Borovets, April 5-8, 2006

# EXPONENTIATION OF $2 \times 2$ AND $3 \times 3$ MATRICES WITHOUT CANONIZATION* 

Angel Popov, Todor Todorov


#### Abstract

We derive explicit formulae for calculating the exponential $e^{A}$ of a given $2 \times 2$ and $3 \times 3$ matrix $A$. They are exclusively in terms of characteristic roots of $A$ and do not involve neither the eigenvectors of $A$, nor the transition matrix associated with a particular canonical basis. In addition, some specific matrices, closely related to $A$ are identified, which allows an easy calculation of some invariant subspaces and canonical bases of $A$, if needed. We believe that our approach has advantage (especially if applied by non-mathematicians) over the more conventional methods based on the construction of canonical bases. We support this point with several examples.


1. Introduction. The exponential $e^{A}$ of a square matrix $A$ and the related oneparametric family $e^{t A}$ are important concept in mathematics, e.g. the description of oneparameter subgroups of $G L_{n}$ allows an useful geometrical visualization and interpretation in the Lie structures associated. Here, one example on an undergraduate college level:

Let $\frac{d}{d t} \vec{x}(t)=A \vec{x}(t)$ be a system of first-order constant-coefficient ordinary differential equations with initial condition $\vec{x}(0)=\vec{x}_{0} \in \mathbb{R}^{n}$, where $A$ is a $n \times n$ matrix (with real entries). Then the solution of the system is given by $\vec{x}(t)=e^{t A} \vec{x}_{0}$.

The exponential $e^{A}$ of a given $n \times n$ real or complex matrix $A$ is defined by the Taylor expansion $e^{A}=\sum_{n=0}^{\infty} \frac{A^{n}}{n!}$, which converges absolutely for all complex matrices $A$. So defined matrix exponential satisfies the following properties:

- If the matrices $A$ and $B$ commute, i.e. $A B=B A$, then $e^{A+B}=e^{A} e^{B}$.
- If the matrices $A$ and $B$ are similar, i.e. $B=T^{-1} A T$, then $e^{B}=T^{-1} e^{A} T$.
- The exponential of a cell-diagonal matrix is the diagonal matrix of the exponents of the cells, i.e. if $A=\operatorname{diag},\left(A_{1}, A_{2}, \ldots\right)$, then $e^{t A}=\operatorname{diag}\left(e^{t A_{1}}, e^{t A_{2}}, \ldots\right)$.

For some specific kind of matrices (scalar, nilpotent, idempotent) the exponential can be easily evaluated. Here we list some explicit formulae we shall use later:

Lemma 1.1. For the exponential of a given (real) matrix we have:
(i) If $A=I$ is the identity matrix, then $e^{t A}=e^{t} I$.
(ii) If $N$ is nilpotent with $N^{m+1}=0$, then $e^{t N}=I+\frac{t N}{1!}+\cdots+\frac{t^{m} N^{m}}{m!}$.

[^0](iii) If $P^{2}=P$, i.e. $P$ is a projector, then $e^{t P}=I+\left(e^{t}-1\right) P=I-P+e^{t} P$.
(iv) If $J^{2}=-I$, then $e^{t J}=(\cos t) I+(\sin t) J$.
(v) If $J^{2}=I$, then $e^{t J}=(\cosh t) I+(\sinh t) J=\frac{1}{2}\left[e^{t}(I+J)+e^{-t}(I-J)\right]$.

Proof. The properties (i)-(ii) are obvious, and if $P$ is a projector, we calculate

$$
e^{t P}=\sum_{n=0}^{\infty} \frac{t^{n} P^{n}}{n!}=I+\sum_{n=1}^{\infty} \frac{t^{n}}{n!} P=I+\left(e^{t}-1\right) P
$$

which proves (iii).
If $J^{2}=-I$, then we have:

$$
\begin{gathered}
e^{t J}=\sum_{n=0}^{\infty} \frac{t^{n} J^{n}}{n!}=\sum_{n=0}^{\infty} \frac{t^{2 n} J^{2 n}}{(2 n)!}+\sum_{n=0}^{\infty} \frac{t^{2 n+1} J^{2 n+1}}{(2 n+1)!}= \\
=\sum_{n=0}^{\infty} \frac{(-1)^{n} t^{2 n}}{(2 n)!} I+\sum_{n=0}^{\infty} \frac{(-1)^{n} t^{2 n+1}}{(2 n+1)!} J=(\cos t) I+(\sin t) J,
\end{gathered}
$$

as required. By analogy one proves (v):

$$
\begin{aligned}
& e^{t J}=\sum_{n=0}^{\infty} \frac{t^{n} J^{n}}{n!}=\sum_{n=0}^{\infty} \frac{t^{2 n} J^{2 n}}{(2 n)!}+\sum_{n=0}^{\infty} \frac{t^{2 n+1} J^{2 n+1}}{(2 n+1)!}= \\
& \sum_{n=0}^{\infty} \frac{t^{2 n}}{(2 n)!} I+\sum_{n=0}^{\infty} \frac{t^{2 n+1}}{(2 n+1)!} J=(\cosh t) I+(\sinh t) J
\end{aligned}
$$

as required.
In Section 3 we shall use an easy generalization of the last two properties:
Lemma 1.2. For the exponential of a given (real) matrix we have:
(i) Let $J^{2 n}=(-1)^{n} Q$, and $J^{2 n-1}=(-1)^{n-1} J, n=1,2, \ldots$. Then

$$
e^{t J}=I-Q+(\cos t) Q+(\sin t) J
$$

(ii) Let $J^{2 n}=Q$, and $J^{2 n-1}=J, n=1,2, \ldots$ Then
$e^{t J}=I-Q+(\cosh t) Q+(\sinh t) J=I-Q+e^{t}(Q+J)+e^{-t}(Q-J)$.

