le cnam

Analyse Mathématique pour l'Ingénieur

Cours 8 Introduction à l'optimisation

- Formule de Taylor dans un contexte très simple

On se donne une fonction φ définie sur l'intervalle $[0,\,1]$ et assez régulière. On a la relation $\varphi(1) = \varphi(0) + \varphi'(0) + \int_0^1 (1-t)\,\varphi''(t)\,\mathrm{d}t$.

- Formule de Taylor à l'ordre deux

On se donne un entier naturel $n \geq 1$, $x_0 \in \mathbb{R}^n$, $h \in \mathbb{R}^n$ et f définie au voisinage du segment $[x_0,\, x_0+h] = x_0 + [0,\,1]\,h = \{x = x_0 + \theta\,h,\, 0 \leq \theta \leq 1\}$. On suppose f deux fois continuement différentiable dans ce voisinage ; les dérivéss partielles premières sont continuement différentiables et les dérivées partielles secondes sont supposées continues. Alors on a la formue de Taylor avec reste intégral : $f(x_0+h) = f(x_0) + \mathrm{d}f(x_0).h + \int_0^1 (1-\theta)\,\mathrm{d}^2 f(x_0+\theta\,h).(h,\,h)\,\mathrm{d}\theta$.

- Formule de Taylor à l'ordre trois

On se donne $x_0 \in \mathbb{R}^n$, $h \in \mathbb{R}^n$ et f définie au voisinage du segment $[x_0,\, x_0+h]$. On suppose la fonction f trois fois continuement différentiable dans ce voisinage ; les dérivéss partielles premières sont deux fois continuement différentiables et les dérivées partielles secondes sont une fois continuement différentiables. Alors on a la formue de Taylor a l'ordre trois :
$f(x_0+h) = f(x_0) + \mathrm{d}f(x_0).h + \frac{1}{2}\,\mathrm{d}^2 f(x_0).(h,\,h) + \frac{1}{2}\int_0^1 (1-\theta)^2\,\mathrm{d}^3 f(x_0+\theta\,h).(h,\,h,\,h)\,\mathrm{d}\theta$.

- Formule de Taylor avec reste intégral

Si, toutes choses égales par ailleurs, la fonction f est $(n+1)$ fois continuement différentiable au voisinage du segment $[x_0,\, x_0+h]$, on a la formule de Taylor avec reste intégral à l'ordre n:

$$f(x_0+h) = f(x_0) + \mathrm{d}f(x_0).h + \cdots + \tfrac{1}{k!}\,\mathrm{d}^k f(x_0).(h,\,\ldots,\,h) + \cdots + \tfrac{1}{n!}\,\mathrm{d}^n f(x_0).(h,\,\ldots,\,h) \\ + \tfrac{1}{n!}\int_0^1 (1-\theta)^n\,\mathrm{d}^{n+1} f(x_0+\theta\,h).(h,\,\ldots,\,h)\,\mathrm{d}\theta.$$

- Condition nécessaire de minimum d'ordre un

On se donne un espace de Banach E, $a \in E$, une application f de E dans $\mathbb{R}$ définie au voisinage de a de sorte que f est minimale au poit a: il existe $r > 0$ tel que pour tout $x \in B(a,\,r)$, $f(x) \geq f(a)$. Alors si f est différentiable au point a, on a $\mathrm{d}f(a) = 0$.

Attention. Si le minimum n'est pas relatif à une boule $B(a,\,r)$ au voisinage du point a, la condition précédente est en défaut en général. Par exemple, la fonction f définie de l'intervalle $[0,\,1]$ et à valeurs réelles par la relation $f(x) = x$ a clairement un minimum pour $x = 0$. Pourtant on n'a pas $f'(0) = 0$!

