Lecture 20: Sylvester's law and inner products

Last lecture, we showed that any symmetric bilinear form can be diagonalized as long as $\mathrm{char}(\mathbb{F}) \neq 2$. In the case that the characteristic is 2, we cannot necessarily find an orthogonal basis: consider

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

In a field with characteristic 2, $1 = -1$, so this is a symmetric matrix and thus defines a symmetric bilinear form f on $\mathbb{F}^2$, where $\mathbb{F} = \mathbb{Z}_2$, by $f(v,w) = w^t A v$. However for any $v = ae_1 + be_2$,

$$f(v,v) = a^2 f(e_1,e_1) + abf(e_1,e_2) + abf(e_2,e_1) + b^2 f(e_2,e_2) = 0 .$$

So if $\{v,w\}$ is an orthogonal basis, we must have $f(v,v) = f(w,w) = f(v,w) = f(w,v) = 0$, implying f is 0, a contradiction.

In the case of $\mathbb{C}^n$, the diagonalization result allows us actually to get a matrix with only 1's and 0's. To do this, take an orthogonal basis B for a symmetric form f and label it $\{v_1,\ldots,v_n\}$. Now define

$$w_i = \begin{cases} \frac{v_i}{\sqrt{f(v_i,v_i)}} & \text{if } f(v_i,v_i) \neq 0 \\ 0 & \text{otherwise} \end{cases} .$$

Then $\{w_1,\ldots,w_n\}$ is still orthogonal relative to f but satisfies $f(w_i,w_i) = 0$ or 1. This actually works in any field $\mathbb{F}$ in which each element has a square root. This of course does not work in $\mathbb{R}^n$, but we have a separate result for that.

Theorem 0.1 (Sylvester's law of inertia). *Let $f \in Sym(\mathbb{R}^n, \mathbb{R})$. Then there exists a basis B such that $[f]_B^B$ is diagonal with only 1's, -1's and 0's. Furthermore, if B' is another basis such that $[f]_{B'}^{B'}$ is in this form, the number of 1's, -1's and 0's respectively is the same.*

Proof. For the first part, just take a basis $B = \{v_1,\ldots,v_n\}$ from the last theorem and define $w_i = v_i/\sqrt{|f(v_i,v_i)|}$ when $f(v_i,v_i) \neq 0$. For the second part, define $S_0(B)$ as the set of vectors v in B such that $f(v,v) = 0$, $S_+(B)$ as the set of $v \in B$ such that $f(v,v) > 0$ and $S_-(B)$ as the set of $v \in B$ such that $f(v,v) < 0$. Last, define their spans as $V_0(B), V_+(B)$ and $V_-(B)$. Since each vector of B falls into one of these categories,

$$V = V_0(B) \oplus V_+(B) \oplus V_-(B) .$$

We have a similar decomposition for B'.

Note that if $v \in V_+(B)$ we can write it as $v = \sum_{i=1}^t a_i v_i$, where $v_1,\ldots,v_t$ are the elements of $S_+(B)$. Then using orthogonality,

$$f(v,v) = \sum_{i=1}^t a_i^2 f(v_i,v_i) > 0 .$$

The same is clearly true for $V_+(B')$. Now assume for a contradiction that $\#S_+(B) > \#S_+(B')$. Then by the two subspace dimension theorem, writing $V_\leq(B') = V_0(B') \oplus V_-(B')$,

$$\dim\ (V_+(B) \cap V_\leq(B')) + \dim\ (V_+(B) + V_\leq(B')) = \dim\ V_+(B) + \dim\ V_\leq(B')\ .$$

Using $\dim\ (V_+(B) + V_\leq(B')) \leq \dim\ V$ and $\dim\ V_+(B) > \dim\ V_+(B')$,

$$\dim\ (V_+(B) \cap V_\leq(B')) \geq \dim\ V_+(B') + \dim\ V_\leq(B') - \dim\ V > 0\ .$$

Therefore there is a vector $v \in V_+(B) \cap V_\leq(B')$. But such a vector must have $f(v,v) > 0$ and $f(v,v) \leq 0$, a contradiction. So $\#S_+(B) \leq \#S_+(B')$. Reversing the roles of B and B' gives

$$\#S_+(B) = \#S_+(B')\ .$$

An almost identical argument gives $\#S_-(B) = \#S_-(B')$. This means we must have also $\#S_0(B) = \#S_0(B')$, since both bases have the same number of elements. □

Some remarks are in order here.

1. The space $V_0(B)$ is unique; that is, for another basis B' giving f the matrix form above, we have $V_0(B) = V_0(B')$. This is because they are both equal to $N(f)$.

