Table des matières

I Conditions d'optimalité 2

1 Existence d'un minimum 2

2 Conditions nécessaires d'optimalité 4

2.1 Plusieurs notions de dérivabilité . . . 4
2.2 Quelques rappels d'analyse convexe . . . 5
2.3 Conditions d'optimalité dans un ouvert . . . 9
2.4 Théorème de Kuhn et Tucker . . . 10
2.5 Cas des contraintes qualifiées . . . 12

3 Problèmes convexes et dualité 13

3.1 Dualité . . . 15

II Programmation linéaire, algotihme du simplexe 18

1 Introduction 18

2 Solutions de base d'un problème sous forme standard 19

3 Algorithme du simplexe 21

3.1 Privot à partir d'une base réalisable : critère de Dantzig . . . 23
3.2 Détermination d'une première base réalisable . . . 26

4 Dualité en programmation linéaire 28

Première partie
Conditions d'optimalité

On considère le problème d'optimisation suivant :

$$\inf_{u\in\mathcal{U}_{ab}} J(u)\ (\mathcal{P}_{\mathcal{U}_{ab}})$$

Ici, $J : \mathbb{R}^d \to \mathbb{R}$ et $\mathcal{U}_{ab}$ est un sous-ensemble non vide de $\mathbb{R}^d$.
J s'appelle la fonction coût.
$\mathcal{U}_{ab}$ s'appelle l'ensemble admissible.

1 Existence d'un minimum

Définition:

Soit $J : \mathbb{R}^d \to \mathbb{R}$ et $\mathcal{U}_{ad}$ un sous-ensemble non vide.
On dit que $l \in [-\infty, +\infty[$ est l'infimum de J sur $\mathcal{U}_{ad}$ si :

1. $J(u) \geq l\ \forall u \in \mathcal{U}_{ad}$
2. $\exists (u_n)_n \subset \mathbb{R}^d;\ u_n \in \mathcal{U}_{ad}$ et $J(u_n) \to l$

On note $l = \inf_{u\in\mathcal{U}_{ad}} J(u)$ et les suites vérifiant $(u_n)_n \subset \mathcal{U}_{ad}$ et $J(u_n) \xrightarrow[n\to+\infty]{} \inf_{\mathcal{U}_{ad}} J(u)$ sont appelées suites minimisantes.

Remarque : L'infimum existe toujours. Il est fini si et seulement si J est minorée.

Définition:

Soit $J : \mathbb{R}^d \to \mathbb{R}$ et $\mathcal{U}_{ab}$ un sous-ensemble non vide.
On dit que $l \in \mathbb{R}$ est le minimum de J sur $\mathcal{U}_{ad}$ (si cette valeur existe) si on a :

1. $J(u) \geq l\ \forall u \in \mathcal{U}_{ad}$
2. $\exists \bar{u} \in \mathcal{U}_{ad}$ tel que $J(\bar{u}) = l$

On dit alors que J atteint son minimum sur $\mathcal{U}_{ad}$ et on note $l = \min_{u\in\mathcal{U}_{ad}} J(u)$

Remarque :

1. Le minimum n'existe pas toujours
2. Par abus de langage, on appelle aussi minimum le point $\bar{u}$ qui vérifie $J(u) \geq J(\bar{u})\ \forall u \in \mathcal{U}_{ad}$ ($\bar{u}$ est l'argument du minimum).

Définition:

On dit que $J : \mathbb{R}^d \to \mathbb{R}$ est coercive si :

$$\lim_{\|u\|\to+\infty} = +\infty$$

Remarque : En dimension finie, toutes les normes sont équivalentes. Il suffit donc de le vérifier pour la norme la plus "facile"

☕Exemple :

1. Soit A une matrice symétrique de taille d, $b \in \mathbb{R}^d$ et c un réel.
   On considère l'application :
   $$\begin{array}{rcl} J : \mathbb{R}^d & \to & \mathbb{R} \\ u & \mapsto & \langle Au, u \rangle + \langle b, u \rangle + c \end{array}$$
   J est coercive si et seulement si A est définie positive.
   Si A est symétrique, on a :
   $$\lambda_{min} \|u\|^2 \leq \langle Au, u \rangle \leq \lambda_{max} \|u\|^2$$
   où λ_{min} est la plus petite valeur propre, et λ_{max} est la plus grande valeur propre.
2. Toute fonction minorée par une fonction coercive est coercive.

ℹPropriété:

On suppose que $J : \mathbb{R}^d \to \mathbb{R}$ prend la forme :
$$\forall u = (u_1, ..., u_d) \in \mathbb{R}^d, J(u) = \sum_i J_i(u_i)$$
avec $J_i : \mathbb{R} \to \mathbb{R}$ coercive et minorée. Alors J est coercive.

Démonstration :
$\forall i \in \{1, ..., d\}$, J_i est minorée par une constante m_i.
On note $m = \max_{i \in \{1,...,d\}} |m_i|$.

Soit $M > 0$ fixé. $\forall i \in \{1, ..., d\}$, $\exists R_i > 0;\ \forall |u_i| > R_i$, on a $J_i(u_i) > M + md$ (car J_i coercive).
Posons $R = \max_{i \in \{1,...,d\}} R_i$

Soit $u \in \mathbb{R}^d$; $\|u\|_\infty \geq R$.
$$\Rightarrow \exists j \in \{1, ..., d\};\ |u_j| \geq R \geq R_j$$
On a donc $J_j(u_j) \geq M + md$
Ainsi :
$$\begin{array}{rcl} J(u) & = & \sum_i J_i(u_i) \\ & = & J_j(u_j) + \sum_{i \neq j} J_i(u_i) \\ & \geq & M + md + \sum_{j \neq j} m_i \\ & \geq & M + \sum_{i \neq j} (m + m_i) \\ & \geq & M \end{array}$$
On a montré que :
$$\forall M > 0, \exists R; \forall u;\ \|u\|_\infty \geq R \Rightarrow J(u) \geq M$$
ie $\lim_{\|u\|_\infty \to \infty} J(u) = +\infty$
Par conséquent, J est coercive.

Théorème:

Soit $J : \mathbb{R}^d \to \mathbb{R}$ une fonction continue et $\mathcal{U}_{ad}$ un ensemble fermé non vide. On suppose que :
- Soit J est coercive
- Soit $\mathcal{U}_{ad}$ est borné

Alors J atteint son minimum sur $\mathcal{U}_{ad}$.

Propriété:

Soit $\mathcal{U}_{ad}$ un ouvert fermé de $\mathbb{R}^d$ et soit $J : \mathbb{R}^d \to \mathbb{R}$ une fonction continue.
On suppose qu'il existe $u_0 \in \mathcal{U}_{ad}$ tel que :

$$J(u_0) < J(u) \ \forall u \in \partial\mathcal{U}_{ad}$$

où $\partial\mathcal{U}_{ad}$ est la frontière de $\mathcal{U}_{ad}$.
Alors J atteint son minimum sur $\mathcal{U}_{ad}$.

Démonstration :
$\overline{\mathcal{U}_{ad}}$ est un compact et J est continue, donc :

$$\exists \bar{u} \in \overline{\mathcal{U}_{ad}}; \ J(\bar{u}) \leq J(u) \forall u \in \overline{\mathcal{U}_{ad}}$$

Montrons que $\bar{u} \in \mathcal{U}_{ad}$.
Si $\bar{u} \notin \mathcal{U}_{ad}$, alors $\bar{u} \in \partial\mathcal{U}_{ad}$ et par hypothèse, on a $J(\bar{u}) > J(u_0)$ avec $u_0 \in \mathcal{U}_{ad}$. Contradiction.
Donc $\bar{u} \in \mathcal{U}_{ad}$ et $J(\bar{u}) \leq J(u) \forall u \in \mathcal{U}_{ad}$.

2 Conditions nécessaires d'optimalité

2.1 Plusieurs notions de dérivabilité

Définition: Dérivées directionnelle

Soit H_1 et H_2 deux espaces de Hilbert et $f : H_1 \to H_2$.
On appelle dérivée directionnelle de f au point $x \in H_1$ dans la direction $d \in H_1$, notée $f'(x, d)$, la limite (si elle existe) :

$$f'(x, d) = \lim_{\varepsilon \to 0^+} \frac{f(x + \varepsilon d) - f(x)}{\varepsilon}$$

Définition: Gâteaux-différentiabilité

On dit que f est Gâteaux-différentiable en $x \in H_1$ si f admet des dérivées directionnelles au point x dans toutes les directions et si l'application

$$d \in H_1 \mapsto f'(x, d)$$

est linéaire continue.

On note alors $f'(x)$ cette application :

$$f'(x,d) = f'(x)d \ \forall d \in H_1$$

On dit que f est Gâteaux-différentiable en tout point $x \in H_1$.

Définition:

Soit $f : H \to \mathbb{R}$ une fonction Gâteaux-différentiable en $x \in H$.
On note $\nabla f(x)$ (appelé gradient de f au point x) l'unique élément de H tel que

$$f'(x)d = \langle \nabla f(x), d \rangle$$

Définition:

On dit que $f : H_1 \to H_2$ est Fréchet-différentiable en x s'il existe une application linéaire continue de H_1 dans H_2 tel que :

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - Lh}{\|h\|} = 0$$

L'opérateur L est appelé la dérivée de f en a.

Propriété:

Soient H_1 et H_2 deux espaces de Hilbert et $f : H_1 \to H_2$. On suppose que f est Fréchet-différentiable en $x \in H_1$ avec une dérivée L. Alors f est Gâteaux-différentiable et $L = f'(x)$.

Démonstration :
Soit $d \neq 0 \in H_1$. $\forall h > 0$:

$$\begin{array}{rcl} \frac{f(x+hd)-f(x)-Lhd}{h\|d\|} & \to & 0 \\ \Rightarrow \quad \frac{f(x+h)-f(x)}{h} - Ld & \to & 0 \\ \Rightarrow \quad \frac{f(x+h)-f(x)}{h} & \to & Ld \end{array}$$

f admet une dérivée directionnelle dans la direction d, et on a $f'(x,d) = Ld$, ie $f'(x, \bullet) = L$ est linéaire continue.

