REVIEW

Matrices and Gaussian Elimination

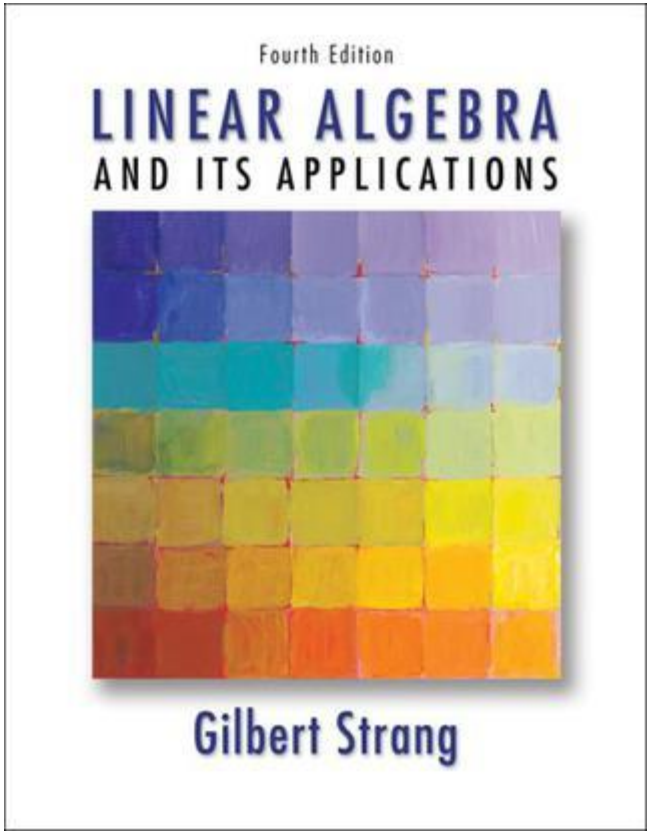
Vector Spaces

Orthogonality

Determinants

Eigenvalues and Eigenvectors

Positive Definite Matrices



Matrices and Gaussian Elimination

Linear systems $\xrightarrow{\text{Gaussian Eliminations}}$ Triangular systems

Solutions:
Inconsistent 0
Consistent $\begin{cases} \infty \\ 1 \end{cases}$

One linear system = Two triangular system

Matrix $\boldsymbol{A}$ $\xrightarrow[r_i \leftrightarrow r_j,\, r_i \to k r_i,\, k \neq 0,\, r_j \to r_j + k r_i]{\text{Elementary row operations}}$ Row echelon form $\boldsymbol{U}$ or Reduced row echelon form $\boldsymbol{R}$

Matrix operations
$\boldsymbol{A}+\boldsymbol{B}, k\boldsymbol{A}$
$\boldsymbol{AB}, \boldsymbol{A}^{-1}, \boldsymbol{A}^{T}$

rank

LU factorization
LDU factorization

Determine $\boldsymbol{A}$ is invertible;
Find $\boldsymbol{A}^{-1}$;
rank($\boldsymbol{A}$);
$\boldsymbol{A} = \boldsymbol{LU}$; $\boldsymbol{A} = \boldsymbol{LDU}$; …
Solve $\boldsymbol{Ax} = \boldsymbol{0}$; $\boldsymbol{Ax} = \boldsymbol{b}$;
Find bases;
Dependency
… … …

Matrix

- $\boldsymbol{A}+\boldsymbol{B}, k\boldsymbol{A}, \boldsymbol{AB}$
- $\boldsymbol{A}^T$
- four fundamental subspaces
- Elementary operations
- Rank
- Matrix factorization

$$\boldsymbol{A}=\boldsymbol{LU}$$
$$\boldsymbol{A}=\boldsymbol{LDU}$$
$$\boldsymbol{PA}=\boldsymbol{LDU}$$

- $\cdots$

Square matrix

- Invertible: $\boldsymbol{A}^{-1}$

Vector Spaces

Vector Spaces

Vectors:

Operations $\begin{cases}\text{Addition;} \\ \text{Scalar} \\ \quad \text{multiplication.}\end{cases}$

Linear combinations;
Linear dependence and
linear independence.

Subspaces

Spanning sets and
spanned subspaces

basis:
$\begin{cases}\text{Spanning set} \\ \text{Linear independent}\end{cases}$

A matrix $A_{m\times n}$:
row space
column space
nullspace
Left nullspace

basis?
dimension?

Linear transformation T: $\mathbb{R}^n \rightarrow \mathbb{R}^m$ $\longleftrightarrow$ $\boldsymbol{A}_{m\times n}$

Rotation
Reflection
Projections

$T(\boldsymbol{u}) = \boldsymbol{A}\boldsymbol{u}$
range$(T) = C(\boldsymbol{A})$
kernel$(T) = N(\boldsymbol{A})$

Vector Spaces and Vectors

- Vector space
- Subspaces: Spanning sets and spanned subspaces
- Orthogonal Subspaces
- Orthogonal complement
- Linear combinations
- Linear dependence and linear independence.
- basis
- Inner product
- Length
- Cosines
- Orthogonal vectors

3.1 Orthogonal Subspaces

$\boldsymbol{C}(A) =$ column space of A; dimension $\underline{r}$.	$\subseteq \mathbf{R}^m$
$\boldsymbol{N}(A) =$ nullspace of A; dimension $\underline{n-r}$.	$\subseteq \mathbf{R}^n$
$\boldsymbol{C}(A^{\mathrm{T}}) =$ row space of A; dimension $\underline{r}$.	$\subseteq \mathbf{R}^n$
$\boldsymbol{N}(A^{\mathrm{T}}) =$ left nullspace of A; dimension $\underline{m-r}$.	$\subseteq \mathbf{R}^m$

