Chapitre 7

Programmation Linéaire

Un problème de programmation linéaire (PPL) consiste en l'optimisation d'une forme linéaire sur un sous-ensemble de $\mathbb{R}^d$ défini par un ensemble d'équations et inéquations (au sens large : $\leq$ ou $\geq$) affines. Toute solution de cet ensemble d'(in)équations est dite *admissible* pour le PPL. La *valeur* d'un PPL admettant une solution admissible est l'optimum de la forme linéaire associée sur l'ensemble des solutions admissibles. Une solution est dite *optimale* si elle est admissible et optimise la forme linéaire associée. Tout PPL peut se ramener de manière équivalente à une forme *canonique*

$$\begin{array}{rcl} \min \mathbf{c}\mathbf{x} & & \\ A\mathbf{x} & \geq & \mathbf{b} \\ \mathbf{x} & \geq & \mathbf{0}. \end{array}$$

ou à une forme *standard*

$$\begin{array}{rcl} \min \mathbf{c}\mathbf{x} & & \\ A\mathbf{x} & = & \mathbf{b} \\ \mathbf{x} & \geq & \mathbf{0}. \end{array} \qquad \text{(E)}$$

(Bien sûr les matrices $A, \mathbf{c}$ et $\mathbf{b}$ ne sont pas les mêmes dans les deux formes !). Il suffit de remarquer pour cela que toute inéquation $\mathbf{a}\mathbf{x} \geq \mathbf{b}$ peut s'écrire $\mathbf{a}\mathbf{x} - x_0 = \mathbf{b}$ où x_0 est une variable supplémentaire non négative et réciproquement que toute équation $\mathbf{a}\mathbf{x} = \mathbf{b}$ est équivalente aux deux inéquations $\mathbf{a}\mathbf{x} \geq \mathbf{b}$ et $-\mathbf{a}\mathbf{x} \geq -\mathbf{b}$. De plus, on peut écrire $\mathbf{x} = \mathbf{x}^+ - \mathbf{x}^-$ où $\mathbf{x}^+$ et $\mathbf{x}^-$ sont deux vecteurs non négatifs.

7.1 Dualité de la programmation linéaire

Théorème 7.1 (de dualité de la programmation linéaire) *Pour des matrices de dimensions appropriées, on a*

$$\max\{\mathbf{c}\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}\} = \min\{\mathbf{y}\mathbf{b} \mid \mathbf{y} \geq \mathbf{0}, \mathbf{y}A = \mathbf{c}\}$$

pourvu que le min et le max soient pris sur des ensembles non vides.

Preuve : On pose $X = \{A\mathbf{x} \leq \mathbf{b}\}$ et $Y = \{\mathbf{y} \geq \mathbf{0}, \mathbf{y}A = \mathbf{c}\}$. On a

$$\forall \mathbf{x} \in X, \forall \mathbf{y} \in Y \, : \, A\mathbf{x} \leq \mathbf{b} \implies \mathbf{y}A\mathbf{x} \leq \mathbf{y}\mathbf{b} \implies \mathbf{c}\mathbf{x} \leq \mathbf{y}\mathbf{b}$$

D'où $\max \mathbf{c}X \leq \min Y\mathbf{b}$. Il suffit donc de montrer l'existence de $\mathbf{x} \in X$ et $\mathbf{y} \in Y$ tels que $\mathbf{c}\mathbf{x} \geq \mathbf{y}\mathbf{b}$. Ce qui s'écrit encore

$$\exists\, \mathbf{x}, \mathbf{y} \text{ tel que } \begin{bmatrix} A & \mathbf{0} \\ \mathbf{0} & -I \\ \mathbf{0} & A^t \\ \mathbf{0} & -A^t \\ -\mathbf{c} & \mathbf{b}^t \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y}^t \end{bmatrix} \leq \begin{bmatrix} \mathbf{b} \\ \mathbf{0} \\ \mathbf{c}^t \\ -\mathbf{c}^t \\ 0 \end{bmatrix}$$

Par le lemme de Farkas 6.21, la non-existence de tels $\mathbf{x}, \mathbf{y}$ implique

$$\exists\, [\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3}, \mathbf{u_4}, u_5] \geq \mathbf{0} \text{ tel que } [\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3}, \mathbf{u_4}, u_5] \begin{bmatrix} A & \mathbf{0} \\ \mathbf{0} & -I \\ \mathbf{0} & A^t \\ \mathbf{0} & -A^t \\ -\mathbf{c} & \mathbf{b}^t \end{bmatrix} = \mathbf{0} \text{ et}$$

$$[\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3}, \mathbf{u_4}, u_5] \begin{bmatrix} \mathbf{b} \\ \mathbf{0} \\ \mathbf{c}^t \\ -\mathbf{c}^t \\ 0 \end{bmatrix} < 0$$

ce qui s'écrit encore $\mathbf{u_1}A = u_5\mathbf{c}$, $(\mathbf{u_4} - \mathbf{u_3})A^t = u_5\mathbf{b}^t - \mathbf{u_2}$ et $\mathbf{u_1}\mathbf{b} < (\mathbf{u_4} - \mathbf{u_3})\mathbf{c}^t$.

– Ou bien $u_5 = 0$ et on en déduit l'existence de $\mathbf{u_1} \geq \mathbf{0}$ et de $\mathbf{v}(= \mathbf{u_4} - \mathbf{u_3})$ tels que

$$\mathbf{u_1}A = \mathbf{0}, \; \mathbf{v}A^t \leq \mathbf{0}, \; \mathbf{u_1}\mathbf{b} < \mathbf{v}\mathbf{c}^t$$

Comme X est non vide, on a pour un certain $\mathbf{x}$: $\mathbf{u_1}\mathbf{b} \geq \mathbf{u_1}A\mathbf{x} = 0$. De même Y est non vide et pour un certain $\mathbf{y} \geq \mathbf{0}$: $\mathbf{v}\mathbf{c}^t = \mathbf{v}A^t\mathbf{y}^t \leq 0$. D'où $\mathbf{u_1}\mathbf{b} \geq \mathbf{v}\mathbf{c}^t$, une contradiction.

– Ou bien $u_5 > 0$ et on en déduit l'existence de $\mathbf{y}(= \mathbf{u_1}/u_5) \geq \mathbf{0}$ et de $\mathbf{x}(= (\mathbf{u_4} - \mathbf{u_3})^t/u_5)$ tels que

$$\mathbf{y}A = \mathbf{c}, \; A\mathbf{x} \leq \mathbf{b}, \; \mathbf{y}\mathbf{b} < \mathbf{c}\mathbf{x}$$

ce qui contredit l'hypothèse de non-existence ci-dessus !

