Cours d'analyse pour économistes

Informations utiles

- Adresse mail :
  - aymeric.baradat@ens.fr
  - aymeric.baradat@polytechnique.edu

Informations utiles

- Adresse mail :
  - aymeric.baradat@ens.fr
  - aymeric.baradat@polytechnique.edu
- Page web : http ://www.math.ens.fr/∼baradat/

Informations utiles

- Adresse mail :
  - aymeric.baradat@ens.fr
  - aymeric.baradat@polytechnique.edu
- Page web : http ://www.math.ens.fr/∼baradat/
- Bureau : bureau C21 dans l'espace Cartan

Informations utiles

- Adresse mail :
  - aymeric.baradat@ens.fr
  - aymeric.baradat@polytechnique.edu
- Page web : http ://www.math.ens.fr/∼baradat/
- Bureau : bureau C21 dans l'espace Cartan
- Organisation de l'année TD, DM ...

Plan du cours

1. Introduction aux fonctions de plusieurs variables.

Plan du cours

1. Introduction aux fonctions de plusieurs variables.
2. Rappels et compléments d'analyse réelle.

Plan du cours

1. Introduction aux fonctions de plusieurs variables.
2. Rappels et compléments d'analyse réelle.
3. Structure euclidienne de $\mathbb{R}^d$.

Plan du cours

1. Introduction aux fonctions de plusieurs variables.
2. Rappels et compléments d'analyse réelle.
3. Structure euclidienne de $\mathbb{R}^d$.
4. Topologie de $\mathbb{R}^d$.

Plan du cours

1. Introduction aux fonctions de plusieurs variables.
2. Rappels et compléments d'analyse réelle.
3. Structure euclidienne de $\mathbb{R}^d$.
4. Topologie de $\mathbb{R}^d$.
5. Compléments d'algèbre linéaire.

Plan du cours

1. Introduction aux fonctions de plusieurs variables.
2. Rappels et compléments d'analyse réelle.
3. Structure euclidienne de $\mathbb{R}^d$.
4. Topologie de $\mathbb{R}^d$.
5. Compléments d'algèbre linéaire.
6. Calcul différentiel pour les fonctions de plusieurs variables.

Plan du cours

1. Introduction aux fonctions de plusieurs variables.
2. Rappels et compléments d'analyse réelle.
3. Structure euclidienne de $\mathbb{R}^d$.
4. Topologie de $\mathbb{R}^d$.
5. Compléments d'algèbre linéaire.
6. Calcul différentiel pour les fonctions de plusieurs variables.
7. Éléments de géométrie.

Plan du cours

1. Introduction aux fonctions de plusieurs variables.
2. Rappels et compléments d'analyse réelle.
3. Structure euclidienne de $\mathbb{R}^d$.
4. Topologie de $\mathbb{R}^d$.
5. Compléments d'algèbre linéaire.
6. Calcul différentiel pour les fonctions de plusieurs variables.
7. Éléments de géométrie.
8. Optimisation sous contraintes.

Chapitre 1 : Introduction aux fonctions de plusieurs variables

Fonctions de plusieurs variables

Une fonction de plusieurs variables, c'est quoi ?

Une application f à $\boldsymbol{n}$ **variables réelles** et à **valeurs réelles** est une application qui prend n nombres réels $x_1, \ldots, x_n$ et qui rend comme valeur un nombre réel noté $f(x_1, \ldots, x_n)$.

Une fonction de plusieurs variables, c'est quoi ?

Une application f à $\boldsymbol{n}$ **variables réelles** et à **valeurs réelles** est une application qui prend n nombres réels $x_1, \ldots, x_n$ et qui rend comme valeur un nombre réel noté $f(x_1, \ldots, x_n)$.

Souvent, on verra les nombres $x_1, \ldots, x_n$ comme un n-uplet $(x_1, \ldots, x_n) \in \mathbb{R}^n$.

Une fonction de plusieurs variables, c'est quoi ?

Selon les cas, les nombres $x_1, \ldots, x_n$ pourront prendre n'importe quelle valeur, on notera alors

$$f : \begin{array}{rcl} \mathbb{R}^n & \to & \mathbb{R} \\ (x_1, \ldots, x_n) & \mapsto & f(x_1, \ldots, x_n). \end{array}$$

Exemples :

- $$f : \begin{array}{rcl} \mathbb{R}^2 & \to & \mathbb{R} \\ (x, y) & \mapsto & x \cos y. \end{array}$$

- $$f : \begin{array}{rcl} \mathbb{R}^2 & \to & \mathbb{R} \\ (x, y) & \mapsto & \begin{cases} x \cos y & \text{si } x \leq y, \\ \exp(x + y^2) & \text{si } x > y. \end{cases} \end{array}$$

- $$f : \begin{array}{rcl} \mathbb{R}^n & \to & \mathbb{R} \\ (x_1, \ldots, x_n) & \mapsto & \sum_{i=1}^{n} \log(1 + (x_{i+1} - x_i)^2), \end{array}$$

Une fonction de plusieurs variables, c'est quoi ?

Exemples :

- $$\begin{aligned} f : \quad & \mathbb{R}^2 \to \mathbb{R} \\ & (x, y) \mapsto x \cos y. \end{aligned}$$

- $$\begin{aligned} f : \quad & \mathbb{R}^2 \to \mathbb{R} \\ & (x, y) \mapsto \begin{cases} x \cos y & \text{si } x \leq y, \\ \exp(x + y^2) & \text{si } x > y. \end{cases} \end{aligned}$$

- $$\begin{aligned} f : \quad & \mathbb{R}^n \to \mathbb{R} \\ & (x_1, \ldots, x_n) \mapsto \sum_{i=1}^{n} \log(1 + (x_{i+1} - x_i)^2), \end{aligned}$$

  avec la convention $n + 1 = 1$ dans la somme.

