mpc2w7.tex
Week 7. 23.11.2011

Eigenvalues and Eigenvectors.
For a square matrix A, if

$$ax = \lambda x$$

for a *non-zero* vector x and scalar λ, x is called an *eigenvector*, with *eigenvalue* λ (eigen = proper, German; other terms are proper value/vector, characteristic value/vector). Then

$$Ax - \lambda x = Ax - \lambda Ix = (A - \lambda I)x = 0.$$

But $x \neq 0$. Since $x = 0$ is trivially also a solution (which is why we require x *non-zero* for an eigenvector), the solution is non-unique. So the matrix $A - \lambda I$ is singular, i.e. has determinant 0:

$$|A - \lambda I| = 0.$$

This is called the *characteristic equation* (or eigenequation) of A. It is a polynomial equation of degree n, the order of A ($n \times n$), in λ. So by the Fundamental Theorem of Algebra (Ch. II), there are n roots $\lambda_1, \ldots, \lambda_n$, possibly complex, counted according to multiplicity.

The characteristic equation $|A - \lambda I| = 0$ is

$$\begin{vmatrix} \lambda - a_{11} & a_{12} & & a_{1n} \\ a_{21} & \lambda - a_{22} & & \\ \vdots & & \ddots & \\ a_{n1} & & & \lambda - a_{nn} \end{vmatrix} = 0,$$

or

$$\lambda^n - (a_{11} + \ldots + a_{nn})\lambda^{n-1} + \ldots + (-)^n |A| = 0$$

(for the constant term, put $\lambda = 0$ to get $|A|$). Write

$$tr\ A := a_{11} + \ldots + a_{nn}$$

for the *trace* of A, the sum of the diagonal elements. Thus the ch. equation is

$$\lambda^n - trA.\lambda^{n-1} + \ldots + (-)^n detA = 0.$$

If all the eigenvalues are *distinct*, one can show that their eigenvectors are *linearly independent* (no sum of multiples can vanish unless all coefficients vanish), meaning that the matrix $P = [p_1, \dots, p_n]$ they form is non-singular. If the eigenvalues are not distinct, we may not be able to find a full set of linearly independent eigenvectors, and then P is not similar to a diagonal matrix.
Example.

$$A = \begin{pmatrix} 4 & -2 \\ 3 & -1 \end{pmatrix}; \quad \text{ch. equation} \begin{vmatrix} 4-\lambda & -2 \\ 3 & -1-\lambda \end{vmatrix} = 0,$$

$(\lambda-4)(\lambda+1) = 0, \qquad \lambda^2-3\lambda+2 = 0, \qquad (\lambda-2)(\lambda-1) = 0, \qquad \lambda = 1 \text{ or } 2.$

$\lambda = 1$: $Ax = x$,

$$\begin{pmatrix} 4 & -2 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

This gives two equations, both

$$3x_1 - 2x_2 = 0$$

(check!). So $x_1 = 2x_2/3$. We can take $x_2 = 3$ (though any other non-zero choice is possible), and then $x_1 = 2$. So the e-vector for the e-value $\lambda = 1$ is

$$x = \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

Similarly, $\lambda = 2$ gives two equations, both (check!)

$$2x_1 - 2x_2 = 0.$$

We can take $x_1 = 1$, and then $x_2 = 1$. So the e-vector for the e-value $\lambda = 2$ is

$$x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

So

$$P = \begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix};$$

$|P| = 1$,

$$P^{-1} = -\begin{pmatrix} 1 & -1 \\ -3 & 2 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 3 & -2 \end{pmatrix}$$

(check). So

$$P^{-1}AP = \begin{pmatrix} -1 & 1 \\ 3 & -2 \end{pmatrix} \begin{pmatrix} 4 & -2 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 3 & -2 \end{pmatrix} \begin{pmatrix} 2 & 2 \\ 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} = D,$$

the diagonal matrix of eigenvalues.

Symmetry.

We quote: if A is *symmetric* (i.e., $A = A^T$ – symmetry about the diagonal, interchanging rows and columns has no effect)

(i) the eigenvalues λ_i are real;

(ii) the eigenvectors p_i are orthogonal (i.e., if $P := [p_1, \ldots, p_n]$, $P^{-1} = P^T$).

