Linear algebra for AI and ML

September - 9

(Lecture - 10)

Sensitivity analysis:

A : square ; A is invertible.

$$Ax = b$$

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

$$\boxed{\frac{\|\delta x\|_2}{\|x\|_2} \leq \underbrace{\boxed{\|A\|_2 \|A^{-1}\|_2}}_{= k_2(A)} \frac{\|\delta b\|_2}{\|b\|_2}}$$

If $k_2(A) = \|A\|_2 \|A^{-1}\|_2$ is "small", then

$\frac{\|\delta b\|_2}{\|b\|_2}$ "small" $\Rightarrow$ $\frac{\|\delta x\|_2}{\|x\|_2}$ is also small.

This implies that the solution is not sensitive, in this case, to perturabations in b.

On the other hand, if $\underbrace{\kappa_2(A)}$ is "large", then small perturbations $\frac{\|\delta b\|_2}{\|b\|_2}$ may result in "large" perturabations $\frac{\|\delta x\|_2}{\|x\|_2}$.

(Refer to the example we discussed. We will calculate $\kappa_2(A)$ where $A = \begin{bmatrix} 1000 & 999 \\ 999 & 998 \end{bmatrix}$.)

This implies, the solution x may be very sensitive to "small" perturabations $\frac{\|\delta b\|_2}{\|b\|_2}$.

Notice: $k_2(I) = \|I\|_2 \|I^{-1}\|_2$

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

$= 1$

$\|I\|_2 = \max_{x \neq 0} \dfrac{\|Ix\|_2}{\|x\|_2}$

$= \max_{x \neq 0} 1$

$= 1$

for any invertible matrix $A \in \mathbb{R}^{n \times n}$, note that

$I = AA^{-1}$

$\Rightarrow 1 = \|I\|_2 = \|AA^{-1}\|_2 \leq \|A\|_2 \|A^{-1}\|_2 = k_2(A)$

$\Rightarrow \boxed{k_2(A) \geq 1}$

"Best possible" condition number is 1.

(Note that orthogonal matrices have "best possible" condition number).

Geometrical Interpretation.

$$\kappa_2(A) = \underbrace{\|A\|_2}\, \|A^{-1}\|_2$$

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

$$= \max_{x \neq 0} \left\| A \frac{x}{\|x\|_2} \right\|_2$$

for $x \neq 0$

$\dfrac{x}{\|x\|_2}$ is a unit norm vector.

$$\alpha = \|x\|_2$$

$$= \max_{x \neq 0} \frac{\|Ax\|_2}{\alpha}$$

$$= \max_{x \neq 0} \left\| \frac{1}{\alpha} Ax \right\|_2$$

$$= \max_{x \neq 0} \left\| A \frac{x}{\alpha} \right\|_2$$

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

$$= \max_{x \neq 0} \left\| A \frac{x}{\|x\|_2} \right\|_2$$

$$\boxed{\|A\|_2 = \max_{\|y\|_2 = 1} \|Ay\|_2}$$

maximum magnification of A

$$\max\operatorname{mag}(A) = \max_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \underbrace{\max_{\|x\|_2 = 1} \|Ax\|_2}$$

minimum magnification of A

$$\min\operatorname{mag}(A) = \min_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \underbrace{\min_{\|x\|_2 = 1} \|Ax\|_2}$$

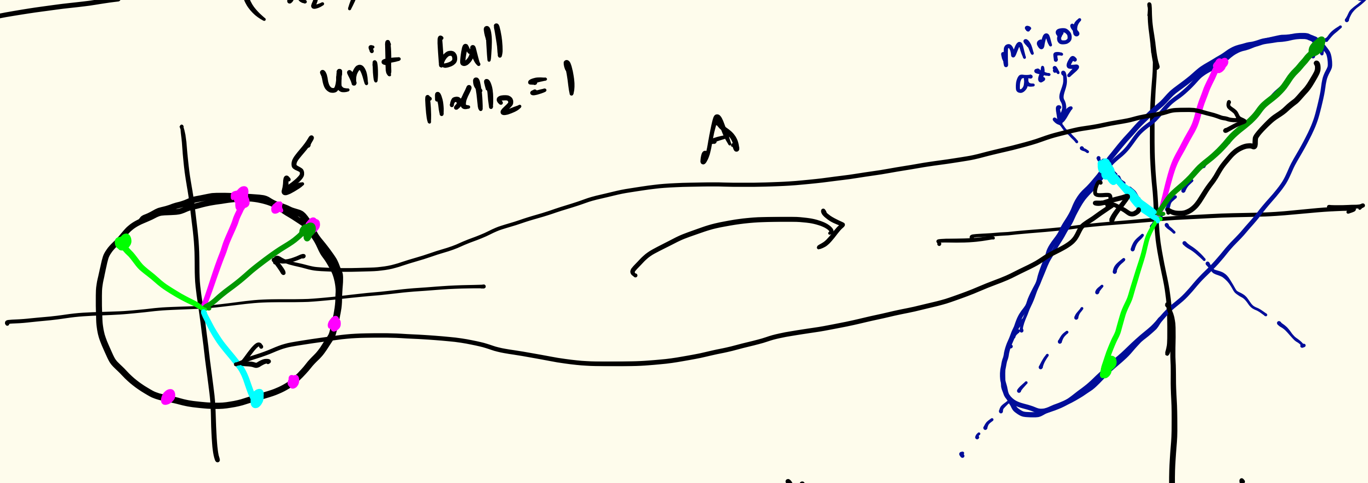
By definition, $\|A\|_2 = \max\operatorname{mag}(A)$