2.2 Quelques rappels d'analyse convexe

Définition:

Soit C un sous ensemble d'un espace vectoriel. On dit que C est convexe si

$$\forall x, y \in C, \forall \alpha \in [0,1], \alpha x + (1-\alpha)y \in C$$

Définition:

Soit $f : H \to \mathbb{R}$. On appelle épigraphe de f, noté $\text{epi}(f)$, l'ensemble :

$$\text{epi}(f) = \{(\alpha, x) \in \mathbb{R} \times H ; \alpha \geq f(x)\}$$

Définition:

Soit $f : H \to \mathbb{R}$. On dit que f est convexe si $\text{epi}(f)$ est convexe.

Propriété:

$f : H \to \mathbb{R}$ est convexe si et seulement si $\forall x, y \in H$, $\forall \alpha \in [0,1]$, on a :

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y)$$

Démonstration :
On suppose que f est convexe. Soient $x, y \in H$ et $\alpha \in [0,1]$.
On a $(f(x), x)$ et $(f(y), y) \in \text{epi}(f)$. Donc $\alpha(f(x), x) + (1-\alpha)(f(y), y) \in \text{epi}(f)$.

$$\Rightarrow f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y)$$

Réciproquement, on suppose $\forall x, y \in H$, $\forall \alpha \in [0,1]$,

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y)$$

Soient (α, x) et $(\beta, y) \in \text{epi}(f)$ et $\lambda \in [0,1]$.

$$\alpha \geq f(x) \text{ et } \beta \geq f(y)$$

On a $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) \leq \lambda\alpha + (1-\lambda)\beta$

$$\Rightarrow (\lambda\alpha + (1-\lambda)\beta, \lambda x + (1-\lambda)y) \in \text{epi}(f)$$

Donc $\text{epi}(f)$ est convexe.

Définition:

Soit $f : H \to \mathbb{R}$. On dit que f est strictement convexe si $\forall x, y \in H$, tel que $x \neq y$, $\forall \alpha \in [0,1]$, on a :

$$f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha) f(y)$$

Définition:

Soit $f : H \to \mathbb{R}$. On dit que f est α-convexe si $\forall x, y \in H,\ \forall \lambda \in [0,1]$, on a :

$$\frac{\alpha}{2}\lambda(1-\lambda)\|x-y\|^2 + f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$$

Propriété: Convexité et dérivée première

Soit $f : H \to \mathbb{R}$ différentiable. On a équivalence entre les propositions suivantes :

1. f est convexe
2. $f(y) \geq f(x) + f'(x, y-x)$
3. $(f'(y) - f'(x))(y-x) \geq 0$

Propriété: α-convexité et dérivée première

Soit $f : H \to \mathbb{R}$ différentiable. On a équivalence entre les propositions suivantes :

1. f est α-convexe
2. $f(y) \geq f(x) + f'(x).(y-x) + \frac{\alpha}{2}\|x-y\|^2$
3. $(f'(y) - f'(x)).(y-x) \geq \alpha\|x-y\|^2$

Démonstration :
$(1) \to (2)$
On suppose que f est α-convexe. Par définition, on a, pour $\lambda = \frac{1}{2^k}$:

$$f\left(\left(1-\frac{1}{2^k}\right)x + \frac{1}{2^k}y\right) \leq \left(1-\frac{1}{2^k}\right)f(x) + \frac{1}{2^k}f(y) - \frac{\alpha}{2^{k+1}}\left(1-\frac{1}{2^k}\right)\|x-y\|^2$$

On a alors :

$$2^k\left[f\left(\left(1-\frac{1}{2^k}\right)x + \frac{1}{2^k}y\right) - f(x)\right] \leq f(y) - f(x) - \frac{\alpha}{2}\left(1-\frac{1}{2^k}\right)\|x-y\|^2$$

Lorsque $k \to +\infty$:

$$f'(x).(y-x) \leq f(y) - f(x) - \frac{\alpha}{2}\|x-y\|^2$$

$(2) \to (3)$

$$f(y) \geq f(x) + f'(x)(y-x) - \frac{\alpha}{2}\|x-y\|^2$$

$$f(x) \geq f(y) + f'(y)(x-y) - \frac{\alpha}{2}\|x-y\|^2$$

On fait la somme :

$$0 \geq (-f'(x) + f'(y)).(x-y) + \alpha\|x-y\|^2$$

(3) $\rightarrow$ (1)
Soient $x, y \in H$ et $\lambda \in [0,1]$. On introduit :

$$\begin{array}{rcl} \phi : \mathbb{R} & \rightarrow & \mathbb{R} \\ t & \mapsto & f(x + t(y-x)) \end{array}$$

ϕ est dérivable et

$$\phi'(t) = f'(x + t(y-x)).(y-x)$$

Soit $t > s$.

$$\begin{array}{rcl} \phi'(t) - \phi'(s) & = & [f'(x + t(y-x)) - f'(x + s(y-x))].(y-x) \\ & \geq & \dfrac{1}{t-s}\alpha\|(t-s)(x-y)\|^2 \\ & \geq & \alpha(t-s)\|x-y\|^2 \end{array}$$

On intègre de $t = \lambda$ à $t = 1$ et de $s = 0$ à $s = \lambda$:

$$\begin{array}{rccc} & \lambda(\phi(1) - \phi(\lambda)) - (1-\lambda)(\phi(\lambda) - \phi(0)) & \geq & \alpha\|y-x\|^2 \left[\frac{\lambda}{2}(1-\lambda^2) - \frac{\lambda^2}{2}(1-\lambda)\right] \\ \Leftrightarrow & \lambda\phi(1) + (1-\lambda)\phi(0) - \phi(\lambda) & \geq & \alpha\|x-y\|^2 \frac{\lambda}{2}(1-\lambda^2-\lambda+\lambda^2) \\ \Leftrightarrow & \lambda f(y) + (1-\lambda)f(x) & \geq & \frac{\alpha}{2}\|x-y\|^2\lambda(1-\lambda) + f((1-\lambda)x + \lambda y) \end{array}$$

Donc f α-convexe.

2.3 Conditions d'optimalité dans un ouvert

Théorème:

Soit $J : \mathbb{R}^n \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$. Soit $\mathcal{U}_{ad}$ un ouvert de $\mathbb{R}^n$.
Si J atteint un minimum local en $\bar{u} \in \mathcal{U}_{ad}$, alors

$$\nabla J(\bar{u}) = 0$$

Remarque : Il existe également une condition du second ordre (si J est $\mathcal{C}^2$) : la matrice Hessienne est positive.

Démonstration :
Soit $d \in \mathbb{R}^n$. Comme $\mathcal{U}_{ad}$ est ouvert et $\bar{u} \in \mathcal{U}_{ad}$, il existe $h_0 > 0$; $\forall h \in]0, h_0]$, $\bar{u} + hd \in \mathcal{U}_{ad}$. Donc $J(\bar{u} + hd) \geq J(\bar{u})$.
Or :

$$\begin{aligned} & J(\bar{u}) \leq J(\bar{u} + hd) = J(\bar{u}) + \langle \nabla J(\bar{u}), hd \rangle + o(h) \\ \Rightarrow \quad & \langle \nabla J(\bar{u}), hd \rangle + o(h) \geq 0 \\ \Rightarrow \quad & \langle \nabla J(\bar{u}), h \rangle + o(1) \geq 0 \\ \Rightarrow \quad & \langle J(\bar{u}), d \rangle \geq 0 \ \forall d \in \mathbb{R}^n \ (h \to 0) \end{aligned}$$

En remplaçant d par $-d$:

$$\langle \nabla J(\bar{u}), d \rangle \leq 0 \ \forall d \in \mathbb{R}^n$$

D'où :

$$\langle \nabla J(\bar{u}), d \rangle = 0 \ \forall d \in \mathbb{R}^n$$
$$\Rightarrow \nabla J(\bar{u}) = 0$$

Théorème:

Soit $\mathcal{U}_{ad}$ un ensemble convexe de $\mathbb{R}^n$ et J une application différentiable de $\mathbb{R}^n$ dans $\mathbb{R}$. Si $\bar{u}$ est un point de minimum de J sur $\mathcal{U}_{ad}$, alors :

$$J'(\bar{u})(u - \bar{u}) \geq 0 \ \forall u \in \mathcal{U} \qquad (*)$$

Réciproquement, si $\bar{u}$ vérifie $(*)$, et si J convexe, alors $\bar{u}$ est un point de minimum de J.

