Exact solutions of some nonlinear differential equations and integrable discretization

Kenji TOMINAGA* · Norimasa SHIBUTANI*

Abstract:

We obtain a new method to solve the systems of some nonlinear differential equations exactly. Using an integrable discretization of the differential equations, we obtain an iteration, which solves the differential equations numerically. Our iteration converges more rapidly than Euler's method.

Key Words: Exact solution; Integrable diseretization

1. Exact solutions of ordinary differential equations

The inner product between $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{R}^N$ is defined by $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^T \boldsymbol{y}$, where $\boldsymbol{x}^T$ is the transposed vector of $\boldsymbol{x}$. We consider the following nonlinear differential equations:

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{a} + A\boldsymbol{x} + \langle \boldsymbol{b}, \boldsymbol{x} \rangle \boldsymbol{x}, \qquad \boldsymbol{x}(0) = \boldsymbol{x}_0 \tag{1.1}$$

where $\boldsymbol{a}, \boldsymbol{b} \in \boldsymbol{R}^N$ are constant vectors, A is an $N \times N$ real matrix. We introduce a new method to solve the differential equations (1.1) exactly.

Definition 1. *Let A be an $(N+1)\times(N+1)$ real matrix such that*

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1(N+1)} \\ \vdots & \ddots & \vdots \\ a_{(N+1)1} & \cdots & a_{(N+1)(N+1)} \end{pmatrix}.$$

We suppose that $det(A) \neq 0$, $N \geq 1$. Then we define the nonlinear mapping $\varphi_A : \boldsymbol{R}^N \to \boldsymbol{R}^N$ as follows.

For $\boldsymbol{x} = (x_1, \cdots, x_N)^T \in \boldsymbol{R}^N$, we set $\boldsymbol{y} = \varphi_A(\boldsymbol{x}) = (y_1, \cdots, y_N)^T$. The each components of $\boldsymbol{y}$ are defined by

$$y_j \equiv \frac{\sum_{k=1}^N a_{jk} x_k + a_{j(N+1)}}{\sum_{k=1}^N a_{(N+1)k} x_k + a_{(N+1)(N+1)}},$$

for $j = 1, \cdots, N$.

Lemma 1. *Let A, B be $(N+1)\times(N+1)$ real matrices. We suppose that $det(A) \neq 0$, $det\,(B) \neq 0$, $N \geq 1$. Then we have*

$$\varphi_A(\varphi_B(\boldsymbol{x})) = \varphi_{AB}(\boldsymbol{x}).$$

* 教養部

平成14年 9 月25日受理

Proof. Let $A = (a_{ij})$, $B = (b_{ij})$, $\varphi_B(\boldsymbol{x}) = (y_1,\cdots,y_N)^T$, $\varphi_A(\varphi_B(\boldsymbol{x})) = (z_1,\cdots,z_N)^T$. Then

$$z_i = \frac{\sum_{k=1}^N a_{ik}y_k + a_{i(N+1)}}{\sum_{k=1}^N a_{(N+1)k}y_k + a_{(N+1)(N+1)}} = \frac{\sum_{k=1}^N a_{ik}\dfrac{\sum_{j=1}^N b_{kj}x_j + b_{k(N+1)}}{\sum_{j=1}^N b_{(N+1)j}x_j + b_{(N+1)(N+1)}} + a_{i(N+1)}}{\sum_{k=1}^N a_{(N+1)k}\dfrac{\sum_{j=1}^N b_{kj}x_j + b_{k(N+1)}}{\sum_{j=1}^N b_{(N+1)j}x_j + b_{(N+1)(N+1)}} + a_{(N+1)(N+1)}}$$

$$= \frac{\sum_{k=1}^N a_{ik}(\sum_{j=1}^N b_{kj}x_j + b_{k(N+1)}) + a_{i(N+1)}(\sum_{j=1}^N b_{(N+1)j}x_j + b_{(N+1)(N+1)})}{\sum_{k=1}^N a_{(N+1)k}(\sum_{j=1}^N b_{kj}x_j + b_{k(N+1)}) + a_{(N+1)(N+1)}(\sum_{j=1}^N b_{(N+1)j}x_j + b_{(N+1)(N+1)})}$$

$$= \frac{\sum_{j=1}^N \sum_{k=1}^{N+1} a_{ik}b_{kj}x_j + \sum_{k=1}^{N+1} a_{ik}b_{k(N+1)}}{\sum_{j=1}^N \sum_{k=1}^{N+1} a_{(N+1)k}b_{kj}x_j + \sum_{k=1}^{N+1} a_{(N+1)k}b_{k(N+1)}}.$$

