Méthode affine primale
Introduction

Fabian Bastin
DIRO
Université de Montréal

Méthode affine primale

Section basée sur les notes de Gilles Savard :
`http://www.iro.umontreal.ca/~marcotte/Ift6511/Pts_interieurs.pdf`.
Voir aussi `https://nptel.ac.in/courses/106108056/module9/LinearProgramming_IV.pdf`

- Présentée uniquement comme introduction aux méthodes de points intérieurs.
- Méthode simple et relative efficacité pratique.
- La complexité polynomiale n'est pas connue pour l'approche primale (ou duale).
- Analyse de convergence complexe.
- Publié pour la première en 1967 par I. I. Dikin, mais ignoré jusqu'au milieu des années 1980.

Principe

L'approche consiste à calculer une direction de descente (i.e. qui permet de réduire la valeur de l'objectif) qui n'approche pas trop rapidement de la frontière.

Trois étapes :

1. calcul de la direction de descente,
2. calcul du pas,
3. calcul de la transformation affine.

Intérieur de l'ensemble des contraintes

Considérons l'ensemble réalisable défini de manière général par des contraintes d'égalité et des contraintes d'inégalité :

$$\mathcal{A} = \left\{ x \;\middle|\; \begin{cases} g_i(x) = 0, & i = 1, \ldots, m \\ h_j(x) \geq 0, & j = 1, \ldots, n \end{cases} \right\}$$

L'intérieur de $\mathcal{A}$, noté notamment $int\mathcal{A}$, est

$$int\mathcal{A} = \left\{ x \;\middle|\; \begin{cases} g_i(x) = 0, & i = 1, \ldots, m \\ h_j(x) > 0, & j = 1, \ldots, n \end{cases} \right\}$$

Avec un problème linéaire sous forme standard, nous avons

$$\mathcal{A} = \{x \,|\, Ax = b, \; x \geq 0\}$$

et

$$int\mathcal{A} = \{x \,|\, Ax = b, \; x > 0\}.$$

Formulation

Considérons à nouveau le primal

$$\begin{aligned} \min_x \; & c^T x \\ \text{t.q. } & Ax = b \\ & x \geq 0, \end{aligned}$$

avec A de rang plein. On suppose aussi que $\mathcal{A}$ et $int\mathcal{A}$ sont non vides.

Direction de descente

Soit x_c le point courant t.q.

$$Ax_c = b, \;\; x_c > 0.$$

On cherche

$$x^+ = x_c + \alpha \Delta x \quad \text{ t.q. } \quad c^T x^+ \leq c^T x_c, \;\; Ax^+ = b.$$

Le déplacement doit donc vérifier

$$\begin{aligned} c^T \Delta x &\leq 0 \\ Ax^+ &= A(x_c + \alpha \Delta x) = b \end{aligned}$$

Direction de descente

Sous l'hypothèse que $\alpha > 0$, Δx doit être dans le noyau de A, i.e.

$$\Delta x \in \mathcal{N}(A) = \{x \in \mathbb{R}^n \mid Ax = 0\}.$$

La direction de plus forte descente est donnée par

$$\begin{aligned} \min_{\Delta x} \; & c^T \Delta x \\ \text{t.q.} \; & A\Delta x = 0 \\ & \|\Delta x\| = 1. \end{aligned}$$

La solution est

$$\frac{\mathrm{proj}_A(-c)}{\|\mathrm{proj}_A(-c)\|} = \Delta x,$$

$\mathrm{proj}_A(\cdot)$ étant la matrice orthogonale de projection sur le noyau de A.

Direction de descente

Comme A est de plein rang et en oubliant la normalisation, il est possible de montrer que

$$\Delta x = \text{proj}_A(-c) = -(I - A^T(AA^T)^{-1}A)c.$$

Note : une matrice orthogonale de projection P sur le noyau de A vérifie les propriétés suivantes :

$$\begin{aligned} AP &= 0 \\ P &= P^T \\ P^2 &= P \qquad \text{(matrice idempotente).} \end{aligned}$$

Notons $P = I - A^T(AA^T)^{-1}A$. Alors

$$c^T\Delta x = -c^T Pc = -c^T P^2 c = -\|Pc\|_2^2 \leq 0.$$

Longueur de pas

Le taux de décroissance est constant dans la direction Δx. Le pas maximal est limité par les contraintes de non-négativité. De plus, la transformation affine qui nous permettra de centrer le point n'est pas définie sur la frontière.

Il suffit de choisir

$$\alpha = \gamma \alpha_{\mathsf{max}},$$

avec $0 < \gamma < 1$, et

$$\alpha_{\mathsf{max}} = \min_{\Delta x_i < 0} \frac{-(x_c)_i}{\Delta x_i}.$$

En pratique, on choisit $\gamma = 0.995$. Si $\Delta x \geq 0$ et $\Delta x \neq 0$, le problème est non borné.

