Linear algebra and applications

Jitkomut Songsiri

Department of Electrical Engineering
Faculty of Engineering
Chulalongkorn University

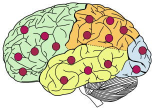
CUEE

Ja [illegible] 5

Outline

How to read this handout

1. the note is used with lecture in EE205 (you cannot master this topic just by reading this note) – class activities include
   - graphical concepts, math derivation of details/steps in between
   - computer codes to illustrate examples
2. always read 'textbooks' after lecture
3. pay attention to the symbol ✎; you should be able to prove such ✎ result
4. each chapter has a list of references; find more formal details/proofs from in-text citations
5. almost all results in this note can be Googled; readers are encouraged to 'stimulate neurons' in your brain by proving results without seeking help from the Internet first
6. typos and mistakes can be reported to `jitkomut@gmail.com`



Matrices

Vector notation

n-vector x:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

- also written as $x = (x_1, x_2, \ldots, x_n)$
- set of n-vectors is denoted $\mathbf{R}^n$ (Euclidean space)
- x_i: ith **element** or **component** or **entry** of x
- it is common to denote x as a column vector
- $x^T = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}$ is then a row vector

Special vectors

standard unit vector in $\mathbf{R}^n$ is a vector with all zero element except one element which is equal to one

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

ones vector is the n-vector with all its elements equal to one, denoted as $\mathbf{1}$

stacked vectors: if b, c, d are vectors (can be different sizes)

$$a = \begin{bmatrix} b \\ c \\ d \end{bmatrix}, \quad \text{or} \quad a = (b, c, d)$$

is the *stacked (or concatenated) vector* of b, c, d

Linear combination of vectors

if $a_1, a_2, \ldots, a_m$ are n-vectors, and $\alpha_1, \ldots, \alpha_m$ are scalars, the n-vector

$$\beta_1 a_1 + \beta_2 a_2 + \cdots + \beta_m a_m$$

is called a **linear combination** of the vectors $a_1, \ldots, a_m$

special linear combinations

- any n-vector a can be expressed as $a = a_1 e_1 + a_2 e_2 + \cdots + a_n e_n$
- the linear combination with $\beta_1 = \cdots = \beta_m = 1$ given by $a_1 + \cdots + a_m$ is the **sum** of the vectors
- the linear combination with $\beta_1 = \cdots = \beta_m = 1/m$ given by $(a_1 + \cdots + a_m)/m$ is the **average** of the vectors
- when the coefficients are non-negative and sum to one, *i.e.*, $\beta_1 + \cdots + \beta_m = 1$, the linear combination is called a **convex combination** or **weighted average**

Inner products

definition: the inner product of two n-vectors x, y is

$$x_1y_1 + x_2y_2 + \cdots + x_ny_n$$

also known as the **dot product** of vectors x, y

notation: $x^T y$

properties ✎

- $(\alpha x)^T y = \alpha(x^T y)$ for scalar α
- $(x + y)^T z = x^T z + y^T z$
- $x^T y = y^T x$

Examples

- unit vector: $e_i^T a = a_i$ the inner product of a vector with e_i gives the ith element of a
- sum: $\mathbf{1}^T a = a_1 + a_2 + \cdots + a_n$
- average: $(\mathbf{1}/n)^T a = (a_1 + \cdots + a_n)/n$
- sum of squares: $a^T a = a_1^2 + a_2^2 + \cdots + a_n^2$
- selective sum: let b be a vector all of whose entries are either 0 or 1; then $b^T a$ is the sum of elements in a for which $b_i = 1$

$$b = (0,1,0,0,1), \quad b^T a = a_2 + a_5$$

- polynomial evaluation: let c be the n-vector represents the coefficients of polynomial p with degree $n-1$

$$p(x) = c_1 + c_2 x + \cdots + c_{n-1}x^{n-2} + c_n x^{n-1}$$

let t be a number and $z = (1, t, t^2, \ldots, t^{n-1})$ then $c^T z = p(t)$

Euclidean norm

$$\|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = \sqrt{x^T x}$$

properties

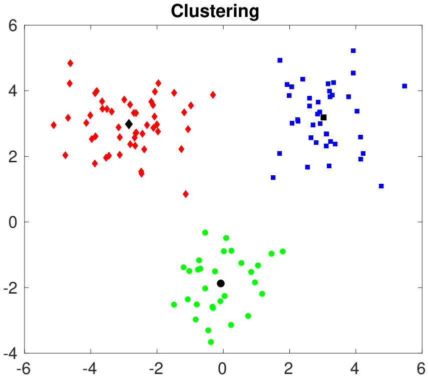
- also written $\|x\|_2$ to distinguish from other norms
- $\|\alpha x\| = |\alpha| \|x\|$ for scalar α
- $\|x + y\| \leq \|x\| + \|y\|$ (triangle inequality)
- $\|x\| \geq 0$ and $\|x\| = 0$ only if $x = 0$

