Chapitre IV

Optimisation avec contraintes

Dans cette leçon, nous voyons les principes d'optimisation avec contrainte :

- comment les contraintes sont définies et caractérisées,
- les conditions d'optimalité pour un programme avec contrainte,
- la méthode du simplexe (problème linéaire sous contraintes linéaires),
- le gradient projeté (projection de gradient du l'hyperplan tangent)
- le Lagrangien augmenté (pénalisation des contraintes avec les multiplicateurs de Lagrange)

Rappel : la quasi-totalité des transparents de ce cours provient du cours d'optimisation de Max Cerf (Université Paris 6 / Ariane Espace).

1 Introduction

Problème non linéaire sous contraintes

$$\min_{x \in R^n} f(x) \quad \text{sous} \quad \begin{cases} c_E(x) = 0 \\ c_I(x) \le 0 \end{cases}$$ $\rightarrow$ Problème noté (PO)

Catégories de problèmes

- **Programmation linéaire** $\rightarrow$ Fonctions f, c_E, c_I linéaires
- **Programmation non linéaire** $\rightarrow$ Fonctions f, c_E, c_I quelconques

Traitement des contraintes

- Méthodes de contraintes actives $\rightarrow$ Identification des inégalités actives
  Transformation en un **problème avec contraintes égalité**
  Respect des contraintes à chaque itération
- Méthodes de point intérieur $\rightarrow$ Fonction barrière (pénalisation intérieure)
  Suivi d'un **chemin central intérieur aux contraintes**
- Méthodes de pénalisation $\rightarrow$ Critère augmenté (pénalisation extérieure)
  Transformation en un **problème sans contraintes**

Problème non linéaire sous contraintes

$$\min_{x \in R^n} f(x) \quad \text{sous} \quad \begin{cases} c_E(x) = 0 \\ c_I(x) \le 0 \end{cases} \qquad \rightarrow \text{ Problème noté (PO)}$$

Classification des méthodes

	Méthode primale	Méthode primale-duale	Méthode duale
Problème traité	problème primal	problème primal	problème dual
Objectif	**min f** - méthode directe - point stationnaire	**solution KKT** - méthode indirecte - point stationnaire	**max w** - méthode indirecte - point col
Itérations	admissibles	admissibles ou non	non admissibles
Variables	primales x	primales x , duales λ	primales x , duales λ
Algorithmes	- simplexe (LP) - gradient projeté - pénalisation	- point intérieur (LP, NLP) - séquentiel quadratique	- simplexe dual (LP) - lagrangien augmenté - Uzawa

2 Contraintes

2.1 Définition

a) Solution admissible

Qu'est-ce qu'une solution admissible ?

Solution admissible

x solution admissible de (PO) $\Leftrightarrow$ x satisfait les contraintes $\begin{cases} c_E(x) = 0 \\ c_I(x) \le 0 \\ x \in X \end{cases}$
(ou point admissible)

Ensemble admissible

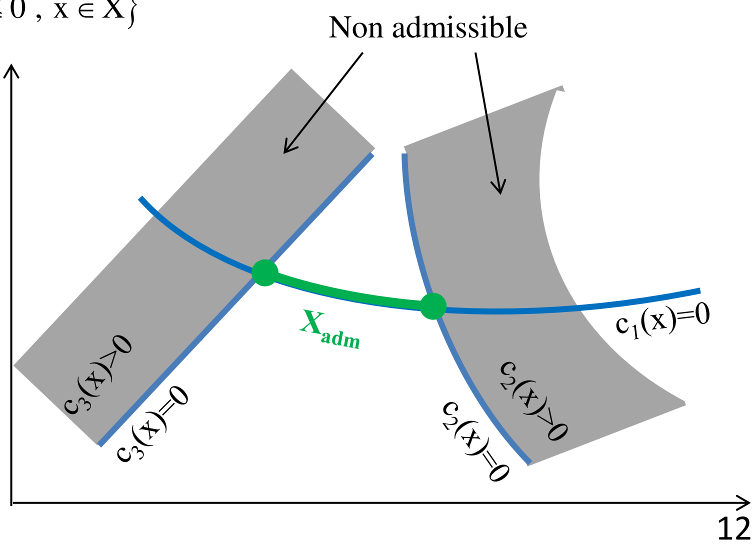
$$X_{adm} = \{x \in R^n \,/\, c_E(x) = 0,\ c_I(x) \le 0\ ,\ x \in X\}$$

Illustration dans R^2

- $c_1(x) = 0 \ \rightarrow$ courbe
- $c_2(x) \le 0 \ \rightarrow$ région du plan
- $c_3(x) \le 0 \ \rightarrow$ région du plan

Dans R^n

- $c(x) = 0 \ \rightarrow$ **hypersurface** (dimension n-1)
- $a^T x = 0 \ \rightarrow$ **hyperplan** $\perp a \in R^n$ (linéaire)



Erratum : $c_3(x) \ge 0$ (aussi une région du plus pour une contrainte ≥ 0)

b) Contrainte active

Une contrainte active est une contrainte qui atteinte (on est sur le contour de la contrainte).

Contrainte active

Une contrainte du problème (PO) est **active** (ou **saturée**) en x si elle s'annule en x.

Ensemble des contraintes actives

$$C_{act}(x) = \{j \,/\, c_{Ej}(x) = 0\} \bigcup \{j \,/\, c_{Ij}(x) = 0\}$$

- Contrainte égalité c_E : x admissible $\Rightarrow c_E(x) = 0 \Rightarrow c_E$ active en x
- Contrainte inégalité c_I : x admissible $\Rightarrow c_I(x) \le 0 \Rightarrow c_I$ active en x si $c_I(x) = 0$
  c_I inactive en x si $c_I(x) < 0$

Intérêt

- Les contraintes inégalité inactives n'ont pas d'influence sur la solution x* du problème (PO). On peut les ignorer, si on identifie l'ensemble $C_{act}(x^*)$. Mais x* n'est pas connu au départ …
- **Le problème (PO) est équivalent au problème $(PO)_{act}$ réduit aux contraintes actives prises comme des contraintes égalité.**

$$\min_{x \in R^n} f(x) \text{ sous } \begin{cases} c_E(x) = 0 \\ c_I(x) \le 0 \end{cases} \Leftrightarrow \min_{x \in R^n} f(x) \text{ sous } c_j(x) = 0,\ j \in C_{act}(x^*) \quad \text{noté} \quad \boxed{\min_{x \in R^n} f(x) \text{ sous } c(x) = 0}$$

Exemple :

$$\min_{x \in R} x^2 + 1 \text{ sous } \begin{cases} x \ge 1 \\ x \le 2 \end{cases} \rightarrow \quad x^* = 1$$

1. Minimum <u>sans contrainte</u>

$$\min_{x \in R} x^2 + 1 \rightarrow \quad x^* = 0$$

- Respecte la contrainte $x \le 2$
- Ne respecte pas la contrainte $x \ge 1$

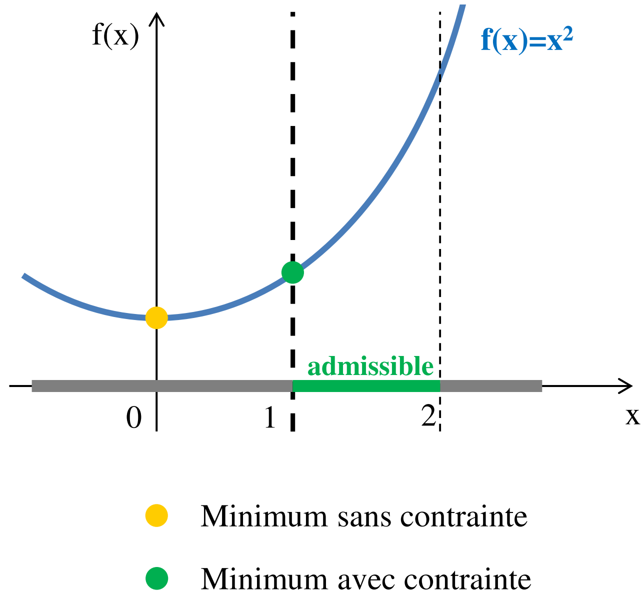
$\rightarrow$ Activation de la contrainte $x = 1$

2. Minimum <u>avec contrainte active x = 1</u>

$$\min_{x \in R} x^2 + 1 \text{ sous } x = 1 \rightarrow \quad x^* = 1$$

- Respecte la contrainte $x \le 2$
- Respecte la contrainte $x \ge 1$

$\rightarrow$ Solution du problème



3. Bilan : 1 contrainte active $x \ge 1$ $\rightarrow$ transformée en égalité
   1 contrainte inactive $x \le 2$ $\rightarrow$ ignorée

Donc, si une contrainte est active, la solution vérifie cette contrainte transformée en égalité.

c) Point intérieur

Un point intérieur est un point à l'intérieur de l'espace des valeurs admissibles (contraintes comprises).

Point intérieur

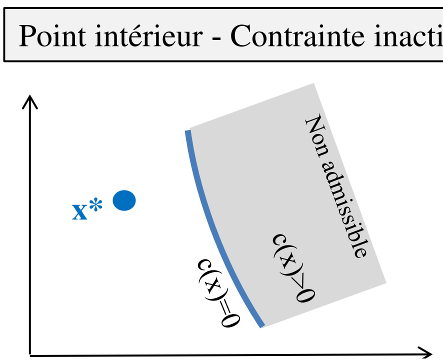
y point intérieur à Y

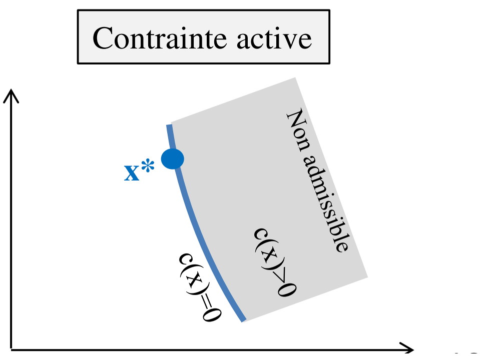
$\Leftrightarrow$ Il existe un voisinage de y contenu dans Y : $\exists \varepsilon > 0 \ / \ \forall z, \|z - y\| \le \varepsilon \ , \ z \in Y$

Un problème avec contraintes égalité n'admet pas de point intérieur

Solution intérieure aux contraintes

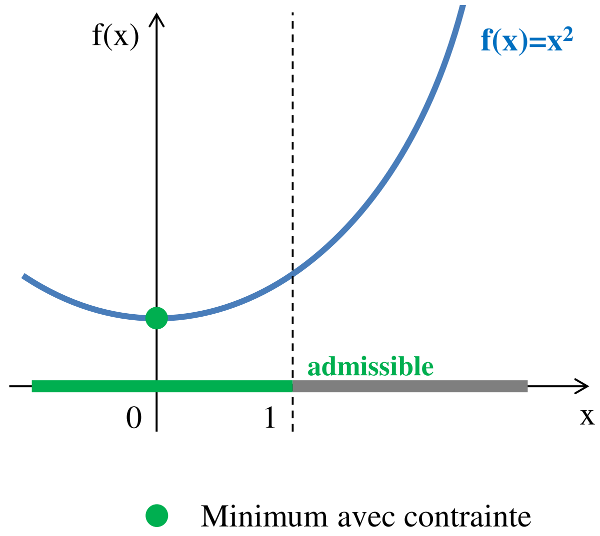
- x^* minimum local du problème avec contraintes inégalité $\min_{x \in R^n} f(x)$ sous $c_I(x) \le 0$
- Si x^* est un point intérieur, alors
  x^* minimum local du problème sans contraintes $\min_{x \in R^n} f(x)$ $\rightarrow$ plus simple





Exemple :

$$\min_{x\in R} x^2+1 \text{ sous } x\le 1 \;\rightarrow\; x^*=0$$



1. <u>Ensemble admissible</u>

   $X_{adm}=\{x\in R\,/\,x\le 1\} \;=\;]-\infty,1]$

2. <u>Ensemble intérieur à la contrainte</u>

   $X_{int}=\{x\in R\,/\,x<1\} \;=\;]-\infty,1[$

   $X_{int}=X_{adm}-\{1\}$

   $x\in X_{int}$ $\rightarrow$ voisinage de x inclus dans X_{int}
   $\rightarrow$ intervalle ouvert

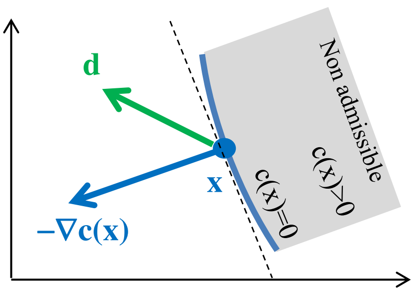
3. Solution : $x^*=0$

   $x^*\in X_{int}$ intérieur à la contrainte $\rightarrow$ <u>contrainte inactive</u>

d) Direction admissible

Direction admissible

$d\in R^n$ direction de déplacement à partir de $x\in X_{adm}$ point admissible



Définition : d direction admissible

$$\Leftrightarrow \exists\varepsilon>0\,/\,\forall s,\; 0<s\le\varepsilon \;\Rightarrow\; x+sd\in X_{adm}$$

On peut se déplacer d'au moins ε suivant d à partir de x en restant admissible

- Contrainte égalité : $\nabla c_E(x)^T d=0 \rightarrow$ tangent
- Contrainte inégalité : $\nabla c_I(x)^T d\le 0 \rightarrow$ intérieur

Ensemble convexe

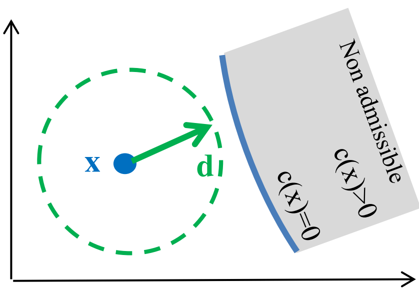
X_{adm} convexe, $y\neq x,\; x,y\in X_{adm} \;\Rightarrow\; [x,y]\subset X_{adm}$

$\Rightarrow$ $d=y-x$ est une direction admissible à partir de x

Point intérieur

x point intérieur à X_{adm}

$\Rightarrow$ Toute direction $d\in R^n$ est admissible à partir de x



Pour une contrainte

- d'égalité : il faut suivre la surface correspondant à l'égalité, la direction est tangente (orthogonale à la normale à la surface)
- d'inégalité : vers l'intérieur du domaine de la contrainte.

2.2 Contraintes linéaires

Dans le cas des contraintes linéaires :

- une contrainte linéaire définit un espace délimité par des hyperplans (= polytope).
- la forme standard d'une contrainte linéaire permet de la réexprimer sous une forme où toutes les contraintes sont des contraintes d'égalités et les variables positives.
- sous cette forme, chaque résolution partielle permet d'obtenir l'un des sommets du polytopes.
- les directions admissibles peuvent être alors engendrées à partir de ces sommets.

Des contraintes linéaires sont sous la forme standard s'il n'y a que des contraintes d'égalité, avec des variables positives.

Contraintes linéaires

Contraintes linéaires sous forme standard : $\begin{cases} Ax = b \\ x \geq 0 \end{cases}$

$X_{adm} = \{x \in R^n / Ax = b,\ x \geq 0\}$

Direction admissible

d $\in R^n$ direction admissible à partir de x point admissible $\Leftrightarrow \begin{cases} Ad = 0 \\ d_i \geq 0 \text{ si } x_i = 0 \end{cases}$

Preuve :

Pour s>0 petit, on doit avoir : $(x+sd) \in X_{adm} \Leftrightarrow \begin{cases} A(x+sd) = b \\ x+sd \geq 0 \end{cases} \Leftrightarrow \begin{cases} Ad = 0 \quad \text{car } Ax = b \\ x+sd \geq 0 \end{cases}$

Si $x_i > 0$*, alors* $x_i + sd_i > 0$ *pour s assez petit*

Si $x_i = 0$*, alors* $x_i + sd_i \geq 0$ *si* $d_i \geq 0$

Combinaison de directions admissibles

Toute combinaison linéaire à coefficients positifs de directions admissibles est une direction admissible.

Preuve : Une combinaison linéaire à coefficients positifs vérifie également $\begin{cases} Ad = 0 \\ d_i \geq 0 \text{ si } x_i = 0 \end{cases}$

a) Direction admissible

Un polytope est une zone de l'espace délimitée par un ensemble d'hyperplan.

Définition

Polytope P dans R^n $\quad \boxed{P = \{x \in R^n \ / \ Ax \le b\}} \quad A \in R^{m \times n}, \ b \in R^m$

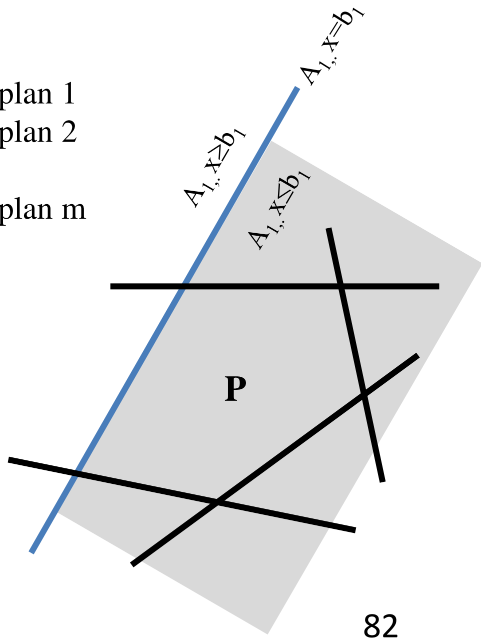
Interprétation géométrique

$$Ax = b \iff \begin{cases} A_{1,1}x_1 + A_{1,2}x_2 + \cdots + A_{1,n}x_n = b_1 & \to \text{hyperplan 1} \\ A_{2,1}x_1 + A_{2,2}x_2 + \cdots + A_{2,n}x_n = b_2 & \to \text{hyperplan 2} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ A_{m,1}x_1 + A_{m,2}x_2 + \cdots + A_{m,n}x_n = b_m & \to \text{hyperplan m} \end{cases}$$

Chaque hyperplan j sépare R^n en 2 sous-espaces : $\begin{cases} A_{j,.}x \le b_j \\ A_{j,.}x \ge b_j \end{cases}$

P = ensemble de points de R^n délimité par m hyperplans

→ Polytope dans R^2 = **polygone**
→ Polytope dans R^3 = **polyèdre**



82

Noter qu'un polytope est **toujours convexe**.

On obtient la forme standard en :

- supprimant les contraintes redondantes (exemple : $-1 \le x \le 12$ et $x \ge 5$ sont redondants, à réduire en $5 \le x \le 12$),
- en ajoutant des variables ou en effectuant des changements de variable,

de manières à ce que l'ensemble des variables exprimant les contraintes soient positives.

Forme standard

Polytope P dans R^n sous forme standard

$$\boxed{P = \{x \in R^n \ / \ Ax = b,\ x \geq 0\}}$$

$A \in R^{m \times n}$, $b \in R^m$

A de rang plein (élimination contraintes redondantes $\to \tilde{A}$)

Passage sous forme standard

- Contrainte inégalité : Transformation en contrainte égalité
  Ajout d'une **variable d'écart** positive

$$c(x) \leq b \iff \begin{cases} c(x)+z=b \\ z \geq 0 \end{cases} \iff \begin{cases} c'(x)=b \\ z \geq 0 \end{cases} \quad \text{avec} \quad c'(x)=c(x)+z$$

$$c(x) \geq b \iff \begin{cases} c(x)-z=b \\ z \geq 0 \end{cases} \iff \begin{cases} c'(x)=b \\ z \geq 0 \end{cases} \quad \text{avec} \quad c'(x)=c(x)-z$$

- Contraintes de bornes : Changement de variable $\to$ borne inférieure
  Ajout d'une **variable d'écart** positive $\to$ borne supérieure

$$x_l \leq x \leq x_u \iff 0 \leq x - x_l \leq x_u - x_l \iff \begin{cases} x' = x - x_l\,,\ x' \geq 0 \\ x' \leq x_u - x_l \end{cases} \iff \begin{cases} x' = x - x_l\,, & x' \geq 0 \\ x' + z = x_u - x_l\,, & z \geq 0 \end{cases}$$

- Variable libre : Différence de 2 variables positives

$x \in R \iff x = z - y,\ y, z \geq 0$ 83

Mise sous forme standard

- Problème linéaire (P)

$$\min_{x_1,x_2,x_3} x_1 + 2x_2 + 3x_3 \quad \text{sous} \begin{cases} -x_1 + 3x_2 = 5 \\ 2x_1 - x_2 + 3x_3 \geq 6 \\ x_1 \in R,\ x_2 \geq 1,\ x_3 \leq 4 \end{cases}$$

- Changement de variables pour les bornes

$$\begin{cases} x_1 \in R \\ x_2 \geq 1 \\ x_3 \leq 4 \end{cases} \iff \begin{cases} x_1 = z_1 - y_1 & \to y_1, z_1 \geq 0 \\ x_2' = x_2 - 1 & \to x_2' \geq 0 \\ x_3' = 4 - x_3 & \to x_3' \geq 0 \end{cases}$$

- Variables d'écart pour les contraintes inégalité

$$2x_1 - x_2 + 3x_3 \geq 6 \iff 2x_1 - x_2 + 3x_3 - z_2 = 6 \quad \to z_2 \geq 0$$

- Problème équivalent à (P) sous forme standard

$$\min_{y_1,z_1,z_2,x_2',x_3'} z_1 - y_1 + 2x_2' - 3x_3' + 14 \quad \text{sous} \begin{cases} y_1 - z_1 + 3x_2' = 2 \\ 2z_1 - 2y_1 - x_2' - 3x_3' - z_2 = -5 \end{cases}$$

Exercice 24: Forme standard

Mettre sous forme standard les problèmes linéaires suivants :

1. $\min_{x_1,x_2} 3x_1 - x_2$ sous $\begin{cases} x_1 \geq 1 \\ 0 \leq x_2 \leq 5 \end{cases}$

2. $\min_{x_1,x_2} x_2 - 2x_1$ sous $\begin{cases} 2 \leq x_1 \leq 8 \\ x_2 \leq x_1 \leq x_2 + 2 \\ x_2 \geq 0 \end{cases}$

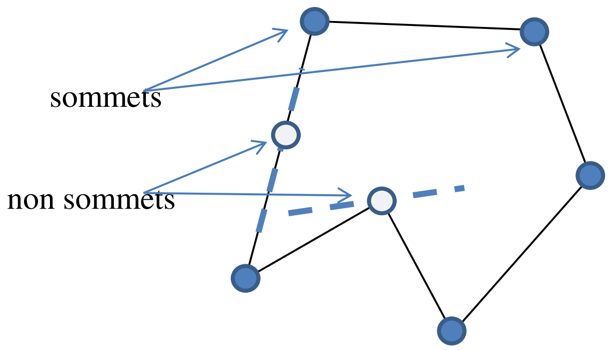
Sommet

Polytope P dans R^n sous forme standard

$P = \{x \in R^n \ / \ Ax = b, \ x \geq 0\}$ $A \in R^{m \times n}, \ b \in R^m$

A de rang plein : rang(A)=r=m≤n

> **Définition**
>
> $x \in P$ est un sommet de P
>
> $\Leftrightarrow$
>
> On ne peut pas trouver $y,z \in P$, différents de x tels que x soit combinaison convexe de y et z
> i.e. $x = \lambda y + (1-\lambda)z$ avec $0<\lambda<1$



Existence

Tout polytope non vide possède au moins un sommet.

Sur la figure ci-dessus, les non-sommets sont obtenus comme une combinaison linéaire de deux autres sommets (premier : le sommet de l'arête, le second à partir des deux sommets distants)

Une base est obtenue en prenant *m* colonnes parmi les *n* possibles dans la matrice *A* du système linéaire.

La résolution du système linéaire fournit un point correspondant à l'intersection de *m* parmi *n* hyperplans possibles.

Si ce point est admissible (car la solution peut être l'extérieur du domaine), constitue un sommet du polytope.

Base

Polytope P dans R^n sous forme standard

$P = \{x \in R^n \ / \ Ax = b, \ x \geq 0\}$ $A \in R^{m \times n}, \ b \in R^m$

- A est de rang plein r=m≤n $\Rightarrow$ Il existe m colonnes indépendantes.
- On choisit une sous-matrice $B \in R^{m \times m}$ de rang plein (parmi C_n^m combinaisons possibles)

$$AE = \begin{pmatrix} \overset{m}{B} & \overset{n-m}{N} \end{pmatrix} = \begin{array}{c} \\ 1 \\ \vdots \\ m \end{array} \begin{array}{c} \overbrace{\begin{array}{ccc} 1 & \cdots & m \end{array}}^{B} \quad \begin{array}{ccccccc} m+1 & \cdots & k-1 & \overset{N}{k} & k+1 & \cdots & n \end{array} \\ \left(\begin{array}{ccc|ccc|c|ccc} \times & \cdots & \times & \times & \cdots & \times & \times & \times & \cdots & \times \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \times & \cdots & \times & \times & \cdots & \times & \times & \times & \cdots & \times \end{array}\right) \end{array}$$

E matrice de permutation des colonnes de A : $EE^T=I$

$A_{.,k} =$ $k^{ème}$ colonne de AE

- B = **matrice de base** $\rightarrow B \in R^{m \times m}$ inversible
- N = **matrice hors base** $\rightarrow N \in R^{m \times (n-m)}$

Ci-dessous, si *B* est constitué par le colonne d'indice $(i_1, \ldots, i_n)$ de *A*, alors x_b est construit comme

$x_B = \begin{bmatrix} x_{i_1} & \dots & x_{i_n} \end{bmatrix}$.

Identification des sommets

Polytope P dans R^n sous forme standard

$P = \{x \in R^n \ / \ Ax = b, \ x \geq 0\}$ $\quad A \in R^{m\times n}, \ b \in R^m$

- Choix d'une base $B \in R^{m\times m}$

$$AE = \begin{array}{cc} m & n-m \end{array} \atop \begin{pmatrix} B & N \end{pmatrix} \qquad E^T x = \begin{pmatrix} x_B \\ x_N \end{pmatrix}$$

- $x_B \in R^m$ = variables en base (ou liées ou dépendantes)
  $x_N \in R^{n-m}$ = variables hors base (ou libres ou indépendantes)

- **Point admissible** : $x \in P \Leftrightarrow \begin{cases} Ax = b \\ x \geq 0 \end{cases} \Leftrightarrow \begin{cases} Bx_B + Nx_N = b \\ x \geq 0 \end{cases} \Leftrightarrow \boxed{\begin{cases} x_B = B^{-1}(b - Nx_N) \\ x \geq 0 \end{cases}}$

- Identification des sommets

Tout point x tel que : $\begin{cases} x_B = B^{-1}b \geq 0 \\ x_N = 0 \end{cases} \Rightarrow E^T x = \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}$ est un sommet du polytope.

$\rightarrow$ « **solutions de base** »

Donc un point obtenu est admissible s'il est dans le domaine des contraintes sous forme normale, à savoir $x_B \geq 0$ et $x_N = 0$.

Pour résumer les propriétés d'une solution de base :

Solution de base

Polytope P dans R^n sous forme standard

$P = \{x \in R^n \ / \ Ax = b, \ x \geq 0\}$ $\quad A \in R^{m\times n}, \ b \in R^m$

A de rang plein : rang(A)=r=m≤n

Définition

$x \in R^n$ est une **solution de base** de P

$\Leftrightarrow$

Il existe m indices $i_1,\dots,i_m$ tels que

- La matrice $B \in R^{m\times m}$ composée des colonnes $i_1,\dots,i_m$ de A est de rang plein
- Les n-m composantes x_i, $i \neq i_1,\dots,i_m$ sont nulles $\rightarrow x_N = 0$
- x vérifie Ax=b $\rightarrow x_B = B^{-1}b$

$\Rightarrow E^T x = \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}$

Solution de base admissible

Une solution de base x est **admissible** si toutes ses composantes sont positives ($x \in P$).

x vérifie également $x \geq 0 \Rightarrow x_B = B^{-1}b \geq 0$

$\rightarrow$ **Base admissible ou réalisable**, solution de base admissible ou réalisable

Solution de base dégénérée

Une solution de base x est **dégénérée** si plus de n-m composantes de x sont nulles.

x_N=0 par définition (n-m composantes) $\Rightarrow x_B$ comporte des composantes nulles

Lien avec les sommets du polytopes :

Lien sommet – solution de base

Polytope P dans R^n sous forme standard

$P = \{x \in R^n \ / \ Ax = b, \ x \geq 0\}$ $\quad A \in R^{m\times n}, \ b \in R^m$

A de rang plein : rang(A)=r=m≤n

- I=indices des composantes nulles en $x^* \in P$: $I^* = \{i \,/\, x_i^* = 0\}$
  (= contraintes inégalités actives)
- S=variété linéaire définie par : $S^* = \{x \in R^n \ / \ Ax = b, \ x_i = 0, \forall i \in I^*\}$

$x^* \in R^n$ est un **sommet** de P $\Leftrightarrow$ $S^* = \{x^*\}$ $\Leftrightarrow$ x* est une **solution de base admissible** de P

Lien sommet – contraintes actives

$x^* \in P$ est un **sommet** de P $\Leftrightarrow$ Au moins **n contraintes sont actives** en x*

- m contraintes égalité : $Ax^* = b$
- n-m contraintes inégalité : $x_N^* = 0$
  Les m contraintes inégalité sur x_B peuvent être actives ou non :
  $x_B^* = B^{-1}b \geq 0 \quad \rightarrow$ dégénérescence

Recherche des solutions de base

Polytope P dans R^4 sous forme standard

$$P = \{(x_1, x_2, x_3, x_4) \,/\, Ax = b, \ x \geq 0\} \quad \text{avec} \quad A = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{pmatrix}, \ b = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

- On utilise les contraintes pour réduire le problème à (x_1,x_2)

$$Ax = b \Leftrightarrow \begin{cases} x_1 + x_2 + x_3 = 1 \\ x_1 - x_2 + x_4 = 1 \end{cases} \Leftrightarrow \begin{cases} x_3 = 1 - x_1 - x_2 \\ x_4 = 1 - x_1 + x_2 \end{cases}$$

$$x \geq 0 \quad \Rightarrow \begin{cases} x_3 \geq 0 \\ x_4 \geq 0 \end{cases} \Rightarrow \begin{cases} x_1 + x_2 \leq 1 \\ x_1 - x_2 \leq 1 \end{cases}$$

- Polytope P' réduit dans R^2

$$P' = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \,/\, \begin{cases} x_1 + x_2 \leq 1 \\ x_1 - x_2 \leq 1 \end{cases}, \ \begin{cases} x_1 \geq 0 \\ x_2 \geq 0 \end{cases} \right\}$$

- Les polytopes P dans R^4 et P' dans R^2 sont équivalents : $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \in P \Leftrightarrow \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in P'$

  Remarque : P = forme standard de P'

Recherche des solutions de base

- Représentation de P' dans R^2

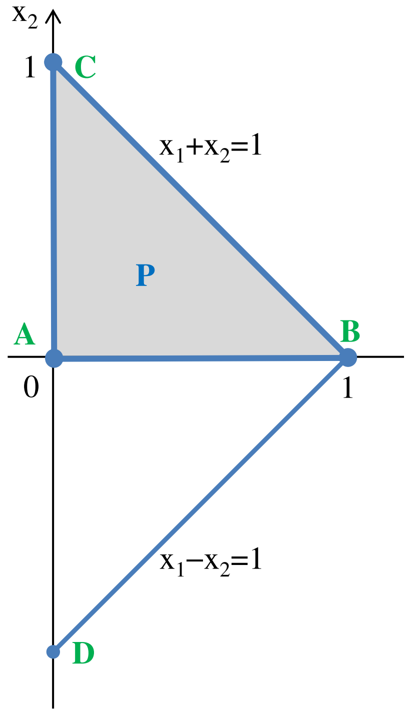
$$P' = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} / \begin{cases} x_1 + x_2 \le 1 \\ x_1 - x_2 \le 1 \end{cases}, \begin{cases} x_1 \ge 0 \\ x_2 \ge 0 \end{cases} \right\}$$

$\rightarrow$ représentation des valeurs possibles de (x_1,x_2) pour $(x_1,x_2,x_3,x_4) \in P$

- Contraintes de P : $Ax = b \Leftrightarrow \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$
  4 variables – 2 contraintes

- **Base de P**
  $\rightarrow$ choisir 2 colonnes indépendantes de A
  $\rightarrow$ 6 combinaisons possibles

- **Solution de base**
  $\rightarrow$ fixer les 2 variables hors base x_N à 0
  $\rightarrow$ calculer les 2 variables de base x_B pour vérifier Ax=b
  $\rightarrow$ base admissible si $x_B \ge 0$



Recherche des solutions de base

Examen des 6 bases possibles de P

$$A = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad x_B = B^{-1}b$$

- **Base (x_1,x_2)** : $B = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, $B^{-1} = \begin{pmatrix} 0,5 & 0,5 \\ 0,5 & -0,5 \end{pmatrix}$, $x_B = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

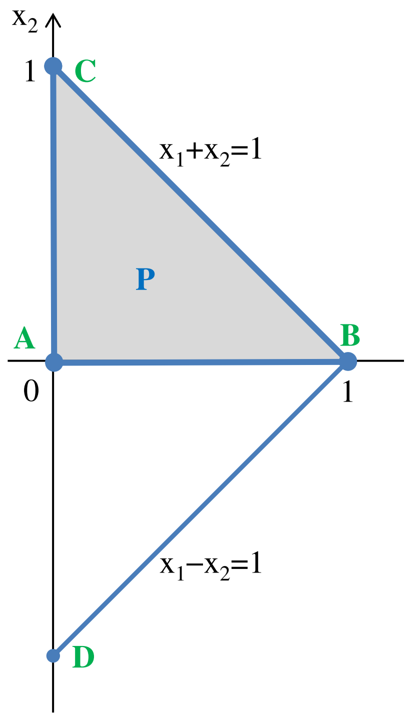
  $x = (1\ \ 0\ \ 0\ \ 0)$ admissible $\rightarrow$ **point B**

- **Base (x_1,x_3)** : $B = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, $B^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}$, $x_B = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

  $x = (1\ \ 0\ \ 0\ \ 0)$ admissible $\rightarrow$ **point B**

- **Base (x_1,x_4)** : $B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, $B^{-1} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$, $x_B = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

  $x = (1\ \ 0\ \ 0\ \ 0)$ admissible $\rightarrow$ **point B**



Recherche des solutions de base

Examen des 6 bases possibles de P

$$A = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad x_B = B^{-1}b$$

- **Base (x_2,x_3)** : $B = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$, $B^{-1} = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}$, $x_B = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$

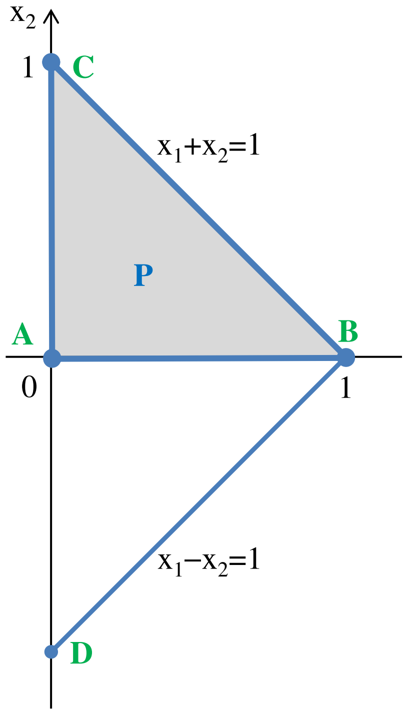
  $x = (0 \ \ -1 \ \ 2 \ \ 0)$ non admissible → **point D**

- **Base (x_2,x_4)** : $B = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$, $B^{-1} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, $x_B = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

  $x = (0 \ \ 1 \ \ 0 \ \ 2)$ admissible → **point C**

- **Base (x_3,x_4)** : $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $B^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $x_B = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

  $x = (0 \ \ 0 \ \ 1 \ \ 1)$ admissible → **point A**



95

Exercice 25: Recherche des solutions de base

$$\min_{x_1,x_2} x_1 + x_2 \text{ sous } \begin{cases} x_1 \leq 1 \\ x_2 \leq 1 \end{cases}$$

1. Donner la forme standard de ce problème.
2. Écrire le polytope des contraintes sous forme matricielle.
3. Combien y-a-t-il de bases possibles ?
4. Examiner l'ensemble des solutions de base. On précisera à chaque fois si la solution est admissible.
5. Représenter les contraintes sur un dessin, et l'ensemble des solutions trouvées à la question précédente.

On se déplace à partir d'un sommet à partir d'une solution de base.

Direction de déplacement à partir d'un sommet

Polytope P dans R^n sous forme standard

$P = \{x \in R^n \ / \ Ax = b,\ x \ge 0\}$ $\quad A \in R^{m\times n}$ de rang plein, $b \in R^m$

$x \in R^n$ solution de base admissible de P : $\boxed{E^T x = \begin{pmatrix} x_B \\ x_N \end{pmatrix} = \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix} \ge 0}$

$d \in R^n$ direction de déplacement : $\boxed{E^T d = \begin{pmatrix} d_B \\ d_N \end{pmatrix}}$

Direction admissible

d direction admissible en x $\Leftrightarrow \begin{cases} Ad = 0 \\ d_i \ge 0 \text{ si } x_i = 0 \end{cases}$ (contraintes linéaires)

$$Ad = 0 \ \Leftrightarrow\ Bd_B + Nd_N = 0 \ \Leftrightarrow\ d_B = -B^{-1}Nd_N$$

d direction admissible en x $\Leftrightarrow \begin{cases} d_B = -B^{-1}Nd_N \\ d_N \ge 0 \text{ car } x_N = 0 \\ d_{Bi} \ge 0 \text{ si } \quad x_{Bi} = 0 \quad \text{(solution de base dégénérée)} \end{cases}$

→ « **directions de base** »

La direction à prendre dans la base est $d_B = -B^{-1}Nd_N$.

Attention, si la base est dégénérée (= l'une des coordonnées de la solution de base nulle), alors la coordonnée de la direction doit être positive (sens de $d_{B_i} \ge 0$ si $x_{B_i} = 0$).

La $k^{ème}$ direction de base consiste, à partir d'une solution de base admissible, de prendre la $k^{ème}$ direction associé à une variable hors base.

Direction de base

$x \in R^n$ solution de base admissible de P : $E^T x = \begin{pmatrix} x_B \\ x_N \end{pmatrix} = \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix} \ge 0$

k = indice d'une variable hors-base

d^k = $k^{ème}$ direction de base en x : $\quad E^T d^k = \begin{pmatrix} d_B^k \\ d_N^k \end{pmatrix}$ noté $\begin{pmatrix} d_B \\ d_N \end{pmatrix}$

- Les **composantes d_N** sur les variables hors base sont toutes nulles, sauf sur la variable x_k

$$E^T \begin{pmatrix} 0 \\ d_N \end{pmatrix} = e_k = \begin{pmatrix} \overset{1}{0} & \cdots & \overset{k-1}{0} & \overset{k}{1} & \overset{k+1}{0} & \cdots & \overset{n}{0} \end{pmatrix}^T$$

- Les **composantes d_B** sur les variables en base vérifient la $1^{ère}$ condition de direction admissible

$$Ad = 0 \ \Rightarrow\ Bd_B + Nd_N = 0 \ \Rightarrow\ d_B = -B^{-1}Nd_N = -B^{-1} \sum_{j \text{ hors base}} A_{.,j} d_j$$

$$\Rightarrow\ d_B = -B^{-1}A_{.,k} \qquad (A_{.,k} = k^{ème} \text{ colonne de AE})$$

- **Définition**
  La $k^{ème}$ direction de base en x est : $E^T d^k = \begin{pmatrix} d_B^k \\ d_N^k \end{pmatrix} = \begin{pmatrix} -B^{-1}A_{.,k} \\ 0 \end{pmatrix} + e_k = \begin{pmatrix} \overset{1}{d_1} \cdots \overset{m}{d_m} & \overset{m+1}{0} \cdots \overset{k-1}{0} & \overset{k}{1} & \overset{k+1}{0} \cdots \overset{n}{0} \end{pmatrix}^T$
- Interprétation géométrique : directions de base = arêtes du polytope en x

= dans un base admissible, la direction de base est obtenue en prenant chaque colonne de A hors base (notée $A_{.k}$) et à calculer $-B^{-1}.A_{.k}$, les autres colonnes à 0 sauf 1 à la $k^{ème}$ colonne.