Ou parfois, $(x_1, \ldots, x_n)$ devra appartenir à un certain sous-ensemble E de $\mathbb{R}^n$. On dira que l'ensemble de définition de f est E et on notera

$$f : \begin{aligned} E &\to \mathbb{R} \\ (x_1, \ldots, x_n) &\mapsto f(x_1, \ldots, x_n). \end{aligned}$$

Exemples :

- $E := \{(x, y) \in \mathbb{R}^2 \,|\, x > 0\} = \mathbb{R}_+^* \times \mathbb{R}$ et

$$\begin{aligned} f : \quad E &\to \mathbb{R} \\ (x, y) &\mapsto y \log x. \end{aligned}$$

- Si on veux donner la quantité $f(k, w)$ produite d'un bien avec une quantité k de capital et une quantité w de travail, on choisit en général $E = (\mathbb{R}_+)^2$.

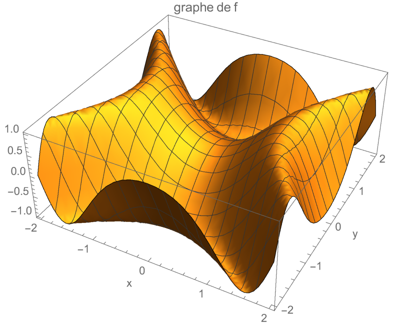
Représentation des fonctions de plusieurs variables

Comment représente-t-on une fonction de plusieurs variables ?

Pour la représentation des fonctions de $\mathbb{R}^2$ dans $\mathbb{R}$ (fonctions de deux variables), il y a deux façons classiques : la représentation en 3D et la représentation par lignes de niveau.

Ici, on représente

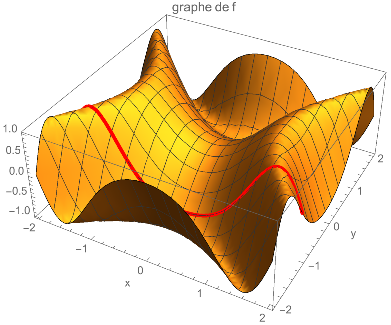
$$\begin{aligned} f: \quad & \mathbb{R}^2 \to \mathbb{R} \\ & (x, y) \mapsto \sin(x^2 - y^2). \end{aligned}$$



Ici, on représente

$$f: \begin{aligned} \mathbb{R}^2 &\to \mathbb{R} \\ (x, y) &\mapsto \sin(x^2 - y^2). \end{aligned}$$

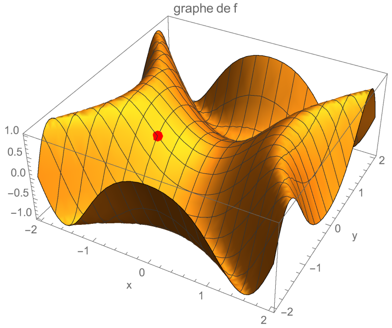
La ligne rouge représente la projection de la droite $y = -1$, $z = 0$ sur le graphe.



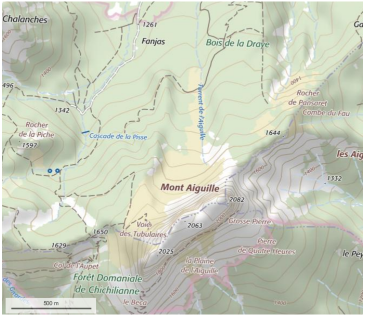
Ici, on représente

$$f: \quad \mathbb{R}^2 \to \mathbb{R}$$
$$(x, y) \mapsto \sin(x^2 - y^2).$$

Le point rouge est la projection sur le graphe du point $(-0.75, -0.5, 0)$. La valeur de f en $(-0.75, -0.5)$ est représenté par la hauteur de ce point.

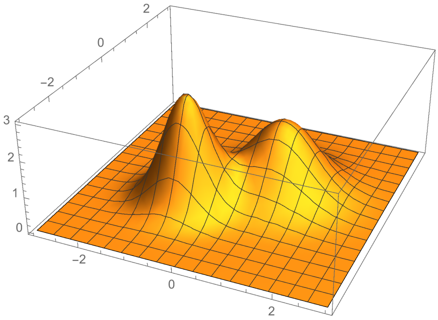


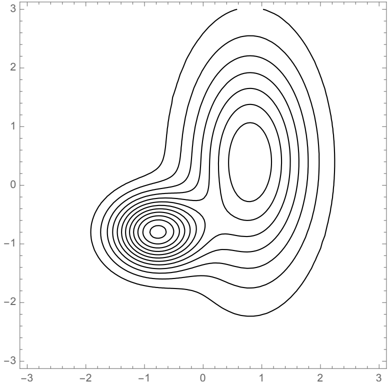
On a l'habitude de voir ces représentations dans les cartes de randonnée : on représente l'altitude en fonction de la latitude et de la longitude à l'aide de lignes de niveau.



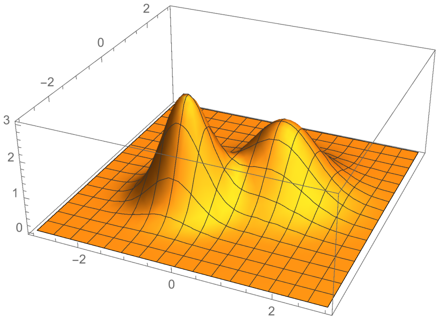
Représentations en 3D et par lignes de niveau de la même fonction, en l'occurence le Mont Aiguille, dans le Vercors.

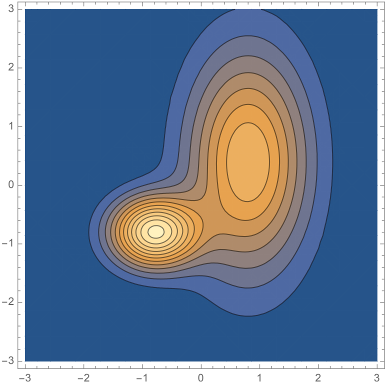
Traçons la même fonction en 3D et en lignes de niveau.





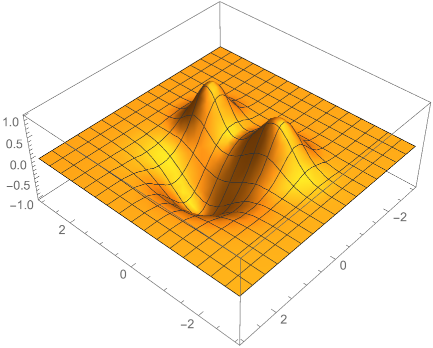
Traçons la même fonction en 3D et en lignes de niveau.