Note: For orthogonal matrices, $\max\operatorname{mag}(Q)$ as well as $\min\operatorname{mag}(Q)$ are 1.

$A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ $\quad$ $(A \in \mathbb{R}^{n \times n}$, invertible$)$

$x \longmapsto Ax$

$\boxed{n=2}$ $\quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \longmapsto A\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix}$



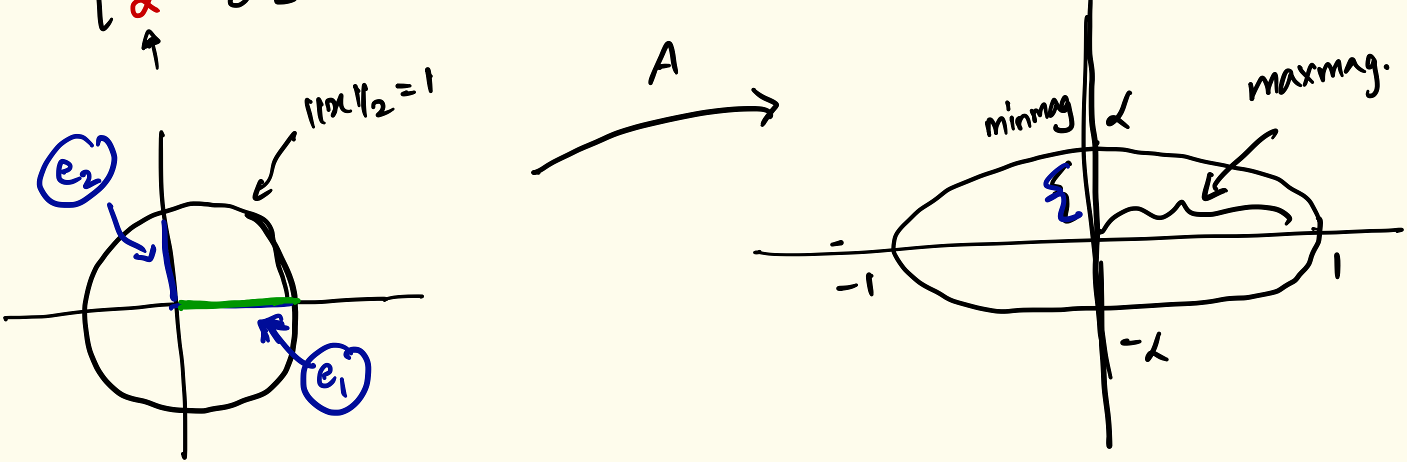
Geometrically, $\text{maxmag}(A)$ = half length of the major axis.

$\text{minmag}(A)$ = half length of the minor axis.

$A = \begin{bmatrix} 0 & 1 \\ \alpha & 0 \end{bmatrix}$

$\alpha \neq 0$; $\alpha > 0$

$\alpha < 1$



domain

$x_1^2 + x_2^2 = 1$

$$Ax = \begin{bmatrix} 0 & 1 \\ \cancel{0}\,\alpha & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ \cancel{0+}\,\alpha\, x_1 \end{bmatrix}$$

$Ae_1 = \begin{bmatrix} 0 \\ \alpha \end{bmatrix}$

$Ae_2 = \begin{bmatrix} \cancel{0+1} \\ 0 \end{bmatrix} \alpha = \begin{bmatrix} \alpha \\ 0 \end{bmatrix}$