5

LINEAR ALGEBRA AND MATRIX ANALYSIS

Notations:

- $\mathbb{F}$: real or complex numbers.
- $\mathbb{Q}$: rational numbers.
- $\mathbb{Z}$: integer numbers.
- $\mathbb{P}$: positive numbers.
- $\mathcal{P}_n$: polynomial of degree of n.
- $\mathbb{N}$: natural numbers.
- $\mathcal{R}(A)$: the range of matrix A.
- $\mathcal{N}(A)$: the null space of matrix A.
- V: vector space.
- $det(A)$: the determinant of matrix A.
- $rank(A)$: the rank of matrix A.
- $\rho(A)$: the spectral radius of matrix A.
- $Tr(A)$: the trace of matrix A.

5.1 Theory for system of linear equations

5.1.1 Overview

In this section, we study the properties of solutions to a system of linear equations in the form of

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{12}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{12}x_2 + \cdots + a_{mn}x_n &= b_m. \end{aligned}$$

which can be written more compactly via matrix and vector notations, given by

$$Ax = b,$$

where $A \in \mathbb{R}^{m\times n}, x \in \mathbb{R}^n, b \in \mathbb{R}^m$. Systems with $b = 0$ are called **homogeneous systems**, and systems with $b \neq 0$ are called **non-homogeneous systems**.

In the following sections, we will first go over solution properties for homogeneous systems and non-homogeneous systems. Then we will examine matrix approach to solving systems of linear equations. Finally, we examine the numerical error in computational approach to these systems and characterize errors by condition number.

5.1.2 Homogeneous systems

Lemma 5.1.1 (solutions to homogeneous systems). *Given a system of linear equations given by $Ax = 0, A \in \mathbb{R}^{m\times n}$, there are exactly two possibilities:*

1. *unique solution of zero $x = 0$.*
2. *infinitely many solutions (including $x = 0$).*

Moreover, if $m < n$, there are exactly ***one possibility: infinitely many solutions****.*

Proof. (1) 0 vector is always one solution; (2) For the proof of infinitely many solutions, it can be showed using linear map theory(rank-nullity theorems) later. A can be viewed as a linear map from larger space to smaller space, and the null space of A has dimensionality greater than 0. □

Example 5.1.1. Consider $Ax = 0$ with $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

There is only one unique zero solution of $x = (0, 0)$.

Example 5.1.2. Consider $Ax = 0$ with $A = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. Clearly, x of the form $x = (\beta, 0), \beta \in \mathbb{R}$ is a solution. Therefore $Ax = 0$ has infinitely many solutions.

Example 5.1.3. Consider $Ax = 0$ with $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. Clearly, x of the form $x = (0, \alpha, \beta), \alpha, \beta \in \mathbb{R}$ is a solution. Therefore $Ax = 0$ has infinitely many solutions.

Lemma 5.1.2 (solutions form subspace nature). *Consider a system of linear equations given by $Ax = 0, A \in \mathbb{R}^{m \times n}$. if it has more than one solutions, then it has infinitely many solutions and all the solutions form a subspace, called **null space**.*

Proof. Use linearity of A. Let x_1 and x_2 be the solutions such that $Ax_1 = Ax_2 = 0$. Then for any $a_1, a_2 \in \mathbb{R}$, $a_1 x_1 + a_2 x_2$ is a solution. □

5.1.3 Non-homogeneous systems

Homogeneous systems always have at least one zero solution. But for nonhomogeneous systems, it is possible that no solution exists. A system of m linear equations in n unknowns,i.e., $Ax = b$ is said to be a **consistent** system if it possesses at least one solution; If there are no solutions, the system is said to be **inconsistent** system.

Example 5.1.4. Consider $Ax = 0$ with

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, b = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Clearly, x of the form $x = (\beta, 0), \beta \in \mathbb{R}$ is a solution. Therefore this is a inconsistent system.

From vector space perspective, $Ax = b$ is consistent if b lies in the subspace spanned by columns in A. We have following summary.

Lemma 5.1.3 (consistence criterion). *The system of linear equations $Ax = b, A \in \mathbb{R}^{m\times n}$ is consistent if one of the following is satisfied:*

- *$rank[A|b] = rank[A]$.*
- *A is full row rank, then $rank[A|b] = rank[A] = m$.*
- *b can be constructed via a linear combination of basic columns in A.*

It is also straight forward to arrive at the following properties regarding solutions to non-homogeneous systems.

Theorem 5.1.1 (solutions to non-homogeneous systems). *Consider a system of linear equations given by $Ax = b, A \in \mathbb{R}^{m\times n}$, there are exactly three possibilities:*[1]

1. *one unique solution if $Ax = 0$ only has zero vector as the solution.*
2. *no solutions if inconsistent.*
3. *infinitely many solutions if $Ax = 0$ has infinitely many solutions; that is, the solution will form an **affine set/space**.*

Corollary 5.1.1.1 (uniqueness criterion). *The system of linear equations $Ax = b, A \in \mathbb{R}^{m\times n}$ has a unique solution if the following conditions are satisfied:*

- *$rank[A|b] = rank[A] = n$.*
- *the homogeneous system $Ax = 0$ only have the trivial solution of $x = 0$.*

5.1.4 Overdetermined vs. underdetermined systems

In $Ax = b, A \in \mathbb{R}^{m\times n}$, it is **overdetermined** if $m > n$; it is underdetermined if $m < n$. Note that **overdetermined/underdetermined is not related to consistence**. Therefore, usually, there are six types of linear equation systems:

1. underdetermined and consistent
2. underdetermined and inconsistent
3. exactly determined and consistent
4. exactly determined and consistent
5. overdetermined and consistent
6. overdetermined and inconsistent

Lemma 5.1.4 (solution properties of overdetermined system). *An overdetermined system $Ax = b$ will have three possibilities:*

1. *no solution if inconsistent;*
2. *unique solution if consistent $Ax = 0$ only has the trivial solution;*
3. *infinitely many solutions if consistent $Ax = 0$ has infinitely many solutions.*

Lemma 5.1.5 (solution properties of underdetermined system). *An underdetermined system $Ax = b$ will have twe possibilities:*

1. *no solution if inconsistent;*
2. *infinitely many solutions if consistent*

Proof. note that $Ax = 0, m < n$ has infinitely many solution. □

5.1.5 Solution methods

Lemma 5.1.6 (orthogonal projection and rank-nullity decomposition). *Let $A \in \mathbb{R}^{m\times n}, rank(A) = n \leq m$. Define*

$$P \triangleq A(A^TA)^{-1}A^T.$$

It follows that

- *$P \in \mathbb{R}^{m\times m}$ and $rank(P) = n$.*
- *P is an orthogonal projection matrix and $Px \in \mathcal{R}(A)$.*
- *$I - P$ is an orthogonal projection matrix, $rank(I - P) = m - n$, and $(I - P)x \in \mathcal{R}(A)$.*
- *consider $b \in \mathbb{R}^m$, and let $b = b_{\mathcal{R}} + b_{\mathcal{N}}$ be the rank-nullity decomposition [Corollary 5.4.4.1] such that $b_{\mathbb{R}} \perp b_{\mathcal{N}}$. Then*

$$b_{\mathcal{R}} = Pb, b_{\mathcal{N}} = (I - P)b.$$

Proof. See subsection 5.5.2,Lemma 5.14.2. □

Lemma 5.1.7 (tall thin matrix, full column rank). *Let $A \in \mathbb{R}^{m\times n}, rank(A) = n \leq m$ and $b \in \mathbb{R}^m$. The minimum 2-norm error solution to $Ax = b$ is given by*

$$x^* = (A^TA)^{-1}A^Tb.$$

- *If* $b \in \mathcal{R}(A)$, *then* $\min_{x\in\mathbb{R}^n} \|Ax - b\|_2^2 = 0$;*that is* $Ax^* = b$.
- *If* $b \notin \mathcal{R}(A)$, *then* $\min_{x\in\mathbb{R}^n} \|Ax - b\|_2^2 = \|b_\mathcal{N}\|_2^2$;*that is* $Ax^* = b_\mathcal{R} \neq b$.

Proof. (1) Note that x^* can be obtained by using the first-order optimality condition of $\min_{x\in\mathbb{R}^n} \|Ax - b\|_2^2$. Note that when $b \in \mathcal{R}(A)$, $Pb = b$. (2) When $b \notin \mathcal{R}(A)$, we have

$$\|Ax - b\|_2^2 = \|Ax - b_\mathcal{R} - b_\mathcal{N}\|_2^2 = \|Ax - b_\mathcal{R}\|_2^2 + \|b_\mathcal{N}\|_2^2$$

where we have used the property that $b_\mathcal{N} \in \mathcal{N}(A^T)$, and $\mathcal{N}(A^T) \perp \mathcal{R}(A)$. □

Lemma 5.1.8 (fat short matrix, full row rank). *Let* $A \in \mathbb{R}^{m\times n}, rank(A) = m \leq n$ *and* $b \in \mathbb{R}^m$. *Further assumes that* A *can be partitioned as* $A = [A_1 A_2]$ *such that* A_1 *contains* m *linearly independent columns.*

- *The solution set to* $Ax = b$ *is given by*

$$x = x_p + Ny.$$

where $N \in \mathbb{R}^{m\times(n-m)}$ *with columns being the basis of* $\mathcal{N}(A)$ *and* x_p *a particular solution given by*

$$x_p = \begin{bmatrix} x_p^1 \\ 0 \end{bmatrix}, x_p^1 = (A_1^T A_1)^{-1} A_1^T b.$$

- *The solution with the minimum 2-norm length* x_m *is given by*

$$x_m = (I - P_\mathcal{N})x_p = P_\mathcal{R} x,$$

where x *is an arbitrary element in the solution set,* $P_\mathcal{N} = N(N^T N)^{-1} N^T x, P_\mathcal{R}$ *is the projection matrix onto* $\mathcal{R}(A^T)$. *Note that* $P_\mathcal{R} \neq A_1(A_1^T A_1)^{-1} A_1^T$. *This solution is equivalent to the unique minimizer of*

$$\min_{x\in\mathbb{R}^n} \|x\|_2^2, \text{ subject to } Ax = b.$$

- ***(Psedudoinverse:)****The solution with the minimum 2-norm length* x_m *is given by*

$$x_m = A^T (AA^T)^{-1} b.$$

Proof. (1)Partition the original linear equation as

$$[A_1\ A_2] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = b.$$

Then $A_1x_1 = b - A_2x_2$. Set $x_2 = 0$ and we use the full column rank thin tall linear equation result to solve x_1. (2) To calculate the minimum length element, we seek the solution to

$$\min_y \frac{1}{2}\|x + Ny\|_2^2.$$

The first order condition gives

$$y^* = -(N^TN)^{-1}N^Tx.$$

(3) It can be showed that x_m is the solution by

$$Ax_m = AA^T(AA^T)^{-1}b = b.$$

To show x_m is the solution with minimum length, we can solve

$$\min_{y\in\mathbb{R}^n} f(y) = \left\|A^T(AA^T)^{-1}b + Ny\right\|^2.$$

The objective function given by (let $R = A^T(AA^T)^{-1}$)

$$\begin{aligned} f &= p^TR^TRp + y^TN^TNy + 2y^TN^TRp \\ &= p^TR^TRp + y^TN^TNy \end{aligned}$$

which will achieve minimum value at $y = 0$. That is, x_m is the minimum length solution.

□

Corollary 5.1.1.2 (tall thin matrix, not full column rank). *Let $A \in \mathbb{R}^{m\times n}$, $rank(A) = r < n \leq m$ and $b \in \mathbb{R}^m$. Further assumes that A can be partitioned as $A = [A_1\ A_2]$ such that A_1 contains r linearly independent columns. The minimum 2-norm error solution to $Ax = b$ has the following properties:*

- *If $b \in \mathcal{R}(A)$, the the solution to $Ax = b$ exists but always not unique.*
- *If $b \in \mathcal{R}(A)$, then $\min_{x\in\mathbb{R}^n}\|Ax - b\|_2^2 = 0$;there exists a set of minimizers x^* such that $Ax^* = b$. The full set of solution/minimizers are $x = x_p + Ny$, where*

$$x_p = \begin{bmatrix} x_p^1 \\ 0 \end{bmatrix}, x_p^1 = (A_1^TA_1)^{-1}A_1^Tb.$$

where $N \in \mathbb{R}^{m\times(n-r)}$ with columns being the basis of $\mathcal{N}(A)$.

Corollary 5.1.1.3 (fat short matrix, not full row rank). *Let $A \in \mathbb{R}^{m\times n}, rank(A) = r < m \leq n$ and $b \in \mathbb{R}^m$. Further assumes that A can be partitioned as $A = [A_1\ A_2]$ such that A_1 contains r linearly independent columns. The minimum 2-norm error solution to $Ax = b$ has the following properties:*

- *If $b \in \mathcal{R}(A)$, the the solution to $Ax = b$ exists but always not unique.*
- *If $b \in \mathcal{R}(A)$, then $\min_{x\in\mathbb{R}^n} \|Ax - b\|_2^2 = 0$;there exists a set of minimizers x^* such that $Ax^* = b$. The full set of solution/minimizers are $x = x_p + Ny$, where*

$$x_p = \begin{bmatrix} x_p^1 \\ 0 \end{bmatrix}, x_p^1 = (A_1^T A_1)^{-1} A_1^T b.$$

- *If $b \notin \mathcal{R}(A)$, then $\min_{x\in\mathbb{R}^n} \|Ax - b\|_2^2 = \|b_{\mathcal{N}}\|_2^2$;there exists a set of minimizer x^* such that $Ax^* = b_{\mathcal{R}} \neq b$. The full set of minimizers are $x = x_p + Ny$, where*

$$x_p = \begin{bmatrix} x_p^1 \\ 0 \end{bmatrix}, x_p^1 = (A_1^T A_1)^{-1} A_1^T b_{\mathcal{R}}.$$

Lemma 5.1.9 (minimum error minimum length solution via SVD theory). *Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. Let A have SVD decomposition Theorem 5.9.1 given by*

$$A = [U_1\ U_2] \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}.$$

Then the minimizers of

$$\min_{x\in\mathbb{R}^n} \|Ax - b\|_2^2$$

is a set

$$x^* = V_1 \Sigma^{-1} U_1^T b + V_2 y.$$

Among the set, the element with the minimum 2-norm length is

$$x_m^* = V_1 \Sigma^{-1} U_1^T b.$$

Proof. (1)

$$\begin{aligned}\|Ax-b\|_2^2 &= \left\|[U_1\ U_2]Ax-[U_1\ U_2]b\right\|_2^2 \\ &= \left\|\begin{bmatrix}\Sigma & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}V_1^T\\ V_2^T\end{bmatrix}x-\begin{bmatrix}U_1^Tb\\ U_2^Tb\end{bmatrix}\right\|_2^2 \\ &= \left\|\begin{bmatrix}\Sigma V_1^Tx-U_1^Tb\\ -U_2^Tb\end{bmatrix}\right\|_2^2\end{aligned}$$

We can solve x from $\Sigma V_1^T x = U_1^T b$. Also note that $V_2 y$ does not contribute the objective function. (2) Use the minimum length result in Lemma 5.1.8, the minimum length is given by

$$x_m^* = V_1V_1^Tx^* = V_1V_1^TV_1\Sigma^{-1}U_1^Tb + V_1V_1^TV_2y = V_1\Sigma^{-1}U_1^Tb,$$

where we use the fact that $V_1V_1^T$ is the orthogonal projection matrix to $\mathcal{R}(A^T)$, and $V_1^TV_2 = 0$. □

Theorem 5.1.2 (solution for general linear system, recap). *Let $A \in \mathbb{R}^{m\times n}$ with SVD $A = U\Lambda V^T$ and rank$(A) = r$. Let $A^+ = V\Lambda^+U^T$ be its pseudoinverse. If the linear system $Ax = b$ has solution, then the solution is given by*

$$x = A^+b + (I_n - A^+A)z, z \in \mathbb{R}^n.$$

where $I_n - A^+A$ being the $\mathcal{N}(A)$ basis matrix. Among all solutions, the minimum norm/length solution is A^+b.

If $Ax = b$ does not have a solution, then

$$x = A^+b + (I_n - A^+A)z, z \in \mathbb{R}^n.$$

is the solution set of minimum error, with A^+b being the minimum norm/length solution.

Proof. Theorem 5.14.2. □

5.1.6 Error bounds in numerical solutions

5.1.6.1 *Condition number*

In real world, systems of linear equations are solved by computers. Since computers use finite bit binary representation and thus inherently inaccurate in the computation process. By characterizing error in terms of matrix properties, we will be cautious when we use computers to solve certain types of linear equations.

We first introduce the concept of condition number associated with a matrix.

Definition 5.1.1 (condition number). *For a square matrix A, we define the condition number as*

$$cond(A) = \begin{cases} \|A\| \left\|A^{-1}\right\| \\ \infty, A \textbf{ is singular} \end{cases},$$

where $\|\cdot\|$ is the matrix norm [section 5.13].

Based on properites of matrix norm [section 5.13], we can easily derive the following properties:

- $cond(I) = 1$
- $cond(A) \geq 1$
- $cond(\alpha A) = cond(A), \forall \alpha \neq 0$
- If D is diagonal, then
$$cond(D) = \frac{\max|d_{ii}|}{\min|d_{ii}|}$$

5.1.6.2 *Error bounds*

Consider using a computer to solve $Ax = b$. Let $\tilde{b}$ be the computer approximate representation of b, and $\tilde{x}$ be the computer solution such that $A\tilde{x} = \tilde{b}$.

Then condition number can help us characterize the upper bound of $\|x - \tilde{x}\|$, given by the following theorem.

Theorem 5.1.3. *If A is nonsingular, $Ax = b$, and $A\tilde{x} = \tilde{b}$, then*

$$\frac{\|\tilde{x} - x\|}{x} \leq cond(A)\frac{\left\|\tilde{b} - b\right\|}{b}$$

Proof. Let $\Delta x = \tilde{x} - x, \Delta b = \tilde{b} - b$, then

$$\frac{\|\Delta x\|}{\|x\|} = \frac{A^{-1}\Delta b}{b} \leq \frac{\|\Delta b\|\left\|A^{-1}\right\|}{\|x\|} \leq cond(A)\frac{\left\|\tilde{b} - b\right\|}{b}$$

where we use $\|b\| \leq \|A\|\|x\|$ in the last step. □

5.2 Vector space theory

5.2.1 Vector space

Definition 5.2.1 (vector space). *Let $\mathbb{F}$ be a field.* ***A vector space over a field*** *$\mathbb{F}$ is a set V together with two operation called addition and scalar multiplication The addition is a function $+ : V \times V \to V$ such that $x + y = x + y \in V, x, y \in V$; the scalar multiplication is a function $\times : \mathbb{F} \times V \to V$ such that $\lambda \times x = \lambda x, \lambda \in F, x \in V$. The addition and the scalar multiplication are required to satisfy the following axioms.*

1. *(addition associativity) $(u + v) + w = u + (v + w), \forall u, v, w \in V$*
2. *(addition community) $u + v = v + u, \forall u, v \in V$*
3. *(additive identity) there exist an element $0 \in V$ such that $0 + v = v, \forall v \in V$*
4. *(additive inverse element) for each $v \in V$, there exists a $u \in V$, such that $u + v = 0$*
5. *(scalar multiplication identity) $1u = u, \forall u \in V$*
6. *(associativity of scalar multiplication) $r(su) = (rs)u, \forall r, s \in \mathbb{F}, u \in V$*
7. *(distributivity of scalar sums) $(r + s)u = ru + su, \forall r, s \in \mathbb{F}, u \in V$*
8. *(distributivity of vector sums) $r(u + v) = ru + rv, , \forall r \in \mathbb{F}, u, v \in V$*

Note that elements of V are called vectors and elements of $\mathbb{F}$ are called scalars.

Example 5.2.1 (common vector space examples).

- Let F be a field, then F is a vector space over F with addition and multiplication defined in F. Therefore, $\mathbb{R}$ is a vector space over R; $\mathbb{Q}$ is a vector space over $\mathbb{Q}$.
- Let F be a field, and $F' \subseteq F$ is a vector space over F'. Therefore, $\mathbb{R}$ is a vector space over $\mathbb{Q}$; $\mathbb{C}$ is a vector space over $\mathbb{R}$.
- $\mathbb{R}^n$ is vector space over $\mathbb{R}$.
- The set
$$M_{m\times n}F = \{m \times n \text{ matrices with entries in } F\}$$
is an F-vector space equipped with component-wise addition and common scalar multiplication between scalar and matrices.

Example 5.2.2 (function spaces as vector space). Note that the first two are infinite dimensional vector space.[2]

- Let $p >= 1$ be a real number. The function space $L^p([a,b])$ of all real or complex-valued measurable functions defined by

$$L^p([a,b]) = \{f : [a,b] \to \mathbb{F} | \int_a^b |f(t)|^p dt < \infty\}$$

is a vector space with point-wise addition and scalar multiplication.
- The function space $\mathcal{C}([a,b])$ of all continuous, real- or complex-valued functions defined on the interval $[a,b]$:

$$\mathcal{C}([a,b]) = \{f : [a,b] \to \mathbb{F} | f is continuous\}$$

is a vector space with point-wise addition and scalar multiplication.
- Let $p >= 1$ be a real number. The function space $l^p(\mathbb{Z})$ of infinite, pth order summable sequences or discrete function $f[n]$ (i.e. f only take discrete integer as argument) defined by

$$l^p(\mathcal{Z}) = \{f([n]) | \sum_{n=-\infty}^{n=+\infty} |f[n]^p < \infty\}$$

$$L^p([a,b]) = \{f : [a,b] \to \mathbb{F} | \int_a^b |f(t)|^p dt < \infty\}$$

is a vector space with component-wise addition and scalar multiplication.

Other similar examples include:

- $\mathcal{C}^k(\Omega)$ of all k times continuously differentiable functions
- $\mathcal{C}^\infty(\Omega)$ of all smooth functions

5.2.2 Subspace

Definition 5.2.2 (closed). *A subset U of a vector space V is said to be* ***closed under addition and scalar multiplication*** *if*

1. $u_1 + u_2 \in U, \forall u_1, u_2 \in U$
2. $\lambda u \in U, \forall u \in U, \lambda \in F$

Definition 5.2.3 (subspace). *A subset U of a vector space V is called a subspace of V if U is itself a vector space relative to addition and scalar multiplication inherited from V.*

Theorem 5.2.1 (subspace). *If V is a vector space and U a subset of V which is* ***nonempty, and closed*** *under addition and scalar multiplication, the U is subspace of V.*

Proof. The key is to prove the existence of additive identity and inverse. Since the set U is nonempty and closed, then $0u = 0 \Rightarrow 0$ exists. The additive inverse: $-1u$. For other properties of the addition and multiplication operation, they will inherit from V. □

Theorem 5.2.2 (subspace condition, alternative). *[3, p. 18] If V is a vector space and U a subset of V, then U is the subspace if and only if it is closed under addition and scalar multiplication and U contains the additive identity 0 of V.*

Example 5.2.3.

- If V is a vector space, then V is a subspace of V. The set $\{0\}$ is called the **zero subspace** of V.
- $\mathbb{R}^2$ is vector space over $\mathbb{R}$. The set L defined by
$$L = \{(x, y) \in \mathbb{R}^2 | y = mx\}, m \in \mathbb{R}^{++}$$
is a subspace of $\mathbb{R}^2$.
- $\mathbb{R}^2$ is vector space over $\mathbb{R}$. The set L defined by
$$L = \{(x, y) \in \mathbb{R}^2 | y = x + 1\}$$
is not a subspace of $\mathbb{R}^2$. Take $(x_1, y_1) \in L, (x_2, y_2) \in L$, then
$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2) \notin L,$$
since $y_1 + y_2 \neq x_1 + x_2 + 1$.

5.2.3 Sum and direct sum

Definition 5.2.4 (sum of subsets). *Suppose $U_1, U_2, ..., U_m$ are subsets of V. The* ***sum*** *of $U_1, U_2, ..., U_m$ is the set*
$$U_1 + U_2 + + ... + U_m = \{u_1 + u_2 + ... + u_m : u_1 \in U_1, u_2 \in U_2, ..., u_m \in U_m\}$$

Lemma 5.2.1 (sum of subspace is the smallest containing subspace). *[3, p. 20] Suppose $U_1, U_2, ..., U_m$ are subspaces of V. Then sum $U_1 + U_2 +, ..., U_m$ is the smallest subspace of V containing $U_1, U_2, ..., U_m$.*

Proof: First, $U_1 + U_2 +, ..., U_m$ is a subspace; second, every subspace in V containing $U_1, U_2, ..., U_m$ will contain $U_1 + U_2 +, ..., U_m$.

Definition 5.2.5 (direct sum). *Suppose $U_1, U_2, ..., U_m$ are subspaces of V: The sum $U_1 + U_2 +, ..., U_m$ is called a direct sum if each element of $U_1 + U_2 +, ..., U_m$ can be written* ***uniquely*** *as $u_1 + u_2 + ... + u_m, u_i \in U_i, \forall i$. Then $U_1 + U_2 +, ..., U_m$ will be written as $U_1 \oplus U_2 \oplus, ..., U_m$.*

Remark 5.2.1 (not every sum is direct sum). Not every sum of subspaces are direct sum due to the possible linear dependence of basis of subspaces.

Lemma 5.2.2 (direct sum criterion for sum to be direct sum). *[3, p. 23]Suppose $U_1, U_2, ..., U_m$ are subspaces of V. Then sum $U_1 + U_2 +, ..., U_m$ is a direct sum if and only if the only way to write 0 as a sum $u_1 + u_2 + ... + u_m, u_i \in U_i$ is to set each $u_j = 0$.*

Or equivalently, sum $U_1 + U_2 +, ..., U_m$ is a direct sum if and only if $u_1, u_2, ..., u_m, u_i \in U_i$ are linearly independent.

Proof. (1) suppose $U_1 + U_2 +, ..., U_m$ is a direct sum, then the definition of direct sum implies that there is an unique way to write 0 as a sum of $u_1 + u_2 + ... + u_m$, since $u_i = 0$ will satisfy, then this is the unique way. (2) Let $u \in U_1 + U_2 +, ..., U_m$ be expressed as $u = u_1 + u_2 + ... + u_m = v_1 + v_2 + ... + v_m$(i.e. two ways). Now we prove that if the only way to write 0 as a sum $u_1 + u_2 + ... + u_m$ is to set each $u_j = 0$ will imply that $u_i = v_i$ in the above expression. We have $0 = \sum(u_i - v_i), (u_i - v_i) \in U_i$, because the only way is to set $u_i - v_i = 0$, then we have $u_i = v_i$. □

Corollary 5.2.2.1. *Suppose U and W are subspaces of V. Then V and W is a direct sum if and only if $U \cap W = \{0\}$.*

Theorem 5.2.3 (dimensions of a sum). *[3, p. 47][1, p. 214] If U_1 and U_2 are subspaces of a finite-dimensional vector space, then*

$$dim(U_1 + U_2) = dim(U_1) + dim(U_2) - dim(U_1 \cap U_2)$$

Moreover, if it is a direct sum,

$$dim(U_1 \oplus U_2) = dim(U_1) + dim(U_2)$$

Proof. Let $\{z_1, z_2, ..., z_t\}$ be the basis of $U_1 \cap U_2$, let $B_X = \{z_1, ..., z_t, x_1, ..., x_m\}$ be the basis of U_1, let $B_Y = \{z_1, ..., z_t, y_1, ..., y_m\}$ be the basis of U_2. Then $B_X \cup B_Y$ will span $U_1 + U_2$. It can be shown that $B_X \cup B_Y$ are linearly independent □

5.2.4 Basis and dimensions

Definition 5.2.6 (linear independence). *We say that vectors $v_1, v_2, ..., v_n \in V$ are linearly independent if the **only** solution of $\sum_{i=1}^{n} c_i v_i = 0$ is $c_i = 0, \forall i$.*

Definition 5.2.7 (span). *The set of all linear combinations of a list of vectors $v_1, v_2, ..., v_n$ in V is called the **span** of $v_1, v_2, ..., v_n$, denoted as $span(v_1, v_2, ..., v_n)$, given as*

$$span(v_1, v_2, ..., v_n) = \{a_1 v_1 + a_2 v_2 + ... + a_n v_n : a_1, a_2, ..a_n \in \mathbb{F}\}$$

Lemma 5.2.3 (polynomials are linear independent).

- *The polynomials $1, t, t^2, ..., t^n$ are linearly independent.*
- *The matrix*

$$P = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \ddots & \vdots & \\ 1 & x_m & x_m^2 & \cdots & x_m^n \end{bmatrix}, m \geq (n+1).$$

and the matrix P has independent columns.

