1.0.3 Extrema de f sur un compact $K \subset \mathbb{R}^2$

La marche à suivre pour étudier les extrema d'une fonction différentiable sur un compact de $\mathbb{R}^2$ est la suivante.

Soit f une fonction différentiable définie sur un compact K de $\mathbb{R}^2$. Comme f est différentiable, elle est continue. Elle est donc bornée sur K et atteint ses bornes.

En pratique, la fonction f sera donnée par une formule valable sur un certain ouvert sous-ensemble de $\mathbb{R}^2$ et le compact K sera inclus dans cet ensemble de définition.

On mène l'étude des extrema de f en plusieurs étapes. La première est d'étudier l'existence d'extrema locaux de f à l'intérieur de K. C'est pour cette étude qu'on utilisera le développement de Taylor à l'ordre 2 donné ci-dessus.

Mais cette étude n'est pas suffisante. Il faut aussi regarder ce qui se passe sur le bord de K. Pour cela on procède autrement. Comme par exemple dansce qui suit

1.0.23 EXEMPLE. Etude de $f(x,\ y) = x^2 - y^2$ sur $K = \{(x,\ y) \in \mathbb{R}^2 / x^2 + y^2 \leq 1\}$. On procède de la manière suivante:

(i) On cherche les points critiques et les extrema locaux dans Int(K).

On trouve un seul point stationnaire en $(0,0)$. Mais en $(0,0) f$ a un point selle. La fonction n'a donc pas d'extremum à l'intérieur de K. Mais comme K est compact et f est continue sur K, f est bornée sur K et atteint ses bornes sur K. Ce sera donc sur le bord de K.

(ii) On analyse f sur le bord $\partial K = K \setminus Int(K)$.

Une possibilité ici est de paramétrer le bord de K : le cercle de rayon 1 centré en $(0,0)$, dont une paramétrisation est donnée par $x = \cos t$, $y = \sin t$, $t \in [0, 2\pi]$.
On obtient: $g(t) = f(\cos t,\ \sin t) = \cos^2 t - \sin^2 t = \cos(2t)$. On peut alors étudier les variations de la fonction g sur l'intervalle $[0, 2\pi]$. On obtient qu'elle atteint son maximum qui vaut 1, aux points $t = 0, \pi$ et 2π et son minimum qui vaut -1 aux points $t = \frac{\pi}{2}$ et $\frac{2\pi}{3}$.

La fonction f atteint donc sur K son maximum 1 aux points $(1,\ 0)$ et $(-1, 0)$ et son minimum -1 aux points $(0, 1)$ et $(0,\ -1)$.

1.0.4 Extrema liés (multiplicateur de Lagrange)

Une seule contrainte

Soit $n \geq 1$ un entier. On notera $X = (x_1, \ldots, x_n)$ un point de $\mathbb{R}^n$.

Il s'agit de trouver les extrema de $f(X)$ lorsque X appartient à une hypersurface S définie par $g(X) = c$ i.e. $S = \{X \in \mathbb{R}^n \mid g(X) = c\}$.

1.0.24 Exercice Maximiser $x^2 y^2 z^2$ lorsque $x^2 + y^2 + z^2 = 1$.

1.0.25 DÉFINITION
Un point X_0 *est un minimum* (*resp. maximum*) local pour f, lié à la contrainte $g(X) = c$ (i.e. sur S) si:

(i) $g(X_0) = c$

(ii) Il existe $r > 0$ tel que $f(X_0) \leq f(X)(resp. f(X_0) \geq f(X))$ pour tout $X \in S \cap B(X_0,\ r)$.

1.0.26 THÉORÈME (DE LAGRANGE)

Soit $f(X)$ et $g(X)$ de classe C^1 sur un ouvert de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$.
Soit X_0 un point de S tel que:

(a) $\nabla g(X_0) \neq \mathbf{0}$ (i.e. X_0 n'est pas un point critique de g)

(b) f admet un *extremum local* en X_0 sur S

Alors il existe un nombre réel λ tel que:

$$\nabla f(X_0) = \lambda \nabla g(X_0).$$

Le nombre réel λ est appelé multiplicateur de Lagrange.

