Introduction à l'optimisation

MU4MEN01 : CALCUL SCIENTIFIQUE – 3 ECTS

RESPONSABLE D'UE ET COURS : FLORENCE OSSART

III. Minimisation sous contraintes-égalités

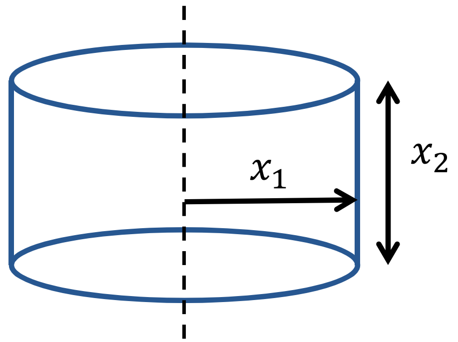
Exemple 1 (déjà vu dans le cours n°1)

→ Situation :

- On veut produire des boîtes cylindriques de volume V_0 donné en utilisant le moins de matière possible. Il faut donc minimiser la surface du cylindre, à volume donné.

→ Formalisation mathématique :

- Variable de décision : $\boldsymbol{x} = (x_1, x_2)$
- Fonction coût : $\mathrm{J}(\boldsymbol{x}) = 2 \times \pi {x_1}^2 + 2\pi x_1 x_2$
- Contraintes : $\mathrm{V}(\boldsymbol{x}) = \pi {x_1}^2 x_2 = V_0$; $x_1 \geq 0$; $x_2 \geq 0$



→ Problème de minimisation sous contrainte :

Trouver $\boldsymbol{x}^*$ tel que :
$J(\boldsymbol{x}^*) = \inf \{ J(\boldsymbol{x}), \boldsymbol{x} \in \mathbb{R}^{+2} \}$
soumis à $C(\boldsymbol{x}) = \pi {x_1}^2 x_2 - V_0 = 0$

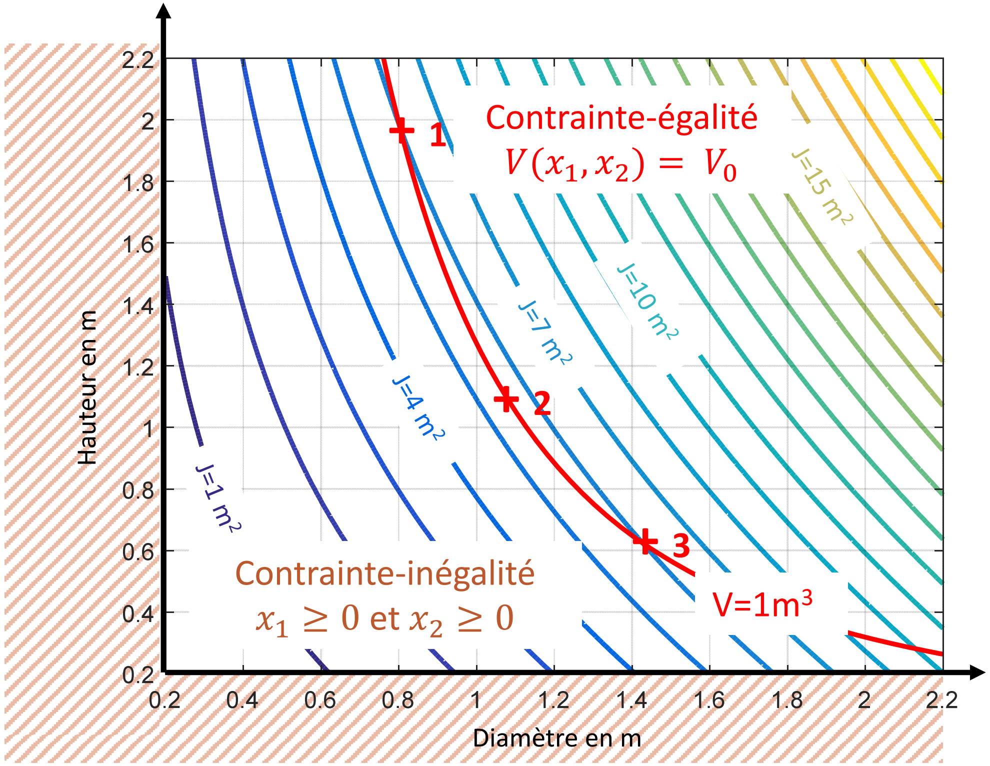
Exemple 1, résolution analytique

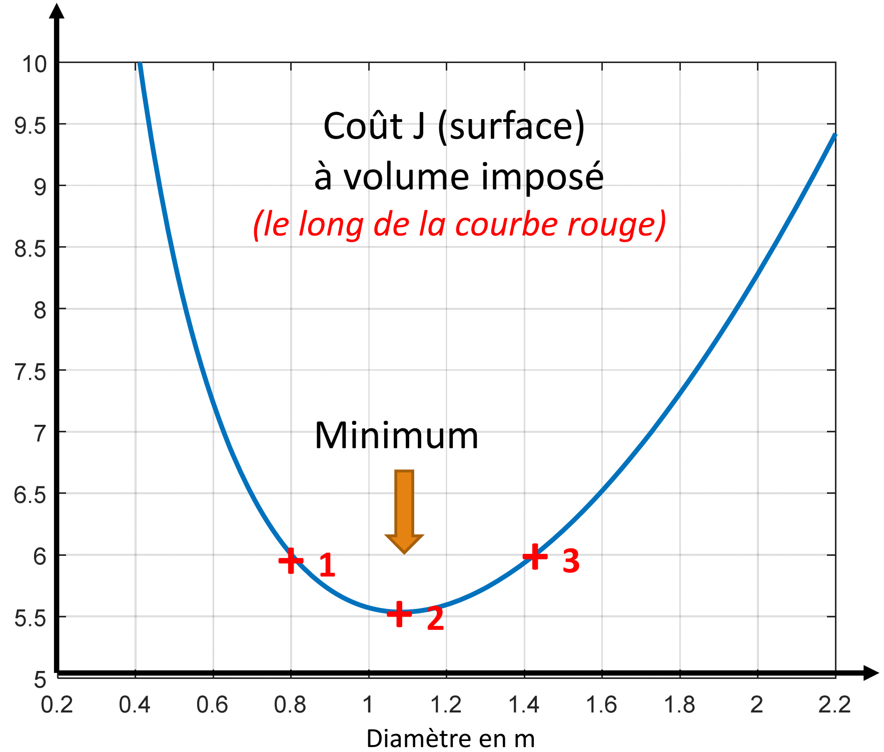
→ Expressions simples => on peut procéder par substitution :

