ECE205A Matrix Analysis (Fall 2020) December 11, 2020

Final Review

Instructor: Dr. M. R. Rajati TA: Xin Jiang

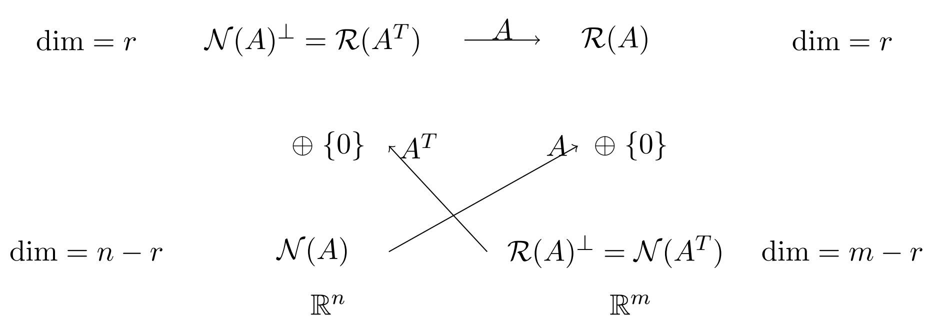
1 Linear transformation

1.1 Linear transformation

- Definition.
- Linear transformation and matrix–vector product.
- Change of basis.
- Composition of linear functions.
- The space of linear functions (HW5P7).
- Adjoint (HW6P8, 10). See also Problem 1.

1.2 Four fundamental subspaces

Consider a general matrix $A \in \mathbb{R}^{m\times n}$ with $\mathbf{rank}(A) = r$. When thought of as a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$, many properties of A can be developed in terms of the four fundamental subspaces.



1.3 Injection, surjection, and bijection

Injection. The linear map $\mathcal{L}\colon \mathcal{V} \to \mathcal{W}$ is called *injective* (or *one-to-one*, *monic*) if $\mathcal{N}(\mathcal{L}) = \{0\}$. In particular, when $\mathcal{L}$ is characterized by the matrix $A \in \mathbb{R}^{m\times n}$, the following conditions are equivalent.

- A is injective, one-to-one, or monic.
- $\mathcal{N}(A) = \{0\}$, or $\mathbf{rank}(A) = n$; *i.e.*, $Ax = Ay$ implies $x = y$, or equivalently, $Ax \neq Ay$ if $x \neq y$.
- A has linearly independent columns.
- A is left invertible; *i.e.*, there exists a linear map $A^{-L}\colon \mathcal{W} \to \mathcal{V}$ such that $A^{-L} \circ A = \mathcal{I}_{\mathcal{V}}$.
- The Gram matrix $A^T A$ is nonsingular.
- Linear equations $Ax = b$ has at most one solution for every right-hand side $b \in \mathbb{R}^m$.

A matrix with these properties must be tall or square ($m \geq n$). Its pseudo-inverse, given by $A^\dagger = (A^T A)^{-1} A^T$, is also a left inverse of A, and the left inverse has the form

$$A^{-L} = A^\dagger + Y(I - AA^\dagger),$$

where $Y \in \mathbb{R}^{n\times m}$ is arbitrary. The solution for $Ax = b$ (if exists) is $x = A^\dagger b$.

Surjection. The linear map $\mathcal{L}\colon \mathcal{V} \to \mathcal{W}$ is called *surjective* (or *onto*, *epic*) if $\mathcal{R}(\mathcal{L}) = \mathcal{W}$. In particular, when $\mathcal{L}$ is characterized by the matrix $A \in \mathbb{R}^{m\times n}$, the following conditions are equivalent.

- A is surjective, onto, or epic.
- $\mathcal{R}(A) = \mathbb{R}^m$, or $\mathbf{rank}(A) = m$.
- A has linearly independent rows.
- A is right invertible; *i.e.*, there exists a linear map $A^{-R}\colon \mathcal{W} \to \mathcal{V}$ such that $A \circ A^{-R} = \mathcal{I}_{\mathcal{W}}$.
- The Gram matrix AA^T is nonsingular.
- Linear equations $Ax = b$ has at least one solution for every right-hand side $b \in \mathbb{R}^n$.

A matrix with these properties must be wide or square ($m \leq n$). Its pseudo-inverse, given by $A^\dagger = A^T (AA^T)^{-1}$, is also a right inverse of A, and the right inverse has the form

$$A^{-R} = A^\dagger + (I - A^\dagger A)Y,$$

where $Y \in \mathbb{R}^{n\times m}$ is arbitrary. Thus, $x = A^{-R} b$ is the general solution for $Ax = b$.