Il faut également vérifier que cette $k^{ème}$ direction de base est admissible.

Direction de base admissible

$x \in R^n$ solution de base admissible de P, k = indice d'une variable hors-base

La $k^{ème}$ direction de base d^k en x vérifie par définition : $\begin{cases} Ad = 0 \\ d_N \geq 0 \end{cases}$

Pour que d^k soit une direction admissible, il faut également vérifier : $d_{Bi} \geq 0$ si $x_{Bi} = 0$

Cas d'une base non dégénérée

x solution de base admissible non dégénérée ($x_B > 0$)
$\Downarrow$
Toutes les directions de base en x sont admissibles

Combinaison de directions de base

$x \in R^n$ solution de base admissible de P

Toute direction admissible d en x est combinaison linéaire des directions de base d^k en x

$$d = \sum_{k \text{ hors base}} \alpha_k d^k \qquad \text{avec } d^k = k^{ème} \text{ direction de base en x}$$

Les directions admissibles à partir d'une solution de base admissible peuvent être construite comme une combinaison linéaire des $k^{ème}$ direction de base.

Méthode pratique de recherche des directions de base :

Recherche des directions de base

Polytope P dans R^4 sous forme standard

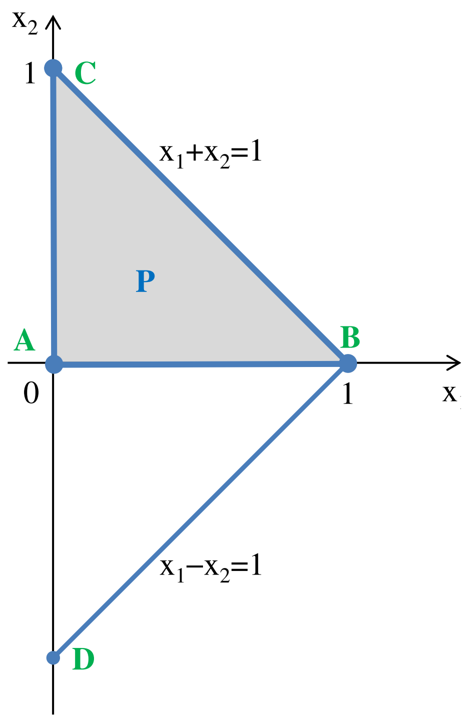
$$P = \{(x_1, x_2, x_3, x_4) \,/\, Ax = b,\ x \geq 0\}$$

$$\text{avec} \quad A = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{pmatrix},\ b = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

- **Direction de base en une solution de base admissible**
  $\rightarrow$ choisir une variable hors base (k)
  $\rightarrow$ fixer la composante hors base correspondante d_{Nk} à 1
  $\rightarrow$ fixer les autres composantes hors base d_N à 0
  $\rightarrow$ calculer les composantes en base d_B par $-B^{-1}A_{.,k}$

- Si la base est non dégénérée, la direction est admissible.
  Sinon, il faut vérifier $d_B \geq 0$ sur les composantes $x_B = 0$

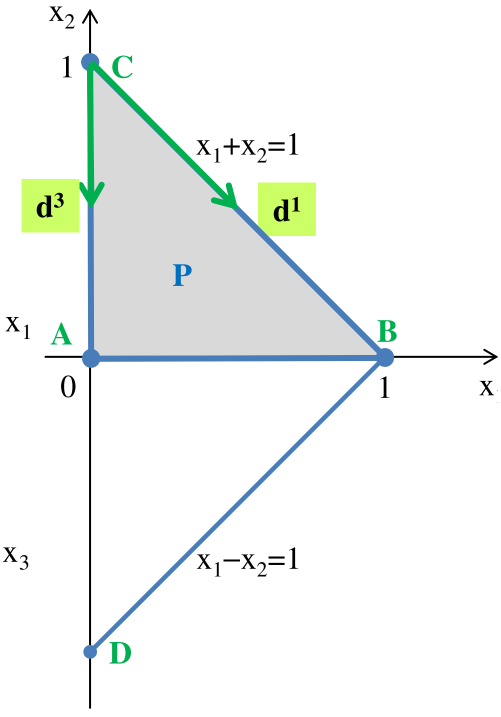
- Sommets de P
  $\rightarrow$ 2 variables hors base à chaque sommet
  $\rightarrow$ 2 directions de base (= arêtes du polytope)



Recherche des directions de base

Examen de directions de base de P

$$A = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{pmatrix}$$



- **Base admissible (x_2,x_4)** : $x = (0 \;\; 1 \;\; 0 \;\; 2)$ $\rightarrow$ **point C**

$$B = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}, \; B^{-1} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

- **Direction de base d^1** correspondant à la variable hors base x_1

$$d_B = -B^{-1}A_{.,1} = -\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ -2 \end{pmatrix}$$

$d^1 = (1 \;\; -1 \;\; 0 \;\; -2)$ $\rightarrow$ admissible

- **Direction de base d^3** correspondant à la variable hors base x_3

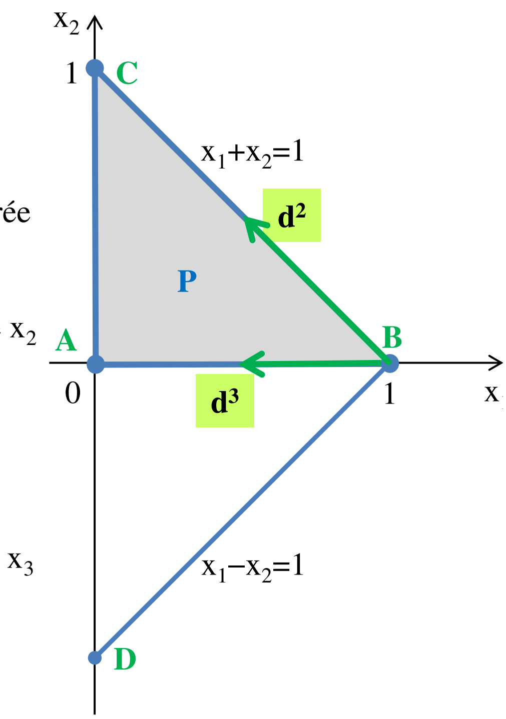
$$d_B = -B^{-1}A_{.,3} = -\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

$d^3 = (0 \;\; -1 \;\; 1 \;\; -1)$ $\rightarrow$ admissible

Recherche des directions de base

Examen de directions de base de P

$$A = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{pmatrix}$$



- **Base admissible (x_1,x_4)** : $x = (1 \;\; 0 \;\; 0 \;\; 0)$ $\rightarrow$ **point B**
  base dégénérée

$$B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \; B^{-1} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$$

- **Direction de base d^2** correspondant à la variable hors base x_2

$$d_B = -B^{-1}A_{.,2} = -\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

$d^2 = (-1 \;\; 1 \;\; 0 \;\; 2)$ $\rightarrow$ admissible

- **Direction de base d^3** correspondant à la variable hors base x_3

$$d_B = -B^{-1}A_{.,3} = -\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$d^3 = (-1 \;\; 0 \;\; 1 \;\; 1)$ $\rightarrow$ admissible

Recherche des directions de base

Examen de directions de base de P

$$A = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{pmatrix}$$

- **Base admissible (x_1,x_2)** : $x = (1\ \ 0\ \ 0\ \ 0) \rightarrow$ **point B**
  base dégénérée

$$B = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},\ B^{-1} = \begin{pmatrix} 0{,}5 & 0{,}5 \\ 0{,}5 & -0{,}5 \end{pmatrix}$$

- **Direction de base d^3** correspondant à la variable hors base x_3

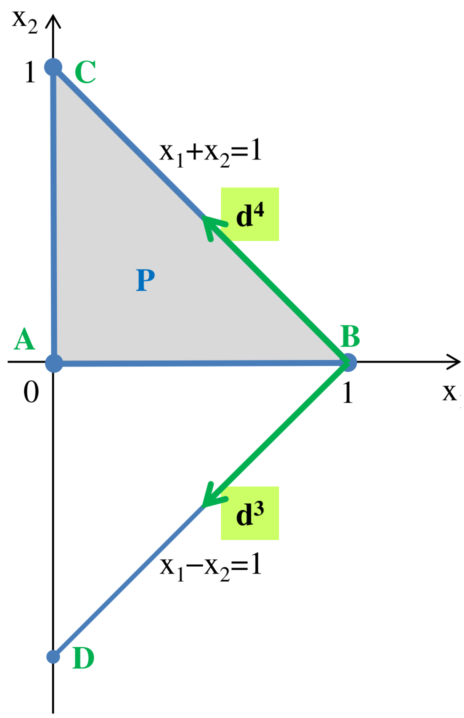
$$d_B = -B^{-1}A_{.,3} = -\begin{pmatrix} 0{,}5 & 0{,}5 \\ 0{,}5 & -0{,}5 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -0{,}5 \\ -0{,}5 \end{pmatrix}$$

$d^3 = (-0{,}5\ \ -0{,}5\ \ 1\ \ 0) \rightarrow$ non admissible (base dégénérée)

- **Direction de base d^4** correspondant à la variable hors base x_4

$$d_B = -B^{-1}A_{.,4} = -\begin{pmatrix} 0{,}5 & 0{,}5 \\ 0{,}5 & -0{,}5 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -0{,}5 \\ 0{,}5 \end{pmatrix}$$

$d^4 = (-0{,}5\ \ 0{,}5\ \ 0\ \ 1) \rightarrow$ admissible



EXERCICE 26: Directions admissibles

On reprend les solutions des bases trouvées dans l'exercice 25 précédent.

1. Rechercher l'ensemble des directions de base pour chaques bases admissibles.
   On indiquera si la base est dégénérée, et si les directions trouvées sont admissibles ou non.
2. Représenter les contraintes, les solutions et les directions admissibles sur un dessin.

2.3 Contraintes non linéaires

a) Indépendance linéaire

Contraintes linéaires

Pour des contraintes linéaires $Ax=b$, $A \in R^{m \times n}$, si A est de rang déficient : rang(A) = r < m,
on peut toujours extraire de A une sous-matrice $\tilde{A} \in R^{r \times n}$ de rang plein : rang($\tilde{A}$) = r, telle que

$\tilde{A}x = \tilde{b} \Leftrightarrow Ax = b \quad \rightarrow$ élimination des contraintes redondantes (cf §1.2.1)

Contraintes non linéaires

Pour des contraintes non linéaires, on considère un modèle linéaire local.

- x_0 point admissible : $\begin{cases} c_E(x_0) = 0 \\ c_I(x_0) \le 0 \end{cases} \Leftrightarrow c(x_0) = 0$ (contraintes actives en x_0)
- Contraintes actives linéarisées : $\hat{c}_0(x) = c(x_0) + \nabla c(x_0)^T (x - x_0)$ avec $c(x_0)=0$

  $\hat{c}_0(x) = 0 \Leftrightarrow \nabla c(x_0)^T x = \nabla c(x_0)^T x_0 \Leftrightarrow Ax = b$

On se ramène au cas de contraintes linéaires avec $\boxed{A = \nabla c(x_0)^T}$ (gradient des contraintes actives)

Condition d'indépendance linéaire

Les contraintes sont dites **linéairement indépendantes** en x_0
si les gradients des contraintes actives sont linéairement indépendants en x_0.
$\Leftrightarrow$ La matrice jacobienne des contraintes actives $J(x_0)=\nabla c(x_0)$ est de rang plein.

Exemple :

Indépendance linéaire

- 1 contrainte égalité + 1 contrainte inégalité dans R^2

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in R^2 \qquad \begin{cases} c_1(x) = x_2 - x_1^2 = 0 \\ c_2(x) = x_1^2 + (x_2 - 1)^2 - 1 \le 0 \end{cases}$$

- En $x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

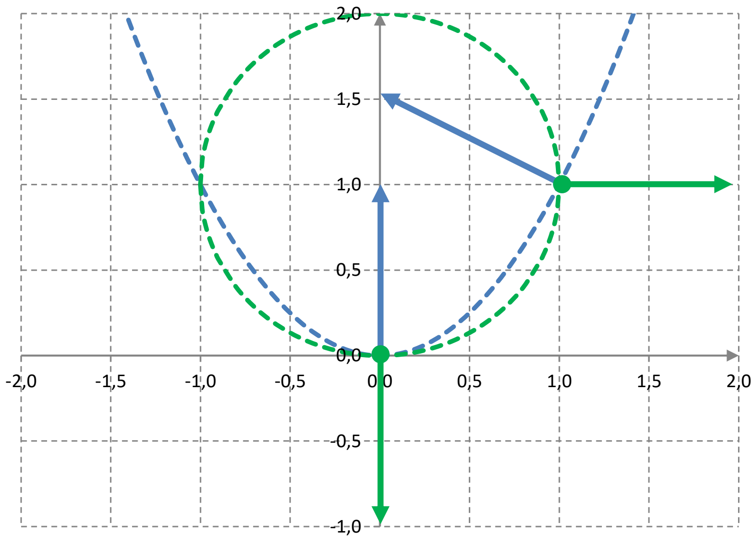
  $\nabla c_1 = \begin{pmatrix} -2 \\ 1 \end{pmatrix} \qquad \nabla c_2 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$

  $\rightarrow$ **linéairement indépendants**

- En $x = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

  $\nabla c_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad \nabla c_2 = \begin{pmatrix} 0 \\ -2 \end{pmatrix}$

  $\rightarrow$ **linéairement dépendants**



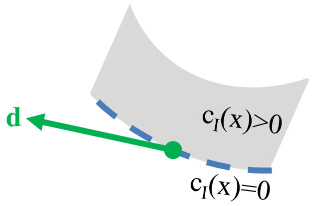
b) Direction admissible

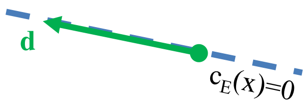
Définition générale

- x point admissible : $\begin{cases} c_E(x_0) = 0 \\ c_I(x_0) \le 0 \end{cases} \Leftrightarrow c(x_0) = 0$ (contraintes actives en x_0)
- **d direction admissible à partir de x** $\Leftrightarrow \exists \eta > 0 / \forall s,\ 0 < s < \eta,\ x + sd$ admissible
  On peut se déplacer sur un segment de longueur ε suivant d à partir de x en restant admissible.

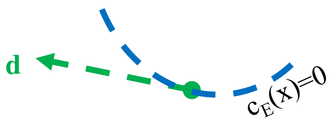
Applicabilité

- Applicable aux contraintes inégalité et aux contraintes égalité linéaires





- Inapplicable aux contraintes égalité non linéaires



$\rightarrow$ Définition à partir de suites de points admissibles

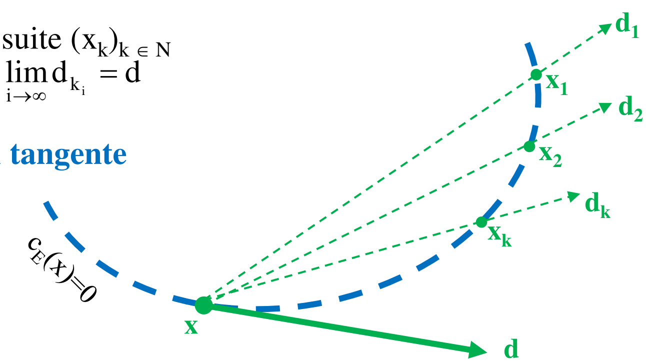
Suite de points admissibles

x point admissible

Définition : $(x_k)_{k\in N}$ **suite admissible** en x $\Leftrightarrow \begin{cases} \forall k,\ x_k \neq x \\ \lim\limits_{k\to\infty} x_k = x \\ \exists k_0 / \forall k \ge k_0,\ x_k \text{ admissible} \end{cases}$

Direction admissible à la limite

- On considère la **suite des directions d_k reliant x_k à x** : $d_k = \dfrac{x_k - x}{\|x_k - x\|}$
- Définition
  d direction admissible à la limite en x pour la suite $(x_k)_{k\in N}$
  $\Leftrightarrow$ Il existe une sous-suite $\left(d_{k_i}\right)_{i\in N}$ telle que : $\lim\limits_{i\to\infty} d_{k_i} = d$
- **Direction admissible à la limite = direction tangente**

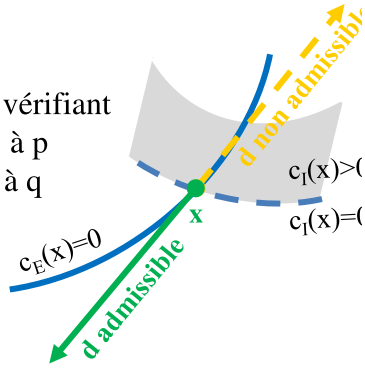


c) Cône des directions

Définition

x point admissible

Le **cône des directions D(x) en x** est l'ensemble des directions $d \in R^n$ vérifiant



- $\nabla c_{Ej}(x)^T d = 0$ pour toutes les contraintes égalité $c_{Ej}(x) = 0$, j=1 à p
- $\nabla c_{Ij}(x)^T d \le 0$ pour les contraintes inégalité actives : $c_{Ij}(x) = 0$, j=1 à q

$d \in D(x)$	→ direction **tangente** aux contraintes égalité → direction **intérieure** aux contraintes inégalité actives

Propriété

Toute direction admissible à la limite en x appartient au cône des directions en x

Preuve :

(x_k) *suite admissible de limite x* $\Rightarrow \begin{cases} c_E(x_k) = 0 \\ c_I(x_k) \le 0 \end{cases} \rightarrow$ *directions* $d_k = \dfrac{x_k - x}{\|x_k - x\|}$

c contrainte active en x : c(x)=0 $\quad c(x_k) = c(x) + \nabla c(x)^T (x_k - x) + o(\|x_k - x\|)$

$$\nabla c(x)^T d_k = \frac{c(x_k) - c(x)}{\|x_k - x\|} - \frac{o(\|x_k - x\|)}{\|x_k - x\|} \Rightarrow \nabla c(x)^T d = \lim_{k\to\infty} \frac{c(x_k)}{\|x_k - x\|} \rightarrow \begin{cases} = 0 \ \ (\textit{égalité}) \\ \le 0 \ \ (\textit{inégalité}) \end{cases}$$

d) Qualification

Caractérisation des directions admissibles

- Le cône des directions D(x) au point x admissible est simple à manipuler en pratique.

$$d \in D(x) \Leftrightarrow \begin{cases} \nabla c_E(x)^T d = 0 & \rightarrow \text{pour toutes les contraintes égalité} \\ \nabla c_I(x)^T d \le 0 & \rightarrow \text{pour les contraintes inégalité actives en x} \end{cases}$$

- Toutes les directions admissibles à la limite en x appartiennent à D(x), mais D(x) peut contenir également des directions non admissibles.
  → D(x) ne caractérise pas les directions admissibles.

Qualification des contraintes

Les contraintes vérifient la **condition de qualification** au point admissible x si toute direction du cône D(x) est admissible à la limite.

→ Condition très importante dans les algorithmes

Conditions suffisantes de qualification des contraintes

- Contraintes linéaires : Ax=b

  ou

- Contraintes **linéairement indépendantes en x** : $\nabla c(x)$ de rang plein
  → réalisable simplement en pratique par extraction d'une sous-matrice de rang plein

2.4 Déplacement admissible

a) Principes

Problème sous contraintes non linéaires

$$\min_{x \in R^n} f(x) \text{ sous } \begin{cases} c_E(x) = 0 \\ c_I(x) \le 0 \end{cases}$$

- On cherche à construire un déplacement p admissible et améliorant à partir d'un point initial x_0. On se ramène à un problème avec contraintes égalité (contraintes actives en x_0).

$$\min_{x \in R^n} f(x) \text{ sous } c(x) = 0 \quad \rightarrow \text{ m contraintes actives en } x_0$$

- Les n composantes du déplacement p doivent vérifier : $\begin{cases} c(x_0 + p) = 0 \\ f(x_0 + p) < f(x_0) \end{cases}$

Méthodes possibles

- Elimination directe
  On exprime m variables à partir des n−m autres à partir des contraintes.
  On substitue dans l'expression de f $\rightarrow$ problème sans contraintes

- Réduction généralisée
  On linéarise les contraintes en x_0.
  On applique la méthode de réduction des contraintes linéaires (matrices Y et Z).
  On corrige le déplacement pour prendre en compte les non-linéarités.

b) Méthode d'élimination directe

Principe

$$\min_{x \in R^n} f(x) \text{ sous } c(x) = 0 \quad \rightarrow \text{ m contraintes actives}$$

- Les contraintes sont de la forme : $c(x) = c(x_{lié}, x_{libre}),\ x_{lié} \in R^m, x_{libre} \in R^{n-m}$

- Si l'on sait résoudre : $c(x_{lié}, x_{libre}) = 0 \Leftrightarrow x_{lié} = \psi(x_{libre})$

  le problème devient : $\min_{x_{libre} \in R^{n-m}} \varphi(x_{libre})$ avec $\varphi(x_{libre}) \underset{def}{=} f(x_{lié}, x_{libre}) = f(\psi(x_{libre}), x_{libre})$

  $\rightarrow$ problème de dimension n-m, sans contrainte

Difficultés

- Il faut faire attention au **domaine de définition** des variables (contraintes implicites)
  $\rightarrow$ voir exemples

- Il faut disposer de l'expression analytique des fonctions (rarement réalisé en pratique)

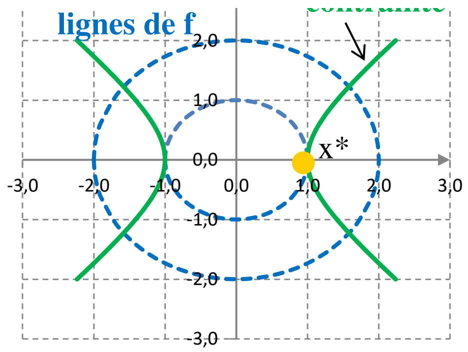
Elimination directe

- Exemple 1 : $\min\limits_{x_1,x_2} x_1^2 + x_2^2$ sous $x_1^2 - x_2^2 = 1$

  Elimination de x_1 : $x_1^2 - x_2^2 = 1 \Rightarrow x_1^2 = 1 + x_2^2$

  $\rightarrow \min\limits_{x_2} 1 + 2x_2^2 \Rightarrow x_2 = 0$

  Solution correcte : $\begin{cases} x_1 = 1 \\ x_2 = 0 \end{cases}$



- Exemple 2 : $\min\limits_{x_1,x_2} x_1^2 + x_2^2$ sous $x_1^2 + 4x_2^2 = 1$

  Elimination de x_1 : $x_1^2 + 4x_2^2 = 1 \Rightarrow x_1^2 = 1 - 4x_2^2$

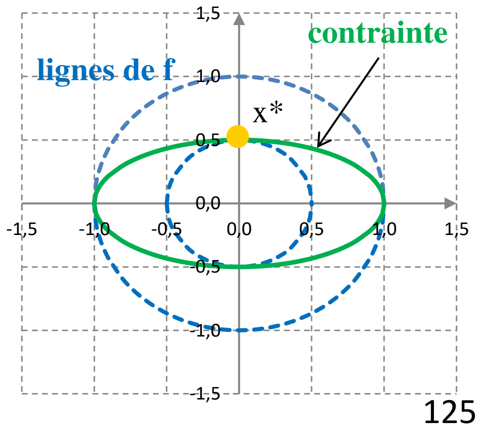
  $\rightarrow \min\limits_{x_2} 1 - 3x_2^2 \Rightarrow x_2 = \pm\infty$

  Solution incorrecte

  Contrainte implicite : $x_1^2 \geq 0 \Rightarrow 1 - 4x_2^2 \geq 0$

  $\Rightarrow -\dfrac{1}{2} \leq x_2 \leq \dfrac{1}{2}$

  $\rightarrow$ à prendre en compte explicitement dans la résolution



125

3 Condition d'optimalité

3.1 Dualité

a) Dualité critère-contrainte

Problème avec contraintes égalité

$\min\limits_{x \in R^n} f(x)$ sous $c(x) = 0$ $\rightarrow$ m contraintes d'égalité (= contraintes actives)

Dualité

Difficulté de résolution due aux 2 objectifs antagonistes :

- Minimiser le critère f(x) $\rightarrow \min\limits_{x} f(x)$
- Satisfaire les contraintes c(x)=0 $\rightarrow \min\limits_{x} \|c(x)\|$

  $\rightarrow$ **Dualité critère-contraintes**

Méthodes duales

Prise en compte des contraintes avec pondération dans la fonction coût

- Critère augmenté $\rightarrow$ pondération = pénalisation des contraintes
- Lagrangien $\rightarrow$ pondération = multiplicateurs de Lagrange
- Lagrangien augmenté $\rightarrow$ pondération = pénalisation + multiplicateurs

$\rightarrow$ Problème sans contraintes plus simple

Réglages des pondérations / **Equivalence au problème avec contraintes**

Critère augmenté

Problème avec contraintes égalité

$\min\limits_{x \in R^n} f(x)$ sous $c(x) = 0$ → m contraintes d'égalité (= contraintes actives)

Critère augmenté

$$\boxed{f_\rho(x) = f(x) + \frac{1}{2}\rho\|c(x)\|^2}$$

ρ = **coefficient de pénalisation** > 0 → Pénalise la violation des contraintes
→ Pondération critère-contraintes

Problème pénalisé sans contraintes

$\min\limits_{x \in R^n} f_\rho(x)$ → Problème **équivalent au problème avec contraintes**
si la pénalisation ρ est assez grande

Problème pénalisé avec contraintes

$\min\limits_{x \in R^n} f_\rho(x)$ sous $c(x) = 0$ → Problème équivalent au problème avec contraintes
Renforce le poids des contraintes dans le critère

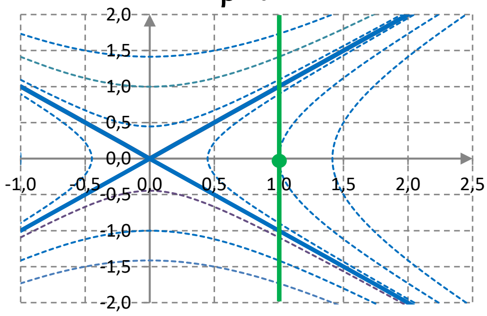
Critère augmenté

$$\min_{x_1, x_2} \frac{1}{2}\left(x_2^2 - x_1^2\right) \text{ sous } x_1 = 1$$

$$f(x) = \frac{1}{2}\left(x_2^2 - x_1^2\right) \Rightarrow \boxed{x^* = (1 \;\; 0)}$$

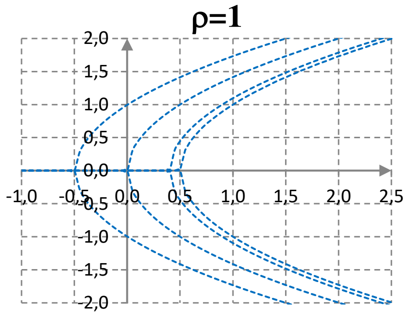
$$c(x) = x_1 - 1$$

ρ=0

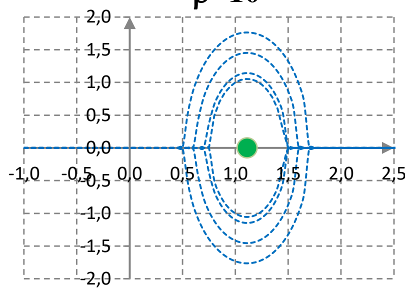


$$f_\rho(x_1, x_2) = \frac{1}{2}\left(x_2^2 - x_1^2\right) + \frac{1}{2}\rho(x_1 - 1)^2 \Rightarrow \boxed{x^*(\rho) = \left(\frac{\rho}{\rho - 1} \;\; 0\right)}$$

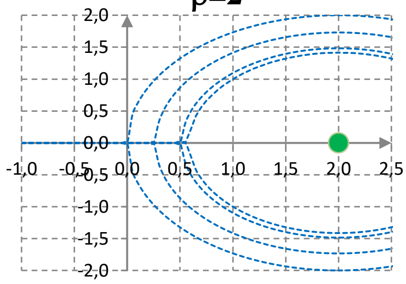
ρ=1



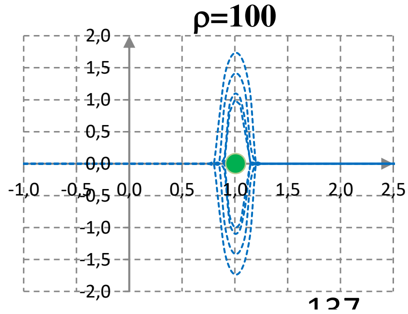
ρ=10



ρ=2

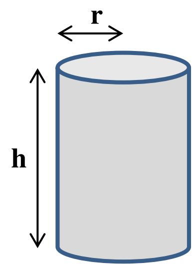


ρ=100



Problème de la boîte

- Réaliser une boîte cylindrique de volume donné V_0 et de surface S minimale
  Dimensions : hauteur = h , rayon = r
  Surface : $S = 2\pi r^2 + 2\pi rh$
  Volume : $V = \pi r^2 h$



- Formulation avec contrainte (en divisant S et V par π)

$$\min_{h,r} f(h,r) = 2r^2 + 2rh \quad \text{sous} \quad c(h,r) = r^2 h - 2v_0 = 0 \qquad \text{avec } V_0 = 2\pi v_0$$

Solution exacte

- La contrainte permet d'éliminer h : $h = \dfrac{2v_0}{r^2}$

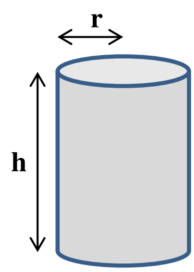
  En reportant dans la fonction coût : $\min_r \left(2r^2 + \dfrac{4v_0}{r}\right) \Rightarrow \boxed{r = v_0^{1/3}} \Rightarrow \boxed{h = 2v_0^{1/3}}$

- Valeur optimale du coût : $\boxed{f = 6v_0^{2/3}}$

Solution par pénalisation

- Formulation avec contrainte

$$\min_{h,r} f(h,r) = 2r^2 + 2rh \quad \text{sous} \quad c(h,r) = r^2 h - 2v_0 = 0$$



- <u>Formulation pénalisée</u>

$$\min_{h,r} f_\rho(h,r) = 2r^2 + 2rh + \frac{1}{2}\rho(r^2 h - 2v_0)^2$$

Conditions de minimum : $\begin{cases} \dfrac{\partial f_\rho}{\partial h} = 0 \Rightarrow 2r + \rho r^2 (r^2 h - 2v_0) = 0 \\ \dfrac{\partial f_\rho}{\partial r} = 0 \Rightarrow 4r + 2h + 2\rho rh(r^2 h - 2v_0) = 0 \end{cases}$

La 1ère équation donne : $\rho r(r^2 h - 2v_0) = -2$

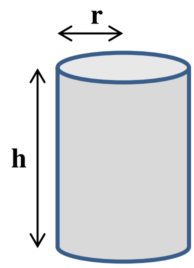
En remplaçant dans la 2ème équation : $h = 2r$

On obtient une équation en r : $\boxed{\rho r(r^3 - v_0) + 1 = 0}$ → à résoudre numériquement

Application numérique

- Volume : $v_0 = 1000$
- Solution exacte : $r = 10$, $h = 20$, $f = 600$
- Solution pénalisée pour ρ allant de 10^{-3} à 10^{2}.

Résolution numérique de : $\rho r(r^3 - v_0) + 1 = 0 \;\rightarrow\; r\;,\; h = 2r$



ρ	r	h	f	c
0.001	9.641582	19.28316	557.761	792.565
0.01	9.966442	19.93288	595.980	979.933
0.1	9.996664	19.99333	599.600	997.999
1	9.999667	19.99933	599.960	999.800
10	9.999967	19.99993	599.996	999.980
100	9.999997	19.99999	600.000	999.998

→ Pénalisation faible
Contrainte mal respectée
Écart important à la solution exacte

→ Pénalisation forte
Contrainte bien respectée
Écart faible à la solution exacte

Lagrangien augmenté

Problème avec contraintes égalité et inégalité

$$\min_{x \in R^n} f(x) \text{ sous } \begin{cases} c_E(x) = 0 \\ c_I(x) \le 0 \end{cases} \quad (PO)$$

→ p contraintes d'égalité
→ q contraintes d'inégalité

Multiplicateurs de Lagrange

1 multiplicateur par contrainte

- $\lambda \in R^p$ → multiplicateurs des contraintes d'égalité
- $\mu \in R^q$ → multiplicateurs des contraintes d'inégalité

Fonction de Lagrange (ou lagrangien)

Le **lagrangien** du problème (PO) est la fonction L de R^{n+p+q} dans R

$x \in R^n, \lambda \in R^p, \mu \in R^q \mapsto L(x, \lambda, \mu) \in R$

$$\boxed{L(x,\lambda,\mu) = f(x) + \lambda^T c_E(x) + \mu^T c_I(x)} \quad \Leftrightarrow \quad L(x,\lambda,\mu) = f(x) + \sum_{j=1}^{p} \lambda_j c_{Ej}(x) + \sum_{j=1}^{q} \mu_j c_{Ij}(x)$$

→ multiplicateurs ≈ coefficients de pénalisation des contraintes
→ interprétation comme des sensibilités aux niveaux des contraintes

Problème pénalisé avec contraintes égalité

- Critère augmenté : coefficient de pénalisation $\rho > 0$

$$\min_{x \in R^n} f_\rho(x) \text{ sous } c(x)=0 \qquad \text{avec} \quad f_\rho(x) = f(x) + \frac{1}{2}\rho\|c(x)\|^2$$

- Lagrangien du problème pénalisé avec contraintes

$$\boxed{\begin{aligned} L_\rho(x,\lambda) &= f_\rho(x) + \lambda^T c(x) \\ &= f(x) \ + \lambda^T c(x) + \frac{1}{2}\rho\|c(x)\|^2 \\ &= L(x,\lambda) \qquad\quad + \frac{1}{2}\rho\|c(x)\|^2 \end{aligned}}$$

L_ρ = **lagrangien augmenté** = lagrangien initial + pénalisation des contraintes

- Utilisation du lagrangien augmenté
  - Démonstration des conditions suffisantes d'optimalité
  - Algorithme de lagrangien augmenté = suite de minimisations sans contraintes

Lagrangien augmenté

- Fonction de 2 variables

$$\min_{x_1,x_2} \frac{1}{2}\left(x_2^2 - x_1^2\right) \text{ sous } x_1 = 1 \qquad \rightarrow \text{ minimum en } x^* = (1 \quad 0), \ \lambda^* = 1$$

- **Critère augmenté**

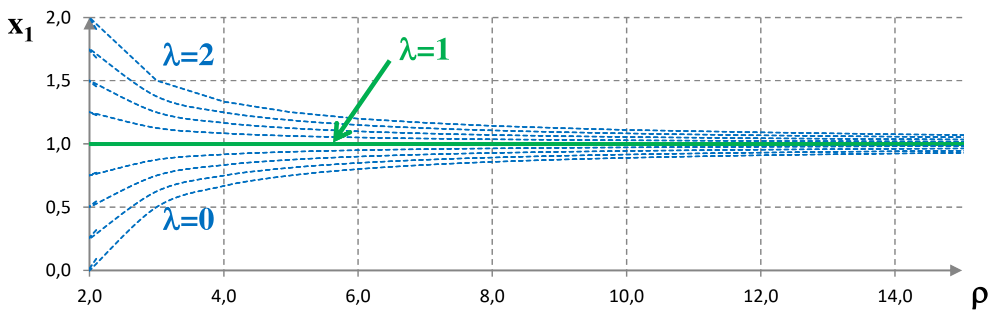
$$f_\rho(x_1,x_2) = \frac{1}{2}\left(x_2^2 - x_1^2\right) + \frac{1}{2}\rho(x_1-1)^2 \qquad \rightarrow \text{ minimum en } \boxed{x^*(\rho) = \left(\frac{\rho}{\rho-1} \quad 0\right)}$$

- **Lagrangien augmenté**

$$L_\rho(x_1,x_2,\lambda) = \frac{1}{2}\left(x_2^2 - x_1^2\right) + \lambda(x_1-1) + \frac{1}{2}\rho(x_1-1)^2 \qquad \rightarrow \text{ minimum en } \boxed{x^*(\rho,\lambda) = \left(\frac{\rho-\lambda}{\rho-1} \quad 0\right)}$$

Pour $\lambda = \lambda^* = 1$, le minimum sans contrainte du lagrangien augmenté est la solution x^* du problème initial.

$$x^*(\rho,\lambda^*) = \left(\frac{\rho-1}{\rho-1} \quad 0\right) = (1 \quad 0)$$



143

Exercice : calculer le minimum pour chacun des cas ci-dessus (utiliser la condition d'ordre 1, sur le

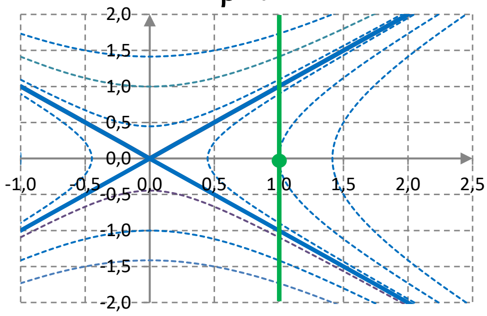
Lagrangien, le critère augmenté, et le Lagrangien augmenté).

