DM559

Linear and Integer Programming

Lecture 2

Matrices and Vectors

Marco Chiarandini

Department of Mathematics & Computer Science
University of Southern Denmark

Outline

Outline

Matrices and Vectors

Definition (Matrix)

A matrix is a rectangular array of numbers or symbols. It can be written as

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

- We denote this array by a single letter A or by (a_{ij}) and
- we say that A has m rows and n columns, or that it is an $m \times n$ matrix.
- The size of A is $m \times n$.
- The number a_{ij} is called the (i, j) entry or scalar.

- A square matrix is an $n \times n$ matrix.
- The diagonal of a square matrix is the list of entries $a_{11}, a_{22}, \ldots, a_{nn}$
- The diagonal matrix is a matrix $n \times n$ with $a_{ij} = 0$ if $i \neq j$(ie, a square matrix with all the entries which are not on the diagonal equal to 0):

$$\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{mn} \end{bmatrix}$$

Definition (Equality)

Two matrices are equal if they have the same size and if corresponding entries are equal. That is, if $A = (a_{ij})$ and $B = (b_{ij})$ are both $m \times n$ matrices, then:

$$A = B \iff a_{ij} = b_{ij} \qquad 1 \leq i \leq m, 1 \leq j \leq n$$

Matrix Addition and Multiplication

Definition (Addition)

If $A = (a_{ij})$ and $B = (b_{ij})$ are both $m \times n$ matrices, then

$$A + B = (a_{ij} + b_{ij}) \qquad 1 \leq i \leq m, 1 \leq j \leq n$$

Definition (Scalar Multiplication)

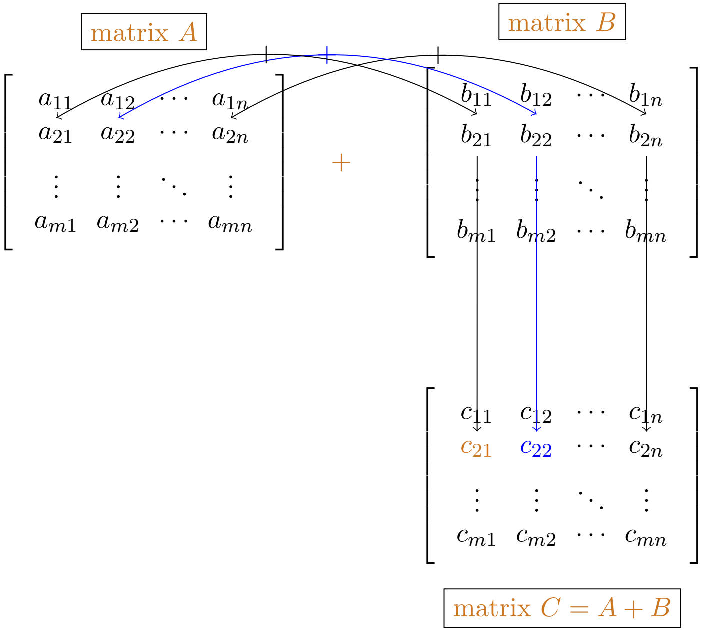
If $A = (a_{ij})$ is an $m \times n$ matrix and $\lambda \in \mathbb{R}$, then

$$\lambda A = (\lambda a_{ij}) \qquad 1 \leq i \leq m, 1 \leq j \leq n$$

Eg:

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

$$-2A =?$$


matrix A
matrix B
$+$
matrix $C = A + B$

Matrix Multiplication

Two matrices can be multiplied together, depending on the size of the matrices

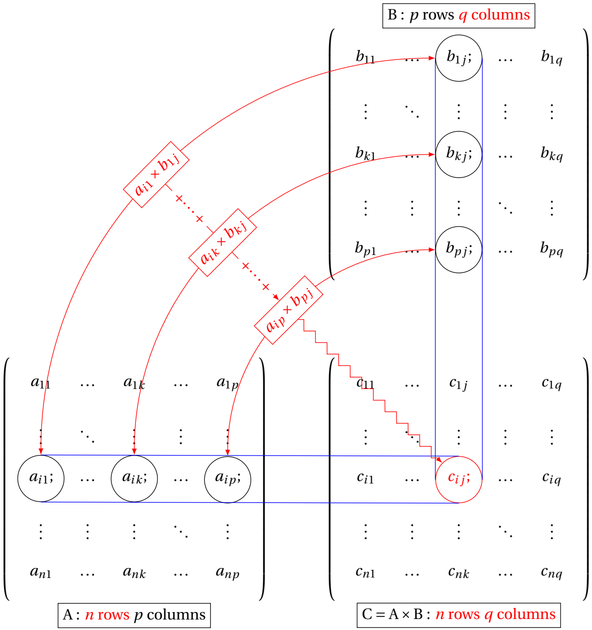
Definition (Matrix Multiplication)

If A is an $m \times n$ matrix and B is an $n \times p$ matrix, then the product is the matrix $AB = C = (c_{ij})$ with

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

$$\begin{bmatrix} & & & \\ & & & \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ & & & \end{bmatrix} \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{bmatrix}$$

What is the size of C?



