1212: Linear Algebra II

Dr. Vladimir Dotsenko (Vlad)

Lecture 9

Diagonalisation of real symmetric matrices

Yesterday we proved that every real symmetric matrix A has a real eigenvector $v \in \mathbb{R}^n$, so that $Av = cv$ for some $c \in \mathbb{R}$.

Theorem 1. *For a real symmetric matrix, there exists a basis of eigenvectors that is orthonormal with respect to the standard scalar product on $\mathbb{R}^n$.*

Proof. We begin the proof of a very useful observation which we shall use several times. Namely for two vectors x and y, the scalar product (x, y) can be identified with the 1×1-matrix $y^T \cdot x$. Therefore, for every matrix B we have

$$(Bx, y) = y^T Bx = (B^T y)^T x = (x, B^T y).$$

In particular, for $B = A$, we have $(Ax, y) = (x, Ay)$. This is very useful, because it allows to generalise the notion of a symmetric matrix to that of a symmetric linear transformation of a Euclidean vector space, that is the transformation A for which $(Ax, y) = (x, Ay)$ for all vectors x and y. (If we introduce a coordinate system corresponding to an orthonormal basis, the vector space gets identified with $\mathbb{R}^n$ with the usual scalar product, and a symmetric linear transformation is represented by a symmetric matrix).

Let us proceed with the proof. Consider a real eigenvector v of A, and consider the vector space $U = \operatorname{span}(v)^{\perp}$. We have $V = \operatorname{span}(V) \oplus U$, so $\dim(U) = \dim(V) - 1$. Let us show that U is an invariant subspace of A. Let $u \in U$, so that $(u, v) = 0$, and let us show that $(Au, v) = 0$. We have $(Au, v) = (u, Av) = (u, cv) = c(u, v) = 0$. Therefore, U is an invariant subspace; the scalar product of V makes it an Euclidean space, and A is a symmetric linear transformation of that space. It follows by induction on dimension that the theorem holds. $\square$

Example 1. For the transformation whose matrix relative to the standard orthonormal (with respect to the standard scalar product) basis of $\mathbb{R}^4$ is

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

let us find an orthonormal basis of eigenvectors.

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.

Orthogonal matrices

Definition 1. An $n \times n$-matrix A is said to be orthogonal if $A^T A = I$. (Or, equivalently, if $A^T = A^{-1}$).

Note that another way to state the same is to remark that the columns of A form an orthonormal basis. Indeed, the entries of $A^T A$ are pairwise scalar products of columns of A.

Theorem 2. *A matrix A is orthogonal if and only if it does not change the scalar product, that is for all $x, y \in \mathbb{R}^n$ we have*

$$(Ax, Ay) = (x, y).$$

Proof. As we proved earlier, $(Ax, Ay) = (x, A^T Ay)$. Clearly, $(x, A^T Ay) = (x, y)$ for all x, y if and only if $(x, (A^T A - I)y) = 0$ for all x, y, and the latter happens only for $A^T A - I = 0$. $\square$

This latter result, again, has the advantage of being coordinate-independent: it allows to define an orthogonal linear transformation of an arbitrary Euclidean space as a transformation for which $(Ax, Ay) = (x, y)$ for all vectors x, y. This means that such a transformation preserves geometric notions like lengths and angles between vectors.

Comparing determinants of $A^T A$ and I, we conclude that $\det(A)^2 = 1$, so $\det(A) = \pm 1$. Intuitively, orthogonal matrices with $\det(A) = 1$ are transformations that can distinguish between left and right, or clockwise and counterclockwise (like rotations), and orthogonal matrices with $\det(A) = -1$ are transformations that swap clockwise with counterclockwise (like mirror symmetry).

Example 2. Let A be an orthogonal 2×2-matrix with $\det(A) = 1$. We have $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with $a^2 + c^2 = 1$, $b^2 + d^2 = 1$, $ab + cd = 0$. There exist some angle α such that $a = \cos\alpha$, $c = \sin\alpha$, and the vector $\begin{pmatrix} b \\ d \end{pmatrix}$ is an orthogonal vector of length 1, so $\begin{pmatrix} b \\ d \end{pmatrix} = \pm \begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix}$. Because of the determinant condition,

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix},$$

which is the matrix of the rotation through α about the origin.