

IFT 2505

Programmation Linéaire

Fabian Bastin
DIRO
Université de Montréal

<http://www.iro.umontreal.ca/~bastin/ift2505.php>

Automne 2012

Aspects similaires à la programmation linéaire :

- algorithmes itératifs
- pour un itéré donné, calcul d'une direction de recherche, puis d'une longueur de pas le long de cette direction.

Considérons le programme

$$\begin{aligned} \min_x & c^T x \\ \text{s.à. } & Ax = b \\ & x \geq 0. \end{aligned}$$

Définissons les ensembles

$$\mathcal{F}_P \stackrel{\text{def}}{=} \{x \mid Ax = b, x \geq 0\}$$

$$\mathring{\mathcal{F}}_P \stackrel{\text{def}}{=} \{x \mid Ax = b, x > 0\}$$

Prblème barrière

On suppose que $\mathring{\mathcal{F}}_P$ est non vide et que l'ensemble de solutions optimales pour ce problème est borné.

Soit $\mu \geq 0$. Problème barrière (PB) :

$$\min_x c^T x - \mu \sum_{j=1}^n \log x_j$$

$$\text{s.à. } Ax = b$$

$$x > 0.$$

Prbème barrière vs Problème linéaire

Si $\mu = 0$, on retrouve le problème original en permettant à x d'avoir des composantes nulles :

$$\min_x c^T x - \mu \sum_{j=1}^n \log x_j$$

s.à. $Ax = b$

$x \geq 0.$

Soit $x(\mu)$, la solution au PB étant donné μ . En faisant varier μ continûment vers 0, nous obtenons le chemin central primal.

Si $\mu \rightarrow \infty$, la solution s'approche du centre analytique de la région réalisable, lorsque celle-ci est bornée : le terme barrière prédomine alors dans l'objectif.

Comme $\mu \rightarrow 0$, ce chemin converge vers le centre analytique de la face optimale $\{x \mid c^T x = z^*, Ax = b, x \geq 0\}$, où z^* est la valeur optimale du programme linéaire.

L'idée de base est de résoudre une succession de problèmes barrières pour des valeurs décroissantes de μ .

Fonction lagrangienne

Réécrivons la contrainte $Ax = b$ sous la forme $Ax - b = 0$, et introduisons un vecteur y , en associant une composante y_i à une contrainte $\sum_{j=1}^n a_{ij}x_j = b_j$.

y_i joue à peu près le même rôle que les variables duales précédentes, mais est appelé à présent multiplicateur de Lagrange. Lagrangien :

$$L(x) = c^T x - \mu \sum_{j=1}^n \log x_j - y^T (Ax - b).$$

On cherche à minimiser cette fonction en annulant son gradient.

Minimisation du lagrangien

$$\begin{aligned}\nabla_x L(x) &= 0 \\ \Leftrightarrow c_j - \frac{\mu}{x_j} - y^T a_j &= 0, \quad j = 1, \dots, n \\ \Leftrightarrow \mu X^{-1} \mathbf{1} + A^T y &= c.\end{aligned}$$

En notant $s_j = \mu/x_j$, l'ensemble des conditions d'optimalité (primales, duales, minimisation du lagrangien) s'écrit :

$$\mathbf{x} \circ \mathbf{s} = \mu \mathbf{1}$$

$$\mathbf{Ax} = \mathbf{b}$$

$$\mathbf{A}^T \mathbf{y} + \mathbf{s} = \mathbf{c}.$$

Lien avec le dual

y est une solution duale réalisable et $c - A^T y > 0$.

