Optimisation

Fabrice Rossi

Télécom ParisTech

Plan

Plan

Introduction

TELECOM ParisTech

Régression linéaire multiple

- Expliquer y par une combinaison linéaire des $(x_j)_{1 \leq j \leq p}$

$$\hat{y} = \sum_{j=1}^{n} a_j x_j = \langle \mathbf{a}, \mathbf{x} \rangle$$

- Comment choisir les a_j ?

Régression linéaire multiple

- Expliquer y par une combinaison linéaire des $(x_j)_{1\leq j\leq p}$

$$\hat{y} = \sum_{j=1}^{n} a_j x_j = \langle \mathbf{a}, \mathbf{x} \rangle$$

- Comment choisir les a_j ?
- Minimisation de l'erreur quadratique sur un ensemble d'exemples $(y_i, \mathbf{x}_i)_{1\leq i\leq n}$

$$\begin{aligned}\mathbf{a}^* &= \arg\min_{\mathbf{a}} \sum_{i=1}^{n} (y_i - \langle \mathbf{a}, \mathbf{x}_i \rangle)^2 \\ &= \arg\min_{\mathbf{a}} \|\mathbf{y} - \mathbf{a}^T X\|_2^2,\end{aligned}$$

avec $X = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$

Régression contrainte

- Pas de contrôle de **a** :
  - explosion possible des coefficients
  - problème mal conditionné
  - trop de coefficients non nuls

Régression contrainte

- Pas de contrôle de **a** :
  - explosion possible des coefficients
  - problème mal conditionné
  - trop de coefficients non nuls
- Régression *ridge* :
  - contrainte sur $\|\mathbf{a}\|_2 = \sqrt{\sum_{j=1}^{p} a_j^2}$

$$\begin{array}{rl} \text{minimiser} & \|\mathbf{y} - \mathbf{a}^T X\|_2^2 \\ \text{avec} & \|\mathbf{a}\|_2 \leq \lambda \end{array}$$

Régression contrainte

- Pas de contrôle de **a** :
  - explosion possible des coefficients
  - problème mal conditionné
  - trop de coefficients non nuls
- Régression *ridge* :
  - contrainte sur $\|\mathbf{a}\|_2 = \sqrt{\sum_{j=1}^{p} a_j^2}$

$$\begin{array}{rl} \text{minimiser} & \|\mathbf{y} - \mathbf{a}^T X\|_2^2 \\ \text{avec} & \|\mathbf{a}\|_2 \leq \lambda \end{array}$$

- Régression parcimonieuse (*lasso*) :
  - contrainte sur $\|\mathbf{a}\|_1 = \sum_{j=1}^{p} |a_j|$

$$\begin{array}{rl} \text{minimiser} & \|\mathbf{y} - \mathbf{a}^T X\|_2^2 \\ \text{avec} & \|\mathbf{a}\|_1 \leq \lambda \end{array}$$

Modèle de Markowitz en finance

- Objectif : investissement dans un portefeuille d'actifs avec minimisation du risque sous contrainte de rendement
- Modèle :
  - r_i rendement de l'actif i
  - évolution stochastique : $\mathbf{r}$ est de moyenne μ et de covariance Σ
  - proportion des actifs dans le portefeuille, $\mathbf{x}$ ($\sum_i x_i = 1 = \mathbf{1}^T\mathbf{x}$)
  - espérance de rendement : $\mu^T\mathbf{x}$
  - variance du rendement : $\mathbf{x}^T\Sigma\mathbf{x}$

Modèle de Markowitz en finance

- Objectif : investissement dans un portefeuille d'actifs avec minimisation du risque sous contrainte de rendement
- Modèle :
  - r_i rendement de l'actif i
  - évolution stochastique : $\mathbf{r}$ est de moyenne μ et de covariance Σ
  - proportion des actifs dans le portefeuille, $\mathbf{x}$ ($\sum_i x_i = 1 = \mathbf{1}^T\mathbf{x}$)
  - espérance de rendement : $\mu^T\mathbf{x}$
  - variance du rendement : $\mathbf{x}^T\Sigma\mathbf{x}$
- Problème

$$\begin{array}{rl} \text{minimiser} & \mathbf{x}^T\Sigma\mathbf{x} \\ \text{avec} & \mathbf{1}^T\mathbf{x} = 1 \\ & \mu^T\mathbf{x} \geq r_{\min} \end{array}$$

Forme générale

- un problème d'optimisation $(\mathcal{P})$ est défini par

$$\begin{array}{rl} \text{minimiser sur } \mathbb{R}^n & J(\mathbf{x}) \\ \text{avec} & g_i(\mathbf{x}) \leq 0\,,\, 1 \leq i \leq p \\ & h_j(\mathbf{x}) = 0\,,\, 1 \leq j \leq q \end{array}$$

Forme générale

- un problème d'optimisation $(\mathcal{P})$ est défini par

$$\begin{array}{rl} \text{minimiser sur } \mathbb{R}^n & J(\mathbf{x}) \\ \text{avec} & g_i(\mathbf{x}) \leq 0\,, 1 \leq i \leq p \\ & h_j(\mathbf{x}) = 0\,, 1 \leq j \leq q \end{array}$$

- vocabulaire :
  - J (à valeur dans $\mathbb{R} \cup \{\infty\}$) est la **fonction de coût**, la **fonction objectif** ou encore le **critère**
  - les g_i sont les **contraintes d'inégalité**
  - les h_j sont les **contraintes d'égalité**
  - l'**ensemble des contraintes** est

$$\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^n | g_i(\mathbf{x}) \leq 0\,, 1 \leq i \leq p \text{ et } h_j(\mathbf{x}) = 0\,, 1 \leq j \leq q\}$$

ensemble des points admissibles ou **réalisables**

Optima locaux et globaux

- optimum local : meilleure valeur *localement* au sens de la métrique de l'espace et de l'ensemble des contraintes
- formellement :
  - $\mathcal{C} \subset X$ espace métrique et J de $\mathcal{C}$ dans $\mathbb{R}$
  - $x^* \in \mathcal{C}$ réalise un minimum local de J sur $\mathcal{C}$ ssi $\exists$ une boule ouverte B centrée en x^* telle que $\forall x \in B \cap \mathcal{C}$, $J(x) \geq J(x^*)$
  - inégalité *stricte* pour $x \neq x^* \rightarrow$ minimum local *strict*