- Utiliser la contrainte pour exprimer x2 en fonction de x1
- Reporter dans l'expression du coût
- Minimiser par rapport à x1
- Résultat : $x_1^* = \sqrt[3]{\frac{V_0}{2\pi}}$ $\quad x_2^* = 2x_1^*$ $\quad J^* = 3\sqrt[3]{2\pi\, V_0^2}$
- Pour $V_0 = 1\, m^3$, on obtient : $x_1^* = 0{,}54\, m$; $x_2^* = 1{,}08\, m$ $\quad J^* = 5{,}54\, m^2$

Exemple 1, représentation graphique

→ Représentation du coût (isovaleurs de la surface) et des contraintes :





Position du problème

→ Formulation d'un problème de minimisation avec contrainte-égalité :

Trouver $\boldsymbol{x}^*$ tel que :

$J(\boldsymbol{x}^*) = \inf \{ J(\boldsymbol{x}) , \boldsymbol{x} \in \mathbb{R}^n \}$ n variables

soumis à $C_j(\boldsymbol{x}) = 0, j = 1, m$ m contraintes

Problème : La condition $\nabla J(x^*) = 0$ ne permet pas de déterminer les points critiques, car elle ne prend pas en compte les contraintes

Méthode de substitution

→ n variables, m contraintes :

=> m relations entre les n variables (1 contrainte => 1 relation)

=> on peut exprimer m variables en fonction des (n-m) autres variables

=> on peut exprimer J en fonction de (n-m) variables

=> on est ramené à un problème de minimisation sans contraintes dans $\mathbb{R}^{n-m}$

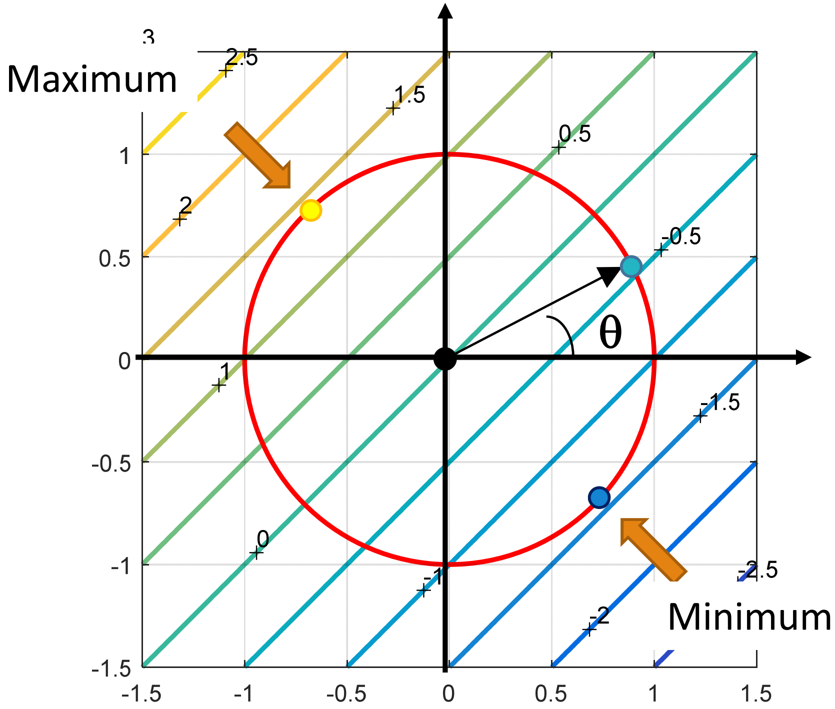
→ Conceptuellement simple, mais l'expression de m variables en fonction des (n-m) autres variables devient vite très compliquée, voire impossible, dès que m augmente

→ Intérêt pratique très limité

Exemple 2

Minimiser $J(x_1, x_2) = x_2 - x_1$,

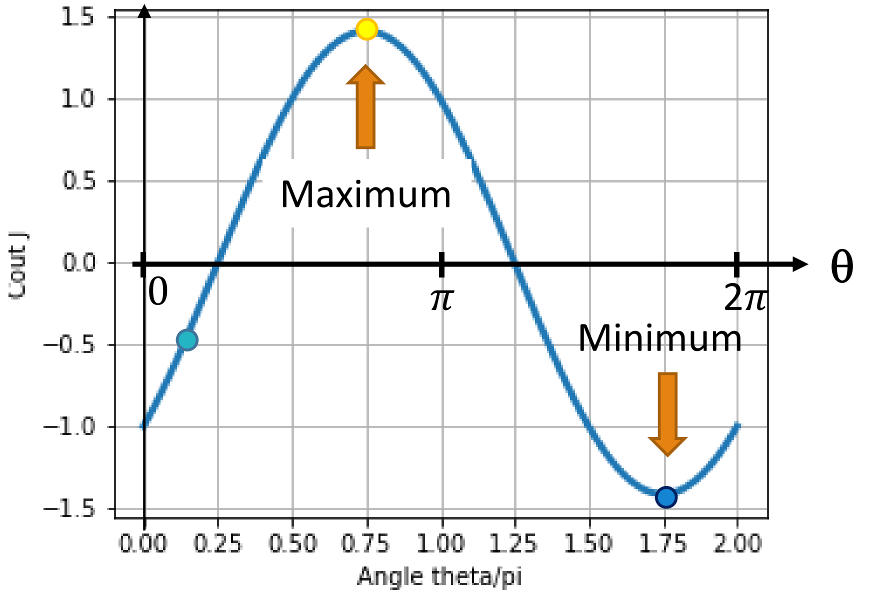
sous contrainte $C(x_1, x_2) = {x_1}^2 + {x_2}^2 - 1 = 0$