This is the i-th element of $\varphi_{AB}(\boldsymbol{x})$. □

We consider the following autonomous differential equations:

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{F}(\boldsymbol{x}), \quad \boldsymbol{x}(0) = \boldsymbol{x}_0, \tag{1.2}$$

where $\boldsymbol{F}: \boldsymbol{R}^N \to \boldsymbol{R}^N$ is a smooth vector field. The solution $\boldsymbol{x}(t) \equiv \boldsymbol{u}(t, \boldsymbol{x}_0)$ has the following property:

$$\boldsymbol{u}(0, \boldsymbol{x}_0) = \boldsymbol{x}_0, \tag{1.3}$$

$$\boldsymbol{u}(t+s, \boldsymbol{x}_0) = \boldsymbol{u}(s, \boldsymbol{u}(t, \boldsymbol{x}_0)). \tag{1.4}$$

Lemma 2. *Let $\boldsymbol{u}(t, \boldsymbol{x})$ be the smooth function on $\boldsymbol{R}\times\boldsymbol{R}^N$ into $\boldsymbol{R}^N$ which has the properties (1.3) and (1.4). If we define*

$$\boldsymbol{F}(\boldsymbol{x}) \equiv \frac{\partial \boldsymbol{u}}{\partial t}(0, \boldsymbol{x}),$$

then $\boldsymbol{x}(t) \equiv \boldsymbol{u}(t, \boldsymbol{x}_0)$ is the solution of (1.2).

Proof.

$$\frac{d\boldsymbol{x}}{dt} = \lim_{s\to 0}\frac{\boldsymbol{u}(t+s, \boldsymbol{x}_0) - \boldsymbol{u}(t, \boldsymbol{x}_0)}{s} = \lim_{s\to 0}\frac{\boldsymbol{u}(s, \boldsymbol{u}(t, \boldsymbol{x}_0)) - \boldsymbol{u}(t, \boldsymbol{x}_0)}{s}$$

$$= \lim_{s\to 0}\frac{\boldsymbol{u}(s, \boldsymbol{x}) - \boldsymbol{u}(0, \boldsymbol{x})}{s} = \frac{\partial \boldsymbol{u}}{\partial t}(0, \boldsymbol{x}) = \boldsymbol{F}(\boldsymbol{x}). \quad \square$$

Theorem 1. *Let A be an $N\times N$ real matrix such that*

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{NN} \end{pmatrix},$$

where we suppose that $N \geq 1$. Let $\boldsymbol{a} = (a_1,\cdots,a_N)^T$, $\boldsymbol{b} = (b,\cdots,b_N)^T$ be the N real constant vectors. We consider the following differential equations:

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{a} + A\boldsymbol{x} + \langle \boldsymbol{b}, \boldsymbol{x}\rangle \boldsymbol{x}, \qquad \boldsymbol{x}(0) = x_0. \tag{1.5}$$

We define the $(N+1)\times(N+1)$ *real matrix* B *by*

$$B \equiv \begin{pmatrix} a_{11} & \cdots & a_{1N} & a_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{N1} & \cdots & a_{NN} & a_N \\ -b_1 & \cdots & -b_N & 0 \end{pmatrix} = \begin{pmatrix} A & \boldsymbol{a} \\ -\boldsymbol{b}^T & 0 \end{pmatrix}.$$

Then

$$\boldsymbol{x}(t) \equiv \varphi_{e^{tB}}(\boldsymbol{x}_0)$$

is the solution of (1.5).