Proof. (1) Suppose we have some weights $C_0, C_1, ..., C_n \in \mathbb{R}$ such that

$$C_0 + C_1 t + C_2 t^2 + \cdots + C_n t^n = 0, \forall t$$

If some of the coefficients are non-zero, then this is a polynomial with degree $\leq n$, therefore has at most n roots(based on the fundamental theorem of algebra [Theorem 4.3.6]. However, here has ∞ roots(the above equation holds for all t). Therefore, all coefficients have to be 0. (2) Suppose we have some weights $C_0, C_1, ..., C_n \in \mathbb{R}$ such that

$$C_0 + C_1 t + C_2 t^2 + \cdots + C_n t^n = 0, t = x_1, x_2, ..., x_m.$$

If some of the coefficients are non-zero, then this is a polynomial with degree $\leq n$, therefore has at most n roots(based on the fundamental theorem of algebra [Theorem 4.3.6]. However, here has $m \geq n+1$ roots(the above equation holds for $t = x_1, x_2, ..., x_m$). Therefore, all coefficients have to be 0. □

Definition 5.2.8 (basis and dimension).

- *A linearly independent set of vectors that span the vector space V is called the **basis** for V.*
- *The cardinality of the basis is called the **dimension** of the vector space.*

Definition 5.2.9 (finite and infinite dimensional vector space). *If a vector space has a finite-sized basis, then it is a **finite dimensional vector space**; otherwise it is a infinite dimensional vector space,i.e., no finite-sized basis can be found to span the vector space.*

Example 5.2.4.

- $\mathbb{R}^n$ has dimension n since it has standard basis $\{e_1, e_2, ..., e_n\}$.
- When $\mathbb{C}$ is a vector space over $\mathbb{R}$, it has basis $\{1, i\}$ and dimension 2; When $\mathbb{C}$ is a vector space over $\mathbb{C}$, it has basis $\{1\}$ and dimension 1; When $\mathbb{C}$ is a vector space over $\mathbb{Q}$, it has dimension ∞(we can finite-sized basis to span C when the scalars are taken from $\mathbb{Q}$).

Lemma 5.2.4 (Basis is maximum linearly independent set). *Suppose that $S = \{v_1, v_2, v_3, \ldots, v_t\}$ is a finite set of vectors which spans the vector space V. Then any set of $t+1$ or more vectors from V is linearly dependent.*

Proof. Let A be the basis of size m, let B be the basis of size n. Suppose $m > n$, we have $A = BM, M \in \mathbb{F}^{n,m}$ since A can be spanned by B. Consider the linear equation $Mx = 0$, it will has a non-trivial solution since there are more variables than equations. Then $Ax = BMx = 0$, that is a non-trival linear combination of A columns leads to 0. Therefore, columns in A are linearly dependent. □

Theorem 5.2.4 (All basis have the same length). *Let V be a finite-dimensional vector space, then all its basis, that is, linearly independent set that span the space, will have the same length.*

Proof. Let A be the basis of size m, let B be the basis of size n. Suppose $m > n$, then based on above lemma, A cannot be an independent set. □

Theorem 5.2.5 (extending linearly independent set). *Let V be a vector space, let $S = \{v_1, v_2, ..., v_m\}$ be a linearly independent set. Then if $v \in V$, but $v \notin span(S)$, then the set $S \cup v$ is linearly independent.*

Proof. Suppose linearly dependent, then v can be expressed as linear combination, which contradicts $v \notin span(S)$. $\square$

Theorem 5.2.6 (proper subspace has smaller dimension). *If U, W are subspaces of vector space V, and $U \subsetneq W$, then $dim(U) < dim(W)$.*

Proof. First, we must have $dim(U) \leq dim(W)$ since basis of W will span U. To show $dim(U) < dim(W)$, suppose $dim(U) = dim(W) = t$, then there exist t linearly independent vectors in U and therefore in W, therefore spanning W, therefore $U = W$, which is a contradiction. $\square$

Theorem 5.2.7 (equal dimension implies equal subspace). *If U, W are subspaces of vector space V, and $U \subseteq W$, if $dim(U) = dim(W)$, then $U = W$.*

Proof. suppose $U \neq W$, then there exists $w \in W$ such that $w \notin U$, let B be the basis of U, then $B \cup w (\in W)$ will form a linear independent set(from above theorem), then $dim(W) > dim(U)$, which is contradiction. $\square$

Corollary 5.2.7.1. *If U is a subspace of vector space V, and $dim(U) = dim(V)$, then $U = V$.*

Remark 5.2.2. The above two theorem will be important in proving 'onto' property of linear maps.

5.2.5 Complex vector space vs. real vector space

Lemma 5.2.5 (same dimensionality of real and complex vector space). *The complex vector space $\mathbb{C}^n$ has a basis $\{v_i\}$ of size n with all real entries. In other words, $\mathbb{C}^n$ and $\mathbb{R}^n$ have the same dimensionality.*

Proof. Consider a standard basis $E = \{e_i\}$ that can span $\mathbb{R}^n$. For a complex number c, it can always be written as $c = a + bi, a, b \in \mathbb{R}^n$, then $c = Ev_a + Ev_b i = E(v_a + v_b i)$. $\square$

Remark 5.2.3. This lemma shows that for a complex vector space $\mathbb{C}^n$, it is always possible to choose a set of real-valued basis.

Lemma 5.2.6 (convert complex-valued basis to real-valued basis). *Given a complex-valued basis $\{u_i\}$ for $\mathbb{C}^n$, then its conjugate $\{\overline{u_i}\}$ is also a basis. Moreover, a real-valued basis can be created by $\{u_i + \overline{u_i}\}$*

Proof. Suppose $\{\overline{u_i}\}$ is not linearly independent, then there exist nonzero $a_1, a_2, ..., a_n$ such that

$$a_1\overline{u_1} + a_2\overline{u_2}... + a_n\overline{u_n} = 0.$$

then this set of coefficients $\{\overline{a_1}, ..., \overline{a_n}\}$ will also make

$$\overline{a_1}u_1 + \overline{a_2}u_2... + \overline{a_n}u_n = 0.$$

□

Note 5.2.1 (caution!).

- $\mathbb{C}^n$ over $\mathbb{R}$ is a vector space, but this vector space cannot have real-valued basis. (because complex-valued vectors cannot be spanned)
- Moreover, the dimensionality is $2n$, with basis $\{e_1, ie_1, ...\}$

5.3 Linear maps & linear operators

5.3.1 Basic concepts of linear maps

Notations in this section

- $\mathbb{F}$ denotes $\mathbb{R}$ or $\mathbb{C}$
- V and W denote vector spaces over $\mathbb{F}$
- $\mathcal{L}(V,W)$ denotes all linear maps from V to W.

Definition 5.3.1 (linear map). *A **linear map** from V to W is a function $T : V \to W$ with the following properties:*

$$T(au+bv) = aT(u) + bT(v), \forall a,b \in \mathbb{F}, \forall u \in U, v \in W$$

Lemma 5.3.1 (linear map maps zero element to zero element). *Let $T \in \mathcal{L}(V,W)$, then $T(0_V) = 0_W$.*

Proof. $T(0_V) = T(a + -a) = T(a) - T(a) = 0_W$. □

Remark 5.3.1. This lemma provides a necessary condition for us to judge whether a function is a linear map or not.

Example 5.3.1. Let V,W be $\mathbb{R}$. Then the map $T(x) = 5x+3$ is not a linear map but $T(x) = 5x$ is a linear map.

Definition 5.3.2 (null space). *For $T \in \mathcal{L}(V,W)$, the null space of T is*

$$\mathcal{N}(T) = \{x \in V : Tx = 0\}$$

Lemma 5.3.2 (null space as a subspace). *The null space of a linear map $T \in \mathcal{L}(V,W)$ is a subspace of V.*

Proof. directly from linearity of T. □

Lemma 5.3.3 (zero null space is equivalent to 1-1). *Let $T \in \mathcal{L}(V, W)$, then T is injective(1-1) if and only if $\mathcal{N}(T) = \{0\}$*

Proof: (1) Suppose $Tx = Ty \Rightarrow T(x-y) = 0$, if $\mathcal{N}(T) = \{0\}$, then $x = y$, therefore 1-1; (2) The converse(1-1 implies nullity): let $v \in \mathcal{N}(T)$, then $T(v) = 0 = T(0) \Rightarrow v = 0$ due to 1-1. (another proof, suppose $dim(\mathcal{N}(T)) > 0$, then $T(x-y)$ cannot lead to $x = y$).

Definition 5.3.3 (range). *For $T \in \mathcal{L}(V, W)$, the range of T is defined as*

$$\mathcal{R}(T) = \{Tv, \forall v \in V\}$$

Definition 5.3.4 (surjective,onto). *A function $T : V \to W$ is called surjective/onto if its range equals W.*

Lemma 5.3.4. *The range of a linear map $T \in \mathcal{L}(V, W)$ is a subspace of W.*

Proof. directly from linearity of T. □

Lemma 5.3.5 (surjective criterion). *$T \in \mathcal{L}(V, W)$ is surjective if $dim(\mathcal{R}(T)) = dim(W)$*

Proof. directly from theorems in subspace that equality in dimension leads to equality in subspacesTheorem 5.2.7. □

Example 5.3.2 (Examples of linear maps/operators). Examples are [2]

- Operator $T : L^2([a, b]) \to L^2([a, b])$ defined as $Tf(t) = tf(t)$
- The differentiation operator $D : \mathcal{C}^1(\mathbb{R}) \to \mathcal{C}(\mathbb{R})$
- The integration operator $T : \mathcal{C}(\mathbb{R}) \to \mathcal{C}^1(\mathbb{R})$
- The trace operator.
- The convolution operator $H : L^1(\mathbb{R}) \to L^1(\mathbb{R})$
- The Laplacian $\Delta : \mathcal{C}^\infty(\mathbb{R}^n) \to \mathcal{C}^\infty(\mathbb{R}^n)$

  Counter-examples are:

- The determinant operator.

5.3.2 Fundamental theorem of linear maps

Theorem 5.3.1 (fundamental theorem of linear maps, Rank-nullity theorem). *[3, p. 62]Suppose V is finite-dimensional and $T \in \mathcal{L}(V, W)$, then $\mathcal{R}(T)$ is finite dimensional and*

$$dim(V) = dim(\mathcal{N}(T)) + dim(\mathcal{R}(T))$$

Proof. Denote $dim(V) = m + n, dim(\mathcal{N}(T)) = m$. Let $u_1, u_2, ..u_m$ be the basis of $\mathcal{N}(T)$, let $u_1, u_2, .., u_m, v_1, v_2, ..., v_n$ be the basis of V. Then for any $v \in V, v = \sum_{i=1}^{m} a_i u_i + \sum_{j=1}^{n} b_j v_j$, $Tv \in \mathcal{R}(T), Tv = \sum_{i=1}^{n} b_j T v_j$, therefore the $\mathcal{R}(T) \subset span(Tv_1, Tv_2, ..., Tv_n)$. Further, we need to prove $Tv_1, Tv_2, ..., Tv_n$ are linearly independent set: suppose it is not, then there are nonzero coefficients c_is such that

$$\sum_{i=1}^{n} c_i T v_i = 0 = T(\sum_{i=1}^{n} c_i v_i) = 0$$

which suggest $\sum_{i=1}^{n} c_i v_i \in \mathcal{N}(T)$, however, by assumption v_i is linearly independent of basis of $\mathcal{N}(T)$, therefore $\sum_{i=1}^{n} c_i v_i = 0$, however contradict v_i are linear independent. □

Remark 5.3.2 (relation to fundamental theorem of linear algebra)**.** For matrix, $dim(\mathcal{R}(A)) = dim(\mathcal{R}(A^T))$, therefore it is consistent with fundamental theorem of linear algebra.

Corollary 5.3.1.1.

- ***mapping into smaller space implies injective(non 1-1):*** *Suppose V and W are finite-dimensional vector space such that $dim(V) > dim(W)$, then no linear maps T from V to W is injective, that is $dim(\mathcal{N}(T)) > 0$.*
- ***mapping into larger space implies non-surjective(non onto):*** *Suppose V and W are finite-dimensional vector space such that $dim(V) < dim(W)$, then no linear maps T from V to W is surjective, that is $dim(\mathcal{R}(T)) < dim(W)$.*

Proof. (1)

$$\begin{aligned} dim(\mathcal{N}(T)) &= dim(V) - dim(\mathcal{R}(T)) \\ &\geq dim(V) - dim(W) > 0 \end{aligned}$$

where we use $dim(\mathcal{R}(T)) \leq dim(W)$. (2)

$$\begin{aligned} dim(\mathcal{R}(T)) &= dim(V) - dim(\mathcal{N}(T)) \\ &\leq dim(V) < dim(W) \end{aligned}$$

where we use $dim(\mathcal{N}(T)) \geq 0$. □

Remark 5.3.3 (application in linear equation theory). A simple application is underdetermined linear homogeneous system has infinitely many solutions.

Theorem 5.3.2 (existence of inverse linear map). *[3, p. 80]*

- *A linear map $T : V \to V$ has the following equivalent statement:*
  - *T has an inverse $T^{-1} : V \to V$*
  - *T is onto, i.e. $\mathcal{R}(T) = V$ [Lemma 5.3.3].*
  - *T is 1-1; or equivalently $\mathcal{N}(T) = \{0\}$.*
- *$T \in \mathcal{L}(V, W)$ has an inverse $T^{-1} \in \mathcal{L}(W, V)$ if and only if T is onto and 1-1.*

Proof. (1) (a) implies (b)(c) The existence of inverse requires that T is 1-1 and onto. (b) implies (c) use rank-nullity theorem [Theorem 5.3.1] such that $dim(V) = dim(\mathcal{N}(T)) + dim(\mathcal{R}(T))$ to prove. If T is 1-1, then $\mathcal{N}(T) = \{0\}$ [Lemma 5.3.3]. Therefore $dim(\mathcal{R}(T)) = dim(V) \Leftrightarrow \mathcal{R}(T) = W$. (c) implies (a)(b) use rank-nullity again we get $\dim(\mathcal{N}(T)) = 0$, implying 1-1. 1-1 and onto implies the existence of inverse. (2) forward: if T is onto and 1-1 then T^{-1} exists be definition. converse: if T^{-1} exists, then T^{-1} is a linear map [Lemma 5.3.6], therefore $T^{-1} \in \mathcal{L}(W, V)$. $\square$

5.3.3 Isomorphism

An special type of linear maps between V and W is **isomorphism**, whose is an bijective linear map.

Definition 5.3.5 (isomorphism). *An **isomorphism** between two vector spaces $V$and W is a map $f : V \to W$ that*

1. *f is one-to-one and onto (bijective)*
2. *preserves structures: If $v_1, v_2 \in V$ then*

$$f(v_1 + v_2) = f(v_1) + f(v_2)$$

and if $v \in V$ and $r \in \mathbb{F}$, then

$$f(rv) = rf(v)$$

We say V and W are **isomorphic** to each other if there exists an isomorphism $T : V \to W$.

Lemma 5.3.6 (basic properties of isomorphisms). *Consider a linear transformation T from V to W. We assume T is bijection such that T^{-1} exists.*

- *If T is an isomorphism, then so is T^{-1}.*

- *A linear transformation T from V to W is an isomorphism if and only if*

$$\mathcal{N}(T) = \{0\}, \mathcal{R}(T) = W.$$

- *Consider an isomorphism T from V to W. If $f_1, f_2, ..., f_n$ is a basis of V, then $T(f_1), T(f_2), ..., T(f_n)$ is a basis of W.*
- *If V and W are isomorphic, then $dim(V) = dim(W)$.*

Proof. (1) We first how that T^{-1} is linear. Consider f, g in V and k and m in $\mathbb{F}$. Then

$$\begin{aligned} T^{-1}(kf + mg) &= T^{-1}(TT^{-1}(kf) + TT^{-1}(mg)) \\ &= T^{-1}T(T^{-1}(kf) + T^{-1}(mg)) \\ &= T^{-1}(kf) + T^{-1}(mg) \\ &= kT^{-1}(f) + mT^{-1}(g) \end{aligned}$$

From the definition of function map, we know that $\mathcal{R}(T^{-1}) = V$. and T^{-1} is 1-1. (2) see Theorem 5.3.2. (3) For any g in W, there exists $T^{-1}(g)$ in V such that

$$T^{-1}(g) = c_1 f_1 + c_2 f_2 + \cdots + c_n f_n,$$

because f_is span V. Applying T on both sides we get

$$g = c_1 T(f_1) + c_2 T(f_2) + \cdots + c_n T(f_n),$$

that is, $T(f_1), T(f_2), ..., T(f_n)$ span W. To show that $T(f_1), T(f_2), ..., T(f_n)$ are linear independent, we consider a relation

$$b_1 T(f_1) + b_2 T(f_2) + \cdots + b_n T(f_n) = 0,$$

or

$$T(b_1 f_1 + b_2 f_2 + \cdots + b_n f_n) = 0.$$

Since $\mathcal{N}(T) = \{0\}$, we have

$$b_1 f_1 + b_2 f_2 + \cdots + b_n f_n = 0.$$

Further because $f_1, f_2, ... f_n$ are linear independent, we have $b_1 = ... = b_n = 0$. Therefore, $T(f_1), T(f_2), ..., T(f_n)$ are linear independent. (4) directly from (3). Note that dimension is the cardinality of the basis. $\square$

5.3.4 Coordinate map properties

An important isomorphism is the **coordinate map**. Let $B = \{b_1, b_2, ..., b_n\}$ be a basis of a vector space V such that **any** element $x \in V$ can be represented by

$$x = x_1 b_1 + x_2 b_2 + \cdots + x_n b_n.$$

Then the **coordinate vector** of x with respect to basis B is defined be a vector in $\mathbb{R}^n$, denoted by $[x]_B$, such that

$$[x]_B = \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix}.$$

Note that we have

$$x = B[x]_B,$$

where here we use B to denote the matrix with columns of $b_1, b_2, ..., b_n$.

The **coordinate map** associated with basis B of vector space V is a linear map $\phi_B : V \to \mathbb{R}^n$.The **inverse coordinate map** is such that

$$\phi^{-1}([x]_B) = x_1 b_1 + x_2 b_2 + ... + x_n b_n : \mathbb{R}^n \to V.$$

Clearly, coordinate map has following properties:

Lemma 5.3.7 (coordinate map properties).

- *Let V be a finite dimensional vector space with basis B. Then the coordinate map ϕ_B is an isomorphism.*
- *Any finite dimensional vector space V is isomorphic to the Euclidean space $\mathbb{R}^{dim(V)}$.*

Example 5.3.3.

- The coordinate map associated with the basis $\{1, t, t^2, ..., t^n\}$ of the polynomial vector space is the isomorphism

$$a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n \to (a_0, a_1, ..., a_n)^T.$$

- The coordinate map associated with the basis

$$\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\},$$

  is the isomorphism

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \to (a, b, c, d)^T$$

5.3.5 Change of basis and similarity

5.3.5.1 *Change of basis for coordinate vector*

Lemma 5.3.8 (change of basis for vector representation). *Given a basis B_1 and a basis B_2, a coefficient vector v_1 with respect to B_1 has the coefficient vector v_2 given as*

$$B_1 v_1 = B_2 v_2 \Leftrightarrow v_2 = B_2^{-1} B_1 v_1$$

specifically, if $B_1 = E$,i.e., B_1 is the standard basis I, then

$$v_2 = B_2^{-1} v_1$$

Note:
When we write $v \in V$ as a tuple $(v_1, v_2, ..., v_n)$, we are **implicitly use the standard basis as the basis**.

5.3.5.2 *Change of basis for linear maps*

5.3.6 Linear maps and matrices

Suppose $T \in \mathcal{L}(V)$ and $v_1, v_2, ..., v_n$ is a basis of V. The matrix $M(T)$ of T with respect to this basis is required to satisfied

$$T(v_j) = \sum_{i=1}^{n} m_{ij} v_i.$$

Collecting terms of m_{ij} into a matrix M, we call M as the the matrix representation T.

Theorem 5.3.3 (change of basis for linear operator, similarity transform). *Suppose $T \in \mathcal{L}(V)$ has a matrix representation M_1 with respect to B_1, then the matrix representation M_2 with respect to B_2 is given as*

$$M_2 = (B_2^{-1}B_1)M_1(B_2^{-1}B_1)^{-1} = B_2^{-1}B_1M_1(B_1^{-1}B_2)$$

specifically, if $B_1 = E$,i.e., B_1 is the standard basis, then

$$M_2 = B_2^{-1}M_1B_2$$

and

$$M_1 = B_2M_2B_2^{-1}$$

where M_1 is the matrix representation in standard basis.

Proof. Because M_1 will map a vector with respect to B_1 to a vector with respect to B_1, we need to first transform the input vector with respect to B_2 to be with respect to B_1 and transform the input vector with respect to B_1 to be with respect to B_2. □

Remark 5.3.4 (interpret matrix diagonalizing). Consider a square matrix A can be written as (via eigendecomposition)

$$A = P\Lambda P^{-1} \Leftrightarrow P^{-1}\Lambda P$$

then we interpret P as the new basis, and in this new basis representation, the linear operator has diagonal representation.

Remark 5.3.5. Change of basis will not affect linear mapping, which is in nature an associative relationship between input space and output space.

Lemma 5.3.9 (change of basis for subspace representation). *Let A and B be two $n \times p$ matrices, both with full rank and $\mathcal{R}(A) = \mathcal{R}(B)$. Then there exists $A = BC$, with C being the $p \times p$ nonsingular matrix.*

Proof. Because $\mathcal{R}(A) = \mathcal{R}(B)$, then for each column b_i of B, it should be able to write as

$$b_i = Ac_i, c_i \in \mathbb{R}^p, i = 1, 2, ..., p.$$

Therefore, $B = AC$. To show $C \in \mathbb{R}^{p\times p}$ is nonsingular, we use the matrix product inequality

$$p = rank(A) = rank(AC) \leq \min(rank(A), rank(C)) \implies rank(C) = p.$$

□

5.3.6.1 *Similarity*

Definition 5.3.6 (similarity of matrices). *Two square matrices $A, B \in \mathbb{R}^{n\times n}$ are said to be **similar**, denote by $A \sim B$, if there exists an invertible matrix $P \in \mathbb{R}^{n\times n}$ such that*

$$A = PBP^{-1}.$$

Lemma 5.3.10 (similarity is an equivalence relation). *Matrices similarity is an equivalence relation; that is,*

- *(reflexivity) $A \sim A$.*
- *(symmetric) If $A \sim B$, then $B \sim A$.*
- *(transitivity) If $A \sim B, B \sim C$, then $A \sim C$.*

Proof. (1) $A = I^{-1}AI$. (2) $A \sim B$ implies $A = PBP^{-1}$, which further implies $B = P^{-1}AP$. Since P^{-1} is invertible, we have $B \sim A$. (3) Suppose $A = PBP^{-1}, B = QCQ^{-1}$, then $A = PQCQ^{-1}P^{-1} = GQG^{-1}$ where $G = PQ$. Therefore, $A \sim C$. □

5.4 Fundamental theorems of ranks and linear algebra

5.4.1 Basics of ranks

Definition 5.4.1 (rank and nullity).

- *The **rank of a matrix** is the dimensionality of its column space/range, i.e., the number of linearly independent columns or the number of linearly independent rows(as we will see in Theorem 5.4.3).*
- *The **nullity of a matrix** is the dimensionality of its null space.*

Lemma 5.4.1 (rank of matrix products).

- *Let A and B be square matrices of the same size, then*

$$rank(AB) = \dim(\mathcal{R}(AB)) \leq \dim(\mathcal{R}(A)) = rank(A)$$

$$dim(\mathcal{N}(AB)) \geq \dim(\mathcal{N}(A))$$

- *For any compatible matrix A, B,*

$$rank(AB) \leq \min(rank(A), rank(B)).$$

- *Let $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{n\times k}$. If $rank(B) = n$, then*

$$rank(AB) = rank(A).$$

- *Let $A \in \mathbb{R}^{m\times n}$ and $C \in \mathbb{R}^{l\times m}$. If $rank(C) = m$, then*

$$rank(CA) = rank(A).$$

Proof. (1)Let $y \in \mathcal{R}(AB)$, then there exists a x, such that $y = ABx = Az, z = Bx$. Then y is also in $\mathcal{R}(A)$, and therefore $\mathcal{R}(AB) \subseteq \mathcal{R}(A)$ and thus $dim(\mathcal{R}(AB)) \leq dim(\mathcal{R}(A))$. The inequality for null space dimensionality can be proved via Theorem 5.4.2. (2) Note that similar to (1), we have $rank(AB) \leq rank(A)$. Take the transpose, we have $rank(AB) = rank(B^T A^T) \leq rank(B^T) = rank(A)$, where we use the fundamental theorem of ranks [Theorem 5.4.3] such that $rank(A) = rank(A^T)$. (3) If $y \in \mathcal{R}(A)$, then there exists a $b \in \mathbb{R}$ such that $y = Ab$. Because B is of full row rank, then there exists a $z \in \mathbb{R}$ such that $b = Bz$. Therefore, $y = ABz$. As a result, we have proved $\mathcal{R}(A) \subseteq \mathcal{R}(AB)$. In (1), we prove $\mathcal{R}(AB) \subseteq \mathcal{R}(A)$. Eventually, we have $rank(AB) = A$. (4)

$$rank(CA) = rank(A^T C^T) = rank(A^T) = rank(A).$$

□

Theorem 5.4.1 (rank sum inequality). *Let* $A, B \in \mathbb{F}^{n\times n}$, *then*

- $\mathcal{R}(A+B) \subset \mathcal{R}(A) + \mathcal{R}(B)$
- $rank(A+B) \leq rank(A) + rank(B)$
- $rank(A+B) \geq \left| rank(A) - rank(B) \right|$

Proof. (1) Let $x = (A+B)y$, for some $y \in \mathbb{R}^n$, then $x = Ay + By$, indicating that $x \in \mathcal{R}(A) + \mathcal{R}(B)$. (2) $dim(\mathcal{R}(A+B)) = rank(A+B) = dim(\mathcal{R}(A)) + dim(\mathcal{R}(B)) - dim(\mathcal{R}(A) \cap \mathcal{R}(B)) \leq rank(A) + rank(B)$ from Theorem 5.2.3. (3) use the fact the $rank(B) = rank(-B)$ and (2). □

5.4.2 Fundamental theorem of ranks

Theorem 5.4.2 (The rank-nullity theorem). *Let* $A \in \mathbb{F}^{m\times n}$, *then*

$$dim(\mathcal{N}(A)) + dim(\mathcal{R}(A)) = n.$$

Proof. see linear map theory part Theorem 5.3.1. □

Lemma 5.4.2 (orthogonality between row space and null space). *[4, p. 102] For any matrix* A, *the subspace* $\mathcal{R}(A^T)$ *is orthogonal to* $\mathcal{N}(A)$ *and* $\mathcal{R}(A^T) \cap \mathcal{N}(A) = \{0\}$..

Proof. Let $x \in \mathcal{N}(A)$ and $y \in \mathcal{R}(A^T)$, then

$$x^T y = x^T A^T z = (Ax)^T z = 0;$$

therefore the subspace $\mathcal{R}(A^T)$ is orthogonal to $\mathcal{N}(A)$. Let $x \in \mathcal{N}(A)$, let $x \in \mathcal{R}(A^T)$, then $x^T x = 0 \Rightarrow x = 0$. □

Lemma 5.4.3 (rank of matrix $A^T A$).

- *For any matrix* A, $\mathcal{N}(A) = \mathcal{N}(A^T A)$.
- $A^T A$ *is of full rank if and only if* A *of full column rank.*
- $\dim(\mathcal{R}(A)) = dim(\mathcal{R}(A^T A))$; *or equivalently,* $rank(A) = rank(A^T A)$.

Proof. (1)(a) Consider any $x \in \mathcal{N}(A)$, we have $Ax = 0$ thus $A^T Ax = 0$. Therefore, $\mathcal{N}(A) \subseteq \mathcal{N}(A^T A)$; (b) Let $x \in \mathcal{N}(A^T A)$, that is

$$A^T Ax = 0 \implies x^T A^T Ax = 0 \implies (Ax)^T (Ax) = 0 \implies Ax = 0.$$

Therefore $\mathcal{N}(A^T A) \subseteq \mathcal{N}(A)$. Combine (a) and (b), we have $\mathcal{N}(A) = \mathcal{N}(A^T A)$. (2) Note that $A^T A$ is a square matrix. If A is full column rank, we have $\mathcal{N}(A) = 0$ then $\mathcal{N}(A^T A) = \mathcal{N}(A) = 0$. Therefore $A^T A$ is full column rank. (3) From rank-nullity theorem [Theorem 5.4.2], we have $\dim(\mathcal{R}(A)) = n - \dim(\mathcal{N}(A)) = n - dim(\mathcal{N}(A^T A)) = \dim(\mathcal{R}(A^T A))$. □

Theorem 5.4.3 (fundamental theorem of ranks). *[4, p. 132] For any matrix A*

$$dim(\mathcal{R}(A)) = dim(\mathcal{R}(A^T));$$

or equivalently, the column rank equals the row rank,

$$rank(A) = rank(A^T).$$

Proof. Note that for any matrix A, we have $\mathcal{R}(AA^T) \subseteq \mathcal{R}(A)$, which implies $rank(A^T A) \leq rank(A^T)$. From above theorem, we know that $rank(A) = rank(A^T A) \leq rank(A^T)$, which says the rank of any matrix is less or equal than its transpose. Then $dim(\mathcal{R}(A^T)) \leq dim(\mathcal{R}((A^T)^T)) = dim(\mathcal{R}((A))$, contradiction. Therefore, we have to have $dim(\mathcal{R}(A)) = dim(\mathcal{R}(A^T))$. □

Remark 5.4.1. Note that when we decompose a matrix, its sum of rank of the decomposed matrix will increase, i.e.,

$$rank(A + B) \leq rank(A) + rank(B)$$

(for example $rank(A + A) < rank(A) + rank(A) = 2rank(A)$) and the equality only holds when $\mathcal{R}(A) \cap \mathcal{R}(B) = \emptyset$.

