4.4 Conditions d'optimalité dans le cas général.

4.4.1 Cône des directions admissibles.

Soit $U \subset \mathbb{R}^n$ une partie fermée de $\mathbb{R}^n$.

Définition: $C_U(x)$

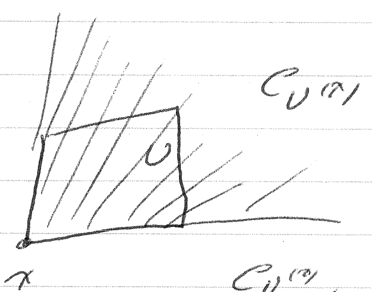
Fixons un pt. $x \in U$.

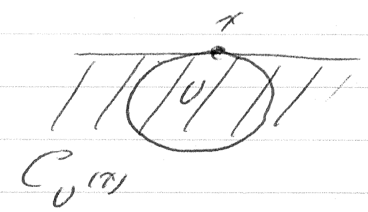
$h \in C_U(x)$ s'il existe une courbe $x(t) \in U$ définie pour des valeurs positives de $t \geq 0$ et $x(0) = x$ de sorte que

$$\frac{dx}{dt}(0) = \lim_{\substack{t \to 0 \\ t > 0}} \frac{x(t) - x}{t} = h$$

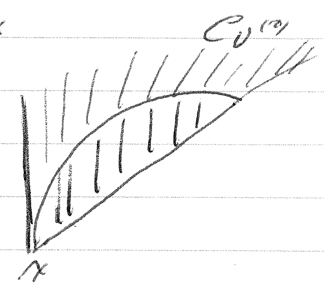
Exemples:

1)

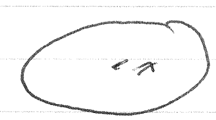




3)



4)



$C_U(x) = \mathbb{R}^n$

Propriétés de $C_U(x)$:

- $C_U(x)$ est un cône de sommet 0, fermé.
- Si U est convexe, on a que

$$C_U(x) = \overline{\{\lambda(y - x) \mid y \in U, \lambda > 0\}}$$

Voici la condition d'optimalité dans le cas général.

Soit $f : U \to \mathbb{R}$ une fonction numérique de classe C^1 définie sur un fermé de $\mathbb{R}^n$.

Si a est un pt. minimisant de f,

$$(\nabla f(a), h) \geq 0 \qquad \forall h \in C_U(a).$$

Preuve: $h \in C_U(a) \Rightarrow \exists\, x(t) \in U$ t.q.

$$x(0) = a$$

$$\frac{dx}{dt}(0) = h$$

On a que: $f(x(t)) \geq f(a)$

$$f(x(t)) - f(a) \geq 0$$

$$\Rightarrow \frac{f(x(t)) - f(a)}{t} \geq 0 \qquad \forall t > 0$$

$$\Rightarrow \lim_{\substack{t \to 0 \\ t > 0}} \frac{f(x(t)) - f(a)}{t} \geq 0$$

$= (\nabla f(a), h) \geq 0$ à cause de la règle de dérivation des fcts. composées.

Donc

$$(\nabla f(a), h) \geq 0 \qquad \forall h \in C_U(a)$$

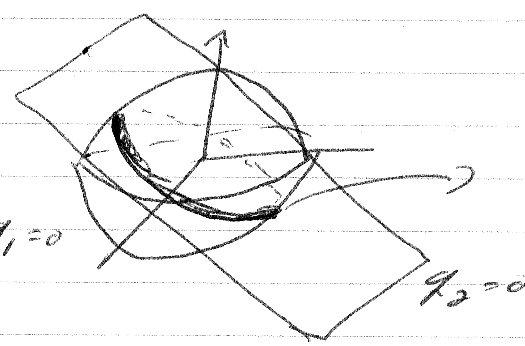
$$\Leftrightarrow \nabla f(a) \in (C_U(a))^*$$

4.4.2 Contraintes de type égalité

Soient $g_1, g_2, \ldots, g_m$ m fonctions numériques de classe C^1.

On pose
$$U = \{ x \in \mathbb{R}^n \mid g_1(x) = g_2(x) = \cdots = g_m(x) = 0 \}$$

Exemple:
$$g_1(x) = x_1^2 + x_2^2 + x_3^2 - 1$$
$$g_2(x) = x_1 + x_2 + x_3$$



U = courbe d'intersection
= cercle non planaire de rayon 1 et centre $(0,0,0)$

En général, U n'est pas convexe mais toujours fermé!