B : p rows q columns
A : n rows p columns
C = A × B : n rows q columns
$a_{i1} \times b_{1j}$
$a_{ik} \times b_{kj}$
$a_{ip} \times b_{pj}$
+ ... +
+ ... +

$$AB = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \\ 1 & 2 & 4 \\ 2 & 2 & -1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 1 & 1 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 5 & 3 \\ 1 & 14 \\ 9 & -1 \end{bmatrix}$$

$$(2)(3) + (0)(1) + (1)(-1) = 5$$

The motivation behind this definition will become clear later. It is exactly what is needed in our study of linear algebra

- $AB \neq BA$ in general, ie, not commutative
  try with the example of previous slide...

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

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

Matrix Algebra

Matrices are useful because they provide compact notation and we can perform algebra with them

Bear in mind to use only operations that are defined. In the following rules, the sizes are dictated by the operations being defined.

- commutative $A + B = B + A$. Proof?
- associative:
  - $(A + B) + C = A + (B + C)$
  - $\lambda(AB) = (\lambda A)B = A(\lambda B)$
  - $(AB)C = A(BC)$

Size?

- distributive:
  - $A(B + C) = AB + AC$
  - $(B + C)A = BA + CA$
  - $\lambda(A + B) = \lambda A + \lambda B$

Why both first two rules?

Definition (Zero Matrix)

A zero matrix, denoted 0, is an $m \times n$ matrix with all entries zero:

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

- additive identity: $A - A = 0$
  - $A + 0 = A$
  - $A - A = 0$
  - $0A = 0, A0 = 0$

Definition (Identity Matrix)

The $n \times n$ identity matrix, denoted I_n or I is the diagonal matrix with $a_{ii} = 1$: zero:

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

- multiplicative identity (like 1 does for scalars)
  - $AI = A$ and $IA = A$

A of size $m \times n$.
What size is I?

Exercise: $3A + 2B = 2(B - A + C)$

Matrix Inverse

- If $AB = AC$ can we conclude that $B = C$?

$$A = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & -1 \\ 3 & 5 \end{bmatrix}, \qquad C = \begin{bmatrix} 8 & 0 \\ -4 & 4 \end{bmatrix}$$

$$AB = AC = \begin{bmatrix} 0 & 0 \\ 4 & 4 \end{bmatrix}$$

- $A + 5B = A + 5C \implies B = C$
  addition and scalar multiplication have inverses ($-A$ and $1/c$)

- Is there a multiplicative inverse?

Definition (Inverse Matrix)

The $n \times n$ matrix A is invertible if there is a matrix B such that

$$AB = BA = I$$

where I is the $n \times n$ identity matrix. The matrix B is called the inverse of A and is denoted by A^{-1}.

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

Theorem

If A is an $n \times n$ invertible matrix, then the matrix A^{-1} is unique.

Proof: Assume A has two inverses B, C so $AB = BA = I$ and $AC = CA = I$. Consider the product CAB:

$$CAB = C(AB) = CI = C \qquad \text{associativity} + AB = I$$

$$CAB = (CA)B = IB = B \qquad \text{associativity} + CA = I$$

- If a matrix has an inverse we say that it is invertible or non-singular
  If a matrix has no inverse we say that it is non-invertible or singular
  Eg:

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

- If

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \qquad ad - bc \neq 0$$

  then A has the inverse

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \qquad ad - bc \neq 0$$

- The scalar $ad - bc$ is called determinant of A and denoted $|A|$.

Matrix Inverse

Back to the question:

- If $AB = AC$ can we conclude that $B = C$?
  If A is invertible then the answer is yes:

$$A^{-1}AB = A^{-1}AC \implies IB = IC \implies B = C$$

- But $AB = CA$ then we cannot conclude that $B = C$.