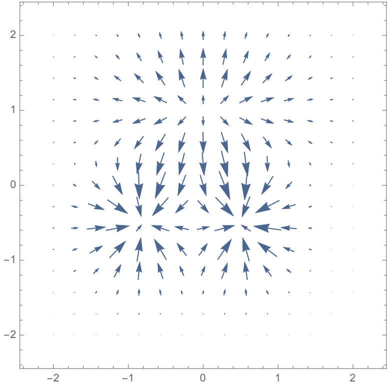




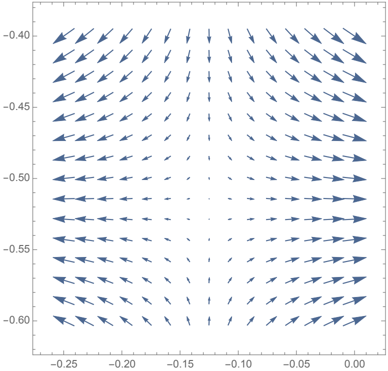
Il y a une dernière façon qui n'est pas très courante pour représenter les fonctions de deux variables. Elle consiste en chaque point à représenter une flèche pointant vers la ligne de plus grande pente, et d'autant plus grande que la pente est raide (en fait, on trace le gradient de la fonction).

Traçons la même fonction en 3D et par champ de gradients.





Si on zoom autour du col, on obtient le graphe suivant.



Dérivation des fonctions de plusieurs variables

Comment dérive-t-on une fonction de plusieurs variables ?

Si f est une fonction de $\mathbb{R}$ dans $\mathbb{R}$, et si $x_0 \in \mathbb{R}$, la dérivée $f'(x_0)$ donne le **taux d'accroissement** de f en x_0.

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

Si f est une fonction de $\mathbb{R}$ dans $\mathbb{R}$, et si $x_0 \in \mathbb{R}$, la dérivée $f'(x_0)$ donne le **taux d'accroissement** de f en x_0.

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

Si on avance d'un petit incrément h à partir de x_0, f augmente du petit incrément $f'(x_0) \times h$:

Comment dérive-t-on une fonction de plusieurs variables ?

Si f est une fonction de $\mathbb{R}$ dans $\mathbb{R}$, et si $x_0 \in \mathbb{R}$, la dérivée $f'(x_0)$ donne le **taux d'accroissement** de f en x_0.

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

Si on avance d'un petit incrément h à partir de x_0, f augmente du petit incrément $f'(x_0) \times h$:

$$f(x_0 + h) \approx f(x_0) + f'(x_0) \times h,$$

ou encore si x proche de x_0,

$$f(x) \approx f(x_0) + f'(x_0) \times (x - x_0).$$

Comment dérive-t-on une fonction de plusieurs variables ?

Si f est une fonction de $\mathbb{R}$ dans $\mathbb{R}$, et si $x_0 \in \mathbb{R}$, la dérivée $f'(x_0)$ donne le **taux d'accroissement** de f en x_0.

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

Si on avance d'un petit incrément h à partir de x_0, f augmente du petit incrément $f'(x_0) \times h$:

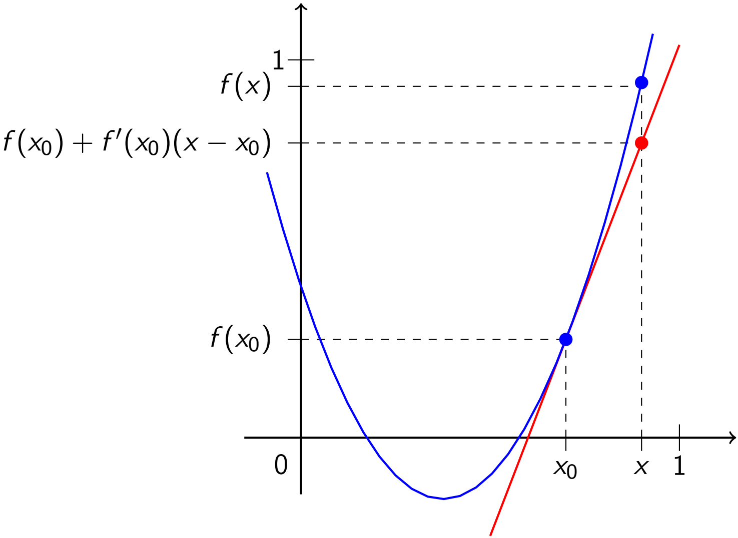
$$f(x_0 + h) \approx f(x_0) + f'(x_0) \times h,$$

ou encore si x proche de x_0,

$$f(x) \approx f(x_0) + f'(x_0) \times (x - x_0).$$

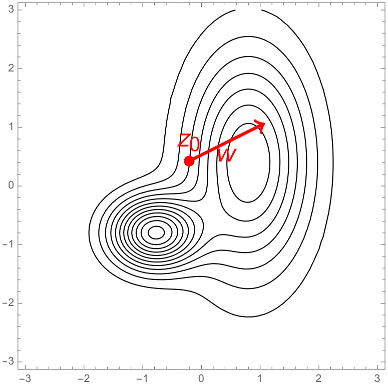
Remarque : $y = f(x_0) + f'(x_0) \times (x - x_0)$ est l'équation de la tangente au graphe de f en x_0.

Comment dérive-t-on une fonction de plusieurs variables ?



Pour les fonctions de plusieurs variables, on peut calculer le taux d'accroissement dans une certaine direction. Illustrons ça en dimension 2. Imaginons que l'on veuille connaître le taux d'accroissement de la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ au point $z_0 = (x_0, y_0) \in \mathbb{R}^2$ dans la direction $w = (u, v) \in \mathbb{R}^2$.

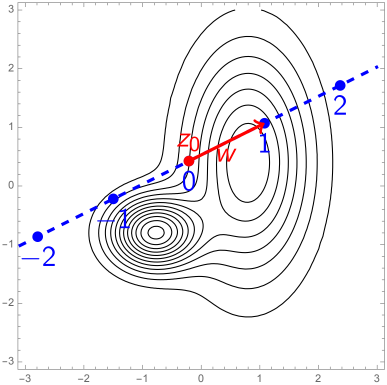
Comment dérive-t-on une fonction de plusieurs variables ?



