Vectors and matrices: review

- vectors
- vector operations
- matrices
- matrix operations
- determinant and inverse
- linear equations

Vector

a *column vector* is an ordered list of scalars or numbers, represented by:

$$a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \quad \text{or} \quad a = (a_1, \ldots, a_n)$$

- a_i is the the ith *entry* (or element, coefficient, component) of vector a
- i is the *index* of the ith element a_i
- number of elements n is the *size* (*length, dimension*) of the vector
- a vector of size n is called an n-vector; $\mathbb{R}^n$ denote the set of real vectors of size n
- two vectors a, b are equal, denoted $a = b$, if the have the same size and corresponding entries are all equal

Example

$$a = \begin{bmatrix} 1 \\ -2 \\ 3.3 \\ 0.3 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ -2 \\ 3.3 \end{bmatrix}$$

- a is a 4-vector, b is a 3-vector
- third component of a is $a_3 = 3.3$
- a_5, b_4 does not make sense
- a is not equal to b since their dimension is different

Row vector and transpose

an *row* vector b of size n with entries $b_1, \ldots, b_n$ has the form:

$$b = \begin{bmatrix} b_1 & b_2 & \ldots & b_n \end{bmatrix}$$

- all vectors are column vectors unless otherwise stated
- other notation exists, *e.g.*, $b = [b_1, b_2, \ldots, b_n]$

Transpose: the *transpose* of an n-column vector a is the row vector a^T:

$$a^T = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}^T = \begin{bmatrix} a_1 & a_2 & \ldots & a_n \end{bmatrix}$$

- $(.)^T$ is transpose operation
- $(a^T)^T = a$ (transpose of row vector is a column vector)

Block vectors, subvectors

Stacking

- vectors can be *stacked* (*concatenated*) to create larger vectors
- stacking vectors b, c, d of size m, n, p gives an $(m+n+p)$-vector

$$a = \begin{bmatrix} b \\ c \\ d \end{bmatrix} = (b, c, d) = \left(b_1, \ldots, b_m, c_1, \ldots, c_n, d_1, \ldots, d_p\right)$$

- we say that b, c, and d are *subvectors* or *slices* of a
- example: if $b = 1$, $c = (2, -1)$, $d = (4, 2, 7)$, then $(b, c, d) = (1, 2, -1, 4, 2, 7)$

Subvectors slicing

- colon (:) notation can be used to define subvectors (slices) of a vector
- for vector a, we define $a_{r:s} = (a_r, \ldots, a_s)$
- example: if $a = (1, -1, 2, 0, 3)$, then $a_{2:4} = (-1, 2, 0)$

Special vectors

Zero vector and ones vector

$$0 = (0, 0, \ldots, 0), \quad \mathbf{1} = (1, 1, \ldots, 1)$$

size follows from context (if not, we add a subscript and write $0_n, \mathbf{1}_n$)

Unit vectors

- there are n *unit vectors* of size n, written $e_1, e_2, \ldots, e_n$

$$(e_i)_j = \begin{cases} 1 & j = i \\ 0 & j \neq i \end{cases}$$

- the ith unit vector is zero except its ith element which is 1
- example: for $n = 3$,

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

- the size of e_i follows from context (or should be specified explicitly)

Outline

- vectors
- **vector operations**
- matrices
- matrix operations
- determinant and inverse
- linear equations

Addition and subtraction

for n-vectors a and b,

$$a+b=\begin{bmatrix} a_1+b_1 \\ a_2+b_2 \\ \vdots \\ a_n+b_n \end{bmatrix}, \quad a-b=\begin{bmatrix} a_1-b_1 \\ a_2-b_2 \\ \vdots \\ a_n-b_n \end{bmatrix}$$

Example

$$\begin{bmatrix} 0 \\ 7 \\ 3 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 9 \\ 3 \end{bmatrix}$$