Démonstration :
Soit $u \in \mathcal{U}_{ad}$. On a $\forall h \in [0, 1]$, $\bar{u} + h(u - \bar{u}) = (1 - h)\bar{u} + hu \in \mathcal{U}_{ad}$ (par convexité de $\mathcal{U}_{ad}$).
Donc

$$J(\bar{u} + h(u - \bar{u})) \geq J(\bar{u})$$
$$\Rightarrow \forall u \in \mathcal{U}_{ad}, \ \forall h \in [0, 1], \ \frac{J(\bar{u} + h(u - \bar{u})) - J(\bar{u})}{h} \geq 0$$
$$\Rightarrow J'(\bar{u})(u - \bar{u}) \geq 0 \ (h \to 0^+)$$

Réciproquement, si J est convexe, et si $\bar{u}$ vérifie $(*)$, alors $\forall u \in \mathcal{U}_{ad}$:

$$0 \leq J'(\bar{u})(u - \bar{u}) \leq J(u)_J(\bar{u})$$

car J convexe, d'où :

$$J(u) \geq J(\bar{u}) \ \forall u \in \mathcal{U}_{ad}$$

Remarque:

On considère le cas où $\mathcal{U}_{ad}$ est un sous-espace affine de $\mathbb{R}^n$. En particulier, :

$$\mathcal{U}_{ad} = \mathcal{P} + \bar{u}$$

où $\mathcal{P}$ est un espace vectoriel.
La condition $(*)$ se réécrit :

$$J'(u)v \geq 0 \ \forall v \in \mathcal{P}$$

Si $v \in \mathcal{P}$ alors $-v \in \mathcal{P}$, donc :

$$J'(u).v = 0, \ \forall v \in \mathcal{P}$$

ie $\nabla J(\bar{u}) \in \mathcal{P}^{\perp}$

En particulier, si $\mathcal{P}$ est défini comme une intersection (finie) d'hyperplan ($a_i \in \mathbb{R}^n$) :

$$\mathcal{P} = \{x, \ \langle a_i, x\rangle = 0 \ \forall i = 1, ..., d\}$$

alors $\mathcal{P}^{\perp}$ est engendré par la famille $(a_i)_{i=1...d}$ La condition d'optimalité s'écrit :

$$\exists (\lambda_i)_{i=1,\ldots,d}; \ \nabla J(\bar{u}) + \sum_{i=1}^{d} \lambda_i a_i = 0$$

Les λ_i sont appelés multiplicateurs de Lagrange.

2.4 Théorème de Kuhn et Tucker

On suppose que la contrainte $\mathcal{U}_{ad}$ s'écrit :

$$\mathcal{U}_{ad} = \left\{ u \in \mathbb{R}^n; \ \begin{array}{ccc} g_i(u) & \leq & 0 & \forall i \in I \\ h_j(u) & = & 0 & \forall j \in J \end{array} \right\}$$

om $I = \{1, ..., l\}$ et $J = \{1, ..., m\}$.
On suppose que les fonctions g_i et h_j sont $\mathcal{C}^1$ et pour $u \in \mathcal{U}_{ad}$, on note $I(u)$ l'ensemble des contraintes saturées, ie :

$$I(u) = \{i \in I; \ g_i(u) = 0\}$$

Théorème:

Si $\bar{u}$ est un point de minimum local de J sur $\mathcal{U}_{ad}$ alors il exsite $p_0 \in \mathbb{R}^+$, $p \in \mathbb{R}^l_+$, $q \in \mathbb{R}^m$, tel que :

$$\left\{ \begin{array}{ccc} \sum_{i\in I} p_i g_i(\bar{u}) & = & 0 \quad \text{(condition d'exclusion)} \\ (p_0, p, q) & \neq & 0 \\ p_0 \nabla J(\bar{u}) + \sum_{i\in I} p_0 \nabla g_i(\bar{u}) + \sum_{j\in J} q_j \nabla h_j(\bar{u}) & = & 0 \quad \text{(condition nécessaire)} \end{array} \right.$$

Démonstration :
Soit $r > 0$ tel que J atteigne un minimum sur $\overline{B(\bar{u}, r)}$ en $\bar{u}$.

$$\min_{u \in \overline{B(\bar{u}, r)}} \left\{ J(u) + \|u - \bar{u}\|^2 + \frac{N}{2} \left(\sum_{i\in I} \max(g_i(u), 0)^2 + \sum_{j\in J} \max(h_j(u), 0)^2 \right) \right\}$$

Le minimum est atteint en $\bar{u}_N$.

Comme J est continue sur $\overline{B(\bar{u},r)}$, elle est bornée. On note $M = \|J\|_{L^\infty(\overline{B(\bar{u},r)})}$.

Donc :

$$\sum_{i\in I} \max(g_i(u),0)^2 + \sum_{j\in J} \max(h_j(u),0)^2 \leq \frac{2}{N}\left(J(\bar{u}) - J(\bar{u}_n) - \|\bar{u} - \bar{u}_N\|^2\right)$$

$$\sum_{i\in I} \max(g_i(\bar{u}_N),0)^2 + \sum_{j\in J} \max(h_j(\bar{u}_N),0)^2 \leq \frac{2}{N}(2M + r) \tag{$*$}$$

On a aussi :

$$J(\bar{u}_N) + \|\bar{u} - \bar{u}_N\|^2 \leq J(\bar{u}) \tag{$**$}$$

Comme $\overline{B(\bar{u},r)}$ est compacte, quitte à extraire une sous-suite, on peut supposer que $\bar{u}_N \to u^* \in \overline{B(\bar{u},r)}$. En prenant la limite dans $(*)$ et dans $(**)$.

$$\sum_{i\in I} \max(g_i(u^*),0)^2 + \sum_{j\in J} \max(h_j(u^*),0)^2 = 0$$

et

$$J(u^*) + \|\bar{u} - u^*\|^2 \leq J(\bar{u})$$

Donc $\forall i,\ g_i(u^*) \leq 0$ et $\forall j,\ h_j(u^*) \leq 0$.

$$\Rightarrow u^* \in \mathcal{U}_{ad}$$

$$\Rightarrow u^* = \bar{u}$$

Pour N assez grand, $\bar{u}_N \in B(\bar{u},r)$. On en déduit donc que :

$$\nabla J(\bar{u}_N) + 2\|\bar{u}_N - \bar{u}\| + N\left(\sum_{i\in I} \max(g_i(\bar{u}_N),0)\nabla g_i(\bar{u}_N) + \sum_{j\in J} h_j(\bar{u}_N)\nabla h_j(\bar{u}_N)\right) = 0 \tag{$*$}$$

On pose

$$\rho_N = \left[1 + N^2\sum_{i\in I} \max(0, g_i(\bar{u}_N))^2 + N^2\sum_{j\in J} h_j(\bar{u}_N)^2\right]^{\frac{1}{2}}$$

On pose :

$$p_0^N = \frac{1}{\rho_N}$$

$$p_i^N = N p_0^N \max(0, g_i(\bar{u}_N))$$

$$q_j^N = N p_0^N h_j(\bar{u}_N))$$

Le vecteur $\left(p_0^N, p^N, q^N\right) \in \mathbb{R}^{p+m+1}$ est de norme 1.

Quitte à extraire une sous-suite, on peut supposer que $(p_0^N, p^N, q^N) \to (p_0, p, q)$ avec $\|(p_0, p, q)\| = 1$.

En utilisant le fait que $\bar{u}_N \to \bar{u}$ et en divisant $(*)$ par ρ_N puis en passant à la limite, on a :

$$p_0\nabla J(\bar{u}) + \sum_{i\in I} p_i\nabla g_i(\bar{u}) + \sum_{j\in J} q_j\nabla h_j(\bar{u}) = 0$$

Il reste à montrer que :

$$\sum_{i\in I} p_i g_i(\bar{u}) = 0$$

Or, $\forall i,\ p_i g_i(\bar{u}) < 0$. Il faut donc montrer que :

$$p_i g_i(\bar{u}) = 0 \forall i$$

ie :

$$p_i = 0 \forall i \notin I(\bar{u})$$

Si $i \notin I(\bar{u}) < 0$ alors $g_i(\bar{u}) < 0$. Donc $g_i(\bar{u}_N) < 0$ pour N assez grand. Donc $p_i = 0$ pour N assez grand.

$$\Rightarrow p_i = 0 \ \forall i \notin I(\bar{u})$$

$$\Rightarrow \sum_{i \in I} p_i g_i(\bar{u}) = 0$$

Remarque :

1. Le vecteur (p_0, p, q) est appelé le multiplicateur de Lagrange généralisé associé à $\bar{u}$
2. On appelle lagrangien généralisé :

$$L(u, p_0, p, q) = p_0 J(u) + \sum_{i \in I} p_i g_i(u) + \sum_{j \in J} q_j h_j(u)$$

La condition nécessaire d'optimalité se réécrit :

$$\nabla_u L(\bar{u}, p_0, p, q) = 0$$

2.5 Cas des contraintes qualifiées

Définition: Contraintes qualifiées

On dit que les contraintes sont qualifiées en un point $\bar{u}$ de $\mathcal{U}_{ad}$ si les conditions suivantes sont vérifiées :

1. $\{\nabla h_1(\bar{u}), ..., \nabla h_n(\bar{u})\}$
2. $\exists v \in \mathbb{R}^n$ tel que :

$$\langle \nabla h_j(\bar{u}), v \rangle = 0 \ \forall j = 1, ..., n$$

et

$$\langle \nabla g_i(\bar{u}), v \rangle < 0 \ \forall i \in I(\bar{u})$$

Théorème:

Si $\bar{u}$ est un point de minimum de J sur $\mathcal{U}_{ad}$ et si les contraintes sont qualifiées en $\bar{u}$, alors $\exists \lambda \in \mathbb{R}^l_+$, $\mu \in \mathbb{R}^n$ tel que :

1. $$\sum_{i \in I} \lambda_i g_i(\bar{u}) = 0 \text{ (condition d'exclusion)}$$
2. $$\nabla J(\bar{u}) + \sum_{i \in I} \lambda_i \nabla g_i(\bar{u}) + \sum_{j \in J} \mu_j \nabla h_j(\bar{u}) = 0 \text{ (condition d'optimalité)}$$

Démonstration :

D'après le théorème précédent, il existe $p_0 \in \mathbb{R}_+$, $p \in \mathbb{R}^l_+$ et $q \in \mathbb{R}^m$ tel que :

$$\begin{cases} \sum_{i \in I} p_i g_i(\bar{u}) & = & 0 \\ (p_0, p, q) & \neq & 0 \\ p_0 \nabla J(\bar{u}) + \sum_{i \in I} p_0 \nabla g_i(\bar{u}) + \sum_{j \in J} q_j \nabla h_j(\bar{u}) & = & 0 \end{cases}$$

Montrons que $p_0 \neq 0$.

Par l'absurde, on suppose $p_0 = 0$.

