Symmetric Bilinear and Quadratic Forms

A *symmetric bilinear form* on a vector space V over a field F is a map $A : V \times V \to F$ satisfying:

- (Symmetry) $A(v, w) = A(w, v)$ for all v and w in V.
- (Left-Additivity) $A(v + w, u) = A(v, u) + A(w, u)$ for all v, w and u in V. Note that, by symmetry this implies right-additivity.
- (Left-Linearity) $A(av, w) = aA(v, w)$ for all v, w in V and a in F. Note that, by symmetry this implies right-linearity.

For a matrix A we define the matrix A^t (as above) to be the matrix whose (i, j)-th entry is the (j, i)-th entry of A. So if A has m rows and n columns, then A^t has n rows and m columns.

Exercise: Check that if A is an $m \times n$ matrix and B is a $n \times k$ matrix, then $(AB)^t = B^t A^t$.

If we consider V as the space of column vectors and define the matrix A with (i, j)-th entry $A(e_i, e_j)$ (where e_i denotes the column vector with all entries 0 except the i-th row containing a 1), then:

Exercise: Check the identity $A(v, w) = v^t Aw$ where we think of the right-hand side, which is a 1×1 matrix, as an element of F. (Here the superscript t denotes the transpose of the matrix.)

We note that $A(e_i, e_j) = A(e_j, e_i)$ by symmetry. Hence, the matrix A is *symmetric*; its (i, j)-th entry is the same as its (j, i)-th entry, or written differently $A^t = A$.

Conversely, given a symmetric matrix A, we can define $A(v, w) = v^t Aw$ and see that this is a symmetric bilinear form.

A *quadratic form* on a vector space V over a field F (of characteristic different from 2; i.e. $1/2 \in F$) is a map $Q : V \to F$ such that:

- $Q(av) = a^2 Q(v)$ for all a in F and v in V.
- $B(v, w) = (Q(v + w) - Q(v) - Q(w))/2$ is a (symmetric) bilinear form on V. Note that the symmetry property is automatic from the commutativity of addition in V.

Exercise: Given a symmetric bilinear form A, we can define a quadratic form $Q(v) = A(v, v)$. Calculate the bilinear form B associated with Q.

More generally, we can give such a definition for *any* matrix A. In that case $Q(v) = v^t Av$ is a 1×1 matrix and so is its own transpose. i.e. we see that $(v^t Av)^t = v^t A^t v$. So we have $v^t A^t v = v^t Aw$.

Exercise: Given an *arbitrary* matrix A consider the quadratic form $Q(v) = v^t Av$. Calculate the symmetric bilinear form B associated with Q.

Diagonalisation of a quadratic form

Given a non-zero quadratic form Q and a vector v such that $Q(v) \neq 0$. Since v is perforce non-zero, we can make a change of basis where $v = e_1$. In this basis, the quadratic form looks like

$$Q(v_1, \ldots, v_n) = a_1 v_1^2 + \sum_{i=2}^{n} b_i v_1 v_i + \sum_{i,j=2}^{n} a_{i,j} v_i v_j$$

where $a_1 = Q(1, 0, \ldots, 0) \neq 0$.

By the process of "completion of the square", we can now introduce the new variable $w_1 = v_1 + \sum_{i=2}^{n} \frac{b_i}{2a} v_i$. In terms of the variables $(w_1, v_2, \ldots, v_n)$ the quadratic form takes the form

$$Q(w_1, v_2, \ldots, v_n) = a_1 w_1^2 + \sum_{i,j=2}^{n} b_{i,j} v_i v_j a w_1^2 + Q'(v_2, \ldots, v_n)$$

Proceeding inductively with the quadratic form $Q'(v_2, \ldots, v_n)$ in fewer variables, we can this find a suitable *linear* change of variables so that the quadratic form looks line

$$Q(w_1, \ldots, w_n) = a_1 w_1^2 + \cdots + a_n w_n^2$$

in terms of these new variables.

Suppose that the linear change of variables is given in the form

$$\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = S \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

for an $n \times n$ matrix S. Furthermore, if the quadratic form is given by $Q(v) = v^t A v$, then

$$A = S^t \operatorname{diag}(a_1, \ldots, a_n) S$$

where the middle term is the diagonal matrix. In other words, this operation of diagonalisation can also be written as $T^t A T$ is a diagonal matrix where $T = S^{-1}$ is an invertible matrix.

