Vectors & Matrices (Review)

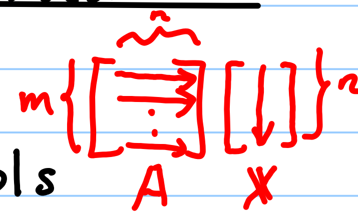
Note Title

☆ Matrix - Vector Multiplication

$\mathbf{x} \in \mathbb{R}^{n}$

$A = (a_{ij}) \in \mathbb{R}^{m \times n}$

i.e., m rows $\times$ n cols



Now, $\mathbf{y} = \underline{A\mathbf{x}} \in \mathbb{R}^{m}$

Very important to notice that $\mathbf{y}$ is a linear combination of the column vectors of A.

$$\mathbf{y} = A\mathbf{x} = [\mathbf{a}_1 | \mathbf{a}_2 | \cdots | \mathbf{a}_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$\mathbf{a}_j$ is the jth col. vector of A.
$\mathbf{a}_j \in \mathbb{R}^m$

$\mathbf{a}_j$ is often written as $\underline{\mathbf{a}_{\cdot j}}$

$$\mathbf{y} = x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 + \cdots + x_n \mathbf{a}_n$$

$$y_i = x_1 a_{i1} + x_2 a_{i2} + \cdots + x_n a_{in}$$
$$= \sum_{j=1}^{n} x_j a_{ij} = \sum_{j=1}^{n} a_{ij} x_j, \quad 1 \le i \le m$$

Thm Let $F_A : \mathbb{R}^n \to \mathbb{R}^m$ be a map defined as

$$F_A : \underset{\in \mathbb{R}^n}{\mathbf{x}} \longmapsto \underset{\in \mathbb{R}^m}{A\mathbf{x}}$$

Then, F_A is a linear map.
i.e., $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $\forall \alpha \in \mathbb{R}$,

$$\begin{cases} F_A(\mathbf{x}+\mathbf{y}) = F_A(\mathbf{x}) + F_A(\mathbf{y}) \\ F_A(\alpha\mathbf{x}) = \alpha F_A(\mathbf{x}) \end{cases}$$

Conversely, for any linear map $F: \mathbb{R}^n \to \mathbb{R}^m$, there exists a unique matrix $A \in \mathbb{R}^{m\times n}$ s.t. $F = F_A$.

(Proof) It's easy to prove F_A is a linear map. ($A(\mathbf{x}+\mathbf{y}) = A\mathbf{x} + A\mathbf{y}$, $A(\alpha\mathbf{x}) = \alpha A\mathbf{x}$ using the definition of a matrix-vector product.)

Showing the converse is more challenging.

Let F be a linear map

Let $\{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$ be the canonical basis of $\mathbb{R}^n$, i.e., $\mathbf{e}_j = \begin{array}{c} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \end{array}\begin{array}{l} \\ \\ \\ \leftarrow j \\ \\ \\ \\ \end{array}$, $1 \le j \le n$

Set $F(\mathbf{e}_j) = \mathbf{a}_j \in \mathbb{R}^m$

Let $A = [\mathbf{a}_1 \cdots \mathbf{a}_n] \in \mathbb{R}^{m\times n}$

Now pick any $\mathbf{x} \in \mathbb{R}^n$, we can always write

$$\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$$

$$\begin{aligned} \text{Then } F(\mathbf{x}) &= F(x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n) \\ &= F(x_1\mathbf{e}_1) + \cdots + F(x_n\mathbf{e}_n) \\ &= x_1F(\mathbf{e}_1) + \cdots + x_nF(\mathbf{e}_n) \\ &= x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n \\ &= A\mathbf{x} = F_A(\mathbf{x}) \end{aligned}$$

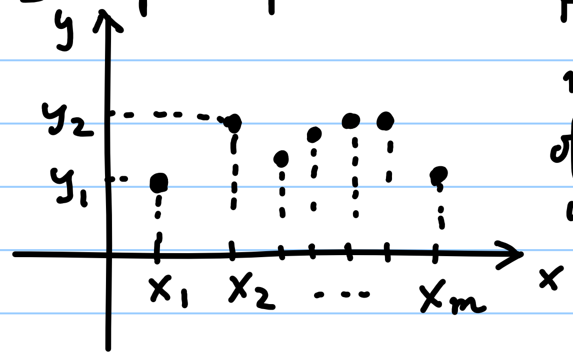
About the uniqueness, let $A, B \in \mathbb{R}^{m\times n}$

Suppose $F_A = F_B$. Then

$F_A(\mathbf{e}_j) = \mathbf{a}_j = F_B(\mathbf{e}_j) = \mathbf{b}_j$, $1 \le j \le n \Rightarrow A = B$ //

Example : A Vandermonde Matrix

Let $\{x_1, \cdots, x_m\}$ be a set of sample points.



