

Optimisation

Formulaire pour le test du 18.04.07

1 Algorithmes itératifs

Problème usuel :	$(P) \quad \begin{array}{ll} \min & f(\underline{x}) \\ \text{s.c.} & \underline{x} \in S = \{\underline{x} \mid \underline{g}(\underline{x}) \geq 0\} \\ & \underline{x} \in \mathcal{A} \end{array}$
------------------	--

- notions :
 - algo itératif : engendre suite de pts $\underline{x}^0, \dots, \underline{x}^k \in \mathbb{R}^n$
 - algo de descente si $f(\underline{x}^{k+1}) < f(\underline{x}^k) \forall k$
 - algo globalement convergent si \forall suite $(\underline{x}^k)_{k \in \mathbb{N}}$ avec $\underline{x}^{k+1} \in A(\underline{x}^k)$, \exists sous-suite $(\underline{x}^k)_{k \in L}$ qui converge vers une solution de (P)
 - algo $A : X \rightarrow Y$ fermé si $\forall (\underline{x}^k)_{k \in \mathbb{N}} \subset X$ avec $\underline{x}^k \rightarrow \underline{x} \in X$, et $\forall (\underline{y}^k)_{k \in \mathbb{N}} \subset A(\underline{x}^k)$ telle que $\underline{y}^k \rightarrow \underline{y}$, on a $\underline{y} \in A(\underline{x})$
- définition d'une fonction de descente : $\Gamma \subseteq X$ ensemble de solutions de (P) satisfaisant une CN d'optimum, $A : X \rightarrow Y$ algo. Alors $Z : X \rightarrow \mathbb{R}$ continue est une fonction de descente de A pour l'ensemble Γ si
 - $\underline{x} \notin \Gamma, \underline{y} \in A(\underline{x}) \Rightarrow Z(\underline{y}) < Z(\underline{x})$,
 - $\underline{x} \in \Gamma, \underline{y} \in A(\underline{x}) \Rightarrow Z(\underline{y}) \leq Z(\underline{x})$.
 Si f est continue, alors elle peut servir de fonction de descente.
- théorème de convergence globale : Γ et A comme avant, $(\underline{x}^k)_{k \in \mathbb{N}}$ suite de points avec $\underline{x}^{k+1} \in A(\underline{x}^k)$, $\underline{x}^k \in K \subseteq X$ compact. Supposons qu'il existe une fonction de descente Z de A pour Γ , et que A est fermé hors de Γ . Alors toute limite d'une sous-suite de $(\underline{x}^k)_{k \in \mathbb{N}}$ est dans Γ .

1.1 Méthode du gradient

Le but est de trouver $\min\{f(\underline{x})\}$ avec $\underline{x} \in \mathbb{R}^n$. On suppose $f \in C^1$ et $f(\underline{x}) \rightarrow \infty$ si $|\underline{x}| \rightarrow \infty$. Algo :

1. choisir \underline{x}^0 et poser $k = 0$
2. chercher α_k tel que $f(\underline{x}^k - \alpha_k \nabla f(\underline{x}^k)) = \min_{\alpha \geq 0} f(\underline{x}^k - \alpha \nabla f(\underline{x}^k))$
3. Tant que $\nabla f(\underline{x}^k) \neq 0$, poser $\underline{x}^{k+1} = \underline{x}^k - \alpha_k \nabla f(\underline{x}^k)$, $k = k + 1$ et aller à 2.

1.2 Remarques sur la vitesse de convergence

Soit $\bar{\underline{x}}$ la limite d'une suite $(\underline{x}^k)_{k \in \mathbb{N}}$ engendrée par l'algo A.

