Formula for matrix exponential

In this challenge/extra-credit problem, we give a formula for the exponential $\exp(tA)$ for a square matrix $A \in \mathrm{M}_n(\mathbb{C})$, in the generic case when the characteristic polynomial of A has n distinct roots, and the "generic exceptional case" when the characteristic polynomial of A has one double root and $n-2$ simple roots.

This formula must have appeared in the literature, but I have not seen them before. Be sure to tell me if you figure out how to do them.

The interest of these formulas are mostly theoretical, in that you can actually write down a closed formula. It is unclear to me whether they offer any real computational edge for n large (say $n \geq 4$), for either exact or numerical computations. Yes you can copy them to your cheat sheet and use them in exams. I am not sure whether such formulas helps to improve your score.

1. Suppose that $A \in \mathrm{M}_n(\mathbb{C})$ is a square matrix whose characteristic polynomial

$$f(x) := \det(x \cdot \mathrm{I}_n - A)$$

has n *mutually distinct* roots $\lambda_1, \ldots, \lambda_n$: $f(x) = \prod_{i=1}^n (x - \lambda_i)$. Define

$$E(t, x; \lambda_1, \ldots, \lambda_n) := \sum_{i=1}^n e^{\lambda_i t} \cdot \frac{\prod_{1 \leq j \leq n,\, j \neq i}(x - \lambda_j)}{f'(\lambda_i)} = \sum_{i=1}^n e^{\lambda_i t} \cdot \frac{\prod_{1 \leq j \leq n,\, j \neq i}(x - \lambda_j)}{\prod_{1 \leq j \leq n,\, j \neq i}(\lambda_i - \lambda_j)}$$

Note that $E(t, x; \lambda_1, \ldots, \lambda_n)$ can be regarded as a polynomial in x whose coefficients are functions in t, depending on n parameters $\lambda_1, \ldots, \lambda_n$; the parameters are assumed to be mutually distinct. Show that

$$\exp(tA) = E(t, A; \lambda_1, \ldots, \lambda_n) = \sum_{i=1}^n f'(\lambda_i)^{-1} \cdot e^{\lambda_i t} \cdot \prod_{1 \leq j \leq n,\, j \neq i} (A - \lambda_i).$$

Hints: (i) You might want to start with the special cases (1a) and (1b) below.
(ii) The Cayley–Hamilton theorem, which says that $f(A) = \prod_{i=1}^n (A - \lambda_i \cdot \mathrm{I}_n) = 0$, is helpful. One consequence of the Cayley–Hamilton theorem is that each monomial $A^n, A^{n+1}, A^{n+2}, \ldots$ in A can be expressed as a polynomial in A of degree at most $n-1$. Therefore one would expect to get a formula for $\exp(tA)$ as a polynomial in A of degree at most $n-1$ depending on the characteristic polynomial $f(x)$ of A, and the coefficients of this polynomial formula would involve the exponentials $e^{\lambda_i t}$. You want to show that $E(t, x; \lambda_1, \ldots, \lambda_n)$ is the sought-after formula.
(iii) Finally notice that $E(t, x; \lambda_1, \ldots, \lambda_n)$ is the unique polynomial of degree at most $n-1$ in x such that $E(t, \lambda_i; \lambda_1, \ldots, \lambda_n) = e^{\lambda_i t}$ for $i = 1, \ldots, n$.

Illustrations with special cases

(1a) When $n = 2$, the formula says that

$$\exp(tA) = e^{\lambda_1 t} \cdot (\lambda_1 - \lambda_2)^{-1} \cdot (A - \lambda_2) + e^{\lambda_1 t} \cdot (\lambda_1 - \lambda_2)^{-1} \cdot (A - \lambda_2)$$

