Exponential Matrix and Stability of Systems.

M.V. Safonov
School of Mathematics, University of Minnesota, Minneapolis, MN, 55455.

1 Exponential Matrix.

Definition 1.1. For \( n \times n \) matrix \( A \), the exponential matrix function

\[
e^{tA} = I + \frac{t}{1!}A + \frac{t^2}{2!}A^2 + \cdots = \sum_{k=0}^{\infty} \frac{t^k}{k!}A^k.
\]  

(1.1)

Lemma 1.2. \( X = X(t) = e^{tA} \) is the unique solution of the Cauchy problem

\[
X' = \frac{dX}{dt} = AX, \quad X(0) = I.
\]  

(1.2)

Proof. It is easy to see that

\[
\frac{d}{dt}e^{tA} = A + \frac{t}{1!}A^2 + \cdots = Ae^{tA},
\]

so that \( X = e^{tA} \) is a solution of (1.2). Moreover, we can treat \( n \times n \)-matrix function \( X \) as a vector functions with values in \( \mathbb{R}^{n^2} \) (or \( \mathbb{C}^{n^2} \)), we only need to rewrite the matrix equation \( X' = AX \) in the vector form \( X' = BX \) with a \( n^2 \times n^2 \)-matrix \( B \). Then the uniqueness for the Cauchy problem in the vector form implies the uniqueness for the problem (1.2).

Definition 1.3. If \( AV = \lambda V \), for some vector \( V \neq 0 \), then \( \lambda \) is an eigenvalue of \( A \), and \( V \) is an eigenvector corresponding to \( \lambda \).

We have \( AV = \lambda V \iff (A - \lambda I)V = 0 \). The last system has nontrivial solutions \( V \neq 0 \iff |A - \lambda I| = 0 \). We introduce the characteristic polynomial of \( A \) by the formula \( p(\lambda) = |\lambda I - A| \).

Now we can conclude:

(i) The eigenvalues of \( A \) are roots of the characteristic equation

\[
p(\lambda) = |\lambda I - A| = 0.
\]  

(1.3)

(ii) For each eigenvalue \( \lambda \), the corresponding eigenvectors \( V \) are nontrivial solutions of the system

\[
(A - \lambda I)V = 0.
\]  

(1.4)

Lemma 1.4. Let \( A \) be a constant \( n \times n \) matrix, and let \( AV = \lambda V \) for some vector \( V \). Then the matrix function \( U = U(t) = e^{\lambda t}V \) satisfies \( U' = AU \).
Proof. \( U'(t) = (e^\lambda t)' V = e^\lambda t V e^\lambda A V = AU(t) \). \( \square \)

Lemma 1.5. For distinct eigenvalues \( \lambda_1, \lambda_2, \ldots, \lambda_m \) of the matrix \( A \), the corresponding eigenvectors \( V_1, V_2, \cdots, V_m \) are linearly independent.

Proof. For \( m = 1 \), this is trivial:
\[
V_1 \neq 0, \quad c_1 V_1 = 0 \quad \Rightarrow \quad c_1 = 0.
\]

Now suppose this statement is true for some \( m = k \), i.e. \( V_1, V_2, \cdots, V_k \) are linearly independent. We will prove that it remains true for \( m = k + 1 \), i.e. the equality
\[
c_1 V_1 + \cdots + c_k V_k + c_{k+1} V_{k+1} = 0
\]
holds only in case \( c_1 = \cdots = c_k = c_{k+1} = 0 \). Multiplying (1.5) by the matrix \( A \) and using the equalities \( AV_j = \lambda_j V_j \), we get
\[
c_1 \lambda_1 V_1 + \cdots + c_k \lambda_k V_k + c_{k+1} \lambda_{k+1} V_{k+1} = 0.
\]

Now subtract (1.5) multiplied by \( \lambda_{k+1} \). This gives us
\[
c_1 (\lambda_1 - \lambda_{k+1}) V_1 + \cdots + c_k (\lambda_k - \lambda_{k+1}) V_k = 0.
\]

By our assumption, \( V_1, V_2, \cdots, V_k \) are linearly independent. Therefore,
\[
c_1 (\lambda_1 - \lambda_{k+1}) = \cdots = c_k (\lambda_k - \lambda_{k+1}) = 0.
\]

Since all the eigenvalues are distinct, this implies \( c_1 = \cdots = c_k = 0 \). Now from (1.5) it follows \( c_{k+1} = 0 \). This proves our statement for \( m = k + 1 \), and by induction, it is true for arbitrary \( m \). \( \square \)

Another way of computation \( e^{tA} \) is based on the following famous result.

