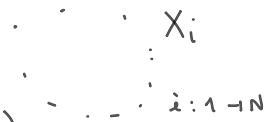
§3. Least squares optimization

80% of optimizations — AI
— computer graphics

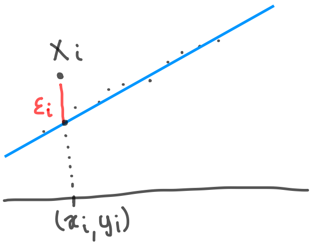
⇝ another name : regressions.

I - Least squares problems

Definition $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is defined like $\boxed{f(x) = \sum_{i=1}^{N} \left(f_i(x) \right)^2}$



Plane fitting on 3D pts (§1)



Eq. ~~$ax+by+cz+d=0$~~

$z = a_1 x + a_2 y + a_3$
(Cartesian eq)

$$f(X) = \sum_i \underbrace{\left(a_1 x_i + a_2 y_i + a_3 - z_i \right)}_{\varepsilon_i}{}^2$$

$$X = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

↓

linear least square fitting

$(\varepsilon_i$ are linear$)$
$\varepsilon_i = f_i(x)$

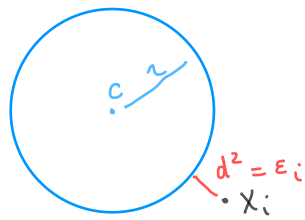
↑
here

Circles fitting in 2D

$n = 3 \quad \longleftrightarrow \quad X = (c, r)$

$$f(x) = \sum \left(\underbrace{\|X_i - c\|^2 - r^2}_{\varepsilon_i} \right)^2$$

$$x = \begin{pmatrix} c \\ r \end{pmatrix}$$



Eq. of the model
$\|X - c\|^2 - r^2 \quad (= 0)$

↓

non linear least square fitting

We have dedicated algorithms :

⟶ Gauss - Newton (approx of ∇f / $H(f)$) ⊛

⟶ $\boxed{\text{Levenberg - Marquardt}}$ (**)

$$f(x) = \sum_i f_i(x)^2$$

$$\nabla f(x) = \sum_i 2 . f_i(x) . \nabla f_i(x)$$

$H(f)(x)$ — can be approx

$\lfloor \Longrightarrow$ Newton - Raphson

II. Linear least squares

$$f(x) = \sum_i \big(\overbrace{\underbrace{f_i(x)}_{\downarrow \text{ linear / affine}}}^{\varepsilon_i}\big)^2$$

Linear
$\downarrow$
model f_i with matrices

1 Modelling

1) Write the error vector (vector of errors at each point)

$$\vec{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_N \end{pmatrix} \longleftarrow \text{linear + const} \;/\; X$$

2) This vector can be written like

$$\vec{\varepsilon} = \underset{\text{matrix}}{A} X + \overset{\text{vector (constant part of } \vec{\varepsilon})}{B}$$

$\longrightarrow$ build A, B

3) $$f(x) = \sum \varepsilon_i^2 = \|\vec{\varepsilon}\|^2 = \|AX + B\|^2$$

$\leadsto$ minimization ...

ex of plane fitting

$$\underset{\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}}{f(X)} = \sum_i \big(\underbrace{a_1 x_i + a_2 y_i + a_3 - z_i}_{\varepsilon_i}\big)^2$$

$$\vec{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_N \end{pmatrix} = \begin{pmatrix} a_1 x_1 + a_2 y_1 + a_3 - z_1 \\ \vdots \\ a_1 x_N + a_2 y_N + a_3 - z_N \end{pmatrix} = \overset{A}{\begin{pmatrix} x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots \\ x_N & y_N & 1 \end{pmatrix}} \overset{X}{\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}} + \overset{B}{\begin{pmatrix} -z_1 \\ \vdots \\ -z_N \end{pmatrix}}$$

$A \downarrow$ size $N \times 3$ ($3 = m$)

$B \downarrow$ $N \times 1$

$f(x) = \|AX+B\|^2$

② Optimization v.1 : normal equations   :-(AVOID!

$\Updownarrow$

necessary cond. for a minimum

$$\boxed{\text{if } X^* \text{ is the min of } f \implies \nabla f(X^*) = 0}$$

$f(x) = \|AX+B\|^2 \longrightarrow \langle AX+B, AX+B \rangle = (AX+B)^t \times (AX+B)$

cf. TP3

$\nabla(\|X\|^2) = 2.I.X = 2X$

$\|X\|^2 = X^t.X = X^t.I.X$

$$\Big(\overbrace{\|\quad\|^2}^{u} \circ \overbrace{(AX+B)}^{v}\Big)'$$

$u : X \mapsto \|X\|^2$

$\nabla(u \circ v) = \nabla v^t \times (\nabla u \circ v)$

$\nabla f(x) = \underbrace{A^t}_{\nabla v^t} \times \underbrace{2(AX+B)}_{\nabla u(v)}$

