Part 4

1 The Eigenvalue Problem

1.1 Introductory

Consider an $n \times n$ real or complex matrix

$$A = [\,\mathbf{c}_1 : \mathbf{c}_2 : \cdots : \mathbf{c}_n]$$

and its action on a vector

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

We have

$$A\mathbf{x} = [\,\mathbf{c}_1 : \mathbf{c}_2 : \cdots : \mathbf{c}_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + \cdots + x_n\mathbf{c}_n$$

So, in particular,

$$A \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \mathbf{c}_1, \qquad A \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{bmatrix} = \mathbf{c}_2, \qquad \&c$$

When D is *diagonal*

$$D = \mathrm{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

we have

$$D\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix} = \lambda_1\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix} \qquad \dots \qquad D\begin{bmatrix}0\\0\\\vdots\\1\end{bmatrix} = \lambda_n\begin{bmatrix}0\\0\\\vdots\\1\end{bmatrix}$$

ie

$$D\mathbf{e}_j = \lambda_j\mathbf{e}_j \qquad \text{where } \mathbf{e}_j = [0 \quad \cdots \quad 0 \quad \underbrace{1}_{j^{th} place} \quad 0 \quad \cdots \quad 0]^T$$

Question

In what other circumstances might $A\mathbf{x}$ be proportional to $\mathbf{x}$?

This happens *trivially* when $\mathbf{x} = 0$: so we ask

Given a matrix A can we find a vector $\mathbf{x}$ $(\neq 0)$ and a number λ such that

$$A\mathbf{x} = \lambda\,\mathbf{x} \qquad ?$$

A pair $(\lambda, \mathbf{x})$ such that $A\mathbf{x} = \lambda\,\mathbf{x}$ is called an *eigenpair*: the vector $\mathbf{x}$ is an *eigenvector* of A *corresponding to* the *eigenvalue* λ [or *a* λ-*eigenvector*]. Note that $(\lambda, \mu\mathbf{x})$ will again be an eigenpair for any $\mu \neq 0$.

E1

For $B = \frac{1}{8}\begin{bmatrix}11 & 13\\13 & 11\end{bmatrix}$ we have

$$B\begin{bmatrix}1\\1\end{bmatrix} = 3\begin{bmatrix}1\\1\end{bmatrix} \qquad \& \qquad B\begin{bmatrix}1\\-1\end{bmatrix} = -\frac{1}{4}\begin{bmatrix}1\\-1\end{bmatrix}$$

while for $C = \frac{1}{4}\begin{bmatrix}12 & 0\\13 & -1\end{bmatrix}$ we have

$$C\begin{bmatrix}1\\1\end{bmatrix} = 3\begin{bmatrix}1\\1\end{bmatrix} \qquad \& \qquad C\begin{bmatrix}0\\1\end{bmatrix} = -\frac{1}{4}\begin{bmatrix}0\\1\end{bmatrix}$$

Now the equation $A\mathbf{x} = \lambda\,\mathbf{x}$ above is equivalent to

$$(A - \lambda I)\mathbf{x} = 0$$

and this has a *nontrivial* solution precisely when the matrix

$$A - \lambda I$$

fails to be *invertible*: *ie* when $A - \lambda I$ is *singular*.

We attack the eigenvalue problem by first seeking the eigenvalues [if any]: and for this we need some criterion — such a criterion can be stated in terms of *determinants*.

1.2 Determinants

1.2.1 Introduction

Each *square* matrix A has a *determinant*, a scalar denoted either by $\det(A)$ or $|\,A\,|$.

The determinant has various interpretations and uses. It provides a criterion for the invertibility of a matrix. In particular, it is the tool for finding *eigenvalues* and then *eigenvectors*.

A particular three dimensional geometrical fact is that if three (column) vectors u, v, w specify the sides of a parallelipiped, relative to the origin, and if we form the 3×3 matrix $A = [\mathbf{u}\,\mathbf{v}\,\mathbf{w}]$, then $|\det(A)\,|$ is the measure of the volume of the parallelipiped.

1.2.2 Definition & properties

Determinants are defined *inductively*: those of matrices of size $n \times n$ are defined in terms of determinants of matrices of size $n-1 \times n-1$.

Throughout this section A will be an $n \times n$ matrix over a field $\mathbf{F}$ ($\mathbf{R}$ or $\mathbf{C}$).