5.4.3 Fundamental theorem of linear algebra

Theorem 5.4.4 (fundamental theorem of linear algebra). *[5][1, p. 178]For any given matrix $A \in \mathbb{R}^{m \times n}$, it holds that $\mathcal{N}(A) \perp \mathcal{R}(A^T)$ and $\mathcal{N}(A^T) \perp \mathcal{R}(A)$, hence*

$$\mathcal{N}(A) \oplus \mathcal{R}(A^T) = \mathbb{R}^n, \mathcal{N}(A^T) \oplus \mathcal{R}(A) = \mathbb{R}^m$$

and

$$rank(A) = rank(A^T),$$
$$dim(\mathcal{N}(A)) + rank(A) = n,$$
$$dim(\mathcal{N}(A^T)) + rank(A^T) = m.$$

Proof. (1) we can always decompose $R^n = \mathcal{N}(A) \oplus \mathcal{N}(A)^{\perp}$ due to orthogonal complement theorem [Theorem 5.5.4]. Therefore, we want to show $\mathcal{N}(A)^{\perp} = \mathcal{R}(A^T)$. (a) First $\mathcal{R}(A^{\perp}) \perp \mathcal{N}(A)$, therefore, $\mathcal{R}(A^{\perp}) \subseteq \mathcal{N}(A)^{\perp}$. Let $z \in \mathcal{R}(A^{\perp})$. Then there exists a $y \in \mathbb{R}^m$ such that $z = A^T y$. Let $m \in \mathcal{N}(A)$ such that $Am = 0$. We have $m^T z = m^T A^T y = 0$. (b) Because of $dim(\mathcal{R}(A^T)) = r = n - dim(\mathcal{N}(A))$ due to rank-nullity theorem [Theorem 5.4.2], then $\mathcal{R}(A^T) = \mathcal{N}(A)^{\perp}$(see theorem for subspaces that equal dimensionality implies equality Theorem 5.2.7).

(2) Others can be proved similarly. □

Remark 5.4.2 (interpretation).

- we can always decompose $R^n = \mathcal{N}(A) \oplus \mathcal{N}(A)^{\perp}$ and $R^m = \mathcal{R}(A) \oplus \mathcal{R}(A)^{\perp}$ due to orthogonal complement theorem in Hilbert space [Theorem 5.5.4].
- For $x = Ay$, if we transpose the equation, we have $x^T = y^T A^T$, where x^T, y^T are still the same vector in the output/input space. By informal symmetric argument, we need to have $\mathcal{N}(A)^{\perp} = \mathcal{R}(A^T), \mathcal{R}(A)^{\perp} = \mathcal{N}(A^T)$.

Remark 5.4.3 (How to calculate different subspace).

- $\mathcal{R}(A)$ can be directly obtained from linear independent columns in A.
- $\mathcal{R}(A^T) = \mathcal{N}(A)^{\perp}$ can be directly obtained from linearly independent rows in A.
- $\mathcal{N}(A)$ can be calculated from solution space $Ax = 0$.
- $\mathcal{N}(A^T)$ can be calculated from solution space $A^T x = 0$.

Corollary 5.4.4.1 (range-null decomposition). *Given matrix $A \in \mathbb{R}^{m\times n}$, we decompose* ***uniquely*** *any $p \in \mathbb{R}^m$ as*

$$p = p_N + p_{N^{\perp}}$$

where $p_N \in \mathcal{N}(A), p_{N^{\perp}} \in \mathcal{R}(A^T)$ and $p_{N^{\perp}} = A^T y$ for some $y \in \mathbb{R}^m$

Proof. Note that $\mathcal{N}(A)$ and $\mathcal{R}(A^T)$ are orthogonal complementary. □

5.5 Complementary subspaces and projections

5.5.1 General complementary subspaces

Definition 5.5.1 (complementary subspaces). *[1, p. 392]Subspaces X, Y of a vector space V are said to be complementary if*

$$V = X + Y, X \cap Y = 0$$

we can also denoted as

$$V = X \oplus Y$$

Definition 5.5.2 (angle between complementary subspaces). *[1, p. 389] The angle between two complementary subspaces X and Y such that*

$$\mathbb{R}^n = X \oplus Y$$

can be defined as

$$\cos(\theta) = \max_{u \in X, v \in Y} \frac{v^T u}{\|v\| \|u\|} = \max_{u \in X, v \in Y, \|v\|=1, \|u\|=1} v^T u$$

Remark 5.5.1. It is easy to see

1. as two complementary subspace between orthogonal complementary, the angle is 90, since $v^T u = 0$.
2. In 3D, for an origin-passing line and a hyperplane, the angle is given by the angle between the line director and a vector in the plane that maximize the dot product.

Theorem 5.5.1. *[1, p. 383] For a vector space V with subspaces X and Y having respective basis B_X and B_Y, the following statement are equivalent:*

- $V = X \oplus Y$
- $B_X \cap B_Y = \emptyset$ *and* $B_X \cup B_Y$ *is a basis for* V.

Proof. Straight forward from the property of direct sum, see subsection 5.2.3. □

Definition 5.5.3 (projection along subspace). *Suppose $V = X \oplus Y$ such that for every $v \in V$, v can be decomposed uniquely as $v = x + y, x \in X, y \in Y$.*

- *The vector x is called the projection of v onto X along Y.*
- *The vector y is called the projection of v onto Y along X.*

Remark 5.5.2. Only for complementary subspaces we can define projection.

Definition 5.5.4 (projector). *[1, p. 386] Given two complementary subspaces X, Y of vector space V such that for every $v \in V$, we have unique decomposition of $v = x + y$. Then a linear operator $P(v) = x$ is called the projector onto X along Y.*

Theorem 5.5.2 (basic properties of projector). *[1, p. 386] Given a projector P onto subspace $\mathcal{X}$ along subspace $\mathcal{Y}$, then*

- $P^2 = P$
- *The range of P si the fixed point set of P, that is $P(x) = x, \forall x \in \{x = P(v), v \in V\}$*
- *$I - P$ is the complementary projector onto Y along X*
- *The matrix representation for projectors in $V = \mathbb{F}^n$ is given as*

$$P = [X|0][X|Y]^{-1} = [X|Y] \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} [X|Y]^{-1}$$

where X, Y are basis of subspace $\mathcal{X}$ and $\mathcal{Y}$.

Proof. (1) $P(P(v)) = P(x) = x = P(v)$; (2) $P(P(v)) = P(x) = x = P(v), \forall v \in V$, that is, the range $P(v)$ is the fixed point set.(3) $(I - P)(v) = v - x = y$;
(4) In a vector space spanned by basis $X \cup Y$, the matrix representation of P is

$$\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$$

then we can use change of basis theorem to prove it(see Theorem 5.12.2). □

Theorem 5.5.3 (projector and idempotent). *[1, p. 387] Let P be a linear operator on V such that $P^2 = P$, then we have*

- *$\mathcal{R}(P)$ and $\mathcal{N}(P)$ are complementary subspaces, that is*

$$V = \mathcal{R}(P) \oplus \mathcal{N}(P)$$

- *P is a projector onto $\mathcal{R}(P)$ along $\mathcal{N}(P)$.*

Proof. (1) At first $V = \mathcal{R}(P) + \mathcal{N}(P)$ Let any $x \in V$, we have

$$x = Px + (I - P)x$$

where $Px \in \mathcal{R}(P), (I - P) \in \mathcal{N}(P)$ Then we can show that $\mathcal{R}(P) \cap \mathcal{N}(P) = 0$: let $x \in \mathcal{R}(P) \cap \mathcal{N}(P) = 0$, then $x = Pv, Px = 0$, which implies

$$x = Pv = P^2 v = Px = 0$$

Therefore

$$V = \mathcal{R}(P) \oplus \mathcal{N}(P)$$

(2) $Pv = x$ where $x \in \mathcal{R}(P)$ and $v = x + y, x \in \mathcal{R}(P), y \in \mathcal{N}(P)$. Therefore by definition P is a projector onto $\mathcal{R}(P)$ along $\mathcal{N}(P)$. □

Corollary 5.5.3.1 (The range and null space of a projection matrix). *For a projection matrix P, the range space is the column space of P, and the null space is the column space of $I - P$.*

Proof. $P(I - P)x = (P - P^2)x = 0, \forall x$. □

5.5.2 Orthogonal complementary spaces and projections

Definition 5.5.5 (orthogonal complement). *[1, p. 404] For a subset M in an inner product space V, we define $M^{\perp}$ to be the set*

$$M^{\perp} = \{x \in V : \langle x, m \rangle = 0, \forall m \in M\}$$

$M^{\perp}$ is known as the orthogonal complement.

Lemma 5.5.1 (orthogonal complement forms a subspace). *[1, p. 404] For a* ***subset*** *M in an inner product space V, $M^{\perp}$ is a subspace of V,* ***no matter*** *M* ***is a subspace or not.***

Proof. By the definition of $M^{\perp}$, it is easy to show that $M^{\perp}$ contains 0, and it is closed under addition and multiplication. □

Theorem 5.5.4 (orthogonal decomposition in finite linear space). *[1, p. 404] For a **subspace** M in an inner product space V, then we have*

- $V = M \oplus M^{\perp}$, *that is, for any vector* $v \in V$, *we have a unique decomposition* $v = m + n, m \in M, n \in M^{\perp}$
- $dim(M^{\perp}) = dim(V) - dim(M)$
- $M^{\perp^{\perp}} = M$

Proof. (1) First $M \cap M^{\perp} = 0$, let $x \in M \cap M^{\perp}$, then $\langle x, x \rangle = 0 \Rightarrow x = 0$. Second, $S = M \oplus M^{\perp} \subseteq V$. Suppose S is a proper subset, and $V = span(B_M, B_{M^{\perp}}, q)$. When we use Gram-Smith procedure to $B_M \cup B_{M^{\perp}} \cup q$, we will yield $q' = 0$ because q' has to be orthogonal to M, then $q \subset M^{\perp}$. Therefore

$$V = M \oplus M^{\perp}.$$

(2) Note that

$$dim(V) = dim(M) + dim(M^{\perp}) - dim(M \cap M^{\perp}),$$

then use $M \cap M^{\perp} = 0$ in (1). (3) Let $m \in M$, then $m \perp M^{\perp}$, that is $M \subset M^{\perp^{\perp}}$. Because $dim(M^{\perp^{\perp}}) = dim(M)$, we have $M^{\perp^{\perp}} = M$ via Theorem 5.2.7. □

Definition 5.5.6 (orthogonal projection). *[1, p. 429] For any inner product space V and a subspace M, we have $V = M \oplus M^{\perp}$, therefore $v = m + n, m \in M, n \in M^{\perp}, \forall v \in V$.Therefore, we can define an linear operator P such that $P(v) = m$, then P is called the orthogonal projector onto M (along $M^{\perp}$.)*

Theorem 5.5.5 (orthogonal projector representation). *[1, p. 430] Let $\mathcal{M}$ be an r dimensional subspace of $\mathbb{R}^n$, let M and N be the basis of $\mathcal{M}$ and $\mathcal{N} = M^{\perp}$. Then we have*

- $P_M = M(M^T M)^{-1} M^T$
- *If M is **orthonormal basis** such that $M^T M = I$, then $P_M = MM^T$*

Proof. (1) From Theorem 5.5.2, we know that

$$P_M = [M|0][M|N]^{-1}$$

we can verify that

$$\begin{bmatrix} (M^T M)^{-1} M^T \\ N^T \end{bmatrix} [M|N] = I$$

because $M^T N = 0, N^T M = 0$. Then we have

$$[M|N]^{-1} = \begin{bmatrix} (M^T M)^{-1} M^T \\ N^T \end{bmatrix}$$

then we have

$$P_M = [M|0][M|N]^{-1} = M(M^T M)^{-1} M^T.$$

(2) use $M^T M = I$. □

Theorem 5.5.6 (characterization of orthogonal projector). *[1, p. 433] Suppose $P \in \mathbb{R}^{n\times n}$ satisfying $P^2 = P$, that is, P is a projector; then P is a orthogonal projector if*

- *P is symmetric matrix; moreover, if P is an orthogonal projector, then P is symmetric.*
- *$\mathcal{R}(P) \perp \mathcal{N}(P)$*
- *$\|P\|_2 = 1$*

Proof. (1) If P is orthogonal projector, then P has matrix representation $P_M = M(M^T M)^{-1} M^T$ with respect to some basis M, it is easy to show that P_M is symmetric. (2) First, let $x \in \mathcal{R}(P)$, then $x = Py$. Let $z \in \mathcal{N}(P)$, then $x^T z = y^T P z = 0$. Therefore, $\mathcal{R}(P) \perp \mathcal{N}(P)$. (3) If P satisfies $P^2 = P, P^T = P$, we can use spectral decomposition of P Theorem 5.5.7 to prove. For example, $\|P\|_2 = \lambda_{max} = 1$ [Corollary 5.8.4.4]. □

Theorem 5.5.7 (spectral properties of orthogonal projector). *Let real matrix P be an orthogonal projector (that is, $P^2 = P, P^T = P$), then we have*

- *The only possible eigenvalues are 1 and 0.*
- *$\mathcal{R}(P)$ are the eigenspace associated with eigenvalue 1; that is, Columns of P are and only are the eigenvectors associated with eigenvalue 1.(Note that P is not necessarily full rank, and therefore some columns are the linear combination of the other columns.)*
- *$\mathcal{R}(I-P)$ are the eigenspace assciated with eigenvalue 0*
- *The algebraic multiplicity of 1 equals $rank(P)$, the algebraic multiplicity of 0 is $rank(I-P) = dim(\mathcal{N}(P))$.*
- *$Tr(P) = rank(P)$.*
- *The diagonal entries of P are all between 0 and 1 inclusively.*

Proof. (1)Let λ be a eigenvalue of P for the eigenvector v, then $Pv = \lambda v \Rightarrow P^2 v = \lambda^2 v = Pv = \lambda v$. Therefore, λ satisfy $\lambda^2 = \lambda$, which yields $\lambda = 1$ or $\lambda = 0$. (2)$P^2 = P$ suggests columns of P are the eigenvectors of eigenvalue 1. Let v be the eigenvector associated with eigenvalue 1, then $Pv = v$, suggesting $v \in \mathcal{R}(P)$. Therefore $\mathcal{R}(P) = \mathcal{N}(I-P)$(the

eigenspace associated with eigenvalue 1). Similarly we can prove (3). (4) Since P is real matrix, from Theorem 5.8.3, we know that the algebraic multiplicity equals the geometric multiplicity. Therefore, $\mathcal{N}(P)$(the null space corresponds to eigenvalue 0) has dimensionality equaling the number of independent columns of $I-P$, i.e., $rank(P)$. Similarly, we can show algebraic multiplicity of 1 equals $rank(P)$.
(5) Directly from (2) since $Tr(P) = \sum_i \lambda_i$. (6) First, all diagonal entries will be non-negative [Lemma 5.12.3]. Second,

$$\begin{aligned} & P = P^2 \\ \implies & P_{ii} = \sum_{j=1}^{n} P_{ij}^2 \\ & = P_{ii}^2 + \sum_{j\neq i} P_{ij}^2 \\ & \geq P_{ii}^2 \end{aligned}$$

which implies that $0 \leq P_{ii} \leq 1, i = 1, 2, ..., n$. □

Remark 5.5.3 (orthogonal projector and positive semidefinite matrix). Orthogonal projectors is a subset of positive semidefinite matrix; particularly, a positive semidefinite matrix with only 0 and 1 eigenvalue is orthogonal projector.

Example 5.5.1 (elementary orthogonal projector). Let $u \in \mathbb{R}^n$. Then the matrix $P_u = \frac{uu^T}{u^T u}$ and $I - P_u$ are called **elementary orthogonal projectors** associated with u.

P_u is an $n \times n$ matrix of rank one that is symmetric and idempotent, i.e., $P_u^T = P_u, P_u^2 = P_u$.

Lemma 5.5.2 (uniqueness of orthogonal projectors with the same column basis). *Let A and B be two $n \times p$ matrices, both with the full column rank and such that $\mathcal{R}(A)$ and $\mathcal{R}(B)$. Then*

$$P_A = A(A^T A)^{-1} A^T = B(B^T B)^{-1} B^T = P_B.$$

Proof. Since $\mathcal{R}(A) = \mathcal{R}(B)$, there exists a $p \times p$ nonsingular matrix C such that $A = BC$ [Lemma 5.3.9]. We have

$$\begin{aligned} P_A &= A(A^TA)^{-1}A^T \\ &= BC(C^TB^TBC)^{-1}C^TB^T \\ &= BCC^{-1}(B^TB)^{-1}C^{-T}C^TB^T \\ &= BCC^{-1}(B^TB)^{-1}C^{-T}C^TB^T \\ &= B(B^TB)^{-1}B^T \\ &= P_B \end{aligned}$$

□

5.5.3 Decomposition of orthogonal projectors

Lemma 5.5.3 (decomposition of orthogonal projector). *[4, p. 222] Let $A \in \mathbb{R}^{m\times n}$ with full column rank. Let X be partitioned as $A = [A_1 \quad A_2]$. Let P_A, P_{A_1}, P_{A_2} be the orthogonal projectors associated with A, A_1, A_2. It follows that the following statements are equivalent*

- $A_1A_2^T = 0,$
- $P_A = P_{A_1} + P_{A_2}$, *that is*
$$A(A^TA)^{-1}A^T = A_1(A_1^TA_1)^{-1}A_1^T + A_2(A_2^TA_2)^{-1}A_2^T.$$
- $$P_{A_1}P_{A_2} = P_{A_2}P_{A_1} = 0.$$

Proof. (1) to (2)

$$\begin{aligned} P_A &= A(A^TA)^{-1}A^T \\ &= [A_1 \ A_2]\begin{bmatrix} A_1^TA_1 & A_1^TA_2 \\ A_2^TA_1 & A_2^TA_2 \end{bmatrix}[A_1 \ A_2]^T \\ &= [A_1 \ A_2]\begin{bmatrix} A_1^TA_1 & 0 \\ 0 & A_2^TA_2 \end{bmatrix}[A_1 \ A_2]^T \\ &= A_1(A_1^TA_1)^{-1}A_1^T + A_2(A_2^TA_2)^{-1}A_2^T \end{aligned}$$

(2) to (3) Note that

$$\begin{aligned}(P_{A_1} + P_{A_2})^2 &= P_{A_1}^2 + P_{A_2}^2 + P_{A_2}P_{A_1} + P_{A_1}P_{A_2} \\ &= P_{A_1} + P_{A_2} + P_{A_2}P_{A_1} + P_{A_1}P_{A_2} \\ &= P_{A_1} + P_{A_2} \\ \implies & P_{A_2}P_{A_1} + P_{A_1}P_{A_2} = 0.\end{aligned}$$

Left multiply $P_{A_2}P_{A_1} + P_{A_1}P_{A_2} = 0$ we get $P_{A_1}P_{A_2}P_{A_1} + P_{A_1}P_{A_2} = 0$; right multiply $P_{A_2}P_{A_1} + P_{A_1}P_{A_2} = 0$ we get $P_{A_2}P_{A_1} + P_{A_1}P_{A_2}P_{A_1} = 0$; from the two equations, we get

$$P_{A_1}P_{A_2} = P_{A_2}P_{A_1}.$$

Plus $P_{A_2}P_{A_1} + P_{A_1}P_{A_2} = 0$, we get

$$P_{A_1}P_{A_2} = P_{A_2}P_{A_1} = 0.$$

(3) to (1):

$$A_1P_{A_1}P_{A_2}A_2 = A_1A_2 = 0.$$

□

Theorem 5.5.8 (decomposition of orthogonal projector). *[4, p. 224] Let $X \in \mathbb{R}^{m\times n}$ with full column rank. Let X be partitioned as $X = [X_1\ X_2]$. Let $Z = (I - P_{X_1})X_2$. It follows that the following statements are equivalent*

- $X_1^TZ = 0$,
- $\mathcal{R}(X) = \mathcal{R}([X_1\ Z])$
- $P_X = P_{X_1} + P_Z$; *that is,*

$$X(X^TX)^{-1}X^T = X_1(X_1^TX_1)^{-1}X_1 + (Z)(ZZ^T)^{-1}(Z)^T.$$

Proof. (1)

$$X_1^T(I - P_{X_1})X_2 = (X_1^T - X_1^T)X_2 = 0.$$

(2) (a) Let $u \in \mathcal{R}(X)$, then exists vectors α, β such that

$$\begin{aligned}u &= X_1\alpha + X_2\beta \\ &= X_1\alpha + (I - P_{X_1} + P_{X_1})X_2\beta \\ &= X_1\alpha + P_{X_1}X_2\beta + Z\beta \\ &= X_1\alpha + X_1(X_1^TX_1)^{-1}X_1^TX_2\beta + Z\beta \\ &= X_1(\alpha + (X_1^TX_1)^{-1}X_1^TX_2\beta) + Z\beta\end{aligned}$$

therefore, $u \in \mathcal{R}([X_1\ Z])$. (b) Let $u \in \mathcal{R}([X_1\ Z])$. then exists vectors α, β such that

$$\begin{aligned} u &= X_1\alpha + Z\beta \\ &= X_1\alpha + (I - P_{X_1})X_2\beta \\ &= X_1\alpha - P_{X_1}X_2\beta + X_2\beta \\ &= X_1\alpha - X_1(X_1^TX_1)^{-1}X_1^TX_2\beta + X_2\beta \\ &= X_1(\alpha - (X_1^TX_1)^{-1}X_1^TX_2\beta) + X_2\beta \end{aligned}$$

(3) use Lemma 5.5.3. □

Corollary 5.5.8.1 (low rank update of orthogonal projector). *[6, p. 173] Let $X \in \mathbb{R}^{m\times n}$ with full column rank and $X = [X_1, X_2, ..., X_n]$. Let $W = [X_2, X_3, ..., X_n]$.*

Define

$$H = X(X^TX)^{-1}X^T, M = I - H, G = W(W^TW)^{-1}W^T, N = I - G.$$

It follows that

- $$H = G + \frac{(NX_1)(NX_1)^T}{X_1NX_1}$$
- $$M = N - \frac{(NX_1)(NX_1)^T}{X_1NX_1}$$

Proof. (1)(informal) Use the property $G^2 = G, G^T = G, N^2 = N, N^T = N, GN = 0$, we can show that

$$(G + \frac{(NX_1)(NX_1)^T}{X_1NX_1})^2 = G + \frac{(NX_1)(NX_1)^T}{X_1NX_1},$$

and it is also symmetric. (formal proof using Theorem 5.5.8) (2)

$$I - H = I - G - \frac{(NX_1)(NX_1)^T}{X_1NX_1}$$

□

Remark 5.5.4 (interpretation). We are augmenting G with a basis X_1 projected in the space of complementing W via NX_1. The additional projector associated with NX_1 is given by

$$\frac{(NX_1)(NX_1)^T}{(NX_1)^T(NX_1)} = \frac{(NX_1)(NX_1)^T}{X_1^TN^TNX_1} = \frac{(NX_1)(NX_1)^T}{X_1^TNX_1}$$

5.6 Orthonormal basis and projections

Definition 5.6.1 (orthonormal basis). *For inner product space, a basis is orthonormal if each vector has unit length, and orthogonal to other vectors.*

Remark 5.6.1. Only in inner product space, we define orthogonality by inner product; in ordinary vector space, we do not have the concept of orthogonality.

Lemma 5.6.1 (representing vectors using orthonormal basis). *Let $\{e_i\}$ be a orthonormal basis for V, then for all $v \in V$, it can be represented as*

$$v = \langle v, e_1 \rangle e_1 + ... + \langle v, e_n \rangle e_n$$

Proof. Let $v = \sum_i a_i e_i$ and use inner product to determine a_i. $\square$

5.6.1 Gram-Schmidt Procedure

5.6.2 Orthogonal-triangular decomposition

Theorem 5.6.1 (QR decomposition). *[5] Suppose $A \in \mathbb{R}^{m\times n}, m \geq n$. Then we have*

- *there exists a orthonormal matrix $Q \in \mathbb{R}^{m\times m}$ and upper triangular matrix $R \in \mathbb{R}^{m\times n}$ such that*
$$A = QR$$
- *If $\hat{Q} \in \mathbb{R}^{n\times n}$ and $\hat{R} \in \mathbb{R}^{n\times n}$, then*
$$A = QR = [\hat{Q}, N]\begin{bmatrix} \hat{R} \\ 0 \end{bmatrix} = \hat{Q}\hat{R}$$
where $\hat{Q} \in \mathbb{R}^{m\times n}$ consists of the basis of $\mathcal{R}(A)$, N consists of the basis of $\mathcal{N}(A^T)$ and $\hat{R} \in \mathbb{R}^{n\times n}$.
- *We can choose R to have nonnegative diagonal entries*
- *If A is of full rank, we can choose R with positive diagonal entries, in which case the economical form $\hat{Q}$ and $\hat{R}$ will be unique.*
- *If A is square nonsingular, then $A = QR$ is unique.*

Proof. (1)(2) Consider the Gram-Smith process for the columns of matrix A given as

$$q_1 = a_1, p_1 = q_1/\|1_1\|$$
$$q_i = a_i - \sum_{j=1}^{i-1} \left\langle a_i, p_j \right\rangle p_j, p_i = p_i/\|q_i\|, i = 2, ..., n$$

or

$$a_1 = r_{11}p_1$$
$$a_j = \sum_{i=1}^{j} r_{ij}p_i, j = 2, ..., n$$
$$r_{ii} = \|q_i\|, r_{ij} = \left\langle a_j, p_i \right\rangle$$

in which orthonormal basis $p_1, ..., p_n$ for the column space $span(a_1, ..., a_n)$ will be produced. We can see that $a_i \in span(p_1, ..., p_i)$, and therefore in matrix form we have

$$A = \hat{Q}\hat{R}$$

where $\hat{Q} \in \mathbb{R}^{m\times n}$ will consist of $p_1, ..., p_n$ as columns and $R \in \mathbb{R}^{n\times n}$ will an upper triangular matrix. The complete form Q can be augmented with basis of $\mathcal{N}(A^T) = \mathcal{R}(A)^\perp$, such that $Q \in \mathbb{R}^{m\times m}$ consist of the complete orthonormal basis of $\mathbb{R}^m$. (3)(4)(5) If $a_1, ..., a_n$ are linearly independent, from GS process, the matrix $\hat{Q}, \hat{R}$ are uniquely determined, and the diagonal entries of R is always positive. If $a_1, ..., a_n$ are linearly dependent, there will exist scenario that

$$a_k \in span(p_1, ..., p_{k-1})$$

and we can set $r_{kk} = 0$. □

5.6.3 Orthonormal basis for linear operators

Lemma 5.6.2 (existence of orthonormal basis). *[3, p. 185]Every finite-dimensional inner product space has an orthonormal basis.*

Proof. Because V has a basis, then we can use Gram-Schmidt procedure to make it orthonormal. □

Lemma 5.6.3 (existence of upper triangular matrix with respect to orthonormal basis). *[3, p. 186] Suppose $T \in \mathcal{L}(V)$. If T has an upper triangular matrix with respect to* *some basis, then T has an upper-triangular matrix with respect to some orthonormal basis of V.*

Proof. note that the Gram-Schmidt matrix is upper triangular. □

Theorem 5.6.2 (Schur's theorem). *Suppose V is a finite-dimensional complex vector space and $T \in \mathcal{L}(V)$. Then T has an upper-triangular matrix with respect to some orthonormal basis of V.*

Proof. directly from above theorem and the existence of upper triangular matrix theorem**??**. □

5.6.4 Riesz representation theorem

Theorem 5.6.3. *[3, p. 188][7, p. 345] Suppose V is **finite** dimensional and ϕ is a linear functional on V. Then there is a **unique** vector $u \in V$ such that*

$$\phi(v) = \langle u, v \rangle, \forall v \in V$$

Moreover,

$$\|\phi\| = \|u\|$$

Proof. (1)Let $e_1, e_2, ..., e_n$ be the orthonormal basis of V. Then

$$\begin{aligned}\phi(v) &= \phi(\langle v, e_1 \rangle e_1 + \langle v, e_2 \rangle e_2, ...) \\ &= \langle v, e_1 \rangle \phi(e_1) + \langle v, e_2 \rangle \phi(e_2) + ... \\ &= \left\langle v, \overline{\phi(e_1)} e_1 + \overline{\phi(e_2)} e_2 + ... \right\rangle\end{aligned}$$

Therefore, we can let $u = \overline{\phi(e_1)} e_1 + \overline{\phi(e_2)} e_2 + ...$ Uniqueness is easy. (2)

$$\|\phi\| = \sup_{\|v\|=1} \langle u, v \rangle = \|u\| \|v\|$$

where we use the fact that $\langle u, v \rangle^2 \leq \|u\|^2 \|v\|^2$, and the equality can be achieved. □

5.7 Eigenvectors and eigenvalues of Matrices: general theory

5.7.1 Existence and properties of eigenvalues

Definition 5.7.1 (characteristic equation). *For a square matrix A, the equation*

$$det[A - \lambda I] = 0$$

is called the characteristic equation of A. The resulting polynomial in λ is called characteristic polynomial.