Optima locaux et globaux

- optimum local : meilleure valeur *localement* au sens de la métrique de l'espace et de l'ensemble des contraintes
- formellement :
  - $\mathcal{C} \subset X$ espace métrique et J de $\mathcal{C}$ dans $\mathbb{R}$
  - $x^* \in \mathcal{C}$ réalise un minimum local de J sur $\mathcal{C}$ ssi $\exists$ une boule ouverte B centrée en x^* telle que $\forall x \in B \cap \mathcal{C}$, $J(x) \geq J(x^*)$
  - inégalité *stricte* pour $x \neq x^* \rightarrow$ minimum local *strict*
- $x^* \in \mathcal{C}$ réalise un minimum global de J sur $\mathcal{C}$ ssi $\forall x \in \mathcal{C}$, $J(x) \geq J(x^*)$

Optima locaux et globaux

- optimum local : meilleure valeur *localement* au sens de la métrique de l'espace et de l'ensemble des contraintes
- formellement :
  - $\mathcal{C} \subset X$ espace métrique et J de $\mathcal{C}$ dans $\mathbb{R}$
  - $x^* \in \mathcal{C}$ réalise un minimum local de J sur $\mathcal{C}$ ssi $\exists$ une boule ouverte B centrée en x^* telle que $\forall x \in B \cap \mathcal{C}$, $J(x) \geq J(x^*)$
  - inégalité *stricte* pour $x \neq x^* \rightarrow$ minimum local *strict*
- $x^* \in \mathcal{C}$ réalise un minimum global de J sur $\mathcal{C}$ ssi $\forall x \in \mathcal{C}$, $J(x) \geq J(x^*)$
- Propriété : minimum de $J \Leftrightarrow$ maximum de $-J$

État de l'art

Résolution informatique

- problème de moindres carrés
  - J est quadratique et il n'y a pas de contraintes
  - résolution facile (inverse ou pseudo-inverse de matrice)
- programmation linéaire :
  - $J(\mathbf{x}) = \mathbf{c}^T\mathbf{x}$ (ou affine)
  - contraintes linéaires (ou affines)
  - résolution relativement facile
- programmation convexe :
  - les fonctions J et g_i sont convexes et les h_i affines
  - résolution possible
- cas général :
  - résolution complète (c.-à-d. trouver un minimum **global**) très difficile (coût exponentiel en la taille du problème)
  - résolution locale envisageable

Plan

Outils fondamentaux

- deux outils mathématiques fondamentaux pour l'analyse des problèmes d'optimisation
- différentiabilité :
  - approximation linéaire locale
  - tendances locales, par exemple plus grande pente
  - caractérisation des optima
- convexité :
  - pour les ensembles : on peut se promener sur un segment
  - pour les fonctions : quand on se promène sur un segment, la fonction reste « sous le segment »
  - existence d'optima
- lien important entre les deux propriétés : une fonction convexe est minorée globalement par son approximation linéaire locale

Rappels

- espace vectoriel normé :
  - espace vectoriel
  - norme (définie positive, sous-additive et homogène)
- espace de Banach : espace vectoriel normé complet
- espace de Hilbert :
  - espace vectoriel
  - produit scalaire (bilinéaire, symétrique, défini positif)
  - complet

Différentiabilité

Ordre un

- J est définie sur ouvert U d'un Banach X et à valeurs dans $\mathbb{R}$
- J est **différentiable** (au sens de **Fréchet**) en $x \in U$ s'il existe une forme linéaire continue DJ_x telle que

$$\lim_{\|h\|_X \to 0} \frac{J(x+h) - J(x) - DJ_x(h)}{\|h\|_X} = 0,$$

c'est-à-dire $J(x+h) = J(x) + DJ_x(h) + o(\|h\|_X)$

Différentiabilité

Ordre un

- J est définie sur ouvert U d'un Banach X et à valeurs dans $\mathbb{R}$
- J est **différentiable** (au sens de **Fréchet**) en $x \in U$ s'il existe une forme linéaire continue DJ_x telle que

$$\lim_{\|h\|_X \to 0} \frac{J(x+h) - J(x) - DJ_x(h)}{\|h\|_X} = 0,$$

c'est-à-dire $J(x+h) = J(x) + DJ_x(h) + o(\|h\|_X)$
- si X est un espace de Hilbert, on définit le **gradient** de J en x comme l'élément de X, noté $\nabla J(x)$, tel que

$$DJ_x(h) = \langle \nabla J(x), h \rangle_X$$

Théorème de représentation de Riesz ($X' = X$)

- si $J(x) = \langle a, x \rangle_X$, $\nabla J(x) = a$
- dans $\mathbb{R}^n$, si $J(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$, alors $\nabla J(\mathbf{x}) = (A + A^T)\mathbf{x}$
- cas particulier, si $J(\mathbf{x}) = \|\mathbf{x}\|^2$, alors $\nabla J(\mathbf{x}) = 2\mathbf{x}$
- dans $\mathbb{R}$, $\nabla J(x) = J'(x)$
- dans $\mathbb{R}^n$,

$$\nabla J(\mathbf{x}) = \left(\frac{\partial J}{\partial x_1}(\mathbf{x}), \ldots, \frac{\partial J}{\partial x_n}(\mathbf{x})\right)^T$$