Il est pratique d'introduire la fonction
$$g : \mathbb{R}^n \to \mathbb{R}^m$$
$$x \mapsto (g_1(x), g_2(x), \ldots, g_m(x))$$

$$\Rightarrow U = \{ x \in \mathbb{R}^n \mid g(x) = (0, 0, \ldots, 0) = 0 \}$$

Définition: matrice Jacobienne en un pt x donné
$$\left(Dg(x)\right)_{ij} = \left(\frac{\partial g_i}{\partial x_j}(x)\right)_{m \times n}$$

Exemple : voir l'ex. précédent

$$Dg(x_1, x_2, x_3) = \begin{pmatrix} 2x_1 & 2x_2 & 2x_3 \\ 1 & 1 & 1 \end{pmatrix}_{2\times 3}$$

Retournons au calcul de $C_U(x)$ avec $x \in U$.

$$h \in C_U(x) \Rightarrow \begin{array}{l} g_1(x(t)) = 0 \\ g_2(x(t)) = 0 \\ \vdots \\ g_m(x(t)) = 0 \end{array}$$

$$\Rightarrow \frac{d}{dt} g_i(x(t)) = 0$$

$$\parallel$$

$$\left(Dg_i(x(t)), \frac{dx}{dt}\right) = 0$$

En $t=0$

$$\Rightarrow \left(Dg_i(x), h\right) = 0 \qquad \forall i = 1, \dots m$$

Ceci signifie que :

$$h \in \operatorname{Ker} Dg(x)$$

On acceptera, sans démonstration (consulter le livre de M. Delfour) que

$$C_U(x) = \operatorname{Ker} Dg(x)$$

si x est un pt. régulier de g, i.e.

$$rg(Dg(x)) = m$$

Remarque : x est un pt. régulier de g

ssi

$\nabla g_1(x), \dots, \nabla g_m(x)$ sont linéairement indépendants.

Exemple : toujours pour l'exemple précédent

$x \in U$ n'est pas régulier $\Leftrightarrow$ $x_1 = x_2 = x_3$

Or $x_1 = x_2 = x_3 \Rightarrow x_1 = 0 = x_2 = x_3$

$\Rightarrow x \notin U$

donc, tous les pts de U sont réguliers.

D'un pt. de vue géométrique, cela signifie qu'il y a une droite tangente en tout pt x de U.

Maintenant, caractérisons le pt. minimisant du problème

$$\min_{g(x)=0} f(x)$$

Faisons l'hypothèse que le pt. minimisant $a \in U$ est régulier.

$$\Rightarrow C_U(a) = \operatorname{Ker} Dg(a).$$

La condition d'optimalité s'écrit :

$$(\nabla f(a), h) \geq 0 \quad \forall h \in \operatorname{Ker} Dg(a)$$

$$\Rightarrow (\nabla f(a), h) = 0 \quad \forall h \in \operatorname{Ker} Dg(a)$$

car le noyau est un sous-espace vectoriel de $\mathbb{R}^n$.

$$\Rightarrow \nabla f(a) \in \left[\operatorname{Ker} Dg(a)\right]^{\perp}$$

$$\parallel$$

$$\nabla f(a) \in \operatorname{Im} \left(Dg(a)\right)^{t}$$

$\Leftrightarrow$ il existe $\lambda \in \mathbb{R}^m$ de sorte que

$$\nabla f(a) + \sum_{i=1}^{m} \lambda_i \nabla g_i(a) = 0$$

En résumé, le pt $a \in U$ doit être solution du système

$$(*) \begin{cases} \nabla f(x) + \displaystyle\sum_{i=1}^{m} \lambda_i \nabla g_i(x) = 0 \\ g_i(x) = 0 & \forall i = 1, \dots m \end{cases}$$

Ceci constitue un système non linéaire de $n+m$ équations à $n+m$ inconnues (x, λ).

Il est coutume d'introduire la fonction dite Lagrangienne

$$L(x, \lambda) = f(x) + (\lambda, g(x))$$

La condition d'optimalité $(*)$ est équivalente à

$$\nabla L(x, \lambda) = 0$$

$$\Leftrightarrow \begin{cases} D_x L = 0 \\ \nabla_\lambda L = 0 \end{cases}$$

C'est la méthode des multiplicateurs de Lagrange.