Properties: for vectors a, b of equal size

- commutative: $a+b=b+a$
- associative: $a+(b+c)=(a+b)+c$

Scalar-vector multiplication

for vector $a \in \mathbb{R}^n$ and scalar β:

$$\beta a = (\beta a_1, \beta a_2, \ldots, \beta a_n)$$

Properties: for vectors a, b of equal size, scalars β, γ

- commutative: $\beta a = a\beta$
- associative: $(\beta\gamma)a = \beta(\gamma a)$, we write as $\beta\gamma a$
- distributive with scalar addition: $(\beta + \gamma)a = \beta a + \gamma a$
- distributive with vector addition: $\beta(a + b) = \beta a + \beta b$

Linear combination

a *linear combination* of vectors $a_1, \ldots, a_m$ is a sum of scalar-vector products

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

- scalars $\beta_1, \ldots, \beta_m$ are the *coefficients* of the linear combination
- example: any n-vector b can be written as

$$b = b_1 e_1 + \cdots + b_n e_n$$

Inner product

if a and b are n-vectors, then the *inner product* or *dot product* of a, b is

$$\langle a, b\rangle = a^T b = a_1 b_1 + \cdots + a_n b_n$$

- a scalar
- for example

$$\begin{bmatrix} 1 \\ -2 \\ 0.5 \end{bmatrix}^T \begin{bmatrix} -2 \\ 6 \\ 4 \end{bmatrix} = (2)(-2) + (-2)(6) + (0.5)(4) = -14$$

Outline

- vectors
- vector operations
- **matrices**
- matrix operations
- determinant and inverse
- linear equations

Matrix

a *matrix* is a rectangular array of scalars or elements written as

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

- numbers in array are the *elements* (*entries*, *coefficients*, *components*)
- a horizontal set of elements is called a *row* and a vertical set is called a *column*
- a_{ij} is the i, j element of A (i is row index, j is column index)
- *size* (*dimensions*) of the matrix is $m \times n =$ (#rows) $\times$ (#columns)
- a matrix of size $m \times n$ is called an $m \times n$ matrix
- $\mathbb{R}^{m \times n}$ is set of $m \times n$ matrices with real elements
- elements a_{ii} are called principal or *main diagonal* of the matrix

Example

$$A = \begin{bmatrix} 0 & 1 & -2.3 & 0.1 \\ 1.3 & 4 & -0.1 & 0 \\ 4.1 & -1 & 0 & 1.7 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & -3 \\ 12 & 0 \end{bmatrix}$$

- A is a 3×4 matrix, B is 2×2
- the matrix A has four columns; B has two rows
- for example, $a_{23} = -0.1$, $a_{22} = 4$, but a_{41} is meaningless
- in A, the row index of the entry with value -2.3 is 1; its column index is 3

Matrix shapes

Scalar: a 1×1 matrix is a scalar

Row and column vectors

- a $1 \times n$ matrix is called a row vector
- an $n \times 1$ matrix is called a column vector (or just vector)

Tall, wide, square matrices: an $m \times n$ matrix is

- tall if $m > n$
- wide if $m < n$
- square if $m = n$

Matrix equality

$A = B$ means:

- A and B have the same size
- the corresponding entries are equal

for example,

- $$\begin{bmatrix} -2 \\ 3.3 \end{bmatrix} \neq \begin{bmatrix} -2 & -3.3 \end{bmatrix}$$

  since the dimensions don't agree

- $$\begin{bmatrix} -2 \\ 3.3 \end{bmatrix} \neq \begin{bmatrix} -2 \\ 3.1 \end{bmatrix}$$

  since the 2nd components don't agree

Columns and rows

an $m \times n$ matrix can be viewed as a matrix with row/column vectors

Columns representation

$$A = [a_1 \; a_2 \; \cdots \; a_n] \qquad a_j = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

each a_j is an m-vector (the jth column of A)

Rows representation

$$A = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \qquad b_i = [a_{i1} \; \cdots \; a_{in}]$$

each b_i is a $1 \times n$ row vector (the ith row of A)