Bijection. The linear map $\mathcal{L}\colon \mathcal{V} \to \mathcal{V}$ is *bijective* (or *invertible*) if and only if its is one-to-one and onto. In particular, the *square* matrix $A\colon \mathbb{R}^n \to \mathbb{R}^n$ is invertible or *nonsingular* if and only if any of the following conditions are valid.

- A is left invertible.
- The columns of A are linearly independent.
- A is right invertible.
- The rows of A are linearly independent.
- Linear equations $Ax = b$ has a unique solution $x = A^{-1} b$ for every right-hand side $b \in \mathbb{R}^n$.

1.4 Linear systems

We study the consistency and structure of the solutions for the (vector) linear system

$$Ax = b, \quad \text{where } A \in \mathbb{R}^{m \times n},\ b \in \mathbb{R}^m. \tag{1}$$

If the linear system (1) is solvable, we call (1) is *consistent*; otherwise, it is *inconsistent.*

Arbitrary right-hand side. We first review the consistency of (1) for any right-hand side.

- There exists at least one solution to (1) for all $b \in \mathbb{R}^m$ if and only if $\mathcal{R}(A) = \mathbb{R}^m$.
- There exists at most one solution to (1) for all $b \in \mathbb{R}^m$ if and only if $\mathcal{N}(A) = \{0\}$.
- There exists a unique solution to (1) for all $b \in \mathbb{R}^m$ if and only if A is nonsingular.

Particular right-hand side. We examine the consistency of (1) for a particular right-hand side.

- There exists (at least) one solution to (1) if and only if $b \in \mathcal{R}(A)$, or equivalently, (HW6P3)

  $$\textbf{rank}(\begin{bmatrix} A & b \end{bmatrix}) = \textbf{rank}(A).$$

  Moreover when it occurs, x is a solution to (1) if and only if

  $$x = x^{\bullet} + z,$$

  where $z \in \mathcal{N}(A)$ is an element of the nullspace of A.
- The solution to (1) (if exists) is unique if and only if $\mathcal{N}(A) = \{0\}$.
- There is a nontrivial solution for the homogeneous system $Az = 0$ if and only if $\mathcal{N}(A) \neq \{0\}$.

1.5 Projection

- Projection is a linear transformation.
- $P \in \mathbb{R}^{n \times n}$ is a projection iff $P^2 = P$.
- $P \in \mathbb{R}^{n \times n}$ is a projection iff $I - P$ is a projection.
- P is an orthogonal projection iff $P^2 = P = P^T$.
- Orthogonal projections on fundamental subspaces.

  $$\mathcal{P}_{\mathcal{R}(A)} = AA^{\dagger} \quad \mathcal{P}_{\mathcal{N}(A)} = I - A^{\dagger}A \quad \mathcal{P}_{\mathcal{R}(A)^{\perp}} = I - AA^{\dagger} \quad \mathcal{P}_{\mathcal{N}(A)^{\perp}} = A^{\dagger}A$$

2 Moore–Penrose pseudo-inverse

Definition. Algebraic; Penrose conditions (HW7P1, 3, 4, 5, 6); and limit characteristic (HW7P2).

Inverse, pseudo-inverse, left/right inverse.

- $A \in \mathbb{R}_n^{n\times n}$ is square and nonsingular, $A^\dagger = A^{-1} = A^{-L} = A^{-R}$.
- $A \in \mathbb{R}_n^{m\times n}$ is a tall matrix with linearly independent columns.

pseudo-inverse	unique	$A^\dagger = (A^TA)^{-1}A^T$
left-inverse	not unique	$A^{-L} = A^\dagger + Y(I - AA^\dagger)$
sol. for $Ax = b$	if exists	$x = A^\dagger b$

  (In the table, Y is an arbitrary $n \times m$ matrix.)

- $A \in \mathbb{R}_m^{m\times n}$ is a wide matrix with linearly independent rows.

pseudo-inverse	unique	$A^\dagger = A^T(AA^T)^{-1}$
right-inverse	not unique	$A^{-R} = A^\dagger + (I - A^\dagger A)Y$
sol. for $Ax = b$	always exists	$x = A^{-R}b$

  (In the table, Y is an arbitrary $n \times m$ matrix.)

Pseudo-inverse and matrix decomposition.

