LA

Linear Algebra

LA.6 MATRIX FUNCTIONS

Consider a scalar function $f(x) : \mathcal{R} \to \mathcal{R}$ admitting the power expansion

$$f(x) = \sum_{i=0}^{\infty} c_i\, x^i. \qquad \text{(LA.6.1)}$$

Definition LA.6.1 – The matrix function $f(A)$ of the square $(n \times n)$ matrix A is defined as

$$f(A) = \sum_{i=0}^{\infty} c_i\, A^i. \qquad \text{(LA.6.2)}$$

This definition leads to the following properties.

Property LA.6.1 – The multiplication of a square matrix by any of its functions is commutative, i.e.

$$A\, f(A) = f(A)\, A. \qquad \text{(LA.6.3)}$$

Property LA.6.2 – The spectrum of $f(A)$ is given by $f(\lambda_1), \dots, f(\lambda_n)$ where λ_1, $\dots, \lambda_n$ denotes the spectrum of A.

COMPUTATION OF MATRIX FUNCTIONS

A possible way to compute matrix functions relies on the interpolating polynomial algorithm. Consider, to this purpose, the minimal polynomial of A,

$$m(\lambda) = \lambda^{\mu} + a_1\, \lambda^{\mu-1} + \dots + a_{n-1}\, \lambda + a_n; \qquad \text{(LA.6.4)}$$

then for any eigenvalue of A, λ_j, and for any integer k,

$$\lambda_j^{\mu+k} = -a_1\, \lambda_j^{\mu+k-1} - \dots - a_{n-1}\, \lambda_j^{k+1} - a_n\, \lambda_j^{k}. \qquad \text{(LA.6.5)}$$

By substituting (LA.6.5) in (LA.6.1) it is thus possible to eliminate all powers greater or equal to μ and obtain the relation

$$f(\lambda_j) = \sum_{i=0}^{\mu-1} \gamma_i \lambda_j^i \tag{LA.6.6}$$

for suitable values of the coefficients γ_i. The polynomial (LA.6.6) is called an *interpolating polynomial* of $f(x)$ because it assumes the same value as $f(x)$ in correspondence of the μ scalars $\lambda_1, \ldots, \lambda_\mu$. When the zeros of $m(\lambda)$ are distinct it is possible to write the system of equations

$$V\gamma = f \tag{LA.6.7}$$

where V is the Vandermonde matrix

$$V = \begin{bmatrix} 1 & \lambda_1 & \ldots & \lambda_1^{\mu-1} \\ 1 & \lambda_2 & \ldots & \lambda_2^{\mu-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_\mu & \ldots & \lambda_\mu^{\mu-1} \end{bmatrix} \tag{LA.6.8}$$

$$\gamma = \left[\, \gamma_0, \gamma_1, \ldots, \gamma_{\mu-1} \,\right]^T \tag{LA.6.9}$$

$$f = \left[\, f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_\mu) \,\right]^T; \tag{LA.6.10}$$

because of the nonsingularity of V, the vector of coefficients γ is given by

$$\gamma = V^{-1} f. \tag{LA.6.11}$$

When the multiplicity of one or more zeros of $m(\lambda)$ is greater than 1, it is still possible to write μ independent equations observing that any multiple zero of a polynomial annihilates also its derivatives up to its multiplicity minus one. Assuming, for instance, the multiplicity of λ_1 equal to 2, matrix (LA.6.8) will be substituted by

$$V = \begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \ldots & \lambda_1^{\mu-1} \\ 0 & 1 & 2\lambda_1 & \ldots & (\mu-1)\lambda_1^{\mu-2} \\ 1 & \lambda_2 & \lambda_2^2 & \ldots & \lambda_2^{\mu-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_{\mu-1} & \lambda_{\mu-1}^2 & \ldots & \lambda_{\mu-1}^{\mu-1} \end{bmatrix} \tag{LA.6.12}$$

and f (LA.6.10) by

$$f = \left[\, f(\lambda_1), f'(\lambda_1), f(\lambda_2) \ldots, f(\lambda_{\mu-1}) \,\right]^T. \tag{LA.6.13}$$

Since the minimal polynomial of A is annihilated by A, also $f(A)$ can be expressed as

$$f(A) = \sum_{i=0}^{\mu-1} \gamma_i \, A^i . \tag{LA.6.14}$$

Once that the coefficients γ_i have been computed by means of (LA.6.11) it is thus possible to obtain $f(A)$ using relation (LA.6.14).

ALGORITHMS BASED ON SIMILARITY TRANSFORMATIONS

Consider the similarity transformation

$$B = T^{-1} A \, T \tag{LA.6.15}$$

where T is a nonsingular $(n \times n)$ matrix. Since

$$B^k = T^{-1} A^k \, T, \tag{LA.6.16}$$

from definition (LA.6.2) it follows that

$$f(B) = T^{-1} f(A) \, T \tag{LA.6.17}$$

or

$$f(A) = T \; f(B) \, T^{-1}. \tag{LA.6.18}$$

A class of algorithms to compute matrix functions relies on transformations leading to matrices whose structure (triangular, companion, Jordan form) allows an easier computation of $f(B)$; the desired function $f(A)$ is then computed by means of (LA.6.18).

The functions of matrices in the Jordan form can be obtained easily; considering, for instance, the function $f(A) = e^{At}$, for the Jordan matrix

$$J = \begin{bmatrix} J_{11} & 0 & \dots & 0 \\ 0 & J_{12} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_{hr} \end{bmatrix} \tag{LA.6.19}$$

where the generic $(k \times k)$ Jordan block J_{ik} associated with λ_i is

$$J_{ik} = \begin{bmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_i \end{bmatrix}, \tag{LA.6.20}$$

we obtain

$$e^{Jt} = \begin{bmatrix} e^{J_{11}t} & 0 & \dots & 0 \\ 0 & e^{J_{12}t} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{J_{hr}t} \end{bmatrix} \tag{LA.6.21}$$

where the exponentials of the Jordan blocks exhibit the following structure

$$e^{J_{ik}t} = \begin{bmatrix} e^{\lambda_i t} & te^{\lambda_i t} & t^2 e^{\lambda_i t}/2 & \dots \\ 0 & e^{\lambda_i t} & te^{\lambda_i t} & \dots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & e^{\lambda_i t} \end{bmatrix}. \tag{LA.6.22}$$

Simple expressions can be given also for $f(A) = A^k$:

$$J^k = \begin{bmatrix} J_{11}^k & 0 & \dots & 0 \\ 0 & J_{12}^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_{hr}^k \end{bmatrix} \tag{LA.6.23}$$

$$J_{ij}^k = \begin{bmatrix} \lambda_i^k & k\lambda_i^{k-1} & k(k-1)\lambda_i^{k-2} & \dots \\ 0 & \lambda_i^k & k\lambda_i^{k-1} & \dots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_i^k \end{bmatrix}. \tag{LA.6.24}$$