Matrix exponential

Enrico Bertolazzi

Integration lectures for the Course:

Numerical Methods for Dynamical System and Control

Accademic Year 2009/2010

Contents

1 Computing matrix exponential for diagonalizable matrices 3

2 Computing matrix exponential for general square matrices 4

2.1 Using Jordan normal form 4

2.2 Using Cayley–Hamilton theorem 7

2.3 Using numerical integration 10

2.4 Using Pade approximation and squaring 12

The matrix exponential

Consider the Taylor series of exponential

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \cdots + \frac{x^p}{p!} + \cdots$$

given a square matrix $\boldsymbol{A}$ we can define the matrix exponential as follows

$$e^{\boldsymbol{A}} = \sum_{k=0}^{\infty} \frac{1}{k!}\boldsymbol{A}^k = \boldsymbol{I} + \boldsymbol{A} + \frac{1}{2}\boldsymbol{A}^2 + \frac{1}{6}\boldsymbol{A}^3 + \cdots + \frac{1}{p!}\boldsymbol{A}^p + \cdots \qquad (1)$$

The first question is: when the series (1) is convergent? To respond to the question we recall the following facts:

Remark 1 (convergence criterion) here we recall some classical convergence criterion:

Comparison. If $\sum_{k=0}^{\infty} b_k$ is convergent and $|a_k| \leq b_k$ for all $k \geq n_0$ then $\sum_{k=0}^{\infty} a_k$ is absolutely convergent.

d'Alembert's ratio test. Consider the series $\sum_{k=0}^{\infty} a_k$ and the limit

$$L = \lim_{k \to \infty} \frac{|a_{k+1}|}{|a_k|}$$

then

- If the limit L exists and $L < 1$ the series converges absolutely.
- If the limit L exists and $L > 1$ the series diverges.

If the limit does not exist of is equal to 1 the series can be convergent or divergent.

Root test. Consider the series $\sum_{k=0}^{\infty} a_k$ and the limit

$$L = \limsup_{k \to \infty} \sqrt[k]{|a_k|}$$

then

- If $L < 1$ the series converges absolutely.
- If $L > 1$ the series diverges.

If the limit is equal to 1 the series can be convergent or divergent.

Theorem 1 *The series* (1) *is convergent for all square matrix* $\boldsymbol{A} \in \mathbb{R}^{n \times n}$. *Moreover*

$$\left\| e^{\boldsymbol{A}} \right\|_F \leq n e^{\|\boldsymbol{A}\|_F} \tag{2}$$

where

$$\|\boldsymbol{A}\|_F = \sqrt{\sum_{i,j=1}^{n} A_{i,j}^2}$$

is the Frobenius matrix norm.

Proof Consider the series

$$\sum_{k=0}^{\infty} a_k \qquad \text{where} \qquad a_k = \frac{1}{k!}(\boldsymbol{A}^k)_{ij}$$

i.e. a_k is the (i, j) component of the matrix $\frac{1}{k!}\boldsymbol{A}^k$. It is easy to verify that

$$|A_{l,m}| \leq \|\boldsymbol{A}\|_F\,, \qquad \left\|\boldsymbol{A}^k\right\|_F \leq \|\boldsymbol{A}\|_F^k$$

and thus

$$\sum_{k=0}^{\infty} a_k = \sum_{k=0}^{\infty} \frac{1}{k!}(\boldsymbol{A}^k)_{ij} \leq \sum_{k=0}^{\infty} \frac{1}{k!}\left\|\boldsymbol{A}^k\right\|_F \leq \sum_{k=0}^{\infty} \frac{1}{k!}\|\boldsymbol{A}\|_F^k = e^{\|\boldsymbol{A}\|_F}$$

in conclusion the series (1) is convergent for each component and inequality (2) is trivially verified.