Différentielle au sens de Gâteaux

- différentielle au sens de Gâteaux : dérivée directionnelle linéaire continue
- J est définie sur un ouvert U d'un Banach X et à valeurs dans $\mathbb{R}$
- J est **Gâteaux-différentiable** en $x \in U$ ssi la dérivée directionnelle

$$J'(x,h) = \lim_{t \to 0^+} \frac{J(x+th) - J(x)}{t},$$

existe pour tout $h \neq 0$ et si $h \mapsto J'(x,h)$ est une forme linéaire continue

Propriétés

- linéaire n'est pas automatique : $J(u,v) = \frac{u^2v^2}{(u^2+v^2)^{\frac{3}{2}}}$ en $(0,0)$ avec $J(u,v) = 0$
- la notion de gradient s'applique aussi dans le cas Gâteaux-différentiable : $J'(x,h) = \langle \nabla J(x), h \rangle_X$ (pour X un Hilbert)
- Fréchet différentiable implique Gâteaux différentiable
- réciproque fausse :
  - $J(x,y) = \frac{x^6}{(y-x^2)^2+x^8}$ avec $J(0,0) = 0$
  - J n'est pas continue en $(0,0)$, mais est Gâteaux différentiable
  - principe de l'exemple : approche linéaire ok, mais approche « curviligne » discontinue

Deuxième ordre

- différentielle : approximation linéaire locale
- différentielle à l'ordre deux : approximation quadratique locale
- J est définie sur un ouvert U d'un Banach X et à valeurs dans $\mathbb{R}$
- J est différentiable à l'ordre deux en $x \in U$ si elle est différentiable dans un voisinage de x et si $u \mapsto DJ_u$ est différentiable en x
- Remarque :
  - DJ_u est à valeurs dans X' le dual (topologique) de X
  - différentielle d'une fonction à valeurs dans un Banach V : il suffit de considérer $\|J(x+h) - J(x) - DJ_x(h)\|_V$ (qui doit être $o(\|h\|_X)$)
  - Gâteaux ou Fréchet

Deuxième ordre

- plus simplement, J est Fréchet différentiable à l'ordre deux en $x \in U$ ssi il existe une application linéaire continue DJ_x et une application bilinéaire symétrique continue D^2J_x telles que

$$J(x+h) = J(x) + DJ_x(h) + \frac{1}{2}D^2J_x(h,h) + o(\|h\|_X^2)$$

- dans un Hilbert
  - $DJ_x(h) = \langle \nabla J(x), h\rangle_X$
  - $D^2J_x(h,h) = \left\langle \nabla^2 J(x)(h), h\right\rangle_X$ pour un endomorphisme symétrique continu $\nabla^2 J(x)$ (appelé le **Hessien** par abus de langage)

Deuxième ordre

Exemples

- dans $\mathbb{R}^n$, la matrice Hessienne est donnée par

$$\nabla^2 J = \begin{pmatrix} \frac{\partial^2 J}{\partial x_1^2} & \frac{\partial^2 J}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 J}{\partial x_1 \partial x_n} \\ \frac{\partial^2 J}{\partial x_2 \partial x_1} & \frac{\partial^2 J}{\partial x_2^2} & \cdots & \frac{\partial^2 J}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 J}{\partial x_n \partial x_1} & \frac{\partial^2 J}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 J}{\partial x_n^2} \end{pmatrix}$$

- dans $\mathbb{R}^n$, si $J(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$, alors $\nabla^2 J(\mathbf{x}) = (A + A^T)$

Convexité

- convexe : contient les segments entre ses points
- formellement, $\mathcal{C} \subset X$ (un espace vectoriel) est **convexe** ssi

$$\forall x, y \in \mathcal{C},\ \forall t \in [0, 1],\ (1 - t)x + ty \in \mathcal{C}$$

Convexité

- convexe : contient les segments entre ses points
- formellement, $\mathcal{C} \subset X$ (un espace vectoriel) est **convexe** ssi

$$\forall x, y \in \mathcal{C},\ \forall t \in [0, 1],\ (1 - t)x + ty \in \mathcal{C}$$

- une fonction J de $\mathcal{C} \subset X$ dans $\mathbb{R} \cup \{\infty\}$ est **convexe** ssi :
  - $\mathcal{C}$ est convexe et,
  - $\forall x, y \in \mathcal{C},\ \forall t \in [0, 1]$,

$$J((1 - t)x + ty) \leq (1 - t)J(x) + tJ(y)$$

  - convexité **stricte** si quand $x \neq y$ et $\forall t \in]0, 1[$,

$$J((1 - t)x + ty) < (1 - t)J(x) + tJ(y)$$

- J est (strictement) concave si $-J$ est (strictement) convexe

Exemples

- remarque : J est convexe ssi toutes les fonctions $g(\lambda) = J(x + \lambda h)$ sont convexes
- les fonctions affines $J(\mathbf{x}) = \mathbf{a}^T\mathbf{x} + b$ sont convexes
- sur $\mathbb{R}^n$:
  - toute norme est convexe (inégalité triangulaire + homogénéité)
  - $J(\mathbf{x}) = \max_i x_i$ est convexe
  - $|x|$ est convexe
  - que dire de $J(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} + \mathbf{b}^T\mathbf{x} + c$ avec A symétrique ?

