Université de Nice
Département de Mathématiques

Année 2007-2008
Licence MI/SM 1e année

Analyse : notes du cours 4
Extrema d'une fonction de deux variables

1. Convexité :

Définition : On dit qu'une fonction dérivable $x \mapsto f(x)$ est *convexe* sur un intervalle $I =]a, b[\subseteq \mathbb{R}$ si sa dérivée est croissante sur I (et *strictement convexe* si elle est strictement croissante). On dit qu'elle est *concave* si sa dérivée est décroissante (et *strictement concave* si elle est strictement décroissante[1]).

Ainsi $x \mapsto x^2$ et $x \mapsto \frac{1}{x}$ sont convexes sur $\mathbb{R}$ et $]0, +\infty[$ respectivement. La fonction $x \mapsto x^3$ n'est ni convexe ni concave sur $\mathbb{R}$ mais elle est convexe sur $[0, +\infty[$ et concave sur $]-\infty, 0]$. Notons qu'une fonction f est concave si et seulement si la fonction $-f$ est convexe.

Voici une propriété caractéristique de la convexité que nous pourrons généraliser aux fonctions de deux variables.

Proposition 1 *Une fonction f dérivable est convexe si et seulement si son graphe est situé au-dessus de toutes ses tangentes, c'est-à-dire si, en tout point x_0, on a pour tout x :*

$$f(x) \geq f(x_0) + f'(x_0)(x - x_0). \tag{1}$$

Cette inégalité qui s'écrit aussi $f(x) \geq L(x)$, où $L(x)$ est la linéarisée de f au point x_0, montre que l'approximation d'une fonction convexe par sa linéarisée est toujours une sous-estimation de la valeur exacte. La stricte convexité correspond à la même définition, mais avec une inégalité stricte.

Preuve : Pour prouver qu'une fonction convexe vérifie bien cette inégalité (1), posons $g(x) = f(x) - L(x)$ et montrons que g est une fonction positive ou nulle. On a $g(x_0) = 0$ et d'autre part $g'(x) = f'(x) - f'(x_0)$. Comme f' est croissante, on a $f'(x) - f'(x_0) \leq 0$ si $x \leq x_0$ et $f'(x) - f'(x_0) \geq 0$ si $x \geq x_0$. Donc la fonction g est décroissante pour $x \leq x_0$ et croissante pour $x \geq x_0$. Elle reste donc toujours positive.

Réciproquement, pour montrer que f' est croissante, considérons deux points x_0 et x_1 de I tels que $x_0 \leq x_1$ et montrons que $f'(x_0) \leq f'(x_1)$. L'inégalité (1) au point x_0 est satisfaite si l'on prend pour x la valeur x_1,

$$f(x_1) \geq f(x_0) + f'(x_0)(x_1 - x_0)$$

et la même inégalité au point x_1 cette fois est satisfaite si l'on prend pour x la valeur x_0

$$f(x_0) \geq f(x_1) + f'(x_1)(x_0 - x_1).$$

Il en résulte que d'une part $f'(x_0) \leq \frac{f(x_1)-f(x_0)}{x_1-x_0}$ et d'autre part $\frac{f(x_1)-f(x_0)}{x_1-x_0} \leq f'(x_1)$ (en se souvenant que $x_0 \leq x_1$). D'où $f'(x_0) \leq f'(x_1)$. $\square$

La caractérisation précédente se généralise facilement aux fonctions de deux variables :

Définition : On dit qu'une fonction $(x, y) \mapsto f(x, y)$ dérivable[2] est *convexe* si son graphe est situé au-dessus de tous ses plans tangents, c'est-à-dire si en tout point (x_0, y_0), on a pour tout (x, y) :

$$f(x, y) \geq f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)\,(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)\,(y - y_0)$$

Ici encore la *stricte convexité* correspondra à l'inégalité stricte et une fonction dérivable sera *concave* si son graphe est situé en-dessous de tous ses plans tangents (c'est-à-dire lorsqu'on aura la même inégalité mais inversée). A nouveau, une fonction f est concave lorsque la fonction $-f$ est convexe.

Géométriquement, une fonction convexe a un graphe *dont la concavité est orientée vers le haut*, comme le bol $f(x, y) = x^2 + y^2$ alors qu'une fonction concave a un graphe dont la concavité est orientée vers le bas comme le chapeau $f(x, y) = \sqrt{9 - x^2 - y^2}$. Un plan $f(x, y) = ax + by + c$ est à la fois concave et convexe (mais pas strictement) et une selle $f(x, y) = x^2 - y^2$ n'est ni concave ni convexe.

