LINEAR SYSTEMS (034032)

TUTORIAL 10

1 Topics

Linear algebra revision, matrix functions, Cayley–Hamilton theorem, matrix calculus.

2 Background

2.1 Linear Algebra Revision

Let $A, B \in \mathbb{R}^{n\times n}$ be real valued square $n \times n$ matrices.

2.1.1 Eigenvalues and Characteristic Polynomial

The *characteristic polynomial* of A is:

$$\chi_A(\lambda) := \det(\lambda I - A) = \lambda^n + \chi_{n-1}\lambda^{n-1} + \cdots + \chi_1\lambda + \chi_0 = 0$$

The solutions to the polynomial $\lambda_i \in \mathbb{C}$ are called the *eigenvalues*. The set of all eigenvalues of matrix A is known as its *spectrum*, denoted $\text{spec}(A)$.

2.1.2 Eigenvalues and Eigenvector

A vector $0 \neq \eta_i \in \mathbb{C}^n$ is a *right eigenvector* of A associated with the eigenvalue λ_i if

$$(\lambda_i I - A)\eta_i = 0$$

and $0 \neq \tilde{\eta}_i \in \mathbb{C}^n$ is a *left eigenvector* associated with the eigenvalue λ_i if

$$\tilde{\eta}_i'(\lambda_i I - A) = 0$$

2.1.3 Multiplicity

The *algebraic multiplicity* of an eigenvalue λ_i is the number of times it appears in the characteristic polynomial. Its *geometric multiplicity* is defined as $n - \text{rank}(\lambda_i I - A)$, where n is the the dimension of matrix A and $\text{rank}(A)$ is the rank of A (the number of linearly independent rows or columns in it).

2.1.4 Similarity

We say that a matrix A is *similar* to a matrix B (and vice versa) if there exists a nonsingular transformation matrix T, such that

$$A = TBT^{-1}$$

2.1.5 Diagonalization

We say that a matrix A is *diagonalizable* if there exists a matrix T and a diagonal matrix Λ_A such that

$$A = T\Lambda_A T^{-1},$$

where the diagonal elements of Λ_A are eigenvalues of A and the columns of T are the right eigenvectors of A, i.e.

$$T = \begin{bmatrix} \eta_1 & \eta_2 & \cdots & \eta_n \end{bmatrix} \in \mathbb{C}^{n\times n}$$

or, conversely, we can write the transformation as

$$A = \tilde{T}^{-1}\Lambda_A \tilde{T},$$

where the rows of $\tilde{T}$ are the left eigenvectors of A (transposed of course),

$$\tilde{T} = \begin{bmatrix} \tilde{\eta}_1' \\ \tilde{\eta}_2' \\ \vdots \\ \tilde{\eta}_n' \end{bmatrix} \in \mathbb{C}^{n\times n}$$

If η_i and $\tilde{\eta}_i$ are normalized, then $\tilde{T} = T^{-1}$.

Theorem 1. *A matrix A is diagonalizable iff the geometric multiplicity of each its eigenvalue is equal to its algebraic multiplicity. Otherwise, there exists an eigenvalue $\lambda_i \in \mathrm{spec}(A)$ such that its geometric multiplicity is smaller than its algebraic multiplicity, and A is called* defective.

2.1.6 Real "Diagonalization" of a Matrix with Complex Eigenvalues

If there is a pair of complex eigenvalues $\lambda_i, \overline{\lambda_i} = \sigma \pm \omega \mathrm{j}$, then we also have two complex conjugate eigenvectors,

$$\eta_i, \overline{\eta_i} = \alpha \pm \mathrm{j}\beta$$

where $\alpha, \beta \in \mathbb{R}^n$ are real valued vectors.

We now show a special representation of the matrix called "real diagonalization of a matrix with complex eigenvalues." We first define two linear combinations of our eigenvectors,

$$\frac{\eta_i + \bar{\eta}_i}{2} = \alpha \quad \text{and} \quad \frac{\eta_i - \bar{\eta}_i}{2\mathrm{j}} = \beta$$

Defining

$$T = \begin{bmatrix} \alpha & \beta \end{bmatrix} \quad \text{and} \quad \hat{\Lambda}_A = \begin{bmatrix} \sigma & \omega \\ -\omega & \sigma \end{bmatrix}$$

we get that

$$A = T\hat{\Lambda}_A T^{-1}.$$

Note that when doing this form of "diagonalizing" of the entire matrix A, we get Λ_A with either real eigenvalues or blocks of 2×2 $\hat{\Lambda}_A$s on the main diagonal:

$$\Lambda_A = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 & \cdots \\ 0 & \sigma_2 & \omega_2 & 0 & \cdots \\ 0 & -\omega_2 & \sigma_2 & 0 & \cdots \\ 0 & 0 & 0 & \lambda_3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