1 Computing matrix exponential for diagonalizable matrices

Let be $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ symmetric, then the matrix has a complete set of linear independent eigenvectors $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, ..., $\boldsymbol{v}_n$:

$$\boldsymbol{A}\boldsymbol{v}_k = \lambda_k \boldsymbol{v}_k, \qquad k = 1, 2, \ldots, n.$$

Thus, defining the matrix $\boldsymbol{T} = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n]$ whose columns are the eigenvectors we have

$$\boldsymbol{A}\boldsymbol{T} = [\boldsymbol{A}\boldsymbol{v}_1, \boldsymbol{A}\boldsymbol{v}_2, \ldots, \boldsymbol{A}\boldsymbol{v}_n] = [\lambda_1\boldsymbol{v}_1, \lambda_2\boldsymbol{v}_2, \ldots, \lambda_n\boldsymbol{v}_n] = \boldsymbol{T}\boldsymbol{\Lambda}$$

and thus $\boldsymbol{A} = \boldsymbol{T}\boldsymbol{\Lambda}\boldsymbol{T}^{-1}$ where

$$\boldsymbol{\Lambda} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}.$$

Using $\boldsymbol{A} = \boldsymbol{T}\boldsymbol{\Lambda}\boldsymbol{T}^{-1}$ we can write

$$e^{\boldsymbol{A}} = \sum_{k=0}^{\infty} \frac{1}{k!}\boldsymbol{A}^k = \sum_{k=0}^{\infty} \frac{1}{k!}(\boldsymbol{T}\boldsymbol{\Lambda}\boldsymbol{T}^{-1})^k = \boldsymbol{T}\left(\sum_{k=0}^{\infty} \frac{1}{k!}\boldsymbol{\Lambda}^k\right)\boldsymbol{T}^{-1} = \boldsymbol{T}e^{\boldsymbol{\Lambda}}\boldsymbol{T}^{-1},$$

and hence

$$e^{\boldsymbol{A}} = \boldsymbol{T}\begin{pmatrix} e^{\lambda_1} & & & \\ & e^{\lambda_2} & & \\ & & \ddots & \\ & & & e^{\lambda_n} \end{pmatrix}\boldsymbol{T}^{-1}$$

2 Computing matrix exponential for general square matrices

2.1 Using Jordan normal form

Let be $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ then the matrix exponential can be computed starting from Jordan normal form (or Jordan canonical form):

Theorem 2 (Jordan normal form) *Any square matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is similar to a block diagonal matrix $\boldsymbol{J}$, i.e. $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \boldsymbol{J}$ where*

$$\boldsymbol{J} = \begin{pmatrix} \boldsymbol{J}_1 & & & \\ & \boldsymbol{J}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{J}_m \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{J}_k = \begin{pmatrix} \lambda_k & 1 & & \\ & \lambda_k & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_k \end{pmatrix}$$

The column of $\boldsymbol{T} = [\boldsymbol{t}_{1,1}, \boldsymbol{t}_{1,2}, \ldots, \boldsymbol{t}_{m,n_m}, \boldsymbol{t}_{m,n_m-1}]$ are generalized eigenvectors, i.e.

$$\boldsymbol{A}\boldsymbol{t}_{k,j} = \begin{cases} \lambda_k \boldsymbol{t}_{k,j} & \text{if } j = 1 \\ \lambda_k \boldsymbol{t}_{k,j} + \boldsymbol{t}_{k,j-1} & \text{if } j > 1 \end{cases} \tag{3}$$

Using Jordan normal form $\boldsymbol{A} = \boldsymbol{T}\boldsymbol{J}\boldsymbol{T}^{-1}$ we can write

$$\begin{aligned}
e^{\boldsymbol{A}} &= \sum_{k=0}^{\infty} \frac{1}{k!}\boldsymbol{A}^k = \sum_{k=0}^{\infty} \frac{1}{k!}(\boldsymbol{T}\boldsymbol{\Lambda}\boldsymbol{T}^{-1})^k \\
&= \boldsymbol{T} \begin{pmatrix} \sum_{k=0}^{\infty} \frac{1}{k!}\boldsymbol{J}_1^k & & & \\ & \sum_{k=0}^{\infty} \frac{1}{k!}\boldsymbol{J}_2 & & \\ & & \ddots & \\ & & & \sum_{k=0}^{\infty} \frac{1}{k!}\boldsymbol{J}_m \end{pmatrix} \boldsymbol{T}^{-1} \\
&= \boldsymbol{T} \begin{pmatrix} e^{\boldsymbol{J}_1} & & & \\ & e^{\boldsymbol{J}_2} & & \\ & & \ddots & \\ & & & e^{\boldsymbol{J}_m} \end{pmatrix} \boldsymbol{T}^{-1}
\end{aligned}$$

