

Gradiente Projetado

L1

$\min f(x)$ s.a. $x \in C$,

onde $C \subset \mathbb{R}^n$ é convexo e fechado.

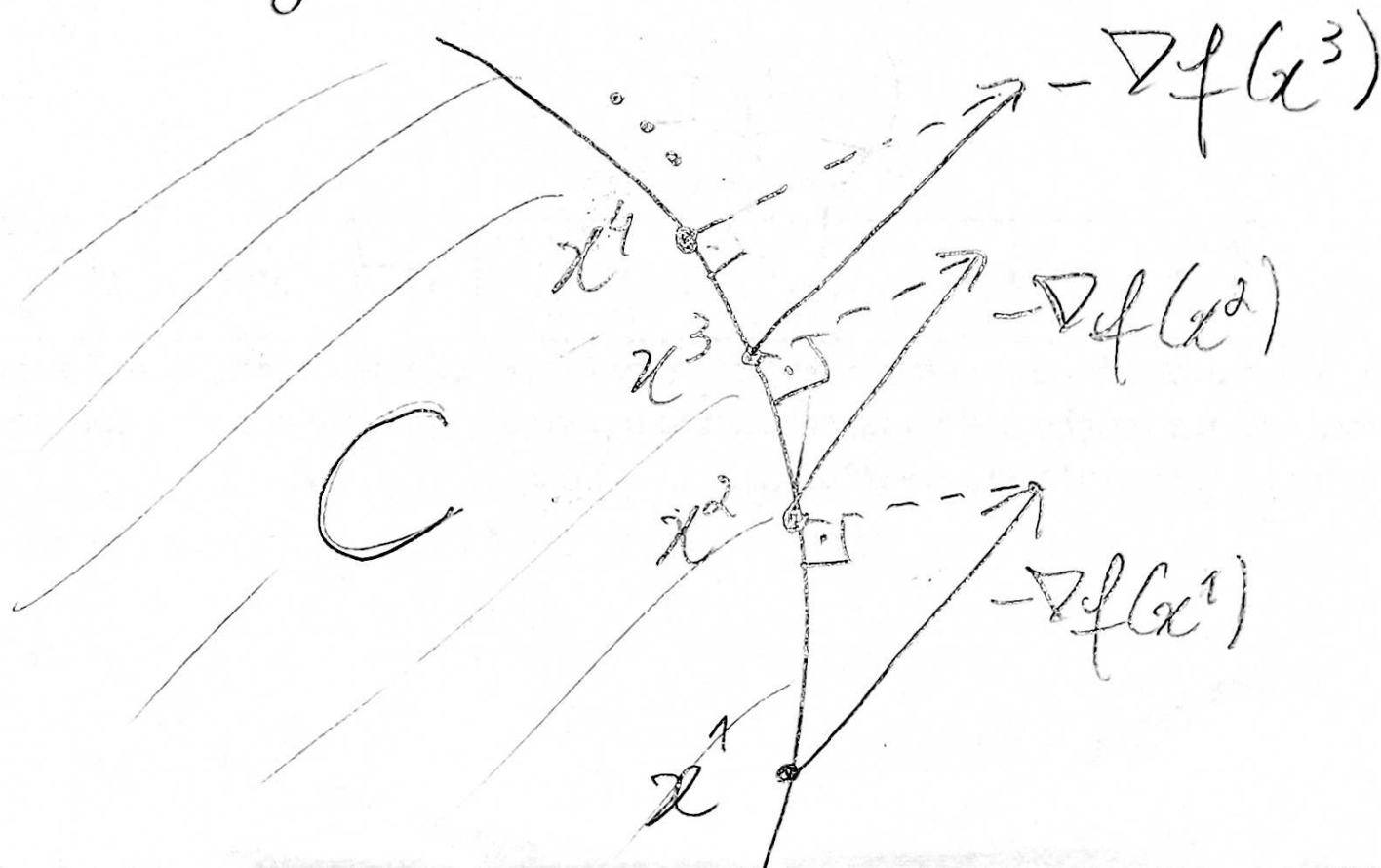
Ideia: Adaptar o método do gradiente

$$x^{k+1} = x^k - t_k \nabla f(x^k)$$

para garantir viabilidade:

