

LECTURE 6: VECTORS AND MATRICES

Juan Dodyk

WashU

PLAN

Vectors

1. Vectors in \mathbb{R}^n
2. Sum, inner product
3. Norm, triangle inequality, Cauchy-Schwarz

Matrices

1. Matrices in $\mathbb{R}^{n \times m}$
2. Sum, scalar product, matrix product
3. Vectors as matrices
4. Transpose



VECTORS

\mathbb{R}^n is the set of **vectors** (x_1, \dots, x_n) , which are lists of real numbers x_1, \dots, x_n .

\mathbb{R}^1 is just \mathbb{R} .

\mathbb{R}^2 is the set of pairs (x_1, x_2) . We can represent them as points (x, y) in the plane.

If we have n people and we have asked their age, we can represent that data as a vector in \mathbb{R}^n .

The zero vector $\mathbf{0} \in \mathbb{R}^n$ is defined as $(0, 0, \dots, 0)$.

Some people use bold to emphasize that something is a vector, e.g., $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{0} \in \mathbb{R}^n$. Others use a little arrow, e.g., $\vec{x} \in \mathbb{R}^n$, $\vec{0} \in \mathbb{R}^n$. I'll use bold here, but in general I don't think it's necessary.

OPERATIONS

Given $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ we can take their sum or difference $\mathbf{x} \pm \mathbf{y}$, which is the vector

$$\mathbf{x} \pm \mathbf{y} = (x_1 \pm y_1, x_2 \pm y_2, \dots, x_n \pm y_n).$$

Example.

- $(1, -1, 0) + (2, 3, -2) = (3, 2, -2)$.
- $(1, 0) + (0, 1, 2)$ is not defined.

Given $\mathbf{x} \in \mathbb{R}^n$ and $a \in \mathbb{R}$ we define the **scalar multiplication** as $a\mathbf{x} \in \mathbb{R}^n$ given by

$$a\mathbf{x} = (ax_1, ax_2, \dots, ax_n).$$

Example. $2 \cdot (1, 2, 3) = (2, 4, 6)$.

INNER PRODUCT

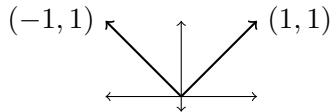
Given $x, y \in \mathbb{R}^n$ we define their **inner product** (or **dot product**, or **scalar product**) as the number

$$\mathbf{x} \cdot \mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

Example. $(1, -1, 0) \cdot (2, 3, -2) = 2 - 3 + 0 = -1$.

If $\mathbf{x} \cdot \mathbf{y} = 0$ we say that the vectors are **orthogonal**, or **perpendicular**.

Example. $(1, 1) \cdot (-1, 1) = 0$.



Properties.

- $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$

- $\mathbf{x} \cdot \mathbf{0} = 0$

- $\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}$

- $\mathbf{x} \cdot (c\mathbf{y}) = c\mathbf{x} \cdot \mathbf{y}$ if $c \in \mathbb{R}$.

EUCLIDEAN NORM

Given $\mathbf{x} \in \mathbb{R}^n$ we define its (Euclidean) **norm** as

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = \sqrt{\mathbf{x} \cdot \mathbf{x}}.$$

It measures the length of the line from the point $\mathbf{0}$ to the point \mathbf{x} .

Example. If $\mathbf{x} = (1, 1)$, $\|\mathbf{x}\| = \sqrt{1^2 + 1^2} = \sqrt{2}$. This is Pythagoras' theorem.

Properties. We have

1. If $a \in \mathbb{R}$, $\|a\mathbf{x}\| = |a|\|\mathbf{x}\|$.
2. Triangle inequality: if $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ then $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.
3. Cauchy-Schwarz inequality: $|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\|\|\mathbf{y}\|$.

If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ we define their **distance** as $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$.

EXERCISE

QUESTION 1

Let $\mathbf{x} = (1, 1)$, $\mathbf{y} = (-1, 2)$, $\mathbf{a} = (1, 2, 3)$, $\mathbf{b} = (-1, 0, 2)$. Calculate, if possible,

1. $\mathbf{x} + \mathbf{y}$

2. $2 \cdot \mathbf{y}$

3. $\mathbf{x} \cdot \mathbf{y}$

4. $\|\mathbf{x} - \mathbf{y}\|$

5. $\mathbf{x} + \mathbf{a}$

6. $\mathbf{a} + \mathbf{b}$

7. $\|\mathbf{a} - \mathbf{b}\|$

8. $|\mathbf{a} \cdot \mathbf{b}|$

9. $\mathbf{y} \cdot \mathbf{a}$

MATRICES

A **matrix** $\mathbf{A} \in \mathbb{R}^{m \times n}$ is an $m \times n$ table of numbers, with m rows and n columns:

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

Examples. $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$, $\begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \in \mathbb{R}^{1 \times 3}$, $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \in \mathbb{R}^{3 \times 1}$.