- Condition nécessaire de minimum d'ordre deux

On se donne un espace de Banach E, $a \in E$, $r > 0$ et f une application définie de $B(a,\,r)$ et à valeurs dans $\mathbb{R}$, continuement différentiable dans la boule $B(a,\,r)$ et deux fois différentiable au point a. On suppose f minimale au poit a: pour tout $x \in B(a,\,r)$, $f(x) \geq f(a)$. Alors on a l'inégalité $\mathrm{d}^2 f(a).(h,\,h) \geq 0$ pour tout $h \in E$.

François Dubois, novembre 2021.

En dimension finie n, la matrice hessienne $\big(\frac{\partial^2 f}{\partial x_i \partial x_j}\big)(a)$ au point de minimum a est une matrice symétrique positive.
Par exemple, avec $E = \mathbb{R}^2$ et $f(x, y) = \frac{1}{2}(x^2+y^2)$, on a $\mathrm{d}f(0, 0) = 0$ et $\mathrm{d}^2 f(0, 0)$ représenté par la matrice hessienne $H(0, 0) = \mathrm{I}_2$. Cette matrice est effectivement symétrique et positive.

- Multiplicateurs de Lagrange

On se place dans le contexte où $E = \mathbb{R}^n$. On se donne une fonction régulière g de $\mathbb{R}^n$ dans $\mathbb{R}$ et on suppose que l'ensemble K défini par $K = \{x \in \mathbb{R}^n, g(x) = 0\}$ est non vide. On se donne une fonction f de $\mathbb{R}^n$ dans $\mathbb{R}$ définie dans un voisinage de K. On se donne un point $a \in K$ qui satisfait donc à la condition $g(a) = 0$. On suppose que le point a est un point régulier, c'est à dire $\mathrm{d}g(a) \neq 0$: il existe un indice j compris entre 1 et n tel que $\big(\frac{\partial g}{\partial x_j}\big)(a) \neq 0$.
Par exemple avec $n = 2$ et $g(x, y) = x^2 + y^2 - 1$, tous les points de K sont réguliers.
Alors si a est un minimum local de la fonction f: $f(x) \geq f(a)$ pour tout $x \in K$ au voisinage de a, il existe $\lambda \in \mathbb{R}$ tel que $\mathrm{d}f(a) + \lambda\, \mathrm{d}g(a) = 0$. Le nombre λ est appelé multiplicateur de Lagrange associé a la contrainte $g(x) = 0$.
Les équations $\mathrm{d}f(a) + \lambda\, \mathrm{d}g(a) = 0$ sont parfois appelées équations d'Euler-Lagrange du problème d'optimisation sous contrainte. On peut les écrire aussi à l'aide de dérivées partielles :
$\big(\frac{\partial f}{\partial x_j} + \lambda \frac{\partial g}{\partial x_j}\big)(a) = 0$ pour tout entier j tel que $1 \leq j \leq n$.
La preuve de ce résultat utilise de façon essentielle le théorème des fonctions implicites.

- Lagrangien pour la recherche des points extrémaux.

Ce qui vient d'être proposé dans le cas d'un minimum d'étend naturellement au cas d'un maximum. La condition s'écrit encore $g(a) = 0$ et $\mathrm{d}f(a) + \lambda\, \mathrm{d}g(a) = 0$ si le point a est régulier.
En pratique, on forme le lagrangien $\mathscr{L}(x, \lambda) \equiv f(x) + \lambda\, g(x)$. Puis on écrit la condition pour que le point de minimum ou de maximum satisfasse aux contraintes, c'est a dire $\frac{\partial \mathscr{L}}{\partial \lambda} \equiv g(x) = 0$.
Enfin la condition de minimum ou de maximum sous contrainte prend la forme
$\frac{\partial \mathscr{L}}{\partial x} \equiv \mathrm{d}f(x) + \lambda\, \mathrm{d}g(x) = 0$.
Si on recherhe par exemple les points extrémaux de la fonction de deux variables réelles
$f(x, y) = x^2 - y^2$ avec la contrainte $x^2 + y^2 = 1$. On a donc
$\mathscr{L}(x, y, \lambda) = x^2 - y^2 - \lambda\,(x^2 + y^2 - 1)$. La suite de la résolution consiste à dériver le lagrangien par rapport à λ et par rapport à x. On annule ensuite ces deux dérivées partielles ; c'est un exercice laissé au lecteur. Il importe enfin de savoir quel est le type d'extremum pour les les points stationnaires obtenus : minimum, maximum ou point selle.