Thus, the problem is to find the matrix exponential of a Jordan block

$$\begin{aligned}
\boldsymbol{J}_\lambda &= \begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix} = \lambda \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix} \qquad (4) \\
&= \lambda\boldsymbol{I} + \boldsymbol{N}
\end{aligned}$$

The matrix $\boldsymbol{N}$ has the property:

$$\boldsymbol{N}^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ & 0 & \ddots & 1 \\ & & \ddots & 0 \\ & & & 0 \end{pmatrix}$$

and in general $\boldsymbol{N}^k$ as ones on the k-th upper diagonal and is the null matrix if $k \geq n$ the dimension of the matrix. Using (4) we have

$$\begin{aligned}
e^{\boldsymbol{J}_\lambda} &= \sum_{k=0}^{\infty} \frac{1}{k!} \boldsymbol{J}_\lambda^k = \sum_{k=0}^{\infty} \frac{1}{k!} (\lambda \boldsymbol{I} + \boldsymbol{N})^k = \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{j=0}^{k} \binom{k}{j} \lambda^{k-j} \boldsymbol{N}^j \\
&= \sum_{k=0}^{\infty} \sum_{j=0}^{k} \frac{1}{(k-j)!j!} \lambda^{k-j} \boldsymbol{N}^j \\
&= \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} \frac{1}{(k-j)!j!} \lambda^{k-j} \boldsymbol{N}^j \mathbb{1}_{k-j} \qquad \left[\quad \mathbb{1}_i = \begin{cases} 1 & \text{if } i \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad \right] \\
&= \sum_{j=0}^{\infty} \frac{1}{j!} \boldsymbol{N}^j \sum_{k=0}^{\infty} \frac{1}{(k-j)!} \lambda^{k-j} \mathbb{1}_{k-j} \\
&= \sum_{j=0}^{\infty} \frac{1}{j!} \boldsymbol{N}^j \sum_{k=0}^{\infty} \frac{1}{k!} \lambda^k = e^\lambda \sum_{j=0}^{n-1} \frac{1}{j!} \boldsymbol{N}^j
\end{aligned}$$

or explicit

$$\begin{aligned}
e^{\boldsymbol{J}_\lambda} &= e^\lambda \left(\boldsymbol{I} + \frac{1}{1!} \boldsymbol{N} + \frac{1}{2!} \boldsymbol{N}^2 + \cdots + \frac{1}{(n-1)!} \boldsymbol{N}^{n-1} \right), \\
&= e^\lambda \begin{pmatrix} 1 & 1/1! & & 1/(n-1)! \\ & 1 & \ddots & \\ & & \ddots & 1/1! \\ & & & 1 \end{pmatrix}
\end{aligned}$$

2.2 Using Cayley–Hamilton theorem

Theorem 3 (Cayley–Hamilton) *Let $\boldsymbol{A}$ a square matrix and $\Delta(\lambda) = |\boldsymbol{A} - \lambda \boldsymbol{I}|$ its characteristic polynomial then $\Delta(\boldsymbol{A}) = \boldsymbol{0}$.*

Consider a $n \times n$ square matrix $\boldsymbol{A}$ and a polynomial $p(x)$ and $\Delta(x)$ be the characteristic polynomial of $\boldsymbol{A}$. Then write $p(x)$ in the form

$$p(x) = \Delta(x)q(x) + r(x),$$

where $q(x)$ is found by long division, and the remainder polynomial $r(x)$ is of degree less than n. Now consider the corresponding matrix polynomial $p(\boldsymbol{A})$:

$$p(\boldsymbol{A}) = q(\boldsymbol{A})\Delta(\boldsymbol{A}) + r(\boldsymbol{A}),$$

But Cayley-Hamilton states that $\Delta(\boldsymbol{A}) = \boldsymbol{0}$, therefore $p(\boldsymbol{A}) = r(\boldsymbol{A})$. In general we can deduce that

$$\frac{1}{k!}\boldsymbol{A}^k = r_k(\boldsymbol{A}),$$

where $r_k(x)$ is the remainder of long division of $x^k/k!$ by $\Delta(x)$, i.e. $x^k/k! = \Delta(x)q_k(x) + r_k(x)$ and thus the matrix exponential can be formally written as

$$e^{\boldsymbol{A}} = \sum_{k=0}^{\infty} \frac{1}{k!}\boldsymbol{A}^k = \sum_{k=0}^{\infty} r_k(\boldsymbol{A}),$$

and thus $e^{\boldsymbol{A}}$ is a polynomial of $\boldsymbol{A}$ of degree less than n, i.e.

