**Optimisation Différentiable**

**Théorie et Algorithmes**

**Partie II**

## Exemple de résumé du cours

J. Ch. GILBERT

5 janvier 2021

## Informations pratiques

- Objectif du cours : l'optimisation
  - aspects théoriques : pénalisation, sous-différentiel...
  - aspects pratiques : pénalisation, SQP, OL...
- Organisation :
  - Partie II : 7 séances, dont 1 pour l'examen.
  - CM : 5 séances d'1h15++,
  - TD + TP : 5+2 séances d'2h00--,
  - TP : suite projet d'optimisation (Matlab/Scilab),
  - travail personnel.
- Supports de cours
  - syllabus [pdf] : ne pas voir les sections avec ⊖,
  - planches [pdf] : points importants du cours [§],
  - notes manuscrites [SP] : 1 document par séance,
  - exercices : en TD, dans le syllabus.
- Contrôle des connaissances
  - TP : rapport et code incrémental à remettre,
  - Séance 7 : résolution de problèmes (3h00).

## Plan du cours II

8. Conjugaison
   TP4
9. Sous-différentiabilité
   TD5 (sous-différentiabilité)
10. Pénalisation
    TD6 (pénalisation)
11. Optimisation quadratique successive (OQS/SQP)
    TP5
12. Optimisation linéaire : simplexe et PI
    TD7 (consolidation)
13. TD8 (consolidation)
    TD9 (consolidation)
14. Contrôle des connaissances

## VIII $\boxed{\textbf{Conjugaison}}$ (§ 3.5)

### Enveloppe convexe fermée d'un ensemble (§ 2.5.5)

Soient $\mathbb{E}$ un e.v. avec $\langle \cdot, \cdot \rangle$ et $P \subset \mathbb{E}$.

- **Dfn**. L'enveloppe convexe de $P$, $\operatorname{co} P$, est le plus petit convexe contenant $P$.

$$P \text{ fermé} \quad \not\Rightarrow \quad \operatorname{co} P \text{ fermé}.$$

- **Dfn**. L'enveloppe convexe fermée de $P$, $\overline{\operatorname{co}} P$, est le plus petit convexe fermé contenant $P$.
- **Dfn**. Un demi-espace fermé de $\mathbb{E}$ :

$$H^-(\xi, \alpha) := \{x \in \mathbb{E} : \langle \xi, x \rangle \leqslant \alpha\},$$

où $\xi \in \mathbb{E}$ est *non nul* et $\alpha \in \mathbb{R}$.

- **Prop**.

| |
|---|
| $\overline{\operatorname{co}} P$ est l'intersection de tous les demi-espaces fermés contenant $P$. |



### Enveloppe supérieure de fonctions réelles (§ 3.4.2)

- Enveloppe supérieure d'une famille de $f_i : \mathbb{E} \to \overline{\mathbb{R}}$, $i \in I$ (quelconque) :

$$\left(\sup_{i \in I} f_i\right)(x) := \sup_{i \in I} \big(f_i(x)\big).$$

| |
|---|
| • epi $(\sup_{i \in I} f_i) = \bigcap_{i \in I} (\operatorname{epi} f_i)$ . |
| • $f_i$ convexes $\Longrightarrow$ $\sup_{i \in I} f_i$ convexe. |
| • $f_i$ fermées $\Longrightarrow$ $\sup_{i \in I} f_i$ fermée. |

## Fonction conjuguée (§ 3.5)

Soient $\mathbb{E}$ un espace euclidien (prod. scal. $\langle\cdot,\cdot\rangle$) et $f:\mathbb{E}\to\mathbb{R}\cup\{+\infty\}$.

- **Dfn**. Conjuguée $f^*:\mathbb{E}\to\mathbb{R}\cup\{+\infty\}$ de $f$ :

$$f^*(x^*) := \sup_{x\in\mathbb{E}}\Big(\langle x^*,x\rangle - f(x)\Big).$$

  Biconjuguée $f^{**}:\mathbb{E}\to\mathbb{R}\cup\{+\infty\}$ de $f$ :

$$f^{**}(x) := \sup_{x^*\in\mathbb{E}}\Big(\langle x^*,x\rangle - f^*(x^*)\Big).$$