interpretation

- $\|x\|$ measures the *magnitude* or length of x
- $\|x - y\|$ measures the *distance* between x and y

Cluster centroid

given three clusters of data points



it can be shown that the representative is in fact, the **centroid** of the group

$$z_j = \text{argmin}_z \ \|x_1 - z\|^2 + \cdots + \|x_N - z\|^2$$

$$z_j = \text{centroid} = \frac{1}{N} \sum_{i \in \text{Group } j} x_i$$

(the average of all points in group G_j)

the black marker is the representative of a cluster, defined by the point that has the smallest sum of distance to all points in a cluster

Inner product and norm of stacked vectors

inner product of stacked vectors

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}^T \begin{bmatrix} a \\ b \\ c \end{bmatrix} = x^T a + y^T b + z^T c$$

norm of a stacked vector

$$\left\| \begin{bmatrix} x \\ y \\ z \end{bmatrix} \right\|^2 = \|x\|^2 + \|y\|^2 + \|z\|^2$$

norm of a distance

$$\|x - y\|^2 = (x - y)^T (x - y) = \|x\|^2 + \|y\|^2 - 2x^T y$$

Cauchy-Schwarz inequality

for $a, b \in \mathbf{R}^n$

$$|a^T b| \leq \|a\|_2 \|b\|_2$$

example: for $a_1, \ldots, a_n \in \mathbf{R}$ with $a_1 + \cdots + a_n = 1$ show that

$$a_1^2 + a_2^2 + \cdots + a_n^2 \geq \frac{1}{n}$$

CS-inequality can be used to verify the triangle inequality

$$\|a+b\|^2 = \|a\|^2 + 2a^T b + \|b\|^2 \leq \|a\|^2 + 2\|a\|\|b\| + \|b\|^2 = (\|a+b\|)^2$$

angle between vectors: gives a similarity degree of two vectors

$$\cos\theta = \frac{a^T b}{\|a\|\|b\|}$$

Matrix notation

an $m \times n$ matrix A is defined as

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \quad \text{or} \quad A = [a_{ij}]_{m\times n}$$

- a_{ij} are the **elements**, or **coefficients**, or **entries** of A
- set of $m \times n$-matrices is denoted $\mathbf{R}^{m\times n}$
- A has m rows and n columns (m, n are the **dimensions**)
- the (i, j) entry of A is also commonly denoted by A_{ij}
- A is called a **square** matrix if $m = n$

Special matrices

zero matrix: $A = 0$

$$A = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

$a_{ij} = 0$, for $i = 1, \ldots, m, j = 1, \ldots, n$

identity matrix: $A = I$

$$A = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

a square matrix with $a_{ii} = 1, a_{ij} = 0$ for $i \neq j$

diagonal matrix: a square matrix with $a_{ij} = 0$ for $i \neq j$

$$A = \begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & a_n \end{bmatrix}$$

triangular matrix: a square matrix with zero entries in a triangular part

upper triangular

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

lower triangular

$$A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

Multiplication

product of $m \times r$-matrix A with $r \times n$-matrix B:

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ir}b_{rj} = \sum_{k=1}^{r} a_{ik}\, b_{kj}$$

dimensions must be compatible: $\#$ of columns in $A = \#$ of rows in B

- $(AB)_{ij}$ is the dot product of the i^{th} row of A and the j^{th} column of B
- $AB \neq BA$ in general ! (even if the dimensions make sense)
- there are exceptions, *e.g.*, $AI = IA$ for all square A
- $A(B+C) = AB + AC$

Matrix transpose

the transpose of an $m \times n$-matrix A is

$$A^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}$$

properties ✎

- A^T is $n \times m$
- $(A^T)^T = A$
- $(\alpha A + B)^T = \alpha A^T + B^T, \quad \alpha \in \mathbf{R}$
- $(AB)^T = B^T A^T$
- a square matrix A is called **symmetric** if $A = A^T$, *i.e.*, $a_{ij} = a_{ji}$