Lagrangien augmenté

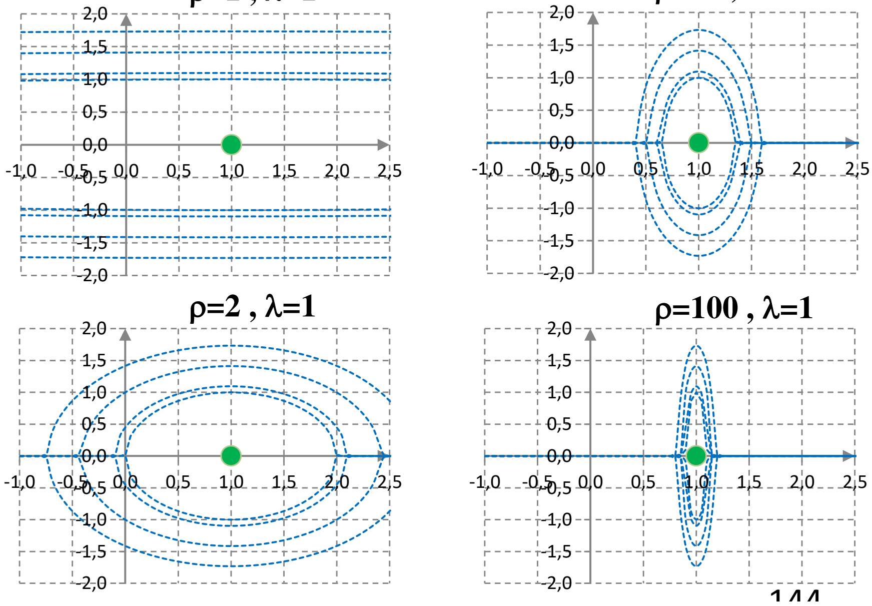
$$\min_{x_1,x_2} \frac{1}{2}\left(x_2^2 - x_1^2\right) \text{ sous } \ x_1 = 1$$

$$f(x) = \frac{1}{2}\left(x_2^2 - x_1^2\right) \quad \Rightarrow \boxed{x^* = (1 \ \ 0)}$$

$$c(x) = x_1 - 1 \qquad \lambda^* = 1$$



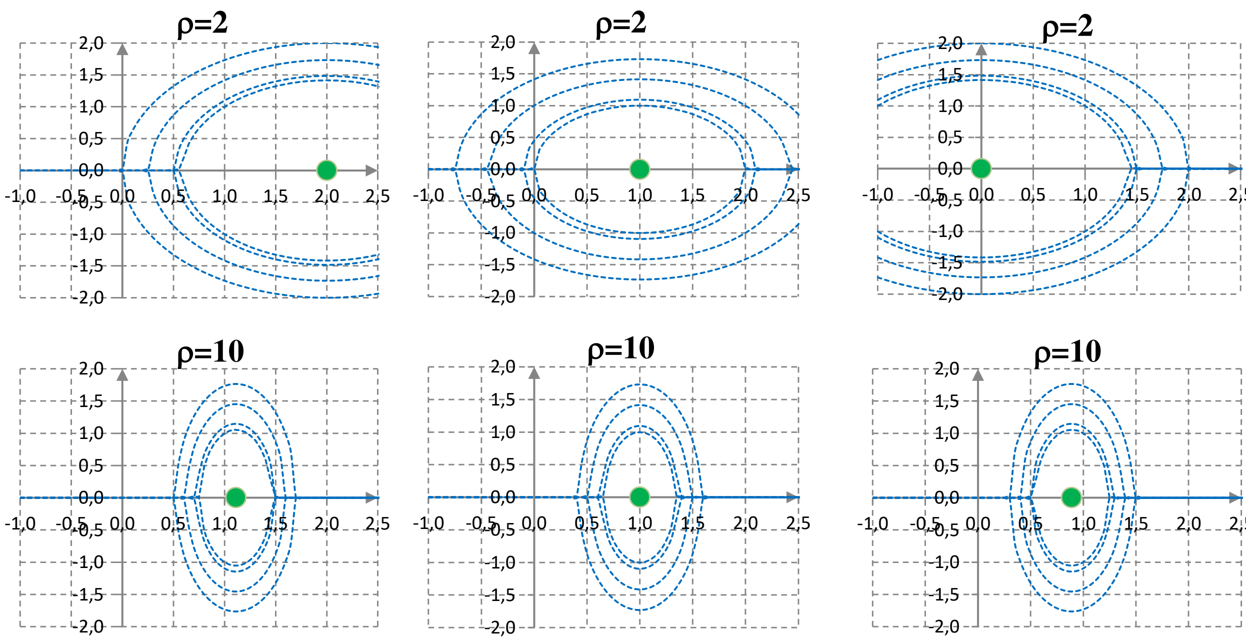
$$L_\rho(x_1, x_2, \lambda) = \frac{1}{2}\left(x_2^2 - x_1^2\right) + \lambda(x_1 - 1) + \frac{1}{2}\rho(x_1 - 1)^2 \quad \Rightarrow \boxed{x^*(\rho) = \begin{pmatrix} \frac{\rho-\lambda}{\rho-1} & 0 \end{pmatrix}}$$



144

Lagrangien augmenté

lignes de niveau pour $\lambda = 0$ | lignes de niveau pour $\lambda = 1$ | lignes de niveau pour $\lambda = 10$



145

b) Problème dual

Fonction duale

La **fonction duale** du problème (PO) est la fonction w de R^{p+q} dans R

$\boxed{w(\lambda,\mu) = \min_{x \in R^n} L(x,\lambda,\mu)}$ → Minimisation du lagrangien à λ et μ fixés

x = **variables primales**

λ et μ = **variables duales**

- Domaine de w : $X_w = \{\lambda \in R^p, \mu \in R^q \ / \ w(\lambda,\mu) > -\infty\}$ → w bornée

Concavité - Convexité

- La fonction duale w est concave
- Le domaine X_w est convexe

Preuve : on note : $\gamma=(\lambda,\mu)$

- $L(x,\alpha\gamma_1+(1-\alpha)\gamma_2) = \alpha L(x,\gamma_1) + (1-\alpha)L(x,\gamma_2)$ *car L linéaire en λ et μ*

$\Rightarrow w(\alpha\gamma_1+(1-\alpha)\gamma_2) \geq \alpha w(\gamma_1) + (1-\alpha)w(\gamma_2)$ *pour le minimum/x de chaque membre*

→ *w concave*

- *Si* γ_1 *et* $\gamma_2 \in X_w$, $w(\alpha\gamma_1+(1-\alpha)\gamma_2) \geq \alpha w(\gamma_1) + (1-\alpha)w(\gamma_2) > -\infty$

$\Rightarrow \alpha\gamma_1+(1-\alpha)\gamma_2 \in X_w$

→ X_w *convexe*

Problème dual

- Problème primal : $\min_{x \in R^n} f(x)$ sous $\begin{cases} c_E(x) = 0 \\ c_I(x) \leq 0 \end{cases}$
- Fonction duale : $\boxed{w(\lambda,\mu) = \min_{x \in R^n} L(x,\lambda,\mu)}$ → $x(\lambda,\mu) \ / \ \nabla_x L(x,\lambda,\mu) = 0$
- Domaine de w : $X_w = \{\lambda \in R^p, \mu \in R^q \ / \ w(\lambda,\mu) > -\infty\}$ → w bornée
- **Problème dual** : $\boxed{\max_{\lambda \in R^p, \mu \in R^q} w(\lambda,\mu) \text{ sous } (\lambda,\mu) \in X_w \ , \ \mu \geq 0}$

$\Leftrightarrow \max_{x \in R^n, \lambda \in R^p, \mu \in R^q} L(x,\lambda,\mu)$ sous $\begin{cases} \nabla_x L(x,\lambda,\mu) = 0 \ \rightarrow \ x(\lambda,\mu) \\ (\lambda,\mu) \in X_w \ , \ \mu \geq 0 \end{cases}$ → dual de Wolfe

Borne sur la fonction duale

- x* solution du problème primal
- $(\lambda,\mu) \in X_w$, $\mu \geq 0$

$\Rightarrow \boxed{w(\lambda,\mu) \leq f(x^*)}$

Preuve :

$$w(\lambda,\mu) = \min_{x \in R^n} L(x,\lambda,\mu) \leq L(x^*,\lambda,\mu) = f(x^*) + \lambda^T c_E(x^*) + \mu^T c_I(x^*)$$

$$= f(x^*) + \mu^T c_I(x^*) \quad \textit{car } x^* \textit{ admissible} \Rightarrow c_E(x^*) = 0$$

$$\leq f(x^*) \quad \textit{car } x^* \textit{ admissible} \Rightarrow c_I(x^*) \leq 0 \ \textit{ et } \mu \geq 0$$

Fonction duale

- **Problème primal** : $\min\limits_{x_1,x_2} \frac{1}{2}\left(x_2^2 - x_1^2\right)$ sous $x_1 = 1$
- Lagrangien : $L(x,\lambda) = \frac{1}{2}\left(x_2^2 - x_1^2\right) + \lambda\left(x_1 - 1\right)$
- Solution : $\boxed{x^* = \begin{pmatrix}1\\0\end{pmatrix},\quad \lambda^* = 1}$

- **Fonction duale** : $w(\lambda) = \min\limits_{x} L(x,\lambda) \Rightarrow \frac{\partial L}{\partial x} = 0 \Rightarrow \begin{cases}-x_1 + \lambda = 0\\ x_2 = 0\end{cases} \Rightarrow \begin{cases}x_1 = \lambda\\ x_2 = 0\end{cases}$

$$\Rightarrow w(\lambda) = \frac{1}{2}\lambda^2 - \lambda \quad \text{avec} \quad \begin{cases}x_1 = \lambda\\ x_2 = 0\end{cases}$$

- **Problème dual** : $\max\limits_{\lambda} w(\lambda) \Rightarrow \frac{\partial w}{\partial \lambda} = 0 \Rightarrow \lambda = 1$
- Solution : $\boxed{\lambda^* = 1,\quad x^* = \begin{pmatrix}1\\0\end{pmatrix}}$

Dualité faible

Théorème de la dualité faible

- x^* solution du problème primal
- λ^*,μ^* solution du problème dual

$\Rightarrow \boxed{w(\lambda^*,\mu^*) \le f(x^*)}$

Preuve : $\forall\lambda, \forall\mu \ge 0,\ w(\lambda,\mu) \le f(x^*) \Rightarrow w(\lambda^*,\mu^*) \le f(x^*)$

Dualité et admissibilité

- Si le problème primal est non borné, le problème dual est non admissible.
- Si le problème dual est non borné, le problème primal est non admissible.

Preuve : en utilisant $w^*(\lambda^*,\mu^*) \le f(x^*)$
Existence de solutions x^*, λ^*, μ^* $\Rightarrow$ *fonctions bornées*

Saut de dualité

Le **saut de dualité** est la différence entre la solution du problème primal et du problème dual.

$$\delta = f(x^*) - w(\lambda^*,\mu^*) \ge 0$$

Dans le cas général δ n'est pas nul, il n'est pas équivalent de minimiser f ou maximiser w.

Dualité forte

Point col

$(x^*,\lambda^*,\mu^*\geq 0)$ est un **point col** (ou **point selle**) du lagrangien si

$$\forall(x,\lambda,\mu\geq 0),\ \begin{cases} L(x^*,\lambda,\ \mu) \leq L(x^*,\lambda^*,\mu^*) & \to \text{ maximisation de L par rapport à } (\lambda,\mu) \\ L(x^*,\lambda^*,\mu^*) \leq L(x,\ \ \lambda^*,\mu^*) & \to \text{ minimisation de L par rapport à } (x) \end{cases}$$

Caractérisation

$(x^*,\lambda^*,\mu^*\geq 0)$ est un point col du lagrangien si et seulement si $\begin{cases} L(x^*,\lambda^*,\mu^*) = \min\limits_x L(x,\lambda^*,\mu^*) \\ c_E(x^*) = 0 \\ c_I(x^*) \leq 0 \\ \mu^* c_I(x^*) = 0 \end{cases}$

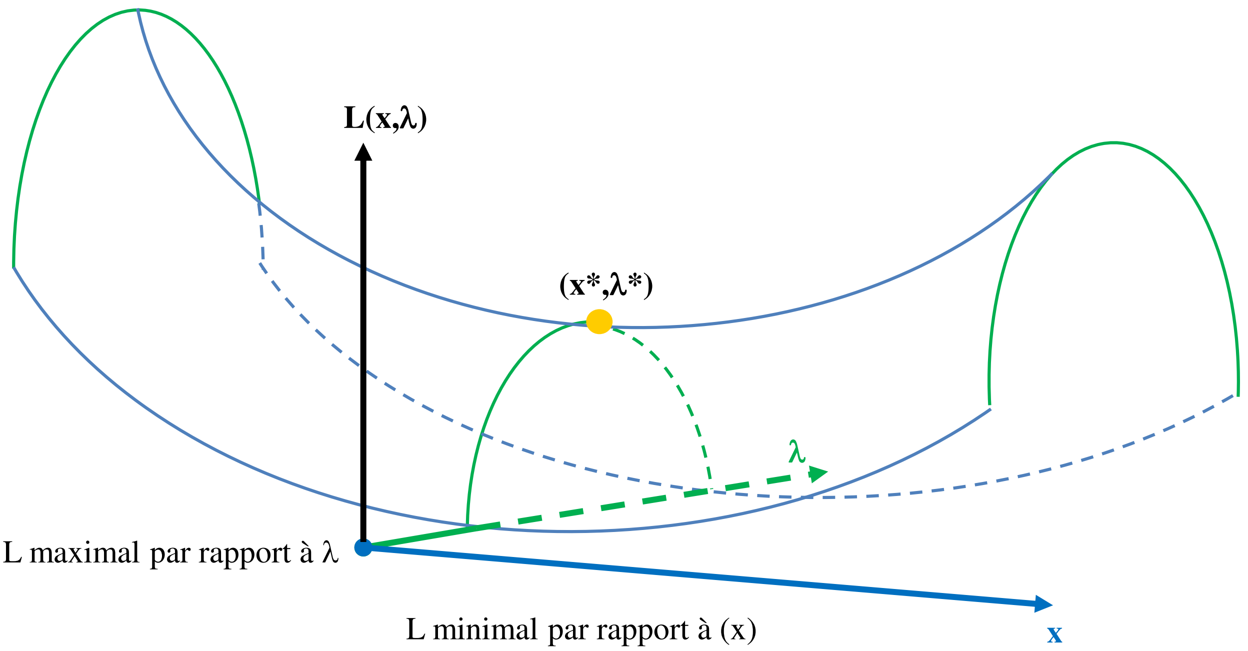
Théorème de la dualité forte

Le lagrangien admet un point col (x^*,λ^*,μ^*) si et seulement si le saut de dualité est nul.

$$\boxed{(x^*,\lambda^*,\mu^*\geq 0) \text{ un point col} \quad \Leftrightarrow \quad w(\lambda^*,\mu^*) = f(x^*)}$$

Il est alors équivalent de minimiser f(x) ou maximiser $w(\lambda,\mu)$.

Point col (ou point selle)



Optimum global

Si ($x^*,\lambda^*,\mu^*\geq 0$) est un point col du lagrangien : $\begin{cases} x^* \to \min\limits_x L(x,\lambda^*,\mu^*) \\ c_E(x^*)=0\ ,\ c_I(x^*)\leq 0 \\ \mu^* c_I(x^*)=0 \end{cases}$

alors x^* est un **optimum global** du problème primal : $\min\limits_{x\in R^n} f(x)$ sous $\begin{cases} c_E(x)=0 \\ c_I(x)\leq 0 \end{cases}$

En pratique
- Si le lagrangien admet un point col, on peut obtenir l'optimum global x^*.
- Pour un problème non convexe, il n'existe en général pas de point col.

Exemple $\min\limits_x f(x)=-x^2$ sous $\begin{cases} 2x-1\leq 0 \\ 0\leq x\leq 1 \end{cases}$ $\to$ solution : $x^* = 0.5$

$L(x,\mu)=-x^2+\mu(2x-1)$ avec $\begin{cases} 0\leq x\leq 1 \\ \mu\geq 0 \end{cases}$

Point col : $\left.\begin{cases} \left.\begin{matrix} \min\limits_x L(x,\mu) \\ 0\leq x\leq 1 \end{matrix}\right\} \to \begin{matrix} x=0 \text{ si } \mu>0.5 \\ x=1 \text{ si } \mu<0.5 \end{matrix} \\ 2x-1\leq 0 \quad \to \quad x=0 \\ \mu(2x-1)=0 \quad \to \quad \mu=0<0.5 \quad \to \quad x=1 \\ \mu\geq 0 \end{cases}\right\} \to$ Il n'existe pas de point col.

c) Programmation linéaire

Problème primal

$\min\limits_x c^T x$ sous $\begin{cases} Ax=b \\ x\geq 0 \end{cases}$ $\to$ problème linéaire sous forme standard

$\Leftrightarrow \min\limits_x c^T x$ sous $\begin{cases} b-Ax=0 \\ -x\leq 0 \end{cases}$ $\begin{matrix} \to \text{ multiplicateur } \lambda \\ \to \text{ multiplicateur } \mu \end{matrix}$

- Fonction de Lagrange : $L(x,\lambda,\mu)= c^T x+\lambda^T(b-Ax)+\mu^T(-x)$
  $=\left(c-A^T\lambda-\mu\right)^T x+\lambda^T b$ $\to$ linéaire en x
- **Fonction duale** : $w(\lambda,\mu)=\min\limits_x L(x,\lambda,\mu)$
- Domaine de définition : $X_w=\{(\lambda,\mu)\ /\ w(\lambda,\mu)>-\infty\}$
- La fonction duale n'est définie que si $L(x, \lambda, \mu)$ est borné inférieurement.
  $L(x, \lambda, \mu)$ est linéaire en x $\to$ Le coefficient de x doit être nul.

$(\lambda,\mu)\in X_w \Rightarrow c-A^T\lambda-\mu=0$

$\Rightarrow L(x,\lambda,\mu)= \lambda^T b \quad \Rightarrow \boxed{w(\lambda,\mu)=\lambda^T b}$

Problème dual

$$\max_{\lambda,\mu} w(\lambda,\mu) \text{ sous } \begin{cases}(\lambda,\mu)\in X_w \\ \mu \geq 0\end{cases}$$

$$\Leftrightarrow \max_{\lambda,\mu} \lambda^T b \quad \text{sous } \begin{cases}c - A^T\lambda - \mu = 0 \\ \mu \geq 0\end{cases} \quad \rightarrow \text{ ne dépend pas de } \mu$$

$$\Leftrightarrow \max_{\lambda} b^T\lambda \quad \text{sous } \ c - A^T\lambda \geq 0 \quad \rightarrow \text{ nouveau problème linéaire en } \lambda$$

- Le problème dual est également un problème linéaire dont la variable est λ.
  On met le problème dual sous forme standard en notant la variable y au lieu de λ

  $$\min_y -b^T y \ \text{ sous } \ A^T y - c \leq 0 \qquad \rightarrow \text{ multiplicateur } \nu$$

  On peut ensuite définir les fonctions associées à ce problème linéaire.

- Fonction de Lagrange notée $L_d(y, \nu)$: $L_d(y,\nu) = -b^T y + \nu^T(A^T y - c)$
  $= (A\nu - b)^T y - \nu^T c \qquad \rightarrow$ bornée si $A\nu - b = 0$

- **Fonction duale notée $w_d(\nu)$** : $\quad w_d(\nu) = \min_y L_d(y,\nu) = -\nu^T c \quad$ si $A\nu - b = 0$

Problème dual du problème dual

Le problème dual admet lui-même pour dual :

$$\max_{\nu} w_d(\nu) \quad \text{sous } \begin{cases}\nu \in X_{w_d} \\ \nu \geq 0\end{cases}$$

$$\Leftrightarrow \max_{\nu} -\nu^T c \ \text{ sous } \begin{cases}A\nu = b \\ \nu \geq 0\end{cases} \Leftrightarrow \min_x c^T x \ \text{ sous } \begin{cases}Ax = b \\ x \geq 0\end{cases} \rightarrow \text{ identique au problème primal}$$

- **Le problème dual du problème dual est le problème primal.**

- Pour un problème linéaire, il est équivalent de résoudre le problème primal ou problème dual.
  Les solutions du problème primal et du problème dual ont le même coût $\rightarrow$ **dualité forte**

Solutions possibles

		Dual		
		Optimum fini	Optimum infini	Sans solution
Primal	Optimum fini	**dualité forte**	impossible	impossible
	Optimum infini	impossible	impossible	**dualité faible**
	Sans solution	impossible	**dualité faible**	contraintes incompatibles

Correspondances primal – dual

- Problème primal (P) sous forme standard : $(P)\quad \min_x c^T x \quad \text{sous} \begin{cases} Ax = b \\ x \ge 0 \end{cases}$

- Problème dual (D) du problème (P) : $(D)\quad \max_y b^T y \quad \text{sous} \quad A^T y \le c$

- Le nombre de variables de (P) est égal au nombre de contraintes de (D).
- Le nombre de contraintes de (P) est égal au nombre de variables de (D).
- La matrice des contraintes de (D) est la transposée de la matrice des contraintes de (P).
- Une variable $x_j \ge 0$ de coût c_j donne une contrainte $\le$ de niveau c_j : $\begin{cases} c_j x_j \\ x_j \ge 0 \end{cases} \rightarrow \sum_{i=1}^{m} a_{ij} y_i \le c_j$
- Une contrainte = de niveau b_i donne une variable $y_j \in R$ de coût b_i : $\sum_{j=1}^{n} a_{ij} x_j = b_i \rightarrow \begin{cases} b_i y_i \\ y_i \in R \end{cases}$

$\rightarrow$ généralisation à un problème linéaire quelconque (signe des variables, sens des contraintes)

Correspondance primal-dual

- **Problème primal (P)**

$$\min_{\substack{x_1 \in R^{n1} \\ x_2 \in R^{n2} \\ x_3 \in R^{n3}}} c_1^T x_1 + c_2^T x_2 + c_3^T x_3 \quad \text{sous} \begin{cases} A_1 x_1 + B_1 x_2 + C_1 x_3 = b_1, \ b_1 \in R^{m_1} & \rightarrow m_1 \text{ égalités} \\ A_2 x_1 + B_2 x_2 + C_2 x_3 \le b_2, \ b_2 \in R^{m_2} & \rightarrow m_2 \text{ inégalités inférieur} \\ A_3 x_1 + B_3 x_2 + C_3 x_3 \ge b_3, \ b_3 \in R^{m_3} & \rightarrow m_3 \text{ inégalités supérieur} \\ x_1 \ge 0 & \rightarrow n_1 \text{ variables positives} \\ x_2 \le 0 & \rightarrow n_2 \text{ variables négatives} \\ x_3 \in R^{n_3} & \rightarrow n_3 \text{ variables libres} \end{cases}$$

- **Problème dual (D)**

$$\max_{\substack{y_1 \in R^{m1} \\ y_2 \in R^{m2} \\ y_3 \in R^{m3}}} b_1^T y_1 + b_2^T y_2 + b_3^T y_3 \quad \text{sous} \begin{cases} A_1^T y_1 + A_2^T y_2 + A_3^T y_3 \le c_1, \ c_1 \in R^{n_1} & \rightarrow n_1 \text{ inégalités inférieur} \\ B_1^T y_1 + B_2^T y_2 + B_3^T y_3 \ge c_2, \ c_2 \in R^{n_2} & \rightarrow n_2 \text{ inégalités supérieur} \\ C_1^T y_1 + C_2^T y_2 + C_3^T y_3 = c_3, \ c_3 \in R^{n_3} & \rightarrow n_3 \text{ égalités} \\ y_1 \in R^{m_1} & \rightarrow m_1 \text{ variables libres} \\ y_2 \le 0 & \rightarrow m_2 \text{ variables négatives} \\ y_3 \ge 0 & \rightarrow m_3 \text{ variables positives} \end{cases}$$

Correspondance primal-dual

Preuve

- *Lagrangien du problème primal (P)*

$$\begin{array}{lll} L(x_1,x_2,x_3,\lambda_1,\lambda_2,\lambda_3,\mu_1,\mu_2) = & c_1^T x_1 + c_2^T x_2 + c_3^T x_3 & \\ & + \lambda_1^T (b_1 - A_1 x_1 - B_1 x_2 - C_1 x_3) & \rightarrow m_1 \text{ multiplicateurs } \lambda_1 \\ & + \lambda_2^T (A_2 x_1 + B_2 x_2 + C_2 x_3 - b_2) & \rightarrow m_2 \text{ multiplicateurs } \lambda_2 \geq 0 \\ & + \lambda_3^T (b_3 - A_3 x_1 - B_3 x_2 - C_3 x_3) & \rightarrow m_3 \text{ multiplicateurs } \lambda_3 \geq 0 \\ & - \mu_1^T x_1 + \mu_2^T x_2 & \rightarrow n_1 \text{ multiplicateurs } \mu_1 \geq 0 \\ & & \rightarrow n_2 \text{ multiplicateurs } \mu_2 \geq 0 \end{array}$$

- *On regroupe les termes en x_1,x_2,x_3 :*

$$\begin{array}{lll} L(x_1,x_2,x_3,\lambda_1,\lambda_2,\lambda_3,\mu_1,\mu_2) = & b_1^T\lambda_1 - b_2^T\lambda_2 + b_3^T\lambda_3 & \\ & + \left(c_1 - A_1^T\lambda_1 + A_2^T\lambda_2 - A_3^T\lambda_3 - \mu_1\right)^T x_1 & \rightarrow \text{linéaire en } x \\ & + \left(c_2 - B_1^T\lambda_1 + B_2^T\lambda_2 - B_3^T\lambda_3 + \mu_2\right)^T x_2 & \\ & + \left(c_3 - C_1^T\lambda_1 + C_2^T\lambda_2 - C_3^T\lambda_3\right)^T x_3 & \end{array}$$

- *La fonction duale est définie par :* $w(\lambda,\mu) = \min_x L(x,\lambda,\mu)$ $\rightarrow$ *bornée si les coefficients de x_1,x_2,x_3 sont nuls*

Correspondance primal-dual

- $$\begin{array}{ll} L(x_1,x_2,x_3,\lambda_1,\lambda_2,\lambda_3,\mu_1,\mu_2) = & b_1^T\lambda_1 - b_2^T\lambda_2 + b_3^T\lambda_3 \\ & + \left(c_1 - A_1^T\lambda_1 + A_2^T\lambda_2 - A_3^T\lambda_3 - \mu_1\right)^T x_1 \\ & + \left(c_2 - B_1^T\lambda_1 + B_2^T\lambda_2 - B_3^T\lambda_3 + \mu_2\right)^T x_2 \\ & + \left(c_3 - C_1^T\lambda_1 + C_2^T\lambda_2 - C_3^T\lambda_3\right)^T x_3 \end{array}$$

- *L bornée* $\Rightarrow \begin{cases} c_1 - A_1^T\lambda_1 + A_2^T\lambda_2 - A_3^T\lambda_3 - \mu_1 = 0 \\ c_2 - B_1^T\lambda_1 + B_2^T\lambda_2 - B_3^T\lambda_3 + \mu_2 = 0 \\ c_3 - C_1^T\lambda_1 + C_2^T\lambda_2 - C_3^T\lambda_3 = 0 \end{cases}$ *avec* $\begin{cases} \mu_1,\mu_2 \geq 0 \\ \lambda_2,\lambda_3 \geq 0 \end{cases}$

- *En posant :* $\begin{cases} y_1 = \lambda_1 \\ y_2 = -\lambda_2 \leq 0 \\ y_3 = \lambda_3 \geq 0 \end{cases} \Rightarrow \begin{cases} A_1^T y_1 + A_2^T y_2 + A_3^T y_3 = c_1 + \mu_1 \leq c_1 & \text{car } \mu_1 \geq 0 \\ B_1^T y_1 + B_2^T y_2 + B_3^T y_3 = c_2 - \mu_2 \geq c_2 & \text{car } \mu_2 \geq 0 \\ C_1^T y_1 + C_2^T y_2 + C_3^T y_3 = c_3 \end{cases}$

- *Fonction duale :* $w(y_1,y_2,y_3) = L(\lambda_1,\lambda_2,\lambda_3,\mu_1,\mu_2) \Rightarrow w(y_1,y_2,y_3) = b_1^T y_1 + b_2^T y_2 + b_3^T y_3$

- *Problème dual :* $\max_{y_1,y_2,y_3} w(y_1,y_2,y_3) = b_1^T y_1 + b_2^T y_2 + b_3^T y_3$ *sous* $\begin{cases} A_1^T y_1 + A_2^T y_2 + A_3^T y_3 \leq c_1 \\ B_1^T y_1 + B_2^T y_2 + B_3^T y_3 \geq c_2 \\ C_1^T y_1 + C_2^T y_2 + C_3^T y_3 = c_3 \end{cases}$

Correspondance primal-dual

- **Problème primal (P)**

$$(P) \quad \min_{x_1,x_2,x_3} x_1+2x_2+3x_3 \quad \text{sous} \begin{cases} -x_1+3x_2 = 5 \\ 2x_1-x_2+3x_3 \geq 6 \\ x_3 \leq 4 \\ x_1 \geq 0,\ x_2 \leq 0,\ x_3 \in R \end{cases}$$

- **Problème dual (D)**

$$(D) \quad \max_{y_1,y_2,y_3} 5y_1+6y_2+4y_3 \quad \text{sous} \begin{cases} -y_1+2y_2 \leq 1 \\ 3y_1-y_2 \geq 2 \\ 3y_2+y_3 = 3 \\ y_1 \in R,\ y_2 \geq 0,\ y_3 \leq 0 \end{cases}$$

$$\Leftrightarrow \quad \min_{y_1,y_2,y_3} -5y_1-6y_2-4y_3 \quad \text{sous} \begin{cases} y_1-2y_2 \geq -1 \\ -3y_1+y_2 \leq -2 \\ -3y_2-y_3 = -3 \\ y_1 \in R,\ y_2 \geq 0,\ y_3 \leq 0 \end{cases}$$

- Problème dual du dual : on retrouve le problème primal (P)

3.2 Conditions nécessaires

Problème avec contraintes

$$\min_{x \in R^n} f(x) \text{ sous } \begin{cases} c_E(x)=0 \\ c_I(x) \leq 0 \end{cases} \quad \begin{array}{l} \rightarrow \text{ p contraintes d'égalité} \\ \rightarrow \text{ q contraintes d'inégalité} \end{array}$$

Conditions nécessaires

x^* minimum local $\Rightarrow$ $\nabla f(x^*)^T d \geq 0$ pour toute direction d admissible à la limite en x^*

Méthode directe

Nécessite de connaître l'ensemble des directions admissibles en x^*

- Cas de contraintes linéaires
  → Définition des directions admissibles à partir des directions de base (§1.2.2)
- Cas de contraintes non linéaires
  → Définition des directions admissibles à la limite
  → Pas de caractérisation des directions admissibles dans le cas général
  sauf hypothèse de qualification des contraintes : cône des directions (§1.3.1)

Méthode indirecte

A partir des multiplicateurs de Lagrange

→ **Conditions d'optimalité dans le cas général**

Problème avec contraintes

$$\min_{x\in R^n} f(x) \text{ sous } \begin{cases} c_E(x)=0 & \rightarrow \text{ p contraintes d'égalité} \\ c_I(x)\le 0 & \rightarrow \text{ q contraintes d'inégalité} \end{cases}$$

Conditions nécessaires

Hypothèse : Contraintes linéairement indépendantes en x*

x* minimum local $\Rightarrow$ Il existe un unique $\lambda^* \in R^p$ et un unique $\mu^* \in R^q$ tels que :

- **Ordre 1** :

$$\begin{cases} \nabla_x L(x^*,\lambda^*,\mu^*)=0 & \rightarrow \text{ conditions nécessaires d'ordre 1} \\ \nabla_\lambda L(x^*,\lambda^*,\mu^*)=0 & \rightarrow \text{ contraintes égalité } c_E(x^*)=0 \\ \nabla_\mu L(x^*,\lambda^*,\mu^*)\le 0 & \rightarrow \text{ contraintes inégalité } c_I(x^*)\le 0 \\ \mu^* \ge 0 & \\ \mu^* c_I(x^*)=0 & \rightarrow \textbf{ conditions complémentaires} \end{cases}$$

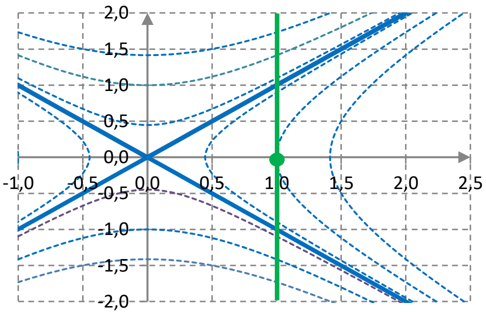
- **Ordre 2** : Pour toute direction d tangente aux contraintes actives (c(x*)=0) :

$$d^T\nabla^2_{xx}L(x^*,\lambda^*,\mu^*)d \ge 0 \quad \rightarrow \text{ conditions nécessaires d'ordre 2}$$

$$\forall d \ / \ d^T\nabla c(x^*)=0$$

$\rightarrow$ **Conditions nécessaires de Karush-Kuhn-Tucker (conditions KKT)**
(1939) (1951)

Conditions nécessaires



$$\min_{x_1,x_2} \frac{1}{2}\left(x_2^2 - x_1^2\right) \text{ sous } x_1 \le 1$$

Lagrangien : $L(x_1,x_2,\mu)=\frac{1}{2}\left(x_2^2-x_1^2\right)+\mu(x_1-1)$

- **Conditions nécessaires d'ordre 1**

$$\begin{cases} -x_1+\mu=0 \\ x_2=0 \\ x_1\le 1 \\ \mu\ge 0 \\ \mu(x_1-1)=0 \end{cases} \rightarrow \text{ vérifiées en } x^*=\begin{pmatrix}1\\0\end{pmatrix},\ \mu^*=1$$

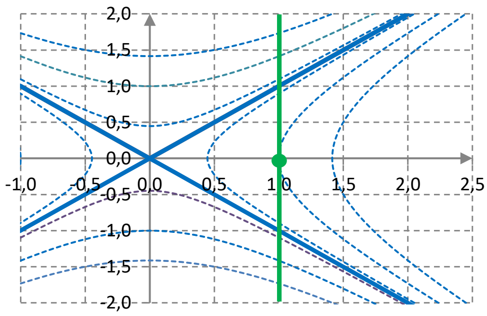
- **Conditions nécessaires d'ordre 2**

d direction tangente aux contraintes actives : $d^T\nabla c(x^*)=0 \Rightarrow \begin{pmatrix}d_1\\d_2\end{pmatrix}^T\begin{pmatrix}1\\0\end{pmatrix}=0 \Rightarrow d_1=0$

$$d^T\nabla^2_{xx}L(x^*,\mu^*)d=\begin{pmatrix}d_1\\d_2\end{pmatrix}^T\begin{pmatrix}-1&0\\0&1\end{pmatrix}\begin{pmatrix}d_1\\d_2\end{pmatrix}=-d_1^2+d_2^2=d_2^2\ge 0$$ qui est vérifié pour tout d

$\boxed{x^*=\begin{pmatrix}1\\0\end{pmatrix},\ \mu^*=1}$ vérifie les conditions nécessaires d'ordre 1 et 2.

Conditions nécessaires



$$\min_{x_1,x_2} \frac{1}{2}\left(x_2^2 - x_1^2\right) \text{ sous } x_1 \le 1$$

Lagrangien : $L(x_1, x_2, \mu) = \frac{1}{2}\left(x_2^2 - x_1^2\right) + \mu(x_1 - 1)$

- **Conditions nécessaires d'ordre 1**

$$\begin{cases} -x_1 + \mu = 0 \\ x_2 = 0 \\ x_1 \le 1 \\ \mu \ge 0 \\ \mu(x_1 - 1) = 0 \end{cases} \rightarrow \text{ vérifiées en } \quad x^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix},\ \mu^* = 0$$

- **Conditions nécessaires d'ordre 2**

Aucune contrainte n'est active en x_1=0. Pour toute direction d on doit avoir :

$$d^T \nabla^2_{xx} L(x^*,\mu^*) d = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}^T \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = -d_1^2 + d_2^2 \ge 0 \qquad \text{qui n'est pas vérifié pour } d = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$x^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $\mu^* = 0$ vérifie les conditions nécessaires d'ordre 1, **mais pas d'ordre 2**.

Multiplicateur du critère

$$\min_{x \in R^n} f(x) \text{ sous } c(x) = 0 \qquad \rightarrow \text{ m contraintes d'égalité (= contraintes actives)}$$

- Dans le cas général, il faut définir le lagrangien avec un **multiplicateur λ_0** sur le critère.

$L(x,\lambda,\lambda_0) = \lambda_0 f(x) + \lambda^T c(x)$ avec $\lambda \in R^m$: multiplicateurs des contraintes
$\lambda_0 \in R$: multiplicateur du critère $\rightarrow$ $\boldsymbol{\lambda_0 \ge 0}$

- Les conditions KKT donnent un système de n+m équations à n+m+1 inconnues.

$$\nabla_x L(x,\lambda,\lambda_0) = 0 \rightarrow n$$
$$\nabla_\lambda L(x,\lambda,\lambda_0) = 0 \rightarrow m$$

- Dans le **cas normal**, le système admet une solution pour toute valeur $\lambda_0 > 0$.
Les multiplicateurs sont définis à une constante multiplicative près $\rightarrow$ on choisit $\boldsymbol{\lambda_0 = 1}$.

- Dans le **cas anormal**, le système n'admet une solution que si $\boldsymbol{\lambda_0 = 0}$.
Ce cas correspond à un ensemble admissible réduit à des **points isolés**.
La solution satisfait les contraintes, mais la valeur du critère n'est pas « minimisable ».
$\rightarrow$ équivaut à une valeur infinie des multiplicateurs λ

Multiplicateur du critère

$$\min_{x_1,x_2} x_1 \text{ sous } x_1^2+x_2^2=0 \quad \rightarrow \text{ solution unique vérifiant la contrainte}$$

Lagrangien : $L(x_1,x_2,\lambda,\lambda_0)=\lambda_0 x_1+\lambda(x_1^2+x_2^2)$ avec $\lambda_0 \geq 0$

- Conditions nécessaires d'ordre 1

$$\begin{cases}\lambda_0+2\lambda x_1=0\\ 2\lambda x_2=0\\ x_1^2+x_2^2=0\end{cases} \quad \rightarrow \text{ système de 3 équations à 4 inconnues}$$

- Solution : $\begin{cases}x_1=0\\ x_2=0\end{cases} \Rightarrow \begin{cases}\lambda_0=0\\ \lambda \text{ quelconque}\end{cases}$ → **cas anormal**

  La contrainte n'est satisfaite qu'en un point isolé (0 , 0).
  Le multiplicateur λ_0 du critère est nul (⇔ critère indifférent, pas de minimisation possible).

- Si l'on écrit le lagrangien sans le multiplicateur λ_0, les conditions KKT sont

$$\begin{cases}1+2\lambda x_1=0\\ 2\lambda x_2=0\\ x_1^2+x_2^2=0\end{cases} \quad \rightarrow \text{ solution } \begin{cases}x_1=0\\ x_2=0\end{cases} \Rightarrow \lambda=\infty \quad (\Leftrightarrow \text{ critère indifférent})$$

181

3.3 Conditions suffisantes

Problème avec contraintes

$$\min_{x\in R^n} f(x) \text{ sous } \begin{cases}c_E(x)=0\\ c_I(x)\leq 0\end{cases} \quad \begin{array}{l}\rightarrow \text{ p contraintes d'égalité}\\ \rightarrow \text{ q contraintes d'inégalité}\end{array}$$

Conditions suffisantes

S'il existe $x^*\in R^n$, $\lambda^*\in R^p$, $\mu^*\in R^q$ tels que :

- **Ordre 1** :

$$\begin{cases}\nabla_x L(x^*,\lambda^*,\mu^*)=0 & \rightarrow \text{ conditions d'ordre 1}\\ \nabla_\lambda L(x^*,\lambda^*,\mu^*)=0 & \rightarrow \text{ contraintes égalité } c_E(x^*)=0\\ \nabla_\mu L(x^*,\lambda^*,\mu^*)\leq 0 & \rightarrow \text{ contraintes inégalité } c_I(x^*)\leq 0\\ \mu^*\geq 0 & \\ \mu^* c_I(x^*)=0 & \rightarrow \text{ conditions complémentaires}\\ \mu_k^*>0 \text{ si } c_{Ik}(x^*)=0 & \rightarrow \textbf{ contraintes actives : multiplicateur > 0}\end{cases}$$

- **Ordre 2** : Pour toute direction d tangente aux contraintes actives ($c(x^*)=0$) :

$$d^T\nabla_{xx}^2 L(x^*,\lambda^*,\mu^*)d>0 \quad \rightarrow \text{ conditions d'ordre 2}$$
$$\forall d \,/\, d^T\nabla c(x^*)=0$$

⇒ x^* est un minimum local strict

Remarque : Pas d'hypothèse de qualification des contraintes dans les conditions suffisantes

Eléments de la démonstration

Cas de contraintes égalité : c(x)=0

On suppose que (x^*,λ^*) vérifie les conditions suffisantes.

- *On considère le problème sans contrainte*

$$\boxed{\min_{x\in R^n} L_\rho(x,\lambda^*) = L(x,\lambda^*) + \frac{1}{2}\rho\|c(x)\|^2 = f(x) + \lambda^{*T} c(x) + \frac{1}{2}\rho\|c(x)\|^2}$$

$L_\rho(x,\lambda)$ = ***lagrangien augmenté***

$\rho>0$ = *pénalisation de la violation des contraintes*

- $\nabla_x L_\rho(x^*,\lambda^*) = \nabla f(x^*) + \nabla c(x^*)\lambda^* + \rho\nabla c(x^*)c(x^*)$

$= \nabla f(x^*) + \nabla c(x^*)\lambda^*$ $\quad\rightarrow$ *car x^* admissible*

$= \nabla_x L(x^*,\lambda^*) = 0$ $\quad\rightarrow$ *par hypothèse sur x^*,λ^**

- $\nabla^2_{xx} L_\rho(x^*,\lambda^*) = \nabla^2_{xx} L(x^*,\lambda^*) + \rho\nabla c(x^*)\nabla c(x^*)^T$ $\quad\rightarrow$ *définie positive pour ρ assez grand*

$\Rightarrow$ *x^* est un minimum local du lagrangien augmenté $L_\rho(x,\lambda^*)$ pour $\lambda=\lambda^*$.*

- *Au voisinage de x^* :* $L_\rho(x^*,\lambda^*) \le L_\rho(x,\lambda^*) \Rightarrow f(x^*,\lambda^*) \le f(x,\lambda^*)$, $\forall x \,/\, c(x)=0$

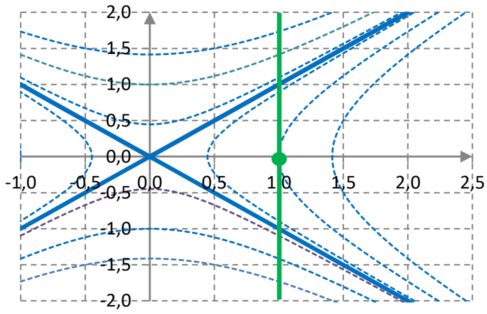
$\Rightarrow$ *x^* est un minimum local de f*

3.4 Exemples

Conditions suffisantes

$$\min_{x_1,x_2} \frac{1}{2}\left(x_2^2 - x_1^2\right) \text{ sous } x_1 \le 1$$

$x^* = \begin{pmatrix}1\\0\end{pmatrix}$, $\mu^* = 1$ vérifie les conditions nécessaires



- **Conditions suffisantes d'ordre 1**

Contrainte active $\rightarrow$ multiplicateur > 0

$x^* - 1 = 0 \qquad \mu^* = 1 > 0$

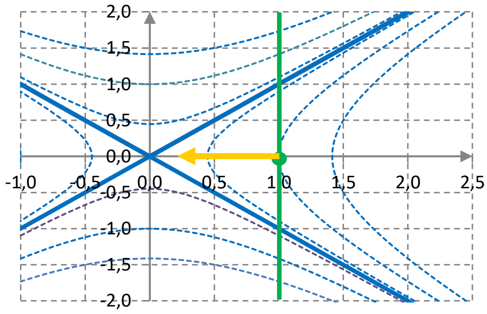
- **Conditions suffisantes d'ordre 2**

d direction tangente aux contraintes actives : $d^T\nabla c(x^*) = 0 \Rightarrow \begin{pmatrix}d_1\\d_2\end{pmatrix}^T \begin{pmatrix}1\\0\end{pmatrix} = 0 \Rightarrow d_1 = 0$

$$d^T\nabla^2_{xx}L(x^*,\mu^*)d = \begin{pmatrix}d_1\\d_2\end{pmatrix}^T \begin{pmatrix}-1 & 0\\0 & 1\end{pmatrix}\begin{pmatrix}d_1\\d_2\end{pmatrix} = -d_1^2 + d_2^2 = d_2^2 > 0 \text{ car } d = \begin{pmatrix}0\\d_2\end{pmatrix} \ne 0 \Rightarrow d_2 \ne 0$$

$\boxed{x^* = \begin{pmatrix}1\\0\end{pmatrix}, \mu^* = 1}$ vérifie les conditions suffisantes d'ordre 1 et 2 $\rightarrow$ minimum local strict.