Proof. The proof is analogous to the proof of the last three properties of Lemma 1.1 and we leave it to the reader.

Some of the relations, just described, can be found in many books on linear algebra and applications, see Artin [1] and Malcev [2]. The calculation of the exponential of an arbitrary matrix however is much more complicated. The technology uses some of the canonical forms of $A$, say the Jordan one. One has to find a transition matrix $T$, such that $C=T^{-1} A T=\operatorname{diag}\left(C_{1}, C_{2}, \ldots\right)$ with Jordan cells $C_{1}, C_{2}, \ldots$ on the diagonal and then $e^{A}=T \operatorname{diag}\left(e^{C_{1}}, e^{C_{2}}, \ldots\right) T^{-1}$. Next, each Jordan cell $C_{k}, k=1,2, \ldots$ has the form $C_{k}=\lambda_{k} I+N$, where $N$ is nilpotent and one evaluates $e^{C_{k}}=e^{\lambda_{k}} \sum_{n \geq 0} \frac{N^{n}}{n!}$, where the last sum is finite, because $N$ is nilpotent.

The procedure discussed is universal, but it has some misgivings:

- The algorithm for calculating the transition matrix is heavy and time consuming. It is especially complicated in the case of multiple characteristic roots, when the construction of the chain of generalized eigenvectors requires skills well beyond in a typical course
of linear algebra and applications.
- The original framework $\mathbb{R}^{n}$ must be extended to $\mathbb{C}^{n}$, which usually is confusing for non-mathematicians.

In this article we derive some explicit formulae for evaluating $e^{t A}$ for a given $2 \times 2$ and $3 \times 3$ matrix $A$ with real entries (in Section 2 and Section 3, respectively). They are in terms of characteristic roots of $A$ and do not involve neither the eigenvectors of $A$, nor the transition matrix associated with a particular canonical basis. We believe that our approach is suitable for handbooks in the sense that one can work it with limited background in linear algebra, in particular without having a slightest idea about change of bases and the "ransition matrices".

Our consideration essentially exploits the characteristic polynomial $f_{A}(\lambda)=(-1)^{n}$ det $(A-\lambda I)$ of the matrix $A$, and the Cayley-Hamilton theorem, which says that $f_{A}(A)=0$.

As a corollary of the main results in each section we identify some matrices closely related to $A$, which allows easy calculations of some invariant subspaces of $A$ and an efficient construction of canonical bases of $A$.

The eigenspace of $A$ corresponding to the particular eigenvalue $c \in \mathbb{C}$ will be denoted by $E_{\lambda=c}(A)$.

We recommend our approach for teaching exponential of matrices in a course on linear algebra and application to differential equations, and to support this point we present several examples.
2. Exponential of $2 \times 2$ matrices. Here we derive formulae for evaluating the exponential of a given $2 \times 2$ matrix. They are easy to memorize and simple to use. One can work with a limited background in linear algebra, and in addition the formulae involve only linear operations between the matrices. This point is illustrated with several examples of first-order linear systems of ordinary differential equations.

We would like to mention that the results of Case 1 and Case 2 of Theorem 2.1 had been exploited by Paul Bamberg and Shlomo Sternberg [3] in similar context.

An easy interpretation of the main result in Corollary 2.2 allows a direct construction of a canonical basis and the transition matrix associated.

The main result of the section is
Theorem 2.1. Let $A$ be a given $2 \times 2$ real matrix with characteristic roots $\lambda_{1}$ and $\lambda_{2}$. Then

Case 1. If $\lambda_{1}=\lambda_{2}=\lambda_{0}$ (it is real), then $A=\lambda_{0} I+N$, where $N=A-\lambda_{0} I$ is nilpotent with $N^{2}=0$. Consequently, for every $t \in \mathbb{R}$ we have

$$
\begin{equation*}
e^{t A}=e^{\lambda_{0} t}(I+t N) \tag{1}
\end{equation*}
$$

Case 2. If $\lambda_{1,2}=\alpha \pm i \omega$ for some $\alpha, \omega \in \mathbb{R}, \omega \neq 0$, then $A=\alpha I+\omega J$, where the matrix $J=\frac{1}{\omega}(A-\alpha I)$ satisfies $J^{2}=-I$. Consequently, for every $t \in \mathbb{R}$ we have

$$
\begin{equation*}
e^{t A}=e^{\alpha t}[\cos (\omega t) I+\sin (\omega t) J] \tag{2}
\end{equation*}
$$

