Master MIMSE - Année 2

Optimisation Quadratique

Optimisation quadratique avec contraintes

Conditions de Karush-Kuhn-Tucker

Optimisation quadratique sous contraintes

- On veut

$$\begin{array}{rl} \min & f(x) \\ s.c. & g_j(x) \leq 0,\ \forall j = 1, \ldots, p \\ & h_i(x) = 0,\ i = 1, \ldots, m \\ & x \in \Re^n. \end{array}$$

  avec f et g_k linéaires et/ou quadratiques,
- donc c^∞.
- On cherche s'il existe une solution *analytique*
- c'est-à-dire des conditions nécessaires et/ou suffisantes pour caractériser un point optimal.
- Analogie avec $\nabla f(x^*) = 0$.
- Conditions locales (voisinage de x^*)
- donc optimum (ou non) local.
- Karush : thèse de Master, 1939,
- Kuhn-Tucker : "Nonlinear Programming", 1951
- Origine : Lagrange, 1761.

Solution dans l'intérieur

- Si un optimum $\hat{\boldsymbol{x}}$ est dans l'intérieur de l'espace des solutions,
- aucune contrainte d'inégalité n'est *active* :

$$\forall \boldsymbol{k},\ \boldsymbol{g_k}(\hat{\boldsymbol{x}}) > \boldsymbol{0}$$

- les conditions en $\hat{\boldsymbol{x}}$ sont les mêmes que pour l'optmisation sans contraintes.
- $\hat{\boldsymbol{x}}$ peut s'obtenir numériquement par les méthodes d'optimisation sans contraintes.

Contraintes d'égalité

- Conditions de Lagrange (1761) : *Théorème :*
- Si $\bar{x}$ est un mimimum local de f
- sous les contraintes $h_i = 0$ (pour $i = 1, \ldots, m$)
- et si les $\nabla h_i(\bar{x})$ sont linéairement indépendants
- alors il existe $\bar{\lambda} = (\bar{\lambda}_1, \ldots, \bar{\lambda}_m)$ tels que

$$\nabla f(\bar{x}) + \sum_{i=1}^{m} \bar{\lambda}_i \nabla h_i(\bar{x}) = 0.$$

- Indépendance des $\nabla h_i(\bar{x})$ entraîne l'unicité de $\bar{\lambda}$.

Interprétation géométrique

- Appelons
$$\boldsymbol{T}(\bar{\boldsymbol{x}}) = \{\boldsymbol{d} | \nabla \boldsymbol{h}_i(\bar{\boldsymbol{x}})^T \boldsymbol{d} = 0\}$$
- Ensemble des directions tangentes à $\boldsymbol{\mathcal{S}}$ en $\bar{\boldsymbol{x}}$,
- $\bar{\boldsymbol{x}} + \boldsymbol{T}(\bar{\boldsymbol{x}})$ espace affine tangent à $\boldsymbol{\mathcal{S}}$ en $\bar{\boldsymbol{x}}$,
- approximation au premier ordre de $\boldsymbol{\mathcal{S}}$ au voisinage de $\bar{\boldsymbol{x}}$.
- Si $\nabla \boldsymbol{f}(\bar{\boldsymbol{x}})^T \boldsymbol{d} < 0$ pour un $\boldsymbol{d} \in \boldsymbol{T}(\bar{\boldsymbol{x}})$,
- "direction de descente" dans $\boldsymbol{T}(\bar{\boldsymbol{x}})$,
- l'approximation au premier ordre de $\boldsymbol{f}$ au voisinage (approché) de $\bar{\boldsymbol{x}}$ n'est pas optimisée,
- $\bar{\boldsymbol{x}}$ ne peut être optimum.
- *cf.* la relation fondamentale de la recherche linéaire en dimension 1.

Interprétation analytique

- On appelle **lagrangien** associé au problème d'optimisation de f sous les contraintes h_i

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

- Pour x faisable, pour tout λ, $\mathcal{L}(x, \lambda) = f(x)$
- (c'est équivalent d'optimiser f et $\mathcal{L}(., \lambda)$).
- Condition : $\nabla_x \mathcal{L}(\bar{x}, \bar{\lambda}) = 0$.
- De plus $\nabla_\lambda \mathcal{L}(\bar{x}, \bar{\lambda}) = (h_1(\bar{x}), \ldots, h_m(\bar{x})) = 0$ pour $\bar{x}$ faisable,
- d'où $\nabla \mathcal{L}(\bar{x}, \bar{\lambda}) = 0.$
- Extension du $\nabla f = 0$ du cas sans contraintes au cas avec contraintes.
- **Attention !** en général $(\bar{x}, \bar{\lambda})$ n'est pas un optimum,
- mais plutôt un point-selle !