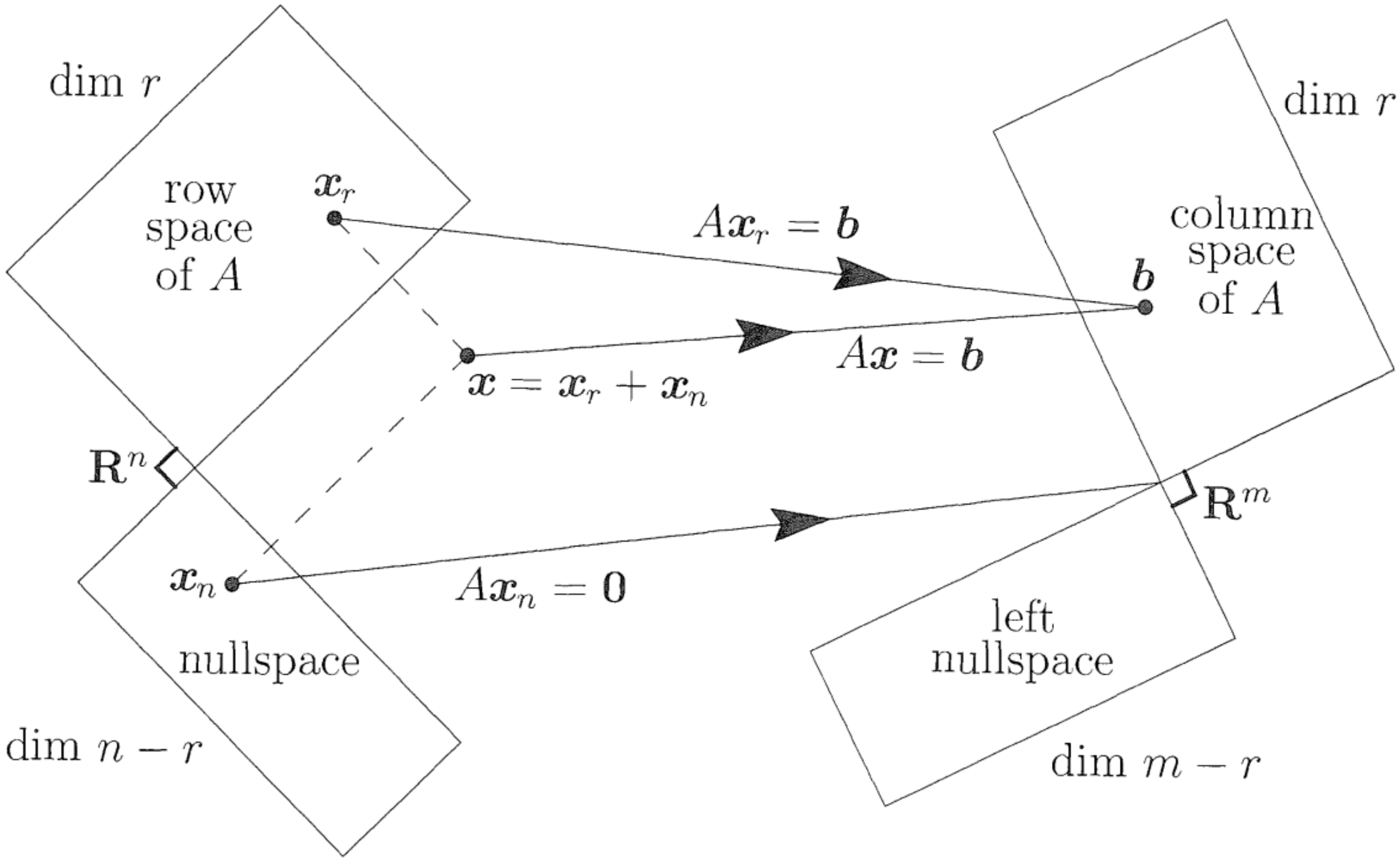
$$r + (n-r) = n$$

$$r + (m-r) = m$$

Theorem *Let* $\boldsymbol{A}$ *be a matrix. Then the row space* $\mathrm{C}(\boldsymbol{A}^{\mathrm{T}})$ *is orthogonal to the nullspace* $\mathrm{N}(\boldsymbol{A})$, *and the column space* $\mathrm{C}(\boldsymbol{A})$ *is orthogonal to the left nullspace* $\mathrm{N}(\boldsymbol{A}^{\mathrm{T}})$. *Moreover, if* $\boldsymbol{A}$ *has size* $m\times n$, *then*

$$\mathrm{N}(\boldsymbol{A}) = (\mathrm{C}(\boldsymbol{A}^{\mathrm{T}}))^{\perp}, \quad \textit{and} \quad \mathrm{N}(\boldsymbol{A}^{\mathrm{T}}) = (\mathrm{C}(\boldsymbol{A}))^{\perp}.$$

In other words: The nullspace is the *orthogonal complement* of the row space in $\mathbf{R}^n$. The left nullspace is the *orthogonal complement* of the column space in $\mathbf{R}^m$.

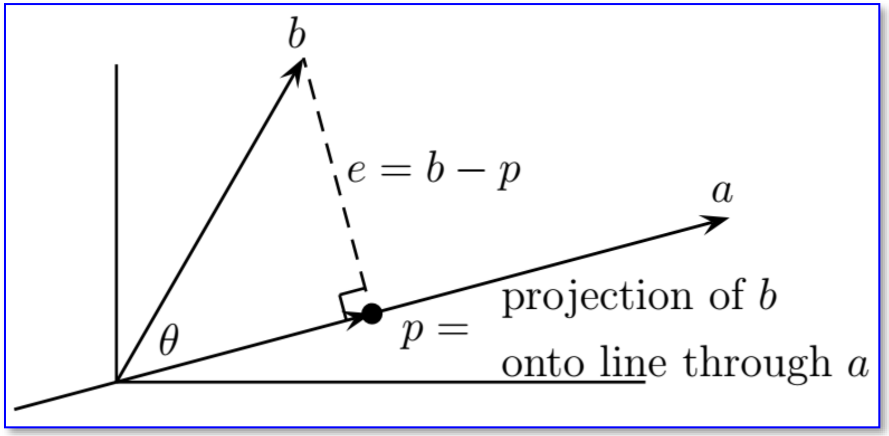


The true action $Ax = A(x_{\text{row}} + x_{\text{null}})$ of any m by n matrix.

3.2 Projection onto a Line

Proposition *The projection* $\operatorname{proj}_{\boldsymbol{a}}$ *satisfies*

$$\operatorname{proj}_{\boldsymbol{a}}(\boldsymbol{b}) = \frac{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}}\boldsymbol{a}.$$



Projection onto a line is carried out by a ***projection matrix*** $\boldsymbol{P}$:

$$\boldsymbol{P} = \frac{\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}}.$$

$\boldsymbol{P}$ is a matrix of rank 1, and as a linear transformation, it transforms a vector $\boldsymbol{b}$ to its projection $\operatorname{proj}_{\boldsymbol{a}}(\boldsymbol{b}) = \boldsymbol{P}\boldsymbol{b}$.