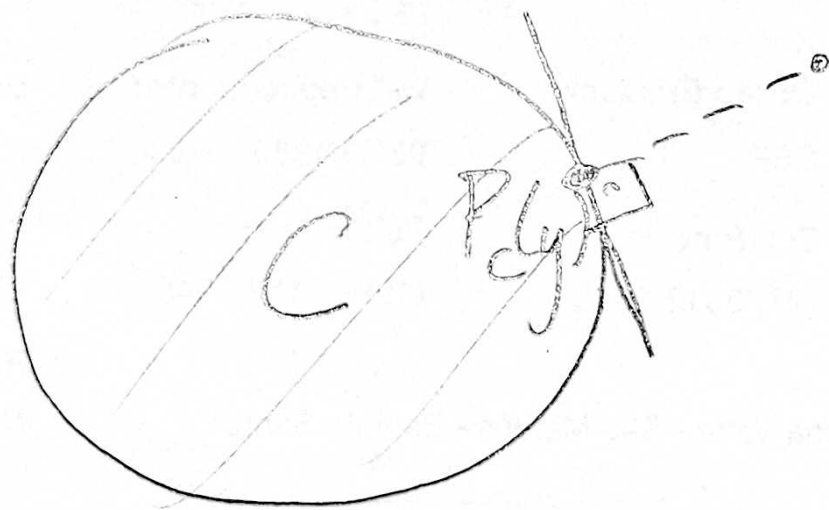
$$x^{k+1} \in C.$$

Como? Calcule x^{k+1} e
"corrija" a viabilidade projetando
no conjunto C . L2



O que é projetar sobre C ? 13

• A projeção de um ponto $y \in \mathbb{R}^n$ em C é o ponto $P_C(y) \in C$ mais
próximo a y .



$P_C(y)$ é
a projeção
de y sobre
 C .

ou seja, $x^* = P_C(y)$ é a solução ⁴
do problema

$$\min_x \frac{1}{2} \|x - y\|^2 \text{ s.a. } x \in C.$$


Teorema: Se $C \neq \emptyset$ é convexo e
fechado então $P_C(y)$ está bem
definida (a projeção é única para
cada $y \in \mathbb{R}^m$).

De fato, o problema de projeção [5]
 $\min_x p(x) = \frac{1}{2} \|x - y\|^2$ s.a. $x \in C$ é
viável, dado que $C \neq \emptyset$. Agora,
 $p(x) = \frac{1}{2} x^t I x - y^t x + \text{cte}(y)$ é
uma quadrática estritamente convexa
(sua Hessiana é $I > 0$). Logo
este problema admite minimizador
global.

Mas como este problema é convexo [6
(pois C é conjunto convexo), então
admite único minimizador global

x^* : Se \bar{x} e \tilde{x} fossem minimizadores
distintos então $\frac{1}{2}\bar{x} + \frac{1}{2}\tilde{x} \in C$ e

$$P\left(\frac{1}{2}\bar{x} + \frac{1}{2}\tilde{x}\right) < \frac{1}{2}P(\bar{x}) + \frac{1}{2}P(\tilde{x}) = P^*,$$

uma contradição. 

$$P_C(y) = \arg \min_{x \in C} \frac{1}{2} \|x - y\|^2, \quad (*) \quad [7]$$

$C \neq \emptyset$ convexo,
fechado.

Como calcular $P_C(y)$?

- Resolvendo (*) via KKT, para conjuntos C específicos.

(Lembre-se que KKT é suficiente para otimalidade em problemas convexos).

$$1) C = \{x \in \mathbb{R}^n; Ax = b\}, \quad \text{L2}$$

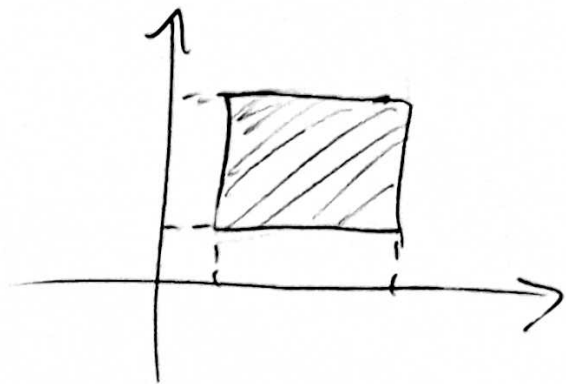
A $m \times n$, posto $A = m$.

