

Calcul Différentiel

Dérivées Partielles et Gradient

Soit $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$.

— Dérivée partielle par rapport à x_i :

$$\frac{\partial f}{\partial x_i}(x) = \lim_{t \rightarrow 0} \frac{f(x + te_i) - f(x)}{t}$$

— Gradient (∇f) : Vecteur des dérivées partielles.

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right)^T$$

— Différentielle ($D_x f$) : Application linéaire t.q.

$$f(x+h) - f(x) = D_x f(h) + o(\|h\|)$$

Pour f scalaire : $D_x f(h) = \langle \nabla f(x), h \rangle$.

Hessienne et Théorème de Schwarz

— Théorème de Schwarz : Si f est C^2 , alors :

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$$

— Matrice Hessienne ($H_f(x)$) : Matrice symétrique des dérivées secondes.

$$[H_f(x)]_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(x)$$

Développements Limités (Taylor)

Soit f de classe C^2 sur U .

— Ordre 1 : $f(x+h) = f(x) + \langle \nabla f(x), h \rangle + o(\|h\|)$

— Ordre 2 :

$$f(x+h) = f(x) + \langle \nabla f(x), h \rangle + \frac{1}{2} \langle H_f(x)h, h \rangle + o(\|h\|^2)$$

Extrema Sans Contraintes

Conditions d'Existence

Théorème des bornes atteintes :

Si $K \subset \mathbb{R}^n$ est compact (fermé et borné) et f continue, alors f atteint son minimum et son maximum sur K .

Conditions Nécessaires (Ordre 1)

Point critique : x est critique si $\nabla f(x) = 0$.

Si f admet un extremum local en x , alors x est un point critique.

Conditions Suffisantes (Ordre 2)

Soit x un point critique ($\nabla f(x) = 0$). On étudie $H_f(x)$.

— $H_f(x)$ définie positive (valeurs propres > 0) \implies Minimum local strict.

— $H_f(x)$ définie négative (valeurs propres < 0) \implies Maximum local strict.

— Valeurs propres de signes opposés \implies Point selle (pas d'extremum).

Cas $n = 2$ (Notations de Monge) :

Posons $r = \frac{\partial^2 f}{\partial x^2}, s = \frac{\partial^2 f}{\partial x \partial y}, t = \frac{\partial^2 f}{\partial y^2}$.

— $rt - s^2 < 0 \implies$ Point selle.

— $rt - s^2 > 0$ et $r > 0 \implies$ Minimum.

— $rt - s^2 > 0$ et $r < 0 \implies$ Maximum.

Optimisation sous Contraintes

Contraintes d'Égalité (Lagrange)

Minimiser $f(x)$ sous $g_1(x) = \dots = g_m(x) = 0$.

Théorème des Multiplicateurs de Lagrange :

Si \tilde{x} est un extremum local et que les vecteurs $\{\nabla g_i(\tilde{x})\}$ sont linéairement indépendants (contraintes qualifiées), alors $\exists \lambda_1, \dots, \lambda_m \in \mathbb{R}$ tels que :

$$\nabla f(\tilde{x}) = \sum_{i=1}^m \lambda_i \nabla g_i(\tilde{x})$$

Le Lagrangien est défini par :

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

Condition : $\nabla_x \mathcal{L} = 0$ et $\nabla_\lambda \mathcal{L} = 0$.

Contraintes d'Inégalité (KKT)

Minimiser $f(x)$ sous :

- $g_i(x) = 0$ ($1 \leq i \leq m$)
- $h_j(x) \leq 0$ ($1 \leq j \leq p$)

Conditions de Karush-Kuhn-Tucker (KKT) :

Si \tilde{x} est un minimum local et que les contraintes sont qualifiées, alors $\exists \lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^p$ tels que :

1. Stationnarité : $\nabla f(\tilde{x}) = \sum \lambda_i \nabla g_i(\tilde{x}) + \sum \mu_j \nabla h_j(\tilde{x})$
2. Admissibilité : $g_i(\tilde{x}) = 0$ et $h_j(\tilde{x}) \leq 0$
3. Signe : $\mu_j \leq 0$ (pour un min)
4. Complémentarité : $\mu_j h_j(\tilde{x}) = 0$ ($\mu_j = 0$ si contrainte inactive).