Remarque sur la condition d'ordre 2



$$\min_{x_1,x_2} \frac{1}{2}\left(x_2^2 - x_1^2\right) \text{ sous } x_1 \le 1 \quad \rightarrow x^* = \begin{pmatrix}1\\0\end{pmatrix},\ \mu^* = 1$$

$d = \begin{pmatrix}-1\\0\end{pmatrix}$ est une direction admissible en $x^* = \begin{pmatrix}1\\0\end{pmatrix}$

mais la condition d'ordre 2 ne porte que
sur les <u>directions tangentes</u> aux contraintes actives.

Importance de la condition de complémentarité

$$\min_{x_1,x_2} \frac{1}{2}\left(x_2^2 - x_1^2\right) \text{ sous } x_1 \le 0$$

$x^* = \begin{pmatrix}0\\0\end{pmatrix}$, $\mu^* = 0$ vérifie les conditions suffisantes d'ordre 1 et 2
sauf la condition de complémentarité

Si la contrainte inégalité est active, le multiplicateur doit être strictement positif.

$x_1^* = 0$ est active et $\mu^* = 0$ n'est pas strictement positif
$\rightarrow$ x^* n'est pas un minimum local (f décroit suivant x_1 <0)

3.5 Interprétation géométrique

Interprétation

- Condition complémentaire

$$\mu_j c_{Ij}(x) = 0, \quad j = 1,\dots,q \quad \Rightarrow \begin{cases} \mu_j = 0 & \rightarrow \text{ sensibilité nulle} \\ \text{ou} & \\ c_{Ij}(x) = 0 & \rightarrow \text{ contrainte active} \end{cases}$$

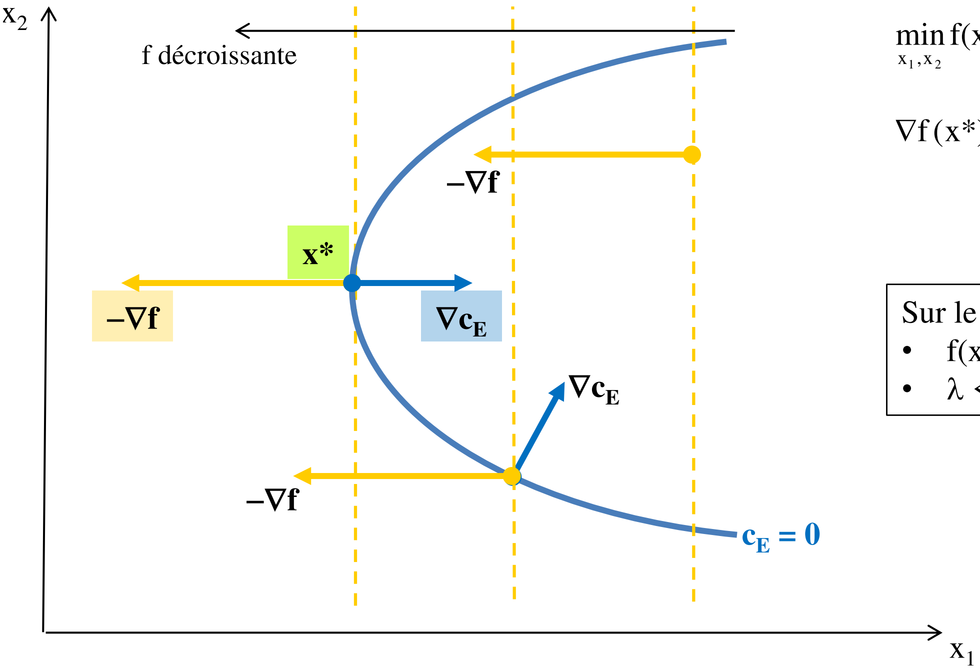
- Condition d'ordre 1

$$\begin{aligned} \nabla_x L(x,\lambda,\mu) = 0 \ &\Rightarrow \ \nabla f(x) + \nabla c_E(x).\lambda + \nabla c_I(x).\mu = 0 \\ &\Rightarrow -\nabla f(x) = \nabla c_E(x).\lambda + \nabla c_I(x).\mu \\ &\Rightarrow -\nabla f(x) = \nabla c(x).\nu \quad \rightarrow \text{ contraintes actives } c(x) \end{aligned}$$

La direction $-\nabla f(x)$ est la direction de plus forte descente en x.
Les directions $\nabla c(x)$ sont orthogonales à l'hyperplan tangent aux contraintes actives en x.
Equation de l'hyperplan tangent aux contraintes actives en x : $d^T \nabla c(x) = 0$

$\rightarrow$ Les déplacements admissibles (dans l'hyperplan tangent) sont orthogonaux au gradient.
$\rightarrow$ **Déplacements suivant les lignes de niveau de f, sans diminution du critère**.

Fonction de 2 variables – 1 contrainte égalité

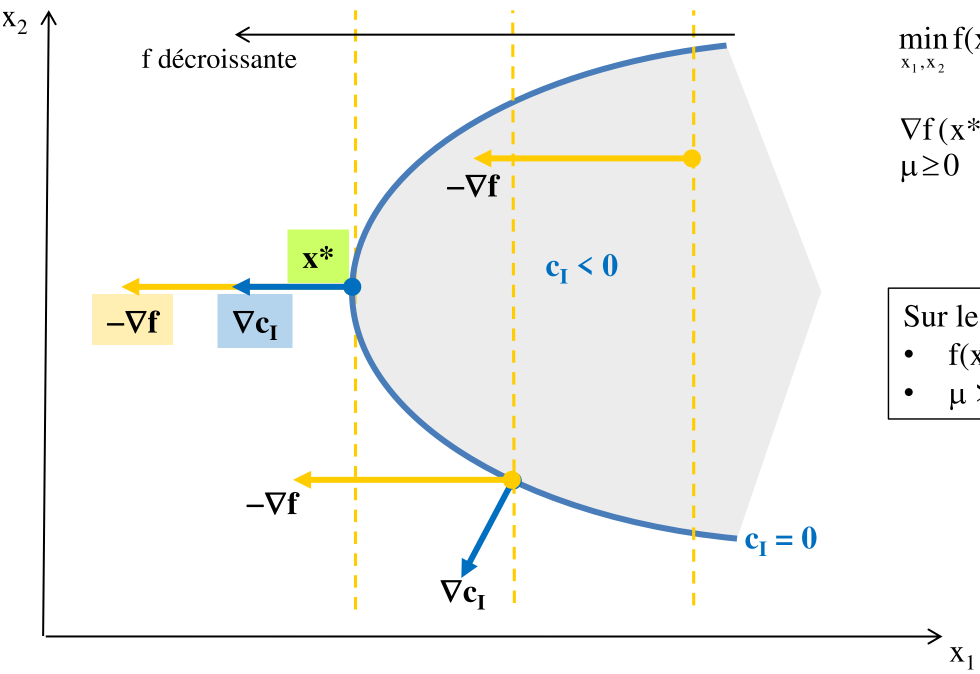


$$\min_{x_1,x_2} f(x_1,x_2) \text{ sous } c_E(x_1,x_2)=0$$

$$\nabla f(x^*)+\lambda \nabla c_E(x^*)=0$$

Sur le schéma :
- $f(x_1, x_2) = x_1$ à minimiser
- $\lambda < 0 \quad (\lambda \approx -2)$

Fonction de 2 variables – 1 contrainte inégalité

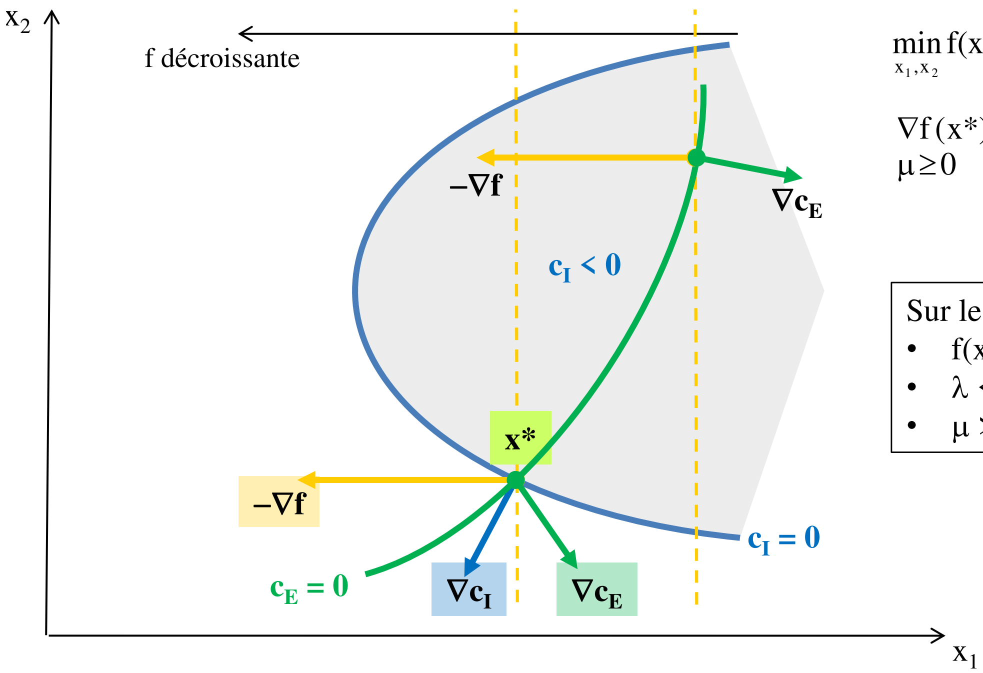


$$\min_{x_1,x_2} f(x_1,x_2) \text{ sous } c_I(x_1,x_2)\le 0$$

$$\nabla f(x^*)+\mu \nabla c_I(x^*)=0$$
$$\mu \ge 0$$

Sur le schéma :
- $f(x_1, x_2) = x_1$ à minimiser
- $\mu > 0 \quad (\mu \approx 2)$

Fonction de 2 variables – 1 contrainte égalité – 1 contrainte inégalité

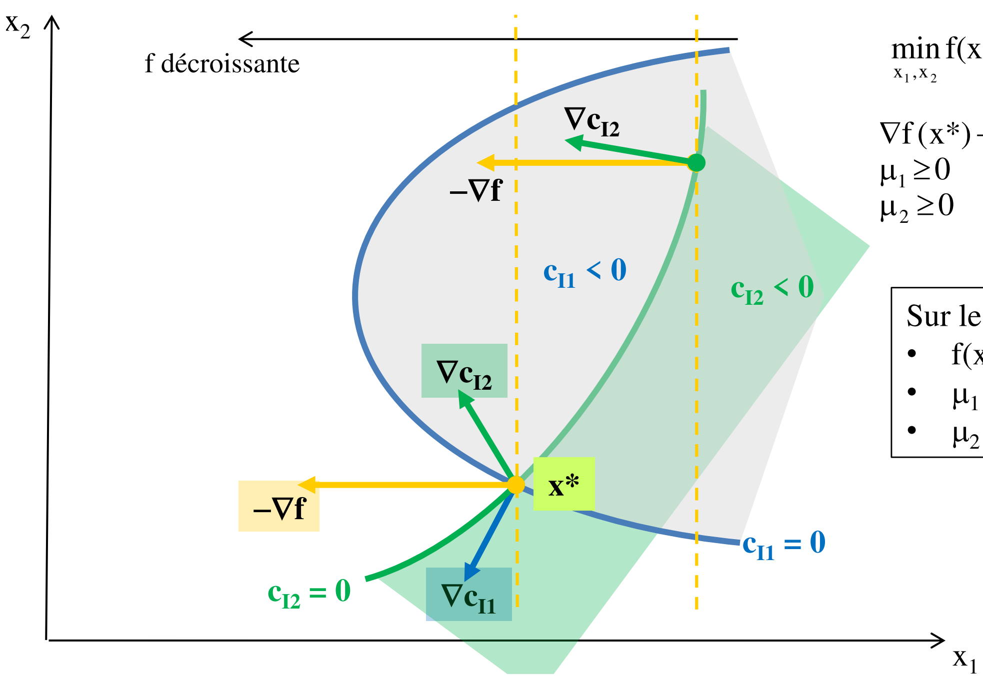


$$\min_{x_1,x_2} f(x_1,x_2) \text{ sous } \begin{cases} c_E(x_1,x_2)=0 \\ c_I(x_1,x_2)\le 0 \end{cases}$$

$$\nabla f(x^*)+\lambda\nabla c_E(x^*)+\mu\nabla c_I(x^*)=0$$
$$\mu\ge 0$$

Sur le schéma :
- $f(x_1, x_2) = x_1$ à minimiser
- $\lambda < 0 \quad (\lambda \approx -1.5)$
- $\mu > 0 \quad (\mu \approx \ 1.5)$

Fonction de 2 variables – 2 contraintes inégalité

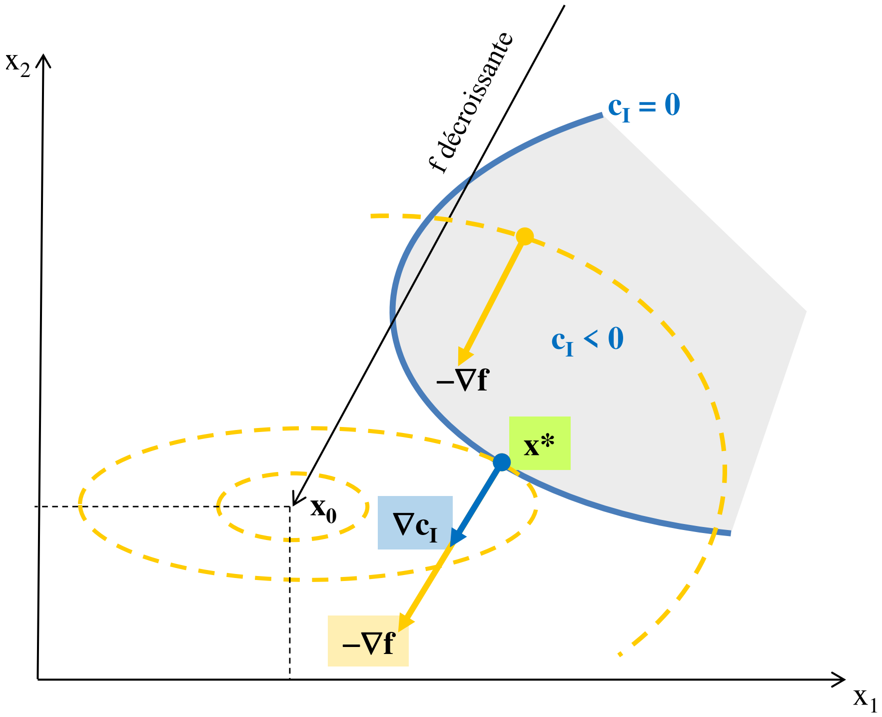


$$\min_{x_1,x_2} f(x_1,x_2) \text{ sous } \begin{cases} c_{I1}(x_1,x_2)\le 0 \\ c_{I2}(x_1,x_2)\le 0 \end{cases}$$

$$\nabla f(x^*)+\mu_1\nabla c_{I1}(x^*)+\mu_2\nabla c_{I2}(x^*)=0$$
$$\mu_1\ge 0$$
$$\mu_2\ge 0$$

Sur le schéma :
- $f(x_1, x_2) = x_1$ à minimiser
- $\mu_1 > 0 \quad (\mu_1 \approx \ 1.5)$
- $\mu_2 > 0 \quad (\mu_2 \approx \ 1.5)$

Fonction de 2 variables – 1 contrainte inégalité



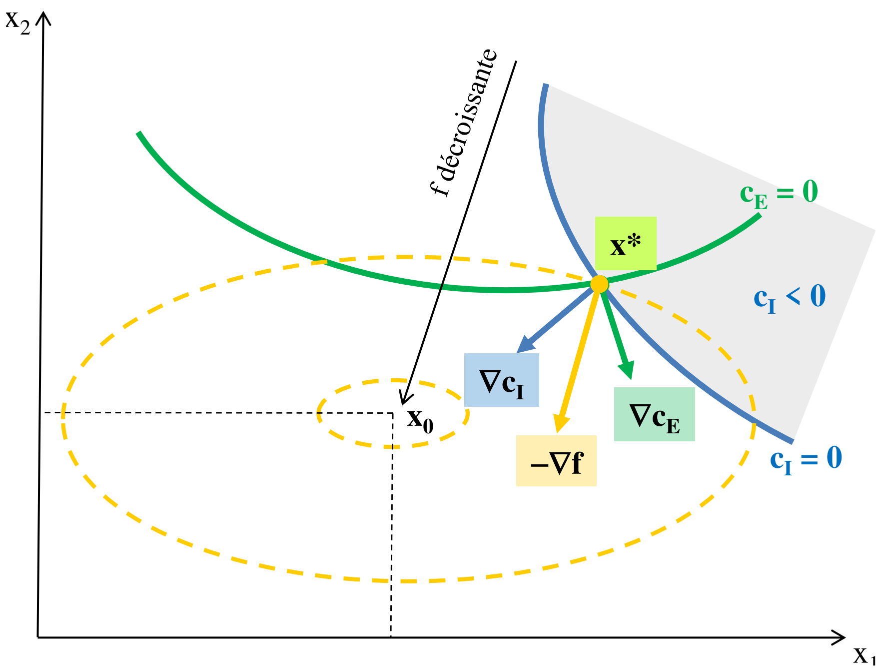
$$\min_{x_1,x_2} f(x_1,x_2) \text{ sous } c_I(x_1,x_2) \le 0$$

$$\nabla f(x^*) + \mu \nabla c_I(x^*) = 0$$
$$\mu \ge 0$$

Sur le schéma :

- $f(x_1, x_2)$ quadratique
- x_0 = minimum sans contrainte
- $\mu > 0 \quad (\mu \approx 2)$

Fonction de 2 variables – 1 contrainte égalité – 1 contrainte inégalité



$$\min_{x_1,x_2} f(x_1,x_2) \text{ sous } \begin{cases} c_E(x_1,x_2) = 0 \\ c_I(x_1,x_2) \le 0 \end{cases}$$

$$\nabla f(x^*) + \lambda \nabla c_E(x^*) + \mu \nabla c_I(x^*) = 0$$
$$\mu \ge 0$$

Sur le schéma :

- $f(x_1, x_2)$ quadratique
- x_0 = minimum sans contrainte
- $\lambda > 0 \quad (\lambda \approx 1.0)$
- $\mu > 0 \quad (\mu \approx 1.0)$

3.6 Méthode pratique

Problème avec contraintes

$$\min_{x\in R^n} f(x) \text{ sous } \begin{cases} c_E(x)=0 \\ c_I(x)\le 0 \end{cases} \qquad \begin{aligned} &\rightarrow \text{ p contraintes d'égalité} \\ &\rightarrow \text{ q contraintes d'inégalité} \end{aligned}$$

La résolution analytique ou numérique nécessite d'**identifier les contraintes actives**.
On se ramène à un problème avec contraintes égalité plus simple.
→ résolution des conditions KKT d'ordre 1
→ vérification des conditions réduites d'ordre 2

Identification des contraintes actives

- Résolution analytique → problème combinatoire (conditions complémentaires)
- Résolution numérique → mise à jour itérative de l'ensemble des contraintes actives

Stratégie itérative d'identification

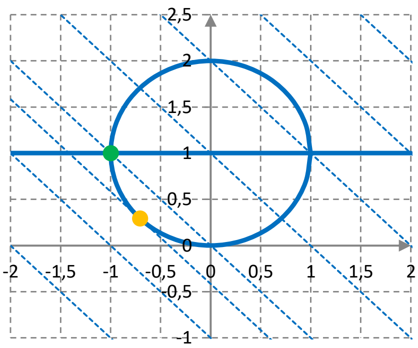
- On cherche un déplacement à partir du point courant sans tenir compte des contraintes inégalité
- Le déplacement peut rendre actives certaines contraintes inégalité.
- On reprend la recherche en ajoutant la première contrainte inégalité activée.

→ résolution d'une succession de problèmes avec contraintes égalité

$$\min_{x\in R^n} f(x) \text{ sous } c(x)=0 \qquad \rightarrow \text{ m contraintes actives}$$

Problème avec 2 contraintes inégalité

$$\min_{x_1,x_2} f(x)=x_1+x_2 \quad \text{sous} \quad \begin{cases} c_1(x)=x_1^2+(x_2-1)^2-1\le 0 \\ c_2(x)=1-x_2\le 0 \end{cases}$$



- Lagrangien : $L(x,\mu)=f(x)+\mu_1 c_1(x)+\mu_2 c_2(x)$
$$= x_1+x_2+\mu_1\left(x_1^2+(x_2-1)^2-1\right)+\mu_2(1-x_2)$$

- Conditions KKT d'ordre 1

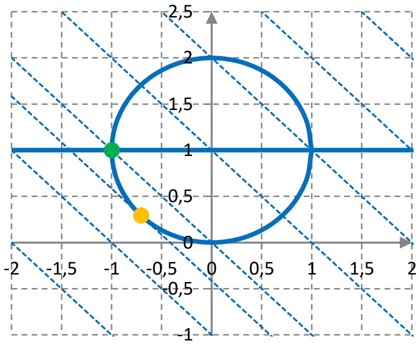
$$\begin{cases} 1+2\mu_1 x_1=0 \\ 1+2\mu_1(x_2-1)-\mu_2=0 \\ x_1^2+(x_2-1)^2-1\le 0 \\ 1-x_2\le 0 \\ \mu_1 c_1(x)=0 \\ \mu_2 c_2(x)=0 \\ \mu_1,\mu_2\ge 0 \end{cases} \qquad \rightarrow \text{ conditions complémentaires : } \textbf{4 combinaisons possibles}$$

Identification des contraintes actives

Problème combinatoire : il faut essayer les 4 possibilités $\begin{cases} \mu_1=0 \text{ ou } c_1(x)=0 \\ \mu_2=0 \text{ ou } c_2(x)=0 \end{cases}$

Problème avec 2 contraintes inégalité

$$\min_{x_1,x_2} f(x) = x_1 + x_2 \quad \text{sous} \quad \begin{cases} c_1(x) = x_1^2 + (x_2-1)^2 - 1 \le 0 \\ c_2(x) = 1 - x_2 \le 0 \end{cases}$$



- Si $\mu_1 = 0$ $\rightarrow$ incompatible équation $1 + 2\mu_1 x_1 = 0$

$\Rightarrow$ $c_1(x) = 0$ $\rightarrow$ c_1 contrainte active

- Si $\mu_2 = 0$ $\Rightarrow$ $1 + 2\mu_1(x_2 - 1) = 0$
  $\rightarrow$ incompatible équations $\begin{cases} 1 - x_2 \le 0 \\ \mu_1 \ge 0 \end{cases}$

$\Rightarrow$ $c_2(x) = 0$ $\rightarrow$ c_2 contrainte active

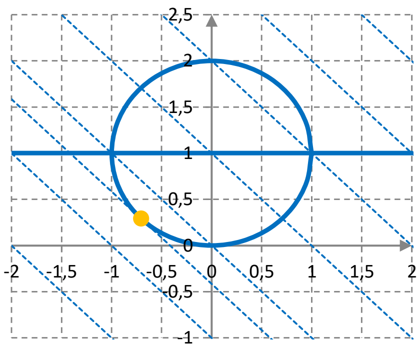
- Combinaison retenue $\begin{cases} c_1(x) = 0 \\ c_2(x) = 0 \end{cases} \Rightarrow \begin{cases} x_1 = \pm 1 \\ x_2 = 1 \end{cases} \Rightarrow \begin{cases} \mu_1 = \mp 0.5 \text{ et } \mu_1 \ge 0 \rightarrow x_1 = -1 \\ \mu_2 = 1 \end{cases}$

- Solution : $\begin{cases} x_1 = -1 \\ x_2 = 1 \\ \mu_1 = 0.5 \\ \mu_2 = 1 \end{cases}$ **Vérification condition d'ordre 2 :** cône admissible vide (2 contraintes actives)

  $\rightarrow$ **minimum local**

Changement de sens contrainte 2

$$\min_{x_1,x_2} f(x) = x_1 + x_2 \quad \text{sous} \quad \begin{cases} c_1(x) = x_1^2 + (x_2-1)^2 - 1 \le 0 \\ c_2(x) = x_2 - 1 \le 0 \end{cases}$$



- Lagrangien : $L(x,\mu) = f(x) + \mu_1 c_1(x) + \mu_2 c_2(x)$
  $= x_1 + x_2 + \mu_1\left(x_1^2 + (x_2-1)^2 - 1\right) + \mu_2(x_2 - 1)$

- Conditions KKT d'ordre 1

$$\begin{cases} 1 + 2\mu_1 x_1 = 0 \\ 1 + 2\mu_1(x_2-1) + \mu_2 = 0 \\ x_1^2 + (x_2-1)^2 - 1 \le 0 \\ x_2 - 1 \le 0 \\ \mu_1 c_1(x) = 0 \\ \mu_2 c_2(x) = 0 \\ \mu_1, \mu_2 \ge 0 \end{cases}$$

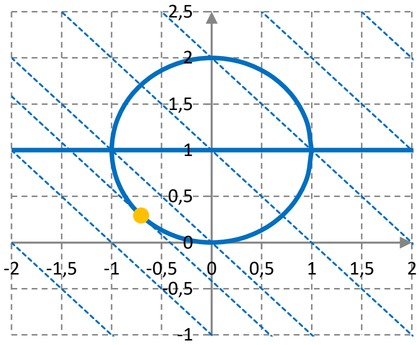
$\rightarrow$ conditions complémentaires : **4 combinaisons possibles**

Identification des contraintes actives

Problème combinatoire : il faut essayer les 4 possibilités $\begin{cases} \mu_1 = 0 \text{ ou } c_1(x) = 0 \\ \mu_2 = 0 \text{ ou } c_2(x) = 0 \end{cases}$

Changement de sens contrainte 2

$$\min_{x_1,x_2} f(x) = x_1 + x_2 \quad \text{sous} \quad \begin{cases} c_1(x) = x_1^2 + (x_2-1)^2 - 1 \le 0 \\ c_2(x) = x_2 - 1 \le 0 \end{cases}$$



- Si $\mu_1 = 0 \quad \rightarrow$ incompatible équation $\quad 1 + 2\mu_1 x_1 = 0$

  $\Rightarrow \; c_1(x) = 0 \quad \rightarrow c_1$ contrainte active

- Si $c_2(x) = 0 \;\Rightarrow\; \begin{cases} x_1 = \pm 1 \\ x_2 = 1 \end{cases} \;\Rightarrow\; \begin{cases} \mu_1 = \mp 0.5 \\ \mu_2 = -1 \end{cases}$

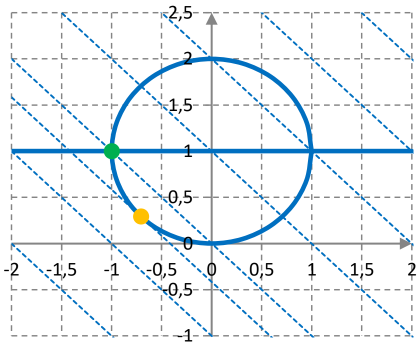
  $\rightarrow$ incompatible condition $\quad \mu_2 \ge 0$

- Si $\mu_2 = 0 \;\Rightarrow\; \begin{cases} 1 + 2\mu_1 x_1 = 0 \\ 1 + 2\mu_1(x_2 - 1) = 0 \\ x_1^2 + (x_2-1)^2 - 1 = 0 \end{cases} \;\Rightarrow\; \begin{cases} x_1 = -1/(2\mu_1) \\ x_2 = 1 - 1/(2\mu_1) \\ \mu_1 = 1/\sqrt{2} > 0 \end{cases} \;\Rightarrow\; \begin{cases} x_1 = -1/\sqrt{2} \\ x_2 = 1 - 1/\sqrt{2} \\ \mu_1 = 1/\sqrt{2} \end{cases}$

- **Vérification condition d'ordre 2** : $\quad \nabla^2_{xx} L(x,\mu) = \begin{pmatrix} 2\mu_1 & 0 \\ 0 & 2\mu_1 \end{pmatrix} > 0 \quad \rightarrow$ **minimum local**
  (1 contrainte active)

Passage contrainte 1 en égalité

$$\min_{x_1,x_2} f(x) = x_1 + x_2 \quad \text{sous} \quad \begin{cases} c_1(x) = x_1^2 + (x_2-1)^2 - 1 = 0 \\ c_2(x) = 1 - x_2 \le 0 \end{cases}$$



- Lagrangien : $L(x,\mu) = f(x) + \lambda_1 c_1(x) + \mu_2 c_2(x)$
  $\qquad\qquad\qquad = x_1 + x_2 + \lambda_1\left(x_1^2 + (x_2-1)^2 - 1\right) + \mu_2(1 - x_2)$

- Conditions KKT d'ordre 1

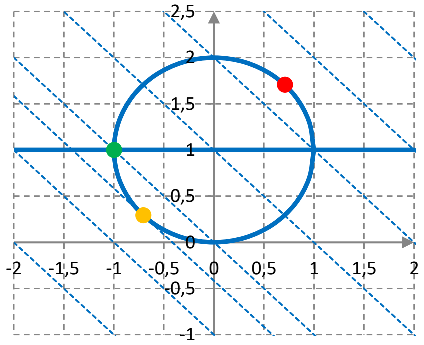
$$\begin{cases} 1 + 2\lambda_1 x_1 = 0 \\ 1 + 2\lambda_1(x_2 - 1) - \mu_2 = 0 \\ x_1^2 + (x_2-1)^2 - 1 = 0 \\ 1 - x_2 \le 0 \\ \mu_2 c_2(x) = 0 \quad \rightarrow \text{conditions complémentaires : } \textbf{2 combinaisons possibles} \\ \mu_2 \ge 0 \end{cases}$$

Identification des contraintes actives

Problème combinatoire : il faut essayer les 2 possibilités $\mu_2 = 0$ ou $c_2(x) = 0$

Passage contrainte 1 en égalité

$$\min_{x_1,x_2} f(x) = x_1 + x_2 \quad \text{sous} \quad \begin{cases} c_1(x) = x_1^2 + (x_2 - 1)^2 - 1 = 0 \\ c_2(x) = 1 - x_2 \le 0 \end{cases}$$



- Si $\mu_2 = 0 \Rightarrow \begin{cases} 1 + 2\lambda_1 x_1 = 0 \\ 1 + 2\lambda_1 (x_2 - 1) = 0 \\ x_1^2 + (x_2 - 1)^2 - 1 = 0 \end{cases} \rightarrow \lambda_1 \ne 0$

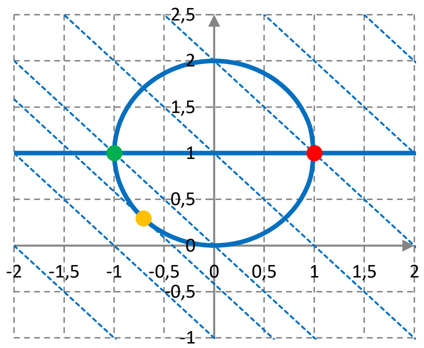
$$\Rightarrow \begin{cases} x_1 = -1/(2\lambda_1) \\ x_2 = 1 - 1/(2\lambda_1) \\ \lambda_1 = \pm 1/\sqrt{2} \end{cases}$$

$$1 - x_2 \le 0 \Rightarrow \begin{cases} x_1 = 1/\sqrt{2} \\ x_2 = 1 + 1/\sqrt{2} \\ \lambda_1 = -1/\sqrt{2} \end{cases}$$

- **Vérification condition d'ordre 2** : $\nabla^2_{xx} L(x,\mu) = \begin{pmatrix} 2\lambda_1 & 0 \\ 0 & 2\lambda_1 \end{pmatrix} < 0$ $\rightarrow$ maximum local
  (1 contrainte active) $\rightarrow$ **solution rejetée**

Passage contrainte 1 en égalité

$$\min_{x_1,x_2} f(x) = x_1 + x_2 \quad \text{sous} \quad \begin{cases} c_1(x) = x_1^2 + (x_2 - 1)^2 - 1 = 0 \\ c_2(x) = 1 - x_2 \le 0 \end{cases}$$



- Si $c_2(x) = 0 \Rightarrow \begin{cases} x_1 = \pm 1 \\ x_2 = 1 \end{cases} \Rightarrow \begin{cases} \lambda_1 = \mp 0.5 \\ \mu_2 = 1 \end{cases}$

- **Vérification condition d'ordre 2**

$$\nabla^2_{xx} L(x,\mu) = \begin{pmatrix} 2\lambda_1 & 0 \\ 0 & 2\lambda_1 \end{pmatrix}$$

Cône admissible vide : $\nabla c_1 = \begin{pmatrix} 2x_1 \\ 2(x_2 - 1) \end{pmatrix} = \begin{pmatrix} 2x_1 \\ 0 \end{pmatrix}$, $\nabla c_2 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$ $\rightarrow$ 2 contraintes actives

$\rightarrow$ **2 minima locaux** : $\begin{cases} x_1 = -1 \\ x_2 = 1 \\ \lambda_1 = 0.5 \\ \mu_2 = 1 \end{cases} \rightarrow f(x) = 0$ et $\begin{cases} x_1 = 1 \\ x_2 = 1 \\ \lambda_1 = -0.5 \\ \mu_2 = 1 \end{cases} \rightarrow f(x) = 2$

Problème avec contraintes actives

$\min_{x \in R^n} f(x)$ sous $c(x)=0$ $\quad\rightarrow$ m contraintes actives

Résolution des conditions KKT

On cherche $x^* \in R^n$ et $\lambda^* \in R^m$ vérifiant les conditions KKT.

- **Condition nécessaire du 1er ordre**

$$\begin{cases} \nabla_x L(x^*,\lambda^*) = 0 & \rightarrow \text{ n équations} \\ \nabla_\lambda L(x^*,\lambda^*) = 0 & \rightarrow \text{ m équations} \end{cases}$$

Les n équations $\nabla_x L(x^*,\lambda^*)$ permettent d'exprimer $x^* \in R^n$ en fonction de $\lambda^* \in R^m$
On remplace ensuite $x^*(\lambda^*)$ dans les m équations $\nabla_\lambda L(x^*,\lambda^*)$.
$\rightarrow$ **système de m équations à m inconnues $\lambda^* \in R^m$**

- **Condition nécessaire du 2ème ordre**

Il faut vérifier que : $\begin{cases} d^T \nabla^2_{xx} L(x^*,\lambda^*) d \geq 0 \\ \forall d \; / \;\; d^T \nabla c(x^*) = 0 \end{cases}$ $\rightarrow$ hessien du lagrangien semi-défini positif sur le cône admissible

Condition difficile à vérifier sous cette forme $\rightarrow$ passage au **hessien réduit**

Problème avec contraintes actives

$\min_{x \in R^n} f(x)$ sous $c(x)=0$ $\quad\rightarrow$ m contraintes actives

Problème équivalent

- Les conditions nécessaires de minimum de f sous contraintes sont :

$$\begin{cases} \nabla_x L(x^*,\lambda^*) = 0 \\ \nabla_\lambda L(x^*,\lambda^*) = 0 \\ d^T \nabla^2_{xx} L(x^*,\lambda^*) d \geq 0 \,,\; \forall d \; / \;\; d^T \nabla c(x^*) = 0 \end{cases}$$

- On observe qu'il s'agit également des conditions nécessaires du problème :

$\min_{x \in R^n} L(x,\lambda^*)$ sous $c(x)=0$

- Il est équivalent de minimiser f(x) ou $L(x,\lambda^*)$, **si l'on connaît λ^***.

$$\boxed{\min_{x \in R^n} f(x) \text{ sous } c(x)=0 \;\Leftrightarrow\; \min_{x \in R^n} L(x,\lambda^*) \text{ sous } c(x)=0}$$

- On écrit les conditions nécessaires sur le modèle quadratique-linéaire local, puis on applique la technique de réduction des contraintes linéaires.