Block matrix and submatrices

- a *block* matrix is a rectangular array of matrices
- elements in the array are the *blocks* or *submatrices* of the block matrix

Example: a 2×2 block matrix

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

- submatrices can be referred to by their block row and column (C is $1,2$ block of A)
- dimensions of the blocks must be compatible
- if the blocks are

$$B = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 2 & 3 \\ 5 & 4 & 7 \end{bmatrix}, \quad D = \begin{bmatrix} 1 \end{bmatrix}, \quad E = \begin{bmatrix} -1 & 6 & 0 \end{bmatrix}$$

then

$$A = \begin{bmatrix} 2 & 0 & 2 & 3 \\ 1 & 5 & 4 & 7 \\ 1 & -1 & 6 & 0 \end{bmatrix}$$

Slice of matrix

$$A_{p:q,r:s} = \begin{bmatrix} a_{pr} & a_{p,r+1} & \cdots & a_{ps} \\ a_{p+1,r} & a_{p+1,r+1} & \cdots & a_{p+1,s} \\ \vdots & \vdots & & \vdots \\ a_{qr} & a_{q,r+1} & \cdots & a_{qs} \end{bmatrix}$$

- an $(q - p + 1) \times (s - r + 1)$ matrix
- obtained by extracting from A elements in rows p to q and columns r to s
- from last page example, we have

$$A_{2:3,3:4} = \begin{bmatrix} 1 & 4 \\ 5 & 4 \end{bmatrix}$$

Special matrices

Zero matrix

- matrix with $a_{ij} = 0$ for all i, j
- notation: 0 or $0_{m \times n}$ (if dimension is not clear from context)
- example:

$$0_{2\times 3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Identity matrix

- square matrix with $a_{ij} = 1$ if $i = j$ and $a_{ij} = 0$ if $i \neq j$
- notation: I or I_n (if dimension is not clear from context)
- columns of I_n are unit vectors $e_1, e_2, \ldots, e_n$; for example,

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} e_1 & e_2 & e_3 \end{bmatrix}$$

Structured matrices

matrices with special patterns or structure arise in many applications

Diagonal matrix

- square with $a_{ij} = 0$ for $i \neq j$
- represented as $A = \mathrm{diag}(a_1, \ldots, a_n)$ where a_i are diagonal elements

$$\mathrm{diag}(0.2, -3, 1.2) = \begin{bmatrix} 0.2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 1.2 \end{bmatrix}$$

Lower triangular matrix: square with $a_{ij} = 0$ for $i < j$

$$\begin{bmatrix} 4 & 0 & 0 \\ 3 & -1 & 0 \\ -1 & 5 & -2 \end{bmatrix}, \quad \begin{bmatrix} 4 & 0 & 0 \\ 0 & -1 & 0 \\ -1 & 0 & -2 \end{bmatrix}$$

Upper triangular matrix: square with $a_{ij} = 0$ for $i > j$

(a triangular matrix is **unit** upper/lower triangular if $a_{ii} = 1$ for all i)

Transpose of a matrix

transpose of an $m \times n$ matrix A is the $n \times m$ matrix $(A^T)_{ij} = a_{ji}$:

- example:

$$\begin{bmatrix} 0 & 4 & 7 \\ 0 & 3 & 1 \end{bmatrix}^T = \begin{bmatrix} 0 & 0 \\ 4 & 3 \\ 7 & 1 \end{bmatrix}$$

- rows and columns of A are transposed in A^T

Properties

- $\left(A^T\right)^T = A$
- the transpose of a block matrix (shown for a 2×2 block matrix)

$$\left[\begin{array}{cc} A & B \\ C & D \end{array} \right]^T = \left[\begin{array}{cc} A^T & C^T \\ B^T & D^T \end{array} \right]$$

  - A, B, C, and D are matrices with compatible sizes
  - concept holds for any number of blocks

