Appendix B: Matrices, eigenvalues, and eigenvectors

This appendix covers the simple algebra of matrices, and some properties of eigenvalues and eigenvectors. The calculation of eigenvalues and eigenvectors is the main topic of Chapter 27.

Basic matrix algebra

For the most part, we will only need to consider the algebra of 2×2 matrices,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Addition of two matrices is component by component, so that

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \end{pmatrix},$$

while multiplication is given by

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} \end{pmatrix}.$$

We can write this more compactly by saying that

$$[\mathbb{A}\mathbb{B}]_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} = \sum_{k=1}^{2} a_{ik}b_{kj},$$

where $[\mathbb{A}\mathbb{B}]_{ij}$ is the entry in the ith row and jth column of the matrix $\mathbb{A}\mathbb{B}$.

One special matrix is the identity matrix,

$$\mathbb{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

which has the property that $\mathbb{I}\mathbb{A} = \mathbb{A}\mathbb{I} = \mathbb{A}$ for any 2×2 matrix $\mathbb{A}$.

A matrix $\mathbb{A}$ is said to be *invertible*, or *non-singular*, if there is another matrix $\mathbb{A}^{-1}$ such that

$$\mathbb{A}^{-1}\mathbb{A} = \mathbb{A}\mathbb{A}^{-1} = \mathbb{I}.$$

The matrix

$$\mathbb{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{B.1}$$

is invertible if and only if its determinant, $\det(\mathbb{A})$, given by

$$\det(\mathbb{A}) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

is not equal to zero, and then

$$\mathbb{A}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Matrices and vectors

Multiplication of vectors by matrices

In general we can calculate the product $\mathbb{A}\mathbb{B}$ when $\mathbb{A}$ is an $n \times m$ matrix and $\mathbb{B}$ is an $m \times k$ matrix (the columns of $\mathbb{A}$ have to match the rows of $\mathbb{B}$). In particular this allows us to calculate $\mathbb{A}\mathbf{x}$ if $\mathbb{A}$ is a 2×2 matrix and $\mathbf{x} = (x_1, \ x_2)$ is a two component vector,

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix}.$$

We can also write this more compactly as

$$[\mathbb{A}\mathbf{x}]_i = a_{i1}x_1 + a_{i2}x_2, \tag{B.2}$$

where $\mathbf{v}_i$ indicates the ith component of the vector $\mathbf{v}$. (Note that this means in particular that if $[\mathbf{v}_1 \ \ \mathbf{v}_2]$ is a matrix with columns made from the vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ then for a 2×2 matrix $\mathbb{A}$ we have

$$\mathbb{A}[\mathbf{v}_1 \ \ \mathbf{v}_2] = [\mathbb{A}\mathbf{v}_1 \ \ \mathbb{A}\mathbf{v}_2],$$

which is used repeatedly in Chapters 28–30.)

Solution of simultaneous equations

The simultaneous linear equations

$$\begin{aligned} ax_1 + bx_2 &= c_1 \\ cx_1 + dx_2 &= c_2 \end{aligned}$$

can be rewritten as the matrix equation

$$\mathbb{A}\mathbf{x} = \mathbf{c},$$

where $\mathbb{A}$ is defined as in (B.1), $\mathbf{x} = (x_1,\ x_2)$ and $\mathbf{c} = (c_1,\ c_2)$. This equation has a unique solution if and only if $\mathbb{A}$ is invertible, and then the solution is given by multiplying both sides by $\mathbb{A}^{-1}$,

$$\mathbf{x} = \mathbb{A}^{-1}\mathbf{c}.$$

It follows that $\mathbb{A}\mathbf{x} = \mathbf{0}$ can have a non-zero solution for $\mathbf{x}$ only if $\mathbb{A}$ is not invertible.

Eigenvalues and eigenvectors

If $\mathbf{v} \neq \mathbf{0}$ and

$$\mathbb{A}\mathbf{v} = \lambda\mathbf{v}$$

then λ is an eigenvalue of $\mathbb{A}$ and $\mathbf{v}$ is the corresponding eigenvector. The calculation of eigenvalues and eigenvectors for 2×2 matrices is treated in detail in Chapter 27.

Linear independence of eigenvectors

The eigenvectors corresponding to two distinct eigenvalues are linearly independent; if

$$\alpha\mathbf{v}_1 + \beta\mathbf{v}_2 = \mathbf{0} \tag{B.3}$$

then we can multiply both sides by $\mathbb{A}$ to obtain

$$\mathbb{A}(\alpha\mathbf{v}_1 + \beta\mathbf{v}_2) = \mathbf{0}.$$

Since $\mathbb{A}\mathbf{v}_j = \lambda_j\mathbf{v}_j$ we have

$$\alpha\lambda_1\mathbf{v}_1 + \beta\lambda_2\mathbf{v}_2 = \mathbf{0}. \tag{B.4}$$

While (B.3) requires that $\mathbf{v}_1 = -\beta\mathbf{v}_2/\alpha$, the second equation (B.4) says that $\mathbf{v}_1 = (\lambda_2/\lambda_1) \times (-\beta\mathbf{v}_2/\alpha)$. Since $\lambda_2 \neq \lambda_1$ these cannot both be true unless $\beta = 0$, in which case $\alpha = 0$ also, and $\mathbf{v}_1$ and $\mathbf{v}_2$ are therefore linearly independent.