[1] En fait, la notion de convexité existe plus généralement, pour toute fonction dérivable ou non (voir par exemple `http ://fr.wikipedia.org/wiki/fonction_convexe`)

[2] On appelle ici *dérivable* une fonction qui possède deux dérivées partielles. Contrairement au cas des fonctions d'une seule variable, cette dérivabilité là n'entraine pas que la fonction puisse être suffisamment bien approchée par sa linéarisée. Pour que ce soit le cas, il faudra demander à la fonction d'être *différentiable* (voir par exemple `http ://fr.wikipedia.org/wiki/différentielle`).

2. Conditions d'optimalité du second ordre

On a vu que pour trouver les extrema d'une fonction $x \mapsto f(x)$, la première chose à faire est de rechercher ses points critiques. Parmi ces points critiques, certains sont l'argument d'un maximum local, d'autres l'argument d'un minimum local et d'autres ni l'un ni l'autre. Pour savoir dans quel cas l'on se trouve, il suffit en fait de connaître la concavité de la fonction, c'est-à-dire de savoir distinguer si elle est convexe (c'est la cas d'un minimum) , concave (c'est le cas d'un maximum) ou ni l'un ni l'autre. Pour les fonctions deux fois dérivables, la croissance stricte de f' correspond au fait que $f'' > 0$ et sa décroissance stricte à $f'' < 0$. Il en résulte le critère suivant :

Proposition 2 Condition d'optimalité du second ordre : *Soit $x \mapsto f(x)$ une fonction deux fois dérivable sur $]a, b[$ et soit c un point de $]a, b[$:*

- *si $f'(c) = 0$ et $f''(c) > 0$ alors c est un minimum local de f.*
- *si $f'(c) = 0$ et $f''(c) < 0$ alors c est un maximum local de f.*

Le corollaire suivant est souvent utile dans les applications :

Corollaire 3 *Si une fonction strictement convexe $f :]a, b[\mapsto \mathbb{R}$ possède un point critique, ce point est l'argument de son minimum global et il est unique.*

Dans le cas des fonctions de deux variables, c'est une quantité appelée le *déterminant Hessien*, qui se calcule facilement à partir des dérivées partielles secondes, qui va jouer le rôle de la dérivée seconde f''. Mais auparavant étendons aux fonctions de deux variables les définitions d'extréma locaux et globaux.

Définition : On dit qu'une fonction $f(x, y)$ définie sur un domaine $\mathcal{D} \subseteq \mathbb{R}^2$ possède un *maximum local* au point $(x^*, y^*) \in \mathcal{D}$ s'il existe un rectangle $I \times J =]x^* - \varepsilon, x^* + \varepsilon[\times]x^* - \delta, x^* + \delta[$ contenu dans $\mathcal{D}$ sur lequel $f(x, y) \leq f(x^*, y^*)$ et si cette inégalité est satisfaite en tout point $(x, y) \in \mathcal{D}$ il s'agit d'un *maximum global.* On a évidemment une définition analogue pour *minimum local* et *minimum global.*

Proposition 4 Condition d'optimalité du second ordre : *Soit $(x, y) \mapsto f(x, y)$ une fonction deux fois dérivable sur domaine ouvert $\mathcal{D} =]a, b[\times]a', b'[$ dont les 4 dérivées partielles secondes sont continues et soit (x_c, y_c) un point de $\mathcal{D}$. Soit*

$$D = D(x, y) = \frac{\partial^2 f}{\partial x^2}(x, y)\frac{\partial^2 f}{\partial y^2}(x, y) - \left(\frac{\partial^2 f}{\partial x \partial y}(x, y)\right)^2.$$

- *Si $Gradf(x_c, y_c) = 0$ et $D(x_c, y_c) > 0$ avec $\frac{\partial^2 f}{\partial x^2}(x_c, y_c) > 0$ alors (x_c, y_c) est un minimum local de la fonction f.*
- *Si $Gradf(x_c, y_c) = 0$ et $D(x_c, y_c) > 0$ avec $\frac{\partial^2 f}{\partial x^2}(x_c, y_c) < 0$ alors (x_c, y_c) est un maximum local de la fonction f.*

