Matrix

CE40282-1: Linear Algebra
Hamid R. Rabiee and Maryam Ramezani
Sharif University of Technology

Basic Notation

- By $A \in \mathbb{R}^{m \times n}$ we denote a matrix with m rows and n columns, where the entries of A are real numbers.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} | & | & & | \\ a^1 & a^2 & \cdots & a^n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} — & a_1^T & — \\ — & a_2^T & — \\ & \vdots & \\ — & a_m^T & — \end{bmatrix}$$

Matrices Equality

- Two matrices are equal if they have the same size ($m \times n$) and entries corresponding to the same position are equal

For $A = [a_{ij}]_{m\times n}$ and $B = [b_{ij}]_{m\times n}$,

$A = B$ if and only if $a_{ij} = b_{ij}$ for $1 \leq i \leq m$, $\quad 1 \leq j \leq n$

Matrix Operations

- Matrix-Matrix addition
- Scalar-Matrix multiplication
- Matrix-Vector multiplication
- Matrix-Matrix multiplication

Matrix-Matrix Addition

- (just like vectors) we can add or subtract matrices of the same size:

$$(A+B)_{ij} = A_{ij} + B_{ij}, \quad i = 1, \ldots, m, \quad j = 1, \ldots, n$$

- Properties:
  - Commutative $A + B = B + A$
  - Associative $A + (B + C) = (A + B) + C$
  - Addition with zero $A + 0 = A$
  - Transpose $(A + B)^T = A^T + B^T$

Scalar-Matrix Multiplication

- Example $2\begin{bmatrix} 1 & -1 & 2 \\ -3 & 0 & 4 \end{bmatrix} = \begin{bmatrix} 2 & -2 & 4 \\ -6 & 0 & 8 \end{bmatrix}$
- Properties:
  - Associative $(\alpha\beta)A = \alpha(\beta A)$
  - Distributive property of scalar multiplication over real-number addition $(\alpha + \beta)A = \alpha A + \beta A$
  - Distributive property of scalar multiplication over matrix addition $\alpha(A + B) = \alpha A + \alpha B$
  - $0A = 0 \qquad 1A = A$
  - Transpose $(\alpha A)^T = \alpha A^T$

Review: Vector-Vector Product

- *inner product* or *dot product*

$$x^T y \in \mathbb{R} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \sum_{i=1}^{n} x_i y_i .$$

- *outer product*

$$xy^T \in \mathbb{R}^{m \times n} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix} = \begin{bmatrix} x_1 y_1 & x_1 y_2 & \cdots & x_1 y_n \\ x_2 y_1 & x_2 y_2 & \cdots & x_2 y_n \\ \vdots & \vdots & \ddots & \vdots \\ x_m y_1 & x_m y_2 & \cdots & x_m y_n \end{bmatrix} .$$

Matrix-Vector Multiplication

- If we write A by rows, then we can express Ax as,

$A \in \mathbb{R}^{m \times n}$

$$y = Ax = \begin{bmatrix} — & a_1^T & — \\ — & a_2^T & — \\ & \vdots & \\ — & a_m^T & — \end{bmatrix} x = \begin{bmatrix} a_1^T x \\ a_2^T x \\ \vdots \\ a_m^T x \end{bmatrix}.$$

$$a_i^T x = \sum_{j=1}^{n} a_{ij} x$$

- If we write A by columns, then we have:

$$y = Ax = \begin{bmatrix} | & | & & | \\ a^1 & a^2 & \cdots & a^n \\ | & | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a^1 \end{bmatrix} x_1 + \begin{bmatrix} a^2 \end{bmatrix} x_2 + \ldots + \begin{bmatrix} a^n \end{bmatrix} x_n .$$

y is a ***linear combination*** of the *columns* of A.

- columns of A are linearly independent if Ax = 0 implies x = 0

Matrix-Vector Multiplication

$A \in \mathbb{R}^{m \times n}$

It is also possible to multiply on the left by a row vector.

- If we write A by columns, then we can express $x^\top A$ as,

$$y^T = x^T A = x^T \begin{bmatrix} | & | & & | \\ a^1 & a^2 & \cdots & a^n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} x^T a^1 & x^T a^2 & \cdots & x^T a^n \end{bmatrix}$$

- expressing A in terms of rows we have:

$$y^T = x^T A = \begin{bmatrix} x_1 & x_2 & \cdots & x_m \end{bmatrix} \begin{bmatrix} — & a_1^T & — \\ — & a_2^T & — \\ & \vdots & \\ — & a_m^T & — \end{bmatrix}$$

$$= x_1 \begin{bmatrix} — & a_1^T & — \end{bmatrix} + x_2 \begin{bmatrix} — & a_2^T & — \end{bmatrix} + \ldots + x_m \begin{bmatrix} — & a_m^T & — \end{bmatrix}$$

y^T is a linear combination of the *rows* of A.