Block matrix notation

example: 2×2-block matrix A

$$A = \begin{bmatrix} B & C \\ D & E \end{bmatrix}$$

for example, if B, C, D, E are defined as

$$B = \begin{bmatrix} 2 & 1 \\ 3 & 8 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 1 & 7 \\ 1 & 9 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 0 & 1 \end{bmatrix}, \quad E = \begin{bmatrix} -4 & 1 & -1 \end{bmatrix}$$

then A is the matrix

$$A = \begin{bmatrix} 2 & 1 & 0 & 1 & 7 \\ 3 & 8 & 1 & 9 & 1 \\ 0 & 1 & -4 & 1 & -1 \end{bmatrix}$$

note: dimensions of the blocks must be compatible

Column and Row partitions

write an $m \times n$-matrix A in terms of its columns or its rows

$$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} = \begin{bmatrix} b_1^T \\ b_2^T \\ \vdots \\ b_m^T \end{bmatrix}$$

- a_j for $j = 1, 2, \ldots, n$ are the columns of A
- b_i^T for $i = 1, 2, \ldots, m$ are the rows of A

example: $A = \begin{bmatrix} 1 & 2 & 1 \\ 4 & 9 & 0 \end{bmatrix}$

$$a_1 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \quad a_2 = \begin{bmatrix} 2 \\ 9 \end{bmatrix}, \quad a_3 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad b_1^T = \begin{bmatrix} 1 & 2 & 1 \end{bmatrix}, \quad b_2^T = \begin{bmatrix} 4 & 9 & 0 \end{bmatrix}$$

Matrix-vector product

product of $m \times n$-matrix A with n-vector x

$$Ax = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{bmatrix}$$

- dimensions must be compatible: $\#$ columns in $A = \#$ elements in x

if A is partitioned as $A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$, then

$$Ax = a_1x_1 + a_2x_2 + \cdots + a_nx_n$$

- Ax is a linear combination of the column vectors of A
- the coefficients are the entries of x

Product with standard unit vectors

post-multiply with a column vector

$$Ae_k = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{bmatrix} = \text{the } k\text{th column of } A$$

pre-multiply with a row vector

$$\begin{aligned} e_k^T A &= \begin{bmatrix} 0 & 0 & \cdots & 1 & \cdots & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \\ &= \begin{bmatrix} a_{k1} & a_{k2} & \cdots & a_{kn} \end{bmatrix} = \text{the } k\text{th row of } A \end{aligned}$$

Trace

definition: trace of a square matrix A is the sum of the diagonal entries in A

$$\mathbf{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}$$

example:

$$A = \begin{bmatrix} 2 & 1 & 4 \\ 0 & -1 & 5 \\ 3 & 4 & 6 \end{bmatrix}$$

trace of A is $2 - 1 + 6 = 7$

properties ✎

- $\mathbf{tr}(A^T) = \mathbf{tr}(A)$
- $\mathbf{tr}(\alpha A + B) = \alpha\,\mathbf{tr}(A) + \mathbf{tr}(B)$
- $\mathbf{tr}(AB) = \mathbf{tr}(BA)$

Inverse of matrices

definition: a *square* matrix A is called **invertible** or **nonsingular** if there exists B s.t.

$$AB = BA = I$$

- B is called an **inverse** of A
- it is also true that B is invertible and A is an inverse of B
- if no such B can be found A is said to be **singular**

assume A is invertible

- an inverse of A is unique
- the inverse of A is denoted by A^{-1}

Facts about invertible matrices

assume A, B are invertible

facts ✎

- $(\alpha A)^{-1} = \alpha^{-1}A^{-1}$ for nonzero α
- A^T is also invertible and $(A^T)^{-1} = (A^{-1})^T$
- AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$
- $(A+B)^{-1} \neq A^{-1} + B^{-1}$

✌ **Theorem:** for a square matrix A, the following statements are equivalent

1. A is invertible
2. $Ax = 0$ has only the trivial solution $(x = 0)$
3. the reduced echelon form of A is I
4. A is invertible if and only if $\det(A) \neq 0$

Inverse of 2×2 matrices

the matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

is invertible if and only if

$$ad - bc \neq 0$$

and its inverse is given by

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

example:

$$A = \begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix}, \quad A^{-1} = \frac{1}{7} \begin{bmatrix} 3 & -1 \\ 1 & 2 \end{bmatrix}$$