Theorem 1.6 (Cayley-Hamilton). Let \( A = (a_{ij}) \) be a \( n \times n \) matrix, and let \( p(\lambda) = |\lambda I - A| \). Then \( p(A) = 0 \).

Proof. Denote
\[
p(\lambda) = |\lambda I - A| = \sum_{j=0}^{n} p_j \lambda^j, \quad (p_n = 1).
\]

Similarly to the geometric series for large \( \lambda > 0 \), we have a convergent series
\[
(\lambda I - A)^{-1} = \frac{1}{\lambda} \cdot \left( I - \frac{1}{\lambda} \cdot A \right)^{-1} = \frac{1}{\lambda} \cdot \sum_{k=0}^{\infty} \lambda^{-k} A^k.
\]

On the other hand, the elements of this function are co-factors of \( \lambda I - A \) divided by \( p(\lambda) = |\lambda I - A| \). Therefore, the product
\[
p(\lambda) \cdot (\lambda I - A)^{-1} = \frac{1}{\lambda} \cdot \sum_{j,k} p_j \lambda^{j-k} A_k
\]
is a polynomial of \( \lambda \) with matrix coefficient, so that the coefficients of all negative powers of \( \lambda \) must be 0. In particular, the coefficient of \( \lambda^{-1} \), which corresponds to a part of the above sum for \( j = k \), is
\[
0 = \sum_{j} p_j A^j = p(A).
\]
\( \square \)
Let $\lambda_1, \lambda_2, \cdots, \lambda_s$ denote all the distinct eigenvalues of a matrix $A$ (real and complex). Then we can write the characteristic polynomial of $A$ in the form
\[
p(\lambda) = |\lambda I - A| = \prod_{k=1}^{s} (\lambda - \lambda_k)^{m_k}, \tag{1.6}\]
where $m_k$ is the multiplicity of $\lambda_k$. We have
\[
\frac{1}{p(\lambda)} = \sum_{k=1}^{s} \frac{a_k(\lambda)}{(\lambda - \lambda_k)^{m_k}}, \tag{1.7}\]
where $a_k(\lambda)$ is a polynomial of degree $\leq m_k - 1$ for each $k$. Therefore,
\[
1 = \sum_{k=1}^{s} a_k(\lambda)p_k(\lambda), \quad I = \sum_{k=1}^{s} a_k(A)p_k(A), \tag{1.8}\]
where
\[
p_k(\lambda) = \frac{p(\lambda)}{(\lambda - \lambda_k)^{m_k}} = \prod_{j \neq k} (\lambda - \lambda_j)^{m_j}. \tag{1.9}\]

**Theorem 1.6.** Let $\lambda_1, \lambda_2, \cdots, \lambda_s$ be all the distinct eigenvalues of a matrix $A$ with the characteristic polynomial $p(\lambda)$ in (1.6). Then
\[
e^{tA} = \sum_{k=1}^{s} e^{\lambda_k t} a_k(A)p_k(A) \sum_{j=0}^{m_k-1} \frac{t^j}{j!} (A - \lambda_k I)^j, \tag{1.10}\]
where $a_k$ are polynomials of degree $\leq m_k - 1$, which are defined from the decomposition (1.7), and $p_k$ are polynomials in (1.9).