□

Le théorème de dualité de la programmation linéaire a de nombreuses formulations équivalente dont :

Corollaire 7.2 *Pour des matrices de dimensions appropriées, on a*

$$\max\{\mathbf{c}\mathbf{x} \,|\, \mathbf{x} \geq \mathbf{0}, A\mathbf{x} \leq \mathbf{b}\} = \min\{\mathbf{y}\mathbf{b} \,|\, \mathbf{y} \geq \mathbf{0}, \mathbf{y}A \geq \mathbf{c}\} \tag{7.1}$$
$$\max\{\mathbf{c}\mathbf{x} \,|\, \mathbf{x} \geq \mathbf{0}, A\mathbf{x} = \mathbf{b}\} = \min\{\mathbf{y}\mathbf{b} \,|\, \mathbf{y}A \geq \mathbf{c}\} \tag{7.2}$$

pourvu dans chaque égalité que le min et le max soient pris sur des ensembles non vides.

7.1.1 Application de la dualité

Comme application de la dualité de la programmation linéaire nous donnons une preuve du classique théorème de flot maximum - coupe minimale.

Soit $G = (S, A)$ un graphe orienté. On note respectivement $o(a)$ et $t(a)$ le sommet origine et le sommet terminaison de l'arc a. On se donne pour chaque arc $a \in A$, une *capacité* $c_a \geq 0$, et on considère deux sommets distingués $s, p \in S$, appelés respectivement la *source* et le *puits*. Un *flot* $\mathbf{f} : A \to \mathbb{R}, a \mapsto f_a$ est *admissible* pour $(G, \mathbf{c}, s, p)$ s'il vérifie les deux conditions suivantes :

1. $\forall a \in A : 0 \leq f_a \leq c_a$
2. $\forall v \in S \setminus \{s, p\} : \sum_{a:o(a)=v} f_a = \sum_{a:t(a)=v} f_a$ (condition de conservation du flot, dite loi des noeuds)

La *valeur* $v(\mathbf{f})$ d'un flot $\mathbf{f}$ est la différence entre le flot sortant de s et le flot entrant en s :

$$v(\mathbf{f}) = \sum_{a:o(a)=s} f_a - \sum_{a:t(a)=s} f_a$$

Une *coupe* pour $(G, \mathbf{c}, s, p)$ est une partition $(V, S \setminus V)$ de S telle que $s \in V$ et $p \notin V$. La *capacité* $c(V, W)$ d'une coupe (V, W) est la somme des capacités des arcs sortant de V :

$$c(V, W) = \sum_{\substack{a:o(a)\in V,\\ t(a)\in W}} c_a$$

On définit le flot d'une coupe (V, W) par

$$f(V, W) = \sum_{\substack{a:o(a)\in V,\\ t(a)\in W}} f_a - \sum_{\substack{a:o(a)\in W,\\ t(a)\in V}} f_a$$

et le flot d'un sommet v par

$$f(v) = \sum_{a:o(a)=v} f_a - \sum_{a:t(a)=v} f_a$$

Ainsi pour tout flot admissible $f(s) = v(\mathbf{f})$ et $f(u) = 0$ si $u \neq s, p$. On vérifie aisément que pour toute coupe (V, W) :

$$f(V, W) = \sum_{v\in V} f(v) = f(s) = v(\mathbf{f}) \tag{7.3}$$

Théorème 7.3 (flot max - coupe min, Ford - Fulkerson, 1956) *La valeur maximale des flots admissibles est égale à la capacité minimale des coupes.*

Nous donnons ci-dessous une preuve consistant à appliquer la dualité à un programme linéaire équivalent au calcul du flot maximal[1], puis à montrer que l'on peut se restreindre

1. Il existe des preuves directes de ce résultat sans utiliser la dualité de la programmation linéaire.

aux solutions entières du problème dual. Ce dernier programme entier s'interprète alors comme le calcul de la coupe minimale.

En interprétant $\mathbf{f}$ et $\mathbf{c}$ comme des vecteurs de $\mathbb{R}^A$ de composantes respectives f_a et c_a, le calcul du flot maximal est modélisé par le PPL suivant :

$$\begin{aligned} &\max v(\mathbf{f}) \\ &\qquad \mathbf{f} \leq \mathbf{c} \\ &\qquad \forall v \in S \setminus \{s,p\} : \sum_{a:o(a)=v} f_a = \sum_{a:t(a)=v} f_a \\ &\qquad \mathbf{f} \geq \mathbf{0} \end{aligned} \tag{7.4}$$

Soit $\Pi_{s,p}$ l'ensemble des chemins simples de G joignant s à p. On note $\mathbf{x}$ un vecteur de composantes indexées par les chemins de $\Pi_{s,p}$ et on considère le PPL

$$\begin{aligned} &\max \sum_{\pi \in \Pi_{s,p}} x_\pi \\ &\qquad \forall a \in A : \sum_{\pi : a \in \pi} x_\pi \leq c_a \\ &\qquad \mathbf{x} \geq \mathbf{0} \end{aligned} \tag{7.5}$$

Lemme 7.4 *La valeur du PPL (7.4) est égale à la valeur du PPL (7.5).*

Preuve : Notons que les deux PPLs admettent $\mathbf{0}$ comme solution admissible et sont trivialement bornés. Soit $\mathbf{x} \in \mathbb{R}^{\Pi_{s,p}}$ une solution admissible pour (7.5). On vérifie aisément que le flot $f_a = \sum_{\pi:a\in\pi} x_\pi$ est admissible pour (7.4) et que $v(\mathbf{f}) = \mathbb{1}\mathbf{x}$. On en déduit valeur(7.5) $\leq$ valeur(7.4). Inversement, soit $\mathbf{f}$ un flot admissible pour (7.4) avec $v(\mathbf{f}) > 0$. Soit G_f le sous-graphe de G restreint aux arcs a tels que $f_a > 0$, et soit S_f l'ensemble des sommets atteignables depuis s dans G_f. Si $p \notin S_f$, alors S_f induit une coupe de flot négatif ou nul (les arcs sortants ont un flot nul). D'après l'équation (7.3), ceci est en contradiction avec l'hypothèse $v(\mathbf{f}) > 0$. Donc $p \in S_f$, i.e. il existe un chemin $\pi \in \Pi_{s,p}$ dont tous les arcs ont un flot strictement positif. Soit x_π la valeur minimale du flot sur les arcs de π. On considère le flot $\mathbf{f}'$:

$$f'_a = \begin{cases} f_a - x_\pi & \text{si } a \in \pi \\ f_a & \text{sinon} \end{cases}$$