Case 3. If $\lambda_{1} \neq \lambda_{2}$ (real), then $A=\alpha I+\beta J$, where $\alpha=\frac{\lambda_{1}+\lambda_{2}}{2}, \beta=\frac{\lambda_{1}-\lambda_{2}}{2}$, and the matrix $J=\frac{1}{\beta}(A-\alpha I)$ satisfies $J^{2}=I$. Consequently, for every $t \in \mathbb{R}$ we have

$$
\begin{equation*}
e^{t A}=e^{\alpha t}[\cosh (\beta t) I+\sinh (\beta t) J]=\frac{1}{2}\left[e^{\lambda_{1} t}(I+J)+e^{\lambda_{2} t}(I-J)\right] \tag{3}
\end{equation*}
$$

Proof. The characteristic polynomial of $A$ is $f_{A}(\lambda)=\lambda^{2}-\left(\lambda_{1}+\lambda_{2}\right) \lambda+\lambda_{1} \lambda_{2}$, and by the Cayley-Hamilton theorem we have in each of the cases:

Case 1. $A^{2}-2 \lambda_{0} A+\lambda_{0}^{2} I=0$, i.e. $N^{2}=\left(A-\lambda_{0} I\right)^{2}=0$, and we evaluate $e^{t A}=e^{\lambda_{0} t I} e^{t N}$ by (i) and (ii) of Lemma 1.1.

Case 2. $A^{2}-2 \alpha A+\left(\alpha^{2}+\omega^{2}\right) I=0$, i.e. $(A-\alpha I)^{2}=-\omega^{2} I$, which implies $J^{2}=-I$. Now we evaluate $e^{t A}=e^{\alpha t I} e^{\omega t J}$ by (i) and (iv) of Lemma 1.1.

Case 3. $A^{2}-2 \alpha A+\left(\alpha^{2}-\beta^{2}\right) I=0$, i.e. $(A-\alpha I)^{2}=\beta^{2} I$, which implies $J^{2}=I$, and we complete the calculations by (i) and (v) of Lemma 1.1.

Note that the formulae (1) - (3) use characteristic roots of $A$ only, and the matrices $N$ and $J$ depend linearly on $A$. In fact, one can obtain (1) - (3) even without characteristic roots by examining the square $B^{2}$ of the matrix $B=A-a I$, with $a=\frac{1}{2} \operatorname{tr}(A)$. In addition, eigenvectors and canonical bases for $A$ can be easily determined in terms of $N$ and $J$ - the next result follows directly from the above considerations.

Corollary 2.2. Under the notations of the previous theorem we have:
Case 1. If $N \neq 0$, then $E_{\lambda=\lambda_{0}}(A)=\operatorname{Im}(N)$, i.e. either of the non-zero columns of $N$ is an eigenvector, and a Jordan basis of $A$ is any pair $\vec{u}, N \vec{u}$, with $N \vec{u} \neq 0$.

Case 2. The matrix $A$ is diagonalizable over the complex numbers with eigenvectors $(I \mp i J) \vec{u}, \vec{u} \neq 0$, associated with the eigenvalues $\alpha \pm i \omega$. There are no eigenvectors over the reals and the matrix $A$ has a conformal canonical form $\left(\begin{array}{cc}\alpha & \omega \\ -\omega & \alpha\end{array}\right)$ in the basis $\vec{u}, J \vec{u}$ for every $\vec{u} \neq \overrightarrow{0}$.

Case 3. We have $E_{\lambda=\lambda_{1}}(A)=\operatorname{Im}(I+J)$, and $E_{\lambda=\lambda_{2}}(A)=\operatorname{Im}(I-J)$. The matrix $A$ is diagonalizable, and any two non-zero vectors $(I+J) \vec{u},(I-J) \vec{v}$ constitute a basis of eigenvectors of $A$.

The following examples, tested by the second author in a course on linear algebra and ordinary differential equations illustrate the results above.

Example 1. Find the solution of the initial value problem

$$
x^{\prime}=3 x+2 y, \quad y^{\prime}=-8 x-5 y
$$

with initial condition $x(0)=1, y(0)=-1$.
We have $A=\left(\begin{array}{cc}3 & 2 \\ -8 & -5\end{array}\right)$, and $\vec{x}_{0}=\binom{1}{-1}$. The characteristic polynomial $f_{A}(\lambda)=$ $\lambda^{2}+2 \lambda+1$ has a multiple root $\lambda_{1}=\lambda_{2}=-1$ (Case 1 ). We calculate $N=A+I=$ $\left(\begin{array}{cc}4 & 2 \\ -8 & -4\end{array}\right)$. Check that $N^{2}=0$. Now applying (1) we have:

$$
\vec{x}(t)=e^{-t}(I+t N) \vec{x}_{0}=e^{-t}\left(\begin{array}{cc}
1+4 t & 2 t \\
-8 t & 1-4 t
\end{array}\right)\binom{1}{-1}=e^{-t}\binom{1+2 t}{-1-4 t} .
$$

Thus, $x(t)=(1+2 t) e^{-t}, y(t)=-(1+4 t) e^{-t}$.
Remark. If a canonical form of the matrix $A$ is needed, then the vectors $\vec{u}=(0,1)$ and $N \vec{u}=(2,-4)$ constitute a Jordan basis by Corollary 2.2, Case 1.

