JOURNAL OF APPROXIMATION THEORY 2, 237–240 (1969)

An L_q Approximate Solution of the Riccati Matrix Equation[1]

M. S. HENRY[2] AND F. MAX STEIN

Department of Mathematics, Colorado State University, Fort Collins, Colorado 80521

1. INTRODUCTION

If A, B, C, and D are $n \times n$ matrices whose elements are continuous functions of the real variable x over the interval $[a, a + \alpha_1]$, the *Riccati matrix differential equation* is defined to be

$$R[W(x)] \equiv W'(x) + W(x)A(x) + D(x)W(x) + W(x)B(x)W(x) = C(x). \quad (1)$$

It is assumed that the elements of A, B, C, and D satisfy a Lipschitz condition over the interval $[a, a + \alpha_1]$. Associated with (1) is the initial condition

$$W(a) = W_a, \quad (2)$$

where W_a is an $n \times n$ nonsingular matrix.

Associated with (1) and (2) is the linear $2n \times 2n$ matrix integral equation, the *associated Riccati system*,

$$L[\phi(x)] \equiv \phi(x) - \int_a^x K(t)\,\phi(t)\,dt = Y_a, \quad (3)$$

where

$$K(x) = \begin{bmatrix} A(x) & B(x) \\ C(x) & -D(x) \end{bmatrix} \quad \text{and} \quad Y_a = \begin{bmatrix} W_a^{-1} & 0 \\ I & W_a^{-1} \end{bmatrix}. \quad (4)$$

The system (1) and (2) is known to have a unique solution under the conditions given, and (3) likewise has a unique solution in a region to be specified later, see [1]. We designate this solution by

$$Y(x) = \begin{bmatrix} Y_{11}(x) & Y_{12}(x) \\ Y_{21}(x) & Y_{22}(x) \end{bmatrix}. \quad (5)$$

In this paper we shall be concerned with approximating the solution of the *nonlinear system* (1) and (2) by *best approximating* the solution $Y(x)$ of the

[1] This work was partially supported by AFOSR Grant 185-67.

[2] Now at Montana State University.

linear system (3) with a matrix (13) whose elements are polynomials, in the sense that the L_q norm

$$\left\{\int_a^b \|L[Y(x)] - L[P_k(x)]\|^q \, dx\right\}^{1/q} \tag{6}$$

be a minimum. (The absolute value function used in the integrand in (6) will be defined in the next section.) With such a best approximation, we return to the nonlinear system (1) and (2), obtain an approximate solution of it, and estimate its deviation from the exact solution.

2. Distance Functions

We shall use three types of distance functions connected with matrices. For any $r \times r$ matrix $H(x) = (h_{ij}(x))$ whose elements are continuous functions on $[a,b]$, we define the *absolute value function* to be

$$\|H(x)\| = \sum_{i,j=1}^{r} |h_{ij}(x)|. \tag{7}$$

Using (7), we set

$$\|H(x)\|_m = \max_{a \leqslant x \leqslant b} \|H(x)\|, \tag{8}$$

a *maximum norm* for matrices of this type. We shall define the *integral q norm*, or L_q *norm*, for the matrix $H(x)$, to be

$$\|H(x)\|_i = \left\{\int_a^b \|H(x)\|^q \, dx\right\}^{1/q}. \tag{9}$$

Using the notation of the preceding paragraph, (3) is known to have a unique solution $Y(x)$ in a generalized rectangle

$$\|\phi - Y_a\| \leqslant c, \qquad a \leqslant x \leqslant a + \alpha_2 = b, \qquad \alpha_2 \leqslant \alpha_1. \tag{10}$$

If $Y(x)$ in (5) is such that $Y_{11}(x)$ has an inverse for $a \leqslant x \leqslant b$ (it will, for example, if $c < 1/\|W_a\|$), then it can be shown, see [*4*], that

$$W(x) = Y_{21}(x)\, Y_{11}^{-1}(x) \tag{11}$$

is the unique solution of (1) and (2) in the generalized rectangle

$$\|W - W_a\| \leqslant \|Y_{21}\|_m \|Y_{11}^{-1}\|_m + \|W_a\|, \qquad a \leqslant x \leqslant b. \tag{12}$$

3. The Best Approximation to $L[Y(x)]$

For our approximating functions we choose matrix polynomials of the class

$$P_k(x) = \sum_{i=0}^{k} x^i \sum_{m,j=1}^{2n} c_{mj}^i E_{mj}, \tag{13}$$

where the E_{mj} are $2n \times 2n$ matrices having 1 as the (m,j) element and zeros elsewhere. The coefficients c_{mj}^i of $P_k(x)$ are required to be such that $P_k(a) = Y_a$. Since L in (3) is a linear operator, it can be shown that

$$\inf_{c_{mji}} \|L[Y(x)] - L[P_k(x)]\|_l \tag{14}$$

is attained for some matrix polynomial

$$P_k{}^*(x) = \begin{bmatrix} P_{11}^k(x) & P_{12}^k(x) \\ P_{21}^k(x) & P_{22}^k(x) \end{bmatrix}, \tag{15}$$

for each fixed k. That is, $L[P_k{}^*(x)]$ is the *best L_q approximation* to $L[Y(x)]$, for fixed k.