Clairement, $\mathbf{f}'$ est un flot admissible et $v(\mathbf{f}') < v(\mathbf{f})$. Puisque π a un unique arc sortant de s at aucun arc entrant dans s on a

$$v(\mathbf{f}) = v(\mathbf{f}') + x_\pi$$

On peut itérer ce procédé en partant de $\mathbf{f}'$ au lieu de $\mathbf{f}$. Puisque le graphe $G_{f'}$ contient au moins un arc de moins que G_f, la valeur du flot obtenu par itération du procédé doit s'annuler au bout d'un nombre fini k d'étapes. On a alors obtenu une famille de chemins distincts $\pi, \pi', \ldots, \pi^{(k)}$ de $\Pi_{s,p}$ et des valeurs $x_\pi, x_{\pi'}, \ldots, x_{\pi^{(k)}}$ correspondantes. En posant

$x_\gamma = 0$ pour les chemins γ hors de cette famille, on obtient une solution admissible $\mathbf{x}$ pour (7.5) telle que

$$v(\mathbf{f}) = v(\mathbf{f^{(k)}}) + \sum_{i=0}^{k} x_{\pi^{(i)}} = \mathbb{1}\mathbf{x}$$

(car $v(\mathbf{f^{(k)}}) = 0$). On en déduit valeur(7.4) $\leq$ valeur(7.5), et finalement valeur(7.5) $=$ valeur(7.4). □

Par l'équation (7.1) de dualité de la programmation linéaire, le PPL (7.5) a même valeur que le PPL suivant, où l'on a posé $v(\mathbf{y}) = \sum_{a\in A} c_a y_a$:

$$\begin{aligned} &\min v(\mathbf{y}) \\ &\qquad \forall \pi \in \Pi_{s,p} : \sum_{a:a\in\pi} y_a \geq 1 \\ &\qquad \mathbf{y} \geq \mathbf{0} \end{aligned} \tag{7.6}$$

Lemme 7.5 *Le PPL (7.6) a même valeur que sa restriction* $(7.6)_{\mathbb{Z}}$ *à des vecteurs y entiers.*

On en déduit une

Preuve du théorème 7.3 : Par les lemmes 7.4 et 7.5, il suffit de montrer que la valeur de $(7.6)_{\mathbb{Z}}$ est la capacité minimale de toute coupe. Soit (V, W) une coupe. On pose $\mathbf{y} = (y_a)_{a\in A}$ avec

$$y_a = \begin{cases} 1 & \text{si } o(a) \in V \text{ et } t(a) \in W \\ 0 & \text{sinon} \end{cases}$$

Alors $\mathbf{y}$ est une solution admissible pour $(7.6)_{\mathbb{Z}}$ et $v(\mathbf{y}) = c(V, W)$. On en déduit que la valeur de $(7.6)_{\mathbb{Z}}$ est inférieure à la capacité minimale des coupes. Inversement, soit $\mathbf{y}$ une solution admissible pour $(7.6)_{\mathbb{Z}}$. On définit V comme l'ensemble des sommets atteignables depuis s en utilisant des arcs tels que $y_a = 0$. La condition $\sum_{a:a\in\pi} y_a \geq 1$ indique que $p \notin V$, donc V induit une coupe. De plus,

$$c(V, S \setminus V) = \sum_{\substack{a:o(a)\in V,\\ t(a)\notin V}} c_a \leq \sum_{\substack{a:o(a)\in V,\\ t(a)\notin V}} c_a y_a \leq v(y)$$

Il suit que la capacité minimale des coupes est majorée par la valeur de $(7.6)_{\mathbb{Z}}$, d'où l'égalité de ces deux grandeurs. □

Il ne reste plus qu'à donner une

Preuve du lemme 7.5 : Une première preuve utilise le fait que la matrice des inéquations du PPL (7.6) est totalement unimodulaire (cf [MG07, p.144-145]). On en donne une preuve directe. Notons que $\mathbf{y} = \mathbf{1}$ est une solution admissible pour les PPLs (7.6) et $(7.6)_{\mathbb{Z}}$. Par restriction de l'espace des solutions, valeur(7.6) est majorée par valeur$(7.6)_{\mathbb{Z}}$.

Inversement, soit $\mathbf{y}$ une solution optimale pour (7.6). On définit V comme l'ensemble des sommets atteignables depuis s en utilisant des arcs tels que $y_a = 0$. Comme dans la preuve précédente, V induit une coupe et on pose $\mathbf{y}^{\mathbb{Z}} = (y_a^{\mathbb{Z}})_{a \in A}$ avec

$$y_a^{\mathbb{Z}} = \begin{cases} 1 & \text{si } o(a) \in V \text{ et } t(a) \notin V \\ 0 & \text{sinon} \end{cases}$$

$\mathbf{y}^{\mathbb{Z}}$ est clairement une solution admissible pour $(7.6)_{\mathbb{Z}}$.
Soit $\alpha = \min\{y_a \mid o(a) \in V, t(a) \notin V\}$. Si $\alpha = 1$ alors $v(\mathbf{y}) \geq v(\mathbf{y}^{\mathbb{Z}})$ et $\mathbf{y}^{\mathbb{Z}}$ est optimale. Sinon, on pose $\mathbf{y}' = (y'_a)_{a \in A}$ avec $y'_a = \frac{y_a - \alpha y_a^{\mathbb{Z}}}{1-\alpha}$. Clairement $\mathbf{y}' \geq \mathbf{0}$. Tout chemin joignant s à p contient au moins un arc de la coupe (entre V et son complémentaire). On écrit $\pi = \pi_{su} \cdot \pi_{up}$ où u est le dernier sommet dans V le long de π. On pose $\pi' = \pi'_{su} \cdot \pi_{up}$ où π'_{su} est un chemin joignant s à u par des arcs tels que $y_a = 0$. Puisque $y_a = 0$ implique $y'_a = 0$, on a

$$\sum_{a \in \pi} y'_a \geq \sum_{a \in \pi'} y'_a$$

Or π' contient un unique arc entre V et son complémentaire, d'où $\sum_{a \in \pi'} y_a^{\mathbb{Z}} = 1$. Par ailleurs $\sum_{a \in \pi'} y_a \geq 1$ et on en déduit aisément $\sum_{a \in \pi'} y'_a \geq 1$, ce qui montre que $\mathbf{y}'$ est une solution admissible pour (7.6).

Donc $\mathbf{y} = \alpha \mathbf{y}^{\mathbb{Z}} + (1-\alpha)\mathbf{y}'$ est une combinaison convexe de deux solution admissibles. Ces deux solutions sont donc optimales, d'où $v(\mathbf{y}^{\mathbb{Z}}) = v(\mathbf{y})$. Il suit que valeur$(7.6)_{\mathbb{Z}}$ est majorée par valeur(7.6), d'où l'égalité entre ces deux grandeurs. □

7.2 Algorithmes du simplexe

On considére ici un PPL sous forme standard

$$\begin{aligned} &\min \mathbf{cx} \\ A\mathbf{x} &= \mathbf{b} \\ \mathbf{x} &\geq \mathbf{0}. \end{aligned} \qquad \text{(E)}$$

dont la matrice A est de dimension $m \times n$. On supposera que A est de rang m (il y a donc plus d'inconnues que d'équations) c'est-à-dire qu'aucune des m équations n'est impliquée par les autres. Il est facile de supprimer de telles équations en temps polynomial ou de vérifier que le système n'a pas de solution. On appelle *solution ou point admissible* de (E) tout $\mathbf{x}$ vérifiant les contraintes $A\mathbf{x} = \mathbf{b}$ et $\mathbf{x} \geq \mathbf{0}$. L'ensemble des solutions de (E) est un polyèdre noté P.

