Advanced Matrix Algebra and Applications

— Swanand Khare

Books :

1) Fundamentals of matrix computations
— David watkins

2) Matrix Algebra...
— S. Boyd

Lecture 1 :

Inner product. (Dot product)

$\mathbb{R}^n$ / $\mathbb{R}^2$ / $\mathbb{R}^3$

$x, y \in \mathbb{R}^n$ $\quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$

$$\langle x, y \rangle = x^T y = \sum_{i=1}^{n} x_i y_i$$

$$\langle x, x \rangle = x^T x = \sum_{i=1}^{n} x_i^2 = \|x\|_2^2$$

$$\cos\theta = \frac{\langle x, y \rangle}{\|x\| \|y\|}$$

C.S. inequality $\Rightarrow -1 \leq \cos\theta \leq 1$

$x, y \in \mathbb{R}^n$ are such that

$\langle x, y \rangle = x^T y = 0$

LU

Cholesky

QR

SVD $\cancel{x}$

$Ax = b$ System of linear equations.

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

$b \in \mathbb{R}^{m}$

Q1: Does there exist $x \in \mathbb{R}^n$ s.t. $Ax = b$??

$$A = \begin{bmatrix} | & | & & | \\ A_1 & A_2 & \cdots & A_n \\ | & | & & | \end{bmatrix} ; \quad x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

$$Ax = \begin{bmatrix} A_1 & \cdots & A_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 A_1 + x_2 A_2 + \cdots + x_n A_n$$

Col span $(A) = \{ \alpha_1 A_1 + \alpha_2 A_2 + \cdots + \alpha_n A_n \mid$
$\alpha_1, \ldots \alpha_n \in \mathbb{R}$ and
$A_1, \ldots, A_n$ are columns
of $A \}$

- very easy to prove that col span(A) is a subspace of $\mathbb{R}^m$

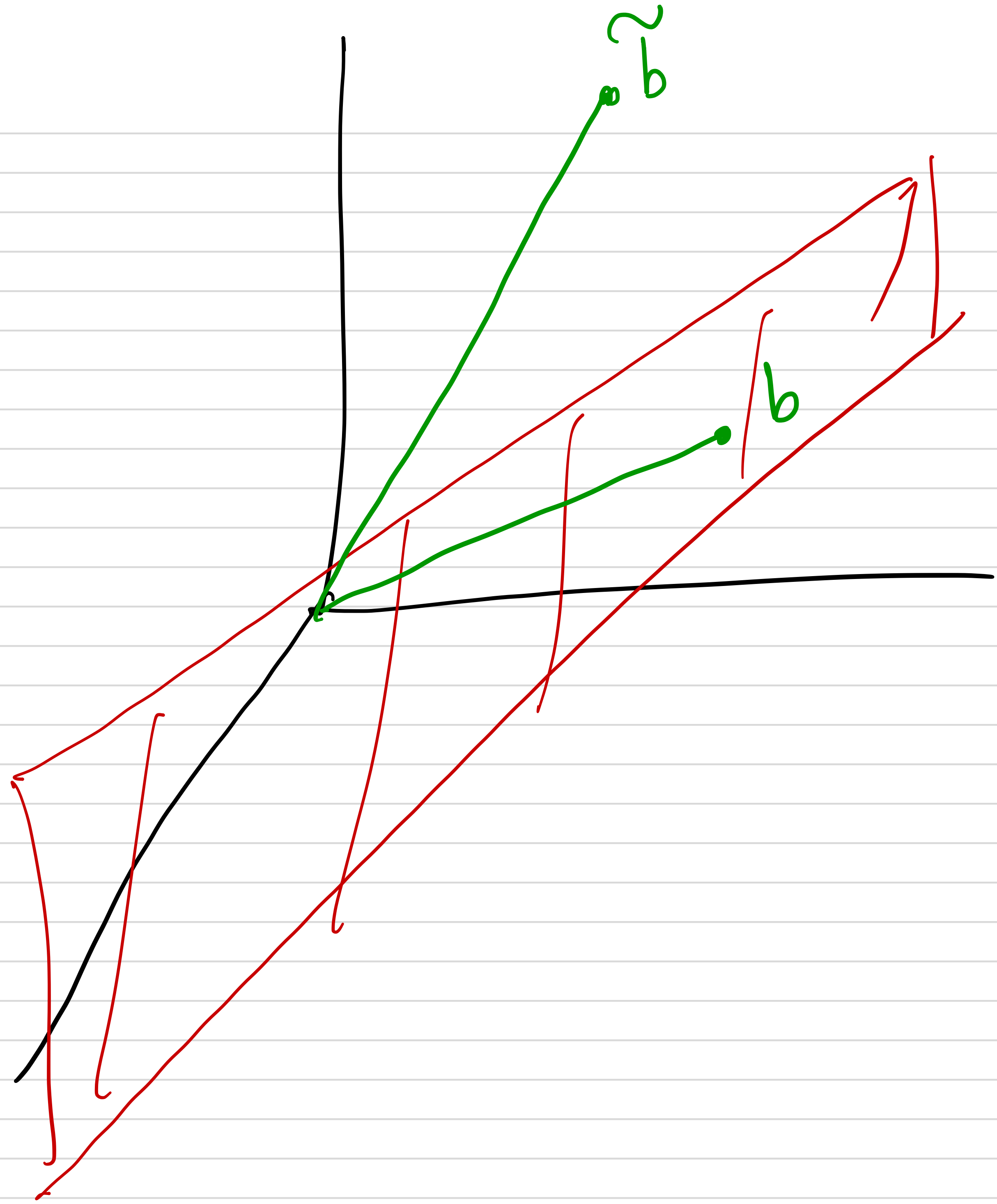
If $b \in \mathbb{R}^m$ is in the col span (A), then b can be written as a linear combination of A_i's.