The Cayley-Hamilton Theorem.

A *matrix polynomial* is the result of replacing the coefficients in a polynomial $p(x)$ by matrices. Each such matrix polynomial is of the form

$$B(\lambda) = B_0 + B_1\lambda + \ldots + B_r\lambda^r,$$

where each B_k is a matrix of constants and $B_r = 0$ (unless $B(\lambda) \equiv 0$).

Lemma.

If $B(\lambda)$ is a matrix polynomial and $C = B(\lambda)(A - \lambda I)$ is a constant matrix, $C = 0$.

Proof.

Expanding $B(\lambda)(A - \lambda I)$ gives highest term $-\lambda^{r+1}B_r$. There is nothing to cancel this if $B_r \neq 0$, so $B(\lambda) \equiv 0$, so $C = 0$. //

Theorem (Cayley-Hamilton Theorem). A matrix satisfies its own characteristic equation. That is, if the ch. equation is

$$f(\lambda) := |A - \lambda I| = b_0 + b_1\lambda + \ldots + b_{n-1}\lambda^{n-1} + (-)^n\lambda^n = 0,$$

then

$$f(A) := b_0 + b_1A + \ldots + b_{n-1}A^{n-1} + (-)^nA^n = 0.$$

Proof. In $|A - \lambda I| = 0$, the elements are polynomials in λ. So the minors are also polynomials in λ. The elements of the adjoint $C(\lambda) := adj(A - \lambda I)$ are such minors, so

$$C(\lambda) = C_0 + C_1\lambda + \ldots + C_{n-1}\lambda^{n-1},$$

say. But by definition of the adjoint,

$$C(\lambda)(A - \lambda I) = adj(A - \lambda I)(A - \lambda I) = |A - \lambda I|I = f(\lambda)I. \qquad (1)$$

Now

$$A^i - \lambda^i I = (A - \lambda I)(A^{i-1} + \lambda A^{i-2} + \ldots + \lambda^{i-1} I)$$

(check by multiplying out the RHS and cancelling terms in pairs). So

$$f(A) - f(\lambda)I = \sum_0^n b_i A^i - \sum_0^n b_i \lambda^i I = \sum_0^n b_i (A^i - \lambda^i I)$$

$$= \sum_0^n b_i (A - \lambda I)(A^{i-1} + \ldots + \lambda^{i-1} I) = D(\lambda)(A - \lambda I), \qquad (2)$$

say. By (1) and (2),

$$f(\lambda) = [C(\lambda) + D(\lambda)](A - \lambda I).$$

By the Lemma, $f(A) = 0$ (as the LHS is independent of λ –constant in λ). //

Example (to show that this has practical value, and can save work!). Find A^5, where

$$A = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}.$$

The 'obvious' way is by three matrix multiplications: $A^2 = AA$; $A^4 = A^2 A^2$; $A^5 = AA^4$. Compare this with the following. The characteristic equation is

$$|A - \lambda I| = \begin{vmatrix} 2 - \lambda & 3 \\ 3 & 5 - \lambda \end{vmatrix} = \lambda^2 - 7\lambda + 1 = 0.$$

By the Cayley-Hamilton Theorem,

$$A^2 - 7A + I = 0, \qquad A^2 = 7A - I, \qquad A = 7I - A^{-1},$$

$$A^{-1} = 7I - A = \begin{pmatrix} 5 & -3 \\ -3 & 2 \end{pmatrix}$$

(easy anyway: $|A| = 10 - 9 = 1$, so A^{-1} is the transposed matrix of cofactors, so we can read it off from A by sight – and you should check this). Then

$$A^4 = (A^2)^2 = (7A - I)^2 = 49A^2 - 14A + I = 49(7A - I) - 14A + I = 329A - 48I,$$

(343 - 14 = 326),

$$A^5 = A.A^4 = 329A^2 - 48A = 329(7A - I) - 48A = 2255A - 329I$$

(2303 - 48 = 2255),

$$A^5 = 2255 \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix} - 329 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 4181 & 6765 \\ 6765 & 10946 \end{pmatrix}.$$