Properties of the Inverse

Let A be invertible $\implies$ A^{-1} exists

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

- $(\lambda A)^{-1} = \frac{1}{\lambda} A^{-1}$
  the inverse of the matrix (λA) is a matrix C that satisfies $(\lambda A) C = C(\lambda A) = I$. Using matrix algebra:

$$(\lambda A)\left(\frac{1}{\lambda}A^{-1}\right) = \lambda\frac{1}{\lambda}AA^{-1} = I \text{ and } \left(\frac{1}{\lambda}A^{-1}\right)(\lambda A) = \frac{1}{\lambda}\lambda A^{-1}A = I$$

- $(AB)^{-1} = B^{-1}A^{-1}$

Powers of a matrix

For A an $n \times n$ matrix and $r \in \mathbb{N}$

$$A^r = \underbrace{AA \dots A}_{r \text{ times}}$$

For the associativity of matrix multiplication:

- $(A^r)^{-1} = (A^{-1})^r$
- $A^r A^s = A^{r+s}$
- $(A^r)^s = A^{rs}$

Transpose Matrix

Definition (Transpose)

The transpose of an $m \times n$ matrix A is the $n \times m$ matrix B defined by

$$b_{ij} = a_{ji} \qquad \text{for } i = 1, \ldots, n \text{ and } j = 1, \ldots, m$$

It is denoted A^T

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

$$\begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$$

Note that if D is a diagonal matrix: $D^T = D$

Properties of the transpose

- $(A^T)^T = A$
- $(\lambda A)^T = \lambda A^T$
- $(A+B)^T = A^T + B^T$
- $(AB)^T = B^T A^T$ (consider first which matrix sizes make sense in the multiplication, then rewrite the terms)
- if A is invertible, $(A^T)^{-1} = (A^{-1})^T$

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

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

Symmetric Matrix

Definition (Symmetric Matrix)

A matrix A is symmetric if it is equal to its transpose, $A = A^T$.
(only square matrices can be symmetric)

Vectors

- An $n \times 1$ matrix is a column vector, or simply a vector:

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

  The numbers $v_1, v_2, \ldots$ are known as the components (or entries) of v.

- A row vector is a $1 \times n$ matrix
- We write vectors in lower boldcase type (writing by hand we can either underline them or add an arrow over $\mathbf{v}$).
- Addition and scalar multiplication are defined for vectors as for $n \times 1$ matrices:

$$\mathbf{v} + \mathbf{w} = \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{bmatrix} \qquad \lambda\mathbf{v} = \begin{bmatrix} \lambda v \\ \lambda v \\ \vdots \\ \lambda v \end{bmatrix}$$

- For a fixed n, the set of vectors together with the operations of addition and multiplication form the set $\mathbb{R}^n$, usually called Euclidean space
- For vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ in $\mathbb{R}^n$ and scalars $\alpha_1, \alpha_2, \ldots, \alpha_k$ in $\mathbb{R}$, the vector

$$\mathbf{v} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_k \mathbf{v}_k$$

  is known as linear combination of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$
- A zero vector is denoted by $\mathbf{0}$;
  $\mathbf{0} + \mathbf{v} = \mathbf{v} + \mathbf{0} = \mathbf{v}$;
  $0\mathbf{v} = \mathbf{0}$

- The matrix product of $\mathbf{v}$ and $\mathbf{w}$ cannot be calculated
- The matrix product of $\mathbf{v}^T\mathbf{w}$ gives an 1×1 matrix
- The matrix product of $\mathbf{v}\mathbf{w}^T$ gives an $n \times n$ matrix

Inner product of two vectors

Definition (Inner product)

Given

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \qquad \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$$

the inner product denoted $\langle \mathbf{v}, \mathbf{w} \rangle$, is the real number given by

$$\langle \mathbf{v}, \mathbf{w} \rangle = \left\langle \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}, \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \right\rangle = v_1 w_1 + v_2 w_2 + \ldots + v_n w_n = \mathbf{v}^T \mathbf{w}$$

It is also called scalar product or dot product (and written $\mathbf{v} \cdot \mathbf{w}$).

$$\mathbf{v}^T\mathbf{w} = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = v_1 w_1 + v_2 w_2 + \ldots + v_n w_n =$$

Theorem

The inner product

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n, \qquad \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$$

satisfies the following properties for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ *and for all* $\alpha \in \mathbb{R}$*:*

- $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$
- $\alpha \langle \mathbf{x}, \mathbf{y} \rangle = \langle \alpha\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \alpha\mathbf{y} \rangle$
- $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$
- $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ *and* $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ *if and only if* $\mathbf{x} = \mathbf{0}$

Note: vectors from different Euclidean spaces live in different 'worlds'

Vectors and Matrices

Theorem

Let A be an $m \times n$ matrix

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

and denote the columns of A by the column vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$, so that

$$\mathbf{a}_i = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{bmatrix}, \qquad i = 1, \ldots, n.$$

Then if $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ is any vector in $\mathbb{R}^n$

$$A\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \ldots + x_n\mathbf{a}_n$$

(vector $A\mathbf{x}$ in $\mathbb{R}^m$ as a linear combination of the column vectors of A)

Resume

1. Matrices and Vectors