3.3 Least Squares

Theorem. *If a system* $\boldsymbol{Ax} = \boldsymbol{b}$ *is inconsistent* (*has no solution*), *its least-squares solution minimizes* $\left|\left|\boldsymbol{Ax} - \boldsymbol{b}\right|\right|^2$:

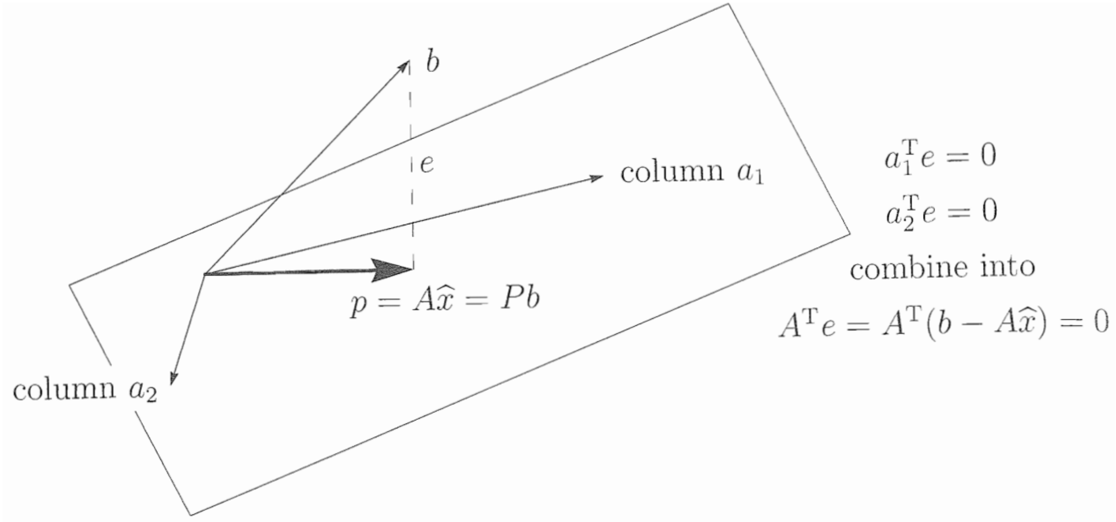
$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\widehat{\boldsymbol{x}} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}. \qquad \textbf{(Normal equations)}$$

Moreover, if $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ *is invertible, then*

$$\widehat{\boldsymbol{x}} = \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}. \qquad \textbf{(Best estimate)}$$

The projection of $\boldsymbol{b}$ *onto the column space is the nearest point* $\boldsymbol{A}\widehat{\boldsymbol{x}}$:

$$\boldsymbol{p} = \boldsymbol{A}\widehat{\boldsymbol{x}} = \boldsymbol{A}\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}. \qquad \textbf{(Projection)}$$



Projection onto the column space of a 3 by 2 matrix

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

- Consistent
  - It has a unique solution.
  - It has infinitely many solutions.

$$\boldsymbol{x} = \boldsymbol{x}_p + \boldsymbol{x}_n$$

where $\boldsymbol{x}_n \in \boldsymbol{N}(A)$.

- Inconsistent
  - $\widehat{\boldsymbol{x}}$: Least Squares solutions
  - The best $\widehat{\boldsymbol{x}}$ is the vector that minimizes the squared error

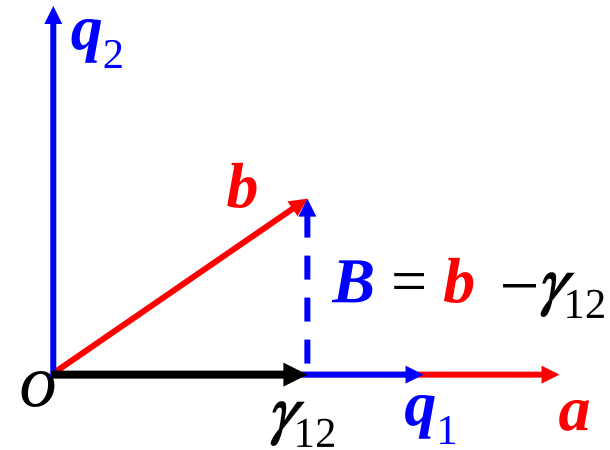
$$E^2 = ||\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}||^2.$$

3.4 Gram-Schmidt Orthogonalization

Convert a skewed set of axes into a perpendicular set

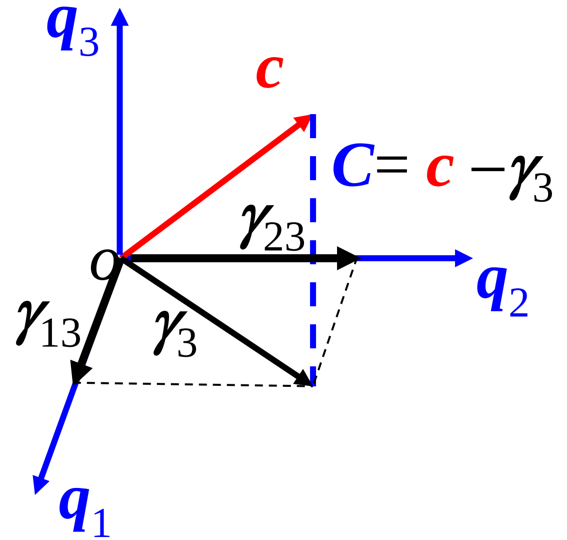
In $\mathbf{R}^n$, we try to make the independent vectors $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ orthonormal.

Let $\boldsymbol{q}_1 = \boldsymbol{a}/\|\boldsymbol{a}\|$ to make it a unit vector.



Take $\boldsymbol{B} = \boldsymbol{b} - \boldsymbol{\gamma}_{12} = \boldsymbol{b} - (\boldsymbol{q}_1^{\mathrm{T}}\boldsymbol{b})\boldsymbol{q}_1$,
and $\boldsymbol{q}_2 = \boldsymbol{B}/\|\boldsymbol{B}\|$.

The vector $\boldsymbol{c}$ will not be in the plane of $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$, which is the plane of $\boldsymbol{a}$ and $\boldsymbol{b}$.

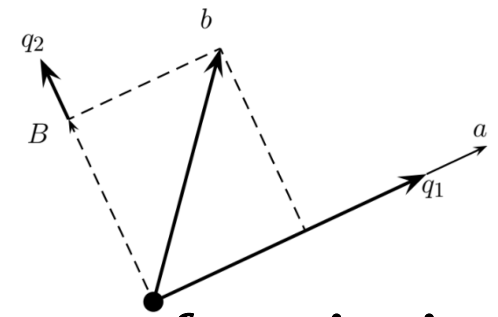


Take

$$\boldsymbol{C} = \boldsymbol{c} - \boldsymbol{\gamma}_3 = \boldsymbol{c} - (\boldsymbol{q}_1^{\mathrm{T}}\boldsymbol{c})\boldsymbol{q}_1 - (\boldsymbol{q}_2^{\mathrm{T}}\boldsymbol{c})\boldsymbol{q}_2,$$

and $\boldsymbol{q}_3 = \boldsymbol{C}/\|\boldsymbol{C}\|$.

