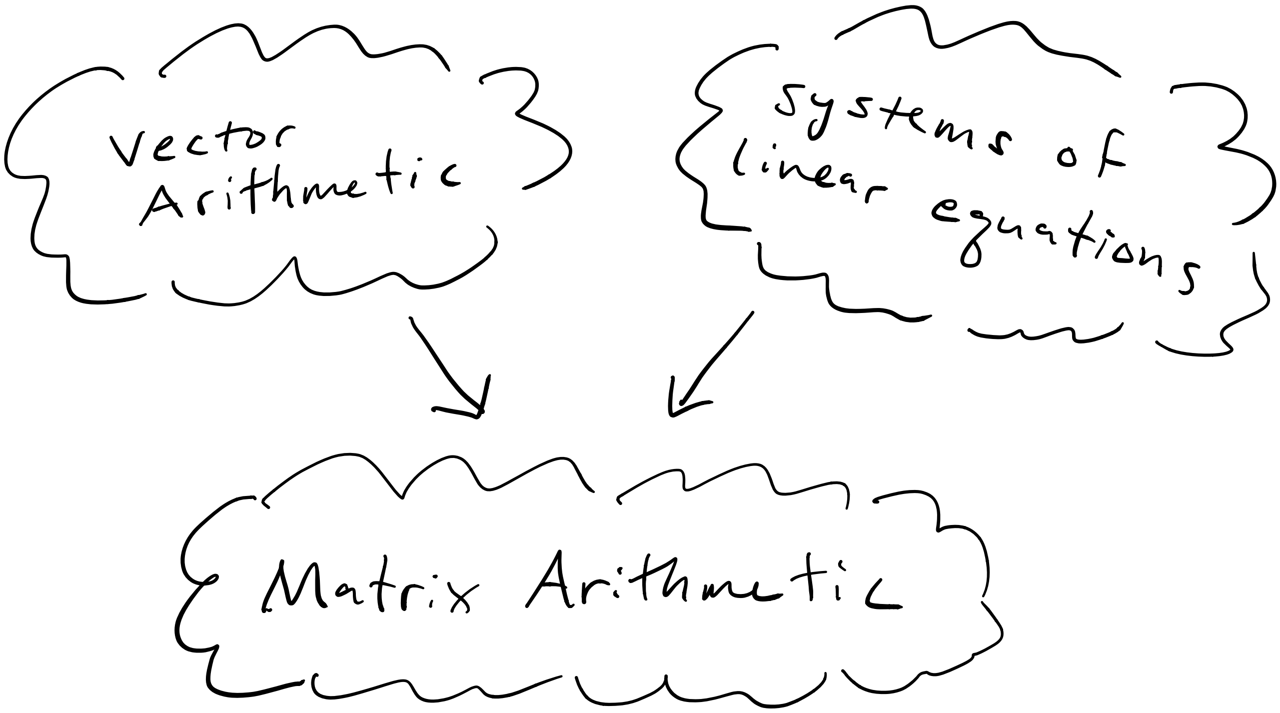
New Topic : Matrix Arithmetic



The goal is to turn all problems about linear systems into mechanical computations with matrices.

PREVIEW: The method of "least squares approximation" is one of the most important applications of linear algebra. It can be summarized with the following symbolic calculation:

$$A\vec{x} = \vec{b}$$

$$A^T A \vec{x} = A^T \vec{b}$$

$$\vec{x} = (A^T A)^{-1} A^T \vec{b}$$

[You will understand this later.]

The Key Definition:

Consider a system of m linear equations in n unknows:

$$\begin{cases} a_{11} x_1 + a_{12} x_2 + \cdots + a_{1n} x_n = b_1 \\ a_{21} x_1 + a_{22} x_2 + \cdots + a_{2n} x_n = b_2 \\ \quad \vdots \qquad\qquad\qquad\qquad\qquad\quad \vdots \\ a_{m1} x_1 + a_{m2} x_2 + \cdots + a_{mn} x_n = b_m \end{cases}$$

This takes too long to write, so we replace this with symbolic notation

$$\text{“} A \vec{x} = \vec{b} \text{”}$$

Let me explain: Let A be the $m \times n$ matrix of coefficients,

$$A = m\left\{ \overbrace{\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}}^{n} \right.$$

"A rectangular array of numbers with m rows & n columns."