$2x \ldots.$

$\nabla f(x) = 0$

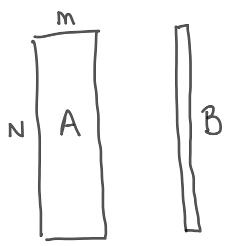
$\hookrightarrow$

$$\boxed{A^tAX + A^tB = 0}$$

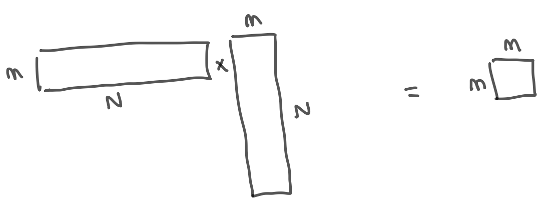
linear system

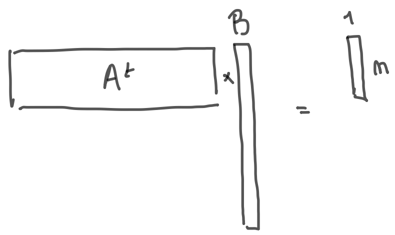
called normal equations

linear system of size m



$A^t \times A$

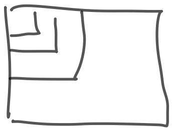




$A^t \cdot A$ — symmetrical

definie positive ? — def $\forall X \neq 0 \quad X^t \cdot \text{matrix} \cdot X > 0$ ←

— prop all eigen > 0

—  minor determinants are > 0

$$X^t \cdot A^t \cdot A \cdot X$$
$$\parallel$$
$$(AX)^t \cdot AX = \|AX\|^2 > 0 \quad \text{if}$$

BUT

$$\boxed{A^t A \cdot X + A^t B = 0}$$

numerically stable ?

$$\text{cond}(A^t A) = \|A^t A\| \cdot \|(A^t A)^{-1}\|$$

10 000 unstable

$\text{cond}(A)^2$

~100 (stable)

conditionning of normal equations is BAD !

③ Right solution : using QR decomposition / SVD

Householder — $A = Q \cdot R$ — Q orthogonal, R upper triangular

more efficient — $A = U \Sigma V^t$ — Σ "diag", U, V orthogonal:

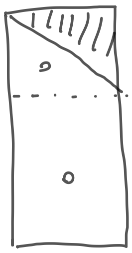
$$U^{-1} = U^t$$
$$\|UX\| = \|X\|$$

$$f(x) = \|AX + B\|^2$$

Apply QR decomposition to A :

$$A = Q \cdot R \quad \text{with}$$

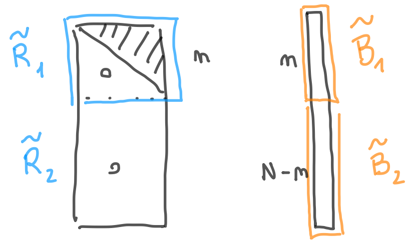
Q orthogonal — $Q^{-1} = Q^t$

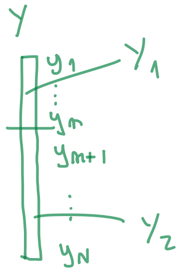
$R =$ 

$Q \cdot Q^t = I$

$$f(x) = \| QRX + B \|^2 = \| Q(RX + Q^t B) \|^2 = \| RX + Q^t B \|^2$$

$$\boxed{\| Q \dots \| = \| \dots \| \quad \text{Q orthogonal}}$$

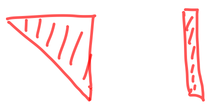




$$\|Y\|^2 = \sum_{i=1}^{N} y_i^2 = \sum_{i=1}^{m} y_i^2 + \sum_{i=m+1}^{N} y_i^2 = \|Y_1\|^2 + \|Y_2\|^2$$

min ?

$$f(X) = \overbrace{\| \tilde{R}_1 X + \tilde{B}_1 \|^2}^{\geq 0} + \underbrace{\| \tilde{B}_2 \|^2}_{\text{constant} / X}$$



$\downarrow$

min $\equiv$ solution of

$$\underbrace{\boxed{\tilde{R}_1 X + \tilde{B}_1 = 0}}_{m \times m \text{ system}}$$

QR - solution

min — X^* which is a solution of $\boxed{\tilde{R}_1 X^* + \tilde{B}_1 = 0}$

— $f(X^*) = \| \tilde{B}_2 \|^2$

$AX + B$ — $N \times m$ "system" — more eq than unknowns

$\longrightarrow$ no solution ...

$A = Q \cdot R$

Q: orthogonal

$$\text{cond}(A) = \text{cond}(R) = \text{cond}(\tilde{R}_1)$$

$$R = \begin{bmatrix} \tilde{R}_1 \\ 0 \end{bmatrix}$$