   Proof. We will show $V_0(B) = N(f)$. The same proof shows $V_0(B') = N(f)$. Let $v \in S_0(B)$. Since B is orthogonal, if w is another element of B (not equal to v) then $f(v,w) = 0$. However we also have $f(v,v) = 0$ since $v \in S_0(B)$. This means $L_f(v)$ kills all basis elements and must be 0, giving $v \in N(f)$. Therefore

   $$S_0(B) \subset N(f) \Rightarrow V_0(B) = \mathrm{Span}(S_0(B)) \subset \mathrm{Span}(N(f)) = N(f)\ .$$

   However,

   $$\#S_0(B) = \dim\ V - [\#S_+(B) + \#S_-(B)] = \dim\ V - \mathrm{rank}\ f = \dim(N(f))\ .$$

   Thus $V_0(B) = N(f)$. □

2. The spaces $V_-(B)$ and $V_+(B)$ are not unique; they just have to have the same dimensions as $V_-(B')$ and $V_+(B')$ respectively, if B' is another basis that puts f into the form of the theorem. As an example, take $f \in Bil(\mathbb{R}^2, \mathbb{R})$ with matrix in the standard basis

   $$[f]_B^B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\ .$$

   Then take $v_1 = (2, \sqrt{3})$ and $v_2 = (\sqrt{3}, 2)$. Since $f((a,b),(c,d)) = ac - bd$,

   $$f(v_1, v_1) = (2)(2) - (\sqrt{3})(\sqrt{3}) = 1\ ,$$
   $$f(v_1, v_2) = (2)(\sqrt{3}) - (\sqrt{3})(2) = 0\ ,$$
   $$f(v_2, v_2) = (\sqrt{3})(\sqrt{3}) - (2)(2) = -1\ .$$

   This means if $B' = \{v_1, v_2\}$ then $[f]_{B'}^{B'} = [f]_B^B$ but the spaces $V_+(B)$ and $V_+(B')$ are not the same (nor are $V_-(B)$ and $V_-(B')$).

Sequilinear and Hermitian forms

One motivation for considering symmetric bilinear forms is to try to abstract the standard dot product. In this respect, we know that in $\mathbb{R}^n$, the quantity

$$\|\vec{a}\| = \sqrt{\vec{a} \cdot \vec{a}} = \sqrt{a_1^2 + \cdots + a_n^2}$$

measures the "length" of the vector $\vec{a} = (a_1, \ldots, a_n)$. If we try to give this same definition for complex vectors, we get

$$(i, \ldots, i) \cdot (i, \ldots, i) = -n \ ,$$

which is bad because we should have $\vec{a} \cdot \vec{a} \geq 0$ for all vectors. This is the motivation for introducing a different dot product on $\mathbb{C}^n$: it is given by

$$\vec{a} \cdot \vec{b} = a_1 \overline{b}_1 + \cdots + a_n \overline{b}_n \ ,$$

where $\overline{z}$ represents the complex conjugate $x - iy$ of a complex number $z = x + iy$. Note that this dot product is no longer bilinear because of the conjugate; however, it is *sesquilinear*.

Definition 0.2. *If V is a vector space over $\mathbb{C}$, a function $f : V \times V \to \mathbb{C}$ is called sesquilinear if*

1. *for each fixed $w \in V$, the function $v \mapsto f(v, w)$ is linear and*
2. *for each fixed $v \in V$, the function $w \mapsto f(v, w)$ is anti-linear; that is, for $w_1, w_2 \in V$ and $c \in \mathbb{C}$,*

$$f(v, cw_1 + w_2) = \overline{c} f(v, w_1) + f(v, w_2) \ .$$

If, in addition, $f(v, w) = \overline{f(w, v)}$ for all $v, w \in V$, we call f Hermitian.

The theory of Sesquilinear and Hermitian forms parallels that of bilinear and symmetric forms. We will not give the proofs of the following statements, as they are quite similar to before:

- If f is a sesquilinear form and B is a basis of V then there is a matrix of f relative to B as before:

  $$\text{for } v, w \in V, \ f(v, w) = \overline{[w]}_B^t [f]_B^B [v]_B \ .$$

  Here, the (i, j)-th entry of $[f]_B^B$ is $f(v_1, v_j)$, where $B = \{v_1, \ldots, v_n\}$, as before.

- The function $L_f(v)$ is no longer linear, it is anti-linear. Although $R_f(w)$ is linear, the map $w \mapsto R_f(w)$ (from V to V^*) is no longer an isomorphism. It is a bijective anti-linear function.

- We have the polarization formula

  $$4f(u, v) = f(u + v, u + v) - f(u - v, u - v) + if(u + iv, u + iv) - if(u - iv, u - iv) \ .$$

  This implies that if $f(v, v) = 0$ for all v, then $f = 0$.

There is a corresponding version of Sylvester's law:

Theorem 0.3 (Sylvester for Hermitian forms)**.** *Let f be a Hermitian form on a finite-dimensional complex vector space V. There is a basis B of V such that $[f]_B^B$ is diagonal with only 0's, 1's and -1's. Furthermore, the number of each does not depend on B as long as the matrix is in diagonal form.*

The proof is the same.