2.2 Matrix Functions

Let $f : \mathbb{R} \to \mathbb{R}$ be analytic, i.e. representable via its power series

$$f(x) = \sum_{j=0}^{\infty} f_j x^j = f_0 + f_1 x + f_2 x^2 + \cdots$$

Its matrix version $f : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ is defined as

$$f(A) := \sum_{j=0}^{\infty} f_j A^j$$

for the very same scalar coefficients f_j. We can use two methods to calculate $f(A)$:

1. Via **diagonalization**: If A is diagonalizable, then

$$A = T\Lambda_A T^{-1} \qquad \Longrightarrow \qquad f(A) = Tf(\Lambda_A)T^{-1} = T^{-1}\begin{bmatrix} f(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & f(\lambda_n) \end{bmatrix} T$$

   If A has a pair of complex eigenvalues, we can use the "real diagonalization" form for that spesific block:

$$A = T\hat{\Lambda}_A T^{-1} \qquad \Longrightarrow \qquad f(A) = Tf(\hat{\Lambda}_A)T^{-1}$$

2. Via **Cayley–Hamilton**: By Cayley–Hamilton arguments we know there is $\{g_i\}_{i=0}^{n}$ such that

$$f(A) = \sum_{j=0}^{\infty} f_j A^j = \sum_{i=0}^{n-1} g_i A^i$$

   In order to find the coefficients g_i, define $g(x) := \sum_{i=0}^{n-1} g_i x^i$ and, assuming that all the eigenvalues of A are simple, obtain that

$$\begin{bmatrix} g_0 & g_1 & \cdots & g_{n-1} \end{bmatrix} = \begin{bmatrix} f(\lambda_1) & f(\lambda_2) & \cdots & f(\lambda_n) \end{bmatrix} V^{-1}$$

   where V is the (invertible) Vandermonde matrix

$$V = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \vdots & \vdots & & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{bmatrix}$$

   If there is an eigenvalue λ_i of A, with the algebraic multiplicity $\mu_i > 1$, then we need to add extra conditions, viz.

$$\left.\frac{\mathrm{d}^j f(x)}{\mathrm{d}x^j}\right|_{x=\lambda_i} = \left.\frac{\mathrm{d}^j g(x)}{\mathrm{d}x^j}\right|_{x=\lambda_i}, \qquad \forall j \in \mathbb{Z}_{1..\mu_i-1}$$

2.2.1 Matrix Exponential

The matrix exponential is a particular case of matrix functions, which is of a special interest. Let $f(x) = \exp(x) = \mathrm{e}^x$. Its power series is

$$\mathrm{e}^x = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \cdots$$

The form of the matrix version of this function that we are interested in is $\exp(At) = \mathrm{e}^{At}$, where $t \in \mathbb{R}$. The solution to the matrix function is as before. Specifically,

1. Via diagonalization,

$$\mathrm{e}^{At} = T\mathrm{e}^{\Lambda_A t}T^{-1} = T\begin{bmatrix} \mathrm{e}^{\lambda_1 t} & & 0 \\ & \ddots & \\ 0 & & \mathrm{e}^{\lambda_n t} \end{bmatrix} T^{-1}$$

   Or for the "real diagonalization of complex eigenvalues,"

$$\mathrm{e}^{At} = T\mathrm{e}^{\hat{A}t}T^{-1} = T\mathrm{e}^{\sigma t}\begin{bmatrix} \cos(\omega t) & \sin(\omega t) \\ -\sin(\omega t) & \cos(\omega t) \end{bmatrix} T^{-1}$$

2. Via Cayley–Hamilton:

$$\begin{bmatrix} g_0 & g_1 & \cdots & g_{n-1} \end{bmatrix} = \begin{bmatrix} \mathrm{e}^{\lambda_1 t} & \mathrm{e}^{\lambda_2 t} & \cdots & \mathrm{e}^{\lambda_n t} \end{bmatrix} V^{-1}$$

   where V is the Vandermonde matrix.