$$e^{\boldsymbol{A}} = \sum_{k=0}^{n-1} a_k \boldsymbol{A}^k,$$

Consider now an eigenvector $\boldsymbol{v}$ with the corresponding eigenvalue λ, then

$$e^{\boldsymbol{A}}\boldsymbol{v} = \sum_{k=0}^{\infty} \frac{1}{k!}\boldsymbol{A}^k\boldsymbol{v} = \sum_{k=0}^{\infty} \frac{1}{k!}\lambda^k\boldsymbol{v} = e^{\lambda}\boldsymbol{v}$$

analogously

$$\sum_{k=0}^{n-1} a_k \boldsymbol{A}^k \boldsymbol{v} = \left(\sum_{k=0}^{n-1} a_k \lambda^k \right) \boldsymbol{v}$$

and thus if we have n distinct eigenvalues λ_j

$$\sum_{k=0}^{n-1} a_k \lambda_j^k = e^{\lambda_j}, \qquad j = 1, 2, \ldots, n \tag{5}$$

so that (5) is an interpolation problem which can be used to compute the coefficients a_k. In the case of multiple eigenvalues we use the corresponding generalized eigenvectors (see equation (3)). For example consider the eigenvectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ such that

$$\boldsymbol{A}\boldsymbol{v}_1 = \lambda \boldsymbol{v}_1, \qquad \boldsymbol{A}\boldsymbol{v}_2 = \lambda \boldsymbol{v}_2 + \boldsymbol{v}_1,$$

then we have

$$\begin{aligned} \boldsymbol{A}^2 \boldsymbol{v}_2 &= \lambda \boldsymbol{A}\boldsymbol{v}_2 + \boldsymbol{A}\boldsymbol{v}_1, \\ &= \lambda(\lambda \boldsymbol{v}_2 + \boldsymbol{v}_1) + \lambda \boldsymbol{v}_1, \\ &= \lambda^2 \boldsymbol{v}_2 + 2\lambda \boldsymbol{v}_1, \end{aligned}$$

and again

$$\begin{aligned} \boldsymbol{A}^3 \boldsymbol{v}_2 &= \boldsymbol{A} \left(\lambda^2 \boldsymbol{v}_2 + 2\lambda \boldsymbol{v}_1 \right), \\ &= \lambda^2 \boldsymbol{A}\boldsymbol{v}_2 + 2\lambda \boldsymbol{A}\boldsymbol{v}_1, \\ &= \lambda^2 (\lambda \boldsymbol{v}_2 + \boldsymbol{v}_1) + 2\lambda \boldsymbol{A}\boldsymbol{v}_1, \\ &= \lambda^3 \boldsymbol{v}_2 + 3\lambda^2 \boldsymbol{v}_1, \end{aligned}$$

and in general

$$\boldsymbol{A}^k \boldsymbol{v}_2 = \lambda^k \boldsymbol{v}_2 + k\lambda^{k-1} \boldsymbol{v}_1, \tag{6}$$

using (6) in matrix exponential we have

$$
\begin{aligned}
e^{\boldsymbol{A}}\boldsymbol{v}^2 &= \sum_{k=0}^{\infty} \frac{1}{k!}\boldsymbol{A}^k\boldsymbol{v}_2 = \sum_{k=0}^{\infty} \frac{1}{k!}\left(\lambda^k\boldsymbol{v}_2 + k\lambda^{k-1}\boldsymbol{v}_1\right), \\
&= \sum_{k=0}^{\infty} \frac{1}{k!}\lambda^k\boldsymbol{v}_2 + \sum_{k=0}^{\infty} \frac{1}{k!}k\lambda^{k-1}\boldsymbol{v}_1, \\
&= \sum_{k=0}^{\infty} \frac{1}{k!}\lambda^k\boldsymbol{v}_2 + \sum_{k=0}^{\infty} \frac{1}{(k-1)!}\lambda^{k-1}\boldsymbol{v}_1, \\
&= e^{\lambda}\boldsymbol{v}_1 + e^{\lambda}\boldsymbol{v}_2
\end{aligned}
\tag{7}
$$