For each i, j we define the *minor* M_{ij} of the *entry* a_{ij} to be the determinant of the matrix obtained by *deleting* the i^{th} row and j^{th} column of A: *eg* if

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

then

$$M_{23}(A) = \det\left(\begin{bmatrix} 1 & 2 \\ 7 & 8 \end{bmatrix}\right).$$

Further, we define the *cofactor* A_{ij} of the entry a_{ij} by

$$A_{ij} = (-1)^{i+j} M_{ij}.$$

The *sign pattern* arising from the $(-1)^{i+j}$ is

$$\begin{bmatrix} + & - & + & - & & \dots \\ - & + & - & + & & \dots \\ + & - & + & - & & \dots \\ \vdots & & & \ddots & & \vdots \end{bmatrix}$$

$\mathbf{D}_1$ We now define the determinant of a 1×1 matrix to be its sole entry:

$$\det([a]) = |[a]| = a.$$

NB Do not confuse 'determinant' and 'absolute value'.

$\mathbf{D}_n$ Next we define the determinant of the $n \times n$ matrix A in terms of the first-row entries and their cofactors:

$$\det(A) = |\, A \,| = \sum_{j=1}^{n} a_{1j} A_{1j}.$$

This is called the *cofactor expansion* by the 1^{st} row.

Filling out the inductive steps, we can establish that the determinant of A can be written as the sum of $n!$ terms, each of which is the product of n entries of A, one from each row and one from each column, with a $-$ sign introducing half of these terms: a large number of terms, even when n is only moderate in size. One of these terms is $a_{11}\, a_{22}\, \cdots\, a_{nn}$, the product of the diagonal terms.

Property 1

$$\det(I_n) = 1$$

for any n.

Proof. Induction.

E

$$\begin{aligned} \det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) &= a \times \text{cofactor}_{11} + b \times \text{cofactor}_{12} \\ &= a(-1)^{1+1} \det([d]) + b(-1)^{1+2} \det([c]) \\ &= ad - bc. \end{aligned}$$

Property 2 The cofactor expansions by all other rows give the same result:

$$\det(A) = |\,A\,| = \sum_{j=1}^{n} a_{ij}A_{ij}$$

for *any* i.

Eg for the matrix

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

the first row cofactor expansion gives

$$\begin{aligned} A_{11} &= (-1)^2 \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} = 5 \times 9 - 6 \times 8 = -3 \\ A_{12} &= (-1)^3 \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} = -(4 \times 9 - 6 \times 7) = 6 \\ A_{13} &= (-1)^4 \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} = 4 \times 8 - 5 \times 7 = -3 \end{aligned}$$

Hence

$$\det(A) = 1(-3) + 2(6) + 3(-3) = 0.$$

To compare, the second row cofactor expansion gives

$$\det(A) = (-1)^{2+1}4 \begin{vmatrix} 2 & 3 \\ 8 & 9 \end{vmatrix} + (-1)^{2+2}5 \begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix} + (-1)^{2+3}6 \begin{vmatrix} 1 & 2 \\ 7 & 8 \end{vmatrix} = \ldots = 0.$$

Property 3

$$\det\left(A^{\mathrm{T}}\right) = \det(A)\,.$$

Proof. Induction.

It follows from this last result that all the properties stated in terms of *rows* will also hold for *columns*.

We therefore have *cofactor expansions by columns* as well as rows: for *any* j

$$\det(A) = |\,A\,| = \sum_{i=1}^{n} a_{ij}A_{ij}.$$

Property 4 If any row [or column] of A is the zero vector, then $\det(A) = 0$.

Property 5 If A is upper or lower triangular then

$$\det(A) = \prod_{i=1}^{n} a_{ii}$$

eg

$$\begin{vmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 1 & -1 & 2 \end{vmatrix} = 1 \begin{vmatrix} 3 & 0 \\ -1 & 2 \end{vmatrix} = 1(3 \times 2 - (-1) \times 0) = 6 = 1 \times 3 \times 2.$$

Property 6

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

Property 7 If A is invertible then $|\,A\,| \neq 0$ and

$$\left|\,A^{-1}\,\right| = |\,A\,|^{-1}$$

Property 8 If $A = \begin{bmatrix} R+S \\ R_2 \\ \vdots \\ R_n \end{bmatrix}$ where R, S, $R_2 \ldots R_n$ are row vectors, then

$$\det(A) = \det\left(\begin{bmatrix} R \\ R_2 \\ \vdots \\ R_n \end{bmatrix}\right) + \det\left(\begin{bmatrix} S \\ R_2 \\ \vdots \\ R_n \end{bmatrix}\right);$$

similarly for any row which is the sum of two row vectors: and similarly for columns.