V. FOURIER SERIES AND TRANSFORMS

1. FOURIER SERIES

Recall (Ch. II)

$$2\sin mx\cos nx = \sin(m+n)x + \sin(m-n)x \qquad (m, n = 0, 1, 2, \ldots).$$

Integrate from 0 to 2π: if $m \neq n$,

$$2\int_0^{2\pi} \sin mx \cos nx dx = -\frac{1}{(m+n)}[\cos(m+n)x]_0^{2\pi} - \frac{1}{m-n}[\cos(m-n)x]_0^{2\pi}$$

(if $m = n$, the second term is 0, so there is no need to integrate it). So: for m, n integer,

$$\int_0^{2\pi} \sin mx \cos nx dx = 0.$$

Similarly,

$$2\cos mx \cos nx = \cos(m+n)x + \cos(m-n)x.$$

As above, this integrates to 0 by periodicity of sin, unless $m = n$, when $2\cos^2 nx = \cos 2nx + 1$,

$$\int_0^{2\pi} \cos^2 nx dx = \frac{1}{2}\int_0^{2\pi} \cos 2nx dx + \frac{1}{2}.2\pi = \frac{1}{2}[\frac{1}{2n}\sin 2nx]_0^{2\pi} + \pi = \pi.$$

$$2\sin mx \sin nx = \cos(m-n)x - \cos(m+n)x,$$

and similarly

$$\int_0^{2\pi} \sin mx \sin nx dx = 0 \quad (m \neq n), \quad \pi \quad (m = n).$$

Now write, for the *trigonometric functions*,

$$\phi_0(x) = \frac{1}{\sqrt{2\pi}}, \quad \phi_{2n-1}(x) = \frac{\cos nx}{\sqrt{\pi}}, \quad \phi_{2n} = \frac{\sin nx}{\sqrt{\pi}} \quad (n = 1, 2, \ldots).$$

Then

$$\int_0^{2\pi} \phi_m(x)\phi_n(x) dx = \delta_{mn}$$

(writing δ_{mn} for the Kronecker delta – 1 if $m = n$, 0 otherwise).

Similarly, using complex exponentials: if

$$\phi_n(x) = \frac{e^{inx}}{\sqrt{2\pi}} = \frac{\cos nx + i \sin nx}{\sqrt{2\pi}},$$

$$\int_0^{2\pi} \phi_m \overline{\phi_n} = \frac{1}{2\pi} \int_0^{2\pi} e^{i(m-n)} x dx = \delta_{mn}.$$

For functions f, g on $[0, 2\pi]$, integrable (say, continuous, or continuous except for finitely many points), we write

$$(f, g) := \int_0^{2\pi} f(x)\overline{g(x)} dx = \int_0^{2\pi} f\overline{g}$$

(bar = complex conjugate, as in Ch. II), and call this the *inner product* of f and g.

Note. This is the continuous analogue of the inner product or dot product between vectors.

Take $f = g$: $\sqrt{(f, f)}$ is called the *norm* of f, $\|f\|$.:

$$\|f\|^2 := (f, f) = \int_0^{2\pi} f(x)\overline{f(x)} dx = \int |f|^2.$$

Note that $\|f\| \geq 0$, and > 0 unless $f = 0$ (almost) everywhere.

Note. If f is continuous, $\|f\| = 0$ implies $f \equiv 0$. For general f, it implies that $f = 0$ 'almost everywhere' – at 'most' points. To make this precise needs Measure Theory and the Lebesgue Integral, which is (way) beyond our scope, so we do not pursue this.

Also

$$(af + bg, h) = a(f, h) + b(g, h)$$

(here of course f, g, h are functions and a, b are constants) – $(.,.)$ is *linear* in its first argument. Similarly,

$$(h, af + bg) = \overline{a}(h, f) + \overline{b}(h, g)$$

– $(.,.)$ is *antilinear* in its second argument.

Note. One needs complex values for many applications, e.g. Quantum Mechanics. If everything is real, we do not need complex conjugates, and then antilinear is the same as linear.