For example, y_i may be temperature of this room measured at time x_i, $i = 1, \cdots, m$

Assume $x_i \neq x_j$ if $i \neq j$

Now consider a space of polynomials

$$P_{n-1}[x] := \{ p(x) = c_0 + c_1 x + \cdots + c_{n-1} x^{n-1}, \quad c_j \in \mathbb{R}, \ j = 0, \cdots, n-1 \}$$

It is clear that $P_{n-1}[x]$ is a linear (vector) space since

$\forall p, q \in P_{n-1}[x]$, easy to check

$$p + q \in P_{n-1}[x],$$
$$\alpha p \in P_{n-1}[x], \ \forall \alpha \in \mathbb{R}$$

Hence a map from a coefficient vector $\mathbb{c} = [c_0, \cdots, c_{n-1}]^T \in \mathbb{R}^n$ to the vectors of sampled polynomial values $y = [p(x_1), \cdots, p(x_m)]^T \in \mathbb{R}^m$ is linear! Say $y = F(\mathbb{c})$

linear map

So, according to the theorem in the previous page, $\exists A \in \mathbb{R}^{m \times n}$ for such F.

What is this matrix A?

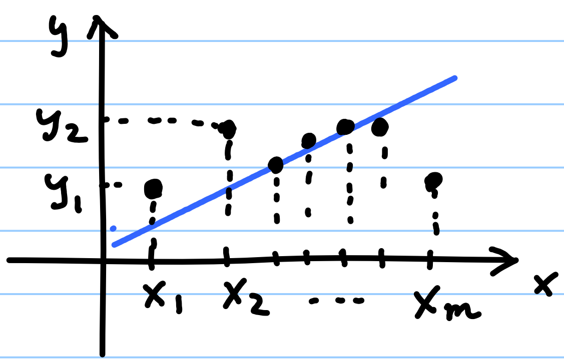
$$A = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ & & \vdots & & \\ 1 & x_m & x_m^2 & \cdots & x_m^{n-1} \end{bmatrix} \in \mathbb{R}^{m \times n}$$

which is called the $m \times n$ Vandermonde matrix.

It's clear that

$$\mathbf{y} = A\mathbf{c} \iff \begin{bmatrix} p(x_1) \\ \vdots \\ p(x_m) \end{bmatrix} = \begin{bmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ & \vdots & & \\ 1 & x_m & \cdots & x_m^{n-1} \end{bmatrix} \begin{bmatrix} c_0 \\ \vdots \\ c_{n-1} \end{bmatrix}$$

This matrix is often used in the least squares polynomial fitting to a set of measurements or noisy data.



In the case of a line fitting, $n = 2$.
But you may have many points, i.e., m: large.

Then, you want to find a line s.t., the size of $\underbrace{\mathbf{y} - A\mathbf{c}}_{\text{residual error}}$ is small.

Note in the case of a line fitting, $\mathbf{c} = \begin{bmatrix} c_0 \\ c_1 \end{bmatrix}$.

We'll discuss this problem later.

☆ Matrix-Matrix Multiplication

$$C = AB$$

$$A \in \mathbb{R}^{m \times k}, \quad B \in \mathbb{R}^{k \times n}$$

$$\Rightarrow C \in \mathbb{R}^{m \times n}$$

Note that

$$[\mathbf{c}_1 \cdots \mathbf{c}_n] = [\mathbf{a}_1 \cdots \mathbf{a}_k][\mathbf{b}_1 \cdots \mathbf{b}_n]$$

i.e., $\mathbf{c}_j = A\mathbf{b}_j \qquad 1 \le j \le n$

So each $\mathbf{c}_j$ is a linear combination of column vectors of A with the coefficient vector $\mathbf{b}_j$.

Example 1. Outer product

Let $\mathbf{u} \in \mathbb{R}^m = \mathbb{R}^{m \times 1}$,
$\mathbf{v} \in \mathbb{R}^n = \mathbb{R}^{n \times 1}$.

Then, the outer product between $\mathbf{u}$ and $\mathbf{v}$ is:

$$\mathbf{u}\mathbf{v}^T = \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} [v_1 \cdots v_n] = \begin{bmatrix} u_1v_1 & \cdots & u_1v_n \\ \vdots & & \vdots \\ u_mv_1 & \cdots & u_mv_n \end{bmatrix} \in \mathbb{R}^{m \times n}$$

This matrix has rank 1 because

$$\mathbf{u}\mathbf{v}^T = [v_1\mathbf{u}, \cdots, v_n\mathbf{u}]$$

i.e., each column is just a scalar multiple of the same vector $\mathbf{u}$.

Example 2.

$$\underset{\mathbb{R}^{m\times n}}{B} = \underset{\mathbb{R}^{m\times n}}{A}R, \quad R \in \mathbb{R}^{n\times n}$$

R : upper triangular

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

all entries below the main diagonal $\equiv 0$

$$[\mathbf{b}_1 \cdots \mathbf{b}_n] = [\mathbf{a}_1 \cdots \mathbf{a}_n] \begin{bmatrix} 1 & \cdots & & 1 \\ 0 & 1 & \cdots & 1 \\ & & \ddots & \vdots \\ 0 & & & 1 \end{bmatrix}$$

So, $\mathbf{b}_j = A\mathbf{r}_j$

$$= [\mathbf{a}_1 \cdots \mathbf{a}_n] \begin{array}{c} \begin{bmatrix} 1 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \end{array} \begin{array}{l} \left.\vphantom{\begin{matrix}1\\ \vdots\\ 1\end{matrix}}\right\} j \\ \vphantom{\begin{matrix}0\\ \vdots\\ 0\end{matrix}} \end{array}$$

$$= \mathbf{a}_1 + \cdots + \mathbf{a}_j = \sum_{k=1}^{j} \mathbf{a}_k$$