Utilisation du théorème

- On a à résoudre :

$$\begin{gathered}\boldsymbol{\nabla} \boldsymbol{f}(\bar{\boldsymbol{x}}) + \sum_{i=1}^{m} \bar{\boldsymbol{\lambda}}_i \boldsymbol{\nabla} \boldsymbol{h}_i(\bar{\boldsymbol{x}}) = 0 \\ \boldsymbol{h}_1(\bar{\boldsymbol{x}}) = 0 \\ \vdots \\ \boldsymbol{h}_m(\bar{\boldsymbol{x}}) = 0\end{gathered}$$

- Pas de condition sur les signes des $\boldsymbol{\lambda}$.
- $\boldsymbol{m} + \boldsymbol{n}$ équations à $\boldsymbol{m} + \boldsymbol{n}$ inconnues,
- pas nécessairement linéaires (si $\boldsymbol{h}_i$ quadratique...)
- ne pas oublier $\boldsymbol{\nabla} \boldsymbol{h}_i$ linéairement indépendants
- “condition de régularité de Lagrange”
- Condition nécessaire, pas suffisante en général.

Un cas particulier quadratique

- Un cas important où les conditions sont suffisantes :
- f quadratique convexe
- et h_i linéaires (affines)
- On veut :

$$\begin{array}{rl} \min & \frac{1}{2}x^T A x + b^T x + c \\ s.c. & Hx = d \\ & x \in \Re^n. \end{array}$$

- On suppose $A \succ 0$ et H de rang $m(< n)$
- donc les ∇h_i sont linéairement indépendants.
- Si en $\bar{x}$ il existe $\bar{\lambda}$ tel que les conditions de Lagrange sont vérifiées,
- alors $\bar{x}$ est un minimum global de f sous les contraintes $h_i(x) = b_i$.
- On peut trouver explicitement $(\bar{x}, \lambda x)$.
- On a $H\bar{x} = d$ et $A\bar{x} + b + H^T\bar{\lambda} = 0$.
- En utilisant que A est inversible et en posant $B = (HA^{-1}H^T)^{-1}$ (inverse d'une matrice définie positive), on obtient :

$$\bar{x} = A^{-1}H^T B(HA^{-1}b + d) - A^{-1}b$$

et

$$\bar{\lambda} = -B(HA^{-1}b + d).$$

Avec des contraintes d'inégalité

- On a maintenant les contraintes sous la forme

$$\boldsymbol{g_j(x) \leq 0,\ j = 1, \ldots, p.}$$

- Soit $\boldsymbol{x}$ faisable, par exemple sur la frontière du domaine,
- on pose $\boldsymbol{J(x) = \{j | g_j(x) = 0\}}$
- ensemble des indices des contraintes *actives*.
- Pour les autres $\boldsymbol{j}$ on a $\boldsymbol{g_j(x) < 0}$ sur tout un voisinage de $\boldsymbol{x}$ donc ne joue aucun rôle.
- *Théorème*
- Si $\boldsymbol{\bar{x}}$ est un minimum local de $\boldsymbol{f}$ sur $\boldsymbol{\mathcal{S}}$, alors il existe des $\boldsymbol{\bar{\mu}_0, \bar{\mu}_1, \ldots, \bar{\mu}_p}$ tels que
  1. $\boldsymbol{\bar{\mu}_0 \nabla f(\bar{x}) + \sum_{j=1}^{p} \bar{\mu}_i \nabla g_j(\bar{x}) = 0,}$
  2. $\boldsymbol{\bar{\mu}_j \geq 0}$ pour tout $\boldsymbol{j \geq 0}$
  3. $\boldsymbol{\bar{\mu}_j g_j(\bar{x}) = 0}$ pour tout $\boldsymbol{j \geq 1}$.

Remarques

- Le fait que $\bar{\mu}_0$ peut être égal à 0 est un vrai problème.
- Si de plus (QC) : il existe d tel que $\nabla g_j(\bar{x})^T d < 0$ pour tout $j \in J(\bar{x})$, alors ("qualification des contraintes") :
- $\bar{\mu}_0 > 0$ et on peut donc reformuler :
  1. $\nabla f(\bar{x}) + \sum_{j=1}^{p} \bar{\mu}_i \nabla g_j(\bar{x}) = 0.$
- Signe des μ_i important,
- "multiplicateurs positivement homogènes"
- La 3ème condition équivaut à : $\bar{\mu}_j = 0 \; \forall j \notin J(\bar{x})$.
- "qualification des contraintes" : il doit exister une direction pointant à l'intérieur du cône convexe qui joue le rôle d'ensemble tangent à $\mathcal{S}$ en $\bar{x}$,
- ou encore : les $\nabla g_j(\bar{x}), \; j \in J(\bar{x})$ sont linéairement indépendants "de façon positivement homogène".
- Conditions nécessaires, pas suffisantes dans le cas général,
- suffisantes si f et les g_j sont convexes ! (optimum global alors).