1. $p_i = 0 \forall i \notin I(\bar{u})$
2. $\sum_{i\in I(\bar{u})} p_i \nabla g_i(\bar{u}) + \sum_{j\in J} q_j \nabla h_j(\bar{u}) = 0$

Soit $v \in \mathbb{R}^n$ tel que $\langle \nabla h_j(\bar{u}), v\rangle = 0 \ \forall j \in J$ et $\langle \nabla h_i(\bar{u}), v\rangle < 0 \ \forall i \in I(\bar{u})$.
On a :

$$\begin{aligned} 0 &= \left\langle \sum_{i\in I(\bar{u})} p_i \nabla g_i(\bar{u}) + \sum_{j\in J} q_j \nabla h_j(\bar{u}), v \right\rangle \\ &= \sum_{i\in I(\bar{u})} \underbrace{p_i}_{\geq 0} \underbrace{\langle \nabla g_i(\bar{u}), v\rangle}_{<0} + \underbrace{\sum_{j\in J} q_j \langle \nabla h_j(\bar{u}), v\rangle}_{=0} \end{aligned}$$

$$\Rightarrow p_i = 0 \ \forall i \in I(\bar{u})$$

$$\Rightarrow p = 0$$

Donc :

$$\sum_{j\in J} q_j \nabla h_j(\bar{u}) = 0$$

Or, $\{\nabla h_1(\bar{u}), ..., \nabla h_m(\bar{u})\}$ forment une famille libre, donc on a forcément $\forall j \in J, \ q_j = 0$, ie :

$$q = 0$$

On en déduit donc que

$$(p_0, p, q) = 0$$

ce qui est absurde.
Donc $p_0 \neq 0$.

On pose $\lambda_i = \frac{p_i}{p_0} > 0$ et $\mu_i = \frac{q_i}{p_0}$. On retrouve ainsi les deux égalités.

Remarque:

1. Le résultat reste vrai sans les hypothèses de qualification si les contraintes sont affines (ie, $\forall i, \forall j$, g_i et h_j sont convexes ou concaves)
2. Un peu de vocabulaire :
   — (λ, μ) est le multiplicateur de Lagrange associé à $\bar{u}$
   — Le lagrangien est défini par :

$$L(u, \lambda, \mu) = J(u) + \sum_{i\in I} \lambda_i g_i(u) + \sum_{j\in J} \mu_j h_j(u)$$

et la condition d'optimalité s'écrit :

$$\nabla_u L(\bar{u}, \lambda, \mu) = 0$$

3 Problèmes convexes et dualité

On considère le problème :

$$\min_{u\in\mathcal{U}_{ad}} J(u)$$

où $\mathcal{U}_{ad} = \{u \in \mathbb{R}^n ; \ g_i(u) \leq 0 \forall i \in I\}$

On suppose que les applications $J, g_1, ..., g_l$ sont convexes et de classe $\mathcal{C}^1$.

Propriété:

On suppose qu'il existe $u_0 \in \mathcal{U}_{ad}$ tel que

$$g_i(u_0) < 0 \ \forall i \in I$$

Alors la contrainte est qualifiée (en $\bar{u}$, $\forall \bar{u} \in \mathcal{U}_{ad}$).

Démonstration :
Soit $u \in \mathcal{U}_{ad}$. On pose $v = u_0 - u$. Soit $j \in I(\bar{u})$. Comme g_i est convexe et de classe $\mathcal{C}^1$, on a :

$$\langle \nabla g_i, v \rangle \leq \underbrace{g_i(u_0)}_{<0} - \underbrace{g_i(\bar{u})}_{=0} < 0$$

La contrainte est donc qualifiée en $\bar{u}$.

Lemme:

Soit F une application de $\mathbb{R}^n$ dans $\mathbb{R}$ convexe et de classe $\mathcal{C}^1$. $u \in \mathbb{R}^n$ est un point de minimum de F sur $\mathbb{R}^n$ si et seulement si $\nabla F(u) = 0$.

Démonstration :
Une première implication a déjà été montrée.
On suppose que $\nabla F(u) = 0$. Comme F est convexe, on a

$$F(v) - F(u) \geq \langle \underbrace{\nabla F(u)}_{=0}, v - u \rangle \forall \ v \in \mathbb{R}^n$$

D'où $F(v) \geq F(u) \ \forall v \in \mathbb{R}^n$, donc u est un minimum de F sur $\mathbb{R}^n$.

Théorème:

Soit $\bar{u}$ un point de $\mathcal{U}_{ad}$ tel que les contraintes soient qualifiées en $\bar{u}$. Alors $\bar{u}$ est un point de minimum de J sur $\mathcal{U}_{ad}$ si et seulement si $\exists \lambda \in \mathbb{R}^l_+$ tel que :

$$\left\{ \begin{array}{rcl} \nabla_u L(\bar{u}, \lambda) & = & 0 \\ \lambda^T g(\bar{u}) & = & 0 \end{array} \right.$$

Démonstration :
L'application $u \mapsto L(u, \lambda)$ est une fonction convexe (en tant que sommes de fonctions convexes), donc elle admet un minimum en $\bar{u}$.

$$\Rightarrow L(u, \lambda) \geq L(\bar{u}, \lambda) \ \forall u \in \mathbb{R}^n$$

Or,

$$\begin{array}{rcl} L(\bar{u}, \lambda) & = & J(\bar{u}) + \underbrace{\lambda^t g(\bar{u})}_{=0} \\ & = & J(\bar{u}) \end{array}$$

soit $u \in \mathcal{U}_{ad}$. alors :

$$\begin{array}{rcl} L(u,\lambda) & = & J(u) + \underbrace{\sum_{i \in I} \underbrace{\lambda_i}_{\geq 0} \underbrace{g_i(u)}_{\leq 0}}_{\leq 0} \\ & = & J(u) \end{array}$$

$$J(u) \geq L(u,\lambda) \geq L(\bar{u},\lambda) = J(\bar{u})$$

Donc $\bar{u}$ est un minimum de J sur $\mathcal{U}_{ad}$.

3.1 Dualité

Définition: Point selle

On dit que $(u,\lambda) \in \mathcal{U}_{ad} \times \mathbb{R}^l_+$ est un point selle de L sur $\mathcal{U}_{ad} \times \mathbb{R}^l_+$ si :

$$\forall \mu \in \mathbb{R}^l_+,\ \forall v \in \mathcal{U}_{ad},\ L(u,\mu) \leq L(u,\lambda) \leq L(v,\lambda)$$

Propriété:

Soit U un ouvert contenant $\mathcal{U}_{ad}$ et (u,λ) un point selle de J sur $U \times \mathbb{R}^l_+$.Alors $u \in \mathcal{U}_{ad}$ et :

$$\left\{ \begin{array}{rcl} \nabla_u L(u,\lambda) & = & 0 \\ \lambda^t g(u) & = & 0 \end{array} \right.$$

Démonstration :
Comme (u,λ) est un point selle, on a

$$\forall \mu \in \mathbb{R}^l_+,\ \forall v \in U,\ L(u,\mu) \leq L(u,\lambda) \leq L(v,\lambda)$$

ie :

$$J(u) + \mu^t g(u) \underset{(1)}{\leq} J(u) + \lambda^t g(u) \underset{(2)}{\leq} J(v) + \lambda^t g(v)$$

$$(1)\ :\ \mu^T g(u) \leq \lambda^T g(u),\ \forall \mu \in \mathbb{R}^l_+$$

Si on prend $\mu = \frac{1}{2}\lambda$ et $\mu = 2\lambda$, on voit bien que $\lambda^t g(u) = 0$.
Ainsi, $\mu^t g(u) \leq 0\ \forall \mu \in \mathbb{R}^l_+$.
Prenons

$$\mu = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

. Ainsi, $g_i(u) \leq 0\ \forall i \in I$.

$$\Rightarrow u \in \mathcal{U}_{ad}$$

(2) : u est un point de minimum de :

$$u \mapsto J(u) + \lambda^t g(u) = L(u, \lambda)$$

sur U. Donc :

$$\nabla_u L(u, \lambda) = 0$$

Lemme:

Soit L une fonction de deux variables u et λ.

$$\inf_u \sup_\lambda L(u, \lambda) \geq \sup_\lambda \inf_u L(u, \lambda)$$

Démonstration :

$$\forall u, \sup_\lambda L(u, \lambda) \geq \sup_\lambda \inf_u L(u, \lambda)$$

Ceci étant vrai pour tout u, c'est également vrai pour celui qui minimise le terme de gauche.

$$\Rightarrow \inf_u \sup_\lambda L(u, \lambda) \geq \sup_\lambda \inf_u L(u, \lambda)$$

Théorème:

On suppose que la contrainte est qualifiée et que $(\mathcal{P}_{\mathcal{U}_{ad}})$ admet une solution.
Alors

$$\min_{u \in \mathcal{U}_{ad}} J(u) = \sup_{\lambda \in \mathbb{R}^n} \inf_{u \in \mathbb{R}^n} L(u, \lambda) = \inf_{u \in \mathbb{R}^n} \sup_{\lambda \in \mathbb{R}^n} L(u, \lambda)$$

De plus, le problème $\sup_{\lambda \in \mathbb{R}^n} \inf_{u \in \mathbb{R}^n} L(u, \lambda)$ admet une solution λ^* et $\inf_{u \in \mathbb{R}^n} \sup_{u \in \mathbb{R}^n} L(u, \lambda)$ admet une solution u^* qui est solution de $(\mathcal{P}_{\mathcal{U}_{ad}})$

Définition:

On note

$$d(\lambda) = \inf_{u \in \mathbb{R}^n} L(u, \lambda)$$

Le problème

$$\sup_{\lambda \in \mathbb{R}^n} d(\lambda)$$

est le problème dual.