- Example for different representations of matrix-vector multiplication

Matrix-Vector Multiplication

- Properties
  - $A(u+v) = Au + Av$
  - $(A+B)u = Au + Bu$
  - $(\alpha A)u = \alpha(Au) = A(\alpha u) = \alpha Au$
  - $0u = 0$
  - $A0 = 0$
  - $Iu = u$

Matrix-Vector Multiplication

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} | & | & & | \\ a^1 & a^2 & \cdots & a^n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} — & a_1^T & — \\ — & a_2^T & — \\ & \vdots & \\ — & a_m^T & — \end{bmatrix}$$

- Column j: $a^j =$
- Row i: $a_i^T =$
- Vector sum of rows of A=
- Vector sum of columns of A=

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

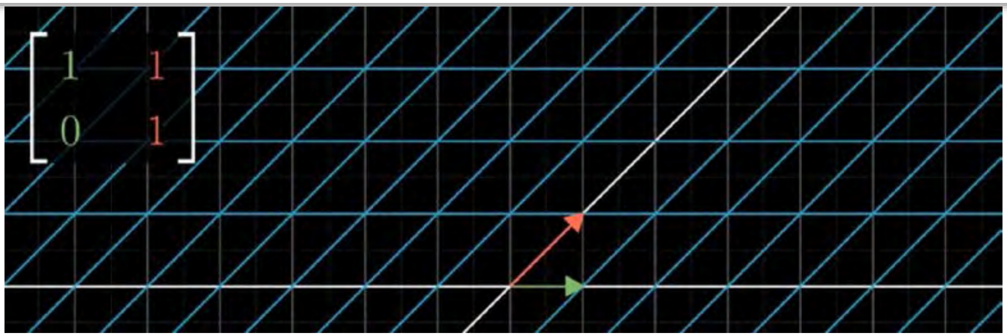
Linear Transformation

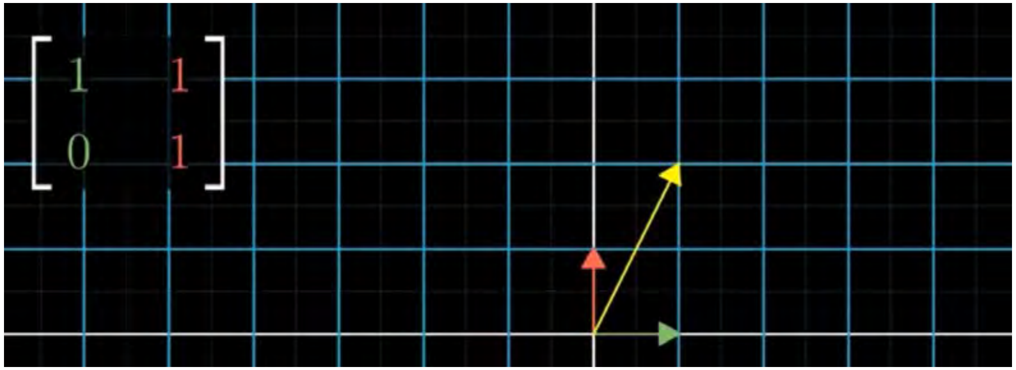
$$L(\vec{\mathbf{v}} + \vec{\mathbf{w}}) = L(\vec{\mathbf{v}}) + L(\vec{\mathbf{w}}) \qquad \text{“Additivity”}$$

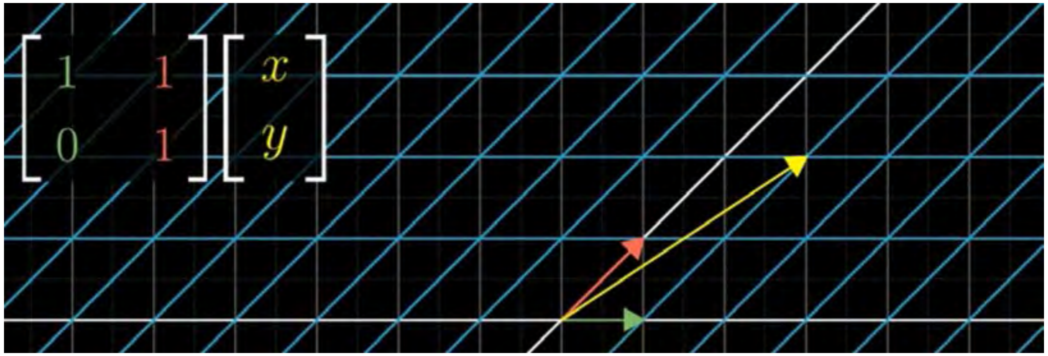
$$L(c\vec{\mathbf{v}}) = cL(\vec{\mathbf{v}}) \qquad \text{“Scaling”}$$

