NOTES ON BILINEAR FORMS

PARAMESWARAN SANKARAN

These notes are intended as a supplement to the talk given by the author at the IMSc Outreach Programme Enriching Collegiate Education-2015.

Symmetric bilinear product

A `symmetric bilinear product` on a (finite dimensional) real vector space V is a mapping $\langle .,. \rangle : V \times V \to \mathbb{R}$, $(u,v) \mapsto \langle u, v\rangle$ which is $\mathbb{R}$-linear in each variable u, v and is `symmetric`, that is, $\langle u, v\rangle = \langle v, u\rangle$. We say that $\langle .,. \rangle$ is `non-degenerate` if $\langle u, v\rangle = 0$ for all $v \in V$ implies that $u = 0$. We say that $\langle .,. \rangle$ is `positive definite` if $\langle u, u\rangle > 0$ if $u \neq 0$. A positive definite symmetric bilinear product on V is also known as an `inner product`. An inner product is evidently non-degenerate.

Fix an ordered basis $\mathcal{B} = \{v_1, \ldots, v_n\}$ of V. Let $A \in M_n(\mathbb{R})$ be a symmetric matrix. Writing any $x \in V$ as $x = \sum_{1\leq j\leq n} x_j v_j$, we may identify x with the column vector $(x_1, \ldots, x_n)^t$. We obtain a symmetric bilinear product on V defined as $\langle x, y\rangle_A := x^t A y$. Note that $\langle v_i, v_j\rangle_A = a_{ij}$ where $A = (a_{ij})$.

Conversely if $\langle .,. \rangle : V \times V \to \mathbb{R}$ is any symmetric bilinear product on V, then $\langle .,. \rangle = \langle .,. \rangle_A$ where $A = (\langle v_i, v_j\rangle)$. The matrix A is called the `matrix of` $\langle .,. \rangle$ with respect to $\mathcal{B}$.

If A is non-singular, then $\langle .,. \rangle_A$ is non-degenerate. Indeed for a non-zero element x in V, there exists z in V such that $x^t z \neq 0$—in fact, we can choose z to be a standard column vector. Since A is non-singular, there exists y such that $z = Ay$. Then $\langle x, y\rangle = x^t A y = x^t z \neq 0$. Conversely, if $\langle .,. \rangle$ is non-degenerate, then its matrix with respect to any basis of V is non-singular.

If $v'_1, \ldots, v'_n$ is another ordered basis $\mathcal{B}'$ for V and if $P = (p_{ij})$ is the change of basis matrix from $\mathcal{B}$ to $\mathcal{B}'$ so that $v'_j = \sum p_{ij} v_i$, then $\sum x_i v_i = x = \sum x'_j v'_j = \sum x'_j p_{ij} v_i$ and so $(x_i) = P(x'_j)$. Let A and A' be the matrix of the same symmetric bilinear product $\langle .,. \rangle$ with respect to $\mathcal{B}$ and $\mathcal{B}'$ respectively, so that $x^t A y = \langle x, y\rangle = (x')^t A' y'$. Thus $(x')^t A' y' = x^t A y = (x')^t P^t A P y'$ for all column vectors $x', y' \in \mathbb{R}^n$. It follows that $A' = P^t A P$; equivalently $A = (P^{-1})^t A' P^{-1}$.

Let $\langle .,. \rangle$ be a fixed symmetric bilinear product on V. For $u, v \in V$ we say that u is `perpendicular` (or `orthogonal`) to v (written $u \perp v$) if $\langle u, v\rangle = 0$. For W a subset of V

we denote by $W^\perp$ the subset $\{v \in V \mid \ v \perp w \ \forall w \in W\}$ of V. Evidently $W^\perp$ is a vector subspace of V. When W is a subspace, we have $\dim W^\perp \geq \dim V - \dim W$. The `null space` of $\langle .,. \rangle$ is the space $V^\perp$ of all vectors v that are orthogonal to the whole of V. The bilinear product $\langle .,. \rangle$ defines a non-degenerate symmetric bilinear product on V/N where N is the null space of $\langle .,. \rangle$.

