学習・研究用テキスト 内点法 (2C)

主パス追跡法

水野 眞治

東京工業大学 大学院社会理工学研究科 経営工学専攻

http://www.me.titech.ac.jp/~mizu_lab/text/

2010 年 11 月 9 日

概要

標準形の線形計画問題を解くための主パス追跡法を解説する．アルゴリズムの計算手順を述べるだけでなく，反復回数が $O(\sqrt{n}L)$ となることを示す．

目次

1 主パス追跡法

n 個の変数を持ち，m 個の等式制約を持つ標準形の線形計画問題を

$$\begin{array}{ll} \text{最小化} & \boldsymbol{c}^T\boldsymbol{x} \\ \text{制約条件} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0} \end{array} \tag{1}$$

とする．上の問題を主問題とすれば，その双対問題は

$$\begin{array}{ll} \text{最大化} & \boldsymbol{b}^T\boldsymbol{y} \\ \text{制約条件} & \boldsymbol{A}^T\boldsymbol{y} + \boldsymbol{z} = \boldsymbol{c} \\ & \boldsymbol{z} \geq \boldsymbol{0} \end{array} \tag{2}$$

となる．

点 $\boldsymbol{x} \in \mathcal{R}^n$ は，$\boldsymbol{x} > \boldsymbol{0}$ を満たすならば，**内点**と呼ばれ，さらに $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ を満たすならば主問題 (1) の**実行可能内点**と呼ばれ，満たさないならば**実行不能内点**と呼ばれる．同様に，点 $(\boldsymbol{y}, \boldsymbol{z}) \in \mathcal{R}^m \times \mathcal{R}^n$ は，$\boldsymbol{z} > \boldsymbol{0}$ を満たすならば，内点と呼ばれ，さらに $\boldsymbol{A}^T\boldsymbol{y} + \boldsymbol{z} = \boldsymbol{c}$ を満たすならば双対問題 (2) の実行可能内点と呼ばれる．線形計画問題 (1) に次の基本的な仮定をおく．

仮定 1.1 行列 $\boldsymbol{A}$ のランクが m である.

1.1 アルゴリズム

線形計画問題の実行可能領域の内部に中心パスと呼ばれるなめらかなパスが存在し，このパスの一方の端点が最適解に収束する．パス追跡法は，中心パスを追跡することにより，最適解の近似解を求める方法である．

n 次元ユークリッド空間の正象限 $\mathcal{R}^n_+ = \{\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^T \in \mathcal{R}^n : \boldsymbol{x} \geq \boldsymbol{0}\}$ の対数障壁関数を

$$p(\boldsymbol{x}) = -\sum_{i=1}^{n} \log x_i \tag{3}$$

とする．これは，$\mathcal{R}^n_+$ の内部を定義域とし，$\boldsymbol{x}$ がその境界に近づくときに発散する狭義凸関数である．線形計画問題 (1) の目的関数にパラメータ $\mu > 0$ を重みとして対数障壁関数 p を加えた問題

$$\begin{array}{ll} \text{最小化} & \boldsymbol{c}^T\boldsymbol{x} - \mu \sum_{i=1}^{n} \log x_i \\ \text{制約条件} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \end{array} \tag{4}$$

を考える．ここで，$p(\boldsymbol{x})$ の定義域は，$\boldsymbol{x} > \boldsymbol{0}$ である．

仮定 1.2 問題 (1) に実行可能内点 $\boldsymbol{x}^0$ と最適解が存在し，最適解の集合が有界である．

このとき，問題 (4) は，凸計画問題であり，唯一つの最適解を持つ．問題 (4) の最適条件は，制約条件 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ のラグランジュ乗数を $\boldsymbol{y} \in \mathcal{R}^m$ とすれば

$$\begin{array}{l} \boldsymbol{c} - \mu\boldsymbol{X}^{-1}\boldsymbol{e} - \boldsymbol{A}^T\boldsymbol{y} = \boldsymbol{0} \\ \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \end{array} \tag{5}$$

と表される．$\boldsymbol{z} = \mu\boldsymbol{X}^{-1}\boldsymbol{e}$ とすれば,

$$\begin{array}{l} \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ \boldsymbol{A}^T\boldsymbol{y} + \boldsymbol{z} = \boldsymbol{c} \\ \boldsymbol{X}\boldsymbol{z} = \mu\boldsymbol{e} \end{array} \tag{6}$$

と書き換えられる．この条件と $\boldsymbol{x} > \boldsymbol{0}$，$\boldsymbol{z} > \boldsymbol{0}$ を満たす解を $(\boldsymbol{x}(\mu), \boldsymbol{y}(\mu), \boldsymbol{z}(\mu))$ とする．$\boldsymbol{x}(\mu)$ は凸計画問題 (4) の唯一つの解であり，解析的中心と呼ばれる．任意の $\mu > 0$ に対して，解析的中心 $\boldsymbol{x}(\mu)$ が唯一つ存在するので，集合 $P = \{\boldsymbol{x}(\mu) : \mu > 0\}$ はパスになる．これを主問題の中心パスと呼ぶ．問題 (6) の条件は，$\mu \to 0$ のときに線形計画問題の最適

条件に一致する．このことから推察できるように，$\mu \to 0$ のとき $\boldsymbol{x}(\mu)$ は主問題 (1) の最適解に収束し，$(\boldsymbol{y}(\mu), \boldsymbol{z}(\mu))$ は双対問題 (2) の最適解に収束する．