Propriétés

- domaine de définition d'une fonction convexe

$$\text{dom}(J) = \{x \in \mathcal{C} | J(x) < \infty\}$$

- J est dite **propre** quand $\text{dom}(J) \neq \emptyset$
- si $\mathcal{C} \subset \mathbb{R}^n$, une fonction convexe propre est continue sur l'intérieur de son domaine
  - dans le cas Banach, il faut supposer que J est majorée et non identiquement égale à $-\infty$ sur un ouvert (non vide) inclus dans $\text{dom}(J)$ pour obtenir la continuité

Caractérisation par les différentielles

- soit J de $\mathcal{C} \subset H$ (Hilbert) dans $\mathbb{R}$ et Gâteaux différentiable sur $\mathcal{C}$ un convexe, J est convexe ssi

$$\forall x, y \in \mathcal{C},\ J(y) \geq J(x) + \langle \nabla J(x), y - x \rangle_H$$

- en d'autres termes : l'approximation linéaire locale est un minorant global

Caractérisation par les différentielles

- soit J de $\mathcal{C} \subset H$ (Hilbert) dans $\mathbb{R}$ et Gâteaux différentiable sur $\mathcal{C}$ un convexe, J est convexe ssi

$$\forall x, y \in \mathcal{C},\ J(y) \geq J(x) + \langle \nabla J(x), y - x \rangle_H$$

- en d'autres termes : l'approximation linéaire locale est un minorant global
- preuve :
  - $\Rightarrow$ passage à la limite sur t positif de

$$\frac{J\left(x + t(y - x)\right) - J(x)}{t} \leq J(x) - J(y)$$

  - $\Leftarrow$ combinaison de deux applications de la minoration

$$\begin{aligned} J(x) &\geq J(x + t(y - x)) - t \left\langle \nabla J(x + t(y - x)), y - x \right\rangle_H \\ J(y) &\geq J(x + t(y - x)) + (1 - t) \left\langle \nabla J(x + t(y - x)), y - x \right\rangle_H \end{aligned}$$

Caractérisation par les différentielles

- autre caractérisation, J est convexe ssi ∇J est un opérateur monotone

$$\forall x, y \in \mathcal{C},\ \langle \nabla J(y) - \nabla J(x), y - x \rangle_H \geq 0$$

- remarque : la convexité stricte est impliquée par une version stricte des conditions (minoration globale et monotonie)
- ordre 2 : si J est C^2, elle est convexe ssi $\nabla^2 J(x)$ est positive en tout $x \in \mathcal{C}$

Exemples

- que dire de $J(x) = \frac{1}{x^2}$?
- maximum « assoupli » : $J(\mathbf{x}) = \log\left(\sum_{i=1}^{n} \exp x_i\right)$
- $J(A) = \log\det A$ est concave sur l'ensemble des matrices symétriques définies positives

Plan

Introduction

Outils mathématiques

- Différentiabilité
- Convexité

Résultats théoriques d'optimisation sans contrainte

- Existence et unicité d'un minimum
- Conditions d'optimalité

Algorithmes

- Fonctions définies sur $\mathbb{R}$
- Fonctions définies sur $\mathbb{R}^n$

Existence d'un minimum

- on s'intéresse à $(\mathcal{P})\ \min_{\mathbf{x}\in\mathbb{R}^n} J(x)$: pas de contrainte
- remarque : $\inf_{\mathbf{x}\in\mathbb{R}^n} J(x)$ est toujours bien défini, le problème est de réaliser un minimum

Existence d'un minimum

- on s'intéresse à $(\mathcal{P})$ $\min_{\mathbf{x}\in\mathbb{R}^n} J(x)$: pas de contrainte
- remarque : $\inf_{\mathbf{x}\in\mathbb{R}^n} J(x)$ est toujours bien défini, le problème est de réaliser un minimum
- J est à valeurs dans $\mathbb{R} \cup \{+\infty\}$
- J est coercitive (*coercive*) si $\lim_{\|x\|\to\infty} J(x) = +\infty$

Existence d'un minimum

- on s'intéresse à $(\mathcal{P})$ $\min_{\mathbf{x}\in\mathbb{R}^n} J(x)$: pas de contrainte
- remarque : $\inf_{\mathbf{x}\in\mathbb{R}^n} J(x)$ est toujours bien défini, le problème est de réaliser un minimum
- J est à valeurs dans $\mathbb{R} \cup \{+\infty\}$
- J est coercitive (*coercive*) si $\lim_{\|x\|\to\infty} J(x) = +\infty$
- Résultat : si J est propre (non identiquement infinie), continue et coercitive alors $(\mathcal{P})$ a au moins une solution

Unicité

- En général, on a des minima, pas un minimum
- Cas particulier : si J est **strictement** convexe, alors $(\mathcal{P})$ a au plus une solution

Unicité

- En général, on a des minima, pas un minimum
- Cas particulier : si J est **strictement** convexe, alors $(\mathcal{P})$ a au plus une solution
- Preuve :
  - soit d un minimum atteint en x et y
  - si $x \neq y$, par convexité stricte, on a

$$J\left(\frac{1}{2}x + \frac{1}{2}y\right) < \frac{1}{2}J(x) + \frac{1}{2}J(y) = d,$$

  ce qui est impossible, puisque d est le minimum de J

Unicité

- En général, on a des minima, pas un minimum
- Cas particulier : si J est **strictement** convexe, alors $(\mathcal{P})$ a au plus une solution
- Preuve :
  - soit d un minimum atteint en x et y
  - si $x \neq y$, par convexité stricte, on a

$$J\left(\frac{1}{2}x + \frac{1}{2}y\right) < \frac{1}{2}J(x) + \frac{1}{2}J(y) = d,$$

  ce qui est impossible, puisque d est le minimum de J
- Remarque :
  - un minimum local d'une fonction convexe est global
  - l'ensemble des points qui réalisent le minimum d'une fonction convexe est convexe