- Interprétation du multiplicateur de Lagrange

On se donne maintenant un domaine d'optimisation K_ξ paramétré par $\xi \in \mathbb{R}^p$. La fonction g est une fonction régulière de $\mathbb{R}^n$ dans $\mathbb{R}^p$ et $K_\xi = \{x \in \mathbb{R}^n, g(x) = \xi\}$. On suppose K_ξ non vide au voisinage de $\xi_0 \in \mathbb{R}^p$. La famille de problèmes de minimisation sous contraintes s'écrit alors : chercher $a_\xi \in K_\xi$ de sorte que $f(x) \geq f(a_\xi)$ pour tout $x \in K_\xi$ voisin de a_ξ. Les équations d'Euler-Lagrange s'écrivent essentiellement toujours de la même façon : $\big(\frac{\partial f}{\partial x_j} + \sum_{k=1}^{p} \lambda_k \frac{\partial g}{\partial x_j}\big)(a_\xi) = 0$ pour tout entier j tel que $1 \leq j \leq n$. Le multiplicateur de lagrange λ est maintenant un vecteur de $\mathbb{R}^p$. De plus, la contrainte $g(a_\xi) = \xi$ donne le paramétrage du problème. Notons $J(\xi) = f(a_\xi)$ la valeur du minimum de la fonction f sur l'ensemble K_ξ. La dérivée partielle de la valeur du minimum dans une variation du paramètre ξ est égale au signe près au multiplicateur de Lagrange : $\frac{\partial J}{\partial \xi_k} + \lambda_k = 0$.

Exercices

- Forme quadratique à plusieurs variables

On se donne un entier $n \geq 1$, une matrice A symétrique réelle à n lignes et n colonnes et b un vecteur de $\mathbb{R}^n$. On note $(x, y) \equiv \sum_{j=1}^{n} x_j y_j$ le produit scalaire de deux vecteurs de $\mathbb{R}^n$. On pose $J(x) = \frac{1}{2}(Ax, x) - (b, x)$. C'est une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$.

a) Montrer que la fonction J est différentiable en tout point x de $\mathbb{R}^n$ et calculer l'action $\mathrm{d}J(x).h$ de la différentielle $\mathrm{d}J(x)$ sur un vecteur $h \in \mathbb{R}^n$ arbitraire.

b) Comment s'exprime la condition $\mathrm{d}J(x) = 0$?

- Point de minimum pour une fonction quadratique

Soit $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ et $b = \begin{pmatrix} -5 \\ 2 \end{pmatrix}$. On note $(.,.)$ le produit scalaire usuel dans $\mathbb{R}^2$. Pour $x \in \mathbb{R}^2$, on pose $J(x) = \frac{1}{2}(x, Ax) - (b, x)$.

a) Pour $h \in \mathbb{R}^2$, calculer $\mathrm{d}J(x).h$.

b) Si il existe x^* tel que pour tout $x \in \mathbb{R}^2$, $J(x) \geq J(x^*)$, quelle équation doit satisfaire le point x^* ?

c) Résoudre cette équation et démontrer que la fonction J est effectivement minimale au point x^* : $J(x) \geq J(x^*)$ pour tout point x de $\mathbb{R}^2$.

- Calcul différentiel [février 2014, avril 2014]

On se donne un entier $n \geq 1$, une matrice A symétrique réelle à n lignes et n colonnes et b un vecteur de $\mathbb{R}^n$. On note $(x, y) \equiv \sum_{j=1}^{n} x_j y_j$ le produit scalaire de deux vecteurs de $\mathbb{R}^n$. On pose $J(x) = \frac{1}{2}(Ax, x) - (b, x)$. C'est une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$.

a) Montrer que la fonction J est différentiable en tout point x , de $\mathbb{R}^n$.

b) Calculer l'action $\mathrm{d}J(x) \bullet h$ de la différentielle $\mathrm{d}J(x)$ sur un vecteur $h \in \mathbb{R}^n$ arbitraire.

c) Comment s'exprime la condition $\mathrm{d}J(x) = 0$?