- **Prop**. Quelle que soit $f:\mathbb{E}\to\overline{\mathbb{R}}$, on a

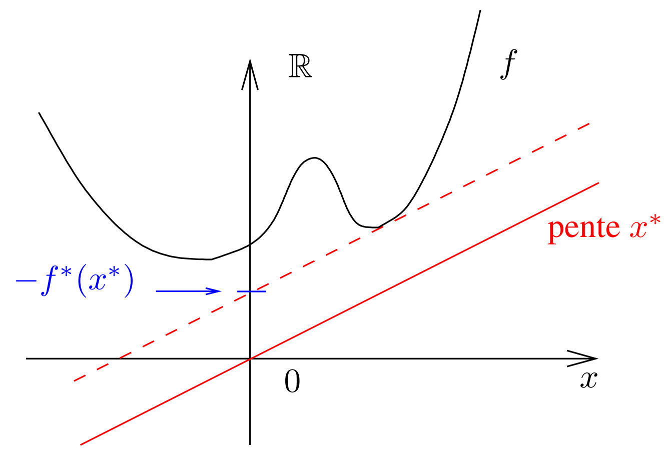
| |
|---|
| 1) $f^{**}\leqslant f$. |

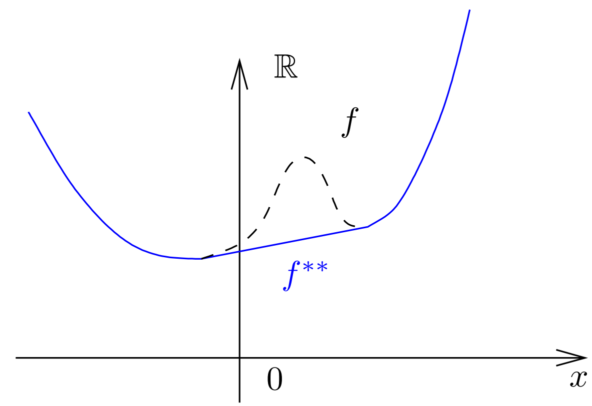
  Si $f$ est propre et a une minorante affine, on a

| |
|---|
| 2) $f^*\in\overline{\mathrm{Conv}}(\mathbb{E})$ et $f^{**}\in\overline{\mathrm{Conv}}(\mathbb{E})$. |
| 3) $f^{**}=$ enveloppe sup. minorantes affines de $f$. |
| 4) $f^{**}=f \iff f\in\overline{\mathrm{Conv}}(\mathbb{E})$. |



## Interprétation de $f^*(x^*)$ et $f^{**}$



## IX Sous-différentiabilité

### Sous-différentiel (§ 3.6)

- **Dfn** Le sous-différentiel $\partial f(x)$ de $f \in \mathrm{Conv}(\mathbb{E})$ en $x \in \mathrm{dom}\, f$ est l'ensemble des $x^* \in \mathbb{E}$ vérifiant les propriétés *équivalentes* suivantes :

  $(S_1)$ $\forall d \in \mathbb{E} : f'(x;d) \geqslant \langle x^*, d\rangle$,
  $(S_2)$ $\forall y \in \mathbb{E} : f(y) \geqslant f(x) + \langle x^*, y-x\rangle$,
  $(S_3)$ $x \in \arg\max_{y\in\mathbb{E}}(\langle x^*, y\rangle - f(y))$,
  $(S_4)$ $f(x) + f^*(x^*) \leqslant \langle x^*, x\rangle$,
  $(S_5)$ $f(x) + f^*(x^*) = \langle x^*, x\rangle$.

  Si $x \notin \mathrm{dom}\, f$, alors $\partial f(x) = \varnothing$ (par définition).

- **Dfn et Prop**

  $f$ est sous-différentiable en $x$
  $\iff \partial f(x) \neq \varnothing$ (définition),
  $\iff \exists\, y \in (\mathrm{dom}\, f)^{\circ} : f'(x; y-x) > -\infty$,
  $\iff f'(x;\cdot)$ ne prend pas la valeur $-\infty$,
  $\iff \exists\, L > 0, \forall\, y \in \mathbb{E} : f(y) \geqslant f(x) - L\|y-x\|$.

  En particulier, $f$ est sous-différentibale en $x \in (\mathrm{dom}\, f)^{\circ}$.



### Propriétés

On suppose que $f \in \mathrm{Conv}(\mathbb{E})$, $x \in \mathrm{dom}\, f$ et $x^* \in \mathbb{E}$.

1) Convexité : $\partial f(x)$ est un convexe fermé.
2) Règle de bascule : si $f \in \overline{\mathrm{Conv}}(\mathbb{E})$,

$$x^* \in \partial f(x) \iff x \in \partial f^*(x^*).$$

3) Sous-différentiel et $f^*$ : si $f \in \overline{\mathrm{Conv}}(\mathbb{E})$,

$$\partial f(x) = \arg\max_{x^*\in\mathbb{E}} \left(\langle x^*, x\rangle - f^*(x^*)\right).$$

4) Optimalité : $\bar{x} \in \arg\min f \iff 0 \in \partial f(\bar{x})$,
   et si $f \in \overline{\mathrm{Conv}}(\mathbb{E})$ : $\arg\min f = \partial f^*(0)$.
4) Formule du max : si $x \in (\mathrm{dom}\, f)^{\circ}$,
   - $\partial f(x)$ compact non vide,
   - $f'(x;d) = \max_{x^*\in\partial f(x)} \langle x^*, d\rangle$ (formule du max).
5) Différentiabilité :

$$f \text{ différentiable en } x \iff \partial f(x) \text{ est un singleton}$$

   et dans ce cas $\partial f(x) = \{\nabla f(x)\}$.