$$x_1 = \cos\theta$$
$$x_2 = \sin\theta$$

Fonction coût le long du cercle rouge :

$$J(\theta) = \sqrt{2}\sin\left(\theta - \frac{\pi}{4}\right)$$



Résolution par les multiplicateurs de Lagrange

→ Condition de *point-critique* sous contrainte

→ x^* est un point critique s'il existe m nombres $\left(\lambda_j\right)_{j=1,m}$, tels que :

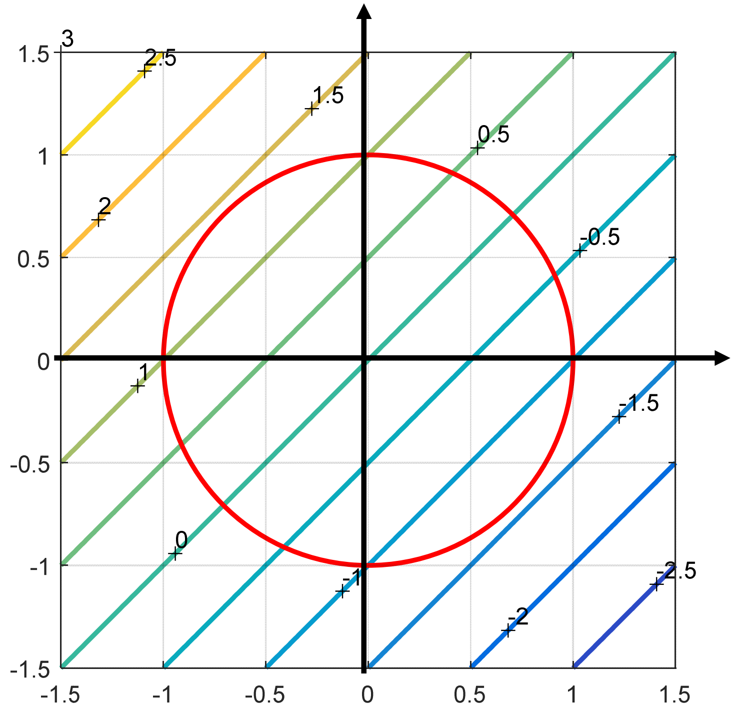
$$\begin{cases} \nabla J(x^*) + \sum_{j=1}^{m} \lambda_j . \nabla C_j(x^*) = 0 \\ C_j(x^*) = 0, j = 1, m \end{cases}$$

→ Système de $n + m$ équations à $n + m$ inconnues => a priori soluble

→ Les coefficients $\left(\lambda_j\right)_{j=1,m}$ sont appelés multiplicateurs de Lagrange

Exemple 2

→ Minimiser $J(x_1, x_2) = x_2 - x_1$, sous contrainte $C(x_1, x_2) = {x_1}^2 + {x_2}^2 - 1 = 0$



$$\nabla J(x) = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \qquad \nabla C(x) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}$$

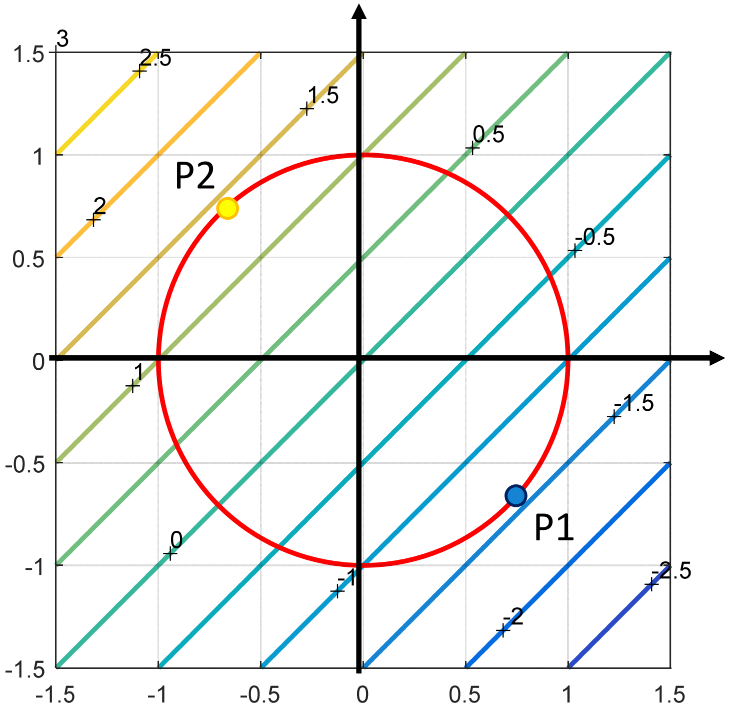
$$\nabla J(x) + \lambda . \nabla C(x) = \begin{pmatrix} -1 + \lambda . 2x_1 \\ 1 + \lambda . 2x_2 \end{pmatrix}$$