Elementary matrices

Definition: a matrix obtained by performing a *single* row operation on the identity matrix I_n is called an **elementary** matrix

examples:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ k & 0 & 1 \end{bmatrix}$$ add k times the first row to the third row of I_3

$$\begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$$ multiply a nonzero k with the second row of I_2

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$ interchange the second and the third rows of I_3

an elementary matrix is often denoted by E

Inverse operations

row operations on E that produces I and vice versa

$I \to E$	$E \to I$
add k times row i to row j	add $-k$ times row i to row j
multiply row i by $k \neq 0$	multiply row i by $1/k$
interchange row i and j	interchange row i and j

$$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ k & 0 & 1 \end{bmatrix} \implies \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -k & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ k & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$E = \begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix} \implies \begin{bmatrix} 1 & 0 \\ 0 & 1/k \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \implies \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Facts ✌

- every elementary matrix is invertible
- the inverse is also an elementary matrix

from the examples in page 28

$$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ k & 0 & 1 \end{bmatrix} \implies E^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -k & 0 & 1 \end{bmatrix}$$

$$E = \begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix} \implies E^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1/k \end{bmatrix}$$

$$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \implies E^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Row operations by matrix multiplication

assume A is $m \times n$ and E is obtained by performing a row operation on I_m

$$EA = \text{the matrix obtained by performing this same row operation on } A$$

example:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & -1 \\ 1 & 1 & 0 \end{bmatrix}$$

- add -2 times the third row to the second row of A

$$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \quad EA = \begin{bmatrix} 1 & 2 & 3 \\ -2 & -1 & -1 \\ 1 & 1 & 0 \end{bmatrix}$$

- multiply 2 with the first row of A

$$E = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad EA = \begin{bmatrix} 2 & 4 & 6 \\ 0 & 1 & -1 \\ 1 & 1 & 0 \end{bmatrix}$$

- interchange the first and the third rows of A

$$E = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad EA = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 1 & 2 & 3 \end{bmatrix}$$

Inverse via row operations

assume A is invertible

- A is reduced to I by a finite sequence of row operations

$$E_1, E_2, \dots, E_k$$

  such that

$$E_k \cdots E_2 E_1 A = I$$

- the reduced echelon form of A is I
- the inverse of A is therefore given by the product of elementary matrices

$$A^{-1} = E_k \cdots E_2 E_1$$

Example

write the augmented matrix $\begin{bmatrix} A & | & I \end{bmatrix}$

$$\left[\begin{array}{ccc|ccc} 2 & 4 & 3 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 & 1 & 0 \\ 1 & 0 & 4 & 0 & 0 & 1 \end{array}\right]$$

and apply row operations until the left side is reduced to I

$$\begin{array}{c} -2R_2 + R_1 \rightarrow R_1 \\ -R_2 + R_3 \rightarrow R_3 \end{array} \qquad \left[\begin{array}{ccc|ccc} 0 & 0 & 1 & 1 & -2 & 0 \\ 1 & 2 & 1 & 0 & 1 & 0 \\ 0 & -2 & 3 & 0 & -1 & 1 \end{array}\right]$$

$$R_1 \leftrightarrow R_2 \qquad \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & -2 & 0 \\ 0 & -2 & 3 & 0 & -1 & 1 \end{array}\right]$$

$$-3R_2 + R_3 \rightarrow R_3 \qquad \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & -2 & 0 \\ 0 & -2 & 0 & -3 & 5 & 1 \end{array}\right]$$

$$R_3/(-2) \to R_3 \quad \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & -2 & 0 \\ 0 & 1 & 0 & \frac{3}{2} & -\frac{5}{2} & -\frac{1}{2} \end{array}\right]$$

$$R_2 \leftrightarrow R_3 \quad \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & \frac{3}{2} & -\frac{5}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & 1 & -2 & 0 \end{array}\right]$$

$$-2R_2 + R_1 \to R_1 \quad \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & -3 & 6 & 1 \\ 0 & 1 & 0 & \frac{3}{2} & -\frac{5}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & 1 & -2 & 0 \end{array}\right]$$

$$-R_3 + R_1 \to R_1 \quad \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -4 & 8 & 1 \\ 0 & 1 & 0 & \frac{3}{2} & -\frac{5}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & 1 & -2 & 0 \end{array}\right]$$

the inverse of A is

$$\begin{bmatrix} -4 & 8 & 1 \\ \frac{3}{2} & -\frac{5}{2} & -\frac{1}{2} \\ 1 & -2 & 0 \end{bmatrix}$$