- Linear Transformation
  - Lines remain lines
  - Origin remains fixed

Linear Transformation







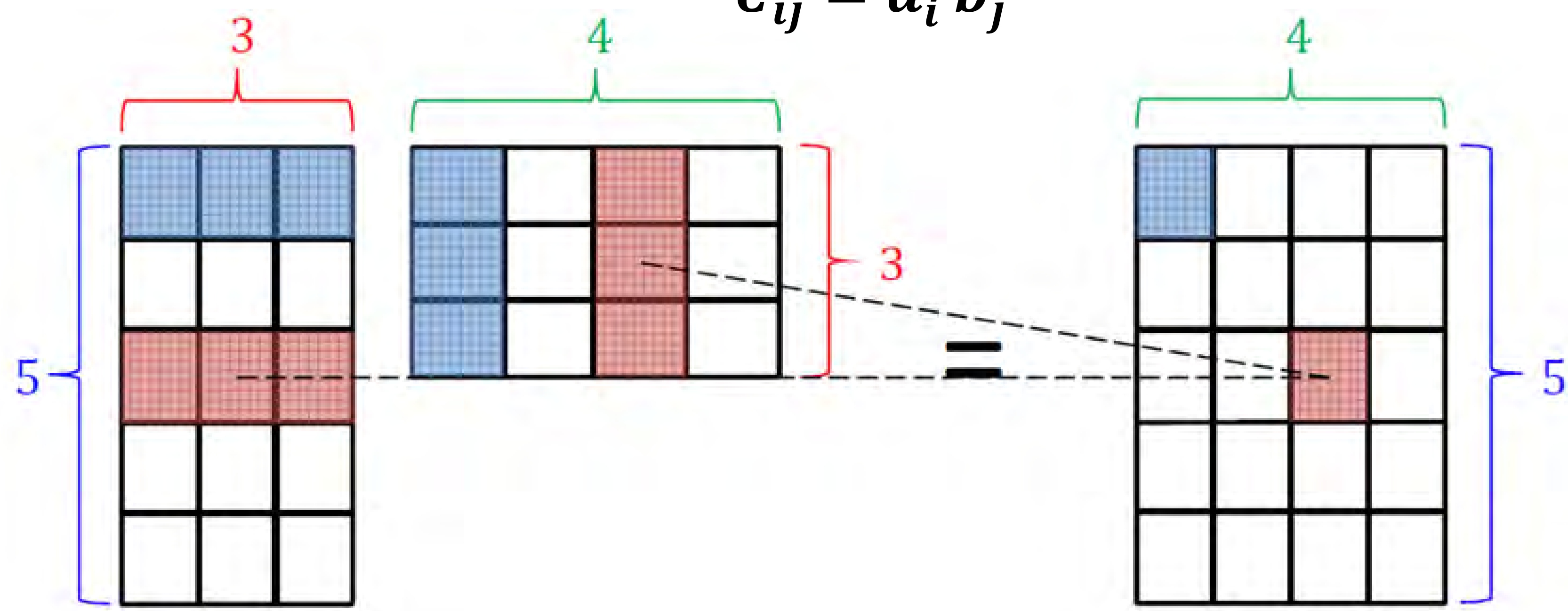
Source: https://www.youtube.com/watch?v=kYB8IZa5AuE&list=PLZHQObOWTQDPD3MizzM2xVFitgF8hE_ab&index=3

Matrix-Matrix Multiplication

- Matrix-matrix: $A \in \mathbb{R}^{m\times k}, B \in \mathbb{R}^{k\times n} \to \mathbb{R}^{m\times n}$
  - $\boldsymbol{a}_i$ rows of A, $\boldsymbol{b}_j$ cols of B

$$\boldsymbol{C = AB} \qquad \boldsymbol{for\ 1 \le i \le m\ ,1 \le j \le n\ inner\ product}$$

$$\boldsymbol{C_{ij} = a_i^T b_j}$$



Matrix-Matrix Multiplication (different views)