- (i) si $\lim_{k \rightarrow \infty} \frac{|\underline{x}^{k+1} - \bar{\underline{x}}|}{|\underline{x}^k - \bar{\underline{x}}|} = \alpha < 1$, convergence linéaire de taux α .
- (ii) si cette limite vaut 0, convergence superlinéaire d'ordre γ avec $\gamma = \max_{\xi \geq 0} \left\{ \lim_{k \rightarrow \infty} \frac{|\underline{x}^{k+1} - \bar{\underline{x}}|}{|\underline{x}^k - \bar{\underline{x}}|^\xi} < \infty \right\}$.
- (iii) si $\lim_{k \rightarrow \infty} \frac{|\underline{x}^{k+1} - \bar{\underline{x}}|}{|\underline{x}^k - \bar{\underline{x}}|^2} < \infty$, convergence quadratique.

1.3 Méthode de Newton

On cherche un minimum de $f(x)$ avec $f : \mathbb{R} \rightarrow \mathbb{R}$ de classe C^2 . Méthode itérative :

1. choisir $x^0 \in \mathbb{R}$, $\varepsilon > 0$ et poser $k = 0$
2. Tant que $|f'(x^k)| > \varepsilon$, faire $x^{k+1} = x^k - \frac{f'(x^k)}{f''(x^k)}$ et poser $k = k + 1$.

Si $f : \mathbb{R}^n \rightarrow \mathbb{R}$, la condition devient $|\nabla f(x^k)| > \varepsilon$ et l'itération $\underline{x}^{k+1} = \underline{x}^k - H(\underline{x}^k)^{-1} \nabla f(\underline{x}^k)$.

1.4 Méthodes d'optimisation de fonctions à une variable

Une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ est unimodale sur l'intervalle $[a, b]$ si $\min_{x \in [a, b]} \{f(x)\} > -\infty$, et si quand \bar{x} est le min de f sur $[a, b]$, on a $\forall x_1, x_2 \in [a, b]$ avec $x_1 < x_2$,

- $x_1 \geq \bar{x} \Rightarrow f(x_1) < f(x_2)$,
- $x_2 \leq \bar{x} \Rightarrow f(x_1) > f(x_2)$.

Méthode itérative pour une minimiser une telle fonction :

1. Partir de $I_0 = [a_0, b_0]$ contenant \bar{x}
2. A partir de $I_k = [a_k, b_k] \ni \bar{x}$, calculer a_{k+1}, b_{k+1} avec $a_k \leq a_{k+1} \leq \bar{x} \leq b_{k+1} \leq b_k$ et $|I_{k+1}| = |b_{k+1} - a_{k+1}| < |I_k|$, $I_{k+1} \subset I_k$.
3. S'arrêter lorsque $|b_k - a_k| < \varepsilon$ donné.

A chaque itération on peut évaluer f en 2 ou 3 nouveaux points. Mieux : on ne prend qu'un seul nouveau point à chaque itération. Si $|I_0| = 1$, la méthode optimale consiste à choisir I_1 avec $|I_1| = 1 - \delta = \frac{\sqrt{5}-1}{2} = 0.618$.

2 Optimisation sans contraintes

Problème (P) $\min_{\underline{x} \in \mathbb{R}^n} \{f(\underline{x})\}$, avec $f \in C^1$ ou $f \in C^2$
--

- | | |
|-----------------|---|
| CN de min local | <ul style="list-style-type: none"> • $\nabla f(\underline{x}) = 0$ • $H(\underline{x})$ semi-définie positive |
| CS de min local | <ul style="list-style-type: none"> • $\nabla f(\underline{x}) = 0$ • $H(\underline{x})$ définie positive |

2.1 Méthodes de gradient

On veut trouver $\min_{\underline{x} \in \mathbb{R}^n} f(\underline{x})$, avec $f \in C^1$. Pour cela on a deux méthodes de gradient.

2.1.1 Méthode à pas prédéterminé (Polyak)

1. partir de \underline{x}^0 avec $\alpha_0, \alpha_1, \dots$ donnés à l'avance
2. tant que "arrêt = faux", poser $\underline{x}^{k+1} = \underline{x}^k - \alpha_k \frac{\nabla f(\underline{x}^k)}{|\nabla f(\underline{x}^k)|}$ et $k = k + 1$
3. arrêter quand "arrêt = vrai" (p.ex. $\max_i |\frac{\partial f}{\partial x^i}| < \varepsilon$, ou $|\nabla f(\underline{x}^k)| < \varepsilon$, ou encore $|f(\underline{x}^{k+1}) - f(\underline{x}^k)| < \varepsilon$).

