26 The Projection Formula

26.1 Review: Symmetric and Hermitian Forms

Last time, we were talking about different kinds of pairings or bilinear forms on vector spaces. In particular, we will be studying two cases in parallel: vector spaces V over $\mathbb{R}$ with symmetric forms on them, and vector spaces over $\mathbb{C}$, with Hermitian forms on them. A Hermitian form is almost symmetric, with a complex conjugate thrown in.

Then, we discussed the idea of vectors being *orthogonal* to each other with respect to the form if the pairing is zero. A form is non-degenerate if and only if the space of vectors orthogonal to the entire vector space V is $\{0\}$, so there are no nonzero vectors orthogonal to all other vectors. Such a vector lies in the kernel of the matrix of the form, so a matrix with nonzero determinant will correspond to a non-degenerate form.

26.2 Orthogonality

Recall this theorem about the restriction of a bilinear form to a subspace. We'll prove it now.

Theorem 26.1
Let $W \subseteq V$. If $\langle \cdot, \cdot \rangle|_W$ is non-degenerate on W, then $V = W \oplus W^\perp$, which means that every vector $v \in V$ is equal to $\vec{w} + \vec{u}$ uniquely, where $w \in W, u \in W^\perp$.

It is possible for the restriction of a non-degenerate form to be degenerate; for example the form $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is non-degenerate but is just given by $A' = 0$ when $W = \operatorname{Span}(\vec{e}_1)$, which is clearly degenerate.

Proof. If $\langle \cdot, \cdot \rangle|_W$ is non-degenerate, then $W \cap W^\perp = \{0\}$. We have $W \oplus W^\perp \subset V$, so it suffices to show that $V \subset W \oplus W^\perp$. Pick a basis of W, $\{w_1, \ldots, w_k\}$, and define a linear transformation

$$\begin{aligned} \varphi : V &\longrightarrow \mathbb{C}^k \\ \vec{v} &\longmapsto (\langle w_1, v, \rangle, \ldots, \langle w_k, v \rangle). \end{aligned}$$

This is a linear transformation just by the properties of a Hermitian form. The kernel is

$$\ker(\varphi) = W^\perp,$$

since $W = \operatorname{Span}\{\vec{w_i}\}$. Also, $\dim \operatorname{im} \varphi \le k = \dim W$, so by the dimension formula,

$$\dim V = \dim \ker \varphi + \dim \operatorname{im} \varphi \le \dim W^\perp + \dim W.$$

Consider the mapping

$$\begin{aligned} W \oplus W^\perp &\longrightarrow V \\ (w, u) &\longmapsto w + u. \end{aligned}$$

It has kernel $\{0\}$, since $W \cap W^\perp = \{0\}$, so

$$\dim W + \dim W^\perp \le \dim V,$$

and thus $\dim W + \dim W^\perp = \dim V$ and therefore $V = W \oplus W^\perp$. □

To emphasize, the geometric version of this with respect to the dot product feels obvious and works in most cases. For general forms, we have to have this condition that our form is non-degenerate on the subspace.

The splitting $V = W \oplus W^\perp$ is helpful, in particular, for inductive arguments, because it is possible to reduce some property of V to being true on W and $W^\perp$.

26.3 Orthogonal Bases

By applying a change of basis, it is always possible to put an arbitrary matrix into Jordan normal form, and if there are distinct eigenvalues, it is in fact possible to diagonalize it. What about the matrix of a bilinear form?

Guiding Question

Given a vector space V and a bilinear form $\langle \cdot, \cdot \rangle$, how simple can we get the form to be?

First, it is always possible to find a basis orthogonal with respect to the bilinear form.

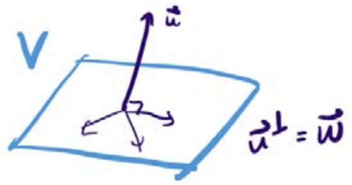
Theorem 26.2

For a symmetric or Hermitian form $\langle \cdot, \cdot \rangle$, the vector space V has an orthogonal basis $\{v_1, \cdots, v_n\}$, which is when $\langle v_i, v_j \rangle = 0$ for $i \neq j$. The matrix for the pairing in the basis will then be diagonal, since it is given by the inner product from the form.

Proof. To prove this, induct on $\dim V = n$.

- **Case 1.** There is some u such that $\langle u, u \rangle \neq 0$. Then, the one-dimensional subspace $W = \mathrm{Span}(u)$, $\langle \cdot, \cdot \rangle|_W$ is non-degenerate.

  By induction, $W^\perp$ has an orthogonal basis $\{v_2, \cdots, v_n\}$, so $\{u, v_2, \cdots, v_n\}$ is an orthogonal basis for V.



- **Case 2.** Otherwise, for every $v \in V$, $\langle v, v \rangle = 0$. This is a very strong constraint on the form, and in fact it forces $\langle v, w \rangle = 0$ for all v, w, which forces *any* basis to be an orthogonal basis. To see this, consider the inner product on a sum of two vectors with itself:

$$\begin{aligned} 0 &= \langle v + w, v + w \rangle \\ &= \langle v, v \rangle + \langle w, w \rangle + \langle v, w \rangle + \langle w, v \rangle \\ &= 2\langle v, w \rangle. \end{aligned}$$

  When $F = \mathbb{R}$, we have $\langle v, w \rangle = 0$, by the symmetry of the form. Otherwise, for $F = \mathbb{C}$, $\mathrm{Re}(\langle v, w \rangle) = 0$, and the same process can also be done for v and iw to show that $\langle v, w \rangle = 0$. Then the inner product is 0 on any two vectors so every basis is orthogonal.