Les points critiques sont les solutions du système :

$$\begin{cases} -1 + \lambda . 2x_1 = 0 \\ 1 + \lambda . 2x_2 = 0 \\ {x_1}^2 + {x_2}^2 - 1 = 0 \end{cases}$$

Exemple 2

→ Minimiser $J(x_1, x_2) = x_2 - x_1$, sous contrainte $\mathrm{C}(x_1, x_2) = {x_1}^2 + {x_2}^2 - 1 = 0$



Après calculs, on obtient :

Point critique 1 P1 :

$\lambda = \frac{\sqrt{2}}{2}$ $\quad x_1 = \frac{\sqrt{2}}{2}$ $\quad x_2 = -\frac{\sqrt{2}}{2}$ $\quad J_1 = -\sqrt{2} \approx -1{,}4$

Point critique 2 P2 :

$\lambda = -\frac{\sqrt{2}}{2}$ $\quad x_1 = -\frac{\sqrt{2}}{2}$ $\quad x_2 = \frac{\sqrt{2}}{2}$ $\quad J_2 = \sqrt{2} \approx 1{,}4$

La comparaison de J_1 et J_2 montre que P1 correspond à un minimum et P2 à un maximum.

Interprétation géométrique

→ Facile dans le cas simple « minimiser $J(x)$, sous contrainte unique $\mathrm{C}(x) = 0$ »

→ x^* est un point critique s'il existe λ, tel que :

$$\begin{cases} \nabla J(x^*) + \lambda . \nabla C(x^*) = 0 \\ C(x^*) = 0 \end{cases}$$

(n+1 équations, n+1 inconnues)

→ Géométriquement : au point critique x^*, $\nabla J(x^*)$ et $\nabla C(x^*)$ sont colinéaires

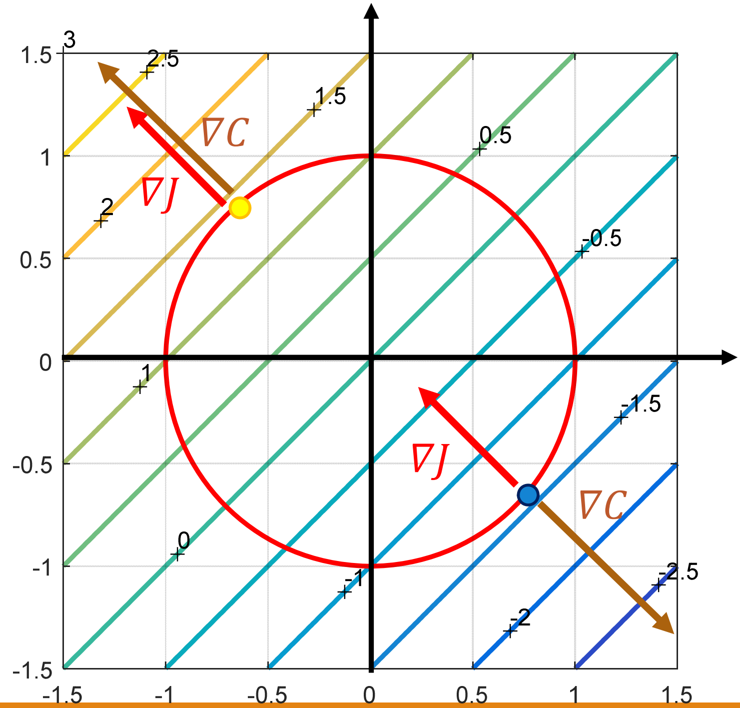
Il n'y a pas d'interprétation du signe de λ sans analyse complémentaire sur la convexité de J et C.

Exemple 2

→ Minimiser $J(x_1, x_2) = x_2 - x_1$, sous contrainte $C(x_1, x_2) = {x_1}^2 + {x_2}^2 - 1 = 0$

→Géométriquement : au point critique x^*, $\nabla J(x^*)$ et $\nabla C(x^*)$ sont colinéaires

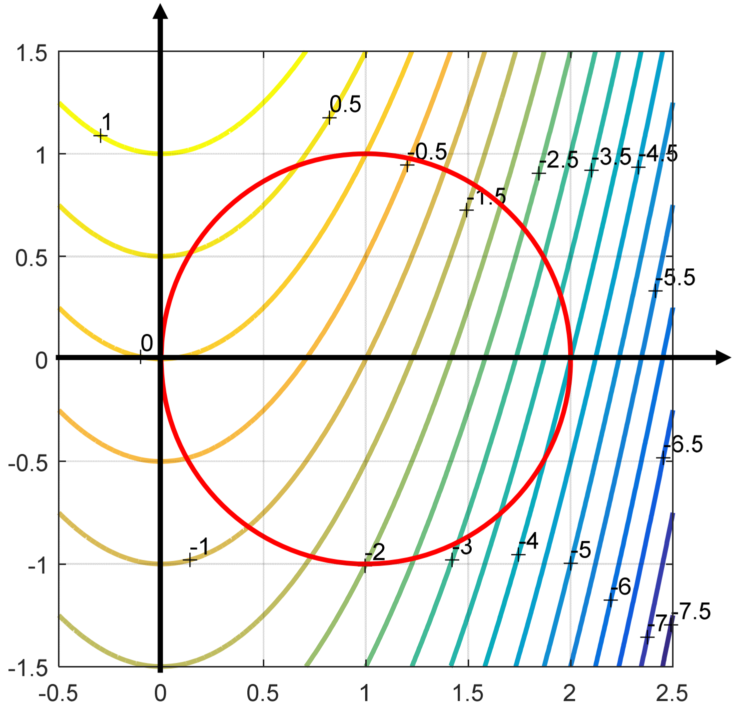
Point critique 1
$\nabla J(x^*)$ et $\nabla C(x^*)$
sont parallèles



