Linear algebra for AI and ML

September - 8

(Lecture - 9)

$Ax = b$

- Existence & uniqueness of soln ✓
- LU / QR – compute x ✓
- Sensitivity analysis.

Matrix norms:

Matrices $\mathbb{R}^{m\times n}$
- view as a vectors in $\mathbb{R}^{mn}$ (vectorization)
- Linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$

$x \in \mathbb{R}^n \longmapsto Ax \in \mathbb{R}^m$

$\|\cdot\| : \mathbb{R}^{n\times n} \longrightarrow \mathbb{R}$

Norm: $\forall\ A, B \in \mathbb{R}^{n\times n}, \alpha, \beta \in \mathbb{R}$

(i) $\|A\| \geq 0$ and $\|A\| = 0 \iff A = 0$

(ii) $\|\alpha A\| = |\alpha|\,\|A\|$ (iii) $\|A + B\| \leq \|A\| + \|B\|$

(iv) $\|AB\| \leq \|A\|\,\|B\|$ . ... (submultiplicativity)

Ex: Frobenius norm:

$$\|A\|_F = \left(\sum_{i=1}^{n} \sum_{j=1}^{n} |A_{ij}|^2 \right)^{1/2}$$

(This norm is basically 2-norm of a vector of size n^2.)

Since 2-norm is a norm on $\mathbb{R}^{n^2}$, first three properties of matrix-norm are automatically satisfied. For submultiplicativity, we use C-S.I.

$$\underbrace{C = AB}$$

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

$A = [a_{ij}]$

$B = [b_{ij}]$

$C = [c_{ij}]$

$$\|AB\|_F^2 = \|C\|_F^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} |c_{ij}|^2$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \boxed{\left|\sum_{k=1}^{n} a_{ik}b_{kj}\right|^2}$$

Cauchy Schwartz inequality

$$\leq \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\sum_{k=1}^{n}|a_{ik}|^2 \sum_{k=1}^{n}|b_{kj}|^2\right)$$

$$= \left(\sum_{i=1}^{n}\sum_{k=1}^{n}|a_{ik}|^2\right)\left(\sum_{j=1}^{n}\sum_{k=1}^{n}|b_{kj}|^2\right)$$

$$= \|A\|_F^2\,\|B\|_F^2 \qquad \blacksquare$$

$\Rightarrow \|\cdot\|_F$ is a matrix norm.

Ex: Viewing the matrix as a linear operator.

$$\|A\|_2 = \max_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} \quad \forall A \in \mathbb{R}^{n \times n}$$

$1 \leq p \leq \infty$

$p = 1, \infty, 2$

(induced-norm)

(operator norm)

Is $\|\cdot\|_2$ a norm??

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

$\|x\|_2$ $\quad\quad$ $\|Ax\|_2$

$$\text{Ratio} = \frac{\|Ax\|_2}{\|x\|_2}$$

← magnification induced by A in x (by action Ax)

$\|A\|_2$: maximum magnification

$A \in \mathbb{R}^{2\times 2}$

$$A = \begin{bmatrix} 10 & 9 \\ 9 & 8 \end{bmatrix} \in \mathbb{R}^{2\times 2}$$

$x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$; $\|x\|_2 = \sqrt{2}$

$\downarrow$

$Ax = \begin{pmatrix} 19 \\ 17 \end{pmatrix}$; $\|Ax\|_2 = \sqrt{19^2 + 17^2} = \sqrt{650}$

$$\frac{\|Ax\|_2}{\|x\|_2} = \frac{\sqrt{650}}{\sqrt{2}} = \sqrt{325} \approx 18$$

$x = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$; $\|x\|_2 = \sqrt{2}$

$Ax = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$; $\|Ax\|_2 = \sqrt{2}$

$$\frac{\|Ax\|_2}{\|x\|_2} = \frac{\sqrt{2}}{\sqrt{2}} = 1$$

Thm: $\|Ax\|_2 \leq \|A\|_2 \, \|x\|_2$

where $\|A\|_2 = \max\limits_{x \neq 0} \dfrac{\|Ax\|_2}{\|x\|_2}$

Pf: For $x = 0$, trivially true.