Théorème de Polyak : Si $\alpha_k \rightarrow 0$ avec $\sum_{k=0}^{\infty} \alpha_k \rightarrow \infty$ et $f \in C^1$ convexe, $\underline{x}^k \in S$ compact $\forall k$, alors on a convergence vers un \underline{x}^* où $\nabla f(\underline{x}^*) = \underline{0}$.

2.1.2 Méthode de la plus grande pente

1. choisir \underline{x}^0 et poser $k = 0$,
2. tant que "arrêt = faux", chercher un $\alpha_k \geq 0$ tel que $f(\underline{x}^k - \alpha_k \nabla f(\underline{x}^k)) = \min_{\alpha \geq 0} f(\underline{x}^k - \alpha \nabla f(\underline{x}^k))$, puis poser $\underline{x}^{k+1} = \underline{x}^k - \alpha_k \nabla f(\underline{x}^k)$ et $k = k + 1$.

Théorème : Si $f \in C^1$ avec $f(\underline{x}) \rightarrow +\infty$ si $|\underline{x}| \rightarrow +\infty$ alors $\forall \underline{x}^0$ la méthode de la plus grande pente converge vers un point \underline{x}^* où $\nabla f(\underline{x}^*) = \underline{0}$.

2.2 Méthode de directions conjuguées

Définition : Soit Q symétrique, $\underline{d}^0, \dots, \underline{d}^k$ sont Q -conjuguées (Q -orthogonales) si $\underline{d}^i \neq 0 \forall i$ et si $\underline{d}^i Q \underline{d}^j = 0 \forall i \neq j$.

1. Soit $\underline{d}^0, \dots, \underline{d}^k$ directions Q -conjuguées ($\underline{d}^i \neq 0$)
 $\underline{x}^{k+1} = \underline{x}^k + \alpha_k \underline{d}^k$ où $\alpha_k = \frac{-\nabla f(\underline{x}^k) \underline{d}^k}{\underline{d}^k Q \underline{d}^k}$.
 $\nabla f(\underline{x}^k) = Q \underline{x}^k - \underline{b}$

2. suite \underline{x}^k engendrée par cet algorithme converge vers la solution unique \underline{x}^* de $Q\underline{x} = \underline{b}$ après n étapes ($\underline{x}^* = \underline{x}^n$)
 si $f(\underline{x}) = \alpha_0 \underline{d}^0 + \dots + \alpha_{n-1} \underline{d}^{n-1}$.
 où $\alpha_i = \frac{\underline{d}^i \underline{b}}{\underline{d}^i Q \underline{d}^i}$ et $\underline{x}^d = \sum_{i=0}^{n-1} \frac{\underline{d}^i \underline{b}}{\underline{d}^i Q \underline{d}^i} \underline{d}^i$.

Théorème du sous-espace croissant : B^k sous-espace engendré par les directions $\underline{d}^0, \dots, \underline{d}^{k-1}$ alors \underline{x}^k minimise $f(\underline{x}) = \frac{1}{2} \underline{x} Q \underline{x} - \underline{b} \underline{x}$ sur la variété linéaire et sur la droite $\underline{x} = \underline{x}^{k+1} + \alpha \underline{d}^{k+1}$.

