmetric if and only if

$$(\forall v, w \in V)\langle v, w\rangle = -\langle w.v\rangle$$

Proof. $\underline{\Rightarrow:}$ Well we then know that

$$0\langle v+w.v+w\rangle = \langle v, v\rangle+\langle w, w\rangle+\langle v, w\rangle+\langle w, v\rangle = \langle v.w\rangle+\langle w, v\rangle$$

So we have $0 = \langle v, w\rangle + \langle w, v\rangle$ and hence $\langle v, w\rangle = -\langle w, v\rangle$

$\underline{\Leftarrow:}$ Well we then know that $\langle v.v\rangle = -\langle v.v\rangle$ and so $2\langle v, v\rangle = 0$. So either $2 = 0$ or $\langle v, v\rangle = 0$. But we are assuming that the characteristic isn't 2 and so we must have $\langle v, v\rangle = 0$.

It is for this last part that we need the characteristic isn't 2. In the case of characteristic 2 we see that the condition that $\langle v, w\rangle = -\langle w, v\rangle$ is the same as saying that $\langle , \rangle$ is symmetric because $a = -a$ for all a. But, as we will see there are in general skew symmetric matrixes over fields of characteristic 2 which are not symmetric.

Hence the above definition is the right one. □

0.3 Matrix Representation of Bilinear Forms

0.3.1 Theorems

The most common examples of bilinear forms are those which act on the space F^n of column vectors as follows. Let A be an $n \times n$ matrix. Then.

$$\langle X, Y \rangle = X^t A Y$$

notice that this is a 1×1 matrix **(work out why on the bo**

The first thing we need to check is that this is in fact a bilinear form.

Matrix Form is Bilinear

Lemma 0.3.1.1. *Let* V *an* n*-dimensional vector space over* F *and let* $X, Y \in V$ *be represented as column vectors relative to some basis. Further let* A *be an* $n \times n$ *matrix in* F*. Then* $\langle X, Y \rangle = X^t AY$ *is a bilinear form.*

Proof. We need to check the following.

- $$< (X_1+X_2), Y >= (X_1+X_2)^t AY = (X_1^t+X_2^t)AY = (X_1^t AY)+(X$$

- $$\langle X, (Y_1+Y_2) \rangle = X^t A(Y_1+Y_2) = (X^t AY_1)+(X^t AY_2) = \langle X, Y_1 \rangle + \langle$$

- $$\langle cX, Y \rangle = (cX^t AY) = (X^t AcY) = c\langle X, Y \rangle = \langle X, cY \rangle$$

And so in fact $\langle , \rangle$ is a linear transform. □

Now, given a finite dimension vector space we want to show that any given bilinear form is of the above form.

Bilinear forms have matrixes

Lemma 0.3.1.2. *Let $\langle , \rangle$ be a bilinear form on V, a finite dimensional vector space and lets let $B = \{b_1, \cdots, b_n\}$ be a basis for V. Then there is a matrix A such that $\langle X, Y\rangle = X^t AY$ where X, Y are considered column vectors relative to the basis B.*

Proof. We want to show that there is a matrix A such that

$$(\forall\{x_i, y_i : i \in n\} \subseteq F)\langle x_1, \cdots, x_n\rangle^t A\langle y_1, \cdots, y_n\rangle = \langle \Sigma_{i\in n} x_i b_i, \Sigma_{i\in n} y_i$$

Well we know that any matrix we come up with will correspond to a bilinear form. So in particular if we can come up with a matrix which agrees with our bilinear form on the basis elements then the bilinear form associated to the matrix must be the one we want.

Specifically what we need is $A = (a_{i,j})$ where $a_{i,j} = \langle v_i, v_j \rangle$. Then by bilinearity $X^t AY = \langle X, Y \rangle$ for all vectors X, Y. □

Definition 0.3.1.3. We say that $A = (\langle b_i, b_j \rangle)$ is the Matrix Associated to the Bilinear form $\langle , \rangle$ relative to the basis $\{b_1, \cdots, b_n\}$.

0.4 Change of Base

One of the most important questions regarding these matrixes is what happens when we change bases. This leads us to the following theorem.

0.4.1 Theorem

Change of Base

Theorem 0.4.1.1. *Let A be the matrix associated to a bilinear form $\langle , \rangle$ with respect to a basis. Then the* *matrixes which represent the same form with respect to different basis are those of the form*

$$QAQ^t$$

for some $Q \in GL_n(F)$.