Inverse of diagonal matrix

$$A = \begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & a_n \end{bmatrix}$$

a diagonal matrix is invertible iff the diagonal entries are all nonzero

$$a_{ii} \neq 0, \quad i = 1, 2, \ldots, n$$

the inverse of A is given by

$$A^{-1} = \begin{bmatrix} 1/a_1 & 0 & \cdots & 0 \\ 0 & 1/a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 1/a_n \end{bmatrix}$$

the diagonal entries in A^{-1} are the inverse of the diagonal entries in A

Inverse of triangular matrix

upper triangular

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

$a_{ij} = 0$ for $i \geq j$

lower triangular

$$A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

$a_{ij} = 0$ for $i \leq j$

a triangular matrix is invertible iff the diagonal entries are all nonzero

$$a_{ii} \neq 0, \quad \forall i = 1, 2, \ldots, n$$

- product of lower (upper) triangular matrices is lower (upper) triangular
- the inverse of a lower (upper) triangular matrix is lower (upper) triangular

Inverse of symmetric matrix

symmetric matrix: $A = A^T$

✎

- for any square matrix A, AA^T and A^TA are always symmetric
- if A is symmetric and invertible, then A^{-1} is symmetric
- if A is invertible, then AA^T and A^TA are also invertible

for a general A, the inverse of A^T is $(A^{-1})^T$ please verify ✎

Determinants

the determinant is a *scalar value* associated with a square matrix A

commonly denoted by $\det(A)$ or $|A|$

determinants of 2×2 matrices:

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

determinants of 3×3 matrices: let $A = \{a_{ij}\}$

$$\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}$$
$$- (a_{31}a_{22}a_{13} + a_{32}a_{23}a_{11} + a_{33}a_{21}a_{12})$$

How to find determinants

for a square matrix of any order, it can be computed by

- cofactor expansion
- performing elementray row operations

Minor and Cofactor

Minor of entry a_{ij}: denoted by M_{ij}

- the determinant of the resulting submatrix after deleting the ith row and jth column of A

Cofactor of entry a_{ij}: denoted by C_{ij}

- $C_{ij} = (-1)^{(i+j)} M_{ij}$

example:

$$A = \begin{bmatrix} 3 & 1 & -2 \\ 5 & 0 & 2 \\ 1 & -1 & 2 \end{bmatrix}, \quad M_{23} = \begin{vmatrix} 3 & 1 \\ 1 & -1 \end{vmatrix} = -4, \quad C_{23} = (-1)^{(2+3)} M_{23} = 4$$

Determinants by Cofactor Expansion

Theorem: the determinant of an $n \times n$-matrix A is given by

$$\begin{aligned}\det(A) &= a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj} \\ \det(A) &= a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}\end{aligned}$$

regardless of which row or column of A is chosen

example: pick the first row to compute $\det(A)$

$$A = \begin{bmatrix} 3 & 1 & -2 \\ 5 & 0 & 2 \\ 1 & -1 & 2 \end{bmatrix}, \quad \det(A) = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13}$$

$$\begin{aligned}\det(A) &= 3(-1)^2 \begin{vmatrix} 0 & 2 \\ -1 & 2 \end{vmatrix} + 1(-1)^3 \begin{vmatrix} 5 & 2 \\ 1 & 2 \end{vmatrix} - 2(-1)^4 \begin{vmatrix} 5 & 0 \\ 1 & -1 \end{vmatrix} \\ &= 3(1)(2) + (-1)(8) - 2(1)(-5) = 8\end{aligned}$$

Basic properties of determinants

✌ let A, B be any square matrices

1. $\det(A) = \det(A^T)$
2. if A has a row of zeros or a column of zeros, then $\det(A) = 0$
3. $\det(\alpha A) = \alpha^n \det(A)$, $\quad \alpha \neq 0$
4. If A has two rows (columns) that are equal, then $\det(A) = 0$
5. $\det(A + B) \neq \det(A) + \det(B)$!
6. $\det(AB) = \det(A)\det(B)$
7. $\det(A^{-1}) = 1/\det(A)$
8. A is invertible if and only if $\det(A) \neq 0$