On définit alors la fonction réelle d'une variable réelle

$$\begin{aligned} \varphi_{z_0,w} : \quad & \mathbb{R} \to \mathbb{R} \\ & t \mapsto f(z_0 + tw) = f(x_0 + tu, y_0 + tv). \end{aligned}$$

Comment dérive-t-on une fonction de plusieurs variables ?



Enfin, on dérive $\varphi_{z_0,w}$ en $t = 0$. On obtient la **différentielle de $\boldsymbol{f}$ au point $\boldsymbol{z_0}$ appliquée à $\boldsymbol{w}$:**

$$\mathrm{d}f[z_0](w) := \varphi'_{z_0,w}(0) = \lim_{t\to 0} \frac{f(z_0 + tw) - f(z_0)}{t}.$$

Enfin, on dérive $\varphi_{z_0,w}$ en $t = 0$. On obtient la **différentielle de $\boldsymbol{f}$ au point $\boldsymbol{z_0}$ appliquée à $\boldsymbol{w}$:**

$$\mathrm{d}f[z_0](w) := \varphi'_{z_0,w}(0) = \lim_{t\to 0} \frac{f(z_0+tw) - f(z_0)}{t}.$$

Si w est petit,

$$f(z_0+w) \approx f(z_0) + \mathrm{d}f[z_0](w),$$

ou encore si $z = (x, y)$ est proche de z_0,

$$f(z) \approx f(z_0) + \mathrm{d}f[z_0](z - z_0).$$

Or pour une grande classe de f, $\mathrm{d}f[z_0]$ est linéaire en w :

$$\mathrm{d}f[z_0](\lambda w_1 + \mu w_2) = \lambda\, \mathrm{d}f[z_0](w_1) + \mu\, \mathrm{d}f[z_0](w_2).$$

Or pour une grande classe de f, $\mathrm{d}f[z_0]$ est linéaire en w :

$$\mathrm{d}f[z_0](\lambda w_1 + \mu w_2) = \lambda\,\mathrm{d}f[z_0](w_1) + \mu\,\mathrm{d}f[z_0](w_2).$$

Donc pour tout $w = (u, v) \in \mathbb{R}^2$,

$$\mathrm{d}f[z_0](w) = u\,\mathrm{d}f[z_0](e_1) + v\,\mathrm{d}f[z_0](e_2),$$

où $e_1 = (1, 0)$ et $e_2 = (0, 1)$.

Or pour une grande classe de f, $\mathrm{d}f[z_0]$ est linéaire en w :

$$\mathrm{d}f[z_0](\lambda w_1 + \mu w_2) = \lambda \, \mathrm{d}f[z_0](w_1) + \mu \, \mathrm{d}f[z_0](w_2).$$

Donc pour tout $w = (u, v) \in \mathbb{R}^2$,

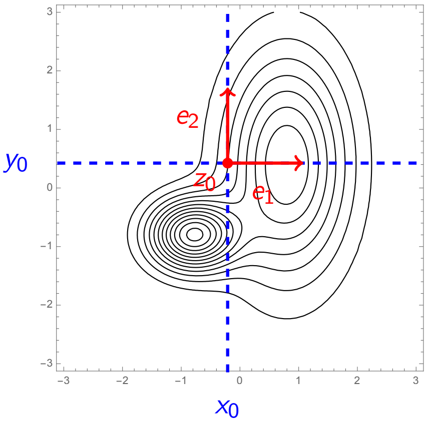
$$\mathrm{d}f[z_0](w) = u \, \mathrm{d}f[z_0](e_1) + v \, \mathrm{d}f[z_0](e_2),$$

où $e_1 = (1, 0)$ et $e_2 = (0, 1)$. **Donc les nombres $\mathbf{d}\boldsymbol{f[z_0](e_1)}$ et $\mathbf{d}\boldsymbol{f[z_0](e_2)}$ suffisent à caractériser les variations de $\boldsymbol{f}$ autour de $\boldsymbol{z_0}$.**

Comment dérive-t-on une fonction de plusieurs variables ?

Ils correspondent aux variations de f quand x varie, mais pas y, et quand y varie, mais pas x respectivement.

Comment dérive-t-on une fonction de plusieurs variables ?



On les appelle les **dérivées partielles de $\boldsymbol{f}$ dans les directions $\boldsymbol{x}$ et $\boldsymbol{y}$**, et on les note

$$\frac{\partial f}{\partial x}(z_0) := \mathrm{d}f[z_0](e_1) \quad \text{et} \quad \frac{\partial f}{\partial y}(z_0) := \mathrm{d}f[z_0](e_2).$$

On les calcule en faisant comme si l'une des coordonnées dans l'expression de f était une constante, et donc comme si f n'était qu'une fonction de l'autre.

Exemple : on définit

$$
\begin{aligned}
f : \quad & \mathbb{R}^2 \to \mathbb{R} \\
& (x, y) \mapsto xy + x^3y^2 + 3x + 2y + 1.
\end{aligned}
$$

Dans ce cas,

$$\frac{\partial f}{\partial x}(x, y) = y + 3x^2y^2 + 3,$$

et

$$\frac{\partial f}{\partial y}(x, y) = x + 2x^3y + 2.$$

Comment dérive-t-on une fonction de plusieurs variables ?

Si on revient sur l'approximation de f, on obtient maintenant que si $w = (u, v)$ est petit,

$$f(z_0 + w) \approx f(z_0) + \frac{\partial f}{\partial x}(z_0) \times u + \frac{\partial f}{\partial y}(z_0) \times v,$$

ou si $z = (x, y)$ est proche de z_0,

$$f(z) \approx f(z_0) + \frac{\partial f}{\partial x}(z_0) \times (x - x_0) + \frac{\partial f}{\partial y}(z_0) \times (y - y_0).$$

Si on revient sur l'approximation de f, on obtient maintenant que si $w = (u, v)$ est petit,

$$f(z_0 + w) \approx f(z_0) + \frac{\partial f}{\partial x}(z_0) \times u + \frac{\partial f}{\partial y}(z_0) \times v,$$

ou si $z = (x, y)$ est proche de z_0,

$$f(z) \approx f(z_0) + \frac{\partial f}{\partial x}(z_0) \times (x - x_0) + \frac{\partial f}{\partial y}(z_0) \times (y - y_0).$$

Remarque : en notant Z la troisième coordonnée d'espace (pas terrible...) l'équation

$$Z = f(z_0) + \frac{\partial f}{\partial x}(z_0) \times (x - x_0) + \frac{\partial f}{\partial y}(z_0) \times (y - y_0)$$

est l'équation du plan tagent à f en z_0.

