KAZAKH-BRITISH TECHNICAL UNIVERSITY

Introduction to Machine Learning Week 4

Olivier JAYLET

School of Information Technology and Engineering

Scalar

Definition : A scalar is a number. Examples of scalars are temperature, distance, speed, or mass – all quantities that have a magnitude but no “direction”, other than perhaps positive or negative.

Vectors

Definitions :

- a list of numbers.
- ordered sequence of numbers
- an element of a vector space

Vectors

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

To obtain a row representation, the vector **x** should be transposed:

$$x^\top = (x_1, \dots, x_n).$$

Vector operations

- **Addition:**

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

$$\mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}.$$

Multiplication by a scalar (number)

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \alpha \in \mathbb{R}, \alpha x = \begin{pmatrix} \alpha x_1 \\ \vdots \\ \alpha x_n \end{pmatrix}.$$

Two types of vector multiplication:

Inner product (dot product)

- Result is a scalar

$$\begin{bmatrix} a_{11} & a_{12} \end{bmatrix} \cdot \begin{bmatrix} b_{11} \\ b_{21} \end{bmatrix} = a_{11}b_{11} + a_{12}b_{21}$$

Outer product

- Result for n-vectors is an n x n matrix

$$\begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} \cdot \begin{bmatrix} b_{11} & b_{12} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} & a_{11}b_{12} \\ a_{21}b_{11} & a_{21}b_{12} \end{bmatrix}$$

Inner product

Inner product (aka dot product) of vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ equals

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x_i y_i.$$

Alternative notation : $x^\intercal y$

Definition : With $m, n \in \mathbb{N}$, a real-valued (m, n) matrix A is an m-n-tuple of elements a_{ij}, $i = 1, ..., m$, $j = 1, ..., n$, which is ordered according to a rectangular scheme consisting of m rows and n columns:

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

By convention, $(1, n)$-matrices are called rows and $(m, 1)$-matrices are called columns.
These special matrices are also called row/column vectors.

Additions

The sum of two matrices $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{m\times n}$ is defined as the element-wise sum, i.e.,

$$A + B := \begin{bmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{bmatrix} \in \mathbb{R}^{m\times n}.$$

Products

Matrices can only be multiplied if their ”neighboring” dimensions match. For instance, an $n \times k$-matrix A can be multiplied with a $k \times m$-matrix B, but only from the left side:

$$\underbrace{A}_{n \times k} \underbrace{B}_{k \times m} = \underbrace{C}_{n \times m}$$

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1p} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{np} \end{bmatrix} = \begin{bmatrix} c_{11} & \cdots & c_{1p} \\ \vdots & \ddots & \vdots \\ c_{m1} & \cdots & c_{mp} \end{bmatrix}$$

The (i, j)-th entry of C is the dot product of the i-th row of A with the j-th column of B:

$$c_{ij} = \sum_{k=1}^{n} a_{ik} \cdot b_{kj}$$

Consider the multiplication of two 2×2 matrices:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}$$

Example :

A 2 x 2 and a 2 x 3 yield a ... x ...

$$\begin{bmatrix} 1 & 4 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 3 & -1 & 5 \\ 6 & 2 & 0 \end{bmatrix} = \begin{bmatrix} 27 & 7 & 5 \\ 12 & 0 & 10 \end{bmatrix}$$

A 3 x 3 and a 3 x 1 result in a ... x ...

$$\begin{bmatrix} 1 & 5 & 0 \\ 0 & 4 & 8 \\ 2 & 7 & 3 \end{bmatrix} \begin{bmatrix} 6 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 11 \\ 20 \\ 25 \end{bmatrix}$$

A 2 x 2 and a 2 x 3 yield a 2 x 3

$$\begin{bmatrix} 1 & 4 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 3 & -1 & 5 \\ 6 & 2 & 0 \end{bmatrix} = \begin{bmatrix} 27 & 7 & 5 \\ 12 & 0 & 10 \end{bmatrix}$$

A 3 x 3 and a 3 x 1 result in a 3 x 1

$$\begin{bmatrix} 1 & 5 & 0 \\ 0 & 4 & 8 \\ 2 & 7 & 3 \end{bmatrix} \begin{bmatrix} 6 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 11 \\ 20 \\ 25 \end{bmatrix}$$

Properties

Matrix multiplication is **not commutative**

- Order matters
- Unlike scalars

In general,

$$A \cdot B \neq B \cdot A$$

If A and/or B is not square then one of the above operations may not be possible anyway.