- *Full-rank decomposition.* Suppose $A \in \mathbb{R}_r^{m\times n}$ has full-rank decomposition $A = BC$, where $B \in \mathbb{R}_r^{m\times r}$ and $C \in \mathbb{R}_r^{r\times n}$. Then $A^\dagger = (BC)^\dagger = C^\dagger B^\dagger = C^T(CC^T)^{-1}(B^TB)^{-1}B^T$.
- *Spectral decomposition.* Suppose $A \in \mathbb{S}^n$ has eigenvaue decomposition $A = Q\Lambda Q^T$. Then $A^\dagger = Q\Lambda^\dagger Q^T$, where $\Lambda^\dagger$ is a diagonal matrix with diagonal elements
$$\lambda_i^\dagger = \begin{cases} \lambda_i^{-1} & \text{if } \lambda_i \neq 0 \\ 0 & \text{elsewhere.} \end{cases}$$
- *Singular value decomposition.* Omitted. Suppose $A \in \mathbb{R}_r^{m\times n}$ has singular value decomposition $A = U\Sigma V^T$, where $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ are orthogonal, and $\Sigma \in \mathbb{R}^{m\times n}$ has diagonal elements $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ and zero elsewhere. Then $A^\dagger = V\Sigma^\dagger U$, where $\Sigma^\dagger \in \mathbb{R}^{n\times m}$ has diagonal elements $\sigma_1^{-1}, \ldots, \sigma_r^{-1}$, and zero elsewhere.
- *Orthogonal equivalence.* Suppose $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ are orthogonal matrices. Let $A \in \mathbb{R}^{m\times n}$, and thus
$$(UAV)^\dagger = V^TA^\dagger U^T.$$

Pseudo-inverse and fundamental subspaces.

- $\mathcal{R}(A^\dagger) = \mathcal{R}(A^T) = \mathcal{R}(A^\dagger A) = \mathcal{R}(A^TA)$.
- $\mathcal{N}(A^\dagger) = \mathcal{N}(A^T) = \mathcal{N}(AA^\dagger) = \mathcal{N}(AA^T) = \mathcal{N}((AA^T)^\dagger)$.
- Let $A \in \mathbb{R}^{n\times p}$ and $B \in \mathbb{R}^{n\times m}$. Then $\mathcal{R}(B) \subseteq \mathcal{R}(A)$ if and only if $AA^\dagger B = B$.
- Let $A \in \mathbb{R}^{p\times n}$ and $B \in \mathbb{R}^{m\times n}$. Then $\mathcal{N}(A) \subseteq \mathcal{N}(B)$ if and only if $BA^\dagger A = B$.

3 Linear least squares problem

3.1 Least squares problem

Least squares problem	minimize $\|Ax - b\|_2^2$
Normal equation	$A^T A\hat{x} = A^T b$
Unique solution	$\hat{x} = A^\dagger b = (A^T A)^{-1} A^T b$

- The uniqueness of the least squares solution requires the left-invertibity of A.
- Three approaches to solve the least squares problem.
- Regularization: motivation and transformation to a least squares problem.
- Linear regression; see Problem 5.

3.2 Least squares model fitting

We choose the model $\hat{f}(x)$ from a family of models

$$\hat{f}(x) = \theta_1 f_1(x) + \cdots + \theta_p f_p(x) = \theta^T F(x).$$

- The *basis functions* f_i are chosen by us, and $F(x) = (f_1(x), \ldots, f_p(x))$ is a p-vector of basis functions. The basis functions usually include a constant function, typically, $f_i(x) = 1$.
- The coefficients $\theta_1, \ldots, \theta_p$ are the model *parameters*.
- It is called a linear-in-parameters model since the model $\hat{f}(x)$ is linear in the parameters θ_i.

To fit the linear-in-parameters model to the data set $(x_1, y_1), \ldots, (x_N, y_N)$, we minimized the sum of ℓ_2-norms of residuals $r_i = \hat{f}(x_i) - y_i$, *i.e.*,

$$\text{minimize} \quad \frac{1}{N} \sum_{i=1}^{N} \|\hat{f}(x_i) - y_i\|_2^2.$$

It is a least squares problem

$$\text{minimize} \quad \|A\theta - y\|_2^2$$

where

$$A = \begin{bmatrix} f_1(x_1) & f_2(x_1) & \cdots & f_p(x_1) \\ f_1(x_2) & f_2(x_2) & \cdots & f_p(x_2) \\ \vdots & \vdots & & \vdots \\ f_1(x_N) & f_2(x_N) & \cdots & f_p(x_N) \end{bmatrix}, \quad \theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_p \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}.$$

