A Galerkin-like scheme to solve Riccati equations encountered in quantum physics

Şuayip Yüzbasi¹ and Murat Karaçayı́r
¹Akdeniz University, Department of Mathematics, Antalya 07058, Turkey
E-mail: syuzbasi@akdeniz.edu.tr

Abstract. In this study, a Galerkin-like approach is applied to numerically solve Riccati differential equations. In this method, inner product is applied to a set of monomials and an expression obtained from the equation in question. The resulting nonlinear system is then solved, yielding a polynomial as the approximate solution. Additionally, the technique of residual correction, whose aim is to increase the accuracy of the approximate solution, is discussed briefly. Lastly, the method and the residual correction technique are illustrated with two examples.

1. Introduction

Many problems of science and engineering utilize differential equations as an important modelling tool. In particular, the Riccati differential equation given by

\[ u'(x) = p(x) + q(x)u(x) + r(x)u^2(x), \quad u(0) = a \]  

plays a significant role in many fields of applied science[1]. To name a few applications of Riccati equations, solitary wave solutions of a partial differential equation are closely related to a projective Riccati equation[2]. The Riccati differential equation also constitute an important tool in present day control theory[3]. Thirdly, applications of Riccati equations can be encountered in Kalman filtering systems such as orbitting satellites[4, 5]. Finally, the application which motivated the writing of this paper is the close relation of a one-dimensional static Schrödinger equation to a Riccati differential equation[6]. Namely, the fundamental solution of the Schrödinger equation given by

\[ i\frac{\partial \psi}{\partial t} + \frac{1}{4}\frac{\partial^2 \psi}{\partial x^2} + tx^2 \psi = 0 \]

can be obtained using the following substitution[7]:

\[ \psi = A(t)e^{iS(x,y,t)} = \frac{1}{\sqrt{2\pi\mu(t)}}e^{i\left(\alpha(t)x^2 + \beta(t)xy + \gamma(t)y^2\right)} \]

Here, the real-valued functions \(\alpha(t), \beta(t)\) and \(\gamma(t)\) satisfy the following system:

\[
\begin{align*}
\alpha'(t) - t + \alpha(t)^2 &= 0, \\
\beta'(t) + \alpha(t)\beta(t) &= 0, \\
\gamma'(t) + \frac{1}{4}\beta(t)^2 &= 0.
\end{align*}
\]
The first equation in this system is a Riccati equation with \( q(t) = 0 \).

Since Riccati differential equations are linked to many important problems of science and engineering, obtaining their solutions is a problem of great importance. As of today, there does not exist an analytical method that has the capability to yield the solution to any given Riccati equation. As a result, numerical techniques or approximate approaches has to be utilized for the purpose of obtaining its solutions. To name a few of such studies, Adomian Decomposition Method was used by El-Tawil et al.[8] and Tsai[9] to solve the Riccati equation approximately. Abbasbandy utilized He’s Variational Iteration Method[10] and Homotopy Perturbation Method[11] to obtain its solutions. Gülsu and Sezer used Taylor matrix method[12] for the same purpose. Yüzbaşı used a collocation method[13] based on Bessel polynomials to obtain approximate solutions of Riccati-type differential equation systems. Finally, Yang et.al. used hybrid functions and Chebyshev polynomials to solve Riccati differential equations.

In this study, our aim is to find an approximate solution to the Riccati differential equation (1) by using a Galerkin-like scheme. The organization of the paper is as follows: The solution method is explained in Section 2. The subject of Section 3 is the technique of residual correction, whose aim is to derive better approximations from already obtained approximate solutions. Section 4 contains numerical examples. Finally, the conclusion of the paper is given in Section 5.

2. Method of Solution

In this section, the procedure we will use to solve Equation (1) is outlined. The same method was used in [14] to obtain approximate solutions of high-order Fredholm integro-differential equations with possibly singular kernel functions.