Soit $B \subset [1, n]$ un sous-ensemble de m indices tels que les vecteurs colonnes de A correspondants soient indépendants, i.e. tels que la matrice A_B des colonnes de A indexées par B soit inversible. On note $N = [1, n] \setminus B$ les indices complémentaires et on décompose chaque vecteur de $\mathbb{R}^n$ en deux morceaux selon les indices de ses composantes $\mathbf{x} = \begin{bmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{bmatrix}$.
L'ensemble B est appelée une *base* de (E) et le vecteur $\begin{bmatrix} A_B^{-1}\mathbf{b} \\ \mathbf{0} \end{bmatrix}$, la *solution basique* associée. On parle de *solution basique admissible (ou réalisable)* (s.b.a.) lorsque $A_B^{-1}\mathbf{b} \geq \mathbf{0}$. La base associée est dite admissible.

Lemme 7.6 *Pour toute base admissible B, il existe un vecteur de $\mathbf{c}'$ tel que la s.b.a. associée à B soit l'unique point qui minimise la forme linéaire $\mathbf{x} \mapsto \mathbf{c}'\mathbf{x}$ sur (E).*

Preuve : Considérer le vecteur $\mathbf{c}' = [\mathbf{c}'_B \; \mathbf{c}'_N] = [\mathbb{0} \; \mathbb{1}]$. □

Théorème 7.7 *Toute solution basique admissible est un sommet de P. Réciproquement, tout sommet de P est la solution basique associée à une base admissible de (E).*

Preuve : Si B est une base admissible alors le lemme 7.6 fournit un hyperplan support de P dont l'intersection avec P est la s.b.a. associée à B. Cette intersection est donc un sommet de P.

Réciproquement, soit $\mathbf{y}$ un sommet de P. Je note B' l'ensemble des indices des coordonnées strictement positives de $\mathbf{y}$. En particulier, $A_{B'}\mathbf{y}_{B'} = \mathbf{b}$. Les colonnes de $A_{B'}$ forment une famille libre. En effet, dans le cas contraire il existe un jeux de coefficients $\mathbf{d}$ non tous nuls tel que $A_{B'}\mathbf{d} = \mathbf{0}$. En choisissant un $\epsilon > 0$ assez petit, on a alors $\mathbf{y}_{B'} \pm \epsilon\mathbf{d} \geq \mathbf{0}$ et $A_{B'}(\mathbf{y}_{B'} \pm \epsilon\mathbf{d}) = \mathbf{b}$. On en déduit que $\mathbf{y}$ est le milieu de deux points distincts dans P, en contradiction avec le lemme 6.26 sur l'extrémalité des sommets d'un polyèdre. On peut maintenant compléter B' en une base B de A de sorte que $\mathbf{y}$ est la s.b.a. associée à B. □

Si B est une base, la projection $\begin{bmatrix}\mathbf{x}_B \\ \mathbf{x}_N\end{bmatrix} \mapsto \mathbf{x}_N$ réalise une bijection, d'inverse $\mathbf{x}_N \mapsto \begin{bmatrix}A_B^{-1}(\mathbf{b} - A_N\mathbf{x}_N) \\ \mathbf{x}_N\end{bmatrix}$, entre P et le polyèdre P_N de $\mathbb{R}^N$: $\{\begin{bmatrix}-A_B^{-1}A_N \\ I_N\end{bmatrix}\mathbf{x}_N \geq \begin{bmatrix}-A_B^{-1}\mathbf{b} \\ \mathbf{0}\end{bmatrix}\}$. Notons que P et P_N ont la même structure combinatoire. De plus, minimiser $\mathbf{c}\mathbf{x} = \mathbf{c}_B\mathbf{x}_B + \mathbf{c}_N\mathbf{x}_N$ sur P est équivalent à minimiser $\mathbf{c}'_N\mathbf{x}_N$ sur P_N avec $\mathbf{c}'_N = \mathbf{c}_N - \mathbf{c}_B A_B^{-1}A_N$. Remarquons également que B est admissible si et seulement si $\mathbf{0} \in P_N$.

Définition 7.8 *Un* pivot *consiste à remplacer un indice d'une base admissible par un indice non-base, de manière à ce que le nouvel ensemble d'indices soit une base admissible. Deux bases qui se déduisent d'un pivot sont dites* adjacentes*. Un pivot est dit* diminuant *si la valeur de la forme linéaire à minimiser diminue lorsqu'on l'évalue en la s.b.a. associée à la base avant et après le pivot.*

Soit B et B' deux bases admissibles adjacentes avec $B' = B - s \cup \{e\}$ où $e \in N$ et $s \in B$ sont les indices *entrant* et *sortant* à échanger lors du pivot correspondant. Le système (E) équivalent au système

$$\begin{array}{rcl} \mathbf{x}_B & = & A_B^{-1}\mathbf{b} - A_B^{-1}A_N\mathbf{x}_N \\ \mathbf{x} & \geq & \mathbf{0}. \end{array} \qquad (E')$$

On pose $\mathbf{p} = A_B^{-1}\mathbf{b}$ et $Q_N = -A_B^{-1}A_N$ de sorte que $\mathbf{p}$ est la s.b.a. associée à B. Chaque équation dans (E') s'écrit ainsi

$$x_i = p_i + q_{ie}x_e + Q_{i,N-e}\mathbf{x}_{N-e} \qquad (7.7)$$

avec $i \in B$. En séparant l'équation pour x_s, les équations de (E') peuvent encore s'écrire

$$\begin{array}{rcl} \mathbf{x}_{B-s} & = & \mathbf{p}_{B-s} + Q_{B-s,e}x_e + Q_{B-s,N-e}\mathbf{x}_{N-e} \\ x_s & = & p_s + q_{se}x_e + Q_{s,N-e}\mathbf{x}_{N-e} \end{array}$$

Soit encore sous forme matricielle

$$\begin{bmatrix} I & -Q_{B-s,e} \\ \mathbb{0} & -q_{se} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{B-s} \\ x_e \end{bmatrix} + \begin{bmatrix} -Q_{B-s,N-e} & 0 \\ -Q_{s,N-e} & -1 \end{bmatrix} \begin{bmatrix} \mathbf{x}_{N-e} \\ x_s \end{bmatrix} = \mathbf{p}$$