If $W \subset V$ is a vector subspace then the restriction of $\langle .,. \rangle$ to $W \times W$ is a symmetric bilinear product on W. This bilinear product on W is non-singular if and only if $W \cap W^\perp = 0$. In turn $W \cap W^\perp = 0$ if and only if $V = W \oplus W^\perp$ (internal direct sum).

Let $V = \mathbb{R}^2$ (regarded as column vectors) and let $\langle x, y \rangle = x_1 y_2 + x_2 y_1$. Then $\langle e_1, e_1 \rangle = 0 = \langle e_2, e_2 \rangle$. Nevertheless, the bilinear product is non-degenerate. Indeed, the matrix of the bilinear product with respect to the basis e_1, e_2 is $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. Since it is invertible, it follows that $\langle .,. \rangle$ is non-degenerate. Note that $\langle e_1 - e_2, e_1 - e_2 \rangle < 0$.

A basis $\mathcal{B} = u_1, \ldots, u_n$ of V is called `orthogonal` if $u_i \perp u_j$ for $i \neq j$. Note that the matrix of $\langle .,. \rangle$ with respect to an orthogonal basis is *diagonal*. There always exists an orthogonal basis $\mathcal{B}$ for any symmetric bilinear product. If it is non-degenerate, one may replace each $v \in \mathcal{B}$ by $v/\sqrt{|\langle v, v \rangle|}$ to obtain, possibly after a rearrangement of the basis elements, an ordered basis with respect to which the matrix of the bilinear product has the form $\left(\begin{smallmatrix} I_r & 0 \\ 0 & -I_s \end{smallmatrix}\right)$, where I_r denotes the identity matrix of order r. The number r is an invariant of the bilinear product. It equals the dimension of V if and only if the bilinear product is positive definite.

If $W \subset V$ is a subspace such that the bilinear product is non-degenerate on W (so that $V = W \oplus W^\perp$ as observed above), then we have the `orthogonal projection` $p : V \to W$ defined by $p|_{W^\perp} = 0$, $p|_W =$ identity. If $\{w_1, \ldots, w_k\}$ is an orthogonal basis for W, then, for any $v \in V$, $p(v) = \sum_{1 \leq j \leq k} \frac{\langle v, w_j \rangle}{\langle w_j, w_j \rangle} w_j$. Indeed $v \mapsto \sum_{1 \leq j \leq k} \frac{\langle v, w_j \rangle}{\langle w_j, w_j \rangle} w_j$ defines a linear map of V to W that vanishes on $W^\perp$ and is identity on W.

Hermitian product

Let V be a (finite dimensional) complex vector space. A `Hermitian product` $\langle .,. \rangle : V \times V \to \mathbb{C}$ is a `sesquilinear map`—conjugate linear in the first argument and complex linear in the second—such that $\langle u, v \rangle = \overline{\langle v, u \rangle}$. It is called `non-degenerate` if $\langle u, v \rangle = 0$ for all $v \in V$ implies $u = 0$. It is called `positive definite` if $\langle u, u \rangle > 0$ for $u \neq 0$ (note that $\langle u, u \rangle$ is real since $\langle u, u \rangle = \overline{\langle u, u \rangle}$). A positive definite Hermitian product is also called an `inner product`. A Hermitian inner product is evidently non-degenerate.

As in the case of symmetric bilinear product on real vector spaces, one has the notion of the `matrix` of a Hermitian product (with respect to an ordered $\mathbb{C}$-basis for V). The matrix A of a Hermitian product is Hermitian, that is, $A^* = A$ where $A^* := \bar{A}^t$. If A, A'

are two matrices of the same Hermitian product with respect to two ordered bases $\mathcal{B}, \mathcal{B}'$ and if P is the change of basis matrix from $\mathcal{B}$ to $\mathcal{B}'$, then $A' = P^*AP$.