(1b) When $n = 3$, the formula says that

$$\exp(tA) = e^{\lambda_1 t} \frac{(A - \lambda_2)(A - \lambda_3)}{(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3)} + e^{\lambda_2 t} \frac{(A - \lambda_1)(A - \lambda_3)}{(\lambda_2 - \lambda_1)(\lambda_2 - \lambda_3)} + e^{\lambda_3 t} \frac{(A - \lambda_1)(A - \lambda_2)}{(\lambda_3 - \lambda_1)(\lambda_3 - \lambda_2)}$$

2. In the formula for $\exp(tA)$ in problem 1, take the limit as λ_n goes to λ_{n-1} while $\lambda_1,\dots,\lambda_{n-2}$ remain fixed. to get a formula for $\exp(tA)$ when the characteristic polynomial of A has a root with multiplicity 2 and $n-2$ simple roots. In other words, define a polynomial $E(t,x;\lambda_1,\dots,\lambda_{n-2}\mid\lambda_{n-1},\lambda_{n-1})$ in x of degree $n-1$ by

$$E(t,x;\lambda_1,\dots,\lambda_{n-2}\mid\lambda_{n-1},\lambda_{n-1}) = \lim_{\lambda_n\to\lambda_{n-1}} E(t,x;\lambda_1,\dots,\lambda_n).$$

Let $g(x) := \prod_{i=1}^{n-2}(x-\lambda_i)$.

(2a) Show that $E(t,x;\lambda_1,\dots,\lambda_{n-2}\mid\lambda_{n-1},\lambda_{n-1})$ defined above as a limit is given explicitly by

$$\begin{aligned} E(t,x;\lambda_1,\dots,\lambda_{n-2}\mid\lambda_{n-1},\lambda_{n-1}) &= \sum_{i=1}^{n-2} e^{\lambda_i t}\,\frac{g(x)\,(x-\lambda_{n-1})^2}{g'(\lambda_i)\,(\lambda_i-\lambda_{n-1})^2} \\ &\quad + e^{\lambda_{n-1}t} g(x)\left[\frac{1}{g(\lambda_{n-1})} + \left(\frac{t}{g(\lambda_{n-1})} - \frac{g'(\lambda_{n-1})}{g(\lambda_{n-1})^2}\right)(x-\lambda_{n-1})\right] \end{aligned}$$

(2b) Suppose that the characteristic polynomial for a square matrix $A\in \mathrm{M}_n(\mathbb{C})$ has one double root λ_{n-1} and $n-2$ simple roots $\lambda_1,\dots,\lambda_{n-2}$. Prove that

$$\exp(tA) = E(t,A;\lambda_1,\dots,\lambda_{n-2}\mid\lambda_{n-1},\lambda_{n-1}).$$

Illustration: when $n=3$ we have

$$E(t,x;\lambda_1\mid\lambda_2,\lambda_2) = e^{\lambda_1 t}\,\frac{(x-\lambda_2)^2}{(\lambda_1-\lambda_2)^2} + e^{\lambda_2 t}(x-\lambda_1)\left[\left(\frac{t}{\lambda_2-\lambda_1} - \frac{1}{(\lambda_2-\lambda_1)^2}\right)(x-\lambda_2)\right]$$

Remark. From an algebraic point of view, what underlies the formulas for $E(t,x;\lambda_1,\dots,\lambda_n)$ and $E(t,x;\lambda_1,\dots,\lambda_{n-2}\mid\lambda_{n-1},\lambda_{n-1})$ in problems 1 and 2 is the partial fraction decomposition. For instance

$$\frac{1}{g(x)(x-\lambda_{n-1}^2)} = \sum_{i=1}^{n-2}\frac{1}{g'(\lambda_i)(x-\lambda_i)} - \frac{g'(\lambda_{n-1})}{g(\lambda_{n-1})^2(x-\lambda_{n-1})} + \frac{1}{g(\lambda_{n-1})(x-\lambda_{n-1})^2}$$

Clearing the denominator, we get

$$1 = \sum_{i=1}^{n-2}\frac{g(x)(x-\lambda_{n-1})^2}{g'(\lambda_i)(x-\lambda_i)} + \left[g(x)\cdot\left(-\frac{g'(\lambda_{n-1})}{g(\lambda_{n-1})^2}(x-\lambda_{n-1}) + \frac{1}{g(\lambda_{n-1})}\right)\right]$$