Problème équivalent

$$\min_{x\in R^n} L(x,\lambda^*) \text{ sous } c(x)=0$$

Modèle quadratique-linéaire

- Modèle quadratique du critère : $\hat{L}(x^*+p) = L(x^*,\lambda^*) + p^T\nabla_x L(x^*,\lambda^*) + \frac{1}{2}p^T\nabla^2_{xx}L(x^*,\lambda^*)p$

  En notant : $\begin{cases} g_L(x^*) = \nabla_x L(x^*,\lambda^*) & \rightarrow \text{ gradient du lagrangien par rapport à x} \\ H_L(x^*) = \nabla^2_{xx}L(x^*,\lambda^*) & \rightarrow \text{ hessien du lagrangien par rapport à x} \end{cases}$

  $$\rightarrow \hat{L}(x^*+p) = L(x^*,\lambda^*) + p^T g_L(x^*) + \frac{1}{2}p^T H_L(x^*)p$$

- Modèle linéaire des contraintes : $\hat{c}(x^*+p) = c(x^*) + \nabla c(x^*)^T p$ avec $c(x^*)=0$

  En notant : $\begin{cases} A = \nabla c(x^*)^T \\ p = Yp_Y + Zp_Z \end{cases}$ avec $\begin{array}{l} AY \text{ inversible} \\ AZ = 0 \text{ (espace nul)} \end{array}$

  $$\nabla c(x^*)^T p = 0 \Leftrightarrow \begin{cases} AYp_Y = 0 \\ p_Z \text{ libre} \end{cases} \Leftrightarrow \begin{cases} p_Y = 0 \\ p_Z \text{ libre} \end{cases} \text{ car AY inversible}$$

- Problème réduit : $\boxed{\min_{p\in R^n} \hat{L}(x^*+p) \text{ sous } \hat{c}(x^*+p)=0 \Leftrightarrow \min_{p_Z\in R^{n-m}} \hat{L}(x^*+Zp_Z)}$

Problème réduit

$$\min_{p_Z\in R^{n-m}} \hat{L}(x^*+Zp_Z) \quad \rightarrow \text{ problème sans contrainte à n−m variables } p_Z$$

avec $\hat{L}(x^*+Zp_Z) = L(x^*,\lambda^*) + p_Z^T Z^T g_L(x^*) + \frac{1}{2}p_Z^T Z^T H_L(x^*)Zp_Z$

Conditions nécessaires de minimum du problème réduit

$$\hat{L}(x^*+Zp_Z) \ge \hat{L}(x^*), \forall p_Z \Rightarrow p_Z^T Z^T g_L(x^*) + \frac{1}{2}p_Z^T Z^T H_L(x^*)Zp_Z \ge 0, \forall p_Z \Rightarrow \begin{cases} Z^T g_L(x^*) = 0 \\ Z^T H_L(x^*)Z \ge 0 \end{cases}$$

- Condition réduite d'ordre 1 : $Z^T g_L(x^*) = 0 \Leftrightarrow Z^T\nabla_x L(x^*,\lambda^*) = Z^T\left(\nabla f(x^*) + \nabla c(x^*)\lambda^{*T}\right) = 0$

  $$\Leftrightarrow Z^T\nabla f(x^*) = 0 \text{ car } \nabla c(x^*)^T Z = 0$$

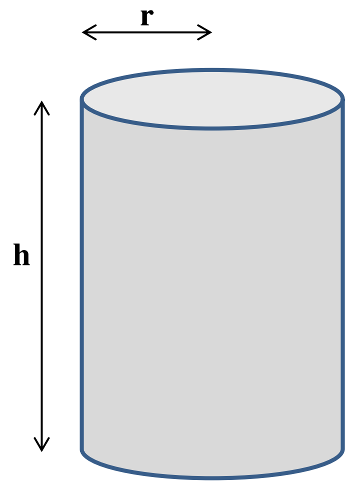
- Condition réduite d'ordre 2 : $Z^T H_L(x^*)Z \ge 0$

- $\begin{cases} g_Z = Z^T g & \rightarrow \textbf{gradient réduit du critère} \\ H_Z = Z^T H_L Z & \rightarrow \textbf{hessien réduit du lagrangien} \end{cases}$ $\quad \begin{array}{l} g(x) = \nabla f(x) \\ H_L(x) = \nabla^2_{xx}L(x,\lambda^*) \end{array}$

- x^* minimum local $\Rightarrow \begin{cases} g_Z(x^*) = 0 & \rightarrow \text{ gradient réduit du critère nul} \\ H_Z(x^*) \ge 0 & \rightarrow \text{ hessien réduit du lagrangien semi-défini positif} \end{cases}$

Problème de la boîte

- Réaliser une boîte cylindrique de volume donné V_0 et de surface minimale
- Dimensions : hauteur = h , rayon = r



Formulation du problème

- Surface : $S = 2\pi r^2 + 2\pi rh$
- Volume : $V = \pi r^2 h$

$\rightarrow \min_{h,r} S(h,r) \quad \text{sous} \quad V(h,r) = V_0$

Résolution

On note : $\mathbf{V_0 = 2\pi v_0}$

- Lagrangien : $L(h,r,\lambda) = 2\pi r^2 + 2\pi rh + \lambda(\pi r^2 h - 2\pi v_0)$
- Conditions KKT

$$\begin{cases} 2\pi r + \lambda\pi r^2 & = 0 \\ 4\pi r + 2\pi h + 2\lambda\pi rh & = 0 \\ \pi r^2 h - 2\pi v_0 & = 0 \end{cases} \Rightarrow \begin{cases} \lambda r + 2 & = 0 \\ 2r + h + \lambda rh & = 0 \\ r^2 h - 2v_0 & = 0 \end{cases} \Rightarrow \begin{cases} \lambda r & = -2 \\ h & = 2r \\ r^3 & = v_0 \end{cases}$$

- Solution : $\begin{cases} r = v_0^{1/3} \\ h = 2v_0^{1/3} \end{cases} \Rightarrow S = 6\pi v_0^{2/3} = 3(2\pi)^{1/3} V_0^{2/3}$

Vérification des conditions réduites

Il faut choisir une base de réduction, puis vérifier les conditions réduites de minimum local.

$$\begin{cases} g_Z(x^*) = Z^T \nabla f(x^*) = 0 & \rightarrow \text{gradient réduit du critère nul} \\ H_Z(x^*) = Z^T \nabla^2_{xx} L(x,\lambda^*) Z \geq 0 & \rightarrow \text{hessien réduit du lagrangien semi-défini positif} \end{cases}$$

- Gradient du critère : $g(h,r) = \nabla_{h,r} f(h,r) = 2\pi \begin{pmatrix} r \\ 2r+h \end{pmatrix}$
- Hessien du lagrangien : $L(h,r,\lambda) = 2\pi r^2 + 2\pi rh + \lambda(\pi r^2 h - 2\pi v_0)$

$$\Rightarrow g_L(h,r) = \nabla_{h,r} L(h,r,\lambda) = \pi \begin{pmatrix} 2r + \lambda r^2 \\ 4r + 2h + 2\lambda rh \end{pmatrix} \ , \quad H_L(h,r) = \nabla^2_{h,r} L(h,r,\lambda) = 2\pi \begin{pmatrix} 0 & 1+\lambda r \\ 1+\lambda r & 2+\lambda h \end{pmatrix}$$

Choix d'une base de réduction

- Contrainte : $c(h,r) = \pi r^2 h - 2\pi v_0 = 0 \Rightarrow \nabla c^T = \begin{pmatrix} \pi r^2 & 2\pi rh \end{pmatrix}$
- **Choix de la base avec la variable h** : $A = \nabla c^T = \begin{pmatrix} \pi r^2 & 2\pi rh \end{pmatrix} = \begin{array}{c} \begin{array}{cc} h & r \end{array} \\ \begin{pmatrix} B & N \end{pmatrix} \end{array}$
- Base de l'espace nul : $Z = \begin{pmatrix} -B^{-1}N \\ I \end{pmatrix} = \begin{pmatrix} -2\pi rh/\pi r^2 \\ 1 \end{pmatrix} = \begin{pmatrix} -2h/r \\ 1 \end{pmatrix}$

Vérification des conditions réduites

- **Gradient réduit du critère** : $g_Z(h,r) = Z^T g(h,r) = 2\pi \begin{pmatrix} -2h/r \\ 1 \end{pmatrix}^T \begin{pmatrix} r \\ 2r+h \end{pmatrix} = 2\pi(2r-h)$

  On vérifie que le gradient réduit est nul : $h = 2r \;\Rightarrow\; g_Z(h,r) = 0$

- **Hessien réduit du lagrangien** : $H_Z(h,r) = Z^T H(h,r) Z = 2\pi \begin{pmatrix} -2h/r \\ 1 \end{pmatrix}^T \begin{pmatrix} 0 & 1+\lambda r \\ 1+\lambda r & 2+\lambda r \end{pmatrix} \begin{pmatrix} -2h/r \\ 1 \end{pmatrix}$

$$\Rightarrow\; H_Z(h,r) = 2\pi\left(2 - 4\frac{h}{r} - 3\lambda h\right)$$

  On vérifie que le hessien réduit est semi-défini positif

$$\begin{cases} \lambda r = -2 \\ h = 2r \end{cases} \Rightarrow\; H_Z(h,r) = 2\pi\left(2 - \frac{h}{r}(4 + 3\lambda r)\right) = 12\pi > 0$$

Résolution par élimination

- Contrainte : $c(h,r) = \pi r^2 h - 2\pi v_0 = 0 \;\Rightarrow\; h = \dfrac{2v_0}{r^2}$
- Elimination de la variable h : $S(h,r) = 2\pi r^2 + 2\pi r h \;\Rightarrow\; S(r) = 2\pi r^2 + \dfrac{4\pi v_0}{r}$
- Gradient : $\dfrac{dS}{dr}(r) = 4\pi r - \dfrac{4\pi v_0}{r^2} = 4\pi r\left(1 - \dfrac{v_0}{r^3}\right)$
- Hessien : $\dfrac{d^2S}{dr^2}(r) = 4\pi + \dfrac{8\pi v_0}{r^3} = 4\pi\left(1 + 2\dfrac{v_0}{r^3}\right)$
- Minimum de S(r) : $\begin{cases} \dfrac{dS}{dr}(r) = 0 \\ \dfrac{d^2S}{dr^2}(r) \geq 0 \end{cases} \Rightarrow \begin{cases} r^3 = v_0 \\ \dfrac{d^2S}{dr^2}(r) = 12\pi > 0 \end{cases}$

Lien avec les conditions réduites

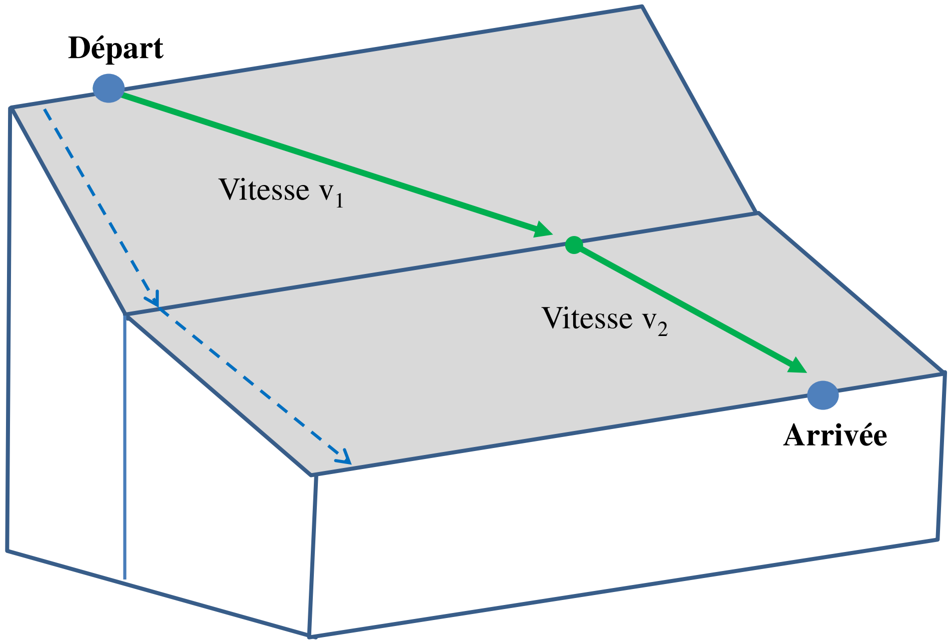
- Gradient réduit

  $g_Z(h,r) = 2\pi(2r-h)$ avec $h = \dfrac{2v_0}{r^2} \;\Rightarrow\; g_Z(h,r) = 2\pi\left(2r - \dfrac{2v_0}{r^2}\right) = 4\pi r\left(1 - \dfrac{v_0}{r^2}\right) = \dfrac{dS}{dr}(r)$

- Hessien réduit

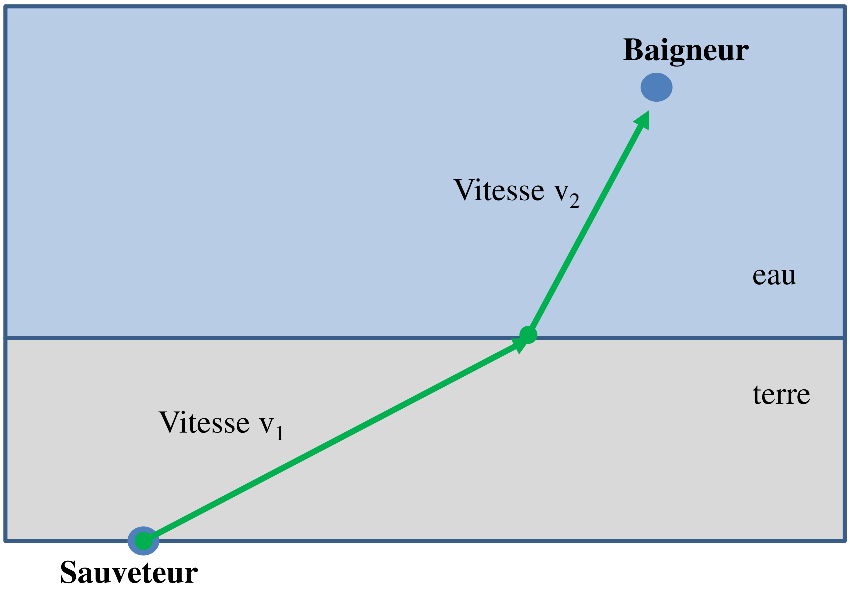
  $\rightarrow$ pas de relation directe entre H_Z et $\dfrac{d^2S}{dr^2}$ (contrainte non linéaire)

Problème du skieur
- Descendre du départ à l'arrivée le plus vite possible
- 2 zones de pentes différentes : vitesse v_1, puis v_2



209

Problème du sauveteur
- Aller secourir le baigneur qui se noie le plus vite possible
- Course sur terre, puis nage dans l'eau : vitesse v_1, puis v_2



Problème du sauveteur

- Données du problème : l_0, l_1, l_2, v_1, v_2

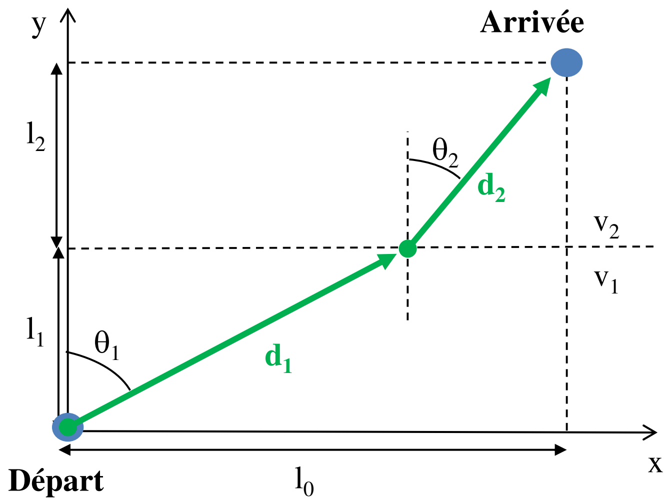
Distance sur terre : $d_1 = \dfrac{l_1}{\cos\theta_1}$

Durée de course : $t_1 = \dfrac{d_1}{v_1}$

Distance dans l'eau : $d_2 = \dfrac{l_2}{\cos\theta_2}$

Durée de nage : $t_2 = \dfrac{d_2}{v_2}$

Distance suivant x : $L = d_1 \sin\theta_1 + d_2 \sin\theta_2$



- Formulation du problème

Variables : θ_1, θ_2
Contrainte : $L = l_0$ → atteindre le point visé
Critère : $T = t_1 + t_2$ → durée totale à minimiser

Problème du sauveteur

- Formulation du problème

$$\min_{\theta_1,\theta_2} T(\theta_1,\theta_2) \text{ sous } L(\theta_1,\theta_2) = l_0$$

$$\Leftrightarrow \min_{\theta_1,\theta_2} T = \frac{l_1}{v_1\cos\theta_1} + \frac{l_2}{v_2\cos\theta_2} \quad \text{sous} \quad L = l_1\tan\theta_1 + l_2\tan\theta_2 = l_0$$

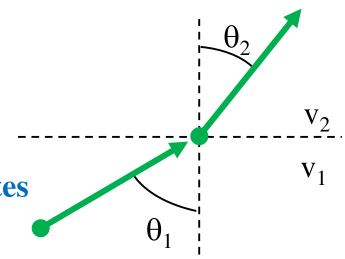
- Résolution du problème

Lagrangien : $$L(\theta_1,\theta_2,\lambda) = \frac{l_1}{v_1\cos\theta_1} + \frac{l_2}{v_2\cos\theta_2} + \lambda\left(l_1\tan\theta_1 + l_2\tan\theta_2 - l_0\right)$$

Conditions KKT : $$\begin{cases} \dfrac{l_1\sin\theta_1}{v_1\cos^2\theta_1} + \lambda l_1 \dfrac{1}{\cos^2\theta_1} = 0 \\ \dfrac{l_2\sin\theta_2}{v_2\cos^2\theta_2} + \lambda l_2 \dfrac{1}{\cos^2\theta_2} = 0 \\ l_1\tan\theta_1 + l_2\tan\theta_2 = l_0 \end{cases} \Rightarrow \begin{cases} \sin\theta_1 + \lambda v_1 = 0 \\ \sin\theta_2 + \lambda v_2 = 0 \\ l_1\tan\theta_1 + l_2\tan\theta_2 = l_0 \end{cases}$$

Problème du sauveteur

- Conditions KKT : $\begin{cases} \sin\theta_1 + \lambda v_1 = 0 \\ \sin\theta_2 + \lambda v_2 = 0 \\ l_1 \tan\theta_1 + l_2 \tan\theta_2 = l_0 \end{cases}$



- θ_1, θ_2 vérifient : $\dfrac{\sin\theta_1}{\sin\theta_2} = \dfrac{v_1}{v_2}$ → **loi de la réfraction de Descartes**

- Pour résoudre complètement

  On exprime θ_1, θ_2 en fonction de λ : $\begin{cases} \sin\theta_1 = -\lambda v_1 & \Rightarrow \cos\theta_1 = \sqrt{1-\lambda^2 v_1^2} \\ \sin\theta_2 = -\lambda v_2 & \Rightarrow \cos\theta_2 = \sqrt{1-\lambda^2 v_2^2} \end{cases}$

  On remplace dans la contrainte : $l_1 \tan\theta_1 + l_2 \tan\theta_2 = l_0 \Rightarrow \dfrac{\lambda l_1 v_1}{\sqrt{1-\lambda^2 v_1^2}} + \dfrac{\lambda l_2 v_2}{\sqrt{1-\lambda^2 v_2^2}} = -l_0$

  On obtient une équation en λ : $\lambda l_1 v_1 \sqrt{1-\lambda^2 v_2^2} + \lambda l_2 v_2 \sqrt{1-\lambda^2 v_1^2} = -l_0 \sqrt{1-\lambda^2 v_2^2}\sqrt{1-\lambda^2 v_1^2}$

  → équation de degré 4

  → solution λ^* → θ_1^*, θ_2^*

a) Cas linéaire

Problème linéaire

$$\min_{x\in R^n} c^T x \ \text{ sous } \begin{cases} Ax = b \\ x \ge 0 \end{cases} \qquad A \in R^{m\times n},\ b \in R^m,\ c \in R^n$$

→ problème linéaire sous forme standard (PL)

Conditions nécessaires d'optimalité à partir du lagrangien

Lagrangien : $L(x,\lambda,s) = c^T x + \lambda^T(b - Ax) - s^T x \Rightarrow \begin{cases} \nabla_x L(x,\lambda,s) = c - A^T\lambda - s \\ \nabla^2_{xx} L(x,\lambda,s) = 0 \end{cases}$

(x,λ,s) minimum local de (PL)

$\Downarrow$

- Condition nécessaire d'ordre 1 : $\begin{cases} c - A^T\lambda - s = 0 \\ s \ge 0 \end{cases}$ → contraintes du problème dual
- Condition nécessaire d'ordre 2 : $\nabla^2_{xx} L(x,\lambda,s) \ge 0$ → vérifiée
- Condition complémentaire : $s_i x_i = 0,\ i = 1,\dots,n$

Problème linéaire

$$\min_{x\in R^n} c^T x \quad \text{sous} \quad \begin{cases} Ax = b \\ x \geq 0 \end{cases} \qquad A \in R^{m\times n},\ b \in R^m,\ c \in R^n$$

$\rightarrow$ problème linéaire sous forme standard (PL)

Conditions nécessaires d'optimalité à partir des dérivées directionnelles

x minimum local de (PL)

$\Downarrow$

Pour toute direction admissible d : $\nabla f(x)^T d \geq 0 \;\Rightarrow\; c^T d \geq 0$

- Toute direction admissible d est combinaison linéaire des directions de base d_j. (contraintes linéaires)

$$d_j = E\begin{pmatrix} d_{jB} \\ d_{jN} \end{pmatrix} \text{ avec } \begin{cases} d_{jB} = -B^{-1}A_j \\ E^T e_j = \begin{pmatrix} 0 \\ d_{jN} \end{pmatrix} \end{cases} \Rightarrow c^T d_j = c_B^T d_{jB} + c_N^T d_{jN} = -c_B^T B^{-1} A_j + c_j$$

- Il suffit de vérifier : $\boxed{c^T d_j \geq 0}$

Coûts réduits

x solution de base admissible : $x = E\begin{pmatrix} x_B \\ x_N \end{pmatrix} = E\begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix} \begin{matrix} \rightarrow m \\ \rightarrow n-m \end{matrix}$

Le **coût réduit** associé à la variable hors base x_j est défini par : $\boxed{\bar{c}_j = c^T d_j = c_j - c_B^T B^{-1} A_j}$

= **dérivée directionnelle** de f suivant la $j^{ème}$ direction de base pour une variable hors base
= 0 par extension pour une variable de base

$$AE = \begin{pmatrix} B & N \end{pmatrix} \Rightarrow B^{-1}AE = \begin{pmatrix} I & B^{-1}N \end{pmatrix} \Rightarrow c_B^T B^{-1} AE = \begin{pmatrix} c_B^T & c_B^T B^{-1} N \end{pmatrix} \quad \Rightarrow \quad \bar{c} = \begin{pmatrix} 0 & \bar{c}_N^T \end{pmatrix}$$

Conditions nécessaires d'optimalité

x* solution de base non dégénérée

x* solution de PL $\quad \Rightarrow \quad \bar{c} \geq 0$

Conditions suffisantes d'optimalité

x* solution de base admissible

$\bar{c} \geq 0 \quad \Rightarrow \quad$ x* solution de PL

Lien entre multiplicateurs et coûts réduits

- Problème linéaire sous forme standard : $\min_{x \in R^n} c^T x$ sous $\begin{cases} Ax = b \\ x \geq 0 \end{cases}$
- Lagrangien : $L(x,\lambda,s) = c^T x + \lambda^T (b - Ax) - s^T x \;\Rightarrow\; \nabla_x L(x,\lambda,s) = c - A^T\lambda - s$
- Conditions d'ordre 1 : $\begin{cases} A^T\lambda + s = c \\ s \geq 0 \;,\; s_i x_i = 0 \;,\; i = 1,\cdots,n \end{cases}$
- Base B : $AE = \begin{pmatrix} B & N \end{pmatrix} \Rightarrow c = \begin{pmatrix} c_B \\ c_N \end{pmatrix}, \; s = \begin{pmatrix} s_B \\ s_N \end{pmatrix}, \; x = \begin{pmatrix} x_B = B^{-1}b \\ x_N = 0 \end{pmatrix}$

$$A^T\lambda + s = c \quad \Leftrightarrow \quad \begin{pmatrix} B^T \\ N^T \end{pmatrix}\lambda + \begin{pmatrix} s_B \\ s_N \end{pmatrix} = \begin{pmatrix} c_B \\ c_N \end{pmatrix}$$

$$s \geq 0 \;,\; s_i x_i = 0 \;,\; i = 1,\cdots,n \quad \Leftrightarrow \quad \begin{cases} s_B \geq 0 \;,\; s_i x_i = 0 \;,\; i \in B \\ s_N \geq 0 \;,\; s_i x_i = 0 \;,\; i \in N \end{cases} \quad \begin{array}{l} \rightarrow \text{vérifié en prenant } s_B = 0 \\ \rightarrow \text{vérifié car } x_N = 0 \end{array}$$

$$\begin{cases} B^T\lambda = c_B \\ N^T\lambda + s_N = c_N \\ s_N \geq 0 \end{cases} \Rightarrow \begin{cases} \lambda = B^{-T}c_B \\ s_N = c_N - \left(B^{-1}N\right)^T c_B = \bar{c}_N \geq 0 \end{cases} \Rightarrow s = \begin{pmatrix} s_B \\ s_N \end{pmatrix} = \begin{pmatrix} 0 \\ \bar{c}_N \end{pmatrix} = \bar{c} \geq 0$$

- **Les coûts réduits sont les multiplicateurs des variables** $\rightarrow$ $\boxed{s = \bar{c} \geq 0}$

3.7 Exercices

Exercice 27: Multiplicateurs de Lagrange (1)

Trouver le maximum de la fonction $f(x, y) = xyz$ respectant les contraintes $(x, y, z) \in \mathbb{R}^3$, $x + y + z \leq 1$ et $x, y, z \geq 0$.

Exercice 28: Multiplicateurs de Lagrange (2)

Trouver le minimum de la fonction $f(x, y) = 2y - x^2$ respectant les contraintes $(x, y) \in \mathbb{R}^2$, $x^2 + y^2 \leq 1$ et $x, y \geq 0$.

Exercice 29: Multiplicateurs de Lagrange (3)

Résoudre le problème de minimisation : $\min x^2 + y^2 - 20x$ tel que $25x^2 + 4y^2 \leq 100$.

Exercice 30: Multiplicateurs de Lagrange (4)

Maximiser $x + y - 2z$ tel que $z \geq x^2 + y^2$ et $x, y, z \geq 0$.

4 Simplexe

4.1 Problème linéaire

a) Forme standard

Problème linéaire sous forme standard

$$\min_x c^T x \text{ sous } \begin{cases} Ax = b \\ x \geq 0 \end{cases} \quad \begin{array}{l} A \in R^{m \times n}, b \in R^m, c \in R^n \\ \text{rang}(A) = m \end{array} \quad \rightarrow \text{ problème noté (PL)}$$

Rappels

- **Base B** = $(A_{j1}, \dots, A_{jm})$ = m colonnes indépendantes de A $\Rightarrow$ B inversible
  $AE = (B \;\; N)$ avec E = matrice de permutation de colonnes ($EE^T = I$)
- **Solution de base** : $\boxed{x = E\begin{pmatrix} x_B \\ x_N \end{pmatrix} \begin{matrix} \to m \\ \to n-m \end{matrix} \quad \text{avec} \begin{cases} Ax = b \\ x_N = 0 \end{cases} \Rightarrow x = E\begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}}$
  admissible (ou réalisable) si $x_B = B^{-1}b \geq 0$
- **Direction de base** d_j, $j \in N$
  $\boxed{d_j = E\begin{pmatrix} d_{jB} \\ d_{jN} \end{pmatrix} \text{ avec } \begin{cases} d_{jB} = -B^{-1}A_j \\ E^T e_j = \begin{pmatrix} 0 \\ d_{jN} \end{pmatrix} \end{cases}}$ $\rightarrow$ pour vérifier Ax=b
  $\rightarrow$ composantes nulles sauf =1 sur la composante j
- **Coût réduit** = dérivée directionnelle suivant la direction de base d_j : $\bar{c}_j = c^T d_j = c_j - c_B^T B^{-1} A_j$
  Coût réduit négatif $\rightarrow$ direction de descente

Problème linéaire sous forme standard

$$\min_x c^T x \text{ sous } \begin{cases} Ax = b \\ x \geq 0 \end{cases} \quad \begin{array}{l} A \in R^{m \times n}, b \in R^m, c \in R^n \\ \text{rang}(A) = m \end{array}$$

Solution

On note P le polytope associé aux contraintes : $P = \{x \in R^n / Ax = b, x \geq 0\}$
Si le problème (PL) admet une solution, alors il existe un sommet optimal x*.

*Preuve : On suppose que PL admet une solution de coût f**
On considère un sommet x du polytope Q inclus dans P :* $Q = \{x \in P / c^T x = f^*\}$
On suppose par l'absurde que x n'est pas un sommet de P.*

$x^* = \alpha^* y + (1-\alpha^*)z$ *avec* $\boxed{y, z \in P}$, $y \neq z, 0 < \alpha^* < 1$

La fonction linéaire $\begin{cases} \varphi(\alpha) = \alpha c^T y + (1-\alpha)c^T z \\ 0 \leq \alpha \leq 1 \end{cases}$ *est minimale en* α^* : $\begin{cases} \varphi(\alpha^*) = c^T x^* = f^* \\ 0 < \alpha^* < 1 \end{cases}$

La fonction φ *est donc constante (sinon elle décroît d'un côté de* α^**)* $\Rightarrow c^T y = c^T z = f^*$
$\Rightarrow y, z \in Q$

On a donc : $x^* = \alpha^* y + (1-\alpha^*)z$ *avec* $\boxed{y, z \in Q}$, $y \neq z, 0 < \alpha^* < 1$
en contradiction avec l'hypothèse que x est un sommet de Q.

b) Recherche systématique

Recherche systématique

Si (PL) admet une solution x*, x* est un sommet du polytope P associé aux contraintes.
On peut donc trouver la solution : en parcourant tous les sommets de P (= bases)
en calculant les solutions de base associées
en conservant la meilleure solution de base réalisable

- Choix de m colonnes parmi les n colonnes de A $\rightarrow$ base B possible
- Vérification que la base est réalisable : B inversible et $B^{-1}b \geq 0$

  $\rightarrow$ Solution de base admissible : $x = E\begin{pmatrix} x_B \\ x_N \end{pmatrix} = E\begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix} \geq 0$
- Valeur du coût associé à la base B : $f = c^T x = c_B^T x_B + c_N^T x_N = c_B^T x_B$
- Sélection de la meilleure solution (f minimal)

Inconvénient

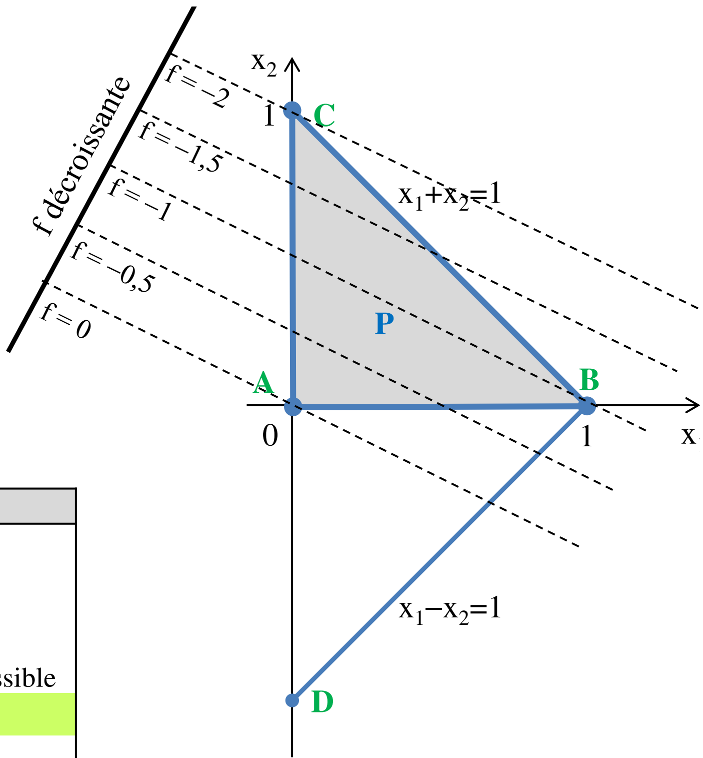
$C_n^m = \dfrac{n!}{m!(n-m)!}$ combinaisons possibles $\rightarrow$ **inapplicable en pratique**

Recherche systématique

$$\min_{x_1,x_2,x_3,x_4} -x_1 - 2x_2 \quad \text{sous} \quad \begin{cases} x_1 + x_2 + x_3 = 1 \\ x_1 - x_2 + x_4 = 1 \\ x_1, x_2, x_3, x_4 \geq 0 \end{cases}$$

- Représentation graphique dans R^2

  $P' = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} / \begin{cases} x_1 + x_2 \leq 1 \\ x_1 - x_2 \leq 1 \end{cases}, \begin{cases} x_1 \geq 0 \\ x_2 \geq 0 \end{cases} \right\}$
- Solution graphique : point C (f = −2)
- Solution par énumération des sommets



Base	x	f
x_1,x_2	$(1\ \ 0\ \ 0\ \ 0) \rightarrow B$	−1
x_1,x_3	$(1\ \ 0\ \ 0\ \ 0) \rightarrow B$	−1
x_1,x_4	$(1\ \ 0\ \ 0\ \ 0) \rightarrow B$	−1
x_2,x_3	$(0\ \ -1\ \ 2\ \ 0) \rightarrow D$	Non admissible
x_2,x_4	**$(0\ \ 1\ \ 0\ \ 2) \rightarrow C$**	**−2**
x_3,x_4	$(0\ \ 0\ \ 1\ \ 1) \rightarrow A$	0

c) Recherche optimisée

Recherche optimisée

On évite l'énumération systématique en parcourant les sommets de façon ordonnée
→ **Méthode du simplexe** = méthode de contraintes actives

Principes

- On se déplace d'une **solution de base admissible** à une autre solution de base admissible.
  → Les solutions non admissibles ne sont pas examinées.

- Les bases successives ne diffèrent que par l'une des variables (**bases adjacentes**)

- Le déplacement d'un sommet à un autre est choisi à partir des **directions de base**
  → Déplacement suivant les arêtes du polytope

- Les **coûts réduits** déterminent les directions de descente possibles.
  → Sélection d'une direction de déplacement (plusieurs règles de sélection possibles)

- Le problème est mis sous **forme canonique** dans la base B
  → Permet de vérifier l'optimalité de la base B
  → Permet de construire le déplacement vers une base adjacente

d) Recherche canonique

Réduction dans la base B

- Forme standard : $\min\limits_{x} c^T x \text{ sous } \begin{cases} Ax = b \\ x \ge 0 \end{cases}$ $\quad A \in R^{m \times n},\ b \in R^m,\ c \in R^n$, $\text{rang}(A) = m$

- Base B : $AE = (B \quad N),\ \ x = E\begin{pmatrix} x_B \\ x_N \end{pmatrix} \begin{matrix} \to m \\ \to n-m \end{matrix} \quad \Rightarrow \begin{cases} Ax = Bx_B + Nx_N \\ c^T x = c_B^T x_B + c_N^T x_N \end{cases}$

- Réduction aux variables hors base

$$\min_{x} c^T x \text{ sous } \begin{cases} Ax = b \\ x \ge 0 \end{cases} \Leftrightarrow \min_{\substack{x_B \in R^m \\ x_N \in R^{n-m}}} c_B^T x_B + c_N^T x_N \text{ sous } \begin{cases} Bx_B + Nx_N = b \Rightarrow x_B = B^{-1}b - B^{-1}Nx_N \\ x_B \ge 0, x_N \ge 0 \end{cases}$$

$$\Leftrightarrow \min_{x_N \in R^{n-m}} c_B^T B^{-1} b + \left(c_N^T - c_B^T B^{-1} N\right) x_N \text{ sous } \begin{cases} x_B = B^{-1}b - B^{-1}Nx_N \ge 0 \\ x_N \ge 0 \end{cases}$$

- **Forme canonique dans la base B**

$$\min_{\substack{x_N \in R^{n-m} \\ x_N \ge 0}} \bar{z} + \bar{c}_N^T x_N \text{ sous } x_B = \bar{b} - B^{-1}Nx_N \ge 0 \qquad \text{avec } \begin{cases} \bar{b} = B^{-1}b \\ \bar{z} = c_B^T \bar{b} \\ \bar{c}_N^T = c_N^T - c_B^T B^{-1} N \end{cases}$$

→ Réduction à n-m variables = variables hors-base x_N

Evaluation de la base B

La solution x* du problème linéaire correspond à un sommet = solution de base admissible
→ Evaluer l'optimalité de la solution de base associée à la base B
→ Construire le déplacement vers une nouvelle base B' meilleure que B

- Forme canonique dans B : $\min\limits_{\substack{x_N \in R^{n-m} \\ x_N \geq 0}} \bar{z} + \bar{c}_N^T x_N$ sous $x_B = \bar{b} - B^{-1} N x_N \geq 0$

- Solution de base associée à B : $x_N = 0 \Rightarrow \begin{cases} z = \bar{z} \\ x_B = \bar{b} \geq 0 \end{cases}$ si B est admissible (ou réalisable)

- Variation du coût : $z(x_N) = \bar{z} + \bar{c}_N^T x_N = \bar{z} + \sum\limits_{j \in N} \bar{c}_j x_j \Rightarrow \dfrac{\partial z}{\partial x_j} = \bar{c}_j$

Optimalité

- Coût réduit $\bar{c}_j$ = dérivée directionnelle suivant la direction de base d_j associée à x_j, $j \in N$
- Si tous les coûts réduits sont positifs ou nuls, la solution est optimale.
- Sinon le coût décroît suivant une direction de base d_j de coût réduit négatif
= direction de descente

4.2 Déplacement

a) Règles de déplacement

Notations

- Matrices : B, N = matrice de base et hors base AE = (B N)
- Par extension : B, N = numéros des variables de base x_B et hors base x_N

- Solution de base associée à B : $E^T x = \begin{pmatrix} x_B \\ x_N \end{pmatrix}$ avec $\begin{cases} Ax = b \\ x_N = 0 \end{cases} \Rightarrow Bx_B = b$

Direction de déplacement

- Si tous les coûts réduits sont positifs ou nuls, la solution courante est optimale.
- Sinon on choisit un indice hors base $e \in N$ de coût réduit strictement négatif : $\boxed{\bar{c}_e < 0}$
→ La direction de base associée d_e est une direction de descente.

- On se déplace à partir de x d'un pas $\alpha \geq 0$ suivant la direction de base d_e → $x' = x + \alpha d_e$
- Le nouveau point $x' = x + \alpha d_e$ doit rester admissible :

$$\begin{cases} Ax' = b \\ x' \geq 0 \end{cases} \Leftrightarrow \begin{cases} A(x + \alpha d_e) = b \\ x + \alpha d_e \geq 0 \end{cases} \Rightarrow \begin{cases} \alpha A d_e = 0 \quad \text{car} \quad Ax = b & \rightarrow \text{vérifié par définition de } d_e \\ x_N + \alpha d_{eN} \geq 0 & \rightarrow \text{vérifié car } x_N = 0 \text{ et } d_{eN} \geq 0 \\ x_B + \alpha d_{eB} \geq 0 & \rightarrow \textbf{limite le déplacement} \end{cases}$$

- Le déplacement est limité par le fait que les variables de base doivent rester positives.

Pas de déplacement

Le déplacement α suivant la direction de base d_e est limité par les contraintes $x_B \geq 0$

$x_B + \alpha d_{eB} \geq 0 \quad \rightarrow$ Borne α_i pour chacune des m variables de base x_i, $i \in B$

- Si la composante d_{eB_i} est positive, le pas n'est pas borné : $\alpha_i = +\infty$
- Si la composante d_{eB_i} est négative, le pas est borné par : $\alpha_i = \dfrac{x_i}{-d_{eB_i}}$ (annulation de x_i)

Déplacement maximal

On note s le numéro de la 1ère variable de base x_i qui s'annule suivant la direction d_e.

$\rightarrow$ Le **pas maximal admissible** suivant la direction d_e est : $\alpha_s = \min\limits_{i \in B} \alpha_i$

- Si $d_{eB} \geq 0$, le pas n'est pas borné suivant d_e
  $\rightarrow$ Le problème PL n'a pas de solution (problème non borné).

- Sinon on réalise le pas maximal α_s suivant la direction d_e : $x' = x + \alpha_s d_e$
  $\rightarrow$ **Changement de base ou pivotage**
  $\rightarrow$ Echange des variables x_s (variable de base sortant de la base courante)
  et x_e (variable hors base entrant dans la nouvelle base)

b) Changement de base

Pivotage

La direction de base d_e associé à la variable hors base x_e est définie par

$$E^T d_e = \begin{pmatrix} d_{eB} \\ d_{eN} \end{pmatrix} \text{ avec } \begin{cases} d_{eB} = -B^{-1}A_e & \rightarrow \text{ pour vérifier } Ax=b \\ E^T e_e = \begin{pmatrix} 0 \\ d_{eN} \end{pmatrix} & \rightarrow \text{ composantes nulles sauf =1 sur la composante e} \end{cases}$$

- Le nouveau point $x' = x + \alpha_s d_e$ est admissible car :
  - la direction d_e est admissible
  - le pas α_s respecte les contraintes $x' \geq 0$

- Variables de base x_i, $i \in B$: $x_i' = x_i + \alpha_s (d_{eB})_i \begin{cases} \geq 0 \text{ si } i \neq s & \rightarrow \text{ car } \alpha_s \leq \alpha \\ = 0 \text{ si } i = s & \rightarrow \text{ par construction du pas } \alpha_s \end{cases}$

- Variables hors base x_j, $j \in N$: $x_j' = x_j + \alpha_s (d_{eB})_j \begin{cases} = 0 & \text{si } j \neq e \\ = \alpha_s \geq 0 & \text{si } j = e \end{cases}$ car $(d_{eB})_j \begin{cases} = 0 \text{ si } j \neq e \\ = 1 \text{ si } j = e \end{cases}$

Nouvelle base

- Nouvelles variables hors base : $\begin{cases} x_j' = 0 \text{ pour } j \in N - \{e\} \\ x_s' = 0 \text{ pour } i = s \end{cases} \Rightarrow N' = N - \{e\} + \{s\}$

- Nouvelles variables de base : $\begin{cases} x_i' \geq 0 \text{ pour } i \in B - \{s\} \\ x_e' \geq 0 \text{ pour } j = e \end{cases} \Rightarrow B' = B - \{s\} + \{e\}$

Variation du coût

- Le nouveau coût est : $c^T x' = c^T(x + \alpha_s d_e) = c^T x + \alpha_s \bar{c}_e \quad \Rightarrow \quad z' = z + \alpha_s \bar{c}_e \le z$

- Si la base n'est pas dégénérée ($x_B > 0$), toutes les directions de base sont admissibles
  → Déplacement non nul possible : $\alpha_s > 0$
  → Le coût décroît strictement : $z' < z$ car on a choisi $e \in N$ tel que $\bar{c}_e < 0$

Méthode pratique

- La nouvelle base ne diffère de la base courante que par une seule variable (= une colonne de A)
  → Limitations des calculs correspondant à un pivotage
  → **Méthode des tableaux**

- Les variables hors base sont constantes ou croissantes suivant les directions de base.
  → Toutes les variables hors base sont candidates pour entrer dans la base.