Proof. From Lemma 1,

$$\boldsymbol{x}(t) = \boldsymbol{u}(t,\ \boldsymbol{x}_0) \equiv \varphi_{e^{tB}}(\boldsymbol{x}_0)$$

has the properties (1.3) and (1.4). Then from Lemma 2, we only calculate $\dfrac{\partial \boldsymbol{u}}{\partial t}(0,\ \boldsymbol{x})$. Let

$$e^{tB} \equiv \begin{pmatrix} \alpha_{11}(t) & \cdots & \alpha_{1(N+1)}(t) \\ \vdots & \ddots & \vdots \\ \alpha_{(N+1)1}(t) & \cdots & \alpha_{(N+1)(N+1)}(t) \end{pmatrix}.$$

For $t=0$, the exponential mapping is the identity matrix, so we have

$$\alpha_{ij}(0) = \delta_{ij},\ \text{for } i,\ j=1,\cdots,N+1.$$

The derivative of the exponential mapping at $t=0$ is the matrix B, so we have

$$\alpha'_{ij}(0) = a_{ij},\ \text{for } i,\ j=1,\cdots,\mathrm{N},\quad \alpha'_{i(N+1)}(0) = a_i,\ \text{for } i=1,\cdots,N,$$
$$\alpha'_{(N+1)j}(0) = -b_j,\ \text{for } j=1,\cdots,N,\quad \text{and } \alpha'_{(N+1)(N+1)}(0) = 0.$$

Let

$$(z_1(t),\cdots,z_N(t))^T \equiv \boldsymbol{u}(t,\ \boldsymbol{x}),\ \text{then } z_j(t) = \frac{\sum_{k=1}^{N}\alpha_{jk}(t)x_k+\alpha_{j(N+1)}(t)}{\sum_{k=1}^{N}\alpha_{(N+1)k}(t)x_k+\alpha_{(N+1)(N+1)}(t)}.$$

Since $\alpha_{ij}(0) = \delta_{ij}$, note that $\sum_{k=1}^{N}\alpha_{(N+1)k}(0)x_k+\alpha_{(N+1)(N+1)}(0) = 1$. Then we have

$$\begin{aligned} z'_j(0) &= \sum_{k=1}^{N}\alpha'_{jk}(0)x_k+\alpha'_{j(N+1)}(0)-\Big(\sum_{k=1}^{N}\alpha_{jk}(0)x_k+\alpha_{j(N+1)}(0)\Big)\Big(\sum_{k=1}^{N}\alpha'_{(N+1)k}(0)x_k+\alpha'_{(N+1)(N+1)}(0)\Big) \\ &= \sum_{k=1}^{N}a_{jk}x_k+a_j+x_j\sum_{k=1}^{N}b_kx_k. \end{aligned}$$

Thus we obtain the conclusion. □

Example 1. As the simplest example of our method we solve the following differential equation:

$$\frac{dx}{dt} = 1+x^2, \quad x(0) = x_0. \tag{1.6}$$

Let

$$B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

We have the exponential mapping:

$$e^{tB} = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}.$$

Thus we obtain the solution of (1.6):

$$x(t) = \varphi_{e^{tB}}(x_0) = \frac{x_0 \cos t + \sin t}{-x_0 \sin t + \cos t} = \frac{x_0 + \tan t}{1 - x_0 \tan t}. \quad \square$$

Example 2. We solve the logistic differential equation:

$$\frac{dx}{dt} = ax(1-x), \quad x(0) = x_0. \tag{1.7}$$

where a is a positive constant. Let

$$B = \begin{pmatrix} a & 0 \\ a & 0 \end{pmatrix}.$$

We have the exponential mapping:

$$e^{tB} = \begin{pmatrix} e^{at} & 0 \\ e^{at}-1 & 1 \end{pmatrix}.$$

Thus we obtain the solution of (1.7):

$$x(t) = \varphi_{e^{tB}}(x_0) = \frac{x_0 e^{at}}{x_0(e^{at}-1)+1}. \quad \square$$

2. Gradient system

Let A be an $N \times N$ real symmetric matrix having eigenvalues such that $\lambda_1 > \lambda_2 \geq \cdots \geq \lambda_{N-1} > \lambda_N$. We are interested in an algorithm to calculate the maximal eigenvalue λ_1. The starting point in the derivation of integrable differential equations is to consider the Rayleigh quotient of A,

$$R_A(\boldsymbol{x}) \equiv \frac{\langle \boldsymbol{x}, A\boldsymbol{x} \rangle}{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}.$$

