

3.4 Newton's method

Near a strict local minimizer \bar{x} we approximate $f \in C^2$ by the quadratic

$$q(x) = f(x_k) + \nabla f(x_k)^T (x - x_k) + \frac{1}{2} (x - x_k)^T \underbrace{H(x_k)}_{\nabla^2 f(x_k)} (x - x_k)$$

$$\nabla q(x) = \nabla f(x_k) + H(x_k) (x - x_k)$$

Since $H(\bar{x})$ pos. def., $H(x_k)$ is pos. def. if x_k is close to \bar{x} by continuity. Hence, $H(x_k)^{-1}$ exists (and is pos. def.!).

Define the next point x_{k+1} by

$$\nabla q(x_{k+1}) = 0 \iff \nabla f(x_k) + H(x_k) (x_{k+1} - x_k) = 0$$

$$\iff x_{k+1} = x_k - \underbrace{H(x_k)^{-1} \nabla f(x_k)}_{d_k}$$

no line search ($\lambda_k = 1$)

Def. d is a **descent direction** at x iff

$$(*) \quad f(x + \lambda d) < f(x) \quad \text{for small } \lambda > 0$$

$$\Rightarrow \quad \frac{f(x + \lambda d) - f(x)}{\lambda} < 0 \quad \text{--- " ---}$$

$$\text{let } \lambda \rightarrow 0 \Rightarrow \quad \underline{\underline{f'(x; d) = \nabla f(x)^T d \leq 0}}$$



CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE

LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS

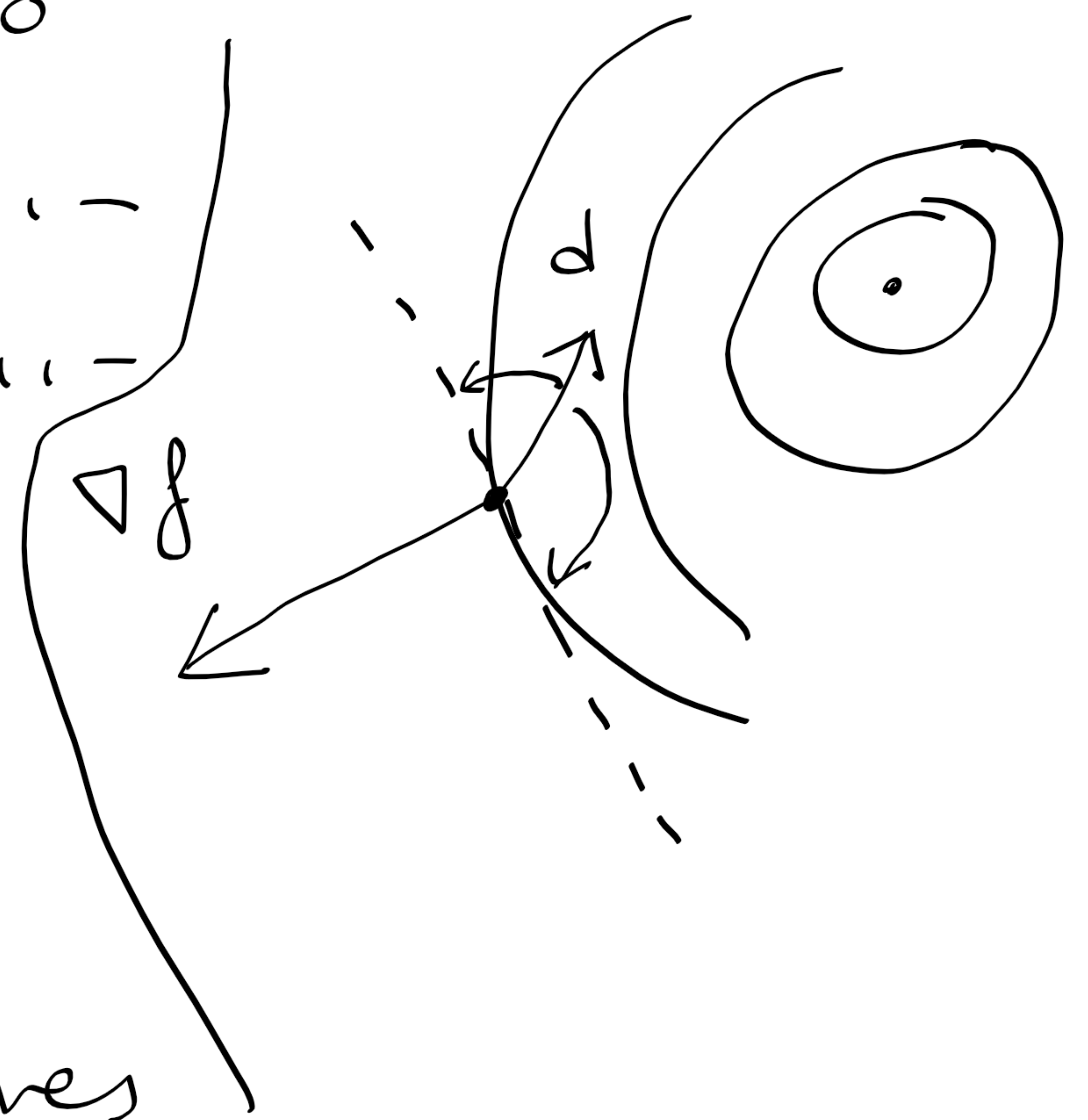
CALL OR WHATSAPP: 689 45 44 70

$$(*) \quad f(x+\lambda d) < f(x) \quad \forall \text{ small } \lambda > 0$$

$$\Rightarrow f(x) + \lambda \nabla f(x)^T d + o(\lambda) < f(x) \quad \dots$$

$$\Rightarrow \lambda (\underbrace{\nabla f(x)^T d}_{\text{fixed}} + \underbrace{o(1)}_{\rightarrow 0}) < 0 \quad \dots$$

sufficient condition $\nabla f(x)^T d < 0$



Newton: $d_k = -H(x_k)^{-1} \nabla f(x_k)$ gives

$$\nabla f(x_k)^T d = - \nabla f(x_k)^T \underbrace{H(x_k)^{-1}}_{\text{pos. def.}} \nabla f(x_k) < 0$$

Thus d_k is a descent direction if $H(x_k)$ is pos. def.

Lemma: H pos. def. with eigenvalues

$$0 < \lambda_{\min} \leq \dots \leq \lambda_{\max} \Rightarrow \lambda_{\min} \|x\| \leq \|Hx\| \leq \lambda_{\max} \|x\|$$

Proof: λ is eigenvalue of $H \Rightarrow \lambda^2$ eigenvalue of H^2

Use spectral thm: $x = Q \hat{x}$, $Q^T H^2 Q = \Lambda^2$

$$\begin{aligned} \|Hx\|^2 &= (Hx)^T Hx = x^T H^2 x = (Q \hat{x})^T H^2 (Q \hat{x}) = \hat{x}^T Q^T H^2 Q \hat{x} \\ &= \hat{x}^T \Lambda^2 \hat{x} = \sum \lambda_i^2 \hat{x}_i^2 \begin{cases} \geq \lambda_{\min}^2 \|\hat{x}\|^2 = \lambda_{\min}^2 \|x\|^2 \\ \leq \lambda_{\max}^2 \|x\|^2 \quad \# \end{cases} \end{aligned}$$

Thm 2: If $f \in C^3$, $\nabla f(\bar{x}) = 0$, $H(\bar{x})$ pos. def. then there exists (\exists) a neighbourhood $\Omega \ni \bar{x}$:

