# Fiche de théorie 12

## Extrema sous contraintes

Soient $f : \mathbb{R}^n \to \mathbb{R}$ et $g : \mathbb{R}^n \to \mathbb{R}$ deux fonctions.

**Définition 1.** *On dit qu'un point $\boldsymbol{x}_0 \in \mathbb{R}^n$ est un maximum (ou minimum) local d'une fonction sous contrainte $g(\boldsymbol{x}) = 0$, si $f(\boldsymbol{x}_0) \geq f(\boldsymbol{x})$ (ou $f(\boldsymbol{x}_0) \leq f(\boldsymbol{x})$) pour tout $\boldsymbol{x} \approx \boldsymbol{x}_0$ tel que $g(\boldsymbol{x}) = 0$.*

Voici quelques remarques :

- L'opposé de "sous contrainte" est "libre" : *extrema libre* signifie "extrema sans contraintes", c'est-à-dire, sur l'ensemble de $\mathbb{R}^n$.
- Notez la différence avec la notion de maximum/minimum local : là, on doit comparer $f(\boldsymbol{x}_0)$ avec les valeurs de $f$ en tous les points voisins $\boldsymbol{x}$, tandis qu'ici $f(\boldsymbol{x}_0)$ est seulement comparé avec les points $\boldsymbol{x}$, où $g(\boldsymbol{x}) = 0$ est satisfaite.
- En particulier, $\boldsymbol{x}_0$ doit lui-même satisfaire $g(\boldsymbol{x}_0) = 0$.
- Pourquoi zéro dans "$g(\boldsymbol{x}) = 0$" ? Juste une convention ! Toute contrainte de la forme $g(\boldsymbol{x}) = c$ peut être réécrite comme $\tilde{g}(\boldsymbol{x}) = 0$ avec une nouvelle fonction $\tilde{g}(\boldsymbol{x}) = g(\boldsymbol{x}) - c$.
- En d'autres termes, nous étudions les extrema sous contraintes d'égalité. Il existe également une théorie similaire des extrema sous contraintes d'inégalité $g(\boldsymbol{x}) \geq 0$, mais nous ne la traiterons pas dans ce cours.
- Interprétation géométrique : puisque $G = \{\boldsymbol{x} : g(\boldsymbol{x}) = 0\} \subset \mathbb{R}^n$ est une certaine surface, les extrema sous contraintes de $f$ sont juste les extrema libres de $f$ restreints à cette surface. C'est-à-dire, les extrema libres de $f : G \to \mathbb{R}$ au lieu de $f : \mathbb{R}^n \to \mathbb{R}$.

Des exemples naturels de problèmes d'extrema sous contraintes proviennent de la théorie de l'utilité, où nous avons généralement des contraintes budgétaires (la quantité d'argent est limitée).

Un autre exemple naturel : maximiser l'aire d'un rectangle si son périmètre est fixe. Ou maximiser le volume d'une boîte si sa surface est fixe.

Ci-dessous, nous allons discuter deux approches distinctes pour résoudre ce problème :

- **méthode de substitution :** traiter la contrainte $g(\boldsymbol{x}) = 0$ comme une équation, à l'aide de laquelle nous pouvons exprimer certaines variables en fonction des autres et l'injecter dans $f$ ; schématiquement,

$$g(x_1, x_2) = 0 \overset{\text{solve}}{\Longrightarrow} x_2 = h(x_1) \implies \text{optimize new function } \tilde{f}(x_1) = f(x_1, h(x_1)).$$

- **méthode des multiplicateurs de Lagrange :** construire une nouvelle fonction $L$ de $n+1$ variables (en introduisant une variable artificielle), appelée le Lagrangien, de sorte que les extrema sous contraintes de $f$ soient les extrema libres de $L$, et optimiser $L$ comme avant.

## Rappel sur les extrema libres

Rappelez-vous que $\boldsymbol{x}_0$ est un point extremal candidat de $f$ si $\operatorname{grad} f(\boldsymbol{x}_0) = \mathbf{0}$. Pour déterminer si $\boldsymbol{x}_0$ est bien un point extremal, nous devons vérifier le type du différentiel du second ordre $d^2 f(\boldsymbol{x}_0)$ ou du Hessien $H_f(\boldsymbol{x}_0)$.

