Understanding the Matrix Exponential

Lecture 8
Math 634
9/17/99

Transformations

Now that we have a representation of the solution of constant-coefficient initial-value problems, we should ask ourselves: "What good is it?" Does the power series formula for the matrix exponential provide an efficient means for calculating exact solutions? Not usually. Is it an efficient way to compute accurate numerical approximations to the matrix exponential? Not according to *Matrix Computations* by Golub and Van Loan. Does it provide insight into how solutions behave? It is not clear that it does. There are, however, transformations that may help us handle these problems.

Suppose that $B, P \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ are related by a similarity transformation; *i.e.*, $B = QPQ^{-1}$ for some invertible Q. Calculating, we find that

$$\begin{aligned} e^B = \sum_{k=0}^{\infty} \frac{B^k}{k!} = \sum_{k=0}^{\infty} \frac{(QPQ^{-1})^k}{k!} &= \sum_{k=0}^{\infty} \frac{QP^kQ^{-1}}{k!} \\ &= Q\left(\sum_{k=0}^{\infty} \frac{P^k}{k!}\right)Q^{-1} = Qe^PQ^{-1}. \end{aligned}$$

It would be nice if, given B, we could choose Q so that P were a diagonal matrix, since

$$e^{\operatorname{diag}\{p_1,p_2,\dots,p_n\}} = \operatorname{diag}\{e^{p_1}, e^{p_2}, \dots, e^{p_n}\}.$$

Unfortunately, this cannot always be done. Over the next few lectures, we will show that what can be done, in general, is to pick Q so that $P = S + N$, where S is a *semisimple* matrix with a fairly simple form, N is a *nilpotent* matrix of a fairly simple form, and S and N commute. (Recall that a matrix is semisimple if it is diagonalizable over the complex numbers and that a matrix is nilpotent if some power of the matrix is 0.) The forms of S and N are simple enough that we can calculate their exponentials fairly easily, and then we can multiply them to get the exponential of $S + N$.

We will spend a significant amount of time carrying out the project described in the previous paragraph, even though it is linear algebra that some of you have probably seen before. Since understanding the behavior of constant coefficient systems plays a vital role in helping us understand more complicated systems, I feel that the time investment is worth it. The particular approach we will take follows chapters 3, 4, 5, and 6, and appendix 3 of Hirsch and Smale fairly closely.

Eigensystems

Given $B \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$, recall that that $\lambda \in \mathbb{C}$ is an *eigenvalue* of B if $Bx = \lambda x$ for some nonzero $x \in \mathbb{R}^n$ or if $\tilde{B}x = \lambda x$ for some nonzero $x \in \mathbb{C}^n$, where $\tilde{B}$ is the *complexification* of B; *i.e.*, the element of $\mathcal{L}(\mathbb{C}^n, \mathbb{C}^n)$ which agrees with B on $\mathbb{R}^n$. (Just as we often identify a linear operator with a matrix representation of it, we will usually not make a distinction between an operator on a real vector space and its complexification.) A nonzero vector x for which $Bx = \lambda x$ for some scalar λ is an *eigenvector*. An eigenvalue λ with corresponding eigenvector x form an *eigenpair* (λ, x).

If an operator $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ is chosen at random, it would almost surely have n distinct eigenvalues $\{\lambda_1, \dots, \lambda_n\}$ and n corresponding linearly independent eigenvectors $\{x_1, \dots, x_n\}$. If this is the case, then A is similar to the (possibly complex) diagonal matrix

$$\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}.$$

More specifically,

$$A = \begin{bmatrix} \\ x_1 & \cdots & x_n \\ \\ \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} \cdot \begin{bmatrix} \\ x_1 & \cdots & x_n \\ \\ \end{bmatrix}^{-1}.$$

If the eigenvalues of A are real and distinct, then this means that

$$tA = \begin{bmatrix} \\ x_1 & \cdots & x_n \\ \\ \end{bmatrix} \cdot \begin{bmatrix} t\lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & t\lambda_n \end{bmatrix} \cdot \begin{bmatrix} \\ x_1 & \cdots & x_n \\ \\ \end{bmatrix}^{-1},$$

and the formula for the matrix exponential then yields

$$e^{tA} = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \cdot \begin{bmatrix} e^{t\lambda_1} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & e^{t\lambda_n} \end{bmatrix} \cdot \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}^{-1}.$$

This formula should make clear how the projections of $e^{tA}x_0$ grow or decay as $t \to \pm\infty$.

The same sort of analysis works when the eigenvalues are (nontrivially) complex, but the resulting formula is not as enlightening. In addition to the difficulty of a complex change of basis, the behavior of $e^{t\lambda_k}$ is less clear when λ_k is not real.