## Calcul

1) Somme

<u>Si</u> • $f_1, \ldots, f_m \in \mathrm{Conv}(\mathbb{E})$,

• $\cap_{i \in [1:p]} (\mathrm{dom}\, f_i)^{\ominus} \neq \varnothing$,

• $\alpha_1, \ldots, \alpha_p \in \mathbb{R}_+$,

<u>alors</u>

$$\partial \left( \sum_{i=1}^{m} \alpha_i f_i \right)(x) = \sum_{i=1}^{m} \alpha_i \partial f_i(x).$$

2) Pré-composition par une application affine

<u>Si</u> • $a : \mathbb{E} \to \mathbb{F}$ affine (i.e., $a(x) = Ax + b$, $\forall\, x \in \mathbb{E}$),

• $g \in \mathrm{Conv}(\mathbb{F})$ telle que $(\mathrm{dom}\, g)^{\ominus} \cap \mathcal{R}(A) \neq \varnothing$,

<u>alors</u>

$$\partial (g \circ a)(x) = A^* \left[ \partial g(a(x)) \right].$$



3) Enveloppe supérieure $f := \sup_{i \in I} f_i$ de $f_i \in \mathrm{Conv}(\mathbb{E})$.

<u>Si</u> • $x \in (\mathrm{dom}\, f)^{\circ}$,

• $I$ est compact,

• $\forall\, x \in \mathrm{dom}\, f$, $i \mapsto f_i(x)$ est s.c.s.,

• $I^0(x) := \{ i \in I : f(x) = f_i(x) \}$,

<u>alors</u>

$$\partial f(x) = \overline{\mathrm{co}} \left( \bigcup_{i \in I^0(x)} \partial f_i(x) \right).$$

4) Fonction marginale de $\varphi : (x, y) \in \mathbb{E} \times \mathbb{F} \mapsto \varphi(x, y)$ :

$$f : \mathbb{E} \to \overline{\mathbb{R}} : x \mapsto f(x) := \inf_{y \in \mathbb{F}} \varphi(x, y).$$

<u>Si</u> • $\varphi \in \mathrm{Conv}(\mathbb{E} \times \mathbb{F})$,

• $f \in \mathrm{Conv}(\mathbb{E})$,

• $x \in \mathrm{dom}\, f$,

• $f(x) = \varphi(x, y_x)$, pour un $y_x \in \mathbb{F}$,

<u>alors</u>

$$\partial f(x) = \{ x^* \in \mathbb{E} : (x^*, 0) \in \partial \varphi(x, y_x) \}.$$

## Exemple 1D

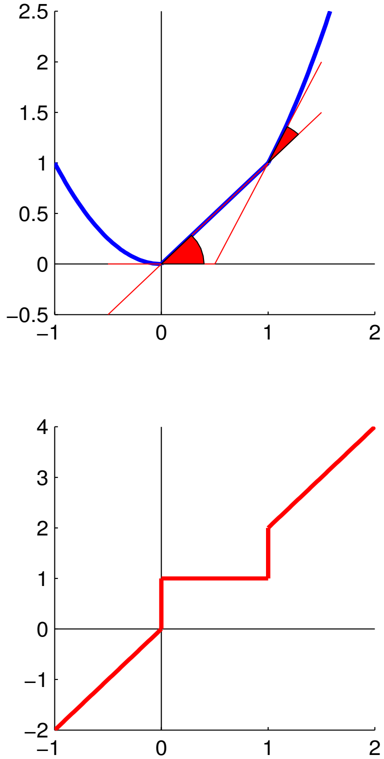



Figure 1: $f(x) = \max(x, x^2)$ et $\partial f(x)$ (en bas)



## Exemple 2D

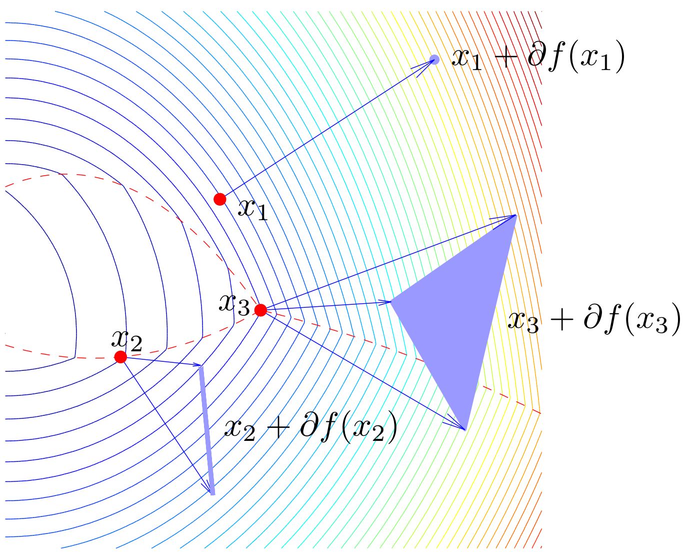



Figure 2: $f = \sup(q_1, q_2, q_3)$ et $\partial f$

## Sous-différentiel de la fonction duale (§ 13.6.1)

Pour le problème (non néc. convexe)

$$\begin{cases} \inf\ f(x) \\ c(x) \leqslant 0 \\ x \in X \end{cases}$$

et une fonction duale

$$\delta(\lambda) = -\inf_{x\in X} \Big(\ell(x,\lambda) := f(x) + \lambda^\top c(x)\Big)$$

propre, on a $\delta \in \overline{\text{Conv}}(\mathbb{R}^m)$ et

$$\boxed{-c\left(\arg\min_{x\in X} \ell(x,\lambda)\right) \subset \partial\delta(\lambda).}$$