3.3 Least norm problem

$$\begin{array}{ll} \text{minimize} & \|x\|_2^2 \\ \text{subject to} & Cx = d \end{array}$$

- The coefficient matrix $C \in \mathbb{R}^{p \times n}$ is a wide matrix ($p < n$) in most applications. So the equation $Cx = d$ is often underdetermined.
- We assume C has linearly independent rows, and thus the optimal solution is given by

$$\hat{x} = C^\dagger d = C^T (CC^T)^{-1} d.$$

4 Eigenvalues and eigenvectors

4.1 Basic definition and properties

Characteristic polynomial and minimal polynomial. $\lambda_1, \ldots, \lambda_\tau$ are distinct eigenvalues of A.

Characteristic polynomial	$\pi(\lambda) = \det(A - \lambda I) = \prod_{i=1}^{\tau}(\lambda - \lambda_i)^{\alpha_i}$
Minimal polynomial	$\phi(\lambda) = \prod_{i=1}^{\tau}(\lambda - \lambda_i)^{\beta_i}$

Multiplicity of eigenvalues.

Eigenvalue (algebraic)	$\pi(\lambda) = \det(A - \lambda I) = 0$	α	repetition of λ
Eigenvector (geometric)	nonzero solutions to $Ax = \lambda x$	γ	$\dim \mathcal{N}(A - \lambda I)$

Diagonalization. Equivalent conditions for a diagonalizable matrix A.

- There exists an $n \times n$ nonsingular matrix X such that $X^{-1}AX = D$ is diagonal.
- A has a complete linearly independent set of eigenvectors.
- For every eigenvalue λ_i, its algebraic and geometric multiplicities are equal, *i.e.*, $\alpha_i = \gamma_i$.

Remark.

- There is no connection between invertibility and diagonalization.
- The matrix A is diagonalizable if and only if the nullity of A is equal to the (algebraic) multiplicity of $\lambda = 0$. (HW8P5)

Eigenvalues and eigenvectors of Hermitian matrices.

- A Hermitian matrix has only real eigenvalues.
- Eigenvectors of a Hermitian matrix are orthogonal to each other. Thus all Hermitian matrices are diagonalizable, and especially, they are unitarily similar to a diagonal matrix. This is called the *spectral decomposition* $A = Q\Lambda Q^H$.
- The diagonal elements of a Hermitian matrix must be real.

4.2 Matrix similarity and equivalence

- $A, B \in \mathbb{C}^{n\times n}$ are *similar* if there is a nonsingular matrix $M \in \mathbb{C}^{n\times n}$ such that $B = M^{-1}AM$.
- $A, B \in \mathbb{C}^{m\times n}$ are *equivalent* if there is nonsingular matrices P and Q such that $B = PAQ$.
- $A, B \in \mathbb{C}^{n\times n}$ are *congruent* if there is a nonsingular matrix $M \in \mathbb{C}^{n\times n}$ such that $B = M^H AM$.

Remark. The table below shows some properties of similarity transformation. In particular, similar matrices share the same eigenvalues. But unfortunately, two matrices can have the same *repeated* eigenvalues and fail to be similar, because they may have different number of linearly independent eigenvectors.

Not changed by M	Changed by M
Eigenvalues	Eigenvectors
Rank	Fundamental subspaces
Trace and determinant	Singular values
Number of linearly independent eigenvectors	
characteristic polynomial $\pi(\lambda)$	
minimal polynomial $\phi(\lambda)$	

4.3 Jordan canonical form

Definition. For any square matrix $A \in \mathbb{C}^{n\times n}$, there is a nonsingular matrix $X \in \mathbb{C}^{n\times n}$ such that

$$X^{-1}AX = J = \mathbf{diag}(J_1, \ldots, J_q),$$

where each of the Jordan block matrices $J_1, \ldots, J_q$ is of the form

$$J_i = \begin{bmatrix} \lambda_j & 1 & 0 & \cdots & \cdots & 0 \\ 0 & \lambda_j & 1 & 0 & & \vdots \\ \vdots & \ddots & \lambda_j & \ddots & \ddots & \vdots \\ & & & \ddots & 1 & 0 \\ \vdots & & & \ddots & \lambda_j & 1 \\ 0 & \cdots & & \cdots & 0 & \lambda_j \end{bmatrix} \in \mathbb{C}^{n_i\times n_i} \tag{2}$$