The special case of symmetric matrices

The transpose of the matrix

$$\mathbb{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

written $\mathbb{A}^T$, is given by

$$\mathbb{A}^T = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix};$$

i.e. $[\mathbb{A}^T]_{ij} = [\mathbb{A}]_{ji}$. A matrix is called *symmetric* if $\mathbb{A} = \mathbb{A}^T$, i.e. if $[\mathbb{A}]_{ij} = [\mathbb{A}]_{ji}$.

A general 2×2 symmetric matrix is of the form

$$\begin{pmatrix} a & b \\ b & d \end{pmatrix}.$$

For such matrices all the eigenvalues are real, and eigenvectors corresponding to distinct eigenvalues are orthogonal.

To see that the eigenvalues are real, suppose that λ is an eigenvalue and $\mathbf{v} = (v_1, \ v_2)$ is the corresponding eigenvector. Then

$$\mathbb{A}\mathbf{v} = \lambda\mathbf{v}, \qquad \text{and} \qquad \mathbb{A}\mathbf{v}^* = \lambda^*\mathbf{v}^*,$$

where the second equation is the complex conjugate of the first. We take the inner (dot) product of the first equation with $\mathbf{v}^*$, and of the second with $\mathbf{v}$,

$$\mathbf{v}^* \cdot \mathbb{A}\mathbf{v} = \lambda|\mathbf{v}|^2 \qquad \text{and} \qquad \mathbf{v} \cdot \mathbb{A}\mathbf{v}^* = \lambda^*|\mathbf{v}|^2. \tag{B.5}$$

Now, the expression on the left-hand side of the first equation in (B.5) is

$$\begin{aligned} \mathbf{v}^* \cdot \mathbb{A}\mathbf{v} = \sum_{i=1}^{2} v_i^*[\mathbb{A}\mathbf{v}]_i &= \sum_{i,j=1}^{2} v_i^* a_{ij} v_j \\ &= \sum_{i,j=1}^{2} v_i^* a_{ji} v_j \\ &= \sum_{i,j=1}^{2} v_j a_{ji} v_i^* \\ &= \sum_{j=1}^{2} v_j[\mathbb{A}\mathbf{v}^*]_j = \mathbf{v} \cdot \mathbb{A}\mathbf{v}^*, \end{aligned}$$

and so is the same as the expression on the left-hand side of the second equation in (B.5). It follows that

$$\lambda|\mathbf{v}|^2 = \lambda^*|\mathbf{v}|^2,$$

i.e. $\lambda = \lambda^*$ and so this eigenvalue is real.

To see that the eigenvectors corresponding to distinct eigenvalues are orthogonal, suppose that $\mathbb{A}\mathbf{v}^{(1)} = \lambda_1\mathbf{v}^{(1)}$ and $\mathbb{A}\mathbf{v}^{(2)} = \lambda_2\mathbf{v}^{(2)}$ with $\lambda_1 \neq \lambda_2$. Then

$$\mathbf{v}^{(1)} \cdot \mathbb{A}\mathbf{v}^{(2)} = \mathbf{v}^{(1)} \cdot \lambda_2\mathbf{v}^{(2)} = \lambda_2(\mathbf{v}^{(1)} \cdot \mathbf{v}^{(2)}).$$

Looking at the left-hand side of this we have

$$\begin{aligned}\mathbf{v}^{(1)} \cdot \mathbb{A}\mathbf{v}^{(2)} = \sum_i v_i^{(1)}[\mathbb{A}\mathbf{v}^{(2)}]_i &= \sum_{i,j=1}^{2} v_i^{(1)} a_{ij} v_j^{(2)} \\ &= \sum_{i,j=1}^{2} v_i^{(1)} a_{ji} v_j^{(2)} \\ &= \sum_{i,j=1}^{2} v_j^{(2)} a_{ji} v_i^{(1)} \\ &= \sum_{j=1}^{2} v_j^{(2)}[\mathbb{A}\mathbf{v}^{(1)}]_j = \mathbf{v}^{(2)} \cdot \mathbb{A}\mathbf{v}^{(1)}.\end{aligned}$$

Now,

$$\mathbf{v}^{(2)} \cdot \mathbb{A}\mathbf{v}^{(1)} = \mathbf{v}^{(2)} \cdot \lambda_1\mathbf{v}^{(1)} = \lambda_1(\mathbf{v}^{(1)} \cdot \mathbf{v}^{(2)}),$$

and since $\mathbf{v}^{(1)} \cdot \mathbb{A}\mathbf{v}^{(2)} = \mathbf{v}^{(2)} \cdot \mathbb{A}\mathbf{v}^{(1)}$ we therefore have

$$\lambda_2(\mathbf{v}^{(1)} \cdot \mathbf{v}^{(2)}) = \lambda_1(\mathbf{v}^{(1)} \cdot \mathbf{v}^{(2)}),$$

i.e.

$$(\lambda_2 - \lambda_1)(\mathbf{v}^{(1)} \cdot \mathbf{v}^{(2)}) = 0.$$

Since $\lambda_2 \neq \lambda_1$ we must have $\mathbf{v}^{(1)} \cdot \mathbf{v}^{(2)} = 0$, i.e. the eigenvectors are orthogonal.