## Signification des multiplicateurs optimaux (§ 4.6.1)

- Problème perturbé : pour $p \in \mathbb{R}^m$, on définit

$$(P_{EI}^p) \quad \begin{cases} \min\ f(x) \\ c_E(x) + p_E = 0 \\ c_I(x) + p_I \leqslant 0. \end{cases}$$

- **Dfn**. La fonction valeur associée à $(P_{EI}^p)$ est $v : p \in \mathbb{R}^m \to \overline{\mathbb{R}}$ définie par

$$v(p) = \inf_{x\in X^p} f(x),$$

où $X^p$ est l'ensemble admissible de $(P_{EI}^p)$.

$$(P_{EI}) \text{ convexe} \implies v \text{ convexe.}$$

- Cas différentiable régulier (rappel).

<u>Si</u> • $(x_*, \lambda_*)$ solution PD de $(P_{EI})$,
· $(\bar{x}(p), \bar{\lambda}(p))$ solution PD de $(P_{EI}^p)$,
· $p \mapsto \bar{x}(p)$ différentiable en 0, $\bar{x}(0) = x_*$,
· $p \mapsto \bar{\lambda}(p)$ continue en 0, $\bar{\lambda}(0) = \lambda_*$,

<u>alors</u> $\lambda_* = \nabla v(0) = \nabla(f \circ \bar{x})(0)$.

- On note
$$\Lambda := \{\lambda \in \mathbb{R}^m : \lambda_I \geqslant 0\}.$$
- **Dfn**. On dit que $(x_*, \lambda_*) \in \mathbb{R}^n \times \Lambda$ est un point-selle de $\ell$ sur $\mathbb{R}^n \times \Lambda$, si $\forall (x, \lambda) \in \mathbb{R}^n \times \Lambda$ :
$$\ell(x_*, \lambda) \leqslant \ell(x_*, \lambda_*) \leqslant \ell(x, \lambda_*).$$
On note
$\Lambda(x_*) := \{\lambda_* : (x_*, \lambda_*)$ est point-selle de $\ell$ sur $\mathbb{R}^m \times \Lambda\}$.

- Cas convexe non différentiable.

| |
|---|
| Si • $x_*$ est solution de $(P_{EI})$,<br>• $v \in \mathrm{Conv}(\mathbb{R}^m)$,<br>alors $\partial v(0) = \Lambda(x_*)$. |

Remarque : Ci-dessus, $\partial v(0)$ peut être vide ! Avec qualification de Slater : $\partial v(0) \neq \varnothing$.



- CN et CS d'existence de solution PD globale.

CN d'optimalité (cas convexe non diff.).

| |
|---|
| Si • $(P_{EI})$ convexe (avec $f$ et $c$ finies),<br>• (Slater) : $c'_E$ surjective, $\exists \hat{x} \in X$ t.q. $c_I(\hat{x}) < 0$,<br>• $x_*$ solution de $(P_{EI})$,<br>alors 1) $v$ est loc. lipschitzienne dans un vois. de 0,<br>2) $\partial v(0) \neq \varnothing$. |

CS d'optimalité globale.
Peu de chance d'être applicable si $(P_{EI})$ non convexe.

| |
|---|
| Si • $(x_*, \lambda_*) \in \mathbb{R}^n \times \Lambda$ est un point-selle de $\ell$ sur $\mathbb{R}^n \times \Lambda$,<br>alors $x_*$ solution (globale) de $(P_{EI})$. |

X **Pénalisation (§ 12)**

- **À quoi ça sert ?**
  En optimisation avec contraintes :
  - pour la théorie: obtenir des propriétés à partir de problèmes approchés sans contrainte,
  - pour l'algorithmique: résoudre un problème avec contraintes « sans trop en faire ».

- **Transformation typique**. Soit $X \subset \mathbb{E}$ (un espace vectoriel). On passe du problème *avec contrainte*

$$(P_X) \quad \inf_{x \in X} f(x)$$

  au problème pénalisé *sans contrainte*

$$(P_r) \quad \inf_{x \in \mathbb{E}} \Big(\Theta_r(x) := f(x) + rp(x)\Big),$$

  où $r \in \mathbb{R}$ est un paramètre de pénalisation et $p : \mathbb{E} \to \overline{\mathbb{R}}$ est une fonction de pénalisation (on va voir ce que c'est).

- **Pénalisation exacte** (notion vague) lorsque $x_*$ est « solution » de $(P)$ « ssi » $x_*$ est « solution » de $(P_r)$.



## Deux résultats généraux (§ 12.1)

- Monotonie en pénalisation

  Si $\forall\, r$ considéré, $(P_r)$ a une solution, notée $\bar{x}_r$.
  Alors lorsque $r$ croît :
  1) $p(\bar{x}_r)$ décroît,
  2) $f(\bar{x}_r)$ croît, si $r \geqslant 0$,
  3) $\Theta_r(\bar{x}_r)$ croît, si $p(\cdot) \geqslant 0$.