and $\sum_{i=1}^{q} n_i = n$. Especially, for real matrices $A \in \mathbb{R}^{n\times n}$, each Jordan block matrix has the form (2) in the case of real eigenvalues $\lambda_j \in \mathbb{R}$, and

$$J_i = \begin{bmatrix} M_j & I_2 & 0 & \cdots & \cdots & 0 \\ 0 & M_j & I_2 & 0 & & \vdots \\ \vdots & \ddots & M_j & \ddots & \ddots & \vdots \\ & & & \ddots & 1 & 0 \\ \vdots & & & \ddots & M_j & I_2 \\ 0 & \cdots & & \cdots & 0 & M_j \end{bmatrix} \in \mathbb{R}^{2n_i\times 2n_i}, \qquad \text{where } M_j = \begin{bmatrix} a_j & b_j \\ -b_j & a_j \end{bmatrix}, \quad I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

in the case of complex conjugate eigenvalues $a_j \pm \hat{\imath} b_j$.

Determination of the JCF. Detailed computation omitted (HW8P4). See also Problem 9.

- The Jordan canonical form J of A can be determined eigenvalue by eigenvalue.
- The algebraic multiplicity α determines the total size of Jordan blocks associated with λ.
- The geometric multiplicity γ indicates the number of Jordan blocks associated with λ.
- The degree β of the associated item in the minimal polynomial is the same as the largest size of Jordan blocks associated with λ.

The knowledge of α, β, and γ cannot determine the JCF (HW9P6, 7, 8). See also Problem 8.

Algorithm 1 Determination of the Jordan canonical form.

Require: A matrix $A \in \mathbb{C}^{n\times n}$ with distinct eigenvalues $\lambda_1, \dots, \lambda_\tau$, each of algebraic multiplicity α_i and geometric multiplicity γ_i.

Ensure: The Jordan canonical form $J \in \mathbb{C}^{n\times n}$, and the transformation matrix $X \in \mathbb{C}^{n\times n}$ with linearly independent columns.

for $i = 1, \dots, \tau$ **do**

  Solve the linear equation

$$(A - \lambda_i I)x^{(1)} = 0.$$

  This step finds all the linearly independent eigenvectors associated with λ_i, and clearly there are $\gamma_i = \dim \mathcal{N}(A - \lambda I)$ of them, denoted as $x_{i,1}^{(1)}, \dots, x_{i,\gamma_i}^{(1)}$.

  for $j = 1, \dots, \gamma_i$ **do**

    for $\ell = 1, \dots, \beta_i$ **do** the following. Solve the linear equation

$$(A - \lambda_i I)x = x_{i,j}^{(\ell)}.$$

      If the solution $x \in \mathcal{R}(A - \lambda_i I)$, denote it as $x_{i,j}^{(\ell+1)}$ and proceed to next ℓ-iteration.

      Otherwise, record the current ℓ as η_j, and then break the current ℓ-iteration.

    end for

  end for

  Arrange all the solutions in a matrix

$$X_i = \begin{bmatrix} x_{i,1}^{(1)} & \cdots & x_{i,1}^{(\eta_1)} \mid x_{i,2}^{(1)} & \cdots & x_{i,2}^{(\eta_2)} \mid \cdots \mid x_{i,\gamma_i}^{(1)} & \cdots & x_{i,\gamma_i}^{(\eta_{\gamma_i})} \end{bmatrix}.$$

  Determine the sizes of the Jordan blocks associated with current λ_i:

$$J_{i,1} \in \mathbb{C}^{\eta_1\times\eta_1},\ J_{i,2} \in \mathbb{C}^{\eta_2\times\eta_2},\ \cdots,\ J_{i,\gamma_i} \in \mathbb{C}^{\eta_{\gamma_i}\times\eta_{\gamma_i}},$$

  and define $J_i = \mathbf{diag}(J_{i,1}, \dots, J_{i,\gamma_i})$.

end for

Define the Jordan canonical form

$$J = \mathbf{diag}(J_1, \dots, J_\tau),$$

and the corresponding nonsingular matrix

$$X = \begin{bmatrix} X_1 & \cdots & X_\tau \end{bmatrix}.$$

5 Positive definite matrices

Positive definite matrices. A symmetric matrix $A \in \mathbb{S}^n$ is *positive definite* if

$$x^T A x > 0 \quad \text{for all } x \neq 0.$$

The following properties are equivalent for a symmetric $n \times n$ matrix A.