Point critique 2
$\nabla J(x^*)$ et $\nabla C(x^*)$
sont anti-parallèles

Exemple 3

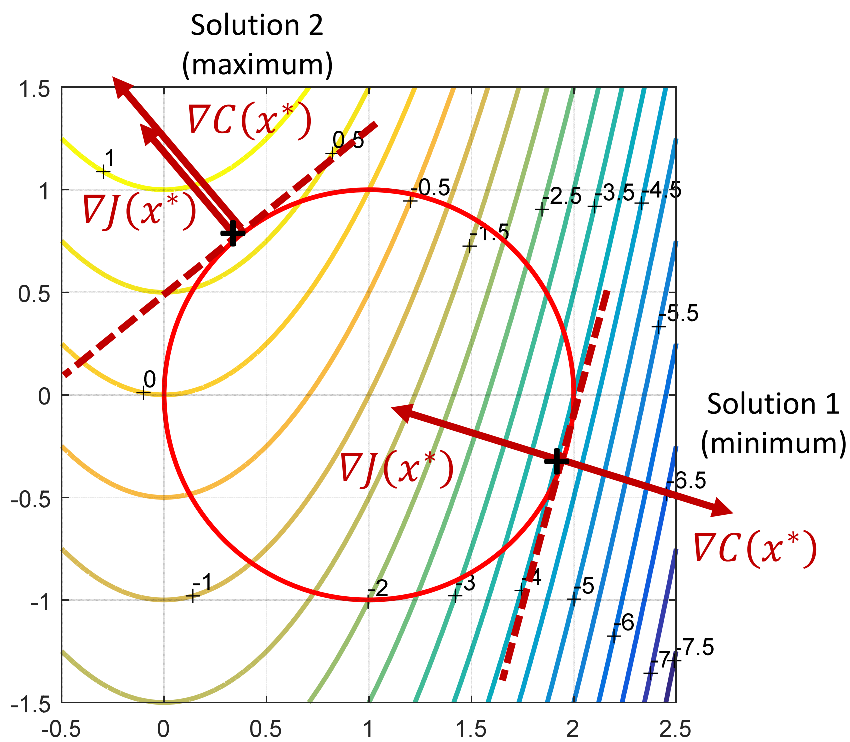
→Minimiser $J(x_1, x_2) = x_2 - {x_1}^2$, sous contrainte $\mathrm{C}(x_1, x_2) = (x_1 - 1)^2 + {x_2}^2 - 1 = 0$



Approche géométrique

→La condition $\nabla J(x^*) + \lambda.\nabla C(x^*) = 0$ exprime qu'au point x^* respectant la contrainte $C(x^*) = 0$, une condition nécessaire pour que J présente un extremum est que : $\nabla J(x^*)$ et $\nabla C(x^*)$ sont colinéaires, et donc qu'au point x^* l'hyperplan tangent à l'isovaleur $\mathrm{J}(x) = J(x^*)$ et l'hyperplan tangent à l'isovaleur $C(x) = 0$ sont confondus.

→ En 2D : recherche de points où la contrainte $C(x) = 0$ est tangente à une isovaleur de J.



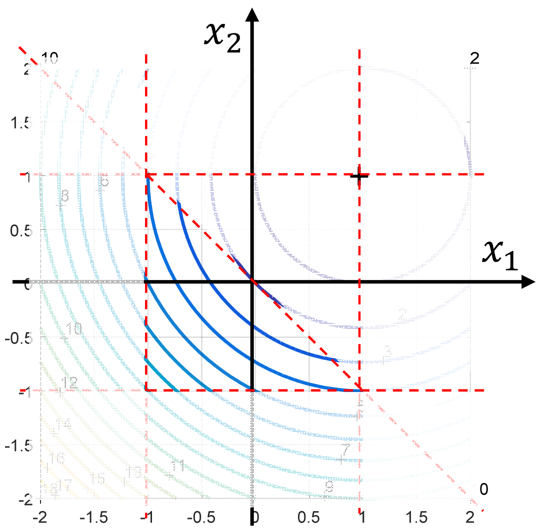
IV. Minimisation sous contraintes-inégalités

OPTIMISATION CONVEXE

Exemple 1

→Minimiser $J(x_1, x_2) = (x_1 - 1)^2 + (x_2 - 1)^2$,

→Sous contraintes : $\begin{cases} {x_1}^2 - 1 \leq 0 \\ {x_2}^2 - 1 \leq 0 \\ x_1 + x_2 \leq 0 \end{cases}$, qui s'écrivent aussi : $\begin{cases} -1 \leq x_1 \leq 1 & \Leftarrow \\ -1 \leq x_2 \leq 1 & \Leftarrow \\ x_1 + x_2 \leq 0 & \Leftarrow \end{cases}$



Isovaleurs de J

En l'absence de contrainte :
J est minimum en $(1,1)$, où elle vaut 0

En présence des contraintes :
J est minimum en $(0,0)$, où elle vaut 2

Problème de minimisation convexe

→ Formulation générale

Trouver $\boldsymbol{x}^*$ tel que :

$J(\boldsymbol{x}^*) = \inf \{ J(\boldsymbol{x}), \boldsymbol{x} \in \mathbb{R}^n \}$ n variables