- Point d'adhérence lorsque $r \downarrow 0$

  Si • $\mathbb{E}$ est un espace topologique,
  • $f : \mathbb{E} \to \overline{\mathbb{R}}$ est propre et s.c.i. (implicitement $X := \operatorname{dom} f$),
  • $p$ est continue,
  • $S := \arg\min \{f(x) : x \in \mathbb{E}\} \neq \varnothing$,
  • $\forall\, r > 0$ petit, $(P_r)$ a une solution, notée $\bar{x}_r$.

  Alors tout point d'adhérence de $\{\bar{x}_r\}_{r \downarrow 0}$ est solution de

$$\inf_{x \in S} p(x).$$

## Pénalisation extérieure (§ 12.2)

- **Exemple**. On veut résoudre

$$(P) \quad \begin{cases} \min\ f(x) \\ c(x) \leqslant 0. \end{cases}$$

On *approche* ce problème par $(r > 0)$

$$(P_r) \quad \min\ f(x) + \frac{r}{2}\|c(x)^+\|_2^2,$$

que l'on résout par un algorithme de descente, pour une suite de $r \to \infty$.



- **Pénalisation quadratique en 1D**

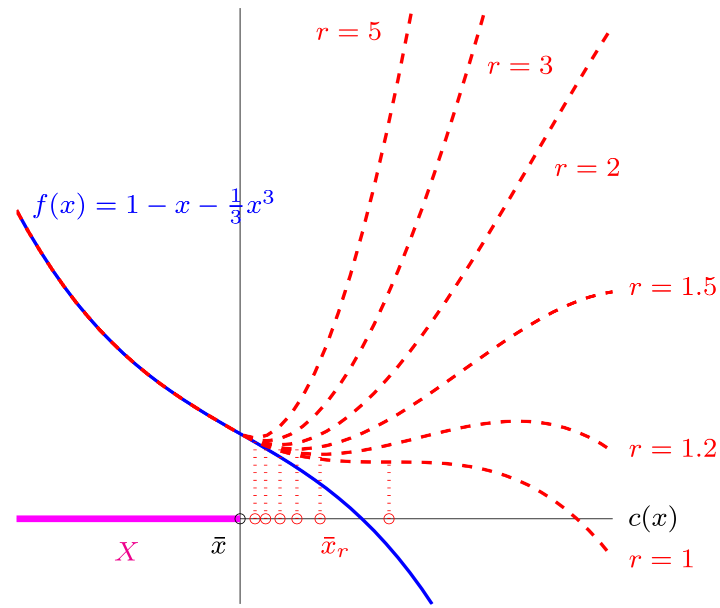



Figure 3: Pénalisation quadratique

- Plus généralement, on suppose que $r \geqslant 0$ et que la fonction de pénalisation vérifie

$$(H_p) \quad \begin{cases} p \text{ est continue sur } \mathbb{E} \\ p(x) \geqslant 0, \ \forall x \in \mathbb{E} \\ p(x) = 0 \iff x \in X. \end{cases}$$

Résultat d'approximation

> Si • $X$ est fermé et non vide,
> - $p : \mathbb{E} \to \mathbb{R}$ vérifie $(H_p)$,
> - $f$ est s.c.i.,
> - $\exists\, r_0 \geqslant 0$ tel que $\Theta_{r_0}(x) \to +\infty$ quand $\|x\| \to \infty$,
>
> alors 1) $\forall r \geqslant r_0$, $(P_r)$ a au moins 1 solution $\bar{x}_r$,
> 2) $\{\bar{x}_r\}_{r\uparrow\infty}$ est bornée,
> 3) tout point d'adhérence de la suite $\{\bar{x}_r\}_{r\uparrow\infty}$ est solution de $(P_X)$.

- **Estimation d'une solution duale.**

> Si • problème $(P_{EI})$ avec $f$ et $c$ différentiables,
> - $\Theta_r(x) = f(x) + \frac{r}{2}\|c(x)^\#\|_2^2$,
> - $\Theta_r(\bar{x}_r) = 0$,
> - $\bar{x}_r \to \bar{x}$,
> - pour $K := E \cup \{i \in I : c_i(\bar{x}) \geqslant 0\}$, $c'_K(\bar{x})$ est surjective,
>
> alors 1) $\exists\, \bar{\lambda}$ tel que $(\bar{x}, \bar{\lambda})$ vérifie (KKT),
> 2) $r\, c(\bar{x}_r)^\# \to \bar{\lambda}$.



- **$\oplus$ et $\ominus$ de la pénalisation extérieure**

$\oplus$ Facile à mettre en œuvre (avec algo. sans contrainte).

$\ominus$ Suite de problèmes non linéaires (bon $r$ inconnu, premier $r$ très grand ne convient pas).

$\ominus$ Le mauvais conditionnement augmente avec $r$ (i.e., les courbes de niveau s'allongent).