To begin with, we assume that the unique solution \( u(x) \) of Equation (1) can be expressed as a power series of the form

\[
u(x) = \sum_{k=0}^{\infty} a_k x^k.
\]

We then truncate this power series after the \((N + 1)\)st term so that

\[
u_N(x) = \sum_{k=0}^{N} a_k x^k = \mathbf{X}_N(x) \cdot \mathbf{A}
\]

where

\[
\mathbf{X}_N(x) = \begin{bmatrix} 1 & x & x^2 & \cdots & x^N \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} a_0 & a_1 & a_2 & \cdots & a_N \end{bmatrix}^T.
\]

Here, the purpose of the method is to determine the unknown coefficients \( a_i \). The derivatives of \( u_N(x) \) can be expressed as a product of matrices with the help of a specially defined matrix. Namely, if we define \( \mathbf{B} \) to be the \((N + 1) \times (N + 1)\) matrix such that \( B_{i,i+1} = i \) for \( i = 1, 2, \ldots, N \) and \( B_{i,j} = 0 \) elsewhere, then the following equality holds:

\[
u_N^{(i)}(x) = \mathbf{X}_N(x) \mathbf{B}^i \mathbf{A} = \begin{bmatrix} 1 & x & x^2 & \cdots & x^N \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & N \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix}
\]

Next, substituting the matrix expressions for \( u_N(x) \) and \( u_N^{(i)}(x) \) into Equation (1) gives us the equation

\[
\mathbf{G}(x) \mathbf{A} = p(x)
\]
where
\[ G(x) = (X_N(x)B - q(x)X_N(x) - r(x)X_{2N}(x)\bar{A})A \]
and where
\[ \bar{A} = \begin{bmatrix} a_0 & 0 & 0 & \ldots & 0 \\ a_1 & a_0 & 0 & \ldots & 0 \\ a_2 & a_1 & a_0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_N & a_{N-1} & a_{N-2} & \ldots & a_0 \\ 0 & a_N & a_{N-1} & \ldots & a_1 \\ 0 & 0 & a_N & \ldots & a_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & a_N \end{bmatrix} \]

Now, in order to convert Equation (2) into a system of nonlinear equations in the unknowns \( a_i \), we take inner product of Equation (2) with the elements of the set \( \Phi = \{1, x, x^2, \ldots, x^N\} \), where the inner product of two functions \( f, g \) from the Hilbert space \( L^2[0, b] \) is defined by
\[ \langle f, g \rangle = \int_0^b f(x)g(x)dx. \]

Each inner product will result in a nonlinear equation in the unknown coefficients \( a_i \). Finally, we are left with a nonlinear system \( WA = G \) where the \((N+1) \times (N+1)\) matrix \( W \) and the column matrix \( G \) of length \( N+1 \) are given by
\[ W_{i,j} = \langle x^{i-1}, G(x) \rangle_{1,j}, \quad G_{i,1} = \langle x^{i-1}, p(x) \rangle. \]

More explicitly, \( W \) and \( G \) are given by
\[
W = \begin{bmatrix}
\langle 1, G(x) \rangle_{1,1} & \langle 1, G(x) \rangle_{1,2} & \ldots & \langle 1, G(x) \rangle_{1,N+1} \\
\langle x, G(x) \rangle_{1,1} & \langle x, G(x) \rangle_{1,2} & \ldots & \langle x, G(x) \rangle_{1,N+1} \\
\langle x^2, G(x) \rangle_{1,1} & \langle x^2, G(x) \rangle_{1,2} & \ldots & \langle x^2, G(x) \rangle_{1,N+1} \\
\vdots & \vdots & \ddots & \vdots \\
\langle x^N, G(x) \rangle_{1,1} & \langle x^N, G(x) \rangle_{1,2} & \ldots & \langle x^N, G(x) \rangle_{1,N+1}
\end{bmatrix},
\]
\[
G = \begin{bmatrix}
\langle 1, p(x) \rangle \\
\langle x, p(x) \rangle \\
\langle x^2, p(x) \rangle \\
\vdots \\
\langle x^N, p(x) \rangle
\end{bmatrix}^T.
\]

Note that \( W \) contains terms which are linear in some \( a_i \); therefore the system \( WA = G \) is nonlinear. Solving it under the initial condition \( u_N(0) = a \), which implies \( a_0 = a \), we compute the matrix \( A \) of unknown coefficients and the approximate solution is then given by
\[ u_N(x) = a_0 + a_1x + \ldots + a_Nx^N. \]

3. Error Estimation and Residual Correction

In this section, the error estimation of our method is performed based on the residual function corresponding to Equation (1). Then, starting with an approximate solution of Equation (1), the way of obtaining a better approximation from the original one is described.