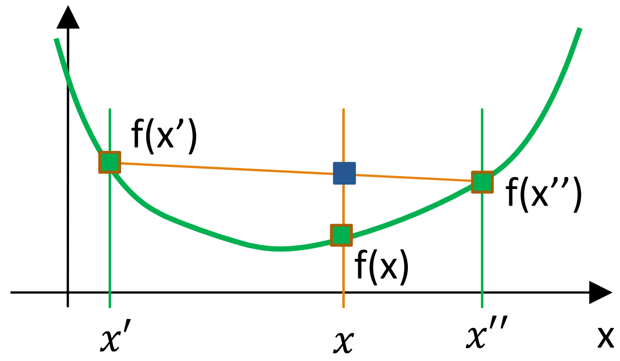
soumis à $C_j(\boldsymbol{x}) \leq 0, j = 1, m$ m contraintes

Problème convexe si J et C_j sont des fonctions convexes

Rappel : fonction convexe

→« Avec les mains » :

- Toute sécante est au-dessus de la courbe



→Définition mathématique :

- Soit $x \in]x', x''[$
- Alors : $\exists \lambda \in]0,1[$ tel que : $x = (1-\lambda).x' + \lambda.x''$

- La fonction f est convexe sur $]x', x''[$ ssi $\forall \mathrm{x} \quad \mathrm{f}(x) \leq (1-\lambda).\mathrm{f}(x') + \lambda.f(x'')$
- La fonction f est strictement convexe ssi $\forall \mathrm{x} \quad \mathrm{f}(x) < (1-\lambda).\mathrm{f}(x') + \lambda.f(x'')$

Propriétés des fonctions convexes

→ f convexe $\Longrightarrow f$ continue

→ Si f est de classe C^2, f strictement convexe $\Longleftrightarrow \forall x \quad H_f(x)$ défini positif

→ La somme de deux fonctions convexes est convexe

→ La composée de deux fonctions convexes est convexe

Rappel : ensemble convexe

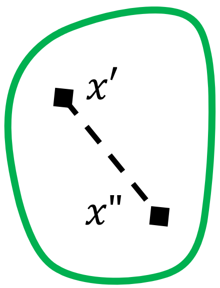
→« Avec les mains » :

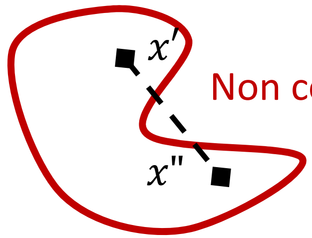
- L'ensemble D est convexe ssi tout segment reliant deux points de D est entièrement contenu dans D

→Définition mathématique :

- Soit $x \in]x', x"[$
- Alors : $\exists\lambda \in]0,1[$ tel que : $x = (1-\lambda).x' + \lambda.x"$
- L'ensemble D est convexe ssi $\forall (x', x") \in D^2 \;\; x \in]x', x"[\; \Longrightarrow x \in D$

Convexe

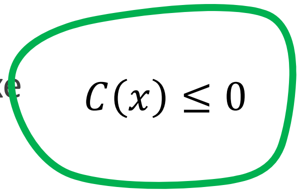




Non convexe

→Propriétés :

- Une intersection d'ensembles convexes est convexe
- L'ensemble défini par $C(x) \leq 0$ est convexe ssi la fonction C est convexe



Problème de minimisation convexe

→ Formulation générale :

Trouver $\boldsymbol{x}^*$ tel que :

$J(\boldsymbol{x}^*) = \inf \{ J(\boldsymbol{x}), \boldsymbol{x} \in \mathbb{R}^n \}$ n variables

soumis à $C_j(\boldsymbol{x}) \leq 0, j = 1, m$ m contraintes

Problème convexe si J et C_j sont des fonctions convexes

Les contraintes définissent un domaine admissible D convexe

Existence et unicité de la solution :

- S'il existe un minimum x^*, alors x^* est un minimum global
- Si J est strictement convexe, alors x^* est un minimum global unique

Cas 1D

→ Formulation du problème :

- J est fonction d'1 variable réelle x
- Minimiser J, sous contrainte $x \in D$, avec $D = [a, b]$

→ Deux situations sont possibles :

- Si la dérivée s'annule en un point de l'intervalle D :

  $\exists a \leq x^* \leq b$ tel que $J'(x^*) = 0$, alors x^* est le minimum de J sur$D$

- Si la dérivée ne s'annule en aucun point de l'intervalle D :
  - *Si* $J'(x) > 0$ *sur* D*, alors* $x^* = a$
  - *Si* $J'(x) < 0$ *sur*D*, alors* $x^* = b$

Exercice : faire un schéma pour chaque cas

Cas 1D

→ Illustration :

Cas général

→ Formulation du problème :

- J est fonction de $\boldsymbol{x} \in \mathbb{R}^n$
- Minimiser J, sous contrainte $\boldsymbol{x} \in D$, où D est un domaine convexe

→ Deux situations sont possibles :

- Si le gradient s'annule en un point du domaine D :

  $\exists\ x^* \in D$ tel que $\nabla J(x^*) = 0$, alors x^* est le minimum de J sur D

- Si le gradient ne s'annule en aucun point de l'intervalle D :
  - *S'il existe, le minimum se trouve sur la frontière du domaine D*

Voir exemple 1

- Il existe différents algorithmes pour traiter le problème

Théorème de Karush-Kuhn-Tucker

→ On introduit les multiplicateurs de Lagrange $\left(\lambda_j\right)_{j=1,m}$

→ Le problème admet une solution ssi le système suivant admet une solution :

$$\begin{cases} \begin{cases} \forall j = 1, m \quad \lambda_j . C_j(x) = 0 & (1) \\ \nabla J(x) + \sum_{j=1}^{m} \lambda_j . \nabla C_j(x) = 0 & (2) \end{cases} \\ \forall j = 1, m \;\; C_j(x) \leq 0 & (3) \\ \forall j = 1, m \quad \lambda_j \geq 0 & (4) \end{cases}$$

