MAU23205 Lecture 15

John Stalker

Trinity College Dublin

15 October 2021

Proof of the formula for $\exp(tA)$, (1/4)

Recall that the companion matrix to a monic polynomial $p(z) = \sum_{i=0}^{m} \alpha_i z^i$, $\alpha_m = 1$ is C, with entries

$$c_{j,k} = \begin{cases} 0 & \text{if } j < m, k \neq j+1, \\ 1 & \text{if } j < m, k = j+1, \\ -\alpha_{k-1} & \text{if } j = m. \end{cases}$$

If x satisfies the differential equation

$$\sum_{i=0}^{m} \alpha_i x^{(i)} = 0$$

then

$$x^{(j)}(t) = \sum_{k=1}^{m} c_{j,k} x^{(k-1)}(t).$$

Proof of the formula for $\exp(tA)$, (2/4)

This holds for any solution, but in particular for the basic solutions x_l, so

$$x_l^{(j)}(t) = \sum_{k=1}^{m} c_{j,k} x_l^{(k-1)}(t).$$

If we let $y_{j,k} = x_k^{(j-1)}$ then this equation is

$$y_{j,l}'(t) = \sum_{k=1}^{m} c_{j,k} y_{k,l}(t).$$

In matrix form $Y'(t) = CY(t)$.
We saw in Lecture 12 that there is also a matrix D such that $Y'(t) = Y(t)D$. The precise form of D was described there, but here we only care that it exists.

Proof of the formula for $\exp(tA)$, (3/4)

If $p(A) = O$ then

$$A^j = \sum_{k=1}^{m} c_{j,k} A^{k-1}.$$

Suppose $\mathbf{v}$ is a row vector. We'll choose a particular $\mathbf{v}$ later. Define

$$\mathbf{u}(t) = \mathbf{v} Y(t) Y(0)^{-1} \qquad W(t) = \sum_{k=1}^{m} u_k(t) A^{k-1}.$$

$$\mathbf{u}'(t) = \mathbf{v} Y'(t) Y(0)^{-1} = \mathbf{v} Y(t) D Y(0)^{-1} = \mathbf{v} Y(t) Y(0)^{-1} C = \mathbf{u}(t) C.$$

$$\begin{aligned} W'(t) &= \sum_{k=1}^{m} u_k'(t) A^{k-1} = \sum_{k=1}^{m} \sum_{j=1}^{m} u_j(t) c_{j,k} A^{k-1} \\ &= \sum_{j=1}^{m} u_j(t) A^j = \sum_{j=1}^{m} u_j(t) A^{j-1} A = W(t) A. \end{aligned}$$

Proof of the formula for $\exp(tA)$, (4/4)

$$\frac{d}{dt}\left(W(t)\exp(-tA)\right) = \left(W(t)A\right)\exp(-tA) + W(t)\left(-A\exp(-tA)\right) = O.$$

So $W(t)\exp(-tA) = W(0)$ and $W(t) = W(0)\exp(tA)$. Choose $v_k = 1$ if $k = 1$ and $v_k = 0$ if $k > 1$. $\mathbf{u}(0) = \mathbf{v}Y(0)Y(0)^{-1} = \mathbf{v}$ and

$$W(0) = \sum_{k=1}^{m} u_k(0)A^{k-1} = \sum_{k=1}^{m} v_k A^{k-1} = I.$$

So $W(t) = \exp(tA)$. $\mathbf{u}(t) = \mathbf{v}Y(t)R$ so

$$u_k(t) = \sum_{i=1}^{m}\sum_{j=1}^{m} v_i x_j^{(i-1)}(t) r_{j,k} = \sum_{j=1}^{m} x_j(t) r_{j,k}.$$

$$\exp(tA) = W(t) = \sum_{k=1}^{m} u_k(t)A^{k-1} = \sum_{k=1}^{m}\sum_{j=1}^{m} x_j(t) r_{j,k} A^{k-1}.$$

Minimal vs characteristic polynomials

The formula

$$\exp(tA) = \sum_{k=1}^{m} \sum_{j=1}^{m} x_j(t) r_{j,k} A^{k-1}$$

computes $\exp(tA)$ from a monic polynomial p such that $p(A) = O$. p could be the characteristic polynomial, but it doesn't have to be.

If A is normal, i.e. if $A^T A = AA^T$, then the minimal polynomial has no repeated roots. Symmetric, antisymmetric and orthogonal matrices are all normal. If we use the minimal polynomial instead of the characteristic polynomial then

- we can avoid positive powers of t in our basic solutions, and
- all of our matrices, except A, are $m \times m$ rather than $n \times n$, where m, the number of distinct roots of p_A, is possibly smaller than n, the degree of p_A.