$\exists\ x_1, x_2, \ldots, x_n \in \mathbb{R}$ s.t.

$$\sum_{i=1}^{n} A_i x_i = b$$

$\Rightarrow \quad Ax = b$

$\Rightarrow$ Solution exists!!



2) Uniqueness

The columns of A are in fact basis vectors for the column space of A.

<u>Ex:</u> Relate the geometrical concepts with algebraic constraints on rank $[A \vdots b]$

Solve:

$$Ax = b$$

Direct methods

Iterative methods

$$\begin{bmatrix} U_{11} & U_{12} & \cdots & U_{1n} \\ 0 & U_{22} & \cdots & U_{2n} \\ 0 & & \ddots & \\ 0 & 0 & & U_{nn}\end{bmatrix}\begin{bmatrix}x_1\\ \vdots \\ x_n\end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_n\end{bmatrix}$$

$$Ux = b$$

$$u_{nn}\, x_n = b_n$$

$$x_n = \frac{1}{u_{nn}}\, b_n$$

$$u_{n-1,n-1}\, x_{n-1} + u_{n-1,n}\, x_n = b_{n-1}$$

$$x_{n-1} = \frac{1}{u_{n-1,n-1}}\left[b_{n-1} - u_{n-1,n}\, x_n \right]$$

Floating Point Operations   Flops

Floating Point Arithmatic

Similarly, $Lx = b$ where L is an lower triangular system can also be solved very easily by forward substitution.

In general, we will try to convert $Ax = b$ problem into an equivalent upper triangular or lower triangular system.

LU, cholesky, QR

---

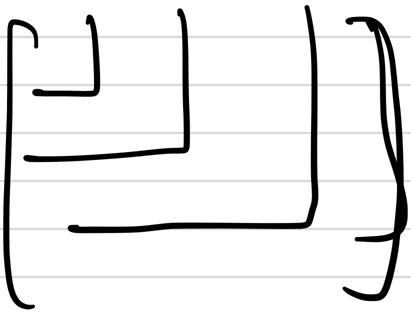
symmetric positive definite matrices.

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

A : symmetric.

$$x^T A x > 0 \quad \forall x \in \mathbb{R}^n$$

$Ax = b$ where symmetric positive definite.

