Analyse 1: Optimisation avec contrainte

Joseph Salmon

Septembre 2014

Exemples de problèmes avec contraintes

En pratique : on optimise souvent avec contraintes (physiques)

- Contrainte de **positivité** : $K = \{x \in \mathbb{R}^d : \forall i \in [\![1, d]\!], x_i \geq 0\}$
- Contrainte de type **simplexe** (pour des probabilités) : $K = \Delta_d = \{x \in \mathbb{R}^d : \sum_{i=1}^d x_i = 1 \text{ et } \forall i \in [\![1, d]\!], x_i \geq 0\}$
- **Moindres carrés contraints** : on cherche x tel que $Ax = b$ avec une contrainte linéaire sur x, *e.g.*, $Bx = 0$ pour une matrice $B \in \mathbb{R}^{m \times d}$

On cherche alors à résoudre

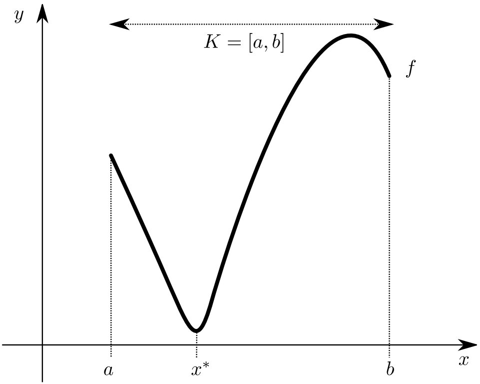
$$\boxed{x^* \in \arg\min_{x \in K} f(x)}$$

oú $K \subset \mathbb{R}^d$ est un ensemble qui encode les contraintes

Condition d'existence d'un minimum II

Théorème de Weierstrass

Si une fonction $f : \mathbb{R}^d \mapsto \mathbb{R}$ est continue sur un ensemble fermée et borné K (*i.e.*, un ensemble **compact**) alors il existe un point x^* qui atteint le minimum : $x^* \in \underset{x \in K}{\arg\min} f(x)$



Condition du premier ordre

Théorème : CNO cas contraint

Si f a un minimum local en x^* sur un convexe K, alors

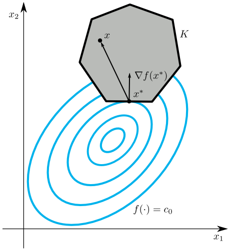
$$\forall x \in K, \langle \nabla f(x^*), x - x* \rangle \geq 0$$

Condition du premier ordre

Théorème : CNO cas contraint

Si f a un minimum local en x^* sur un convexe K, alors

$$\forall x \in K, \langle \nabla f(x^*), x - x* \rangle \geq 0$$

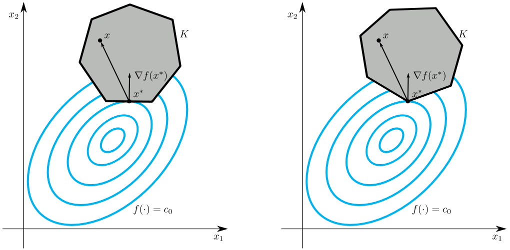


Condition du premier ordre

Théorème : CNO cas contraint

Si f a un minimum local en x^* sur un convexe K, alors

$$\forall x \in K, \langle \nabla f(x^*), x - x* \rangle \geq 0$$

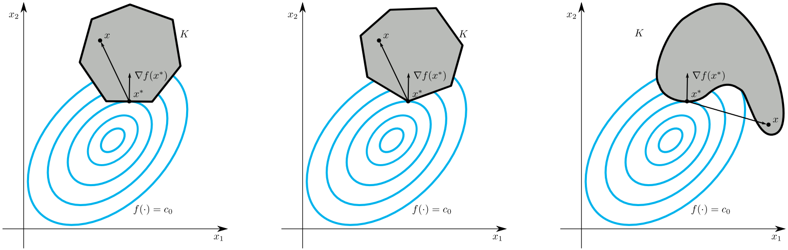


Condition du premier ordre

Théorème : CNO cas contraint

Si f a un minimum local en x^* sur un convexe K, alors

$$\forall x \in K, \langle \nabla f(x^*), x - x* \rangle \geq 0$$



Projection sur les convexes fermés

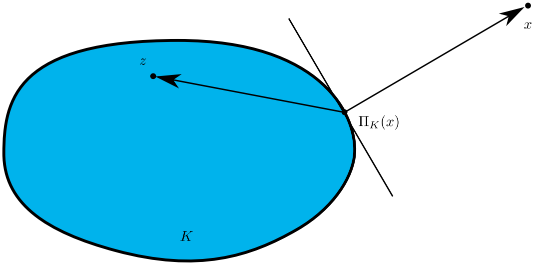
Théorème de projection

Si $K \subset \mathbb{R}^d$ est un convexe fermé non-vide, alors pour tout point $x \in \mathbb{R}^d$ il y a un unique point noté $\Pi_K(x)$ qui satisfait :

$$\Pi_K(x) = \arg\min_{z \in K} \frac{1}{2} \|x - z\|^2$$

De plus un point x^* est solution de ce problème ssi

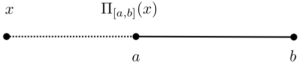
$$\forall z \in K, \langle z - x^*, x - x^* \rangle \leq 0$$

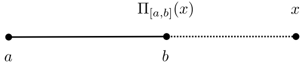


Projection sur les convexes fermés : exemples

- Le projecteur sur un intervalle $[a, b] \subset \mathbb{R}$ est la fonction

$$\Pi_{[a,b]}(x) = \begin{cases} x & \text{si } x \in [a, b] \\ a & \text{si } x < a \\ b & \text{si } x > b \end{cases}$$