Basic properties of determinants

suppose the following is true

- A and B are equal except for the entries in their kth row (column)
- C is defined as that matrix identical to A and B except that its kth row (column) is the sum of the kth rows (columns) of A and B

then we have

$$\det(C) = \det(A) + \det(B)$$

example:

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 1 & 2 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 3 & 0 & 2 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 4 & 2 & 1 \end{bmatrix}$$

$$\det(A) = 0, \quad \det(B) = -1, \quad \det(C) = -1$$

Determinants of special matrices

- the determinant of a diagonal or triangular matrix is given by the product of the diagonal entries
- $\det(I) = 1$

(these properties can be proved from the def. of cofactor expansion)

Determinants under row operations

- multiply k to a row or a column

$$\begin{vmatrix} ka_{11} & ka_{12} & ka_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = k \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

- interchange between two rows or two columns

$$\begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = - \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

- add k times the ith row (column) to the jth row (column)

$$\begin{vmatrix} a_{11} + ka_{21} & a_{12} + ka_{22} & a_{13} + ka_{23} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

Example

B is obtained by performing the following operations on A

$$R_2 + 3R_1 \to R_2, \quad R_3 \leftrightarrow R_1, \quad -4R_1 \to R_1$$

$$A = \begin{bmatrix} 2 & 3 & -2 \\ 3 & 1 & 0 \\ -3 & -3 & 3 \end{bmatrix} \quad \Longrightarrow \quad \det(B) = (-4) \cdot (-1) \cdot 1 \cdot \det(A)$$

the changes of det. under elementary operations lead to obvious facts ✎

- $\det(\alpha A) = \alpha^n \det(A), \quad \alpha \neq 0$
- If A has two rows (columns) that are equal, then $\det(A) = 0$

Determinants of elementary matrices

let B be obtained by performing a row operation on A then

$$B = EA \quad \text{and} \quad \det(B) = \det(EA)$$

$$E = \begin{bmatrix} k & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \det(B) = k\det(A) \quad (\det(E) = k)$$

$$E = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \det(B) = -\det(A) \quad (\det(E) = -1)$$

$$E = \begin{bmatrix} 1 & k & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \det(B) = \det(A) \quad (\det(E) = 1)$$

conclusion: $\det(EA) = \det(E)\det(A)$

Determinants of product and inverse

✌ let A, B be $n \times n$ matrices

- A is invertible if and only if $\det(A) \neq 0$
- if A is invertible, then $\det(A^{-1}) = 1/\det(A)$
- $\det(AB) = \det(A)\det(B)$

Adjugate formula

the adjugate of A is the transpose of the matrix of cofactors from A

$$\mathrm{adj}(A) = \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix}$$

if A is invertible then

$$A^{-1} = \frac{1}{\det(A)} \mathrm{adj}(A)$$

Proof.

- the cofactor expansion using the cofactors from different row is zero

$$a_{i1}C_{k1} + a_{i2}C_{k2} + \ldots + a_{in}C_{kn} = 0, \quad \text{for } i \neq k$$

- $A\,\mathrm{adj}(A) = \det(A) \cdot I$

Cramer's rule

consider a linear system $Ax = b$ when A is **square**

if A is invertible then the solution is unique and given by

$$x = A^{-1}b$$

each component of x can be calculated by using the Cramer's rule

Cramer's rule

$$x_1 = \frac{|A_1|}{|A|}, \quad x_2 = \frac{|A_2|}{|A|}, \quad \ldots \quad , \quad x_n = \frac{|A_n|}{|A|}$$

where A_j is the matrix obtained by replacing b in the jth column of A

(its proof is left as an exercise)

Example

$$A = \begin{bmatrix} 3 & 1 & -2 \\ 5 & 0 & 2 \\ 1 & -1 & 2 \end{bmatrix}, \quad b = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$$

since $\det(A) = 8$, A is invertible and the solution is

$$x = A^{-1}b = \frac{1}{8}\begin{bmatrix} 2 & 0 & 2 \\ -8 & 8 & -16 \\ -5 & 4 & -5 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ -5 \\ -2 \end{bmatrix}$$