using (6) in a polynomial matrix we have

$$
\begin{aligned}
p(\boldsymbol{A})\boldsymbol{v}_2 &= \sum_{k=0}^{m} p_k\boldsymbol{A}^k\boldsymbol{v}_2, \\
&= \sum_{k=0}^{m} p_k\left(\lambda^k\boldsymbol{v}_2 + k\lambda^{k-1}\boldsymbol{v}_1\right), \\
&= p(\lambda)\boldsymbol{v}_2 + p'(\lambda)\boldsymbol{v}_1
\end{aligned}
\tag{8}
$$

from (7) and (8) we have that $p(\lambda) = p'(\lambda) = e^{\lambda}$ for a multiple eigenvalue. In general it can be proved that if λ is an eigenevalue of multiplicity m we have

$$
p(\lambda) = p'(\lambda) = \cdots = p^{(m-1)}(\lambda) = e^{\lambda}.
$$

thus using eigenvalues with their multiplicity we have an Hermite interpolation problem with enough conditions to determine uniquely the polynomial.

Example 1 Consider the matrix

$$
\boldsymbol{A} = \begin{pmatrix} 2 & 0 & 1 & 1 \\ -4 & 4 & 4 & -1 \\ 2 & -1 & 1 & 2 \\ 0 & 0 & 0 & 2 \end{pmatrix}
$$

we have

$$\Delta(\lambda) = |\boldsymbol{A} - \lambda \boldsymbol{I}| = 24 - 44\lambda + 30\lambda^2 - 9\lambda^3 + \lambda^4$$

which can be factorized as

$$\Delta(\lambda) = (\lambda - 2)^3(\lambda - 3)$$

The matrix exponential is a polynomial $p(\boldsymbol{A})$ where $p(x) = p_0 + p_1 x + p_2 x^2 + p_4 x^3$, to determine $p(x)$ we use interpolation conditions:

$$\begin{aligned} p(2) &= \; p_0 + 2p_1 + 4p_2 + 8p_4 &= e^2, \\ p'(2) &= \quad p_1 + 4p_2 + 12p_4 &= e^2, \\ p''(2) &= \qquad 2p_2 + 12p_4 &= e^2, \\ p(3) &= p_0 + 3p_1 + 9p_2 + 27p_4 &= e^2, \end{aligned}$$

which has the solution

$$\begin{aligned} p_0 &= 21\,e^2 - 8\,e^3, & p_1 &= -31\,e^2 + 12\,e^3, \\ p_2 &= \frac{31}{2}\,e^2 - 6\,e^3, & p_3 &= -\frac{5}{2}\,e^2 + \,e^3, \end{aligned}$$

and evaluating $p(\boldsymbol{A})$ we have

$$e^{\boldsymbol{A}} = e^2 \begin{pmatrix} -3 & 2 & 3 & -1/2 \\ -4 & 3 & 0 & 0 \\ -2 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + e^3 \begin{pmatrix} 2 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 2 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

2.3 Using numerical integration

Consider the ODE:

$$\boldsymbol{x}_k' = \boldsymbol{A}\boldsymbol{x}_k, \quad \boldsymbol{x}(0) = \boldsymbol{e}_k = (0, \ldots, 0, \underbrace{1}_{\text{k-position}}, 0, \ldots, 0)^T$$

then the solution is

$$\boldsymbol{x}_k(t) = e^{t\boldsymbol{A}}\boldsymbol{e}_k$$

and collecting the solution for $k = 1, 2, \ldots, n$ we have

$$
\begin{aligned}
(\,\boldsymbol{x}_1(t), \boldsymbol{x}_2(t), \ldots, \boldsymbol{x}_n(t)\,) &= (\,e^{t\boldsymbol{A}}\boldsymbol{e}_1, e^{t\boldsymbol{A}}\boldsymbol{e}_2, \ldots, e^{t\boldsymbol{A}}\boldsymbol{e}_n\,)\,, \\
&= e^{t\boldsymbol{A}}\,(\,\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\,)\,, \\
&= e^{t\boldsymbol{A}}\boldsymbol{I} \\
&= e^{t\boldsymbol{A}},
\end{aligned}
$$

