MA1112: Linear Algebra II

Dr. Vladimir Dotsenko (Vlad)

Lecture 19

The eigenvalue theorem

Our last step would be to prove a theorem relating linear transformations and bilinear forms. Suppose that B is a real symmetric matrix representing a bilinear form b.

Theorem 1. *The matrix B has only real eigenvalues. The signature of b is completely determined by eigenvalues of B: the number n_+ is the number of positive eigenvalues, the number n_- is the number of negative eigenvalues, and the number n_0 is the number of zero eigenvalues.*

Proof. As it often happens in mathematics, we shall find that it is easier to prove a bit more than the theorem actually states. Let us view the matrix B as a matrix of a linear transformation φ of $\mathbb{R}^n$ that multiplies every vector by B, and equip $\mathbb{R}^n$ with the standard scalar product. Let us note that $(x, y) = y^T x$, so

$$(x, \varphi(y)) = (By)^T x = y^T B^T x = y^T B x = (\varphi(x), y)$$

for all vectors x and y. A linear transformation of a Euclidean vector space for which this property hold is said to be symmetric; this is a way to define a property that does not depend on the choice of a basis, which is very advantageous. Note that $(\varphi(x), y)$ is equal to $b(x, y)$, where b is the bilinear form we are considering; also, $q(x) = b(x, x) = (\varphi(x), x)$.

The promised stronger statement that we shall now prove by induction on the dimension of the space claims that every symmetric transformation of an Euclidean vector space has an orthonormal basis of eigenvectors. For a one-dimensional space, every linear transformation is symmetric, and we can take either of the two vectors of length 1 to create an orthonormal basis of eigenvectors. Suppose that V is a Euclidean vector space of dimension n. Let us consider the function $q(x)$ on the unit sphere $|x| = 1$. This function is continuous, and a continuous function reaches its maximal and minimal value on any compact (closed and bounded) set (this will be proved in one of your analysis modules later). Therefore, for all x with $|x| = 1$ we have $m \leqslant q(x) \leqslant M$ for some m and M, and these inequalities become equalities for some x. Now, note that for $x \neq 0$ we have

$$q(x) = (\varphi(x), x) = \left(\varphi\left(|x|\frac{1}{|x|}x\right), |x|\frac{1}{|x|}x\right) = |x|^2 \left(\varphi\left(\frac{1}{|x|}x\right), \frac{1}{|x|}x\right),$$

and the vector $\frac{1}{|x|}x$ is of length 1 for each $x \neq 0$, because $\left(\frac{1}{|x|}x, \frac{1}{|x|}x\right) = \frac{1}{|x|^2}(x, x) = 1$. Therefore, for each $x \neq 0$ we have

$$m(x, x) = m|x|^2 \leqslant q(x) \leqslant M|x|^2 = M(x, x),$$

and by inspection this holds for $x = 0$ also. In particular, this implies that

$$q(x) - M(x, x) \leqslant 0$$

for all x, so the values of x where $q(x) = M(x, x)$ are solutions to the local maximum problem for the function $q(x) - M(x, x)$. This means that the gradient of that function must be equal to zero at each point where maximum is attained. By examining the formula

$$\begin{aligned} f(x) &= (\varphi(x), x) - M(x, x) = \\ &= (b_{11} - M)x_1^2 + 2b_{12}x_1x_2 + \cdots + 2b_{1n}x_1x_n + (b_{22} - M)x_2^2 + 2b_{23}x_2x_3 + \cdots + (b_{nn} - M)x_n^2, \end{aligned}$$

we note that

$$\begin{aligned}
\frac{\partial f}{\partial x_1} &= 2(b_{11}-M)x_1 + 2b_{12}x_2 + \cdots + 2b_{1n}x_n,\\
\frac{\partial f}{\partial x_2} &= 2b_{12}x_1 + 2(b_{22}-M)x_2 + \cdots + 2b_{2n}x_n,\\
&\cdots\\
\frac{\partial f}{\partial x_n} &= 2b_{1n}x_1 + 2b_{2n}x_2 + \cdots + 2(b_{nn}-M)x_n.
\end{aligned}$$

Therefore, the gradient vanishes at $v \neq 0$ if and only if $(\varphi - M\cdot I)(v) = 0$, that is v is an eigenvector of φ with the eigenvalue M; we may assume $|v| = 1$. Let us now consider the vector space $U = \mathrm{span}(v)^{\perp}$. It turns out that U is an invariant subspace of φ. Let $u \in U$, so that $(u, v) = 0$. We have $(\varphi(u), v) = (u, \varphi(v)) = (u, \lambda v) = \lambda(u, v) = 0$, so U is indeed an invariant subspace. The scalar product of V makes U a Euclidean space, and φ is a symmetric linear transformation of that space. By induction, there is an orthonormal basis of eigenvectors of φ in U; adjoining the vector v to it makes it an orthonormal basis of V.