$\mathbf{0} \in \mathbb{R}^{m \times n}$ is the matrix whose entries are all zero.

If we have a dataset of n people with k numerical variables, we can represent the data as a matrix $\mathbf{A} \in \mathbb{R}^{n \times k}$. Each row is a person, each column is a variable.

OPERATIONS

Given two matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$ we can define their sum or difference $\mathbf{A} \pm \mathbf{B}$ in the obvious way:

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

Examples.

- $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix}$
- $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} - \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ 3 \\ 3 \end{pmatrix}$

Properties. We have

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A},$$

$$\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C},$$

and

$$\mathbf{A} - \mathbf{A} = \mathbf{0}.$$

OPERATIONS

We can do scalar multiplication: if $c \in \mathbb{R}$ and $\mathbf{A} \in \mathbb{R}^{m \times n}$ then

$$c\mathbf{A} = \begin{pmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{pmatrix}.$$

Examples.

- $2 \cdot \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}$
- $-\frac{1}{2} \cdot \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ 1 \\ -\frac{3}{2} \end{pmatrix}$

Property. If $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$ and $c \in \mathbb{R}$ then

$$c(\mathbf{A} + \mathbf{B}) = c\mathbf{A} + c\mathbf{B}.$$

MATRIX \times VECTOR MULTIPLICATION

If $\mathbf{x} \in \mathbb{R}^n$ we can see it as a matrix $\mathbf{x} \in \mathbb{R}^{n \times 1}$: $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$.

Given a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ and a vector $\mathbf{x} \in \mathbb{R}^n$ we can multiply them:

$$\mathbf{Ax} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ a_{21}x_1 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix} \in \mathbb{R}^{m \times 1}.$$

Notice that the i th entry of \mathbf{Ax} is the inner product of the i th row of \mathbf{A} with \mathbf{x} .

MATRIX \times VECTOR MULTIPLICATION

Examples.

- $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$

- $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ can't be done.

- $\begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 8 \\ 2 \end{pmatrix}$

EXERCISE

QUESTION 2

Calculate:

1. $\begin{pmatrix} 1 & 0 \\ 0 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

2. $\begin{pmatrix} 1 & 0 \\ 0 & -1 \\ 1 & 1 \end{pmatrix} \left[2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right]$

LINEAR MODELING IS MATRIX MULTIPLICATION

Suppose we have units $i = 1, \dots, n$, a vector of outcomes $\mathbf{y} \in \mathbb{R}^n$, and k independent variables $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n$. We want to approximate the outcome as a linear function of the independent variables. In other words, we want coefficients $\beta_1, \dots, \beta_k \in \mathbb{R}$ such that

$$y_i \approx \beta_1 \mathbf{x}_{1i} + \dots + \beta_k \mathbf{x}_{ki}.$$

We can create a matrix $\mathbf{X} \in \mathbb{R}^{n \times k}$ whose columns are the k variables. If $\boldsymbol{\beta} \in \mathbb{R}^k$ is the vector of coefficients,

$$\mathbf{X}\boldsymbol{\beta} = \begin{pmatrix} \mathbf{x}_{11} & \cdots & \mathbf{x}_{k1} \\ \vdots & \ddots & \vdots \\ \mathbf{x}_{1n} & \cdots & \mathbf{x}_{kn} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} = \begin{pmatrix} \beta_1 \mathbf{x}_{11} + \cdots + \beta_k \mathbf{x}_{k1} \\ \vdots \\ \beta_1 \mathbf{x}_{1n} + \cdots + \beta_k \mathbf{x}_{kn} \end{pmatrix}.$$

Therefore, the objective is to minimize the distance between the outcomes \mathbf{y} and the linear model $\mathbf{X}\boldsymbol{\beta}$, i.e.,

$$\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|.$$

OLS means doing this. The answer is $\hat{\boldsymbol{\beta}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}$. We'll see what this means.

MATRIX MULTIPLICATION

If $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{B} \in \mathbb{R}^{n \times k}$ then we define $\mathbf{AB} \in \mathbb{R}^{m \times k}$ by

$$\begin{aligned}\mathbf{AB} &= \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & \cdots & b_{1k} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nk} \end{pmatrix} \\ &= \begin{pmatrix} a_{11}b_{11} + \cdots + a_{1n}b_{n1} & \cdots & a_{11}b_{1k} + \cdots + a_{1n}b_{nk} \\ \vdots & \ddots & \vdots \\ a_{m1}b_{11} + \cdots + a_{mn}b_{n1} & \cdots & a_{m1}b_{1k} + \cdots + a_{mn}b_{nk} \end{pmatrix}\end{aligned}$$

We take each column $\mathbf{b}_{\bullet 1}, \dots, \mathbf{b}_{\bullet k}$ of \mathbf{B} and form $(\mathbf{Ab}_{\bullet 1} \quad \cdots \quad \mathbf{Ab}_{\bullet k})$.