Example 2. Find the solution of the initial value problem

$$
x^{\prime}=y, \quad y^{\prime}=-5 x-2 y
$$

with initial condition $x(0)=2, y(0)=1$.

We have $A=\left(\begin{array}{cc}0 & 1 \\ -5 & -2\end{array}\right)$, and $\vec{x}_{0}=\binom{2}{1}$. The characteristic polynomial $f_{A}(\lambda)=$ $\lambda^{2}+2 \lambda+5$ has roots $\lambda_{1,2}=-1 \pm 2 i$, i.e. $\alpha=-1, \omega=2$ (Case 2). We evaluate $J=\frac{1}{\omega}(A-\alpha I)=\frac{1}{2}\left(\begin{array}{cc}1 & 1 \\ -5 & -1\end{array}\right)$. Now $J^{2}=-I$ and by (2) we have

$$
\vec{x}(t)=e^{t A} \vec{x}_{0}=e^{-t}\left[\cos (2 t)\left(\begin{array}{ll}
1 & 0 \\
0 & 1
\end{array}\right)+\frac{1}{2} \sin (2 t)\left(\begin{array}{cc}
1 & 1 \\
-5 & -1
\end{array}\right)\right]\binom{2}{1}
$$

which after standard calculations gives the solution

$$
x(t)=e^{-t}\left(2 \cos 2 t+\frac{3}{2} \sin 2 t\right), y(t)=e^{-t}\left(\cos 2 t-\frac{11}{2} \sin 2 t\right)
$$

Remark. The matrix $A$ is diagonalizable over the complex numbers - a canonical basis can be selected, say $(I-i J) \vec{u}=(2+i,-5 i)$, and $(I+i J) \vec{u}=(2-i, 5 i)$ for $\vec{u}=$ $(2,0)$. In the framework of real numbers the matrix $A$ is similar to the conformal matrix $\left(\begin{array}{cc}-1 & 2 \\ -2 & -1\end{array}\right)$, and a canonical basis is $\vec{u}=(2,0), J \vec{u}=(1,-5)$.

Example 3. Find the solution of the initial value problem

$$
x^{\prime}=5 x-y, \quad y^{\prime}=3 x+y
$$

with initial condition $x(0)=1, y(0)=2$.
We have $A=\left(\begin{array}{cc}5 & -1 \\ 3 & 1\end{array}\right)$, and $\vec{x}_{0}=\binom{1}{2}$. The characteristic polynomial $f_{A}(\lambda)=$ $\lambda^{2}-6 \lambda+8$ has roots $\lambda_{1}=4, \lambda_{2}=2$ i.e. $\alpha=3, \beta=1$ (Case 3 ). We evaluate $J=$ $\frac{1}{\beta}(A-\alpha I)=\left(\begin{array}{ll}2 & -1 \\ 3 & -2\end{array}\right)$. Now $J^{2}=I$ and by $(3)$ we have

$$
\vec{x}(t)=e^{t A} \vec{x}_{0}=e^{3 t}\left[\cosh (t)\left(\begin{array}{ll}
1 & 0 \\
0 & 1
\end{array}\right)+\sinh (t)\left(\begin{array}{ll}
2 & -1 \\
3 & -2
\end{array}\right)\right]\binom{1}{2}
$$

which after standard calculations gives the solution

$$
x(t)=e^{3 t} \cosh t=\frac{1}{2}\left(e^{4 t}+e^{2 t}\right), y(t)=e^{3 t}(\cosh t-\sinh t)=\frac{1}{2}\left(e^{4 t}+3 e^{2 t}\right) .
$$

Remark. The matrix $A$ is diagonalizable - a canonical basis can be selected, say $(I+J) \vec{u}=(-1,-1)$, and $(I-J) \vec{u}=(1,3)$ for $\vec{u}=(0,1)$. Note that the last two are the eigenvectors of $A$.
3. Exponential of $\mathbf{3} \times \mathbf{3}$ matrices. In this section we extend the results of the previous one for $3 \times 3$ matrices. We shall use some elementary and well known facts about characteristic polynomials, e.g. if $B=A-c I$, then $f_{B}(\lambda)=f_{A}(\lambda+c)$, i.e. if $\lambda_{1}, \lambda_{2}, \ldots$ are the characteristic roots of $A$, then $\lambda_{1}-c, \lambda_{2}-c, \ldots$ are the characteristic roots of $B=A-c I$.