## Exemple

Dans cette section, nous utilisons la méthode de substitution pour résoudre le problème suivant : maximiser

$$B(q_1, q_2) = 55q_1 + 70q_2 - 2q_1^2 - 3q_1q_2 - 3q_2^2$$

sous contrainte

$$q_1 + q_2 - 13 = 0.$$

Tout d'abord, nous résolvons la contrainte pour l'une des deux variables :

$$q_2 = 13 - q_1,$$

puis nous l'injectons dans $B$ :

$$\begin{aligned}\tilde{B}(q_1) = B(q_1, 13 - q_1) &= 55q_1 + 70(13 - q_1) - 2q_1^2 - 3q_1(13 - q_1) - 3(13 - q_1)^2 \\ &= 403 + 24q_1 - 2q_1^2.\end{aligned}$$

Ensuite, nous maximisons $\tilde{B}$ comme d'habitude :

$$\tilde{B}'(q_1) = 0 \implies 24 - 4q_1 = 0 \implies q_1 = 6.$$

Il reste à trouver $q_2$ :

$$q_2 = 13 - q_1 = 7.$$

Puisque $\tilde{B}'' = -4 < 0$, cette solution est bien un maximum.

## Conditions nécessaires pour un extremum sous contraintes

Le théorème suivant est donné sans preuve :

**Théorème 1.** *Si $\boldsymbol{x}_0$ est un point extremum de $f$ sous contrainte $g(\boldsymbol{x}) = 0$, alors $\operatorname{grad} f$ et $\operatorname{grad} g$ sont codirectionnels ou proportionnels. C'est-à-dire, il existe un nombre $\lambda \in \mathbb{R}$ tel que*

$$\operatorname{grad} f(\boldsymbol{x}_0) = \lambda \operatorname{grad} g(\boldsymbol{x}_0).$$

- $\operatorname{grad} f(\boldsymbol{x}_0) = \lambda \operatorname{grad} g(\boldsymbol{x}_0)$ signifie que les deux vecteurs pointent dans la même direction si $\lambda > 0$ et qu'ils pointent dans les directions opposées si $\lambda < 0$.
- Le théorème précédent n'est qu'une condition nécessaire, pas suffisante.
- La condition du second ordre n'est pas applicable ici.

Le théorème ci-dessus nous dit que pour optimiser $f$ sous contrainte donnée par $g$, nous devons résoudre $n+1$ équations

$$f'_{x_1} = \lambda g'_{x_1}, \quad \ldots, \quad f'_{x_n} = \lambda g'_{x_n}, \quad g(x_1, \ldots, x_n) = 0$$

sur $n+1$ variables

$$x_1, \quad \ldots, \quad x_n, \quad \lambda.$$

# Lagrangien

L'idée de la méthode des multiplicateurs de Lagrange est de réécrire la condition

$$\operatorname{grad} f(\boldsymbol{x}) = \lambda \operatorname{grad} g(\boldsymbol{x})$$

comme suit :

$$\boxed{\operatorname{grad} L(\boldsymbol{x}, \lambda) = 0, \quad \text{where} \quad L(\boldsymbol{x}, \lambda) = f(\boldsymbol{x}) - \lambda\, g(\boldsymbol{x}).}$$

**Définition 2.** *Le Lagrangien du problème d'optimisation pour $f$ avec contrainte donnée par $g$ est une fonction de $n+1$ variables réelles*

$$L(\boldsymbol{x}, \lambda) = f(\boldsymbol{x}) - \lambda g(\boldsymbol{x}).$$

**Remarque 1.** *Lorsque nous écrivons $\operatorname{grad} L(\boldsymbol{x}, \lambda)$, voulons-nous dire le gradient par rapport à $\boldsymbol{x}$, ou par rapport à $\boldsymbol{x}$ et $\lambda$ ? La réponse : ce n'est pas important, car les deux gradients diffèrent d'une (dernière) coordonnée, la dérivée par rapport à $\lambda$, qui donne*

$$\frac{\partial L}{\partial \lambda} = \frac{\partial}{\partial \lambda}(f(\boldsymbol{x}) - \lambda g(\boldsymbol{x})) = -g(\boldsymbol{x}).$$