Cartagena99

CLASES PARTICULARES, TUTORIAS TÉCNICAS ONLINE
 LLAMA O ENVÍA WHATSAPP: 689 45 44 70
 ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
 CALL OR WHATSAPP: 689 45 44 70

$$\underbrace{x_{k+1} - \bar{x}}_{\delta_{k+1}} = \underbrace{x_k - \bar{x}}_{\delta_k} - H_k^{-1} g_k$$

$$(a) \quad \delta_{k+1} = \delta_k - H_k^{-1} g_k$$

Taylor expansion of the gradient at x_k
(exerc. 1.7)

$$\nabla \phi(x) = \nabla \phi(x_k) + H_k (x - x_k) + \mathcal{O}(\|x - x_k\|^2)$$

$$x = \bar{x} \Rightarrow 0 = g_k - H_k \delta_k + \mathcal{O}(\|\delta_k\|^2)$$

$$(a) \Leftrightarrow H_k \delta_{k+1} = H_k \delta_k - g_k$$

$$H_k \delta_{k+1} = \underbrace{H_k \delta_k - H_k \delta_k}_{=0} + \mathcal{O}(\|\delta_k\|^2)$$

$$\Rightarrow \|H_k \delta_{k+1}\| = \mathcal{O}(\|\delta_k\|^2) \leq K \|\delta_k\|^2 \quad \text{for } x_k \in \Omega_1 \ni \bar{x}$$

Lemma gives $\lambda_{\min}(x_k) \|\delta_{k+1}\| \leq \|H_k \delta_{k+1}\| \leq K \|\delta_k\|^2$

Choose $\Omega_2 \ni \bar{x} : 0 < L \leq \lambda_{\min}(x_k)$

$$\text{then } \underline{\|\delta_{k+1}\|} \leq \frac{K}{L} \|\delta_k\|^2 = \underbrace{\frac{K}{L}}_{< M < 1 \text{ in } \Omega_3 \ni \bar{x}} \|\delta_k\| \|\delta_k\|$$

Let $\Omega = \Omega_1 \cap \Omega_2 \cap \Omega_3$

$$\|x_{k+1} - \bar{x}\| < M \|x_k - \bar{x}\| < M^k \|x_1 - \bar{x}\| \rightarrow 0, k \rightarrow \infty \quad \#$$

⊕ Fast convergence

⊖ • Must start close to a minimizer, otherwise $d_k = -H_k^{-1} g_k$ may not be a descent direction.

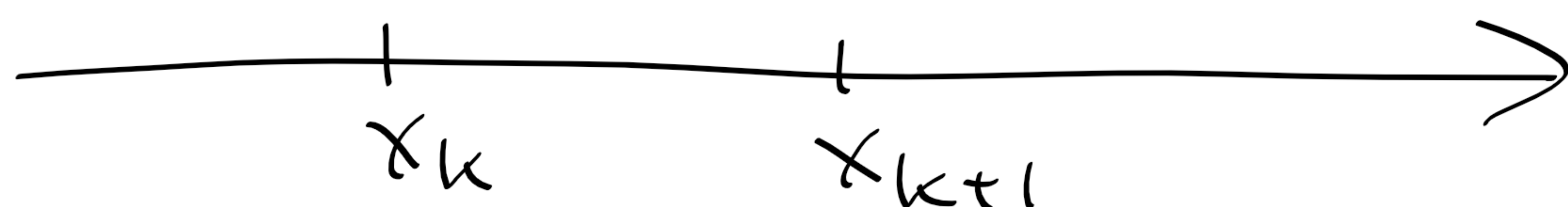
Cartagena99

CLASES PARTICULARES, TUTORIAS TÉCNICAS ONLINE

LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS

CALL OR WHATSAPP: 689 45 44 70



Modified Newton methods

- Use line search in the direction

$$d_k = -H(x_k)^{-1} \nabla f(x_k)$$

- If $H(x_k)$ is not pos. def., then solve

$$(H(x_k) + \varepsilon_k I) d_k = -\nabla f(x_k)$$

for $\varepsilon_k > 0$ that makes $H(x_k) + \varepsilon_k I$ pos. def.

Note: ε_k large \Rightarrow almost steepest descent

In implementations you don't compute inverses, but factorize, e.g. Cholesky:

$$H(x_k) + \varepsilon_k I = LL^T, \quad L = \begin{pmatrix} \times & & \\ & \times & \\ & & \times \end{pmatrix}$$

$$LL^T d = -\nabla f \Leftrightarrow \begin{cases} Lz = -\nabla f \\ L^T d = z \end{cases}$$

- Use $d_k = -B_k \nabla f(x_k)$

with B_k is pos. def. updated on every iteration. No second derivative needed.