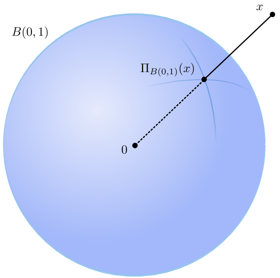




Vérification visuelle

- Le projecteur sur $B(0,1)$ (la boule centrée en 0 et de rayon unité) est la fonction

$$\Pi_{B(0,1)}(x) = \begin{cases} x & \text{si } \|x\| \leq 1 \\ \frac{x}{\|x\|} & \text{si } \|x\| > 1 \end{cases}$$



Optimisation avec contraintes et Lagrangien

En pratique : forme explicite pour les contraintes, avec m contraintes d'égalité, et r contraintes d'inégalité

$$(\mathcal{P}) \qquad \begin{array}{ll} \min & f(x) \\ \text{s. c.} & h_1(x) = 0, \ldots, h_m(x) = 0, \\ & g_1(x) \leq 0, \ldots, g_r(x) \leq 0, \end{array}$$

Définition : Lagrangien

On appelle **Lagrangien** du problème $(\mathcal{P})$ la fonction

$$\mathcal{L}(x, \lambda, \mu) = f(x) + \sum_{i=1}^{m} \lambda_i h_i(x) + \sum_{j=1}^{r} \mu_j g_j(x)$$

avec $\lambda = (\lambda_1, \ldots, \lambda_m)$ et $\mu = (\mu_1, \ldots, \mu_r)$ ayant toutes leurs coordonnées négatives ou nulles.

Conditions de Karush-Khunn-Tucker (KKT)

Théorème : KKT

Si x^* est un minimum local du problème $(\mathcal{P})$, que f, h_i, g_j sont dérivables avec des gradients continus, sous des conditions de qualifications sur x^*, il existe $\lambda^* = (\lambda_1^*, \dots, \lambda_m^*)$ et $\mu^* = (\mu_1^*, \dots, \mu_r^*)$ tel que :

$$
\begin{aligned}
\forall j \in [\![1, r]\!], \quad \mu_j^* &\geq 0, && \\
\nabla_x \mathcal{L}(x^*, \lambda^*, \mu^*) &= 0, && \text{(CNO)} \\
h_1(x^*) = 0, \dots, h_m(x^*) &= 0, && \text{(satisfiabilité)} \\
g_1(x^*) \leq 0, \dots, g_r(x^*) &\leq 0, && \text{(satisfiabilité)} \\
\forall j \in [\![1, r]\!], \quad \mu_j^* g_j(x^*) &= 0. && \text{(complémentarité)}
\end{aligned}
$$

cf. Bertsekas (1999) pour les détails sur la qualification d'un point

Conditions de Slater

Théorème : Slater

Supposons les même hypothèses que précédemment, et que de plus $\forall i \in [\![1, m]\!], h_i$ est affine, et qu'il existe un point $\bar{x}$ vérifiant $\forall j \in [\![1, r]\!], g_j(\bar{x}) < 0$ alors

$$
\begin{aligned}
\forall j \in [\![1, r]\!], \quad \mu_j^* &\geq 0, && \\
\nabla_x \mathcal{L}(x^*, \lambda^*, \mu^*) &= 0, && \text{(CNO)} \\
h_1(x^*) = 0, \ldots, h_m(x^*) &= 0, && \text{(satisfiabilité)} \\
g_1(x^*) \leq 0, \ldots, g_r(x^*) &\leq 0, && \text{(satisfiabilité)} \\
\forall j \in [\![1, r]\!], \quad \mu_j^* g_j(x^*) &= 0. && \text{(complémentarité)}
\end{aligned}
$$

Exemple de résolution I

Objectif quadratique et contrainte affine

$$(\mathcal{P}) \qquad \begin{array}{ll} \min\limits_{x_1,x_2} & \frac{1}{2}(x_1^2 + x_2^2) \\ \text{s. c.} & x_1 + x_2 \leq -2, \end{array}$$

$$\mathcal{L}(x,\mu) = \frac{1}{2}(x_1^2 + x_2^2) + \mu(x_1 + x_2 + 2)$$

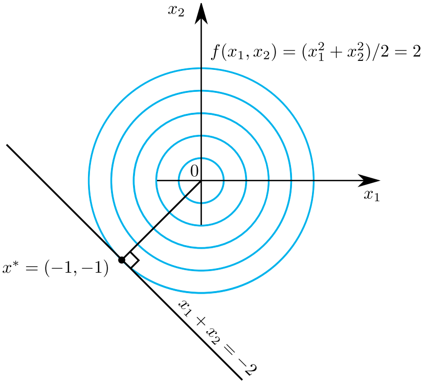
La CNO donne $x_1^* + \mu^* = x_2^* + \mu^* = 0$. Par complémentarité, on peut traiter deux cas exclusif

1. $x_1^* + x_2^* < -2$ et $\mu^* = 0$ (absurde !)
2. $x_1^* + x_2^* = -2$ et $\mu^* = 1$, puis $x_1^* = x_2^* = -1$

Vérification visuelle

Objectif quadratique et contrainte affine

$$
\begin{array}{lll}
 & \displaystyle\min_{x_1,x_2} & \dfrac{1}{2}(x_1^2 + x_2^2) \\
(\mathcal{P}) & \text{s. c.} & x_1 + x_2 \leq -2,
\end{array}
$$



Références I

- D. P. Bertsekas.
  Nonlinear programming.
  Athena Scientific, 1999.