Alg Gradient

$\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \end{pmatrix}$

$$x_{n+1} = x_n - \gamma_n \nabla f(x_n)$$

Dacă γ_n e suficient de mic

atunci $f(x_{n+1}) < f(x_n)$

Cele mai simple funcții în 2D

$x^T A x$, $a > 0$, $b > 0$

$$f(x) = \frac{1}{2} x^T \overbrace{\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}}^{A} x = \frac{1}{2}(a \cdot x_1^2 + b x_2^2)$$

$b >> a$

$$\nabla f(x) = \begin{pmatrix} a x_1 \\ b x_2 \end{pmatrix}$$

$$x_{n+1} = x_n - \gamma_n \nabla f(x_n)$$

efectuează schimbări "mari" în variabila

x_2

Sol. propune: $\underbrace{\begin{pmatrix} \frac{1}{a} & 0 \\ 0 & \frac{1}{b} \end{pmatrix}}_{A^{-1}} \nabla J(x) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$.

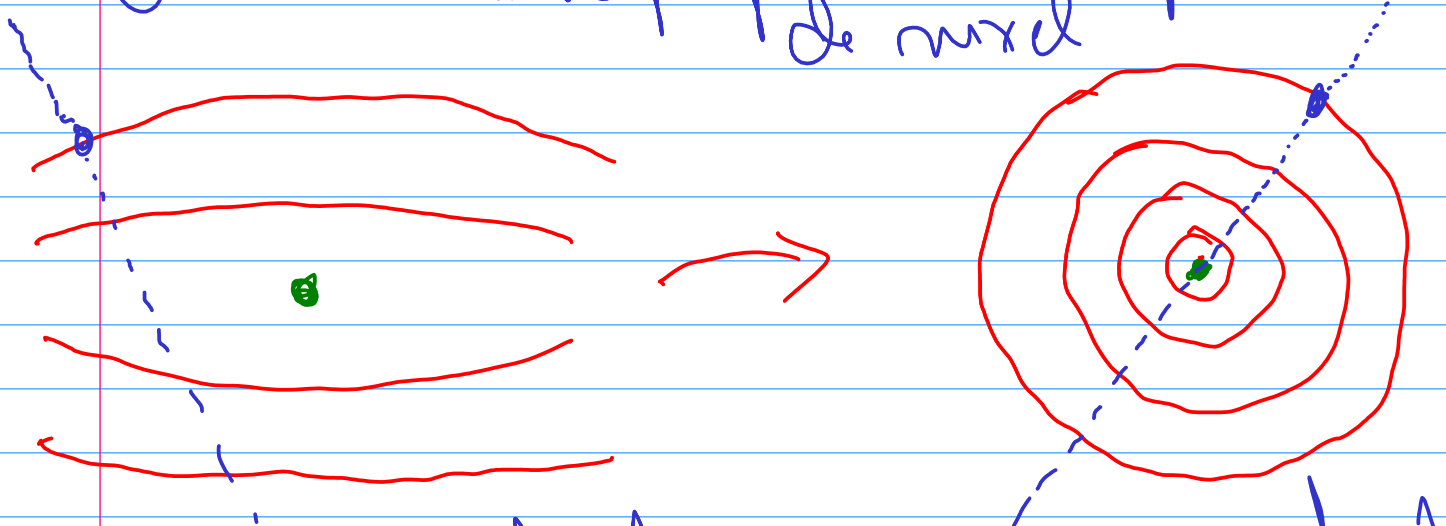
$$A \cdot A^{-1} = A^{-1} \cdot A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$X_{n+1} = X_n - \gamma_n \begin{pmatrix} x_{n,1} \\ x_{n,2} \end{pmatrix}$$

→ dilatare cu factor 10 pe coord 1

→ contracții cu factor 2000 pe coord 2

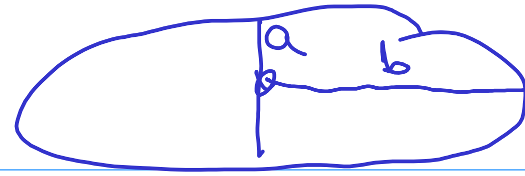
Gradientul este perpendicular pe liniile de nivel



