ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Optimisation non linéaire sans contraintes

Recherche opérationnelle

GC-SIE

Conditions d'optimalité

Fonctions à une variable

- min $f(x)$, $x \in IR$
- Définitions :
- x^* est un **maximum local** de f s'il existe a > 0 tel que $f(x^*) \geq f(x)$ pour tout $x \in]x^*-a,x^*+a[$
- x^* est un **minimum local** de f s'il existe a > 0 tel que $f(x^*) \leq f(x)$ pour tout $x \in]x^*-a,x^*+a[$
- L'intervalle $]x^*-a,x^*+a[$ est appelé un **voisinage** de x^*

Fonctions à une variable

Définitions :

- x^* est un **maximum local strict** de f s'il existe a > 0 tel que
  $f(x^*) > f(x)$ pour tout $x \in]x^*-a,x^*+a[$
- x^* est un **minimum local strict** de f s'il existe a > 0 tel que
  $f(x^*) < f(x)$ pour tout $x \in]x^*-a,x^*+a[$

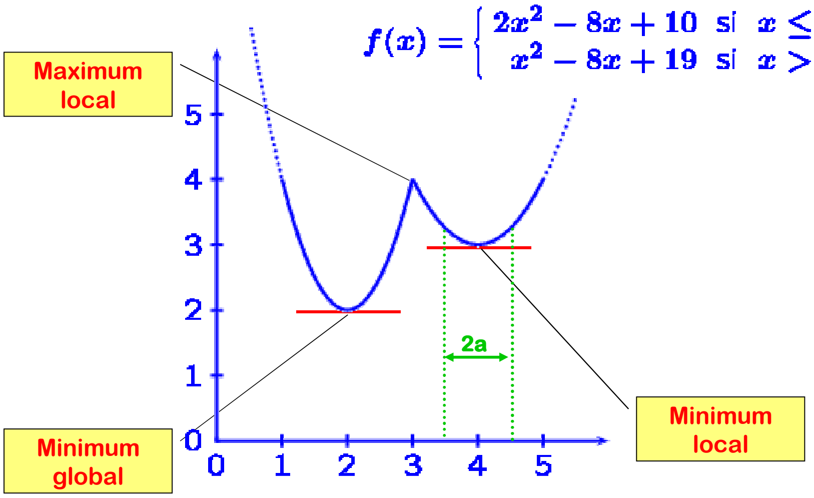
Fonctions à une variable

Définitions :

- x* est un **maximum global** de f si
  $f(x^*) \geq f(x)$ pour tout $x \in$ IR
- x* est un **minimum global** de f si
  $f(x^*) \leq f(x)$ pour tout $x \in$ IR
- Un **extremum** est un minimum ou un maximum.

Fonctions à une variable

$$f(x) = \begin{cases} 2x^2 - 8x + 10 & \text{si } x \leq 3 \\ x^2 - 8x + 19 & \text{si } x > 3 \end{cases}$$



Fonctions à une variable

Définition :

- Un point x où la tangente est horizontale, c'est-à-dire tel que $f'(x)=0$, est appelé un **point critique** ou **point stationnaire**.

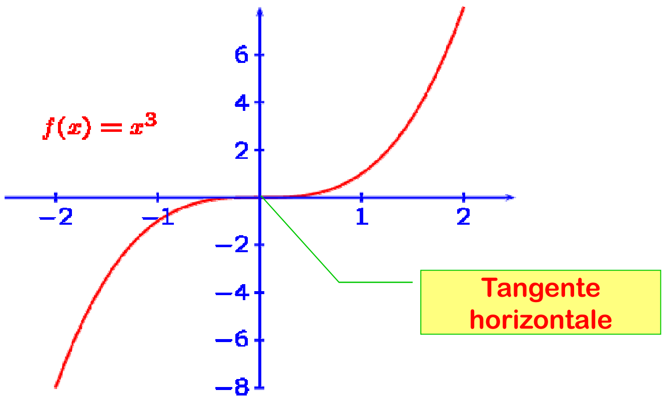
Théorème de Fermat :

- Si une fonction continue f possède un extremum local en x^*, et si $f'(x^*)$ existe, alors $f'(x^*) = 0$.