Transformation affine

- Il s'agit de faire une mise à l'échelle afin que le nouveau point x^+ soit loin de la frontière définie par $x \geq 0$.
- Un point idéal serait le vecteur unitaire e. Ainsi, on cherche une transformation affine qui transforme x^+ en e. La matrice de transformation est simplement l'inverse de la matrice diagonale dont les composantes sont les mêmes que celles de x^+ (qui sont > 0). Notons cette matrice par X. On a alors $X^{-1}x^+ = e$, et notons
$$X^{-1}x = \overline{x}.$$
- Dans l'espace transformé, le programme devient
$$\begin{aligned} \min_{\overline{x}} \ & c^T X \overline{x} = \overline{c}^T \overline{x} \\ \text{t.q. } & AX\overline{x} = \overline{A}\overline{x} = b \\ & \overline{x} \geq 0. \end{aligned}$$

Calcul du nouveau point

Dans l'espace-$\overline{x}$, on obtient

$$\begin{aligned}
\Delta\overline{x} &= \mathrm{proj}_{\overline{A}}(-\overline{c}) \\
&= -(I - \overline{A}^T(\overline{AA}^T)^{-1}\overline{A})\overline{c} \\
&= -(I - XA^T(AX^2A^T)^{-1}AX)Xc
\end{aligned}$$

et

$$\overline{x}^+ = \overline{x}_c + \alpha\Delta\overline{x},$$

d'où

$$\begin{aligned}
x^+ &= X\overline{x}^+ \\
&= x_c + \alpha X\Delta\overline{x} \\
&= x_c - \alpha X(I - XA^T(AX^2A^T)^{-1}AX)Xc.
\end{aligned}$$

Convergence

Supposons A plein rang, qu'il existe un point strictement intérieur et que la fonction objectif est non constante sur le domaine réalisable. Alors

1. Si le problème primal et le problème dual sont non dégénérés, alors pour tout $\gamma < 1$, la suite générée par l'algorithme converge vers une solution optimale.
2. Pour $\gamma \leq 2/3$, la suite produite par l'algorithme converge vers une solution optimale (y compris en présence de dégénérescence).

Preuve : admis !

Critère d'arrêt

Basé sur la satisfaction des conditions d'optimalité.

Considérons d'abord le cas non dégénéré (primal et dual).

Programme dual :

$$\begin{aligned} \max_{y,s}\ & b^T y \\ \text{t.q. } & A^T y + s = c \\ & s \geq 0. \end{aligned}$$

Soit x_k la solution courante et considérons le programme

$$\begin{aligned} \min_{y,s}\ & \|X_k s\| \\ \text{t.q. } & A^T y + s = c \\ & s \geq 0. \end{aligned}$$

Critère d'arrêt

Problème non linéaire !

Il est possible de montrer que la solution est donnée par

$$
\begin{aligned}
y_k &= (AX_k^2A^T)^{-1}AX_k^2c \\
s_k &= c - A^Tyk
\end{aligned}
$$

Dès lors

$$
s_k = c - A^T(AX_k^2A^T)^{-1}AX_k^2c = (I - A^T(AX_k^2A^T)^{-1}AX_k^2)c
$$

et

$$
\begin{aligned}
-X_k^2s_k &= -X_k^2(I - A^T(AX_k^2A^T)^{-1}AX_k^2)c \\
&= -X_k(I - X_kA^T(AX_k^2A^T)^{-1}AX_k)X_kc \\
&= \Delta x_k.
\end{aligned}
$$

Le vecteur dual est obtenu comme sous-produit du calcul de la direction de descente.

Convergence

Sous l'hypothèse de non dégénérescence primale et duale, la solution (y_k, s_k) converge vers la solution optimale duale lorsque x_k converge vers la solution optimale primale.

Dans ce cas, un critère d'arrêt peut être défini sur la norme du vecteur de complémentarité.

Dans le cas dégénéré, la convergence de (y_k, s_k) n'est pas assurée, et un critère d'arrêt classique sur l'amélioration successive de deux itérés est utilisé.

Algorithme

Soit $\gamma \in (0,1)$, $\epsilon > 0$ et un point initial x_0. Poser

$$x_c := x_0$$
$$\Delta c := \epsilon |c^T x_c| + 2$$

Tant que $\Delta c > \epsilon \max\{|c^T x_c|, 1\}$ répéter

$$X := X_c$$
$$\Delta x := -X(I - XA^T(AX^2A^T)^{-1}AX)Xc$$
$$\alpha := \gamma \min_{\Delta x_i < 0} \frac{-(x_c)_i}{\Delta x_i}$$
$$x^+ := x_c + \alpha \Delta x$$
$$\Delta c := c^T x_c - c^T x^+$$
$$x_c := x^+.$$

Algorithme

On peut légèrement simplifier les opérations comme suit.
Soit $\gamma \in (0,1)$, $\epsilon > 0$ et un point initial x_0. Poser

$$x_c := x_0$$
$$\Delta c := \epsilon |c^T x_c| + 2$$

Tant que $\Delta c > \epsilon \max\{|c^T x_c|, 1\}$ répéter

$$X := X_c$$
$$\Delta x := -(I - XA^T(AX^2A^T)^{-1}AX)Xc$$
$$\alpha := \gamma \min_{\Delta x_i < 0} \frac{-1}{\Delta x_i}$$
$$x^+ := x_c + \alpha X \Delta x$$
$$\Delta c := c^T x_c - c^T x^+$$
$$x_c := x^+.$$

Initialisation

Une méthode de phase I peut également être développée pour trouver un point réalisable sans nécessairement qu'il soit solution de base. Voir Vanderbei, Chapitre 21, Section 5.