The main result of this section is
Theorem 3.1. For a given $3 \times 3$ real matrix $A$ with characteristic roots $\lambda_{1}, \lambda_{2}, \lambda_{3}$ we have

Case 1. Let $\lambda_{1}=\lambda_{2}=\lambda_{3}=\lambda_{0}$ (it is real). Then $A=\lambda_{0} I+N$, where $N=A-\lambda_{0} I$ is nilpotent with $N^{3}=0$ Consequently, for every $t \in \mathbb{R}$ we have

$$
\begin{equation*}
e^{t A}=e^{\lambda_{0} t}\left(I+t N+\frac{t^{2}}{2} N^{2}\right) \tag{4}
\end{equation*}
$$

Case 2. Let $\lambda_{1}=\lambda_{2}=\lambda_{0} \neq \lambda_{3}$ (both are real), and $b=\lambda_{3}-\lambda_{0}$. Then $A=$ $\lambda_{0} I+b P+N$, where the matrices

$$
P=\frac{1}{b^{2}}\left(A-\lambda_{0} I\right)^{2}, \text { and } N=A-\lambda_{0} I-b P
$$

satisfies the following relations: $P N=N P=0, P^{2}=P$, and $N^{2}=0$. Consequently, for every $t \in \mathbb{R}$ we have

$$
\begin{equation*}
e^{t A}=e^{\lambda_{0} t}(I-P+t N)+e^{\lambda_{3} t} P \tag{5}
\end{equation*}
$$

Case 3. Let $\lambda_{1,2}=\alpha \pm i \omega$ for some $\alpha, \omega \in \mathbb{R}, \omega \neq 0$, and $b=\lambda_{3}-\alpha$. Then $A=\alpha I+b P+\omega J$, where the matrices

$$
P=\frac{1}{b^{2}+\omega^{2}}\left[(A-\alpha I)^{2}+\omega^{2} I\right], \text { and } J=\frac{1}{\omega}[A-\alpha I-b P]
$$

satisfies the following relations:

$$
P J=J P=0, P^{2}=P, J^{2 n}=(-1)^{n}(I-P), \text { and } J^{2 n-1}=(-1)^{n-1} J, n=1,2, \ldots
$$

Consequently for every $t \in \mathbb{R}$ we have

$$
\begin{equation*}
e^{t A}=e^{\alpha t}[(\cos \omega t)(I-P)+(\sin \omega t) J]+e^{\lambda_{3} t} P . \tag{6}
\end{equation*}
$$

Case 4. Let $\lambda_{1}, \lambda_{2}, \lambda_{3}$ are distinct real, and let denote $\alpha=\frac{\lambda_{1}+\lambda_{2}}{2}, \beta=\frac{\lambda_{1}-\lambda_{2}}{2}$, and $b=\lambda_{3}-\alpha$. Then $A=\alpha I+b P+\beta J$, where the matrices

$$
P=\frac{1}{b^{2}-\beta^{2}}\left[(A-\alpha I)^{2}-\beta^{2} I\right], \text { and } J=\frac{1}{\beta}[A-\alpha I-b P]
$$

satisfies the following relations:

$$
P J=J P=0, P^{2}=P, \quad J^{2 n}=I-P, \quad \text { and } J^{2 n-1}=J, n=1,2, \ldots
$$

Consequently, for every $t \in \mathbb{R}$ we have

$$
\begin{gather*}
e^{t A}=e^{\alpha t}[(\cosh \beta t)(I-P)+(\sinh \beta t) J]+e^{\lambda_{3} t} P=  \tag{7}\\
=\frac{1}{2} e^{\lambda_{1} t}(I+J-P)+\frac{1}{2} e^{\lambda_{2} t}(I-J-P)+e^{\lambda_{3} t} P .
\end{gather*}
$$

Proof. We shall consider the cases consecutively:
Case 1. The characteristic polynomial of the matrix $N=A-\lambda_{0} I$ is $f_{N}(\lambda)=\lambda^{3}$, hence $N^{3}=0$, and now (4) is a direct corollary of Lemma 1.1 (i) - (ii).

Case 2. Denote $B=A-\lambda_{0} I$, and observe that

$$
P=\frac{1}{b^{2}} B^{2}, \text { and } N=B-b P=-\frac{1}{b} B(B-b I)
$$

The characteristic polynomial of the matrix $B$ is $f_{B}(\lambda)=\lambda^{2}(\lambda-b)$, and the CayleyHamilton theorem gives $B^{3}=b B^{2}$ and $B^{4}=b^{2} B^{2}$.

Now we have $P^{2}=b^{-4} B^{4}=b^{-2} B^{2}=P$, as required.
Next we calculate $P N=N P=-b^{-3} B^{2} B(B-b I)=-b^{-3} B f_{B}(B)=0$ and finally $N^{2}=-b^{-2} B^{2}(B-b I)^{2}=-b^{-2} f_{B}(B)(B-b I)=0$, as required.

Now as a direct corollary of Lemma 1.1 (i) - (iii) we have

$$
e^{t A}=e^{\lambda_{0} t I} e^{b t P} e^{t N}=e^{\lambda_{0} t}\left[I-P+e^{b t} P\right](I+t N)
$$

and the one easily obtains (5) after standard calculations.