Symmetric matrices

a square matrix A is *symmetric* if

$$A = A^T$$

- $a_{ij} = a_{ji}$
- examples:

$$\begin{bmatrix} 4 & 3 & -2 \\ 3 & -1 & 5 \\ -2 & 5 & 0 \end{bmatrix}, \quad \begin{bmatrix} 4+3j & 3-2j & 0 \\ 3-2j & -j & -2j \\ 0 & -2j & 3 \end{bmatrix}$$

Outline

- vectors
- vector operations
- matrices
- **matrix operations**
- determinant and inverse
- linear equations

Matrix addition

sum of two $m \times n$ matrices A and B

$$A + B = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{bmatrix}$$

Example

$$\begin{bmatrix} 0 & 4 & 7 \\ 0 & 3 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 2 & 2 \\ 3 & 0 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 6 & 9 \\ 3 & 3 & 5 \end{bmatrix}$$

matrix subtraction is similar:

$$\begin{bmatrix} 1 & 6 \\ 9 & 3 \end{bmatrix} - I = \begin{bmatrix} 0 & 6 \\ 9 & 2 \end{bmatrix}$$

(here we had to figure out that I must be 2×2)

Properties of matrix addition

- *commutativity:* $A + B = B + A$
- *associativity:* $(A + B) + C = A + (B + C)$, , so we can write as $A + B + C$
- *addition with zero matrix:* $A + 0 = 0 + A = A$
- *transpose of sum:* $(A + B)^T = A^T + B^T$

Scalar-matrix multiplication

scalar-matrix product of $m \times n$ matrix A with scalar β is entry-wise

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

for example,

$$(-2)\begin{bmatrix} 1 & 6 & 9 \\ 3 & 6 & 0 \end{bmatrix} = \begin{bmatrix} -2 & -12 & -18 \\ -6 & -12 & 0 \end{bmatrix}$$

Properties: for matrices A, B, scalars β, γ

- *transposition:* $(\beta A)^T = \beta A^T$
- associativity: $(\beta\gamma)A = \beta(\gamma A)$
- *distributivity:* $(\beta + \gamma)A = \beta A + \gamma A$ and $\beta(A + B) = \beta A + \beta B$
- $0 \cdot A = 0$; $1 \cdot A = A$

Matrix-vector product

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

$$Ax = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} b_1^T x \\ \vdots \\ b_m^T x \end{bmatrix}$$

- b_i^T is ith row of A
- dimensions must be compatible (number of columns of A equals the size of x)
- Ax is a linear combination of the columns of A:

$$Ax = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 a_1 + x_2 a_2 + \cdots + x_n a_n$$

each a_i is an m-vector (ith column of A)

Properties of matrix-vector multiplication

for matrix A, vectors u, v and scalar α

- *associativity:* $(\alpha A)u = A(\alpha u) = \alpha(Au)$ (we write αAu)
- *distributivity:* $A(u + v) = Au + Av$ and $(A + A)u = Au + Au$
- *transposition:* $(Au)^T = u^T A^T$

General examples

- $0x = 0$, *i.e.*, multiplying by zero matrix gives zero
- $Ix = x$, *i.e.*, multiplying by identity matrix does nothing
- inner product $a^T b$ is matrix-vector product of $1 \times n$ matrix a^T and n-vector b
- $Ae_j = a_j$, the jth column of A [$(A^T e_i)^T = e_i^T A$ is ith row]
- the m-vector $A\mathbf{1}$ is the sum of the columns of A

Matrix multiplication

product of $m \times n$ matrix A and $n \times p$ matrix B

$$C = AB$$

is the $m \times p$ matrix with i, j element

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

- to get c_{ij} : move along i th row of A, j th column of B
- dimensions must be compatible:

$$\#\text{columns in } A = \#\text{rows in } B$$

- to find i, j entry of the product $C = AB$, you need the ith row of A and the jth column of B
  - form product of corresponding entries, *e.g.*, third component of ith row of A and third component of jth column of B
  - add up all the products