The well-known minimax theorem states that

$$\lambda_1 = \max_{\|\boldsymbol{x}\|=1} R_A(\boldsymbol{x}), \quad \lambda_N = \min_{\|\boldsymbol{x}\|=1} R_A(\boldsymbol{x}),$$

where $\| \boldsymbol{x} \|^2 = \langle \boldsymbol{x}, \boldsymbol{x} \rangle$. One of the simplest strategies for maximizing $R_A(\boldsymbol{x})$ is the method of steepest ascent. At a current point $\boldsymbol{x}$, where $\| \boldsymbol{x} \| = 1$, the function $R_A(\boldsymbol{x})$ increases most rapidly in the direction of the positive gradient: $\nabla R_A(\boldsymbol{x}) = 2A\boldsymbol{x} - 2\langle \boldsymbol{x}, A\boldsymbol{x} \rangle \boldsymbol{x}$, which is restricted to the unit sphere. Thus the maximal eigenvalue λ_1 of A can be calculated through the trajectory of the gradient system:

$$\frac{d\boldsymbol{x}(t)}{dt} = A\boldsymbol{x} - \langle \boldsymbol{x}, A\boldsymbol{x} \rangle \boldsymbol{x}, \quad \boldsymbol{x}(0) = \boldsymbol{x}_0, \quad \| \boldsymbol{x}_0 \| = 1. \tag{2.1}$$

Note that if $\| \boldsymbol{x}_0 \| = 1$, then $\| \boldsymbol{x}(t) \| = 1$. When we discretize (2.1) by Euler's method, there is no conservative quantity. As we see later, the nonlinear equation (2.1) is essentially linear. Therefore, if we discretize the linear equation, then the discretization may have a conservative quantity. This is a fundamental idea of the integrable discretization (cf. [3, 4]).

Let P be an orthogonal matrix which diagonalizes A as $D \equiv P^TAP = \mathrm{diag}\,(\lambda_1, \cdots, \lambda_N)$. Using $\boldsymbol{x} = P\boldsymbol{r}$ and $D = P^TAP$, the equation (2.1) can be transformed into the following equivalent form:

$$\frac{dr_j^2}{dt} = 2\lambda_j r_j^2 - 2r_j^2 \sum_{k=1}^{N} \lambda_k r_k^2, \tag{2.2}$$

where $\boldsymbol{r} = (r_1, \cdots, r_N)^T$. Let $\boldsymbol{y} \equiv (r_1^2, \cdots, r_N^2)^T$, $\boldsymbol{d} \equiv (\lambda_1, \cdots, \lambda_N)^T$. Then the gradient system (2.1) is also equivalent to the following form:

$$\frac{d\boldsymbol{y}(t)}{dt} = 2D\boldsymbol{y} - 2\langle \boldsymbol{d}, \boldsymbol{y} \rangle \boldsymbol{y}, \quad \boldsymbol{y}(0) = \boldsymbol{y}_0, \tag{2.3}$$

where $\boldsymbol{y}_0 = (r_1^2(0), \cdots, r_N^2(0))^T$. We note $\| \boldsymbol{r}(t) \| = 1$.

Theorem 2. *Let A be an $N \times N$ real symmetric matrix. We consider the following nonlinear differential equations*:

$$\frac{d\boldsymbol{x}}{dt} = A\boldsymbol{x} - \langle \boldsymbol{x}, A\boldsymbol{x} \rangle \boldsymbol{x}, \quad \boldsymbol{x}(0) = \boldsymbol{x}_0. \tag{2.4}$$

The solution of (2.4) *is given by*:

$$\boldsymbol{x}(t) = \frac{e^{tA}\boldsymbol{x}_0}{\sqrt{1 - \| \boldsymbol{x}_0 \|^2 + \| e^{tA}\boldsymbol{x}_0 \|^2}}. \tag{2.5}$$

Proof. We solve the equation (2.3). Let

$$B = 2\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_N & 0 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_N & 0 \end{pmatrix}. \tag{2.6}$$