Fonctions à une variable

- La condition $f'(x^*) = 0$ est une **condition nécessaire d'optimalité** pour une fonction différentiable.
- Attention : ce n'est pas une condition suffisante.
- Rappel: Si $P \Rightarrow Q$, alors P est suffisante et Q est nécessaire.
- x^* optimal $\Rightarrow f'(x^*) = 0$

Fonctions à une variable



mais pas un maximum, ni un minimum…

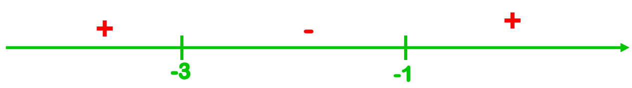
Fonctions à une variable

Test de premier ordre :

- Soit une fonction différentiable f, et x* un point critique. S'il existe a > 0 tel que
  - f'(x) > 0 si x*-a < x < x*
  - f'(x) < 0 si x* < x < x*+a
- Alors x* est un maximum local de f.
- Soit une fonction différentiable f, et x* un point critique. S'il existe a > 0 tel que
  - f'(x) < 0 si x*-a < x < x*
  - f'(x) > 0 si x* < x < x*+a
- Alors x* est un minimum local de f.

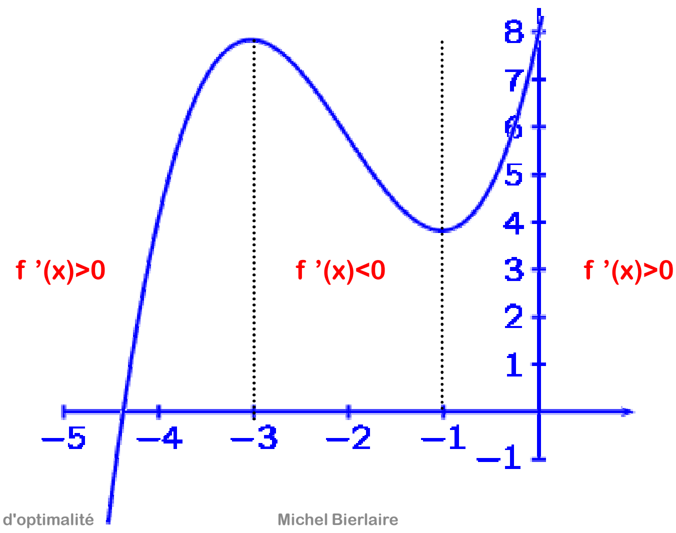
Fonctions à une variable

- Soit $f(x) = x^3 + 6\,x^2 + 9\,x + 8$
- $f\,'(x) = 3\,x^2 + 12\,x + 9 = 3(x+3)(x+1)$
- Points critiques : $x_1 = -3$ et $x_2 = -1$
- Signes de $f\,'(x)$:



- x_1 maximum local
- x_2 minimum local

Fonctions à une variable



Fonctions à une variable

Test de premier ordre :

- Condition suffisante d'optimalité d'un point critique.
- Inconvénient : il faut de l'information sur f à d'autres points que le point critique.

Fonctions à une variable

Test de second ordre :

- Si la fonction f possède une dérivée seconde continue dans un voisinage d'un point critique x^*, alors
  - $f''(x^*) < 0$ est une condition suffisante pour que x^* soit un maximum local, et
  - $f''(x^*) > 0$ est une condition suffisante pour que x^* soit un minimum local.

Fonctions à une variable

- Soit $f(x) = x^3 + 6x^2 + 9x + 8$
- $f'(x) = 3x^2 + 12x + 9 = 3(x+3)(x+1)$
- Points critiques : $x_1 = -3$ et $x_2 = -1$
- $f''(x) = 6x + 12$
- $f''(-3) = -6 < 0 \Rightarrow x_1$ maximum local
- $f''(-1) = 6 > 0 \Rightarrow x_2$ minimum local

Fonctions multivariables

Rappels :

- Soit $f : IR^n \to IR : (x_1,\ldots,x_n)^T \to f(x_1,\ldots,x_n)$
- Si la limite

  $$\lim_{\alpha \to 0} \frac{f(x + \alpha e_i) - f(x)}{\alpha}$$

  existe, elle est appelée la *i* ième **dérivée partielle** de f.
- e_i étant le *i* ième vecteur unité, composé de zéros, sauf la *i* ième composante qui est 1.
- Elle est notée $\frac{\partial f}{\partial x_i}(x)$ ou $\nabla_i f(x)$