Comme les équations de (E) s'obtiennent en multipliant l'équation ci-dessus par la matrice inversible A_B (après réorganisation de ses colonnes), il suit que B' est une base si et seulement si la matrice $\begin{bmatrix} I & -Q_{B-s,e} \\ \mathbb{0} & -q_{se} \end{bmatrix}$ est inversible, c'est-à-dire $q_{se} \neq 0$. Par ailleurs, la s.b.a. $\mathbf{p}'$ associée à B' est par définition telle que $\mathbf{p}'_{N'} = \mathbf{0}$, soit encore $\mathbf{p}'_{N-e} = \mathbf{0}$ et $p'_s = 0$. Ceci implique avec les équations (7.7) que $p'_e = -\frac{p_s}{q_{se}}$ et $p'_i = p_i - q_{ie}\frac{p_s}{q_{se}}$ pour $i \in B - e$. Les conditions $\mathbf{p}, \mathbf{p}' \geq \mathbf{0}$ d'admissibilité des bases B et B' se traduisent donc par les conditions

$$q_{se} < 0 \quad \text{et} \quad \frac{p_s}{q_{se}} = \min_{i \in B : q_{ie} < 0} \frac{p_i}{q_{ie}} \tag{7.8}$$

On remarque finalement que les s.b.a. $\mathbf{p}$, $\mathbf{p}'$ sont distinctes si et seulement si p_s est non nul.

Théorème 7.9 *Deux bases adjacentes correspondent à deux sommets identiques ou adjacents de P.*

Preuve : Avec les notations de la discussion ci-dessus, on sait par les théorème 7.7 que $\mathbf{p}$ et $\mathbf{p}'$ sont des sommets de P. Par ailleurs, en notant π_N la projection de P sur le polyèdre P_N on a $\pi_N(\mathbf{p}) = \mathbf{0}$ et $\pi_N(\mathbf{p}') = \mathbf{p}'_N$ avec $\mathbf{p}'_{N-e} = \mathbf{0}$ et $p'_e = -p_s/q_{se}$. Si p_s est non nul, le segment $\pi_N(\mathbf{p})\pi_N(\mathbf{p}')$ n'est pas réduit à un point et est inclus dans l'arête $\{\mathbf{x}_{N-e} = \mathbf{0}, x_e \geq 0\}$ de l'hyperoctant $\mathbf{x}_N \geq \mathbf{0}$ de $\mathbb{R}^N$. Comme P_N est inclus dans cet hyperoctant, il suit que $\pi_N(\mathbf{p})\pi_N(\mathbf{p}')$ est une arête de P_N, ou encore que $\mathbf{pp}'$ est une arête de P. □

L'algorithme du simplexe est dû à George B. Dantzig [Dan51]. Il consiste à partir d'une base admissible à effectuer une suite de pivots diminuants jusqu'à atteindre l'optimum. Dit autrement, l'algorithme du simplexe consiste à se déplacer le long des arêtes de P en descendant toujours relativement à la forme linéaire associée à minimiser. Si l'on effectue un pivot à partir d'une base admissible B et que l'on écrit (E) sous la forme équivalente (E') ci-dessus avec $f(\mathbf{x}) = \mathbf{c}'_N\mathbf{x}_N$ comme forme à minimiser, l'indice e entrant lors du pivot doit être choisi tel que $c'_e > 0$. En effet, en notant $\mathbf{p}$ et $\mathbf{p}'$ les s.b.a. respectivement avant et après le pivot, on a $f(\mathbf{p}) = 0$ et $f(\mathbf{p}') = c'_e p'_e$ avec $p'_e > 0$ dès que $\mathbf{p}$ et $\mathbf{p}'$ sont distincts.

Cette dernière condition sur e forme avec les conditions (7.8) le principe fondamental de l'algorithme du simplexe. Si on ne peut trouver d'indices entrant et sortant avec ces conditions alors ou bien tous les coefficients de $\mathbf{c}'_N$ sont non positifs et on a atteint le minimum et l'algorithme s'arête. Ou bien, pour un $c'_e > 0$ donné on ne peut trouver $q_{se} < 0$ (i.e. $q_{se} \geq 0$ pour tout $s \in B$) et le programme est non borné. En effet, on peut augmenter x_e à l'infini avec $\mathbf{x}_{N-e} = \mathbf{0}$ en restant dans (E) et faire ainsi tendre f vers l'infini.

Les conditions ci-dessus laissent tout de même un certain choix pour les indices entrant et sortant lors de chaque pivot. Des règles explicites pour effectuer ce choix ont été étudiées depuis les travaux précurseurs de Dantzig. Parmi celles-ci les plus connues sont le choix de l'indice e qui maximise c_e, le choix d'un pivot qui maximise la diminution de f, le choix d'un pivot qui minimise l'angle de l'arête $\mathbf{pp}'$ avec $\mathbf{c}'$, ou encore la règle de Bland. Cette dernière consiste à choisir l'indice minimal parmi les indices entrants possibles et de même pour l'indice sortant. On peut montrer que la règle de Bland assure que l'algorithme du simplexe termine en un nombre fini d'étapes. Il se peut en effet que l'application d'une règle donne $\mathbf{p}' = \mathbf{p}$ pour plusieurs pivots de suite et que l'on retombe sur une même base induisant une boucle infinie dans l'algorithme.

Détermination d'une base admissible Pour initialiser l'algorithme du simplexe, il faut connaître une base admissible, c'est-à-dire un sommet de l'espace des solutions ou montrer que cet espace est vide. Pour cela on résoud le PPL suivant :

$$\begin{aligned} \min \mathbb{1}\mathbf{y} & \\ A\mathbf{x} + \mathbf{y} &= \mathbf{b} \\ \mathbf{x}, \mathbf{y} &\geq \mathbf{0}. \end{aligned} \qquad \text{(E0)}$$

En ayant pris soin de multiplier chaque ligne de $A\mathbf{x} = \mathbf{b}$ par ± 1 pour obtenir $\mathbf{b} \geq \mathbf{0}$. Le point $\begin{bmatrix}\mathbf{x}\\ \mathbf{y}\end{bmatrix} = \begin{bmatrix}0\\ \mathbf{b}\end{bmatrix}$ est triviallement solution de (E0) et (E) est non vide si et seulement $\min \mathbb{1}\mathbf{y} = 0$. Il reste à appliquer l'algorithme du simplexe à (E0) pour vérifier cette denière condition et en déduire une base admissible de (E) le cas échéant.