Let $\langle ., . \rangle$ be a Hermitian inner product on a (finite dimensional) complex vector space V. Suppose that $T : V \to V$ is a $\mathbb{C}$-linear transformation. We obtain a linear transformation $T^* : V \to V$, where T^*v for v in V is defined by $\langle T^*v, w\rangle = \langle v, Tw\rangle$ for all $v, w \in V$. Note that $(ST)^* = T^*S^*$ for any two linear transformations S, T of V and that $T \mapsto T^*$ is an `involution`, that is, $(T^*)^* = T$ for all T.

We say that T is `normal` if $TT^* = T^*T$, equivalently $\langle Tv, Tw\rangle = \langle T^*v, T^*w\rangle$ for all $v, w \in V$. We say that T is `Hermitian` if $T = T^*$, equivalently, $\langle Tv, w\rangle = \langle v, Tw\rangle$. We say that T is `unitary` if $TT^* = T^*T = I$, the identity transformation, equivalently $\langle Tv, Tw\rangle = \langle v, w\rangle$ for all $v, w \in V$.

These notions carry over to elements of $M_n(\mathbb{C})$ where A^* is defined as $\bar{A}^t$. Thus A is `Hermitian` if $A^* = A$, `unitary` if $AA^* = A^*A = I$, and `normal` if $AA^* = A^*A$. Note that a real $n \times n$ matrix is Hermitian if and only if it is symmetric and is unitary if and only if it is `orthogonal` (that is $AA^t = A^tA = I$).

The two versions—in terms of linear transformation and matrices—are related by considering the matrix of a transformation with respect to an ordered basis $\mathcal{B}$ of V with respect to which the matrix of the Hermitian inner product is the identity.

The following are some basic properties of Hermitian matrices.

• *The eigenvalues of a Hermitian (or a real symmetric) transformation are all real.*
Proof. Suppose that A is Hermitian and v is an eigenvector corresponding to an eigenvalue λ of A. Then $\bar{\lambda}\langle v, v\rangle = \langle \lambda v, v\rangle = \langle Av, v\rangle = \langle v, Av\rangle = \lambda\langle v, v\rangle$. But $\langle v, v\rangle \neq 0$ since $v \neq 0$, by positive definiteness of $\langle ., .\rangle$. Hence $\bar{\lambda} = \lambda$.

• *The eigenvalues of a unitary matrix are of absolute value* 1.
Proof. Let $Av = \lambda v$ with $v \neq 0$ and A being unitary. Then $\langle v, v\rangle = \langle Av, Av\rangle = \langle \lambda v, \lambda v\rangle = \bar{\lambda}\lambda\langle v, v\rangle$. Cancelling $\langle v, v\rangle$ we obtain that $||\lambda||^2 = \bar{\lambda}\lambda = 1$.

Theorem (Spectral theorem for normal matrices) *Let T be a normal matrix in $M_n(\mathbb{C})$. Then there exists a $n \times n$ unitary matrix P such that P^*TP is diagonal.* □

Reference

Chapter 8 of **M. Artin,** *Algebra*, 2nd ed., Pearson, New Delhi (2011).

Problems

(1) Find an orthogonal basis of $\mathbb{R}^2$ for the symmetric bilinear given by the matrix (a) $\left(\begin{smallmatrix} 2 & 1 \\ 1 & 3 \end{smallmatrix}\right)$, (b) $\left(\begin{smallmatrix} 1 & 2 \\ 2 & 0 \end{smallmatrix}\right)$.

(2) Find the orthogonal projection of the vector $(2,3,4)^t \in \mathbb{R}^3$ onto the xy-plane where the symmetric bilinear product is given by the matrix $\begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 3 \end{pmatrix}$.

(3) Prove that the maximum of entries of a positive definite matrix A is attained on the diagonal.

(4) Suppose that A is a complex $n \times n$ matrix such that x^*Ax is real for all $x \in \mathbb{C}^n$. Is A Hermitian?