We have the exponential mapping:

$$e^{tB}=\begin{pmatrix} e^{2\lambda_1 t} & 0 & \cdots & 0 & 0 \\ 0 & e^{2\lambda_2 t} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & e^{2\lambda_N t} & 0 \\ e^{2\lambda_1 t}-1 & e^{2\lambda_2 t}-1 & \cdots & e^{2\lambda_N t}-1 & 1 \end{pmatrix}.$$

The solution of (2.3) is $\boldsymbol{y}(t)=\varphi_{e^{tB}}(\boldsymbol{y}_0)$. For each components of $\boldsymbol{y}(t)$, we have

$$y_j(t)=\frac{e^{2\lambda_j t}y_j(0)}{\Sigma_{k=1}^N(e^{2\lambda_k t}-1)y_k(0)+1},$$

where $\boldsymbol{y}_0=(y_1(0),y_2(0),\ \cdots,\ y_N(0))^T$. The denominator is as follows:

$$\begin{aligned} &1-\sum_{k=1}^N y_k(0)+\sum_{k=1}^N e^{2\lambda_k t}y_k(0)=1-\sum_{k=1}^N r_k^2(0)+\sum_{k=1}^N e^{\lambda_k t}r_k(0)e^{\lambda_k t}r_k(0) \\ &=1-\langle \boldsymbol{r}(0),\ \boldsymbol{r}(0)\rangle+\langle e^{tD}\boldsymbol{r}(0),\ e^{tD}\boldsymbol{r}(0)\rangle \\ &=1-\langle P^T\boldsymbol{x}_0,\ P^T\boldsymbol{x}_0\rangle+\langle Pe^{tD}P^T\boldsymbol{x}_0,\ Pe^{tD}P^T\boldsymbol{x}_0\rangle \\ &=1-\|\boldsymbol{x}_0\|^2+\|e^{tA}\boldsymbol{x}_0\|^2. \end{aligned}$$

Thus we have the solution of (2.4):

$$y_j(t)=r_j^2(t)=\frac{(e^{\lambda_j t}r_j(0))^2}{1-\|\boldsymbol{x}_0\|^2+\|e^{tA}\boldsymbol{x}_0\|^2}.$$

Hence, it follows that

$$r_j(t)=\frac{e^{\lambda_j t}r_j(0)}{\sqrt{1-\|\boldsymbol{x}_0\|^2+\|e^{tA}\boldsymbol{x}_0\|^2}},$$

where the sign is uniquely determined by the initial condition. Thus we have

$$\boldsymbol{r}(t)=\frac{e^{tD}\boldsymbol{r}(0)}{\sqrt{1-\|\boldsymbol{x}_0\|^2+\|e^{tA}\boldsymbol{x}_0\|^2}}.$$

Using again $\boldsymbol{r}=P^T\boldsymbol{x}$ and $PDP^T=A$, we obtain the exact solution (2.5). □

Corollary 1. *The solution of* (2.1) *is given by*

$$\boldsymbol{x}(t)=\frac{e^{tA}\boldsymbol{x}_0}{\|e^{tA}\boldsymbol{x}_0\|}. \tag{2.7}$$

Proof. Since $\|\boldsymbol{x}_0\|=1$, it is obvious from Theorem 2. □

3. Integrable discretization

As we showed in previous section, the nonlinear differential equations (2.1) are essentially linear. That is, we only find the exponential mappping e^{tB}, where B is in (2.6). This is the solution of a linear differential equation:

$$\frac{d\boldsymbol{x}}{dt} = B\boldsymbol{x}, \quad \boldsymbol{x}(0) = \boldsymbol{x}_0. \tag{3.1}$$

In this section it will be shown that if we discretize the linear equation (3.1) using the forward Euler method, then the power method with a shift of the origin is derived. If we use the backward Euler method, then the inverse iteration (also called fractional iteration) algorithm is obtained.