Case 3. Denote $B=A-\alpha I$ and observe that

$$
P=\frac{1}{b^{2}+\omega^{2}}\left(B^{2}+\omega^{2} I\right), \text { and } J=\frac{1}{\omega}(B-b P)
$$

The characteristic polynomial of $B$ is $f_{B}(\lambda)=\left(\lambda^{2}+\omega^{2}\right)(\lambda-b)$ and the CayleyHamilton theorem gives

$$
B^{3}+\omega^{2} B=b\left(B^{2}+\omega^{2} I\right), \text { and } B^{4}+\omega^{2} B^{2}=b^{2}\left(B^{2}+\omega^{2} I\right)
$$

Next we have

$$
P B=B P=\frac{1}{b^{2}+\omega^{2}}\left(B^{3}+\omega^{2} B\right)=\frac{b}{b^{2}+\omega^{2}}\left(B^{2}+\omega^{2} I\right)=b P
$$

which implies $B^{2} P=P B^{2}=b^{2} P$.
Now we check that $P$ is a projector, namely:

$$
P^{2}=\frac{1}{b^{2}+\omega^{2}}\left(B^{2}+\omega^{2} I\right) P==\frac{1}{b^{2}+\omega^{2}}\left(B^{2} P+\omega^{2} P\right)=\frac{1}{b^{2}+\omega^{2}}\left(b^{2} P+\omega^{2} P\right)=P
$$

and also $P J=J P=\omega^{-1}(B-b P) P=\omega^{-1}(b P-b P)=0$.
Finally, we have $J^{2}=\omega^{-2}(B-b P)^{2}=\omega^{-2}\left(B^{2}-2 b B P+b^{2} P\right)=$

$$
\omega^{-2}\left[\left(b^{2}+\omega^{2}\right) P-\omega^{2} I-b^{2} P\right]=-(I-P), \text { as required. }
$$

The rest of the relations for the powers of $J$ follow immediately. At the end as a direct corollary of Lemma 1.1 (i) - (iii) and Lemma 1.2 (i) we have

$$
e^{t A}=e^{\alpha t I} e^{b t P} e^{\omega t J}=e^{\alpha t}\left(I-P+e^{b t} P\right)[P+(\cos \omega t)(I-P)+(\sin \omega t) J]
$$

which easily leads to (6) after standard manipulations.
Case 4. Denote, as before $B=A-\alpha I$, and observe that

$$
P=\frac{1}{b^{2}-\beta^{2}}\left(B^{2}-\beta^{2} I\right), \text { and } J=\frac{1}{\beta}(B-b P) .
$$

In this case the characteristic polynomial of $B$ is $f_{B}(\lambda)=\left(\lambda^{2}-\beta^{2}\right)(\lambda-b)$ and the Cayley-Hamilton theorem gives

$$
B^{3}-\beta^{2} B=b\left(B^{2}-\beta^{2} I\right), \text { and } B^{4}-\beta^{2} B^{2}=b^{2}\left(B^{2}-\beta^{2} I\right)
$$

Now the relations needed between $B, P$ and $J$ follow exactly the same steps as in the Case 3. At the end with the help of Lemma 1.1 (i) - (iii), and Lemma 1.2 (ii) we evaluate:

$$
e^{t A}=e^{\alpha t I} e^{b t P} e^{\omega t J}=e^{\alpha t}\left(I-P+e^{b t} P\right)[P+(\cosh \omega t)(I-P)+(\sinh \omega t) J],
$$

which after standard calculations proves (7).
The formulae (4) - (7) show that one can evaluate the exponential of a $3 \times 3$ matrix without eigenvectors technique. However, the eigenvectors, canonical forms and bases can be easily extracted from the matrices $N, P$ and $J$ introduced above. The next result follows easily from the above considerations.

Corollary 3.2. Under the notations of the previous theorem we have:
Case 1. If $N^{2} \neq 0$, then $E_{\lambda=\lambda_{0}}(A)=\operatorname{Im}\left(N^{2}\right)$. In particular either of the non-zero columns of $N^{2}$ is an eigenvector of A, a Jordan basis can be selected as $\vec{u}, N \vec{u}, N^{2} \vec{u}$, provided that $N^{2} \vec{u} \neq 0$ and $A$ is similar to a Jordan cell of order 3 corresponding to the root $\lambda_{0}$. If $N=0$, then $A=\lambda_{0} I$. If $N^{2}=0$ and $N \neq 0$, then $E_{\lambda=\lambda_{0}} \nsupseteq \operatorname{Im}(N)$. To complete a Jordan basis we select an additional eigenvector by an easy inspection of the columns of $N$.

Case 2. If $N \neq 0$, then $E_{\lambda=\lambda_{0}}(A)=\operatorname{Im}(N)$ and $E_{\lambda=\lambda_{3}}(A)=\operatorname{Im}(P)$. In particular either of the non-zero columns of $N$ and $P$ is an eigenvector of $A$ and a Jordan basis can be selected as $\left(A-\lambda_{0} I\right)\left(A-\lambda_{3} I\right) \vec{u},\left(A-\lambda_{3} I\right) \vec{u}, P \vec{v}$, provided that they are not zero. If $N=0$, then $A$ is diagonalizable.

