MAU 23205 Lecture 6

John Stalker

Trinity College Dublin

24 September 2021

Matrix algebra

We can think of vectors in $\mathbf{R}^n$ as $n \times 1$ matrices and scalars in $\mathbf{R}$ as 1×1 matrices. So it's enough to consider matrices and matrix valued functions. Vectors and vector valued functions are just a special case.

From this point of view, scalars should multiply from the right, e.g. the eigenvalue equation is

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

If you want them to multiply from the right, to get $A\mathbf{v} = \lambda\mathbf{v}$, then you need to identify scalars with scalar multiples of the identity matrix, i.e. interpret λ as

$$\begin{bmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda \end{bmatrix}.$$

Square matrices

For square matrices, say $n \times n$, we can form the trace

$$\operatorname{tr}(A) = \sum_{j=1}^{n} a_{j,j}$$

and determinant

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \prod_{j=1}^{n} a_{j,\sigma(j)}.$$

A is invertible if and only if $\det(A) \neq 0$. There's also the characteristic polynomial

$$p_A(z) = \det(zI - A).$$

Then $p_A(z) = z^n - \operatorname{tr}(A) z^{n-1} + \cdots + (-1)^n \det(A)$.

Polynomials of matrices

Addition and multiplication of polynomials are defined as follows.

$$\sum_{k=0}^{m} \alpha_k z^j + \sum_{k=0}^{n} \beta_k z^j = \sum_{k=0}^{\max(m,n)} (\alpha_k + \beta_k) z^j$$

$$\left(\sum_{k=0}^{m} \alpha_k z^k\right) \left(\sum_{k=0}^{n} \beta_k z^k\right) = \sum_{k=0}^{m+n} \left(\sum_{i+j=k} \alpha_i \beta_j\right) z^k.$$

The commutative, associative and distributive laws hold.
$(p+q)(z) = p(z) + q(z)$ and $(pq)(z) = p(z)q(z) = q(z)p(z)$ as functions.
We can also apply polynomials to matrices. If $p(z) = \sum_{j=0}^{n} \alpha_j z^j$ then $p(A) = \sum_{j=0}^{n} \alpha_j A^j$ Powers are defined inductively by $A^0 = I$ and $A^{j+1} = A^j A$.
Just as for evaluation at scalars, $(p+q)(A) = p(A) + q(A)$ and $(pq)(A) = p(A)q(A) = q(A)p(A)$.

Cayley-Hamilton

The Cayley-Hamilton Theorem says that

$$p_A(A) = 0.$$

Formally, $p_A(A) = \mathbf{det}(AI - A) = \mathbf{det}(0) = 0$. p_A is a monic polynomial which is zero when evaluated at A .
There is a monic polynomial q such that $q(A) = 0$ and q is of smaller degree than any other such polynomial. It might or might not be the characteristic polynomial. If p is any multiple of the minimal polynomial then $p(A) = 0$.
The characteristic polynomial has a unique factorisation into powers of distinct monic irreducible polynomials: $p_A = \prod_{j=1}^{m} r_j^{f_j}$.
The minimal polynomial factors as $q = \prod_{j=1}^{m} r_j^{e_j}$ where $1 \leq e_j \leq f_j$. If all the f's are 1 then so are the e's and the characteristic polynomial is minimal.

Companion matrices

For any monic polynomial $p(z) = \sum_{j=0}^{n} \alpha_j z^j$ the matrix

$$C_p = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -\alpha_0 & -\alpha_1 & -\alpha_2 & \cdots & -\alpha_{n-1} \end{bmatrix}$$

is called the *companion matrix* of p. The characteristic and minimal polynomials of C_p are both p.

Vandermonde determinants

A matrix of the form

$$V = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \lambda_3 & \cdots & \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2 & \cdots & \lambda_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \lambda_3^{n-1} & \cdots & \lambda_n^{n-1} \end{bmatrix}$$

is called a Vandermonde matrix.

$$\det(V) = \prod_{1 \le j < k \le n} (\lambda_k - \lambda_j).$$

The determinant is non-zero, and hence V is invertible, if and only if the λ's are all distinct.

Transposes

As usual, A^T is defined by $a^T_{i,j} = a_{j,i}$. Transposes reverse order: $(AB)^T = B^T A^T$. A square matrix A is symmetric if $A^T = A$, orthogonal if $A^T A = I$ and normal if $A^T A = AA^T$. You can write quadratic polynomials in terms of symmetric matrices:

$$\sum_{j=1}^{n} \sum_{k=1}^{n} a_{j,k} x_j x_k = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}^T \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

This is true even if A is not symmetric, but not very interesting since you get the same polynomial if you replace A by the symmetric matrix $\frac{1}{2}(A + A^T)$.

More about quadratic polynomials

Note that if you write your polynomials as

$$\sum_{1 \le j \le k \le n} c_{j,k} x_j x_k$$

then the symmetric matrix which represents it is

$$a_{j,k} = \begin{cases} \frac{1}{2} c_{j,k} & \text{if } j < k, \\ c_{j,k} & \text{if } j = k, \\ \frac{1}{2} c_{k,j} & \text{if } j > k. \end{cases}$$

Sylvester's Criterion for the quadratic polynomial to be non-negative everywhere and positive except at 0 is that the *principle minors* $\begin{bmatrix} a_{1,1} \end{bmatrix}$, $\begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}$, $\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{2,1} & a_{3,2} & a_{3,3} \end{bmatrix}$, ... all have positive determinants.