It is worth noting that this is uses $T^t A T$ and *not* $T^{-1} A T$. Secondly, this diagonalisation is possible over *any* field of characteristic different different from 2, without assuming that the characteristic polynomial of A has roots over the field!

While the above procedure looks complex, it becomes clear if one applies it to a specific example.

Rank, nullity and Signature

Given a bilinear symmetric form $A(v, w) = v^t A w$, we have seen that a change of basis of the vector space replaces A by $T^t A T$ for an invertible matrix T. It is clear that the *rank* of A is invariant under such a replacement.

Further, we have seen that, for a suitable choice of T, the matrix $T^t A T$ is a diagonal matrix. We note right away that the rank of A is the same as the number of non-zero diagonal entries.

If we are working over the field of real numbers $\mathbb{R}$, we can make a further simplification as follows. Given a quadratic form as a diagonal form $Q(v_1, \ldots, v_n) = \sum a_i v_i^2$. We note that, for each i, the non-negative real number $|a_i|$ has a non-negative square root b_i; let s_i be 0, 1 or -1 according as a_i is 0, positive or negative respectively. Using the variables $w_i = b_i v_i$, we see that the quadratic form Q becomes $\sum s_i w_i^2$ in terms of these new variables. In other words, for a *suitable* change of basis matrix T over real numbers, the diagonal matrix $T^t A T$ has entries which are 0, 1 or -1. Note that there may be many such matrices T. What can we say about the number of 0's, 1's and -1's? Do these depend on the choice of T?

As seen above, the number of 0 entries is $n - r$ where r is the rank of the matrix A; the number $n - r$ is sometimes called the *nullity* of the bilinear form. Thus, this number is independent of the choice of T.

We will se later that the number of 1's and -1's are also independent of the choice of T. If there are n 0's, p 1's and q -1's, we can say that the form is of type (n, p, q). The term *non-degenerate* is also used for qudratic forms where $n = 0$. The term *signature* is used for (p, q) (or sometimes, when the rank $(= p + q)$ is known, the same term is used signature for $p - q$).

A form is said to be *positive semi-definite* if $q = 0$, and *positive definite* if $n = q = 0$. A standard example of a positive definite form is given by the identity matrix; in this case, the symmetric bilinear form is just the familiar "dot-product" of vectors. At the other extreme, a form for which $n = p = 0$ is called *negative definite*.

Gram-Schmidt Process

Given a basis $v_1, \ldots, v_n$ of $\mathbb{Q}^n$ (or F^n for a subfield F of the field $\mathbb{R}$ of real numbers). We wish to make a find a way to "convert" this to an orthogonal basis $f_1, \ldots, f_n$.

We put $f_1 = v_1$. Next, we form vectors of the form

$$w_k = v_k + a_{1,k} v_1 \text{ for all } k \geq 2$$

so that $(w_k)^t v_1 = 0$ for $k \geq 2$. To find w_k we write the above conditions as

$$(v_k)^t v_1 + a_{1,k} (v_1)^t v_1 = 0 \text{ for all k} \geq 2$$

These can be solved as $a_{1,k} = -((v_k)^t v_1)/((v_1)^t v_1)$. Hence,

$$[f_1, w_2, \dots, w_n] = [v_1, v_2, \dots, v_n] \begin{pmatrix} 1 & a_{1,2} & a_{1,3} & \dots & a_{1,n} \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

We can then solve the problem by induction. We replace the vectors $w_2, \dots, w_n$ by linear combinations $f_2, \dots, f_n$. that are orthogonal. By induction, we can say that this change of basis is is given by a matrix equation

$$[f_1, f_2, \dots, f_n] = [f_1, w_2, \dots, w_n] \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & b_{2,3} & \dots & b_{2,n} \\ 0 & 0 & 1 & \dots & b_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

Combining the two we get an equation of the form

$$[f_1, d_2, \dots, f_n] = [v_1, v_2, \dots, v_n]U$$

where U is a matrix which has 1's along the diagonal and zeroes the diagonal; i.e. is of the form $1 + N$ where N is strictly upper triangular.

Exercise: Show that the inverse of a strictly upper triangular matrix U is also strictly upper triangular. (Hint: The strictly strictly upper triangular matrix $N = (U - 1)$ is nilpotent. Show that $U = 1 + N$ has an inverse of the form $1 - N + N^2 - \cdots + (-1)^{n-1} N^{n-1}$.)

