Notes 22 – Symmetric and Orthogonal Matrices

In this lecture, we focus attention on symmetric matrices, whose eigenvectors can be used to construct orthogonal matrices. Determinants will then help us to distinguish those orthogonal matrices that define rotations.

L22.1 Orthogonal eigenvectors. Recall the definition of the *dot* or *scalar product* of two column vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^{n,1}$. Without writing out their components, we can nonetheless assert that

$$\mathbf{v} \cdot \mathbf{w} = \mathbf{v}^\top \mathbf{w}. \tag{1}$$

Recall too that a matrix S is *symmetric* if $S^\top = S$ (this implies of course that it is square).

Lemma. *Let $\mathbf{v}_1, \mathbf{v}_2$ be eigenvectors of a symmetric matrix S corresponding to* distinct *eigenvalues λ_1, λ_2. Then $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$.*

Proof. First note that

$$(S\mathbf{v}_1) \cdot \mathbf{v}_2 = (S\mathbf{v}_1)^\top \mathbf{v}_2 = \mathbf{v}^\top{}_1 S^\top \mathbf{v}_2 = \mathbf{v}^\top{}_1 (S\mathbf{v}_2) = \mathbf{v}_1 \cdot (S\mathbf{v}_2).$$

This is true for *any* vectors $\mathbf{v}_1, \mathbf{v}_2$, but the assumptions $S\mathbf{v}_1 = \lambda_1 \mathbf{v}_1$ and $S\mathbf{v}_2 = \lambda_2 \mathbf{v}_2$ tell us that

$$\lambda_1 \mathbf{v}_1 \cdot \mathbf{v}_2 = \lambda_2 \mathbf{v}_1 \cdot \mathbf{v}_2,$$

and the result follows. QED

Example. To begin with the 2×2 case, consider the symmetric matrix

$$A = \begin{pmatrix} 5 & 2 \\ 2 & 8 \end{pmatrix}.$$

It is easy to check that its eigenvalues are 9 and 4, and that respective eigenvectors are

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \qquad \mathbf{v}_2 = \begin{pmatrix} -2 \\ 1 \end{pmatrix}.$$

As predicted by the Lemma, $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$. Given this fact, we can normalize $\mathbf{v}_1, \mathbf{v}_2$ to manufacture an *orthonormal* basis

$$\mathbf{f}_1 = \tfrac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \qquad \mathbf{f}_2 = \tfrac{1}{\sqrt{5}} \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

of eigenvectors, and use *these* to define the matrix

$$P = \tfrac{1}{\sqrt{5}} \left(\begin{array}{c|c} 1 & -2 \\ 2 & 1 \end{array} \right).$$

With this choice,

$$P^\top P = \tfrac{1}{5} \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \left(\begin{array}{c|c} 1 & -2 \\ 2 & 1 \end{array} \right) = \tfrac{1}{5} \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix} = I_2. \tag{2}$$

Another way of expressing this relationship is

$$P^{-1} = \tfrac{1}{\sqrt{5}} \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} = P^\top,$$

and also $PP^\top = I$. It is easy to verify that

$$P^{-1}AP = \tfrac{1}{5}\begin{pmatrix}1 & 2\\ -2 & 1\end{pmatrix}\begin{pmatrix}5 & 2\\ 2 & 8\end{pmatrix}\begin{pmatrix}1 & -2\\ 2 & 1\end{pmatrix} = \begin{pmatrix}9 & 0\\ 0 & 4\end{pmatrix}.$$

Definition. *A matrix $P \in \mathbb{R}^{n,n}$ is called* orthogonal *if it satisfies one of the equivalent conditions: (i) $P^\top P = I_n$, (ii) $PP^\top = I_n$, (iii) P is invertible and $P^{-1} = P^\top$.*

This definition was first given in L9.3 in the context of orthonormal bases. Let us explain why the three conditions are indeed equivalent. As (1) and (2) make clear, condition (i) asserts that the *columns* of P are orthonormal. Condition (ii) assserts that the *rows* are orthonormal. A set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of orthonormal vectors is necessarily LI since

$$a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = \mathbf{0}$$

implies (by taking the dot product with each $\mathbf{v}_i$ in turn) that $a_i = 0$; thus either (i) or (ii) implies that P is invertible. It follows that both (i) and (ii) are equivalent to (iii).