1. The quadratic form $x^T A x$ is positive except at $x = 0$.
2. All the eigenvalues are positive.
3. All the leading principal minors have positive determinant; and they are positive definite.
4. The matrix A can be factored as $A = M^T M$ where $M \in \mathbb{R}^{n \times n}$ is nonsingular. Note that the factor M may not be unique.
5. It has a unique Cholesky factorization $A = LL^T$.

Positive semidefinite matrices. A symmetric matrix $A \in \mathbb{S}^n$ is *positive semidefinite* if

$$x^T A x \geq 0 \quad \text{for all } x.$$

The following properties are equivalent for a symmetric $n \times n$ matrix A.

1. The quadratic form $x^T A x$ is always nonnegative.
2. All the eigenvalues are nonnegative.
3. All the principal minors have nonnegative determinants; and they are positive semidefinite.
4. The matrix A can be factored as $A = M^T M$ with $M \in \mathbb{R}^{k \times n}$ and $k \geq \mathbf{rank}(A) = \mathbf{rank}(M)$. Note that the factor M may not be unique.
5. There is a permutation matrix $P \in \mathbb{R}^{n \times n}$ such that $P^T A P$ has a Cholesky factorization $PAP^T = LL^T$, and L might be singular.

6 Functions of matrices

A Jordan-based approach. Let $A \in \mathbb{C}^{n\times n}$ and $J = \mathbf{diag}(J_1, \ldots, J_q) = X^{-1}AX$ be its JCF where each Jordan block J_i has the form (2). The matrix function $f(A)$ is defined by

$$f(A) = X\,\mathbf{diag}(F_1, \ldots, F_q)X^{-1},$$

where

$$F_i = \begin{bmatrix} f(\lambda_i) & f^{(1)}(\lambda_i) & \cdots & \cdots & \dfrac{f^{(n_i-1)}(\lambda_i)}{(n_i-1)!} \\ 0 & f(\lambda_i) & \ddots & \cdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & f^{(1)}(\lambda_i) \\ 0 & \cdots & \cdots & \cdots & f(\lambda_i) \end{bmatrix}, \quad i = 1, 2, \ldots, q,$$

assuming that all the required derivative evaluations exist. (HW9P5)

An eigenvector approach. If the matrix A is diagonalizable: $A = X\,\mathbf{diag}(\lambda_1, \ldots, \lambda_n)X^{-1}$, then

$$f(A) = X\,\mathbf{diag}\big(f(\lambda_1), \ldots, f(\lambda_n)\big)X^{-1}.$$

The Taylor series representation. Assume that f is analytic in a neighborhood of $z_0 \in \mathbb{C}$, *i.e.*

$$f(z) = \sum_{k=0}^{\infty} \frac{f^{(k)}(z_0)}{k!}(z - z_0)^k, \quad \text{for } |z - z_0| < r. \tag{3}$$

Let $A \in \mathbb{C}^{n\times n}$ and suppose $|\lambda - z_0| < r$ for all $\lambda \in \Lambda(A)$. Then

$$f(A) = \sum_{k=0}^{\infty} \frac{f^{(k)}(z_0)}{k!}(A - z_0 I)^k.$$

(This definition is rarely used in computation, but Problem 12 would be a good example.)

Interpolation method. We can regroup the Taylor series (3) of $f(z)$ so that

$$f(z) = \pi(z)\sum_{k=0}^{\infty} a_k z^k + r(z),$$

where $\pi(\lambda) = \prod_{i=1}^{\tau}(\lambda - \lambda_i)^{\alpha_i}$ is the characteristic polynomial of A, and $r(z)$ is a polynomial of degree at most $n-1$. Thus we can define

$$r(z) = c_0 + c_1 z + \cdots + c_{n-1}z^{n-1},$$

where $c_0, \ldots, c_{n-1}$ are n constants to be determined. In fact, they are the unique solution of the n linear equations

$$r^{(k)}(\lambda_i) = f^{(k)}(\lambda_i), \quad \text{for } k = 0, 1, \ldots, \alpha_i - 1, \text{ and } i = 1, \ldots, \tau.$$

The superscript (k) denotes the kth derivative. This is true because $\pi^{(k)}(\lambda_i) = 0$ for $k = 0, 1, \ldots, \alpha_i - 1$. By solving the coefficients of the polynomial r, we determine the function $f(A) = r(A)$ by the Cayley–Hamilton theorem. (HW9P3, 4, 9)

7 Linear differential equations

homogeneous LDE	$\dot{x}(t) = Ax(t)$	$x(t) = e^{(t-t_0)A}x_0$
non-homogeneous LDE	$\dot{x}(t) = Ax(t) + Bu(t)$	$x(t) = e^{(t-t_0)A}x_0 + \int_{t_0}^{t} e^{(t-s)A}Bu(s)\,ds$

Modal decomposition. The solution $x(t)$ is a weighted sum of its modal directions.

