UE Modélisation :

Topic 4 :

Newton
-Raphson

& Reg.
Logistique

Autour de l'algorithme de Newton - Raphson

I

Objectif : Etant donné $f : \mathbb{R}^n \to \mathbb{R}^n$, trouver un zéro i.e $x^* \in \mathbb{R}^n$ tel que

$f(x^*) = 0 \quad (\neq)$

Remarque : Les problèmes d'optimisation sont aussi des problèmes de recherche de zéro :

Si $\varphi : \mathbb{R}^n \to \mathbb{R}$ (Vraisemblance ou Loss)

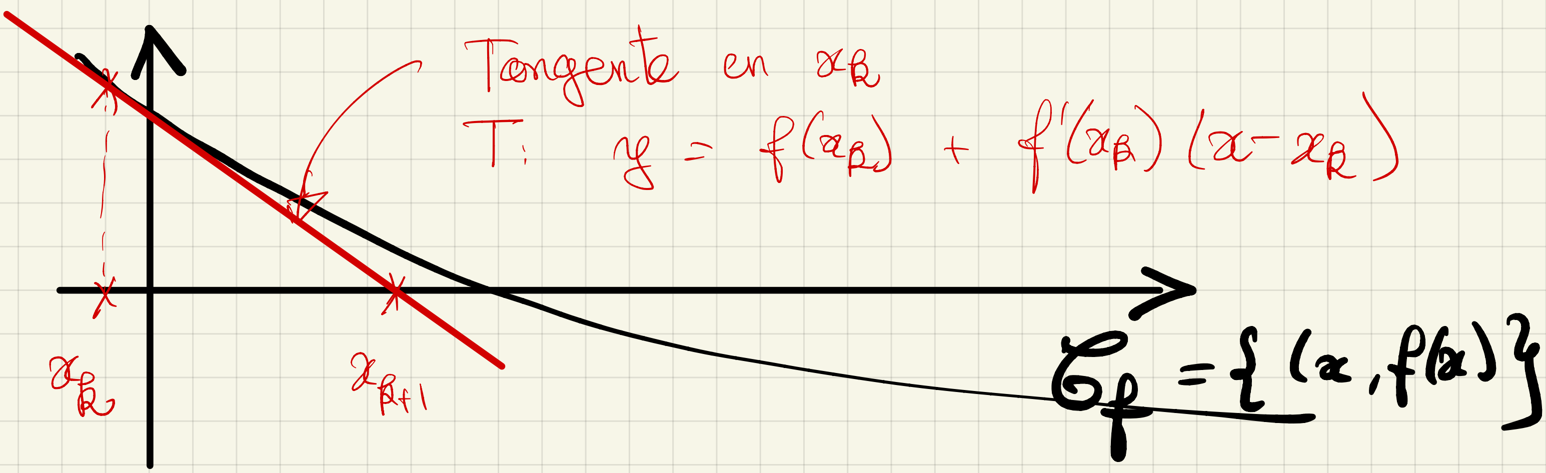
Alors (O) $x^* = \underset{x \in \mathbb{R}^n}{\mathrm{Argmax}} / \mathrm{Argmin}\ \varphi(x)$

$\Longrightarrow \nabla\varphi(x^*) = 0$ } Condition de point critique / Condition de premier ordre.

En posant $f := \nabla\varphi$, on transforme le problème d'optimisation (O) en problème de recherche de zéro. (Z)

I - En dimension 1 :

En France, cette méthode est appelée la méthode de la sécante à cause du dessin :