Examples

example 1:

$$\begin{bmatrix} 1 & 6 \\ 9 & 3 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} -6 & 11 \\ -3 & -3 \end{bmatrix}$$

for example, to get 1, 1 entry of product:

$$c_{11} = a_{11}b_{11} + a_{12}b_{21} = (1)(0) + (6)(-1) = -6$$

example 2:

$$\begin{bmatrix} -1.5 & 3 & 2 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} -1 & -1 \\ 0 & -2 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 3.5 & -4.5 \\ -1 & 1 \end{bmatrix}$$

Special cases of matrix multiplication

- scalar-vector product (with scalar on right!) $x\alpha$
- matrix-vector multiplication Ax
- outer product of m-vector a and n-vector b

$$ab^T = \begin{bmatrix} a_1 b_1 & a_1 b_2 & \cdots & a_1 b_n \\ a_2 b_1 & a_2 b_2 & \cdots & a_2 b_n \\ \vdots & \vdots & & \vdots \\ a_m b_1 & a_m b_2 & \cdots & a_m b_n \end{bmatrix}$$

- multiplication by identity $AI = A$ and $IA = A$

Matrix powers

- if matrix A is square, then product AA makes sense, and is denoted A^2
- more generally, k copies of A multiplied together gives A^k:

$$A^k = \underbrace{AA \cdots A}_{k}$$

  by convention we set $A^0 = I$
- (non-integer powers like $A^{1/2}$ are tricky — that's an advanced topic)
- we have $A^k A^l = A^{k+l}$

Properties of matrix-matrix product

- associativity: $(AB)C = A(BC)$ so we write ABC
- associativity with scalar multiplication: $(\gamma A)B = \gamma(AB) = \gamma AB$
- distributivity with sum:

$$A(B+C) = AB + AC, \quad (A+B)C = AC + BC$$

- transpose of product: $(AB)^T = B^T A^T$
- **not** commutative: $AB \neq BA$ in general; for example,

$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \neq \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

  order of multiplication is important
- $0A = 0$, $A0 = 0$ (here 0 can be scalar, or a compatible matrix)
- $IA = A$, $AI = A$

Product of block matrices

block-matrices can be multiplied as regular matrices

Example: product of two 2×2 block matrices

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} W & Y \\ X & Z \end{bmatrix} = \begin{bmatrix} AW + BX & AY + BZ \\ CW + DX & CY + DZ \end{bmatrix}$$

if the dimensions of the blocks are compatible

Column and row representations

Column representation

- A is $m \times p$, B is $p \times n$ with columns b_i

$$AB = A \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix} = \begin{bmatrix} Ab_1 & Ab_2 & \cdots & Ab_n \end{bmatrix}$$

- so AB is ‘batch’ multiply of A times columns of B

Row representation

- with a_i^T the rows of A

$$AB = \begin{bmatrix} a_1^T B \\ a_2^T B \\ \vdots \\ a_m^T B \end{bmatrix} = \begin{bmatrix} \left(B^T a_1\right)^T \\ \left(B^T a_2\right)^T \\ \vdots \\ \left(B^T a_m\right)^T \end{bmatrix}$$

- row i is $(B^T a_i)^T$

Outline

- vectors
- vector operations
- matrices
- matrix operations
- **determinant and inverse**
- linear equations

Matrix determinant

if A is an $n \times n$ matrix, then the ij**th submatrix** of A, denoted by A_{ij}, is the $(m-1) \times (m-1)$ obtained by deleting row i and column j of A; for example,

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}, \qquad A_{12} = \begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix}, \quad A_{32} = \begin{bmatrix} 1 & 3 \\ 4 & 6 \end{bmatrix}$$

Determinant: the *determinant* of a matrix is computed a follows; pick any value of i ($i = 1, 2, \ldots, n$) and compute

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} \det(A_{ij}) a_{ij}$$

- $\det(A_{ij})$ is called the *minor* of element a_{ij}
- $(-1)^{i+j} \det(A_{ij})$ is called the *cofactor* of element a_{ij}