Quasi-Newton methods

The logo for Cartagena99 features the text 'Cartagena99' in a stylized, teal-colored font. The '99' is significantly larger and more prominent than the word 'Cartagena'. The text is set against a light blue background with a white, cloud-like shape behind it. Below the text is a horizontal orange and yellow gradient bar.

CLASES PARTICULARES, TUTORIAS TÉCNICAS ONLINE

LLAMA O ENVÍA WHATSAPP: 689 45 44 70

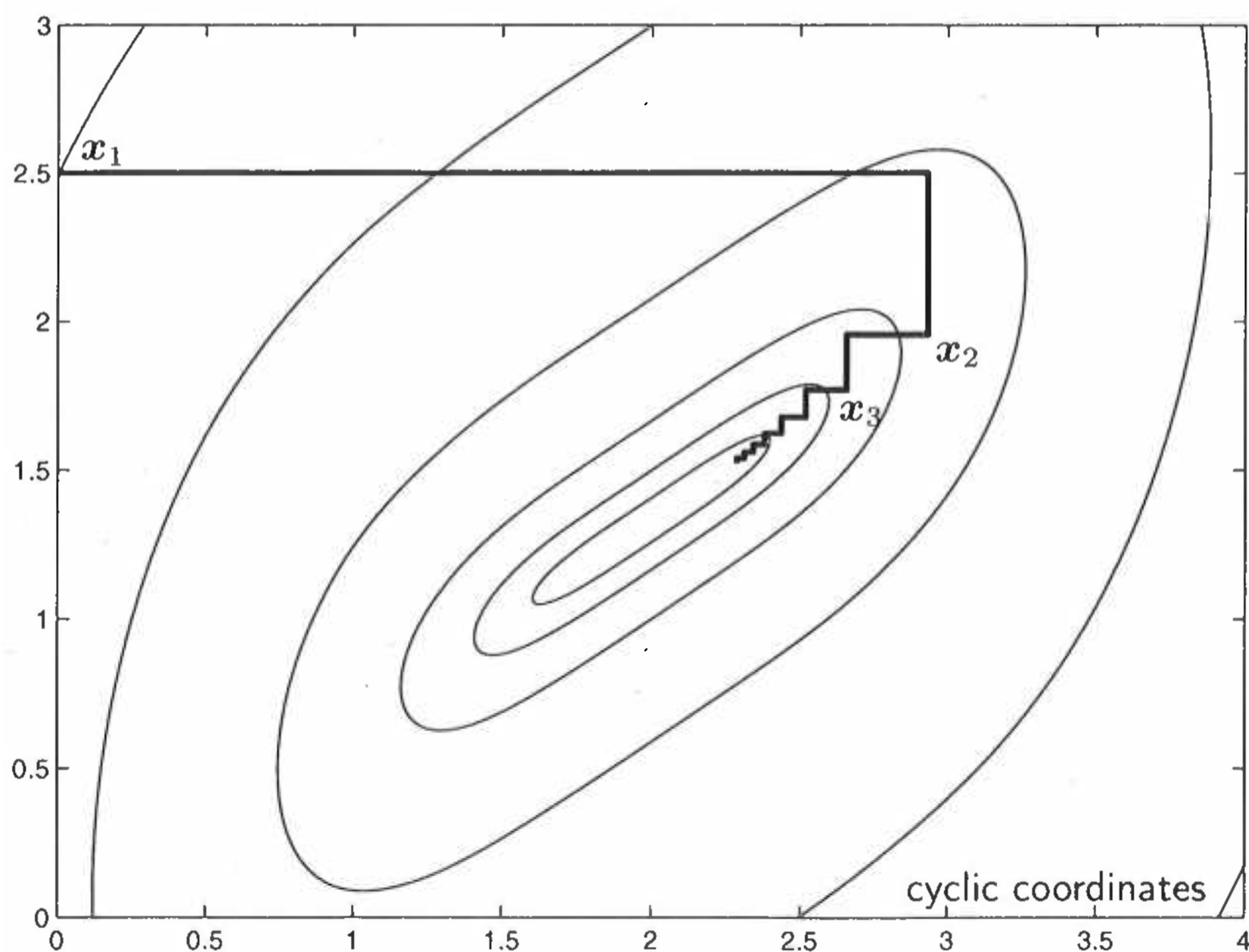
ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS

CALL OR WHATSAPP: 689 45 44 70

3.2 Cyclic-coordinates search

Use canonical base $e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, $e_2, \dots, e_n, e_1, \dots$ as search directions plus line search.