EXERCÍCIO: Use as condições KKT para mostrar que

$$P_C(y) = (I - A^t(AA^t)^{-1}A)y + A^t(AA^t)^{-1}b$$

(compare com a expressão (2.2) do livro de Cima, página 57)

$$2) C = \{x \in \mathbb{R}^n; l_i \leq x_i \leq u_i, \forall i\} \\ \text{(caixa, } \underline{l_i < u_i, \forall i})$$



Problema projeção:

$$\min \frac{1}{2} \|x - y\|^2 \text{ s.a. } -x + l \leq 0, x - u \leq 0$$

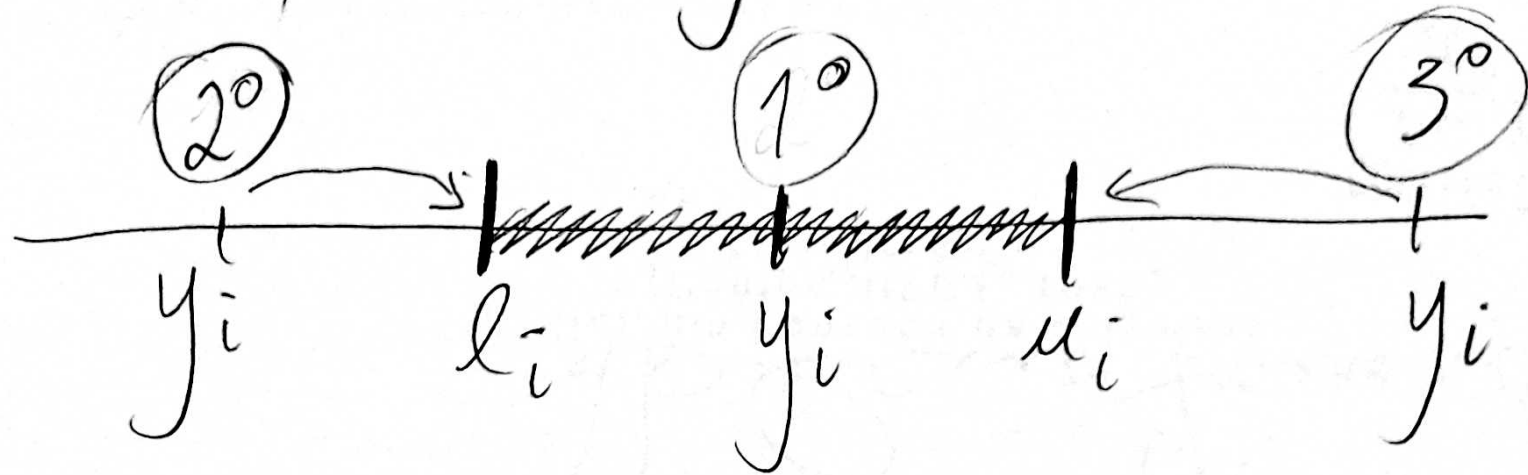
$$\underline{\text{KKT}}: (x^* - y) - \mu^l + \mu^u = 0 \quad (1)$$

$$\mu^l \geq 0, \mu^u \geq 0, l \leq x^* \leq u \quad (2)$$

$$\mu_i^l (-x_i^* + l_i) = \mu_i^u (x_i^* - u_i) = 0, \forall i \quad (3)$$

3 casos para y_i :

10



• $l_i < y_i < u_i$. A fim de satisfazer
KKT, basta tomar $x_i^* = y_i$ e
 $\mu_i^l = \mu_i^u = 0$.

• $y_i \leq l_i$: tomar $x_i^* = l_i$, 11

$$\mu_i^l = x_i^* - y_i \geq 0 \quad \text{e} \quad \mu_i^u = 0.$$

• $y_i \geq u_i$: tomar $x_i^* = u_i$, $\mu_i^l = 0$
e $\mu_i^u = y_i - x_i^* \geq 0$.