Let $\vec{x} \in \mathbb{R}^n$ be the $n \times 1$ column vector of unknowns,

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Let $\vec{b} \in \mathbb{R}^m$ be the $m \times 1$ column vector of constants,

$$\vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Emphasize the Shapes:

$$\underset{\text{cancel}}{(m \times \cancel{n})\ (\cancel{n} \times 1)} \qquad (m \times 1)$$

$$m\left\{\underbrace{\begin{pmatrix} & & \\ & A & \\ & & \end{pmatrix}}_{n}\right. \left.\underbrace{\begin{pmatrix} \\ \vec{x} \\ \\ \end{pmatrix}}_{1}\right\} n = \left.\underbrace{\begin{pmatrix} \vec{b} \end{pmatrix}}_{1}\right\} m$$

Let's unpack this definition.

Given an $m \times n$ matrix A

and an $n \times 1$ column vector $\vec{x}$,

let

$$A = \left.\begin{pmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & & | \end{pmatrix}\right\} m$$

where $\vec{c}_1, \dots, \vec{c}_n \in \mathbb{R}^m$ are the $m \times 1$ column vectors of A.

Then the $m \times 1$ column $A\vec{x}$ is defined as follows:

$$A\vec{x} = \begin{pmatrix} | & | & & | \\ \vec{c}_1 & \vec{c}_2 & \cdots & \vec{c}_n \\ | & | & & | \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$:= x_1 \vec{c}_1 + x_2 \vec{c}_2 + \cdots + x_n \vec{c}_n$$

"a linear combination of the columns of A with coefficients given by the entries of $\vec{x}$."

Example: $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix}$ & $\vec{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$.

By definition:

$$A\vec{x} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$= x\begin{pmatrix}1\\1\end{pmatrix} + y\begin{pmatrix}1\\2\end{pmatrix} + z\begin{pmatrix}1\\3\end{pmatrix}$$

$$= \begin{pmatrix}x\\x\end{pmatrix} + \begin{pmatrix}y\\2y\end{pmatrix} + \begin{pmatrix}z\\3z\end{pmatrix}$$

$$= \begin{pmatrix}x+y+z\\x+2y+3z\end{pmatrix}.$$

Observe that the matrix equation

$$\begin{pmatrix}1 & 1 & 1\\1 & 2 & 3\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}2\\5\end{pmatrix}$$

expresses the linear system

$$\begin{cases} x+y+z = 2, \\ x+2y+3z = 5. \end{cases}$$

Special Case: What if the matrix A has just one row?

Let A be a $1 \times n$ row vector:

$$A = \underbrace{(a_1 \ a_2 \ \cdots \ a_n)}_{n} \big\} 1$$

Then for any $n \times 1$ column vector $\vec{x} \in \mathbb{R}^n$ we define

$$A\vec{x} = (a_1 \ a_2 \ \cdots \ a_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$= x_1 a_1 + x_2 a_2 + \cdots + x_n a_n$$

The result is a "1×1 matrix" i.e., just a scalar. In fact this is the <u>dot product</u>!

Jargon: For any matrix $A = (a_{ij})$ of shape $m \times n$, we define the transpose matrix of shape $n \times m$:

$$A^T = (a_{ji})$$

In other words,

ij entry of A^T = ji entry of A

We observe that $(A^T)^T = A$.

Example:

Let $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix}$

then $A^T = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}$

Here 3 is the $(2,3)$ entry of A,
but it is the $(3,2)$ entry of A^T.