- Plusieurs règles de choix sont possibles pour la variable entrant dans la base.
  → **Règles de pivotage**

- L'algorithme nécessite une base initiale admissible.
  → Etape préliminaire de détermination de la **base initiale**

Règles de pivotage

- Choix de la variable entrante → différents choix possibles
- Détermination de la variable sortante → imposé

Variable entrante

- La variable hors base entrant dans la base doit avoir un **coût réduit négatif**.

- Choix de la variable de plus petit indice
  → **Règle de Bland** (évite le cyclage pouvant se produire lorsque une base est dégénérée)

- Choix de la variable de coût réduit le plus négatif (plus forte descente)
  → **1ère règle de Dantzig**

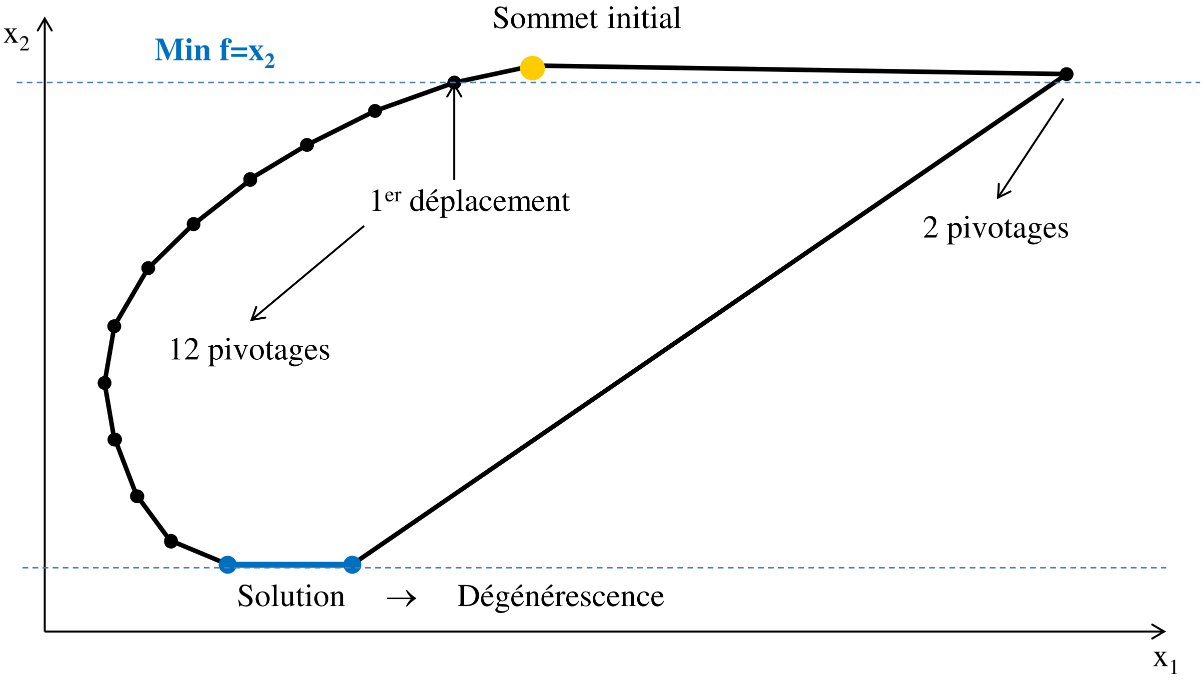
- Choix de la variable conduisant à la plus forte diminution de la fonction coût

- Choix aléatoire avec une probabilité proportionnelle au coût réduit

Variable sortante

- La variable de base sortant de la base est la 1ère à s'annuler suivant la direction de base choisie
  → **2ème règle de Dantzig**

Illustration



Forme canonique

On écrit le problème sous forme canonique dans la base B.

- Formulation matricielle

$$\min_{\substack{x_N \in R^{n-m} \\ x_N \geq 0}} \bar{z} + \bar{c}_N^T x_N \quad \text{sous} \quad x_B = \bar{b} - B^{-1} N x_N \geq 0 \qquad \text{avec} \begin{cases} \bar{b} = B^{-1} b \\ \bar{z} = c_B^T \bar{b} \\ \bar{c}_N^T = c_N^T - c_B^T B^{-1} N \end{cases}$$

- Formulation explicite en fonction des variables hors base x_j , $j \in N$

$$\boxed{\min_{x_N \geq 0} \bar{z} + \sum_{j \in N} \bar{c}_j x_j \quad \text{sous} \quad x_i = \bar{b}_i - \sum_{j \in N} \bar{a}_{ij} x_j \geq 0,\ i \in B \quad \text{avec} \begin{cases} \bar{b} = B^{-1} b \\ B^{-1} N \underset{\text{noté}}{=} (\bar{a}_{ij})_{i \in B, j \in N} \\ \bar{z} = \sum_{i \in B} c_i \bar{b}_i \\ \bar{c}_j = c_j - \sum_{i \in B} c_i \bar{a}_{ij} \ ,\ j \in N \end{cases}}$$

c) Formule de pivotage

Changement de base

- Le pivotage consiste à remplacer la variable hors base x_e (entrante) : $B' = B - \{s\} + \{e\}$
  par la variable de base x_s (sortante) : $N' = N - \{e\} + \{s\}$

- Forme canonique dans la **base B** : $z = \bar{z} + \sum_{j \in N} \bar{c}_j x_j \qquad x_i = \bar{b}_i - \sum_{j \in N} \bar{a}_{ij} x_j \ , \ i \in B$

  $\rightarrow$ Expression en fonction des variables x_j, $j \in N$

- Forme canonique dans la **base B'** : $z = \bar{z}' + \sum_{j \in N'} \bar{c}_j' x_j \qquad x_i = \bar{b}_i' - \sum_{j \in N'} \bar{a}_{ij}' x_j \ , \ i \in B'$

  $\rightarrow$ Expression en fonction des variables x_j, $j \in$ **N'=N−{e}+{s}**

Pour passer de la forme canonique dans la base B à la forme canonique dans la base B', il faut :
- exprimer x_e en fonction de x_s,
- remplacer x_e dans les expressions du coût z et des variables de base x_i, $i \in$ **B'=B−{s}+{e}**

On obtient les formules de pivotage.

Expression de x_e en fonction de x_s

- x_s est dans l'ancienne base B : $\quad i=s \in B \quad \rightarrow \quad x_s = \bar{b}_s - \sum_{j \in N} \bar{a}_{sj} x_j$

$$x_s = \bar{b}_s - \sum_{j \in N} \bar{a}_{sj} x_j = \bar{b}_s - \bar{a}_{se} x_e - \sum_{j \in N - \{e\}} \bar{a}_{sj} x_j \quad \Rightarrow \quad \boxed{x_e = \frac{\bar{b}_s}{\bar{a}_{se}} - \frac{1}{\bar{a}_{se}} x_s - \sum_{j \in N - \{e\}} \frac{\bar{a}_{sj}}{\bar{a}_{se}} x_j}$$

- x_e est dans la nouvelle base B' : $\quad i=e \in B' \quad \rightarrow \quad \boxed{x_e = \bar{b}_e' - \sum_{j \in N'} \bar{a}_{ej}' x_j}$

En identifiant les coefficients :
$$\begin{cases} \bar{b}_e' = \dfrac{\bar{b}_s}{\bar{a}_{se}} \\ \bar{a}_{es}' = \dfrac{1}{\bar{a}_{se}} & \rightarrow j = s \\ \bar{a}_{ej}' = \dfrac{\bar{a}_{sj}}{\bar{a}_{se}} & \rightarrow j \in N - \{e\} \end{cases}$$

On exprime ensuite les autres variables de base x_i , $i \in B$-{s} en remplaçant x_e.

Expression des autres variables de base

- x_i est dans l'ancienne base B : $i \in B\text{-}\{s\} \;\rightarrow\; x_i = \bar{b}_i - \sum_{j \in N} \bar{a}_{ij} x_j$

$$x_i = \bar{b}_i - \bar{a}_{ie} x_e - \sum_{j \in N-\{e\}} \bar{a}_{ij} x_j = \bar{b}_i - \bar{a}_{ie} \left(\frac{1}{\bar{a}_{se}} \bar{b}_s - \frac{1}{\bar{a}_{se}} x_s - \sum_{j \in N-\{e\}} \frac{\bar{a}_{sj}}{\bar{a}_{se}} x_j \right) - \sum_{j \in N-\{e\}} \bar{a}_{ij} x_j$$

$$\Rightarrow \boxed{x_i = \bar{b}_i - \frac{\bar{a}_{ie}}{\bar{a}_{se}} \bar{b}_s + \frac{\bar{a}_{ie}}{\bar{a}_{se}} \bar{b}_s x_s - \sum_{j \in N-\{e\}} \left(\bar{a}_{ij} - \frac{\bar{a}_{ie}}{\bar{a}_{se}} \bar{a}_{sj} \right) x_j}$$

- x_i reste dans la nouvelle base B' : $i \in B'$ $\rightarrow$ $\boxed{x_i = \bar{b}_i' - \sum_{j \in N'} \bar{a}_{ij}' x_j}$

En identifiant les coefficients : $\begin{cases} \bar{b}_i' = \bar{b}_i - \dfrac{\bar{a}_{ie}}{\bar{a}_{se}} \bar{b}_s \\ \bar{a}_{is}' = -\dfrac{\bar{a}_{ie}}{\bar{a}_{se}} & \rightarrow j = s \\ \bar{a}_{ij}' = \bar{a}_{ij} - \dfrac{\bar{a}_{ie}}{\bar{a}_{se}} \bar{a}_{sj} & \rightarrow j \in N - \{e\} \end{cases}$

Expression du coût

- Dans l'ancienne base B : $z = \bar{z} + \sum_{j \in N} \bar{c}_j x_j$

$$z = \bar{z} + \bar{c}_e x_e + \sum_{j \in N-\{e\}} \bar{c}_j x_j = \bar{z} + \bar{c}_e \left(\frac{1}{\bar{a}_{se}} \bar{b}_s - \frac{1}{\bar{a}_{se}} x_s - \sum_{j \in N-\{e\}} \frac{\bar{a}_{sj}}{\bar{a}_{se}} x_j \right) + \sum_{j \in N-\{e\}} \bar{c}_j x_j$$

$$\Rightarrow \boxed{z = \bar{z} + \bar{c}_e \frac{\bar{b}_s}{\bar{a}_{se}} - \frac{\bar{c}_e}{\bar{a}_{se}} x_s + \sum_{j \in N-\{e\}} \left(\bar{c}_j - \bar{c}_e \frac{\bar{a}_{sj}}{\bar{a}_{se}} \right) x_j}$$

- Dans la nouvelle base B' : $\boxed{z = \bar{z}' + \sum_{j \in N'} \bar{c}_j' x_j}$

En identifiant les coefficients : $\begin{cases} \bar{z}' = \bar{z} + \bar{c}_e \dfrac{\bar{b}_s}{\bar{a}_{se}} \\ \bar{c}_s' = -\dfrac{\bar{c}_e}{\bar{a}_{se}} & \rightarrow j = s \\ \bar{c}_j' = \bar{c}_j - \bar{c}_e \dfrac{\bar{a}_{sj}}{\bar{a}_{se}} & \rightarrow j \in N - \{e\} \end{cases}$

Récapitulatif

- Nouvelles variables de base : $i \in B' = B - \{s\} + \{e\}$

$$i = e \rightarrow \begin{cases} \bar{b}_e' = \dfrac{\bar{b}_s}{\bar{a}_{se}} & \\ \bar{a}_{es}' = \dfrac{1}{\bar{a}_{se}} & \rightarrow j = s \\ \bar{a}_{ej}' = \dfrac{\bar{a}_{sj}}{\bar{a}_{se}} & \rightarrow j \in N - \{e\} \end{cases} \qquad i \in B - \{s\} \rightarrow \begin{cases} \bar{b}_i' = \bar{b}_i - \dfrac{\bar{a}_{ie}}{\bar{a}_{se}} \bar{b}_s & \\ \bar{a}_{is}' = -\dfrac{\bar{a}_{ie}}{\bar{a}_{se}} & \rightarrow j = s \\ \bar{a}_{ij}' = \bar{a}_{ij} - \dfrac{\bar{a}_{ie}}{\bar{a}_{se}} \bar{a}_{sj} & \rightarrow j \in N - \{e\} \end{cases}$$

- Nouveau coût

$$\begin{cases} \bar{z}' = \bar{z} + \bar{c}_e \dfrac{\bar{b}_s}{\bar{a}_{se}} & \\ \bar{c}_s' = -\dfrac{\bar{c}_e}{\bar{a}_{se}} & \rightarrow j = s \\ \bar{c}_j' = \bar{c}_j - \bar{c}_e \dfrac{\bar{a}_{sj}}{\bar{a}_{se}} & \rightarrow j \in N - \{e\} \end{cases}$$

On dispose dans un tableau les éléments nécessaires au pivotage $\rightarrow$ **tableau du simplexe**.

d) Méthode des tableaux

Tableau du simplexe

- On écrit le problème sous forme canonique dans la base B.

$$\min_{x_N \geq 0} z = \bar{z} + \bar{c}_N^T x_N \quad \text{sous} \quad x_B = \bar{b} - B^{-1} N x_N \geq 0 \qquad \text{avec} \begin{cases} \bar{b} = B^{-1} b \\ \bar{z} = c_B^T \bar{b} \\ \bar{c}_N^T = c_N^T - c_B^T B^{-1} N \end{cases}$$

- La solution de base associée à B est : $\begin{cases} x_B = \bar{b} \\ x_N = 0 \\ z = \bar{z} \end{cases}$

- Le **tableau du simplexe** est : $T = \begin{array}{|c|c|} \hline B^{-1}A & \bar{b} \\ \hline \bar{c}^T & -\bar{z} \\ \hline \end{array} = \begin{array}{cccccc} x_1 & & x_j & & x_n & \\ \hline B^{-1}A_1 & \cdots & B^{-1}A_j & \cdots & B^{-1}A_n & x_B \\ \hline \bar{c}_1 & \cdots & \bar{c}_j & \cdots & \bar{c}_n & -c_B^T x_B \\ \hline \end{array}$

  $A_j = j^{\text{ème}}$ colonne de A
  $AE = (B \quad N)$ avec E = matrice de permutation de colonnes $\Rightarrow B^{-1}A = (I \quad B^{-1}N)\, E^T$

- En permutant les colonnes : $T = \begin{array}{cc|c} x_B & x_N & \\ \hline I & B^{-1}N & \bar{b} \\ \hline 0 & \bar{c}_N^T & -\bar{z} \\ \hline \end{array} \begin{array}{l} \\ \rightarrow x_B + B^{-1} N x_N = \bar{b} \\ \rightarrow \bar{c}_N^T x_N = z - \bar{z} \end{array}$

Description du tableau

Le tableau du simplexe est noté T(i,j) : i=1 à m+1 , j=1 à n+1

- T(1:m,1:n) : **Matrice $B^{-1}A$** $\rightarrow m \times n$

$$B^{-1}A = \begin{pmatrix} I & B^{-1}N \end{pmatrix} E^T \quad \text{en plaçant les variables de base en premier}$$

- T(m+1,1:n) : **Coûts réduits** $\rightarrow 1 \times n$

$$\bar{c}^T = c - c_B^T B^{-1}A \Rightarrow \begin{cases} \bar{c}_B^T = 0 \\ \bar{c}_N^T = c_N^T - c_B^T B^{-1}N \end{cases}$$

- T(1:m,n+1) : **Solution de base** $\rightarrow m \times 1$

$$x = \begin{pmatrix} x_B \\ x_N \end{pmatrix} = \begin{pmatrix} \bar{b} \\ 0 \end{pmatrix}$$

- T(m+1,n+1) : **Opposé du coût** $\rightarrow 1 \times 1$

$$-z = -\bar{z} = -c_B^T \bar{b}$$

$$\begin{array}{ccccc|c}
x_1 & \cdots & x_j & \cdots & x_n & \\
\hline
\vdots & & \vdots & & \vdots & \bar{b}_1 \\
B^{-1}A_1 & \cdots & B^{-1}A_j & \cdots & B^{-1}A_n & \vdots \\
\vdots & & \vdots & & \vdots & \bar{b}_m \\
\hline
\bar{c}_1 & \cdots & \bar{c}_j & \cdots & \bar{c}_n & -\bar{z}
\end{array}$$

$$T = \begin{array}{cc|c}
x_B & x_N & \\
\hline
I & B^{-1}N & \bar{b} \\
\hline
0 & \bar{c}_N^T & -\bar{z}
\end{array}
\quad
\begin{array}{l}
\\
\rightarrow x_B + B^{-1}Nx_N = \bar{b} \\
\rightarrow \bar{c}_N^T x_N = z - \bar{z}
\end{array}$$

Utilisation du tableau

Le tableau du simplexe permet de :

- Repérer les **variables de base** $\rightarrow$ colonnes = matrice identité, coûts réduits nuls
- Vérifier si la base est **admissible** $\rightarrow$ valeurs positives ou nulles des variables de base
- Vérifier si la base est **optimale** $\rightarrow$ valeurs strictement positives des coûts réduits
- Sélectionner un **pivotage** pour passer à une base adjacente meilleure
- Mettre à jour la forme canonique dans la **nouvelle base**

Méthode de pivotage

- On choisit une variable hors base de coût réduit négatif $\rightarrow$ **colonne e**
- On examine la variation des variables de base suivant la direction d_e

$$x_i = \bar{b}_i - \bar{a}_{ie} x_e , i \in B \qquad \text{s'annule pour :} \quad x_e = \frac{\bar{b}_i}{\bar{a}_{ie}}$$

- La première variable de base à s'annuler sort de la base $\rightarrow$ **ligne s**

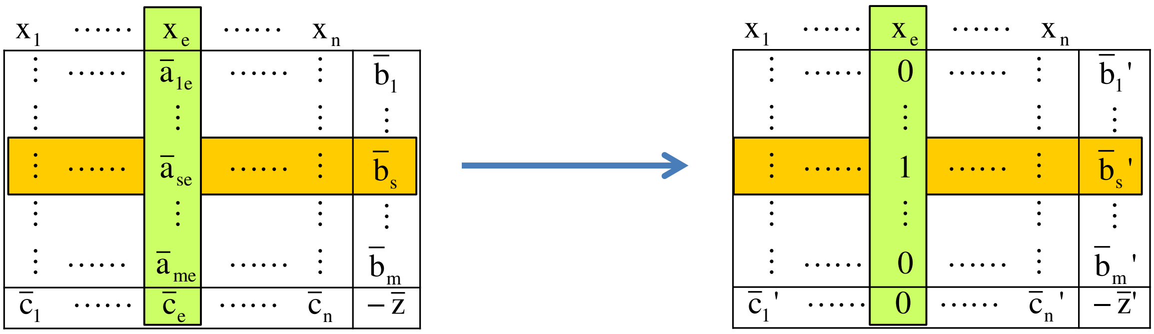
- Le **pivotage e-s** consiste à faire apparaître une colonne de la matrice identité en colonne e.
  - $\rightarrow$ forme canonique dans la nouvelle base
  - $\rightarrow$ par combinaison linéaire des lignes du tableau

Réalisation du pivotage

La variable x_e entre dans la nouvelle base B'=B−{s}+{e}

Pour faire apparaître une colonne de la matrice identité en colonne e,
on réalise des **combinaisons linéaires des lignes du tableau**, y compris la dernière colonne.

- Division de la ligne s par le **pivot** $\bar{a}_{se} \rightarrow \bar{a}_{se}' = 1$
- Addition de la ligne s aux autres lignes pour annuler les coefficients dans la colonne e
- Annulation du coût réduit dans la colonne e



Algorithme de pivotage

1. **Choix du pivot**

- Variable hors base entrante x_e = 1er coût réduit négatif
- Pas maximal admissible pour chaque variable de base : $\alpha_i = \dfrac{T(i,n+1)}{T(i,e)},\ i \in B,\ \text{si } T(i,e) > 0$
- Variable de base sortante x_s : $\alpha_s = \min\limits_{\substack{i \in B \\ T(i,e)>0}} \alpha_i$

2. **Réalisation du pivotage**

- Pivot = T(s,e)
- Lignes i=1,…,m+1, i≠s $\quad T(i,k) = T(i,k) - \dfrac{T(i,e)}{T(s,e)} T(s,k),\ k = 1,\cdots,n+1$
- Ligne s du pivot $\quad T(s,k) = \dfrac{T(s,k)}{T(s,e)},\ k = 1,\cdots,n+1$

→ méthode similaire à la méthode du pivot de Gauss

Méthode des tableaux

- Problème linéaire à 3 variables x_1, x_2, x_3

$$\min_{x_1,x_2,x_3} -10x_1-12x_2-12x_3 \text{ sous } \begin{cases} x_1+2x_2+2x_3 \le 20 \\ 2x_1+\ x_2+2x_3 \le 20 \\ 2x_1+2x_2+\ x_3 \le 20 \\ x_1,x_2,x_3 \ge 0 \end{cases}$$

- **Forme standard**

$\rightarrow$ **Variables d'écart** x_4, x_5, x_6 positives

$$\min_{x_1,x_2,x_3,x_4,x_5,x_6} -10x_1-12x_2-12x_3 \text{ sous } \begin{cases} x_1+2x_2+2x_3+x_4 = 20 \\ 2x_1+\ x_2+2x_3+x_5 = 20 \\ 2x_1+2x_2+\ x_3+x_6 = 20 \\ x_1,x_2,x_3,x_4,x_5,x_6 \ge 0 \end{cases}$$

Méthode des tableaux

- **Tableau du simplexe**

x_1	x_2	x_3	x_4	x_5	x_6		
1	2	2	1	0	0	20	x_4
2	1	2	0	1	0	20	x_5
2	2	1	0	0	1	20	x_6
-10	-12	-12	0	0	0	0	-z

Base initiale admissible
(x_4 , x_5 , x_6)

- Solution de base **non optimale** : coûts réduits négatifs (= directions de descente)
- Variable entrante : 1er coût réduit négatif $\rightarrow$ x_1
- Variable sortante : 1ère variable de base à s'annuler $\rightarrow$ x_5
- Pivot : $\bar{a}_{51} = 2$

Méthode des tableaux

- **1^er pivotage** : entrée x_1, sortie x_5

x_1	x_2	x_3	x_4	x_5	x_6			Pas
1	2	2	1	0	0	20	x_4	$\rightarrow s_{14}=20$
2	1	2	0	1	0	20	x_5	$\rightarrow s_{15}=10$
2	2	1	0	0	1	20	x_6	$\rightarrow s_{16}=10$
-10	-12	-12	0	0	0	0	z	

x_1	x_2	x_3	x_4	x_5	x_6		
0	1.5	1	1	-0.5	0	10	x_4
1	0.5	1	0	0.5	0	10	x_1
0	1	-1	0	-1	1	0	x_6
0	-7	-2	0	5	0	100	z

Nouvelle base
(x_1 , x_4 , x_6)

Méthode des tableaux

- **2^ème pivotage** : entrée x_2, sortie x_6

x_1	x_2	x_3	x_4	x_5	x_6			Pas
0	1.5	1	1	-0.5	0	10	x_4	$\rightarrow s_{24}=20/3$
1	0.5	1	0	0.5	0	10	x_1	$\rightarrow s_{21}=20$
0	1	-1	0	-1	1	0	x_6	$\rightarrow s_{26}=0$
0	-7	-2	0	5	0	100	z	

x_1	x_2	x_3	x_4	x_5	x_6		
0	0	2.5	1	1	-1.5	10	x_4
1	0	1.5	0	1	-0.5	10	x_1
0	1	-1	0	-1	1	0	x_2
0	0	-9	0	-2	7	100	z

Nouvelle base
(x_1 , x_2 , x_4)

Méthode des tableaux

- **$3^{ème}$ pivotage** : entrée x_3, sortie x_4

x_1	x_2	x_3	x_4	x_5	x_6			Pas
0	0	2.5	1	1	-1.5	10	x_4	$\rightarrow s_{34}=4$
1	0	1.5	0	1	-0.5	10	x_1	$\rightarrow s_{31}=20/3$
0	1	-1	0	-1	1	0	x_2	$\rightarrow s_{32}=+\infty$
0	0	-9	0	-2	7	100	z	

x_1	x_2	x_3	x_4	x_5	x_6			
0	0	1	0.4	0.4	-0.6	4	x_3	Nouvelle base
1	0	0	-0.6	0.4	0.4	4	x_1	$(x_1\ ,x_2\ ,x_3)$
0	1	0	0.4	-0.6	0.4	4	x_2	
0	0	0	3.6	1.6	1.6	136	z	

- **Solution optimale** : $\bar{c} \geq 0 \quad \rightarrow \begin{cases} x^* = \begin{pmatrix} 4 & 4 & 4 & 0 & 0 & 0 \end{pmatrix}^T \\ z^* = -136 \end{cases}$

Méthode des tableaux

Récapitulatif des itérations

k	B	c			x_1	x_2	x_3	x_4	x_5	x_6	d_B			s_{max}	e	s	z
0	4 5 6	-10	-12	-12	0	0	0	20	20	20	-1	-2	-2	10	1	5	0
1	1 4 6	-7	-2	5	10	0	0	10	0	0	-1.5	-0.5	-1	0	2	6	-100
2	1 2 4	-9	-2	7	10	0	0	10	0	0	-2.5	-1.5	1	4	3	4	-100
3	1 2 3	3.6	31.6	1.6	4	4	4	0	0	0							-136

Commentaires

- La mise sous forme standard nécessite d'introduire des variables supplémentaires.
  - $\rightarrow$ Variables d'écart positives
- On dispose directement d'une base initiale admissible formée des variables d'écart.
  - $\rightarrow$ Ce n'est pas toujours le cas.
  - $\rightarrow$ Phase préliminaire pour construire une base initiale
- Certains pivotages ne réduisent pas le coût (exemple : pivotage numéro 2).
  - $\rightarrow$ Base dégénérée + risque de cyclage (= retrouver une base précédente)
- La solution optimale ne comporte que les variables initiales.
  - $\rightarrow$ Ce n'est pas toujours le cas.
  - $\rightarrow$ Des pivotages supplémentaires peuvent être nécessaires.

4.3 Simplexe dual

a) Problème dual

Méthode du simplexe dual

L'algorithme du simplexe dual consiste à appliquer la méthode du simplexe **au problème dual**.

- **Correspondances primal (P) – dual (D)** :

Le tableau est utilisable dans les 2 sens :
de (P) vers (D)
ou de (D) vers (P) car le dual de (D) est (P).

Primal (P)		Dual (D)
$\min\limits_{x\in R^n} c^T x$	1	$\max\limits_{y} b^T y$
$Ax = b$	m	$y \in R$
$Ax \le b$	m	$y \ge 0$
$x \ge 0$	n	$A^T y \le c$
$x \in R$	n	$A^T y = c$

- Forme canonique de (P) dans la base B

$$\min_{x_N \ge 0} z = \bar{z} + \bar{c}_N^T x_N \quad \text{sous} \quad \begin{cases} x_B + B^{-1} N x_N = \bar{b} \\ x \ge 0 \end{cases} \quad \text{avec} \quad \begin{cases} \bar{b} = B^{-1} b \\ \bar{z} = c_B^T \bar{b} \\ \bar{c}_N^T = c_N^T - c_B^T B^{-1} N \end{cases}$$

Pour appliquer la méthode du simplexe au problème dual, on doit écrire la forme canonique du problème dual dans la base B.

511

b) Forme canonique

Forme canonique du dual

- On part de la forme canonique du problème primal (P) dans la base B.

$$(P) \quad \min_{x_B, x_N} \bar{z} + \bar{c}_N^T x_N \quad \text{sous} \quad \begin{cases} x_B + B^{-1} N x_N = \bar{b} & \rightarrow \text{ m contraintes} \\ x_B, x_N \ge 0 & \rightarrow \text{ n variables} \end{cases}$$

- On peut considérer les variables de base x_B comme des variables d'écart positives.
On obtient un problème (P') ne portant que sur les variables hors base x_N.

$$(P') \quad \min_{x_N} \bar{c}_N^T x_N \quad \text{sous} \quad \begin{cases} B^{-1} N x_N \le \bar{b} & \rightarrow \text{ m contraintes} \\ x_N \ge 0 & \rightarrow \text{ n–m variables} \end{cases}$$

- On écrit (P') comme un problème de maximisation,
pour obtenir un problème de minimisation en passant au dual.

$$(P') \quad \max_{x_N} -\bar{c}_N^T x_N \quad \text{sous} \quad \begin{cases} B^{-1} N x_N \le \bar{b} & \rightarrow \text{ m contraintes} \\ x_N \ge 0 & \rightarrow \text{ n–m variables} \end{cases}$$

- On passe au dual (D') de (P')
$\rightarrow$ en utilisant le tableau de correspondances dans le sens de (D) vers (P)

Forme canonique du dual

- Le dual (D') de (P') s'écrit :

$$(P') \quad \max_{x_N} -\bar{c}_N^T x_N \quad \text{sous} \quad \begin{cases} B^{-1} N x_N \le \bar{b} \\ x_N \ge 0 \end{cases} \quad \begin{aligned} &\rightarrow \text{ m contraintes} \\ &\rightarrow \text{ n−m variables} \end{aligned}$$

$$(D') \quad \min_{y_B} \bar{b}^T y_B \quad \text{sous} \quad \begin{cases} \left(B^{-1} N\right)^T y_B \le -\bar{c}_N \\ y_B \ge 0 \end{cases} \quad \begin{aligned} &\rightarrow \text{ n−m contraintes} \\ &\rightarrow \text{ m variables} \end{aligned}$$

- On met (D') sous forme standard avec des variables d'écart y_N positives.
  On obtient un problème (D) à n variables.

$$\boxed{(D) \quad \min_{y_B, y_N} \bar{b}^T y_B \quad \text{sous} \quad \begin{cases} y_N - \left(B^{-1} N\right)^T y_B = \bar{c}_N \\ y_B, y_N \ge 0 \end{cases} \quad \begin{aligned} &\rightarrow \text{ n−m contraintes} \\ &\rightarrow \text{ n variables} \end{aligned}}$$

- Le problème (D) est sous **forme canonique dans la base B** :
  - variables de base → y_N
  - variables hors base → y_B

  → notations inversées par rapport au problème primal

  On peut écrire le tableau simplexe pour le problème (D) et appliquer les règles de pivotage.

c) Tableau dual

Tableau simplexe du dual

- Forme canonique de (D) dans la base B.

$$(D) \quad \min_{y_B, y_N} \bar{b}^T y_B \quad \text{sous} \quad \begin{cases} y_N - \left(B^{-1} N\right)^T y_B = \bar{c}_N \\ y_B, y_N \ge 0 \end{cases} \quad \begin{aligned} &\rightarrow \text{ n−m contraintes} \\ &\rightarrow \text{ n variables} \end{aligned}$$

- **Tableau T_D du simplexe de (D) dans la base B** :

$$T_D = \begin{array}{c} \begin{array}{cc} y_N & \quad y_B \end{array} \\ \begin{array}{|cc|c|} \hline I & -(B^{-1}N)^T & \bar{c}_N^T \\ \hline 0 & \bar{b} & -\bar{z} \\ \hline \end{array} \end{array}$$

  Variables de base : y_N → valeurs $\bar{c}_N$
  Variables hors base : y_B → coûts réduits $\bar{b}$
  Matrice des contraintes : $-A^T$ → $-(B^{-1}N)^T$

- La solution de base associée à la base B est : $\begin{cases} y_N = \bar{c}_N \\ y_B = 0 \end{cases}$

- La base B est admissible si $y_N = \bar{c}_N \ge 0$ → **base dual-admissible**

- On applique les règles de pivotage du simplexe :
  - variable hors base entrante : coût réduit négatif
  - variable de base sortante : première variable à s'annuler

d) Pivotage

Pivotage sur le tableau dual

Les notations sont inversées par rapport au problème primal
- indices B → variables hors base
- indices N → variables de base

$$T_D = \begin{array}{c} \quad y_N \qquad\quad y_B \qquad\qquad \\ \begin{array}{|cc|c|} \hline I & -(B^{-1}N)^T & \bar{c}_N^T \\ \hline 0 & \bar{b} & -\bar{z} \\ \hline \end{array} \end{array}$$

1. **Choix du pivot**
- Variable hors base entrante y_e = 1[er] coût réduit négatif $\boxed{\bar{b}_e < 0\,,\, e \in B}$
- Pas maximal admissible pour chaque variable de base : $\alpha_i = \dfrac{\bar{c}_{Ni}}{-\bar{a}_{ie}}$, $i \in N$, si $-\bar{a}_{ie} > 0$
- Variable de base sortante y_s : $\alpha_s = \min\limits_{\substack{i \in N \\ \bar{a}_{ie} < 0}} \alpha_i$

Ligne s de la variable sortante : $\boxed{s \in N \;\rightarrow\; \min\limits_{\substack{i \in N \\ \bar{a}_{ie} < 0}} \dfrac{\bar{c}_{Ni}}{-\bar{a}_{ie}} \;\Leftrightarrow\; \max\limits_{\substack{i \in N \\ \bar{a}_{ie} < 0}} \dfrac{\bar{c}_{Ni}}{\bar{a}_{ie}}}$

2. **Réalisation du pivotage**
- Pivot = $\bar{a}_{se} < 0$
- Elimination pour faire apparaître des zéros sur la colonne e du pivot

Pivotage sur le tableau primal
- On observe que le pivotage dual peut être réalisé à partir du tableau primal **sans écrire explicitement le tableau dual**.

$$T_D = \begin{array}{c} \quad y_N \qquad\quad y_B \qquad\qquad \\ \begin{array}{|cc|c|} \hline I & -(B^{-1}N)^T & \bar{c}_N^T \\ \hline 0 & \bar{b} & -\bar{z} \\ \hline \end{array} \end{array}$$

→ n−m contraintes
→ n−m variables de base y_N

$$T_P = \begin{array}{c} \quad x_B \qquad\quad x_N \qquad\qquad \\ \begin{array}{|cc|c|} \hline I & B^{-1}N & \bar{b} \\ \hline 0 & \bar{c}_N^T & -\bar{z} \\ \hline \end{array} \end{array}$$

→ m contraintes
→ m variables de base x_B

- Choisir la 1[ère] variable de base négative x_e : $\boxed{\bar{b}_e < 0\,,\, e \in B}$ → ligne e → variable **sortante**
- Déterminer la 1[ère] variable hors base x_s à s'annuler : $\boxed{s \;\rightarrow\; \max\limits_{\substack{j \in N \\ \bar{a}_{ej} < 0}} \dfrac{\bar{c}_{Nj}}{\bar{a}_{ej}}}$ → colonne s → variable **entrante**
- Effectuer le pivotage e–s de façon usuelle.

e) Comparaison simplexe primal-dual

Comparaison simplexe primal et dual

- L'algorithme du simplexe primal maintient une base **primal–admissible** : $\bar{b} \geq 0$
  L'optimum est atteint lorsque les coûts réduits sont positifs ou nuls : $\bar{c}_N \geq 0$

- L'algorithme du simplexe dual maintient une base **dual–admissible** : $\bar{c}_N \geq 0$
  L'optimum est atteint lorsque les variables de base sont positives ou nulles : $\bar{b} \geq 0$

Intérêt du simplexe dual

L'algorithme du simplexe dual est adapté si l'on dispose d'une base dual–admissible.

Ceci se produit lorsque l'on modifie un problème linéaire déjà résolu par le simplexe primal

- en ajoutant des contraintes au problème
- en modifiant les seuils des contraintes
- en fixant des variables à une valeur différente de la solution

Ces modifications : - ne changent pas les coûts réduits ($\rightarrow \bar{c}_N \geq 0$)
- rendent certaines variables de base négatives

$\rightarrow$ La solution de base n'est plus primal–admissible, mais reste dual–admissible.

Application : problèmes de programmation linéaire mixte (entiers et réels)

f) Exemple

Simplexe dual

- Problème linéaire à 5 variables x_1,x_2,x_3,x_4,x_5

$$\min_{x_1,x_2,x_3,x_4,x_5} \quad x_1+2x_2+2x_3+3x_4+x_5 \quad \text{sous} \begin{cases} x_1+x_2=1 \\ -x_2-x_3+x_5=0 \\ -x_1+x_3+x_4=0 \\ x_1,x_2,x_3,x_4,x_5 \geq 0 \end{cases}$$

- On choisit comme **base initiale (x_2,x_3,x_4)**.

La solution de base associée est : $\begin{cases} x_1=0 \\ x_5=0 \end{cases} \Rightarrow \begin{cases} x_2=1 \\ x_3=-1 \\ x_4=1 \end{cases}$ $\rightarrow$ base non primal admissible

- La matrice des contraintes est : $A=\begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 & 1 \\ -1 & 0 & 1 & 1 & 0 \end{pmatrix}$ $\quad b=\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$

- Pour construire le tableau du simplexe, il faut mettre le problème sous **forme canonique** dans la base (x_2,x_3,x_4) en faisant apparaître des zéros par élimination dans les colonnes 2, 3 et 4.

Simplexe dual

- Problème linéaire à 5 variables x_1,x_2,x_3,x_4,x_5

$$\min_{x_1,x_2,x_3,x_4,x_5} \quad x_1+2x_2+2x_3+3x_4+x_5 \quad \text{sous} \begin{cases} x_1+x_2=1 \\ -x_2-x_3+x_5=0 \\ -x_1+x_3+x_4=0 \\ x_1,x_2,x_3,x_4,x_5 \geq 0 \end{cases}$$

- Tableau de départ

x_1	x_2	x_3	x_4	x_5	
1	1	0	0	0	1
0	-1	-1	0	1	0
-1	0	1	1	0	0
1	2	2	3	1	0

$\rightarrow$ contraintes

$\rightarrow$ coût

- On fait apparaître : - une matrice identité sur les colonnes de x_2 , x_3 , x_4
  - des zéros sur les coûts de x_2 , x_3 , x_4

Simplexe dual

x_1	x_2	x_3	x_4	x_5	
1	1	0	0	0	1
0	-1	-1	0	1	0
-1	0	1	1	0	0
1	2	2	3	1	0

Elimination x_2 $\longrightarrow$

x_1	x_2	x_3	x_4	x_5	
1	1	0	0	0	1
1	0	-1	0	1	1
-1	0	1	1	0	0
-1	0	2	3	1	-2

x_1	x_2	x_3	x_4	x_5	
1	1	0	0	0	1
1	0	-1	0	1	1
-1	0	1	1	0	0
-1	0	2	3	1	-2

Elimination x_3 $\longrightarrow$

x_1	x_2	x_3	x_4	x_5	
1	1	0	0	0	1
-1	0	1	0	-1	-1
0	0	0	1	1	1
1	0	0	3	3	0

x_1	x_2	x_3	x_4	x_5	
1	1	0	0	0	1
-1	0	1	0	-1	-1
0	0	0	1	1	1
1	0	0	3	3	0

Elimination x_4 $\longrightarrow$

x_1	x_2	x_3	x_4	x_5	
1	1	0	0	0	1
-1	0	1	0	-1	-1
0	0	0	1	1	1
1	0	0	0	0	-3

Simplexe dual

- Tableau du simplexe dans la base (x_2,x_3,x_4).

	x_1	x_2	x_3	x_4	x_5		
c =	1	2	2	3	1		
	1	1	0	0	0	1	x_2
	-1	0	1	0	-1	-1	x_3
	0	0	0	1	1	1	x_4
	1	0	0	0	0	-3	-z

→

c_B	c_N		
I	$B^{-1}N$	$B^{-1}b$	x_B
0	$c_N - c_B{}^T B^{-1}N$	$-c_B{}^T B^{-1}b$	-z

- On vérifie bien que la dernière ligne correspond à $\begin{cases} \bar{c}_N = c_N - c_B^T B^{-1} N \\ -z = -c_B^T B^{-1} b \end{cases}$

- La base est : - non admissible pour le primal ($x_3 < 0$)
  - admissible pour le dual ($\bar{c}_N \geq 0$)

  On peut appliquer l'algorithme dual du simplexe pour résoudre le problème.