Finally, let us prove the statement about signature. Let $v_1, \dots, v_n$ be an orthonormal basis of eigenvectors of φ. Then $b(v_i, v_j) = v_j^T B v_i = v_j^T \lambda_i v_i = \lambda_i (v_i, v_j)$, therefore, relative to that basis, the matrix of b is diagonal with eigenvalues on the diagonal. We conclude that

$$q(x_1 v_1 + \cdots + c_n v_n) = \lambda_1 x_1^2 + \cdots + \lambda_n x_n^2,$$

which proves the claim. □

Example 1. For the bilinear form whose matrix relative to the standard basis of $\mathbb{R}^4$ is

$$\begin{pmatrix} 7 & 1 & -1 & -3\\ 1 & 7 & -3 & -1\\ -1 & -3 & 7 & 1\\ -3 & -1 & 1 & 7 \end{pmatrix},$$

we can attempt to find a canonical form by finding an orthonormal basis of eigenvectors of the corresponding matrix.

The eigenvalues of this matrix are 4, 8, and 12. Orthonormal basis of eigenvectors:

$$\frac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\1\\0\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\0\\1\end{pmatrix}, \frac{1}{2}\begin{pmatrix}-1\\1\\-1\\1\end{pmatrix}, \frac{1}{2}\begin{pmatrix}-1\\-1\\1\\1\end{pmatrix}.$$

It is not unique; one can choose an arbitrary orthonormal basis in the plane spanned by the first two of them.

Note that eigenvectors corresponding to different eigenvalues are automatically orthogonal:

$$c_1(v_1, v_2) = (c_1 v_1, v_2) = (Av_1, v_2) = (v_1, Av_2) = (v_1, c_2 v_2) = c_2(v_1, v_2),$$

which for $c_1 \neq c_2$ implies that $(v_1, v_2) = 0$. Therefore, the only things we need to do is normalise the eigenvectors for the eigenvalues 8 and 12, since each of these has a one-dimensional space of eigenvectors, and find an orthonormal basis of the solution set of $(A - 4I)x = 0$, which can be obtained from any basis of that space by Gram–Schmidt orthogonalisation.

Let us give an example for how the proof of the Jacobi theorem from the previous lecture applies in practice.

Example for the Jacobi theorem

Suppose that a bilinear form has the matrix $\begin{pmatrix} 3 & 1 & -1 \\ 1 & 1 & -2 \\ -1 & -2 & 1 \end{pmatrix}$ relative to a basis e_1, e_2, e_3.

Let us compute the determinants $\Delta_1, \Delta_2, \Delta_3$. We have $\Delta_1 = 3$, $\Delta_2 = 2$, $\Delta_3 = -7$. Therefore, the conditions of the Jacobi theorem are satisfied.

We are looking for a basis of the form

$$\begin{aligned} f_1 &= \alpha_{11} e_1, \\ f_2 &= \alpha_{12} e_1 + \alpha_{22} e_2, \\ f_3 &= \alpha_{13} e_1 + \alpha_{23} e_2 + \alpha_{33} e_3, \end{aligned}$$

imposing equations $A(e_i, f_j) = 0$ for $i < j$, and $A(e_i, f_i) = 1$ for all i. This means that

$$1 = A(e_1, f_1) = 3\alpha_{11},$$

$$\begin{aligned} 0 &= A(e_1, f_2) = 3\alpha_{12} + \alpha_{22}, \\ 1 &= A(e_2, f_2) = \alpha_{12} + \alpha_{22}, \end{aligned}$$

$$\begin{aligned} 0 &= A(e_1, f_3) = 3\alpha_{13} + \alpha_{23} - \alpha_{33}, \\ 0 &= A(e_2, f_3) = \alpha_{13} + \alpha_{23} - 2\alpha_{33}, \\ 1 &= A(e_3, f_3) = -\alpha_{13} - 2\alpha_{23} + \alpha_{33}. \end{aligned}$$

Solving these linear equations, we get $\alpha_{11} = \frac{1}{3}$, $\alpha_{12} = -\frac{1}{2}$, $\alpha_{22} = \frac{3}{2}$, $\alpha_{13} = \frac{1}{7}$, $\alpha_{23} = -\frac{5}{7}$, $\alpha_{33} = -\frac{2}{7}$, so the corresponding change of basis is

$$\begin{aligned} f_1 &= \frac{1}{3} e_1, \\ f_2 &= -\frac{1}{2} e_1 + \frac{3}{2} e_2, \\ f_3 &= \frac{1}{7} e_1 - \frac{5}{7} e_2 - \frac{2}{7} e_3. \end{aligned}$$