Ch III - Recherche linéaire

I] Préambule

Méthode de descente

$$\begin{cases} k \\ d_k \ (= -\nabla f(x_k)) & : \text{Direct° de descente} \\ t_k > 0 & : \text{Pas de descente} \quad (f(x_k + t_k d_k) \leq f(x_k)) \\ x_{k+1} = x_k + t_k d_k \\ \text{Arrêt (Critère } \|x_{k+1} - x_k\| < \underbrace{TOL}_{10^{-8}}) \end{cases}$$

Objectif : $x_k \xrightarrow[k \to +\infty]{} x^*$ $\quad (x^* \text{ réalise } \min_x f(x))$

Exemple :

$$f(x) = \frac{1}{2}(Ax, x) - (b, x)$$

A sym. déf. pos

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

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

t_k ?

$$f(x + t_k d) = \min_{t > 0} (x + td) \quad (*)$$

(POSSIBLE et FAISABLE)

$$\varphi :]0, +\infty[\to \mathbb{R}$$
$$t \longmapsto \varphi(t) = f(x + td)$$

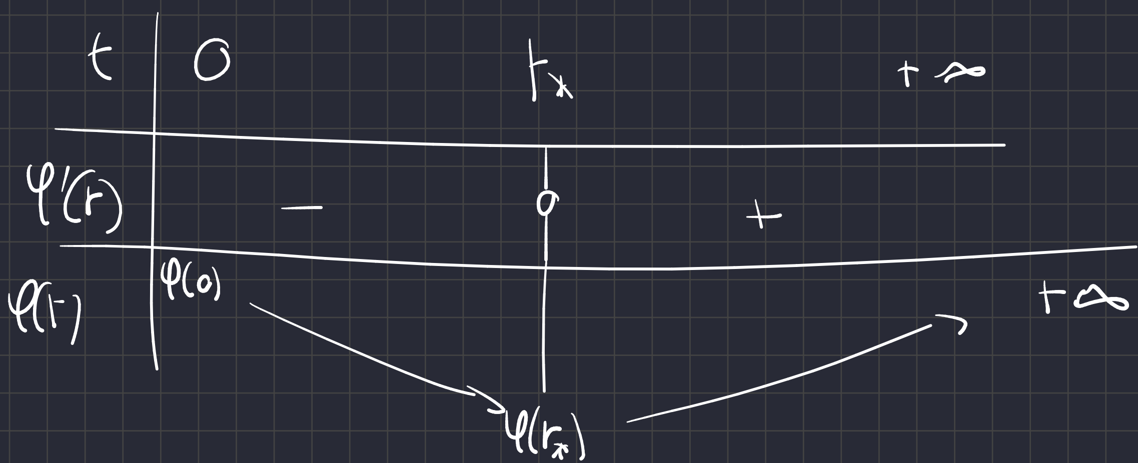
$$(*) \Leftrightarrow \varphi(t_k) = \min_{t > 0} \varphi(t)$$

$$\varphi(t) = f(x+td)$$
$$= \frac{1}{2}\left(A(x+td),\ x+td\right) - (b, x+td)$$
$$\left(= \varphi(0) + \varphi'(0)t + \frac{1}{2}\varphi''(0)t^2\right)$$
$$= f(x) + (Ax-b, d)t + \frac{1}{2}(Ad, d)t^2$$
$$= f(x) + (\nabla f(x), d)t + \frac{1}{2}(Hf(x)d, d)t^2$$

Rmq :
$\nabla f(x) = Ax - b$ $\qquad$ $\varphi''(0) = \frac{1}{2}(Ad, d) \geq 0$
$Hf(x) = A.$ $\qquad$ $\varphi'(0) = (Ax-b, d)$

$$\varphi'(t_*) = 0 \iff \varphi'(0) + \varphi''(0)t_* = 0$$
$$\iff t_* = -\frac{\varphi'(0)}{\varphi''(0)} = \frac{(b-Ax, d)}{(Ad, d)}$$

$$\boxed{t_* = \frac{\|d\|^2}{(Ad, d)}} > 0$$



Exemple : Algo à gradient optimal :

$$\left[\begin{array}{l} k \\ d_k = -\nabla f(x_k) = (b - Ax_k) \\ t_k = \dfrac{\|d_k\|^2}{(Ad_k, d_k)} \\ x_{k+1} = x_k + t_k d_k \end{array}\right.$$

Arrêt

Rmq : Avec f quelconque : ça ne marche pas

II] Principe

d_k : directo de descente

$t_k > 0$? $\varphi(t_k) < \varphi(0)$, $(\varphi(t) = f(x + t d_k))$

$(f(x_k + t_k d_k) < f(x_k))$

Descente assez substancielle (qd j'ai de la marge).