□

We can simplify the basis even further.

Corollary 26.3

In fact, V has an orthogonal basis $\{v_1, \cdots, v_k\}$ where $\langle v_i, v_i \rangle = 1, -1$, or 0.

Proof. Take an orthogonal basis $\{x_1, \cdots, x_k\}$. Consider $\langle x_i, x_i \rangle$, which is a real number.

- If the pairing is 0, then let $v_i = x_i$.
- Otherwise, we can normalize and take $v_i = \frac{1}{\sqrt{|\langle x_i, x_i \rangle|}} x_i$; then $\langle v_i, v_i \rangle = \frac{\langle x_i, x_i \rangle}{|\langle x_i, x_i \rangle|}$, so it will be 1 or -1 depending on the sign of $\langle x_i, x_i \rangle$.

□

In particular, if $\langle \cdot, \cdot \rangle$ is non-degenerate, only ± 1 occur. Also, if $\langle \cdot, \cdot \rangle$ is positive definite, by definition, $\langle v, v \rangle > 0$ if $v \neq 0$, so only $+1$s occur, so in that basis, the form looks just like the dot product or the standard Hermitian product.

The following claim will be shown in the upcoming problem set.

Claim 26.4 (Sylvester's Law)**.** *In fact, given V and $\langle \cdot, \cdot \rangle$, the number of 1s, the number of -1s, and the number of 0s that occur in the diagonal form are determined by V and $\langle \cdot, \cdot \rangle$, and not by the choice of orthogonal basis.*[92] *This is called* Sylvester's Law*, and the number of 1s, -1s, and 0s is called the* signature *of the form.*

For example, in the form used in special relativity, $\begin{pmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}$, the signature is $(3, 1, 0)$.

In matrix form, the corollary states that for a symmetric matrix $A \in \mathrm{Mat}_{n \times n}(\mathbb{R})$, for which $A^T = A$, there exists some matrix $P \in GL_n(\mathbb{R})$ such that $P^T A P$ is a diagonal matrix with all $1s$, $-1s$, or 0s on the diagonal:

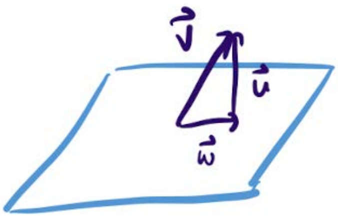
$$P^T A P = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & -1 & & & \\ & & & \ddots & & \\ & & & & 0 & \\ & & & & & \ddots \end{pmatrix}.$$

If A is positive definite, which is when $x^T A x > 0$, there exists P such that $P^T A P = I_n$ implies that $A = Q^T Q$, where $Q = P^{-1}$.

The statement is similar for complex matrices, where we replace the transposes with adjoints.

26.4 Projection Formula

Consider a vector space V and a form $\langle \cdot, \cdot \rangle$, as well as a subspace W for which $\langle \cdot, \cdot \rangle|_W$ is non-degenerate. By Theorem 26.1, $V = W \oplus W^\perp$ such that $v = w + u$.



Guiding Question
How can we compute w and u?

To do so, we use the orthogonal projection. We want a map

$$\begin{aligned} \pi : V &\longrightarrow W \\ v &\longmapsto w, \end{aligned}$$

so that $v = \pi(v) \perp W$.[93]

[92] They are similar to eigenvalues in that while there are many choices of orthogonal basis, the *number* of 1s, -1s, and 0s are not dependent on the particular basis.

[93] This is an extremely useful application of linear algebra! In geometric situations, the vector w is the vector closest to v of the vector in the plane, and perhaps these vectors are in a vector space of data points. Finding a formula for w explicitly is called least-squares regression.

Assuming there exists an orthogonal basis $\{w_1, \cdots, w_k\}$ for W, the formula for π is simple.[94] The vector can be written as

$$v = cw_1 + \cdots + cw_k + u,$$

where $u \perp W$. Then for all i,

$$\langle w_i, v\rangle = 0 + \cdots + 0 + c_i\langle w_i, w_i\rangle + 0 + \cdots + 0,$$

so

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

It is not possible for $\langle w_i, w_i\rangle = 0$, because the form would be degenerate. In fact, this formula is useful when $W = V$, because it provides a formula for the coordinates of some vector with respect to the orthogonal basis.

Example 26.5

Let $V = \mathbb{R}^3$ and $\langle\cdot,\cdot\rangle$ be the dot product. Then let W be the span of $\vec{w}_1 = (1,1,1)^T$ and $\vec{w}_1 = (1,1,-2)^T$. The pairings are $\langle w_1, w_1\rangle = 3$, $\langle w_2, w_2\rangle = 6$, $\langle w_1, v\rangle = 6$, and $\langle w_2, v\rangle = -3$. The projection of $(1,2,3)$ is then

$$\pi(v) = \frac{6}{3}w_1 - \frac{1}{2}w_2 = \begin{pmatrix} 3/2 \\ 3/2 \\ 3 \end{pmatrix}.$$

To verify, $v - \pi(v) = (-1/2, 1/2, 0)$, which is orthogonal both to w_1 and w_2.

[94]Once we've developed the machinery for bilinear forms, these ideas become a lot simpler!