QR Factorization



$$\boldsymbol{a} = (\boldsymbol{q}_1^{\mathrm{T}}\boldsymbol{a})\boldsymbol{q}_1,$$
$$\boldsymbol{b} = (\boldsymbol{q}_1^{\mathrm{T}}\boldsymbol{b})\boldsymbol{q}_1 + (\boldsymbol{q}_2^{\mathrm{T}}\boldsymbol{b})\boldsymbol{q}_2,$$
$$\boldsymbol{c} = (\boldsymbol{q}_1^{\mathrm{T}}\boldsymbol{c})\boldsymbol{q}_1 + (\boldsymbol{q}_2^{\mathrm{T}}\boldsymbol{c})\boldsymbol{q}_2 + (\boldsymbol{q}_3^{\mathrm{T}}\boldsymbol{c})\boldsymbol{q}_3.$$

If we express that in matrix form we have ***the new factorization*** ***A*** = ***QR***:

$$\boldsymbol{A} = [\boldsymbol{a} \quad \boldsymbol{b} \quad \boldsymbol{c}] = [\boldsymbol{q}_1 \quad \boldsymbol{q}_2 \quad \boldsymbol{q}_3]\begin{bmatrix} \boldsymbol{q}_1^{\mathrm{T}}\boldsymbol{a} & \boldsymbol{q}_1^{\mathrm{T}}\boldsymbol{b} & \boldsymbol{q}_1^{\mathrm{T}}\boldsymbol{c} \\ & \boldsymbol{q}_2^{\mathrm{T}}\boldsymbol{b} & \boldsymbol{q}_2^{\mathrm{T}}\boldsymbol{c} \\ & & \boldsymbol{q}_3^{\mathrm{T}}\boldsymbol{c} \end{bmatrix} = \boldsymbol{QR}.$$

- ***R*** is *upper triangular* because of the way Gram-Schmidt was done.
- ***Q*** has orthonormal columns.

4.1-4.3 Formulas and Properties of Determinants

The definition of determinant is expanded along **row 1**. Actually it can be extended along ***any row, or any column***, resulting in same value of the determinant.

Theorem *The determinant of* $\boldsymbol{A}$ *can be calculated by expanding along* ***row i***,

$$|\boldsymbol{A}| = (-1)^{i+1}a_{i1}|\boldsymbol{A}_{i1}| + (-1)^{i+2}a_{i2}|\boldsymbol{A}_{i2}| + \cdots + (-1)^{i+n}a_{in}|\boldsymbol{A}_{in}|,$$

and by expanding along ***column j***,

$$|\boldsymbol{A}| = (-1)^{1+j}a_{1j}|\boldsymbol{A}_{1j}| + (-1)^{2+j}a_{2j}|\boldsymbol{A}_{2j}| + \cdots + (-1)^{n+j}a_{nj}|\boldsymbol{A}_{nj}|.$$

Note: The determinant of the submatrix $\boldsymbol{A}_{ij}$ with the correct sign is also called the **cofactor (代数余子式)**, denoted by $C_{ij} = (-1)^{i+j}|\boldsymbol{A}_{ij}|$.

Pay attention to the sign!

For example,

$$D = \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix}, \quad \boxed{\begin{matrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{matrix}}$$

Summary of

The properties of Determinant

(可用于计算)

转置不改

换行反号

因子能提

行列可拆

倍加不变

三角化法
(Using elementary operations to find determinants)

4.4 Applications of Determinants

Define the **cofactor matrix** as

$$C=\begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{bmatrix}, \quad \text{and} \quad A^{*}=C^{\mathrm{T}}=\begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix}.$$

A^{*} is called the **adjoint matrix (adjugate matrix, 伴随矩阵)** of A.

The method to find the inverse of a matrix (求逆矩阵的**伴随矩阵法**):

(1) Calculate the determinant of $A=[a_{ij}]$;

(2) If $|A|=0$, then A is not invertible;

(3) If $|A|\neq 0$, then find the cofactor of each entry a_{ij} and the adjoint matrix of A, denoted by A^{*}, and finally we get $\boxed{A^{-1}=\dfrac{1}{|A|}A^{*}.}$

5.1 Eigenvalues and Eigenvectors

*A **Process** for finding eigenvalues and eigenvectors of a matrix **A**:*

1. ***Compute the determinant of*** $A - \lambda I$.

With λ subtracted along the diagonal, this determinant is a polynomial of degree n. It starts with $(-\lambda)^n$.

2. ***Find the roots of this polynomial***.

The n roots are the eigenvalues of A.

3. ***For each eigenvalue solve the equation*** $(A - \lambda I)x = 0$.

Since the determinant is zero, there are solutions other than $x = 0$. Those are the eigenvectors.

$$A \rightarrow |A - \lambda I| = 0 \rightarrow (A - \lambda_i I)x = 0$$

求特征值λ_i

求特征向量

5.2 Diagonalization of a Matrix

algebraic multiplicity ≥ geometric multiplicity

(几何重数总是不超过代数重数)

The matrix ***A*** is diagonalizable if and only if *algebraic multiplicity = geometric multiplicity for each eigenvalue* λ_i.

(矩阵*A* 可以对角化 当且仅当 对于每一个特征值 λ_i都有:

其代数重数与几何重数相等)

5.5 Complex Matrices

实数域 R		复数域 C	
Name	**Definition**	**Name**	**Definition**
Real symmetric matrix	$\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$	*Hermitian matrix*	$\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{A}$
Real normal matrix	$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$	*Normal matrix*	$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$
Orthogonal matrix	$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{I}$	*Unitary matrix*	$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{I}$