III] Intervalle de sécurité

Déf : (Domaine de travail $]0, +\infty[$)

$[a, b] \subset]0, +\infty[$: Intervalle de sécurité s'il permet de classer $t \in]0, +\infty[$ (pas de descente) de la façon suivante

1. Si $t < a$, alors t est trop petit
2. Si $a \leq t \leq b$: OK
3. Si $t > b$: alors t est trop grand.

Typiquement $[a,b]$ inaccessible !

Algo de base :

0° $[\alpha,\beta] =]0,+\infty[$

1° Choisir $t \in [\alpha,\beta)$

2° Si t trop petit, $\alpha = t$, revenir en 1

3° Si t trop grand, $\beta = t$,

4° Si t convient, STOP

V) Caractérisation de l'intervalle de sécurité

(méthodes de recherche)

1) Méthode d'Armijo

$\alpha = 0$, $0 < m < 1$

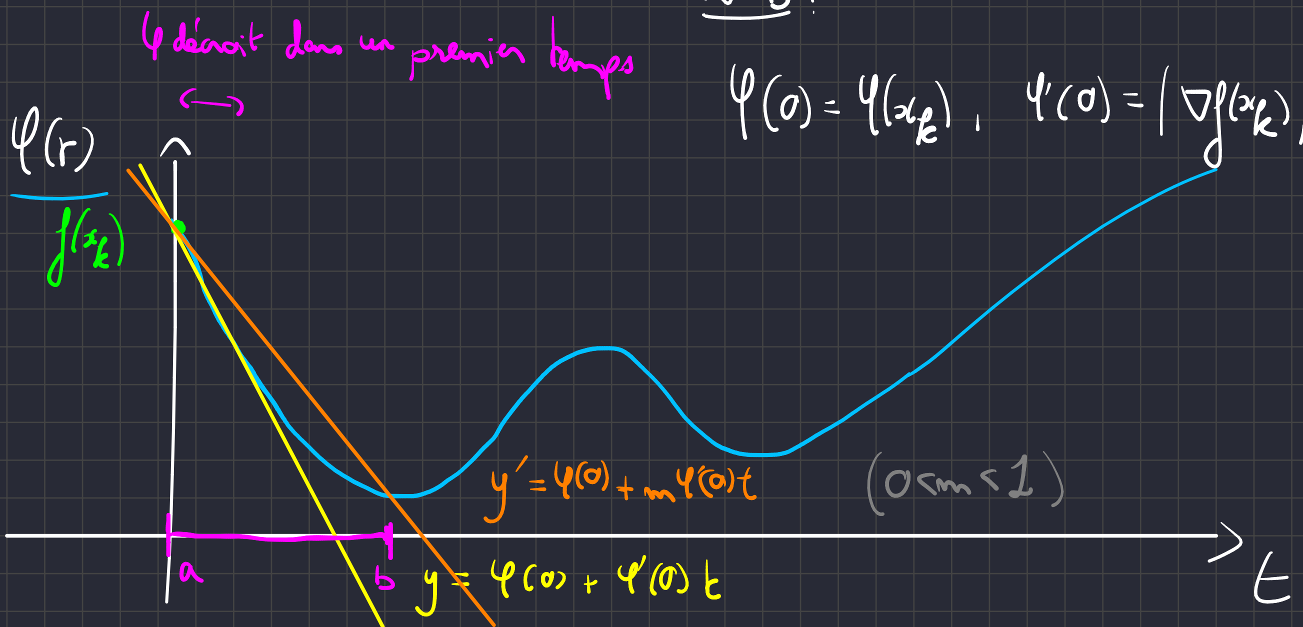
Règle

Si $\varphi(t) \leq \varphi(0) + m\varphi'(0)t$, alors t convient

Si $\varphi(t) > \varphi(0) + m\varphi'(0)t$, alors t trop grand

$\varphi(t) = f(x_k + t_k d_k)$

N-B : $\varphi(0) = \varphi(x_k)$, $\varphi'(0) = (\nabla f(x_k), d_k)$



Point faible d'Armijo :

Il tolère des petites valeurs de t_k.
Ça risque d'actionner le critère d'arrêt alors que $\|\nabla f(x_k)\|$ est assez grand...

