The Matrix Exponential

Gordon K. Smyth

March 17, 1997

One of the most frequently occurring matrix functions is the *matrix exponential*, defined for a square matrix A by

$$e^{At} = \sum_{j=0}^{\infty} \frac{(At)^j}{j!}.$$

The matrix exponential arises from the differential equation

$$\mathbf{x}'(t) = A\mathbf{x}(t), \quad t \geq 0$$

where $\mathbf{x}(t) = \{x_1(t), \ldots, x_n(t)\}^T$ is a vector-valued function of the real argument t . If $\mathbf{x}(t)$ satisfies the initial value condition $\mathbf{x}(0) = \mathbf{v}$, then it can be shown [2] that the solution for $\mathbf{x}(t)$ is

$$\mathbf{x}(t) = e^{At}\mathbf{v}.$$

The linear differential equation above arises in homogeneous continuous time Markov chains, for which $x_j(t)$ is the probability that the Markov chain is in state j at time t, and A is the matrix of instantaneous transition rates. Another application is to compartment models, which are common in pharmacokinetics, where the exchange of materials in biological systems is studied [3]. A system is divided into compartments, and it is assumed that the rates of flow of drugs between compartments follow first order kinetics, so that the rate of transfer to a receiving compartment is proportional to the concentration in the supplying compartment. In this case, $x_j(t)$ is the concentration of the drug in the jth compartment at time t, while the elements of A are the transfer rates between the compartments. See [4] and [1, Chapter 5].

If the matrix A is diagonalizable, then the matrix exponential can be viewed as transforming the eigenvalues while leaving the eigenvectors unchanged. Suppose that

$$A = UDU^{-1}$$

where $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ is a diagonal matrix of eigenvalues, and the columns of U are the corresponding eigenvectors. Then

$$\exp(At) = Ue^{Dt}U^{-1}$$

where e^{Dt} represents the diagonal matrix with diagonal elements $e^{\lambda_1 t}, \ldots, e^{\lambda_n t}$.

Numerous algorithms for computing $\exp(At)$ have been proposed, but most of them are of dubious numerical quality, as pointed out in the survey article by Moler and Van Loan [5]. In Markov chain contexts the actual matrix exponential is not required, only the product of $\exp(At)$ and the initial value $\mathbf{v}$. This is significant, as A may be of large dimension but consist mainly of zeros. The computation of $e^{At}\mathbf{v}$ in the Markov chain context is discussed in detail in [7] and [6]. Software associated with [6] is available from `http://www.maths.uq.oz.au/expokit/`.

References

[1] Bates, D. M., and Watts, D. G. (1988). *Nonlinear regression analysis and its applications.* Wiley, New York.

[2] Coddington, E. A., and Levinson, N. (1955). *The theory of ordinary differential equations.* McGraw-Hill, New York.

[3] Godfrey, K. (1983). *Compartmental models and their application.* Academic Press, New York.

[4] Jennrich, R. I., and Bright, P. B. (1976). Fitting systems of linear differential equations using computer generated exact derivatives. *Technometrics*, **18**, 385–392.

[5] Moler, C. B., and Van Loan, C. F. (1978). Nineteen dubious ways to compute the exponential of a matrix. *SIAM Review*, **20**, 801–836.

[6] Sidje, R. B. (1997). EXPOKIT: software package for computing matrix exponentials. *ACM Transactions of Mathematical Software.* To appear.

[7] Stewart, W. J. (1994). *Introduction to the numerical solution of Markov chains.* Princeton University Press, Princeton, New Jersey.