Case 3. Over the complex numbers the matrix $A$ is diagonalizable with eigenvalues $\lambda_{1,2}=\alpha \pm i \omega$ and $\lambda_{3}$ and eigenvectors $\left(J \mp i J^{2}\right) \vec{u}, P \vec{v}, \vec{u} \neq 0, P \vec{v} \neq 0$. Over the reals the only eigenspace is $E_{\lambda=\lambda_{3}}(A)=\operatorname{Im}(P)$. A canonical form in the basis $J^{2} \vec{u}, J \vec{u}, P \vec{v}$ has a conformal cell like in the Case 2 of Corollary 2.2.

Case 4. Here we have $E_{\lambda=\lambda_{1}}(A)=\operatorname{Im}\left(J+J^{2}\right), E_{\lambda=\lambda_{2}}(A)=\operatorname{Im}\left(J-J^{2}\right)$ and $E_{\lambda=\lambda_{3}}(A)=\operatorname{Im}(P)$. Thus, the matrix $A$ is diagonalizable in the basis consisting of the eigenvectors listed above.

Here, several examples illustrate the results of Theorem 3.1 and Corollary 3.2 in the first three cases (either a multiple, or complex characteristic roots).

Example 4. The matrix $A=\left(\begin{array}{ccc}2 & -1 & 2 \\ 5 & -3 & 3 \\ -1 & 0 & -2\end{array}\right)$ has a multiple characteristic root $\lambda_{1}=\lambda_{2}=\lambda_{3}=-1$ (Case 1).

We evaluate

$$
N=A+I=\left(\begin{array}{ccc}
3 & -1 & 2 \\
5 & -2 & 3 \\
-1 & 0 & -1
\end{array}\right), N^{2}=\left(\begin{array}{ccc}
2 & -1 & 1 \\
2 & -1 & 1 \\
-2 & 1 & -1
\end{array}\right), N^{3}=0
$$

Then according to (4) we have

$$
e^{t A}=e^{\lambda_{0} t}\left(I+t N+\frac{t^{2}}{2} N^{2}\right)=e^{-t}\left(\begin{array}{ccc}
1+3 t+t^{2} & -t-t^{2} / 2 & 2 t+t^{2} / 2 \\
5 t+t^{2} & 1-2 t-t^{2} / 2 & 3 t+t^{2} / 2 \\
-t-t^{2} & t^{2} / 2 & 1-t-t^{2} / 2
\end{array}\right)
$$

In addition, $E_{\lambda=-1}(A)=\operatorname{Im}\left(N^{2}\right)$ and a Jordan basis can be easily selected, say $\vec{u}=(0,0,1), N \vec{u}=(2,3,-1), N^{2} \vec{u}=(1,1,-1)$.

Example 5. The matrix $A=\left(\begin{array}{ccc}1 & -3 & 4 \\ 4 & -7 & 8 \\ 6 & -7 & 7\end{array}\right)$ has a multiple characteristic root $\lambda_{1}=$ $\lambda_{2}=-1=\lambda_{0}$ and $\lambda_{3}=3$ (Case 2), and $b=4$.

We determine the matrices $B=A+I, P=\frac{1}{16} B^{2}$ and $N=B-4 P$ as prescribed:

$$
B=\left(\begin{array}{lll}
2 & -3 & 4 \\
4 & -6 & 8 \\
6 & -7 & 8
\end{array}\right), P=\left(\begin{array}{lll}
1 & -1 & 1 \\
2 & -2 & 2 \\
2 & -2 & 2
\end{array}\right), N=\left(\begin{array}{lll}
-2 & 1 & 0 \\
-4 & 2 & 0 \\
-2 & 1 & 0
\end{array}\right)
$$

Then according to (5) we have

$$
e^{t A}=e^{-t}\left(\begin{array}{ccc}
-2 t & 1+t & -1 \\
-2-4 t & 3+2 t & -2 \\
-2-2 t & 2+t & -1
\end{array}\right)+e^{3 t}\left(\begin{array}{ccc}
1 & -1 & 1 \\
2 & -2 & 2 \\
2 & -2 & 2
\end{array}\right)
$$

In addition, $E_{\lambda=-1}(A)=\operatorname{Im}(N), E_{\lambda=3}(A)=\operatorname{Im}(P)$ and a Jordan basis can be easily selected, say for $\vec{u}=\vec{v}=(1,0,0)$ we get

$$
\left(A-\lambda_{0} I\right)\left(A-\lambda_{3} I\right) \vec{u}=(8,16,8),\left(A-\lambda_{3} I\right) \vec{u}=(-1,2,3), P \vec{v}=(1,2,2)
$$

Example 6. The matrix $\left(\begin{array}{ccc}-2 & 3 & 1 \\ -6 & -1 & 4 \\ -10 & 0 & 7\end{array}\right)$ has characteristic roots $\lambda_{1,2}=1 \pm 2 i$ and $\lambda_{3}=2$ (Case 3), and $b=1$.