Consider the following approximation of (3.1) by the forward Euler method:

$$\boldsymbol{x}(n+1)-\boldsymbol{x}(n) \equiv \epsilon B\boldsymbol{x}(n), \tag{3.2}$$

where ϵ is a stepsize and we denote $\boldsymbol{x}(\epsilon n)$ by $\boldsymbol{x}(n)$. Then the exponential mapping e^{tB} is approximated by

$$e^{\epsilon B} \simeq I+\epsilon B = \begin{pmatrix} 1+2\epsilon\lambda_1 & 0 & \cdots & 0 & 0 \\ 0 & 1+2\epsilon\lambda_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1+2\epsilon\lambda_N & 0 \\ 2\epsilon\lambda_1 & 2\epsilon\lambda_2 & \cdots & 2\epsilon\lambda_N & 1 \end{pmatrix},$$

where I is the identity matrix. It suffices to solve for $\boldsymbol{y}(1)$ only, so we have

$$y_j(1) = \frac{(1+2\epsilon\lambda_j)y_j(0)}{1+\Sigma_{k=1}^N 2\epsilon\lambda_k y_k(0)} = \frac{(1+2\epsilon\lambda_j)y_j(0)}{\Sigma_{k=1}^N (1+2\epsilon\lambda_k)y_k(0)}, \tag{3.3}$$

since $\Sigma_{k=1}^N y_k(0) = 1$. Hence we have

$$y_j(2) = \frac{(1+2\epsilon\lambda_j)y_j(1)}{\Sigma_{k=1}^N (1+2\epsilon\lambda_k)y_k(1)} = \frac{(1+2\epsilon\lambda_j)\dfrac{(1+2\epsilon\lambda_j)y_j(0)}{\Sigma_{k=1}^N(1+2\epsilon\lambda_k)y_k(0)}}{\Sigma_{k=1}^N(1+2\epsilon\lambda_k)\dfrac{1+2\epsilon\lambda_k y_k(0)}{\Sigma_{k=1}^N(1+2\epsilon\lambda_k)y_k(0)}} = \frac{(1+2\epsilon\lambda_j)^2 y_j(0)}{\Sigma_{k=1}^N(1+2\epsilon\lambda_k)^2 y_k(0)}.$$

The denominator is as follows:

$$\sum_{k=1}^N (1+2\epsilon\lambda_k)^2 y_k(0) = \sum_{k=1}^N (1+2\epsilon\lambda_k)^2 r_k^2(0) = \| (I+2\epsilon D)\boldsymbol{r}(0) \|^2.$$

So we have

$$r_j(2) = \frac{(1+2\epsilon\lambda_j)r_j(0)}{\| (I+2\epsilon D)\boldsymbol{r}(0) \|},$$

where the sign is uniquely determined by the initial condition. Using again $\boldsymbol{r} = P^T\boldsymbol{x}$ and $PDP^T = A$, we obtain

$$\boldsymbol{x}(2) = \frac{(I+2\epsilon A)\boldsymbol{x}(0)}{\| (I+2\epsilon A)\boldsymbol{x}(0) \|}.$$

In general we have following iteration

$$\boldsymbol{x}(n+1) = \frac{(I+\epsilon A)\boldsymbol{x}(n)}{\|(I+\epsilon A)\boldsymbol{x}(n)\|}.$$

Thus the power method with a shift of the origin is obtained (cf. [5,7]).

In a similar way, we discretize the equation (3.1) by using the backward Euler method:

$$\boldsymbol{x}(n+1) - \boldsymbol{x}(n) \equiv \epsilon B\boldsymbol{x}(n+1). \tag{3.4}$$

Then the exponential mapping e^{tB} is approximated by

$$\begin{aligned} e^{\epsilon B} &\simeq (I-\epsilon B)^{-1} \\ &= \begin{pmatrix} 1-2\epsilon\lambda_1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1-2\epsilon\lambda_N & 0 \\ -2\epsilon\lambda_1 & \cdots & -2\epsilon\lambda_N & 1 \end{pmatrix}^{-1} = \begin{pmatrix} (1-2\epsilon\lambda_1)^{-1} & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & (1-2\epsilon\lambda_N)^{-1} & 0 \\ 2\epsilon\lambda_1(1-2\epsilon\lambda_1)^{-1} & \cdots & 2\epsilon\lambda_N(1-2\epsilon\lambda_N)^{-1} & 1 \end{pmatrix}. \end{aligned}$$