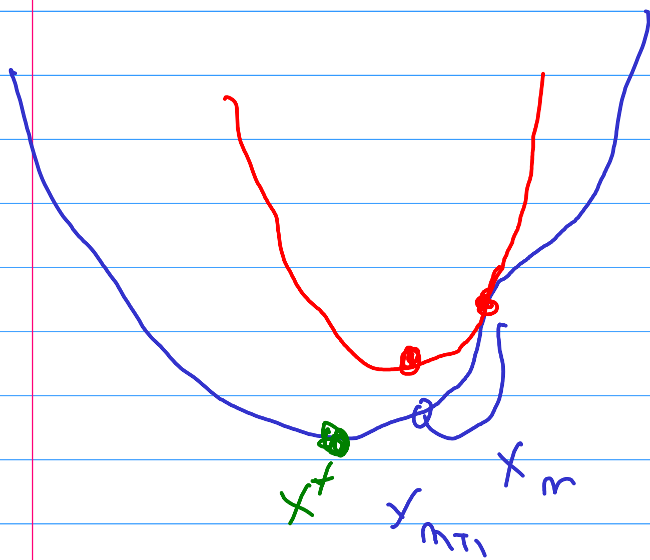
cond > 1

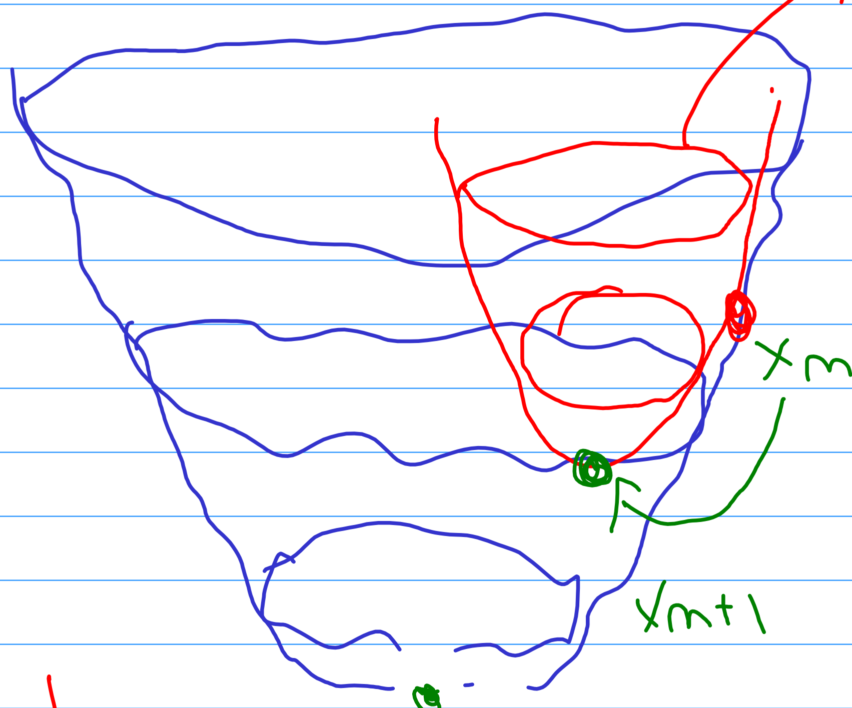
cond = 1

Newton



$\frac{b}{a}$ = condition number > 1





→ aprox de gradul 2

↓

funcție pătratică

↓

o minimizăm

Newton

$$x_{n+1} = x_n - [D^2 f(x_i)]^{-1} \nabla f(x_i)$$

$$D^2 f(x_i) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right)$$

Quad convergence: $e_{n+1} \leq C e_n^2$

next error	current error
0,01	0,1
0,0001	0,01

Dimensiuni mari: n

- costul de stocare al $D^2 f(x)$ este $O(n^2)$.
- $[D^2 f(x)]^{-1} \nabla f(x_i) \to$ rezolvare unui sistem

$Ax = b$: numpy

x = np.linalg.linsolve(A, b)

$O(n^2 \to n^3)$

- Nu calculăm niciodată inversa de matrici în dimensiuni mari!