Simplexe dual

- Tableau du simplexe dans la base (x_2,x_3,x_4).

x_1	x_2	x_3	x_4	x_5		
1	1	0	0	0	1	x_2
-1	0	1	0	-1	-1	x_3
0	0	0	1	1	1	x_4
1	0	0	0	0	-3	-z

Base dual–admissible (x_2,x_3,x_4)

- Solution de base **non optimale** : variables de base négatives

- Variable **sortante** : 1ère variable de base négative → x_3 $\qquad \bar{b}_e < 0 \,,\, e \in B$

- Variable **entrante** : 1er coût réduit à s'annuler → x_5 $\qquad s \rightarrow \max\limits_{\substack{j \in N \\ \bar{a}_{ej} < 0}} \dfrac{\bar{c}_{Nj}}{\bar{a}_{ej}}$

- Pivot : $\bar{a}_{35} = -1$

Simplexe dual

- **1er pivotage** : entrée x_5, sortie x_3

x_1	x_2	x_3	x_4	x_5		
1	1	0	0	0	1	**x_2**
-1	0	1	0	-1	-1	**x_3**
0	0	0	1	1	1	**x_4**
1	0	0	0	0	-3	-z

x_1	x_2	x_3	x_4	x_5		
1	1	0	0	0	1	**x_2**
1	0	-1	0	1	1	**x_5**
-1	0	1	1	0	0	**x_4**
1	0	0	0	0	-3	-z

Nouvelle base (x_2,x_4,x_5)
- primal–admissible $\bar{b} \geq 0$
- dual–admissible $\bar{c}_N \geq 0$

$\rightarrow$ **optimale**

Solution : $x^* = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \end{pmatrix}$
$z^* = 3$

5 Gradient projeté

5.1 Principes

Problème avec contraintes égalité

$$\min_{x \in R^n} f(x) \text{ sous } \begin{cases} c_E(x) = 0 \\ c_I(x) \leq 0 \end{cases} \Leftrightarrow \min_{x \in R^n} f(x) \text{ sous } c(x) = 0 \rightarrow \text{m contraintes actives}$$

Etapes principales

A chaque itération :
- Construction d'une direction de descente d_k à partir du point x_k
- Réglage du pas de déplacement s_k suivant d_k

Direction de descente

On construit la direction d_k dans l'hyperplan tangent aux contraintes (= espace nul) en x_k
- **Gradient projeté** $\rightarrow$ projection du gradient sur l'hyperplan tangent
- **Gradient réduit** $\rightarrow$ réduction du gradient sur une base de l'espace nul

Pas de déplacement
- Recherche linéaire suivant d_k $\rightarrow$ pas s_k
- Restauration de l'admissibilité $\rightarrow$ méthode de Newton
- Règles d'acceptation du pas $\rightarrow$ Armijo, Goldstein, Wolfe

5.2 Direction de déplacement

a) Hyperplan tangent aux contraintes

Problème avec contraintes linéaires

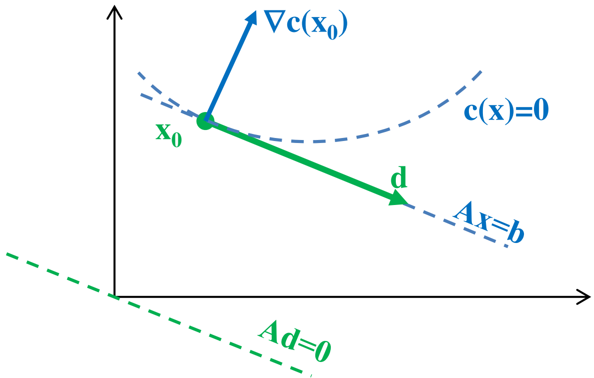
$$\min_{x \in R^n} f(x) \text{ sous } Ax = b$$

- x_0 point admissible $\rightarrow Ax_0 = b$
- d déplacement admissible à partir de x_0 $\rightarrow A(x_0 + d) = b$

$\left.\begin{matrix} Ax_0 = b \\ A(x_0+d) = b \end{matrix}\right\} \rightarrow Ad = 0$

- $Ad = 0$ définit l'espace nul des contraintes
  = hyperplan des contraintes

- Le déplacement $d \in R^n$ est admissible si
  d est dans l'hyperplan des contraintes.



Problème avec contraintes non linéaires

$$\min_{x \in R^n} f(x) \text{ sous } c(x) = 0$$

- On définit l'**espace nul tangent** ou **hyperplan tangent** en x_0 avec $A = \nabla c(x_0)^T$ $\rightarrow Ad = 0$

- On cherche un déplacement $d \in R^n$ dans l'hyperplan tangent : $\nabla c(x_0)^T d = 0$
  Un déplacement complémentaire est ensuite nécessaire pour restaurer l'admissibilité.

b) Gradient projeté

Définition

$$\min_{x \in R^n} f(x) \text{ sous } c(x) = 0$$

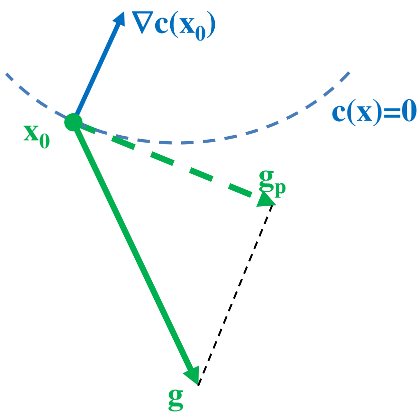
Le gradient projeté est la **projection orthogonale** du gradient de f sur l'hyperplan tangent.

Expression du gradient projeté

- Hyperplan tangent aux contraintes en x_0 admissible
  $Ad = 0$ avec $A = \nabla c(x_0)^T$

- Matrice de projection sur l'hyperplan tangent
  $P = I - A^T\left(AA^T\right)^{-1}A$

- **Notations**
  $g(x_0)$ gradient de f en x_0 $\rightarrow g(x_0) = \nabla f(x_0)$
  $g_p(x_0)$ gradient projeté $\rightarrow g_p(x_0) = Pg(x_0)$
  $\rightarrow \boxed{g_p = \left(I - A^T\left(AA^T\right)^{-1}A\right)g \quad \text{avec} \quad A = \nabla c(x_0)^T}$



- g_p vérifie : $\begin{cases} Ag_p = 0 & \rightarrow g_p \in \text{hyperplan tangent} \\ g_p^T\left(g - g_p\right) = 0 & \rightarrow g - g_p \perp \text{hyperplan tangent} \end{cases}$ car $\begin{cases} P^T = P \\ P^2 = P \end{cases}$

Direction de descente

- La direction du gradient projeté est la **direction de plus forte pente dans l'hyperplan tangent**
 = direction dans l'hyperplan qui maximise la dérivée directionnelle de f

Preuve
La direction d dans l'hyperplan maximisant la dérivée directionnelle de f est solution de

$$\min_{d\in R^n} g^t d \quad sous \quad \begin{cases} Ad = 0 & \rightarrow d \in \text{hyperplan tangent} \\ \|d\| = 1 \Leftrightarrow d^T d = 1 & \rightarrow \text{norme} = 1 \end{cases}$$

$$Lagrangien : L(d,\lambda,\mu) = g^t d + \lambda^T Ad + \mu\left(d^T d - 1\right) \quad \text{avec } \lambda \in R^m\ ,\ \mu \in R$$

$$Conditions\ KKT : \begin{cases} g + A^T\lambda + 2\mu d = 0 \rightarrow & d = -\left(g + A^T\lambda\right)/(2\mu) \\ Ad = 0 & \rightarrow\ Ag + AA^T\lambda = 0 \ \rightarrow\ \lambda = -\left(AA^T\right)^{-1} Ag \\ \|d\| = 1 & \rightarrow\ 2\mu = \pm\left\|g + A^T\lambda\right\| \end{cases}$$

d est bien un vecteur normé colinéaire à $\left(I - A^T\left(AA^T\right)^{-1}A\right)g$

- La méthode du gradient projeté équivaut à la méthode de plus forte pente appliquée dans l'espace nul des contraintes $\rightarrow$ méthode d'ordre 1 peu efficace
 $\rightarrow$ amélioration par méthode de quasi-Newton

c) Gradient réduit

Définition

$$\min_{x\in R^n} f(x) \text{ sous } c(x) = 0$$

Le gradient réduit est le gradient de la **fonction réduite** sur une base de l'espace nul tangent.

Expression du gradient réduit

- Base Z de l'espace nul tangent aux contraintes : $AZ = 0$ avec $A = \nabla c(x_0)^T$
- Décomposition du déplacement : $p = Yp_Y + Zp_Z$ avec $\begin{cases} AZ = 0 \\ AY \text{ inversible} \end{cases}$
- Déplacement admissible : $Ap = 0 \Rightarrow \begin{cases} p_Y = 0 \\ p = Zp_Z \end{cases}$
- Fonction réduite f_r : $\min_{p\in R^n} f(x_0 + p)$ sous $A(x_0 + p) = b \Leftrightarrow \min_{p_Z\in R^{n-m}} f_r(p_Z) = f\left(x_0 + Zp_Z\right)$
- **Notations**
 $g(x_0)$ gradient de f en x_0 $\rightarrow g(x_0) = \nabla f(x_0)$ $\rightarrow g \in R^n$
 $g_r(x_0)$ gradient réduit $\rightarrow g_r(x_0) = \nabla f_r(p_Z{=}0)$ $\rightarrow g_r \in R^{n-m}$ (m = nombre de contraintes)
 $f_r(p_Z) = f\left(x_0 + Zp_Z\right) \Rightarrow \nabla f_r(p_Z) = Z^T \nabla f\left(x_0 + Zp_Z\right) \rightarrow \boxed{g_r = Z^T g}$ en $p_Z = 0$
 $\rightarrow$ g_r est le gradient de la fonction réduite f_r (= fonction de n−m variables p_Z)

Direction de descente

- Le déplacement à partir du point x_0 admissible est décomposé en

$$p = Yp_Y + Zp_Z \quad \text{avec} \quad \begin{cases} AZ = 0 \\ AY \text{ inversible} \end{cases} \qquad Ap = 0 \;\Rightarrow\; \begin{cases} p_Y = 0 \\ p = Zp_Z \end{cases}$$

- Le gradient réduit g_r donne la direction **de plus forte pente suivant les variables p_Z**.
  La direction de déplacement dans R^n est : **$d = Zg_r$** .

- On peut choisir les matrices Y et Z
  - à partir de matrices orthogonales → factorisation QR de A
  - à partir d'une base de A → $B \in R^{m\times m}$ (= m colonnes indépendantes de A)

Gradient réduit sur une base B de A

- $$AE = \begin{array}{c} \begin{array}{cc} m & n-m \end{array} \\ \begin{pmatrix} B & N \end{pmatrix} \end{array} \quad \Rightarrow \quad g = \begin{pmatrix} g_B \\ g_N \end{pmatrix} \qquad Y = \begin{array}{c} m \\ \begin{pmatrix} B^{-1} \\ 0 \end{pmatrix} \end{array} \begin{array}{l} m \\ n-m \end{array} \qquad Z = \begin{array}{c} n-m \\ \begin{pmatrix} -B^{-1}N \\ I \end{pmatrix} \end{array} \begin{array}{l} m \\ n-m \end{array}$$

  (E = matrice de permutation de colonnes de A)

- Le gradient réduit par rapport à la base B est : $\boxed{g_r = Z^T g = g_N - \left(B^{-1}N\right)^T g_B}$

d) Direction de déplacement

Problème avec contraintes égalité

$$\min_{x\in R^n} f(x) \text{ sous } c(x) = 0 \qquad \rightarrow \text{ m contraintes actives}$$

- On construit la direction de déplacement $d \in R^n$ dans l'hyperplan tangent aux contraintes en x_0
  $Ad = 0$ avec $A = \nabla c(x_0)^T$
  → 2 méthodes de construction de la direction d

- **Méthode du gradient projeté**
  La direction d est celle du gradient projeté : $d = g_p$

  $\boxed{d = Pg \quad \text{avec} \quad P = I - A^T\left(AA^T\right)^{-1}A}$ (P = matrice de projection sur l'hyperplan tangent)

- **Méthode du gradient réduit**
  La direction d est obtenue à partir du gradient de la fonction réduite : $d = Zg_r$ avec $g_r = Z^Tg$

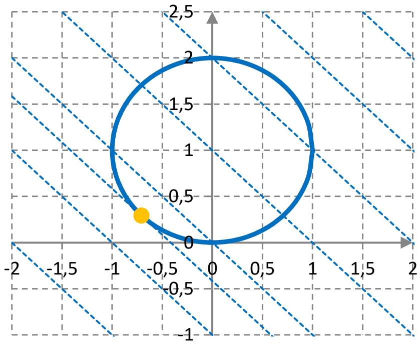
  $\boxed{d = ZZ^Tg \quad \text{avec} \quad AZ=0}$ (Z = base de l'hyperplan tangent)

- On cherche un pas de déplacement suivant –d pour minimiser f.
  Un déplacement complémentaire est nécessaire pour restaurer l'admissibilité.

Exemple

- Minimisation de $f(x) = x_1 + x_2$ sous $c(x) = x_1^2 + (x_2 - 1)^2 - 1 = 0$

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



- Changement de variables (coordonnées polaires)

$$\begin{cases} x_1 = r\cos\theta \\ x_2 = r\sin\theta + 1 \end{cases} \rightarrow \begin{cases} f(r,\theta) = r(\cos\theta + \sin\theta) + 1 \\ c(r,\theta) = r^2 - 1 \end{cases}$$

- Elimination variable r

$r = 1 \rightarrow f(\theta) = \cos\theta + \sin\theta + 1$

- **Minimum**

$$\left.\begin{cases} f'(\theta) = -\sin\theta + \cos\theta = 0 \rightarrow \tan\theta = 1 & \rightarrow \theta = \dfrac{\pi}{4} \text{ ou } \dfrac{\pi}{4} + \pi \\ f''(\theta) = -\cos\theta - \sin\theta \geq 0 \rightarrow -\cos\theta(1 + \tan\theta) \geq 0 & \rightarrow -\cos\theta \geq 0 \end{cases}\right\} \rightarrow \theta^* = \frac{\pi}{4} + \pi$$

$$\rightarrow \begin{cases} x_1^* = \quad -1/\sqrt{2} \approx -0.70711 \\ x_2^* = 1 - 1/\sqrt{2} \approx \quad 0.29289 \end{cases}$$

Gradient projeté

- Minimisation de $f(x) = x_1 + x_2$ sous $c(x) = x_1^2 + (x_2 - 1)^2 - 1 = 0$

Point admissible x_0 (r_0=1, θ_0) $\rightarrow \nabla f(x_0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\nabla c(x_0) = \begin{pmatrix} 2x_1 \\ 2(x_2 - 1) \end{pmatrix} = 2r_0 \begin{pmatrix} \cos\theta_0 \\ \sin\theta_0 \end{pmatrix}$

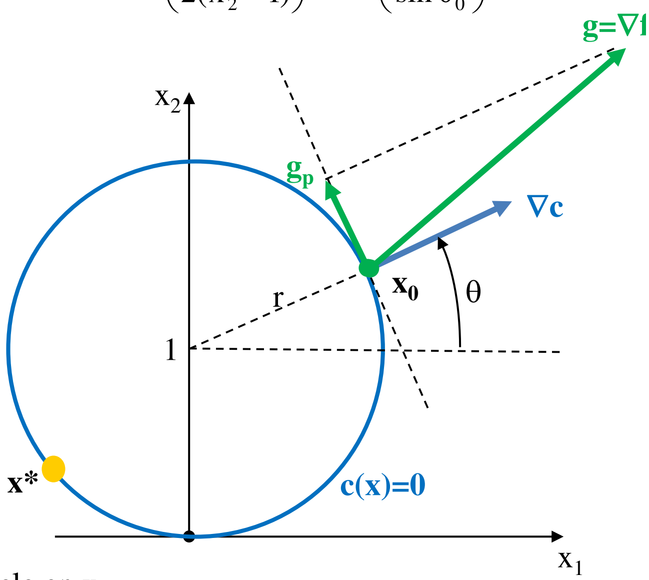
- Gradient projeté au point x_0

$$g_p = Pg \quad \text{avec} \quad \begin{cases} A = \nabla c(x_0)^T \ , \ g = \nabla f(x_0) \\ P = I - A^T (AA^T)^{-1} A \end{cases}$$

$$A = 2r_0 (\cos\theta_0 \quad \sin\theta_0)$$

$$P = \begin{pmatrix} \sin^2\theta_0 & -\sin\theta_0\cos\theta_0 \\ -\sin\theta_0\cos\theta_0 & \cos^2\theta_0 \end{pmatrix}$$

$$\rightarrow \boxed{g_p = (\cos\theta_0 - \sin\theta_0) \begin{pmatrix} -\sin\theta_0 \\ \cos\theta_0 \end{pmatrix}}$$



- Direction de descente au point x_0

$$\boxed{d = -\frac{g_p}{\|g_p\|} = \pm \begin{pmatrix} -\sin\theta_0 \\ \cos\theta_0 \end{pmatrix}} \rightarrow \textbf{tangente} \text{ au cercle en } x_0$$

Gradient réduit

- Minimisation de $f(x) = x_1 + x_2$ sous $c(x) = x_1^2 + (x_2 - 1)^2 - 1 = 0$

  Point admissible x_0 (r_0=1, θ_0) $\rightarrow$ $\nabla f(x_0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\nabla c(x_0) = \begin{pmatrix} 2x_1 \\ 2(x_2 - 1) \end{pmatrix} = 2r_0 \begin{pmatrix} \cos\theta_0 \\ \sin\theta_0 \end{pmatrix}$

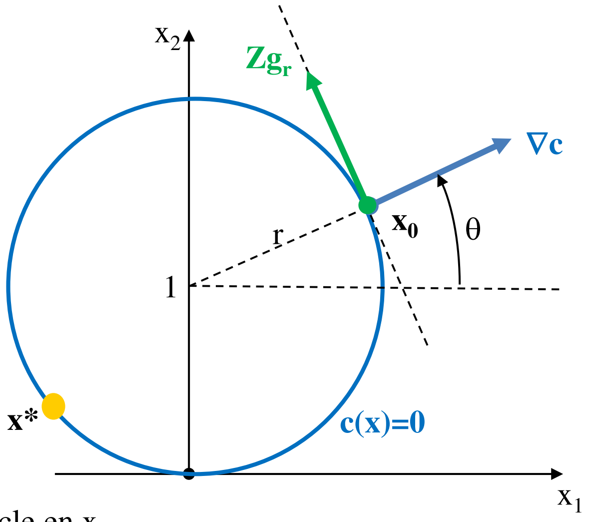
- Gradient réduit au point x_0

$$g_r = Z^T g \quad \text{avec} \quad \begin{cases} A = \nabla c(x_0)^T \ , \ g = \nabla f(x_0) \\ AE = \begin{pmatrix} B & N \end{pmatrix} , \ Z = \begin{pmatrix} -B^{-1}N \\ I \end{pmatrix} \end{cases}$$

$$A = 2r_0 \begin{pmatrix} \cos\theta_0 & \sin\theta_0 \end{pmatrix} \qquad Z = \begin{pmatrix} -\tan\theta_0 \\ 1 \end{pmatrix}$$

$$\rightarrow \boxed{g_r = 1 - \tan\theta_0}$$

$$\rightarrow \ Zg_r = \frac{\cos\theta_0 - \sin\theta_0}{\cos\theta_0} \begin{pmatrix} -\sin\theta_0 \\ \cos\theta_0 \end{pmatrix}$$



- Direction de descente au point x_0

$$\boxed{d = -\frac{Zg_r}{\|Zg_r\|} = \pm \begin{pmatrix} -\sin\theta_0 \\ \cos\theta_0 \end{pmatrix}} \rightarrow \text{tangente au cercle en } x_0$$

596

5.3 Restauration

a) Point initial

Itérations admissibles

La méthode du gradient projeté ou réduit construit une suite de solutions admissibles
$\rightarrow$ point initial admissible
$\rightarrow$ restauration de la faisabilité à chaque itération

Point initial

- On peut construire un point initial admissible du problème $\min\limits_{x \in R^n} f(x)$ sous $\begin{cases} c_E(x) = 0 \\ c_I(x) \leq 0 \end{cases}$

  en résolvant le **problème préliminaire sans contrainte** $\min\limits_{x \in R^n} \|c_E(x)\|_2 + \|\max(0, c_I(x))\|_2$

- La solution x_0 de ce problème préliminaire est admissible si le coût est nul.

$$\|c_E(x_0)\|_2 + \|\max(0, c_I(x_0))\|_2 = 0 \ \Rightarrow \ \begin{cases} c_E(x_0) = 0 \\ \max(0, c_I(x_0)) = 0 \end{cases} \ \Rightarrow \ \begin{cases} c_E(x_0) = 0 \\ c_I(x_0) \leq 0 \end{cases}$$

Restauration de la faisabilité

- La direction de descente d est dans l'hyperplan tangent aux contraintes au point courant.
- Si les contraintes ne sont pas linéaires, un pas s suivant d donne un point non admissible.
  $\rightarrow$ **Il faut restaurer la faisabilité avant d'évaluer si le pas s est acceptable**.

b) Itérations admissibles

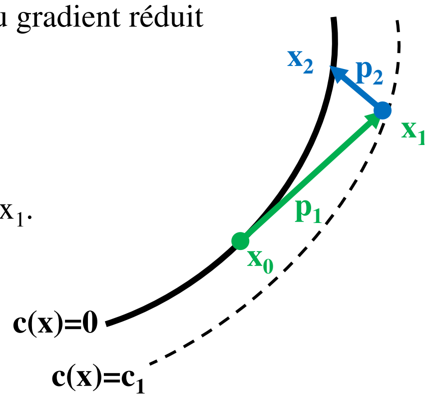
Déplacement admissible

$\min_{x \in R^n} f(x)$ sous $c(x) = 0$ $\quad \to$ m contraintes actives

On construit le déplacement p à partir du point initial x_0 en 2 étapes : $\mathbf{p = p_1 + p_2}$

- Le déplacement p_1 est suivant la direction de descente d dans l'hyperplan tangent : $p_1 = -sd$
  $d \in R^n$ = direction construite à partir du gradient projeté ou du gradient réduit
  $s > 0$ = pas de déplacement suivant $-d$ (pour minimisation)
  On obtient un point $\mathbf{x_1 = x_0 + p_1}$ dans l'hyperplan tangent
  $\to$ non admissible si contraintes non linéaires

- Le déplacement p_2 restaure un point admissible à partir du point x_1.
  $\to$ linéarisation des contraintes en x_1
  $\to$ résolution d'un système sous-déterminé
  On obtient un point $\mathbf{x_2 = x_1 + p_2}$ admissible.



Recherche linéaire

- On évalue **le point x_2** correspondant au pas s de recherche linéaire suivant d.
- Le pas s est modifié par dichotomie jusqu'à trouver un point $x_2(s)$ acceptable.
  $\to$ règles d'Armijo, Goldstein, Wolfe,…

c) Méthode de restauration

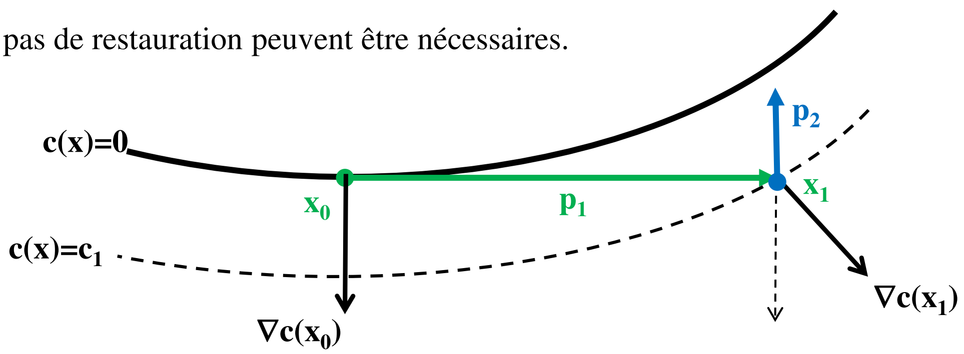
Méthode de restauration

Le déplacement p_2 doit vérifier : $\quad A_1 p = b_1 \quad$ avec $\begin{cases} A_1 = \nabla c(x_1)^T \approx \nabla c(x_0)^T = A_0 \\ b_1 = -c(x_1) = -c_1 \end{cases}$

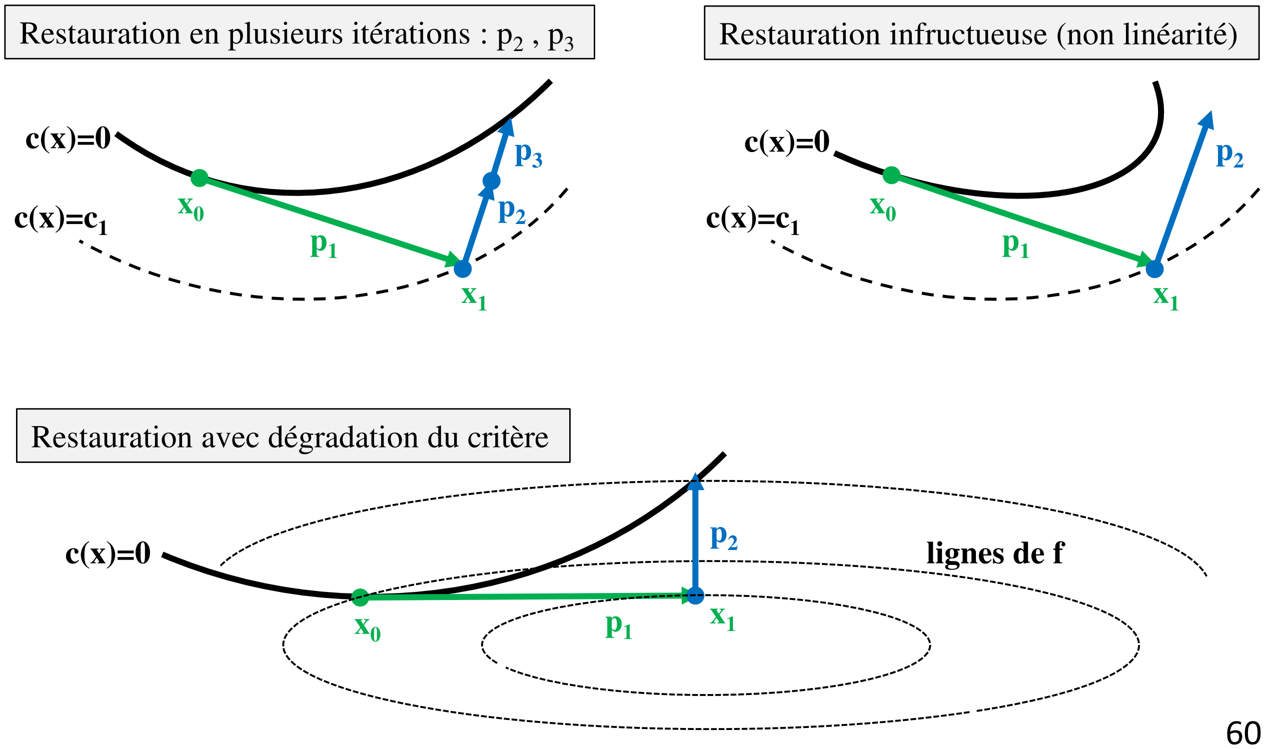
- **Solution de norme minimale** $\to$ projection sur l'hyperplan tangent aux contraintes actives (cf §1.2.4)
  $\min_{p \in R^n} \|p\|$ sous $A_1 p = b_1 \quad \to \quad p_2 = A_1^T \left(A_1 A_1^T\right)^{-1} b_1$

- **Solution de base** $\to$ pour ne pas dégrader la minimisation apportée par p_1 (cf §1.2.3)
  $A_1(Yp_Y + Zp_Z) = b_1 \;\Rightarrow\; p_Y = (A_1 Y)^{-1} b_1 \;\to\; p_2 = Y(A_1 Y)^{-1} b_1$

- Plusieurs pas de restauration peuvent être nécessaires.



Illustrations



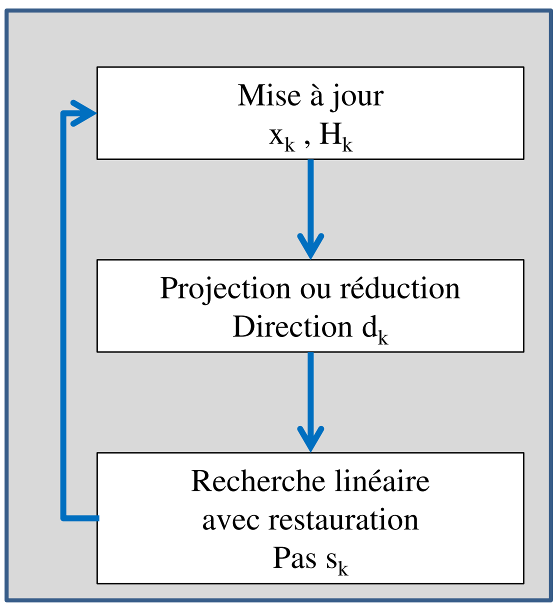
601

5.4 Algorithme

a) Algorithme de gradient projeté/réduit

Algorithme de gradient projeté/réduit

Les principales étapes d'une itération de gradient projeté/réduit sont
- construire la direction de descente au point courant
- effectuer une recherche linéaire avec restauration



Direction de descente
- Sélection des contraintes actives
- Projection ou réduction dans l'hyperplan tangent
- Mise à jour du hessien (quasi-Newton)

Recherche linéaire
- Méthode de dichotomie sur le pas de déplacement
- Restauration avant évaluation du pas
- Règles d'acceptation (Armijo,…)

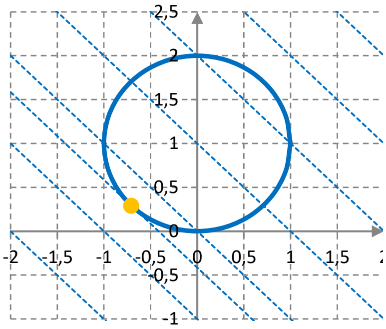
Principales difficultés
- Amélioration critère → grands pas
- Restauration contraintes → petits pas

→ difficultés sur problèmes très non-linéaires
→ réglages à adapter au cas par cas

b) Exemple

Exemple

- Minimisation de $f(x) = x_1 + x_2$ sous $c(x) = x_1^2 + (x_2 - 1)^2 - 1 = 0$
- Solution : $\begin{cases} x_1{}^* = -1/\sqrt{2} \approx -0.70711 \\ x_2{}^* = 1 - 1/\sqrt{2} \approx 0.29289 \end{cases}$



Itérations

- Point courant : $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} r\cos\theta \\ r\sin\theta + 1 \end{pmatrix}$
- Descente : $x' = x - s_1 d_1$ avec $d_1 = \pm\begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix} \leftarrow g_p = (\cos\theta - \sin\theta)\begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$

  $\rightarrow$ **pas s_1 suivant le gradient projeté**
- Restauration : $x'' = x' - s_2 d_2$ avec $d_2 = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \leftarrow \nabla c(x) = \begin{pmatrix} 2x_1 \\ 2(x_2 - 1) \end{pmatrix}$

  $\rightarrow$ **pas s_2 suivant le gradient des contraintes**
- Réglage des pas : s_2 est calculé pour restaurer $c(x'') = 0$

  s_1 est choisi pour vérifier une décroissance suffisante $f(x'') < f(x)$

Exemple

- Minimisation de $f(x) = x_1 + x_2$ sous $c(x) = x_1^2 + (x_2 - 1)^2 - 1 = 0$
- Point initial : $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0.1 \\ 1 \end{pmatrix}$ $\rightarrow$ Restauration initiale : $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

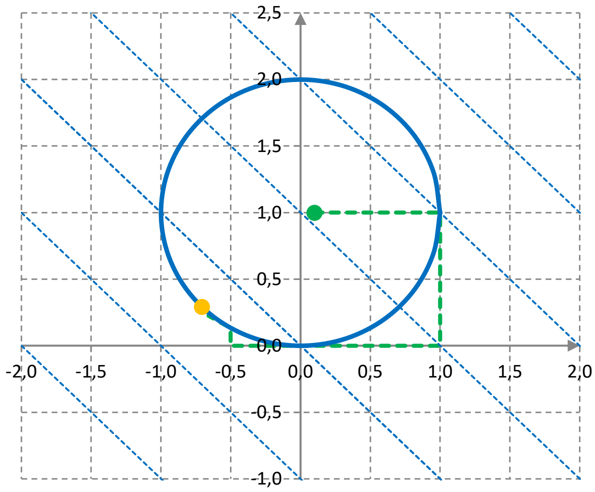
Itération	x_1	x_2	f(x)	c(x)	Descente s_1	x_1'	x_2'	c(x')	Restauration s_2
1	**0,10000**	**1,00000**	**1,10000**	**-0,99000**	0,00000	0,10000	1,00000	-0,99000	4,50000
2	1,00000	1,00000	2,00000	0,00000	1,00000	1,00000	0,00000	1,00000	-0,50000
3	0,00000	0,00000	0,00000	0,00000	0,50000	-0,50000	0,00000	0,25000	-0,06699
4	-0,50000	0,13397	-0,36603	0,00000	0,18301	-0,65849	0,22548	0,03349	-0,00844
5	-0,65005	0,24011	-0,40994	0,00000	5,492E-02	-0,69178	0,27581	3,016E-03	-7,547E-04
6	-0,69080	0,27696	-0,41385	0,00000	1,612E-02	-0,70246	0,28809	2,599E-04	-6,497E-05
7	-0,70237	0,28819	-0,41418	0,00000	4,722E-03	-0,70573	0,29150	2,230E-05	-5,576E-06
8	-0,70572	0,29151	-0,41421	0,00000	1,383E-03	-0,70670	0,29249	1,913E-06	-4,783E-07
9	-0,70670	0,29249	-0,41421	0,00000	4,051E-04	-0,70699	0,29277	1,641E-07	-4,103E-08
10	-0,70699	0,29277	-0,41421	0,00000	1,187E-04	-0,70707	0,29286	1,408E-08	-3,520E-09
11	-0,70707	0,29286	-0,41421	0,00000	3,475E-05	-0,70710	0,29288	1,208E-09	-3,020E-10
12	**-0,70710**	**0,29288**	**-0,41421**	**0,00000**					

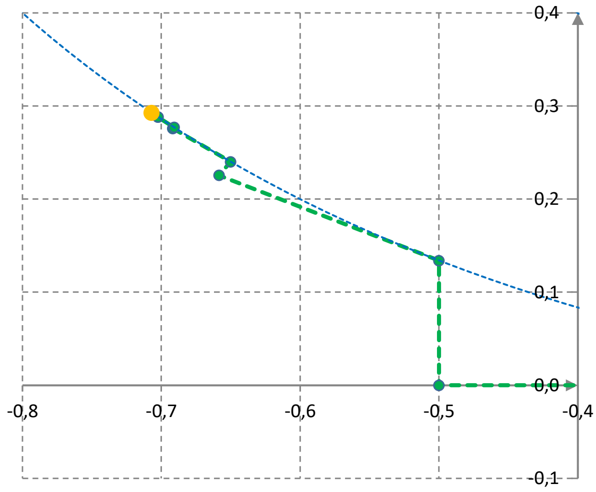
Exemple

Minimisation de $f(x) = x_1 + x_2$ sous $c(x) = x_1^2 + (x_2 - 1)^2 - 1 = 0$

Point initial **Point initial admissible** Solution

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0.1 \\ 1 \end{pmatrix} \xrightarrow{\text{restauration initiale}} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \xrightarrow{\text{itérations}} \begin{pmatrix} x_1 * \\ x_2 * \end{pmatrix} = \begin{pmatrix} -1/\sqrt{2} \approx -0.70711 \\ 1 - 1/\sqrt{2} \approx \ 0.29289 \end{pmatrix}$$





6 Lagrangien augmenté

6.1 Principes

Problème avec contraintes égalité

$$\min_{x \in R^n} f(x) \text{ sous } \begin{cases} c_E(x) = 0 \\ c_I(x) \le 0 \end{cases} \Leftrightarrow \min_{x \in R^n} f(x) \text{ sous } c(x) = 0 \quad \rightarrow \text{ contraintes actives}$$

La difficulté de résolution vient des 2 objectifs antagonistes :

- Minimiser le critère f(x)
- Satisfaire les contraintes c(x)=0

Méthodes de pénalisation

Les contraintes sont ajoutées à la fonction coût avec une pondération :

- Critère augmenté → pondération = pénalisation des contraintes
- Lagrangien → pondération = multiplicateurs de Lagrange
- Lagrangien augmenté → pondération = pénalisation + multiplicateurs

→ On se ramène à un **problème sans contraintes** plus simple

Les difficultés viennent du réglage de la pondération :

- Le problème pénalisé sans contraintes doit être équivalent au problème avec contraintes.
- Le problème pénalisé est mal conditionné lorsque la pénalisation est grande.

6.2 Pénalisation

a) Critère augmenté

Problème avec contraintes égalité

$$\min_{x\in R^n} f(x) \ \text{sous} \ \begin{cases} c_E(x)=0 \\ c_I(x)\le 0 \end{cases} \Leftrightarrow \min_{x\in R^n} f(x) \ \text{sous} \ c(x)=0 \quad \rightarrow \text{contraintes actives}$$

On note x* la solution du problème avec contraintes.

Critère augmenté

On ajoute au critère un terme positif fonction de la violation des contraintes
avec un coefficient de pénalisation $\rho > 0$ $\rightarrow$ 2 méthodes usuelles de pénalisation

- **Pénalisation en norme 2** (pénalisation quadratique)

$$f_\rho(x)=f(x)+\frac{1}{2}\rho\left(\|c_E(x)\|_2^2+\|\max(0,c_I(x))\|_2^2\right) \quad \Leftrightarrow \quad f_\rho(x)=f(x)+\frac{1}{2}\rho\|c(x)\|_2^2$$

- **Pénalisation en norme 1**

$$f_\rho(x)=f(x)+\rho\left(\|c_E(x)\|_1+\|\max(0,c_I(x))\|_1\right) \quad \Leftrightarrow \quad f_\rho(x)=f(x)+\rho\|c(x)\|_1$$

Problème sans contraintes

$$\min_{x\in R^n} f_\rho(x) \quad \rightarrow \text{solution } x_\rho$$

b) Pénalisation quadratique

Problème pénalisé l_2

$$\min_{x\in R^n} f_\rho(x) \quad \text{avec} \quad f_\rho(x)=f(x)+\frac{1}{2}\rho\left(\|c_E(x)\|_2^2+\|\max(0,c_I(x))\|_2^2\right)$$

$$\Leftrightarrow \quad f_\rho(x)=f(x)+\frac{1}{2}\rho\|c(x)\|_2^2 \quad \rightarrow \text{contraintes actives}$$

Le critère l_2 est différentiable deux fois pour un problème avec contraintes égalité.
On peut appliquer les algorithmes d'optimisation sans contraintes à base de gradient.