The right hand side of the above equality is a sum of $n-1$ polynomials, with the last one grouped under a pair of square brackets. Let's call them $p_1(x),\dots,p_{n-2}(x),p_{n-1}(x)$. The $n-1$ matrices $p_1(A),\dots,p_{n-2}(A),p_{n-1}(A)$ are the "projections" to the one-dimensional eigenspace for the simple eigenvalues $\lambda_1,\dots,\lambda_{n-2}$ and the two-dimensional generalized eigenspace for λ_{n-1}. Clearly the sum of these matrices $p_j(A)$ is I_n. The Cayley–Hamilton theorem implies that $p_i(A)p_j(A)=0$ if $i\neq j$, and $p_j(A)^2 = p_j(A)^2$ for all $j=1,\dots,n-2$. The image (range) of $p_1(A),\dots,p_{n-2}(A)$ are the one-dimensional eigenspaces for $\lambda_1,\dots,\lambda_{n-2}$, and the image of p_{n-1} is the generalized eigenspace corresponding to the double root λ_{n-1} of $\det(x\cdot\mathrm{I}_n-A)$. In the next problem we give an algorithm for computing $\exp(tA)$ in all cases based on the knowledge of all roots of the characteristic polynomial of $\exp(tA)$. (Of course finding all roots of $\det(x\cdot\mathrm{I}_n-A)$ is arguably the hardest part for computing either the exponential $\exp(tA)$ or the Jordan form of A for a "randomly picked" A.)

3. We first describe a general algorithm/procedure.

- The input is a polynomial $f(x)$ with coefficients in $\mathbb{C}$ already factored in the form

$$f(x) = \prod_{i=1}^{r} (x-\lambda_i)^{e_i} = (x-\lambda_1)^{e_1} \cdots (x-\lambda_r)^{e_r},$$

  where $\lambda_1, \ldots, \lambda_r$ are mutually distinct complex numbers and $e_1, \ldots, e_r \geq 1$ are positive integers. Let $n := e_1 + \ldots + e_r = \deg(f(x))$

- The output is a function $E(t,x;\lambda_1,\ldots,\lambda_r;e_1,\ldots,e_r)$ in variables t and x, which depends on the paratmeters $\lambda_1,\ldots,\lambda_r$ and $e_1,\ldots,e_r$. It will be a polynomial in x of degree $\leq n-1$ whose coefficents are functions which can be written as polynomials in $t, \lambda_1, \ldots, \lambda_r$ and $e^{\lambda_1 t}, \ldots, e^{\lambda_r t}$.

Algorithmic definition of $E(t,x;\lambda_1,\ldots,\lambda_r;e_1,\ldots,e_r)$.

Step 1. Compute the partial fraction decomposition of the rational function $\frac{1}{f(x)}$:

$$\frac{1}{f(x)} = \sum_{i=1}^{r} \frac{h_i(x)}{(x-\lambda_i)^{e_i}},$$

where $h_1(x), \ldots, h_r(x)$ are polynomials with $\deg(h_i(x)) \leq e_i - 1$ for each $i = 1, \ldots, r$.

One way to compute the $h_i(x)$'s is as follows. For each i let

$$g_i(x) := \frac{f(x)}{(x-\lambda_i)^{e_i}},$$

a polynomial in x which does not vanish at λ_i. Then

$$h_i(x) = a_{i,e_{i-1}}(x-\lambda_i)^{i-1} + \cdots + a_{i,e_1}(x-\lambda_i) + a_{i,0}\,,$$

where

$$a_{i,0} = \frac{1}{g_i(\lambda_i)},\ a_{i,1} = \frac{d}{dx}\left(\frac{1}{g_i(x)}\right)\bigg|_{x=\lambda_i},\ a_{i,2} = \frac{1}{2!}\,\frac{d^2}{dx^2}\left(\frac{1}{g_i(x)}\right)\bigg|_{x=\lambda_i},$$
$$a_{i,3} = \frac{1}{3!}\,\frac{d^3}{dx^3}\left(\frac{1}{g_i(x)}\right)\bigg|_{x=\lambda_i}, \ldots,\ a_{i,e_{i-1}} = \frac{1}{e_{i-1}!}\,\frac{d^{e_i-1}}{dx^{e_{i-1}}}\left(\frac{1}{g_i(x)}\right)\bigg|_{x=\lambda_i}$$