Démonstration :
Comme les contraintes sont qualifiées et que le problème admet au moins une solution u^*, on a :

$$\left\{ \begin{array}{rcl} \nabla J(u^*) + \sum_{i \in I} \lambda_i^* \nabla g_i(u^*) & = & 0 \\ \sum_{i \in I} \lambda_i^* g_i(u^*) & = & 0 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{rcl} \nabla_u L(u^*, \lambda^*) & = & 0 \\ {\lambda^*}^t g(u^*) & = & 0 \end{array} \right.$$

Comme J, g_1, ..., g_l sonr convexes et $\lambda_i^* \geq 0$, on a $u \mapsto L(u, \lambda^*)$ convexe.
Donc u^* est un point de maximum de $u \mapsto L(u, \lambda^*)$.

$$\sup_{\lambda \in \mathbb{R}_+^l} \inf_{u \in \mathbb{R}^n} L(u, \lambda) \geq \inf_{u \in \mathbb{R}^n} L(u, \lambda^*) = L(u^*, \lambda^*)$$

De plus,

$$\begin{aligned} L(u^*, \lambda^*) &= J(u^*) + \underbrace{\sum_{i \in I} \lambda_i^* g_i(u^*)}_{=0} \\ &= J(u^*) \end{aligned}$$

$$\Rightarrow L(u^*, \lambda^*) = \min_{u \in \mathcal{U}_{ad}} J(u) \leq \sup_{\lambda \in \mathbb{R}_+^l} \inf_{u \in \mathbb{R}^n} L(u, \lambda)$$

Montrons à présent que :

$$\sup_{\lambda \in \mathbb{R}_+^l} L(u, \lambda) = \begin{cases} J(u) & \text{si} \quad u \in \mathcal{U}_{ad} \\ +\infty & \text{sinon} \end{cases}$$

— Si $u \in \mathcal{U}_{ad}$,

$$\begin{aligned} J(u) = L(u, 0) &\leq \sup_{\lambda \in \mathbb{R}_+^l} L(u, \lambda) \\ &\leq J(u) + \sup_{\lambda \in \mathbb{R}_+^l} \sum_{i \in I} \underbrace{\lambda_i}_{\geq 0} \underbrace{g_i(u)}_{\leq 0} \\ &\leq J(u) \end{aligned}$$

$$\Rightarrow \sup_{\lambda \in \mathbb{R}_+^l} L(u, \lambda) = J(u)$$

— Si $u \notin \mathcal{U}_{ad}$, $\exists i$ tel que $g_i(u) > 0$. On pose

$$\lambda_j^k = \begin{cases} k & \text{si } i = j \\ 0 & \text{sinon} \end{cases}$$

$$\begin{aligned} \sup_{\lambda \in \mathbb{R}_+^l} L(u, \lambda) &\geq \sup_{k \in \mathbb{N}} L(u, \lambda^k) \\ &\geq \sup_{k \in \mathbb{N}} \{J(u) + k g_i(u)\} \\ &\geq J(u) + \underbrace{g_i(u)}_{\geq 0} \underbrace{\sup_{k \in \mathbb{N}} \{k\}}_{+\infty} \end{aligned}$$

$$\Rightarrow \sup_{\lambda \in \mathbb{R}_+^l} L(u, \lambda) = +\infty$$

On en déduit que :

$$\begin{aligned} \inf_{u \in \mathbb{R}^n} \sup_{\lambda \in \mathbb{R}_+^l} L(u, \lambda) &= \inf_{u \in \mathcal{U}_{ad}} J(u) \\ &= L(u^*, \lambda^*) \\ &\leq \sup_{\lambda \in \mathbb{R}_+^l} \inf_{u \in \mathbb{R}^n} L(u, \lambda) \end{aligned}$$

Avec le lemme précédent, on en déduit que :

$$\sup_{\lambda \in \mathbb{R}^n} \inf_{u \in \mathbb{R}^n} L(u, \lambda) = \inf_{u \in \mathbb{R}^n} \sup_{\lambda \in \mathbb{R}^n} L(u, \lambda)$$

Deuxième partie

Programmation linéaire, algotihme du simplexe

1 Introduction

Un problème d'optimisation linéaire est un problème d'optimisation dans lequel le coût et les contraintes sont linéaires (ou plutôt affines).
Il s'agit de trouver les solutions $x \in \mathbb{R}^n$ du problème :

$$\left\{ \begin{array}{ccc} \inf_{x\in\mathbb{R}^n} & \langle c, x\rangle & \\ \text{s. c.} \quad Ax & = & b \\ x & \geq & 0 \end{array} \right. \tag{P_L}$$

où A est une matrice de raille $m \times n$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$.
$x \geq 0$ signifie que toutes les composantes de x sont positives.
Ce problème est dit sous forme standard.

Remarque : On a l'impression que P_L est un cas particulier du problème (sous forme canonique) :

$$\left\{ \begin{array}{ccc} \inf_{x\in\mathbb{R}^n} & \langle c, x\rangle & \\ \text{s. c.} \quad A'x & = & b' \\ Ax & \geq & b \end{array} \right. \tag{1}$$

Mais un problème sous forme canonique peut toujours se ramener à un problème sous forme standard. En effet, la contrainte $A'x = b'$ est équivalent à $A'x \geq b'$ et $-A'x \geq -b'$. Donc 1 est équivalent à :

$$\inf_{x\in\mathbb{R}^n,\ Ax\geq b} \langle c, x\rangle \tag{2}$$

On introduit des variables d'écart $\lambda \in \mathbb{R}^n_+$ tel que $Ax = b + \lambda$. Donc 2 se ramène à :

$$\left\{ \begin{array}{ccc} \inf_{x\in\mathbb{R}^n,\ \lambda\in\mathbb{R}^n_+} & \langle c, x\rangle & \\ \text{s. c.} \quad Ax-\lambda & = & b \\ b & > & 0 \end{array} \right.$$

On décompose x sous la forme :

$$x = x^+ - x^-$$

où $x^+ = \max(0, x) \geq 0$ et $x^- = -\min(0, x) \geq 0$
2 revient donc à résoudre :

$$\left\{ \begin{array}{ccc} \inf_{x^+\in\mathbb{R}^n,x^-\in\mathbb{R}^n,\lambda\in\mathbb{R}^n} & \langle c, x\rangle & \\ \text{s. c.} \quad Ax-\lambda & = & b \\ b & > & 0 \end{array} \right.$$

qui est bien sous forme standard (mais avec plus de variables).

Remarque : On peut supposer sans perte de généralité que toutes les lignes de A sont linéairement indépendantes. Si ce n'est pas le cas, soit certaines contraintes sont redondantes, soit les contraintes sont incompatibles, ie $rg(A) = m \leq n$

Définition:

L'ensemble

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

est appelé l'ensemble des solutions réalisables (ou admissibles).
On appelle sommet (ou point extremal) de X_{ad} un point $x \in X_{ad}$ tel qu'il n'existe pas $\alpha \in]0, 1[$ et $y, z \in X_{ad}$, $y \neq z$ tel que $x = \alpha y + (1 - \alpha)z$.

2 Solutions de base d'un problème sous forme standard

On note $A_1 = \begin{pmatrix} a_{1,1} \\ \vdots \\ a_{m,1} \end{pmatrix}$, ..., $A_n = \begin{pmatrix} a_{1,n} \\ \vdots \\ a_{m,n} \end{pmatrix}$.

Comme $rg(A) = m$, on peut toujours trouver m colonnes de A linéairement indépendantes.
On note

$$\Gamma = \{\gamma : \{1, ..., m\} \rightarrow \{1, ..., n\},\ \text{strictement croissante}\}$$

On définit :

$$A_\gamma = (A_{\gamma(1)} \cdots A_{\gamma(m)})$$

et :

$$\mathcal{B} = \{\gamma \in \Gamma,\ rg(A_\gamma) = m\}$$

Pour $\gamma \in \Gamma$, on définit $\hat{\gamma}$ comme l'unique application strictement croissante de $\{1, ..., n-m\}$ dans $\{1, ..., n\}$ tel que :

$$\gamma(\{1, ..., m\}) \cup \gamma(\{1, ..., n-m\}) = \{1, ..., n\}$$

Définition:

Pour $\gamma \in \mathcal{B}$, la matrice A_γ est appelée base associée à γ.
Les composantes $(x_{\gamma(1)}, ..., x_{\gamma(m)})$ sont appelées les composantes de base, et les composantes $(x_{\hat{\gamma}(1)}, ..., x_{\hat{\gamma}(n-m)})$ sont appelées les composantes hors base.

Pour $\gamma \in \mathcal{B}$, on note

$$x_\mathcal{B} = (x_{\gamma(1)}, ..., x_{\gamma(m)})$$
$$x_N = (x_{\hat{\gamma}(1)}, ..., x_{\hat{\gamma}(n-m)})$$
$$B = A_\gamma$$
$$N = A_{\hat{\gamma}}$$
$$Ax = Bx_\mathcal{B} + Nx_N$$

Comme $rg(B) = rg(A_\gamma) = m$, B est inversible. Donc les contraintes $Ax = b$ peuvent se réécrire :

$$Bx_\mathcal{B} = b - Nx_N$$
$$\Rightarrow x_B = B^{-1}(b - Nx_N)$$

Définition:

Soit $\gamma \in \mathcal{B}$, on appelle solution de base du système $Ax = b$ associé à la base γ la solution x^* définie par :

$$\begin{array}{rcl} x^*_\mathcal{B} & = & B^{-1}b \\ x^*_N & = & 0 \end{array}$$

Définition:

Une solution de base réalisable est une solution de base telle que $x^*_\mathcal{B} \geq 0$ ($\rightarrow x^* \in X_{ad}$).
Dans ce cas, γ est appelée base réalisable. On note $\mathcal{R}$ l'ensemble des bases réallisables.
Enfin on dit que x^* est non dégénéré si $x^*_\mathcal{B} > 0$ (ie $B^{-1}b > 0$).

⤃ *Lemme:*

Les sommets de X_{ad} sont exactement les solutions de base réalisable.