En effet, le dual est

$$\begin{aligned} \max_y \quad & y^T b \\ \text{s.à. } & A^T y \leq c \end{aligned}$$

Comme

$$\begin{aligned} A^T y + s &= c \\ s_j &= \frac{\mu}{x_j} \end{aligned}$$

nous avons

$$A^T y < c.$$

Exemple

Considérons le programme

$$\begin{aligned} & \max x_1 \\ \text{s.à. } & 0 \leq x_1 \leq 1 \\ & 0 \leq x_2 \leq 1. \end{aligned}$$

On réécrit ce système sous forme standard

$$\begin{aligned} & \min -x_1 \\ \text{s.à. } & x_1 + x_3 = 1 \\ & x_2 + x_4 = 1 \\ & x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0. \end{aligned}$$

Conditions d'optimalité

$$x_1 s_1 = \mu$$

$$x_2 s_2 = \mu$$

$$x_3 s_3 = \mu$$

$$x_4 s_4 = \mu$$

$$x_1 + x_3 = 1$$

$$x_2 + x_4 = 1$$

$$y_1 + s_1 = -1$$

$$y_2 + s_2 = 0$$

$$y_1 + s_3 = 0$$

$$y_2 + s_4 = 0$$

Conditions d'optimalité

De

$$y_2 + s_2 = 0$$

$$y_2 + s_4 = 0$$

on a $s_2 = s_4$.

On en déduit aussi

$$x_2 = x_4,$$

et de là

$$x_2 = x_4 = \frac{1}{2}.$$

Conditions d'optimalité

On a aussi

$$\begin{aligned} -1 &= s_1 - s_3 = \frac{\mu}{x_1} - \frac{\mu}{x_3} \\ \Leftrightarrow -1 &= \frac{\mu}{x_1} - \frac{\mu}{1-x_1} \\ \Leftrightarrow -x_1(1-x_1) &= \mu(1-x_1) - \mu x_1 \\ \Leftrightarrow x_1^2 - x_1 &= \mu - 2\mu x_1 \\ \Leftrightarrow x_1^2 - (1-2\mu)x_1 - \mu &= 0. \end{aligned}$$

Le discriminant de cette équation quadratique est

$$\rho = (1-2\mu)^2 + 4\mu = 1 + 4\mu^2,$$

et de là

$$x_1 = \frac{1-2\mu \pm \sqrt{1+4\mu^2}}{2}$$

Pour μ grand, on doit choisir la racine correspondant à '+'.

D'autre part,

$$x \rightarrow \left(1, \frac{1}{2}\right), \text{ comme } \mu \rightarrow 0.$$

On converge vers le centre analytique de la face optimale

$$\{x \mid x_1 = 1, 0 \leq x_2 \leq 1\}$$

plutôt qu'un coin du carré.

De plus,

$$x(\mu) \rightarrow \left(\frac{1}{2}, \frac{1}{2}\right)$$

comme

$$-2\mu + \sqrt{1 + 4\mu^2} = -2\mu + 2\mu \sqrt{\frac{1}{4\mu^2} + 1}.$$

Chemin central

Dès lors, le chemin central est une ligne droite progressant du centre analytique du carré (comme $\mu \rightarrow \infty$) vers le centre analytique de la face optimale (comme $\mu \rightarrow 0$).

Partons à présent du problème dual

$$\begin{aligned} & \max y^T b \\ \text{s.à. } & y^T A + s^T = c^T \\ & s \geq 0. \end{aligned}$$

Le problème barrière associé est

$$\begin{aligned} & \max y^T b + \mu \sum_{j=1}^n \log s_j \\ \text{s.à. } & y^T A + s^T = c^T \\ & s \geq 0. \end{aligned}$$

Chemin dual central

On suppose que $\mathring{\mathcal{F}}_P = \{(y, s) : y^T A + s^T = c^T, s > 0\}$ est non vide, et l'ensemble des solutions optimales du dual est borné.

On obtient le chemin central dual en faisant tendre μ vers 0.

Lagrangien :

$$L(y) = y^T b + \mu \sum_{j=1}^n \log s_j - (y^T A + s^T - c^T)x.$$

Dès lors,

$$\nabla_y L = 0 \Leftrightarrow b_i - a^i x = 0, \quad \forall i,$$

$$\nabla_s L = 0 \Leftrightarrow \frac{\mu}{s_j} - x_j = 0, \quad \forall j.$$

Chemin dual central : conditions d'optimlaité

On obtient les conditions d'optimalité

$$\mathbf{x} \circ \mathbf{s} = \mu \mathbf{1}$$

$$\mathbf{Ax} = \mathbf{b}$$

$$\mathbf{A}^T \mathbf{y} + \mathbf{s} = \mathbf{c}.$$

On retrouve les mêmes conditions que pour le chemin central.