Fonctions multivariables

Rappels:

- Si toutes les dérivées partielles existent, le **gradient de f en x** est le vecteur colonne

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

Fonctions multivariables

Rappels :

- Soit f: $IR^n \to IR^m$. Si chaque composante f_i, i=1,…m, est (continûment) différentiable, alors f est dite (continûment) différentiable.
- La matrice n x m dont la colonne i est le gradient $\nabla f_i(x)$ est la **matrice gradient** de f en x.
- La transposée de la matrice gradient est appelée **matrice jacobienne** ou **Jacobien** de f en x.

Fonctions multivariables

Rappels :

- Soit x et $d \in IR^n$. La **dérivée directionnelle** de f en x dans la direction d est

$$f'(x;d) = \lim_{\alpha \downarrow 0} \frac{f(x+\alpha d) - f(x)}{\alpha}$$

à condition que la limite existe.

- Si toutes les dérivées directionnelles de f en x existent, alors f est (Gateaux) **différentiable** en x.

Fonctions multivariables

Rappels :

- Si f est différentiable sur un ensemble ouvert S, et le gradient $\nabla f(x)$ est une fonction continue de x, alors f est **continûment différentiable** sur S.
- Si toutes les dérivées partielles de f sont continûment différentiables, alors $\partial^2 f/ \partial x_i \partial x_j(x)$ est la i ième dérivée partielle de $\partial f/ \partial x_j$ en x.

Fonctions multivariables

Rappels :

- La matrice symétrique $\nabla^2 f(x)$, dont la cellule (i,j) est $\partial^2 f / \partial x_i \partial x_j(x)$ est appelée la **matrice des dérivées secondes**, ou **matrice hessienne**, ou encore le **Hessien** de f en x.
- Soient $f:IR^k \longrightarrow IR^m$ et $g:IR^m \longrightarrow IR^n$ deux fonctions continûment différentiables, et h leur composée, c-à-d h(x)=g(f(x)).Alors,

$$\nabla h(x) = \nabla f(x)\ \nabla(g(f(x))).$$

- Notamment, $\nabla(f(Ax)) = A^T \nabla f(Ax)$, où A est une matrice.

Fonctions multivariables

Rappels :

- Soit $f:IR^n \longrightarrow IR$ deux fois continûment différentiable sur une sphère ouverte S centrée en x.
- Pour tout d tel que $x+d \in S$, il existe $0 \leq \varepsilon \leq 1$ tel que $f(x+d)=f(x) + d^T \nabla f(x) + ½\ d^T \nabla^2 f(x+\varepsilon d)\ d$.
- Pour tout d tel que $x+d \in S$,

$$f(x+d) = f(x) + d^T \nabla f(x) + ½\ d^T\ \nabla^2 f(x)\ d + o(\|d\|^2)$$

Fonctions multivariables

Rappels : notation o(.)

- Soit p un entier positif
- Soit h: $IR^n \longrightarrow IR^m$
- Alors

$$h(x) = o\left(\|x\|^p\right)$$

  ssi

$$\lim_{x_k \to 0} \frac{h(x_k)}{\|x_k\|^p} = 0$$

pour toute suite $\{x_k\}$, sans élément nul, et convergeant vers 0.

Fonctions multivariables

Conditions nécessaires d'optimalité

- Soit x* un minimum local de f: $IR^n \longrightarrow IR$. Si f est continûment différentiable sur un ouvert S contenant x*, alors

$\nabla f(x^*)=0$.

- Si, de plus, f est deux fois continûment différentiable sur S, alors

$\nabla^2 f(x^*)$ est semi définie positive

$d^T \nabla^2 f(x^*)\, d \geq 0$ pour tout $d \in IR^n$

Fonctions multivariables

Conditions suffisantes d'optimalité

- Soit f: $IR^n \rightarrow IR$ une fonction deux fois continûment différentiable sur un ouvert S. Si $x^* \in S$ satisfait les conditions

$$\nabla f(x^*)=0$$

et

$$\nabla^2 f(x^*) \text{ est définie positive}$$

$$d^T \nabla^2 f(x^*)\, d > 0 \text{ pour tout } d \in IR^n, d \neq 0$$

Alors x^* est un **minimum local strict** de f