Existence et unicité

- Si J est strictement convexe et coercitive alors $(\mathcal{P})$ admet une solution unique
- Cas intéressant, les fonctions **elliptiques** (α-convexes ou fortement convexes) :
  - J est elliptique ssi existe une constante d'ellipticité $\alpha > 0$ telle que $\forall x, y, t \in [0, 1]$,

$$J((1-t)x + ty) \leq (1-t)J(x) + tJ(y) - \frac{\alpha}{2}t(1-t)\|x-y\|^2$$

  - elliptique $\Rightarrow$ strictement convexe et coercitive

Fonctions elliptiques

- Caractérisations à l'ordre 1 :
  - si J est différentiable au sens de Gâteaux, J est α-convexe ssi :
    - $\forall x, y,\ J(y) \geq J(x) + \langle \nabla J(x), y - x\rangle + \frac{\alpha}{2}\|x - y\|^2$
    - $\forall x, y,\ \langle \nabla J(y) - \nabla J(x), y - x\rangle \geq \alpha\|x - y\|^2$
- Caractérisation à l'ordre 2 :
  - si J est C^2, J est α-convexe ssi :

$$\forall x, h,\ \left\langle \nabla J^2(x)h, h\right\rangle \geq \alpha\|h\|^2$$

- Preuves dans le même esprit que pour la caractérisation des fonctions convexes
- Exemple : $J(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} + \mathbf{b}^T\mathbf{x} + c$ avec A symétrique

Condition d'optimalité du 1er ordre

- si J Gâteaux différentiable en x^* qui réalise un minimum local de J, alors $\nabla J(x^*) = 0$
- preuve :
  - pour t assez petit $J(x^*) \leq J(x^* + th)$ et donc $J'(x, h) \geq 0$
  - par linéarité, $\forall h$, $\langle \nabla J(x^*), h \rangle = 0$ et donc $\nabla J(x^*) = 0$
  - la linéarité est indispensable, cf $J(u, v) = \frac{u^2 v^2}{(u^2+v^2)^{\frac{3}{2}}}$
- ce n'est pas une condition suffisante, par ex. $J(x) = x^3$ en zéro
- vocabulaire :
  - $\nabla J(x^*) = 0$ est l'**équation d'Euler**
  - les x^* tels que $\nabla J(x^*) = 0$ sont les **points critiques** ou **stationnaires**

Cas convexe

- si J est Gâteaux différentiable et convexe alors x^* réalise un minimum global si et seulement si $\nabla J(x^*) = 0$
- preuve : une fonction convexe est minorée globalement par son approximation linéaire locale

Cas convexe

- si J est Gâteaux différentiable et convexe alors x^* réalise un minimum global si et seulement si $\nabla J(x^*) = 0$
- preuve : une fonction convexe est minorée globalement par son approximation linéaire locale
- extension :
  - J convexe n'est pas toujours différentiable, par ex. $J(x) = |x|$
  - un **sous-gradient** de J en x est un vecteur v tel que

$$\forall y,\ J(y) - J(x) \geq \langle v, y - x \rangle$$

  - la **sous-différentielle** de J en x est l'ensemble des sous-gradients, notée $\partial J(x)$ (dans $\mathbb{R}^n$, c'est un ensemble non vide, compact et convexe)
  - exemple $\partial |x|(0) = [-1, 1]$
  - CNS d'optimalité $0 \in \partial J(x^*)$

Condition d'optimalité du 2ème ordre

- si J est C^2 et si x^* réalise un minimum de J alors :
  - $\nabla J(x^*) = 0$
  - $\langle \nabla^2 J(x^*)h, h \rangle \geq 0$ pour tout h
- preuve par passage à la limite du développement limité à l'ordre 2
- ce n'est pas une condition suffisante, par ex. $J(x) = x^3$

Condition d'optimalité du 2ème ordre

- si J est C^2 et si x^* réalise un minimum de J alors :
  - $\nabla J(x^*) = 0$
  - $\left\langle \nabla^2 J(x^*)h, h \right\rangle \geq 0$ pour tout h
- preuve par passage à la limite du développement limité à l'ordre 2
- ce n'est pas une condition suffisante, par ex. $J(x) = x^3$
- si J est C^2 et si
  - $\nabla J(x^*) = 0$
  - il existe $\alpha > 0$ tel que $\left\langle \nabla^2 J(x^*)h, h \right\rangle \geq \alpha \|h\|^2$ ($\nabla^2 J(x)$ est défini positif)
  - alors x^* réalise un minimum local (strict) de J
- preuve identique
- **ellipticité locale**
- ce n'est pas une condition nécessaire, par ex. $J(x) = x^4$

Plan

Principes

- algorithmes numériques :
  - pas de résultat exact
  - notion de convergence
  - coût par itération
- hypothèses algorithmiques :
  - fonctions plus ou moins régulières
  - compromis coût d'une itération et nombre d'itérations
- hypothèses mathématiques :
  - pour garantir la convergence
  - convexité plus ou moins forte

Section dorée

- minimisation de f de $\mathbb{R}$ dans $\mathbb{R}$
- outil de base : permet de construire des algorithmes pour J définie sur $\mathbb{R}^n$
- méthode de la section dorée :
  - trois points initiaux tels que $f(x_1) > f(x_2)$ et $f(x_3) > f(x_2)$
  - on cherche le minimum dans $]x_1, x_3[$
  - on évalue f en x_4 choisit dans le plus grand de deux intervalles déterminés par x_2 par exemple $[x_1, x_2]$
  - on fonction de $f(x_4)$:
    - on passe dans $[x_4, x_3]$ si $f(x_4) > f(x_2)$
    - sinon on passe dans $[x_1, x_2]$
  - réduction optimale de la longueur de l'intervalle de recherche si le grand sous intervalle est $\frac{1+\sqrt{5}}{2}$ plus long que le petit (nombre d'or)