Remarque :

Analysons le problème dans $\mathbb{R}^3$: $n = 3$ et $m = 1$

$$\min f(x, y, z)$$
$$g(x, y, z) = 0$$

Si on peut éliminer z au profit de x et y

$$\Rightarrow \quad z = z(x, y)$$

on obtient le problème sans contrainte

$$\min_{(x,y)} f(x, y, z(x, y))$$

Grâce au théorème des fonctions implicites, il est possible d'exprimer

$$z = z(x, y)$$

si

$$g_z \left(= \frac{\partial g}{\partial z} \right) \neq 0$$

$$\Rightarrow \quad z_x = -g_x / g_z \quad \text{et} \quad z_y = -g_y / g_z$$

Dans ce cas, la condition d'optimalité du problème

$$\min f(x, y, z(x, y)) = h(x, y)$$

s'écrit.

$$\begin{cases} h_x = f_x + f_z z_x = 0 \\ h_y = f_y + f_z z_y = 0 \end{cases}$$

$$\Rightarrow \begin{cases} f_x + f_z \left(\frac{-g_x}{g_z} \right) = 0 \\ f_y + f_z \left(-g_y / g_z \right) = 0 \end{cases}$$

En posant $\lambda = -\dfrac{f_z}{g_z}$

Au pt. minimisant P

$\Rightarrow$
$$\begin{cases} f_x(P) + \lambda g_x(P) = 0 \\ f_y(P) + \lambda g_y(P) = 0 \\ f_z(P) + \lambda g_z(P) = 0 \quad (\text{par déf. de } \lambda) \end{cases}$$

$$g(\underbrace{x, y, z}_{P}) = 0$$

Donc résoudre le problème sans contrainte ou le problème contraint (sans éliminer z) est équivalent.

Toutefois, il est important de noter qu'il est, en général, impossible d'éliminer une variable au profit des autres surtout, en présence de plusieurs contraintes.

Exemples:

1. Projection du pt $(3,1,-1)$ sur la sphère

$$\min_{x_1^2+x_2^2+x_3^2=4} \frac{(x_1-3)^2 + (x_2-1)^2 + (x_3+1)^2}{2}$$

$$L(x_1, x_2, x_3, \lambda) = f(x) + \lambda\left(\frac{x_1^2+x_2^2+x_3^2-4}{2}\right)$$

$$\nabla L = 0 \iff \begin{cases} x_1 - 3 = \lambda x_1 \\ x_2 - 1 = \lambda x_2 \\ x_3 + 1 = \lambda x_3 \\ x_1^2 + x_2^2 + x_3^2 = 4 \end{cases}$$

$$\Rightarrow \quad x_1 = \frac{3}{1-\lambda}$$

$$x_2 = \frac{1}{1-\lambda} \qquad \lambda \neq 1$$

$$x_3 = \frac{-1}{1-\lambda}$$

Si $\lambda = 1 \Rightarrow x_1 - 3 = x_1 \Rightarrow$ impossible

$\Downarrow$

$\lambda \neq 1$

$$\Rightarrow \left(\frac{3}{1-\lambda}\right)^2 + \left(\frac{1}{1-\lambda}\right)^2 + \left(\frac{-1}{1-\lambda}\right)^2 = 4$$

$$\Rightarrow \lambda = 1 \pm \frac{\sqrt{11}}{2}$$

Le pt. le plus près correspond à:

$$\left(6/\sqrt{11},\ 2/\sqrt{11},\ -2/\sqrt{11} \right)$$

Exemple 2: $\quad m = 2 \quad$ 2 contraintes

$$\max \; x + 2y + 3z$$
$$x - y + z = 1$$
$$x^2 + y^2 = 1$$

Remarque: en changeant f par $-f$, il est clair qu'un pt. maximisant vérifie le même système

$$\begin{cases} D_x L = 0 \\ D_\lambda L = 0 \end{cases}$$

Le Lagrangien s'écrit:

$$L(x, y, z, \lambda_1, \lambda_2) = x + 2y + 3z + \lambda_1 (x - y + z - 1) + \lambda_2 (x^2 + y^2 - 1)$$

Calcul de $\nabla L = 0$:

$$\begin{cases} 1 + \lambda_1 + 2x\lambda_2 = 0 \\ 2 + -\lambda_1 + 2y\lambda_2 = 0 \\ 3 + \lambda_1 = 0 \\ x - y + z = 1 \\ x^2 + y^2 = 1 \end{cases}$$

$\Rightarrow \lambda_1 = -3 \Rightarrow 2x\lambda_2 = 2 \Rightarrow x = \dfrac{1}{\lambda_2}$