Stability of LTI systems. Definitions of Lyapunov and asymptotical stability, and their representations in terms of $\Lambda(A)$.

- The system is asymptotically stable if $\Re(\lambda_i) < 0$ for all $i = 1, \ldots, n$.
- The system is Lyapunov stable if every eigenvalue λ_i satisfy the following conditions.
  - If λ_i is a simple eigenvalue, then $\Re(\lambda_i) \leq 0$.
  - If λ_i is a repeated eigenvalue, then $\Re(\lambda_i) < 0$.
- The system is unstable if there is at least one eigenvalue λ_i with $\Re(\lambda_i) > 0$.

Lyapunov's indirect method. Use linearization to determine the stability of the original system.

Linearization	asymptotically stable	unstable	Lyapunov stable
Original nonlinear system	asymptotically stable	unstable	unknown

Lyapunov's direct method. Use Lyapunov equation to determine the stability of an LTI system.

Higher-order linear differential equations. Use change of variables to transform a higher-order LDE

$$y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \cdots a_1\dot{y}(t) + a_0 y(t) = u(t)$$

into a first-order linear system

$$\dot{x}(t) = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \\ -a_0 & -a_1 & \cdots & \cdots & -a_{n-1} \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} u(t).$$

8 Kronecker product

Definition and properties.

9 Additional practice problems

Remark. These problems are additional exercises collected by the TA, and they do not indicate the format, structure, or difficulty of the final exam.

We denote by $\mathbb{S}^n$ the set of $n \times n$ real symmetric matrices. In addition, the set of $n \times n$ symmetric positive definite (and positive semidefinite) matrices is denoted by $\mathbb{S}^n_{++}$ (and $\mathbb{S}^n_{+}$).

Problem 1. Let $\mathcal{V}$, $\mathcal{W}_1$, and $\mathcal{W}_2$ be inner product spaces, and let $L_1 \colon \mathcal{V} \to \mathcal{W}_1$, $L_2 \colon \mathcal{V} \to \mathcal{W}_2$ be two linear functions. Define $J_1 = L_1^* \circ L_1$ and $J_2 = L_2^* \circ L_2$. Show that the sum $J = J_1 + J_2$ can be written as a self-adjoint combination $J = L^* \circ L$ for some linear function L.

Problem 2. Let $P \in \mathbb{R}^{n \times n}$ be an orthogonal projection matrix.

(a) Show that $\|x\|_2^2 = \|Px\|_2^2 + \|(I - P)x\|_2^2$ for all $x \in \mathbb{R}^n$.

(b) Show that the matrix $2P - I$ is orthogonal, and the inequalities

$$-\|x\|_2 \|y\|_2 \leq x^T(2P - I)y \leq \|x\|_2 \|y\|_2$$

holds for all n-vectors x and y.

(c) Show that P is positive semidefinite.

(d) Is P always positive definite? If no, list all the positive definite, orthogonal projection matrices.

(e) Show that $\mathbb{R}^n = \mathcal{R}(P) \oplus \mathcal{N}(P)$. Explain whether this direct sum still holds if P is only a projection matrix, but not orthogonal.

Problem 3. Suppose A is an $n \times (n - 1)$ matrix with linearly independent columns, and b is an n-vector with $A^T b = 0$ and $\|b\|_2 = 1$.

(a) Show that the matrix $\begin{bmatrix} A & b \end{bmatrix}$ is nonsingular with inverse $\begin{bmatrix} A^\dagger \\ b^T \end{bmatrix}$.

(b) Let C be any left inverse of A. Show that

$$\begin{bmatrix} C(I - bb^T) \\ b^T \end{bmatrix} \begin{bmatrix} A & b \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & 1 \end{bmatrix}.$$

(c) Use the results of parts (a) and (b) to show that $C(I - bb^T) = A^\dagger$.

Problem 4. Suppose $A, B \in \mathbb{R}^{m \times n}$ have linearly independent columns and

$$AA^T = BB^T.$$

In this problem we show that $B = AQ$ for some orthogonal Q.

(a) Show that the matrix $U = A^\dagger B$ is orthogonal.

(b) Show that the matrix $V = B^\dagger A$ is orthogonal.