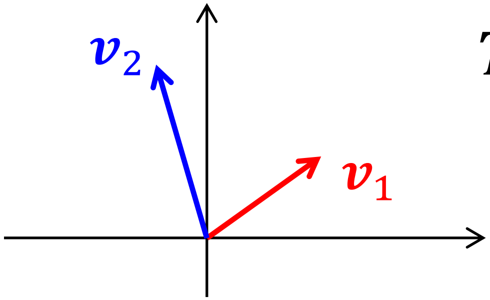
Real versus Complex

$\mathbf{R}^n$ (n real components)	$\leftrightarrow$	$\mathbf{C}^n$ (n complex components)
length: $\|x\|^2 = x_1^2 + \cdots + x_n^2$	$\leftrightarrow$	length: $\|x\|^2 = \lvert x_1\rvert^2 + \cdots + \lvert x_n\rvert^2$
transpose: $A_{ij}^{\mathrm{T}} = A_{ji}$	$\leftrightarrow$	Hermitian transpose: $A_{ij}^{\mathrm{H}} = \overline{A_{ji}}$
$(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}$	$\leftrightarrow$	$(AB)^{\mathrm{H}} = B^{\mathrm{H}}A^{\mathrm{H}}$
inner product: $x^{\mathrm{T}}y = x_1y_1 + \cdots + x_ny_n$	$\leftrightarrow$	inner product: $x^{\mathrm{H}}y = \bar{x}_1y_1 + \cdots + \bar{x}_ny_n$
$(Ax)^{\mathrm{T}}y = x^{\mathrm{T}}(A^{\mathrm{T}}y)$	$\leftrightarrow$	$(Ax)^{\mathrm{H}}y = x^{\mathrm{H}}(A^{\mathrm{H}}y)$
orthogonality: $x^{\mathrm{T}}y = 0$	$\leftrightarrow$	orthogonality: $x^{\mathrm{H}}y = 0$
symmetric matrices: $A^{\mathrm{T}} = A$	$\leftrightarrow$	Hermitian matrices: $A^{\mathrm{H}} = A$
$A = Q\Lambda Q^{-1} = Q\Lambda Q^{\mathrm{T}}$ (real Λ)	$\leftrightarrow$	$A = U\Lambda U^{-1} = U\Lambda U^{\mathrm{H}}$ (real Λ)
skew-symmetric $K^{\mathrm{T}} = -K$	$\leftrightarrow$	skew-Hermitian $K^{\mathrm{H}} = -K$
orthogonal $Q^{\mathrm{T}}Q = I$ or $Q^{\mathrm{T}} = Q^{-1}$	$\leftrightarrow$	unitary $U^{\mathrm{H}}U = I$ or $U^{\mathrm{H}} = U^{-1}$
$(Qx)^{\mathrm{T}}(Qy) = x^{\mathrm{T}}y$ and $\|Qx\| = \|x\|$	$\leftrightarrow$	$(Ux)^{\mathrm{H}}(Uy) = x^{\mathrm{H}}y$ and $\|Ux\| = \|x\|$

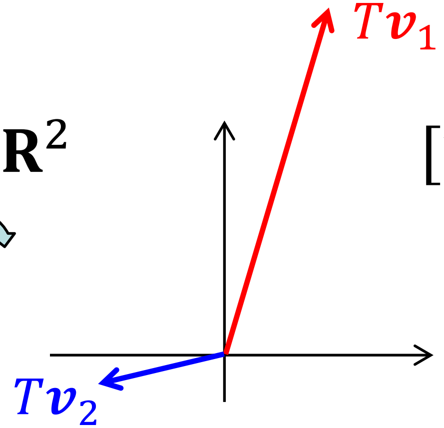
The columns, rows, and eigenvectors of Q and U are orthonormal, and every $\lvert\lambda\rvert = 1$

5.6 Similarity Transformations

Basis $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$

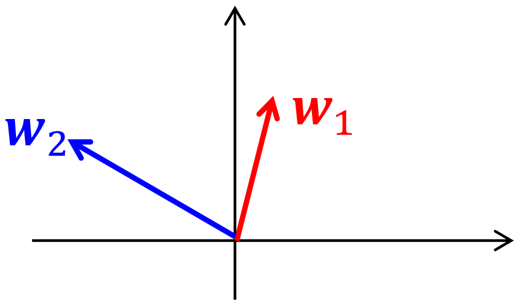


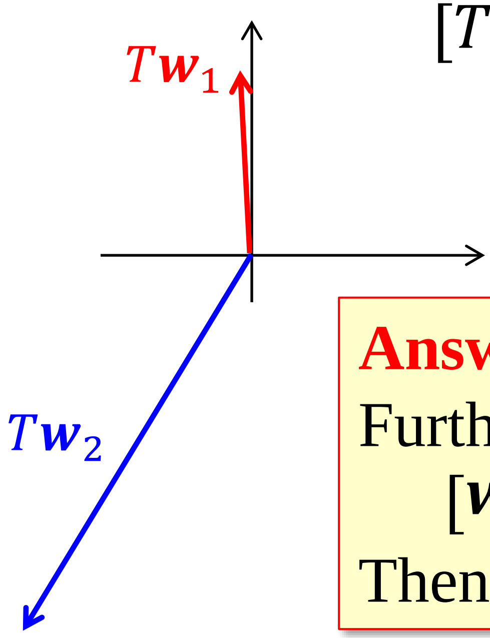
$T\colon \mathbf{R}^2 \to \mathbf{R}^2$



$$[T\boldsymbol{v}_1 \quad T\boldsymbol{v}_2] = [\boldsymbol{v}_1 \quad \boldsymbol{v}_2]\boldsymbol{A}$$

Basis $\{\boldsymbol{w}_1, \boldsymbol{w}_2\}$





$$[T\boldsymbol{w}_1 \quad T\boldsymbol{w}_2] = [\boldsymbol{w}_1 \quad \boldsymbol{w}_2]\boldsymbol{B}$$

Question:

How is $\boldsymbol{A}$ related to $\boldsymbol{B}$?

Answer: $\boldsymbol{A}$ *is similar to* $\boldsymbol{B}$.

Furthermore, if

$$[\boldsymbol{w}_1 \quad \boldsymbol{w}_2] = [\boldsymbol{v}_1 \quad \boldsymbol{v}_2]\boldsymbol{M}$$

Then $\boldsymbol{B} = \boldsymbol{M}^{-1}\boldsymbol{A}\boldsymbol{M}$.

When A is diagonalizable: $\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \boldsymbol{\Lambda}$

$$\boldsymbol{S} = \underbrace{\begin{bmatrix} | & | & \cdots & | \\ \boldsymbol{x}_1 & \boldsymbol{x}_2 & \cdots & \boldsymbol{x}_n \\ | & | & \cdots & | \\ | & | & \cdots & | \end{bmatrix}}_{\textit{independent}}, \quad \boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}.$$

$\boldsymbol{S}$: invertible matrix

$\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$: eigenvectors

$\lambda_1, \dots, \lambda_n$: eigenvalues

When A is real symmetric: $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q} = \boldsymbol{\Lambda}$

$$\boldsymbol{Q} = \underbrace{\begin{bmatrix} | & | & \cdots & | \\ \boldsymbol{x}_1 & \boldsymbol{x}_2 & \cdots & \boldsymbol{x}_n \\ | & | & \cdots & | \\ | & | & \cdots & | \end{bmatrix}}_{\textit{orthonormal}}, \quad \boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}.$$

$\boldsymbol{Q}$: orthogonal matrix

$\lambda_1, \dots, \lambda_n$: real

When A is Hermitian: $\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U} = \boldsymbol{\Lambda}$

$$\boldsymbol{U} = \underbrace{\begin{bmatrix} | & | & \cdots & | \\ \boldsymbol{x}_1 & \boldsymbol{x}_2 & \cdots & \boldsymbol{x}_n \\ | & | & \cdots & | \\ | & | & \cdots & | \end{bmatrix}}_{\textit{orthonormal}}, \quad \boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}.$$