- ⊕ Simple, no derivatives needed, can be used on non-differentiable functions.
- ⊖ Slow, can get stuck.



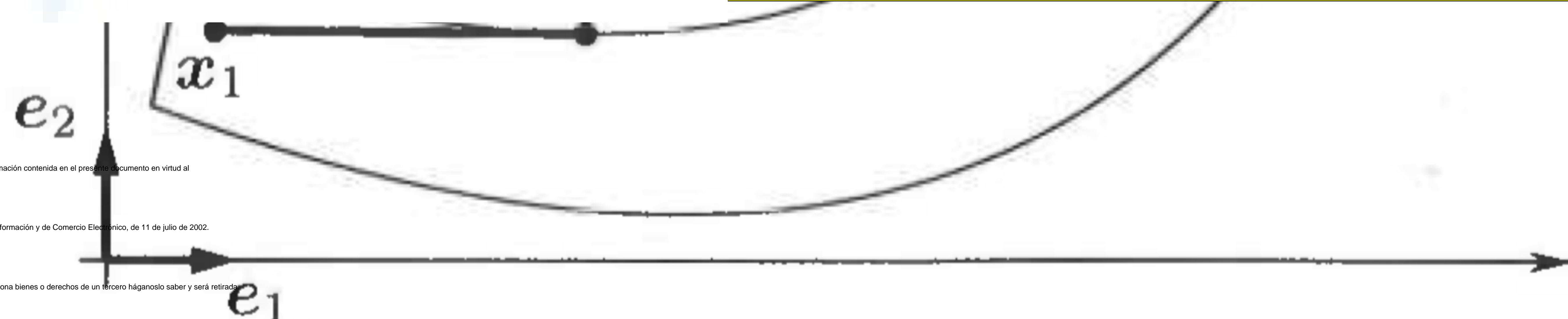