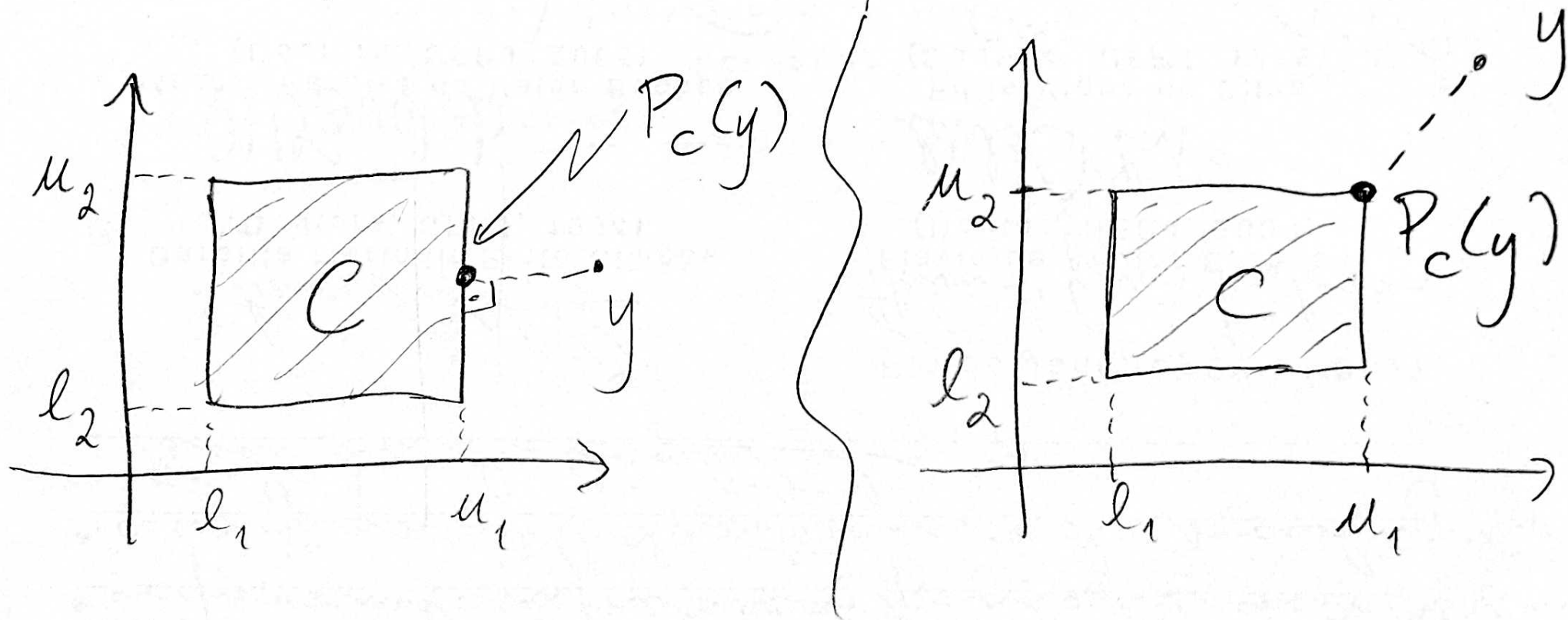
Em resumo:

$$[P_c(y)]_i = \begin{cases} l_i & \text{se } y_i \leq l_i \\ y_i & \text{se } l_i < y_i < u_i \\ u_i & \text{se } y_i \geq u_i \end{cases}$$

Em implementações, fazemos, $\forall i$, 12

$$[P_C(y)]_i = \min\{u_i, \max\{l_i, y_i\}\}$$

(verifique que essa é a projeção)



Como para? Ou, o que é 13
KKT para o problema original

min $f(x)$

s.a. $l_i \leq x_i \leq u_i, \forall i$?

Resposta: x^* é KKT se, e somente se,
 $P_c(x^* - \nabla f(x^*)) - x^* = 0$.
(EXERCÍCIO)

Quais direções tomar?