Les deux dérivées partielles de f forment un vecteur de $\mathbb{R}^2$ que l'on appelle le gradient de f et que l'on note

$$\nabla f(z_0) := \begin{pmatrix} \dfrac{\partial f}{\partial x}(z_0) \\ \dfrac{\partial f}{\partial y}(z_0) \end{pmatrix}.$$

Les deux dérivées partielles de f forment un vecteur de $\mathbb{R}^2$ que l'on appelle le gradient de f et que l'on note

$$\nabla f(z_0) := \begin{pmatrix} \dfrac{\partial f}{\partial x}(z_0) \\ \dfrac{\partial f}{\partial y}(z_0) \end{pmatrix}.$$

- Il est orthogonal aux lignes de niveau.

Comment dérive-t-on une fonction de plusieurs variables ?

Les deux dérivées partielles de f forment un vecteur de $\mathbb{R}^2$ que l'on appelle le gradient de f et que l'on note

$$\nabla f(z_0) := \begin{pmatrix} \dfrac{\partial f}{\partial x}(z_0) \\ \dfrac{\partial f}{\partial y}(z_0) \end{pmatrix}.$$

- Il est orthogonal aux lignes de niveau.
- Il est dans la ligne de plus grande pente.

Les deux dérivées partielles de f forment un vecteur de $\mathbb{R}^2$ que l'on appelle le gradient de f et que l'on note

$$\nabla f(z_0) := \begin{pmatrix} \dfrac{\partial f}{\partial x}(z_0) \\ \dfrac{\partial f}{\partial y}(z_0) \end{pmatrix}.$$

- Il est orthogonal aux lignes de niveau.
- Il est dans la ligne de plus grande pente.
- Il est d'autant plus grand que la pente est raide.

Comment trouve-t-on les extrema d'une fonction de plusieurs variables ?

Par ailleurs, si z_0 est un point de maximum ou minimum local, le taux d'accroissement de f en z_0 est nul dans toutes les directions, ce qui donne le système

$$\begin{cases} \dfrac{\partial f}{\partial x}(z_0) = 0, \\ \dfrac{\partial f}{\partial y}(z_0) = 0, \end{cases}$$

ou encore

$$\nabla f(z_0) = 0.$$

Comment trouve-t-on les extrema d'une fonction de plusieurs variables ?

Par ailleurs, si z_0 est un point de maximum ou minimum local, le taux d'accroissement de f en z_0 est nul dans toutes les directions, ce qui donne le système

$$\begin{cases} \dfrac{\partial f}{\partial x}(z_0) = 0, \\ \dfrac{\partial f}{\partial y}(z_0) = 0, \end{cases}$$

ou encore

$$\nabla f(z_0) = 0.$$

C'est ce système que l'on écrit lorsque l'on cherche les points d'optimum de f.

Comment trouve-t-on les extrema d'une fonction de plusieurs variables ?

Par ailleurs, si z_0 est un point de maximum ou minimum local, le taux d'accroissement de f en z_0 est nul dans toutes les directions, ce qui donne le système

$$\begin{cases} \dfrac{\partial f}{\partial x}(z_0) = 0, \\ \dfrac{\partial f}{\partial y}(z_0) = 0, \end{cases}$$

ou encore

$$\nabla f(z_0) = 0.$$

C'est ce système que l'on écrit lorsque l'on cherche les points d'optimum de f.

Si z_0 annule le gradient, on dit que c'est un **point critique** de f.

En particulier les minima locaux ou les maxima locaux sont des points critiques.

Généralisation en plus grande dimension.

Si maintenant f est une fonction réelle de n variables réelles. On note $x = (x_1, \dots, x_n)$ ou $y = (y_1, \dots, y_n)$ les points de $\mathbb{R}^n$ et $v = (v_1, \dots, v_n)$ les directions dans $\mathbb{R}^n$.

Généralisation en plus grande dimension.

Si maintenant f est une fonction réelle de n variables réelles. On note $x = (x_1, \ldots, x_n)$ ou $y = (y_1, \ldots, y_n)$ les points de $\mathbb{R}^n$ et $v = (v_1, \ldots, v_n)$ les directions dans $\mathbb{R}^n$.
Cette fois f admet n dérivées partielles en x

$$\frac{\partial f}{\partial x_1}(x), \quad \ldots \quad , \frac{\partial f}{\partial x_n}(x)$$

Généralisation en plus grande dimension.

Si maintenant f est une fonction réelle de n variables réelles. On note $x = (x_1, \dots, x_n)$ ou $y = (y_1, \dots, y_n)$ les points de $\mathbb{R}^n$ et $v = (v_1, \dots, v_n)$ les directions dans $\mathbb{R}^n$.
Cette fois f admet n dérivées partielles en x

$$\frac{\partial f}{\partial x_1}(x), \quad \dots \quad , \frac{\partial f}{\partial x_n}(x)$$

Qui se calculent à chaque fois en fixant toutes les coordonnées sauf une.

Elles forment le gradient de f en x

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

Elles forment le gradient de f en x

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

- Il est orthogonal aux surfaces de niveau.

Elles forment le gradient de f en x

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

- Il est orthogonal aux surfaces de niveau.
- Il est dans la ligne de plus grande pente.

Généralisation en plus grande dimension.

Elles forment le gradient de f en x

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

- Il est orthogonal aux surfaces de niveau.
- Il est dans la ligne de plus grande pente.
- Il est d'autant plus grand que la pente est raide.

Généralisation en plus grande dimension.