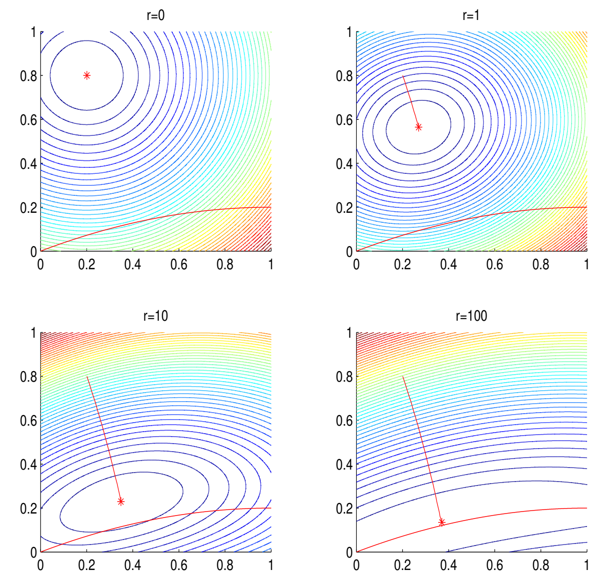



Figure 4: Chemin des minimiseurs

## Pénalisation lagrangienne (§ 12.4.1)

On considère le problème $(P_{EI})$.

Si • $(P_{EI})$ est convexe,
• $\bar{x}$ est solution de $(P_{EI})$,
• $f$ et $c$ sont différentiables en $\bar{x}$,
• $\exists\, \bar{\lambda}$ tel que $(\bar{x}, \bar{\lambda})$ vérifie (KKT),

alors $\ell(\cdot, \bar{\lambda})$ a un minimum *global* en $\bar{x}$.



## Pénalisation lagrangienne augmentée (§ 12.4)

### Exemple en 1D

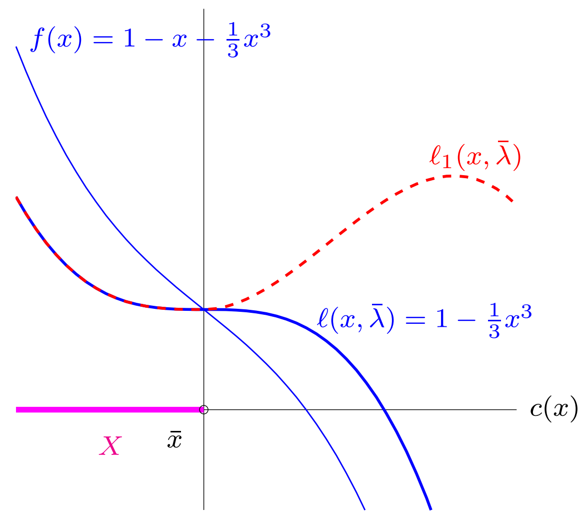



Figure 5: Pénalisation lagrangienne augmentée

- **Le lagrangien augmenté de $(P_E)$**

$$\ell_r(x,\lambda) = f(x) + \lambda^\top c(x) + \frac{r}{2}\,\|c(x)\|_2^2.$$

- **Lemme de Finsler**

<u>Si</u> • $A$ et $M$ symétriques de même ordre,
• $A \succcurlyeq 0$,
• $v^\top M v > 0$ pour tout $v \in \mathcal{N}(A) \setminus \{0\}$,
<u>alors</u> $\exists\, \bar{r} \in \mathbb{R}$ tel que $\forall\, r \geqslant \bar{r}$, on a $M + rA \succ 0$.

- **Exactitude du lagrangien augmenté**.

<u>Si</u> • $f$ et $c$ sont deux fois différentiables en $\bar{x}$,
• $(\bar{x}, \bar{\lambda})$ vérifie KKT et CS2,
<u>alors</u> $\exists\, \bar{r} \in \mathbb{R}$, $\forall\, r \geqslant \bar{r}$, $\bar{x}$ est un minimum *local strict* de $\ell_r(\cdot, \bar{\lambda})$.



## L'algorithme des multiplicateurs (ou du LA)

C'est une méthode de dualité qui cherche à déterminer un multiplicateur optimal $\bar{\lambda}$, un facteur de pénalisation $r$ et à minimiser $\ell_r(\cdot, \bar{\lambda})$.

De $(\lambda_k, r_k) \in \mathbb{E} \times \mathbb{R}$ à $(\lambda_{k+1}, r_{k+1}) \in \mathbb{E} \times \mathbb{R}$ :

1. *Nouvel itéré primal* :
$$x_k \in \arg\min_{x \in \mathbb{E}} \ell_{r_k}(x, \lambda_k).$$
2. *Test d'arrêt* : si $(x_k, \lambda_k)$ est satisfaisant (vérifie approximativement des conditions d'optimalité), arrêt.
3. *Nouvel itéré dual* :
$$\lambda_{k+1} := \lambda_k + r_k\, c(x_k). \tag{1}$$
4. Mise à jour de $r_{k+1} > r_k$ si nécessaire (heuristique).

La formule (1) est justifiable par

- $\nabla_x \ell(x_k, \lambda_{k+1}) = 0$ (on a bien envie d'annuler $\nabla_x \ell$).
- $-c(x_k) \in \partial\delta(\lambda_{k+1})$, où $\delta$ est la fonction duale, si bien que l'algorithme est une méthode proximale (i.e., de gradient *implicite* puisque $\partial\delta$ est évalué en $\lambda_{k+1}$ et pas en $\lambda_k$) sur $\delta$.