2.3 Méthode du gradient conjugué (cas particulier $f(\underline{x}) = \frac{1}{2} \underline{x} Q \underline{x} - \underline{b} \underline{x}$)

- Partir de $\underline{x}^0 \in \mathbb{R}^n$ et prendre $\underline{d}^0 = -\underline{g}^0 = \underline{b} - Q\underline{x}^0$ où $\underline{g}^k = \nabla f(\underline{x}^k)$.
- $\underline{x}^{k+1} = \underline{x}^k + \alpha_k \underline{d}^k$ où $\alpha_k = \frac{-\underline{g}^k \underline{d}^k}{\underline{d}^k Q \underline{d}^k} = \frac{\underline{g}^k \underline{g}^k}{\underline{d}^k Q \underline{d}^k}$
 $\underline{d}^{k+1} = -\underline{g}^{k+1} + \beta_k \underline{d}^k$ où $\beta_k = \frac{\underline{g}^{k+1} Q \underline{d}^k}{\underline{d}^k Q \underline{d}^k} = \frac{\underline{g}^{k+1} \underline{g}^{k+1}}{\underline{g}^k \underline{g}^k}$
- arrêt quand la condition d'arrêt est satisfaite (par exemple : $\underline{g}^k = 0$).

2.4 Méthode du gradient conjugué (cas général : $f \in C^2$)

(Convergence global car on applique la méthode du gradient à chaque étape)

2.4.1 Algorithme 1

(ne converge pas forcément en n étapes)

- partir de \underline{x}^0 , choisir $\underline{d}^0 = -\underline{g}^0 = -\nabla f(\underline{x}^0)$
- pour $k=0,1,\dots,n-1$:
 - $\underline{x}^{k+1} = \underline{x}^k + \alpha_k \underline{d}^k$ où $\alpha_k = \frac{-\underline{g}^k \underline{d}^k}{\underline{d}^k H(\underline{x}^k) \underline{d}^k}$
 - $\underline{g}^{k+1} = \nabla f(\underline{x}^{k+1})$
 - si $k < n - 1$ alors $\underline{d}^{k+1} = -\underline{g}^{k+1} + \beta_k \underline{d}^k$ où $\beta_k = \frac{\underline{g}^{k+1} H(\underline{x}^k) \underline{d}^k}{\underline{d}^k H(\underline{x}^k) \underline{d}^k}$.
- $\underline{x}^0 = \underline{x}^n$ et faire 1. tant qu'on n'a pas de condition d'arrêt.

Avantage : pas d'optimisation monodimensionnelle.

Désavantage : calcul de $H(\underline{x}^k)$ à chaque étape.

2.4.2 Algorithme 2 (Fletcher - Reeves)

- partir de \underline{x}^0 , choisir $\underline{d}^0 = -\underline{g}^0 = -\nabla f(\underline{x}^0)$.
- pour $k=0,1,\dots,n-1$:
 - $\underline{x}^{k+1} = \underline{x}^k + \alpha_k \underline{d}^k$ (où $f(\underline{x}^k + \alpha_k \underline{d}^k) = \min_{\alpha \geq 0} f(\underline{x}^k + \alpha \underline{d}^k)$).
 - $\underline{g}^{k+1} = \nabla f(\underline{x}^{k+1})$
 - si $k < n - 1$, $\underline{d}^{k+1} = -\underline{g}^{k+1} + \beta_k \underline{d}^k$ où $\beta_k = \frac{\underline{g}^{k+1} \underline{g}^{k+1}}{\underline{g}^k \underline{g}^k}$
- $\underline{x}^0 = \underline{x}^n$ et aller à 1. tant qu'on n'a pas de condition d'arrêt.

Avantage : pas de calcul de $H(\underline{x}^k)$.

Désavantage : optimisation monodimensionnelle.

2.5 Méthode de Newton

$$\underline{x}^{k+1} = \underline{x}^k - H^{-1}(\underline{x}^k) \nabla f(\underline{x}^k) \quad f \in C^2$$

(convergence quadratique. Si f quadratique, convergence en 1 étape)

3 Optimisation avec contraintes

Problème : (P) $S = \{\underline{x} g(\underline{x}) \geq 0 \text{ et } h(\underline{x}) = 0\}$	$\begin{aligned} \min & f(\underline{x}) \\ \text{s.c.} & g(\underline{x}) \geq 0 \\ & h(\underline{x}) = 0 \\ & f, g, h \in C^2 \text{ au moins} \end{aligned}$
---	--