*Lorsque nous fixons cette dérivée à zéro, nous retrouvons simplement la contrainte :*

$$g(\boldsymbol{x}) = 0.$$

*Donc, si par $\operatorname{grad}$ nous entendons le gradient par rapport à $(\boldsymbol{x}, \lambda)$, alors l'équation $\operatorname{grad} L = \mathbf{0}$ englobe également la contrainte $g(\boldsymbol{x}) = 0$. Sinon, nous devons énoncer la contrainte séparément.*

**Remarque 2.** *Le signe moins devant $g(\boldsymbol{x})$ dans la définition de $L$ est conventionnel. Puisque $\lambda$ est un nombre réel arbitraire, nous le problème avec $+\lambda$ au lieu de $-\lambda$ est exactement le même. Vous pouvez l'écrire comme vous le souhaitez.*

*Une certaine différence apparaît dans les problèmes d'optimisation avec des contraintes d'inégalité, que nous ne traitons de toute façon pas dans ce cours.*

## Exemple (suite)

Résolvons maintenant le problème

$$\text{maximize} \quad B(q_1, q_2) = 51q_1 + 70q_2 - 2q_1^2 - 3q_1q_2 - 3q_2^2$$
$$\text{sous contrainte} \quad q_1 + q_2 - 13 = 0$$

en utilisant la méthode des multiplicateurs de Lagrange. Tout d'abord, nous construisons le Lagrangien de ce problème :

$$L(q_1, q_2, \lambda) = 51q_1 + 70q_2 - 2q_1^2 - 3q_1q_2 - 3q_2^2 - \lambda(q_1 + q_2 - 13) = 0.$$

En prenant le gradient, nous obtenons le système d'équations suivant :

$$\begin{cases} 51 - 4q_1 - 3q_2 - \lambda = 0, \\ 70 - 3q_1 - 6q_2 - \lambda = 0, \\ -(q_1 + q_2 - 13) = 0. \end{cases}$$

En résolvant ce système comme d'habitude, nous obtenons

$$q_1 = 6, \quad q_2 = 7, \quad \lambda = 10.$$

**Remarque 3.** *Puisque nous avons introduit $\lambda$ comme un paramètre artificiel, nous n'avons pas besoin de connaître sa valeur. Par conséquent, nous pouvons nous en débarrasser dès que possible lors de la résolution du système.*

## Autre exemple

Considérons le problème d'optimisation suivant :

$$\text{maximize} \quad f(x_1, x_2) = x_1x_2 \quad \text{sous contrainte} \quad 2x_1 + x_2 = 1.$$

Notez que le seul point extremum candidat *libre* $x_1 = x_2 = 0$ de $f$ est un point de selle, mais sous la contrainte $2x_1 + x_2 = 1$ la fonction a un maximum ! Voyons d'abord cela en utilisant la méthode de substitution :

$$2x_1 + x_2 = 1 \implies x_2 = 1 - 2x_1,$$

nous devons donc trouver les extrema libres de

$$\tilde{f}(x_1) = x_1(1 - 2x_1).$$

Par conséquent,

$$\tilde{f}' = 0 \implies 1 - 4x_1 = 0 \implies x_1 = \frac{1}{4} \implies x_2 = \frac{1}{2}.$$

Puisque $\tilde{f}'' = -4 < 0$, ce point est bien un maximum.

## Applications économiques

Soient $q_1$ et $q_2$ les quantités de certains matériaux, à l'aide desquels nous pouvons produire la quantité $Q = Q(q_1, q_2)$ d'un nouvel objet. Si $p_1 = 4$ est le prix d'une unité de matériau 1 et $p_2 = 3$ est le prix d'une unité de matériau 2, et que nous avons besoin de $Q = 9$, alors le problème de minimisation des coûts consiste à minimiser le coût total

$$4q_1 + 3q_2$$

sous contrainte

$$Q(q_1, q_2) = 9.$$

Une hypothèse typique est que $Q$ est de la forme Cobb-Douglas. Par exemple,

$$Q(q_1, q_2) = 6q_1^{1/2}q_2^{3/2}.$$

Alors le problème est bien posé (c'est-à-dire que nous avons suffisamment d'informations pour commencer à le résoudre) :

$$\text{minimize} \quad 4q_1 + 3q_2 \quad \text{under} \quad 6q_1^{1/2}q_2^{3/2} = 9.$$