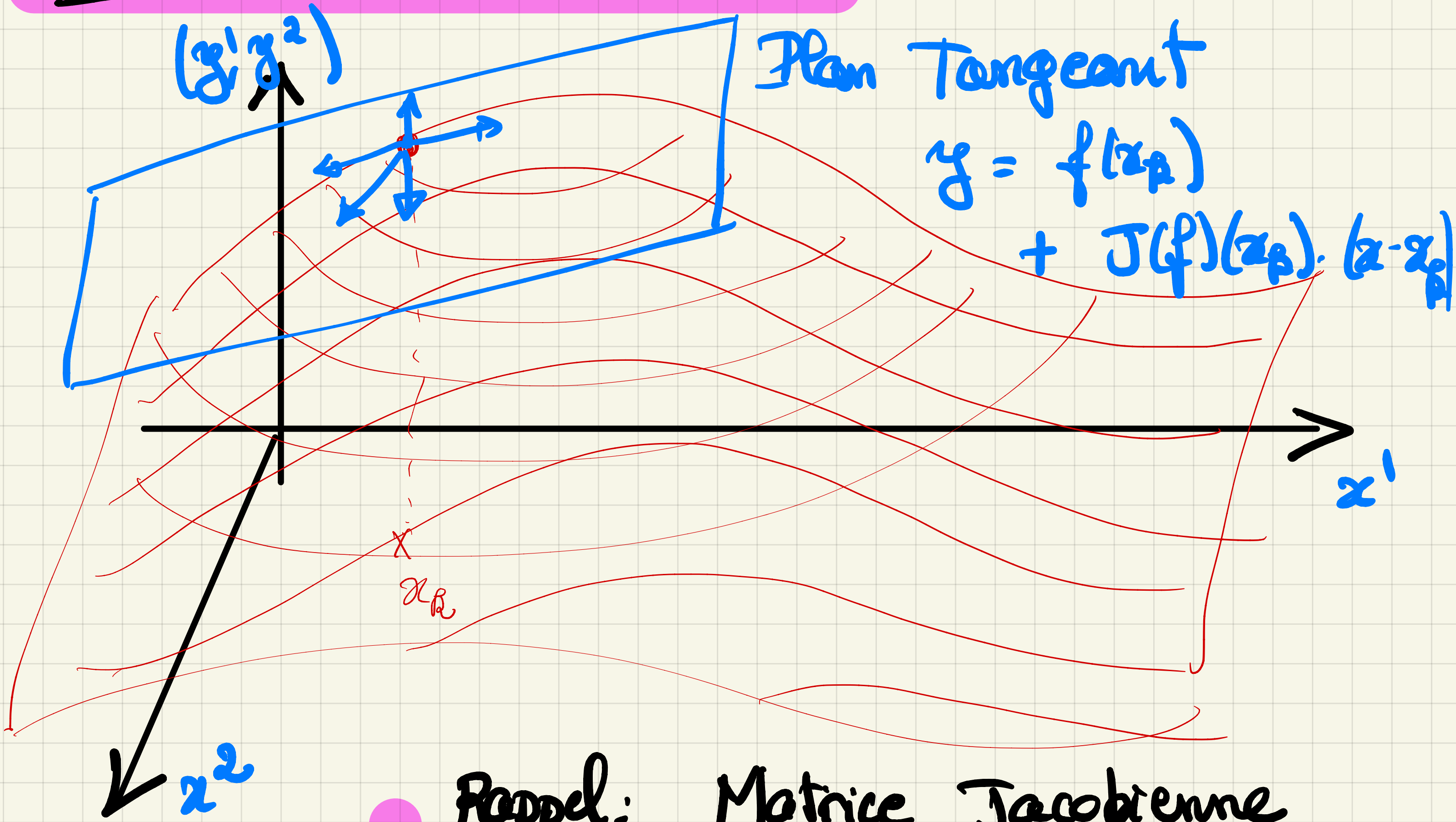
x_{k+1} = Point où T croise $\{y = 0\}$

= (Solution de $0 = f(x_k) + f'(x_k)(x - x_k)$)

$= x_k - \dfrac{f(x_k)}{f'(x_k)}$

C'est la récurrence vue en TP ou trouvée sur Wikipédia.

II - En dimension n :



Rappel : Matrice Jacobienne

$$\boxed{J(f)(x) = \left(\frac{\partial f^i}{\partial x^j}(x) \right)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}}$$

De même,

x_{k+1} = Point où le plan tangeant $T = \left\{ (x,y) \,\middle|\, \begin{array}{l} y = f(x_k) \\ + \cdots \end{array} \right\}$ coupe $\{(x,0)\} = \{y=0\} \subset \mathbb{R}^{2n}$

$= \left(\text{Solution de } f(x_k) + J(f)(x_k).(x - x_k) = 0 \right)$

$$\boxed{x_{k+1} = x_k - J(f)(x_k)^{-1} \cdot f(x_k).}$$

⚠ $\mathbb{R}^n \times \{0\} = \{(x,0)\} = \{y=0\} \subset \mathbb{R}^{2n}$

Commentaire sur le dessin :

$f : \mathbb{R}^n \to \mathbb{R}^n$ avec $n = 2$.

$x = (x^1, x^2) \longmapsto (y^1, y^2) = (f_1(x), f_2(x))$

D'où l'algorithme : ε = Erreur $= 10^{-8}$ par exemple.

n_{max} = Nombre maximal d'itérations $= 100$ par ex.

Tant que $\|x - x_0\|_\infty \geq \varepsilon$, et $k \leq n_{max}$:

$x_0 \leftarrow x$

$x \leftarrow x - J(f)(x)^{-1} \cdot f(x)$.

$k \leftarrow k+1$

Retourner (x, k)

III - Analyse de la convergence.

L'algorithme de Newton-Raphson (N.R) ne converge pas toujours. Cependant, lorsqu'il fonctionne (ou que l'on prend la peine de le faire fonctionner) il est d'une efficacité redoutable.

Thm [Version très affaiblie du thm de Newton-Kantorovich]

Soit $f : \mathbb{R} \to \mathbb{R}$ de classe $\mathcal{C}^2$

avec $f(x^*) = 0$, $f'(x^*) \neq 0$.