Proof.

$$\begin{aligned} \det(A) &= \sum_{j=1}^{n} a_{1j}A_{1j} \\ &= \sum_{j=1}^{n} (r_j + s_j)A_{1j} \\ &= \sum_{j=1}^{n} r_j A_{1j} + \sum_{j=1}^{n} s_j A_{1j}. \end{aligned}$$

1.2.3 Determinants of elementary matrices

Property 9 $\det\left(E_{R_i\longleftrightarrow R_j}\right) = -1$.

Equivalently, if any pair of rows in A is interchanged then the determinant is multiplied by -1 *ie* the determinant *changes sign.*

Property 9′ If any two rows of A are equal then $|\,A\,| = \left|\,E_{R_i\longleftrightarrow R_j}A\,\right| = -\,|\,A\,|$: so $|\,A\,| = 0$.

Property 10 $\det\left(E_{R_i\to\lambda R_i}\right) = \lambda$.

Equivalently, if any single row of A is multiplied by a constant λ then the determinant is multiplied by λ. So

$$\det(\lambda A) = \lambda^n \det(A)$$

Property 11 $\det\left(E_{R_i\to R_i+\lambda R_j}\right) = 1$.

Adding a multiple of one row to another does not change the value of the determinant.

Practical evaluation

Almost *always* by row & column operations, almost *never* by direct cofactor expansions.

E

$$\begin{vmatrix} 1 & 1 & 2 & 6 \\ 2 & -1 & 3 & 1 \\ 1 & 1 & 2 & -3 \\ 0 & 2 & 4 & -2 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 2 & 6 \\ 2 & -1 & 3 & 1 \\ 0 & 0 & 0 & -9 \\ 0 & 2 & 4 & -2 \end{vmatrix} \quad R_3 \to R_3 - R_1$$

$$= 9\begin{vmatrix} 1 & 1 & 2 \\ 2 & -1 & 3 \\ 0 & 2 & 4 \end{vmatrix} = 9\times 2\begin{vmatrix} 1 & 1 & 2 \\ 2 & -1 & 3 \\ 0 & 1 & 2 \end{vmatrix}$$

$$= 18\begin{vmatrix} 1 & 0 & 0 \\ 2 & -1 & 3 \\ 0 & 1 & 2 \end{vmatrix} = 18\begin{vmatrix} -1 & 3 \\ 1 & 2 \end{vmatrix} = 18[-2-3] = -90$$

1.2.4 Determinant as invertibility criterion

We have seen that all *elementary matrices* have *nonzero* determinants.

Suppose that A and B are *row-similar ie* $A \sim B$. Then $A = QB$, where Q is a product of elementary matrices, so is invertible and has nonzero determinant. Now $\det(A) = \det(Q)\det(B)$; so $\det(A) = 0$ if and only if $\det(B) = 0$.

Now A is invertible if and only if A is row-similar to the identity matrix: and singular if and only if it is row-similar to a matrix having at least one zero row.

Since A is invertible precisely when its transpose is invertible we see that A is invertible if and only if it is column-similar to the identity matrix.

Theorem. A is invertible if and only if $|A| \neq 0$.

1.2.5 Adjugates & Cramer's Rule

The *adjugate* [sometimes called *adjoint*] $\mathrm{Adj}(A)$ of A is the *transpose of the matrix of cofactors of* A:

$$[\mathrm{Adj}(A)]_{ij} = A_{ji}.$$

Now

$$[A\,\mathrm{Adj}(A)]_{ik} = \sum_j a_{ij}(\mathrm{Adj}(A))_{jk} = \sum_j a_{ij}A_{kj},$$

and this is the k^{th} row cofactor expansion of $\det(A)$ when $i = k$; while when $i \neq k$ this is the cofactor expansion of the determinant of the matrix obtained by replacing the k^{th} row by the i^{th} row, *ie* of a matrix with *two identical rows*, so $= 0$. So

$$A\,\mathrm{Adj}(A) = \mathrm{Adj}(A)\,A = \det(A)\,I_n.$$

Cramer's Rule is that

$$A^{-1} = \frac{1}{|A|}\,\mathrm{Adj}(A)$$

if A is invertible. It is of little practical use for matrices of any appreciable size.