One way around this is the following. Sort the eigenvalues (and eigenvectors) of A so that complex conjugate eigenvalues $\{\lambda_1, \overline{\lambda}_1, \dots, \lambda_m, \overline{\lambda}_m\}$ come first and are grouped together and so that real eigenvalues $\{\lambda_{m+1}, \dots, \lambda_r\}$ come last. For $k \le m$, set $a_k = \operatorname{Re}\lambda_k \in \mathbb{R}$, $b_k = \operatorname{Im}\lambda_k \in \mathbb{R}$, $y_k = \operatorname{Re} x_k \in \mathbb{R}^n$, and $z_k = \operatorname{Im} x_k \in \mathbb{R}^n$. Then

$$\begin{aligned} Ay_k &= A\operatorname{Re} x_k = \operatorname{Re} Ax_k = \operatorname{Re}\lambda_k x_k = (\operatorname{Re}\lambda_k)(\operatorname{Re} x_k) - (\operatorname{Im}\lambda_k)(\operatorname{Im} x_k) \\ &= a_k y_k - b_k z_k, \end{aligned}$$

and

$$\begin{aligned} Az_k &= A\operatorname{Im} x_k = \operatorname{Im} Ax_k = \operatorname{Im}\lambda_k x_k = (\operatorname{Im}\lambda_k)(\operatorname{Re} x_k) + (\operatorname{Re}\lambda_k)(\operatorname{Im} x_k) \\ &= b_k y_k + a_k z_k. \end{aligned}$$

Using these facts, we have $A = QPQ^{-1}$, where

$$Q = \begin{bmatrix} z_1 & y_1 & \cdots & \cdots & z_m & y_m & x_{m+1} & \cdots & x_r \end{bmatrix}$$

and

$$P = \left[\begin{array}{cc|cc|cc|cc||cccc}
a_1 & -b_1 & 0 & 0 & \cdots & \cdots & 0 & 0 & 0 & \ldots & \cdots & 0 \\
b_1 & a_1 & 0 & 0 & \cdots & \cdots & 0 & 0 & 0 & \ldots & \cdots & 0 \\
\hline
0 & 0 & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\hline
\vdots & \vdots & \ddots & \ddots & \ddots & \ddots & 0 & 0 & \vdots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \ddots & \ddots & \ddots & \ddots & 0 & 0 & \vdots & \vdots & \vdots & \vdots \\
\hline
0 & 0 & \cdots & \cdots & 0 & 0 & a_m & -b_m & 0 & \cdots & \cdots & 0 \\
0 & 0 & \cdots & \cdots & 0 & 0 & b_m & a_m & 0 & \cdots & \cdots & 0 \\
\hline\hline
0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 & \lambda_{m+1} & 0 & \cdots & 0 \\
\vdots & \vdots & \cdots & \cdots & \cdots & \cdots & \vdots & \vdots & 0 & \ddots & \ddots & \vdots \\
\vdots & \vdots & \cdots & \cdots & \cdots & \cdots & \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\
0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 & 0 & \cdots & 0 & \lambda_r
\end{array}\right].$$

In order to compute e^{tA} from this formula, we'll need to know how to compute e^{tA_k}, where

$$A_k = \begin{bmatrix} a_k & -b_k \\ b_k & a_k \end{bmatrix}.$$

This can be done using the power series formula. An alternative approach is to realize that

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} := e^{tA_k} \begin{bmatrix} c \\ d \end{bmatrix}$$

is supposed to solve the IVP

$$\begin{cases} \dot{x} = a_k x - b_k y \\ \dot{y} = b_k x + a_k y \\ x(0) = c \\ y(0) = d. \end{cases} \tag{1}$$

Since we can check that the solution of (1) is

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} e^{a_k t}(c \cos b_k t - d \sin b_k t) \\ e^{a_k t}(d \cos b_k t + c \sin b_k t) \end{bmatrix},$$

we can conclude that

$$e^{tA_k} = \begin{bmatrix} e^{a_k t}\cos b_k t & -e^{a_k t}\sin b_k t \\ e^{a_k t}\sin b_k t & e^{a_k t}\cos b_k t \end{bmatrix}$$

Putting this all together and using the form of P, we see that $e^{tA} = Qe^{tP}Q^{-1}$, where

$$e^{tP} = \left[\begin{array}{c||c} \mathcal{B}_1 & \mathbf{0} \\ \hline\hline \mathbf{0}^T & \mathcal{B}_2 \end{array}\right],$$

$\mathcal{B}_1 =$

$$\left[\begin{array}{cc|cc|cc|cc} e^{a_1 t}\cos b_1 t & -e^{a_1 t}\sin b_1 t & 0 & 0 & \cdots & \cdots & 0 & 0 \\ e^{a_1 t}\sin b_1 t & e^{a_1 t}\cos b_1 t & 0 & 0 & \cdots & \cdots & 0 & 0 \\ \hline 0 & 0 & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \hline \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & 0 & 0 \\ \hline 0 & 0 & \cdots & \cdots & 0 & 0 & e^{a_m t}\cos b_m t & -e^{a_m t}\sin b_m t \\ 0 & 0 & \cdots & \cdots & 0 & 0 & e^{a_m t}\sin b_m t & e^{a_m t}\cos b_m t \end{array}\right],$$

$\mathcal{B}_2 =$

$$\begin{bmatrix} e^{\lambda_{m+1} t} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & e^{\lambda_r t} \end{bmatrix},$$

and $\mathbf{0}$ is a $2m \times (r-m-1)$ block of 0's.

This representation of e^{tA} shows that not only may the projections of $e^{tA}x_0$ grow or decay exponentially, they may also exhibit sinusoidal oscillatory behavior.