(c) Show that U is the inverse of V.

(d) Find an orthogonal matrix Q such that $B = AQ$.

(This result can be extended to two arbitrary tall matrices A and B, not necessarily left-invertible.)

Problem 5. Suppose the set of points $\{(t_i, y_i)\}_{i=1}^m$ are approximated by a function of the form

$$f(t) = \frac{e^{\alpha t+\beta}}{1+e^{\alpha t+\beta}}.$$

Formulate the following problem as a least squares problem: Find values of the parameters α, β such that

$$\frac{e^{\alpha t_i+\beta}}{1+e^{\alpha t_i+\beta}} \approx y_i, \qquad i = 1, \ldots, m.$$

(You can assume that $0 < y_i < 1$ for $i = 1, \ldots, m$.)

Problem 6. Find the optimal solution of the following two optimization problems.

(a)

$$\begin{array}{ll} \text{minimize} & \|x\|_2^2 + \|y\|_2^2 \\ \text{subject to} & A^T x - 2A^T y = b, \end{array}$$

where $x, y \in \mathbb{R}^m$ are the optimization variables, and $A \in \mathbb{R}^{m\times n}$ is left-invertible.

(b)

$$\begin{array}{ll} \text{minimize} & x^T A x \\ \text{subject to} & c^T x = 1, \end{array}$$

where $x \in \mathbb{R}^n$ is the optimization variable, $A \in \mathbb{S}^n_{++}$, and $c \in \mathbb{R}^n$ is a nonzero vector.

Problem 7. In this problem, suppose A is an $n \times n$ symmetric matrix.

(a) Show that the largest and smallest eigenvalues of A are

$$\lambda_{\max} = \max_{x\neq 0} \frac{x^T A x}{x^T x} = \max_{\|x\|_2=1} x^T A x, \qquad \lambda_{\min} = \min_{x\neq 0} \frac{x^T A x}{x^T x} = \min_{\|x\|_2=1} x^T A x.$$

(b) Write the matrix norms $\|A\|_2$ and $\|A\|_F$ as a function of eigenvalues of A, *i.e.*, find functions f and g such that $\|A\|_2 = f(\Lambda(A))$ and $\|A\|_F = g(\Lambda(A))$.

Problem 8. Suppose a matrix $A \in \mathbb{C}^{4\times 4}$ has a characteristic polynomial $\lambda^4 = 0$. Determine all the possible Jordan canonical forms of A.

Problem 9. Compute the Jordan canonical form of the matrix

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

and the associated nonsingular matrix X such that $A = XJX^{-1}$.

Problem 10. Suppose A and B are symmetric positive definite. Show that the eigenvalues of AB are still positive, even when AB is not symmetric.

Problem 11. Let $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{S}^n$. Suppose the matrix $\begin{bmatrix} 0 & A \\ A^T & B \end{bmatrix}$ is positive semidefinite. Show that $A = 0$ and B is positive semidefinite.

Problem 12. Let $A = \begin{bmatrix} I & X \\ 0 & -I \end{bmatrix}$, where $X \in \mathbb{R}^{m\times n}$ is arbitrary. Show that $e^{tA} = \begin{bmatrix} eI & (\sinh 1)X \\ 0 & (1/e)I \end{bmatrix}$.

Problem 13. *(Exponential of skew-symmetric matrices.)* In general, there is no closed-form formula for the exponential e^A of a matrix A, but for skew-symmetric matrices of dimension 2 and 3, there are explicit formulas.

(a) Denote the 2×2 skew-symmetric matrix by $A = \theta J$ with $J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. Show that

$$e^A = (\cos\theta) I_2 + (\sin\theta) J = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

(b) For an $n \times n$ skew-symmetric matrix B, *i.e.*, $B^T = -B$, show that $Q = e^B$ is a rotation matrix, *i.e.*, an orthogonal matrix with $\det Q = 1$.

(Part (a) is a special case of the following result. Let $C = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$ with $\alpha, \beta \in \mathbb{R}$. Then

$$e^{tA} = \begin{bmatrix} e^{\alpha t}\cos\beta t & e^{\alpha t}\sin\beta t \\ -e^{\alpha t}\sin\beta t & e^{\alpha t}\cos\beta t \end{bmatrix}.$$

But the matrix C is no longer skew-symmetric when $\alpha \neq 0$.)

Problem 14. Let $A, B \in \mathbb{R}^{n\times n}$ be two orthogonal matrices. Is $A \otimes B$ also an orthogonal matrix?