$O(n^2)$ stocare

$O(n^3)$ calcul

- calcularea unei inverse

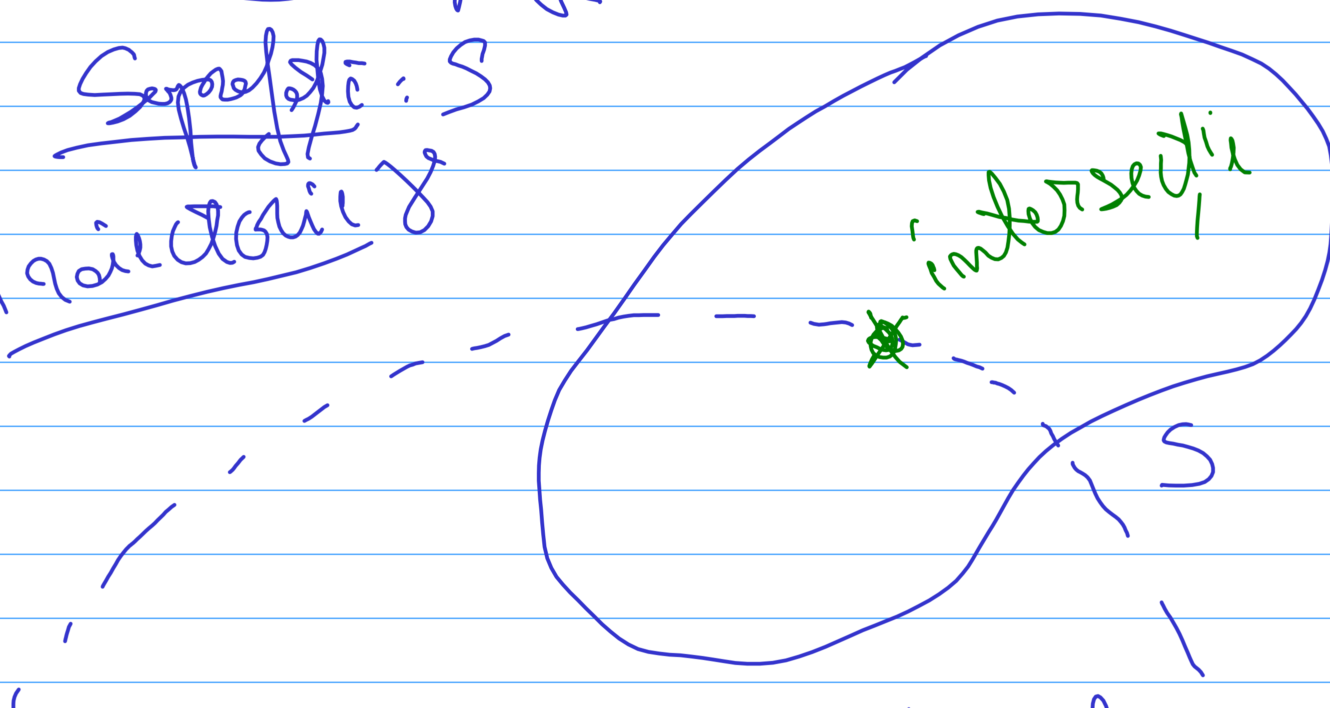
→ n sisteme liniare

$$A\underline{c_i} = e_i = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \to i$$

Probleme specifice

Suprafață: S

Traiectorie γ



→ sistem ecuații neliniare

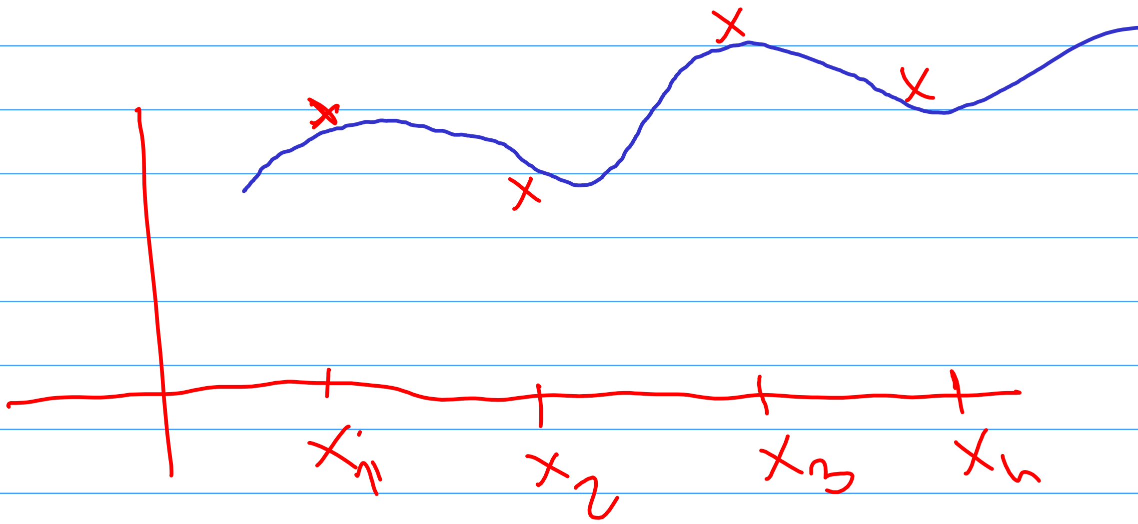
Gauss - Newton

f = sumă de pătrate

Găsirea de parametri într-un model

Date: (x_i, y_i)