Thus the following matricial ODE

$$
\boldsymbol{X}'(t) = \boldsymbol{A}\boldsymbol{X}(t), \qquad \boldsymbol{X}(0) = \boldsymbol{I}, \tag{9}
$$

has the solution:

$$
\boldsymbol{X}(t) = e^{t\boldsymbol{A}}\boldsymbol{I} = e^{t\boldsymbol{A}}.
$$

Using this observation we can use a numerical integrator with step $\Delta t = t/m$

$$
\begin{aligned}
\boldsymbol{X}_0 &= \boldsymbol{I} \\
\boldsymbol{X}_{k+1} &= \boldsymbol{X}_k + \Delta t \boldsymbol{\Phi}(t_k, \boldsymbol{X}_k), \qquad k = 0, 1, \ldots, m-1 \\
e^{t\boldsymbol{A}} &\approx \boldsymbol{X}_m.
\end{aligned}
$$

for example using *explicit* Euler scheme we have

$$
\begin{aligned}
\boldsymbol{X}_0 &= \boldsymbol{I} \\
\boldsymbol{X}_{k+1} &= \boldsymbol{X}_k + \Delta t \boldsymbol{A}\boldsymbol{X}_k = (\boldsymbol{I} + \Delta t \boldsymbol{A})\boldsymbol{X}_k, \qquad k = 0, 1, \ldots, m-1 \\
e^{t\boldsymbol{A}} &\approx \boldsymbol{X}_m = (\boldsymbol{I} + \Delta t \boldsymbol{A})^m.
\end{aligned} \tag{10}
$$

or using *implicit* Euler scheme we have

$$
\begin{aligned}
\boldsymbol{X}_0 &= \boldsymbol{I} \\
\boldsymbol{X}_{k+1} &= \boldsymbol{X}_k + \Delta t \boldsymbol{A}\boldsymbol{X}_{k+1}, \qquad k = 0, 1, \ldots, m-1 \\
e^{t\boldsymbol{A}} &\approx \boldsymbol{X}_m = (\boldsymbol{I} - \Delta t \boldsymbol{A})^{-m}.
\end{aligned}
$$

Remark 2 The computation can be reduced choosing the number of steps m as a power of two $m = 2^p$ is this case the matrix multiplication can be reduced from m to p. For example for Euler method (10) we have:

$$\begin{aligned} \boldsymbol{R}_0 &= \boldsymbol{I} + \Delta t \boldsymbol{A} \\ \boldsymbol{R}_{k+1} &= \boldsymbol{R}_k^2, \qquad k = 0, 1, \ldots, p-1 \\ e^{t\boldsymbol{A}} &\approx \boldsymbol{R}_p. \end{aligned}$$

Remark 3 Choosing $\Delta t = t$ i.e $m = 1$ only one step and using Taylor expansion as advancing numerical scheme we obtain again the taylor series approximation of the matrix exponential

2.4 Using Pade approximation and squaring

Consider the ODE (9) and the Crank–Nicholson approximation we have

$$\begin{aligned} \boldsymbol{X}_0 &= \boldsymbol{I} \\ \boldsymbol{X}_{k+1} &= \boldsymbol{X}_k + \frac{\Delta t}{2} \boldsymbol{A} \left(\boldsymbol{X}_k + \boldsymbol{X}_{k+1} \right), \qquad k = 0, 1, \ldots, m-1 \\ e^{t\boldsymbol{A}} &\approx \boldsymbol{X}_m = \left[\left(\boldsymbol{I} - \frac{\Delta t}{2} \boldsymbol{A} \right)^{-1} \left(\boldsymbol{I} + \frac{\Delta t}{2} \boldsymbol{A} \right) \right]^m. \end{aligned} \tag{11}$$

by choosing $m = 2^P$ equation (11) can be reorganized as

$$\begin{aligned} \boldsymbol{X}_0 &= \left(\boldsymbol{I} - \frac{\Delta t}{2} \boldsymbol{A} \right)^{-1} \left(\boldsymbol{I} + \frac{\Delta t}{2} \boldsymbol{A} \right) \\ \boldsymbol{X}_{k+1} &= \boldsymbol{X}_k^2, \qquad k = 0, 1, \ldots, p-1 \\ e^{t\boldsymbol{A}} &\approx \boldsymbol{X}_p. \end{aligned} \tag{12}$$