1.3 Eigenvalues and Eigenvectors

1.3.1 Characteristic equation

Recall that λ is an eigenvalue [or *characteristic value* or *proper value*] of A when the equation

$$(A - \lambda I_n)\mathbf{x} = 0$$

has a solution for some *nonzero* vector $\mathbf{x}$. We know that this is possible if and only if the matrix $A - \lambda I_n$ is singular: and a criterion for this is that

$$\det(A - \lambda I_n) = 0.$$

This equation is called the *characteristic equation* for A. Since expanding $\det(A - \lambda I_n)$ gives a polynomial in λ of degree precisely n we know that it has up to n distinct solutions: but bear in mind that the zeros of a real polynomial may all be nonreal complex numbers.

E

$$\det\left(\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} - \lambda I_2\right) = \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1 \ :$$

so this real matrix has purely imaginary eigenvalues.

E1 *revisited*
The *characteristic polynomial* for B is

$$\begin{aligned} \det\left(B - \lambda I_2\right) &= \left|\frac{1}{8}\begin{bmatrix} 11 - 8\lambda & 13 \\ 13 & 11 - 8\lambda \end{bmatrix}\right| \\ &= \frac{1}{8^2}\left\{(11 - 8\lambda)^2 - 13^2\right\} \\ &= \frac{1}{64}(11 - 8\lambda - 13)(11 - 8\lambda + 13) \\ &= (\lambda - 3)\left(\lambda + \frac{1}{4}\right) \end{aligned}$$

which shows us that the eigenvalues of B are indeed $-\dfrac{1}{4}$ and 3.

The *characteristic polynomial* for C is

$$\begin{aligned}
\det(C - \lambda I_2) &= \left| \frac{1}{4} \begin{bmatrix} 12-4\lambda & 0 \\ 13 & -1-4\lambda \end{bmatrix} \right| \\
&= \frac{1}{4^2}(12-4\lambda)(-1-4\lambda) \\
&= (\lambda - 3)\left(\lambda + \frac{1}{4}\right)
\end{aligned}$$

which shows us that the eigenvalues of C are also $-\frac{1}{4}$ and 3.

E Calculate the eigenvalues of the matrix

$$A = \begin{bmatrix} -1 & 2 & -3 \\ 0 & 1 & -1 \\ 1 & -1 & 2 \end{bmatrix}.$$

The *characteristic polynomial* for A is

cofactor expansion by *first column*

$$\begin{aligned}
\det(A - \lambda I) &= \begin{vmatrix} -1-\lambda & 2 & -3 \\ 0 & 1-\lambda & -1 \\ 1 & -1 & 2-\lambda \end{vmatrix} \\
&= (-1-\lambda)[(1-\lambda)(2-\lambda) - 1] + 1[2(-1) + 3(1-\lambda)] \\
&= (-1-\lambda)[\lambda^2 - 3\lambda + 1] + 1 - 3\lambda = -\lambda^3 + 2\lambda^2 - \lambda \\
&= -\lambda(\lambda^2 - 2\lambda + 1) = -\lambda(\lambda - 1)^2.
\end{aligned}$$

We see that A has eigenvalues $\lambda = 0$ and $\lambda = 1$ (twice).

Factorising polynomials is a costly business: we should always look out for common factors. Sometimes we can produce them by suitable row or column operations.

