Matrix functions

Let A be a square n-th order matrix. If $P(x) = \alpha_0 + \alpha_1 x + \ldots + \alpha_k x^k$ is a polynomial of degree k, let us define

$$P(A) = \alpha_0 I + \alpha_1 A + \ldots + \alpha_k A^k \tag{1}$$

This definition can, in a natural way, be extended to analytical functions $f(x)$ which can be written as the sum of a power series in an interval $(-R; R)$:

$$f(x) = \alpha_0 + \alpha_1 x + \ldots + \alpha_k x^k + \ldots \quad \left(\alpha_k = \frac{f^{(k)}(0)}{x!}\right)$$

In this case, the value of $f(A)$ is defined by the matrix series

$$f(A) = \alpha_0 + \alpha_1 A + \ldots + \alpha_k A^k + \ldots \tag{2}$$

This definition of $f(A)$, however, is correct only if (2) is convergent. Let us give one of the convergence criteria:

For a series (2) *to be convergent, it is necessary and sufficient for all the eigenvalues of A to lie within the circle of convergence of the series for $f(x)$, i.e. $\mid \lambda_i \mid< R$.*

In particular, if $f(x)$ decomposes into a power series with infinite radius of convergence, then (2) can be used for any matrix A.

The formula (2) is usually used to calculate approximate values of $f(A)$ by replacing the infinite matrix power series by a finite subseries

$$f(A) = \alpha_0 + \alpha_1 A + \ldots + \alpha_k a^k$$

To calculate the exact values of $f(A)$ we use other formulas:

a) If A is a nillpotent matrix, i.e. $A^j = 0$, then also $0 = A^{j+1} = A^{j+2} = \ldots$.

b) If D is a diagonal matrix with the elements $\lambda_1, \lambda_2, \ldots, \lambda_n$ on the main diagonqal, then $f(D)$ is also diagonal with the elements $f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_n)$ on the main diagonal.

$$\begin{pmatrix} f(\lambda_1) & 0 & \cdots & 0 \\ 0 & f(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & f(\lambda_n) \end{pmatrix} \tag{3}$$

c) If we know the value of $f(A)$ and $B = T^{-1}AT$ is a matrix similar to A, then

$$f(B) = T^{-1}f(A)T \tag{4}$$

i.e. $f(B)$ is similar to $f(A)$ with the same transforming matrix T.

d) If all the eigenvalues λ_i of A are mutually different, A can be calculated using the formula

$$f(A) = f(\lambda_1)G_1 + f(\lambda_2)G_2 + \ldots + f(\lambda_n)G_n \tag{5}$$

where the matrices G_i are independent of the choice of f and are defined by the formula

$$G_i = \frac{(A - \lambda_1 I)\ldots(A - \lambda_{i-1}I)(A - \lambda_{i+1}I)\ldots(A - \lambda_n I)}{(\lambda_i - \lambda_1)\ldots(\lambda_i - \lambda_{i-1})(\lambda_i - \lambda_{i+1})\ldots(\lambda_i - \lambda_n)} \tag{6}$$

e) A formula analogous to (6) also exists in the case of multiple eigenvalues of A. Let $\lambda_1, \ldots, \lambda_s$ be the eigenvalues of A with the multiplicities $m_1, \ldots, m_s$ respectively, then

$$f(A) = \sum_{i=1}^{s} \left[f(\lambda_i)G_{i1} + f'(\lambda_i)G_{i2} + \ldots + f^{(m_i-1)}(\lambda_i)G_{im_i} \right] \tag{7}$$

where G_{ij} are matrices independent of the choice of f.

The number of terms in (7) equals the multiplicity of λ_i. For a particular A, however, some of the G_{ij}'s may be zero thus the actual number of terms being less. To calculate the G_{ij}'s a formula similar to (6) could be found. It is, however simpler to use the fact that, in (7), the G_{ij}'s are independent of the choice of f. Selecting several appropriate functions f such as

$$f(x) = 1, f(x) = x - \lambda_1, f(x) = (x - \lambda_1)^2, \ldots$$

and using (7) we get a system od linear matrix equations for the matrices G_{ij}.