Proof. Let P be the element of $GL_n(F)$ which represents the linear transformation which changes the base. So we have $X^* = PX, Y^* = PY$ and $\langle X, Y\rangle = \langle X^*, Y^*\rangle$ (as X, X^* and Y, Y^* are just different representation of the same vectors)

We then know that

$$\langle X, Y\rangle = X^t AY = (P^{-1}X^*)^t A(P^{-1}Y^*) = (X^*)^t (P^{-1})^t AP^{-1}Y^*$$

But we also know that

$$\langle X, Y\rangle = \langle X^*, Y^*\rangle = (X^*)^t A^* Y^*$$

for A^* the matrix representing the bilinear for relative to the new basis. Hence letting $Q = (P^{-1})^t$ we must have $A^* = QAQ^t$. $\square$

0.5 Dot Product

Dot product matrix

Now lets consider what happens to the dot product if we change basis. Recall that

$$(X \cdot Y) = X^t Y$$

And so we have that the matrix associated to the standard dot product is just the identity matrix.

Orthogonal

Definition 0.5.0.2. Recall that a matrix is said to be orthogonal if $P^t P = I$ or $P^{-1} = P^t$.

Lemma 0.5.0.3. *If you change base relative to an orthogonal change of base then the dot product is preserved.*

Proof. Let P be the orthogonal change of base. Well I is the matrix associated with the identity so by previous theorems this means that the matrix associated with the dot product under the new basis is

$$(P^{-1})^t I (P^{-1}) = (P^t)^t I P^t = PP^t = I$$

So changing the basis by an orthogonal matrix preserves the dot product. □

Similarly we have

Lemma 0.5.0.4. *The matrixes which represent the dot product are those of the form PP^t for $P \in GL_n(\mathbb{R})$.*

Proof. By previous theorem. □

Now while this is nice, it doesn't tell us a whole lot as we don't know what matrixes of the form PP^t should look like. So we will have to use other properties of the dot product to help us out.

Recall the three conditions on the dot product which were important. First off was Bilinearity. But this isn't a helpful as we know that X^tAY is bilinear for every A.

The next is symmetry. This is in fact useful.

Definition Symmetric Matrix

Definition 0.5.0.5. We say a matrix is symmetric if $A = A^t$.

Lemma 0.5.0.6. *A bilinear form is symmetric if and only if the matrix associated to it is symmetric.*

Proof. Symmetry is equivalent to

$$\langle X, Y \rangle = X^tAY = Y^tAX = \langle Y, X \rangle$$

but we have $(Y^tAX) = (Y^tAX)^t = X^tA^tY$ because the transpose of a 1×1 matrix is itself.

Hence being a symmetric bilinear for is equivalent to

$$(\forall X, Y) X^t A Y = X^t A^t Y$$

and this is equivalent to $A = A^t$. □

The third condition is that $(X \cdot X) > 0$ if $X \neq 0$ (Positivity).

Positive Definite

Definition 0.5.0.7. We call a bilinear form $\langle , \rangle$ on V Positive Definite if

$$(\forall v \in V, v \neq 0) \langle v, v \rangle > 0$$

0.6 Orthogonal

0.6.1 Definitions

Orthonormal basis

Definition 0.6.1.1. Given a bilinear form $\langle , \rangle$ on a vector space V we say that two vectors $v, w \in V$ are orthogonal ($v \perp w$) if

$$\langle v, w\rangle = 0$$

Let $B = \langle v_1, \cdots, v_n\rangle$ be a basis for V. We then say that B is an orthonormal basis if

$$(\forall i \neq j)\langle v_i, v_j\rangle = 0$$

$$(\forall i)\langle v_i, v_i\rangle = 1$$

0.6.2 Theorems

Lemma 0.6.2.1. *If B is an orthonormal basis for V with respect to $\langle,\rangle$ then the matrix associated to $\langle,\rangle$ relative to B is the identity.*

Proof. Immediate. □

Now we are going to show that for any positive definite bilinear form an orthonormal basis exists.

Orthonormal basis always exist for symmetric

Theorem 0.6.2.2. *Let $\langle , \rangle$ be a positive definite symmetric bilinear form on a finite dimensional vector space V. Then there is an orthonormal basis for V*

Proof. The method we are going to use is called the Gram-Schmidt procedure

We are going to start with a basis $B = (v_1, \cdots, v_n)$

Step 1:

The first step will be to normalize v_1. Now we know that

$$\langle cv_1, cv_1 \rangle = c^2 \langle v_1, v_1 \rangle$$

But, because we know that that $\langle , \rangle$ is positive definite,

$$\langle v_1, v_1 \rangle > 0$$

and so we know $\sqrt{\langle v_1, v_1 \rangle}$ is a real number and hence if we let

$$w_1 = \sqrt{\langle v_1, v_1 \rangle} v_1$$

then we see that $\langle w_1, w_1 \rangle = 1$.