$\Rightarrow 2y\lambda_2 = -5 \Rightarrow y = \dfrac{-5}{2\lambda_2}$

$\Rightarrow x^2 + y^2 = \dfrac{1}{\lambda_2^2} + \dfrac{25}{4\lambda_2^2} = 1$

$\Rightarrow \lambda_2^2 = \dfrac{29}{4} \Rightarrow \lambda_2 = \pm\dfrac{\sqrt{29}}{2}$

ⓐ Si $\lambda_2 = \dfrac{\sqrt{29}}{2} \Rightarrow x = \dfrac{2}{\sqrt{29}}$, $y = \dfrac{-5}{\sqrt{29}}$, $z = 1 - \dfrac{7}{\sqrt{29}}$

ⓑ Si $\lambda_2 = \dfrac{-\sqrt{29}}{2} \Rightarrow x = -\dfrac{2}{\sqrt{29}}$, $y = \dfrac{5}{\sqrt{29}}$, $z = 1 + \dfrac{7}{\sqrt{29}}$

$z = 1 - x + y$

ⓐ $f = 3 - \sqrt{29} \rightarrow$ min

ⓑ $f = 3 + \sqrt{29} \rightarrow$ max

Exemple 3. $\max\limits_{\sum_{i=1}^{N} x_i^2=1} \sum_{i=1}^{N} x_i$ dans $\mathbb{R}^N$

$$f(x) = \sum_{i=1}^{N} x_i$$

$$g(x) = \sum x_i^2 - 1 \qquad m = 1$$

$$\nabla f = (1, 1, \ldots, 1)$$

$$\nabla g = 2(x_1, x_2, \ldots, x_N)$$

La cond. d'opt. s'écrit : $\begin{cases} \nabla f + \lambda \nabla g = 0 \\ g = 0 \end{cases}$

$$\Leftrightarrow \begin{cases} 1 + 2\lambda x_i = 0 \\ \sum x_i^2 = 1 \end{cases} \Rightarrow x_i = -\frac{1}{2\lambda} \quad \lambda \neq 0$$

($\lambda = 0$ impossible)

$$\Downarrow$$

$$\sum_i \frac{1}{4\lambda^2} = 1$$

$$\frac{N}{4\lambda^2} = 1 \Rightarrow \lambda = \pm \frac{\sqrt{N}}{2}$$

(a) $\lambda = \frac{\sqrt{N}}{2} \Rightarrow x_i = \frac{-1}{\sqrt{N}} \Rightarrow f = \frac{-N}{\sqrt{N}} = -\sqrt{N}$

min.

(b) $\lambda = -\frac{\sqrt{N}}{2} \Rightarrow x_i = \frac{1}{\sqrt{N}} \Rightarrow f = \sqrt{N}$ max

Valeur max : $f = \sqrt{N}$ atteint en $x_i = \frac{1}{\sqrt{N}} \quad \forall i$

Exemple 4 :

$$\min_{\|x\|=1} \frac{1}{2}(Ax, x) \qquad \text{dans } \mathbb{R}^n, \quad A \text{ symétrique}$$

$$f(x) = \frac{1}{2}(Ax, x)$$

$$g(x) = \sum_{i=1}^{n} x_i^2 - 1$$

$$\nabla f(x) = Ax \qquad \text{car } A \text{ est sym.}$$

$$\nabla g(x) = 2x = 2(x_1, \ldots, x_n)$$

La cond. d'opt. s'écrit :

$$\Leftrightarrow \begin{cases} Ax + 2\lambda x = 0 \\ \|x\| = 1 \end{cases} \Rightarrow Ax = (-2\lambda)x$$

$\Rightarrow -2\lambda$ est une valeur propre de A

Pour $-2\lambda \in sp(A)$ = l'ensemble des valeurs propres de A (qui est réelle)

$\Rightarrow x$ = vecteur propre correspondant avec $\|x\| = 1$

Donc tous les vecteurs propres de A vérifient la cond. d'optimalité

si x = vecteur lié à λ (valeur propre $= -2\lambda$) $\quad \alpha = -2\lambda$

$$\Rightarrow f(x) = \frac{1}{2}(Ax, x) = \frac{-2\lambda}{2}(x, x) = -\lambda = \frac{\alpha}{2}$$

$$\min f = \min_{\alpha \in sp(A)} \alpha/2 \Rightarrow \text{plus petite valeur propre}$$

$$\max f = \max_{\alpha \in sp(A)} \alpha/2 \Rightarrow \text{plus grande}$$

$$\underset{\text{MIN}}{\alpha_1} \leq \alpha_2 \leq \alpha_3 \leq \ldots \leq \underset{\text{MAX}}{\alpha_n}/2 : \text{ valeurs propres de } A$$