We determine the matrices $B=A-I, P=\frac{1}{5}\left(B^{2}+4 I\right)$ and $J=\frac{1}{2}(B-P)$ as prescribed:

$$
B=\left(\begin{array}{ccc}
-3 & 3 & 1 \\
-6 & -2 & 4 \\
-10 & 0 & 6
\end{array}\right), P=\left(\begin{array}{ccc}
-3 & -3 & 3 \\
-2 & -2 & 2 \\
-6 & -6 & 6
\end{array}\right), \quad J=\left(\begin{array}{ccc}
0 & 3 & -1 \\
-2 & 0 & 1 \\
-2 & 3 & 0
\end{array}\right) .
$$

Then according to (6) we have

$$
e^{t A}=e^{t}\left[\cos 2 t\left(\begin{array}{lll}
4 & 3 & -3 \\
2 & 3 & -2 \\
6 & 6 & -5
\end{array}\right)+\sin 2 t\left(\begin{array}{ccc}
0 & 3 & -1 \\
-2 & 0 & 1 \\
-2 & 3 & 0
\end{array}\right)\right]+e^{3 t}\left(\begin{array}{lll}
-3 & -3 & 3 \\
-2 & -2 & 2 \\
-6 & -6 & 6
\end{array}\right) .
$$

In addition a canonical basis can be easily selected, say $J^{2} \vec{u}=(-3,-3,-6), J \vec{u}=$ $(1,0,1), P \vec{v}=(3,2,6)$, for $\vec{u}=(0,1,0)$ and $\vec{v}=(0,0,1)$.
4. Further applications. For a given square matrix $A$ we can define another elementary functions by the Taylor series, e.g. $\sin (A)=\sum_{n=0}^{\infty}(-1)^{n} \frac{A^{2 n+1}}{(2 n+1)!}, \cosh (A)=$ $\sum_{n=0}^{\infty} \frac{A^{2 n}}{(2 n)!}$, etc. If the series is closely related to the Taylor expansion of $e^{t}$, then one can obtain formulae similar to the results of Theorem 2.1 and Theorem 3.1, e.g.

Corollary. Under the notations of the Theorem 2.1 we have:
Case 1. Let $\lambda_{1}=\lambda_{2}=\lambda_{0}$ (it is real), and $N=A-\lambda_{0} I$. Then

$$
\begin{aligned}
\cos A & =\cos \left(\lambda_{0}\right) I-\sin \left(\lambda_{0}\right) N, \\
\sin A & =\sin \left(\lambda_{0}\right) I+\cos \left(\lambda_{0}\right) N .
\end{aligned}
$$

Case 2. Let $\lambda_{1,2}=\alpha \pm i \omega$ for some $\alpha, \omega \in \mathbb{R}, \omega \neq 0$ and $A=\alpha I+\omega J$. Then

$$
\begin{aligned}
\cos A & =\cos (\alpha) \cosh (\omega) I-\sin (\alpha) \sinh (\omega) J \\
\sin A & =\sin (\alpha) \cosh (\omega) I+\cos (\alpha) \sinh (\omega) J
\end{aligned}
$$

Case 3. Let $\lambda_{1} \neq \lambda_{2}$ (both real), and $A=\alpha I+\beta J$, with $\alpha=\frac{\lambda_{1}+\lambda_{2}}{2}, \beta=\frac{\lambda_{1}-\lambda_{2}}{2}$. Then

$$
\begin{aligned}
\cos A & =\cos (\alpha) \cos (\beta) I-\sin (\alpha) \sin (\beta) J, \\
\sin A & =\sin (\alpha) \cos (\beta) I+\cos (\alpha) \sin (\beta) J .
\end{aligned}
$$

## REFERENCES

[1] M. Artin. Algebra, Prentice Hall, Upper Saddle River, New Jersey, 1991
[2] A. I. Malcev. Foundations of Linear Algebra, W. H. Freeman and Co., San FranciscoLondon, 1963.
[3] P. Bamberg, Sh. Sternberg. A course in Mathematics for Students of Physics, University Pres, New York, 1988.
352

A. Popov<br>American University in Bulgaria<br>2700 Blagoevgrad, Bulgaria<br>e-mail: apopov@aubg.bg

T. Todorov<br>Mathematics Department, Cal Poly<br>San Louis Obispo, CA 93407, USA<br>e-mail: ttodorov@calpoly.edu

# ПРЕСМЯТАНЕ НА ЕКСПОНЕНТАТА НА МАТРИЦИ ОТ РЕД 2 И 3 БЕЗ КАНОНИЗАЦИЯ 

## Ангел Попов, Тодор Тодоров

Получени са експлицитни формули за пресмятане на експонентата на дадена матрица А от ред 2 или 3 . Те използуват съществено характеристичните корени на матрицата, без да се търсят собствени вектори, да се конструира каноничен базис и съответната матрица на прехода. Освен това, се получават някои свързани с А матрици, които директно дават каноничен базис и идентифицират някои инвариантни подпространства на А. Според нас този подход има редица предимства пред по-традиционните методи, базирани на конструирането на каноничен базис, особено ако се прилага от непрофесионалисти в математиката. Като приложение на тази техника са приведени няколко примера.


[^0]:    *Key words: exponential of a matrix, characteristic polynomial, Cayley-Hamilton theorem, canonical basis, transition matrix