Theorem 5.7.1 (existence of roots, Fundamental theorem of algebra, recap). *Every non-constant single-variable polynomials with complex coefficients has at least one complex root, or equivalently, every non-zero single variable, degree n polynomial with complex coefficients has, counted with multiplicity, exactly n roots.*

Proof. See Theorem 4.3.6. □

Theorem 5.7.2 (existence of solution to characteristic polynomial).

- *Any square matrix $A \in \mathbb{C}^{n\times n}$ has n eigenvalues in $\mathbb{C}$, counted with multiplicity.*
- *Every square matrix A has at least one eigenvalue and a corresponding (nonzero) eigenvector.*

Proof. From the fundamental theorem of algebra [Theorem 4.3.6], we know there exist a λ such that $det[A - \lambda I] = 0$. Then from linear equation solution theory, the linear equation $(A - \lambda I)x = 0$ must have $dim(\mathcal{N}(\mathcal{A} - \geq \mathcal{I})) \geq 1$. □

Remark 5.7.1.

- Each distinct eigenvalues $\lambda_i, i = 1, 2, ..., k \leq n$, has an associated *algebraic multiplicity* $\mu_i \geq 1$, and $\sum_{i=1}^{k} \mu_i = n$.
- To each distinct eigenvalues $\lambda_i, i = 1, 2, ..., k$, there corresponds a whole subspace $\phi_i = \mathcal{N}(\lambda_i I - A)$ of eigenvalues associated with this eigenvalue, called eigenspace.

Caution! Possible nonexistence of eigenvalues on $\mathbb{R}^2$
consider the linear map $T : \mathbb{R}^2 \to \mathbb{R}^2$,

$$T(x, y) = (-y, x)$$

with the matrix representation of

$$\begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix}$$

which simply rotates a vector. **We cannot find a scalar in $\mathbb{F}$ such that rotating a vector in $\mathbb{R}^2$ equals its scalar multiplication**. However, if the linear map is from $\mathbb{C}^2$ to $\mathbb{C}^2$, we can find eigenvalue and its corresponding eigenvector. [3, p. 135]

Theorem 5.7.3 (properties of eigenvalues).

- *(invariance under similar transformation) Let A be a squared matrix, let B be any invertible matrix. Then the eigenvalues of A and $T = BAB^{-1}$ are the same.*
- *(transformation under scalar multiplication) Let λ be the eigenvalue of A, then $\alpha\lambda$ will be the eigenvalues of αA, for $\alpha \in \mathbb{R}$.*
- *(transformation under matrix power) Let λ be the eigenvalue of A, then λ^k will be the eigenvalues of A^k, for $k \in \mathbb{Z}_+$.*
- *(transformation under matrix polynomial) Let λ be the eigenvalue of A, then $P(\lambda)$ will be the eigenvalues of $P(A)$, where $P(A) = a_0 A^0 + a_1 A + a_2 A^2 + ... + a_k A^k$.*
- *(eigenvalues of an inverse) If A is invertible, then for an eigenvector v associated with eigenvalue λ, A^{-1} has a corresponding eigenvalue $1/\lambda$, with the same eigenvector.*
- *The sum of the all eigenvalues of A is equal to trace of A (that is, the sum of the diagonal elements of A).*

Proof. (1) $0 = det(BAB^{-1} - I) = det(B)det(A - \lambda I)det(B)^{-1} = det(A - \lambda I)$ (2) Let v be the eigenvector, then

$$Av = \lambda v \implies \alpha A v = \alpha \lambda v$$

therefore $\alpha\lambda$ is the eigenvalue of αA. (3)Let v be the eigenvector, then

$$Av = \lambda v \implies A^2 v = \lambda^2 v \implies A^k v = \lambda^k v$$

therefore λ^k is the eigenvalue of A^k. (4) same as (3). (5) Let v be an eigenvector of A associated with eigenvalue λ, then

$$\begin{aligned} A^{-1}v &= A^{-1}(\lambda \frac{1}{\lambda} v) \\ &= \frac{1}{\lambda} A^{-1}(\lambda v) \\ &= \frac{1}{\lambda} A^{-1} A v \\ &= \frac{1}{\lambda} v \end{aligned}$$

(6) The characteristic polynomial of $A \in \mathbb{R}^{n\times n}$ can be expressed as

$$det(A-\lambda I) = \prod_{i=1}^{n}(\lambda-\lambda_i) = \lambda^n - \lambda^{n-1}\sum_{i=1}^{n}\lambda_i + \cdots + (-1)^n\prod_{i=1}^{n}\lambda_i;$$

on the other hand,

$$= det(A-tI) = (-1)^n\left(t^n - (tr(A))t^{n-1} + \cdots + (-1)^n det(A)\right),$$

we must have $\sum_{i=1}^{n}\lambda_i = Tr(A)$.

□

5.7.2 Properties of eigenvectors

Theorem 5.7.4 (existence of eigenvector in complex field). *In $\mathbb{C}^N$, any square matrix $A \in \mathbb{C}^{N\times N}$ must have **at least** one eigenvector $\mathbb{C}^N$ associated with each distinct eigenvalue in $\mathbb{C}$.*

Proof. Because for any matrix A, we can always have at least one eigenvalues in $\mathbb{C}$, therefore we can always have at least one eigenvector $\mathbb{C}^N$. For each distinct eigenvalue, $\mathcal{N}(A-\lambda I)$ has dimensionality equal or greater than 1(From linear equation solution theory,$A-\lambda I$ is singular, then the linear equation $(A-\lambda I)x = 0$ must have $dim(\mathcal{N}(\mathcal{A}-\geq\mathcal{I})) \geq 1.$). Therefore, $\mathcal{N}(A-\lambda I)$ must have one eigenvector as its basis. Also see Theorem 5.7.2. □

Lemma 5.7.1 (linear independence of eigenvectors). *Let $\lambda_1,\lambda_2,...,\lambda_k$ be distinct eigenvalues of $A \in \mathbb{R}^{n\times n}$ and $k \leq n$, then*

- *the corresponding eigenvectors $e_1,e_2,...,e_k$ are linearly independent.* [a].
- *let μ_i denote the corresponding algebraic multiplicities, and let $\phi_i = \mathcal{N}(\lambda_i I - A)$, and let u^i be any nonzero vectors such that $u^i \in \phi_i, i = 1,2,...k$. Then $u^1,u^2,...,u^k$s are linearly independent.*

[a] in Theorem 5.7.2, every distinct eigenvalue has at least one eigenvector associated with it

Proof. (1)Assume they are linear dependent, without loss of generality, we have

$$e_1 + \sum_{i=2}^{k} a_i e_i = 0$$

where for some $a_i \neq 0$. Multiply A, we have

$$\lambda_1 e_1 + \sum_{i=2}^{k} a_i \lambda_i e_i = 0$$

From the two equations, we have

$$\sum_{i=2}^{k} a_i(\lambda_i - \lambda_1) e_i = 0$$

indicating that $e_2, e_3, ..., e_k$ are linearly dependent. Continue the same argument will lead to the conclusion that e_{k-1} and e_k are linearly dependent(that is, there exists some $\alpha \in \mathbb{R}$ such that $e_{k-1} = \alpha e_k$), which is obvious not true. (2) Directly from (1) because $Au^i = \lambda_i u^i$. □

5.7.3 Right and left eigenvectors

Definition 5.7.2 (Right and left eigenvectors). *[8, p. 82]Given a square matrix A, an eigenvector e_i is right eigenvector if there exists $\lambda_i \in \mathbb{F}$ such that*

$$Ae_i = \lambda_i e_i$$

An eigenvector f_i is left eigenvector if there exists $\lambda_i \in \mathbb{F}$ such that

$$f_i^T A = \lambda_i f_i$$

Lemma 5.7.2 (left/right eigenvectors and symmetry).

- *The left eigenvector of A is the right eigenvector of A^T.*
- *For symmetric matrix A, the left eigenvector is the same as right eigenvector.*

Proof. (1)

$$(f_i^T A)^T = (\lambda_i f_i^T)^T \implies A^T f_i = \lambda_i f_i$$

(2) from (1). □

Lemma 5.7.3 (left and right eigenvalues). *The left and right eigenvalues are identical.*

Proof. $det(A - \lambda I) = det(A^T - \lambda I^T) = det(A^T - \lambda I)$. □

Theorem 5.7.5 (orthogonality of left and right eigenvectors). *[8, p. 83] For any two distinct eigenvalues of a matrix, the left eigenvector of one eigenvalue is orthogonal to the right eigenvalues of the other.*

Proof. Let e_i, f_i be the left and eight eigenvectors of eigenvalue λ_i. Let e_j, f_j be the left and eight eigenvectors of eigenvalue λ_j. Then

$$f_i^T(Ae_j) = \lambda_j f_i^T e_j$$

$$(f_i^T A)e_j = (\lambda_i f_i^T)e_j$$

then $(\lambda_i - \lambda_j)f_i^T e_j = 0 \Rightarrow f_i^T e_j = 0$ □

Theorem 5.7.6 (orthogonality of eigenvectors for symmetric matrix). *For a real-valued symmetric matrix A, eigenvectors of distinct eigenvalues are orthogonal.*

5.7.4 Diagonalizable matrices

Definition 5.7.3 (algebraic multiplicity,geometric multiplicity).

- *The algebraic multiplicity μ_i of eigenvalue λ_i is its multiplicity as a root of the characteristic polynomial.*
- *The geometric multiplicity γ_i of eigenvalue λ_i is the dimensionality of the null space $\mathcal{N}(A - \lambda_i I)$.*

Theorem 5.7.7 (boundedness of geometric multiplicity). *[4, p. 323] For a square matrix A with eigenvalues λ, we have*

$$1 \leq \mu \leq \gamma \leq n$$

that is, the geometric multiplicity μ is bounded by the algebraic multiplicity γ.

Proof. The algebraic part can be obtained from fundamental theorem of algebra. For geometric multiplicity, it will always be greater than 1 because the null space $A - \lambda I$ with

$det(A - \lambda I) = 0$ has non-zero dimensionality. Let $[x_1, ..., x_\mu]$ be the basis of the eigenspace. We can extend this basis to $[x_1, ..., x_\mu, ..., x_n] = P$, then

$$\begin{aligned} P^{-1}AP &= P^{-1}[\lambda x_1, ..., \lambda x_\mu, Ax_{\mu+1}, ..., Ax_n] \\ &= [\lambda e_1, .., \lambda e_\mu, PAx_{\mu+1}, ..., PAx_n] \\ &= \begin{Bmatrix} \lambda I_\mu & B \\ 0 & D \end{Bmatrix} \end{aligned}$$

Then $Det[A - xI] = Det[P^{-1}AP - xI] = (x - \lambda)^\mu det[D - xI]$ which has least μ roots counting multiplicity. □

Theorem 5.7.8 (diagonalizable matrices). *Let $\lambda_i, i = 1, 2, ..., k \leq n$ be the distinct eigenvalues of $A \in \mathbb{R}^{n\times n}$, let μ_i denote the corresponding algebraic multiplicities, and let $\phi_i = \mathcal{N}(\lambda_i I - A)$. Let further U^i be a matrix containing the basis of ϕ_i.*

- *If $dim(U^i) = \nu_i = \mu_i, \forall i$, the matrix A is said to be* ***diagonalizable****.*
- *Assume A is diagonalizable. then*

$$U = [U^1\ U^2\ ...\ U^k]$$

is invertible, and

$$A = U\Lambda U^{-1}$$

where

$$\begin{bmatrix} \lambda_1 I_{\nu_1} & 0 & 0 & \dots \\ 0 & \lambda_2 I_{\nu_2} & 0 & \dots \\ 0 & 0 & \lambda_3 I_{\nu_3} & \dots \\ & \dots & \dots & \dots \end{bmatrix}$$

- *The space $\mathbb{R}^n$ can be decomposed as the direct sum of all eigenspaces.*

Proof. (2) From the linear independence of the eigenvectors associated with distinct eigenvalues [Lemma 5.7.1], U will contain n linear independent columns, therefore invertible. For every column u in U, we have $Au = \lambda u$; therefore $AU = U\Lambda \Leftrightarrow A = U\Lambda U^{-1}$. (3) Use the criterion for direct sum [Lemma 5.2.2] □

Lemma 5.7.4 (enough distinct eigenvalues implies diagonalizability). *A square matrix $A \in \mathbb{F}^{n\times n}$ can be diagonalized if it has n* ***distinct*** *eigenvalues.*

Proof. use algebraic and geometric multiplicity inequality [Theorem 5.7.7]. □

Lemma 5.7.5 (zero eigenvalue and singularity). *A square matrix A is invertible if and only if it has no zero eigenvalues.*

Proof. use the fact of $Det(A) = \prod \lambda_i$ or $dim(\mathcal{N}(A - 0I)) = dim(\mathcal{N}(A)) \geq 1$. □

Remark 5.7.2. If a square matrix has zero eigenvalue, that means the null space in non-trivial. Then the eigenvector corresponding to the zero-eigenvalue span the null space. Since $(A - \lambda I)x = 0 \Rightarrow Ax = 0$.

Remark 5.7.3. **A matrix is diagonalizable does not imply it is invertible**, since it might contain eigenvalue of 0.

5.8 Eigenvalue and eigenvectors of matrices: case studies

5.8.1 Real diagonalizable matrix

Theorem 5.8.1. *Let $A \in \mathbb{R}^{n\times n}$ has n distinct eigenvalues, then the complex eigenvalues come as conjugate pairs, and corresponding eigenvectors are conjugate to each other.*

Proof. Because the characteristic polynomial coefficients are real-valued, then its complex roots come as conjugate pairs(see polynomial theory section). Let V_1 be the eigenvector of eigenvalue λ_1, then

$$AV_1 = \lambda_1 V_1 \Rightarrow \overline{AV_1} = \overline{\lambda_1 V_1} \Rightarrow A\overline{V_1} = \overline{\lambda_1}\overline{V_1}$$

Therefore $V_2 = \overline{V_1}$ is the eigenvector associated with $\lambda_2 = \overline{\lambda_1}$. □

Remark 5.8.1. Note that V_1 must have non-zero imaginary part, otherwise we cannot have $AV_1 = \lambda_1 V_1$, where λ_1 has non-zero imaginary part.

Theorem 5.8.2 (convert complex eigenvector to real eigenvector). *Let $A \in \mathbb{R}^{n\times n}$ has n distinct eigenvalues. Suppose A has a pair of complex conjugated eigenvalue $\lambda_1 = a+bi, \lambda_2 = a-bi, a, b \in \mathbb{R}$, with a pair of corresponding complex conjugated eigenvectors $V_1 = C+Di, V_2 = C-Di, C, D \in \mathbb{R}^n$ then we can create the 2-D real-valued subspace as*

$$A[C,D] = [C,D]\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

Proof. We have

$$A(C+Di) = (a+bi)(C+Di) = (aC-bD) + i(aD+bc)$$

$$A(C-Di) = (a-bi)(C-Di) = (aC-bD) - i(aD+bc)$$

Sum each other, and we get

$$AC = aC - bD$$

Subtract each other, and we get

$$AD = bC + aD$$

□

Remark 5.8.2. **This conversion is only appealing when we want to make everything real in order to interpret its physical meaning.** It is not appealing in its mathematical structure since it makes two 1D subspace become one 2D subspace.

Corollary 5.8.2.1. *Let $A \in \mathbb{R}^{n\times n}$ has n distinct eigenvalues. Then there exists an invertible matrix T such that*

$$T^{-1}AT = \begin{pmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_k & & & \\ & & & D_1 & & \\ & & & & \ddots & \\ & & & & & D_l \end{pmatrix}$$

where D_j has the form of

$$D_j = \begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix}$$

Moreover, ***every item in this decomposition is real-valued****.*

5.8.2 Real symmetric matrix

5.8.2.1 *Spectral properties*

Diagonalization of a matrix can allow decomposition of a matrix into matrices with favorable properties. Although not all the matrices are diagonalizable, real-valued symmetric matrices can always be **diagonalized** and their **eigenvalues are also real**.

The diagonalibility of real symmetric matrices have many important applications. For example, any quadratic function represented by symmetric matrix can always be completed to square forms. The eigendecomposition of real symmetric matrices also lay the foundation of singular value decomposition, one of most elegant theory of matrix analysis.

Theorem 5.8.3 (Eigen-decomposition of a real symmetric matrix). *Let $A \in \mathbb{R}^{n\times n}$ be symmetric, let $\lambda_i, i = 1, ..., k \leq n$ be the distinct eigenvalues of A, and further let μ_i denote the algebraic multiplicity of λ_i and $\phi_i = \mathcal{N}(\lambda_i I - A)$, we have:*

- $\lambda_i \in \mathbb{R}$
- $\phi_i \perp \phi_j$
- *The eigenvectors can be chosen to lie in* $\mathbb{R}^n$
- $dim\phi_i = \mu_i$

Proof. (1)
$$(Ae_i)^H e_i = \overline{\lambda_i} e_i^H e_i = e_i^H (Ae_i) = \lambda_i e_i^H e_i \Rightarrow \lambda_i = \overline{\lambda_i}$$
(2) See self-adjoint linear operator theory and left right eigenvector orthogonality theory [Theorem 5.7.6].
(3) Let V be an eigenvector with λ, then $\overline{V}$ will be an eigenvector associated with with λ since
$$AV = \lambda V \Rightarrow \overline{AV} = \overline{\lambda V} \Rightarrow A\overline{V} = \lambda \overline{V}$$
then we can remove the imaginary part by $V + \overline{V}$,which is also an eigenvector. (4) Consider λ is a eigenvalue with algebraic multiplicity greater than 1. Let $P = [x_1, ..., x_n] = [x_1 X_2]$ be the orthonormal basis, with x_1 being the normalized eigenvector associated with λ. Then
$$P^T A P = \begin{Bmatrix} \lambda & \lambda x_1^T X_2 \\ \lambda X_2^T x_1 & X_2^T A X_2 \end{Bmatrix} = \begin{Bmatrix} \lambda & 0 \\ 0 & X_2^T A X_2 \end{Bmatrix}$$
Let $B = X_2^T A X_2$. Note that $det(A - tI) = det(P^T A P - tI) = (t - \lambda) det(B - tI)$. Since A has λ with multiplicity of γ, B will has λ with multiplicity of $\gamma - 1$. We can continue the same operation on B, and when we reduce one algebraic multiplicity, we get out of one eigenvector. **The key is after the operation, B is still real symmetric, and we can continue the process**. □

Corollary 5.8.3.1 (Spectral theorem for symmetric matrix). *[5]Let $A \in \mathbb{R}^{n\times n}$ be symmetric, let $\lambda_i, i = 1, 2, ..,$ be the eigenvalues of A (counting multiplicities). Then, there exists a set of orthonormal vectors $u_i, i = 1, 2..., n$ such that $Au_i = \lambda_i u_i$. Equivalently, there exists an orthogonal matrix $U = [u_1, ..., u_n], U^T U = UU^T = I$, such that*
$$A = U \Lambda U^T.$$

Remark 5.8.3. The implication is that any symmetric matrix can be decomposed as a weighted sum of simple rank-one matrix.

Remark 5.8.4 (zero eigenvalue issue). A symmetric matrix might contain zero eigenvalues, in this case, there are diagonal entries in Λ that are zeros. Then the corresponding eigenvectors in U will be the basis span the null space.

Example 5.8.1 (zero matrix). Consider a matrix $A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. The eigenvalue of A is 0, whose algebraic multiplicity is 2. Any basis for $\mathbb{R}^2$ are eigenvectors of A.

Note that A is not invertible.

Example 5.8.2 (diagonal matrix). Consider a diagonal matrix $A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & \vdots \\ \vdots & \cdots & \ddots & \vdots \\ 0 & \cdots & \cdots & a_{nn} \end{bmatrix}$. The eigenvalue of A are $a_{11}, ..., a_{nn}$. The standard basis of $\mathbb{R}^n$ are the eigenvectors of A.

5.8.2.2 *Rayleigh quotients*

Theorem 5.8.4 (Rayleigh quotients). *Given a symmetric matrix $A \in \mathbb{R}^{n\times n}$, it holds that*

$$\lambda_{min}(A) \leq \frac{x^T A x}{x^T x} \leq \lambda_{max}(A), \forall x \neq 0, x \in \mathbb{R}^n.$$

Moreover,

$$\lambda_{max}(A) = \max_{\|x\|_2=1} x^T A x$$

$$\lambda_{min}(A) = \min_{\|x\|_2=1} x^T A x$$

and the maximum and minimum value are attained when x is the unit eigenvector of A associated with its largest and smallest eigenvalues of A, respectively.

Proof. (1) Let $x \neq 0$, let $A = U\Lambda U^T$, then

$$x^T A x = x^T U \Lambda U^T x = y^T \Lambda y = \sum_{i=1}^n \lambda_i y_i^2 \leq \lambda_{max} \|y\|_2^2 = \lambda_{max} \left\| U^T x \right\|_2^2 = \lambda_{max} \|x\|_2^2 .$$

Similarly, we can prove another inequality.
(2) (second method) Using constrained optimization theory, we have

$$\max_{x\in\mathbb{R}^n} x^T A x, \ s.t. \ x^T x = 1.$$

The first order KKT condition gives

$$Ax = \lambda x;$$

that is, optimal x should have the same direction of eigenvectors. It is easy to see optimal x should be the eigenvector with the maximum eigenvalue. □

Corollary 5.8.4.1 (generalized Rayleigh quotients). *Given a symmetric matrix $A \in \mathbb{R}^{n\times n}$ and positive symmetric matrix $\Sigma \in \mathbb{R}^{n\times n}$, it holds that*

$$\lambda_{min}(\Sigma^{-1/2}A\Sigma^{-1/2}) \leq \frac{x^T A x}{x^T \Sigma x} \leq \lambda_{max}(\Sigma^{-1/2}A\Sigma^{-1/2}), \forall x \neq 0, x \in \mathbb{R}^n,$$

where $\Sigma = \Sigma^{1/2}\Sigma^{1/2}$, and $\Sigma^{1/2}$ is a positive semi-definite symmetric matrix and the matrix square root of Σ [Theorem 5.12.3].

The maximium/minimum value is achieved at $x^* = \Sigma^{-1/2}u^*$, where u^* is the unit eigenvector associated with the maximum/minimum eigenvalue of matrix $\Sigma^{-1/2}A\Sigma^{-1/2}$.

Moreover, x^* is also the eigenvectors of $\Sigma^{-1}A$ associated with the maximum/minimum eigenvalue. In other words, the matrix $\Sigma^{-1}A$ and $\Sigma^{-1/2}A\Sigma^{-1/2}$ have the same eigenvalues, and their eigenvectors are connected by $x^* = \Sigma^{-1/2}u^*$.

Proof. (1) Note that

$$\begin{aligned}
\frac{x^T A x}{x^T B x} &= \frac{x^T A x}{x^T \Sigma^{1/2}\Sigma^{1/2} x} \\
&= \frac{x^T A x}{x^T \Sigma^{1/2}\Sigma^{1/2} x} \\
&= \frac{u^T \Sigma^{-1/2} A \Sigma^{-1/2} u}{u^T u} \quad (use\ u = \Sigma^{1/2}x)
\end{aligned}$$

Then we use Theorem 5.8.4. (2) To show the connection of eigenvalue problem of $B^{-1}A$, we have

$$\begin{aligned}
\Sigma^{-1/2}A\Sigma^{-1/2}u^* &= \lambda u^* \\
\Sigma^{-1/2}A\Sigma^{-1/2}\Sigma^{1/2}x &= \lambda \Sigma^{1/2}x \\
\Sigma^{-1/2}Ax &= \lambda \Sigma^{1/2}x \\
\Sigma^{-1/2}\Sigma^{-1/2}Ax &= \lambda x \\
B^{-1}Ax &= \lambda x
\end{aligned}$$

□

Corollary 5.8.4.2. *Let A be an $m \times m$ symmetric matrix with eigenvalues $\lambda_1 \geq \lambda \geq \cdots \geq \lambda_m$, and denote the corresponding normalized eigenvectors as $P_1, P_2, ..., P_m$. Then the supremum of*

$$\sum_{i=1}^{r} x_i^T A x_i = Tr(X^T A X),$$

with $X = [x_1, ..., x_r]$, over all sets of $r \leq m$ mutually orthonormal vectors $x_1, ..., x_r$, is equal to $\sum_{i=1}^{r} \lambda_i$ and is attained when $x_i = P_i, i = 1, 2, ..., r$.

Proof. Use the inequality technique similar in Theorem 5.8.4. □

Corollary 5.8.4.3 (maximization lemma). *Given symmetric positive definite matrix $A \in \mathbb{R}^{n \times n}$ and a vector $d \in \mathbb{R}^p$. It follows that*

$$\max_{x \in \mathbb{R}^p} \frac{(d^T x)^2}{x^T A x}, \; st \; x^T A x = 1$$

has maximum value of $d^T A^{-1} d$, which is attained at $x = \frac{A^{-1}d}{\|A^{-1}d\|^2}$.

Proof. The Lagrange is given by

$$L(x) = x^T d d^T x - \lambda(x^T A x - 1).$$

Then first order KKT condition gives

$$d d^T x = \lambda A x \implies A^{-1} d (d^T x) = \lambda x,$$

that is, optimal x should have the same direction of $A^{-1}d$. The rest is straight forward. □

Corollary 5.8.4.4 (matrix 2-norm). *For a matrix A, if we define its norm as*

$$\|A\|_2 = \max_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2}$$

then

$$\|A\|_2 = \sqrt{\lambda_{max}(A^T A)}$$

Moreover, if A is square, then $\|A\|_2 = \lambda_{max}$

Proof. $\|Ax\| = \sqrt{x^T A^T A x}$ and $A^T A$ is a symmetric matrix. Then we can use Theorem 5.8.4. □

Theorem 5.8.5 (connections of spectral properties of XX^T and X^TX). *Let X be a real-valued matrix, then the eigen decomposition of XX^T and X^TX are related. If*

$$XX^T = U\Lambda U^T$$

then

$$X^TX = V\Lambda V^T$$

That is they have the same ***non-zero*** *eigenvalue.Moreover, $u_i = Xv_i/\sqrt{\lambda_i}, v_i = X^Tu_i/\sqrt{\lambda_i}$*

Proof. XX^T is symmetric and therefore can have a eigen-decomposition. Let u_i be an eigenvector XX^T, then

$$XX^Tu_i = \lambda u_i \Rightarrow X^TXX^Tu_i = \lambda_i X^Tu_i$$

therefore X^Tu_i is an eigenvector of X^TX with length

$$\left\|X^Tu\right\| = \sqrt{u_i^TXX^Tu_i} = \sqrt{\lambda_i u_i^Tu_i} = \sqrt{\lambda_i}$$

The rest is straight forward. □

Remark 5.8.5. This theorem is important in proving SVD theorem.

5.8.2.3 *Pointcare inequality*

Theorem 5.8.6 (Pointcare inequality). *[5, p. 126] Let $A \in \mathbb{F}^{n\times n}$ be a symmetric matrix, and let V be any $k, 1 \leq k \leq n$ dimensional subspace of $\mathbb{R}^n$. Then there exist vectors $x, y \in V$, with $\|x\|_2 = \|y\|_2 = 1$, such that*

$$x^TAx \leq \lambda_k(A),$$

$$y^TAy \geq \lambda_{n-k+1}(A)$$

where λ_k is the kth largest eigenvalue(with λ_n being the largest eigenvalue).($\lambda_1, \lambda_2, ..\lambda_n$ are sorted in increasing order.)

Proof: (1)Let $A = U\Lambda U^T$, let $Q = span(u_k, u_{k+1}, ..., u_n)$ and $dim(Q) = n-k+1$, then $V \cap Q$ is not nonempty (since $dim(Q) + dim(V) > n$). Let $x \in V \cap Q$, and x must have representation of $x = \sum_{i=k}^n \eta_i u_i$, let $U_k = [u_k, u_{k+1}, ..., u_n]$then

$$x^TAx = (U_k\eta)^TU\Lambda U^T(U_k\eta) = \sum_{i=k}^n \eta_i^2\lambda_i \leq \lambda_k \sum_{i=k}^n \eta_i^2 = \lambda_k$$

(2) Consider the matrix $-A$, where $\lambda_k(-A) = -\lambda_{n-k+1}(A)$

Remark 5.8.6 (when the equality hold). When we take $V = span(u_1, ..., u_k)$,the equality will hold.