Démonstration :
Soit x^* une solution de base réalisable associée à la base $\gamma \in \mathcal{R}$.
Par l'absurde, on va supposer que x^* n'est pas un sommet de X_{ad}.

$$\forall \theta \in]0,1[, \exists y, z \in X_{ad}, y \neq z; x^* = \theta y + (1-\theta) z$$

$\forall i \in \{0, ..., n-m\}$, on a :

$$x_{\hat{\gamma}(i)} = 0 = \underbrace{\theta}_{>0} \underbrace{y_{\hat{\gamma}(i)}}_{\geq 0} + \underbrace{(1-\theta)}_{>0} \underbrace{z_{\hat{\gamma}(i)}}_{\geq 0}$$

$$\Rightarrow y_{\hat{\gamma}(i)} = z_{\hat{\gamma}(i)} = 0$$

$$\Rightarrow y_N = z_N = 0$$

Or, $y_B = B^{-1}(b - N y_N) = B^{-1} b = z_B = x^*_B$. Donc $y = z = x^*$, ce qui est absurde.
Donc x^* est un sommet de X_{ad}.

Réciproquement, on suppose que x^* est un sommet de X_{ad}. On note k le nombre de composantes non nulles de x^*.

$$\exists \gamma_1 : \{1, ..., k\} \to \{1, ..., n\} \text{ strictement croissante telle que } x^*_{\gamma_1(i)} > 0,\ \forall i = 1, ..., k$$

On veut montrer que les vecteurs $A_{\gamma_1(1)}, ..., A_{\gamma_1(k)}$ sont libres. Si c'est le cas, on pourra compléter cette famille de vecteurs afin d'obtenir une base.
($\Rightarrow$ On définit $\gamma \in \Gamma$ telle que :

- $\gamma_1(\{1, ..., k\}) \subset \gamma(\{1, ..., m\})$
- $\gamma \in \mathcal{B}$)

Par l'absurde, on suppose que $A_{\gamma_1(1)}, ..., A_{\gamma_1(k)}$ sont liés. Alors $\exists y \in \mathbb{R}^n$, $y \neq 0$,

$$\sum_{i=1}^{k} y_{\gamma_1(i)} A_{\gamma_1(i)} = 0$$

et

$$y_{\hat{\gamma}_1(i)} = 0$$

$\forall \varepsilon > 0$, on a

$$A(x + \varepsilon y) = Ax + \varepsilon \underbrace{Ay}_{=0} = b$$

$$A(x - \varepsilon y) = Ax - \varepsilon Ay = b$$

On a $x^*_{\gamma_1(i)} \pm \varepsilon y_{\gamma_1(i)} > 0$, $\forall i = 1, ..., k$ pour ε assez petit.
De plus, $x^*_{\hat{\gamma}_1(i)} \pm \varepsilon y_{\hat{\gamma}_1(i)} = x^*_{\hat{\gamma}_1(i)} = 0$

$$\Rightarrow x^* \pm \varepsilon y \in X_{ad},\ \forall 0 \leq \varepsilon \leq \varepsilon_0$$

Or, $x^* = \frac{1}{2}(x^* + \varepsilon y) + \frac{1}{2}(x^* - \varepsilon y)$. Donc x^* n'est pas un sommet de X_{ad}. Contradiction.

⤃ *Lemme:*

S'il existe une solution optimale de (P_L) alors il existe une solution optimale de base réalisable.

Démonstration :
Soit $x \in X_{ad}$ une solution optimale. On note k le nombre de composantes non nulles de x. Soit $\gamma_1 : \{1, ..., k\} \to \{1, ..., n\}$ strictement croissante telle que $x_{\gamma_1(i)} > 0 \ \forall i = 1, ..., k$.
Si la famille $A_{\gamma_1(1)}, ..., A_{\gamma_1(k)}$ est libre alors x est une solution optimale de base réalisable.
Si la famille $A_{\gamma_1(1)}, ..., A_{\gamma_1(k)}$ est liée, alors $\exists y \in \mathbb{R}^n$ tel que :

$$\sum_{i=1}^{k} y_{\gamma_1(i)} A_{\gamma_1(i)} = 0$$

$$y_{\hat{\gamma}_1(i)} = 0 \ \forall i 1 \leq i \leq n-k$$

$x + \varepsilon y \in X_{ad}$ pour ε assez petit (cf démo précédente).
Comme c est un point de minimum, on a :

$$\langle c, x \rangle \leq \langle c, x \pm \varepsilon y \rangle$$
$$\Rightarrow \pm \langle c, y \rangle \geq 0$$
$$\langle c, y \rangle = 0$$

On pose $\varepsilon_1 = \sup\{\varepsilon > 0; \ x \pm \varepsilon y \in X_{ad}\}$.

$$\exists i \in \{1, ..., k\}, 0 < \varepsilon < \varepsilon_1 \text{ tels que } (x + \varepsilon y)_{\gamma_1(i)} = 0 \text{ ou } (x - \varepsilon y)_{\gamma_1(i)} = 0$$

On pose $z = x + \varepsilon y$ ou $z = x - \varepsilon y$. tel que $z_{\gamma_1(i)} = 0$. z est donc une solution optimale

$$(\langle c, z \rangle = \langle c, x \pm \varepsilon y \rangle = \langle c, x \rangle)$$

qui a au plus $k-1$ composantes non nulles.
On peut refaire la preuve en remplaçant x par z (et en diminuant k) jusqu'à obtenir une famille libre.

Remarque : D'après les deux propositions précédentes, en prenant $c = 0$ (toutes les solutions admissibles sont optimales), si $X_{ad} = \emptyset$, alors X_{ad} possède au moins 1 sommet (ceci n'est pas le cas pour un polyèdre quelconque).

3 Algorithme du simplexe

Idée : Parcourir les sommets du polyèdre en diminuant le coût en passant d'un sommet au suivant.

Lemme:

Soit $\gamma \in \mathcal{R}$ et x^* la solution de base réalisable associée.
Pour $x \in X_{ad}$ réalisable, on a :

$$x_B = x_B^* - B^{-1} N x_N$$
$$c^T x = c^T x^* + d^T x$$

où $d_N^T = c_N^T - c_B^T B^{-1} N$ et $d_B^T = 0$.

Démonstration :

$$\begin{aligned} Ax &= Bx_B + Nx_N = b \\ \to x_B &= B^{-1} b - B^{-1} N x_N \end{aligned}$$

Or, $x_B^* = B^{-1} b$, donc

$$x_B = x_B^* - B^{-1} N x_N$$

$$\begin{aligned} c^T x &= c_B^T x_B + c_N^T x_n \\ &= c_B^T (x_B^* - B^{-1} N x_N) + c_N^T x_N \\ &= c_B^T x_B^* + (c_N^T - c_B^T B^{-1} N) x_N \\ &= c^T x^* + d^T x \end{aligned}$$

Définition:

On dit que le vecteur d est le veteur des prix marginaux associé à la base γ.

Proposition:

Soit γ une base réalisable et x^* la solution de base associée.
Si toutes les composantes de d sont positives, alors x^* est une solution otpimale.
Si de plus, x^* est non dégénérée ($B^{-1}b > 0$) et est une solution optimale, alors les composantes de d sont positives.

Démonstration :
Soit x une solution réalisable. alors

$$c^T x = c^T x^* + \underbrace{d^T}_{\geq 0} \underbrace{x}_{\geq 0} \geq c^T x^*$$

x^* est donc une solution optimale.

Réciproquement, on suppose que x^* est une solution optimale non dégénérée.
Raisonnons par l'absurde : supposons que la ième composante de d est strictement négative.
Pour $\varepsilon > 0$, on définit :

$$\begin{array}{rcl} x_N(\varepsilon) & = & \varepsilon a_i \\ x_B(\varepsilon) & = & B^{-1}b - B^{-1}N x_N(\varepsilon) \end{array}$$

On a :

$$\begin{array}{rcl} Ax(\varepsilon) & = & Bx_B(\varepsilon) + Nx_N(\varepsilon) \\ & = & b \end{array}$$

Comme $\varepsilon > 0$ on a $x_N(\varepsilon) > 0$.
Comme $B^{-1}b > 0$, pour ε assez petit, on a $x_B(\varepsilon) \geq 0$.

$$\Rightarrow x(\varepsilon) > 0$$

$\Rightarrow x(\varepsilon)$ est une solution réalisable.

On a également :

$$\begin{array}{rcl} c^T x(\varepsilon) & = & c^T x^* + d^T x(\varepsilon) \\ & = & c^T x^* + d_N^T x_N(\varepsilon) \\ & = & c^T x^* + \varepsilon \underbrace{d^T_{\hat{\gamma}(i)}}_{<0} \\ & < & c^T x^* \end{array}$$

$\Rightarrow x^*$ n'est pas otimal. Absurde.

3.1 Privot à partir d'une base réalisable : critère de Dantzig

Étant donné un choix de base $\gamma \in B$, tel que x^* soit une solution de base réalisable mais pas optimale, le but est de trouver une nouvelle base $\delta \in B$ tel que la solution de base y^* associée à δ soit réalisable et telle que

$$c^T y^* \leq c^T x^*$$

Cette méthode opère au moyen d'un pivot dans le sens où $\gamma(\{1, ..., m\})$ et $\delta(\{1, ..., m\})$ ne diffèrent que par 1 seul élément.
En pratique, on note

$$E_\gamma = \{j \in \{1, ..., n-m\};\ d_{\hat{\gamma}(j)} < 0\}$$

Si la solution de base x^* associée à γ n'est pas optimale, alors $E_\gamma \neq \emptyset$.
Soit $j^* \in E_\gamma$ (plusieurs choix possibles, on verra ça plus tard). On définit l'ensemble

$$S_{\gamma,j^*} = \left\{i \in \{1, ..., m\};\ (B^{-1}N)_{i,j^*} > 0\right\}$$

Proposition:

Si $S_{\gamma,j^*} = \emptyset$, alors le problème (P_L) n'admet pas de solution car la fonction coût n'est pas bornée inférieurement sur X_{ad}.