Let us consider the residual function
\[ R(x) = u'(x) - p(x) - q(x)u(x) - r(x)u^2(x) \]
of Equation (1). Substituting the approximate solution $u_N(x)$ in place of $u(x)$ we get

$$R_N(x) = u_N'(x) - p(x) - q(x)u_N(x) - r(x)u_N^2(x)$$

(4)

Subtracting Equation (3) from Equation (4) and rearranging yields

$$R_N(x) = -e_N'(x) + q(x)e_N(x) + 2r(x)u_N(x)e_N(x) - r(x)e_N^2(x),$$

(5)

which is similar to Equation (1) with $p(x)$ replaced by $-R_N(x)$, $q(x)$ replaced by $q(x) - 2r(x)u_N(x)$ and $e_N(x) = u_N(x) - u(x)$ is the error function. In addition, since $u(x)$ and $u_N(x)$ both satisfy the initial condition $u(x_0) = a$,

$$e_N(x_0) = 0$$

is the initial condition for Equation (5). Next, the method of Section 2 is applied to solve it for $e_N(x)$ with some choice of the parameter value $M$ and the approximate solution $e_{N,M}(x)$ is obtained. Consequently, this new approximation is used to obtain a corrected solution

$$u_{N,M}(x) = u_N(x) + e_{N,M}(x)$$

of Equation (1). In what follows, $E_{N,M}(x)$ will denote the actual error of $u_{N,M}(x)$ given by $E_{N,M}(x) = u_{N,M}(x) - u(x)$.

4. Numerical Examples

In this section, the method of Section 2 is applied to several examples.

**Example 1.** Our first example is the following equation from [5]:

$$u'(x) = u(x) - 2u^2(x), \quad u(0) = 1.$$  

(6)

Here, the interval of interest is $0 \leq x \leq 0.5$. Equation (6) also has the property of being separable, thus its exact solution can be found to be $u(x) = \frac{e^x}{2e^x + 1}$. We applied the method of Section 2 to this equation with several $N$ values. For example, the approximate solution corresponding to $N = 3$ is obtained as follows:

$$u_3(x) = 1 - 0.9782168543x + 1.1840357357x^2 - 0.7141540751x^3.$$  

Since the exact solution is at hand for this particular problem, we measure the accuracy of the approximate solutions looking at the absolute value of the actual error functions $E_N(x)$. Figure 1 depicts these absolute error functions for $N = 4, 5, 6$. As the figure shows, the absolute error decreases as we increase the parameter $N$. Another means of understanding the effectiveness of a numerical method is comparisons with other methods, therefore such a comparison is given in Table 1. As the table shows, the present scheme gives rise to better results when compared to three other methods.

**Example 2.** Next, we consider the following equation:

$$u'(x) = u(x) - u^2(x), \quad u(0) = 2.$$  

(7)

This time, the problem is considered in the interval $0 \leq x \leq 1$. We applied the present scheme to Equation (7) with the choice of $N = 3, 4, 5, 6, 7$ and we considered the residual functions $R_N(x)$ for these $N$ values. The results are depicted in Table 2. As apparent from the table, the accuracy of the approximate solutions improves with increasing $N$, just like the case in Example 1. We now attempt to improve the approximate solutions using the method of residual correction described in Section 3. For this purpose we choose $u_3(x)$ and $u_6(x)$ as the solutions to be improved. For $N = 3$, let us apply residual correction with $M = 5, 6, 8$ and for $N = 6$, let us apply it with $M = 7, 8, 9$. In order to understand how much improvement is satisfied by residual correction for these parameter values, one can examine Figure 2. As is clear from the figure, residual correction brings about great improvement in the accuracy of the approximate solutions. This is the case both for $N = 3$ and $N = 6$. Furthermore, for fixed $N$, increasing $M$ value yields better and better results.
Table 1. Comparison of the error function with three other methods in Example 1.

| $x$ | Euler method | TMM ($N = 6$) | Tau method | Present ($N = 4$) |
|-----|--------------|---------------|------------|-------------------|
| 0   | 0            | 0             | 0          | 0                 |
| 0.1 | 0.013146     | 2.941E-6      | 5.270E-4   | 3.468E-5          |
| 0.2 | 0.0185470    | 1.487E-4      | 5.282E-4   | 4.794E-5          |
| 0.3 | 0.0204833    | 0.0019719     | 4.001E-4   | 6.858E-5          |
| 0.4 | 0.207262     | 0.0111989     | 3.272E-4   | 9.714E-6          |
| 0.5 | 0.0201354    | 0.0414865     | 2.749E-4   | 1.266E-8          |

Figure 1. Graphics of the absolute errors for $u_N(x)$ for $N = 4, 5, 6$ in Example 1.