$A =$  principal minors

Lecture 2 :

- Gaussian Elimination (pivoting)

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} ; \quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

$$[A \,\vdots\, b]$$

$$\boxed{R_2 - \frac{a_{21}}{a_{11}} R_1}$$

$$\Big\downarrow$$

$$\begin{bmatrix} \diagdown \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \tilde{b}_1 \\ \vdots \\ \tilde{b}_n \end{bmatrix}$$

- LU

$A \in \mathbb{R}^{n \times n}$ non-singular.

$$A = LDU$$

$$\begin{bmatrix} 1 & & 0 \\ & \ddots & \\ * & & 1 \end{bmatrix} \begin{bmatrix} d_1 & & 0 \\ & d_2 \ddots & \\ 0 & & d_n \end{bmatrix} \begin{bmatrix} 1 & & * \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$$

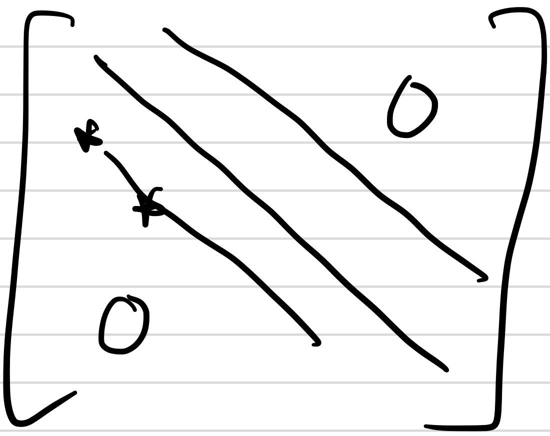
$$Ax = b$$

$$L\underbrace{[Ux]}_{y} = b$$

$$Ly = b$$

$$Ux = y$$

$$Ax = b$$



$A \in \mathbb{R}^{n\times n}$ symmetric

$$A = LL^T$$

A is symmetric p.d.

$$A = LDL^T$$ where D has only positive diagonal entries.

$$A = GG^T$$

where $G = LD^{1/2}$

To solve $Ax = b$

In practice we generally solve a perturbed system.

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

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

Suppose that one can solve $\tilde{A}\tilde{x} = \tilde{b}$ exactly for $\tilde{x}$.

Then we declare $\tilde{x}$ to be the solution to $Ax = b$.

Can we justify this??

Answer: Not always!! (Sensitivity analysis)

Ex: $Ax = b$

$$A = \begin{bmatrix} 1000 & 999 \\ 999 & 998 \end{bmatrix} \quad ; \quad b = \begin{bmatrix} 1999 \\ 1997 \end{bmatrix}$$

Note: $x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is the solution.

$$\tilde{b} = \begin{bmatrix} 1998.99 \\ 1997.01 \end{bmatrix} \qquad \tilde{b} = b + \Delta b$$

$$\Delta b = \begin{bmatrix} -0.01 \\ 0.01 \end{bmatrix} *$$

Now solve:

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

$$\tilde{x} = \begin{bmatrix} 20.97 \\ -18.99 \end{bmatrix} *$$

$$A^{-1} = \begin{bmatrix} -998 & 999 \\ 999 & -1000 \end{bmatrix} \qquad \tilde{b} = \begin{bmatrix} 1999.01 \\ 1997.01 \end{bmatrix}$$

$$\Delta b = \begin{bmatrix} 0.01 \\ 0.01 \end{bmatrix}$$

$$\tilde{x} = A^{-1} \tilde{b}$$

$$= \begin{bmatrix} -998 & 999 \\ 999 & -1000 \end{bmatrix} \begin{bmatrix} 1999.01 \\ 1997.01 \end{bmatrix}$$

$$\tilde{x} = \begin{bmatrix} -998 \times 1999.01 + 999 \times 1997.01 \\ \\ \end{bmatrix}$$

$$\tilde{x} = \begin{bmatrix} 1.01 \\ 0.99 \end{bmatrix}$$

We observe here that a "small" perturbation is b causes a "very large" perturbation in x. $\ast$

Q: Can we quantify this perturbation in x??

Norms :

vector norm.

$\|.\| : \mathbb{R}^n \to \mathbb{R}$

i) $\|x\| \geq 0 \quad \forall x \in \mathbb{R}^n$

and $\|x\| = 0$ iff $x = 0$

ii) $\|\alpha x\| = |\alpha| \, \|x\|$

iii) $\|x + y\| \leq \|x\| + \|y\|$

(•) $\|x\|_2 = \sqrt{\langle x, x \rangle} = \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2}$

(•) $\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p} \qquad p \geq 1$

is a vector norm.