E Find the eigenvalues of the matrix

$$A = \begin{bmatrix} 2 & -1 & 5 \\ 3 & 4 & -1 \\ -2 & -1 & 9 \end{bmatrix}.$$

Here

$$\begin{aligned}
\det(A - \lambda I) &= \begin{vmatrix} 2-\lambda & -1 & 5 \\ 3 & 4-\lambda & -1 \\ -2 & -1 & 9-\lambda \end{vmatrix} \\
&= \begin{vmatrix} 2-\lambda & -1 & 5 \\ 3 & 4-\lambda & -1 \\ -4+\lambda & 0 & 4-\lambda \end{vmatrix} \quad R_3 \to R_3 - R_1 \\
&= (4-\lambda) \begin{vmatrix} 2-\lambda & -1 & 5 \\ 3 & 4-\lambda & -1 \\ -1 & 0 & 1 \end{vmatrix} \\
&= (4-\lambda) \begin{vmatrix} 2-\lambda & -1 & 7-\lambda \\ 3 & 4-\lambda & 2 \\ -1 & 0 & 0 \end{vmatrix} \quad C_3 \to C_3 + C_1 \\
&= (4-\lambda)\,(-1)\,[-2 - (7-\lambda)(4-\lambda)] \\
&= -(4-\lambda)\,[-\lambda^2 + 11\lambda - 30] \\
&= (4-\lambda)\,(5-\lambda)\,(6-\lambda).
\end{aligned}$$

Once we have determined an eigenvalue we can return to the original equation

$$(A - \lambda I_n)\mathbf{x} = 0$$

to find eigenvector(s) corresponding to this eigenvalue.

E Find eigenvectors corresponding to the three eigenvalues of the matrix

$$A = \begin{bmatrix} 2 & -1 & 5 \\ 3 & 4 & -1 \\ -2 & -1 & 9 \end{bmatrix}.$$

We have already established that the eigenvalues are 4, 5 and 6.

For $\lambda = 4$:

$$\begin{aligned} A - 4I &= \begin{bmatrix} -2 & -1 & 5 \\ 3 & 0 & -1 \\ -2 & -1 & 5 \end{bmatrix} \\ &\sim \begin{bmatrix} -2 & -1 & 5 \\ 1 & -1 & 4 \\ 0 & 0 & 0 \end{bmatrix} \begin{matrix} \\ R_2 \to R_2 + R_1 \\ R_3 \to R_3 - R_1 \end{matrix} \\ &\sim \begin{bmatrix} 0 & -3 & 13 \\ 1 & -1 & 4 \\ 0 & 0 & 0 \end{bmatrix} \begin{matrix} R_1 \to R_1 + 2R_2 \\ \\ \\ \end{matrix} \\ &\sim \begin{bmatrix} 3 & 0 & -1 \\ 0 & 3 & -13 \\ 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

If we take x_3 as the free parameter we find that $x_2 = \frac{13}{3}x_3$ and $x_1 = \frac{1}{3}x_3$.

So $[\frac{1}{3}x_3, \frac{13}{3}x_3, x_3]^{\mathrm{T}} = x_3[\frac{1}{3}, \frac{13}{3}, 1]^{\mathrm{T}}$ $[x_3 \neq 0]$ is the general eigenvector corresponding to $\lambda = 4$: each of these is a multiple of $[1, 13, 3]^{\mathrm{T}}$.

The 'simplest' 4-eigenvector is $[1, 13, 3]^{\mathrm{T}}$.

For $\lambda = 5$:

$$\begin{aligned} A - 5I &= \begin{bmatrix} -3 & -1 & 5 \\ 3 & -1 & -1 \\ -2 & -1 & 4 \end{bmatrix} \\ &\sim \begin{bmatrix} -3 & -1 & 5 \\ 0 & -2 & 4 \\ -2 & -1 & 4 \end{bmatrix} \quad R_2 \to R_2 + R_1 \end{aligned}$$

$$
\begin{aligned}
&\sim \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & -2 \\ -2 & -1 & 4 \end{bmatrix} \begin{array}{l} R_1 \to R_1 - R_3 \\ R_2 \to R_2/(-2) \\ \end{array} \\
&\sim \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & -1 & 2 \end{bmatrix} \begin{array}{l} \\ \\ R_3 \to R_3 - 2R_1 \end{array} \\
&\sim \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix} \begin{array}{l} \\ \\ R_3 \to R_3 + R_2 \end{array}
\end{aligned}
$$

If we take x_3 as the free parameter we find that $x_2 = 2x_3$ and $x_1 = x_3$.

So $[x_3, 2x_3, x_3]^{\mathrm{T}} = x_3[1, 2, 1]$ $[x_3 \neq 0]$ is the general eigenvector corresponding to $\lambda = 5$: each of these is a multiple of $[1, 2, 1]^{\mathrm{T}}$.

