33 Orthogonal Diagonalisation

By the end of this section, you should be able to answer the following questions:

- What is a symmetric matrix?
- What is an orthogonal matrix?
- How do you diagonalise symmetric matrices?

Given an $n \times n$ matrix A, we call A *orthogonally diagonalisable* if there exists an orthogonal matrix P such that $P^{-1}AP = \boxed{P^TAP}$ is diagonal. To understand this, we first need to know what is meant by an orthogonal matrix.

33.1 Orthogonal matrices

An *orthogonal* matrix is a real square matrix Q such that the columns of Q are mutually orthogonal unit vectors (i.e. $\boldsymbol{v}_i \cdot \boldsymbol{v}_j = 0$ if $i \neq j$, and $|\boldsymbol{v}_i| = 1$).

Note that mutually orthogonal unit vectors are called *orthonormal*.

An orthogonal matrix is then a real square matrix Q such that $Q^{-1} = Q^T$. Note also that $\det(Q) = \pm 1$.

Q orthogonal $\Longrightarrow QQ^T = I$

$\Longrightarrow \det(QQ^T) = \det(I)$

$\Longrightarrow \det Q \det Q^T = 1$

$\Longrightarrow (\det Q)^2 = 1$

(aside: in $\mathbb{R}^2$, orthog. matrix, $\det = 1 \Leftrightarrow$ rotation.
" " $\det = -1 \Leftrightarrow$ reflection)

33.2 Symmetric matrices

A matrix A is *symmetric* if and only if $A = A^T$. Symmetric matrices are easy to identify due to their "mirror symmetry" about the main diagonal. For example, we can tell by inspection that $A = \begin{pmatrix} -3 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -3 \end{pmatrix}$ is symmetric.

33.2.1 If A is real symmetric, then the eigenvectors corresponding to different eigenvalues are orthogonal.

Proof: $(AB)^T = B^T A^T$.

$$A\underline{v}_1 = \lambda_1 \underline{v}_1 \quad , \quad A\underline{v}_2 = \lambda_2 \underline{v}_2$$

$$\lambda_1 \neq \lambda_2$$

Want to show $\underline{v}_1 \cdot \underline{v}_2 = 0$ $\left[(u_1, v_1, w_1) \begin{pmatrix} u_2 \\ v_2 \\ w_2 \end{pmatrix} \right]$.

or equivalently $\underline{v}_1^T \underline{v}_2 = 0$

Consider
$$\begin{aligned} \lambda_1 \underline{v}_1^T \underline{v}_2 &= (\lambda_1 \underline{v}_1)^T \underline{v}_2 \\ &= (A\underline{v}_1)^T \underline{v}_2 \\ &= \underline{v}_1^T A^T \underline{v}_2 \quad \Big\} \; A = A^T. \\ &= \underline{v}_1^T A \underline{v}_2 \\ &= \lambda_2 \underline{v}_1^T \underline{v}_2 \end{aligned}$$

$$\Longrightarrow (\lambda_1 - \lambda_2)\, \underline{v}_1^T \underline{v}_2 = 0$$

Since $\lambda_1 \neq \lambda_2 \Rightarrow \underline{v}_1^T \underline{v}_2 = 0$ ...

33.2.2 Real symmetric matrices are orthogonally diagonalisable

It is straightforward to show that if a matrix is orthogonally diagonalisable, then it is symmetric: $\exists P$ s.t. $P^{-1} = P^T$ & $P^TAP = D$.

$$\begin{aligned} A^T &= (PDP^T)^T \\ &= (P^T)^T D^T P^T \\ &= PDP^T \\ &= A. \end{aligned}$$

logically equivalent to
"If A is NOT symmetric, then it is NOT orthog. diagonalisable".

In fact, the converse is also true (although difficult to prove), giving us the amazing result:

* An $n \times n$ real matrix is orthogonally diagonalisable if and only if it symmetric. *

The significance of this is that a symmetric matrix is *always* diagonalisable by an orthogonal matrix.

33.2.3 Eigenvectors and eigenvalues

Here we state two results about any symmetric matrix A without proof:

(1) All the eigenvalues of A are real;

(2) A has n linearly independent eigenvectors.

33.2.4 Example

Let $A = \begin{pmatrix} -3 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -3 \end{pmatrix}$ (see previous examples).

We already know the eigenvalues are $-3, -1, -4$ with corresponding eigenvectors

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

Note that A is real symmetric, so v_1, v_2 and v_3 should be pairwise orthogonal.

Note $\underline{v}_i \cdot \underline{v}_j = 0 \quad i \neq j$

$|\underline{v}_1| = \sqrt{1^2+0^2+(-1)^2} = \sqrt{2}$

$|\underline{v}_2| = \sqrt{6}$

$|\underline{v}_3| = \sqrt{3}$

"normalise" eigenvectors (i.e. turn them into unit vectors) $\hat{\underline{v}}_i = \dfrac{\underline{v}_i}{|\underline{v}_i|}$

$\hat{\underline{v}}_1 = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{pmatrix}, \quad \hat{\underline{v}}_2 = \begin{pmatrix} 1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{pmatrix}, \quad \hat{\underline{v}}_3 = \begin{pmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}.$

form an orthonormal set.

$\Rightarrow P = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & -1/\sqrt{3} \\ -1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \end{pmatrix} \quad \left(= \left(\hat{\underline{v}}_1 \middle| \hat{\underline{v}}_2 \middle| \hat{\underline{v}}_3 \right) \right).$

orthogonally diagonalises A, check $P^T = P^{-1}$

$P^T = \begin{pmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{6} & 2/\sqrt{6} & 1/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{3} & 1/\sqrt{3} \end{pmatrix}.$

i.e. check $PP^T = I$