(5) Let $\langle .,. \rangle$ be a non-zero symmetric bilinear product on a real vector space or a Hermitian product on a complex vector space V. Show that there exists a vector v such that $\langle v, v \rangle \neq 0$. Show that $V = U \oplus U^\perp$ where U is the vector subspace of V spanned by v.

(6) Let $\langle .,. \rangle$ be a positive definite Hermitian product on a complex vector space V. Define bilinear maps $(.,.), [.,.] : V \times V \to \mathbb{R}$ (where V is regarded as a real vector space) as the real and imaginary parts of $\langle .,. \rangle$ so that $\langle u, v \rangle = (u, v) + \sqrt{-1}[u, v]$. Show that $(.,.)$ is a positive definite symmetric bilinear product on V and that $[.,]$ is skew symmetric, i.e., $[u, v] = -[v, u]$.

(7) On the vector space $M_n(\mathbb{R})$ define $\langle A, B \rangle$ as $tr(A^t.B)$. Show that this is a positive definite symmetric bilinear product.

(8) Let $W_1, W_2 \subset V$ and let $\langle .,. \rangle$ be a symmetric bilinear product. Show that (a) $(W_1 + W_2)^\perp = W_1^\perp \cap W_2^\perp$, (b) $W \subset W^{\perp\perp}$. When does equality hold in (b)?

(9) Suppose that $\langle .,. \rangle$ is non-degenerate symmetric bilinear product on V. Show that $\dim W \leq (1/2) \dim V$ if $W \subset W^\perp$.

(10) If A, B are symmetric $n \times n$ real matrices which commute, show that there exists a matrix P such that both P^tAP and P^tBP are diagonal. Find two 2×2 symmetric matrices A and B such that AB is not symmetric. (Comment: The product of two commuting symmetric matrices is symmetric.)

Hints/Solutions to Problems

(1) (a) Take, for example, $e_1 = (0,1)^t$ to be the first basis vector (this wouldn't have been a good choice if $\langle e_1, e_1\rangle$ were 0 but that is not the case). Let the second be $ae_1 + be_2$. We want $\langle ae_1 + be_2, e_1\rangle = 2a + b = 0$. Thus, we can take the second vector to be $(1,-2)^t$. (b) Proceeding as in (a), we get $e_1 = (0,1)^t$, $(-2,1)^t$.

(2) We could directly apply the formula in the notes provided we have an orthogonal basis for the xy-plane. To find such a basis, we could take $w_1 = e_1 = (1,0,0)^t$ to be the first basis vector. Let $w_2 = ae_1 + be_2$ be the second, where $e_2 = (0,1,0)^t$. We want $\langle w_2, w_1\rangle = \langle ae_1 + be_2, w_1\rangle = 0$, so we get $a + b = 0$. Thus we can take $w_2 = e_1 - e_2$. Letting $v = (1,2,3)^t$, we have

$$p(v) = \frac{\langle v, w_1\rangle}{\langle w_1, w_1\rangle} w_1 + \frac{\langle v, w_2\rangle}{\langle w_2, w_2\rangle} w_2 = \frac{3}{1} w_1 + \frac{-5}{1} w_2 = 3e_1 - 5(e_1 - e_2) = (-2,5,0)^t$$

(3) In fact, for a symmetric positive definite real matrix $A = (a_{ij})$, the maximum cannot be attained at a non-diagonal element. To see this, just observe that $\langle e_i - e_j, e_i - e_j\rangle = a_{ii} + a_{jj} - 2a_{ij}$, for $i \neq j$. Thus if the maximum were attained at a_{ij}, we would get a contradiction: $\langle e_i - e_j, e_i - e_j\rangle \leq 0$.