Démonstration: On va démontrer le théorème pour la dimension $n = 3$. La démonstration dans le cas général, se fait en suivant le même schema.
Soit $X_0 = (x_0, y_0, z_0)$ vérifiant les conditions (a) et (b).
D'après (a), on peut appliquer le théorème des fonctions implicites.
En effet, $\nabla g(X_0) \neq (0,0,0)$, entraîne que l'une des composantes de $\nabla g(X_0)$ est non nulle; on va supposer sans perdre de généralité que $\dfrac{\partial g}{\partial z}(X_0) \neq 0$. Alors d'après le théorème des fonctions implicites, il existe une fonction φ définie dans un voisinage de (x_0, y_0), telle que $z = \varphi(x, y)$. On aura $(x, y, z) \in S \iff g(x, y, \varphi(x, y)) = c$ au voisinage de X_0.
Comme la fonction $h(x, y) = f(x, y, \varphi(x, y))$ admet un extremum en (x_0, y_0), on a $\nabla h(x_0, y_0) \neq 0$, en utilisant la formule de dérivation des fonctions composées, on aura au point X_0

$$\frac{\partial h}{\partial x}(x_0, y_0) = \frac{\partial f}{\partial x}(X_0) + \frac{\partial \varphi}{\partial x}(x_0, y_0).\frac{\partial f}{\partial z}(X_0) = 0$$
$$\frac{\partial h}{\partial y}(x_0, y_0) = \frac{\partial f}{\partial y}(X_0) + \frac{\partial \varphi}{\partial y}(x_0, y_0).\frac{\partial f}{\partial z}(X_0) = 0$$

De même les dérivées partielles de la relation $g(x, y, \varphi(x, y)) = c$ donnent

$$\frac{\partial g}{\partial x}(X_0) + \frac{\partial \varphi}{\partial x}(x_0, y_0).\frac{\partial g}{\partial z}(X_0) = 0$$
$$\frac{\partial g}{\partial y}(X_0) + \frac{\partial \varphi}{\partial y}(x_0, y_0).\frac{\partial g}{\partial z}(X_0) = 0$$

Comme $\dfrac{\partial g}{\partial z}(X_0) \neq 0$, on peut diviser par cette quantité pour obtenir

$$\frac{\partial \varphi}{\partial x}(x_0, y_0) = -\frac{\frac{\partial g}{\partial x}(X_0)}{\frac{\partial g}{\partial z}(X_0)}, \quad \frac{\partial \varphi}{\partial y}(x_0, y_0) = -\frac{\frac{\partial g}{\partial y}(X_0)}{\frac{\partial g}{\partial z}(X_0)}.$$

D'où

$$\frac{\partial f}{\partial x}(X_0) = \left(\frac{\frac{\partial g}{\partial x}(X_0)}{\frac{\partial g}{\partial z}(X_0)}\right)\frac{\partial f}{\partial z}(X_0) = \left(\frac{\frac{\partial f}{\partial z}(X_0)}{\frac{\partial g}{\partial z}(X_0)}\right)\frac{\partial g}{\partial x}(X_0)$$

$$\frac{\partial f}{\partial y}(X_0) = \left(\frac{\frac{\partial g}{\partial x}(X_0)}{\frac{\partial g}{\partial z}(X_0)}\right)\frac{\partial f}{\partial z}(X_0) = \left(\frac{\frac{\partial f}{\partial z}(X_0)}{\frac{\partial g}{\partial z}(X_0)}\right)\frac{\partial g}{\partial y}(X_0)$$

et $\dfrac{\partial f}{\partial z}(X_0) = \left(\dfrac{\frac{\partial f}{\partial z}(X_0)}{\frac{\partial g}{\partial z}(X_0)}\right)\dfrac{\partial g}{\partial z}(X_0)$.

Ainsi, on a bien $\nabla f(X_0) = \lambda \nabla g(X_0)$ avec $\lambda = \frac{\frac{\partial f}{\partial z}(X_0)}{\frac{\partial g}{\partial z}(X_0)}$. $\square$

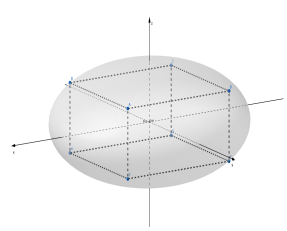
1.0.28 REMARQUE
Si P est un extremum lié, on a $\nabla f(P)$ parallèle à $\nabla g(P)$. La réciproque n'est pas vraie. Nous avons une condition nécessaire mais pas suffisante. C'est l'équivalent de la nullité de la dérivée pour les extrema libres : en un extremum libre la dérivée est nulle mais la dérivée peut être nulle sans que la fonction ait un extremum (penser à $x \mapsto x^3$ en $x = 0$).