The relationship between symmetric and orthogonal matrices is cemented by the

Theorem. *Let $S \in \mathbb{R}^{n,n}$ be a symmetric matrix. Then*
(i) the eigenvalues (or roots of the characteristic polynomial $p(x)$) of S are all real.
(ii) there exists an orthogonal *matrix P such that $P^{-1}SP = P^\top SP = D$.*

Proof. (i) Suppose that $\lambda \in \mathbb{C}$ is a root of $p(x)$. Working over the field $\mathbb{C}$, we can assert that there exists a *complex* eigenvector $\mathbf{v} \in \mathbb{C}^{n,1}$ satisfying $S\mathbf{v} = \lambda\mathbf{v}$. If $\mathbf{v}^\top = (z_1, \ldots, z_n)$ then the *complex conjugate* of this vector is $\overline{\mathbf{v}}^\top = (\overline{z}_1, \ldots, \overline{z}_n)$ and

$$\overline{\mathbf{v}}^\top \mathbf{v} = |z_1|^2 + \cdots + |z_n|^2 > 0,$$

since $\mathbf{v} \neq \mathbf{0}$. Thus

$$\lambda\,\overline{\mathbf{v}}^\top \mathbf{v} = \overline{\mathbf{v}}^\top (S\mathbf{v}) = \overline{S\mathbf{v}}^\top \mathbf{v} = \overline{\lambda}\,\overline{\mathbf{v}}^\top \mathbf{v},$$

and necessarily $\lambda = \overline{\lambda}$ and $\lambda \in \mathbb{R}$.

In the light of (i), part (ii) follows immediately if all the roots of $p(x)$ are distinct. For each repeated root λ, one needs to know that $\mathrm{mult}(\lambda) = \dim E_\lambda$; for if this is true the Lemma permits us to build up an orthonormal basis of eigenvectors. We shall not prove the multiplicity statement (that *is* always true for a symmetric matrix), but a convincing exercise follows. QED

Exercise. Consider again the symmetric matrix

$$A = \begin{pmatrix}-2 & 1 & 1\\ 1 & -2 & 1\\ 1 & 1 & -2\end{pmatrix},$$

and its eigenvectors

$$\mathbf{v}_1 = \begin{pmatrix}1\\1\\1\end{pmatrix}, \qquad \mathbf{v}_2 = \begin{pmatrix}1\\-1\\0\end{pmatrix}, \qquad \mathbf{v}_3 = \begin{pmatrix}1\\0\\-1\end{pmatrix}$$

(found in L18.2), with respective eigenvalues 0 (multiplicity 1) and -3 (multiplicity 2). As predicted by the Lemma, $\mathbf{v}_1\cdot\mathbf{v}_2 = 0 = \mathbf{v}_1\cdot\mathbf{v}_3$. Observe however that $\mathbf{v}_2\cdot\mathbf{v}_3 \neq 0$; show nonetheless that there exists an eigenvector $\mathbf{v}_3'$ with eigenvalue 3 such that $\mathbf{v}_2\cdot\mathbf{v}_3' = 0$. Normalize the vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3'$ so as to obtain an *orthogonal* matrix P for which $P^{-1}AP$ is diagonal. Compute the determinant of P; can the latter be chosen so that $\det P = 1$?

L22.2 Rotations in the plane. We proceed to classify 2×2 orthogonal matrices. Suppose that

$$P = \begin{pmatrix} p & q \\ r & s \end{pmatrix}$$

is orthogonal, so that both columns are mutually orthogonal unit vectors. The only two unit vectors orthogonal to $\begin{pmatrix} p \\ r \end{pmatrix}$ are $\begin{pmatrix} -r \\ p \end{pmatrix}$ and $\begin{pmatrix} r \\ -p \end{pmatrix}$, so one of these must equal $\begin{pmatrix} q \\ s \end{pmatrix}$. This gives us two possibilities:

$$P = \begin{pmatrix} p & -r \\ r & p \end{pmatrix} \qquad \text{or} \qquad P = \begin{pmatrix} p & r \\ r & -p \end{pmatrix}.$$

Since $p^2 + r^2 = 1$, the first matrix has determinant 1 and the second -1.