Step 2. Define/compute polynomials $p_1(x), \ldots, p_r(x)$ of degree $\leq n-1$, given by

$$p_i(x) := \frac{f(x)}{(x-\lambda_i)^{e_i}} \cdot h_i(x) = g_i(x) \cdot h_i(x).$$

Define/compute the functions $E_i(t,x;\lambda_i,e_i)$ for $i = 1, \ldots, r$ given by

$$E_i(t,x;\lambda_i,e_i) := e^{\lambda_i t} p_i(x) \left(1 + t(x-\lambda_i) + \frac{t^2(x-\lambda_i)^2}{2!} + \cdots + \frac{t^{e_i-1}(x-\lambda_i)^{e_i-1}}{(e_i-1)!}\right)$$

(We will think about $E_i(t,x;\lambda_i,e_i)$ as a function in two variables (t,x) depending on two parameters λ_i, e_i. It is a polynomial in x of degree $\leq e_i - 1$ whose coefficients are simple expressions involving $e^{\lambda_i t}$ and λ_i.)

Step 3. The sum of the $E_i(t,x;\lambda_i,e_i)$'s is the formula $E(t,x;\lambda_1,\dots,\lambda_r;e_1,\dots,e_r)$ we want:

$$E(t,x;\lambda_1,\dots,\lambda_r;e_1,\dots,e_r) = E_1(t,x;\lambda_1,e_1) + \cdots + E_r(t,x;\lambda_r,e_r)$$

Comments: In step 1 we gave a formula for the partial fraction decomposition of $\frac{1}{f(x)}$. An alternative way is to think of the coefficients of the polynomial $h_i(x)$ of degree $\leq e_i - 1$ unknowns, so altogether you have $e_1 + \cdots + e_r = n$ unknowns. Multiplying the first equality in step 1 by $f(x)$ and equate the coefficients of $1, x, x^2, \dots, x^{n-1}$ on both sides, you get d *linear* equations in the n unknowns, which is easily solved by elimination.

Challenge problem 3. Prove that for any square matrix $A \in \mathrm{M}_n(\mathbb{C})$ such that

$$\det(x \cdot \mathrm{I}_n - A) = \prod_{i=1}^{r} (x - \lambda_i)^{e_i}$$

where $\lambda_1, \dots, \lambda_r$ are mutually distinct complex numbers and $e_1, \dots, e_r$ are positive integers, we have

$$\exp(tA) = E(t,A;\lambda_1,\dots,\lambda_r;e_1,\dots,e_r).$$

Comments and hints.

(i) The formulas in problems 1 and 2 are both special cases of the formula in problem 3.

(ii) Let $P_i = p_i(A)$ for $i = 1, \ldots, n$, where $p_i(x)$ is the polynomial defined in Step 2. It is not difficulty to see from the Cayley–Hamilton theorem that

$$P_1 + \cdots P_r = \mathrm{I}_n, \quad P_i^2 = P_i \text{ for } i = 1, \ldots, r, \quad P_i \cdot P_j = 0 \text{ if } i \neq j.$$

(iii) The column span of the matrix P_i is the vector subspace V_i of $\mathbb{C}^n_{\mathrm{col}}$ consisting of all generalized eigenvectors of A associated with the eigenvalue λ_i; $\dim_{\mathbb{C}}(V_i) = e_i$ for each i. We know that $P_i^{e_i} \cdot v = \vec{0}$ for every vector $v \in V_i$.