First we prove a lemma, which allows us to use the equalities
\[
e^{tA} = e^{tA}e^{t(A-\lambda A)} = e^{tA}e^{t(A-\lambda A)}. \tag{1.11}\]

**Lemma 1.7.** Let $A$ and $B$ be $n \times n$ matrices satisfying $AB = BA$. Then
\[
e^{tA}e^{tB} = e^{t(A+B)}. \tag{1.12}\]

**Proof of Lemma.** We have
\[
AB = BA \implies A^kB = BA^k \implies e^{tA}B = Be^{tA}.
\]
Therefore, the matrix function $X(t) = e^{tA}e^{tB}$ satisfies
\[
X' = (e^{tA})'e^{tB} + e^{tA}(e^{tB})' = Ae^{tA}e^{tB} + Be^{tA}e^{tB} = (A + B)X, \quad X(0) = I.
\]
By Lemma 1.2, we have $X(t) = e^{t(A+B)}$. \(\square\)
Proof of Theorem 1.6. Using the equalities (1.8) and (1.11) we get

\[ e^{tA} = \sum_{k=1}^{s} a_k(A)p_k(A)e^{\lambda k t} = \sum_{k=1}^{s} e^{\lambda k t} a_k(A)p_k(A)e^{t(A - \lambda_k I)}. \]  

(1.12)

Since \( p_k(A)(A - \lambda_k I)^{m_k} = p(A) = 0 \),

\[ p_k(A)e^{t(A - \lambda_k I)} = p_k(A) \sum_{j=0}^{\infty} \frac{t^j}{j!}(A - \lambda_k I)^j = p_k(A) \sum_{j=0}^{m_k-1} \frac{t^j}{j!}(A - \lambda_k I)^j. \]

Together with (1.12), this gives us the equality (1.10).

Example 1.8. Consider the matrix \( A = \begin{pmatrix} -1 & 1 & -2 \\ 0 & -1 & 4 \\ 0 & 0 & 1 \end{pmatrix} \). We have

\[ p(\lambda) = |\lambda I - A| = (\lambda + 1)^2(\lambda - 1) \implies \lambda_1 = -1, \quad m_1 = 2, \quad \lambda_2 = 1, \quad m_2 = 1, \quad p_1(\lambda) = \lambda - 1, \quad p_2(\lambda) = (\lambda + 1)^2; \]

\[ \frac{1}{p(\lambda)} = \frac{1}{(\lambda + 1)^2(\lambda - 1)} = \frac{1}{4} \left( \frac{1}{\lambda - 1} + \frac{\lambda + 3}{(\lambda + 1)^2} \right) \implies a_1(\lambda) = -\frac{1}{4}(\lambda + 3), \quad a_2(\lambda) = \frac{1}{4}. \]

By (1.10), we have

\[ e^{tA} = -\frac{1}{4} e^{-t}(A + 3I)(A - I)\left[ I + t(A + I) \right] + \frac{1}{4} e^{t}(A + I)^2. \]

Finally,

\[ -\frac{1}{4}(A + 3I)(A - I) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix}, \quad \frac{1}{4}(A + I)^2 = \frac{1}{4} \begin{pmatrix} 0 & 1 & -2 \\ 0 & 0 & 4 \\ 0 & 0 & 2 \end{pmatrix}^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix} \implies e^{tA} = e^{-t} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix} + te^{-t} \begin{pmatrix} 0 & 1 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + e^t \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix}. \]

Example 1.9. Consider the matrix \( A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \). We have

\[ p(\lambda) = \lambda^2 + 1, \quad \lambda_1 = i, \quad p_1(\lambda) = \lambda + i; \quad \lambda_2 = -i, \quad p_2(\lambda) = \lambda - i, \quad m_1 = m_2 = 1; \]

\[ \frac{1}{\lambda^2 + 1} = \frac{a_1}{\lambda - 1} + \frac{a_2}{\lambda + 1}, \quad \text{where} \quad a_1 = \lim_{\lambda \to i} \frac{\lambda - i}{\lambda^2 + 1} = \frac{1}{2i}; \quad a_2 = -\frac{1}{2i}; \]

\[ e^{tA} = \sum e^{\lambda k t} a_k p_k(A) = e^{it} \cdot \frac{1}{2i} \cdot (A + iI) + e^{-it} \cdot \frac{-1}{2i} \cdot (A - iI) = \frac{e^{it} - e^{-it}}{2i} \cdot A + \frac{e^{it} + e^{-it}}{2} \cdot I \]

\[ = \cos t \cdot I + \sin t \cdot A = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}. \]