18.C06 Linear Algebra and Optimization
MIT Fall 2024

Recitation 14

Tuesday Oct 29, 2024

1 Recap

1.1 ODEs and Matrix Exponenentials $\subset$ 18.03

A system of linear ordinary differential equations (ODEs) is described by $dx/dt = Ax$. With an initial value $x(0)$, the solution $x(t)$ is

$$x(t) = e^{At}x(0) = X \underbrace{\begin{pmatrix} e^{\lambda_1 t} & & & \\ & e^{\lambda_2 t} & & \\ & & \ddots & \\ & & & e^{\lambda_m t} \end{pmatrix}}_{e^{\Lambda t}} X^{-1}x(0)\,,$$

where $e^B = I + B + B^2/2 + \cdots + B^n/n! + \cdots$ and the latter equation assumes that A is diagonalizable ($A = X\Lambda X^{-1}$). That is, we first expand $x(0)$ in the basis of eigenvectors to get coefficients $X^{-1}x(0)$, then multiply each coefficient c_k by $e^{\lambda_k t}$, then multiply by the eigenvectors and add them up:

$$x(t) = c_1 e^{\lambda_1 t}x_1 + c_2 e^{\lambda_2 t}x_2 + \cdots + c_m e^{\lambda_m t}x_m$$

Hence:

- A term is exponentially *decaying* if $\operatorname{Re}\lambda_k < 0$, and *growing* if $\operatorname{Re}\lambda_k > 0$. (If *all* the eigenvalues have negative real parts, then $x(t)$ will always decay exponentially to zero for any initial condition.)
- The *dominant* term is the one where λ_k has the *largest real part* (*not* the largest magnitude, unlike matrix powers). For large t, $x(t)$ can be approximated by this dominant term (or terms), unless that coefficient happens to be exactly zero.
- A nonzero imaginary part $\operatorname{Im}\lambda_k \neq 0$ leads to *oscillating* solutions (possibly also decaying/growing depending on the real part).
- $\lambda_k = 0$ corresponds to a *constant* term. If some $\lambda_k = 0$ and all other eigenvalues have negative real parts, then $x(t)$ goes exponentially to this *steady state.*

(Contrast with matrix powers and linear recurrences.)

1.2 Markov matrices

A Markov matrix M is an $m \times m$ matrix with entries ≥ 0 and for which each column sums to 1. If we let $o \in \mathbb{R}^m$ be the vector of all 1's, then the latter property corresponds to $o^T M = o^T$. It follows that

- $o^T(M^n x) = o^T x$: the sum of the entries $o^T x$ is "conserved" by a Markov process (multiplying by M over and over).
- Every eigenvalue must have $|\lambda| \leq 1$.
- There is an eigenvalue $\lambda = 1$.

If the Markov matrix (or some power thereof) additionally has *all positive entries* (> 0, not just ≥ 0), then all of the eigenvalues have $|\lambda| < 1$ except for a *single* (multiplicity 1) eigenvalue $\lambda = 1$. That means $M^n x$ for any initial vector x must asymptotically go to an "steady state" eigenvector of $\lambda = 1$ for $n \to \infty$.

1.3 Hermitian and real-symmetric matrices

For complex matrices and vectors, we **replace transposes with conjugate-transposes ("adjoints")** everywhere in linear algebra (inner products, norms, ...). A matrix with $\overline{A^T} = A$ is called **Hermitian**. For a *real* Hermitian matrix, we can omit the conjugation and get $A^T = A$, a **real-symmetric matrix**. Any Hermitian matrix satisfies:

- It is *always* diagonalizable (even if there are repeated eigenvalues), i.e. there is always a basis of eigenvectors.
- The eigenvalues are always *real*. If A is real, then the eigenvectors are also real.
- One can find an *orthonormal* basis of eigenvectors $X = Q$. (Eigenvectors of distinct λ are always orthogonal, whereas if $N(A - \lambda I)$ has dimension > 1 then we can choose an orthonormal basis of eigenvectors for that eigenvalue.) Hence $A = Q\Lambda\overline{Q^T}$ (and you can omit the conjugate if A is real since in that case Q is a real orthogonal matrix).

This makes eigenvectors particularly useful for Hermitian/real-symmetric matrices. To express any vector x in the eigenvector basis, you can just take dot products with the eigenvectors Q!