<u>Step 2a:</u>

Now we want to look for a linear combination of v_2, w_1 which is orthogonal to w_1. The value is

$$w = v_2 - \langle v_2, w_1 \rangle w_1$$

because

$$\langle w, w_1 \rangle = \langle v_2, w_1 \rangle - \langle v_2, w_1 \rangle \langle w_1, w_1 \rangle = 0$$

<u>Step 2b:</u>

Then normalize w and call that vector w_2

Further $(w_1, w_2, v_3, \ldots, v_n)$ is a basis for V. $\cdots$

<u>Step k a:</u>

Suppose we have defined orthonormal vectors $(w_1, \ldots, w_{k-1})$ and that $(w_1, \ldots, w_{k-1}, v_k, \ldots, v_n)$ is a basis.

Then we want to look for a linear combination of $v_k, w_1, \ldots, w_{k-1}$

which is orthogonal to w_i for all $i < k$. The value is

$$w = v_k - \langle v_k, w_1 \rangle w_1 + \cdots + \langle v_k, w_{k-1} \rangle w_{k-1}$$

because

$$\langle w, w_i \rangle = \langle v_k, w_i \rangle - \langle v_k, w_1 \rangle \langle w_1, w_i \rangle + \cdots + \langle v_k, w_{k-1} \rangle \langle w_{k-1}, w_i \rangle$$

But $\langle w_j, w_i \rangle = 0$ if $i \neq j$ and $\langle w_j, w_j \rangle = 1$ so we have

$$\langle w, w_i \rangle = \langle v_k, w_i \rangle - \langle v_k, w_i \rangle \langle w_i, w_i \rangle = 0$$

Step kb:

Then normalize w and call that vector w_k

Further $v_k \in Span(w_1, \ldots, w_k, v_{k+1}, \ldots, v_n)$ and so $(w_1, \ldots, w_k, v_{k+1},$

is a basis.

Hence after iterating this process n times we see that $(w_1, \ldots, w_n)$ is an orthonormal basis for V. $\square$

We then have the following theorem.

Theorem 0.6.2.3. *The following are equivalent for a real $n \times n$ matrix*

(1) A represents the dot product.

(2) There is an invertible matrix $P \in GL_n(\mathbb{R})$ such that $A = P^t P$

(3) A is symmetric and positive definite.

Proof. We have already shown that (1) and (2) are equivalent.

Further, the fact that (1) $\to$ (3) is by virtue of the fact that the dot product satisfies positivity and symmetry.

So all that is left is to show that (3) $\to$ (2).
Well the first thing to notice is that if A is positive definite then so is the form $\langle X, Y \rangle = X^t AY$. So in particular there is an orthonormal basis B' with respect to $\langle , \rangle$.

Now we know then that with respect to the basis B' the matrix associated to $\langle , \rangle$ is I (because B' is orthonormal). But at the same time we know that if P is the matrix associated to the change of base from B' to the standard basis of $\mathbb{R}^n$ then

$$A = P^t A' P = P^t P$$

and so A satisfies (2). □

0.7 Geometry Associated To A Positive Form

0.7.1 Definitions

Suppose we have a bilinear form $\langle , \rangle$ on a real vector space. Then, it is possible to define the length of a vector as follows.

Euclidian space definition

Definition 0.7.1.1. Let $v \in V$ and let $\langle , \rangle$ be a positive definite bilinear form. Then we can define

$$|v| = \surd(\langle v, v\rangle)$$

We often call a real vector space with a length a Euclidian space.

0.7.2 Theorems

It isn't hard to see the following in equalities

In equalities

Lemma 0.7.2.1. *Let $\langle , \rangle$ be a positive definite bilinear form on a real vector space V. Then we have*

chwarz Inequality) $|\langle v, w\rangle| \leq |v| \cdot |w|$

riangle Inequality) $|v + w| \leq |v| + |w|$

Proof. Immediate. (MAYBE MAKE A HOMEWORK) □

Now given a subspace of W we have that

Restriction

Lemma 0.7.2.2. *Let V be a vector space and $\langle . \rangle$ a bilinear form on V. Then if $W \subseteq V$ is a subspace then $\langle , \rangle$ restricts to a bilinear form on $\langle , \rangle$.*

Further if $\langle , \rangle$ is positive definite or symmetric on V then $\langle , \rangle$ is positive definite or symmetric on W

Proof. Immediate. □

Now this will then allow us to define a relative angle between two vectors. Suppose we have V is a vectors space with a positive definite bilinear form $\langle , \rangle$ and let $v, w \in V$.

Then we can restrict $\langle , \rangle$ to $Span(v, w)$. Now if v, w are linearly dependent we say that the angle is 0.

Otherwise we can define the angel between the two vectors is determined by

$$\langle v, w \rangle = |v| \cdot |w| \cdot cos(\theta)$$

FIGURE OUT WHY

0.8 TODO

- Go through Lang's book on the same topics.