Example

a) for a scalar matrix $A = [a_{11}]$, we have $\det(A) = a_{11}$

b) for a 2×2 matrix, the determinant is

$$\det(A) = \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{21}a_{12}$$

c) for the matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$

– we have for $i = 1$

$$A_{11} = \begin{bmatrix} 5 & 6 \\ 8 & 9 \end{bmatrix}, \quad A_{12} = \begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix}, \quad A_{13} = \begin{bmatrix} 4 & 5 \\ 7 & 8 \end{bmatrix}$$

– thus, the determinant is

$$\begin{aligned} \det(A) &= (-1)^2 a_{11} \det(A_{11}) + (-1)^3 a_{12} \det(A_{12}) + (-1)^4 a_{13} \det(A_{13}) \\ &= a_{11} \det(A_{11}) - a_{12} \det(A_{12}) + a_{13} \det(A_{13}) \\ &= 1(-3) - 2(-6) + 3(-3) = 0 \end{aligned}$$

Properties of determinants

- *multiplication of a single row/column by a constant:* if a single row or column of a matrix, A, is multiplied by a constant, k, forming the matrix, $\tilde{A}$, then

$$\det \tilde{A} = k \det A$$

- *multiplication of all elements by a constant*

$$\det(kA) = k^n \det A$$

- *transpose*

$$\det A^T = \det A$$

- *determinant of the product of square matrices*

$$\det AB = \det A \det B$$

$$\det AB = \det BA$$

Inverse

the matrix A^{-1} is said to be the **inverse** of the $n \times n$ matrix A if it satisfies

$$AA^{-1} = A^{-1}A = I_n$$

- if A has an inverse, it is called *invertible* or *nonsingular*
- invertible matrices must be *square*
- for a non-zero scalar a, the inverse is the number x such that $ax = 1$, which we denote by $x = 1/a = a^{-1}$
- a *square matrix* A is invertible if and only if $\det(A) \neq 0$
- if A doesn't have an inverse, it's called *singular* or *noninvertible*

Example

a) the identity matrix I is invertible, with inverse $I^{-1} = I$ since $(I)I = I$

b) any 2×2 matrix A is invertible if and only if $a_{11}a_{22} \neq a_{12}a_{21}$, with inverse

$$A^{-1} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

for example

$$\begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}^{-1} = \frac{1}{3} \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix}$$

the matrix

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

does not have an inverse; let's see why:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix} = \begin{bmatrix} a - 2b & -a + 2b \\ c - 2d & -c + 2d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

. . . but you can't have $a - 2b = 1$ and $-a + 2b = 0$

c) a diagonal matrix

$$D = \begin{bmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & d_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{nn} \end{bmatrix}$$

is invertible if and only if $d_{ii} \neq 0$ for $i = 1, \ldots, n$, and

$$D^{-1} = \begin{bmatrix} 1/d_{11} & 0 & \cdots & 0 \\ 0 & 1/d_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/d_{nn} \end{bmatrix}$$

Properties of inverse

- $(A^{-1})^{-1} = A$, *i.e.*, inverse of inverse is original matrix (assuming A is invertible)
- $(AB)^{-1} = B^{-1}A^{-1}$ (assuming A, B are invertible)
- $(A^T)^{-1} = (A^{-1})^T$ (assuming A is invertible)
- $(\alpha A)^{-1} = (1/\alpha)A^{-1}$ (assuming A invertible, $\alpha \neq 0$)
- if $y = Ax$, where $x \in \mathbb{R}^n$ and A is invertible, then $x = A^{-1}y$:

$$A^{-1}y = A^{-1}Ax = Ix = x$$

- let A be a square invertible matrix, then

$$(A^p)^{-1} = (A^{-1})^p$$

  for any integer p

Outline

- vectors
- vector operations
- matrices
- matrix operations
- determinant and inverse
- **linear equations**