On peut déduire des dérivées partielles de f en x une approximation de f autour de x. Si v est petit,

$$f(x+v) \approx f(x) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x) \times v_i,$$

ou si y proche de x,

$$f(y) \approx f(x) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x) \times (y_i - x_i).$$

Généralisation en plus grande dimension.

On peut déduire des dérivées partielles de f en x une approximation de f autour de x. Si v est petit,

$$f(x+v) \approx f(x) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x) \times v_i,$$

ou si y proche de x,

$$f(y) \approx f(x) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x) \times (y_i - x_i).$$

En fixant x, et en notant y_{n+1} la $(n+1)$-ième coordonnée dans $\mathbb{R}^{n+1}$, l'équation en y

$$y_{n+1} = f(x) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x) \times (y_i - x_i)$$

est l'équation du plan tangent à f en x.

Généralisation en plus grande dimension.

Aux points d'extremum de f, toutes les dérivées partielles de f s'annulent. Cela donne un système de n équations à n inconnues que l'on doit résoudre pour trouver ces points.

Aux points d'extremum de f, toutes les dérivées partielles de f s'annulent. Cela donne un système de n équations à n inconnues que l'on doit résoudre pour trouver ces points.
Encore une fois, les solutions de ce système s'appellent les **points critiques** de f. Les minima locaux et les maxima locaux sont des points critiques.

Calculs d'ordre 2

Dérivées d'ordre supérieur.

Retournons en dimension 2. On note

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

Retournons en dimension 2. On note

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

Cela forme en chaque point une matrice 2×2 appelée la matrice **hessienne** de f

$$H_f(z_0) := \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(z_0) & \frac{\partial^2 f}{\partial y \partial x}(z_0) \\ \frac{\partial^2 f}{\partial x \partial y}(z_0) & \frac{\partial^2 f}{\partial y^2}(z_0) \end{pmatrix}.$$

Dans notre exemple de tout à l'heure, on trouve

$$\frac{\partial^2 f}{\partial x^2}(x,y) = 6x^2 y, \qquad \frac{\partial^2 f}{\partial y \partial x}(x,y) = 1 + 6x^2,$$
$$\frac{\partial^2 f}{\partial x \partial y}(x,y) = 1 + 6x^2 y, \qquad \frac{\partial^2 f}{\partial y^2}(x,y) = 2x^3.$$

Dérivées d'ordre supérieur.

Théorème (de Schwartz)

Pour une grande classes de fonction f de deux variables

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}.$$

(Il suffit que toutes les dérivées partielles de f d'ordre 2 existent et soient continues.)

Théorème (de Schwartz)

Pour une grande classes de fonction f de deux variables

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}.$$

(Il suffit que toutes les dérivées partielles de f d'ordre 2 existent et soient continues.)

En particulier, H_f est en chaque point **symétrique**.

Dérivées d'ordre supérieur.

En dimension supérieur, si f a n variables, les dérivées partielles de f d'ordre 2 forment une matrice $n \times n$: en tout point x,

$$H_f(x) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2}(x) & \dots & \dfrac{\partial^2 f}{\partial x_n \partial x_1}(x) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(x) & \dots & \dfrac{\partial^2 f}{\partial x_n^2}(x) \end{pmatrix}.$$

Dérivées d'ordre supérieur.

Théorème (de Schwartz)

Pour une grande classes de fonction f de n variables, pour tous indices $i \neq j$ dans $\{1, \ldots, n\}$,

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j}.$$

(Il suffit que toutes les dérivées partielles de f d'ordre 2 existent et soient continues.)

Dérivées d'ordre supérieur.

Théorème (de Schwartz)

Pour une grande classes de fonction f de n variables, pour tous indices $i \neq j$ dans $\{1, \ldots, n\}$,

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j}.$$

(Il suffit que toutes les dérivées partielles de f d'ordre 2 existent et soient continues.)

En particulier, H_f est en chaque point **symétrique**, c'est à dire qu'elle est égale à sa transposée.

Théorème spectral

Cette structure symétrique est très intéressante dans l'étude des points critiques d'une fonction f grâce au théorème spectral.

Théorème (Théorème spectral)

Soit M une matrice carrée symétrique. Alors elle est diagonalisable en base orthonormée.

Le signe des valeurs propres de la hessienne d'une fonction en un de ces points critiques nous renseigne sur le type de point critique auquel on a affaire.

Application du théorème spectral en dimension 2

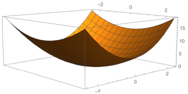
Soit $f : \mathbb{R}^2 \to \mathbb{R}$ et $z = (x, y)$ un point critique de f. On note α et β les valeurs propres de la matrice symétrique $H_f(z)$. Dans ce cas, dans une certaine base, autour de z, f ressemble à

$$\alpha \frac{X^2}{2} + \beta \frac{Y^2}{2}.$$

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ et $z = (x, y)$ un point critique de f. On note α et β les valeurs propres de la matrice symétrique $H_f(z)$. Dans ce cas, dans une certaine base, autour de z, f ressemble à

$$\alpha \frac{X^2}{2} + \beta \frac{Y^2}{2}.$$

• Si $\alpha > 0$ et $\beta > 0$, f ressemble à :
z est un point de minimum local.

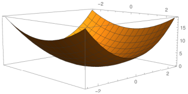


Application du théorème spectral en dimension 2

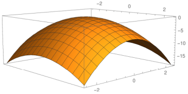
Soit $f : \mathbb{R}^2 \to \mathbb{R}$ et $z = (x, y)$ un point critique de f. On note α et β les valeurs propres de la matrice symétrique $H_f(z)$. Dans ce cas, dans une certaine base, autour de z, f ressemble à

$$\alpha \frac{X^2}{2} + \beta \frac{Y^2}{2}.$$

• Si $\alpha > 0$ et $\beta > 0$, f ressemble à :
z est un point de minimum local.



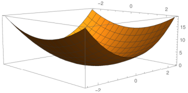
• Si $\alpha < 0$ et $\beta < 0$, f ressemble à :
z est un point de maximum local.



