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

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

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

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

$$AI = A = IA.$$

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

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

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

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

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

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 Product

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

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

$$y = Ax = \begin{bmatrix} — & a_1^T & — \\ — & a_2^T & — \\ & \vdots & \\ — & a_m^T & — \end{bmatrix} x = \begin{bmatrix} a_1^T x \\ a_2^T x \\ \vdots \\ a_m^T x \end{bmatrix}. \qquad a_i^T x = \sum_{j=1}^{n} a_{ij} x$$

- 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.

Matrix-Vector Product

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

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.

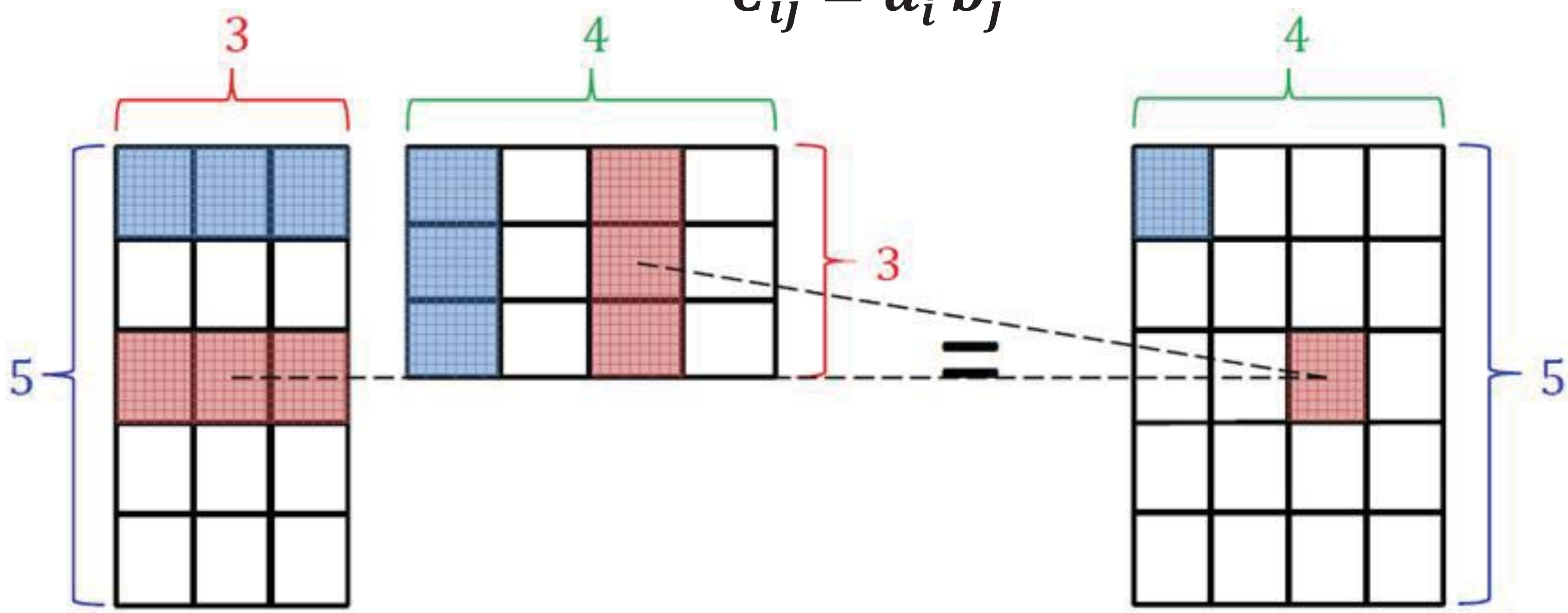
- Example for different representations of matrix-vector multiplication

Matrix-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

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

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



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.

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

Linear Operator

- Matrix-vector: $x \in \mathbb{R}^{n}, M \in \mathbb{R}^{m\times n} \to \mathbb{R}^{m}$

$$\boldsymbol{Mx}$$

$$\boldsymbol{M(\alpha x + \beta y) = \alpha Mx + \beta My}$$
$$\forall \boldsymbol{\alpha, \beta} \in \boldsymbol{R} \quad \forall \boldsymbol{x, y} \in \boldsymbol{R^n}$$

$$\boldsymbol{M\left(\sum_{i=1}^{p} \alpha_i x_i\right) = \sum_{i=1}^{p} \alpha_i M x_i}$$
$$\forall \boldsymbol{\alpha_i} \in \boldsymbol{R} \quad \forall \boldsymbol{x_i}\ \boldsymbol{R^n} \ \ \boldsymbol{1 \le i \le n}$$

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

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

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

The following properties of transposes are easily verified:

- $(A^T)^T = A$
- $(AB)^T = B^T A^T$
- $(A + B)^T = A^T + B^T$

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.