Si x_0 point de départ de l'algorithme de N.R.

avec $|x^* - x_0| \times \left[\sup_{|h| \leq |x^*-x_0|} \frac{1}{|f'(x^*+h)|} \right] \times \left[\sup_{|h| \leq |x^*-x_0|} |f''(x^*+h)| \right] \leq 1$.

Alors l'algorithme de Newton-Raphson converge super-exponentiellement.

$$\begin{cases} |x^* - x_k| \xrightarrow[k \to +\infty]{} 0 \quad \left(|x^*-x_k| \leq \frac{1}{2^k} |x^*-x_0| \right) \\ \& \quad |x^* - x_{k+1}| \underset{k \to +\infty}{\sim} |x^* - x_k|^2 \left| \frac{f''(x^*)}{2 f'(x^*)} \right| \end{cases}$$

Remarque : Si $\begin{cases} \left| \frac{f''(x^*)}{2f'(x^*)} \right| \approx 1 \\ |x^* - x_k| \approx 10^{-1} \end{cases}$

Alors

$|x^* - x_{k+1}| \approx 10^{-2}$

$|x^* - x_{k+2}| \approx 10^{-4}$

$|x^* - x_{k+3}| \approx 10^{-8}$

$|x^* - x_{k+4}| \approx 10^{-16}$

Plus que la précision machine.

Preuve : Argument : Développement de Taylor (-Lagrange)

$$x^* - x_{k+1} = x^* - x_k + \frac{f(x_k)}{f'(x_k)}$$ (Def itérat° N.R)

$$= \frac{f(x_k) - (x_k - x^*) f'(x_k)}{f'(x_k)}$$

$$= \frac{f(x_k) - \overset{0}{\overset{\shortparallel}{f(x^*)}} - (x_k - x^*) f'(x_k)}{f'(x_k)}$$

Taylor-Lagrange $=$ (à l'ordre 2) $\displaystyle - \frac{(x^* - x_k)^2}{2} f''(\xi_k) \Big/ f'(x_k)$

$$f(x^*) = f(x_k) + (x^* - x_k) f'(x_k) + \frac{(x^* - x_k)^2}{2} f''(\xi_k)$$

Sur le segment reliant x^* & x_k — ξ_k ⟹ point intermédiaire

$f(t x^* + (1-t) x_k) = \varphi(t)$. Par TL : $\exists \xi$ point intermédiaire

$$\varphi(1) = \varphi(0) + t\,\varphi'(0) + \frac{t^2}{2} \varphi''(\xi)$$

En tout cas :
$$\boxed{|x^* - x_{k+1}| = |x^* - x_k|^2 \frac{|f''(\xi_k)|}{2\,|f'(x_k)|}}$$
$\exists\, \xi_k$ entre x^* et x_k tel que $\mathcal{F}$ — ($\mathcal{F}$)

* Par hypothèse si $k = 0$, alors

$$|x^* - x_0| \cdot \frac{|f''(\xi_k)|}{2|f'(x_k)|}$$

$$\leq |x^* - x_0| \cdot \frac{1}{2} \cdot \sup_{|h| \leq |x^* - x_0|} |f''(x^* + h)| \times \sup_{|h| \leq |x^* - x_0|} \frac{1}{|f'(x^* + h)|} \leq 1.$$

$$\leq \frac{1}{2}$$

Donc l'hypothèse implique

$$|x^* - x_1| \overset{\star}{=} |x^* - x_0| \times \underbrace{|x^* - x_0| \left| \frac{f''(\xi_k)}{2f'(x_k)} \right|}_{\leq 1/2}$$

$$\leq \frac{|x^* - x_0|}{2}.$$

Ainsi x_1 se rapproche de moitié et satisfait aussi l'hypothèse (Se rendre compte que si un point satisfait l'hypothèse, tout point plus proche de x^* aussi).

Par récurrence immédiate :

$$|x^* - x_{k+1}| \leq \frac{1}{2} |x^* - x_k|$$

$$\rightsquigarrow \quad |x^* - x_k| \leq \frac{1}{2^k} |x^* - x_0| \xrightarrow[k \to +\infty]{} 0$$

Nous avons montré la CV.

$$* \quad \star \quad |x_{k+1} - x^*| = |x_k - x^*|^2 \left| \frac{f''(\xi_k)}{2f'(x_k)} \right|$$

Mais $\begin{cases} x_k \longrightarrow x^* \\ \xi_k \longrightarrow x^* \end{cases}$ car point entre x_k et x^*.

Par continuité de f' et f'' :

$$|x_{k+1} - x^*| \underset{k \to +\infty}{\sim} |x_k - x^*|^2 \left| \frac{f''(x^*)}{2 f'(x^*)} \right|$$

$\boxed{/\!/}$