1.0.29 EXEMPLE. Sur l'exemple précédent on montre la méthode de résolution.

Déterminer le maximum de la distance de l'ellipsoide donné par $\frac{x^2}{2} + \frac{y^2}{2} + \frac{z^2}{2} = 1$ à l'origine.

1.0.30 EXEMPLE. On se propose de déterminer le parallélépipède ayant le plus grand volume, tel que ses côtés a, b et c sont parallèles aux axes Ox, Oy et Oz respectivement et qu'on peut inscrire à l'intérieur de l'ellipsoïde d'équation

$$x^2 + 2y^2 + 9z^2 = 1.$$



On pose $a = 2x$, $b = 2y$ et $c = 2z$ avec $(x,y,z) \in ([0,+\infty[)^3$. On veut alors maximiser la fonction "volume" $f(x,y,z) = abc = (2x)(2y)(2z) = 8xyz$ sur la partie de l'ellipsoïde $S = \{(x,y,z) \in ([0,+\infty[)^3 \,|\, x^2+2y^2+9z^2 = 1\}$. On va utiliser pour cela la méthode des multiplicateurs de Lagrange.

On va résoudre alors le système $\begin{cases} \nabla f(X_0) = (8yz, 8xz, 8xy) = \lambda \nabla g(X_0) = \lambda(2x, 8y, 19z) \\ x^2+2y^2+9z^2 = 1. \end{cases}$

Considérant les deux cas, $\lambda = 0$ ou $\lambda \neq 0$. Si $\lambda = 0$ alors au moins deux des coordonnées x, y et z sont nulles, d'où le volume est nulle, c'est un minimum, mais pas un maximum. On doit donc supposer $\lambda \neq 0$ et toutes les coordonnées non nulles. On peut donc considérer les quotients; on a :

$$\frac{8yz}{8xz} = \frac{y}{x} = \frac{2\lambda x}{8\lambda y} = \frac{x}{4y}, \quad \frac{8yz}{8xy} = \frac{z}{x} = \frac{2\lambda x}{18\lambda z} = \frac{x}{9z}$$

D'où $4y^2 = 9z^2 = x^2$ et la contrainte $x^2+2y^2+9z^2 = 1$, donne alors $x^2+x^2+x^2 = 1$ i.e. $3x^2 = 1$. Comme x est positif, l'unique solution est $x = \frac{1}{\sqrt{3}}$, par suite $y = \frac{1}{2\sqrt{3}}$ et $z = \frac{1}{3\sqrt{3}}$.

Ainsi, le parallélépide est de côtes, $a = \frac{2}{\sqrt{3}}$, $b = \frac{2}{2\sqrt{3}} = \frac{1}{\sqrt{3}}$ et $c = \frac{2}{3\sqrt{3}}$

et le volume maximal est $f(x,y,z) = abc = \frac{2}{\sqrt{3}}.\frac{1}{\sqrt{3}}.\frac{2}{3\sqrt{3}} = \frac{4}{9\sqrt{3}}$.

Généralisation à plusieurs contraintes:

On cherche les extrema d'une fonction f sur l'ensemble S défini par $g_1 = g_2 = \ldots = g_k = 0$, toutes les fonctions considérées étant de classe C^1. Pour que les choses marchent bien il faut faire l'hypothèse suivante : en tout point de S les gradients des fonctions g_i sont linéairement indépendants.

1.0.31 Théorème (multiplicateurs de Lagrange)

Soient f *et* $g_1, \ldots, g_k$ $k+1$ fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$ de classe C^1 telles que les vecteurs $\nabla g_1, \ldots, \nabla g_k$, *soit linéairement indépendants* sur l'ensemble S défini par

$$g_1 = \ldots = g_k = 0.$$

Alors si f *admet un* extrema lié *sur* S *en* X_0 le vecteur, $\nabla f(X_0)$ *est combinaison linéaire des vecteurs* $\nabla g_i(x_0)$ i.e. il existe des nombres réels $\lambda_1, \ldots, \lambda_k$ tels que

$$\nabla f(X_0) = \sum_{i=1}^{k} \lambda_i \nabla g_i(X_0).$$

Les nombres réels $\lambda_1, \ldots, \lambda_k$ sont appelés *des multiplicateurs de Lagrange*.