Remark 5.8.7 (restatement of Rayleigh quotient). When $k = n$, we have

$$x^T Ax \leq \lambda_{max}(A)\|x\|^2$$

and

$$x^T Ax \geq \lambda_{min}(A)\|x\|^2$$

Corollary 5.8.6.1 (minimax principle, Courant-Fisher theorem). *[5, p. 127][9, p. 237] Let $A \in \mathbb{F}^{n\times n}$ be a symmetric matrix, and let V be any subspace of $\mathbb{R}^n$. Then for $k \in \{1, 2, ..., n\}$ it holds that*

$$\lambda_k(A) = \min_{dimV=k} \max_{x\in V,\|x\|_2=1} x^T Ax$$

and

$$\lambda_k(A) = \max_{dimV=n-k+1} \min_{x\in V,\|x\|_2=1} x^T Ax$$

where $\lambda_1, \lambda_2, ..\lambda_n$ are sorted in increasing order.

Proof. Directly from Pointcare inequality. □

Remark 5.8.8 (reduction to Reylergh quotient).

- If we let $dimV = n = k$ in the first equation, we have

$$\lambda_n(A) = \min_{dimV=n} \max_{x\in V,\|x\|_2=1} x^T Ax = \max_{x\in V,\|x\|_2=1} x^T Ax$$

  the minimization operator can be drop because it is the whole ambient space.
- If we let $dimV = n, k = n$ in the second equation, we have

$$\lambda_1(A) = \max_{dimV=n} \min_{x\in V,\|x\|_2=1} x^T Ax = \max_{x\in V,\|x\|_2=1} x^T Ax$$

  the maximization operator can be drop because it is the whole ambient space.

5.8.3 Hermitian matrix

Definition 5.8.1 (conjugate transpose). *Given a matrix $A \in \mathbb{F}^{m\times n}$, the conjugate transpose of A is denoted as A^H such that*

$$(A^H)_{ij} = \overline{A_{ji}}$$

Lemma 5.8.1 (elementary property of conjugate transpose).

- $(A+B)^H = A^H + B^H$
- $(A^H)^H = A$
- $(AB)^H = B^H A^H$
- $(rA)^H = \overline{r} A^H$
- $(A^{-1})^H = (A^H)^{-1}$ *if A is invertiable*

Proof. We only prove (3) and (5). (3): $[(AB)^H]_{ij} = \sum_k \overline{A_{jk}B_{ki}} = \sum_k B^H_{ik}A^H_{kj}$; (5) $A^H(A^{-1})^H = (A^{-1}A)^H = I^H = I$, where we have used (3). □

Lemma 5.8.2 (elementary property of transpose).

- $(A+B)^T = A^T + B^T$
- $(A^H)^T = A$
- $(AB)^T = B^T A^T$
- $(rA)^T = rA^T$
- $(A^{-1})^T = (A^T)^{-1}$ *if A is invertible*

Proof. Similar to above lemma. □

Definition 5.8.2. *A matrix A is Hermitian if $A^H = A$.*

Theorem 5.8.7. *If A is Hermitian, then all eigenvalues of A are real.*

Proof. Same as the real symmetric case. □

Theorem 5.8.8 (spectral theorem for Hermitian matrix). *Let A be a Hermittian matrix, then there exists a unitary matrix U and real diagonal matrix D such that*

$$A = UDU^H$$

Proof. Same as the real symmetric case. □

5.8.4 Matrix congruence

Definition 5.8.3 (congruence). *[9, p. 281] Let $A, B \in \mathbb{F}^{n\times n}$. If there exists a nonsingular matrix S such that*

- *$B = SAS^T$, the B is said to be* ***congruent*** *to A.*
- *$B = SAS^H$, the B is said to be* ****congruent*** *to A.*

Lemma 5.8.3. *Both* ***congruence*** *and* ****congruence*** *are equivalence relationship.*

Proof. This can be easily showed that transitivity, reflectivity and symmetric are satisfied using the property of $(S^T)^{-1} = (S^{-1})^T$ and nonsingular matrix form a group. □

Definition 5.8.4 (inertia). *[9, p. 280] Let $A, B \in \mathbb{F}^{n\times n}$ be* ***Hermitian***. *The* ***inertial*** *of A is the ordered triple*

$$i(A) = (i_+(A), i_-(A), i_0(A)) \in \mathbb{N}^3$$

Definition 5.8.5 (inertia matrix). *[9, p. 282] The inertial matrix for a Hermitian matrix A is defined as*

$$I(A) = I_{i_+} \oplus I_{i_-} \oplus 0_{i_0}$$

Lemma 5.8.4. *Each Hermitian matrix A is* ****congruent*** *to its inertia matrix.*

Proof. Since A is Hermitian, based on spectral decomposed theorem, we have $A = U\Lambda U^H$. We can rewrite Λ as:

$$\Lambda = DI(A)D$$

where

$$D = diag(\lambda_1^{1/2}, ..., \lambda_{i_+}^{1/2}, \lambda_{i_++1}^{1/2}, ..., -\lambda_{i_++i_-}^{1/2}, 1, ..., 1)$$

and $A = U\Lambda U^H = UDI(A)DU^H = SI(A)S^H$ where S is non-singular. Therefore A is *congruence to its inertia matrix. □

Theorem 5.8.9 (Sylvester inertia theorem). *[9, p. 282] Hermitian matrices A, B are *congruence if and only if they have the same inertia.That is*

$$i(A) = i(SAS^H)$$

Proof. (1) forward: If A and B have the same inertial, then both of them will *congruence to the same inertia matrix, and then A and B will be *congruence to each other since *congruence is equilvalence relation.(2) converse: see reference. □

5.8.5 Complex symmetric matrix

Caution!

- Complex symmetric matrices are not Hermitian, and therefore do not admit spectral decomposition.
- Complex symmetric matrices are **fundamentally different from** real symmetric matrices, and therefore do not admit spectral decomposition.

5.8.6 Unitary, orthonormal & rotation matrix

Definition 5.8.6 (unitary matrix). *A complex square matrix U is unitary if*

$$U^H U = UU^H = I$$

A real orthonormal matrix is also unitary.

Theorem 5.8.10. *Any eigenvalue of an unitary matrix has absolute value 1 for real eigenvalue and modulus 1 for complex eigenvalue.*

Proof.

$$Rx = \lambda x \Rightarrow \|Rx\| = \|\lambda x\| = |\lambda|\|x\| = \|x\| \Rightarrow |\lambda| = 1$$

where we have used the preservation of length. □

Definition 5.8.7 (orthonormal matrix). *A real square matrix A is orthonormal matrix if*

$$A^T A = AA^T = I$$

Lemma 5.8.5. *An orthonormal matrix has determinant value of 1 or -1.*

Proof. $Det(AA^T) = 1 = det(A)det(A^T) = det(A)^2$, where we have use $det(A) = det(A^T)$. □

Definition 5.8.8 (rotation matrix). *An orthonormal matrix A is called a rotation matrix if $det(A) = 1$.*

Lemma 5.8.6 (preservation of length). *Let $v \in \mathbb{R}^n$, and $A \in \mathbb{R}^{n\times n}$ is orthonormal matrix, then*

$$\|v\|_2 = \|Av\|_2$$

Proof.

$$\|Av\|_2^2 = v^T A^T A v = v^T v = \|v\|_2^2$$

□

Theorem 5.8.11. *Any eigenvalue of an orthonormal has absolute value 1 for real eigenvalue and modulus 1 for complex eigenvalue.*

Proof.

$$Rx = \lambda x \Rightarrow \|Rx\| = \|\lambda x\| = |\lambda|\|x\| = \|x\| \Rightarrow |\lambda| = 1$$

where we have used the preservation of length. □

Theorem 5.8.12. *Any rotation matrix A in $\mathbb{R}^3$ has a real eigenvalue of 1. The eigenvector of this eigenvalue is called axis of rotation.*

Proof. The characteristic polynomial of A is a 3 degree polynomial with real coefficients, therefore it must have one real eigenvalue. If it has three real eigenvalues, then each of them is 1 or -1. Because the determinant is 1, which is the product the eigenvalues, therefore one eigenvalue must be 1. If it has one pair complex conjugated eigenvalue $a+bi, a-bi$, the real eigenvalue cannot be -1: $(a-bi)(a-bi)(-1) = -a^2 - b^2 < 0$. □

5.9 Singular Value Decomposition theory

5.9.1 SVD fundamentals

Theorem 5.9.1 (complete form SVD). *Any matrix $A \in \mathbb{R}^{m\times n}$ has a factorization given by [Figure 5.9.1]*

$$A = U\Sigma V^T$$

where $U \in \mathbb{R}^{m\times m}, V \in \mathbb{R}^{n\times n}, \Sigma \in \mathbb{R}^{m\times n}$ and Σ is rectangle diagonal matrix. The diagonal entries in Σ, known as ***singular values****, consist of first $r = rank(A)$* ***non-zero, positive, decreasing entries*** *$(\sigma_1, ..., \sigma_r)$, and other zeros.*
Moreover, σ_i^2 is a eigenvalue of matrix AA^T and A^TA; u_i and v_i (columns in U and V) are eigenvectors of AA^T and A^TA, respectively.

Proof. We use following three steps to prove SVD. (1) Consider the matrix AA^T, which is real-valued symmetric and therefore diagonalizable [Theorem 5.8.3]. Let $AA^T = \sum_{i=1}^r \lambda_i u_i u_i^T$, where $\{\lambda_i\}$, $\{u_i\}$ are reversely sorted r non-zero, positive eigenvalues and their eigenvectors of AA^T. Note that A and A^TA have the same rank r [Lemma 5.4.3]; therefore, AA^T only has r non-zero eigenvalues. (2) For each u_i, construct $v_i = \frac{A^Tu_i}{\sqrt{\lambda_i}}$. Now we show that v_i is a unit eigenvector associated with eigenvalue λ_i of A^TA. Note that

$$A^TAv_i = \frac{A^TAA^Tu_i}{\sqrt{\lambda_i}} = \frac{A^T\lambda_i u_i}{\sqrt{\lambda_i}} = \lambda_i v_i,$$

and

$$v_i^Tv_i = \frac{u_i^TA^TAu_i}{\lambda_i} = 1.$$

Therefore, we can write $A^TA = \sum_{i=1}^r \lambda_i v_i v_i^T$. (3) Let U consist of columns $u_1, ..., u_r, u_{r+1}, ..., u_m$, where $u_{r+1}, ..., u_m$ are the basis spanning the $\mathcal{N}(A^TA)$ (or $\mathcal{N}(A^T)$, which is the same since $\mathcal{N}(A^T) = \mathcal{N}(A^TA)$, Lemma 5.4.3). Similarly, let V consist of columns $v_1, ..., v_r, v_{r+1}, ..., v_n$, where $v_{r+1}, ..., v_n$ are the basis spanning $\mathcal{N}(A)$. Note that $u_i^TAv_j = \delta_{ij}\sqrt{\lambda_i}$, therefore

$$U^TAV = \Sigma \implies A = U\Sigma V^T,$$

where Σ is a diagonal matrix with entries of $\sqrt{\lambda_1}, ..., \sqrt{\lambda_r}, 0, ..., 0$. □

Note 5.9.1 (interpretation of blocks).

- The first r columns of U span the range space of A, i.e., $\mathcal{R}(A)$, the last $n-r$ columns span $\mathcal{R}(A)^{\perp}$, and by fundamental theorem of linear algebra, $\mathcal{R}(A)^{\perp} = \mathcal{N}(A^T)$.
- The first r columns of V span a subspace that will contribute to the final result (we denote it as $\mathcal{N}(A)^{\perp}$, and by fundamental theorem of linear algebra $\mathcal{N}(A)^{\perp} = \mathcal{R}(A^T)$), while the last $n-r$ columns span the null space $\mathcal{N}(A)$, which will not contribute to the final result.
- The transformation $y = Ax = U\Sigma V^T x$ can be interpreted in the SVD framework: first map/decompose the vector x into components lying in two subspaces (one space will contribute, and one null space that will not contribute), then only scale the components in the contributing space; finally the scaled components are recovered in the range space of A spanned by the first r columns in U.

Remark 5.9.1 (redundant information in full form SVD).

- If we change the entries in the last $n-r$ columns of V, the resulting matrix from product $U\Sigma V^T$ will not change.
- If we change the entries in the last $m-r$ columns of U, the resulting matrix from product $U\Sigma V^T$ will not change.

Remark 5.9.2 (relationship between U and V). It is a common mistake to think that U and V are orthogonal to each other, i.e. $U^T V = I$. Actually, U and V are orthogonal to each other when A is symmetric. Particularly, we have:

- U consists of the eigenvectors of AA^T, and V consists of the eigenvectors of $A^T A$.
- If A is not square, U and V cannot even multiply together (incompatible sizes).
- If A is symmetric, columns in U and V are eigenvectors of A^2 and A. Therefore, U and V are orthogonal to each other.

Corollary 5.9.1.1 (compact form SVD). *Any matrix $A \in \mathbb{R}^{m\times n}$ has a factorization given by [Figure 5.9.1]*

$$A = \sum_{i=1}^{r} \sigma_i u_i v_i^T = U_r \Sigma_r V_r^T$$

where $U_r \in \mathbb{R}^{m\times r}, V_r \in \mathbb{R}^{r\times n}, \Sigma_r \in \mathbb{R}^{r\times r}$ and $\Sigma = diag(\sigma_1, ..., \sigma_r)$ is diagonal matrix, and the diagonal entries being non-zero/positive decreasing entries.Moreover, $\sigma_i^2 = \lambda_i(AA^T) = \lambda_i(A^T A)$ and u_i and v_i are eigenvectors of $A^T A$ and AA^T.

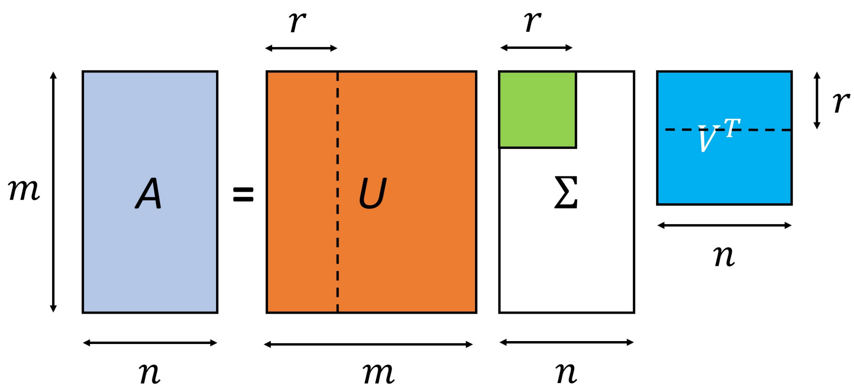
Lemma 5.9.1 (SVD of inverse). *Let A be a invertible matrix with SVD as*

$$A = U\Sigma V^T$$

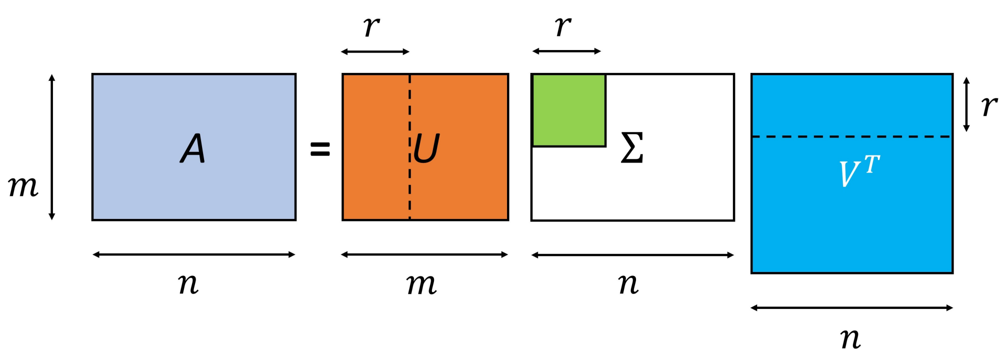
then

$$A^{-1} = V\Sigma^{-1}U^T$$

Proof. $A^{-1} = (U\Sigma V^T)^{-1} = V^{-T}\Sigma^{-1}U^{-1} = V\Sigma^{-1}U^T$. □



(a) Demonstration of SVD for a tall and skinny matrix.



(b) Demonstration of SVD for a short and fat matrix.

Figure 5.9.1: Demonstration of SVD for matrices of different shapes. The dashed lines highlight the compact form SVD.

5.9.2 SVD and matrix norm

The singular values from SVD are closely related to Frobenius norm and 2-norm.

Theorem 5.9.2 (Frobnius norm). *For any matrix $A \in \mathbb{R}^{m \times n}$, then*

$$\|A\|_F^2 = \sum_{i=1}^{r} \sigma_i^2$$

where σ_i are singular values of A.

Proof. $\|A\|_F^2 = Tr(AA^T) = Tr(U\Sigma^2 U^T) = Tr(U^T U\Sigma^2) = Tr(\Sigma^2)$ □

Theorem 5.9.3 (matrix 2-norm). *For any matrix $A \in \mathbb{R}^{m \times n}$, we have*

$$\|A\|_2^2 = \sigma_1^2$$

or

$$\|A\|_2 = \sigma_1$$

where σ_1 is the largest singular value of A.

Particularly, if A is symmetric,

$$\|A\|_2 = \max_i |\lambda_i| .$$

Proof. Because $\|Ax\|_2^2 = x^T A^T Ax$, from Rayleigh quotient theorem [Theorem 5.8.4], we know that for $\|x\| = 1$, the maximum value of $x^T A^T Ax = \lambda_{max}(A^T A) = \sigma_1^2$. □

Lemma 5.9.2 (condition number from SVD). *Let A be a square matrix, then the condition number cond of A:*

$$cond = \frac{\|A\|_2}{\|A^{-1}\|_2} = \frac{\sigma_{max}}{\sigma_{min}}$$

5.9.3 SVD vs. eigendecomposition

SVD and eigendecomposition are two most commonly used matrix decomposition methods. Their key differences and connections include

- Every matrix has SVD but not necessarily eigendecomposition. For example, non-square matrices and non-diagonalizable matrices do not have eigendecomposition.
- For squared and symmetric matrices, SVD and eigendecomposition are closely related, as described by the following Lemma.

Lemma 5.9.3. *If $A \in \mathbb{R}^{n\times n}$ and is symmetric positive semi-definite, then $\lambda_i = \sigma_i$ in sorted order. Moreover, if $A = U\Sigma V^T$ via SVD and $A = W\Lambda W^T$ via eigen-decomposition, then $U = V = W$. (If A is real-symmetric, then $U = V = W$ up to the $\pm$ sign.)*

Proof. From SVD, then U, V are both the eigenvectors of AA; From eigen-decomposition, we have $A = W\Lambda W^T$, and therefore $AA = W\Lambda^2 W^T$, and therefore $W = U = V$. □

- From the derivation of singular value, we know that $\sigma_i^2 = \lambda_i(A^T A) = \lambda_i(A^2) = \lambda_i^2(A)$.

Corollary 5.9.3.1. *If $A \in \mathbb{R}^{n\times n}$ and is symmetric, then $|\lambda_i| = \sigma_i$ in sorted order.*

Caution!
For general square matrix A, the eigenvalue of A might not have simple relations to singular values.

5.9.4 SVD low rank approximation

This section covers the SVD approach to matrix low rank approximation in terms of Frobenius norm and 2-norm.

5.9.4.1 *Frobenius norm low rank approximation*

Lemma 5.9.4 (Unitary invariance of Frobenius norm). *For all $A \in \mathbb{R}^{m\times n}$, $\|A\|_F = \|QAR\|_F$ for Q, R are orthonormal matrices.*

Proof. Use the fact that $\|A\|_F^2 = Tr(AA^T) = Tr(A^T A)$, then $\|QAR\|_F^2 = \left\|QARR^T A^T Q^T\right\| = \left\|Q^T QAA^T\right\| = \left\|A^T A\right\| = \|A\|$. □

Theorem 5.9.4 (Frobenius norm low rank approximation). *Let $A \in \mathbb{R}^{m\times n}$, with $rank(A) = r$, then the minimization problem*

$$\min_{A_k \in \mathbb{R}^{m\times n}, rank(A_k)=k} \|A - A_k\|_F^2$$

with $1 \leq k \leq r$ has the solution

$$A_k^* = \sum_{i=1}^{k} \sigma_i u_i v_i^T$$

with the optimal value of $\sum_{i=k+1}^{r} \sigma_i^2$

Proof. Using the orthonormal invariance of Frobenius norm, we have

$$\left\| \Sigma - U^T A_k V \right\|_F^2$$

Let $Z = U^T A_k V$, Z is better to be diagonal (since off-diagonal terms only make things worst). Then best diagonal matrix Z of rank k can be is first k entries equal σ_i. □

Corollary 5.9.4.1 (rank approximation alternative formulation). *Let S be a matrix of size $m \times n$. Let S have SVD given by*

$$S = U\Sigma V^T.$$

It follows that

- *the value of*
$$\|S - P\|_F^2 = Tr((S-P)(S-P)^T)$$
is minimum among matrices P of the same size but of rank $r \leq rank(S)$, when $P = U_r U_r^T S$, where U_r is $m \times r$ and the columns of U_r are the r normalized eigenvectors of SS^T with the r largest eigenvalues (or the first r columns of U).
- *Alternatively, $P = SV_r V_r^T$, where V_r is $r \times n$ and the columns of V_r are the r normalized eigenvectors of $S^T S$ with the r largest eigenvalues (or the first r columns of V).*

Proof. Note that

$$P = \sum_{i=1}^{r} \sigma_i u_i v_i^T$$

□

Lemma 5.9.5. *For any matrix $A \in \mathbb{R}^{n \times d}$, let $A_k = \sum_{i=1}^{k} \sigma_i u_i v_i^T, k \leq rank(A) = r$, then*

$$\|A - A_k\|_2 = \sigma_{k+1}$$

Proof. Let $A = \sum_{i=1}^{r} \sigma_i u_i v_i^T$ be the SVD of A, then we have

$$A - A_k = \sum_{i=1+k}^{r} \sigma_i u_i v_i^T$$

Based on the definition of 2-norm, we have

$$\|A - A_k\|_2^2 = \max_{\|x\|=1} \left\|(A - A_k)x\right\|_2^2$$

In order to maximize the above, x should lie in the subspace spanned by $v_{k+1}, v_{k+2}, ..., v_r$, and we write $x = \sum_{i=k+1}^r a_i v_i$ then we have

$$\left\|(A - A_k)x\right\|_2^2 = \sum_{i=1+k}^{r} a_i^2 \sigma_i^2 \leq \sigma_{k+1}^2 \sum_{i=k+1}^{r} a_i^2 = \sigma_{k+1}^2$$

and the maximum is attained at $x = v_{k+1}$ □

5.9.4.2 *Two-norm low rank approximation*

Lemma 5.9.6 (Unitary invariance of 2-norm). *For all $A \in \mathbb{R}^{m\times n}$, $\|A\|_2 = \|QAR\|_2$ for Q, R are orthonormal matrices.*

Proof. Use the fact that $\|Ax\|_2^2 = xAA^Tx$, then $\|QARx\|_2^2 = x^tR^TA^TQ^TQARx = x^tR^TA^TARx$, and $\|x\| = \|Rx\|$, therefore

$$\frac{\|Ax\|}{\|x\|} = \frac{\|ARx\|}{\|Rx\|} = \frac{\|Ay\|}{\|y\|}$$

□

Theorem 5.9.5 (matrix 2-norm low rank approximation). *Let $A \in \mathbb{R}^{m\times n}$, with $rank(A) = r$, then minimization problem*

$$\min_{A_k \in \mathbb{R}^{m\times n}, rank(A_k)\leq k} \|A - A_k\|_2^2$$

with $1 \leq k \leq r$ has the solution

$$A_k^* = \sum_{i=1}^{k} \sigma_k u_i v_i^T$$

and

$$\left\|A - A_k^*\right\|_2 = \sigma_{k+1}^2$$

Proof. Since A_k has at most rank k, then its null space has at least dimensionality of $n - k$ based on then rank-nullity theorem. Consider the subspace S spanned by $v_1, v_2, ..., v_{k+1}$, then $S \cap \mathcal{N}(A_k) \neq \emptyset$, based on dimensionality argument ($dim(S) + dim(\mathcal{N} A_k) > n$). Let $x \in S \cap \mathcal{N}(A_k), \|x\| = 1$, then $x = \sum_{i=1}^{k+1} a_i v_i$. We have

$$\left\|(A - A_k)x\right\| = \|Ax\| \geq \sigma_{k+1}$$

where the minimum is attained when $x = v_{k+1}, B = A_k^*$. And therefore

$$\|A - A_k\|_2 \geq \left\|(A - A_k)x\right\| \geq \sigma_{k+1}$$

Note:(1) one might wander why we do not let S spanned by $v_1, v_2, ..., v_{k+2}$ and take $x = v_{k+2}$, we can do so and will obtain a looser bound of

$$\|A - A_k\|_2 \geq \sigma_{k+2}$$

(2) On the other hand, if S spanned by k singular vector, then $\|A - B\| \geq \left\|(A - B)y\right\| \geq \left\|(A - B)\right\| x$, where x is in the subspace spanned by k singular vector, and y is in the subspace spanned by $k+1$ singular vectors. Therefore, we can get tighter bound when use subspace spanned by $k+1$ singular vectors. (3) Therefore, we can get the tightest bound when use subspace spanned by $k+1$ singular vectors. Then it is easy to prove that among all the subspaces span by the $k+1$ singular vectors, the maximum of the minimum in the $k+1$ singular values is σ_{k+1}.

□

5.10 Generalized eigenvectors and Jordan normal forms

5.10.1 Generalized eigenvectors

Definition 5.10.1 (defective eigenvalue). *An eigenvalue λ_i of matrix A is called* ***defective*** *if its geometric multiplicity ,dim$(\mathcal{N}(A-\lambda_i I))$, is strictly less than its algebraic multiplicity.*

Example 5.10.1. Consider the matrix

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

This matrix has eigenvalue $\lambda = 1$ with multiplicity of 2. Note that

$$A - I = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

which is clear that the range space dimension is 1 and therefore the null space dimension is 1.

So the eigenvalue 1 is defective.

Definition 5.10.2 (generalized eigenvector of rank r). *For a given eigenvalue λ, the vector x is a* ***generalized eigenvector of rank*** *r if*

$$(A-\lambda I)^r x = 0$$
$$(A-\lambda I)^{r-1} x \neq 0.$$

Particularly, the eigenvector v is a generalized eigenvector of rank 1 since

$$(A-\lambda I)v = 0, (A-\lambda I)^0 v = v \neq 0.$$

Remark 5.10.1. If a vector u is the generalized eigenvector of rank s, then it cannot be the generalized eigenvector of rank $m > s$, because

$$(A-\lambda I)^m u = 0, (A-\lambda I)^{m-1} u = 0.$$

Definition 5.10.3 (a chain of generalized eigenvectors of length r). *Given an eigenvalue λ, we say that vectors $v_1, v_2, ..., v_r$ form a* ***chain of generated eigenvectors of length r*** *if $v_1 \neq 0$ and*

$$\begin{aligned} v_{r-1} &= (A-\lambda I)v_r \\ v_{r-2} &= (A-\lambda I)v_{r-1} \\ &\vdots \\ v_1 &= (A-\lambda I)v_2 \\ 0 &= (A-\lambda I)v_1. \end{aligned}$$

We can write down the matrix form as

$$A\begin{bmatrix} v_r \\ v_{r-1} \\ \vdots \\ v_1 \end{bmatrix} = \begin{bmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}\begin{bmatrix} v_r \\ v_{r-1} \\ \vdots \\ v_1 \end{bmatrix}.$$

Remark 5.10.2 (generate a chain of generalized eigenvectors of length r)**.** Given an eigenvalue λ and its generalized eigenvector u of rank r, that is

$$\begin{aligned} (A-\lambda I)^r u &= 0 \\ (A-\lambda I)^{r-1} u &\neq 0. \end{aligned}$$

We can define vectors $v_1, v_2, ..., v_r$ as follows

$$\begin{aligned} v_r &= (A-\lambda I)^0 u = u \\ v_{r-1} &= (A-\lambda I)^1 u \\ \vdots v_1 &= (A-\lambda I)^{r-1} u \end{aligned}$$

Theorem 5.10.1 (linear independence among a chain of generalized eigenvectors). *The vectors in a chain of generalized eigenvectors, $v_1, v_2, ..., v_r$ given by,*

$$\begin{aligned} v_{r-1} &= (A-\lambda I)v_r \\ v_{r-2} &= (A-\lambda I)v_{r-1} \\ &\vdots \\ v_1 &= (A-\lambda I)v_2 \\ 0 &= (A-\lambda I)v_1, \end{aligned}$$

are linearly independent.