Linear functions

- $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ means f is a function mapping n-vectors to m-vectors
- value is an m-vector $f(x) = (f_1(x), \ldots, f_m(x))$
- example: $f(x) = (x_1^2, x_2 - x_1, x_2)$ is $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$

Linear function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is *linear* if it satisfies:

- *scaling:* for any n-vector x, any scalar α, $f(\alpha x) = \alpha f(x)$
- *superposition:* for any n-vectors u and v, $f(u + v) = f(u) + f(v)$

Example: $f(x) = y$, where

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad y = \begin{bmatrix} x_3 - 2x_1 \\ 3x_1 - 2x_2 \end{bmatrix}$$

let's check scaling property:

$$f(\alpha x) = \begin{bmatrix} (\alpha x_3) - 2(\alpha x_1) \\ 3(\alpha x_1) - 2(\alpha x_2) \end{bmatrix} = \alpha \begin{bmatrix} x_3 - 2x_1 \\ 3x_1 - 2x_2 \end{bmatrix} = \alpha f(x)$$

Matrix multiplication and linear functions

general example: $f(x) = Ax$, where A is $m \times n$ matrix

- scaling: $f(\alpha x) = A(\alpha x) = \alpha Ax = \alpha f(x)$
- superposition: $f(u + v) = A(u + v) = Au + Av = f(u) + f(v)$

so, matrix multiplication is a linear function

Converse

- every linear function $y = f(x)$, with y an m-vector and x and n-vector, can be expressed as $y = Ax$ for some $m \times n$ matrix A
- you can get the coefficients of A from $a_{ij} = y_i$ when $x = e_j$

Linear equations

an equation in the variables $x_1, \ldots, x_n$ is called *linear* if each side consists of a sum of multiples of x_i, and a constant, *e.g.*,

$$1 + x_2 = x_3 - 2x_1$$

is a linear equation in x_1, x_2, x_3

Systems of linear equations: m linear equations in n variables $x_1, \ldots, x_n$:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

- can express compactly as $Ax = b$
- a_{ij} are the *coefficients*; $A \in \mathbb{R}^{m \times n}$ is the *coefficient matrix*
- $b \in \mathbb{R}^m$ is called the *right-hand side*
- may have no solution, a unique solution, infinitely many solutions

Classification of linear equations

$$Ax = b$$

- *under-determined* if $m < n$ (A wide; more unknowns than equations)
- *square* if $m = n$ (A square)
- *over-determined* if $m > n$ (A tall; more equations than unknowns)

Example

two equations in three variables x_1, x_2, x_3:

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

- step 1: rewrite equations with variables on the lefthand side, lined up in columns, and constants on the righthand side:

$$\begin{array}{rrrcr} 2x_1 & +x_2 & -x_3 & = & -1 \\ 0x_1 & -x_2 & +x_3 & = & -2 \end{array}$$

  (each row is one equation)

- step 2: rewrite equations as a single matrix equation:

$$\begin{bmatrix} 2 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \end{bmatrix}$$

  - ith row of A gives the coefficients of the ith equation
  - jth column of A gives the coefficients of x_j in the equations
  - ith entry of b gives the constant in the ith equation

Solving square linear equations

- suppose we have n linear equations in n variables $x_1, \ldots, x_n$
- compact matrix form: $Ax = b$, where A is an $n \times n$ matrix, and b is an n-vector
- suppose A is invertible, *i.e.*, its inverse A^{-1} exists
- multiply both sides of $Ax = b$ on the left by A^{-1}:

$$A^{-1}(Ax) = A^{-1}b$$

- lefthand side simplifies to $A^{-1}Ax = Ix = x$, so we've solved the linear equations:

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

Square linear equation

set or system of n linear equations with n variables $x_1, \ldots, x_n$:

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
&\vdots \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n
\end{aligned}
$$

- scalars a_{ij} are called *coefficients*
- the numbers b_i are called *right-hand-sides*

Matrix notation

$$Ax = b$$

- the $n \times n$ matrix A is called the *coefficient matrix*
- the m vector b is called the *right-hand side*

