MA1112: Linear Algebra II

Dr. Vladimir Dotsenko (Vlad)

Lecture 20

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

$$(\varphi(x), \varphi(y)) = (x, y).$$

Proof. We have

$$(\varphi(x), \varphi(y)) = (Ax, Ay) = (Ay)^T Ax = y^T A^T Ax = (A^T Ax, y).$$

Clearly, $(A^T Ax, y) = (x, y)$ for all x, y if and only if $(A^T A - I)x, y) = 0$ for all x, y, and the latter happens only for $A^T A - I = 0$. □

This latter result 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 $(\varphi(x), \varphi(y)) = (x, y)$ 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 for an orthogonal matrix A, we have $\det(A)^2 = 1$, 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). This intuition is supported by the following examples.

Example 1. 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.

Example 2. Let us consider an orthogonal 3×3-matrix A with $\det(A) = 1$. We have

$$\chi_A(t) = \det(A - tI) = \det(A) + a_1 t + a_2 t^2 - t^3 = 1 + a_1 t + a_2 t^2 - t^3.$$

Clearly, $\chi_A(0) = 1$, and $\chi_A(t)$ is negative for large t, so by continuity, $\chi_A(t)$ has a real root, and A has a real eigenvalue λ. Let v be such that $Av = \lambda v$. Then $(v, v) = (Av, Av) = (\lambda v, \lambda v) = \lambda^2(v, v)$, so $\lambda^2 = 1$. If $\lambda = 1$, we are done. Otherwise, our operator has an eigenvalue -1, and $\chi_A(t) = (1 + t)(1 + at - t^2)$, where the polynomial $1 + at - t^2$ must have real roots, so all eigenvalues of A in this case must be real. The product of the eigenvalues is equal to the determinant, so all of them cannot be equal to -1, and this operator has an eigenvalue 1.

If $Av = v$, then $\text{span}(v)^\perp$ is an invariant subspace, where our operator defines a linear operator whose matrix relative to an orthonormal basis is orthogonal and has determinant 1, therefore is a rotation. Consequently, our original matrix represents a rotation about the line containing v.

Hermitian vector spaces

We would like to adapt to the case of complex numbers some of the results that we proved before. However, we started with defining scalar products, and for complex numbers, the notion of a positive number does not make sense. So we have to be a bit more imaginative.

Definition 2. A vector space V over complex numbers is said to be a Hermitian vector space if it is equipped with a function (Hermitian scalar product) $V \times V \to \mathbb{C}$, $v_1, v_2 \mapsto (v_1, v_2)$ satisfying the following conditions:

- sesquilinearity: $(v_1 + v_2, v) = (v_1, v) + (v_2, v)$, $(v, v_1 + v_2) = (v, v_1) + (v, v_2)$, $(cv_1, v_2) = c(v_1, v_2)$, and $(v_1, cv_2) = \overline{c}(v_1, v_2)$,
- symmetry: $(v_1, v_2) = \overline{(v_2, v_1)}$ for all v_1, v_2 (in particular, $(v, v) \in \mathbb{R}$ for all v),
- positivity: $(v, v) \geqslant 0$ for all v, and $(v, v) = 0$ only for $v = 0$.

Example 3. Let $V = \mathbb{C}^n$ with the *standard scalar product*

$$\left(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}\right) = x_1\overline{y}_1 + x_2\overline{y}_2 + \cdots + x_n\overline{y}_n.$$

All the three properties from the definition of a Hermitian space are trivially true.

Lemma 1. *For every Hermitian scalar product and every basis $e_1, \ldots, e_n$ of V, we have*

$$(x_1 e_1 + \ldots + x_n e_n, y_1 e_1 + \ldots + y_n e_n) = \sum_{i,j=1}^{n} a_{ij} x_i \overline{y}_j,$$

where $a_{ij} = (e_i, e_j)$.

This follows immediately from the sesquilinearity property of Hermitian scalar products.