Cas général

- **Conditions de Karush-Kuhn-Tucker**
- Si $\bar{x}$ est un minimum local de f sur $\mathcal{S}$,
- si des conditions de qualifications de contraintes sont vérifiées,
- alors il existe $\bar{\lambda}_1, \ldots, \bar{\lambda}_m, \bar{\mu}_1, \ldots, \bar{\mu}_p$ tels que
  1. $\bar{\mu}_0 \nabla f(\bar{x}) + \sum_{i=1}^m \bar{\lambda}_i \nabla h_i(\bar{x}) + \sum_{j=1}^p \bar{\mu}_i \nabla g_j(\bar{x}) = 0,$
  2. $\bar{\mu}_j \geq 0$ pour tout $j \geq 0$
  3. $\bar{\mu}_j g_j(\bar{x}) = 0$ pour tout $j \geq 1$.

- **Qualification des contraintes :**
  1. Il existe d tel que $\nabla h_i(\bar{x})^T d = 0$ pour tout i
  2. et $\nabla g_j(\bar{x})^T d < 0$ pour tout $j \in J(\bar{x})$ et
  3. les $\nabla h_i(\bar{x})$ sont linéairement indépendants.

Remarques

- Résoudre les conditions de KKT : résoudre un système de $n+m+p$ équations à $n+m+p$ inconnues. Avec $\mu_j g_j(x) = 0$, séparer en 2 cas (donc 2^p cas).
- Les $\bar{\lambda}_i$ et les $\bar{\mu}_j$ sont appelés *“multiplicateurs de Lagrange”*.
- On peut définir le lagrangien associé au problème :

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

- Alors on a $\nabla \mathcal{L}(\bar{x}, \bar{\lambda}, \bar{\mu}) = 0$.
- En fait à $(\bar{\lambda}, \bar{\mu})$ fixés, $\bar{x}$ minimise $\mathcal{L}$,
- et à $\bar{x}$ fixé $(\bar{\lambda}, \bar{\mu})$ maximise le lagrangien.
- Si inégalité avec un second membre $g_j(x) \leq b_j$ on a dans le lagrangien $\mu_j(g_j(x) - b_j)$.
- Alors $\mu_j = -\frac{\partial \mathcal{L}}{\partial b_j}$: b_j est le bénéfice de relâcher de 1 la contrainte (“interprétation économique des multiplicateurs de Lagrange”).
- Il existe une multitude de conditions de qualification des contraintes. On peut par exemple demander que (c’est plus fort)

  les $\nabla h_i(\bar{x}), \nabla g_j(\bar{x}),\ j \in J(\bar{x})$ sont linéairement indépendants.

Cas particulier quadratique

- On veut

$$\begin{array}{rl} \min & f(x) \\ s.c. & g_j(x) \leq 0,\ \forall j = 1, \ldots, p \\ & h_i(x) = 0,\ i = 1, \ldots, m \\ & x \in \Re^n. \end{array}$$

avec

- f quadratique convexe,
- g_j quadratiques convexes,
- h_i affines : $h_i(x) = a_i^T x - e_i$.

- Alors, les conditions de KKT sont toujours suffisantes pour que $\bar{x}$ soit un minimum sous contraintes.
- Si QC, conditions nécessaires et suffisantes.
- “KKT convexe”
- Condition de Slater (entraîne la qualification des contraintes) :
  1. les a_i sont linéairement indépendants,
  2. il existe un point à l'intérieur de $\mathcal{S}$: x tel que $h_i(x) = 0$ et $g_j(x) < 0$.
- Conditions QC indépendantes de $\bar{x}$!

Conditions de minimalité du 2nd ordre

Conditions nécessaires

Si $\bar{x}$ est un minimum local et si QC tient en $\bar{x}$, alors il existe des $\bar{\lambda}$ et $\bar{\mu}$ tels que

- $\nabla_x \mathcal{L}(\bar{x}, \bar{\lambda}, \bar{\mu}) = 0$
- $\bar{\mu}_j g_j(\bar{x}) = 0, \forall j$
- $d^T \nabla^2_{xx} \mathcal{L}(\bar{x}, \bar{\lambda}, \bar{\mu}) d \geq 0$ pour tout d tel que
  - $\nabla h_i(\bar{x})^T d = 0$ pour tout i,
  - $\nabla g_j(\bar{x})^T d = 0$ pour tout $j \in J(\bar{x})$ tel que $\bar{\mu}_j > 0$,
  - $\nabla h_i(\bar{x})^T d \leq 0$ pour tout $j \in J(\bar{x})$ tel que $\bar{\mu}_j = 0$.

Conditions suffisantes

Si plus haut on a $d^T \nabla^2_{xx} \mathcal{L}(\bar{x}, \bar{\lambda}, \bar{\mu}) d > 0$ pour d vérifiant les conditions, alors $\bar{x}$ est un minimum strict de f sous les contraintes.