On ne connaît pas à l'heure actuelle de règle permettant de rendre l'algorithme du simplexe polynomial. Un contre-exemple célèbre, le cube de Klee-Minsky montre que la plupart des règles, dont celle de Bland, conduise à un algorithme de complexité exponentielle dans le cas le pire. Récemment un algorithme du simplexe randomisé [KS06] de complexité moyenne polynomiale (mais pas fortement polynomial) a été proposé. Pour en savoir plus sur la programmation linéaire en général, on pourra consulter l'excellente introduction au domaine de J. Matoušek et B. Gärtner [MG07]. La vidéo du cours de Craig Tovey
`www.youtube.com/watch?v=Ci1vBGn9yRc` explique de maniere imagée la raison des termes simplexes et pivot dans la méthode de Dantzig.

7.3 La conjecture de Hirsch

Le *diamètre* d'un polytope est le diamètre de son 1-squelette, c'est-à-dire encore la distance maximale qui sépare deux sommets quelconques dans ce 1-squelette. Puisque la méthode du simplexe consiste essentiellement à se déplacer le long des arêtes du polytope des contraintes, une borne sur le diamètre des d-polytopes à n facettes fournie une borne sur le nombre d'étapes nécessaires pour résoudre un PPL à d inconnues et n inéquations par la méthode du simplexe. En 1957 Warren Hirsch a proposé le résultat suivant.

Conjecture de Hirsch *Le diamètre d'un d-polytope à n facettes est majoré par $n - d$.*

En référence à Hirsch, on note $H(n, d)$ le diamètre maximal de tout d-polytope à n facettes.

La conjecture de Hirsch, longtemps considérée vraie, a été récemment infirmée par Francisco Santos Leal [San12] (cf. `http://personales.unican.es/santosf/Hirsch/`) à l'aide d'un contre-exemple. Il s'agit d'un 43-polytope à 86 facettes dont le diamètre est au moins $44 > 86 - 43$. En pratique F. Santos travaille avec le polytope dual et obtient donc un 43-polytope à 86 sommets dont le diamètre du dual est au moins 44. Sa construction repose sur la notion de prismatoide.

Définition 7.10 *Un* prismatoide *est un polytope dont les sommets sont contenus dans deux facettes ayant des hyperplans supports parallèles. Ces deux facettes sont dites de base.*

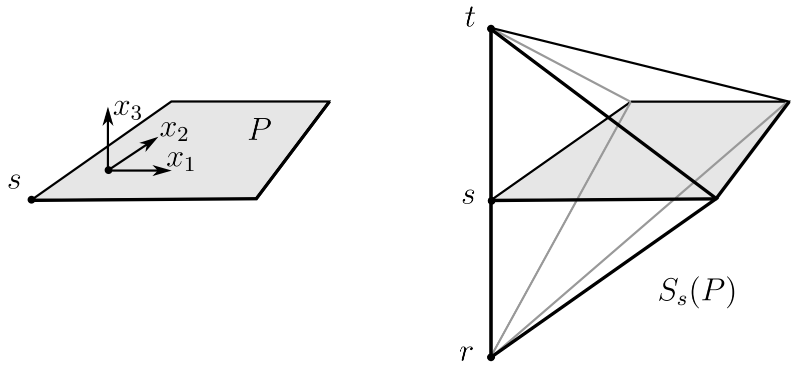

FIGURE 7.1 – Un quadrilatère P (à gauche) et sa suspension par le point s (à droite).

Santos introduit ensuite la *supsension par un point* d'un polytope. Si s est un sommet d'un d-polytope P, sa suspension par le point s, notée $S_s(P)$ est le $(d+1)$-polytope obtenu en plaçant P dans l'hyperplan $x_{d+1} = 0$ de $\mathbb{R}^{d+1}$ et en ajoutant les deux sommets $r := (s, -1)$ et $t := (s, 1) \in \mathbb{R}^{d+1}$, c'est-à-dire en posant $S_s(P) = Conv(P \cup \{r, s\})$. Voir la figure 7.3. Les facettes de $S_s(P)$ sont de la forme $S_s(F)$ pour toute facette F de P contenant s ou de la forme $F * r$ et $F * t$ pour toute autre face de P.

On note $w(P^*)$ le diamètre du dual de P.

Lemme 7.11 *Soit P un d-prismatoide à n sommets. Il existe un $(n-d)$-prismatoide Q à $2(n-d)$ sommets tel que $w(Q^*) \geq w(P^*) + n - 2d$.*

Preuve : On raisonne par récurrence sur $k := n - 2d$. Notons que P ayant ces deux facettes de base P^+ et P^- disjointes, on a $n \geq 2d$ et donc $k \geq 0$. Pour $k = 0$, le lemme est une tautologie. On suppose donc $k > 0$. Il suit que l'une des deux facettes de base, disons P^+, a au moins $d+1$ sommets. Soit s un sommet de P^-. Le polytope $S_s(P)$ a ses sommets contenu dans les faces $S_s(P^-)$ et P^+. En perturbant un sommet de P^+ dans la direction de l'hyperplan support de $S_s(P^-)$, on obtient un d-polytope P_1^+ contenu dans un hyperplan parallèle à $S_s(P^-)$. Le polytope $P_1 := Conv(S_s(P^-) \cup P_1^+)$ est donc un $(d+1)$-prismatoide à $n+1$ sommets. On vérifie que s'il faut traverser $w(P^*)$ facettes pour passer de P^+ à P^- dans P, il faut en traverser au moins $w(P^*)+1$ pour passer de $S_s(P^-)$ à P_1^+ dans P_1 (cf. [San12] pour les détails). Par récurrence, on obtient ainsi des polytopes $P_1, P_2, \ldots, P_{n-2d}$ avec $w(P_i^*) \geq w(P^*) + i$. On remarque finalement que $Q := P_{n-2d}$ convient. □

Santos décrit la liste des 48 sommets d'un certain 5-prismatoide P. Les nombreuses symétries de P et sa taille relativement petite (322 facettes) permettent de vérifier à la main que $w(P^*) = 6$. Le lemme précédent implique alors l'existence d'un 43-polytope Q à 86 sommets tel que $w(Q^*) \geq 44 > 86 - 43$, ce qui achève la construction du contre-exemple. Des contre-exemples plus petits ont été exhibés depuis (voir la page ouaib de F. Santos à ce sujet).

7.4 Programmation linéaire en temps linéaire

Dans les années 1980 N. Megiddo [Meg84] a proposé un algorithme probabiliste de complexité linéaire pour résoudre un PPL en dimension d fixée. Cet algorithme a ensuite grandement été simplifié par Seidel [Sei91]. Nous présentons la méthode de Seidel.