Proof. We consider the linear combination

$$\sum_{i=1}^{r} a_i v_i = 0.(*)$$

Using the chain definition, we have

$$v_i = (A-\lambda I)^{r-i} v_r;$$

Equation (*) becomes

$$\sum_{i=1}^{r} a_i (A-\lambda I)^{r-i} v_r = 0.(**)$$

We multiply $(A-\lambda I)^{r-1}$ to (**), we get

$$a_r (A-\lambda I)^{r-1} v_r = a_r v_1 = 0 \implies a_r = 0.$$

Similarly, we multiply $(A-\lambda I)^{r-2}$ to (**), we get

$$a_{r-1}(A-\lambda I)^{r-2} v_{r-1} = a_{r-1} v_1 = 0 \implies a_{r-1} = 0.$$

We can continue to prove

$$a_1 = a_2 = ... = a_r = 0.$$

Therefore,

$$v_1, v_2, ..., v_r$$

are linearly independent. □

Definition 5.10.4 (generalized eigenvector). *A generalized eigenvector x for eigenvalue λ is a solution to $(A-\lambda I)^k x = 0$.*

Definition 5.10.5 (generalized eigenspace). *The generalized eigenspace $G(\lambda, A)$ is the set of all generalized eigenvectors associated with the eigenvalue λ of matrix A.*

Lemma 5.10.1. *If $x \in \mathcal{N}(A - \lambda I)$ with k be any positive integer, then*

$$x \in \mathcal{N}((A - \lambda I)^k)$$

that is a eigenvector of λ is also a generalized eigenvector of λ.

Proof.

$$(A - \lambda I)^k v = (A - \lambda I)^{k-1}(A - \lambda I)v = 0$$

□

Caution! Possible linear dependence between different generalized eigenvectors.
Let λ_i be a eigenvalue with multiplicity of $k > 1$, then the generalized eigenvector solved from $(A - \lambda_i)^m v_1 = 0$ and $(A - \lambda_i)^n v_2 = 0$, where $m \neq n$, then v_1 might linearly depend on v_2.
Suppose $m < n$, then v_1 is also the solution of $(A - \lambda_i)^n v_2 = 0$.

Theorem 5.10.2 (The dimensionality of generalized eigenspace). *[3, p. 149] Let λ_i be a eigenvalue with algebraic multiplicity of $k > 1$, then the generalized eigenspace associated with λ_i has dimensionality k.*

5.10.2 Upper triangle matrix and nilpotent matrix

Theorem 5.10.3. *[3, p. 149] Let $A \in \mathbb{C}^{n \times n}$. There exists a nonsingular S such that*

$$T = S^{-1}AS$$

*is upper-triangular. Moreover, T has the same eigenvalues as A with the same multiplicity, showing **on the diagonal**.*

Proof. for the existence proof, see ref. □

Lemma 5.10.2. *The finite powers of a upper-triangular matrix (all diagonal entries $a_{11}, a_{22}, ...a_{nn}$ are nonzeros) $A \in \mathbb{F}^{n\times n}$, i.e., A^k, will still be upper triangular((all diagonal entries are nonzero). Moreover, the diagonal terms of A^k is $a_{11}^k, a_{22}^k, ..., a_{nn}^k$*

Proof. can be directly proved via matrix multiplication. □

Definition 5.10.6 (nilpotent matrix). *An nilpotent matrix A is an $n \times n$ matrix such that there exists a finite power $k \leq n$ for which $A^k = 0$.*

Lemma 5.10.3. *If $A \in \mathbb{F}^{n\times n}$ is nilpotent, then $A^n = 0$.*

Proof. by definition, there exists $k \leq n$ for which $A^k = 0$. □

Definition 5.10.7 (strictly upper-triangular matrix). *A matrix $A \in \mathbb{F}^{n\times n}$ is strictly upper-tiangular if all entries on and below the diagonal are 0.*

Lemma 5.10.4. *For a strictly upper-triangular matrix $A \in \mathbb{F}^{n\times n}$, A is nilpotent and $A^n = 0$.*

Proof: consider how A^k acts on standard basis $e_1, e_2, ..., e_n$.

$$\begin{aligned}
Ae_1 &= 0 \\
Ae_2 &\in span(e_1) \\
A^2e_2 &= 0 \\
Ae_3 &\in span(e_1, e_2) \\
A^2e_3 &\in span(e_1) \\
A^3e_3 &= 0 \\
&\cdots
\end{aligned}$$

Lemma 5.10.5. *[3, p. 242] Let $A \in \mathbb{F}^{n\times n}$, then*

- $$\{0\} = \mathcal{N}(A^0) \subseteq \mathcal{N}(A^1) \subseteq \mathcal{N}(A^2) \subseteq \mathcal{N}(A^3) \subseteq \ldots$$
- *If there is an nonnegative integer m such that $\mathcal{N}(A^m) = \mathcal{N}(A^{m+1})$, then $\mathcal{N}(A^m) = \mathcal{N}(A^{m+1}) = \mathcal{N}(A^{m+2}) = \mathcal{N}(A^{m+3})...$*

- *If there is an nonnegative integer m such that $\mathcal{N}(A^m) \subsetneq \mathcal{N}(A^{m+1})$, then for all $k \leq m$*
$$\mathcal{N}(A^k) \subsetneq \mathcal{N}(A^{k+1})$$

Proof. (1) $A^k x = 0 \Rightarrow AA^k x = A^{k+1}x = 0$; (2) Suppose there exist an integer $k > 0$ such that
$$\mathcal{N}(A^m + k) \neq \mathcal{N}(A^{m+1+k})$$
or equivalently
$$\mathcal{N}(A^m + k) \subsetneq \mathcal{N}(A^{m+1+k})$$
then there exists a v such that $A^{m+1+k}x = 0$ but $A^{m+k}x \neq 0$, then for
$$A^{m+1}A^k x = 0, A^m A^k x \neq 0$$
therefore, $\mathcal{N}(A^{m+1}) \subsetneq \mathcal{N}(A^m)$ because $A^k x$ is in the former but not the latter. (3) directly from (2) by contradiction. □

Remark 5.10.3. For similar inequality of range, see [3, p. 251].

Lemma 5.10.6 (null space saturation). *Let $A \in \mathbb{R}^{n\times n}$, then*
$$\mathcal{N}(A^n) = \mathcal{N}(A^{n+1}) = \mathcal{N}(A^{n+2})...$$

Proof. Suppose $\mathcal{N}(A^n) \subsetneq \mathcal{N}(A^{n+1})$, then from above theorem, we have for $k \leq n$, $\mathcal{N}(A^k) \subsetneq \mathcal{N}(A^{k+1})$, as a result, the dimensionality of $\mathcal{N}A^n$ will be $n+1$, which is impossible for a n by n matrix. □

5.10.3 Jordan normal forms

Definition 5.10.8 (Jordan basis, Jordan canonical form). *[3, p. 273] Suppose $T \in \mathcal{L}(V)$. A basis of V is called a* ***Jordan basis*** *for T if with respect to this basis T has a block diagonal matrix*
$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_p \end{pmatrix}$$

where each A_j is an upper triangular matrix of the form

$$\begin{pmatrix} \lambda_j & 1 & & 0 \\ & \ddots & \ddots & 0 \\ & & \ddots & 1 \\ 0 & & & \lambda_j \end{pmatrix}$$

which is known as ***Jordan block***.

Example 5.10.2. The matrix B is a Jordan basis(Jordan canonical form)

$$B = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix}$$

Theorem 5.10.4 (Jordan decomposition). *[3, p. 273]Suppose V is a complex vector space. If $T \in \mathcal{L}(V)$, then there is a basis of V that is a Jordan basis for T. In matrix form, we have*

$$M = PJP^{-1}$$

where J is the Jordan basis and P is the invertible matrix.

Remark 5.10.4. If matrix A is diagonalizable, then Jordan decomposition reduce to eigen-decomposition.

Lemma 5.10.7 (matrix function of Jordan block). *[10][1, p. 600] Let $f(z)$ be an analytical function of a complex argument. Applying the function on a Jordan block $A \in \mathbb{R}^{n\times n}$ given as*

$$\begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & 0 \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}$$

then $f(z)$ is given as

$$\begin{pmatrix} f(\lambda) & f'(\lambda) & & \frac{f^{n-1}(\lambda)}{(n-1)!} \\ & \ddots & \ddots & 0 \\ & & \ddots & \frac{f^{n-1}(\lambda)}{(n-1)!} \\ 0 & & & f(\lambda) \end{pmatrix}$$

Proof. Use Taylor expansion on $f(z)$, which is analytical and Taylor series exists. □

Lemma 5.10.8 (The power of Jordan block). *Let A be a Jordan block given as*

$$\begin{pmatrix} \lambda_j & 1 & & 0 \\ & \ddots & \ddots & 0 \\ & & \ddots & 1 \\ 0 & & & \lambda_j \end{pmatrix}$$

then A^m is given as

$$\begin{pmatrix} \lambda^m & \binom{m}{1}\lambda^{m-1} & \dots & \binom{m}{n-1}\lambda^{m-n+1} \\ & \ddots & \ddots & 0 \\ & & \ddots & \binom{m}{1}\lambda^{m-1} \\ 0 & & & \lambda^m \end{pmatrix}$$

Proof. This can be directly verified. □

Lemma 5.10.9 (The exponential of Jordan block). *Let $A \in \mathbb{R}^{n\times n}$ be a Jordan block given as*

$$\begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & 0 \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}$$

then $\exp(At)$ *is given as*

$$\begin{pmatrix} \exp(\lambda t) & \binom{m}{1}\lambda t \exp(\lambda t) & \dots & \binom{m}{n-1}(\lambda t)^{m-n-1}\exp(\lambda t) \\ & \ddots & \ddots & 0 \\ & & \ddots & \binom{m}{1}\lambda t \exp(\lambda t) \\ 0 & & & \exp(\lambda t) \end{pmatrix}$$

Lemma 5.10.10 (conditional boundedness of matrix exponential). *Given an arbitrary square matrix A, if all its eigenvalue has a negative real part, then* $\exp(At) \to 0$ *as $t \to \infty$.*

Proof. Directly from above. □

Lemma 5.10.11 (conditional boundedness of matrix power). *Given an arbitrary square matrix A, if all its eigenvalue $|\lambda_i| < 1$(absolute sign is interpreted as modulus for complex number), then $\|A^m\|_2$ is bounded for any positive number m. Moreover, $A^m \to 0$ as $m \to \infty$.*

Proof. From Jordan decomposition, $A = PJP^{-1}$, then $A^m = PJ^mP^{-1}$. Note that J^m has diagonal entries of λ_i^m. Therefore, J^m is a matrix with every entry finite(including the corner terms $\binom{m}{n}\lambda^m \to 0$), therefore A^m is bounded for every m and will go to 0. □

5.11 Matrix factorization

5.11.1 Orthogonal-triangular decomposition

Theorem 5.11.1 (QR decomposition properties). *Suppose $A \in \mathbb{R}^{m\times n}, m \geq n$. Then we have*

- *there exists a orthonormal matrix $Q \in \mathbb{R}^{m\times m}$ and upper triangular matrix $R \in \mathbb{R}^{m\times n}$ such that*
$$A = QR$$
- *If $\hat{Q} \in \mathbb{R}^{n\times n}$ and $\hat{R} \in \mathbb{R}^{n\times n}$, then*
$$A = QR = [\hat{Q}, N]\begin{bmatrix}\hat{R}\\ 0\end{bmatrix} = \hat{Q}\hat{R}$$
where $\hat{Q} \in \mathbb{R}^{m\times n}$ consists of the basis of $\mathcal{R}(A)$, N consists of the basis of $\mathcal{N}(A^T)$ and $\hat{R} \in \mathbb{R}^{n\times n}$.
- *We can choose R to have nonnegative diagonal entries*
- *If A is of full rank, we can choose R with positive diagonal entries, in which case the economical form $\hat{Q}$ and $\hat{R}$ will be unique.*
- *If A is square nonsingular, then $A = QR$ is unique.*

Proof. (1)(2) Consider the Gram-Smith process for the columns of matrix A given as

$$\begin{aligned}
q_1 &= a_1, p_1 = q_1/\|1_1\| \\
q_i &= a_i - \sum_{j=1}^{i-1}\left\langle a_i, p_j\right\rangle p_j, p_i = p_i/\left\|q_i\right\|, i = 2, ..., n
\end{aligned}$$

or

$$\begin{aligned}
a_1 &= r_{11}p_1 \\
a_j &= \sum_{i=1}^{j} r_{ij}p_i, j = 2, ..., n \\
r_{ii} &= \left\|q_i\right\|, r_{ij} = \left\langle a_j, p_i\right\rangle
\end{aligned}$$

in which orthonormal basis $p_1, ..., p_n$ for the column space $span(a_1, ..., a_n)$ will be produced. We can see that $a_i \in span(p_1, ..., p_i)$, and therefore in matrix form we have

$$A = \hat{Q}\hat{R}$$

where $\hat{Q} \in \mathbb{R}^{m\times n}$ will consist of $p_1, ..., p_n$ as columns and $R \in \mathbb{R}^{n\times n}$ will an upper triangular matrix. The complete form Q can be augmented with basis of $\mathcal{N}(A^T) = \mathcal{R}(A)^\perp$, such that $Q \in \mathbb{R}^{m\times m}$ consist of the complete orthonormal basis of $\mathbb{R}^m$. (3)(4)(5) If $a_1, ..., a_n$ are linearly independent, from GS process, the matrix $\hat{Q}, \hat{R}$ are uniquely determined, and the diagonal entries of R is always positive. If $a_1, ..., a_n$ are linearly dependent, there will exist scenario that

$$a_k \in span(p_1, ..., p_{k-1})$$

and we can set $r_{kk} = 0$. □

Remark 5.11.1. QR decomposition is the matrix form of Gram-Smith procedures.

5.11.2 LU decomposition

Definition 5.11.1 (LU decomposition with partial pivoting, LUP). *The LU decomposition with partial pivoting for a square matrix A is given as*

$$PA = LU$$

where P is the permutation matrix (to reorder rows in A), L and U are the lower and upper triangle matrix.

Remark 5.11.2 (existence and uniqueness)**.**

- If P is not used to reorder rows, then LU decomposition might not exist.
- Any square matrix A admits an LUP decomposition.
- The LUP decomposition is not unique(for example, set $L' = -L, U' = -U$).

5.11.3 Cholesky decomposition

Definition 5.11.2 (Cholesky decomposition). *The Cholesky decomposition of a Hermitian positive definite matrix A is a decompsition of the form*

$$A = LL^H$$

where L is a lower triangular matrix with real and positive diagonal entries, and L^H denotes the Hermitian.

Remark 5.11.3 (extension to positive semidefinite matrix)**.** If we allow the diagonal entries to be zero, then positive semi-definite matrix also has Cholesky decomposition(might not be unique).

Remark 5.11.4 (QR decomposition vs. Cholesky decomposition,existence and uniqueness)**.**

- Note that any for real symmetric positive definite matrix A, we know that $A = BB^T$ [Theorem 5.12.2]. If we do a unique QR decomposition on B as $B^T = QL^T$, then $BB^T = LL^T$.
- The Cholesky decomposition for positive semidefinite matrices always exists. It is unique only for positive definite matrices.

5.12 Positive definite matrices and quadratic forms

5.12.1 Quadratic forms

Definition 5.12.1 (quadratic forms). *For an $n \times n$ matrix A, the quadratic form associated with A is defined as*

$$x^T Ax = \sum_{i,j} a_{ij} x_i x_j$$

caution! Note that given a quadratic form $\sum_{i,j} a_{ij} x_i x_j$, the matrix A satisfying

$$x^T Ax = \sum_{i,j} a_{ij} x_i x_j$$

is not unique, unless we require A to be symmetric.

Lemma 5.12.1 (every quadratic form is associated with an unique symmetric matrix). *Given a square matrix A with its associated quadratic form $x^T Ax$, there exist an unique symmetric matrix B such that*

$$x^T Ax = x^T Bx$$

where

$$B = \frac{1}{2}(A + A^T)$$

That is, every quadratic form is associated with an unique symmetric matrix.

Proof. We have

$$x^T Ax = (x^T Ax)^T = x^T A^T x$$

then

$$\frac{1}{2} x^T (A + A^T) x = x^T Ax$$

is proved. To show uniqueness, note that by equaling the coefficients, we have $a_{ij} = B_{ij} + B_{ji}$, the symmetry requirement impose $B_{ij} = B_{ji}$ and then therefore given a quadratic form, its associated symmetric matrix is unique. □

Lemma 5.12.2. *Let A be symmetric square matrix. Then $x^T Ax = 0$ for every $x \in \mathbb{R}^n$ if and only if $A = 0$.*

Proof. (1) forward part is straight forword; (2) The converse part: (a)set $x = e_i$, and we get $e_i Ae_i = a_{ii} = 0$; (b) set $x = e_j + e_k$, and we get $x^A x = 2a_{jk} = 0$ □

Example 5.12.1.

- Let $A = \begin{bmatrix} 3 & 0 \\ 0 & 4 \end{bmatrix}$. Then $x^T Ax = 3x_1^2 + 4x_2^2$.
- Let $A = \begin{bmatrix} 3 & -2 \\ -2 & 5 \end{bmatrix}$. Then $x^T Ax = 3x_1^2 + 5x_2^2 - 4x_1 x_2$.

5.12.2 Real symmetric non-negative definite matrix

5.12.2.1 *Characterization*

Definition 5.12.2 (non-negative definite, positive definite).

- *A square matrix $A \in \mathbb{R}^{n\times n}$ is **non-negative definite** if*
$$x^T Ax \geq 0, \forall x \in \mathbb{R}^n.$$
- *A square matrix $A \in \mathbb{R}^{n\times n}$ is **positive definite** if*
$$x^T Ax > 0, \forall x \in \mathbb{R}^n, x \neq 0.$$

Lemma 5.12.3 (characterization by eigenvalues and diagonal entries).

- *Let A be a real symmetric matrix. If A is non-negative definite, then*
  - ***non-negative real eigenvalues.****(necessary and sufficient)*
  - ***non-negative diagonal entries****(necessary conditions)*
- *Let A be a real symmetric matrix. If A is positive definite, then*
  - ***positive real eigenvalues.****(necessary and sufficient)*
  - ***positive diagonal entries****(necessary conditions)*

Proof. (1)(necessary) The proof of eigenvalues are real directly from the results in symmetric matrix. For the non-negativity, let u_i be an unit eigenvector corresponding to eigenvalue λ_i, we have

$$Au_i = \lambda_i u_i \Rightarrow u_i^T A u_i = \lambda_i \geq 0.$$

(sufficient) Let A have eigendecomposition of $A = P\Lambda P^T$. Then for any $x \in \mathbb{R}^n$

$$\begin{aligned} x^T A x &= xP\lambda Px \\ &= y\lambda y \\ &= \sum_{i=1}^{n} y_i^2 \lambda \geq 0 \end{aligned}$$

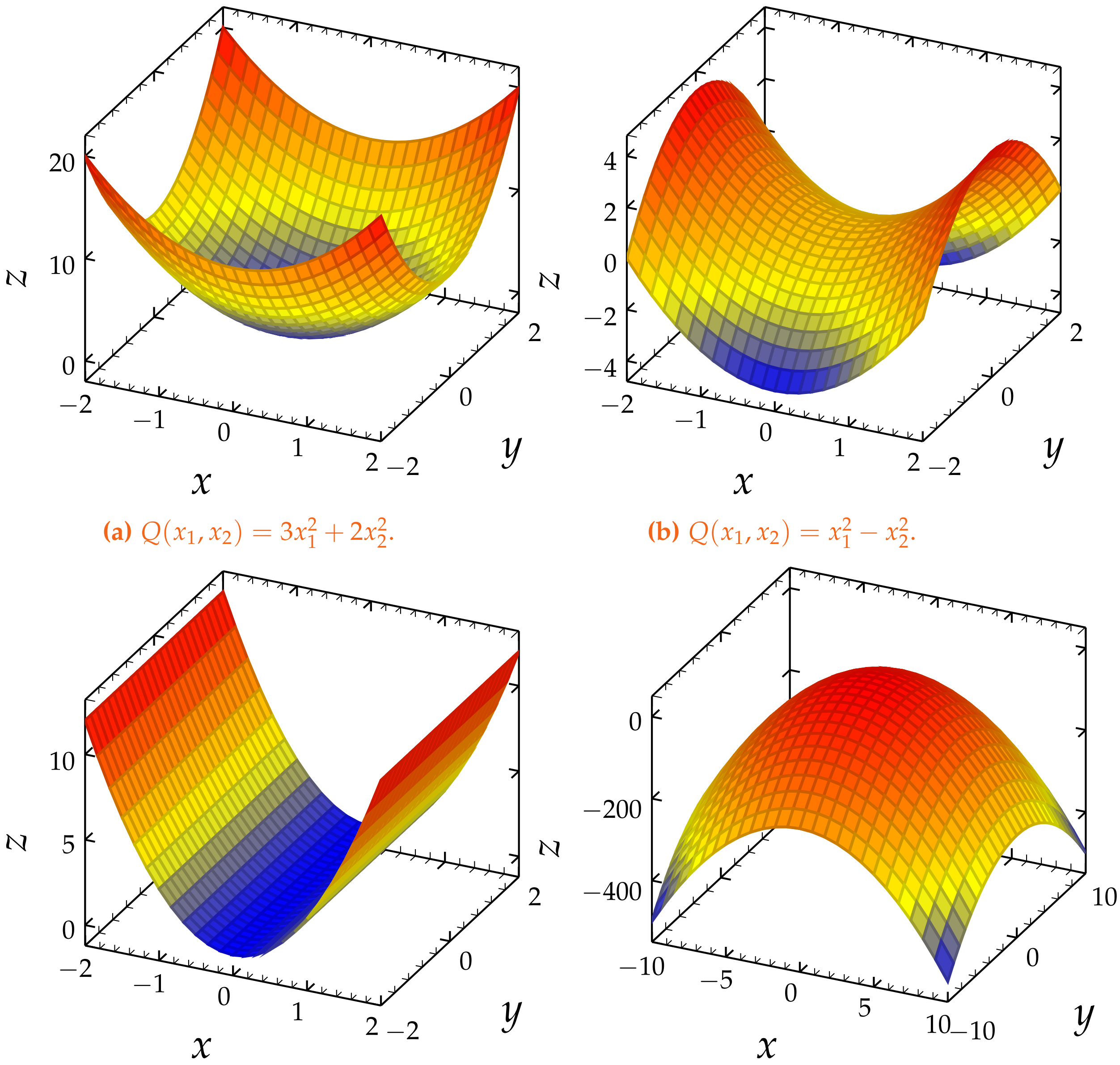
(2) Let e_i be the standard basis, then

$$e_i^T A e_i = a_{ii} \geq 0$$

□

Example 5.12.2. In Figure 5.12.1, we illustrate different quadratic forms.

- $Q(x_1, x_2) = 3x_1^2 + 2x_2^2$, whose eigenvalues are 3 and 2.
- $Q(x_1, x_2) = x_1^2 - x_2^2$, whose eigenvalues are 1 and -1.
- $Q(x_1, x_2) = 3x_1^2$, whose eigenvalues are 3 and 0.
- $Q(x_1, x_2) = -3x_1^2 - 2x_2^2$, whose eigenvalues are -3 and -2.



(a) $Q(x_1, x_2) = 3x_1^2 + 2x_2^2$. **(b)** $Q(x_1, x_2) = x_1^2 - x_2^2$.

(c) $Q(x_1, x_2) = 3x_1^2$. **(d)** $Q(x_1, x_2) = -3x_1^2 - 2x_2^2$.

Figure 5.12.1: Illustration of different quadratic forms.

Lemma 5.12.4 (characterization by submatrix). *Let A be a $n \times n$ symmetric matrix and $Q = x^T Ax, x \in \mathbb{R}^n$. Let A_k be the $k \times k$ submatrix of A such that $A_k = (A)_{1 \leq i \leq k, 1 \leq j \leq k}$. Then the following statements are equivalent:*

- *$Q > 0$ for all $x \in \mathbb{R}^n, x \neq 0$.*
- *All eigenvalues of A are positive.*

- *For each $1 \leq k \leq n$, A_k is positive definite.*
- *$det(A_k) > 0$, for $1 \leq k \leq n$.*

Proof. (1) is equivalent (2) is from Lemma 5.12.3. (1) implies (3): Assume $Q > 0$ for all $x \neq 0$. Then for any $1 \leq k \leq n$,

$$\begin{aligned} 0 &< (x_1, ..., x_k, 0, ..., 0)A(x_1, ..., x_k, 0, ..., 0)^T \\ &= (x_1, ..., x_k)A_k(x_1, ..., x_k)^T \\ &= Q_k \end{aligned}$$

for all $(x_1, ..., x_k) \neq 0$. Therefore, A_k is positive definite. (3) implies (4): A_k has all positive eigenvalues. The determinant is the product of all eigenvalues. (4) implies (1)(2). The determinant is the product of all eigenvalues. We can get that every eigenvalue is positive if $det(A_k) > 0$, for $1 \leq k \leq n$. □

5.12.2.2 *Decomposition and transformation*

Theorem 5.12.1 (preserving positive definiteness). *Let $A \in \mathbb{R}^{n\times n}$ be a symmetric positive definite matrix,let $P \in \mathbb{R}^{n\times k}$, if P has full column rank, then*

$$P^T AP$$

is still symmetric positive definite.

Proof. Since $dim(\mathcal{N}(P)) = 0$, $Px \neq 0, \forall x \neq 0$, and therefore $y^T Ay > 0$, if $y = Px \neq 0$ □

Corollary 5.12.1.1. *Let $A \in \mathbb{R}^{n\times n}$ be a symmetric positive definite matrix,and let $P \in \mathbb{R}^{n\times n}$. if P is nonsingular, then*

$$P^T AP$$

is still symmetric positive definite.

Theorem 5.12.2 (decomposition). *Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix.*

- *A is non-negative definite if and only if there exists $B \in \mathbb{R}^{n\times k}$ such that $A = BB^T$*
- *A is positive definite if and only if there exists nonsingular $B \in \mathbb{R}^{n\times n}$ such that $A = BB^T$*

Note that B is usually not unique.

Proof. (1)(a)forward: If $A = BB^T$, then for any $x \in \mathbb{R}^n$, $x^T Ax = (xB)^T(Bx) = \|Bx\|^2 \geq 0$, and thus A is non-negative definite. (b)converse: Because A is symmetric, we know that it can be diagonalized as

$$A = V\Lambda V^T$$

because A have non-negative eigenvalues, let $B = V\Lambda^{1/2}$ and complete the proof.
(2) similar to (1). □

Remark 5.12.1 (Compare with Cholesky decomposition). Cholesky decomposition is usually decompose a positive symmetric matrix into the product of a **lower triangular matrix** and its conjugate transpose.

Corollary 5.12.2.1. *Orthogonal projectors P are nonnegative/semi-positive definite.*

Proof. P is orthogonal projector and therefore is symmetric and idempotent. That is $P^2 = P$ and $P^T = P$, therefore $P = P^T P$ and thus P is nonnegative definite. □

Lemma 5.12.5. *Let A be a matrix, then AA^T and $A^T A$ has the same non-zero eigenvalues.*

Proof. Let $\lambda \neq 0$ be an eigenvalue of $A^T A$, i.e.

$$A^T Ax = \lambda x$$

for some x. then

$$(AA^T)Ax = A\lambda x = \lambda Ax$$

that is Ax is the eigenvector of AA^T associated with eigenvalue λ. Therefore, λ is also the eigenvalue of AA^T. □

Remark 5.12.2. Both $A^T A$ and AA^T are symmetric, but might have different dimensions.

Lemma 5.12.6. *Given a semi-positive definite symmetric matrix H, $H + aI$ is positive definite for $a > 0$.*

Proof. for H we can decompose as $H = RR^T$, therefore for any nonzero x, $x^T(H + aI)x = xRR^T x + x^T x > 0$. □

5.12.2.3 *Matrix square root*

Theorem 5.12.3 (matrix square root). *Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix. If A is positive definite or non-negative definite, then there exists a* ***positive definite or non-negative definite symmetric matrix*** *B such that $A = B^2$. Moreover, B is uniquely (up to order of eigenvectors) given by*

$$B = U\Lambda^{1/2}U^T$$

where U and Λ are matrices associated with the eigen-decomposition of $A = U\Lambda U^T$.

Proof. It can be verified that $B^2 = A$. To prove the uniqueness, we have (1) B has to be positive definite, because $rank(A) = rank(BB) = rank(B)$

$$B_1 = WD_1W^T, B_2 = VD_2V^T$$

, $B_1^2 = B^2$ implies $WD_1^2W^T = VD_2^2V^T, D_1 = D_2$ □

Remark 5.12.3 (different versions of square root). In engineering applications, there are many definitions of a square root for a matrix. For example, in Cholesky decomposition $A = LL^T$, the triangular matrix L (which is not a symmetric matrix) is usually referred as square root of A. See [11] for summary and discussion.