1. As a set of vector-vector products

$$C = AB = \begin{bmatrix} — & a_1^T & — \\ — & a_2^T & — \\ & \vdots & \\ — & a_m^T & — \end{bmatrix} \begin{bmatrix} | & | & & | \\ b^1 & b^2 & \cdots & b^p \\ | & | & & | \end{bmatrix} = \begin{bmatrix} a_1^T b^1 & a_1^T b^2 & \cdots & a_1^T b^p \\ a_2^T b^1 & a_2^T b^2 & \cdots & a_2^T b^p \\ \vdots & \vdots & \ddots & \vdots \\ a_m^T b^1 & a_m^T b^2 & \cdots & a_m^T b^p \end{bmatrix}$$

2. As a sum of outer products

$$C = AB = \begin{bmatrix} | & | & & | \\ a^1 & a^2 & \cdots & a^n \\ | & | & & | \end{bmatrix} \begin{bmatrix} — & b_1^T & — \\ — & b_2^T & — \\ & \vdots & \\ — & b_n^T & — \end{bmatrix} = \sum_{i=1}^{n} a^i b_i^T$$

Matrix-Matrix Multiplication (different views)

3. As a set of matrix-vector products.

$$C = AB = A \begin{bmatrix} | & | & & | \\ b^1 & b^2 & \cdots & b^p \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ Ab^1 & Ab^2 & \cdots & Ab^p \\ | & | & & | \end{bmatrix}.$$

Here the ith column of C is given by the matrix-vector product with the vector on the right, $c_i = Ab_i$. These matrix-vector products can in turn be interpreted using both viewpoints given in the previous subsection.

4. As a set of vector-matrix products.

$$C = AB = \begin{bmatrix} — & a_1^T & — \\ — & a_2^T & — \\ & \vdots & \\ — & a_m^T & — \end{bmatrix} B = \begin{bmatrix} — & a_1^T B & — \\ — & a_2^T B & — \\ & \vdots & \\ — & a_m^T B & — \end{bmatrix}$$

Matrix-Matrix Multiplication

- Properties:
  - Associative

$$(AB)C = A(BC)$$

  - Distributive

$$A(B+C) = AB + BC$$

  - NOT commutative

$$AB \neq BA$$

    - Dimensions may not even be conformable

Matrix-Matrix Multiplication

- A^k : repeated multiplication of a square matrix

$$A^1 = A, A^2 = AA, \dots, A^k = \underbrace{AA \cdots A}_{k \text{ matrices}}$$

- Properties:
  - $A^j A^k = A^{j+k}$
  - $(A^j)^k = A^{jk}$

where j and k are nonnegative integers and A^0 is assumed to be I

- For diagonal matrices:

$$D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} \Rightarrow D^k = \begin{bmatrix} d_1^k & 0 & \cdots & 0 \\ 0 & d_2^k & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_n^k \end{bmatrix}$$

Note

- Two properties which is held for real numbers, but not for matrices:
  - (1) commutative property of matrix multiplication

$$ab = ba \qquad AB \neq BA$$

  - Example

$$A = \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 2 & -1 \\ 0 & 2 \end{bmatrix}$$

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

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

Note

- (2) cancellation law

$$ac = bc, \qquad c \neq 0 \quad \Rightarrow \quad a = b$$

$AC = BC$ and $C \neq 0$ (C is not a zero matrix)

(1) If C is invertible, then $A = B$

(2) If C is not invertible, then $A \neq B$

  - Example

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

$$AC = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} -2 & 4 \\ -1 & 2 \end{bmatrix} \qquad BC = \begin{bmatrix} 2 & 4 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} -2 & 4 \\ -1 & 2 \end{bmatrix}$$

So, although $AC = BC$, $A \neq B$

Matrix Operations Complexity

- $m \times n$ matrix stored A as $m \times n$ array of numbers
  (for sparse A, store only $\mathbf{nnz}(A)$ nonzero values)
- matrix addition, scalar-matrix multiplication cost mn flops
- matrix-vector multiplication costs $m(2n-1) \approx 2mn$ flops
  (for sparse A, around $2\mathbf{nnz}(A)$ flops)

Transpose

- The *transpose* of a matrix results from "flipping" the rows and columns. Given a matrix $A \in \mathbb{R}^{m \times n}$, its transpose, written $A^T \in \mathbb{R}^{n \times m}$, is the $n \times m$ matrix whose entries are given by

$$(A^T)_{ij} = A_{ji}.$$

- Properties:
  - $(A^T)^T = A$
  - $(A + B)^T = A^T + B^T$
  - $(cA)^T = c(A^T)$
  - $(AB)^T = B^T A^T \longrightarrow (A_1 A_2 A_3 \cdots A_n)^T = A_n^T \cdots A_3^T A_2^T A_1^T$