$\boldsymbol{U}$: unitary matrix

$\lambda_1, \dots, \lambda_n$: real

Some special matrices

Real matrices	Complex matrices	Eigenvalues
Symmetric $\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A}$	Hermitian $\boldsymbol{A}^{\mathrm{H}} = \boldsymbol{A}$	All λ's are real (on the real axis)
Skew-symmetric $\boldsymbol{A}^{\mathrm{T}} = -\boldsymbol{A}$	Skew-Hermitian $\boldsymbol{A}^{\mathrm{H}} = -\boldsymbol{A}$	All λ's are imaginary (including 0 sometimes) (on the imaginary axis)
Orthogonal $\boldsymbol{Q}^{\mathrm{T}} = \boldsymbol{Q}^{-1}$	Unitary $\boldsymbol{U}^{\mathrm{H}} = \boldsymbol{U}^{-1}$	all $\lvert\lambda\rvert = 1$ (on the unit circle)

Theorem *A matrix is diagonalized by a unitary matrix if and only if it is a normal matrix.*
(In other words, *A matrix* $\boldsymbol{A}$ *is a normal matrix if and only if there exists a unitary matrix* $\boldsymbol{U}$ *such that* $\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}$ *is diagonal.*)

The Spectral Theorem for Real Symmetric Matrices

An $n \times n$ real symmetric matrix $\boldsymbol{A}$ ($\boldsymbol{A} \in \mathbf{R}^{n\times n}$ and $\boldsymbol{A} = \boldsymbol{A}^{\mathrm{T}}$) has the following properties:

(一个对称的 $n \times n$ 实矩阵具有下面的特性)

a. $\boldsymbol{A}$ has n real eigenvalues, counting multiplicities. ($\boldsymbol{A}$有n个实特征值, 包含重复的特征值)

b. The dimension of the eigenspace for each eigenvalue λ equals the multiplicity of λ as a root of the characteristic equation. (对于每一个特征值λ, 对应特征子空间的维数等于λ作为特征方程的根的重数, 即: 几何重数=代数重数)

c. The eigenspaces are mutually orthogonal, in the sense that eigenvectors corresponding to different eigenvalues are orthogonal. (特征子空间相互正交, 这种正交性是在特征向量对应不同特征值的意义下成立的)

d. $\boldsymbol{A}$ is orthogonally diagonalizable. ($\boldsymbol{A}$可以正交对角化)

The Spectral Theorem for Hermitian Matrices

An $n \times n$ Hermitian matrix $\boldsymbol{A}$ ($\boldsymbol{A} \in \mathbf{C}^{n\times n}$ and $\boldsymbol{A} = \boldsymbol{A}^{\mathrm{H}}$) has the following properties:

(一个 $n \times n$ 厄米特矩阵具有下面的特性)

a. $\boldsymbol{A}$ has n real eigenvalues, counting multiplicities. ($\boldsymbol{A}$有n个实特征值, 包含重复的特征值)

b. The dimension of the eigenspace for each eigenvalue λ equals the multiplicity of λ as a root of the characteristic equation. (对于每一个特征值λ, 对应特征子空间的维数等于λ作为特征方程的根的重数, 即: 几何重数=代数重数)

c. The eigenspaces are mutually orthogonal, in the sense that eigenvectors corresponding to different eigenvalues are orthogonal. (特征子空间相互正交, 这种正交性是在特征向量对应不同特征值的意义下成立的)

d. $\boldsymbol{A}$ can be diagonalized by a unitary matrix. ($\boldsymbol{A}$可以用酉矩阵对角化)

Similarity Transformations

1. $\boldsymbol{A}$ is ***diagonalizable***: The columns of $\boldsymbol{S}$ are eigenvectors and $\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \boldsymbol{\Lambda}$.

2. $\boldsymbol{A}$ is ***arbitrary***: The columns of $\boldsymbol{M}$ include "generalized eigenvectors" of $\boldsymbol{A}$, and the Jordan form $\boldsymbol{M}^{-1}\boldsymbol{A}\boldsymbol{M} = \boldsymbol{J}$ is *block diagonal*.

3. $\boldsymbol{A}$ is ***arbitrary***: The unitary $\boldsymbol{U}$ can be chosen so that $\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U} = \boldsymbol{T}$ is *triangular*.

4. $\boldsymbol{A}$ is ***normal***, $\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}} = \boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$: then $\boldsymbol{U}$ can be chosen so that $\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U} = \boldsymbol{\Lambda}$.

Special cases of normal matrices, all with orthonormal eigenvectors:

(a) If $\boldsymbol{A} = \boldsymbol{A}^{\mathrm{H}}$ is Hermitian, then all λ_i are real.

(b) If $\boldsymbol{A} = \boldsymbol{A}^{\mathrm{T}}$ is real symmetric, then $\boldsymbol{\Lambda}$ is real and $\boldsymbol{U} = \boldsymbol{Q}$ is orthogonal.

(c) If $\boldsymbol{A} = -\boldsymbol{A}^{\mathrm{H}}$ is skew-Hermitian, then all λ_i are purely imaginary.

(d) If $\boldsymbol{A}$ is orthogonal or unitary, then all $|\lambda_i| = 1$ are on the unit circle.

6.1-6.2 Positive Definiteness

Theorem (Test for *positive definiteness*) Each of the following tests is a necessary and sufficient condition for the real symmetric matrix $\boldsymbol{A}$ to be ***positive definite***:

(正定性判别：以下任何一个都是判定一个实对称矩阵$\boldsymbol{A}$正定的充要条件)

(I) $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} > 0$ for all nonzero real vectors $\boldsymbol{x}$. (Definition)

(II) All the eigenvalues of $\boldsymbol{A}$ satisfy $\lambda_i > 0$.

(III) All the upper left submatrices $\boldsymbol{A}_k$ have positive determinants.

(IV) All the pivots (without row exchanges) satisfy $d_k > 0$.

(V) There is a matrix $\boldsymbol{R}$ with independent columns such that $\boldsymbol{A} = \boldsymbol{R}^{\mathrm{T}}\boldsymbol{R}$.

Theorem Each of the following tests is a necessary and sufficient condition for a symmetric matrix $\boldsymbol{A}$ to be ***positive semidefinite***:

(I’) $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} \geq 0$ for all vectors $\boldsymbol{x}$. (This defines positive semidefinite)

(II’) All the eigenvalues of $\boldsymbol{A}$ satisfy $\lambda_i \geq 0$.

(III’) No principal submatrices have negative determinants.

(判定半正定性时, 不仅要检查左上角各阶子矩阵的行列式, 即顺序主子式, 而且检查所有各阶主子矩阵的行列式即主子式)

(IV’) No pivots are negative.

(V’) There is a matrix $\boldsymbol{R}$, possibly with dependent columns, such that $\boldsymbol{A} = \boldsymbol{R}^{\mathrm{T}}\boldsymbol{R}$.

For an invertible matrix $\boldsymbol{C}$, the linear transformation $\boldsymbol{A} \to \boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$ is called a **congruence transformation (合同变换)**, which transforms the vector $\boldsymbol{y}$ to the vector $\boldsymbol{x} = \boldsymbol{C}\boldsymbol{y}$, and the quadratic form $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ to the quadratic form $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}\boldsymbol{y}$.