Corollary 5.12.3.1 (inverse of matrix square root).

$$(A^{-1})^{1/2} = (A^{1/2})^{-1}$$

Proof. We have $B = A^{1/2}$, $BB = A$,$A^{-1} = B^{-1}B^{-1}$, and therefore $B^{-1} = (A^{-1})^{1/2}$. □

5.12.2.4 *Maximization of quadratic forms*

Theorem 5.12.4 (maximization of quadratic forms on the unit sphere). *Let $B \in \mathbb{R}^{n\times n}$ be a positive semi-definite matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ and associated unit eigenvectors $e_1, e_2, ..., e_n$. Then*

- $$\max_{x\neq 0} \frac{x^TBx}{x^Tx} = \lambda_1,$$

  where $x^* = e_1$.
- $$\min_{x\neq 0, x\perp e_1} \frac{x^TBx}{x^TAx} = \lambda_2,$$

  where $x^* = e_2$.

- *Moreover,*

$$\max_{x\neq 0, x\perp e_1,...,e_k} \frac{x^T Bx}{x^T x} = \lambda_{k+1},$$

where $x^* = e_{k+1}$.

Proof. (1) and (2) are results in Rayleigh quotients theorem [Theorem 5.8.4]. (3) Because $x \perp x_1, ..., x_k$; therefore, $x \in span(e_{k+1}, e_{k+2}, ..., e_n)$. Let

$$x = y_{k+1}e_{k+1} + y_{k+2}e_{k+2} + ... + y_n e_n,$$

we have

$$\frac{x^T Bx}{x^T x} = \frac{\sum_{i=k+1}^n \lambda_i y_i^2}{\sum_{i=k+1}^n y_i^2}.$$

Taking $y_{k+1} = 1, y_{k+2} = ... = y_n = 0$ will give the maximum value of the ratio. Then $x = e_{k+1}$. □

Corollary 5.12.4.1 (maximization of general Quadratic forms on the unit sphere). *Let* $B \in \mathbb{R}^{n\times n}$ *be a positive semi-definite matrix. Let* $A \in \mathbb{R}^{n\times n}$ *be a positive definite matrix with decompostion* $A = \Sigma^{1/2}\Sigma^{1/2}$*, where* $\Sigma^{1/2}$ *is a positive semi-definite symmetric matrix and the matrix square root of A [Theorem 5.12.3].*

Let $\Sigma^{-1/2}B\Sigma^{1/2}$ *have eigenvalues* $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ *and associated unit eigenvectors* $e_1, e_2, ..., e_n$*. Then*

- $$\max_{x\neq 0} \frac{x^T Bx}{x^T Ax} = \lambda_1,$$

  where $x^* = \Sigma^{-1/2}u_1, u_1 = e_1$.
- $$\min_{x\neq 0, x\perp e_1} \frac{x^T Bx}{x^T Ax} = \lambda_2,$$

  where $x^* = \Sigma^{-1/2}u_2, u_2 = e_2$.
- *Moreover,*

$$\max_{x\neq 0, x\perp e_1,...,e_k} \frac{x^T Bx}{x^T Ax} = \lambda_{k+1},$$

where $x^* = \Sigma^{-1/2}u_{k+1}, u_{k+1} = e_{k+1}$.

Proof. Note that

$$\begin{aligned}\frac{x^TBx}{x^TAx} &= \frac{x^TBx}{x^T\Sigma^{1/2}\Sigma^{1/2}x} \\ &= \frac{x^TAx}{x^T\Sigma^{1/2}\Sigma^{1/2}x} \\ &= \frac{u^T\Sigma^{-1/2}B\Sigma^{-1/2}u}{u^Tu} \quad (use\ u = \Sigma^{1/2}x)\end{aligned}$$

Then we use Theorem 5.12.4. □

Example 5.12.3. Consider the quadratic form $Q(x) = 9x_1^2 + 5x_2^2 + 4x_3^2$. Under the constraint $x_1^2 + x_2^2 + x_3^2 = 1$, the maximum is achieved at $x = e_1$, where $e_1 = (1, 0, 0)$ is the unit eigenvector associated of the largest eigenvalue.

5.12.2.5 *Gramian matrix*

Definition 5.12.3 (Gramian matrix). *Let B be a real-valued matrix. The matrix $A = B^TB$ is called a* ***Gramian matrix****.*

Lemma 5.12.7 (properties of Gramian matrix). *Consider a Gramian matrix denoted by X^TX. We have*

- *X^TX is symmetric and $(X^TX)^T = X^TX$.*
- *X^TX is of full rank if and only if X is of full column rank.*
- $$rank(X^TX) = rank(X)$$
- *X^TX is non-negative definite.*
- *X^TX is positive definite if and only if X is of full column rank.*
- $$X^TX = \mathbf{0} \implies X = 0.$$

Proof. (1) straight forward. (2)(3) Lemma 5.4.3. (4) for any vector a, we have

$$a^TX^TXa = (Xa)^T(Xa) \geq 0.$$

(5) If X is of full column rank, then for any $a \neq 0$, $Xa \neq 0$. Therefore

$$a^TX^TXa = (Xa)^T(Xa) > 0.$$

(6) Let $A = X^T X$. If $A_{ii} = 0$, then

$$A_{ii} = \sum_k X_{ki}^2 = 0 \implies X_{ki} = 0 \forall k.$$

For all i, we have $X = \mathbf{0}$. □

5.12.3 Completing the square

Theorem 5.12.5 (completing the square). *[4, p. 407] Let $A \in \mathbb{R}^{n\times n}$ be a symmetric positive definite matrix,let $x, b \in \mathbb{R}^n$, then*

$$x^T Ax - 2b^T x + c = (x - A^{-1}b)^T A(x - A^{-1}b) + c - b^T A^{-1}b$$

Proof. Direct verification. □

Theorem 5.12.6 (completing the square in general cases). *Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix,let $x, b \in \mathbb{R}^n$. For completing the squares for $x^T Ax + bx$, we have the following situations:*

- *If A is non-singular, then the completing square exists and is given as*
- *If A is singular and $b \in \mathcal{R}(A)$, then the completing square exists.*
- *If A is singular and $b \notin \mathcal{R}(A)$, then the completing square does not exists.*

Example 5.12.4 (non-existence of completing squares). Consider the case

$$x_1^2 + x_1 + 2x_2 + 3.$$

We cannot complete the squares.

5.13 Matrix norm and spectral estimation

5.13.1 Basics

Definition 5.13.1 (spectral radius). *[12, p. 8] Let $A \in \mathbb{C}^{n\times n}$ with eigenvalues $\lambda_i, i = 1, 2, ..., n$. Then the spectral radius of the matrix A is defined as*

$$\rho(A) = \max_{1\leq i\leq n} |\lambda_i|$$

Definition 5.13.2 (matrix norm). *Let $A \in \mathbb{C}^{n\times n}$, then the matrix norm induced by the vector norm is*

$$\|A\| = \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|}$$

Theorem 5.13.1 (properties of matrix norm). *[12, p. 9] If A and B are two $n \times n$ complex matrices, then we have*

- $\|A\| > 0$ *unless* $A = 0$.
- $\|aA\| = |a|\|A\|$ *for all* $a \in \mathbb{C}$.
- $\|A + B\| \leq \|A\| + \|B\|$
- $\|A \cdot B\| \leq \|A\|\|B\|$
- $\left\|A^k\right\| \leq \|A\|^k$
- $\|Ax\| \leq \|A\|\|x\|$ *and there exists a nonzero vector such that the equality holds.*

Proof. Straight forward. For (6), because the set $\|x\| = 1$ is compact, therefore $\|A\| = \sup_{\|x\|=1}\|Ax\|$ can be achieved. □

Theorem 5.13.2 (relation between matrix norm and spectral radius). *Let $A \in \mathbb{C}^{n\times n}$*

- $\|A\| \geq \rho(A)$
- $\|A\|_2 = \rho(A^H A)^{0.5}$
- *If A is Hermitian, then* $\|A\|_2 = \rho(A)$

Proof. (1) $\|A\| = \sup_{\|x\|=1}\|Ax\| \geq \|Au\| = |\lambda|$ where u is a unit vector associated with an eigenvalue of A. (2) $\|A\|_2^2 = \sup_{\|x\|=1}\|Ax\|^2 = \lambda_{max}(A^H A)$ where the equality will hold when x is the unit eigenvector of $A^H A$ from Theorem 5.8.4.

(3) use (2), $A^H = A$, and the eigenvalues of A^2 is the square of eigenvalues of A. □

Example 5.13.1. Consider the example

$$A = \begin{bmatrix} 1 & 100 \\ 0 & 1 \end{bmatrix}.$$

$\rho(A) = 1, \|A\| = \rho(A^T A)^{0.5} \approx 100$

Theorem 5.13.3 (Existence of a matrix norm that is arbitrarily close to the spectral radius). *[9, p. 347] Let $A \in \mathbb{C}^{n\times n}$ and $\epsilon > 0$ be given. Then there exists a matrix norm $\|\cdot\|$ such that*

$$\rho(A) \leq \|A\| \leq \rho(A) + \epsilon$$

Remark 5.13.1. The norm can not only be common $L^p, 1 \leq p \leq \infty$ norm, it can also be weighted norm.

Theorem 5.13.4 (convergence). *Let $A \in \mathbb{C}^{n\times n}$. Then*

$$\lim_{n\to\infty} A^n = 0$$

if and only if $\rho(A) < 1$.

Proof. Use Jordan block decomposition. Also see Lemma 5.10.11. □

5.13.2 Singularity from matrix norm and spectral radius

Theorem 5.13.5 (singularity from spectral radius). *Let G be a square matrix such that $\rho(G) < 1$. Then $I - G$ is nonsingular.*

Proof. The eigenvalue of G satisfying the polynomial of $det(\lambda I - G) = 0$, and the eigenvalue of $I - G$ satisfying $det((-\lambda' + 1)I - G) = 0$. Therefore, we have $\lambda' = 1 - \lambda$. Since $|\lambda| < 1$, we must have $|\lambda'| > 0$. Therefore, $I - G$ is nonsingular. □

Remark 5.13.2 (interpretation). We have expansion of $(I - G)^{-1} = I + G + G^2 + ...$ when $G^k \to 0$ *as* $k \to \infty$. For $G^k \to 0$ *as* $k \to \infty$, the condition is $\rho(G) < 1$ [Theorem 5.13.4].

Corollary 5.13.5.1. *Let G be a square matrix such that $\|G\| < 1$. Then $I - G$ is nonsingular.*

Proof. Use $\rho(G) \leq \|G\|$ [Theorem 5.13.2]. □

5.13.3 Gerschgorin theorem

Theorem 5.13.6 (Gerschgorin theorem). *[1, p. 498][12, p. 16][13, p. 120] The eigenvalues of $A \in \mathbb{C}^{n\times n}$ are contained in the union of the n Gerschgorin circles defined by*

$$|z - a_i i| \leq r_i, r_i = \sum_{j=1, j\neq i}^{n} \left|a_{ij}\right|, \text{for } i = 1, 2, ..., n$$

Moreover, since A and A^T have the same eigenvalues, then the eigenvalues of $A \in \mathbb{C}^{n\times n}$ are contained in the union of the n Gerschgorin circles defined by

$$\left|z - a_{jj}\right| \leq r_j, r_j = \sum_{i=1, i\neq j}^{n} \left|a_{ij}\right|, \text{for } j = 1, 2, ..., n$$

Proof. Let x be an eigenvector such that $\|x\|_\infty = 1$. Assume the ith component x_i satisfying $|x_i| = 1$. Then $\lambda x = Ax$ and for the ith row we have $\lambda x_i = \sum_{j=1}^n a_{ij} x_j$ Finally, we have $|\lambda - a_{ii}||x_i| \leq \sum_{j=1, j\neq i}^n \left|a_{ij}\right|$. Therefore, λ is lying within some circle; in otherwise, all λ are lying within the union of all circles. □

Corollary 5.13.6.1 (diagonally dominant matrix property).

- *Any strictly diagonally dominant matrix $A(a_{ii} > \sum_{i=1}^n \left|a_{ij}\right|)$ is nonsingular.*
- *Any **symmetric** and strictly diagonally dominant matrix will be positive definite(and nonsingular).*

Proof. Its eigenvalues are strictly bounded away from and greater 0. □

Corollary 5.13.6.2 (spectral properties of stochastic matrix). *For any stochastic matrix(matrices where row sum is 1), its eigenvalues λ have $|\lambda| \leq 1$.*

Proof.

$$-1 \leq a_{ii} - r_i \leq \lambda \leq a_{ii} + r_i \leq 1$$

where $r_i = \sum_{j\neq i} \left|a_{ij}\right|$ □

5.13.4 Irreducible matrix and stronger results

Definition 5.13.3 (irreducible matrix). *[12, p. 18] For $n \geq 2$, an $n \times n$ complex matrix A is* ***reducible*** *if there exists an $n \times n$ permutation matrix P such that*

$$PAP^T = \begin{bmatrix} A_{1,1} & A_{1,2} \\ 0 & A_{2,2} \end{bmatrix}$$

where $A_{i,j}$ are block matrices. If no such permutation matrix exists, then the matrix is called ***irreducible****.*

Remark 5.13.3 (interpretation). If we view A as the transition matrix of a Markov chain, then A is reducible if there exists absorbing states(once trapped, cannot get out).

Theorem 5.13.7 (characterizing irreducibility using directed graph). *An $n \times n$ complex matrix A is irreducible if and only if its directed graph G is strongly connected; that is, for any other two ordered pair of two nodes i, j, there exists a directed path connecting them.*

Theorem 5.13.8 (Gerschgorin Taussky theorem). *[12, p. 20] Let A be an irreducible $n \times n$ complex matrix. If an eigenvalue λ is the on the boundary of the union of all the circles $|z - a_{ii}| \leq r_i$, then for all the n circles, $|\lambda - a_{ii}| = r_i, \forall i$.*

Proof. See reference. □

Remark 5.13.4. If an eigenvalue A is on the boundary of the circle/interval, and if A is irreducible, then the eigenvalue is on the boundary of all the intervals.

Corollary 5.13.8.1. *[14, p. 197] A matrix A is positive definite if the following* ***all*** *holds:*

- $a_{ii} \geq \sum_{j=1,j\neq i}^{n} \left|a_{ij}\right|, \forall i$
- $0 < a_{ii}, \forall i$
- *There is at least one row where $a_{ii} > \sum_{j=1,j\neq i}^{n} \left|a_{ij}\right|$*
- *A is irreducible.*

Proof. (1)(2) make sure that all eigenvalues are at least non-negative. (4) makes sure that all eigenvalues must be bounded away from 0. □

5.14 Pseudoinverse of matrix

5.14.1 Pseudoinverse for full rank system

Definition 5.14.1 (pseudoinverse for full rank system). *Let $A \in \mathbb{R}^{m\times n}$.*

- *If A has full column rank, then we define its pseudoinverse as*

$$A^+ = (A^TA)^{-1}A^T$$

such that $A^+ \in \mathbb{R}^{n\times m}, A^+A = I_n$.
- *If A has full row rank, then we define its pseudoinverse as*

$$A^+ = A^T(AA^T)^{-1}$$

such that $A^+ \in \mathbb{R}^{n\times m}, AA^+ = I_m$.

Lemma 5.14.1 (basic properties of pseudoinverse). *Let $A \in \mathbb{R}^{m\times n}$ with either $rank(A) = m$ or $rank(A) = n$. It follows that*

- *If $m = n$, then $A^+ = A^{-1}$.*
- *If A has full column rank, then A^+ has full row rank; If A has full row rank, then A^+ has full column rank;*
- *$(A^+)^+ = A$.*
- *Let A has full column rank such that A^T has full row rank, then*

$$(A^T)^+ = (A^+)^T.$$

- *For matrix A with either full column rank or full row rank, we have*

$$A^+AA^+ = A^+, AA^+A = A.$$

- *A^+A and AA^+ are symmetric.*

Proof. (1) If A has full column rank, then

$$(A^TA)^{-1}A^T = A^{-1}A^{-T}A^T = A^{-1}.$$

If A has full row rank, then

$$A^T(AA^T)^{-1} = A^TA^{-T}A^{-1} = A^{-1}.$$

(2) Let A has full column rank, then

$$rank(A^+) = rank((A^TA)^{-1}A^T) = rank((A^TA)^{-1}) = rank(A^TA) = n,$$

where we use results in ranks of matrix products [Lemma 5.4.1].

We can similarly prove the other case. (3) Let A have full column rank, then $A^+ = (A^TA)^{-1}A^T$ has full row rank. Then

$$\begin{aligned}(A^+)^+ &= [(A^TA)^{-1}A^T]^T((A^TA)^{-1}A^T[(A^TA)^{-1}A^T)]^T)^{-1} \\ &= A(A^TA)^{-T}(A^TA)^T \\ &= A\end{aligned}$$

We can similarly prove the other case. (4) straight forward. (5) Let $A^+ = A^T(AA^T)^{-1}$, we have

$$(A^+A)^T = (A^T(AA^T)^{-1}A)^T = A^T(AA^T)^{-1}A = A^+A.$$

We can similarly prove the other case. □

Lemma 5.14.2 (projector properties from pseudoinverse).

- *Let A have full column rank, then $P = AA^+ = A(A^TA)^{-1}A^T$ has the following properties*
  - *P is an orthogonal projector such that $P^T = P, PP = P$.*
  - *P is the orthogonal projector into $\mathcal{R}(A)$; or equivalently,*

    $$PA = A, P^TN_{A^T} = 0,$$

    where N_{A^T} is the basis matrix of $\mathcal{N}(A^T)$.
- *Let A have full row rank, then $Q = A^+A = A^T(AA^T)^{-1}A$ has the following properties*
  - *Q is an orthogonal projector such that $Q^T = Q, QQ = Q$.*
  - *Q is the orthogonal projector into $\mathcal{R}(A^T)$; or equivalently,*

    $$QA^T = A^T, Q^TN_A = 0,$$

    where N_A is the basis matrix of $\mathcal{N}(A)$.

Proof. (1) From Lemma 5.14.1, P is symmetric and

$$PP = AA^+AA^+ = (AA^+A)A^+ = AA^+ = P.$$

(2) $PA = A(A^TA)^{-1}A^TA = A$. Let $y \in \mathcal{N}(A^T)$ such that $A^Ty = 0$. Then

$$P^Ty = Py = A(A^TA)^{-1}A^Ty = 0.$$

(3)(4) Similar to (1)(2). □

Lemma 5.14.3 (pseudoinverse for special matrices).

- *Let A have full column rank and columns are orthonormal $A^T A = I$. Then*
$$A^+ = A^T.$$
- *Let A have full row rank and rows are orthonormal $AA^T = I$. Then*
$$A^+ = A^T.$$
- *Let diagonal matrix $D \in \mathbb{R}^{m\times n}, m \geq n$ with nonzero diagonal elements $d_1, d_2, ..., d_n$, then $D^+ \in \mathbb{R}^{n\times m}$ is diagonal with diagonal elements $1/d_1, 1/d_2, ..., 1/d_n$.*
- *Let diagonal matrix $D \in \mathbb{R}^{m\times n}, m \leq n$ with nonzero diagonal elements $d_1, d_2, ..., d_m$, then $D^+ \in \mathbb{R}^{n\times m}$ is diagonal with diagonal elements $1/d_1, 1/d_2, ..., 1/d_m$.*

Proof. (1)$A^+ = (A^T A)^{-1} A^T = A^T$. (2)(3)(4) straight forward.

□

Theorem 5.14.1 (SVD and pseudoinverse). *Let $A \in \mathbb{R}^{m\times n}$ (full column rank and full row rank)have the SVD [Theorem 5.9.1] given by $A = U\Lambda V^T$, then*
$$A^+ = V\Lambda^+ U^T.$$

Proof. Note that
$$\begin{aligned}
A^+ &= (A^T A)^{-1} A^T = (V\Lambda^2 V^T)^{-1} V\Lambda^T U \\
&= (A^T A)^{-1} A^T = (V\Lambda\Lambda^T V^T)^{-1} V\Lambda^T U^T \\
&= (V(\Lambda\Lambda^T)^{-1} V^T) V\Lambda^T U^T \\
&= V\lambda^+ U^T.
\end{aligned}$$
where we use the that $V\Lambda\Lambda^T V^T$ can be viewed as an eigen-decomposition and its inverse is given by $(V(\Lambda\Lambda^T)^{-1}V^T)$ □

5.14.2 Pseudoinverse for general matrix

Definition 5.14.2 (pseudoinverse for general). *Let $A \in \mathbb{R}^{m\times n}$ and its SVD given by $A = U\Lambda V^T$, then we define the pseudoinverse of A by*
$$A^+ = V\Lambda^+ U^T.$$

where we define Λ^+ as the transpose of Λ and the diagonal elements in Λ^+ is the inverse of the diagonal elements in Λ such that $\Lambda^+\Lambda = I_r \otimes 0_{n-r}, \Lambda\Lambda^+ = I_m \otimes 0_{n-r}$, where

$$I_r \otimes 0_{n-r} =\triangleq \begin{bmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & \ddots & & & & \\ & & & 1 & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{bmatrix}.$$

where there are r elements of 1 in the diagonal.

Note 5.14.1 (existence and uniqueness). Because a unique SVD always exists for any matrix, a unique pseudoinverse always exists for any matrix.

Lemma 5.14.4 (basic properties of pseudoinverse of general matrix). *Let $A \in \mathbb{R}^{m\times n}$ with rank $r \leq \min(m,n)$. Let its SVD be $A = U\Lambda V^T$. It follows that*

- *A^+ has rank r.*
- *$(A^+)^+ = A$.*
- $$(A^T)^+ = (A^+)^T.$$
- $$A^+AA^+ = A^+, AA^+A = A.$$
- *A^+A and AA^+ are symmetric.*

Proof. (1) Note that $V\Lambda^+U^T$ is still SVD form, and it has r non-zero elements in Λ^+. (2)

$$(A^+)^+ = (V\Lambda^+U^T)^+ = U(\Lambda^+)^+V^T = U\Lambda V^T = A.$$

(3)

$$(A^T)^+ = (V\Lambda^TU^T)^+ = (V\Lambda^TU^T)^+ = U(\Lambda^T)^+V^T$$

and

$$(A^+)^T == (V\Lambda^+U^T)^T = U(\Lambda^+)^TV^T.$$

Further note that $(\Lambda^+)^T = (\Lambda^T)^+$. (4)

$$AA^+A = U\Lambda V^T V\Lambda^+ U^T U\Lambda V^T = U\Lambda(I_r \otimes 0_{m-r})V^T = U\Lambda V^T = A.$$

similarly for the other. (5) Note that $A^+A = V\Lambda^+U^TU\Lambda V^T = V(I_r \otimes 0_{n-r})V^T$, a symmetric matrix. $\square$

Lemma 5.14.5 (projector properties from pseudoinverse). *Let $A \in \mathbb{R}^{m\times n}$ and its SVD given by $A = U\Lambda V^T$*

Then $P = AA^+ = U(I_r \otimes 0_{m-r})U^T = U_rU_r^T$ has the following properties:

- *P is an orthogonal projector such that $P^T = P, PP = P$.*
- *P is the orthogonal projector into $\mathcal{R}(A) = \mathcal{R}(U_r)$; or equivalently,*

$$PA = A, PU = U_r.$$

Proof. (1) Note that $AA^+ = U\Lambda V^TV\Lambda^+U^T = U(I_r \otimes 0_{m-r})U^T$, a symmetric matrix. Also from Lemma 5.14.4, P is symmetric and

$$PP = AA^+AA^+ = (AA^+A)A^+ = AA^+ = P.$$

(2)

$$PA = U(I_r \otimes 0_{m-r})U^TU\Lambda V^T = U(I_r \otimes 0_{m-r})\Lambda V^T = U\Lambda V^T = A$$
$$PU = U(I_r \otimes 0_{m-r})U^TU = U_r$$

$\square$

5.14.3 Application in linear systems

Lemma 5.14.6 (solution for full rank system). *Let $A \in \mathbb{R}^{m\times n}$ have either full column rank or full row rank. If the linear system $Ax = b$ has solution, then the solution is given by*

$$x = A^+b + (I_n - A^+A)z, z \in \mathbb{R}^n,$$

where $I_n - A^+A$ being the $\mathcal{N}(A)$ basis matrix. Among all solutions, the minimum norm/length solution is A^+b.

If $Ax = b$ does not have a solution, then

$$x = A^+b + (I_n - A^+A)z, z \in \mathbb{R}^n.$$

is the solution set of minimum error, with A^+b being the minimum nomr/length solution.

Proof. See SVD approach to linear system Lemma 5.1.9 and the relationship between SVD and pseudoinverse [Lemma 5.14.2]. Note that when A has full column rank $A^+A = I_n$. To show $I_n - A^+A$ is the null space basis matrix, we have

$$A(I_n - A^+A) = A - AA^+A = A - A = 0.$$

To show A^+b is of the minimum length, we have

$$\begin{aligned}
&\left\|A^+b + (I_n - A^+A)z\right\|^2 \\
=&\left\|A^+b\right\|^2 + \left\|(I_n - A^+A)z\right\|^2 + 2z(I_n - A^+A)^T(A^+b) \\
=&\left\|A^+b\right\|^2 + \left\|(I_n - A^+A)z\right\|^2 + 2z(A^+ - A^+AA^+)b \\
=&\left\|A^+b\right\|^2 + \left\|(I_n - A^+A)z\right\|^2 + 2z(A^+ - A^+)b \\
=&\left\|A^+b\right\|^2 + \left\|(I_n - A^+A)z\right\|^2 \geq \left\|A^+b\right\|^2
\end{aligned}$$

where we use the basic property $A^+AA^+ = A^+$. □

Remark 5.14.1 (interpretation)**.**

- If A has full column rank, then $A^+ = (A^TA)^{-1}A^T$, AA^+b is the orthogonal projection [Lemma 5.14.2] of b into $\mathcal{R}(A)$. Also, $A^+A = I_n$ implies the null space is 0 dimensional.
- If A is full row rank, then $A^+ = A^T(AA^T)^{-1}$, $AA^+b = I_mb = b$ is the solution. and also the orthogonal projection [Lemma 5.14.2] of b into $\mathcal{R}(A)$. Also, $A^+A = I_n$ implies the null space is 0 dimensional.

Theorem 5.14.2 (solution for general linear system). *Let $A \in \mathbb{R}^{m\times n}$ with SVD $A = U\Lambda V^T$ and rank$(A) = r$. Let $A^+ = V\Lambda^+U^T$ be its pseudoinverse. If the linear system $Ax = b$ has solution, then the solution is given by*

$$x = A^+b + (I_n - A^+A)z, z \in \mathbb{R}^n.$$

where $I_n - A^+A$ being the $\mathcal{N}(A)$ basis matrix. Among all solutions, the minimum norm/length solution is A^+b.

If $Ax = b$ does not have a solution, then

$$x = A^+b + (I_n - A^+A)z, z \in \mathbb{R}^n.$$

is the solution set of minimum error, with A^+b being the minimum norm/length solution.

Proof. See SVD approach to linear system Lemma 5.1.9 and the relationship between SVD and pseudoinverse [Lemma 5.14.5] and above proof. Note that AA^+ is orthogonal projector into $\mathcal{R}(A)$. To show $I_n - A^+A$ is the null space basis matrix, we have

$$A(I_n - A^+A) = A - AA^+A = A - A = 0.$$

□

5.15 Multilinear forms

5.15.1 Bilinear forms

Definition 5.15.1 (bilinear form). *Let V be a vector space over the filed $\mathbb{F}$. The map*

$$\phi : V \times V \to \mathbb{F}$$

is called ***bilinear form*** *on V if for any $u, v, w \in V$ and any scalar $\lambda \in \mathbb{F}$ we have*

- $$\phi(u+v, w) = \phi(u, w) + \phi(v, w), \phi(\lambda v, w) = \lambda \phi(v, w).$$
- $$\phi(u, v+w) = \phi(u, w) + \phi(u, v), \phi(v, \lambda w) = \lambda \phi(v, w).$$

Lemma 5.15.1 (representation of bilinear form). *For any bilinear form ϕ defined on $\mathbb{R}^n$, there exists a matrix $A \in \mathbb{R}^{n\times n}$ such that ϕ can be represented by*

$$\phi(x, y) = x^T A y, \forall x, y \in \mathbb{R}^n.$$

Proof. Let $A_{ij} = \phi(e_i, e_j)$. Then any $x = \sum_{i=1}^n x_i e_i, y = \sum_{j=1}^n y_i e_i$, we have

$$\phi(x, y) = \sum_{i=1}^n \sum_{j=1}^n x_i y_j \phi(e_i, e_j) = \sum_{i=1}^n \sum_{j=1}^n x_i y_i A_{ij} = x^T A y.$$

□

Definition 5.15.2 (symmetric, skew symmetric, and alternating bilinear forms). *Let U be a F-vector space.*

- *A bilinear form ϕ is called* ***symmetric*** *if for any $u_1, u_2 \in U$*
$$\phi(u_1, u_2) = \phi(u_2, u_1).$$
- *A bilinear form ϕ is called* ***skew-symmetric*** *if for any $u_1, u_2 \in U$,*
$$\phi(u_1, u_2) = -\phi(u_2, u_1).$$

- *A bilinear form ϕ is called **alternating** if for any $u \in U$ we have*

$$\phi(u,u) = 0.$$

5.15.2 Multilinear forms

Definition 5.15.3 (k-linear form). *Let V be a vector space over the filed $\mathbb{F}$. The map*

$$\phi : \underbrace{V \times V \cdots V}_{k} \to \mathbb{F}$$

*is called k-**linear form** on V if for any $u_i, v_i \in V, i = 1,2,...,k$ and any scalar $\lambda \in \mathbb{F}$ we have*

- $$\phi(u_1,...,u_i + v_i,...,u_k) = \phi(u_1,...,u_i,...,u_k) + \phi(u_1,...,v_i,...,u_k).$$
- $$\phi(u_1,...,\lambda u_i,... + u_k) = \lambda\phi(u_1,...,u_i,...,u_k).$$

Example 5.15.1. Let $A \in \mathbb{R}^{n\times n}$. Define $\phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ such that

$$\phi(x,y) = x^T Ay, \forall x,y \in \mathbb{R}^n.$$

We can see that

- $$\phi(x+z,y) = (x+z)^T Ay = x^T Ay + z^T Ay = \phi(x,y) + \phi(z,y)$$
  .
- $$\phi(\lambda x,y) = (\lambda x)^T Ay = \lambda x^T Ay = \lambda\phi(x,y).$$

Therefore, ϕ is a bi-linear form on $\mathbb{R}^n$.