Table 2. Absolute residuals for several values of $N$ and at some points in Example 2.

| $x$ | $N = 3$ | $N = 4$ | $N = 5$ | $N = 6$ | $N = 7$ |
|-----|---------|---------|---------|---------|---------|
| 0   | 0       | 0       | 0       | 0       | 0       |
| 0.2 | 0.0014  | 3.1731E-4| 4.5167E-5| 1.0794E-5| 2.9698E-6|
| 0.4 | 0.0071  | 2.4464E-4| 6.0264E-5| 1.8150E-5| 3.7031E-6|
| 0.6 | 0.0034  | 5.3748E-5| 6.9106E-5| 1.7900E-5| 2.8054E-6|
| 0.8 | 0.0054  | 2.3628E-4| 5.6255E-5| 1.4082E-5| 1.8500E-6|
| 1   | 1.3094E-4| 4.7008E-8| 5.7946E-10| 3.2620E-10| 6.0696E-11|

5. Conclusion
In this paper, we have presented a Galerkin-like approach for the approximate solutions of Riccati differential equations. It can be concluded that our approach gives fairly good results despite its simplicity. Another advantage of this approach is that the quality of the approximate solutions improves with increasing $N$ values. Residual error correction technique to improve the accuracy of approximate solutions has also been discussed. Simulation results show that considerable...
improvements in the approximate solutions can be achieved as a result of this technique.

Acknowledgments
The authors are supported by the Scientific Research Project Administration of Akdeniz University.

References
[1] Reid W T 1972 Riccati Differential Equations (New York, Academic Press)
[2] Carinena J F, Marmo G, Perelomov A M and Randa M F 1998 Related operators and exact solutions of Schrödinger equations Int. J. Mod. Phys. A 13 4913-29
[3] Bittanti S, Colaneri P and Guardabassi G 1984 Periodic solutions of periodic Riccati equations IEEE Trans. Autom. Control 29 665-7
[4] Anderson B D O and Moore J B 1979 Optimal Filtering (Englewood Cliffs: Prentice Hall)
[5] Yang C, Hou J and Qin B 2012 Numerical solution of Riccati differential equations by using hybrid functions and tau method International Journal of Mathematical and Computational Sciences 6 216
[6] Lanfear N and Suslov S K 2009 The time-dependent Schrödinger equation, Riccati equation and Airy functions Preprint arXiv:0903.3608
[7] Cordero-Soto R, Lopez R M, Suazo E and Suslov S K 2008 Propagator of a charged particle with a spin in uniform magnetic and perpendicular electric fields Lett. Math. Phys. 84 159-78
[8] El-Tawil M A, Bahnasawi A A and Abdel-Naby A 2004 Solving Riccati differential equation using Adomian’s decomposition method Appl. Math. Comput. 157 503-14
[9] Tsai P Y 2010 An approximate analytic solution of the nonlinear Riccati differential equation J. Franklin Ins. 347 1850-62
[10] Abbasbandy S 2007 A new application of He’s variational iteration method for quadratic Riccati differential equation by using Adomian’s polynomials J. Comput. Appl. Math. 207 59-63
[11] Abbasbandy S 2006 Homotopy perturbation method for quadratic Riccati differential equation and comparison with Adomians decomposition method Appl. Math. Comput. 172 485-90
[12] Gülsu M and Sezer M 2006 On the solution of the Riccati equation by the Taylor matrix method Appl. Math. Comput. 176 414-21
[13] Yüzbaşi Ş 2012 A collocation approach to solve the Riccati-type differential equation systems Int. J. Comput. Math. 89 2180-97
[14] Türkyılmazoğlu M 2014 An effective approach for numerical solutions of high-order Fredholm integro-differential equations Appl. Math. Comput. 227 384-98

Figure 2. Graphics of the absolute residual errors of $u_{NM}(x)$ for several values of $(M,N)$ in Example 2.