Inner dimensions may not agree for both product orders.

Properties

Matrix multiplication **is associative**

- Insertion of parentheses anywhere within a product of multiple terms does not affect the result:

$$(A \cdot B) \cdot C = D$$

$$A \cdot (B \cdot C) = D$$

Matrix multiplication **is distributive**

- Multiplication distributes over addition
- Must maintain correct order, i.e., left- or right-multiplication

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

$$(B + C)A = BA + CA$$

Identity matrix

The matrix version of 1 is the **identity matrix**

- Ones along the diagonal, zeros everywhere else
- Square ($n \times n$) matrix
- Denoted as I or I_n where n is the matrix dimension, e.g.,

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

Left- or right-multiplication by an identity matrix results in that matrix, unchanged :

$$A \cdot I = I \cdot A = A$$

Right-multiplication of an $n \times n$ matrix by an $n \times n$ identity matrix, I_n

$$\begin{bmatrix} 1 & 5 & 0 \\ 0 & 4 & 8 \\ 2 & 7 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 5 & 0 \\ 0 & 4 & 8 \\ 2 & 7 & 3 \end{bmatrix}$$

Same result if we left-multiply by I_n

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 5 & 0 \\ 0 & 4 & 8 \\ 2 & 7 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 5 & 0 \\ 0 & 4 & 8 \\ 2 & 7 & 3 \end{bmatrix}$$

Right-multiplication of an $m \times n$ matrix by an $n \times n$ identity matrix

$$\begin{bmatrix} 1 & 5 & 0 \\ 0 & 4 & 8 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 5 & 0 \\ 0 & 4 & 8 \end{bmatrix}$$

Same result if we left-multiply the $m \times n$ matrix by an $m \times m$ identity matrix

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 5 & 0 \\ 0 & 4 & 8 \end{bmatrix} = \begin{bmatrix} 1 & 5 & 0 \\ 0 & 4 & 8 \end{bmatrix}$$

Divisions

Scalar division that we are accustomed to can be thought of as multiplication by an inverse:

$$a \div b = a \cdot \frac{1}{b} = a \cdot b^{-1}$$

This is how we 'divide' matrices as well

$$A \cdot B \cdot B^{-1} = A$$

Multiplication of a scalar by its inverse is equal to 1.

- For a matrix, the result is the **identity matrix**

$$A \cdot A^{-1} = I = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & 1 & \vdots \\ 0 & \cdots & 1 \end{bmatrix}$$

Matrix inverses

Recall that matrix multiplication is not commutative

- Right- and left-multiplication are different operations

$$A \cdot B \cdot B^{-1} = A \neq B^{-1} \cdot A \cdot B$$

The inverse does not exist for all matrices

- Non-invertible matrices are referred to as **singular**
- Matrix must be **square** for its inverse to exist

Possible to calculate matrix inverses by hand

- Simple for small matrices
- Quickly becomes tedious as matrices get larger

For example, the inverse of a 2 x 2 matrix:

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

For example:

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

$$A^{-1} = \frac{1}{8 - 10} \begin{bmatrix} 4 & -5 \\ -2 & 2 \end{bmatrix} = \begin{bmatrix} -2 & 2.5 \\ 1 & -1 \end{bmatrix}$$

Multiplication of a matrix by its inverse yields the identity matrix

For example:

$$A \cdot A^{-1} = \begin{bmatrix} 2 & 5 \\ 2 & 4 \end{bmatrix} \cdot \begin{bmatrix} -2 & 2.5 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Or, for a 3 x 3:

$$A = \begin{bmatrix} 2 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{bmatrix}, \quad A^{-1} = \begin{bmatrix} 0.5 & 0 & -0.5 \\ 0 & 1 & -1 \\ 0 & 0 & 0.5 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} 0.5 & 0 & -0.5 \\ 0 & 1 & -1 \\ 0 & 0 & 0.5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The transpose of a matrix is that matrix with rows and columns swapped

- First row becomes the first column, second row becomes the second column, and so on

For example:

$$A = \begin{bmatrix} 0 & 9 \\ 2 & 7 \\ 6 & 3 \end{bmatrix}, \quad A^T = \begin{bmatrix} 0 & 2 & 6 \\ 9 & 7 & 3 \end{bmatrix}$$

Row vectors become column vectors and vice versa

$$x = \begin{bmatrix} 7 \\ -1 \\ -4 \end{bmatrix}, \quad x^T = \begin{bmatrix} 7 & -1 & -4 \end{bmatrix}$$

Thank you for your attention !