Now assuming that we are over the field of real numbers, we can find positive square-roots s_j of each of the numbers $(f_j)^t f_j$. If we put $u_j = f_j/s_j$ we have

$$[u_1, \dots, u_n] = [f_1, \dots, f_n]\mathrm{diag}(1/s_1, \dots, 1/s_n)$$

The latter is a diagonal matrix with positive real numbers along the diagonal. Moreover, the matrix $[u_1, \dots, u_n]$ is an orthogonal matrix. Turning this around, we get an expression

$$[v_1, \dots, v_n] = [u_1, \dots, u_n]\mathrm{diag}(s_1, \dots, s_n)U^{-1}$$

Since the only condition we have used is that the vectors v_i are linearly independent (**Exercise**: Where did we use this?), we see that *any* invertible matrix over real numbers can be written as the product of an orthogonal matrix, a diagonal matrix with positive diagonal entries and a matrix of the form $1 + N$ where N is strictly lower triangular.

Canonical Form

Given a symmetric matrix A and an eigenvector v for A. In other words, $v \neq 0$ is such that $Av = av$ for some element a of the field.

If $v^t w = 0$, then

$$v^t(Aw) = (v^t A)w = (A^t v)^t w = (Av)^t w = av^t w = 0$$

In other words, if w is perpendicular to v then so is Aw. In other words, A takes the space $v^\perp = \{w | v^t w = 0\}$ to itself.

Thus, putting v as the first vector in an *orthonormal* basis, we see that the matrix of A takes the block form

$$\begin{pmatrix} a & 0 \\ 0 & A_1 \end{pmatrix}$$

where A_1 is a symmetric matrix of 1 dimension less than A.

If we can repeat this process by finding an eigenvector of A_1 and so on, we ultimately get an orthonormal basis of eigenvectors. This would mean that the canonical form of A is a diagonal matrix. Moreover, the change of basis matrix $G = [v_1, \ldots, v_n]$, so that $G^{-1}AG$ is diagonal would have the property that $v_1, \ldots, v_n$ is an orthonormal basis. In other words, $G^t G = 1$, or equivalently $G^t = G^{-1}$.

However, we have not explained why we can find an eigenvector for a symmetric matrix. For the proof, we need to assume that the field of coefficients is the field $\mathbb{R}$ of real numbers and the proof will use some analysis in $\mathbb{R}^n$.

The set $S^{n-1} = \{v | v^t v = 1\}$ of unit vectors in n-space is a *compact* set. Moreover, the restriction of the quadratic form Q to it is a conditnuously differentiable function. Hence, there is a vector v_0 in S^{n-1} where Q takes a maximum value. The claim is that this is an eigenvector.

If a differentiable function f on S^{n-1} takes a maximum at a vector v_0, then the directional derivative of f along any vector w which is tangent to S^{n-1} at v_0 must be 0. Such a vector w is given by the condition $v_0^t w = 0$. Thus we get the condition

$$(\nabla w f)_{|v_0} = 0 \text{ for all } w \text{ such that } v_0^t w = 0$$

For the quadratic form Q we compute

$$(\nabla w Q)_{|v_0} = \lim_{t \to 0} \frac{Q(v_0 + tw) - Q(v_0)}{t}$$
$$= \lim_{t \to 0} \frac{2tA(v_0, w) + t^2 Q(w)}{t}$$
$$= 2A(v_0, w)$$

The above condition thus becomes

$$A(v_0, w) = 0 \text{ for all } w \text{ such that } v_0^t w = 0$$

Since $v_0 \neq 0$ (it is on S^{n-1}) and $A(v_0, w) = (Av_0)^t w$, this is means that Av_0 is proportional to v_0. In other words, v_0 is an eigenvector.

Combining this existence theorem with the previous argument shows that for every symmetric matrix A over real numbers there is an orthogonal matrix G (i.e. $G^t = G^{-1}$) so that $G^{-1}AG = G^t AG$ is a diagonal matrix.

Note that the number of 0, positive and negative eigenvalues A is exactly the same as the numbers n, p and q defined earlier. This shows that these numbers depend only on A.

Note also that this means that a symmetric matrix over any subfield of the field of real numbers (for example the field of rational numbers) can be diagonalised over the field of real numbers. As shown in earlier sections, this implies that the minimal polynomial of A has distinct roots.