Then we have

$$y_j(1) = \frac{(1-2\epsilon\lambda_j)^{-1}y_j(0)}{1+\Sigma_{k=1}^N 2\epsilon\lambda_k(1-2\epsilon\lambda_k)^{-1}y_k(0)} = \frac{(1-2\epsilon\lambda_j)^{-1}y_j(0)}{\Sigma_{k=1}^N(1+2\epsilon\lambda_k(1-2\epsilon\lambda_k)^{-1})y_k(0)} = \frac{(1-2\epsilon\lambda_j)^{-1}y_j(0)}{\Sigma_{k=1}^N(1-2\epsilon\lambda_k)^{-1}y_k(0)},$$

since $\Sigma_{k=1}^N y_k(0) = 1$. Hence we have

$$y_j(2) = \frac{(1-2\epsilon\lambda_j)^{-2}y_j(0)}{\Sigma_{k=1}^N(1-2\epsilon\lambda_k)^{-2}y_k(0)}.$$

The denominator is as follows:

$$\sum_{k=1}^N(1-2\epsilon\lambda_k)^{-2}y_k(0) = \sum_{k=1}^N(1-2\epsilon\lambda_k)^{-2}r_k^2(0) = \|(I-2\epsilon D)^{-1}\boldsymbol{r}(0)\|^2.$$

So we have

$$r_j(2) = \frac{(1-2\epsilon\lambda_j)^{-1}r_j(0)}{\|(I-2\epsilon D)^{-1}\boldsymbol{r}(0)\|}.$$

Using again $\boldsymbol{r} = P^T\boldsymbol{x}$ and $PDP^T = A$, we obtain the following iteration

$$\boldsymbol{x}(n+1) = \frac{(I-\epsilon A)^{-1}\boldsymbol{x}(n)}{\|(I-\epsilon A)^{-1}\boldsymbol{x}(n)\|}.$$

Thus the inverse iteration is obtained.

We will discretize the equation (3.1) by the second-order forward Runge-Kutta method. We have

$$\boldsymbol{x}(1) - \boldsymbol{x}(0) = \frac{\boldsymbol{k}_1 + \boldsymbol{k}_2}{2}, \tag{3.5}$$

where $\boldsymbol{F}(x) \equiv B\boldsymbol{x}$, $\boldsymbol{k}_1 \equiv \epsilon\boldsymbol{F}(\boldsymbol{x}(0)) = \epsilon B\boldsymbol{x}(0)$ and $\boldsymbol{k}_2 \equiv \epsilon\boldsymbol{F}(\boldsymbol{x}(0)+\boldsymbol{k}_1) = \epsilon B(\boldsymbol{x}(0)+B\boldsymbol{x}(0))$. Hence we have

$$\boldsymbol{x}(1) = (I+\epsilon B+\frac{1}{2}\epsilon^2 B^2)\boldsymbol{x}(0). \tag{3.6}$$

Since

$$B^2 = 4\begin{pmatrix} \lambda_1^2 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2^2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_N^2 & 0 \\ \lambda_1^2 & \lambda_2^2 & \cdots & \lambda_N^2 & 0 \end{pmatrix},$$

the exponential mapping e^{tB} is approximated by

$$e^{\epsilon B} \simeq I+\epsilon B+\frac{\epsilon^2}{2}B^2 = \begin{pmatrix} 1+2\epsilon\lambda_1+2\epsilon^2\lambda_1^2 & 0 & \cdots & 0 & 0 \\ 0 & 1+2\epsilon\lambda_2+2\epsilon^2\lambda_2^2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1+2\epsilon\lambda_N+2\epsilon^2\lambda_N^2 & 0 \\ 2\epsilon\lambda_1+2\epsilon^2\lambda_1^2 & 2\epsilon\lambda_2+2\epsilon^2\lambda_2^2 & \cdots & 2\epsilon\lambda_N+2\epsilon^2\lambda_N^2 & 1 \end{pmatrix}.$$

Using again $\Sigma_{k=1}^N y_k(0) = 1$, we have

$$y_j(1) = \frac{(1+2\epsilon\lambda_j+2\epsilon^2\lambda_j^2)y_j(0)}{\Sigma_{k=1}^N(1+2\epsilon\lambda_k+2\epsilon^2\lambda_k^2)y_k(0)}.$$