Par conséquent, x est une solution réalisable primale et $x > 0$.

Considérons l'ensemble de niveau dual

$$\Omega(z) = \{y \mid c^T - y^T A \geq 0, y^T b = z\},$$

avec $z < z^*$, la valeur optimale du dual.

Le centre analytique $(y(z), s(z))$ de Ω coïncide avec le chemin central dual comme z tend vers z^* .

Chemin dual central : exemple

Reprendons le problème

$$\min -x_1$$

$$\text{s.à. } x_1 + x_3 = 1$$

$$x_2 + x_4 = 1$$

$$x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0.$$

Le dual s'écrit

$$\max y_1 + y_2$$

$$\text{s.à. } y_1 \leq -1$$

$$y_2 \leq 0$$

$$y_1 \leq 0$$

$$y_2 \leq 0.$$

Chemin dual central : exemple

On peut réécrire le dual comme

$$\begin{aligned} & \max y_1 + y_2 \\ \text{s.à. } & y_1 + s_1 - 1 \\ & y_2 + s_2 = 0. \end{aligned}$$

Les conditions d'optimalité sont les mêmes que pour le primal,
aussi

$$x_2 = x_4 = \frac{1}{2},$$

d'où

$$s_2 = s_4 = 2\mu, y_2 = -2\mu.$$

Chemin dual central : exemple

Nous avons également, en se rappelant de la résolution du primal,

$$\begin{aligned}y_1 &= -1 - s_1 \\&= -1 - \frac{\mu}{x_1(\mu)} \\&= -1 - \frac{2\mu}{1 - 2\mu \pm \sqrt{1 + 4\mu^2}}\end{aligned}$$

Comme $\mu \rightarrow 0$, $y_1 \rightarrow -1$, $y_2 \rightarrow 0$.

Il s'agit de l'unique solution du problème linéaire (les deux contraintes sont actives)

Si $\mu \rightarrow \infty$, y est non borné, car l'ensemble dual réalisable est non borné.

Chemin central primal-dual

Hypothèse : la région réalisable du problème de programmation linéaire a un intérieur non vide et un ensemble borné de solutions optimales.

Le dual a un intérieur réalisable non vide, en vertu des conditions (d'optimalité) sur le lagrangien.

Le chemin primal-dual est l'ensemble des vecteurs $(x(\mu), y(\mu), s(\mu))$ satisfaisant les conditions

$$\mathbf{x} \circ \mathbf{s} = \mu \mathbf{1}$$

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

$$\mathbf{A}^T \mathbf{y} + \mathbf{s} = \mathbf{c}$$

$$x \geq 0, s \geq 0$$

$$0 \leq \mu < \infty.$$

Proposition

Si les ensembles réalisables primal et dual ont des intérieurs non vides, alors le chemin central $(x(\mu), y(\mu), s(\mu))$ existe pour tout μ , $0 \leq \mu < \infty$.

De plus, $x(\mu)$ est le chemin central primal, $(y(\mu), s(\mu))$ est le chemin central dual.

$x(\mu)$ et $(y(\mu), s(\mu))$ convergent vers les centres analytiques des faces des solutions optimales primales et duales, respectivement, quand $\mu \rightarrow 0$.

Saut de dualité

Soit $(x(\mu), y(\mu), s(\mu))$ sur le chemin central primal-dual. Nous avons

$$\begin{aligned} c^T x - y^T b &= (A^T y)^T x + s^T x - y^T b \\ &= y^T A x + s^T x - y^T b \\ &= y^T b + s^T x - y^T b \\ &= s^T x \\ &= n\mu. \end{aligned}$$

Comme pour la dualité faible, $c^T x \geq y^T b$, et $n\mu$ est appelé le saut de dualité.

Soit $g = c^T x - y^T b$.

Comme $y^T b \leq z^*$, $z^* \geq c^T x - g$, et donc, étant donné (x, y, s) , on peut mesurer la qualité de x comme $c^T x - z^* \leq g$.