Procedure (12) can be generalized by observing

$$e^{t\boldsymbol{A}} = e^{(t\boldsymbol{A}/m)m} = \left(e^{(t\boldsymbol{A})/m} \right)^m.$$

Thus approximating $e^{(t\boldsymbol{A})/m}$ with a rational polynomial, i.e.

$$e^{(t\boldsymbol{A})/m} \approx P(t\boldsymbol{A}/m)^{-1} Q(t\boldsymbol{A}/m)$$

permits to approximate the exponential as follows

$$
\begin{aligned}
\boldsymbol{X}_0 &= P(t\,2^{-p}\boldsymbol{A})^{-1}Q(t\,2^{-p}\boldsymbol{A}) \\
\boldsymbol{X}_{k+1} &= \boldsymbol{X}_k^2, \qquad k = 0, 1, \ldots, p-1 \\
e^{t\boldsymbol{A}} &\approx \boldsymbol{X}_p.
\end{aligned}
$$

when $p = 0$ the rational polynomial $P(x)/Q(x)$ approximate e^x. The key idea of the squaring algorithm is to choose p large enough to have $\|t2^{-p}\boldsymbol{A}\| \leq C$ where C is a small constant (e.g. 1 or 1/2) where the rational polynomial $P(z)/Q(z)$ is a good approximation of e^z for $z \in \mathbb{C}$ and $|z| \leq C$.

To approximate exponential with a rational polynomial we can use Padé procedure with schematically determine the coefficients of $P(x)$ and $Q(x)$ by matching the product

$$
Q(x)e^x - P(x) = \mathcal{O}(x^r)
$$

with r the maximum possibile.

Example 2 Let $P(x) = 1 + p_1 x$ and $Q(x) = q_0 + q_1 x$ then

$$
(q_0 + q_1 x)\left(1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \mathcal{O}(x^4)\right) - (1 + p_1 x) =
$$

$$
q_0 - 1 + x(q_0 + q_1 - p_1) + \frac{x^2}{2}(q_0 + 2q_1) + \frac{x^3}{6}(q_0 + 3q_1) + \mathcal{O}(x^4)
$$

and matching up to x^3 produce the linear system:

$$
\begin{cases}
q0 = 1 \\
q_0 + q_1 - p_1 = 0 \\
q_0 + 2q_1 = 0 \\
q_0 + 3q_1 = 0
\end{cases}
$$

which has the solution $q_0 = 1$, $q_1 = -1/2$, $p_1 = 1/2$ and the rational polynomial is $P(x)/Q(x) = (1 + x/2)/(1 - x/2)$.

Using (for example) procedure of example 2 we have the followiong table

$\frac{1}{1}$	$\frac{1}{1-z}$	$\frac{1}{1-z+\frac{z^2}{2}}$	$\frac{1}{1-z+\frac{z^2}{2}-\frac{z^3}{6}}$
$\frac{1+z}{1}$	$\frac{1+\frac{z}{2}}{1-\frac{z}{2}}$	$\frac{1+\frac{z}{3}}{1-\frac{2z}{3}+\frac{z^2}{6}}$	$\frac{1+\frac{z}{4}}{1-\frac{3z}{2}+\frac{z^2}{4}-\frac{z^3}{24}}$
$\frac{1+z+\frac{z^2}{2}}{1}$	$\frac{1+\frac{2z}{3}+\frac{z^2}{6}}{1-\frac{z}{3}}$	$\frac{1+\frac{z}{2}+\frac{z^2}{12}}{1-\frac{z}{2}+\frac{z^2}{12}}$	$\frac{1+\frac{2z}{5}+\frac{z^2}{20}}{1-\frac{3z}{5}+\frac{3z^2}{20}-\frac{z^3}{60}}$
$\frac{1+z+\frac{z^2}{2}+\frac{z^3}{6}}{1}$	$\frac{1+\frac{3z}{2}+\frac{z^2}{4}+\frac{z^3}{24}}{1-\frac{z}{4}}$	$\frac{1+\frac{3z}{5}+\frac{3z^2}{20}+\frac{z^3}{60}}{1-\frac{2z}{5}+\frac{z^2}{20}}$	$\frac{1+\frac{2z}{5}+\frac{z^2}{10}+\frac{z^3}{120}}{1-\frac{z}{2}+\frac{z^2}{10}-\frac{z^3}{120}}$