

Vectors & Matrices (Review)

Note Title

4/3/2012

★ Matrix - Vector Multiplication

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

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

i.e., m rows \times n cols

$$m \left\{ \begin{bmatrix} \rightarrow \\ \rightarrow \\ \vdots \\ \rightarrow \end{bmatrix} \begin{bmatrix} \downarrow \\ \downarrow \\ \vdots \\ \downarrow \end{bmatrix} \right\}^n$$

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

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

$$\underline{y} = A \underline{x} = [a_1 | a_2 | \dots | a_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

a_j is the j th col. vector of A .
 $\underbrace{\quad}_{\mathbb{R}^m}$

a_j is often written as $a_{\cdot j}$

$$\underline{y} = \underline{x_1 a_1 + x_2 a_2 + \dots + x_n a_n}$$

$$y_i = x_1 a_{i1} + x_2 a_{i2} + \dots + x_n a_{in}$$

$$= \sum_{j=1}^n x_j a_{ij} = \sum_{j=1}^n a_{ij} x_j, \quad 1 \leq i \leq m$$

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

$$F_A: \underbrace{\mathbb{x}}_{\mathbb{R}^n} \mapsto \underbrace{A \mathbb{x}}_{\mathbb{R}^m}$$

Then, F_A is a linear map.

i.e., $\forall \underline{x}, \underline{y} \in \mathbb{R}^n, \forall \alpha \in \mathbb{R}$.

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

Conversely, for any linear map $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$, 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. ($A(x+y) = Ax + Ay$, $A(\alpha x) = \alpha Ax$ using the definition of a matrix-vector product.)

Showing the converse is more challenging.

Let F be a linear map

Let $\{e_1, \dots, e_n\}$ be the canonical basis of \mathbb{R}^n , i.e., $e_j = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \leftarrow j$, $1 \leq j \leq n$

Set $F(e_j) = a_j \in \mathbb{R}^m$

Let $A = [a_1 \dots a_n] \in \mathbb{R}^{m \times n}$

Now pick any $x \in \mathbb{R}^n$, we can always write $x = x_1 e_1 + \dots + x_n e_n$

$$\begin{aligned} \text{Then } F(x) &= F(x_1 e_1 + \dots + x_n e_n) \\ &= F(x_1 e_1) + \dots + F(x_n e_n) \\ &= x_1 F(e_1) + \dots + x_n F(e_n) \\ &= x_1 a_1 + \dots + x_n a_n \\ &= Ax = F_A(x) \end{aligned}$$

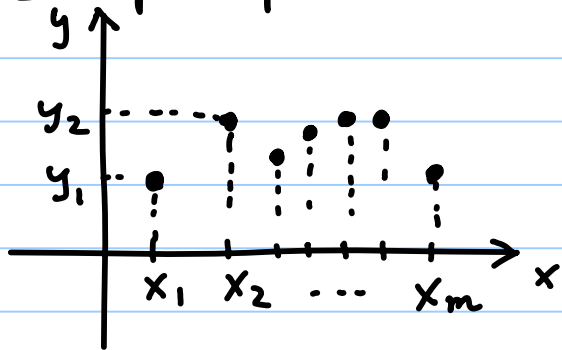
About the uniqueness, let $A, B \in \mathbb{R}^{m \times n}$

Suppose $F_A = F_B$. Then

$$F_A(e_j) = a_j = F_B(e_j) = b_j, 1 \leq j \leq n \Rightarrow A = B$$

Example: A Vandermonde Matrix

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



For example, they may be weights of m people.

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

Now consider a space of polynomials
$$\mathbb{P}_{n-1}[x] := \{p(x) = c_0 + c_1x + \dots + c_{n-1}x^{n-1},$$

$$c_j \in \mathbb{R}, j = 1, \dots, n-1\}$$

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

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

$$p + q \in \mathbb{P}_{n-1}[x],$$

$$\alpha p \in \mathbb{P}_{n-1}[x], \forall \alpha \in \mathbb{R}$$

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

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

What is this matrix A ?

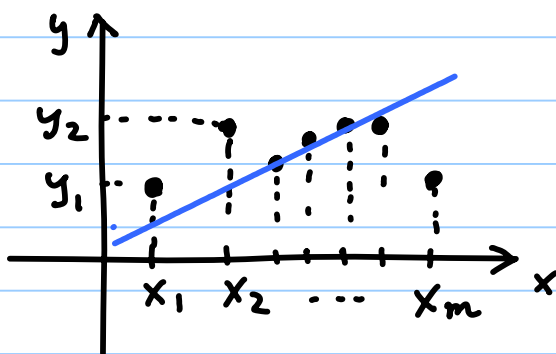
$$A = \begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_m & x_m^2 & \dots & x_m^{n-1} \end{bmatrix} \in \mathbb{R}^{m \times n}$$

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

It's clear that

$$y = A c \Leftrightarrow \begin{bmatrix} p(x_1) \\ \vdots \\ p(x_m) \end{bmatrix} = \begin{bmatrix} 1 & x_1 & \dots & x_1^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_m & \dots & x_m^{n-1} \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_{n-1} \end{bmatrix}$$

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 $y - A c$ is small.
residual error

Note in the case of a line fitting,
 $c = \begin{bmatrix} c_0 \\ c_1 \end{bmatrix}$. We'll discuss this problem later.

★ Matrix-Matrix Multiplication

$$C = A B$$

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

Note that

$$[c_1 \cdots c_n] = [a_1 \cdots a_k] [b_1 \cdots b_n]$$

i.e., $c_j = A b_j \quad 1 \leq j \leq n$

So each c_j is a linear combination of column vectors of A with the coefficient vector b_j .

Example 1. Outer product

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

Then, the outer product between u and v is:

$$u v^T = \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} [v_1 \cdots v_n] = \begin{bmatrix} u_1 v_1 & \cdots & u_1 v_n \\ \vdots & & \vdots \\ u_m v_1 & \cdots & u_m v_n \end{bmatrix}$$

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

This matrix has rank 1 because

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

i.e., each column is just a scalar multiple of the same vector u .

Example 2.

$$B = AR, \quad R: \text{upper triangular}$$

$\mathbb{R}^{m \times n} \quad \mathbb{R}^{m \times n} \quad \in \mathbb{R}^{n \times n}$

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

$$[b_1 \cdots b_n] = [a_1 \cdots a_n] \begin{bmatrix} 1 & \cdots & 1 \\ 0 & 1 & \cdots & 1 \\ \vdots & & \ddots & \vdots \\ 0 & & & 1 \end{bmatrix}$$

So, $b_j = A r_j$

$$= [a_1 \cdots a_n] \begin{bmatrix} 1 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix}_j$$
$$= a_1 + \cdots + a_j = \sum_{k=1}^j a_k$$