using Cramer's rule gives

$$x_1 = \frac{1}{8}\begin{vmatrix} 2 & 1 & -2 \\ 1 & 0 & 2 \\ 2 & -1 & 2 \end{vmatrix}, \quad x_2 = \frac{1}{8}\begin{vmatrix} 3 & 2 & -2 \\ 5 & 1 & 2 \\ 1 & 2 & 2 \end{vmatrix}, \quad x_3 = \frac{1}{8}\begin{vmatrix} 3 & 1 & 2 \\ 5 & 0 & 1 \\ 1 & -1 & 2 \end{vmatrix}$$

which yields

$$x_1 = 1, \quad x_2 = -5, \quad x_3 = -2$$

Pseudo-inverse (Penrose Theorem)

one can have a notion of 'inverse' for a non-square matrix

Penrose's Theorem: given $A \in \mathbf{R}^{m \times n}$, there is exactly one $n \times m$ matrix B such that

1. $ABA = A$ and $BAB = B$
2. both AB and BA are symmetric

definition: the **pseudo inverse** of $A \in \mathbf{R}^{m \times n}$ is the unique $n \times m$ matrix $A^\dagger$ such that

1. $AA^\dagger A = A$ and $A^\dagger AA^\dagger = A^\dagger$
2. both $AA^\dagger$ and $A^\dagger A$ are symmetric

Pseudo-inverse

consider a full rank matrix $A \in \mathbf{R}^{m \times n}$ in three cases

- **tall matrix:** A is full rank $\Leftrightarrow$ columns of A are LI $\Leftrightarrow$ A^TA is invertible

$$((A^TA)^{-1}A^T)A = (A^TA)^{-1}(A^TA) = I$$

  the **pseudo-inverse** of A (or left-inverse) is $A^\dagger = (A^TA)^{-1}A^T$

- **wide matrix:** A is full rank $\Leftrightarrow$ row of A are LI $\Leftrightarrow$ AA^T is invertible

$$A(A^T(AA^T)^{-1}) = (AA^T)(AA^T)^{-1} = I$$

  the **pseudo-inverse** of A (or right-inverse) is $A^\dagger = A^T(AA^T)^{-1}$

- **square matrix:** A is full rank $\Leftrightarrow$ A is invertible and both formula of pseudo-inverses reduce to the ordinary inverse A^{-1}

✎ the pseudo inverses of the three cases have the same dimension ?

Example

$$A = \begin{bmatrix} 0 & 2 & 1 \\ -2 & 1 & -2 \end{bmatrix}, \quad A^\dagger = A^T(AA^T)^{-1} = \begin{bmatrix} 0 & -2/9 \\ 2/5 & 1/9 \\ 1/5 & -2/9 \end{bmatrix}$$

$$A = \begin{bmatrix} -2 & -1 \\ 2 & -1 \\ -1 & 0 \end{bmatrix}, \quad A^\dagger = (A^TA)^{-1}A^T = \begin{bmatrix} -2/9 & 2/9 & 1/9 \\ -1/2 & -1/2 & 0 \end{bmatrix}$$

however, when rentangular A has low rank, we can use SVD to find the pseudo inverse

Softwares (MATLAB)

1. `eye(n)` creates an identity matrix of size n
2. `inv(A)` finds the inverse of A (not used for large dimension)
3. `A\eye(n)` finds the inverse of a square matrix A
4. `pinv(A)` gives a pseudoinverse of A, denoted by $A^\dagger$
   - if A is square, a pseudoinverse is the inverse of A
   - if A is tall, $A^\dagger = (A^TA)^{-1}A^T$ is a left inverse of A
   - if A is fat, $A^\dagger = A^T(AA^T)^{-1}$ is a right inverse of A
5. `x = pinv(A)*b` solves the linear system $Ax = b$
   - if A is square, $x = A^{-1}b$
   - if A is tall, x is the solution to the least-square problem: minimize $\|Ax - b\|_2$
   - if A is fat, x is the least-norm solution that satisfies $Ax = b$
6. `det(A)` finds the determinant of A

Softwares (Python)

1. `numpy.eye` creates an identity matrix
2. `numpy.linalg.inv` finds the inverse of a square matrix A
3. `numpy.linalg.pinv` gives a pseudoinverse of A
4. `numpy.linalg.det` find the determinants of A

References

1. W.K. Nicholson, *Linear Algebra with Applications*, McGraw-Hill, 2006
2. S. Boyd and L. Vandenberghe, *Introduction to Applied Linear Algebra: Vectors, Matrices, and Least squares*, Cambridge, 2018
3. H.Anton and C. Rorres, *Elementary Linear Algebra*, John Wiley, 2011