2.3 Matrix Calculus

The derivative of a matrix $A(t)$ by a scalar $t \in \mathbb{R}$ is done component-wise:

$$\frac{\mathrm{d}}{\mathrm{d}t}(A(t)) = \begin{bmatrix} \dot{a}_{11}(t) & \dot{a}_{12}(t) & \cdots & \dot{a}_{1n}(t) \\ \dot{a}_{21}(t) & \dot{a}_{22}(t) & \cdots & \dot{a}_{2n}(t) \\ \vdots & \vdots & & \vdots \\ \dot{a}_{n1}(t) & \dot{a}_{n2}(t) & \cdots & \dot{a}_{nn}(t) \end{bmatrix}$$

Some properties:

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}(A_1(t)A_2(t)) = \frac{\mathrm{d}}{\mathrm{d}t}(A_1(t))A_2(t) + A_1(t)\frac{\mathrm{d}}{\mathrm{d}t}(A_2(t)) \\
&\frac{\mathrm{d}}{\mathrm{d}t}A^{-1}(t) = -A^{-1}(t)\left(\frac{\mathrm{d}}{\mathrm{d}t}A(t)\right)A^{-1}(t) \\
&\frac{\mathrm{d}}{\mathrm{d}t}(At)^k = A^k(kt^{k-1}) \\
&\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{At} = A\mathrm{e}^{At} = \mathrm{e}^{At}A
\end{aligned}$$

The derivative of $f : \mathbb{R}^m \to \mathbb{R}^n$ by its vector argument $x \in \mathbb{R}^m$ is defined as

$$\frac{\partial f}{\partial x} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_m} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_m} \end{bmatrix} \in \mathbb{R}^{n\times m}$$

3 Problems

Question 1. Consider the matrix

$$A = \begin{bmatrix} 4 & 0 & 1 \\ -1 & -6 & -2 \\ 5 & 0 & 0 \end{bmatrix}$$

1. Find the diagonalizing transformation of A (in the real form if it exists).
2. Calculate the matrix exponent e^{At} using diagonalization.
3. Calculate the matrix exponent e^{At} using Cayley-Hamilton.

Question 2. Let A be the matrix:

$$\begin{bmatrix} -1 & 4 \\ -1 & -1 \end{bmatrix}$$

1. Find the diagonalizing transformation (in the real form if it exists).
2. Find the matrix exponential e^{At}.

Question 3. Let A be the matrix:

$$\begin{bmatrix} 1 & 2 & -8 \\ 0 & -1 & 4 \\ 0 & -1 & -1 \end{bmatrix}$$

1. Find the diagonalizing transformation (in the real form if it exists).
2. Find the matrix exponential e^{At}.

Question 4. Let $A \in \mathbb{R}^{n \times m}$ be a matrix and $x \in \mathbb{R}^m$ be a vector. Calculate the derivative of

$$y = Ax \in \mathbb{R}^n$$

by x.

4 Homework problems

Question 5. Consider:

$$A = \begin{bmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{bmatrix}$$

1. Find the diagonalizing transformation of A (in the real form if it exists).
2. Calculate the matrix exponent using Cayley-Hamilton.

Question 6. Find the derivative of the vector $f \in \mathbb{R}^3$ as a function of $x \in \mathbb{R}^2$:

$$\begin{bmatrix} x_1^3 - 2x_2 + 5\sin(x2) \\ x_2 \mathrm{e}^{x_1} - 3x_2 \\ \frac{x_1}{1 - x_2 - x_2^2} \end{bmatrix}$$

Question 7. Find the derivative of matrix $A \in \mathbb{R}^{3\times 3}$ by scalar $t \in \mathbb{R}$:

$$A = \begin{bmatrix} t & \sin t & \frac{1}{t^2}\mathrm{e}^{-t} \\ \ln t & \frac{cos(t)}{1+sint} & (2t+1)^4(3t-5)^2 \\ \sqrt{3t+t^2} & 3t^2 - 2t + 1 & \ln(\mathrm{e}^t + 1)^2 \end{bmatrix}$$

Question 8. Let $A(t) \in \mathbb{R}^{n\times p}$, $B(t) \in \mathbb{R}^{p\times m}$ and $C \in \mathbb{R}^{n\times n}$ be matrices that depend on scalar $t \in \mathbb{R}$. Prove the following statements:

1. $\frac{\mathrm{d}}{\mathrm{d}t}(A(t)B(t)) = (\frac{\mathrm{d}}{\mathrm{d}t}A(t))B(t) + A(t)(\frac{\mathrm{d}}{\mathrm{d}t}B(t))$.
2. $\frac{\mathrm{d}}{\mathrm{d}t}(C^{-1}(t)) = -C^{-1}(\frac{\mathrm{d}}{\mathrm{d}t}C(t))C^{-1}$.