Journal of Computational Mathematics, Vol.11, No.3, 1993, 225–235.

SQUARE MATRIX PADÉ APPROXIMATION AND CONVERGENCE ACCELERATION OF SEQUENCES*1)

Xu Guo-liang   Zhuang Guo-zhong
(Computing Center, Academic of Sincica, Beijing, China)

Abstract

This paper provides a new method for approximating matrix-valued functions—-square Padé approximation. Some computational methods of the approximants are given. For accelerating matrix sequences, a family of nonlinear extrapolation formulas based on the square Padé approximation is given, a convergence acceleration theorem is proved and numerical examples are presented.

§1. Introduction

For a given formal power series $f(z) = \sum_{i=0}^{\infty} c_i z^i, c_i \in C$, let the polynomials $P_v(z)$ and $Q_u(z)$ be of degree v and u respectively and solve the (v, u) Padé approximation problem:

$$(fQ_u - P_v)(z) = \sum_{i \geq v+u+1} E_i z^i, \quad Q_u(0) = 1,$$

then we call $P_v(z)/Q_u(z)$ the classical Padé approximant. For a formal power series $f(z) = \sum_{i=0}^{\infty} c_i z^i$ with matrix coefficients $c_i \in C^{p\times m}$, we can directly generalize, when $p = m$, the definition of the classical Padé approximation to the matrix case (see [1], [7]). If $c_i \in C^{p\times m}$ is not a square matrix ($p \neq m$), but mu can be divided by p, then we can still define matrix Padé approximant successfully (see [8]). However, in other cases it is difficult to define matrix Padé approximant because of the trouble with matching the number of unknowns to be determined with the number of equations that determine the unknowns. In paper [9], different matrix Padé approximants have been derived. However, these definitions have a defect that all the elements of the numerator (or denominator) don't have the same degree. In this paper, we define a new kind matrix Padé approximant, which can eliminate the above defect. The basic idea is, instead of setting some coefficients of the residual to be zero, we let them satisfy some minimization conditions in Euclid (i.e.,Frobenius) norm. Hence, the approximant derived in this sense is called square matrix Padé approximant. Although other norms can be used, Euclid norm seems to be the most convenient one.

* Received September 6, 1991.
1) The Project Supported by National Science Foundation of China for Youth.

In section 2, we first give the definition of square Padé approximant under two normalization conditions, and then discuss its existence and uniqueness. In section 3, we give some effective computational methods of square matrix Padé approximant by utilizing fully the special structure of the related matrix. In section 4, we derive a set of accelerating formulas for matrix sequences from two different square matrix Padé approximants and obtain an accelerating convergence theorem . In the last section, numerical examples for accelerating the convergence of vector sequences are presented.

§2. Definition, Existence and Uniqueness

Let $F(z) = \sum_{i=0}^{\infty} c_i z^i$, $\quad c_i \in C^{p\times m}$; $\quad H_n^{(p,q)} = \{\sum_{i=0}^{n} a_i z^i : a_i \in C^{p\times q}\}$. If the polynomial $P_v(z) = \sum_{i=0}^{v} A_i z^i \in H_v^{(p,m)}$ and $Q_u(z) = \sum_{i=0}^{u} B_i z^i \in H_u^{(m,m)}$ satisfy

$$F(z)Q_u(z) - P_v(z) = \sum_{i=0}^{\infty} E_i z^i, \quad E_i \in C^{p\times m}$$

and

$$E_i = 0, \qquad i = 0, 1, \cdots, v, \tag{2.1}$$

$$\sum_{i=1}^{k} \|E_{v+i}\|_F = \min, \tag{2.2}$$

$$\sum_{i=1}^{u} \|B_i\|_F = \min, \tag{2.3}$$

under one of the following normalization conditions

$$B_0 = I, \tag{2.4}$$

$$\sum_{i=0}^{u} B_i = I, \tag{2.5}$$

where $\|\cdot\|_F$ denotes the Frobenius norm, for $A = (a_{ij})_{1,1}^{p,q}$, $\|A\|_F = \left(\sum_{i=1}^{p}\sum_{j=1}^{q}|a_{ij}|^2\right)^{1/2}$, we call the approximant $P_v(z)(Q_u(z))^{-1}$ the square matrix Padé approximant of $F(z)$, and denote it by $[v, u, k]_F$.

If $mu = pk$, and the matrix Padé approximant $[v/u]$ (See [8]) exists, then (2.2) becomes $E_{v+i} = 0, i = 1, 2, \cdots, k$. Hence the usual matrix Padé approximant is a special case of our square matrix Padé approximant. If $k < 1$, (2.2) is neglected, (2.3) implies $B_i = 0$ (under the condition (2.4)), thus $Q_u(z) = I$, $P_v(z) = \sum_{i=0}^{v} c_i z^i$. In the following we will always assume $k \geq 1$. The condition (2.3) is proposed so that the solution is unique. From (2.1), we have

$$A_i = \sum_{j=0}^{u} c_{i-j} B_j, \quad i = 0, 1, \cdots, v \tag{2.6}$$