14

- $d^k = -\nabla f(x^k) \rightarrow \text{min. sem restrições.}$

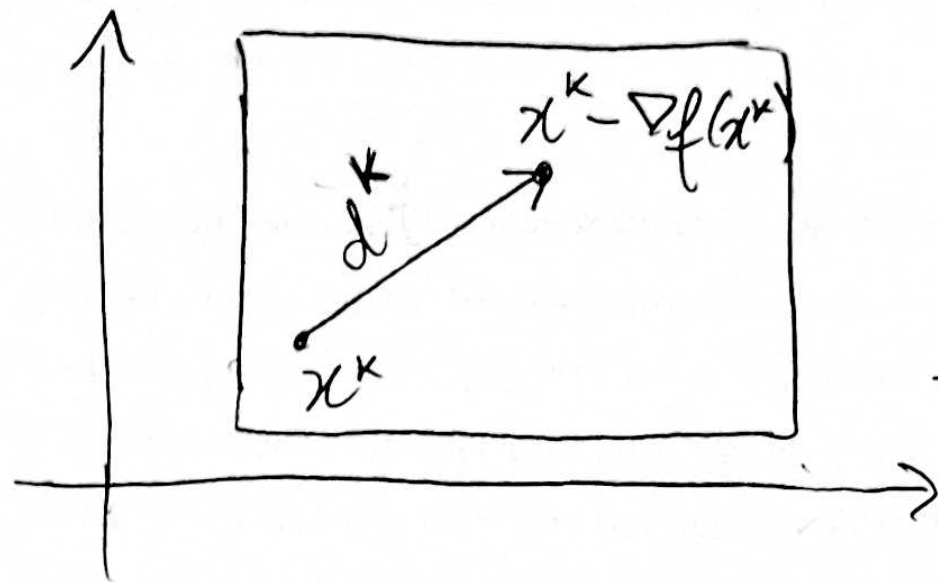
- $\hookrightarrow x^{k+1} = x^k - t_k \nabla f(x^k), t_k > 0.$

- Restrições de caixa:

$$d^k = P_C(x^k - \nabla f(x^k)) - x^k.$$

Observe que se $x^k - \nabla f(x^k) \in C$ então

$$d^k = (x^k - \nabla f(x^k)) - x^k = -\nabla f(x^k)$$

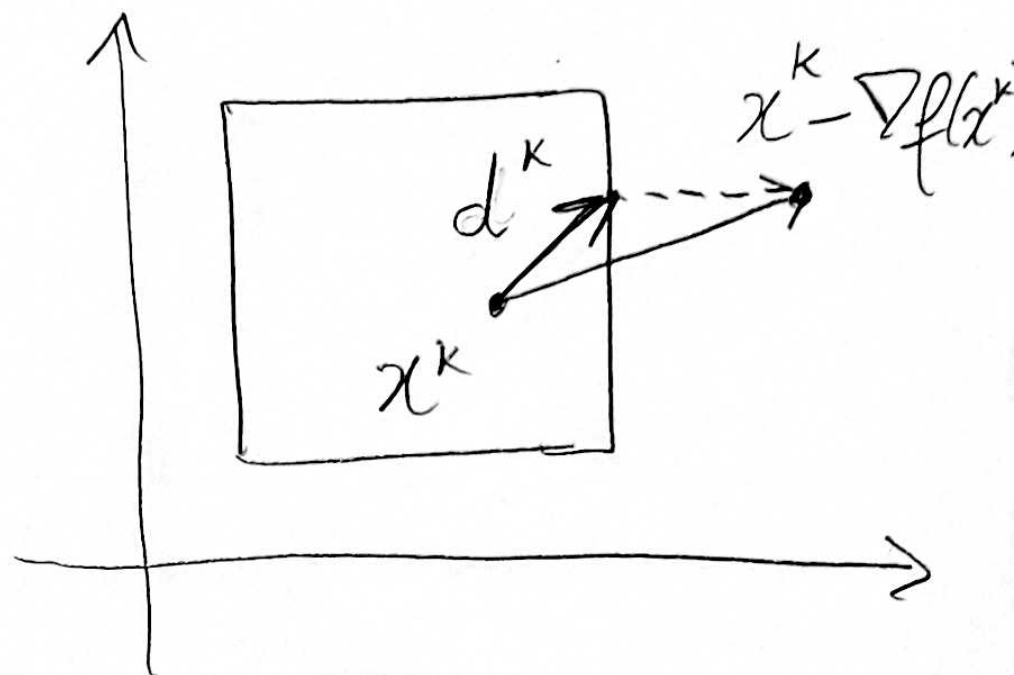


(15)

(passo dentro da
caixa = passo min.
sem restrições).