EXAMPLES

$$\bullet \underbrace{\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}}_{2 \times 2} \underbrace{\begin{pmatrix} 1 & 2 & 3 \\ 1 & -2 & 0 \end{pmatrix}}_{2 \times 3} = \underbrace{\begin{pmatrix} -1 & 2 & 0 \\ 1 & 2 & 3 \end{pmatrix}}_{2 \times 3}$$

$$\bullet \underbrace{\begin{pmatrix} 1 & 2 & 3 \end{pmatrix}}_{1 \times 3} \underbrace{\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}}_{3 \times 1} = \underbrace{\begin{pmatrix} 2 \end{pmatrix}}_{1 \times 1}$$

$$\bullet \underbrace{\begin{pmatrix} 0 & -1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix}}_{3 \times 2} \underbrace{\begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}}_{2 \times 3} = \underbrace{\begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}}_{3 \times 3}$$

$$\bullet \underbrace{\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}}_{3 \times 1} \underbrace{\begin{pmatrix} 1 & 2 & 3 \end{pmatrix}}_{1 \times 3} = \underbrace{\begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ -1 & -2 & -3 \end{pmatrix}}_{3 \times 3}$$

EXERCISE

QUESTION 3

Calculate:

1. $\begin{pmatrix} 1 & 0 \\ 0 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix}$

2. $\begin{pmatrix} 1 & 0 \\ 0 & -1 \\ 1 & 1 \end{pmatrix} \left[2 \begin{pmatrix} 1 & 0 & -1 \\ -1 & 2 & 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \end{pmatrix} \right]$

PROPERTIES

- **Associativity.** If $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{m \times k}$ and $C \in \mathbb{R}^{k \times l}$ then

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

- **Distributivity.** If $A \in \mathbb{R}^{m \times n}$, $B, C \in \mathbb{R}^{m \times k}$ then

$$A(B + C) = AB + AC.$$

If $A, B \in \mathbb{R}^{m \times n}$, $C \in \mathbb{R}^{m \times k}$ then

$$(A + B)C = AC + BC.$$

- **Commutativity with scalars.** $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times k}$ and $c \in \mathbb{R}$ then

$$cAB = A(cB).$$

MULTIPLICATION IS NOT COMMUTATIVE

Example.

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$$

$$\mathbf{AB} = \begin{pmatrix} 2 & 3 \\ -2 & 2 \end{pmatrix}, \quad \mathbf{BA} = \begin{pmatrix} 1 & 7 \\ -1 & 3 \end{pmatrix}$$

so $\mathbf{AB} \neq \mathbf{BA}$.

TRANSPOSE

The **transpose** of $\mathbf{A} \in \mathbb{R}^{m \times n}$ is the matrix $\mathbf{A}^\top \in \mathbb{R}^{n \times m}$ whose rows are the columns of \mathbf{A} .

Sometimes people write \mathbf{A}' instead of \mathbf{A}^\top .

Examples.

- $\begin{pmatrix} 1 \\ 2 \end{pmatrix}^\top = \begin{pmatrix} 1 & 2 \end{pmatrix}$
- $\begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \end{pmatrix}^\top = \begin{pmatrix} 0 & -1 \\ 1 & 0 \\ 2 & 1 \end{pmatrix}$
- $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^\top = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$

PROPERTIES OF THE TRANSPOSE

- $(\mathbf{A}^\top)^\top = \mathbf{A}$.
- $(\mathbf{A} + \mathbf{B})^\top = \mathbf{A}^\top + \mathbf{B}^\top$.
- $(c\mathbf{A})^\top = c\mathbf{A}^\top$.
- $(\mathbf{AB})^\top = \mathbf{B}^\top \mathbf{A}^\top$.

Notice that if $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ then $\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^\top \mathbf{y}$:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}^\top \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x_1 y_1 + \cdots + x_n y_n.$$

EXERCISE

QUESTION 4

Take $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \\ 1 & 1 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & 2 \end{pmatrix}$. Calculate:

1. \mathbf{AB}
2. $\mathbf{B}^\top \mathbf{A}^\top$
3. $\mathbf{A} + 2\mathbf{B}^\top$
4. $\mathbf{A}^\top + \mathbf{B}$