Trace

The ***trace*** of a square matrix $A \in \mathbb{R}^{n\times n}$, denoted $\mathrm{tr}A$, is the sum of diagonal elements in the matrix:

$$\mathrm{tr}A = \sum_{i=1}^{n} A_{ii}.$$

The trace has the following properties:

- For $A \in \mathbb{R}^{n\times n}$, $\mathrm{tr}A = \mathrm{tr}A^T$.
- For $A, B \in \mathbb{R}^{n\times n}$, $\mathrm{tr}(A+B) = \mathrm{tr}A + \mathrm{tr}B$.
- For $A \in \mathbb{R}^{n\times n}$, $t \in \mathbb{R}$, $\mathrm{tr}(tA) = t\,\mathrm{tr}A$.
- For A, B such that AB is square, $\mathrm{tr}AB = \mathrm{tr}BA$.
- For A, B, C such that ABC is square, $\mathrm{tr}ABC = \mathrm{tr}BCA = \mathrm{tr}CAB$, and so on for the product of more matrices.

Inverse

- For$A \in M_{n\times n}$, if there exists a matrix $B \in M_{n\times n}$ such that $AB = BA = I_n$, then:
  - A is invertible (or nonsingular)
  - B is the inverse of A
- A square matrix that does not have an inverse is called noninvertible (or singular)
- Note
  - The definition of the inverse of a matrix is similar to that of the inverse of a scalar, i.e., $c \cdot (1/c) = 1$
  - Since there is no inverse (or said multiplicative inverse for the real number 0, you can “imagine” that noninvertible matrices act a similar role to the real number 0 is some sense

Inverse

- The inverse of A is denoted by A^{-1}
- **Theorem: The inverse of a matrix is unique**
  - Proof?

- Find the inverse of a matrix by the Gauss-Jordan Elimination:

$$[A \quad | \quad I] \xrightarrow{\text{Gauss−Jordan Elimination}} [I \quad | \quad A^{-1}]$$

Gauss-Jordan Elimination for finding the Inverse of a Matrix

- Let A be an $n \times n$ matrix.
  - Adjoin the identity $n \times n$ matrix I_n to A to form the matrix $[A : I_n]$.
  - Compute the reduced echelon form of $[A : I_n]$.
- If the reduced echelon form is of the type $[I_n : B]$, then B is the inverse of A.
- If the reduced echelon form is not of the type $[I_n : B]$, in that the first $n \times n$ submatrix is not I_n, then A has no inverse.

> An $n \times n$ matrix A is invertible if and only if its reduced echelon form is I_n.

A is row equivalent to I_n

Inverse (Example)

$$A = \begin{bmatrix} 1 & 4 \\ -1 & -3 \end{bmatrix}$$

$$AX = I$$

$$\begin{bmatrix} 1 & 4 \\ -1 & -3 \end{bmatrix}\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} x_{11} + 4x_{21} & x_{12} + 4x_{22} \\ -x_{11} - 3x_{21} & -x_{12} - 3x_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

by equating corresponding entries

$$\Rightarrow \begin{array}{rcrcl} x_{11} & + & 4x_{21} & = & 1 \\ -x_{11} & - & 3x_{21} & = & 0 \end{array} \quad (1)$$

$$\begin{array}{rcrcl} x_{12} & + & 4x_{22} & = & 0 \\ -x_{12} & - & 3x_{22} & = & 1 \end{array} \quad (2)$$

This two systems of linear equations have the same coefficient matrix, which is exactly the matrix A.

$$(1) \Rightarrow \left[\begin{array}{rrcr} 1 & 4 & \vdots & 1 \\ -1 & -3 & \vdots & 0 \end{array}\right] \xrightarrow{A_{1,2}^{(1)},\ A_{2,1}^{(-4)}} \left[\begin{array}{rrcr} 1 & 0 & \vdots & -3 \\ 0 & 1 & \vdots & 1 \end{array}\right] \Rightarrow x_{11} = -3, x_{21} = 1$$

$$(2) \Rightarrow \left[\begin{array}{rrcr} 1 & 4 & \vdots & 0 \\ -1 & -3 & \vdots & 1 \end{array}\right] \xrightarrow{A_{1,2}^{(1)},\ A_{2,1}^{(-4)}} \left[\begin{array}{rrcr} 1 & 0 & \vdots & -4 \\ 0 & 1 & \vdots & 1 \end{array}\right] \Rightarrow x_{12} = -4, x_{22} = 1$$