For $\lambda = 6$:

$$
\begin{aligned}
A - 6I &= \begin{bmatrix} -4 & -1 & 5 \\ 3 & -2 & -1 \\ -2 & -1 & 3 \end{bmatrix} \\
&\sim \begin{bmatrix} -1 & -3 & 4 \\ 3 & -2 & -1 \\ -2 & -1 & 3 \end{bmatrix} \begin{array}{l} R_1 \to R_1 + R_2 \\ \\ \end{array} \\
&\sim \begin{bmatrix} 1 & 3 & -4 \\ 0 & -11 & 11 \\ 0 & 5 & -5 \end{bmatrix} \begin{array}{l} R_1 \to -R_1 \\ R_2 \to R_2 + 3R_1 \\ R_3 \to R_3 - 2R_1 \end{array} \\
&\sim \begin{bmatrix} 1 & 3 & -4 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \\
&\sim \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}
\end{aligned}
$$

If we take x_3 as the free parameter we find that $x_2 = 2x_3$ and $x_1 = x_3$.

So $[1, 1, 1]^{\mathrm{T}}$ is an eigenvector corresponding to the eigenvalue 6.

Note that the *row sums* of A are all equal to 6 — SIGNIFICANCE ?

1.3.2 Sums & products of eigenvalues

Given a square matrix A we define its *trace* by

$$\operatorname{tr} A = \sum_{j=1}^{n} a_{jj}$$

ie the *trace is the sum of the diagonal entries.*

Consider the 2×2 case, where $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$:

$$\begin{aligned} \det(A - \lambda I) = \begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} &= \lambda^2 - (a+d)\lambda + ad - bc \\ &= \lambda^2 - (\operatorname{tr} A)\,\lambda + \det A \end{aligned}$$

But

$$\begin{aligned} \det(A - \lambda I) &= (\lambda_1 - \lambda)(\lambda_2 - \lambda) \\ &= \lambda^2 - (\lambda_1 + \lambda_2)\lambda + \lambda_1\lambda_2, \end{aligned}$$

where λ_1 and λ_2 are the two eigenvalues of A. From this we see that

$$\begin{aligned} \det(A) &= \prod\{\lambda_j : \lambda_j \textit{ eigenvalues of } A\} \\ \operatorname{tr} A &= \sum\{\lambda_j : \lambda_j \textit{ eigenvalues of } A\}. \end{aligned}$$

These two properties hold for square matrices *of any size*. First,

$$(\lambda_1 - \lambda)(\lambda_2 - \lambda)\cdots(\lambda_n - \lambda) = (-\lambda)^n - \left(\sum \lambda_j\right)(-\lambda)^{n-1} + \cdots + \prod \lambda_j$$

so $\prod \lambda_j = \det(A - 0I) = \det A$.

Second, the only instances of λ^n and λ^{n-1} in the characteristic polynomial occur in the principal diagonal product term, &

$$(a_{11} - \lambda)(a_{22} - \lambda)\cdots(a_{nn} - \lambda) = (-\lambda)^n - (\operatorname{tr} A)(-1)^{n-1}\lambda^{n-1} + \cdots$$

1.3.3 Other properties

Remark. If $(\lambda, \mathbf{x})$ is an eigenpair of A then $(\lambda + \mu, \mathbf{x})$ is an eigenpair of $A + \mu I$ (for any μ).

Theorem. If λ is an eigenvalue of A then λ^k is an eigenvalue of A^k (for any positive integer k).

Proof. If $\mathbf{x} \neq 0$ is such that
$$A\mathbf{x} = \lambda\mathbf{x}$$
then
$$A^2\mathbf{x} = A(A\mathbf{x}) = A(\lambda\mathbf{x}) = \lambda A\mathbf{x} = \lambda\lambda\mathbf{x} = \lambda^2\mathbf{x},$$
which proves the result for $k = 2$. And so on ...

Theorem. If A is invertible and λ is an eigenvalue of A then $\lambda \neq 0$ and λ^{-1} is an eigenvalue of A^{-1}.

Proof. If A is invertible then $\prod \lambda_j = \det(A) \neq 0$: so no eigenvalue of A can vanish.

If $A\mathbf{x} = \lambda\mathbf{x}$ for some $x \neq 0$ then $A^{-1}A\mathbf{x} = A^{-1}\lambda\mathbf{x}$ *ie* $\mathbf{x} = \lambda A^{-1}\mathbf{x}$: now divide both sides by λ to get $A^{-1}\mathbf{x} = \lambda^{-1}\mathbf{x}$.