Agora, se $x^k - \nabla f(x^k) \notin C$, então
 $x^k + d^k \in C$.

EXERCÍCIO: mostre
que $x^k + d^k \in C$.



Conclusão: pela convexidade de \mathbb{L}
 C , temos $x^k + t_k d^k \in C$,

$\forall t \in [0, 1]$. Então podemos realizar
uma busca linear em $t_k \in [0, 1]$

(Armijo), e definir

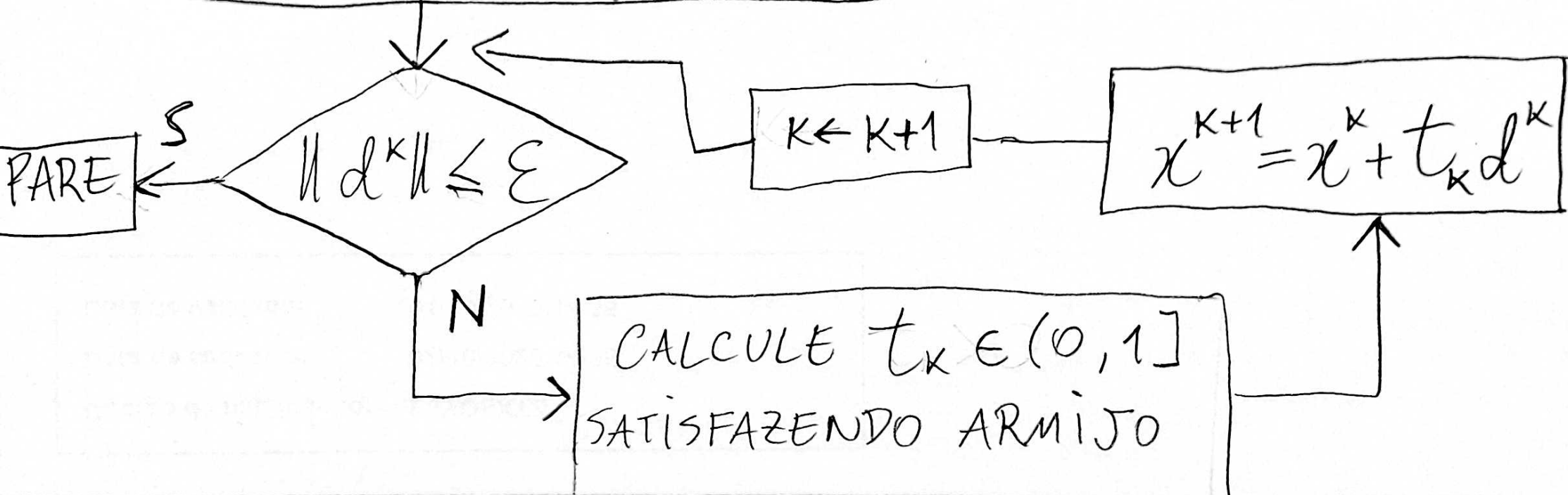
$$x^{k+1} = x^k + t_k d^k \quad (\in C) \quad !!!$$

Método do gradiente projetado

17

DADOS $x^0 \in C$, $\eta \in (0, 1)$, $k = 0$

$$d^k = P_C(x^k - \nabla f(x^k)) - x^k$$



Observações:

18

1) Calcular $t_k \in (0, 1]$ satisfazendo Armijo se faz da mesma maneira que no esquema geral de descida.

Ouás, $d^k = P_C(x^k - \nabla f(x^k)) - x^k$ é uma direção de descida para f a partir de x^k !! (não vou provar :())

2) É exigido que o ponto inicial pertença a C ($x^0 \in C$).

Na prática

- Escolhemos qualquer $x^0 \in \mathbb{R}^n$;
- Atualizamos $x^0 \leftarrow P_C(x^0)$.