Thus

$$X = A^{-1} = \begin{bmatrix} -3 & -4 \\ 1 & 1 \end{bmatrix}$$

Perform the Gauss-Jordan elimination on the matrix A with the same row operations

$$\underbrace{\left[\begin{array}{rr} 1 & 4 \\ -1 & -3 \end{array}\right.}_{A} \begin{array}{c} \vdots \\ \vdots \end{array} \underbrace{\left.\begin{array}{rr} 1 & 0 \\ 0 & 1 \end{array}\right]}_{I} \xrightarrow[A_{1,2}^{(1)},A_{2,1}^{(-4)}]{\text{Gauss-Jordan Elimination}} \underbrace{\left[\begin{array}{rr} 1 & 0 \\ 0 & 1 \end{array}\right.}_{I} \begin{array}{c} \vdots \\ \vdots \end{array} \underbrace{\left.\begin{array}{rr} -3 & -4 \\ 1 & 1 \end{array}\right]}_{A^{-1}}$$

solution for $\begin{bmatrix} x_{11} \\ x_{21} \end{bmatrix}$ solution for $\begin{bmatrix} x_{12} \\ x_{22} \end{bmatrix}$

Inverse

- Properties (If A is an invertible matrix, k is a positive integer, and c is a scalar):
  - A^{-1} is invertible and $(A^{-1})^{-1} = A$
  - A^k is invertible and $(A^k)^{-1} = A^{-k} = (A^{-1})^k$
  - cA is invertible if $c \neq 0$ and $(cA)^{-1} = \frac{1}{c}A^{-1}$
  - A^T is invertible and $(A^T)^{-1} = (A^{-1})^T$
- Theorem: If A and B are invertible matrices of order n, then AB is invertible and $\boldsymbol{(AB)^{-1} = B^{-1}A^{-1}}$
  - Proof?

$$\boldsymbol{(A_1A_2A_3 \cdots A_n)^{-1} = A_n^{-1} \cdots A_3^{-1}A_2^{-1}A_1^{-1}}$$

Inverse

- Theorem: Let $AX = B$ be a system of n linear equations in n variables.
  If A^{-1} exists, the solution is unique and is given by $X = A^{-1}B$.

Elementary Matrices

- An elementary matrix is one that is obtained by performing a single elementary row operation on an identity matrix.

If an elementary row operation is performed on an $m \times n$ matrix A, the resulting matrix can be written as EA, where the $m \times m$ matrix E is created by performing the same row operation on I_m.

- Example

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{bmatrix}.$$

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

Elementary Matrices

Each elementary matrix E is invertible. The inverse of E is the elementary matrix of the same type that transforms E back into I.

- Example

Find the inverse of $E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}$.

Inverse

- Theorem:

An $n \times n$ matrix A is invertible if and only if A is row equivalent to I_n, and in this case, any sequence of elementary row operations that reduces A to I_n also transforms I_n into A^{-1}.

Inverse

- Theorem: If A and B are row equivalent matrices and A is invertible, then B is invertible.

Invertible Matrix

Let A be a square $n \times n$ matrix. Then the following statements are equivalent. That is, for a given A, the statements are either all true or all false.

a. A is an invertible matrix.

b. A is row equivalent to the $n \times n$ identity matrix.

c. A has n pivot positions.

d. The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.

e. The columns of A form a linearly independent set.

f. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.

g. The equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$.

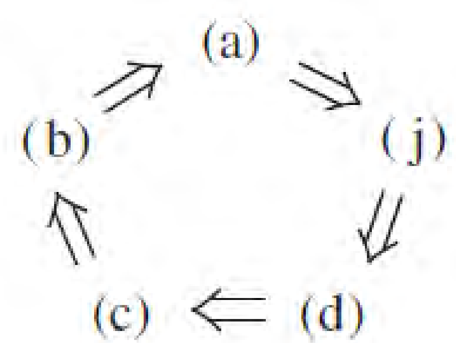
h. The columns of A span $\mathbb{R}^n$.

i. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^n$.

j. There is an $n \times n$ matrix C such that $CA = I$.

k. There is an $n \times n$ matrix D such that $AD = I$.

l. A^T is an invertible matrix.