Démonstration :
Soit $t > 0$. On considère le vecteur x définit par :

$$\begin{array}{rcl} x_{\hat{\gamma}(j)} & = & 0 \text{ si } j \in \{1, ..., n-m\} \backslash \{j^*\} \\ x_{\hat{\gamma}(j^*)} & = & t \\ x_{\gamma(i)} & = & (B^{-1}b)_i - t(B^{-1}N)_{ij^*} \text{ si } i \in \{1, ..., m\} \end{array}$$

On a $Ax = b$ par construction.
De plus, $x_N \geq 0$ et $x_B \geq 0$ car $(B^{-1}N)_{ij^*} \leq 0$ (car $S_{\gamma,j^*} = \emptyset$).
$\Rightarrow$ x est une solution admissible.

De plus :

$$\begin{array}{rcl} c^T x & = & c^T x^* + d^T x \\ & = & c^T x^* + d_N^T x_N \\ & = & c^T x^* + \underbrace{d_{\hat{\gamma}(j^*)}}_{<0} t \end{array}$$

Quand $t \to +\infty$, on a $c^T x \to -\infty$, avec $x \in X_{ad}$. La fonction coût n'est donc pas bornée inférieurement.

On suppose à partir de maintenant que la fonction coût est bornée inférieurement.

$$\Rightarrow S_{\gamma,j^*} \neq \emptyset$$

On choisit $i^* \in S_{\gamma,j^*}$.

Lemme:

Soit $\delta \in \Gamma$ l'unique application telle que

$$\delta(\{1, ..., m\}) = \gamma(\{1, ..., m\}) \backslash \gamma(i^*) \cup \hat{\gamma}(j^*)$$

Alors $\delta \in B$.

Démonstration :
Montrons que :

$$A_{\hat{\gamma}(j^*)} = \sum_{i=1}^{m} (B^{-1}N)_{ij^*} A_{\gamma(i)}$$

$$\begin{aligned}
\sum_{i=1}^{m} (B^{-1}N)_{ij^*} A_{\gamma(i)} &= \sum_{i=1}^{m}\sum_{j=1}^{m} B^{-1}_{ij} N_{jj^*} A_{\gamma(i)} \\
&= \sum_{j=1}^{m} N_{jj^*} \sum_{i=1}^{m} B^{-1}_{ij} \begin{pmatrix} B_{1i} \\ \vdots \\ B_{mi} \end{pmatrix} \\
&= \sum_{j=1}^{m} N_{jj^*} \begin{pmatrix} \sum_{i=1}^{m} B_{1i} B^{-1}_{ij} \\ \vdots \\ \sum_{i=1}^{m} B_{mi} B^{-1}_{ij} \end{pmatrix} \\
&= \sum_{j=1}^{m} N_{jj^*} \begin{pmatrix} (BB^{-1})_{1j} \\ \vdots \\ (BB^{-1})_{mj} \end{pmatrix} \\
&= \sum_{j=1}^{m} N_{jj^*} e_j \\
&= A_{\hat{\gamma}(j^*)}
\end{aligned}$$

Comme $\{A_{\gamma(i)}\}_{i=1}^{m}$ forme une base et que $(B^{-1}N)_{i^*j^*} > 0$, on ne peut pas exprimer $A_{\hat{\gamma}(j^*)}$ comme une combinaison linéaire de $A_{\gamma(i)}$ pour $i = 1, ..., n$ avec $i \neq i^*$.
Donc $((A_{\gamma(i)})_{i=1,...,m,\ i\neq i^*}, A_{\hat{\gamma}(j^*)})$ forme une base.

$$\Rightarrow \delta \in B$$

➾ *Lemme: Critère de Dantzig*

On choisit i^* tel que $\frac{x_{\gamma(i^*)}}{(B^{-1}N)_{i^*j^*}} = t_{j^*} = \min_{i \in S_{\gamma,j^*}} \frac{x_{\gamma(i)}}{(B^{-1}N)_{ij^*}}$. Alors δ défini dans le lemme précédent est une base réalisable. De plus, si y^* désigne la solution de base réalisable associée à δ alors

$$c^T y^* \leq c^T x^*$$

l'inégalité étant stricte si $t_{j^*} > 0$.

Démonstration :
On définit $y \in \mathbb{R}^n$ par :

$$y_{\hat{\gamma}(i)} = 0 \ \forall i \in \{1, ..., n-m\} \backslash \{j^*\}$$

$$y_{\hat{\gamma}(j^*)} = t_{j^*}$$

$$y_{\gamma(i)} = x^*_{\gamma(i)} = t_{j^*}(B^{-1}N)_{i,j^*} \ \forall i \in \{1, ..., m\}$$

où x^* est la solution de base réalisable associée à γ.

$$\begin{aligned}
Ay &= By_B + Ny_N \\
&= B(x^*_B - t_{j^*}(B^{-1}N)_{\bullet,j^*}) + Ny_N \\
&= Bx^*_B \\
&= b
\end{aligned}$$

De plus, $y \geq 0$ (par définition de t_{j^*}).
Donc y est une solution réalisable.

$$y_{\gamma(i^*)} = \hat{x}_{\gamma(i^*)} - t_{j^*}(B^{-1}N)_{i^*j^*} = 0$$

et $y_{\hat{\gamma}(i)} = 0 \ \forall i \in \{1, ..., n-m\} \backslash \{j^*\}$. Donc $y_{\hat{\delta}(j)} = 0$, $\forall j \in \{1, ..., n-m\}$.
Donc y est la solution de base réalisable associée à la base δ. Donc δ est une base réalisable.

$$\begin{aligned} c^T y^* &= c^T y \\ &= c^T x^* + d^T y \\ &= c^T x^* + d_{\hat{\gamma}(j^*)} t_{j^*} \end{aligned}$$

car :

$$\begin{aligned} d^T y &= d_N^T y_N + \underbrace{d_B^T}_{=0} y_B \\ &= d_N^T \begin{pmatrix} 0 \\ \vdots \\ t_{j^*} \\ \vdots \end{pmatrix} \\ &= d_{\hat{\gamma}(j^*)} t_{j^*} \end{aligned}$$

$$\Rightarrow c^T y^* = c^T x^* + \underbrace{d_{\hat{\gamma}(j^*)}}_{<0} \underbrace{t_{j^*}}_{\geq 0} \leq c^T x^*$$

Définition: Critère naturel

On appelle variable entrante selon le critère naturel la variable $x_{\hat{\gamma}(j^*)}$ tel que :

$$d_{\hat{\gamma}(j^*)} = \min_{j \in \{1, ..., n-m\}} d_{\hat{\gamma}(j)}$$

et $j^* = \min\{j, \ d_{\hat{\gamma}(j)} = d_{\hat{\gamma}(j^*)}\}$. On appelle variable sortante selon le critère naturel la variable $x_{\gamma(i^*)}$ tel que :

$$i^* = \min \left\{ i \in S_{\gamma, j^*}, \frac{x_{\gamma(i)}}{(B^{-1}N)_{i,j^*}} \right\} = t_{j^*}$$

Définition: Critère de Bland

On choisit la variable entrante selon le critère de Bland pour que $x_{\hat{\gamma}(j^*)}$ soit telle que :

$$j^* = \min\{j, j \in E_\gamma\}$$

et la variable sortante $x_{\gamma(i^*)}$ telle que :

$$i^* = \min \left\{ i \in S_{\gamma, j^*}, \frac{x_{\gamma(i)}}{(B^{-1}N)_{i,j^*}} \right\} = t_{j^*}$$

3.2 Détermination d'une première base réalisable

On considère le problème :

$$\left\{\begin{array}{lccc} \inf_{x\in\mathbb{R}^q} & c^T x & & \\ \text{s. c.} & Ax & \leq & b \\ & x & \geq & 0 \end{array}\right. \tag{*}$$

$$A = (a_{ij})_{1\leq i\leq n,\ 1\leq j\leq q},\ b \in \mathbb{R}^m,\ c \in \mathbb{R}^q$$

Définition:

Le problème (*) est un problème de première espèce si toutes les composantes de b sont positives.
Dans le cas contraire, ou si le problème n'est pas sous forme canonique, on dit qu'il s'agit d'un problème de deuxième espèce.

On considère pour le moment un problème de première espèce ($b \geq 0$).

$$(*) \Leftrightarrow \left\{\begin{array}{lccc} \inf_{x\in\mathbb{R}^n} & \bar{c}^T\bar{x} & & \\ \text{s. c.} & \bar{A}\bar{x} & = & \bar{b} \\ & \bar{x} & \geq & 0 \end{array}\right.$$

où $\bar{x} = (x_1, ..., x_{q+m})^T$,

$$\bar{A} = \begin{pmatrix} a_{11} & \cdots & a_{1q} & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{nq} & 0 & \cdots & 1 \end{pmatrix}$$

$\bar{c} = (c_1, ..., c_q, 0, ..., 0)^T$
$\bar{b} = b$

Lemme:

On suppose que toutes les composantes de b sont positives. Alors le problème (*) possède comme base réalisable celle obtenue en ne gardant en base que les m variables d'écart.

Démonstration :
On définit γ par :

$$\gamma(\{1, ..., m\}) = \{q+1, ..., q+m\}$$

Alors la matrice B correspondante est la matrice identité qui est inversible.
De plus, $x^* = B^{-1}b = b > 0$
Donc x^* est une solution de base réalisable.