2 Exercises

1. Indicate whether the following statements are True or False.
   (a) If a matrix A is symmetric and invertible, so is A^{-1}.
   (b) For any matrix A, the matrix AA^T is symmetric.

2. A is a real 3×3 matrix. The matrix $B = A + A^T$ has eigenvalues $\lambda_1 = 2$, $\lambda_2 = 0$, and $\lambda_3 = 1$, with corresponding eigenvectors $x_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$, $x_2 = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$, and $x_3 = \begin{pmatrix} 1 \\ 2 \\ -5 \end{pmatrix}$.
   (a) What is the matrix e^B? You can leave your answer as a product of several matrices, as long as you write down each matrix explicitly.
   (b) Let $C = (I - B)(I + B)^{-1}$.
      i. What are the eigenvalues of C? (Not much calculation is needed!)
      ii. Suppose that we compute
      $$y = C^{100} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$
      Give a good approximation for the vector y in terms of a single eigenvector.

3. Say A is a 3×3 real matrix with eigenvalues $\lambda_1 = -1, \lambda_2 = -3 + 4i, \lambda_3 = -3 - 4i$, with corresponding eigenvectors x_1, x_2, x_3. Two eigenvectors of A are $x_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ and $x_2 = \begin{pmatrix} 0 \\ 1 \\ i \end{pmatrix}$.
   (a) What is x_3?
   (b) Which of the following almost certainly has unbounded magnitude (i.e. magnitude blowing up) as $n \to \infty$ or $t \to \infty$? Assume y is chosen at random (e.g. with `randn(3)`).
      a) $A^n y$ as $n \to \infty$
      b) $A^{-n} y$ as $n \to \infty$
      c) The solution of $\frac{dx}{dt} = Ax$ as $t \to \infty$ for the initial condition $x(0) = y$.
      d) The solution of $\frac{dx}{dt} = -Ax$ as $t \to \infty$ for the initial condition $x(0) = y$.
   (c) Write down the exact solution $x(t)$ to $\frac{dx}{dt} = Ax$ for the initial condition $x(0) = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$.

3 Solutions

1. (a) True. We already stated (proved?) in class that $(A^{-1})^T = (A^T)^{-1}$, so the result follows. Since, A is a symmetric matrix, so $A^T = A$, and thus $(A^{-1})^T = (A^T)^{-1} = A^{-1}$ is symmetric.

   To prove this identity explicitly, note that $AA^{-1} = I$, and by transposing both sides, we have
   $$(AA^{-1})^T = I \Rightarrow (A^{-1})^T A^T = I.$$
   Therefore, the inverse of A^T is $(A^{-1})^T$.

   (b) True. $(AA^T)^T = (A^T)^T A^T = AA^T$. Therefore, AA^T is symmetric. (Similarly for $A^T A$.)

2. (a) Here, B is obviously diagonalizable ($B = X\Lambda X^{-1}$) because 3 independent eigenvectors are given, but in fact this must be the case since B is real-symmetric! The matrix exponential is then given by $e^B = Xe^\Lambda X^{-1}$, where
   $$\boxed{e^\Lambda = \begin{pmatrix} e^2 & & \\ & 1 & \\ & & e \end{pmatrix}}.$$
   X is a matrix whose columns are the corresponding eigenvectors. However, since B is a real symmetric matrix, it has orthogonal eigenvectors. We can therefore simply normalize each of the eigenvectors to obtain an orthonormal basis:
   $$q_1 = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \ q_2 = \frac{1}{\sqrt{5}}\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \ q_3 = \frac{1}{\sqrt{30}}\begin{pmatrix} 1 \\ 2 \\ -5 \end{pmatrix}.$$
   Then we have $\boxed{e^B = Qe^\Lambda Q^{-1}}$, where $Q^{-1} = Q^T$and
   $$\boxed{Q = \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{30}} \\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{30}} \\ \frac{1}{\sqrt{6}} & 0 & \frac{-5}{\sqrt{30}} \end{pmatrix}}.$$
   Alternatively, since $X^T X = D$ is a diagonal matrix by orthogonality, we have $X^{-1} = D^{-1}X^T$, so then
   $$e^B = Xe^\Lambda D^{-1} X^T$$
   where $D^{-1} = \begin{pmatrix} 1/6 & & \\ & 1/5 & \\ & & 1/30 \end{pmatrix}$ is just the inverses of the squared lengths. Alternatively, you could compute X^{-1} by the Gauss–Jordan method, but that is a lot more work and is easy to get wrong!