Méthode de Newton

- méthode de recherche d'un zéro d'une fonction f :
  - approximation linéaire de f

$$f(x+h) = f(x) + hf'(x) + o(h)$$

  - au premier ordre, $f(x+h) = 0$ conduit à $h = -\frac{f(x)}{f'(x)}$
  - algorithme : $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$
  - convergence quadratique : $|x_{n+1} - x^*| \leq \mu |x_{n+1} - x^*|^2$
- application à la minimisation de f :
  - on cherche à résoudre l'équation d'Euler $f'(x) = 0$
  - algorithme

$$x_{n+1} = x_n - \frac{f'(x_n)}{f''(x_n)}$$

Algorithme de relaxation

- méthode naïve : optimisations successives par rapport à chaque variable
- algorithme :
  1. point initial $\mathbf{x}^0$ (avec $J(\mathbf{x}^0) < \infty$)
  2. pour $k \geq 1$ croissant
     2.1 pour i allant de 1 à n :

$$x_i^{k+1} = \arg\min_{x \in \mathbb{R}} J(x_1^{k+1}, \ldots, x_{i-1}^{k+1}, x, x_{i+1}^k, \ldots, x_n^k)$$

     2.2 tester la convergence $\|\mathbf{x}^{k+1} - \mathbf{x}^k\| < \epsilon$
- remarque : calcul du minimum dans $\mathbb{R}$ par un algorithme adapté
- Attention : très mauvais algorithme en pratique (mais analysable théoriquement...)

Convergence

- on suppose :
  - J coercive,
  - J C^1 et α-convexe,
  - ∇J M-Lipschitzienne :

$$\forall \mathbf{x}, \mathbf{y}\ \|\nabla J(\mathbf{x}) - \nabla J(\mathbf{y})\| \leq M\|\mathbf{x} - \mathbf{y}\|$$

- alors l'algorithme de relaxation converge vers le minimum de J

Méthodes de descentes

- principe général :
  1. point initial $\mathbf{x}_0$
  2. pour $k \geq 1$ croissant
     2.1 choisir une **direction de descente** $\mathbf{d}_k \neq \mathbf{0}$
     2.2 choisir un **pas de descente** $\rho_k > 0$
     2.3 poser $\mathbf{x}_{k+1} = \mathbf{x}_k + \rho_k \mathbf{d}_k$
     2.4 tester la convergence (par ex. $\|\mathbf{x}_{k+1} - \mathbf{x}_k\| < \epsilon$)

Méthodes de descentes

- principe général :
  1. point initial $\mathbf{x}_0$
  2. pour $k \geq 1$ croissant
     - 2.1 choisir une **direction de descente** $\mathbf{d}_k \neq \mathbf{0}$
     - 2.2 choisir un **pas de descente** $\rho_k > 0$
     - 2.3 poser $\mathbf{x}_{k+1} = \mathbf{x}_k + \rho_k \mathbf{d}_k$
     - 2.4 tester la convergence (par ex. $\|\mathbf{x}_{k+1} - \mathbf{x}_k\| < \epsilon$)
- il faut qu'on puisse descendre :
  - on doit pouvoir trouver ρ_k tel que $J(\mathbf{x}_{k+1}) < J(\mathbf{x}_k)$
  - si J est convexe, $J(\mathbf{x}_{k+1}) \geq J(\mathbf{x}_k) + \rho_k \langle v, \mathbf{d}_k \rangle$ pour tout sous-gradient v : on doit donc avoir $\langle v, \mathbf{d}_k \rangle < 0$ pour au moins un sous-gradient v

Méthodes de descentes

- principe général :
  1. point initial $\mathbf{x}_0$
  2. pour $k \geq 1$ croissant
     2.1 choisir une **direction de descente** $\mathbf{d}_k \neq \mathbf{0}$
     2.2 choisir un **pas de descente** $\rho_k > 0$
     2.3 poser $\mathbf{x}_{k+1} = \mathbf{x}_k + \rho_k \mathbf{d}_k$
     2.4 tester la convergence (par ex. $\|\mathbf{x}_{k+1} - \mathbf{x}_k\| < \epsilon$)
- il faut qu'on puisse descendre :
  - on doit pouvoir trouver ρ_k tel que $J(\mathbf{x}_{k+1}) < J(\mathbf{x}_k)$
  - si J est convexe, $J(\mathbf{x}_{k+1}) \geq J(\mathbf{x}_k) + \rho_k \langle v, \mathbf{d}_k \rangle$ pour tout sous-gradient v : on doit donc avoir $\langle v, \mathbf{d}_k \rangle < 0$ pour au moins un sous-gradient v
- solution classique :
  - algorithme du **gradient**
  - $\mathbf{d}_k = -\nabla J(\mathbf{x}_k)$

Convergence

- L'algorithme du gradient converge sous des conditions assez fortes :
  - J est C^1, coercitive et strictement convexe
  - ∇J est M-Lipschitzienne
  - si $\rho_k \in [\beta_1, \beta_2]$ avec $0 < \beta_1 < \beta_2 < \frac{2}{M}$
- On peut obtenir une preuve plus simple en supposant que J est α-convexe C^1 et de gradient M-Lipschitzien :
  - on suppose alors que $\rho_k \in]0, \frac{2\alpha}{M^2}[$
  - on obtient une vitesse de convergence linéaire

$$\|\mathbf{x}_{k+1} - \mathbf{x}^*\| \leq \sqrt{1 - 2\alpha\rho_k + M^2\rho_k^2}\|\mathbf{x}_k - \mathbf{x}^*\|,$$

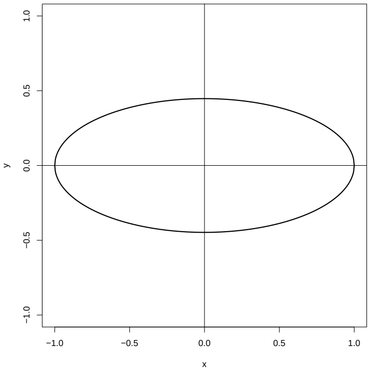
optimale de valeur $\sqrt{1 - \frac{\alpha^2}{M^2}}$ (pour $\rho_k = \frac{\alpha}{M^2}$)