On suppose maintenant que la matrice A est positive : $(h, Ah) \geq 0$, pour tout $h \in \mathbb{R}^n$.

d) Montrer que la fonction J admet un unique point de minimum x^* qui vérifie
$J(x) \geq J(x^*)$ pour tout $x \in \mathbb{R}^n$.

e) Préciser la valeur de x^* lorsque $n = 2$, $A = \begin{pmatrix} 3 & -1 \\ -1 & 1 \end{pmatrix}$ et $b = \begin{pmatrix} 4 \\ -2 \end{pmatrix}$.

f) Préciser la valeur de x^* lorsque $n = 2$, $A = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}$ et $b = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$.

- Optimisation sous contrainte

On cherche à minimiser la fonction $f(x) = (x-2)^2 + y^2 + z^2$ sur l'ellipsoïde E d'équation $x^2 + 2y^2 + 3z^2 = 1$.

a) Les points de l'ellipsoïde sont-ils tous réguliers ?

b) Pourquoi la borne inférieure de la fonction f sur l'ellipsoïde E est-elle atteinte ?

c) Ecrire les équations d'Euler-Lagrange de ce problème.
d) En quels points le minimum de la fonction f est-il atteint ?
e) Reprendre les questions précédentes avec la fonction $g(x, y)) = x^3 + y^3$ sur le cercle K d'équation $x^2 + y^2 = 4$.
f) Reprendre les questions a) à d) avec la fonction $h(x, y) = x^2 + y^2 - 4xy$ sur le cercle L d'équation $x^2 + y^2 = 8$.

• Matrices symétriques

On se donne un entier $n \geq 1$ et une matrice carrée symétrique $A \in \mathscr{M}_n(\mathbb{R})$ à coefficients réels. On désigne par $\| \ \|$ la norme euclidienne sur $\mathbb{R}^n$ et par $(\ , \ )$ le produit scalaire associé. Soit $S = \{x \in \mathbb{R}^n, \|x\| = 1\}$ la sphère unité de $\mathbb{R}^n$.
a) Montrer que le problème de chercher $x \in S$ de sorte que $\forall y \in S, \ (x, Ax) \geq (y, Ay)$ a au moins une solution.
b) Ecrire le lagrangien associé au problème proposé à la question précédente.
c) Montrer que la solution du problème proposé à la question a) définit un vecteur propre de la matrice A.
d) Soit x une solution du problème proposé à la question a). Montrer que si $y \in x^\perp$, alors $Ay \in x^\perp$.
e) En déduire une démonstration par récurrence sur la dimension n du résultat classique qui énonce qu'une matrice symétrique réelle admet une base orthonormée de vecteurs propres.

• Surface maximale

On se donne L strictement positif et on cherche une fonction $y = f(x)$ pour $0 \leq x \leq a$ de sorte que la longueur de la courbe soit égale à L et telle que la surface sous la courbe $y = f(x)$ est maximale.
Déterminer une telle courbe.

• Introduction à l'équation de Hamilton-Jacobi

On cherche à minimiser la fonction différentiable J de $\mathbb{R}^n$ dans $\mathbb{R}$ sous la contrainte $g(x) = \xi$, où g est une fonction différentiable de $\mathbb{R}^n$ dans $\mathbb{R}^p$. On note $f(\xi)$ la valeur minimale de J solution du problème précédent et λ la valeur du multiplicateur de Lagrange associé au point de minimum.
Montrer que la fonction f est différentiable par rapport à ξ et que sa différentielle est égale (au signe près éventuellement !) au multiplicateur de Lagrange.