## Pénalisation non différentiable (§ 12.5)

### Exemple en 1D

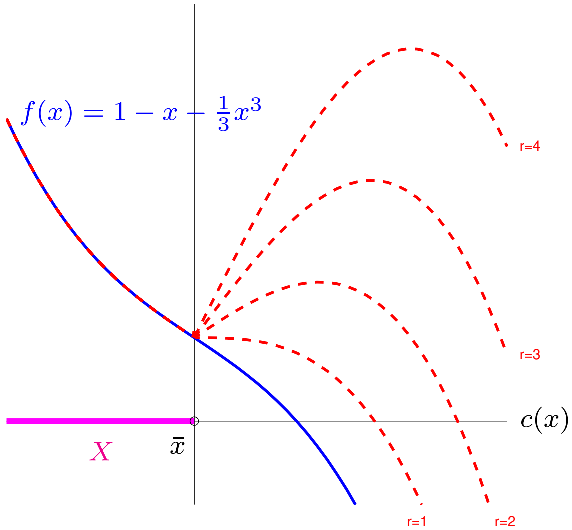



Figure 6: Pénalisation non différentiable

## XI Optimisation quadratique successive (OQS)

**(§ 14)**

Le problème à résoudre en $x \in \mathbb{E}$ :

$$(P_{EI}) \quad \begin{cases} \min f(x) \\ c_E(x) = 0 \in \mathbb{R}^{m_E} \\ c_I(x) \leqslant 0 \in \mathbb{R}^{m_I}. \end{cases}$$

Le lagrangien du problème ($c := (c_E, c_I)$) :

$$\ell(x, \lambda) = f(x) + \lambda^\top c(x).$$



Newton pour résoudre $F(x) = 0$ (rappel)

Si
- $x_*$ vérifie $F(x_*) = 0$,
- $F$ est $C^{1,1}$ dans un voisinage de $x_*$,
- $F'(x_*)$ est inversible,

alors il existe un voisinage $V$ de $x_*$ tel que si $x_1 \in V$, l'algorithme de Newton est bien défini et génère une suite $\{x_k\} \subset V$ qui converge *quadratiquement* vers $x_*$.

Résultat de convergence locale de OQS

Si
- $f$ et $c$ sont $C^{2,1}$ près de $x_* \in \mathrm{Sol}(P_{EI})$,
- $\exists\, \lambda_*$, unique multiplicateur associé à $x_*$,
- les CS2 sont vérifiées en $(x_*, \lambda_*)$,

alors il existe un voisinage $V$ de $(x_*, \lambda_*)$ tel que, si $(x_1, \lambda_1) \in V$, l'algorithme OQS démarrant en $(x_1, \lambda_1)$

1) peut générer une suite $\{(x_k, \lambda_k)\} \subset V$ en calculant à chaque itération des points stationnaires du PQO,
2) cette suite $\{(x_k, \lambda_k)\}$ converge *quadratiquement* vers $(x_*, \lambda_*)$.

## XII $\boxed{\textbf{Optimisation linéaire : simplexe}}$ (§ 15)

- On considère le problème d'optimisation linéaire sur $\mathbb{R}^n$ (forme standard)

$$(P_L) \quad \begin{cases} \min\ c^\top x \\ Ax = b \\ x \geqslant 0, \end{cases}$$

où

- $c \in \mathbb{R}^n$,
- $A$ est $m \times n$ surjective ($m \leqslant n$),
- $b \in \mathbb{R}^m$.

- On note

$$X := \{x \in \mathbb{R}^n : Ax = b,\ x \geqslant 0\}$$

l'ensemble admissible.

- Pour $x \in X$, on note

$$\begin{aligned} I^+(x) &:= \{i : x_i > 0\} \\ I^0(x) &:= \{i : x_i = 0\}. \end{aligned}$$



### Existence de solution (§ 15.2.2)

- **Théor** (existence de solution) :

$(P_L)$ a une solution

$\iff$ $(P_L)$ est réalisable et borné.

### Conditions d'optimalité (§ 15.2.2)

- **Théor** :

$(P_L)$ a une solution

$$\iff \exists y \in \mathbb{R}^m,\ \exists s \in \mathbb{R}^n\ : \begin{cases} A^\top y + s = c, \quad s \geqslant 0, \\ Ax = b, \quad x \geqslant 0, \\ x^\top s = 0. \end{cases}$$

## <u>Algorithme du simplexe</u> (§ 15.4)

(description géométrique)

- **Hypothèse** : $A$ est $m \times n$ surjective.

- **Phase I** : trouver $\hat{x} \in X$, un sommet de $X$.

  On prend $B \supset I^+(\hat{x})$ avec $|B| = m$ et $N = B^c \subset I^0(\hat{x})$.

- **Phase II** : on itère de sommet en sommet.
  Voici une itération.

  – Coût réduit : $r := c_N - A_{\bullet N}^\top A_{\bullet B}^{-\top} c_B$.

  – Optimalité :
    si $r \geqslant 0$, $\hat{x}$ est solution (arrêt);
    sinon $\exists j$ t.q. $r_j < 0$.