Jargon: The symbol $\vec{x}$ always represents a <u>column</u> vector.
If we want to talk about row

vectors then we will write

$$\vec{x}^T = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}^T = (x_1 \; x_2 \; \cdots \; x_n)$$

This allows us to express the dot product of vectors in terms of matrix multiplication:

$$\vec{x} \circ \vec{y} = \vec{x}^T \vec{y}$$

(col) ∘ (col)   (row)(col)

dot product   matrix product

So we never have to use the notation "∘" again!

This gives us another way to

think about the product $A\vec{x}$, in terms of the <u>rows</u> of the matrix A. Let

$$A = \left.\underbrace{\begin{pmatrix} \text{---}\ \vec{r}_1^{\,T}\ \text{---} \\ \text{---}\ \vec{r}_2^{\,T}\ \text{---} \\ \vdots \\ \text{---}\ \vec{r}_m^{\,T}\ \text{---} \end{pmatrix}}_{n}\right\} m$$

where $\vec{r}_1, \ldots, \vec{r}_m \in \mathbb{R}^n$. Then for any column vector $\vec{x} \in \mathbb{R}^n$, I claim that

$$A\vec{x} = \begin{pmatrix} \text{---}\ \vec{r}_1^{\,T}\ \text{---} \\ \vdots \\ \text{---}\ \vec{r}_m^{\,T}\ \text{---} \end{pmatrix} \begin{pmatrix} | \\ \vec{x} \\ | \end{pmatrix}$$

$$= \begin{pmatrix} \vec{r}_1^{\,T}\vec{x} \\ \vdots \\ \vec{r}_m^{\,T}\vec{x} \end{pmatrix}$$

$$= \begin{pmatrix} \vec{r}_1 \cdot \vec{x} \\ \vdots \\ \vec{r}_m \cdot \vec{x} \end{pmatrix}$$

In other words, $A\vec{x}$ is a column vector whose i^{th} entry is

$$\vec{r}_i \cdot \vec{x} = \vec{r}_i^{\,T} \vec{x},$$

the dot product of $\vec{x}$ with the i^{th} row of matrix A.

Example:

$$A = \begin{pmatrix} 1 & 1 & 2 & 2 \\ 1 & 2 & 3 & 4 \end{pmatrix} \quad \& \quad \vec{x} = \begin{pmatrix} -1 \\ 0 \\ 2 \\ 1 \end{pmatrix}.$$

We can compute $A\vec{x}$ by columns:

$$A\vec{x} = (-1)\begin{pmatrix} 1 \\ 1 \end{pmatrix} + 0\begin{pmatrix} 1 \\ 2 \end{pmatrix} + 2\begin{pmatrix} 2 \\ 3 \end{pmatrix} + 1\begin{pmatrix} 2 \\ 4 \end{pmatrix}$$

$$= \begin{pmatrix} -1 \\ -1 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 4 \\ 6 \end{pmatrix} + \begin{pmatrix} 2 \\ 4 \end{pmatrix}$$

$$= \begin{pmatrix} -1+0+4+2 \\ -1+0+6+4 \end{pmatrix} = \begin{pmatrix} 5 \\ 9 \end{pmatrix}.$$

Or we can compute $A\vec{x}$ by rows:

$$A\vec{x} = \begin{pmatrix} (1\ 1\ 2\ 2)\begin{pmatrix} -1 \\ 0 \\ 2 \\ 1 \end{pmatrix} \\ (1\ 2\ 3\ 4)\begin{pmatrix} -1 \\ 0 \\ 2 \\ 1 \end{pmatrix} \end{pmatrix}$$

$$= \begin{pmatrix} -1+0+4+2 \\ -1+0+6+4 \end{pmatrix}$$

$$= \begin{pmatrix} 5 \\ 9 \end{pmatrix} \quad \text{SAME} \checkmark$$

The fact that we can view matrix multiplication in terms of columns <u>or</u> rows is a strength that makes the language more flexible.