Choix du pas

- pas constant
- pas variable :
  - adaptation du pas au problème
  - pas optimal :
    - $\rho_k = \arg\min_{\rho>0} J\left(\mathbf{x}_k - \rho\nabla J(\mathbf{x}_k)\right)$
    - avantage : meilleure réduction possible par itération
    - inconvénient : coût de la recherche
  - pas approximativement optimal :
    - rechercher un ρ_k qui réduit « assez » J
    - comme $J\left(\mathbf{x}_k - \rho\nabla J(\mathbf{x}_k)\right) = J(\mathbf{x}_k) - \rho\|\nabla J(\mathbf{x}_k)\|^2 + o(\rho\|\nabla J(\mathbf{x}_k)\|)$, on peut demander une réduction d'au moins $\alpha\rho\|\nabla J(\mathbf{x}_k)\|^2$ (avec $\alpha < \frac{1}{2}$)

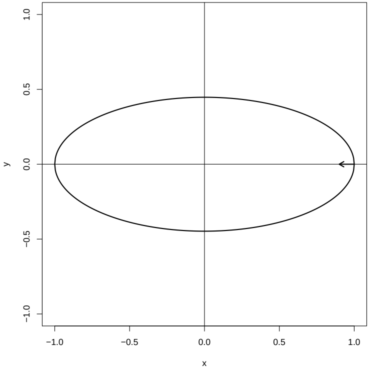
Gradient conjugué

- Analyse de l'algorithme du gradient sur le cas particulier des formes quadratiques $J(\mathbf{x}) = \frac{1}{2}\langle \mathbf{x}, A\mathbf{x}\rangle - \langle \mathbf{b}, \mathbf{x}\rangle$ pour A définie positive
- si A est mal conditionnée, l'algorithme du gradient est mauvais



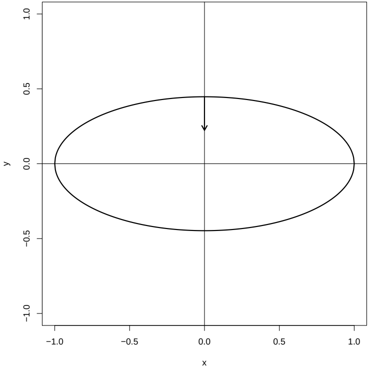
Gradient conjugué

- Analyse de l'algorithme du gradient sur le cas particulier des formes quadratiques $J(\mathbf{x}) = \frac{1}{2}\langle \mathbf{x}, A\mathbf{x}\rangle - \langle \mathbf{b}, \mathbf{x}\rangle$ pour A définie positive
- si A est mal conditionnée, l'algorithme du gradient est mauvais



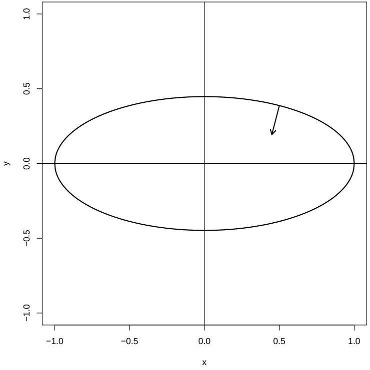
Gradient conjugué

- Analyse de l'algorithme du gradient sur le cas particulier des formes quadratiques $J(\mathbf{x}) = \frac{1}{2}\langle \mathbf{x}, A\mathbf{x}\rangle - \langle \mathbf{b}, \mathbf{x}\rangle$ pour A définie positive
- si A est mal conditionnée, l'algorithme du gradient est mauvais



Gradient conjugué

- Analyse de l'algorithme du gradient sur le cas particulier des formes quadratiques $J(\mathbf{x}) = \frac{1}{2}\langle \mathbf{x}, A\mathbf{x}\rangle - \langle \mathbf{b}, \mathbf{x}\rangle$ pour A définie positive
- si A est mal conditionnée, l'algorithme du gradient est mauvais



Méthode du gradient conjugué

- idée fondamentale : utiliser des directions de descentes conjuguées c.-à-d. orthogonales au sens de A
- autrement dit : ne pas perturber l'optimisation précédente
- principe général :
  - algorithme de descente à pas optimal
  - directions de descente obtenues par récurrence à partir des gradients aux points précédents
- applicable à des fonctions quelconques (pas seulement quadratiques)

Algorithme

1. point initial $\mathbf{x}_0$, gradient initial $\mathbf{g}_0 = A\mathbf{x}_0 - b$, direction de descente initiale $\mathbf{w}_0 = \mathbf{g}_0$ (on suppose que $\mathbf{g}_0 \neq \mathbf{0}$)
2. pour $k \geq 1$ croissant
   2.1 $\rho_{k-1} = \frac{\langle \mathbf{g}_{k-1}, \mathbf{w}_{k-1} \rangle}{\langle A\mathbf{w}_{k-1}, \mathbf{w}_{k-1} \rangle}$
   2.2 $\mathbf{x}_k = \mathbf{x}_{k-1} - \rho_{k-1}\mathbf{w}_{k-1}$
   2.3 $\mathbf{g}_k = A\mathbf{x}_k - b$
   2.4 $\mathbf{g}_k = \mathbf{0}$ fin de l'algorithme
   2.5 sinon :
      2.5.1 $\alpha_k = -\frac{\langle \mathbf{g}_k, A\mathbf{w}_{k-1} \rangle}{\langle A\mathbf{w}_{k-1}, \mathbf{w}_{k-1} \rangle}$
      2.5.2 $\mathbf{w}_k = \mathbf{g}_k + \alpha_k \mathbf{w}_{k-1}$