(•) $\|x\|_1 = \sum_{i=1}^{n} |x_i|$

(•) $\|x\|_\infty = \max_i |x_i|$

Matrix norms:

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

Consider A as a vector in $\mathbb{R}^{n^2}$.

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

is in fact the 2-norm of the vector in $\mathbb{R}^{n^2}$.

Induced matrix norm:

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

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

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

Lecture 4.

Least squares problem.

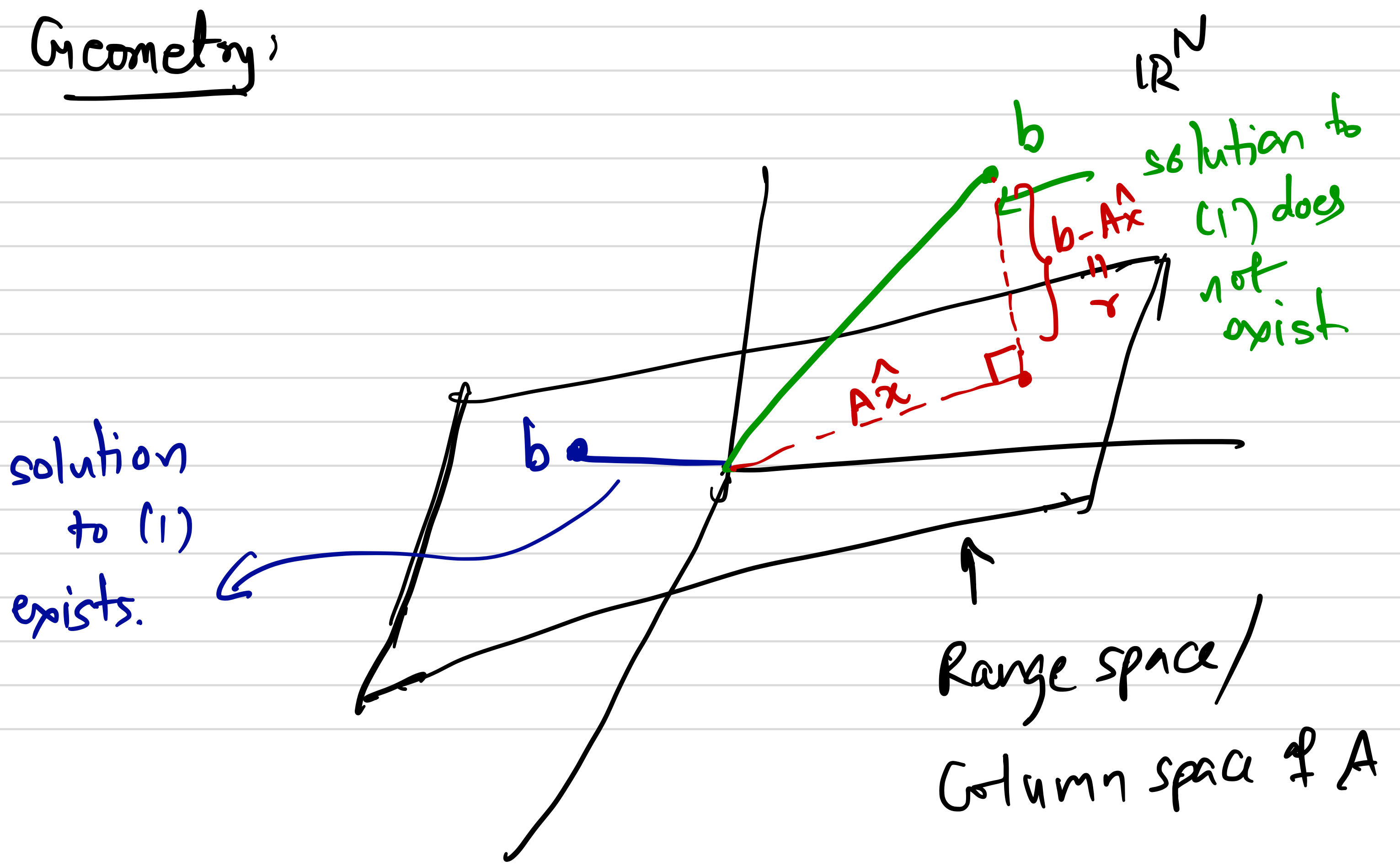
$A \in \mathbb{R}^{N \times n}$ $\qquad N >> n.$ } Given
$b \in \mathbb{R}^{N}$