Useful Matrices

- An $m \times n$ matrix is
  - Tall $m > n$
  - Wide $n > m$
  - Square $m = n$
- Main diagonal of matrix

$$A_{n\times n} = \begin{bmatrix} \ddots \end{bmatrix} \qquad a_{11}, a_{22}, \dots a_{nn}$$

- Anti diagonal of matrix

$$A_{n\times n} = \begin{bmatrix} \iddots \end{bmatrix} \qquad a_{1,n}, a_{2,n-1}, \dots a_{n,1}$$

Useful Matrices

- Identity matrix

$I \in \mathbb{R}^{n\times n}$, is a square matrix with ones on the diagonal and zeros everywhere else. That is,

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

$$I_n = [e_1, e_2, e_3]$$

$$I_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

It has the property that for all $A \in \mathbb{R}^{m\times n}$,

$$AI = A = IA.$$

- Diagonal matrix

a matrix where all non-diagonal elements are 0. $D = \mathrm{diag}(d_1, d_2, \ldots, d_n)$, with

$$D_{ij} = \begin{cases} d_i & i = j \\ 0 & i \neq j \end{cases}$$

Clearly, $I = \mathrm{diag}(1, 1, \ldots, 1)$.

$$A = \mathrm{diag}(a_1, \ldots, a_m) = \begin{bmatrix} a_1 & \cdots & 0 \\ \vdots & a_i & \vdots \\ 0 & \cdots & a_m \end{bmatrix}$$

- Scalar matrix A special kind of diagonal matrix in which all diagonal elements are the same

$$\begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Useful Matrices

- A matrix A over R is called:
  - symmetric if $A^T = A$
  - skew-symmetric if $A^T = -A$
  - $A^T A$ must be symmetric (*A with any size, it is not necessary for A to be a square matrix*)

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & -1 \\ 3 & -1 & 4 \end{bmatrix} \text{ and } B = \begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & -3 \\ -2 & 3 & 0 \end{bmatrix}$$

- A is orthogonal if $AA^T = A^T A = I$

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

Useful Matrices

- Let $A = [a_{ij}]$ be an $n \times n$ matrix with $a_{ij} = \begin{cases} 1 & if\ i = j+1 \\ 0 & other \end{cases}$.

  Then $A^n = 0$ and $A^k \neq 0$ for $1 \leq k \leq n-1$
  - Nilpotent: A for which a positive integer p exists such that $A^p = 0$.
  - Order of nilpotency (degree , index): Least positive integer p for which $A^p = 0$ is called the.

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad C = \begin{bmatrix} 5 & -3 & 2 \\ 15 & -9 & 6 \\ 10 & -6 & 4 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 & 2 & 1 & 6 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} B^2 = \begin{bmatrix} 0 & 0 & 2 & 7 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix};$$

$$B^3 = \begin{bmatrix} 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}; \ B^4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Useful Matrices

- Idempotent: satisfy the condition that $A^2 = A$

Examples of 2×2 idempotent matrices are:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 3 & -6 \\ 1 & -2 \end{bmatrix}$$

Examples of 3×3 idempotent matrices are:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 2 & -2 & -4 \\ -1 & 3 & 4 \\ 1 & -2 & -3 \end{bmatrix}$$

If a matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is idempotent, then

- $a = a^2 + bc$,
- $b = ab + bd$, implying $b(1 - a - d) = 0$ so $b = 0$ or $d = 1 - a$,
- $c = ca + cd$, implying $c(1 - a - d) = 0$ so $c = 0$ or $d = 1 - a$,
- $d = bc + d^2$.

Useful Matrices

- Toeplitz: diagonal-constant matrix: values on diagonals are equal
- A Toeplitz matrix is not necessarily square.

$$A = \begin{bmatrix} a_0 & a_{-1} & a_{-2} & \cdots & \cdots & a_{-(n-1)} \\ a_1 & a_0 & a_{-1} & \ddots & & \vdots \\ a_2 & a_1 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & a_{-1} & a_{-2} \\ \vdots & & \ddots & a_1 & a_0 & a_{-1} \\ a_{n-1} & \cdots & \cdots & a_2 & a_1 & a_0 \end{bmatrix}$$

$$A_{i,j} = A_{i+1,j+1} = a_{i-j}.$$

$$T(b) = \begin{bmatrix} b_1 & 0 & 0 & 0 \\ b_2 & b_1 & 0 & 0 \\ b_3 & b_2 & b_1 & 0 \\ 0 & b_3 & b_2 & b_1 \\ 0 & 0 & b_3 & b_2 \\ 0 & 0 & 0 & b_3 \end{bmatrix}$$

Useful Matrices

- Submatrix of matrix: A matrix obtained by deleting some of the rows and/or columns of a matrix is said to be a submatrix of the given matrix.