More generally, if λ is an eigenvalue of A then $a_k\lambda^k + a_{k-1}\lambda^{k-1} + \ldots + a_1\lambda + a_0$ is an eigenvalue of $a_kA^k + a_{k-1}A^{k-1} + \ldots + a_1A + a_0I$ (for any polynomial of degree k).

1.3.4 Transposes

Since
$$\left| A^{\mathrm{T}} - \lambda I \right| = \left| (A - \lambda I)^{\mathrm{T}} \right| = | A - \lambda I |$$
we see that A and A^{T} have the *same* eigenvalues. This is not to say that they have the same eigenvectors.

1.4 Linear algebra

An expression

$$\alpha_1 \mathbf{u}_1 + \ldots + \alpha_m \mathbf{u}_m$$

is a *linear combination* of the set of vectors $U = \{\mathbf{u}_1, \ldots \mathbf{u}_m\}$ [with coefficients $\alpha_1, \ldots, \alpha_m$, some (or all) of which may be zero].

A linear combination of linear combinations is again a linear combination.

A subset of $\mathbf{F}_n$ is a *linear subspace* if it contains all linear combinations of its elements.

The *span* of a set $U = \{\mathbf{u}_1, \ldots \mathbf{u}_m\}$ of vectors in $\mathbf{F}_n$ is the set of all vectors that can be written as *linear combination*s of the vectors U: *ie* the set

$$\{\alpha_1 \mathbf{u}_1 + \ldots + \alpha_m \mathbf{u}_m : \alpha_j \in \mathbf{F}\}.$$

The span of a set U is a linear subspace and is the smallest linear subspace that contains U.

A set $U = \{\mathbf{u}_1, \ldots \mathbf{u}_m\}$ is *linearly independent* in $\mathbf{F}_n$ if

$$\alpha_1 \mathbf{u}_1 + \ldots + \alpha_m \mathbf{u}_m = 0 \Longrightarrow \alpha_1 = \ldots = \alpha_m = 0 :$$

equivalently, if every vector in the span of $\mathbf{u}_1, \ldots \mathbf{u}_m$ can be written as a linear combination of the $\mathbf{u}_1, \ldots \mathbf{u}_m$ in a *unique* way.

A set $U = \{\mathbf{u}_1, \ldots \mathbf{u}_m\}$ is *linearly dependent* in $\mathbf{F}_n$ if it is not linearly independent: *ie* if there are scalars $\alpha_1, \ldots \alpha_m$ *not all zero* such that $\alpha_1 \mathbf{u}_1 + \ldots + \alpha_m \mathbf{u}_m = 0$.

If U is linearly dependent we can select a proper subset V of U which is linearly independent and has the same span as U.

A *basis* is a linearly independent spanning set.

The *row space* $RS(A)$ of a matrix A is the linear subspace generated by its rows [considered as row vectors]. In particular, $RS(I_n) = \mathbf{F}_n$.

Remark. If $B = EA$ where E is an *ero* then $RS(B) = RS(A)$.

Given a set of row vectors U we can *stack* them to produce an $m \times n$ matrix

$$A_U = \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \vdots \\ \mathbf{u}_m \end{bmatrix}$$

and reduce A_U to reduced echelon form B $(= B_U)$.

If $m > n$ then some row of B must vanish, so some nontrivial linear combination of U must vanish *ie* U is linearly dependent.

If U spans $\mathbf{F}_n$ then B must have at least n nonzero rows, so $m \geq n$.

Putting this all together we get the

Theorem. A subset $U = \{\mathbf{u}_1, \ldots \mathbf{u}_m\}$ of $\mathbf{F}_n$ a basis if it satisfies *any two* of the following conditions, in which case it also satisfies the third.

1. $m = n$
2. U spans $\mathbf{F}_n$
3. U is linearly independent

Transposing ...

Similar results hold for spaces of *column vectors* $\mathbf{F}^n$.

For a square matrix

An $n \times n$ matrix A is *invertible* if and only if the set of *rows* of A is a basis for $\mathbf{F}_n$.

Also, an $n \times n$ matrix A is *invertible* if and only if the set of *columns* of A is a basis for $\mathbf{F}^n$.

Similar matrices

Matrices A and B, square of the same size, are *similar* if there is an invertible matrix S such that $A = SBS^{-1}$.

[*Theorem.* Matrices are similar precisely when they represent the same linear transformation with respect to two bases.]