An example (1/2)

$$A = \begin{bmatrix} 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & 1 \end{bmatrix} \qquad A^T = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 \end{bmatrix}$$

$A^T A = 4I = AA^T$ so A is normal. Its characteristic polynomial is $p_A(z) = (z^2 - 2z + 4)^2$. Its minimal polynomial is $z^2 - 2z + 4$. It's easier to compute that directly, not via the characteristic polynomial.

$$A^2 = \begin{bmatrix} -2 & -2 & -2 & -2 \\ 2 & -2 & -2 & 2 \\ 2 & 2 & -2 & -2 \\ 2 & -2 & 2 & -2 \end{bmatrix} = 2A - 4I.$$

An example (2/2)

The complex roots of $z^2 - 2z + 4$ are $1 + i\sqrt{3}$ and $1 - i\sqrt{3}$. The basic solutions are $x_1(t) = \exp(t)\cos(\sqrt{3}t)$ and $x_2(t) = \exp(t)\sin(\sqrt{3}t)$.

$$Y(0) = \begin{bmatrix} 1 & 0 \\ 1 & \sqrt{3} \end{bmatrix} \qquad Y(0)^{-1} = \begin{bmatrix} 1 & 0 \\ -\frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} \end{bmatrix}$$

$$\begin{aligned} \exp(tA) = \exp(t)\cos(\sqrt{3}t)I - \frac{\sqrt{3}}{3}\exp(t)\sin(\sqrt{3}t)I \\ + \frac{\sqrt{3}}{3}\exp(t)\sin(\sqrt{3}t)A. \end{aligned}$$

Various useful facts about matrix exponentials

$\exp(t\Lambda + tN) = \exp(t\Lambda)\exp(tN) = \exp(tN)\exp(t\Lambda)$ if $\Lambda N = N\Lambda$. I proved this in Lecture 14. If V is invertible then $\exp(tA) = V\exp(tV^{-1}AV)V^{-1}$.

$$\exp(tV^{-1}AV) = \sum_{k=0}^{\infty}\frac{t^k}{k!}(V^{-1}AV)^k = \sum_{k=0}^{\infty}\frac{t^k}{k!}V{-}1A^kV = V^{-1}\exp(tA)V.$$

$$\exp\left(t\begin{bmatrix}\lambda_1 & 0 & \dots & 0\\ 0 & \lambda_2 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & \lambda_n\end{bmatrix}\right) = \begin{bmatrix}\exp(\lambda_1 t) & 0 & \dots & 0\\ 0 & \exp(\lambda_2 t) & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & \exp(\lambda_n t)\end{bmatrix}.$$

If $N^k = 0$ then $\exp(tN) = \sum_{j=0}^{k-1}\frac{t^j}{j!}A^j$. For any real or complex square matrix A there are complex matrices V, N and Λ and a positive integer k such that $N\Lambda = \Lambda N$, V is invertible, Λ is diagonal and $V^{-1}AV = \Lambda + N$, $N^k = 0$. This is essentially the Jordan Normal Form Theorem. You can arrange that N is strictly triangular.

Computing matrices using Jordan normal forms

You can compute matrix exponentials by this method. To compute $\exp(tA)$ you find V, Λ, N and k as on the previous slide. In principle you learn how to do that in Linear Algebra. Then

$$\begin{aligned}
\exp(tA) &= V\exp(tV^{-1}AV)V^{-1} = V\exp(t\Lambda + tN)V^{-1} \\
&= V\exp(tN)\exp(t\Lambda)V^{-1} = \sum_{j=0}^{k-1}\frac{t^j}{j!}VN^k\exp(t\Lambda)V^{-1} \\
&= \sum_{j=0}^{k-1}\frac{t^j}{j!}VN^k\begin{bmatrix} \exp(\lambda_1 t) & 0 & \dots & 0 \\ 0 & \exp(\lambda_2 t) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \exp(\lambda_n t) \end{bmatrix} V^{-1}.
\end{aligned}$$

This is, in fact, the method of matrix exponentiation that's usually taught.

Comparing the methods

Advantages of Jordan form method:

- It's easy to describe.
- It's easy to prove that it works.
- It gives a quick proof that $\exp(tA)$ is a sum of powers of constant matrices times powers of t times $\exp(\lambda t)$, where λ is an eigenvalue of A, i.e. a root of p_A.

Advantages of $\exp(tA) = \sum r_{j,k} x_j(t) A^{k-1}$:

- The computations are easier, particularly if there are repeated or complex roots, or if A is normal. If you try the Jordan form method on an exam you will probably get it wrong.
- It works better numerically. The Jordan form method doesn't really work at all numerically if you have repeated roots. If finding the roots numerically suggests two roots are nearly equal then assume they are equal.