Méthode de résolution

- On résout une suite de problèmes pénalisés avec des valeurs croissantes de la pénalisation ρ.
- Chaque problème k+1 est initialisé avec la solution précédente x_k.
- Problème k avec pénalisation ρ_k : $\min_{x\in R^n} f_{\rho_k}(x) \quad \rightarrow$ solution x_k
- Il faut vérifier que la suite des solutions x_k converge vers la solution x* du problème initial

  $$\lim_{k\to\infty} x_k = x^* \quad \text{si} \quad \lim_{k\to\infty} \rho_k = +\infty$$

  $\rightarrow$ 2 résultats de convergence selon que x_k est une solution exacte ou approchée

Problème pénalisé l_2

- Problème avec contraintes : $\min_{x\in R^n} f(x)$ sous $c(x)=0$ $\rightarrow$ solution x^*
- Problème k avec pénalisation ρ_k : $\min_{x\in R^n} f_{\rho_k}(x)$ $\rightarrow$ solution x_k
  $\lim_{k\to\infty} \rho_k = +\infty$ $\qquad \lim_{k\to\infty} x_k = x_\infty$

Convergence

- Si x_k est le **minimum global exact**, alors $\lim_{k\to\infty} x_k = x^*$
- Si x_k est un **minimum local approché** : $\left\|\nabla f_{\rho_k}(x_k)\right\| \le \varepsilon_k$ avec $\lim_{k\to\infty} \varepsilon_k = 0$
  $\rightarrow$ précision de résolution ε_k décroissante

  alors la limite x_∞ est :
  - soit un point non admissible qui minimise $\|c(x)\|_2^2$
  - soit un point x^* vérifiant les conditions KKT du problème initial

  On a dans ce 2ème cas : $\boxed{\begin{cases}\lim_{k\to\infty} x_k = x^* \\ \lim_{k\to\infty} \rho_k c(x_k) = \lambda^*\end{cases}}$ $\begin{array}{l}\rightarrow \text{minimum local} \\ \rightarrow \text{multiplicateurs des contraintes actives}\end{array}$

  La solution exacte x* n'est obtenue qu'à la limite lorsque la pénalisation ρ tend vers l'infini.

Eléments de la démonstration

$$f_\rho(x) = f(x) + \frac{1}{2}\rho\|c(x)\|^2 \quad \Rightarrow \quad \nabla f_\rho(x) = \nabla f(x) + \rho \nabla c(x) c(x)$$

- *Critère d'arrêt sur* x_k *:* $\left\|\nabla f_{\rho_k}(x_k)\right\| \le \varepsilon_k$

$$\left\|\nabla f_{\rho_k}(x_k)\right\| = \left\|\nabla f(x_k) + \rho_k \nabla c(x_k) c(x_k)\right\| \ge \left\|\rho_k \nabla c(x_k) c(x_k)\right\| - \left\|\nabla f(x_k)\right\| \quad car\ \|a+b\| \ge \|a\| - \|b\|$$

$$\Rightarrow \left\|\rho_k \nabla c(x_k) c(x_k)\right\| \le \varepsilon_k + \left\|\nabla f(x_k)\right\| \quad \Rightarrow \left\|\nabla c(x_k) c(x_k)\right\| \le \frac{\varepsilon_k + \left\|\nabla f(x_k)\right\|}{\rho_k} \xrightarrow[k\to\infty]{} 0$$

$$\Rightarrow \begin{cases} \nabla c(x^*) = 0 \ \rightarrow \ \min\|c(x)\| \\ \qquad ou \\ c(x^*) = 0 \ \rightarrow \ admissible \end{cases} \qquad si\ les\ gradients\ sont\ linéairement\ indépendants$$

- *Multiplicateurs de Lagrange*

$$\left\|\nabla f_{\rho_k}(x_k)\right\| = \left\|\nabla f(x_k) + \rho_k \nabla c(x_k) c(x_k)\right\| \le \varepsilon_k \xrightarrow[k\to\infty]{} 0$$

$$\rightarrow \nabla f(x^*) + \nabla c(x^*)\lambda^* = 0 \qquad avec \quad \lambda^* = \lim_{k\to\infty} \rho_k c(x_k)$$

c) Pénalisation exacte

Problème pénalisé l_1

$$\min_{x \in R^n} f_\rho(x) \quad \text{avec} \quad f_\rho(x) = f(x) + \rho\left(\|c_E(x)\|_1 + \|\max(0, c_I(x))\|_1\right)$$

$$\Leftrightarrow \quad f_\rho(x) = f(x) + \rho\|c(x)\|_1 \qquad \rightarrow \text{contraintes actives}$$

Le critère l_1 n'est pas différentiable.

Méthode de résolution

- On résout une suite de problèmes pénalisés avec des valeurs croissantes de la pénalisation ρ.
- Chaque problème k+1 est initialisé avec la solution précédente x_k.
- Problème k avec pénalisation ρ_k : $\min_{x \in R^n} f_{\rho_k}(x) \qquad \rightarrow$ solution x_k

Convergence

- Si $\rho > \rho^* = \|\lambda^*\|_\infty = \max|\lambda_i|$ alors **x* est un minimum local de f_ρ** avec la pénalisation l_1.
- La pénalisation l_1 est **exacte** si est ρ est supérieur au plus grand multiplicateur.
  $\rightarrow$ <u>ne nécessite pas d'augmenter indéfiniment ρ pour obtenir la solution exacte x*</u>

d) Mise en œuvre

Méthodes avec critère augmenté

- Type de pénalisation
  - Pénalisation l_2 $\rightarrow$ différentiable, mais nécessite une pénalisation forte pour approcher x*
  - Pénalisation l_1 $\rightarrow$ exacte, mais non différentiable
- Réglage de la pénalisation
  - Trop faible $\rightarrow$ risque de divergence (pas de minimum du problème pénalisé)
  - Trop forte $\rightarrow$ mauvais conditionnement, difficultés numériques
- Utilisation du critère augmenté
  - Difficultés pratiques si l'on veut obtenir une bonne précision sur la solution x*
  - Le critère augmenté peut servir de fonction mérite
    dans le cadre d'autres algorithmes pour évaluer la progression vers l'optimum.

Méthodes avec lagrangien augmenté

On cherche à se ramener à une suite de problèmes sans contraintes
- en conservant un critère différentiable
- en évitant le mauvais conditionnement du à une pénalisation trop forte
  $\rightarrow$ utilisation des multiplicateurs de Lagrange pour réduire la pénalisation
  $\rightarrow$ **méthode duale** ou **lagrangienne**

e) Lagrangien augmenté

Problème pénalisé l_2

- La méthode de pénalisation consiste à minimiser le critère augmenté.
  $\min_{x\in R^n} f_\rho(x) = f(x) + \frac{1}{2}\rho\|c(x)\|^2$
- La convergence est obtenue pour des valeurs croissantes de pénalisation.

$$\lim_{k\to\infty} \rho_k = +\infty \quad \to \quad \begin{cases} \lim_{k\to\infty} x_k = x^* & \to \text{ minimum local} \\ \lim_{k\to\infty} \rho_k c(x_k) = \lambda^* & \to \text{ multiplicateurs des contraintes actives} \end{cases}$$

- La solution x_k ne respecte qu'approximativement les contraintes : $c(x_k) \approx \dfrac{\lambda^*}{\rho_k}$
- Pour respecter précisément les contraintes, il faut augmenter fortement la pénalisation.
  $\to$ cause de mauvais conditionnement et de difficultés numériques

- On peut appliquer la méthode de pénalisation **au problème équivalent**

$$\boxed{\min_{x\in R^n} L(x,\lambda^*) \text{ sous } c(x)=0 \iff \min_{x\in R^n} f(x) \text{ sous } c(x)=0} \quad \to \text{ si l'on connait } \lambda^*$$

$$\to \min_{x\in R^n} L_\rho(x,\lambda^*) = L(x,\lambda^*) + \frac{1}{2}\rho\|c(x)\|_2^2 = f(x) + \lambda^{*T} c(x) + \frac{1}{2}\rho\|c(x)\|^2$$

Lagrangien augmenté

- La méthode de lagrangien augmenté consiste à résoudre une suite de problèmes :
  $\min_{x\in R^n} L_{\rho_k}(x,\lambda_k) = L(x,\lambda_k) + \frac{1}{2}\rho_k\|c(x)\|_2^2 = f(x) + \lambda_k^T c(x) + \frac{1}{2}\rho_k\|c(x)\|^2$
  avec ρ_k = valeur de pénalisation du problème k
  λ_k = estimation des multiplicateurs pour le problème k

- Si $\lim_{k\to\infty} \lambda_k = \lambda^*$ et $\lim_{k\to\infty} \rho_k = +\infty$ les problèmes deviennent équivalents.

$$\left.\begin{aligned} &\min_{x\in R^n} L(x,\lambda^*) \text{ sous } c(x)=0 \iff \min_{x\in R^n} f(x) \text{ sous } c(x)=0 \\ \iff &\min_{x\in R^n} L_\rho(x,\lambda^*) \qquad\qquad\quad \iff \min_{x\in R^n} f_\rho(x) \end{aligned}\right\} \to \lim_{k\to\infty} x_k = x^*$$

La solution x_k du problème $\min_{x\in R^n} L_{\rho_k}(x,\lambda_k)$ converge vers la solution x* du problème initial.

- La solution x_k vérifie également : $\nabla_x L_{\rho_k}(x_k,\lambda_k) = \nabla f(x_k) + \nabla c(x_k)\big(\lambda_k + \rho_k c(x_k)\big) = 0$
  à comparer à x* qui vérifie : $\nabla_x L(x^*,\lambda^*) = \nabla f(x^*) + \nabla c(x^*)\lambda^* = 0$

$$\Rightarrow \boxed{\lim_{k\to\infty} \lambda_k + \rho_k c(x_k) = \lambda^*}$$

(même démonstration que pour le critère augmenté avec pénalisation quadratique)

Lagrangien augmenté

- On peut estimer les multiplicateurs à l'itération k.

$$\lim_{k\to\infty} \lambda_k + \rho_k c(x_k) = \lambda^* \quad \rightarrow \quad \lambda^* \approx \lambda_k + \rho_k c(x_k) \quad \text{pour } \rho_k \text{ assez grand}$$

$$\rightarrow \quad \lambda_{k+1} = \lambda_k + \rho_k c(x_k) \quad \text{pour l'itération k+1}$$

- La valeur des contraintes à l'itération k est : $\boxed{c(x_k) \approx \dfrac{\lambda^* - \lambda_k}{\rho_k}} \xrightarrow[\lambda_k \to \lambda^*]{} 0$

  On peut parvenir à respecter les contraintes <u>sans augmenter indéfiniment la pénalisation</u>
  si λ_k est une bonne estimation des multiplicateurs $\rightarrow$ **méthode duale**
  $\rightarrow$ meilleur conditionnement
  $\rightarrow$ convergence plus rapide et précise que la méthode du critère augmenté
  $\rightarrow$ **méthode de lagrangien augmenté** appelée aussi **méthode des multiplicateurs**

Convergence

Pour ρ assez grand, **la solution x* du problème initial est un minimum local du problème**

$$\min_{x\in R^n} L_\rho(x,\lambda^*) = f(x) + \lambda^{*T} c(x) + \frac{1}{2}\rho\|c(x)\|^2 \quad \rightarrow \text{ pénalisation } \textbf{exacte} \text{ si on connaît } \lambda^*$$

$\rightarrow$ <u>ne nécessite pas d'augmenter indéfiniment ρ pour obtenir la solution exacte x*</u>

6.3 Méthode duale

a) Problème dual

Problème avec contraintes égalité

- Problème primal : $\min_{x\in R^n} f(x)$ sous $c(x) = 0$ $\rightarrow$ m contraintes actives
- Lagrangien : $L(x,\lambda) = f(x) + \lambda^T c(x)$ $\rightarrow$ $x = n$ variables primales
  $\rightarrow$ $\lambda = m$ variables duales
- Fonction duale : $w(\lambda) = \min_{x\in R^n} L(x,\lambda)$ $\rightarrow$ minimisation par rapport à x
- Problème dual : $\max_{\lambda\in R^m} w(\lambda)$ $\rightarrow$ maximisation par rapport à λ

Méthode duale (ou lagrangienne)

Une méthode duale consiste à résoudre le problème dual : $\boxed{\max_{\lambda\in R^m} w(\lambda)}$

- Le problème est sans contraintes, mais <u>la fonction $w(\lambda)$ est coûteuse à évaluer</u>.
- La solution (x^*,λ^*) est un point col du lagrangien : $\forall x,\lambda,\ L(x^*,\lambda) \le L(x^*,\lambda^*) \le L(x,\lambda^*)$
  $\rightarrow$ **existence non garantie** (saut de dualité)

b) Maximisation duale

Fonction duale

On note $x_L(\lambda)$ la valeur de x qui minimise $L(x,\lambda)$ à λ fixé.

Condition d'ordre 1 : $\min_{x\in R^n} L(x,\lambda)$ en $x = x_L(\lambda) \Rightarrow \boxed{\nabla_x L(x_L(\lambda),\lambda) = 0}$

- Fonction w : $w(\lambda) = \min_{x\in R^n} L(x,\lambda) = L(x_L(\lambda),\lambda) \rightarrow$ fonction composée
- Gradient de w : $\nabla w(\lambda) = \nabla x_L(\lambda).\underbrace{\nabla_x L(x_L(\lambda),\lambda)}_{=0} + \underbrace{\nabla_\lambda L(x_L(\lambda),\lambda)}_{=c(x_L(\lambda))}$

  $\Rightarrow \boxed{\nabla w(\lambda) = c(x_L(\lambda))}$
- Hessien de w : $\nabla^2 w(\lambda) = \nabla_\lambda c(x_L(\lambda)) = \nabla x_L(\lambda) \nabla c(x_L(\lambda))$

  $\Rightarrow \boxed{\nabla^2 w(\lambda) = -\nabla c(x_L(\lambda))^T \left[\nabla^2_{xx} L(x_L(\lambda),\lambda)\right]^{-1} \nabla c(x_L(\lambda))}$

Preuve : on calcule $\nabla x_L(\lambda)$ à partir de $\nabla_x L(x_L(\lambda),\lambda)=0$

$$\nabla_x L(x_L(\lambda),\lambda) = 0, \forall\lambda \Rightarrow \nabla_\lambda\left[\nabla_x L(x_L(\lambda),\lambda)\right] = 0 \Rightarrow \nabla x_L(\lambda)\nabla^2_{xx} L(x_L(\lambda),\lambda) + \nabla^2_{\lambda x} L(x_L(\lambda),\lambda) = 0$$

$$\Rightarrow \nabla x_L = -\nabla c^T \left[\nabla^2_{xx} L\right]^{-1} \qquad \text{avec } \nabla_\lambda L = c$$

Méthode de maximisation

- La fonction duale est **coûteuse** à évaluer.

  $w(\lambda) = \min_{x\in R^n} L(x,\lambda) = L(x_L(\lambda),\lambda)$ $\rightarrow$ minimisation en chaque valeur λ pour obtenir $x_L(\lambda)$

  $\rightarrow$ on ne peut pas directement maximiser w

- On réalise à chaque itération :
  - une minimisation par rapport à x : $(x_k, \lambda_k) \rightarrow (x_{k+1},\lambda_k)$
  - une maximisation par rapport à λ : $(x_{k+1},\lambda_k) \rightarrow (x_{k+1},\lambda_{k+1})$

Déplacement

- Le déplacement sur x est construit en minimisant $L(x,\lambda_k)$.

  $\min_{x\in R^n} L(x,\lambda_k) \rightarrow x_{k+1} = x_L(\lambda_k)$ $\rightarrow$ minimisation exacte ou approchée

- Le déplacement sur λ est construit à partir du gradient et éventuellement du hessien de w.

  $$\begin{cases} \nabla w(\lambda_k) = c(x_{k+1}) \\ \nabla^2 w(\lambda_k) = -\nabla c(x_{k+1})^T \left[\nabla^2_{xx} L(x_{k+1},\lambda_k)\right]^{-1} \nabla c(x_{k+1}) \end{cases}$$ $\rightarrow$ minimisation approchée (**pas fixé**)

- Méthodes de déplacement sur λ : Uzawa, lagrangien augmenté, Newton

c) Méthode d'Uzawa

Principe

Le déplacement en λ est construit par une **méthode de plus forte pente**.

$$\lambda_{k+1} = \lambda_k + s_k \nabla w(\lambda_k) \Rightarrow \boxed{\lambda_{k+1} = \lambda_k + s_k c(x_{k+1})}$$

L'efficacité de la méthode dépend du réglage du pas s_k.

Méthode d'Uzawa (1958)

- On choisit un **pas constant** : $s_k = s_0$.
- Difficulté : - pas s_0 trop grand $\rightarrow$ oscillations autour de la solution
  - pas s_0 trop petit $\rightarrow$ convergence très lente
- La convergence est linéaire (méthode de plus forte pente).
- La vitesse de convergence dépend du conditionnement du hessien de la fonction duale.

$$\nabla^2 w(\lambda^*) = -\nabla c(x^*)^T \left[\nabla^2_{xx} L(x^*, \lambda^*)\right]^{-1} \nabla c(x^*)$$

Améliorations

- On peut régler le pas à chaque itération par une recherche linéaire suivant $\nabla w(\lambda_k)$.
- On peut appliquer une **méthode de Newton** au problème dual en utilisant $\nabla^2 w(\lambda_k)$.
  $\rightarrow$ méthodes nécessitant une globalisation pour contrôler la convergence
  $\rightarrow$ évaluations coûteuses de $w(\lambda)$

d) Méthode de Newton

Principe

L'itération de Newton sur le problème dual est définie par :

$$\lambda_{k+1} = \lambda_k - \left[\nabla^2 w(\lambda_k)\right]^{-1} \nabla w(\lambda_k) \quad \text{avec} \quad \begin{cases} \nabla w(\lambda_k) = c(x_{k+1}) \\ \nabla^2 w(\lambda_k) = -\nabla c(x_{k+1})^T \left[\nabla^2_{xx} L(x_{k+1}, \lambda_k)\right]^{-1} \nabla c(x_{k+1}) \end{cases}$$

- On applique une **méthode de quasi-Newton** au problème primal : $\min\limits_{x \in R^n} L(x, \lambda_k)$

  pour obtenir une approximation de l'inverse du hessien de L : $H_k \approx \nabla^2_{xx} L(x_{k+1}, \lambda_k)^{-1}$

- On applique ensuite une **méthode de Newton** au problème dual.

$$\lambda_{k+1} = \lambda_k - \left[\nabla^2 w(\lambda_k)\right]^{-1} \nabla w(\lambda_k) \Rightarrow \boxed{\lambda_{k+1} = \lambda_k + \left[\nabla c(x_{k+1})^T H_k \nabla c(x_{k+1})\right]^{-1} c(x_{k+1})}$$

Convergence

- La convergence est superlinéaire (quasi-Newton) ou quadratique (Newton).
- Une globalisation est nécessaire pour vérifier la solution de Newton.
  $\rightarrow$ évaluations supplémentaires de $w(\lambda)$ pouvant être coûteuses
- Le point col n'existe pas nécessairement si le problème n'est pas convexe (saut de dualité)
  $\rightarrow$ création d'un point col par pénalisation

Exemple

- Minimisation de $f(x) = x_1 + x_2$ sous $c(x) = x_1^2 + (x_2 - 1)^2 - 2 = 0$
- Lagrangien : $L(x,\lambda) = x_1 + x_2 + \lambda\left(x_1^2 + (x_2 - 1)^2 - 2\right)$
- Fonction duale : $w(\lambda) = \min_{x \in R^2} L(x,\lambda) = 1 - 2\lambda - \dfrac{1}{2\lambda}$ pour $x_L(\lambda) = \begin{pmatrix} -\dfrac{1}{2\lambda} \\ -\dfrac{1}{2\lambda} + 1 \end{pmatrix}$
- Problème dual : $\max_{\lambda \in R} w(\lambda) \Rightarrow \boxed{\lambda^* = \dfrac{1}{2} \rightarrow x^* = \begin{pmatrix} -1 \\ 0 \end{pmatrix}}$ $\rightarrow$ **point col**
- Gradient : $\nabla w(\lambda) = c\left(x_L(\lambda)\right) = \dfrac{1}{2\lambda^2} - 2$
- Hessien : $\nabla^2 w(\lambda) = -\nabla c\left(x_L(\lambda)\right)^T \left[\nabla^2_{xx} L\left(x_L(\lambda),\lambda\right)\right]^{-1} \nabla c\left(x_L(\lambda)\right) = -\dfrac{1}{\lambda^3}$

avec $\nabla c\left(x_L(\lambda)\right) = -\dfrac{1}{\lambda}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\nabla^2_{xx} L\left(x_L(\lambda),\lambda\right) = 2\lambda\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

Méthode d'Uzawa

- Minimisation de $f(x) = x_1 + x_2$ sous $c(x) = x_1^2 + (x_2 - 1)^2 - 2 = 0$
- Itération à pas s fixé : $x_{k+1} = \begin{pmatrix} -\dfrac{1}{2\lambda_k} \\ -\dfrac{1}{2\lambda_k} + 1 \end{pmatrix}$ $\lambda_{k+1} = \lambda_k + s\left(\dfrac{1}{2\lambda_k^2} - 2\right)$
- **Influence du pas** : s=0,1 ou s=0,2 $\rightarrow$ oscillations autour de la solution

Uzawa pas s = 0,1

Itération	x_1	x_2	λ	s
1	-0,10000	1,00000	1,00000	0,1
2	-0,50000	0,50000	0,85000	0,1
3	-0,58824	0,41176	0,71920	0,1
4	-0,69521	0,30479	0,61587	0,1
5	-0,81186	0,18814	0,54769	0,1
6	-0,91292	0,08708	0,51438	0,1
7	-0,97205	0,02795	0,50335	0,1
8	-0,99334	0,00666	0,50070	0,1
9	-0,99861	0,00139	0,50014	0,1
10	-0,99972	0,00028	0,50003	0,1
11	-0,99994	5,63E-05	0,50001	0,1
12	-0,99999	1,13E-05	0,50000	0,1

Uzawa pas s = 0,2

Itération	x_1	x_2	λ	s
1	-0,10000	1,00000	1,00000	0,2
2	-0,50000	0,50000	0,70000	0,2
3	-0,71429	0,28571	0,50408	0,2
4	-0,99190	0,00810	0,49763	0,2
5	-1,00476	-0,00476	0,50145	0,2
6	-0,99711	0,00289	0,49914	0,2
7	-1,00172	-0,00172	0,50052	0,2
8	-0,99896	0,00104	0,49969	0,2
9	-1,00062	-0,00062	0,50019	0,2
10	-0,99963	0,00037	0,49989	0,2
11	-1,00022	-2,24E-04	0,50007	0,2
12	-0,99987	1,34E-04	0,49996	0,2

Comparaison Uzawa / Newton

- Minimisation de $f(x) = x_1 + x_2$ sous $c(x) = x_1^2 + (x_2 - 1)^2 - 2 = 0$

- Itération k : $x_{k+1} = \begin{pmatrix} -\dfrac{1}{2\lambda_k} \\ -\dfrac{1}{2\lambda_k} + 1 \end{pmatrix}$

$$\lambda_{k+1} = \lambda_k + s\left(\frac{1}{2\lambda_k^2} - 2\right) \rightarrow \textbf{Uzawa}$$

Globalisation $\rightarrow$ pas s

$$\lambda_{k+1} = \lambda_k + s\lambda_k^3\left(\frac{1}{2\lambda_k^2} - 2\right) \rightarrow \textbf{Newton}$$

Uzawa pas s = 0,1

Itération	x_1	x_2	λ	s
1	-0,10000	1,00000	1,00000	0,1
2	-0,50000	0,50000	0,85000	0,1
3	-0,58824	0,41176	0,71920	0,1
4	-0,69521	0,30479	0,61587	0,1
5	-0,81186	0,18814	0,54769	0,1
6	-0,91292	0,08708	0,51438	0,1
7	-0,97205	0,02795	0,50335	0,1
8	-0,99334	0,00666	0,50070	0,1
9	-0,99861	0,00139	0,50014	0,1
10	-0,99972	0,00028	0,50003	0,1
11	-0,99994	5,63E-05	0,50001	0,1
12	-0,99999	1,13E-05	0,50000	0,1

Newton pas s = 1

Itération	x_1	x_2	λ	s
1	-0,10000	1,00000	1,00000	0,25
2	-0,50000	0,50000	0,62500	0,50
3	-0,80000	0,20000	0,53711	1,00
4	-0,93091	0,06909	0,49577	1,00
5	-1,00854	-0,00854	0,49995	1,00
6	-1,00011	-0,00011	0,50000	1,00
7	-1,00000	-1,72E-08	0,50000	1,00
8	-1,00000	0,00E+00	0,50000	1,00

627

e) Méthode de lagrangien augmenté

Principe

La méthode de lagrangien augmenté est une **méthode duale appliquée au critère augmenté**.

- Critère augmenté f_ρ : $f_\rho(x) = f(x) + \frac{1}{2}\rho\|c(x)\|_2^2$
- Lagrangien augmenté L_ρ : $L_\rho(x,\lambda) = f(x) + \lambda^T c(x) + \frac{1}{2}\rho\|c(x)\|_2^2$
- **Fonction duale augmentée w_ρ** : $w_\rho(\lambda) = \min\limits_{x\in R^n} L_\rho(x,\lambda)$

Le problème dual devient : $\max\limits_{\lambda\in R^m} w_\rho(\lambda)$

Fonction duale augmentée

Le gradient et le hessien de w_ρ sont identiques à ceux de w en remplaçant L par L_ρ.

$$\begin{cases} \nabla w_\rho(\lambda) = c(x_{L\rho}(\lambda)) \\ \nabla^2 w_\rho(\lambda) = -\nabla c(x_{L\rho}(\lambda))^T \left[\nabla^2_{xx} L_\rho(x_{L\rho}(\lambda),\lambda)\right]^{-1} \nabla c(x_{L\rho}(\lambda)) \end{cases} \quad \text{avec} \quad \min_{x\in R^n} L_\rho(x,\lambda) \ \text{en} \ x = x_{L\rho}(\lambda)$$

On réalise à chaque itération :
- une minimisation de L_ρ par rapport à x
- un déplacement en λ à partir du gradient et du hessien de w_ρ

Déplacement

- Le déplacement sur x est construit en minimisant $L_\rho(x,\lambda_k)$ avec une précision donnée.

  $\min\limits_{x\in R^n} L_\rho(x,\lambda_k) \;\rightarrow\; x_{k+1} = x_{L\rho}(\lambda_k)$
- Le déplacement sur λ est construit par une méthode de plus forte pente.

  $\lambda_{k+1} = \lambda_k + s_k \nabla w_\rho(\lambda_k) \;\Rightarrow\; \lambda_{k+1} = \lambda_k + s_k c(x_{k+1})$

  Le pas s_k est choisi à partir de la propriété de convergence : $\lim\limits_{k\to\infty} \lambda_k + \rho_k c(x_k) = \lambda^*$

  $s_k = \rho_k \quad \rightarrow \boxed{\lambda_{k+1} = \lambda_k + \rho_k c(x_{k+1})}$

Convergence

- La convergence est linéaire (méthode de plus forte pente).
- La vitesse de convergence dépend du conditionnement des hessiens de L_ρ (primal) et w_ρ (dual).

  $\nabla^2_{xx} L_\rho = \nabla^2_{xx} L + \rho \nabla c . \nabla c^T \xrightarrow[\rho\to\infty]{} \rho \nabla c . \nabla c^T \quad \rightarrow$ mal conditionné si ρ grand

  $\nabla^2 w_\rho = -\nabla c^T \left[\nabla^2_{xx} L_\rho\right]^{-1} \nabla c \xrightarrow[\rho\to\infty]{} -\rho^{-1} I \quad \rightarrow$ bien conditionné si ρ grand
- On peut appliquer une méthode de Newton au dual avec $H_k \approx \left[\nabla^2_{xx} L_\rho\right]^{-1}$ (quasi-Newton primal)

  $\rightarrow$ convergence superlinéaire + globalisation nécessaire
- La pénalisation peut **créer un point col** qui n'existe pas dans le problème initial.

f) Exemple

Exemple

- Minimisation de $f(x) = -x_1 x_2$ sous $c(x) = x_1 + x_2 - 1 = 0$
- Lagrangien : $L(x,\lambda) = -x_1 x_2 + \lambda(x_1 + x_2 - 1)$
- Solution KKT : $\boxed{\lambda^* = \frac{1}{2} \;,\; x_1^* = x_2^* = \frac{1}{2}} \rightarrow L(x^*,\lambda^*) = -\frac{1}{4}$
- Fonction duale : $w(\lambda) = \min\limits_{x\in R^2} L(x,\lambda) = -\infty$ en prenant $\begin{cases} x_1 = -\lambda \\ x_2 \to \pm\infty \text{ (signe de } -\lambda) \end{cases}$

  La fonction duale vaut $-\infty$ pour toute valeur de λ.
- Problème dual : $\max\limits_{\lambda\in R} w(\lambda) = -\infty \quad \rightarrow \; < -\frac{1}{4}$
- Le lagrangien n'admet **pas de point col** (problème non convexe).

  Le problème non pénalisé présente un **saut de dualité**.

  $\rightarrow$ On ne peut pas appliquer une méthode duale (Uzawa) sur ce problème.

  $\rightarrow$ On peut créer un point col par pénalisation (méthode de lagrangien augmenté).

Point col

- Problème pénalisé : $f_\rho(x) = -x_1x_2 + \frac{1}{2}\rho(x_1 + x_2 - 1)^2$ sous $c(x) = x_1 + x_2 - 1 = 0$
- Lagrangien augmenté : $L_\rho(x,\lambda) = -x_1x_2 + \frac{1}{2}\rho(x_1 + x_2 - 1)^2 + \lambda(x_1 + x_2 - 1)$
- Fonction duale : $w_\rho(\lambda) = \min\limits_{x \in R^2} L_\rho(x,\lambda)$
- Minimisation de L_ρ par rapport à x

Ordre 1 : $\nabla_x L_\rho = 0 \;\Rightarrow\; \begin{cases} -x_2 + \rho(x_1 + x_2 - 1) + \lambda = 0 \\ -x_1 + \rho(x_1 + x_2 - 1) + \lambda = 0 \end{cases} \;\Rightarrow\; \boxed{x_1(\lambda) = x_2(\lambda) = \frac{\lambda - \rho}{1 - 2\rho}}$

Ordre 2 : $\nabla^2_{xx} L_\rho > 0 \;\Rightarrow\; \begin{pmatrix} \rho & \rho - 1 \\ \rho - 1 & \rho \end{pmatrix} > 0$

Valeurs propres : $(\rho - \sigma)^2 - (\rho - 1)^2 = 0 \;\Rightarrow\; \begin{cases} \sigma_1 = 1 \\ \sigma_2 = 2\rho - 1 \end{cases} \;\rightarrow\; \boxed{L_\rho \text{ convexe si } \rho > \frac{1}{2}}$

- On obtient si $\rho > \frac{1}{2}$: $w_\rho(\lambda) = -\left(\frac{\lambda - \rho}{1 - 2\rho}\right)^2 + \frac{1}{2}\rho\left(\frac{2\lambda - 1}{1 - 2\rho}\right)^2 + \lambda\frac{2\lambda - 1}{1 - 2\rho}$

Point col

- Fonction duale : $w_\rho(\lambda) = \frac{1}{(1 - 2\rho)^2}\left[-(\lambda - \rho)^2 + \frac{1}{2}\rho(2\lambda - 1)^2 + (1 - 2\rho)(2\lambda^2 - \lambda)\right]$
- Problème dual : $\max\limits_{\lambda \in R} w_\rho(\lambda)$

$$\nabla w_\rho = 0 \;\Rightarrow\; -2(\lambda - \rho) + 2\rho(2\lambda - 1) + (1 - 2\rho)(4\lambda - 1) = 0$$

Solution : $\lambda^* = \frac{1}{2} \;\Rightarrow\; w_\rho(\lambda^*) = -\frac{1}{4}$ et $x_1(\lambda^*) = x_2(\lambda^*) = \frac{1}{2}$

$$w_\rho(\lambda^*) = L_\rho(x^*,\lambda^*) = -\frac{1}{4}$$

On retrouve la solution du problème primal **sans saut de dualité** $\rightarrow$ point col

- La pénalisation permet de **créer un point col** pour $\rho > \frac{1}{2}$ en rendant le problème convexe.

$\rightarrow$ On peut appliquer une méthode duale au problème pénalisé ($\neq$ problème initial).
$\rightarrow$ La méthode de lagrangien augmenté élargit le domaine d'application des méthodes duales.

6.4 Algorithme

a) Algorithme de lagrangien augmenté

Méthode de lagrangien augmenté (ou méthode des multiplicateurs)

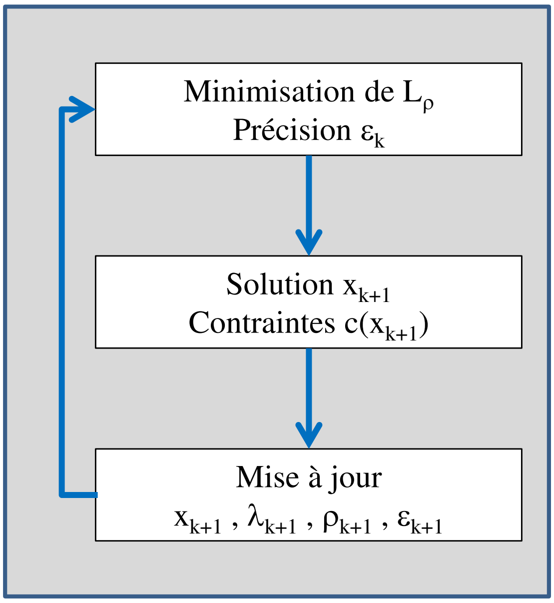
Les principales étapes d'une itération de lagrangien augmenté sont
- minimiser le lagrangien augmenté
- mettre à jour les paramètres de réglage

Minimisation du lagrangien augmenté
- Méthode de quasi-Newton
- Recherche linéaire ou région de confiance
- **Précision d'arrêt sur gradient**

Paramètres de réglage
- Multiplicateurs
- Pénalisation
- Précisions (gradient, contraintes)

Principales difficultés
- Précision contraintes $\rightarrow$ pénalisation forte
- Conditionnement $\rightarrow$ pénalisation faible

$\rightarrow$ convergence précise difficile
$\rightarrow$ réglages à adapter au cas par cas



Méthode de lagrangien augmenté (ou méthode des multiplicateurs)

- Réglages à l'itération k :
  - multiplicateurs λ_k
  - pénalisation ρ_k
  - précision de résolution ε_k
  - précision sur les contraintes η_k

- Minimisation : $\boxed{\min_{x \in R^n} L_{\rho_k}(x, \lambda_k) \rightarrow \textbf{critère d'arrêt} \quad \left\| \nabla_x L_{\rho_k}(x_{k+1}, \lambda_k) \right\| \le \varepsilon_k}$

  initialisation x_k $\rightarrow$ solution x_{k+1} $\rightarrow$ valeurs des contraintes = $c(x_{k+1})$

- Mise à jour des réglages à l'itération k+1 en fonction du respect des contraintes

  - Si $\left\| c(x_{k+1}) \right\| < \eta_k$ $\rightarrow$ **mise à jour des multiplicateurs** $\lambda_{k+1} = \lambda_k + \rho_k c(x_k)$
    (contraintes bien respectées) $\rightarrow$ résolution plus précise $\varepsilon_{k+1} < \varepsilon_k$, $\eta_{k+1} < \eta_k$

  - Si $\left\| c(x_{k+1}) \right\| > \eta_k$ $\rightarrow$ **augmentation de la pénalisation** $\rho_{k+1} > \rho_k \rightarrow \times 10$
    (contraintes mal respectées) $\rightarrow$ résolution moins précise $\varepsilon_{k+1} > \varepsilon_k$, $\eta_{k+1} > \eta_k$

Estimation des multiplicateurs

- La solution (x^*,λ^*) doit vérifier la condition KKT d'ordre 1.

$$\nabla_x L(x^*,\lambda^*) = 0 \quad \Leftrightarrow \quad \nabla f(x^*) + \nabla c(x^*)\lambda^* = 0$$

- On cherche au point x_0 le multiplicateur qui approche « au mieux » la condition KKT.

$$\min_\lambda \left\| \nabla_x L(x_0,\lambda) \right\|^2 \quad \Leftrightarrow \quad \min_\lambda \left\| \nabla c(x_0)\lambda + \nabla f(x_0) \right\|^2$$

- Il s'agit d'un problème de moindres carrés de la forme : $\min_\lambda \left\| A\lambda - b \right\|^2$ avec $\begin{cases} A = \nabla c(x_0) \\ b = -\nabla f(x_0) \end{cases}$

  La solution λ_{MC} vérifie les équations normales : $A^T A \lambda_{MC} = A^T b$

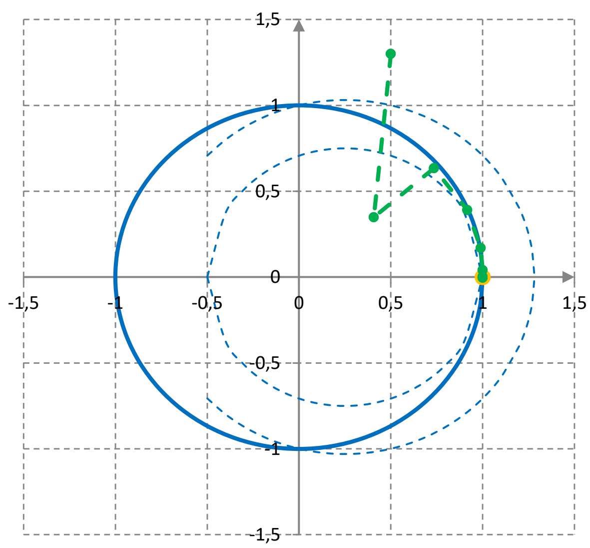
$$\boxed{\lambda_{MC} = -\left[\nabla c(x_0)^T \nabla c(x_0)\right]^{-1} \nabla c(x_0)^T \nabla f(x_0)} \quad \rightarrow \text{ multiplicateurs « des moindres carrés »}$$

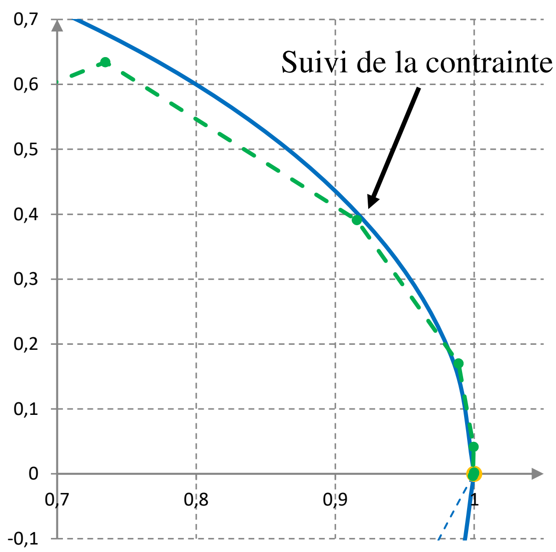
- L'estimation des multiplicateurs par moindres carrés est utile :
  - à la première itération pour **initialiser λ_0**
  - au cours des itérations pour réinitialiser λ_k (si nécessaire).

b) Exemple

Exemple

- Minimisation de $f(x) = 2(x_1^2 + x_2^2 - 1) - x_1$ sous $c(x) = x_1^2 + x_2^2 - 1 = 0$
- Point initial : $x = \begin{pmatrix} 0.5 \\ 1.3 \end{pmatrix}$, $\lambda = 0$ Solution : $x^* = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\lambda = -\dfrac{3}{2}$





Exemple

- Minimisation de $f(x)=2(x_1^2+x_2^2-1)-x_1$ sous $c(x)=x_1^2+x_2^2-1=0$
- Point initial : $x=\begin{pmatrix}0.5\\1.3\end{pmatrix}$, $\lambda=0$ Solution : $x^*=\begin{pmatrix}1\\0\end{pmatrix}$, $\lambda^*=-\dfrac{3}{2}$

Itération	x_1	x_2	λ	ρ	$c(x)$	$\lVert\nabla L_\rho(x,\lambda)\rVert$	Newton
1	**0,50000**	**1,30000**	**0,00000**	1	-0,71238	0,90050	1
2	0,40707	0,34917	-0,71238	10	-0,05788	0,90016	1
3	0,73467	0,63433	-1,29122	10	-0,00905	0,50091	2
4	0,91556	0,39077	-1,38175	10	0,00635	0,41807	2
5	0,98869	0,16985	-1,38175	100	0,00188	0,62061	2
6	0,99953	0,04158	-1,30283	100	-0,00188	0,01728	2
7	0,99905	-0,00320	-1,49103	100	-0,00009	0,00172	1
8	0,99995	0,00171	-1,50003	100	2,06E-06	0,00057	3
9	**1,00000**	**0,00045**	**-1,50003**	100	1,85E-06	0,00031	

pénalisation	précision	précision	itérations Newton