If A and B are similar, with $A = SBS^{-1}$, then

$$\begin{aligned} A^2 &= (SBS^{-1})(SBS^{-1}) = SB(S^{-1}S)BS^{-1} = SB^2S^{-1} \\ A^3 &= (SBS^{-1})(SB^2S^{-1}) = SB(S^{-1}S)B^2S^{-1} = SB^3S^{-1} \\ &\vdots \\ A^k &= SB^kS^{-1} \\ &\vdots \end{aligned}$$

1.5 Diagonalisation

A matrix is *diagonalisable* if it is similar to a diagonal matrix.

Theorem. If $\lambda_1, \ldots, \lambda_k$ are mutually distinct eigenvalues of an $n \times n$ matrix A then their corresponding eigenvectors $\mathbf{u}_1, \ldots \mathbf{u}_k$ are linearly independent. In particular, if A has n distinct eigenvalues then $\mathbf{u}_1, \ldots \mathbf{u}_n$ is a basis for $\mathbf{F}^n$.

Theorem. If A is *symmetric*, *ie* if $A^{\mathrm{T}} = A$, then A has a set of eigenvectors that is a basis for $\mathbf{F}^n$.

Suppose that the $n \times n$ matrix A has n linearly independent eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ corresponding to the (not necessarily distinct) eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$.

Let

$$S = [\mathbf{x}_1 : \mathbf{x}_2 : \ldots : \mathbf{x}_n]$$

be the matrix whose columns are these $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$. Note that

$$\begin{aligned} AS &= A\,[\mathbf{x}_1 : \mathbf{x}_2 : \ldots : \mathbf{x}_n] \\ &= [A\mathbf{x}_1 : A\mathbf{x}_2 : \ldots : A\mathbf{x}_n] \\ &= [\lambda_1\mathbf{x}_1 : \lambda_2\mathbf{x}_2 : \ldots : \lambda_n\mathbf{x}_n] \\ &= [\mathbf{x}_1 : \mathbf{x}_2 : \ldots : \mathbf{x}_n]\,\mathrm{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\} \\ &= SD \end{aligned}$$

where

$$D = \mathrm{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

Since the eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ are *linearly independent* the matrix S is invertible: *ie* S^{-1} exists: and

$$A = SDS^{-1}.$$

In brief, A is diagonalisable.

Conversely, if $A = SDS^{-1}$, where D is diagonal, then $AS = DS$ and the j^{th} column of S is an eigenvector of A corresponding to the eigenvalue d_{jj}.

NB Not all matrices are diagonalisable. *Eg* consider $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. For this matrix $\det(A - \lambda I) = \begin{vmatrix} -\lambda & 1 \\ 0 & -\lambda \end{vmatrix} = \lambda^2$. The eigenvalues are 0 (twice). *Check* that $(A - 0I)\mathbf{x} = 0$ only for $\mathbf{x}$ of the form $\alpha \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

If A *were* diagonalisable, *ie* if there were matrices D and S, with S invertible, such that $S^{-1}AS = D$, then the diagonal entries of D would be the eigenvalues of A: so here $D = 0$, from which $A = 0$.

E1 *revisited*

With $S = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ [note that $S^T S = SS^T = 2I$] we have

$$BS = S \begin{bmatrix} 3 & 0 \\ 0 & -\frac{1}{4} \end{bmatrix}$$

and thus

$$B = S \begin{bmatrix} 3 & 0 \\ 0 & -\frac{1}{4} \end{bmatrix} S^{-1}$$

where

$$S^{-1} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \frac{1}{2} S$$

while, putting $T = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$, we have

$$CT = T \begin{bmatrix} 3 & 0 \\ 0 & -\frac{1}{4} \end{bmatrix}$$

and thus

$$C = T \begin{bmatrix} 3 & 0 \\ 0 & -\frac{1}{4} \end{bmatrix} T^{-1}$$

where

$$T^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$$

Diagonalisation & powers

It is easy to compute the powers of a diagonalisable matrix: if $A = SDS^{-1}$ where

$$D = \mathrm{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix},$$

then

$$A^k = SD^kS^{-1}$$

where

$$D^k = \mathrm{diag}\{\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k\} = \begin{bmatrix} \lambda_1^k & & & \\ & \lambda_2^k & & \\ & & \ddots & \\ & & & \lambda_n^k \end{bmatrix}.$$