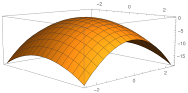
Soit $f : \mathbb{R}^2 \to \mathbb{R}$ et $z = (x, y)$ un point critique de f. On note α et β les valeurs propres de la matrice symétrique $H_f(z)$. Dans ce cas, dans une certaine base, autour de z, f ressemble à

$$\alpha \frac{X^2}{2} + \beta \frac{Y^2}{2}.$$

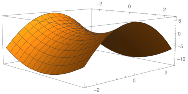
- Si $\alpha > 0$ et $\beta > 0$, f ressemble à :
z est un point de minimum local.



- Si $\alpha < 0$ et $\beta < 0$, f ressemble à :
z est un point de maximum local.



- Si $\alpha > 0$ et $\beta < 0$ (ou l'inverse), f ressemble à :
z est un point point selle.

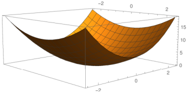


Application du théorème spectral en dimension 2

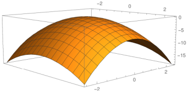
Soit $f : \mathbb{R}^2 \to \mathbb{R}$ et $z = (x, y)$ un point critique de f. On note α et β les valeurs propres de la matrice symétrique $H_f(z)$. Dans ce cas, dans une certaine base, autour de z, f ressemble à

$$\alpha \frac{X^2}{2} + \beta \frac{Y^2}{2}.$$

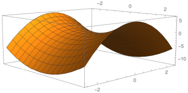
- Si $\alpha > 0$ et $\beta > 0$, f ressemble à :
z est un point de minimum local.



- Si $\alpha < 0$ et $\beta < 0$, f ressemble à :
z est un point de maximum local.



- Si $\alpha > 0$ et $\beta < 0$ (ou l'inverse), f ressemble à :
z est un point point selle.



- Si α ou β est nulle, alors on ne peut rien dire.

Optimisation sous contraintes d'égalité

But du jeu

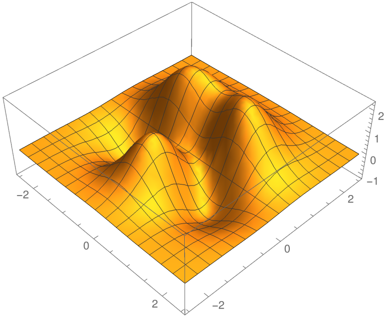
Cette fois on se donne deux fonctions f et g de $\mathbb{R}^2$ dans $\mathbb{R}$ et un réel c. Le but est d'optimiser (*i.e.* de trouver les points de maximum et de minimum) de f sur l'ensemble

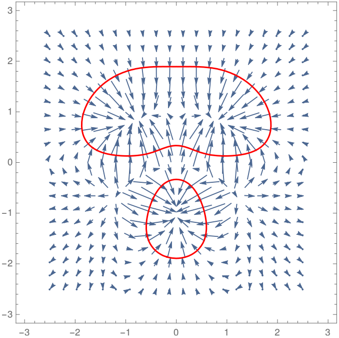
$$\{(x, y) \in \mathbb{R}^2 \text{ tels que } g(x, y) = c\}.$$

Ensemble satisfaisant la contrainte

Si tout se passe bien, cet ensemble est une (ou plusieurs) courbe(s), partout orthogonale au gradient de g.

Si tout se passe bien, cet ensemble est une (ou plusieurs) courbe(s), partout orthogonale au gradient de g.





Équations d'optimalité

Si $z = (x, y)$ est un point d'optimalité sous contraintes, alors $\nabla f(z)$ doit être orthogonal à la courbe des contraintes. Comme c'est aussi le cas pour $\nabla g(z)$, **les gradients de $\boldsymbol{f}$ et $\boldsymbol{g}$ en $\boldsymbol{z}$ sont colinéaires**. En d'autres termes, il existe $\lambda \in \mathbb{R}$ tel que

$$\nabla f(z) = \lambda \nabla g(z).$$

Équations d'optimalité

Si $z = (x, y)$ est un point d'optimalité sous contraintes, alors $\nabla f(z)$ doit être orthogonal à la courbe des contraintes. Comme c'est aussi le cas pour $\nabla g(z)$, **les gradients de $\boldsymbol{f}$ et $\boldsymbol{g}$ en $\boldsymbol{z}$ sont colinéaires**. En d'autres termes, il existe $\lambda \in \mathbb{R}$ tel que

$$\nabla f(z) = \lambda \nabla g(z).$$

Le réel λ s'appelle le **multiplicateur de Lagrange** associé à la contrainte $g = c$.

Équations d'optimalité

(λ, x, y) est donc solution du système de trois équations à trois inconnues

$$\begin{cases} g(x,y) = c, \\ \dfrac{\partial f}{\partial x}(x,y) = \lambda \dfrac{\partial g}{\partial x}(x,y), \\ \dfrac{\partial f}{\partial y}(x,y) = \lambda \dfrac{\partial g}{\partial y}(x,y). \end{cases}$$

Quand on cherche des optima sous contrainte, on cherche donc à résoudre ce système. Les (x, y) pour lesquels il existe λ tel que (λ, x, y) soit solution sont candidats pour être des points d'optimum sous contrainte.

Formulation Lagrangienne

On définit de la façon suivante le **Langrangien** associé à notre problème d'optimisation. C'est une fonction des trois variables (λ, x, y), et pour tout (λ, x, y) :

$$L(\lambda, x, y) := f(x, y) + \lambda(c - g(x, y)).$$

Remarquons que

$$\begin{cases} \dfrac{\partial L}{\partial \lambda}(\lambda, x, y) = c - g(x, y), \\ \dfrac{\partial L}{\partial x}(\lambda, x, y) = \dfrac{\partial f}{\partial x}(x, y) - \lambda \dfrac{\partial g}{\partial x}(x, y), \\ \dfrac{\partial L}{\partial y}(\lambda, x, y) = \dfrac{\partial f}{\partial y}(x, y) - \lambda \dfrac{\partial g}{\partial y}(x, y). \end{cases}$$

Les solutions des équations d'optimalité sont donc des points critiques de L.

