Modified on Monday 31st August, 2020, 10:59:3110:59

Aug 31, 2020

- Introduction. Vectors and matrices. [1.1, 1.2, 1.3]
- We have seen vectors and matrices in Calculus, *e.g.*, matrix of partial derivatives of $f = (f_1, f_2, \dots, f_m) : \mathbb{R}^n \to \mathbb{R}^m$.

$$Df = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

  Let's review some essential definitions related to vectors and matrices.

- $\mathbb{R}^n = \{(x_1, x_2, \dots, x_n) : x_1, x_2, \dots, x_n \in \mathbb{R}\}$
- An $m \times n$ real matrix $A = (a_{ij})$ is an array of real numbers a_{ij} of the form

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

  The number a_{ij} is called the (i, j)-entry of A.

- A vector $\mathbf{x} \in \mathbb{R}^n$ may be regarded as an $n \times 1$ matrix, and call it a column vector $\mathbf{x}$.

$$\mathbf{x} = (x_1, x_2, \dots, x_n) \qquad \Longleftrightarrow \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

- Rows R_i and columns C_j of an $m \times n$ matrix $A = (a_{ij})$

$$R_i = \begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix}, \quad C_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{bmatrix}$$

  Then A can be written as below.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_n \end{bmatrix} = \begin{bmatrix} C_1 & C_2 & \cdots & C_n \end{bmatrix}$$

- The transpose of an $m \times n$ matrix $A = (a_{ij})$ is an $n \times m$ matrix $A^T = (b_{ij})$ with $b_{ij} = a_{ji}$ for all i and j.
- The zero vector in $\mathbb{R}^n$: $(0, 0, \dots 0)$
- The $m \times n$ zero matrix: $\mathbf{0}_{m\times n} = (a_{ij})$ with $a_{ij} = 0$ for all i and j.

- At this point, we may think about the difference between a vector in $\mathbb{R}^{100}$ and a matrix of size 10×10. Both of them are sequences of length 100.

  The contexts in which they are discussed are different.
- Vectors in various contexts
  - a finite sequence
  - direction, velocity
  - experimental data
  - storage and warehousing
  - score table
  - sound, time series, stock prices
- Matrices in various contexts
  - data: preferences of users, purchase record of customers, health data, climate data, etc.
  - picture
  - graph with vertices end edges
  - information
- We need mathematical operations on information, picture or data to manipulate, encode, decode, compress, store, confirm, . . . , transform efficiently and consistently.
- Operations in $\mathbb{R}^n$: addition, scalar multiplication

  For $\alpha \in \mathbb{R}$ and $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, $\mathbf{u} = (u_i) = (u_1, \ldots, u_n)$ and $\mathbf{v} = (v_i) = (v_1, \ldots, v_n)$,

$$\alpha\mathbf{u} = \alpha(u_1, \ldots, u_n) = (\alpha u_1, \ldots, \alpha u_n)$$

$$\mathbf{u} + \mathbf{v} = (u_1, \ldots, u_n) + (v_1, \ldots, v_n) = (u_1 + v_1, \ldots, u_n + v_n)$$

- Dot product, orthogonality in $\mathbb{R}^n$

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + \cdots + u_n v_n$$

  Two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ are said to be orthogonal if and only if $\mathbf{u} \cdot \mathbf{v} = 0$.
- Operations for matrices: addition, scalar multiplication

  Let $A = (a_{ij})$ and $B = (b_{ij})$ be matrices of the same size. $\alpha \in \mathbb{R}$.

$$\alpha A = \alpha(a_{ij}) = (\alpha a_{ij})$$

$$A + B = (a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij})$$

- Matrix product

  Let $A = (a_{ij})$ be an $m \times n$ matrix and $B = (b_{ij})$ an $n \times p$ matrix. The product of A and B is the $m \times p$ matrix defined by $AB = C = (c_{ij})$, where $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$ for all i and j.

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mp} \end{bmatrix}$$

  The (i, j)-entry of AB is the matrix product of the ith row of A and the jth column of B.

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

- Let R_i and C_j be rows and columns of A and R'_i and C'_j be the rows and columns of B.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_m \end{bmatrix} = \begin{bmatrix} C_1 & C_2 & \cdots & C_n \end{bmatrix}$$

$$B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix} = \begin{bmatrix} R'_1 \\ R'_2 \\ \vdots \\ R'_n \end{bmatrix} = \begin{bmatrix} C'_1 & C'_2 & \cdots & C'_p \end{bmatrix}$$

  Then we have
$$AB = (c_{ij}) = C_1R'_1 + C_2R'_2 + \cdots + C_nR'_n,$$
  where $C_kR'_k$ is an $m \times p$ matrix for all k, and
$$AB = (c_{ij}) = (R_iC'_j),$$
  where $c_{ij} = R_iC'_j \in \mathbb{R}$.

- Dot product as a matrix product

  For two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ regarded as column vectors, we have $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T\mathbf{v}$.

  Compare $\mathbf{u}^T\mathbf{v}$ with $\mathbf{u}\mathbf{v}^T$. The former is a real number, a 1×1 matrix, but the latter is an $n \times n$ matrix.

- What is the complexity of the matrix product? Let A, B be $n \times n$ real matrices.
  - How many additions and multiplications of real numbers, do we need to compute the matrix product AB?
  - How many additions and multiplications of real numbers, do we need to compute $AB\mathbf{v}$ for a vector $\mathbf{v} \in \mathbb{R}^n$?

- If you have a large matrix M, say of size $10^9 \times 10^4$, but don't have enough space to save, then what can we do?

  Can we express M as a sum of simpler matrices M_i's,
$$M = M_1 + M_2 + M_3 + \cdots,$$
  so that the partial sum $\sum_{i=1}^{k} M_k$ is a good approximation of M?

- Let M be a photo taken by your high resolution digital camera. M is a large matrix. How can you approximate it by a matrix of less information content?

- There is a picture, i.e. a large matrix of size $10^5 \times 10^4$, which may contain a cat.

  How can an AI determine whether it contains a cat or not? Of course, linear algebra alone can't solve this problem but the role of linear algebra is significant.

- Suppose you have a customer data as an $10^8 \times 10^4$ matrix M, whose rows are indexed by customers and columns by types of preferences. As is the case for a data in real life, the data M is not complete. It may have only partial responses to preferences.

  What can we do using this data?

  If you want to predict a new customer's preferences from M, assuming that his/her responses are not complete, what properties of M will be useful?

- The standard norm of $\mathbf{v} = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

  The distance between $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$

$$\|\mathbf{u} - \mathbf{v}\|$$

- We will later define some useful norm and distance for matrices.

Next time, we will study matrix algebra and linear systems.