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

$$[1], [2], \begin{bmatrix} 1 \\ 0 \end{bmatrix}, [1\ 5], \begin{bmatrix} 1 & 5 \\ 0 & 2 \end{bmatrix}, A. \begin{bmatrix} 1 & 4 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 0 & 2 \end{bmatrix}$$

Useful Matrices

- Zero or null Matrix

If $A \in M_{m\times n}$, and c is a scalar,

then (1) $A + 0_{m\times n} = A$

※ So, $\mathbf{0}_{m\times n}$ is also called the additive identity for the set of all $m \times n$ matrices

(2) $A + (-A) = 0_{m\times n}$

※ Thus , $-A$ is called the additive inverse of A

(3) $cA = 0_{m\times n} \Rightarrow c = 0$ or $A = 0_{m\times n}$

All above properties are very similar to the counterpart properties for the real number 0

Useful Matrices

- Block Matrix whose entries are matrices, such as

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

Submatrix or block of A

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

then

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

- Matrices in each block row must have same height (row dimension)
- Matrices in each block column must have same width (column dimension)
- Note: A is not a square matrix but it is a block square matrix.

Useful Matrices

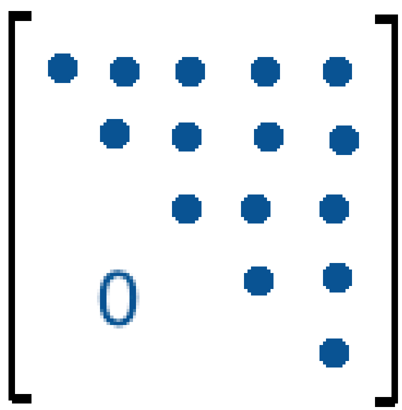
- Block Matrix
  - Transpose of block matrix $\begin{bmatrix} A & B \\ C & D \end{bmatrix}^T = \begin{bmatrix} A^T & C^T \\ B^T & D^T \end{bmatrix}$
  - Multiplication

$$A = \left[\begin{array}{ccc|cc} 2 & -3 & 1 & 0 & -4 \\ 1 & 5 & -2 & 3 & -1 \\ \hline 0 & -4 & -2 & 7 & -1 \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \left[\begin{array}{cc} 6 & 4 \\ -2 & 1 \\ -3 & 7 \\ \hline -1 & 3 \\ 5 & 2 \end{array}\right] = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

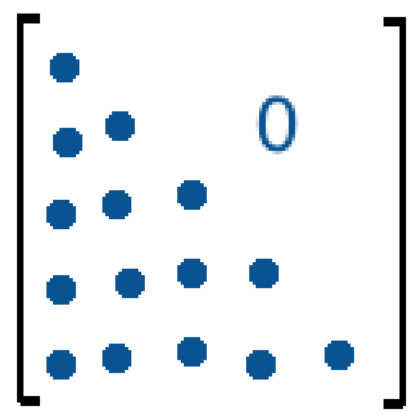
$$AB = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} A_{11}B_1 + A_{12}B_2 \\ A_{21}B_1 + A_{22}B_2 \end{bmatrix} = \left[\begin{array}{cc} -5 & 4 \\ -6 & 2 \\ \hline 2 & 1 \end{array}\right]$$

Useful Matrices

- Triangular matrix
  - Upper triangular $a_{ij} = 0, \quad i > j$
  - Lower triangular $a_{ij} = 0, \quad i < j$



Upper Triangular Matrix



Lower Triangular Matrix

Useful Matrices

- Sparse matrix
  - Density of matrix $A_{m \times n}$
  - Density of identity matrix?
  - Sparse matrix has low density

$$1 \geq \frac{\mathbf{nnz}(A)}{mn}$$

Vec Operator

- The vec-operator applied on a matrix A stacks the columns into a vector

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \qquad \mathrm{vec}(\mathbf{A}) = \begin{bmatrix} A_{11} \\ A_{21} \\ A_{12} \\ A_{22} \end{bmatrix}$$

- Properties:

$$\begin{aligned}
\mathrm{vec}(\mathbf{AXB}) &= (\mathbf{B}^T \otimes \mathbf{A})\mathrm{vec}(\mathbf{X}) \\
\mathrm{Tr}(\mathbf{A}^T\mathbf{B}) &= \mathrm{vec}(\mathbf{A})^T\mathrm{vec}(\mathbf{B}) \\
\mathrm{vec}(\mathbf{A} + \mathbf{B}) &= \mathrm{vec}(\mathbf{A}) + \mathrm{vec}(\mathbf{B}) \\
\mathrm{vec}(\alpha\mathbf{A}) &= \alpha \cdot \mathrm{vec}(\mathbf{A}) \\
\mathbf{a}^T\mathbf{XBX}^T\mathbf{c} &= \mathrm{vec}(\mathbf{X})^T(\mathbf{B} \otimes \mathbf{ca}^T)\mathrm{vec}(\mathbf{X})
\end{aligned}$$