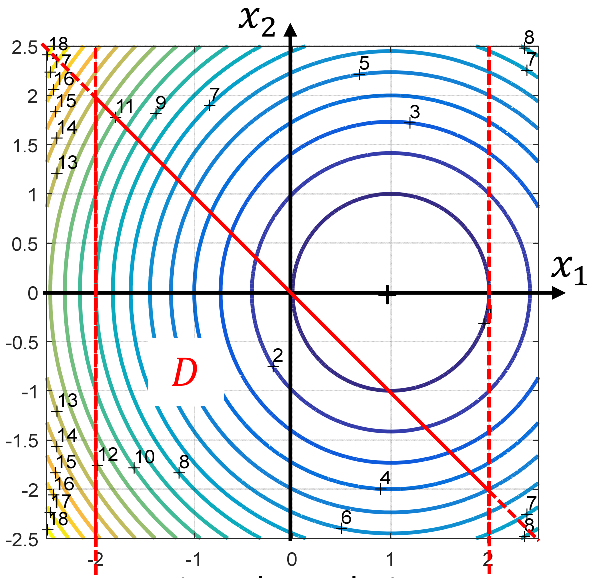
→(1) & (2) : système de $n + m$ équations à $n + m$ inconnues à résoudre

→(3) & (4) : conditions à vérifier a posteriori

Exemple 2

→Minimiser $J(x_1, x_2) = (x_1 - 1)^2 + {x_2}^2$,

→Sous contraintes : $\begin{cases} C_1(x) \leq 0 \\ C_2(x) \leq 0 \end{cases}$, où $\begin{cases} C_1(x) = x_1 + x_2 \\ C_2(x) = {x_1}^2 - 4 \end{cases}$



Isovaleurs de J

En l'absence de contrainte :
J est minimum en $(1{,}0)$, où elle vaut 0

En présence des contraintes :
On applique l'algorithme de KKT

florence.ossart@sorbonne-universite.fr

Exemple 2

→Le problème admet une solution ssi le système suivant admet une solution :

$$\begin{cases} \begin{cases} \forall j = 1, m \quad \lambda_j . C_j(x) = 0 \\ \nabla J(x) + \sum_{j=1}^{m} \lambda_j . \nabla C_j(x^*) = 0 \end{cases} \\ \forall j = 1, m \;\; C_j(x) \leq 0 \\ \forall j = 1, m \quad \lambda_j \geq 0 \end{cases}$$

→Pour le cas considéré : 2 contraintes, 2 multiplicateurs λ_1 et λ_2

Exemple 2

→ Système à résoudre :

$$\begin{cases} \begin{cases} \lambda_1 . C_1(x) = 0 \\ \lambda_2 . C_2(x) = 0 \\ \nabla J(x) + \lambda_1 . \nabla C_1(x) + \lambda_2 . \nabla C_2(x) = 0 \end{cases} & \text{(I)} \\ C_1(x) \leq 0 \\ C_2(x) \leq 0 \\ \lambda_1 \geq 0 \\ \lambda_2 \geq 0 & \text{(II)} \end{cases}$$

Étape 1 :
4 équations, 4 inconnues
=> On résout (I)

Étape 2 :
=> On vérifie si les solutions du sous-problème respectent (II)

→ $J(x_1, x_2) = (x_1 - 1)^2 + {x_2}^2$ $C_1(x) = x_1 + x_2$ $C_2(x) = {x_1}^2 - 4$

→ $\nabla J = \begin{pmatrix} 2(x_1 - 1) \\ 2x_2 \end{pmatrix}$ $\nabla C_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ $\nabla C_2 = \begin{pmatrix} 2x_1 \\ 0 \end{pmatrix}$

Exemple 2

→Système (I) :

$$\rightarrow \begin{cases} \lambda_1.(x_1 + x_2) = 0 & (1) \\ \lambda_2.({x_1}^2 - 4) = 0 & (2) \\ \begin{pmatrix} 2(x_1 - 1) \\ 2x_2 \end{pmatrix} + \lambda_1.\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda_2.\begin{pmatrix} 2x_1 \\ 0 \end{pmatrix} = 0 \end{cases} \Leftrightarrow \begin{cases} 2x_1 + \lambda_1 + 2\lambda_2 x_1 = 2 \\ 2x_2 + \lambda_1 = 0 \end{cases}$$

→4 cas à envisager pour satisfaire (1) et (2):

- $\lambda_1 = 0 \ / \ \lambda_1 \neq 0$
- $\lambda_2 = 0 \ / \ \lambda_2 \neq 0$

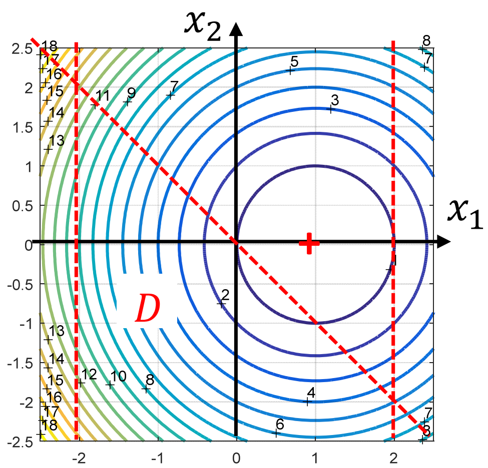
Exemple 2

→Système (I) :

$$\rightarrow \begin{cases} \lambda_1.(x_1 + x_2) = 0 & (1) \\ \lambda_2.({x_1}^2 - 4) = 0 & (2) \\ \begin{pmatrix} 2(x_1 - 1) \\ 2x_2 \end{pmatrix} + \lambda_1.\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda_2.\begin{pmatrix} 2x_1 \\ 0 \end{pmatrix} = 0 \end{cases} \Leftrightarrow \begin{cases} 2x_1 + \lambda_1 + 2\lambda_2 x_1 = 2 \\ 2x_2 + \lambda_1 = 0 \end{cases} \quad (3)$$