3.1 Méthodes des directions admissibles

1. Partir de $\underline{x}^k \in S$ (ensemble des solutions admissibles)
2. Ayant obtenu \underline{x}^k , on cherche une direction \underline{d}^k telle que :
 - (a) $\nabla f(\underline{x}^k)\underline{d}^k < 0$
 - (b) $\exists \epsilon > 0$ tel que $\underline{x}^k + \alpha \underline{d}^k \in S$ ($0 \leq \alpha \leq \epsilon$)

Déterminer α_k tel que $f(\underline{x}^k + \alpha_k \underline{d}^k) = \min_{0 \leq \alpha \leq \epsilon} f(\underline{x}^k + \alpha \underline{d}^k)$
 Poser $\underline{x}^{k+1} = \underline{x}^k + \alpha_k \underline{d}^k$

3. Répéter jusqu'à une condition d'arrêt.

3.1.1 Cas particulier : méthode de Zoutendijk

Les contraintes $g(\underline{x}) \geq 0$ deviennent $\underline{a}^i \underline{x} \geq b_i$ et on définit $I = \{\underline{a}^i \underline{x} = b_i\}$.
 Pour trouver une solution admissible en \underline{x} , on résout :

$$(P) \begin{cases} \min & f(\underline{x}) * d & (\underline{x} \text{ fixé}) \\ \text{s.c.} & \nabla g_i(\underline{x})\underline{d} \geq 0 & i \in I(\underline{x}) \\ & -\epsilon_i \leq d_i \leq \epsilon_i \\ & \sum_i \epsilon_i = 1 \end{cases}$$

3.1.2 Cas des contraintes quelconques

Hypothèse : en chaque \underline{x} , les gradients des contraintes actives sont linéairements indépendantes.
 Résoudre :

$$(P) \begin{cases} \min & \epsilon \\ \text{s.c.} & \nabla f(\underline{x})\underline{d} \geq \epsilon & (\underline{x} \text{ fixé}) \\ & g(\underline{x}) + \nabla g(\underline{x})\underline{d} \geq -\epsilon \underline{u} & \text{où } \underline{u} = (1, \dots, 1) \end{cases}$$

Si on trouve $\epsilon < 0$, on a une direction \underline{d} telle que $\nabla f(\underline{x})\underline{d} < 0$ et $g_i(\underline{x}) + \nabla g_i(\underline{x})\underline{d} > 0$.

3.2 Méthodes des plans sécants

Hypothèses : S est fermé et convexe, f est convexe (On peut toujours supposer f linéaire).

1. A l'étape k , on a un polyèdre P_k tel que $P_k \supseteq S$ et on résout le P.L suivant :

$$\begin{cases} \min & \underline{c}\underline{x} \\ \text{s.c.} & \underline{x} \in P_k \end{cases}$$

2. Soit \underline{x}^k l'optimum de 1. . On a 2 cas :
 - $\underline{x}^k \in S$, STOP.
 - $\underline{x}^k \notin S$, on cherche un hyperplan $H_k = \{\underline{x} | \underline{a}^k \underline{x} = b_k\}$ séparant \underline{x}^k et S , i.e $S \subseteq \{\underline{x} | \underline{a}^k \underline{x} \leq b_k\}$, $\underline{x}^k \in \{\underline{x} | \underline{a}^k \underline{x} > b_k\}$, $P_{k+1} = P_k \cap \{\underline{x} | \underline{a}^k \underline{x} \leq b_k\}$.

3.2.1 Algorithme de Kelley

Hypothèses : g_i concaves et donc $S = \{\underline{x} | g_i(\underline{x}) \geq 0, i = 1, \dots, m\}$ convexe.