4. An Approximation to $W(x)$

Let $Q_k(x)$ be an arbitrary $2n \times 2n$ matrix polynomial of degree k, of the type given in (13), such that $Q_k(a) = Y_a$. Then if

$$\| Y(x) - Q_k(x)\| \leqslant \epsilon_k, \tag{16}$$

where ϵ_k is a positive constant independent of x, the following theorem can be proved.

Theorem. *If $W(x)$ is the unique solution of* (1) *and* (2) *in the generalized rectangle* (12), *then*

(i) $\|W(x) - P_{21}^k(x)[P_{11}^k(x)]^{-1}\| \leqslant s_1 k^{2/q} \epsilon_k, \qquad k \geqslant k_0, \qquad q > 1,$

s_1 being a constant, and

(ii) $P_{21}^k(a)[P_{11}^k(a)]^{-1} = W_a, \qquad k \geqslant k_0.$

Proof. First, it can be shown in a manner similar to that used by Oberg [3], that if k is sufficiently large, then $P_k{}^*(x)$, the minimizing matrix polynomial (15), is such that

$$\| Y(x) - P_k{}^*(x)\| \leqslant sk^{2/q} \epsilon_k, \tag{17}$$

where s is a constant independent of k and x, and $q > 1$ is the q given in (6). The result (17) and the assumed Lipschitz conditions on the matrices $A, B, C,$ and D imply that

$$\lim_{k\to\infty} \| Y_{11}(x) - P_{11}^k(x)\| = 0, \tag{18}$$

since $Q_k(x)$ in (16) may be chosen so that ϵ_k is of the order $O(1/k^2)$, see [2]. Therefore for k large enough, $P_{11}^k(x)$ has an inverse.

In the remainder of the discussion we assume that $k \geqslant k_0$, where k_0 is large enough to insure both that (17) holds and also that $P_{11}^k(x)$ has an inverse. Then

$$\begin{aligned} &\| Y_{21} Y_{11}^{-1} - P_{21}^k (P_{11}^k)^{-1} \|_m \\ &= \| Y_{21} Y_{11}^{-1} - P_{21}^k Y_{11}^{-1} + P_{21}^k Y_{11}^{-1} - P_{21}^k (P_{11}^k)^{-1} \|_m \\ &\leqslant \| Y_{11}^{-1} \|_m \| Y_{21} - P_{21}^k \|_m + \| P_{21}^k \|_m \| (P_{11}^k)^{-1} - (Y_{11}^{-1}) \|_m \\ &\leqslant \| Y_{11}^{-1} \|_m \| Y_{21} - P_{21}^k \|_m + 2 \| P_{21}^k \|_m \| Y_{11}^{-1} \|_m{}^2 \| Y_{11} - P_{11}^k \|_m . \end{aligned} \tag{19}$$

The last inequality in (19), which is given without proof, follows from the fact that the linear space of $n \times n$ matrices whose elements are continuous functions over $[a,b]$ is a Banach algebra [5]. Thus by the use of the submatrices of (17) in (19), result (i) of the Theorem is obtained, where s_1 is a constant.

Since we require that $P_k(a) = Y_k$, by examining the $n \times n$ submatrices on each side of the equation, we have for $k \geqslant k_0$ that

$$P_{21}^k(a) \, [P_{11}^k(a)]^{-1} = I(W_a^{-1})^{-1} = W_a,$$

and conclusion (ii) of the Theorem is satisfied.

5. Conclusion

We observe that from the matrix polynomials $P_k{}^*(x)$ which minimize the L_q norm in (14) we obtain a sequence of matrices of *rational functions* which satisfy inequality (i). If we require A, B, C, and D to be $n \times n$ matrices whose elements are merely continuous over the interval $[a, a + \alpha_1]$ instead of satisfying a Lipschitz condition over this interval, then the Theorem is still valid for $q > 2$.

References

1. E. A. Coddington and N. Levinson, "Theory of Ordinary Differential Equations." McGraw-Hill, New York, 1955.
2. D. Jackson, The Theory of Approximation, *Am. Math. Soc. Colloq. Pub.*, Vol. XI, New York, 1930.
3. E. N. Oberg, The approximate solution of integral equations. *Bull. Am. Math. Soc.* **41** (1935), 276–284.
4. W. T. Reid, Riccati matrix differential equatinos and non-oscillatory criteria for associated linear differential systems. *Pacific J. Math.* **13** (1963), 665–685.
5. G. F. Simmons, "Introduction to Topology and Modern Analysis." McGraw-Hill, New York, 1963.