Definition 5.15.4 (symmetric, skew symmetric, and alternating). *Let U be a F-vector space.*

- *A k-linear form ϕ is called **symmetric** if for any $u_1,u_2,...,u_k \in U$ and any permutation $\sigma \in S_k$ we have*
$$\phi(u_{\sigma(1)}, u_{\sigma(2)},...,u_{\sigma(k)}) = \phi(u_1,u_2,...,u_k).$$

- *A k-linear form ϕ is called* ***skew-symmetric*** *if for any $u_1, u_2, ..., u_k \in U$ and any permutation $\sigma \in S_k$ we have*

$$\phi(u_{\sigma(1)}, u_{\sigma(2)}, ..., u_{\sigma(k)}) = sign(\sigma)\phi(u_1, u_2, ..., u_k).$$

or equivalently(swap will change sign),

$$\phi(u_1, ..., u_i, ..., u_j, ..., u_k) = -\phi(u_1, ..., u_j, ..., u_i, ..., u_k).$$

- *A k-linear form ϕ is called* ***alternating*** *if for any $u_1, u_2, ..., u_k \in U$ we have*

$$\phi(u_1, u_2, ..., u_k) = 0,$$

whenever $u_i = u_j, i \neq j$.

Remark 5.15.1 (simplified condition for checking skew-symmetric). Note that the sufficient condition for checking a k-linear form is to examine all permutations σ. A simplified condition is to only check whether a simple swap will change sign.

To show that these two conditions are equivalently, we have

- $$\begin{aligned} &\phi(u_{\sigma(1)}, u_{\sigma(2)}, ..., u_{\sigma(k)}) = sign(\sigma)\phi(u_1, u_2, ..., u_k) \\ \implies &\phi(u_1, ..., u_i, ..., u_j, ..., u_k) = -\phi(u_1, ..., u_j, ..., u_i, ..., u_k) \end{aligned}$$

  since a simple swap has sign -1.
- We note that any permutation can be decomposed as compositions of simple swap. And the sign of the permutation equals the number of simple swaps. Define

  $$\sigma \circ \phi(u_1, u_2, ..., u_k) = \phi(u_{\sigma(1)}, u_{\sigma(2)}, ..., u_{\sigma(k)}),$$

  and suppose we have decomposition $\sigma = \sigma_1 \circ \sigma_2 \cdots \sigma_m$, where σ_is are simple swaps. Then

  $$\begin{aligned} \phi(u_{\sigma(1)}, ..., u_{\sigma(k)}) &= \sigma_1 \circ \sigma_2 \cdots \circ \sigma_m \phi(u_1, ..., u_k) \\ &= (-1)^m \phi(u_1, ..., u_k) \\ &= sign(\sigma)\phi(u_1, ..., u_k) \end{aligned}$$

Lemma 5.15.2 (skew symmetric and alternating are equivalent). *Let U be a F-vector space. Let ϕ be a k-linear form. Then ϕ is alternating if and only if ϕ is skew-symmetric.*

Proof. (1)(alternating implies skew-symmetric) For all $x, y \in U$, we have

$$\begin{aligned} 0 = \phi(..., x+y, ..., x+y, ...) &= \phi(..., x, ..., x, ...) + \phi(..., y, ..., y, ...) \\ &\quad + \phi(..., y, ..., x, ...) + \phi(..., x, ..., y, ...) \\ &= 0 + 0 + \phi(..., y, ..., x, ...) + \phi(..., x, ..., y, ...) \\ &= \phi(..., y, ..., x, ...) + \phi(..., x, ..., y, ...) \\ \implies \phi(..., y, ..., x, ...) &= -\phi(..., x, ..., y, ...) \end{aligned}$$

(2)(skew-symmetric implies alternating) For any $x \in U$, the skew-symmetry properties implies that

$$\phi(..., x..., x, ...) = -\phi(..., x..., x, ...);$$

rearrange and we will get

$$2\phi(..., x..., x, ...) = 0.$$

□

Example 5.15.2. Consider the following bilinear forms on $\mathbb{R}^4$. Let $x, y \in \mathbb{R}^4$

- $f(x, y) = x_1y_2 - x_2y_1 + x_1y_1$ is not alternating since

$$f(x, x) = x_1x_2 - x_2x_2 + x_1^2 = x_1^2 \geq 0.$$

  That is $f(x, x)$ does not equal 0 for all $x \in \mathbb{R}^4$.
- $g(x, y) = x_1y_3 - x_3y1$ is alternating since

$$g(x, x) = x_1x_3 - x_3x_1 = 0 \forall x \in \mathbb{R}^4.$$

5.16 Determinant

5.16.1 Basic properties

Definition 5.16.1 (determinant). *[4, p. 279] The **determinant** of an $n \times n$ matrix $A = a_{ij}$ is defined by*

$$det(A) = \sum_{\sigma \in S_n} sign(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} = \sum_{\sigma \in S_n} sign(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n},$$

where we are summing up all $n!$ permutation in the symmetric group S_n.

Example 5.16.1. Consider a 2×2 matrices. Let

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

then

$$|A| = \sigma(1,2)a_{11}a_{22} + \sigma(2,1)a_{12}a_{21} = a_{11}a_{22} - a_{12}a_{21}.$$

Example 5.16.2. Consider a 3×3 matrices. Let

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

then

$$\begin{aligned} |A| &= \sigma(1,2,3)a_{11}a_{22}a_{33} + \sigma(1,3,2)a_{11}a_{23}a_{32} + \sigma(2,1,3)a_{12}a_{21}a_{33} \\ &+ \sigma(2,3,1)a_{12}a_{21}a_{33} + \sigma(3,1,2)a_{13}a_{21}a_{32} + \sigma(3,2,1)a_{13}a_{22}a_{31} \\ &= a_{11}a_{22}a_{33} + -a_{11}a_{23}a_{32} + -a_{12}a_{21}a_{33} \\ &+ a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} + -a_{13}a_{22}a_{31} \end{aligned}$$

Theorem 5.16.1 (the equivalence of $detA$ and $detA^T$.). *The **determinant** of an $n \times n$ matrix $A = a_{ij}$ is defined by* *where we are summing up all $n!$ permutation in the symmetric group S_n.*

- $$det(A) = \sum_{\sigma \in S_n} sign(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} = \sum_{\sigma \in S_n} sign(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n},$$
- $$det(A) = \sum_{\sigma \in S_n} sign(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} = \sum_{\sigma \in S_n} sign(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n},$$

Lemma 5.16.1 (determinant and multilinear forms).

- *For any matrix $A \in \mathbb{R}^{n\times n}$, $A = [a_1, a_2, ..., a_n]$,the determinant of A given by*
$$det(A) \triangleq det(a_1, a_2, ..., a_n)$$
is the n-linear form mapping from $\mathbb{R}^{n\times n}$ to $\mathbb{R}$.
- *$det(A)$ is both alternating and skew-symmetric; Specifically,*
  - *(skew-symmetric) For any $a_1, a_2, ..., a_k \in \mathbb{R}^n$ we have*
  $$det(u_{\sigma(1)}, u_{\sigma(2)}, ..., u_{\sigma(n)}) = sign(\sigma)\phi(u_1, u_2, ..., u_k).$$
  - *(alternating) For any $a_1, a_2, ..., a_k \in \mathbb{R}^n$ we have*
  $$det(u_1, u_2, ..., u_n) = 0,$$
  whenever $u_i = u_j, i \neq j$.

Proof. (1) We can show that (a)

$$\begin{aligned} & det(a_1, ..., a_i + b_i, ..., a_n) \\ & = \sum_{\sigma \in S_n} sign(\sigma) a_{1\sigma(1)} \cdots (a_{1\sigma(i)} + b_{1\sigma(i)}) a_{n\sigma(n)} \\ & = \sum_{\sigma \in S_n} sign(\sigma) a_{1\sigma(1)} \cdots (a_{1\sigma(i)}) a_{n\sigma(n)} + \sum_{\sigma \in S_n} sign(\sigma) a_{1\sigma(1)} \cdots (b_{1\sigma(i)}) a_{n\sigma(n)} \\ & = det(a_1, ..., a_i, ..., a_n) + det(a_1, ..., b_i, ..., a_n) \end{aligned}$$

(b)

$$\begin{aligned} & det(a_1, ..., \lambda a_i, ..., a_n) \\ & = \sum_{\sigma \in S_n} sign(\sigma) a_{1\sigma(1)} \cdots (\lambda a_{1\sigma(i)}) a_{n\sigma(n)} \\ & = \lambda \sum_{\sigma \in S_n} sign(\sigma) a_{1\sigma(1)} \cdots (a_{1\sigma(i)}) a_{n\sigma(n)} \\ & = \lambda det(a_1, ..., a_i, ..., a_n) \end{aligned}$$

(2) We only need to show skew-symmetric since skew-symmetric and alternating are equivalent [Lemma 5.15.2]. □

Theorem 5.16.2 (determinants for matrices after column(row) operation). *[4, p. 282] Let B be the matrix obtained from an $n \times n$ matrix $A = [a_1, a_2, ..., a_n]$ by applying one of the three elementary column(row) operations:*

- *(type I) interchange two columns(rows) of A, then*
$$detB = -detA.$$
- *(type II): multiply a column(row) of A by a scalar α, then*
$$detB = \alpha detA.$$
- *(type III): add α times a given column(row) of A to another column(row), then*
$$detB = detA.$$

Proof. (1)(2) Since *det* is skew-symmetric [Lemma 5.16.1],

$$det(B) = det(a_1, ..., a_j, ..., a_i, ...a_n) = -det(a_1, ..., a_i, ..., a_j, ...a_n) = det(A).$$

and

$$det(B) = det(a_1, ..., \alpha a_i ... a_n) = \alpha det(a_1, ..., a_i, ...a_n) = \alpha det(A).$$

(3)

$$det(B) = det(a_1, ..., a_j, ..., a_i, ...a_n) = -det(a_1, ..., a_i, ..., a_j, ...a_n) = det(A).$$

and

$$\begin{aligned} det(B) &= det(a_1, ..., a_i + \alpha a_j, , ...a_n) \\ &= det(a_1, ..., a_i, ..., a_n) + det(a_1, ..., \alpha a_j, ..., a_j, ...a_n) \\ &= det(A) + 0 \\ &= det(A) \end{aligned}$$

□

Lemma 5.16.2 (determinant of triangular matrix).

- *If $A \in \mathbb{R}^{n \times n}$ is an upper(lower) triangular matrix, then $det(A)$ is the product of the diagonal entries.*
- *For the identity matrix I_m, $det(I_m) = 1$.*

Proof. From the definition

$$det(A) = \sum_{\sigma \in S_n} sign(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)},$$

we know that the only nonzero terms in the summation is the permutation such that

$$\sigma(i) = i, i = 1, 2...., n.$$

Then

$$det(A) = a_{11} \cdots a_{nn}.$$

□

Lemma 5.16.3 (determinant and invertibility of a matrix).

- *Consider a matrix $A \in \mathbb{R}^{n \times n}$. A is invertible if and only if $det(A) \neq 0$.*
- *Consider a upper(or lower) triangular matrix $A \in \mathbb{R}^{n \times n}$. A is invertible if and only if **all diagonal entries** in A are nonzero.*

Proof. (1)A matrix A is invertible if and only if by performing elementary row operations we can reduce to an upper triangular matrix B whose diagonal entries are nonzero,i.e., $detB \neq 0$. (2)Note that for a triangular matrix, its determinant is the product of its diagonal entries [Lemma 5.16.2]. □

Lemma 5.16.4 (determinant of matrix product).

- *If A, B are $n \times n$ matrices, then*

$$det(AB) = det(A)det(B).$$

- *If A is invertible, then*

$$detA^{-1} = \frac{1}{detA}.$$

Proof. (1) Note that for the product $(AB)_{ij} = \sum_{k=1}^{n} A_{ik}B_{kj}$. The column j of AB is given by

$$\sum_{k=1}^{n} a_k B_{kj}.$$

Therefore,

$$\begin{aligned}det(AB) &= det(\sum_{i_i=1}^{n} a_{i_1} B_{i_1 1}, ..., \sum_{i_n=1}^{n} a_{i_n} B_{i_n n}) \\ &= \sum_{i_1}^{n} ... \sum_{i_n=1}^{n} B_{i_1 1} ... B_{i_n n} det(a_1, a_2, ..., a_n) \\ &= \sum_{i_1}^{n} ... \sum_{i_n=1}^{n} B_{i_1 1} ... B_{i_n n} det(a_{i_1}, a_{i_2}, ..., a_{i_2})\end{aligned}$$

Because of the alternating properties of determinant, the only non-zero terms in the above summation correspond to choices of pairwise distinct indices $i_1, ..., i_n$. For such a choice, the sequence $i_1, ..., i_n$ describes a permutation from S_n. We then have

$$\begin{aligned}det(AB) &= \sum_{i_1}^{n} ... \sum_{i_n=1}^{n} B_{i_1 1} ... B_{i_n n} det(a_{i_1}, a_{i_2}, ..., a_{i_2}) \\ &= \sum_{\sigma \in S_n} B_{\sigma(1)1} ... B_{\sigma(n)n} det(a_{\sigma(1)}, ..., a_{\sigma(n)}) \\ &= \sum_{\sigma \in S_n} B_{\sigma(1)1} ... B_{\sigma(n)n} sign(\sigma) det(a_1, ..., a_n) \\ &= det(B) det(A),\end{aligned}$$

where we use the skew-symmetry property of determinant [Lemma 5.16.1] such that

$$det(a_{\sigma(1)}, ..., a_{\sigma(n)}) = sign(\sigma) det(a_1, ..., a_n).$$

(2) If

$$det A A^{-1} = det A det A^{-1} = 1 \implies det A^{-1} = \frac{1}{det A}.$$

□

Lemma 5.16.5 (determinant of block matrix).

- *Let*

$$M = \begin{bmatrix} A & 0 \\ 0 & I_m \end{bmatrix}, A \in \mathbb{R}^{n \times n}, .$$

Then

$$det(M) = det(A)$$

- *Let*
$$M = \begin{bmatrix} A & 0 \\ C & D \end{bmatrix}.$$
Then
$$det(M) = det(A)det(D).$$
- *Let*
$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$
Then
$$det(M) = det(A)det(D - CA^{-1}B).$$

Proof. (1) From the definition of determinant, we have

$$det(M) = \sum_{\sigma \in S_{m+n}} sign(\sigma) M_{1\sigma(1)} \cdots M_{n+m\sigma(n+m)}.$$

The non-zero terms in the above summation correspond to the choices where $\sigma(i) = i, i = n+1, ..., n+m$. Then we can simplify

$$det(M) = \sum_{\sigma \in S_n} sign(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} = det(A).$$

(2) Note that

$$\begin{bmatrix} A & 0 \\ C & D \end{bmatrix} = \begin{bmatrix} A & 0 \\ C & I_m \end{bmatrix} \begin{bmatrix} I_n & 0 \\ 0 & D \end{bmatrix};$$

Then

$$det \begin{bmatrix} A & 0 \\ C & I_m \end{bmatrix} \begin{bmatrix} I_n & 0 \\ 0 & D \end{bmatrix} = det \begin{bmatrix} A & 0 \\ C & I_m \end{bmatrix} det \begin{bmatrix} I_n & 0 \\ 0 & D \end{bmatrix} = det(A)det(C)$$

(3) Note that

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} A & 0 \\ C & I_m \end{bmatrix} \begin{bmatrix} I_n & A^{-1}B \\ 0 & D - CA^{-1}B \end{bmatrix};$$

then use (2). □

5.16.2 Vandermonde matrix and determinant

Definition 5.16.2. *For any list of complex numbers $(x_1, x_2, ..., x_n)$, the associated following matrix*

$$V_n(x_1, x_2, ..., x_n) = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{pmatrix},$$

is called ***Vandermonde matrix***.

Lemma 5.16.6 (determinant of Vandermonde matrix). *Consider a Vandermonde matrix associated with n complex numbers $(x_1, x_2, ..., x_n)$. It follows that*

- $$det V_n(x_1, x_2, ..., x_n) = \prod_{1 \leq i < j \leq n} (x_j - x_i).$$

  note that there are $n(n-1)/2$ terms in the product.
- *If $x_1, x_2, ..., x_n$ are not pairwise distinct, then*

  $$det V_n(x_1, x_2, ..., x_n) = 0.$$

Example 5.16.3.

- $$det V_2(x_1, x_2) = det \begin{pmatrix} 1 & 1 \\ x_1 & x_2 \end{pmatrix} = (x_2 - x_1).$$
- $$det V_3(x_1, x_2, x_3) = det \begin{pmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ x_1^2 & x_2^2 & x_3^2 \end{pmatrix} = (x_2 - x_1)(x_3 - x_2)(x_3 - x_1).$$

Lemma 5.16.7 (application example: existence of polynomial passing points). *If* $(x_1, y_1), ..., (x_n, y_n)$ *are* ***distinct*** *complex number pairs. Then there exists a polynomial of degree* $\leq n-1$ *uniquely determined by the conditions*

$$P(x_1) = y_1, P(x_2) = y_2, ..., P(x_n) = y_n.$$

Proof. Consider a polynomial with degree less than $n-1$ given by

$$P(x) = a_0 + a_1 x + \cdots + a_{n-1} x^{n-1},$$

where the coefficients $a_0, a_1, ..., a_{n-1}$ are to be determined. The conditions

$$P(x_1) = y_1, P(x_2) = y_2, ..., P(x_n) = y_n,$$

gives the following linear systems

$$\begin{aligned} a_0 + a_1 x_1 + a_2 x_1^2 + ... + a_{n-1} x_1^{n-1} &= y_1 \\ a_0 + a_1 x_2 + a_2 x_2^2 + ... + a_{n-1} x_2^{n-1} &= y_2 \\ &\cdots\cdots \\ a_0 + a_1 x_n + a_2 x_n^2 + ... + a_{n-1} x_n^{n-1} &= y_n, \end{aligned}$$

which can be written as matrix form as

$$\underbrace{\begin{pmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ 1 & x_2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{pmatrix}}_{V_n(x_1, x_2, ..., x_n)^T} \cdot \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

Because $x_1, x_2, ..., x_n$ are distinct, then from Lemma 5.16.6, $detV_n^T = detV_n \neq 0$; that is we can uniquely solve for $a_0, a_1, ..., a_{n-1}$. $\square$

5.17 Numerical iteration analysis

5.17.1 Numerical linear equation solution

5.17.1.1 *Goals and general principles*

This section we introduce methods to solve linear system $Ax = b$, where A is nonsingular, using iterative method. The **general principle** is to decompose $A = M - N$ with M being nonsingular, then

$$Mx = b + Nx \Rightarrow x = M^{-1}(b + Nx).$$

When the matrix M satisfies center conditions on its spectrum and norm, iteration will converge to the solution.

Theorem 5.17.1 (general convergence condition). *[15, p. 614] The iteration $x = M^{-1}(b + Nx)$ converges to $x^* = A^{-1}b$ for all initial starting vector x^0 if $\rho(G) < 1$, where $G = M^{-1}N$*

Proof. Let T be the operator such that $x = M^{-1}(b + Nx)$. Then $Tx^* = x^*$ indicting that x^* is the fixed point. $Tx - Tx^* = G(x - x^*) \Rightarrow T^n x - T^n x^* \to 0$ as $n \to \infty$ since $\rho(G) < 1$ implies $G^n \to 0$. (We can also use Theorem 5.13.3 to show there is a matrix norm such that $\rho(G) < 1 \implies \|G\| < 1$.) □

In the following, we will go over two classical algorithms, Jacobin algorithm and Gauss Seidel algorithm, that employs such principle.

5.17.1.2 *Jacobi algorithm*

Consider the equations $Ax = b$. The **Jacobi algorithm** represents $A = D + L + U$, and the iteration is given by

$$x = D^{-1}(b - (L + U)x).$$

The convergence of this algorithm under certain condition is given by

Lemma 5.17.1 (sufficient condition for convergence). *[15, p. 615] The Jacobi algorithm will converge for any initial x^0 if A is strictly diagonally dominant.*

Proof. It can be showed that the row sum of the M matrix ($M = D^{-1}(L + U)$)is less than 1 and the diagonal entry of M is 0. Then we can use the Gerschgorin theorem [Theorem 5.13.6] to show the $\rho(M) < 1$. □

5.17.1.3 *Gauss Seidel algorithm*

Consider the equations $Ax = b$. The **Jacobi algorithm** represents $A = D + L + U$, and the iteration is given by

$$x = (D+L)^{-1}(b - Ux).$$

Lemma 5.17.2 (sufficient condition for convergence). *[13, p. 122] The Gauss Seidel algorithm will converge for any initial x^0 if A is strictly diagonally dominant.*

Proof. Let $M = (D+L)^{-1}U$. Let x be an eigenvector of M such that $\|x\|_\infty = 1$. Assume the ith component x_i satisfying $|x_i| = 1$. Then $Ux = \lambda(D+L)$ and for the ith row we have

$$\sum_{j<i} a_{ij}x_j = \lambda(a_{ii} + \sum_{j>i} a_{ij}x_j)$$

Further we have

$$|\lambda| = \left|\frac{\sum_{j<i} a_{ij}x_j}{(a_{ii} + \sum_{j>i} a_{ij}x_j)}\right| \leq \frac{\sum_{j<i}\left|a_{ij}\right|}{a_{ii} - \sum_{j>i}\left|a_{ij}\right|} < 1.$$

Therefore, $\rho(M) < 1$. □

5.17.2 Power method for eigen-decomposition

In this section, we introduce Power method, a widely used method to compute top eigenvector u_1 of a matrix A. The algorithm starts with an initial guess $u^0 \in \mathbb{R}^N$ that has nonzero projection a_1 on u_1 and then carries out the following iteration

$$u^{k+1} = \frac{Au^k}{\|Au^k\|}.$$

where k is the iteration number. As k is sufficiently large, u^k will converge to u_1.

The following theorem examines the condition for convergence and the convergence speed.

Theorem 5.17.2 (power method for top eigenvector). *[16, p. 115][15, p. 451] Let $A \in \mathbb{R}^{N\times N}$ be a real symmetric positive definite matrix with eigenvector $\{u_1, ..., u_N\}$ and* *eigenvalues $\{\lambda_1, ..., \lambda_N\}$ sorted in descending order. Assume $\lambda_1 > \lambda_2$ and let $u^0 \in \mathbb{R}^N$ be an arbitrary vector has nonzero projection a_1 on u_1. Consider the sequence of vectors*

$$u^{k+1} = \frac{Au^k}{\left\|Au^k\right\|}.$$

We have

- *u^k converges to $\frac{a_1}{|a_1|}u_1$ with rate $\frac{\lambda_2}{\lambda_1}$. That is, there exist a constant $C > 0$ such that for all $k \geq 0$,*

$$\left\|u^k - \frac{a_1}{|a_1|}u_1\right\| \leq C(\frac{\lambda_2}{\lambda_1})^k.$$

Proof. Note that The iterate u^k is a multiple of $A^k u^0$ with length 1.Let $u^0 = sum_{i=1}^N a_1 u_1$. Let A^k has eigendecomposition of $A^k = \sum_{i=1}^N \lambda_i^k u_i u_i^T$, then

$$\begin{aligned} u^k &= \frac{A^k u^0}{\left\|A^k u^0\right\|} \\ &= \frac{\sum_{i=1}^N \lambda_i^k a_i u_i}{\sqrt{\lambda_i^{2k} a_i^2}} \\ &\leq \frac{\lambda_1^k a_1}{\lambda_1^k |a_1|} + \frac{\sum_{i=2}^N \lambda_i^k a_i u_i}{\lambda_1^k |a_1|} \end{aligned}$$

Then

$$\left\|u^k - \frac{a_1}{|a_1|}u_1\right\| \leq C\frac{\lambda_2^k}{\lambda_1^k}.$$

□

Remark 5.17.1 (not a contraction mapping). It can be showed that $\frac{a_1}{|a_1|}u_1$ is a fixed point for the mapping

$$T(u) = \frac{Au}{\left\|Au^k\right\|}.$$

However, for any vector on the subspace spanned by u_1, the mapping will not shrink it; therefore, the mapping is not a contraction.

Corollary 5.17.2.1 (extension to real symmetric matrix). *Let $A \in \mathbb{R}^{N\times N}$ be a real symmetric matrix with eigenvector $\{u_1, ..., u_N\}$ and eigenvalues $\{\lambda_1, ..., \lambda_N\}$ sorted in descending order.Then the sequence of vectors generated by*

$$u^{k+1} = \frac{Au^k}{\|Au^k\|}$$

will converge to the eigenvector(up to scale) associated with eigenvalue with largest absolute value.

We can also extend the power method to the top d eigenvectors.

Methodology 5.17.1 (power method for top d eigenvectors, orthogonal iteration). *[16, p. 115][15, p. 454] Let $A \in \mathbb{R}^{N\times N}$ be a real symmetric positive definite matrix with eigenvector $\{u_1, ..., u_N\}$ and eigenvalues $\{\lambda_1, ..., \lambda_N\}$ sorted in descending order. Assume that $\lambda_d > \lambda_{d+1}$ and let $U^0 \in \mathbb{R}^{N\times d}$ be an arbitrary matrix whose column space is not orthogonal to the subspace $\{u_1, ..., u_d\}$ spanned by the top d eigenvectors of A. Consider the sequence of the matrices*

$$U^{k+1} = AU^k(R^k)^{-1}$$

where $Q^kR^k = AU^k$ is the QR decomposition of AU^k. We have

- *U^k converges to a matrix U whose columns are the top d eigenvectors of A with rate of convergence $\frac{\lambda_{d+1}}{\lambda_d}$.*

5.18 Notes on bibliography

For comprehensive treatment on both theory and applications of linear algebra and functional analysis on signal processing, see [17].

For introductory level treatment in linear algebra, see [1]. For intermediate to advanced treatment, see [3][4][9][18].

For positive matrix theory, see [8][9].

For numerical linear algebra, see [13][15]

BIBLIOGRAPHY

1. Meyer, C. D. *Matrix analysis and applied linear algebra* (Siam, 2000).
2. Krim, H. & Hamza, A. B. *Geometric methods in signal and image analysis* (Cambridge University Press, 2015).
3. Axler, S. J. *Linear algebra done right* (Springer).
4. Banerjee, S. & Roy, A. *Linear algebra and matrix analysis for statistics* (CRC Press, 2014).
5. Calafiore, G. C. & El Ghaoui, L. *Optimization Models* (Cambridge university press, 2014).
6. Theil, H. *Principles of econometrics* (1971).
7. Johnsonbaugh, R. & Pfaffenberger, W. Foundations of mathematical analysis (2010).
8. Luenberger, D. *Introduction to dynamic systems: theory, models, and applications* (Wiley, 1979).
9. Horn, R. A. & Johnson, C. R. *Matrix analysis* (Cambridge university press, 2012).
10. Wikipedia. *Matrix function — Wikipedia, The Free Encyclopedia* [Online; accessed 14-August-2016]. 2016.
11. Rhudy, M., Gu, Y., Gross, J. & Napolitano, M. R. Evaluation of matrix square root operations for UKF within a UAV GPS/INS sensor fusion application. *International Journal of Navigation and Observation* **2011** (2012).
12. Varga, R. S. *Matrix iterative analysis* (Springer Science & Business Media, 2009).
13. Saad, Y. *Iterative methods for sparse linear systems* (Siam, 2003).
14. Holmes, M. *Introduction to Numerical Methods in Differential Equations* ISBN: 9780387681214 (Springer New York, 2007).
15. Golub, G. H. & Van Loan, C. F. *Matrix Computations* (JHU Press, 2013).
16. Ma, Y. & Vidal, R. Generalized principal component analysis. *Unpublished Notes* (2002).
17. Moon, T. K. S. & Wynn, C. *Mathematical methods and algorithms for signal processing* **621.39: 51 MON** (2000).
18. Hom, R. A. & Johnson, C. R. Topics in matrix analysis. *Cambridge UP, New York* (1991).