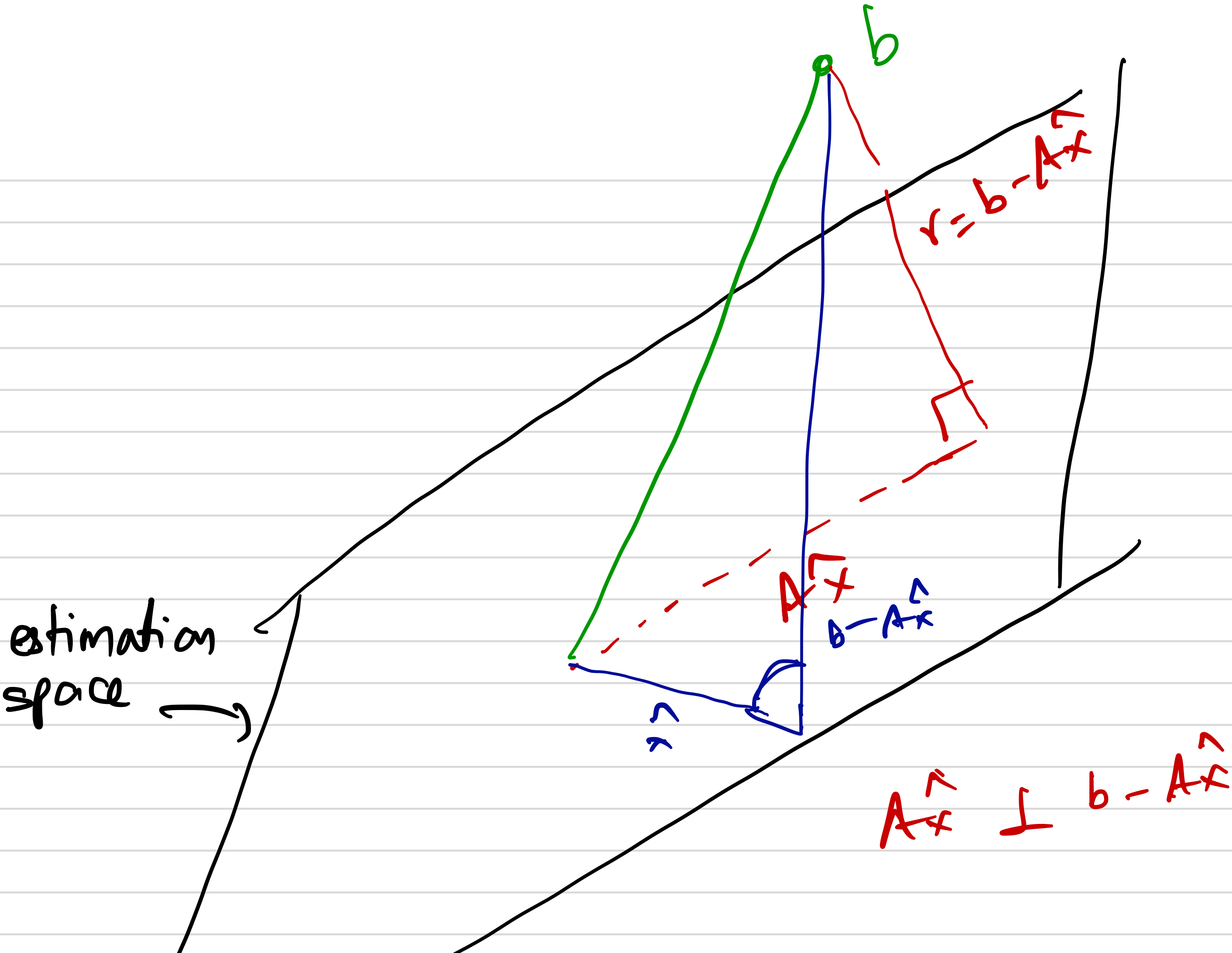
Solve: $\quad A\mathbf{x} = b$

$$\begin{bmatrix} & & \\ & A & \\ & & \\ & & \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} \\ b \\ \\ \\ \end{bmatrix} \quad \text{— (1)}$$

over-determined system of linear eq^{n}s.