Note that $\Sigma_{k=1}^N y_k(1) = 1$, then

$$y_j(2) = \frac{(1+2\epsilon\lambda_j+2\epsilon^2\lambda_j^2)y_j(1)}{\Sigma_{k=1}^N(1+2\epsilon\lambda_k+2\epsilon^2\lambda_k^2)y_k(1)} = \frac{(1+2\epsilon\lambda_j+2\epsilon^2\lambda_j^2)^2 y_j(0)}{\Sigma_{k=1}^N(1+2\epsilon\lambda_k+2\epsilon^2\lambda_k^2)^2 y_k(0)}.$$

So we have

$$r_j(2) = \frac{(1+2\epsilon\lambda_j+2\epsilon^2\lambda_j^2)r_j(0)}{\|(I+2\epsilon D+2\epsilon^2 D^2)\boldsymbol{r}(0)\|},$$

where the sign is uniquely determined by the initial condition. Using again $\boldsymbol{r} = P^T\boldsymbol{x}$ and $PDP^T = A$, we obtain

$$\boldsymbol{x}(2) = \frac{(I+2\epsilon A+2\epsilon^2 A^2)\boldsymbol{x}(0)}{\|(I+2\epsilon A+2\epsilon^2 A^2)\boldsymbol{x}(0)\|}.$$

In general if we rewrite 2ϵ by ϵ, we have following iteration

$$\boldsymbol{x}(n+1) = \frac{(I+\epsilon A+(\epsilon^2/2)A^2)\boldsymbol{x}(n)}{\|(I+\epsilon A+(\epsilon^2/2)A^2)\boldsymbol{x}(n)\|}, \quad \boldsymbol{x}(0) = \boldsymbol{x}_0.$$

Finally we solve the following differential equations numerically:

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{a}+A\boldsymbol{x}+\langle \boldsymbol{b},\ \boldsymbol{x}\rangle \boldsymbol{x}, \quad \boldsymbol{x}(0) = \boldsymbol{x}_0. \tag{3.7}$$

where A is an $N\times N$ real matrix.

Let

$$B = \begin{pmatrix} A & \boldsymbol{a} \\ -\boldsymbol{b}^T & 0 \end{pmatrix}.$$

We solve the linear differential equations $d\boldsymbol{x}/dt = B\boldsymbol{x}$ by Euler's method, then we have $\boldsymbol{x}(n+1) = (I_1+\epsilon B)\boldsymbol{x}$, where I_1 is the $(N+1)\times(N+1)$ identity matrix. Also it is the approximation of $e^{\epsilon B}$, that is,

$$e^{\epsilon B} \simeq I_1+\epsilon B+\begin{pmatrix} I+\epsilon A & \epsilon \boldsymbol{a} \\ -\epsilon \boldsymbol{b}^T & 1 \end{pmatrix},$$

where I is the $N\times N$ identity matrix. So we have a new algorithm which solves the equation (3.7) numerically;

$$\boldsymbol{x}(n+1) = \frac{(I+\epsilon A)\boldsymbol{x}(n)+\epsilon \boldsymbol{a}}{1-\epsilon\langle \boldsymbol{b},\ \boldsymbol{x}(n)\rangle}, \quad \boldsymbol{x}(0) = \boldsymbol{x}_0.$$

Especially in one dimensional case, it is known as Riccati difference equation (cf. [2]). We consider the following differential equation:

$$\frac{dx}{dt} = a+bx+cx^2, \quad x(0) = x_0.$$

Then our iteration is as follows:

$$x(n+1) = \frac{(1+\epsilon b)x(n)+\epsilon a}{1-\epsilon cx(n)}, \quad x(0) = x_0.$$

Example 3. As the simplest example of our iteration we solve the following differential equation (see Example 1):

$$\frac{dx}{dt} = 1+x^2, \quad x(0) = 1. \tag{3.8}$$

Euler's method yields that:

$$x(n+1) = x(n)+\epsilon(1+x(n)^2).$$

Our method yields that:

$$x(n+1) = \frac{x(n)+\epsilon}{1-\epsilon x(n)}.$$

Table 1 shows the numerical results in which $\epsilon = 0.1$, table 2 shows the results in which $\epsilon = 0.01$. Our method converges more rapidly than Euler's method, especially near the critical point $t = \pi/4 = 0.785398\cdots$, it is evident.