For any $x \neq 0$,

$$\boxed{\frac{\|Ax\|_2}{\|x\|_2}} \leq \max_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \|A\|_2$$

$$\Rightarrow \frac{\|Ax\|_2}{\|x\|_2} \leq \|A\|_2$$

$$\Rightarrow \|Ax\|_2 \leq \|A\|_2 \, \|x\|_2 \quad x$$

This result helps in proving submultiplicative property for $\|\cdot\|_2$ norm.

To prove: $\|A\|_2 = \max_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2}$ is a norm.

i) $\|A\|_2 \geq 0 \quad \forall A$

$$\|A\|_2 = \underbrace{\max_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} \geq 0}$$

in fact, if $A \neq 0$, then $\exists$ a nonzero vector $\hat{x}$ s.t. $A\hat{x} \neq 0$.

$$\|A\|_2 = \max_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} \geq \frac{\|A\hat{x}\|_2}{\|\hat{x}\|_2} > 0$$

ii) $\|\alpha A\|_2 = |\alpha| \|A\|_2$

iii) triangle inequality:

$$\|A+B\|_2 = \max_{x\neq 0} \frac{\|(A+B)x\|_2}{\|x\|_2} = \max_{x\neq 0} \frac{\|Ax+Bx\|_2}{\|x\|_2}$$

triangle inequality of $\|\cdot\|_2$ vector norm

$$\leq \max_{x\neq 0} \frac{\|Ax\|_2+\|Bx\|_2}{\|x\|_2}$$

$$\leq \max_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2} + \max_{x\neq 0} \frac{\|Bx\|_2}{\|x\|_2}$$

$$= \|A\|_2 + \|B\|_2$$

iv) Replace x by Bx in the previous thm:

$$\|ABx\|_2 \leq \|A\|_2 \|Bx\|_2 \leq \|A\|_2 \|B\|_2 \|x\|_2$$

$$\|AB\|_2 = \max_{x\neq 0} \frac{\|ABx\|_2}{\|x\|_2} \leq \|A\|_2 \|B\|_2$$

$\blacksquare$

Ex: $I \in \mathbb{R}^{n \times n}$ identity matrix.

$$\|I\|_F = \sqrt{n}$$

$$\|I\|_2 = 1$$

$$\|Q\|_2 = 1$$ (where $Q \in \mathbb{R}^{n \times n}$ is orthogonal).

$\forall x \in \mathbb{R}^n$; $\|Qx\|_2 = \|x\|_2$

$$\|Q\|_2 = \max_{x \neq 0} \frac{\|Qx\|_2}{\|x\|_2} = \max_{x \neq 0} \frac{\|x\|_2}{\|x\|_2} = \max_{x \neq 0} 1$$

$$= 1$$

Suppose $A \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^n$ and A is <u>invertible</u>.

We know that $\exists$ a unique x s.t.

$$Ax = b \quad \text{—(1)}$$

Assume that b is perturbed to $b + \delta b$.

then $\exists$ a unique $\hat{x}$ s.t.

$$A\hat{x} = b + \delta b \quad \text{—(2)}$$

suppose define $\delta x \in \mathbb{R}^n$ s.t $\hat{x} = x + \delta x$

$$A(x + \delta x) = b + \delta b$$

$$\Rightarrow \quad Ax + A\delta x = b + \delta b$$

From (1) $\Rightarrow \quad A\delta x = \delta b \Rightarrow \delta x = A^{-1}\delta b \quad \text{—(3)}$

From (3),

$$\|\delta x\|_2 = \|A^{-1}\delta b\|_2$$

$$\|\delta x\|_2 \leq \|A^{-1}\|_2 \|\delta b\|_2 \quad \text{--- (4)}$$

Further, from (1), $b = Ax$

$$\Rightarrow \|b\|_2 = \|Ax\|_2$$

$$\Rightarrow \|b\|_2 \leq \|A\|_2 \|x\|_2$$

$$\Rightarrow \frac{1}{\|x\|_2} \leq \|A\|_2 \frac{1}{\|b\|_2} \quad \text{--- (5)}$$

From (4) & (5);

$$\boxed{\frac{\|\delta x\|_2}{\|x\|_2} \leq \boxed{\|A\|_2 \|A^{-1}\|_2} \frac{\|\delta b\|_2}{\|b\|_2}}$$

$\kappa_2(A)$: Condition number.