On doit maintenant traiter des problèmes de deuxième espèce.
On commence par le cas d'un problème sous forme standard.

$$\left\{\begin{array}{lccc} \inf_{x\in\mathbb{R}^n} & \langle c, x\rangle & & \\ \text{s. c.} & Ax & = & b \\ & x & \geq & 0 \end{array}\right. \tag{1}$$

On peut supposer que toutes les composantes de b sont positives (sinon, il suffit de multiplier la ligne correspondante de A par -1).
On introduit alors des variables "fictives" $y_1, ..., y_m$ et on considère le problème :

$$\left\{\begin{array}{lccc} \inf_{x\in\mathbb{R}^n} & y_1 + ... + y_m & & \\ \text{s. c.} & Ax + y & = & b \\ & x, y & \geq & 0 \end{array}\right. \tag{2}$$

Si le problème (1) admet une solution réalisable, alors le minimum dans (2) est nul et inversement. L'avantage est que le problème (2) admet une solution de base réalisable triviale (x=0,y=b).

On peut alors appliquer l'algorithme du simplexe pour trouver une solution de base réalisable optimale. Si le minimum obtenu est 0 ($y_1 = ... = y_m = 0$), alors le x correspondant est une solution de base réalisable pour (1).

On considère maintenant un problème sous forme canonique avec b qui n'est pas positif.

$$\begin{cases} \inf_{x\in\mathbb{R}^n} & \langle c, x\rangle \\ \text{s. c.} & Ax \leq b \\ & x \geq 0 \end{cases}$$

Ce problème est équivalent à :

$$\begin{cases} \inf_{x\in\mathbb{R}^{q+m}} & \bar{c}^T\bar{x} \\ \text{s. c.} & \bar{A}\bar{x} = \bar{b} \\ & \bar{x} \geq 0 \end{cases} \tag{1}$$

où $\bar{x} = (x_1, ..., x_{q+m})^T$,

$$\bar{A} = \begin{pmatrix} a_{11} & \cdots & a_{1q} & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{nq} & 0 & \cdots & 1 \end{pmatrix}$$

$\bar{c} = (c_1, ..., c_q, 0, ..., 0)^T$
$\bar{b} = b$

On considère le problème d'initialisation :

$$\begin{cases} \inf_{x\in\mathbb{R}^{q+m+1}} & x_{q+m+1} \\ \text{s. c.} & \bar{\bar{A}}x = b \\ & x \geq 0 \end{cases} \tag{2}$$

$$\bar{\bar{A}} = \begin{pmatrix} a_{11} & \cdots & a_{1q} & 1 & \cdots & 0 & -1 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{nq} & 0 & \cdots & 1 & -1 \end{pmatrix}$$

Le problème (1) admet une solution réalisable si et seulement si le minimum du problème (2) est 0. Pour réaliser le problème (2) par la méthode du simplexe, il suffit de trouver une première base réalisable pour (2).

Pour ce faire, on définit $i_0 \in \{1, ..., m\}$ tel que

$$b_{i_0} = \min_{i\in\{1,...,m\}} b_i < 0$$

Pour les variables de base, on choisit :

$$\{x_{q+1}, ..., x_{q+m+1}\}\backslash\{x_{q+i_0}\}$$

Le système $\bar{\bar{A}}x = b$ se traduit par :

$$x_{q+i} - x_{q+m+1} = b_i, \ \forall i \neq i_0$$

$$x_{q+m+1} = -b_{i_0} > 0$$

$\forall i \neq i_0$, $x_{q+i} = b_i - b_{i_0} \geq 0$. On a donc bien une base réalisable.

4 Dualité en programmation linéaire

On condidère un problème sous forme canonique :

$$\left\{\begin{array}{lccc} \inf_{x\in\mathbb{R}^n} & & \langle c, x\rangle & \\ \text{s. c.} & Ax & \leq & b \\ & x & \geq & 0 \end{array}\right. \tag{P}$$

On suppose que (P) possède au moins une solution optimale x^*.
La méthode du simplexe permet à chaque itération d'améliorer une borne supérieure sur le coût $c^T x^*$. Peut-on obtenir une borne inférieure ? Pour cela, on cherche un esolution du problème dual.
Pour ce faire, on part des m contraintes d'inégalités :

$$\sum_{j=1}^{q} a_{ij}x_j \leq b_i,\ i = 1, ..., m$$

On prend $y_i > 0$:

$$\begin{aligned} \sum_{i=1}^{m} y_i \sum_{j=1}^{q} a_{ij}x_j &\leq \sum_{i=1}^{m} y_i b_i \\ \sum_{j=1}^{q} \left(\sum_{i=1}^{m} -y_i a_{ij}\right) x_j &\geq -\sum_{i=1}^{m} y_i b_i \end{aligned}$$

Ainsi,

$$-\sum_{i=1}^{m} y_i a_{ij} \leq c_j \Rightarrow \sum_{j=1}^{q} c_j x_j \geq -\sum_{i=1}^{m} y_i b_i$$

Donc $c^T x^* \geq -b^T y,\ \forall y \geq 0,\ \ y \in \mathbb{R}^m$ vérifiant

$$-\sum_{i=1}^{m} y_i a_{ij} \leq c_j,\ \forall j \in \{1, ..., q\} \Leftrightarrow -A^T y \leq c$$

Définition:

Le problème dual de (P) est le problème suivant :

$$\left\{\begin{array}{lccc} \max_{y\in\mathbb{R}^m} & & -\langle b, y\rangle & \\ \text{s. c.} & -A^T y & \leq & c \\ & y & \geq & 0 \end{array}\right. \tag{D}$$

Théorème:

Si x est une solution réalisable de (P), et si y est une solution réalisable de (D), alors

$$x^T c \geq -b^T y$$

En particulier, si $c^T x = -b^T y$, alors x est une solution optimale de (P) et y est une solution optimale de (D).

Corollaire:

Si la fonction coût n'est pas minorée sur l'ensemble admissible, alors le problème dual n'admet pas de solution réalisable.
Inversement, si le problème dual n'est pas majoré sur son ensemble admissible, alors le problème primal n'admet pas de solution réalisable.

Remarque : La réciproque est fausse en général.

Théorème:

Le problème primal (P) possède une solution optimale x^* si et seulement si le problème dual possède une solution optimale y^*. Dans ce cas, on a $c^Tx^* = -b^Ty^*$.

Démonstration :
Il suffit de montrer que si $x^* = (x_1^*, ..., x_q^*)$ est une solution optimale de (P), alors il existe une solution réalisable $y^* = (y_1^*, ..., y_n^*)$ de (D) tel que :

$$c^Tx^* = -b^Ty^*$$

On part du problème primal. On y introduit les variables d'écart :

$$x_{q+i} = b_i - \sum_{j=1}^{q} a_{ij}x_j$$

Le problème sous forme standrard est :

$$\left\{\begin{array}{lrcl} \inf_{x\in\mathbb{R}^{q+m}} & & \langle \bar{c}, x\rangle & \\ \text{s. c.} & \bar{A}x & = & b \\ & x & \geq & 0 \end{array}\right. \tag{3}$$

avec $\bar{c} = (c_1, ..., c_q, 0, ..., 0)^T$ et :

$$\bar{A} = \begin{pmatrix} a_{11} & \cdots & a_{1q} & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{nq} & 0 & \cdots & 1 \end{pmatrix}$$

On note γ la base réalisable associée à :

$$\bar{x}^* = \left(x_1^*, .., x_q^*, b_1 - \sum_{j=1}^{q} a_{1j}x_j^*, ..., b_m - \sum_{j=1}^{q} a_{mj}x_j^* \right)$$

On note d le vecteur des prix marginaux. ($d_B = 0, d_N^T = c_N^T - c_B^T B^{-1}N$)
Soit $\bar{x} \in \mathbb{R}^{q+m}$ tel que $\bar{A}\bar{x} = b$. On a :

$$\begin{array}{rcl} c^T\bar{x} & = & c^Tx^* + d^T\bar{x} \\ & = & c^Tx^* + \displaystyle\sum_{i=1}^{q} d_i\bar{x}_i + \sum_{i=1}^{m} d_{q+i}\bar{x}_{q+i} \end{array}$$

Or, $\bar{x}_{q+i} = b_i - \sum_{j=1}^{q} a_{ij}\bar{x}_j$

$$\begin{array}{rcl} \Rightarrow c^Tx & = & c^Tx^* + \displaystyle\sum_{i=1}^{q} d_i\bar{x}_i + \sum_{i=1}^{m} d_{q+i}(b_i - \sum_{j=1}^{q} a_{ij}\bar{x}_j) \\ & = & c^Tx^* + \displaystyle\sum_{i=1}^{m} d_{q+i}b_i + \sum_{j=1}^{q} \left(d_j - \sum_{i=1}^{m} a_{ij}d_{q+i} \right) \bar{x}_j \end{array}$$

On choisit maintenant $\bar{x}$ tel que $\bar{x}_1 = ... = \bar{x}_q = 0$.

$$\Rightarrow c^T x^* = -\sum_{i=1}^{m} d_{q+i} b_i \tag{$**$}$$

On choisit maintenant $\bar{x}$ tel que $\bar{x}_j = 0$, $\forall j \in \{1, ..., q\} \backslash \{i\}$, $\bar{x}_i = 1$.

$$c_i = d_i - \sum_{k=1}^{m} a_{ki} d_{q+k}$$

Comme x^* est optimale, on aque le vecteur $d \leq 0$.

$$-\sum_{k=1}^{m} a_{ki} d_{q+k} = c_i - d_i \leq c_i \tag{$\overset{*}{**}$}$$

On définit maintenant y^* par :

$$y_i^* = d_{q+i}, \ i = 1, ..., m$$

En particulier, $y^* \geq 0$, et

$$\begin{aligned} (\overset{*}{**}) \quad &\Leftrightarrow \quad -\sum_{k=1}^{p} a_{ki} y_k^* \leq c_i \\ &\Leftrightarrow \quad -A^T y \leq c \end{aligned}$$

On en déduit que y^* est admissible pour le problème dual.
De plus, $(**)$ implique :

$$c^T x^* = -b^T y^*$$

Définition:

On dit qu'une base γ est primal-réalisable si $B^{-1} b \geq 0$.
De même, on dit qu'elle est duale-réalisable si $d_N^T = c_N^T - c_B^T B^{-1} N \geq 0$.