param. — măsurători



Model: $x \longmapsto y(x)$

$$y(x_i) \approx y_i$$

y poate să depindă de param

$$y(\underline{s_i}, x)$$

$$\min_{s} \sum_{i=1}^{m} \left(y(s, x_i) - y_i\right)^2$$

sumă de pătrate

Least squares

$$f(x) = \sum_{i=1}^{m} r_i(x)^2$$

m functii
n coord in x

$$\text{Jac} \quad J(x) = \begin{pmatrix} \dfrac{\partial r_i}{\partial x_j} \\ \\ \\ \end{pmatrix}$$

$\partial r_i \to$ linii

$\partial x_j \to$ coloane

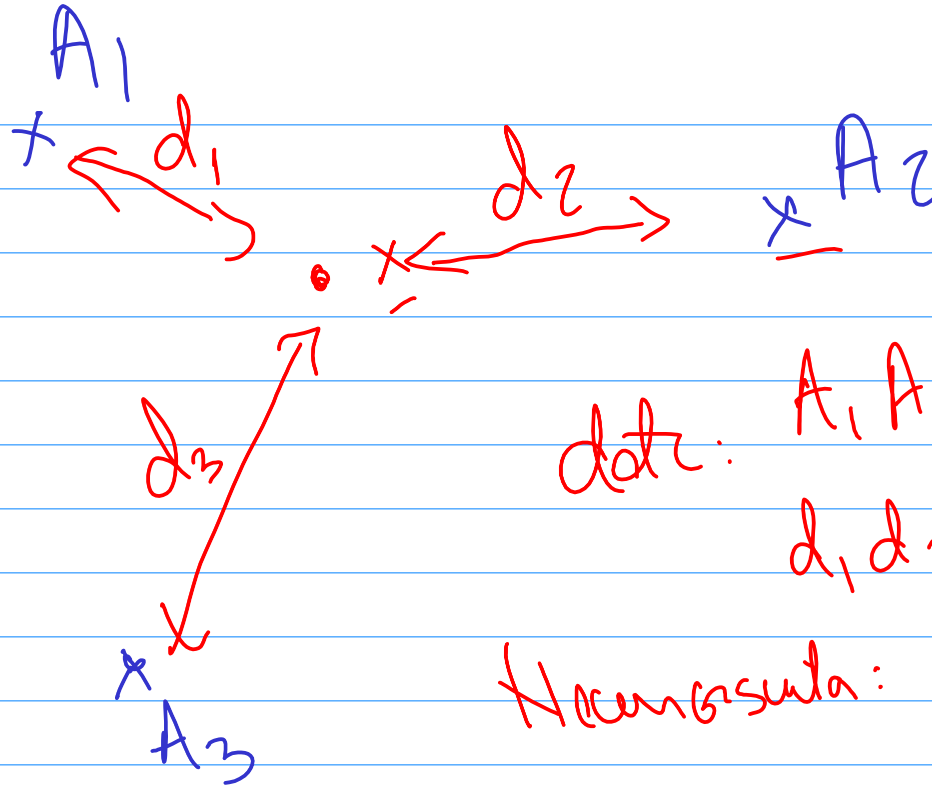
$$r = \begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix}$$

$$\nabla f = 2 J(x)^T \cdot r$$

$$\nabla^2 f(x) = \underbrace{2 J(x)^T J(x)}_{\text{incomplete Hessian}} + \text{"small"} \ldots$$

Gauss-Newton

$$x_{i+1} = x_i - \underbrace{\gamma_i}_{p-S} \underbrace{\left(J(x_i)^T J(x_i) \right)^{-1}}_{\text{incomplete Newton}} \overbrace{J(x)^T r}^{\text{grad}}$$



date: $A_1 A_2 A_3$

$d_1 d_2 d_3$

Necunoscuta: x