1. Soit P polyèdre tel que $P \supseteq S$ où $\underline{c}\underline{x}$ est borné inférieurement, on résout le P.L suivant :

$$\begin{cases} \min & \underline{c}\underline{x} \\ \text{s.c.} & \underline{x} \in P \end{cases}$$

2. Soit \underline{x}^* l'optimum de 1. . On a 2 cas :
 - $\underline{x}^* \in S$, STOP.
 - $\underline{x}^* \notin S$, soit i tel que $g_i(\underline{x}^*) = \min_j g_j(\underline{x}^*) < 0$.
 $P' = P \cap \{\underline{x} | g_i(\underline{x}^*) + \nabla g_i(\underline{x}^*)(\underline{x} - \underline{x}^*)\}$ et recommencer en 1. avec $P = P'$.

3.3 Méthode des pénalités extérieures

Soit $\mu_k, k=1,2,\dots$ tel que $\mu_k \geq 0$ et $\mu_{k+1} > \mu_k \forall k$. Soit $D(\underline{x}, \mu_k) = f(\underline{x}) + \mu_k P(\underline{x})$ où $P(\underline{x})$ est la fonction de pénalité extérieure.

1. $k=1, \mu_1 < 0$
2. Pour chaque k , trouver $\underline{x}^k = \operatorname{argmin}_{\underline{x}} D(\underline{x}, \mu_k)$.
Tant que l'on a pas une condition d'arrêt, choisir μ_{k+1} (ex : $\mu_{k+1} = \alpha \mu_k$ avec $\alpha > 1$) et aller en 2.

3.4 Méthode de barrière (pénalité intérieure)

Hypothèses : S à un intérieur $S^0 \neq \emptyset$ "arbitrairement proche de tout point de S ". S est compact et f est supposée continue.

Soit $D(\underline{x}, \mu_k) = f(\underline{x}) + \frac{1}{\mu_k} B(\underline{x})$, où $B(\underline{x})$ est la fonction barrière.

1. $k=1, \mu_1 < 0$
2. Pour chaque k , trouver $\underline{x}^k = \operatorname{argmin}_{\underline{x} \in S^0} D(\underline{x}, \mu_k)$.
Tant que l'on a pas une condition d'arrêt, choisir μ_{k+1} (ex : $\mu_{k+1} = \alpha \mu_k$ avec $\alpha > 1$) et aller en 2.

3.5 Méthodes lagrangiennes

1. Partir de $\underline{\lambda}^0 \geq 0$
2. A l'étape k , calculer $w(\underline{\lambda}^k) = \min_{\underline{x} \in A} (f(\underline{x}) - \underline{\lambda}^k g(\underline{x}))$. Soit $\underline{x}^k = \operatorname{argmin}(w(\underline{\lambda}^k))$.
3. $-g(\underline{x}^k)$ est un sous gradient de $w(\lambda)$ en $\underline{\lambda}^k$. Soit $\underline{\lambda}^{k+1} = \operatorname{argmax}_{\lambda \geq 0} w(\lambda)$ (on le trouve à partir de $\underline{\lambda}^k$ en se déplaçant dans la direction de $-g(\underline{x}^k)$).
4. Si on n'a pas de condition d'arrêt $k=k+1$ et aller en 2.

3.5.1 Méthode de Danzig

1. A l'étape k , on a déjà introduit les contraintes associées aux points $\underline{x}^1, \dots, \underline{x}^k$ et on résoud le P.L suivant :

$$P(k) \left| \begin{array}{l} \max \quad z \\ \text{s.c.} \quad f(\underline{x}^i) - \gamma^i (x - \underline{x}^i) \geq z \quad i = 1, \dots, k \\ \gamma^i \geq 0 \end{array} \right.$$

2. Soit \underline{z}^{k+1} la solution optimale de $P(k)$, alors si :
 - si $\underline{z}^{k+1} < f(\underline{x}^{k+1})$, on cherche un sous gradient γ^{k+1} de f en \underline{x}^{k+1} et on rajoute la contrainte $z \leq f(\underline{x}^{k+1}) - \gamma^{k+1}(x - \underline{x}^{k+1})$ on résoud de nouveau le P.L en 1.
 - si $\underline{z}^{k+1} = f(\underline{x}^{k+1})$, STOP.