2) Règle de Goldstein

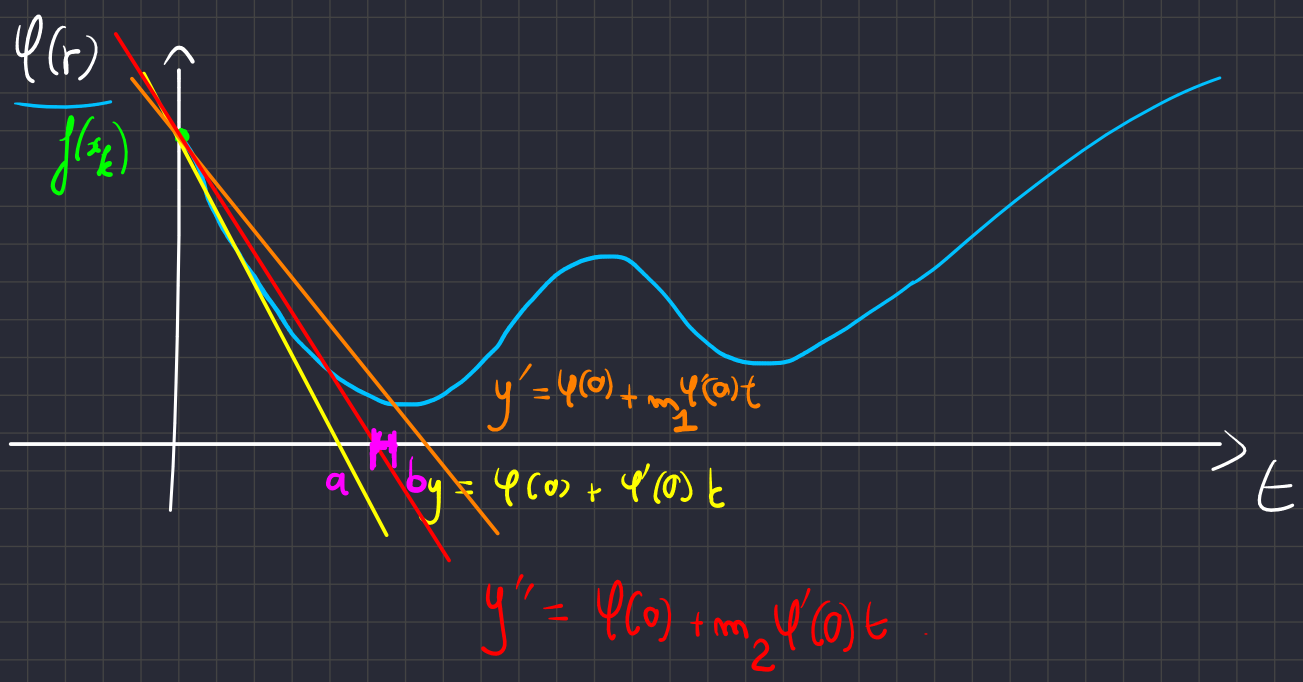
Soient $0 < m_1 < m_2 < 1$

Règle :

Si $\varphi(t) < \varphi(0) + m_2 \varphi'(0) t \Rightarrow t$ trop petit

Si $\varphi(t) > \varphi(0) + m_1 \varphi'(0) t \Rightarrow t$ trop grand

Si $\varphi(0) + m_2 \varphi'(0) t \leq \varphi(t) \leq \varphi(0) + m_1 \varphi'(0) t \Rightarrow$ OK



Goldstein : Correction d'Armijo.

Prop :

① Pour une forme quadratique, Goldstein accepte le pas optimal !
↳ Cf TD3

Thm :

Si $f : \mathbb{R}^n \to \mathbb{R}$ est coercive $\left(f(x) \xrightarrow[\|x\| \to +\infty]{} +\infty\right)$, différentiable

Soit l'algorithme $x_{k+1} = x_k - r_k \nabla f(x_k)$

(r_k selectionné par Goldstein)

Alors $x_k \xrightarrow[k \to +\infty]{} x_*$, et $\nabla f(x_*) = 0$

3) Règle de Wolfe

$0 < m_1, m_2 < 1$

Si $\varphi(r) > \varphi(0) + m_1 \varphi'(0) t \Rightarrow t$ trop grand

Si $\varphi(t) \leq \varphi(0) + m_1 \varphi'(0) r$, $\varphi'(r) < m_2 \varphi'(0) \Rightarrow t$ trop petit

Si $\varphi(r) \leq \varphi(0) + m_1 \varphi'(0) r$, $\varphi'(r) \geq m_2 \varphi'(0) \Rightarrow t$ convient.