Soit $\mathcal{H}$ une famille finie de demi-espaces de $\mathbb{R}^d$, et soit h une forme linéaire, possiblement nulle. On fixe également un point O de $\mathbb{R}^d$. Rappelons que par le théorème de Minkowski-Weil le polyèdre $P = \cap\mathcal{H}$ est la somme de Minkowski de son cône de récession $\vec{P}$ et d'un polytope. Si P est non vide, il y a deux cas possible.

1. Ou bien h atteint une valeur maximale finie $h_{\max}$ sur P et donc sur une face $F = P \cap \{h = h_{\max}\}$. On a par exemple $F = P$ si $h = 0$.
2. Ou bien h est non majorée sur P, ce qui équivaut à l'existence d'une direction $\mathbf{v}$ de $\vec{P}$ telle que $h(\mathbf{v}) > 0$. De manière encore équivalente, le vecteur de $\vec{P}$ à distance minimale du vecteur dual de h, vu comme point de $\mathbb{R}^d$, est non nul.

Dans le premier cas on définit la *solution* du problème de programmation linéaire donné par $(\mathcal{H}, h, O)$ comme le point de F à distance (euclidienne) minimale de O. Dans le second cas, on définit la solution comme le vecteur de $\vec{P}$ à distance minimale du vecteur dual de h. Il est facile de voir que la solution est définie de manière unique dans les deux cas (cf. exercice ci-après).

Si $P = \cap\mathcal{H}$ est non vide et si la solution du problème associé à $(\mathcal{H}, h, O)$ est un point p, on peut ainsi conclure que h est maximale en p sur P. Si la solution est un vecteur $\mathbf{v}$ on conclut que h est non majorée sur P dans la direction $\mathbf{v}$, i.e. que $h(\mathbf{v}) > 0$.

Exercice 7.12 *On considère dans $\mathbb{R}^d$ un point O et un polyèdre non vide P. Montrer qu'il existe un unique point de P qui minimise la distance euclidienne à O.*

Lemme 7.13 *Soit $\mathcal{H}$ une famille finie de demi-espaces vectoriels dans $\mathbb{R}^d$. Soit E un hyperplan vectoriel valide pour le cône $C = \cap\mathcal{H}$ et soit $F = E \bigcap \cap\mathcal{H}$, la face de C correspondante. Il existe une sous famille $\mathcal{H}' \subset \mathcal{H}$ d'au plus $d - \dim F$ demi-espaces telle que E est valide pour le cône $\cap\mathcal{H}'$.*

Preuve : Soit $\mathbf{v}$ la normale à E opposée à C. Pour tout demi-espace $H \in \mathcal{H}$, on note $H = \{\mathbf{x} \mid \mathbf{v_H x} \leq 0\}$ son équation et $H^0 := \{\mathbf{x} \mid \mathbf{v_H x} = 0\}$ son hyperplan bordant.

On suppose dans un premier temps que C est de dimension d. Par la proposition 6.58 de dualité des cônes, il existe une sous-famille de $\mathcal{K} \subset \mathcal{H}$ telle que

$$\mathbf{v} = \sum_{H \in \mathcal{K}} \lambda_H \mathbf{v_H} \tag{7.9}$$

où les λ_H sont strictement positifs et où $\{\mathbf{v_H}\}_{H\in\mathcal{K}}$ est de rang au plus $d - \dim F$. Si cette famille contient plus d'éléments que son rang, alors elle est liée et on peut écrire $\sum_{H\in\mathcal{K}} \mu_H \mathbf{v_H} = \mathbf{0}$ pour des coefficients μ_H non tous nuls. On peut choisir un réel α tel que les coefficients $\lambda_H + \alpha\mu_H$ sont positifs et au moins l'un est nul. Ceci permet de réduire $\mathcal{K}$ d'un élément au moins. Par récurrence on peut supposer que $\mathcal{K}$ contient au plus $d - \dim F$. L'équation (7.9) montre que l'on peut prendre $\mathcal{H}' = \mathcal{K}$.

Si C est de dimension $k < d$, alors par le lemme 6.46, on peut extraire une sous-famille $\mathcal{H}_\mathcal{C} \subset \mathcal{H}$ de taille $d - k$ telle que l'espace engendré par C soit $vec(C) = \cap_{H\in\mathcal{H}_\mathcal{C}} H^0$. En appliquant ce qui précède aux traces de E et $\mathcal{H}$ dans $vec(C)$, on en déduit une sous-famille $\mathcal{K} \subset \mathcal{H}$ de taille au plus $k - \dim F$ telle que $E \cap vec(C)$ est valide pour le cône $\cap\mathcal{K} \cap vec(C)$. On peut alors prendre $\mathcal{H}' = \mathcal{H}_\mathcal{C} \cup \mathcal{K}$. □

Lemme 7.14 *On considère dans $\mathbb{R}^d$ un point O et une famille finie $\mathcal{H}$ de demi-espaces vectoriels. On peut extraire une sous-famille $\mathcal{H}' \subset \mathcal{H}$ d'au plus d demi-espaces telle que la distance de O à $\cap\mathcal{H}$ et la distance de O à $\cap\mathcal{H}'$ soient minimisées par le même point.*

Preuve : On pose $C = \cap\mathcal{H}$. Soit p le point de C qui minimise la distance euclidienne à O. Si $p = O$ on peut prendre $\mathcal{H}' = \emptyset$ (on convient que $\cap\mathcal{H} = \mathbb{R}^d$ si $\mathcal{H} = \emptyset$). Supposons donc $p \neq O$. Par convexité de C, il suit aisément que l'hyperplan E orthogonal à Op passant par p est valide pour C. Par le lemme précédent, on peut choisir une sous-famille $\mathcal{H}' \subset \mathcal{H}$ d'au plus d demi-espaces telle que E est valide pour le cône $C' = \cap\mathcal{H}'$. Il suit que $p \in C \subset C'$ minimise la distance de O à C'. □

Théorème 7.15 *Soit $\mathcal{H}$ une famille de n demi-espaces de $\mathbb{R}^d$, et soit $h : \mathbf{x} \mapsto \mathbf{cx}$ une forme linéaire sur $\mathbb{R}^d$. On fixe également une origine O dans $\mathbb{R}^d$. Il existe un algorithme randomisé de complexité moyenne $O(2^d(d!)n)$ qui détermine si $\cap\mathcal{H}$ est vide, et dans la négative calcule une solution du problème de programmation linéaire défini par $(\mathcal{H}, h, O)$.*

Preuve : On suppose, par une double récurrence sur d et n qu'il existe un algorithme randomisé répondant aux exigences du théorème. On note $T(d, n)$ sa complexité moyenne. En dimension $d = 1$, l'intersection $\cap\mathcal{H}$ est un segment possiblement vide ou (semi)infini qui se détermine aisément en temps $O(n)$. On en déduit ensuite tout aussi aisément la solution du problème en temps constant. On a ainsi $T(1, n) = O(n)$.