(4) Put $A = A_1 + iA_2$. The imaginary part of $(x_1 - ix_2)^t(A_1 + iA_2)(x_1 + ix_2)$ is

$$-x_2^t A_1 x_1 + x_1^t A_1 x_2 + x_1^t A_2 x_1 + x_2^t A_2 x_2$$

Putting $x_2 = 0$, we get $x_1^t A_2 x_1 = 0$ for all x_1, which means A_2 is skew-symmetric. But now we also have $x_1^t A_1 x_2 - x_2^t A_1 x_1 = 0$. But $x_1^t A_1 x_2 - x_2^t A_1 x_1 = x_1^t A_1 x_2 - x_1^t A_1^t x_2 = x_1^t(A_1 - A_1^t)x_2$. So $x_1^t(A_1 - A_1^t)x_2 = 0$ for all x_1, x_2. So $A_1 - A_1^t = 0$; in other words, A_1 is symmetric. Thus A is Hermitian.

(5) Since $\langle ., .\rangle$ is non-zero, there exists $u_1, u_2 \in V$ such that $\langle u_1, u_2\rangle =: \lambda \neq 0$. In case V is a complex vector space, replacing u_2 by $\bar{\lambda} u_2$ if necessary, we may (and do) assume, $\langle u_1, u_2\rangle$ is real. Thus $\langle u_2, u_1\rangle = \langle u_1, u_2\rangle$. If $\langle u_i, u_i\rangle \neq 0$ for some $i \leq 2$, take $v = u_i$. If $\langle u_i, u_i\rangle = 0$ for $i = 1, 2$, set $v = u_1 + u_2$. Then $\langle v, v\rangle = 2\langle u_1, u_2\rangle \neq 0$.

Since U is the one-dimensional vector space spanned by v, if $U \cap U^\perp \neq 0$, then $v \perp v$ and so $\langle v, v\rangle = 0$, contrary to our choice of v. So we must have $U \cap U^\perp = 0$. This implies that $V = U \oplus U^\perp$.

(6) Fix notation as in the solution above of (4).That A_2 is skew-symmetric and A_1 is symmetric follows from (4). The real part of $(x_1 - ix_2)^t(A_1 + iA_2)(x_1 + ix_2)$ is

$$x_1^t A_1 x_1 - x_1^t A_2 x_2 + x_2^t A_2 x_1 + x_2^t A_1 x_2$$

Putting $x_2 = 0$, we get $x_1^t A_2 x_1 \geq 0$ for all x_1 with equality only if $x_1 = 0$, which means that A_1 is positive definite.

(7) Writing $A = (a_1, \ldots, a_n)$ where a_i is the i^{th} column of A, we see that $\text{trace}(A^t A) = a_1^t a_1 + \cdots + a_n^t a_n \geq 0$ with equality holding only if each $a_i = 0$ (in other words,

only if $A = 0$). This proves that the form $\langle A, B\rangle = \text{trace}(B^t A)$ is positive definite. It is evidently symmetric.

(8) (a) and (b) follow readily from the definitions. To see when equality holds in (b), first observe the following: for any subspace W of V, we have (1) $\dim W^\perp = \dim V - \dim W + \dim(W \cap V^\perp)$ and (2) $W^\perp \supseteq V^\perp$. Now, using these, we get $\dim W^{\perp\perp} = \dim W + \dim V^\perp - \dim(W \cap V^\perp)$. Thus for $W^{\perp\perp}$ to equal W, it is necessary and sufficient that $V^\perp = W \cap V^\perp$, or equivalently $V^\perp \subseteq W$. Two cases where this condition holds are: $V^\perp = 0$ (the form is non-degenerate); $W = V$.

(9) This follows from the equality $\dim W^\perp = \dim V - \dim W$ when the form is non-degenerate (see the solution to the previous item).

(10) In fact, we can find an orthogonal matrix P with the desired property. By using the spectral theorem for real symmetric matrices once, we may assume that A is diagonal. Then B is block diagonal symmetric: the block sizes are the multiplicities of the entries of A. Now we apply the spectral theorem to each block of B. Since each corresponding block of A is a scalar matrix, A will not be disturbed when we diagonalize B.

For the second part: If $A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, then $AB = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$.