Note : Pour un maximum, $\mu_j \geq 0$.

Qualification des contraintes

— Indépendance linéaire : Les gradients des contraintes actives sont linéairement indépendants.

— Qualification linéaire : Si les contraintes sont affines ($Ax \leq b$), la qualification est automatique.

Dualité

Problème Primal : $p^* = \inf_x \sup_{\lambda, \mu} \mathcal{L}(x, \lambda, \mu)$.

Problème Dual : $d^* = \sup_{\lambda, \mu} \inf_x \mathcal{L}(x, \lambda, \mu)$ avec $\mu \leq 0$.

— Dualité Faible : $d^* \leq p^*$ (toujours vrai).

— Dualité Forte : $d^* = p^*$. Vrai si le problème est convexe et satisfait la condition de Slater (il existe un point strictement admissible $h_j(x) < 0$).

Convexité

Définitions et Propriétés

Un ensemble U est convexe si $\forall x, y \in U, t \in [0, 1], (1-t)x + ty \in U$. Une fonction $f : U \rightarrow \mathbb{R}$ est convexe si :

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y)$$

Caractérisations ($f \in C^1$ ou C^2) :

1. Ordre 1 : $f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle$.

(Le graphe est au-dessus de la tangente).

2. Ordre 2 : $H_f(x)$ est une matrice positive ($\forall h, \langle H_f(x)h, h \rangle \geq 0$).

Théorème fondamental :

Pour une fonction convexe sur un ouvert convexe :

— Tout minimum local est un minimum global.

— Si $\nabla f(x) = 0$, alors x est un minimum global.

— Si f est strictement convexe, le minimum (s'il existe) est unique.

Opérations conservant la convexité

— Somme de fonctions convexes.

— Multiplication par un scalaire positif.

— Composition $g \circ f$ si f convexe et g convexe croissante.

— Maximum de fonctions convexes.

Optimisation Quadratique

$$f(x) = \frac{1}{2} \langle Ax, x \rangle + \langle b, x \rangle.$$

$\nabla f(x) = Ax + b$ (si A symétrique).

$$H_f(x) = A.$$

f convexe $\iff A$ positive.

— Si A définie positive, unique min global en $x = -A^{-1}b$.

Algorithmes Numériques

Méthode de Newton

Pour résoudre $\nabla f(x) = 0$. Convergence quadratique (rapide) mais locale et coûteuse (inversion de matrice).

$$x_{k+1} = x_k - [H_f(x_k)]^{-1} \nabla f(x_k)$$

Descente de Gradient

Pour minimiser f . On suit la direction opposée au gradient.

$$x_{k+1} = x_k - t_k \nabla f(x_k)$$

— Pas constant : $t_k = t$. Converge si $t < 2/L$ (L = constante de Lipschitz du gradient).

— Pas optimal : t_k minimise $\phi(t) = f(x_k - t \nabla f(x_k))$.

— Cas quadratique (A def. pos.) : Convergence globale garantie.

Méthode d'Uzawa

Pour optimisation quadratique sous contraintes $Hx \leq d$. Itération sur le multiplicateur dual μ .

$$\begin{cases} Ax_k + b - H^T \mu_k = 0 & (\text{Minimisation en } x) \\ \mu_{k+1} = P_{\mathbb{R}_-}(\mu_k - \rho(Hx_k - d)) & (\text{Projection}) \end{cases}$$

Converge si $0 < \rho < 2\lambda_{\min}(A)/\|H\|^2$.

Théorèmes Classiques

Fonctions Implicites

Si $g(x, y) = 0$ et $\frac{\partial g}{\partial y} \neq 0$, alors on peut exprimer localement $y = \varphi(x)$.

Général : Si la jacobienne partielle est inversible, on peut paramétriser la surface de contrainte.

Inversion Locale

Si f est C^1 et $D_x f$ est inversible en a , alors f est un difféomorphisme local autour de a .

Espace Tangent

Pour une sous-variété définie par $g(x) = 0$, l'espace tangent en x est l'orthogonal du gradient :

$$T_x \mathcal{S} = (\text{Vect}\{\nabla g_1(x), \dots, \nabla g_m(x)\})^\perp$$