→Cas 1:

- $\lambda_1 = 0$ & $\lambda_2 = 0$: (1) et (2) satisfaites
- $(3) \Longrightarrow \begin{cases} 2x_1 = 2 \\ 2x_2 = 0 \end{cases}$, d'où $\begin{cases} x_1 = 1 \\ x_2 = 0 \end{cases}$
- Mais $C_1(x) = x_1 + x_2 = 1$,
- La contrainte $C_1(x) \leq 0$ n'est pas respectée
- Solution à rejeter



Exemple 2

→Système (I) :

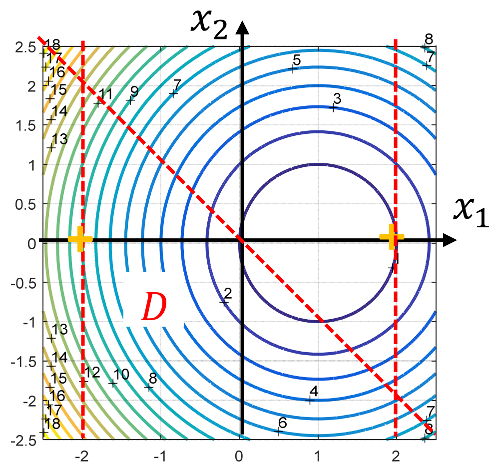
$$\rightarrow \begin{cases} \lambda_1.(x_1 + x_2) = 0 & (1) \\ \lambda_2.({x_1}^2 - 4) = 0 & (2) \\ \begin{pmatrix} 2(x_1 - 1) \\ 2x_2 \end{pmatrix} + \lambda_1.\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda_2.\begin{pmatrix} 2x_1 \\ 0 \end{pmatrix} = 0 \end{cases} \Leftrightarrow \begin{cases} 2x_1 + \lambda_1 + 2\lambda_2 x_1 = 2 \\ 2x_2 + \lambda_1 = 0 \end{cases} \quad (3)$$

→<u>Cas 2:</u>

- $\lambda_1 = 0$ & $\lambda_2 \neq 0$: (2) $\Longrightarrow$ $x_1 = -2$ ou $x_1 = +2$
- (3) $\Longrightarrow \begin{cases} 2x_1 + 2\lambda_2 x_1 = 2 \\ 2x_2 = 0 \end{cases}$, d'où $\begin{cases} \lambda_2 = \frac{1-x_1}{x_1} \\ x_2 = 0 \end{cases}$
- On trouve 2 résultats, mais aucun ne convient :

$$\begin{matrix} x_1 = -2 \\ x_2 = 0 \end{matrix} \in D \qquad \text{ou} \qquad \begin{matrix} x_1 = +2 \\ x_2 = 0 \end{matrix} \notin D\ !$$

$$\lambda_2 = -\frac{3}{2} < 0\ ! \qquad\qquad \lambda_2 = -\frac{1}{2} < 0\ !$$



Exemple 2

→Système (I) :

$$\rightarrow \begin{cases} \lambda_1.(x_1+x_2)=0 & (1) \\ \lambda_2.({x_1}^2-4)=0 & (2) \\ \begin{pmatrix} 2(x_1-1) \\ 2x_2 \end{pmatrix} + \lambda_1.\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda_2.\begin{pmatrix} 2x_1 \\ 0 \end{pmatrix} = 0 \end{cases}$$

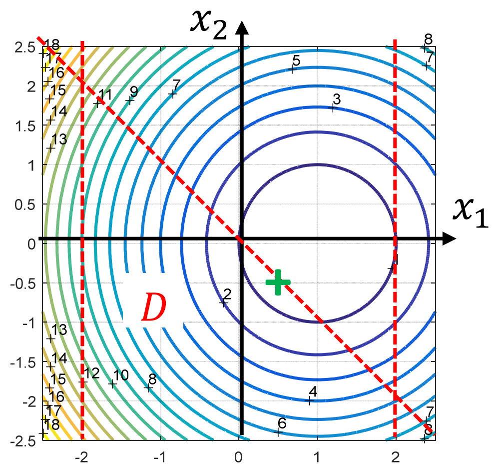
$$\Leftrightarrow \begin{cases} 2x_1+\lambda_1+2\lambda_2 x_1 = 2 \\ 2x_2+\lambda_1 = 0 \end{cases} \quad (3)$$

→<u>Cas 3:</u>

- $\lambda_2 = 0$ & $\lambda_1 \neq 0 : (1) \Longrightarrow x_2 = -x_1$
- $(3) \Longrightarrow \begin{cases} 2x_1+\lambda_1 = 2 \\ -2x_1+\lambda_1 = 0 \end{cases}$, d'où $\begin{cases} 4x_1 = 2 \\ \lambda_1 = 2x_1 \end{cases}$
- 1 résultat qui convient, c'est LA solution:

$$\begin{matrix} \boldsymbol{x_1 = 1/2} \\ \boldsymbol{x_2 = -1/2} \end{matrix} \in D$$

$$\boldsymbol{\lambda_1 = +1} \quad > 0$$



Signification des multiplicateurs

→ Un multiplicateur de Lagrange λ_j nul signifie que la contrainte associée est inopérante

→ Les valeurs de λ_j positives correspondent à des minima (la fonction coût augmente quand on se déplace vers l'intérieur du domaine)

- NB : les valeurs positives correspondent à des maxima

→ Plus la valeur de λ_j est grande, plus l'influence de la contrainte associée est grande

- Modifier la contrainte agira fortement sur le minimum