   (b) If $C = (I - B)(I + B)^{-1}$ then:

      i. The eigenvalues of C are just $\frac{1-\lambda_i}{1+\lambda_i}$, i.e. $\frac{1-2}{1+2} = -\frac{1}{3}$, 1 and $\frac{1-1}{1+1} = 0$, with the same corresponding eigenvectors x_i

ii. The vector y, where

$$y = C^{100}\begin{pmatrix}1\\1\\1\end{pmatrix}.$$

will be almost parallel to the eigenvector corresponding to the largest-magnitude eigenvalue. The largest-magnitude eigenvalue of C is 1, with normalized eigenvector q_2, and so $y = C^{100}\begin{pmatrix}1\\1\\1\end{pmatrix} \approx \frac{\alpha}{\sqrt{5}}\begin{pmatrix}-2\\1\\0\end{pmatrix}$, where

$$\alpha = q_2^T\begin{pmatrix}1\\1\\1\end{pmatrix} = \frac{-1}{\sqrt{5}}$$

so that

$$\boxed{y \approx \frac{1}{5}\begin{pmatrix}2\\-1\\0\end{pmatrix} = q_2q_2^T\begin{pmatrix}1\\1\\1\end{pmatrix} = \frac{x_2x_2^T}{x_2^Tx_2}\begin{pmatrix}1\\1\\1\end{pmatrix}}.$$

3. (a) Since A is real, its complex eigenvectors must come in complex-conjugate pairs. So

$$x_3 = \overline{x}_2 = \begin{pmatrix}0\\1\\-i\end{pmatrix}.$$

(b) Notice that the eigenvalues of A satisfy $|\lambda| \geq 1$ and $\mathrm{Re}[\lambda] < 0$. This is enough information for us to answer the first 4 parts of this question.

a) This has unbounded magnitude. If we write $y = c_1x_1 + c_2x_2 + c_3x_3$, then

$$A^ny = c_1\lambda_1^nx_1 + c_2\lambda_2^nx_2 + \lambda_3^nc_3x_3$$

and λ_2^n, λ_3^n become larger and larger in magnitude as $n \to \infty$ (you can see this by writing those eigenvalues in polar form). Since y was chosen at random, c_2, c_3 are likely nonzero.

b) The magnitude of this vector will stay bounded as $n \to \infty$ (though it may not converge to any vector in particular). Remember the eigenvalues of A^{-1} are $1/\lambda_i$ and $|1/\lambda_i| \leq 1$. So writing

$$A^{-n}y = c_1\lambda_1^{-n}x_1 + c_2\lambda_2^{-n}x_2 + \lambda_3^{-n}c_3x_3$$

we see that the second and last term will decay as $n \to \infty$ (e.g. by writing those eigenvalues in polar form). The first term will always have the same magnitude.

c) The solution to this equation is

$$x(t) = e^{At}y = c_1e^{\lambda_1t}x_1 + c_2e^{\lambda_2t}x_2 + c_3e^{\lambda_3t}x_3$$

and it has bounded magnitude as $t \to \infty$. This is because $\mathrm{Re}[\lambda_j] < 0$ for all eigenvalues, so e^{λ_jt} always approaches zero as $t \to \infty$ (you can see this by writing $\lambda_j = a + ib$).

d) The eigenvalues of $-A$ are $-\lambda_j$, so they all have positive real parts. This means that the solution

$$x(t) = e^{-At}y = c_1e^{-\lambda_1 t}x_1 + c_2e^{-\lambda_2 t}x_2 + c_3e^{-\lambda_3 t}x_3$$

will have unbounded magnitude as $t \to \infty$, since each term has magnitude which blows up.

(c) As above, the general solution to $\frac{dx}{dt} = Ax$ is

$$x(t) = e^{At}x(0) = c_1e^{-1t}x_1 + c_2e^{(-3+4i)t}x_2 + c_3e^{(-3-4i)t}x_3$$

where c_1, c_2, c_3 are some constants depending on $x(0)$. Because the initial conditions are real, we expect $c_2 = \overline{c_3}$.
Setting $t = 0$, we get

$$x(0) = c_1x_1 + c_2x_2 + c_3x_3 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}.$$

Eyeballing, we see that this is true if $c_1 = c_2 = c_3 = 1$, so the exact solution is

$$x(t) = e^{-1t}x_1 + e^{(-3+4i)t}x_2 + e^{(-3-4i)t}x_3.$$