Theorem (*The Principal Axes Theorem*, 主轴定理)

Let $\boldsymbol{A}$ be an $n \times n$ real symmetric matrix. Then there is an orthogonal change of variable, $\boldsymbol{x} = \boldsymbol{Q}\boldsymbol{y}$ (i.e., $\boldsymbol{Q}$ is an *orthogonal* matrix), that transforms the quadratic form $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ into a quadratic form $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{y}$ with no cross-product term (不含交叉乘积项) (i.e., $\boldsymbol{\Lambda}$ is a diagonal matrix).

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{y} = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 \text{ (二次型的标准形)},$$

where $\lambda_1, \cdots, \lambda_n$ are eigenvalues of $\boldsymbol{A}$, and their orthonormal eigenvectors go into the columns of $\boldsymbol{Q}$. $(\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q} = \boldsymbol{\Lambda})$

Matrix A	Operations (变换)	Matrix B	Invariants (不变量)
$\boldsymbol{A}$ is any $m \times n$ matrix	Elementary operations	$\boldsymbol{B} = \boldsymbol{PAQ}$ (where $\boldsymbol{P}$ and $\boldsymbol{Q}$ are invertible $m \times m$ and $n \times n$ matrices)	Rank
$\boldsymbol{A}$ is any $n \times n$ matrix	Similarity transformation (相似变换)	$\boldsymbol{B} = \boldsymbol{M}^{-1}\boldsymbol{AM}$ (where $\boldsymbol{M}$ is an invertible $n \times n$ matrix)	Eigenvalues; Determinant; Trace; Rank
$\boldsymbol{A}$ is any real symmetric $n \times n$ matrix	Congruence transformation (合同变换)	$\boldsymbol{B} = \boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC}$ (where $\boldsymbol{C}$ is an invertible $n \times n$ matrix)	Symmetry; Rank; Number of positive eigenvalues, negative eigenvalues, and zero eigenvalues

6.3 Singular Value Decomposition

Facts about $A^{\mathrm{T}}A$ and AA^{T} ($A \in \mathrm{R}^{m\times n}$)

- $\operatorname{rank}\left(A^{\mathrm{T}}A\right) = \operatorname{rank}\left(AA^{\mathrm{T}}\right) = \operatorname{rank}(A) = r$.
- $A^{\mathrm{T}}A$ and AA^{T} are real symmetric (degree n and m respectively), and positive semidefinite. ($A^{\mathrm{T}}A$ and AA^{T}的特征值为非负实数)
- The eigenvalues of $A^{\mathrm{T}}A$ and AA^{T} :
  - $A^{\mathrm{T}}A$ has n eigenvalues $\lambda_1, \dots, \lambda_n$, then
  $$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0 = \lambda_{r+1} = \cdots = \lambda_n.$$
  - AA^{T} has m eigenvalues $\mu_1, \dots, \mu_m$, then
  $$\mu_1 \geq \mu_2 \geq \cdots \geq \mu_r > 0 = \mu_{r+1} = \cdots = \mu_m.$$
  - We have the following conclusion: $\lambda_i = \mu_i > 0,\ i = 1, \dots, r$. ($A^{\mathrm{T}}A$ and AA^{T}的非零特征值集合相同)
  - **Definition**: $\sigma_i = \sqrt{\lambda_i} = \sqrt{\mu_i} > 0\ (i = 1, \dots, r)$ are called the **singular values (奇异值)** of A.

Theorem ***(Singular Value Decomposition -- “SVD”)***

Any $m \times n$ real matrix with rank r can be factored into

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}},$$

(orthogonal) (rectangular diagonal) (orthogonal)

where $\boldsymbol{U}$ is orthogonal of degree m, $\boldsymbol{V}$ is orthogonal of degree n, and $\boldsymbol{\Sigma}$ is diagonal (but rectangular: $m \times n$).

Further, the columns of $\boldsymbol{U}$ are eigenvectors of $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$, the columns of $\boldsymbol{V}$ are eigenvectors of $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$, and the r positive entries $\sigma_1, \ldots, \sigma_r$ (called ‘singular values’) on the diagonal of $\boldsymbol{\Sigma}$ are the square roots of the nonzero eigenvalues of both $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$ and $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$.

The factorization is called a **singular value decomposition (奇异值分解)**, or **SVD** for short.

$$\boldsymbol{A}_{m\times n} = \boldsymbol{U}_{m\times m} \left[\begin{array}{cccc:c} \sigma_1 & & & & \\ & \sigma_2 & & & \mathbf{0}_{r\times(n-r)} \\ & & \ddots & & \\ & & & \sigma_r & \\ \hdashline & \mathbf{0}_{(m-r)\times r} & & & \mathbf{0}_{(m-r)\times(n-r)} \end{array}\right]_{m\times n} (\boldsymbol{V}_{n\times n})^T$$

(orthogonal) (rectangular diagonal) (orthogonal)

where $\sigma_1, \sigma_2, \cdots, \sigma_r$ are the square roots of the nonzero eigenvalues of both $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$ and $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$.

- $\boldsymbol{U}$: the columns are orthonormal eigenvectors for $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$.
- $\boldsymbol{V}$: the columns are orthonormal eigenvectors for $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$.

$$\boldsymbol{\Sigma} = \left[\begin{array}{cccc|c} \sigma_1 & & & & \\ & \sigma_2 & & & \mathbf{0}_{r\times(n-r)} \\ & & \ddots & & \\ & & & \sigma_r & \\ \hline & \mathbf{0}_{(m-r)\times r} & & & \mathbf{0}_{(m-r)\times(n-r)} \end{array}\right]_{m\times n}$$

$$\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{\mathrm{T}} = \begin{bmatrix} \sigma_1^2 & & & & \\ & \ddots & & & \\ & & \sigma_r^2 & & \\ & & & 0 & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix}_{m\times m} \qquad \boldsymbol{\Sigma}^{\mathrm{T}}\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & & & & \\ & \ddots & & & \\ & & \sigma_r^2 & & \\ & & & 0 & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix}_{n\times n}$$

Diagonal entries are eigenvalues for $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$

Diagonal entries are eigenvalues for $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$

How to construct the matrix V

$\boldsymbol{V}$: the columns are orthonormal eigenvectors for $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$.

Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ are columns of $\boldsymbol{V}$, and let

$$\boldsymbol{V} = [\boldsymbol{v}_1 \quad \cdots \quad \boldsymbol{v}_r \quad \boldsymbol{v}_{r+1} \quad \cdots \quad \boldsymbol{v}_n] = [\boldsymbol{V}_r \vdots \boldsymbol{V}_{n-r}]$$

Then $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \boldsymbol{V}\left(\boldsymbol{\Sigma}^{\mathrm{T}}\boldsymbol{\Sigma}\right)\boldsymbol{V}^{\mathrm{T}}$ becomes

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}[\boldsymbol{v}_1 \quad \cdots \quad \boldsymbol{v}_r \quad \boldsymbol{v}_{r+1} \quad \cdots \quad \boldsymbol{v}_n]$$

$$= [\boldsymbol{v}_1 \quad \cdots \quad \boldsymbol{v}_r \quad \boldsymbol{v}_{r+1} \quad \cdots \quad \boldsymbol{v}_n]\begin{bmatrix} \sigma_1^2 & & & & & \\ & \ddots & & & & \\ & & \sigma_r^2 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix}$$

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{v}_1 = \sigma_1^2\boldsymbol{v}_1, \cdots, \boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{v}_r = \sigma_r^2\boldsymbol{v}_r, \boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{v}_{r+1} = \mathbf{0}, \cdots, \boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{v}_n = \mathbf{0}$$

$\boldsymbol{v}_1, \cdots, \boldsymbol{v}_r$ are eigenvectors of $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ belonging to nonzero eigenvalues $\sigma_1^2, \cdots, \sigma_r^2$ respectively.

$\boldsymbol{v}_{r+1}, \cdots, \boldsymbol{v}_n$ are eigenvectors of $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ belonging to $\lambda = 0$.

How to construct the matrix U

$$\boldsymbol{A}_{m\times n} = \boldsymbol{U}_{m\times m}\left[\begin{array}{cccc|c} \sigma_1 & & & & \\ & \sigma_2 & & & \mathbf{0}_{r\times(n-r)} \\ & & \ddots & & \\ & & & \sigma_r & \\ \hline & \mathbf{0}_{(m-r)\times r} & & & \mathbf{0}_{(m-r)\times(n-r)} \end{array}\right]_{m\times n} (\boldsymbol{V}_{n\times n})^{\mathrm{T}}$$

$$\Longrightarrow \boldsymbol{A}_{m\times n}\boldsymbol{V}_{n\times n} = \boldsymbol{U}_{m\times m}\,\boldsymbol{\Sigma}_{m\times n}$$

$$\boldsymbol{A}[\boldsymbol{v}_1 \quad \cdots \quad \boldsymbol{v}_r \quad \boldsymbol{v}_{r+1} \quad \cdots \quad \boldsymbol{v}_n] = [\boldsymbol{u}_1 \quad \cdots \quad \boldsymbol{u}_r \quad \boldsymbol{u}_{r+1} \quad \cdots \quad \boldsymbol{u}_m]\begin{bmatrix} \boldsymbol{\Sigma}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

Comparing the first r columns of each side, we see that

$$\boldsymbol{A}\boldsymbol{v}_j = \sigma_j\boldsymbol{u}_j, j = 1,\cdots,r \Longrightarrow \boldsymbol{u}_j = \frac{1}{\sigma_j}\boldsymbol{A}\boldsymbol{v}_j, j = 1,\cdots,r$$

It follows from that each $\boldsymbol{u}_j, j = 1,\cdots,r$, is in the column space of $\boldsymbol{A}$.

The dimension of the column space is r, so $\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_r$ form an orthonormal basis for $C(\boldsymbol{A})$. The vector space $\mathrm{C}(\boldsymbol{A})^{\perp} = \mathrm{N}\left(\boldsymbol{A}^{\mathrm{T}}\right)$ has dimension $m-r$. $\boldsymbol{u}_{r+1},\cdots,\boldsymbol{u}_m$ is an orthonormal basis for $\mathrm{N}\left(\boldsymbol{A}^{\mathrm{T}}\right)$.

Type	Form	The Factors	Notes
***LU* factorization** *(Gaussian elimination)*	$\boldsymbol{PA} = \boldsymbol{LU}$ ($\boldsymbol{A}$ is any $m \times n$ matrix)	$\boldsymbol{P}$: permutation matrix $\boldsymbol{L}$: lower triangular matrix with unit diagonal $\boldsymbol{U}$: $m \times n$ echelon matrix (When $m = n$, $\boldsymbol{U}$ is upper triangular.)	The permutation matrix $\boldsymbol{P}$ is needed when there are row exchanges during the row reduction. (Otherwise, $\boldsymbol{A} = \boldsymbol{LU}$) $\boldsymbol{PA} = \boldsymbol{LDU}$ if $\boldsymbol{U}$ is upper triangular with unit diagonal.
***QR* factorization** *(Gram-Schmidt orthogonalization)*	$\boldsymbol{A} = \boldsymbol{QR}$ ($\boldsymbol{A}$ is any $m \times n$ matrix with independent columns)	$\boldsymbol{Q}$: matrix with orthonormal columns (When $m = n$, $\boldsymbol{Q}$ becomes an orthogonal matrix.) $\boldsymbol{R}$: upper triangular and invertible	When $m = n$, any invertible matrix can be factorized as a product of an orthogonal matrix and an upper triangular matrix.
Singular Value Decomposition *(SVD)*	$\boldsymbol{A} = \boldsymbol{U\Sigma V}^{\mathrm{T}}$ ($\boldsymbol{A}$ is any $m \times n$ matrix with rank r)	$\boldsymbol{U}$: $m \times m$ orthogonal matrix $\boldsymbol{V}$: $n \times n$ orthogonal matrix $\boldsymbol{\Sigma}$ is diagonal (but rectangular: $m \times n$).	The columns of $\boldsymbol{U}$ are eigenvectors of $\boldsymbol{AA}^{\mathrm{T}}$, the columns of $\boldsymbol{V}$ are eigenvectors of $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$, and the r positive entries on the diagonal of $\boldsymbol{\Sigma}$ are the square roots of the nonzero eigenvalues of both $\boldsymbol{AA}^{\mathrm{T}}$ and $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$.

For any rectangular diagonal matrix $\boldsymbol{\Sigma}$:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & \end{bmatrix}_{m\times n}, \quad \boldsymbol{\Sigma}^+ = \begin{bmatrix} 1/\sigma_1 & & & \\ & \ddots & & \\ & & 1/\sigma_r & \\ & & & \end{bmatrix}_{n\times m},$$

$$\boldsymbol{x}^+ = \boldsymbol{\Sigma}^+\boldsymbol{b} = \begin{bmatrix} b_1/\sigma_1 \\ \vdots \\ b_r/\sigma_r \\ \\ \end{bmatrix}_{n\times 1}. \qquad \textit{and obviously } (\boldsymbol{\Sigma}^+)^+ = \boldsymbol{\Sigma}.$$

Theorem *If* $\boldsymbol{A} = \boldsymbol{U\Sigma V}^T$ *(the SVD), then its* ***pseudoinverse*** *is*

$$\boldsymbol{A}^+ = \boldsymbol{V\Sigma}^+\boldsymbol{U}^T.$$

The End