Supposons donc $d > 1$. Si $n = 1$, alors $\mathcal{H}$ contient un unique demi-espace de la forme $H = \{\mathbf{x} \mid \mathbf{c_1 x} \leq b_1\}$. La solution s'obtient comme suit.
– Si $h(\mathbf{c_1}) < 0$, alors h est non bornée sur H et la solution est le vecteur $\mathbf{c}$ dual de h.
– Sinon, soit $\mathbf{c_H}$ la projection orthogonale de $\mathbf{c}$ sur l'hyperplan vectoriel bordant H.
  – Si $\mathbf{c_H} \neq \mathbf{0}$, alors h est non bornée sur H et la solution est le vecteur $\mathbf{c_H}$.
  – Sinon h est bornée sur H. Si $O \in H$, la solution est le point O, sinon c'est la projection de O sur l'hyperplan affine bordant H.
Ce calcul peut triviallement s'effectuer en temps $O(d)$, d'où $T(d, 1) = O(d)$.

Si $n > 1$, on suppose que les demi-espaces $H_1, H_2, \ldots, H_n$ de $\mathcal{H}$ sont donnés dans un ordre aléatoire pour la loi uniforme. On note respectivement $H_i^0, \vec{H}_i$, et $\vec{H}_i^0$ l'hyperplan bordant H_i, le demi-espace vectoriel translaté de H_i, et son hyperplan vectoriel bordant. Par hypothèse de récurrence, on peut calculer la solution du problème défini par $(\mathcal{H} \setminus \{H_n\}, h, O)$ en temps moyen $T(d, n-1)$. Plusieurs cas se présentent.

1. Si $\cap_{i<n} H_i$ est vide, alors c'est le cas pour $\cap\mathcal{H}$.
2. Si la solution de $(\mathcal{H}\setminus\{H_n\}, h, O)$ est un vecteur (donc $(\mathcal{H}\setminus\{H_n\}, h, O)$ est non borné) contenu dans $\vec{H}_n$, alors c'est aussi la solution du problème défini par $(\mathcal{H}, h, O)$ (qui est donc également non borné).
3. Si la solution de $(\mathcal{H} \setminus \{H_n\}, h, O)$ est un point contenu dans H_n, alors c'est aussi la solution du problème défini par $(\mathcal{H}, h, O)$.
4. Sinon, la solution de $(\mathcal{H}, h, O)$ est la solution du problème $(\mathcal{H}', h', O')$ où
   – $h'(\mathbf{x}) = \mathbf{c}'\mathbf{x}$ et $\mathbf{c}'$ est la projection orthogonale de $\mathbf{c}$ sur $\vec{H}_n^0$,
   – $\mathcal{H}' = \{H_j \cap H_n^0\}_{1 \leq j < n}$,
   – O' est la projection orthogonale de O sur H_n^0.

Les tests des cas 2 et 3 s'effectuent aisément en temps $O(d)$.

Dans le cas 4, $P' = \cap\mathcal{H}'$ est un polyèdre de dimension $d-1$ défini par l'intersection de $n-1$ demi-espaces dans l'hyperplan H_n^0. Le calcul d'une base orthonormée[2] de H_n^0 peut s'effectuer en temps $O(d^3)$, puis l'équation dans cette base de chaque demi-espace $H_j \cap H_n^0$ peut s'obtenir en temps $O(d^2)$. Ce qui permet de décrire le problème $(\mathcal{H}', h', O')$ en temps $O(d^3 n)$. Par hypothèse de récurrence, on peut ensuite déterminer en temps moyen $T(d-1, n-1)$ la solution de $(\mathcal{H}', h', O')$ et donc de $(\mathcal{H}, h, O)$.

Analysons la probabilité de se retrouver dans le cas 4. On suppose P non vide. Ou bien h est bornée sur P de valeur maximale h_{max}. On note p la solution de $(\mathcal{H}, h, O)$, et on pose $E = \{h = h_{max}\}$. Soit O_H la projection orthogonale de O sur E.

- Si $h \neq 0$, par le lemme 7.13, on peut extraire une sous-famille $\mathcal{K} \subset \mathcal{H}$ de taille au plus d telle que l'hyperplan E est valide pour $\cap\mathcal{K}$. Puisque $p \in E$, p est aussi le point de la trace de P sur E le plus proche de O_H. Par le lemme 7.14 appliqué dans E, on peut extraire une sous-famille $\mathcal{K}' \subset \mathcal{H}$ de taille au plus $d-1$ telle que le point le plus proche de O_H dans $E \bigcap \cap\mathcal{K}'$ est p. On en déduit que si $H_n \notin \mathcal{K} \cup \mathcal{K}'$, alors h était borné sur $Q := \bigcap_{1 \leq j \leq n-1} H_i$ et la solution du problème associé était déjà p, ce qui correspond

2. Une base orthonormée permet de calculer la norme euclidienne d'un vecteur en temps $O(d)$ à partir de ses coordonnées dans la base.

au cas 2. Dit autrement, la probabilité de se retrouver dans le cas 4 est majorée par la probabilité que H_n soit dans $\mathcal{K} \cup \mathcal{K}'$. Cette probabilité est donc majorée par $2d/n$.

- Si $h = 0$, alors $E = \mathbb{R}^d$, $O_H = O$ et p est par définition le point de P le plus proche de O. Par le lemme 7.14, on peut extraire une sous-famille de taille au plus d dans $\mathcal{H}$ telle que le point de le plus proche de l'intersection de cette sous-famille est p. Un raisonnement analogue à ce qui précède montre que la probabilité de se retrouver dans le cas 4 est majorée par d/n.

Ou bien h est non bornée sur P et un raisonnement analogue montre que la probabilité de se retrouver dans le cas 4 est aussi majorée par d/n.

On en conclut que $T(d, n)$ vérifie la récurrence :

$$T(d,n) \leq \begin{cases} O(n) & \text{si } d = 1, \\ O(d) & \text{si } n = 1, \\ T(d, n-1) + O(d) + \frac{2d}{n}(O(d^3 n) + T(d-1, n-1)) & \text{sinon.} \end{cases}$$

La dernière inéquation s'écrit encore.

$$T(d,n) \leq T(d, n-1) + O(d^4) + \frac{2d}{n}(T(d-1, n-1))$$

Un calcul simple montre que $T(d, n) = O(2^d(d!)n)$. En effet, supposons que $T(d, m) \leq a(d)2^d(d!)m$ pour une certaine fonction $a(d)$ et tout $m < n$. En reportant dans l'équation de récurrence ci-dessus, il apparaît que l'inégalité s'étend à $m = n$ si on choisit $a(d) = O(\sum_{1 \leq k \leq d} \frac{d^4}{2^d(d!)})$. Cette série étant convergente, on a $a(d) = O(1)$. $\square$