Geometry:





Find $\hat{x}$ such that

$$A\hat{x} \perp b - A\hat{x}$$

$$\Rightarrow \langle A\hat{x},\, b - A\hat{x} \rangle = 0$$

$$\Rightarrow \hat{x}^T A^T (b - A\hat{x}) = 0$$

$$\Rightarrow \hat{x}^T A^T b = \hat{x}^T A^T A \hat{x}$$

why??

$$\Rightarrow \boxed{A^T b = A^T A \hat{x}}$$

← Normal Eq^{ns}.

$A^T A \in \mathbb{R}^{n \times n}$
$A^T b \in \mathbb{R}^{n \times 1}$

Solving the normal eq^{ny}:

$$(A^T A)\,\hat{x} = A^T b$$

we get the LS solution to the system $Ax = b$

In particular, if $A^T A$ is invertible, then $\boxed{\hat{x} = (A^T A)^{-1} A^T b}$

The matrix $(A^T A)^{-1} A^T$ is known as pseudo-inverse of A and is denoted as A^+.

Q: Why is this solution called Least Squares (LS) ??

$$\min_{x \in \mathbb{R}^n} \|b - Ax\|_2$$

$$\min_{x \in \mathbb{R}^n} \|b - Ax\|_2^2 \quad (**)$$

$$\|b - Ax\|_2^2 = \langle b - Ax, b - Ax \rangle$$

$$= (b - Ax)^T (b - Ax)$$

$$= (b^T - x^T A^T)(b - Ax)$$

$$= b^T b - \underline{x^T A^T b - b^T A x} +$$

$$x^T A^T A x$$

$$= x^T A^T A x - \underline{2 x^T A^T b} + b^T b$$

$$\min_{x \in \mathbb{R}^n} \underbrace{x^T A^T A x - 2 x^T A^T b + b^T b}_{\text{cost } f^n}$$

$$\nabla_x (\text{cost } f^n) = 0$$

$$2 A^T A x - 2 A^T b = 0$$

$$\Rightarrow \quad A^T A x = A^T b$$

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

$$\langle b - Ax, b - Ax \rangle$$

$$b = \begin{bmatrix} \\ \\ \\ \end{bmatrix}_{N \times 1} \quad Ax = \begin{bmatrix} \\ \\ \\ \end{bmatrix}_{N \times 1}$$

$$r = b - Ax = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_N \end{bmatrix}_{N \times 1}$$

$$\langle b - Ax, b - Ax \rangle = \langle r, r \rangle$$

$$= \sum_{i=1}^{n} r_i^2$$

Given A & b, to solve the LS problem, one needs to construct normal equations and then solve them simultaneously.

$$\underbrace{A^T A}\,\hat{x} = \underbrace{A^T b}$$

i) computational complexity of computing $A^T A$.

ii) $$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & \varepsilon \end{bmatrix}$$

$$A^T A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & \varepsilon \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & \varepsilon \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 2 \\ 2 & 2+\varepsilon^2 \end{bmatrix} \quad (A)$$

Numerical approach to solve LS problem.

$$Ax = b \qquad A \in \mathbb{R}^{N \times n}$$

$$b \in \mathbb{R}^{N \times 1}$$

To find x s.t. $\|b - Ax\|_2^2$ is minimized.

$$\|b - Ax\|_2^2 = \langle b - Ax, b - Ax \rangle$$

$$= \langle Q(b - Ax), Q(b - Ax) \rangle$$

$$= (b - Ax)^T \underbrace{Q^T Q}_{I} (b - Ax)$$

If Q is such that $Q^T Q = I$

$$\|Q^T (b - Ax)\|_2^2 = \|b - Ax\|_2^2$$

$$\|$$

$$\|Q^T b - \underbrace{Q^T A}_{R = \begin{bmatrix} \diagdown \end{bmatrix}} x\|_2^2 = \|Q^T b - Rx\|_2^2$$