Cartagena99

CLASES PARTICULARES, TUTORIAS TÉCNICAS ONLINE

LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS

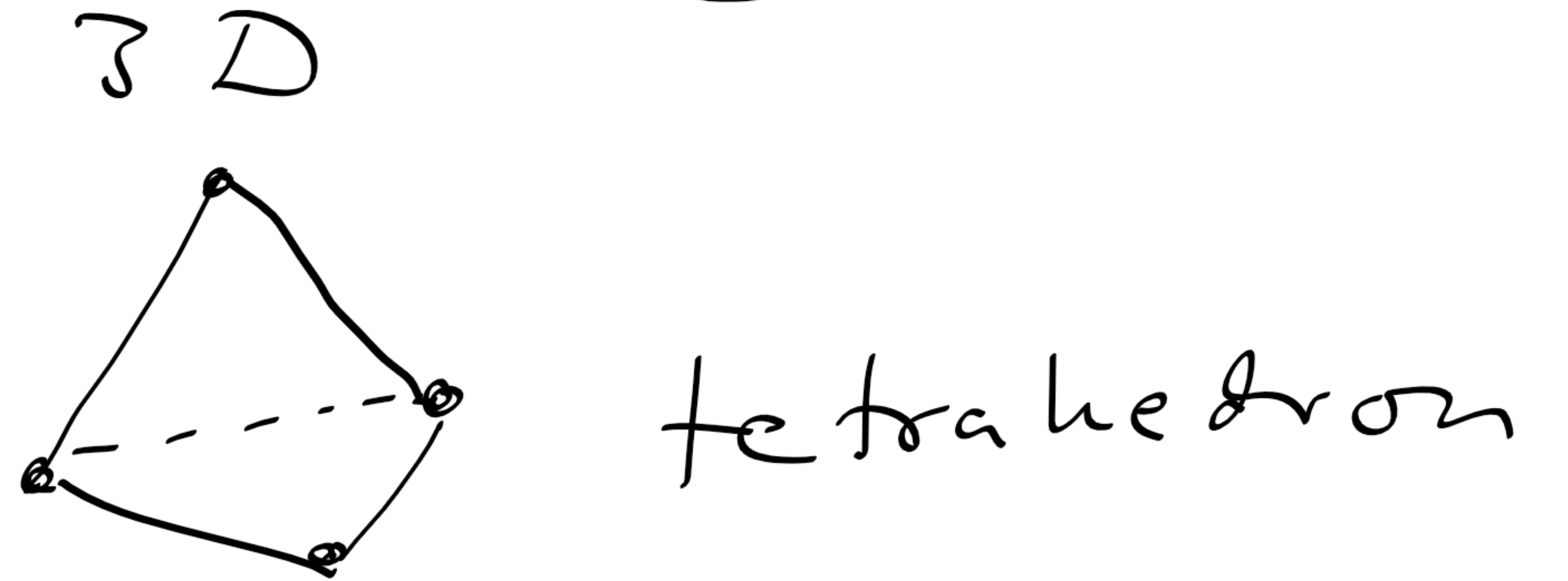
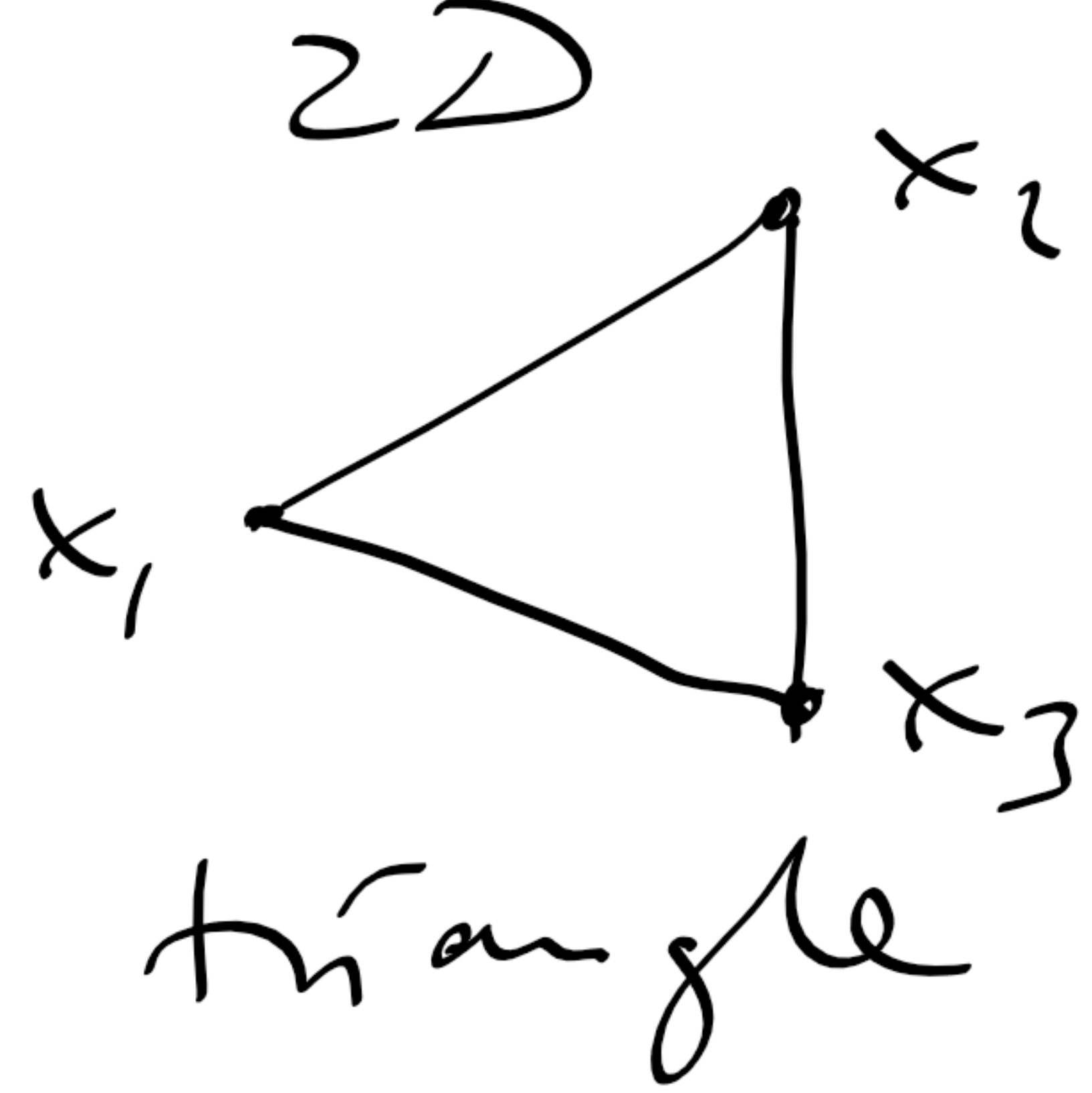
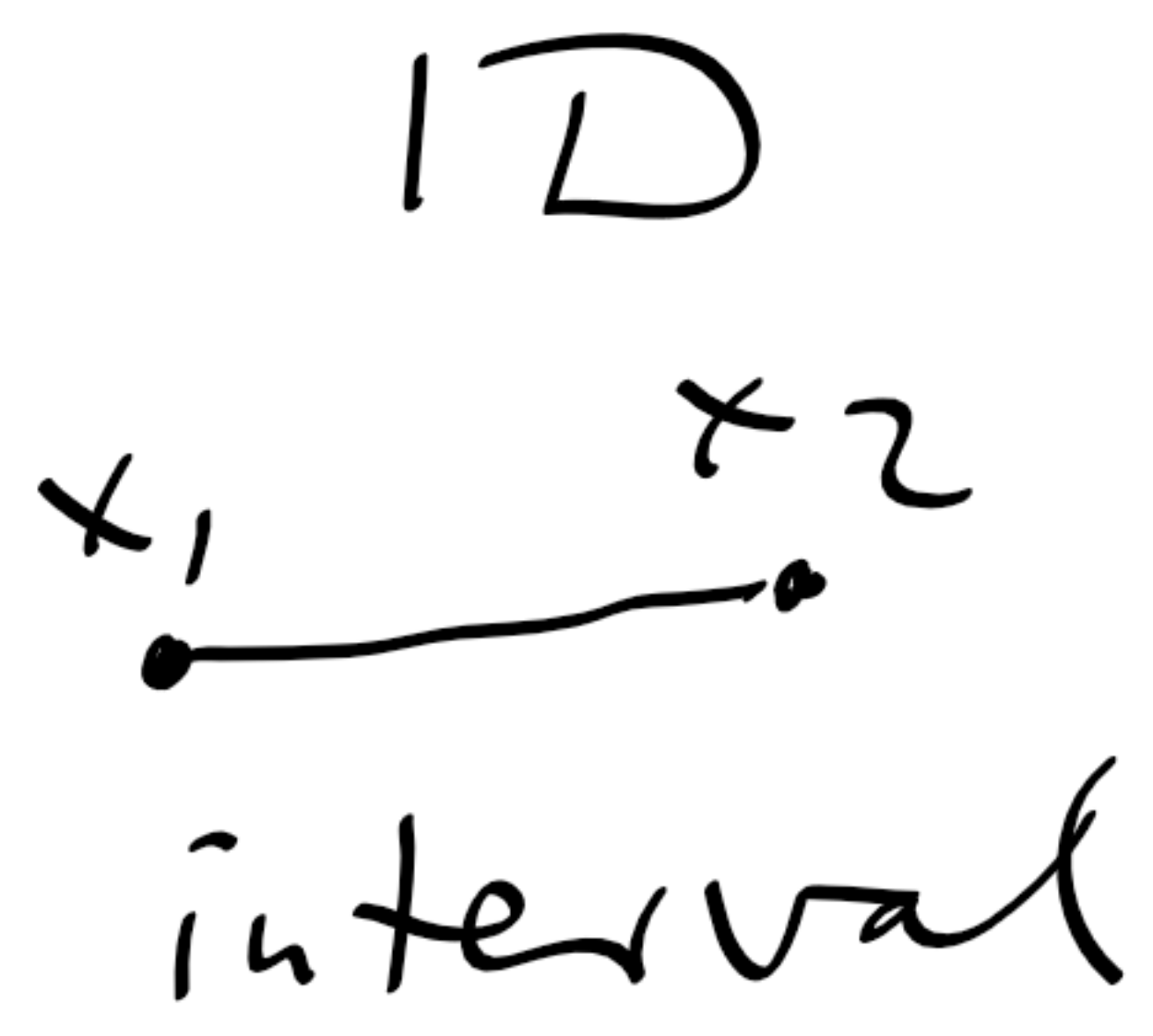
CALL OR WHATSAPP: 689 45 44 70