Cramer's rule

if $\det(A) \neq 0$, then the square linear system $Ax = b$ has a unique solution

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

we can find the solution using *Cramer's formula*

$$x_k = \frac{|D_k|}{|A|}, \quad k = 1, 2, \ldots, n$$

- D_k is the matrix obtained replacing the kth column of A by b
- from Cramer's formula (with some algebra), we have

$$A^{-1} = \frac{1}{\det A} \underbrace{\begin{bmatrix} \det A_{11} & \det A_{21} & \cdots & \det A_{n1} \\ \det A_{12} & \det A_{22} & \cdots & \det A_{n1} \\ \vdots & \vdots & \cdots & \vdots \\ \det A_{1n} & \det A_{2n} & \cdots & \det A_{nn} \end{bmatrix}}_{\text{adj } A}$$

A_{ij}, is the $(m-1) \times (m-1)$ obtained by deleting row i and column j of A

Example: Cramer's rule

$$\begin{aligned} 0.3x_1 + 0.52x_2 + x_3 &= -0.01 \\ 0.5x_1 + x_2 + 1.9x_3 &= 0.67 \\ 0.1x_1 + 0.3x_2 + 0.5x_3 &= -0.44 \end{aligned}$$

the determinant can be written as

$$|A| = \begin{vmatrix} 0.3 & 0.52 & 1 \\ 0.5 & 1 & 1.9 \\ 0.1 & 0.3 & 0.5 \end{vmatrix}$$

the minors are:

$$A_{11} = \begin{vmatrix} 1 & 1.9 \\ 0.3 & 0.5 \end{vmatrix} = 1(0.5) - 1.9(0.3) = -0.07$$

$$A_{12} = \begin{vmatrix} 0.5 & 1.9 \\ 0.1 & 0.5 \end{vmatrix} = 0.5(0.5) - 1.9(0.1) = 0.06$$

$$A_{13} = \begin{vmatrix} 0.5 & 1 \\ 0.1 & 0.3 \end{vmatrix} = 0.5(0.3) - 1(0.1) = 0.05$$

Example: Cramer's rule

$$|A| = 0.3(-0.07) - 0.52(0.06) + 1(0.05) = -0.0022$$

Solution using Cramer's rule

$$x_1 = \frac{\begin{vmatrix} -0.01 & 0.52 & 1 \\ 0.67 & 1 & 1.9 \\ -0.44 & 0.3 & 0.5 \end{vmatrix}}{-0.0022} = \frac{0.03278}{-0.0022} = -14.9$$

$$x_2 = \frac{\begin{vmatrix} 0.3 & -0.01 & 1 \\ 0.5 & 0.67 & 1.9 \\ 0.1 & -0.44 & 0.5 \end{vmatrix}}{-0.0022} = \frac{0.0649}{-0.0022} = -29.5$$

$$x_3 = \frac{\begin{vmatrix} 0.3 & 0.52 & -0.01 \\ 0.5 & 1 & 0.67 \\ 0.1 & 0.3 & -0.44 \end{vmatrix}}{-0.0022} = \frac{-0.04356}{-0.0022} = 19.8$$

Linear equations with non-inveretible matrix

when A isn't invertible, *i.e.*, inverse doesn't exist

- one or more of the equations is redundant (*i.e.*, can be obtained from the others)
- the equations are inconsistent or contradictory

in practice: A isn't invertible means you've set up the wrong equations, or don't have enough of them

Solving linear equations in practice

- to solve $Ax = b$ (*i.e.*, compute $x = A^{-1}b$) by computer, we don't compute A^{-1}, then multiply it by b (but that would work!)
- practical methods compute $x = A^{-1}b$ directly, via specialized methods (studied in numerical linear algebra)
- standard methods, that work for any (invertible) A, require about n^3 multiplies & adds to compute $x = A^{-1}b$
- but modern computers are very fast, so solving say a set of 1000 equations in 1000 variables takes only a second or so, even on a small computer
- . . . which is simply amazing