  – Direction de déplacement $d$ :
    $d_N = e_N^j$ et $d_B = -A_{\bullet B}^{-1} A_{\bullet N} e_N^j$.

  – Si $d_B \geqslant 0$, $(P_L)$ est non borné (arrêt).

  – Nouveau sommet $\hat{x}^+$ :
    prendre le plus grand $\alpha \geqslant 0$ tel que

$$\hat{x}^+ := \hat{x} + \alpha d \in X$$

    ($\alpha > 0$ si $\hat{x}$ est non dégénéré).

XII' $\boxed{\textbf{Optimisation linéaire : points intérieurs}}$

**(§ 16)**

## La une du New York Times

(19 novembre 1984)

Voici comment l'algorithme de Karmarkar (le premier algorithme de points intérieurs efficace) était « révélé » au grand public :

> *« The discovery, which is to be formally published next month, is already circulating rapidly through the mathematics world. It has also set off a deluge of inquiries from brokerage houses, oil companies and airlines, industries with millions of dollars at stake in problems known as linear programming. »*



## Notations

Le problème et son dual

$$(P)\ \begin{cases} \inf\ c^\top x \\ Ax = b \\ x \geqslant 0, \end{cases} \qquad (D)\ \begin{cases} \sup\ b^\top y \\ A^\top y + s = c \\ s \geqslant 0. \end{cases}$$

Ensembles admissibles

$$\mathcal{F}_P := \{x \in \mathbb{R}^n : Ax = b,\ x \geqslant 0\}$$

$$\mathcal{F}_D := \{(y,s) \in \mathbb{R}^{m+n} : A^\top y + s = c,\ s \geqslant 0\}.$$

Intérieurs relatifs

$$\mathcal{F}_P^o := \{x \in \mathbb{R}^n : Ax = b,\ x > 0\}$$

$$\mathcal{F}_D^o := \{(y,s) \in \mathbb{R}^{m+n} : A^\top y + s = c,\ s > 0\}.$$

Conditions d'optimalité

$$\begin{cases} Ax = b, \quad x \geqslant 0 \\ A^\top y + s = c, \quad s \geqslant 0 \\ x^\top s = 0. \end{cases}$$

## Le chemin central primal-dual (§ 16.1)

- Conditions d'optimalité perturbées par $\mu > 0$ :

$$(\mathrm{KKT}_\mu) \quad \begin{cases} Ax = b & (x > 0) \\ A^\top y + s = c & (s > 0) \\ Xs = \mu\, e. \end{cases}$$

- **Existence et unicité** :

$\boxed{\begin{array}{l}\underline{\text{Si}}\ \ \mathcal{F}_P^o \neq \varnothing, \mathcal{F}_D^o \neq \varnothing \text{ et } \mu > 0, \\ \underline{\text{alors}}\ \ (\mathrm{KKT}_\mu) \text{ a une solution unique.}\end{array}}$

Si $\mu = 0$, il y a existence mais pas néc. unicité !

- **Dfn** : le chemin central est l'ensemble des solutions de $(\mathrm{KKT}_\mu)$ :

$$\{(x_\mu, y_\mu, s_\mu) : \mu > 0\}.$$



## Algorithme PD de suivi de chemin

(éléments constitutifs, § 16.2)

- On suppose $A$ surjective.
- Soit $z = (x, y, s)$ l'itéré courant admissible.
- Choix de

$$\mu := \frac{x^\top s}{n}.$$

- Facteur de réduction $\sigma \in ]0, 1[$ de $\mu$.
- Direction de Newton $d_z = (d_x, d_y, d_s)$

$$\begin{pmatrix} 0 & A^\top & I \\ A & 0 & 0 \\ S & 0 & X \end{pmatrix} \begin{pmatrix} d_x \\ d_y \\ d_z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \sigma\mu e - Xs \end{pmatrix},$$

où $X = \mathrm{diag}(x_i)$, $S = \mathrm{diag}(s_i)$, $e = (1 \ \cdots \ 1)^\top$.

$$(x, s) > 0 \implies \text{SL inversible.}$$

- Nouvel itéré $z^+ = z + \alpha d_z$.
- Contrôle du pas $\alpha$ pour que ($\gamma \simeq 10^{-3}$)

$$z^+ \in V_{-\infty}(\gamma) := \{z \in \mathcal{F}^o : Xs \geqslant \gamma\mu e\}.$$

## Algorithme PD de suivi de chemin

(algorithme des grands déplacements, § 16.3.3)

- **Paramètres** :
  $0 < \sigma_{\min} < \sigma_{\max} < 1$ et $\gamma \in ]0, 1[$.
- **Donnée** :
  $z = (x, y, s) \in V_{-\infty}(\gamma)$ primal-dual admissible (i.e., $Ax = b$ et $A^\top y + s = c$).

- **Une itération** :
  - Choix de $\sigma \in ]0, 1[$.
  - Calcul de la direction de Newton $d_z$.
  - Choisir $\alpha \in [0, 1]$ le plus grand possible pour que $z + \alpha d_z \in V_{-\infty}(\gamma)$.
  - Nouvel itéré $z^+ := z + \alpha d_z$.