Let us focus attention on the first case. There exists a unique angle θ such that $\cos\theta = p$ and $\sin\theta = r$. We denote the resulting matrix P by

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \tag{3}$$

to emphasize that it is a function of θ. Observe that

$$R_\theta \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{pmatrix}.$$

The right-hand side is the image of the vector $\begin{pmatrix} x \\ y \end{pmatrix}$ under a rotation about the origin by an angle θ. To summarize:

Proposition. *Any 2×2 orthogonal matrix with determinant 1 has the form (3), and represents a rotation in $\mathbb{R}^2$ by an angle θ anti-clockwise, with centre the origin.*

Exercise. Use standard trigonometric identities to verify that

$$R_\theta R_\phi = R_{\theta+\phi}.$$

Deduce that $R_{2\theta} = (R_\theta)^2$ and $R_{-\theta} = (R_\theta)^{-1}$.

We next extend some of these results to bigger orthogonal matrices.

L22.3 Properties of orthogonal matrices. Recall that $(AB)^T = B^T A^T$ is always true.

Lemma. *If A, B are orthogonal matrices of the same size, then A^{-1} and AB are also orthogonal.*

Proof. Suppose that $A^\top A = I$ and $B^\top B = I$. Then $A^\top = A^{-1}$; this implies that

$$(A^{-1})^\top = A, \quad \text{and so} \quad (A^{-1})^\top A^{-1} = I,$$

so A^{-1} is orthogonal. Moreover,

$$(AB)^\top (AB) = B^\top A^\top AB = B^\top IB = B^\top B = I,$$

as required. QED

In order to compute the determinant of an orthogonal matrix, we need the following fundamental result that we quote without proof.

Binet's Theorem. *If A, B are square matrices of the same size, then*

$$\det(AB) = (\det A)(\det B).$$

Suppose that $P^\top P = I$. Then

$$1 = \det I = \det(P^\top P) = \det(P^\top)\det P = (\det P)^2,$$

since $\det(P^\top) = \det P$.

Corollary. *Any orthogonal matrix has determinant equal to 1 or -1.*

Example. Suppose that P is an orthogonal matrix with $\det P = 1$. Then

$$\begin{aligned}\det(P-I) = \det(P^\top)\det(P-I) &= \det\left((P^\top)(P-I)\right)\\ &= \det\left(P^\top P - P^\top\right)\\ &= \det(I - P^\top)\\ &= \det\left((I-P)^\top\right)\\ &= \det(I-P) = (-1)^n \det(P-I).\end{aligned}$$

If n is odd, it follows that $\det(P-I)=0$, so 1 is a root of the characteristic polynomial. Therefore 1 is an eigenvalue of P, and there exists $\mathbf{v} \in \mathbb{R}^{3,1}$ such that $P\mathbf{v} = \mathbf{v}$ (and $\mathbf{v} \neq \mathbf{0}$).

This example can be used to show that any 3×3 orthogonal matrix P with $\det P = 1$ represents a rotation of $\mathbb{R}^3$ about an axis passing through the origin. For given such a rotation, one can choose an orthonormal basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ of $\mathbb{R}^3$ such that $\mathbf{v}_3$ points in the direction of the axis of rotation. It then follows (from an understanding of what is meant by a rotation of a rigid body in space, and referring to (3)) that the rotation is described by a *linear* mapping $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ whose matrix with respect to the basis is

$$M_\theta = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

For example, $\mathbf{v} \mapsto M_\theta \mathbf{v}$ is itself a rotation about the z-axis.

L22.4 Further exercises.

1. For which values of θ does the rotation matrix (3) have a real eigenvalue?

2. Show that if an $n \times n$ matrix S is both symmetric *and* orthogonal then $S^2 = I$. Deduce that the eigenvalues of S are 1 or -1.

3. An *isometry* of $\mathbb{R}^n$ is any mapping $f : \mathbb{R}^n \to \mathbb{R}^n$ such that $|f(\mathbf{v}) - f(\mathbf{w})| = |\mathbf{v} - \mathbf{w}|$ for all $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$. Show that such a mapping is necessarily injective. Now suppose that $\mathbf{v}_0 \in \mathbb{R}^{n,1}$ is a fixed vector and that $P \in \mathbb{R}^{n,n}$ is an orthogonal matrix. Set $g(\mathbf{v}) = \mathbf{v}_0 + P\mathbf{v}$. Verify that g is an isometry; is it surjective?

4. [Uses the complex field $\mathbb{C}$.] Find a matrix $Y \in \mathbb{C}^{2,2}$ for which $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = Y^{-1}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} Y$.