3.7 Nelder-Mead simplex method

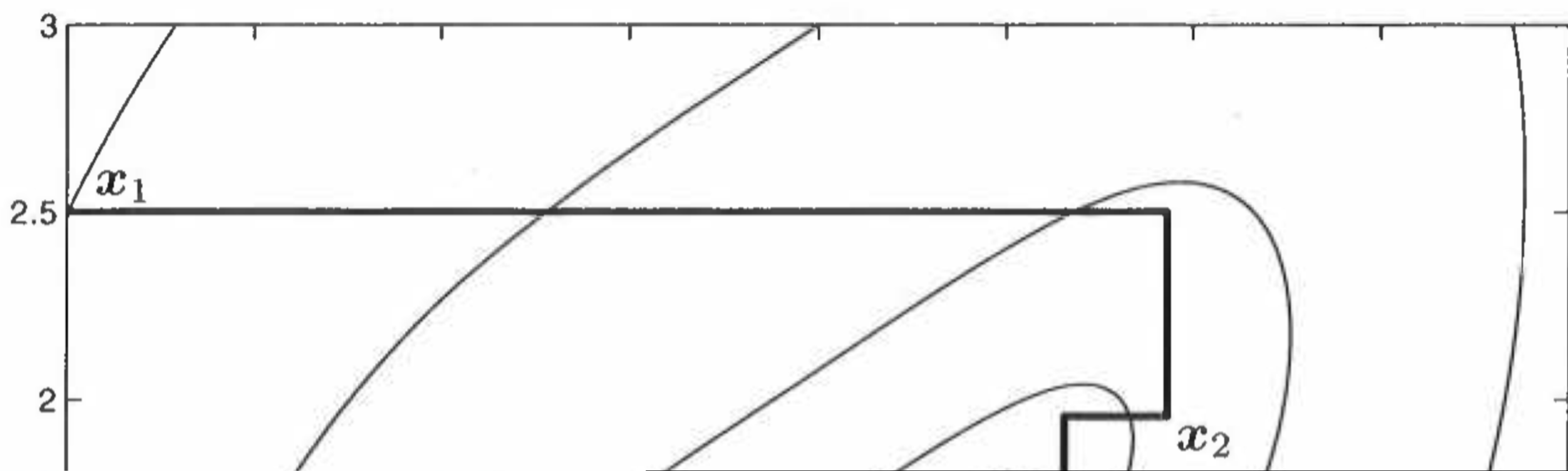
A simplex in \mathbb{R}^n is a set of $n+1$ points x_1, \dots, x_{n+1} such that the vectors

$x_i - x_1, i = 2, \dots, n+1,$ are linearly independent.



Compares function values $f(x_i)$

- ⊕ Robust, no derivatives needed
- ⊖ Slow



Cartagena99

CLASES PARTICULARES, TUTORIAS TÉCNICAS ONLINE

LLAMA O ENVÍA WHATSAPP: 689 45 44 70

...

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS

CALL OR WHATSAPP: 689 45 44 70

