Sep 29, 2025

Last time

We would like an algorithm to be

$$\begin{cases} \text{accurate} \\ \text{fast} \end{cases}$$

Ex. Inner product between $v, w \in \mathbb{R}^n$, $v = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$, $w = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$

$$\langle v, w \rangle = v^T \cdot w = [v_1 \cdots v_n] \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$$

$$= \sum_{i=1}^{n} v_i w_i$$

pseudo-code

```
p = 0;
for i = 1:n
    p = p + v(i) * w(i);
end
```

$\overset{\}}{2}$ — 2n FLOPS

Ex. Matrix-vector multiplication

$A \in \mathbb{R}^{m \times n}$, $v \in \mathbb{R}^{n \times 1}$

$$A \cdot v = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \qquad r_1 = \begin{bmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1n} \end{bmatrix}, \quad r_2 = \begin{bmatrix} a_{21} \\ \vdots \\ a_{2n} \end{bmatrix}$$

$$= \begin{bmatrix} \langle r_1, v\rangle \\ \langle r_2, v\rangle \\ \vdots \\ \langle r_m, v\rangle \end{bmatrix} \begin{matrix} \to 2n \text{ FLOPS} \\ \to 2n \text{ FLOPS} \\ \\ \to 2n \text{ FLOPS} \end{matrix} \qquad r_m = \begin{bmatrix} a_{m1} \\ \vdots \\ a_{mn} \end{bmatrix}$$

$\Rightarrow 2mn$ FLOPS . When $n = m$, $2n^2$ FLOPS

$v^T w$ needs $2n$ FLOPS

Av needs $2n^2$ FLOPS $\qquad 2n < 2n^2 \quad n > 1$

Vector inner product computation scales linearly with n

(If you double n, the number of FLOPS gets doubled, i.e., the number of FLOPS gets multiplied by 2)

Matrix-vector multiplication scales quadratically with n

(If you double n, the number FLOPS gets multiplied 2^2)

Fact: $A \cdot B$, $A, B \in \mathbb{R}^{n \times n}$, FLOPS $= 2n^3$

This scales cubically with n

Observation: Both $2n$ and $3n$ scale linearly with n.

Both $2n^2$ and $5n^2$ scale quadratically with n.

The constant prefactor does not matter as much.

Big-O notation

$v^T w$ needs $2n$ Flops, we say it is $O(n)$ as $n \to \infty$

(Big-O of n as n goes to infinity)

Similarly

$A \cdot v$ is $O(n^2)$ as $n \to \infty$

$A \cdot B$ is $O(n^3)$ as $n \to \infty$

In general, if n is large, $O(n)$ is faster than $O(n^2)$

$O(n^2)$ is faster than $O(n^3)$

Linear Algebra Review

Matrix-vector multiplication Ax , $A = \begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 3 & 7 \end{bmatrix}$, $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

Two ways

M1. By rows $\begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 + 3x_2 \\ 2x_1 + 4x_2 \\ 3x_1 + 7x_2 \end{bmatrix}$

☆ Inner products of rows of A with x $\rightarrow$ useful for computation

M2. By columns. $A = [a_1 \ \ a_2]$ $a_1 = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}$, $a_2 = \begin{bmatrix} 3 \\ 4 \\ 7 \end{bmatrix}$

$$Ax = \begin{bmatrix} a_1 & a_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 a_1 + x_2 a_2$$

$$= x_1 \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 4 \\ 7 \end{bmatrix}$$

☆ Linear combination of the columns of A.

→ useful for understanding.

Ax is a linear combination of columns of A.

Definition. All combinations of the columns of fill out the column space of A. We use $C(A)$ to denote the column space of A.

$\dim(C(A))$ is the rank of A.

Implication: 1. $Ax = b$ has a solution x iff b is in the column space $C(A)$

2. rank-1 matrix.

Let $v = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{R}^{n\times 1}$, $w = \begin{bmatrix} w_1 \\ \vdots \\ w_m \end{bmatrix} \in \mathbb{R}^{m\times 1}$

Define $A = \underline{v \cdot w^T} \in \mathbb{R}^{n\times m}$ is called a rank-1 matrix.

b.c. $Ax = \underbrace{v}\cdot\underbrace{(w^T x)}_{\in \mathbb{R}}$ is a multiple of v.

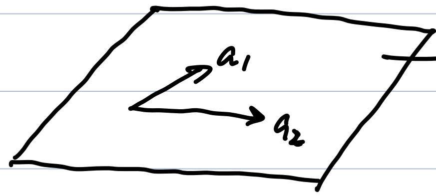
$x \in \mathbb{R}^{m\times 1}$

"low-rank" matrix. $A = v^{(1)} \cdot w^{(1)T} + v^{(2)} \cdot w^{(2)T} + \cdots + v^{(r)} \cdot w^{(r)T}$

for a small r.

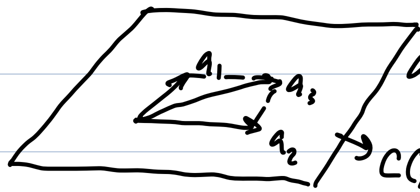
Ex. Visualize the column space of $A = \begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 3 & 7 \end{bmatrix} = [a_1 \ a_2]$

$a_1, a_2 \in \mathbb{R}^3$



All linear combinations of a_1 and a_2 form a plane in $\mathbb{R}^3$ (The plane spanned by a_1 and a_2)

Ex. Visualize the column space of $A = \begin{bmatrix} 2 & 3 & 5 \\ 2 & 4 & 6 \\ 3 & 7 & 10 \end{bmatrix} = [a_1 \ a_2 \ a_3]$



$a_3 = a_1 + a_2$

$C(A)$ = the plane spanned by a_1 and a_2.

a_3 is a "dependent" column that does not go beyond the original plane.

In order for a_1, a_2, a_3 to span the whole $\mathbb{R}^3$, they have to be "linearly independent"

Def. a_1, a_2, a_3 are linearly independent if none of them can be expressed as a linear combination of the other vectors.

Equivalently, the only combination that gives the zero vector is $0 \cdot a_1 + 0 \cdot a_2 + 0 \cdot a_3$

[Exercise. prove the equivalence of the two statements]

Implication.

a_1, a_2, a_3 linearly independent $\Leftrightarrow$ $C(A) = \mathbb{R}^3$

$\Leftrightarrow$ $Ax = b$ has a solution x for every $b \in \mathbb{R}^3$

[In fact, the solution x is unique]

Why is x unique?

Def. All solutions to $Ax = 0$ forms the null space of A, denoted by $N(A)$

Its dimension is called nullity of A.