Algorithme

1. point initial $\mathbf{x}_0$, gradient initial $\mathbf{g}_0 = A\mathbf{x}_0 - b$, direction de descente initiale $\mathbf{w}_0 = \mathbf{g}_0$ (on suppose que $\mathbf{g}_0 \neq \mathbf{0}$)
2. pour $k \geq 1$ croissant
   - 2.1 $\rho_{k-1} = \frac{\langle \mathbf{g}_{k-1}, \mathbf{w}_{k-1} \rangle}{\langle A\mathbf{w}_{k-1}, \mathbf{w}_{k-1} \rangle}$
   - 2.2 $\mathbf{x}_k = \mathbf{x}_{k-1} - \rho_{k-1}\mathbf{w}_{k-1}$
   - 2.3 $\mathbf{g}_k = A\mathbf{x}_k - b$
   - 2.4 $\mathbf{g}_k = \mathbf{0}$ fin de l'algorithme
   - 2.5 sinon :
     - 2.5.1 $\alpha_k = -\frac{\langle \mathbf{g}_k, A\mathbf{w}_{k-1} \rangle}{\langle A\mathbf{w}_{k-1}, \mathbf{w}_{k-1} \rangle}$
     - 2.5.2 $\mathbf{w}_k = \mathbf{g}_k + \alpha_k \mathbf{w}_{k-1}$

- ρ_k est un pas optimal
- α_k est tel que $\langle \mathbf{w}_k, A\mathbf{w}_{k-1} \rangle = 0$

Propriétés

- pour tout $l < k$:
  - $\langle \nabla J(\mathbf{x}_k), \nabla J(\mathbf{x}_l) \rangle = 0$
  - $\langle \nabla J(\mathbf{x}_k), \mathbf{w}_l \rangle = 0$
  - $\langle \mathbf{w}_k, A\mathbf{w}_l \rangle = 0$
- l'algorithme trouve le minimum de $J(\mathbf{x}) = \frac{1}{2} \langle \mathbf{x}, A\mathbf{x} \rangle - \langle \mathbf{b}, \mathbf{x} \rangle$ en au plus n itérations (n est l'ordre de A)
- erreurs numériques : critère d'arrêt sur $\|\nabla J(\mathbf{x}_k)\|$

Cas non linéaire

1. point initial $\mathbf{x}_0$, gradient initial $\mathbf{g}_0 = \nabla J(\mathbf{x}_0)$, direction de descente initiale $\mathbf{w}_0 = \mathbf{g}_0$ (on suppose que $\mathbf{g}_0 \neq \mathbf{0}$)
2. pour $k \geq 1$ croissant
   - 2.1 $\rho_{k-1} = \arg\min_{\rho>0} J(\mathbf{x}_{k-1} - \rho\mathbf{w}_{k-1})$
   - 2.2 $\mathbf{x}_k = \mathbf{x}_{k-1} - \rho_{k-1}\mathbf{w}_{k-1}$
   - 2.3 $\mathbf{g}_k = \nabla J(\mathbf{x}_k)$
   - 2.4 $\mathbf{g}_k = \mathbf{0}$ fin de l'algorithme
   - 2.5 sinon :
     - 2.5.1 choisir un α_k
     - 2.5.2 $\mathbf{w}_k = \mathbf{g}_k + \alpha_k\mathbf{w}_{k-1}$

- Diverses formules pour le calcul de α_k, par exemple $\alpha_k = \frac{\|\mathbf{g}_k\|^2}{\|\mathbf{g}_{k-1}\|^2}$ (Fletcher-Reeves)
- « Redémarrage » : $\mathbf{w}_k = \mathbf{g}_k$ de temps en temps

TELECOM ParisTech

Méthode de Newton

- on cherche un zéro de ∇J, or

$$\nabla J(\mathbf{x}+\mathbf{h}) = \nabla J(\mathbf{x}) + \nabla^2 J(\mathbf{x})\mathbf{h} + o(\|\mathbf{h}\|)$$

  donc au premier ordre

$$\mathbf{h} = -\nabla^2 J(\mathbf{x})^{-1} \nabla J(\mathbf{x})$$

- correspond aussi au minimum de l'approximation au deuxième ordre

$$J(\mathbf{x}+\mathbf{h}) = J(\mathbf{x}) + \mathbf{h}^T \nabla J(\mathbf{x}) + \frac{1}{2}\mathbf{h}^T \nabla^2 J(\mathbf{x})\mathbf{h} + o(\|\mathbf{h}\|^2)$$

- si $\nabla^2 J(\mathbf{x})$ est définie positive, $\mathbf{h}$ est une direction de descente

$$\left\langle \nabla J(\mathbf{x}), -\nabla^2 J(\mathbf{x})^{-1} \nabla J(\mathbf{x}) \right\rangle < 0$$

Algorithmes

- Newton pur :
  1. point initial $\mathbf{x}_0$
  2. pour $k \geq 1$ croissant
     - 2.1 calculer $\mathbf{d}_k = -\nabla^2 J(\mathbf{x_k})^{-1} \nabla J(\mathbf{x_k})$
     - 2.2 tester la convergence par $\langle \nabla J(\mathbf{x_k}), \nabla^2 J(\mathbf{x_k})^{-1} \nabla J(\mathbf{x_k}) \rangle < \epsilon$
     - 2.3 poser $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{d}_k$
- variante dite « gardée » :
  - mise à jour par $\mathbf{x}_{k+1} = \mathbf{x}_k + \rho_k \mathbf{d}_k$
  - recherche de $\rho_k \leq 1$ par réduction suffisante
- on peut montrer la convergence de l'algorithme gardé sous les hypothèses classiques (α-convexité) et si le Hessien est Lipschitzien

Résumé

- algorithme du gradient :
  - coût en $O(n)$
  - grand nombre d'itérations
- gradient conjugué :
  - coût en $O(n)$
  - nombre d'itérations beaucoup plus faible que le gradient simple
- méthode Newton :
  - coût en $O(n^3)$
  - nombre d'itérations faible (convergence quadratique près du minimum)
- solutions intermédiaires :
  - méthodes de quasi-Newton (par ex. BFGS)
  - approximation de $\nabla^2 J(\mathbf{x})^{-1}$
  - coût en $O(n^2)$ ou même $O(n)$
  - convergence rapide