Optimisation sous contraintes d'inégalité

But du jeu

On se donne de nouveau deux fonctions f et g de $\mathbb{R}^2$ dans $\mathbb{R}$ et un réel c. Le but est de **maximiser** de f sur l'ensemble

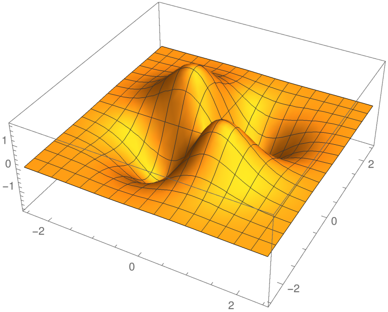
$$\{(x, y) \in \mathbb{R}^2 \text{ tels que } g(x, y) \leq c\}.$$

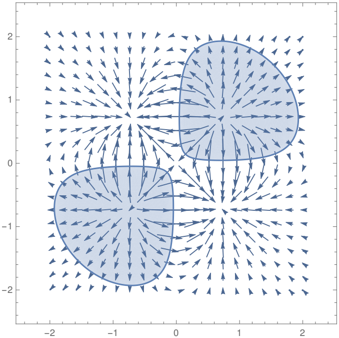
Ensemble satisfaisant la contrainte

Cet ensemble se décompose en son **intérieur** (où $g < c$) et son **bord** (où $g = c$).

Ensemble satisfaisant la contrainte

Cet ensemble se décompose en son **intérieur** (où $g < c$) et son **bord** (où $g = c$).





Remarque : le gradient est non seulement orthogonal au courbes où $g = c$, mais en plus il pointe toujours vers l'extérieur du domaine. C'est normal car g doit passer de valeurs inférieures à c à l'intérieur du domaine à des valeurs supérieures à c à l'extérieur.

Équations de maximalité

Soit (x, y) un point de maximum sous contrainte. Si $g(x, y) < c$, alors on doit avoir

$$\nabla f(x, y) = 0.$$

Équations de maximalité

Soit (x, y) un point de maximum sous contrainte. Si $g(x, y) < c$, alors on doit avoir

$$\nabla f(x, y) = 0.$$

Si $g(x, y) = c$ alors (x, y) est un point d'optimum sous contrainte d'égalité, et alors il existe $\lambda \in \mathbb{R}$ tel que

$$\nabla f(x, y) = \lambda \nabla g(x, y).$$

Équations de maximalité

Soit (x, y) un point de maximum sous contrainte. Si $g(x, y) < c$, alors on doit avoir

$$\nabla f(x, y) = 0.$$

Si $g(x, y) = c$ alors (x, y) est un point d'optimum sous contrainte d'égalité, et alors il existe $\lambda \in \mathbb{R}$ tel que

$$\nabla f(x, y) = \lambda \nabla g(x, y).$$

En fait, on se rend compte que dans ce cas là l'on doit avoir $\lambda \geq 0$: en effet, il faut que $\nabla f(z)$ pointe vers l'extérieur du domaine, soit dans le même sens que $\nabla g(z)$. Sinon, en suivant $\nabla f(z)$ à partir de z, on trouverait des plus grandes valeurs de f tout en restant dans l'ensemble satisfaisant les contraintes.

Équations de maximalité

Soit (x,y) un point de maximum sous contrainte. Si $g(x,y) < c$, alors on doit avoir

$$\nabla f(x,y) = 0.$$

Si $g(x,y) = c$ alors (x,y) est un point d'optimum sous contrainte d'égalité, et alors il existe $\lambda \in \mathbb{R}$ tel que

$$\nabla f(x,y) = \lambda \nabla g(x,y).$$

En fait, on se rend compte que dans ce cas là l'on doit avoir $\lambda \geq 0$: en effet, il faut que $\nabla f(z)$ pointe vers l'extérieur du domaine, soit dans le même sens que $\nabla g(z)$. Sinon, en suivant $\nabla f(z)$ à partir de z, on trouverait des plus grandes valeurs de f tout en restant dans l'ensemble satisfaisant les contraintes. (Ce serait $\lambda \leq 0$ si on cherchait des points de minimum sous contrainte.)

Équations de maximalité

En joignant ces deux observations, on obtient le système de trois équations à trois inconnues

$$\begin{cases} \lambda(c - g(x,y)) = 0, \\ \dfrac{\partial f}{\partial x}(x,y) = \lambda \dfrac{\partial g}{\partial x}(x,y), \\ \dfrac{\partial f}{\partial y}(x,y) = \lambda \dfrac{\partial g}{\partial y}(x,y). \end{cases}$$

Équations de maximalité

En joignant ces deux observations, on obtient le système de trois équations à trois inconnues

$$\begin{cases} \lambda(c - g(x,y)) = 0, \\ \dfrac{\partial f}{\partial x}(x,y) = \lambda \dfrac{\partial g}{\partial x}(x,y), \\ \dfrac{\partial f}{\partial y}(x,y) = \lambda \dfrac{\partial g}{\partial y}(x,y). \end{cases}$$

auquel on ajoute les deux inéquations

$$\begin{cases} g(x,y) \leq c, \\ \lambda \geq 0. \end{cases}$$

Équations de maximalité

En joignant ces deux observations, on obtient le système de trois équations à trois inconnues

$$\begin{cases} \lambda(c - g(x,y)) = 0, \\ \dfrac{\partial f}{\partial x}(x,y) = \lambda \dfrac{\partial g}{\partial x}(x,y), \\ \dfrac{\partial f}{\partial y}(x,y) = \lambda \dfrac{\partial g}{\partial y}(x,y). \end{cases}$$

auquel on ajoute les deux inéquations

$$\begin{cases} g(x,y) \leq c, \\ \lambda \geq 0. \end{cases}$$

Quand on cherche des optima sous contrainte, on cherche donc à résoudre ce système. Les (x,y) pour lesquels il existe λ tel que (λ, x, y) soit solution sont candidats pour être des points d'optimum sous contrainte.

Formulation Lagrangienne

En formant le même **Lagrangien** que précédemment, pour tout (λ, x, y) :

$$L(\lambda, x, y) := f(x, y) + \lambda(c - g(x, y)),$$

le système devient :

$$\begin{cases} g(x, y) \leq c \text{ et } \lambda \geq 0, \\ \lambda \dfrac{\partial L}{\partial \lambda}(\lambda, x, y) = 0, \\ \dfrac{\partial L}{\partial x}(\lambda, x, y) = 0, \\ \dfrac{\partial L}{\partial y}(\lambda, x, y) = 0. \end{cases}$$