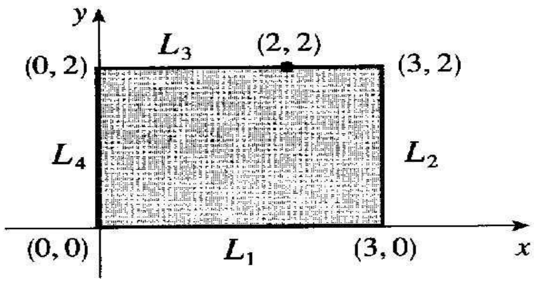
3. Détermination des extrema globaux par la méthode à trois étapes. Comme pour les fonctions d'une variables, on dispose d'une méthode simple pour trouver la plus grande et la plus petite valeur d'une fonction $(x, y) \mapsto f(x, y)$ dérivable lorsque le domaine $\mathcal{D}$ sur lequelle elle est définie est un rectangle (avec son bord), un polygône (avec son bord) ou plus généralement un domaine borné[3] défini par des inégalités larges portant sur des polynômes. Pour localiser ces deux extréma globaux, il suffit en effet de procéder en trois étapes :

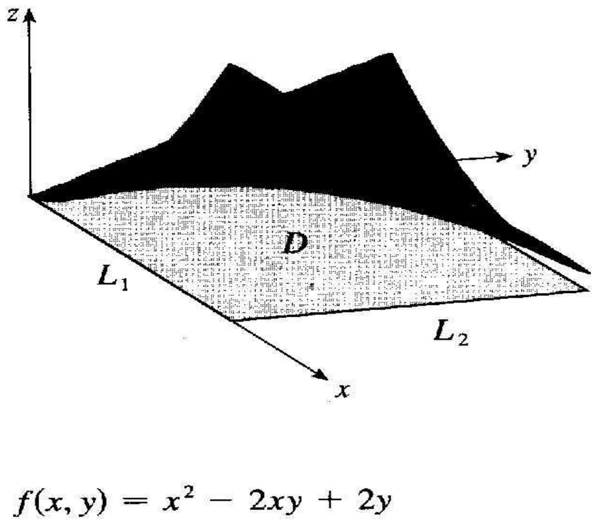
1. Trouver tous les points critiques de f (points où le gradient est nul) et calculer la valeur de f en ces points.
2. Trouver les extrema de f sur le bord (en restriction à chaque segment du bord, f est une fonction d'une seule variable à laquelle on peut appliquer la méthode à trois étapes).
3. La plus grande et la plus petite valeurs trouvées aux deux étapes précédentes sont les extrema cherchés.

[3] Une partie de $\mathbb{R}^2$ est dite *bornée* s'il existe un (possiblement très grand) rectangle $[-M, M] \times [-M, M]$ qui la contient.

Exemple : Prenons l'exemple de la fonction $f(x,y) = x^2 - 2xy + 2y$ sur le rectangle $\mathcal{D} = \{(x,y), 0 \leq x \leq 3, 0 \leq y \leq 2\}$.

1. Pour trouver les points critiques on calcule tout d'abord les deux dérivées partielles $\frac{\partial f}{\partial x}(x,y) = 2x - 2y$ et $\frac{\partial f}{\partial y}(x,y) = -2x + 2$. Puis on constate que ces deux dérivées partielles ne s'annulent simultanément qu'en un seul point $(x_c, y_c) = (1,1)$ et qu'en ce point on a $f(1,1) = 1$.
2. Pour trouver les extrema de f sur le bord du rectangle, on étudie successivement la restriction de f à chacun des 4 cotés (voir la figure).
   - Sur L_1 : la fonction s'écrit $f(x,0) = x^2$ avec $0 \leq x \leq 3$, c'est une fonction croissante de minimum $f(0,0) = 0$ et de maximum $f(3,0) = 9$.
   - Sur L_2 : la fonction s'écrit $f(3,y) = 9 - 4y$ avec $0 \leq y \leq 2$, c'est une fonction décroissante de minimum $f(3,2) = 1$ et de maximum $f(3,0) = 9$.
   - Sur L_3 : la fonction s'écrit $f(x,2) = x^2 - 4x + 4$ avec $0 \leq x \leq 3$, c'est une fonction décroissante puis croissante (utiliser la méthode à trois étapes appliquée à cette fonction) de minimum $f(2,2) = 0$ et de maximum $f(0,2) = 4$.
   - Sur L_4 : la fonction s'écrit $f(0,y) = 2y$ avec $0 \leq y \leq 2$, c'est une fonction croissante de minimum $f(0,0) = 0$ et de maximum $f(0,2) = 4$.
3. En résumé, la plus grande et la plus petite valeurs trouvées aux deux étapes précédentes sont donc respectivement $f(3,0) = 9$ et $f(0,0) = f(2,2) = 0$, toutes deux atteintes au bord du domaine.





$f(x, y) = x^2 - 2xy + 2y$