Le Lagrangien de ce problème est donné par

$$L(q_1, q_2, \lambda) = 4q_1 + 3q_2 - \lambda(6q_1^{1/2}q_2^{3/2} - 9).$$

Les conditions du premier ordre se lisent :

$$\operatorname{grad} L = \mathbf{0} \quad \text{or} \quad \begin{cases} 4 - 3\lambda q_1^{-1/2}q_2^{3/2} = 0 \\ 3 - 9\lambda q_1^{1/2}q_2^{1/2} = 0 \\ -6q_1^{1/2}q_2^{3/2} + 9 = 0. \end{cases}$$

Tout d'abord, nous nous débarrassons de $\lambda$ :

$$\lambda = \frac{4}{3}q_1^{1/2}q_2^{-3/2} = \frac{1}{3}q_1^{-1/2}q_2^{-1/2},$$

par conséquent

$$4q_1 = q_2.$$

En injectant cela à la place de $q_2$ dans la contrainte, nous obtenons

$$-6q_1^{1/2}(4q_1)^{3/2} + 9 = 0 \implies -48q_1^2 + 9 = 0 \implies q_1 = \pm\frac{\sqrt{3}}{4}.$$

Notez que la solution négative n'a pas de sens dans notre interprétation du problème, nous sommes donc laissés avec la solution suivante :

$$q_1 = \frac{\sqrt{3}}{4} \quad \text{and} \quad q_2 = \sqrt{3}.$$

Pour trouver $Q$, nous injectons $q_1$ et $q_2$ dans $Q(q_1, q_2)$ :

$$Q = 9.$$

### Solution alternative

Résolvons également le même problème par substitution :

$$6q_1^{1/2}q_2^{3/2} = 9 \implies q_1 = \frac{9}{4}q_2^{-3},$$

donc, au lieu de $4q_1 + 3q_2$ sous la contrainte $6q_1^{1/2}q_2^{3/2} = 9$, nous devons minimiser

$$\tilde{f}(q_1) = 4 \cdot \frac{9}{4}q_2^{-3} + 3q_2.$$

En différenciant, nous obtenons

$$-27q_2^{-4} + 3 = 0 \implies q_2 = \sqrt{3} \implies q_1 = \frac{\sqrt{3}}{4}.$$

## Optimisation sous plusieurs contraintes d'égalité

La méthode des multiplicateurs de Lagrange fonctionne également pour les problèmes d'optimisation sous plusieurs contraintes d'égalité :

$$\text{maximize/minimize} \quad f(\boldsymbol{x}) \quad \text{under constraints} \quad g_1(\boldsymbol{x}) = g_2(\boldsymbol{x}) = \cdots = g_m(\boldsymbol{x}) = 0.$$

Nous avons juste besoin de construire la fonction Lagrangienne comme suit :

$$L(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) - \boldsymbol{\lambda}^\top \boldsymbol{g}(\boldsymbol{x})$$

où $\boldsymbol{g}(x) = (g_1(\boldsymbol{x}), \dots, g_m(\boldsymbol{x}))^\top$ ou, en d'autres termes,

$$L(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) - \sum_{i=1}^{m} \lambda_i g_i(\boldsymbol{x}).$$

Notez que $L$ dans ce cas est une fonction de $n + m$ variables.

Le reste de la solution se déroule comme avant.

## Conditions du second ordre

Les conditions du second ordre dans les problèmes contraints sont plus compliquées que simplement vérifier que le Hessien est défini positif ou négatif. Nous ne traitons pas ces conditions dans ce cours ! En particulier, cette section n'est pas incluse dans l'examen.

Pour ceux qui sont intéressés, nous devons vérifier que le Hessien de $f$ *restreint à l'espace tangent à la contrainte* $F = \{\boldsymbol{v} : \boldsymbol{v}^\top \operatorname{grad} g(\boldsymbol{x}_0) = 0\}$ est défini positif ou négatif. En d'autres termes, que $\boldsymbol{v}^\top H(\boldsymbol{x}_0)\, \boldsymbol{v} < 0$ ou $> 0$ pour tout $\boldsymbol{v} \in F$ non nul.