パラメータ $\beta \in (0,1)$ を使って関係式 (6) を緩めることにより，解析的中心 $\boldsymbol{x}(\mu)$ の近傍

$$N(\mu,\beta) = \{\boldsymbol{x} | \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{A}^T\boldsymbol{y} + \boldsymbol{z} = \boldsymbol{c}, \|\boldsymbol{X}\boldsymbol{z} - \mu\boldsymbol{e}\| \leq \beta\mu, \boldsymbol{x} > \boldsymbol{0}, \boldsymbol{z} > \boldsymbol{0}\}$$

を定義する．また，初期実行可能内点 $\boldsymbol{x}^0$ が，ある μ_0 と $\beta \in (0,1)$ に対して，近傍 $N(\mu_0,\beta)$ に含まれるとする．パス追跡アルゴリズムでは，この初期点 $\boldsymbol{x}^0$ とパラメータ値 $\mu_0 > 0$ が与えられたとき，数列 $\{\mu_k\}$ が 0 に収束するように各反復で μ_k を更新し，その μ_k に対する解析的中心の近傍 $N(\mu_k,\beta)$ 上の点 $\boldsymbol{x}^k$ を求める．このとき，μ_k が十分小さくなれば，得られた点 $\boldsymbol{x}^k$ は問題 (1) の近似解となる．

アルゴリズム 1.3 主問題のパス追跡法は，次のステップからなる．

ステップ 0 ある $\beta \in (0,1)$ と $\mu_0 > 0$ に対して，初期実行可能内点を $\boldsymbol{x}^0 \in N(\mu_0,\beta)$ とする．$\gamma \in (0,1)$，$k = 0$ とする．

ステップ 1 点 $\boldsymbol{x}^k$ から，$\mu = \gamma\mu_k$ のときの問題 (4) の近似解 $\boldsymbol{x}^{k+1} \in N(\mu,\beta)$ を求める．

ステップ 2 $\mu_{k+1} = \mu$ とする．k を一つ増加して，ステップ 1 へ行く．

1.2 点の更新と反復回数

ここでは，第 k 反復におけるパラメータ $\mu_k > 0$ と解析的中心の近傍上の点 $\boldsymbol{x}^k \in N(\mu_k,\beta)$ が与えられていると仮定し，パラメータを $\mu = \gamma\mu_k$ と更新したときの近傍上の点 $\boldsymbol{x}^{k+1} \in N(\mu,\beta)$ の求め方を示す．まずはじめに，考え方と計算方法を示した後に，次節で $\boldsymbol{x}^{k+1}$ が近傍上にあることを示す．

解析的中心は，問題 (4)，つまり

$$\begin{array}{ll} \text{最小化} & \boldsymbol{c}^T\boldsymbol{x} - \mu\sum_{i=1}^n \log x_i \\ \text{制約条件} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \end{array} \tag{7}$$

の解である．これは，非線形計画問題であるので，簡単に解くことはできない．そこで，現在の点 $\boldsymbol{x}^k$ から，最適化問題に対するニュートン法を使うことにする．すなわち，目的関数を 2 次関数で近似し，その 2 次計画問題を解くことにする．目的関数を

$$f(\boldsymbol{x}) = \boldsymbol{c}^T\boldsymbol{x} - \mu\sum_{i=1}^n \log x_i$$

とおき，それを点 $\boldsymbol{x}^k$ で 2 次近似すれば

$$
\begin{aligned}
& f(\boldsymbol{x}^k) + \nabla f(\boldsymbol{x}^k)^T(\boldsymbol{x}-\boldsymbol{x}^k) + \frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}^k)^T\nabla^2 f(\boldsymbol{x}^k)(\boldsymbol{x}-\boldsymbol{x}^k) \\
&= f(\boldsymbol{x}^k) + (\boldsymbol{c}-\mu\boldsymbol{X}_k^{-1}\boldsymbol{e})^T(\boldsymbol{x}-\boldsymbol{x}^k) + \frac{\mu}{2}(\boldsymbol{x}-\boldsymbol{x}^k)^T\boldsymbol{X}_k^{-2}(\boldsymbol{x}-\boldsymbol{x}^k)
\end{aligned}
$$

となる．ここで，$\boldsymbol{d} = \boldsymbol{x}-\boldsymbol{x}^k$ とおき，定数項 $f(\boldsymbol{x}^k)$ を除いた部分を $\boldsymbol{d}$ の 2 次関数と見て

$$
g(\boldsymbol{d}) = (\boldsymbol{c}-\mu\boldsymbol{X}_k^{-1}\boldsymbol{e})^T\boldsymbol{d} + \frac{\mu}{2}\boldsymbol{d}^T\boldsymbol{X}_k^{-2}\boldsymbol{d}
$$

とおく．このとき，$\boldsymbol{A}\boldsymbol{d} = \boldsymbol{A}(\boldsymbol{x}-\boldsymbol{x}^k) = \boldsymbol{b}-\boldsymbol{b} = \boldsymbol{0}$ であるから，(7) を 2 次近似した問題は,

$$
\begin{array}{ll}
\text{最小化} & (\boldsymbol{c}-\mu\boldsymbol{X}_k^{-1}\boldsymbol{e})^T\boldsymbol{d} + \frac{\mu}{2}\boldsymbol{d}^T\boldsymbol{X}_k^{-2}\boldsymbol{d} \\
\text{制約条件} & \boldsymbol{A}\boldsymbol{d} = \boldsymbol{0}
\end{array}
\tag{8}
$$

となる．これは，等式制約のみの凸 2 次計画問題であるから，その最適解は，ラグランジュ関数

$$
L(\boldsymbol{d},\boldsymbol{y}) = (\boldsymbol{c}-\mu\boldsymbol{X}_k^{-1}\boldsymbol{e})^T\boldsymbol{d} + \frac{\mu}{2}\boldsymbol{d}^T\boldsymbol{X}_k^{-2}\boldsymbol{d} - \boldsymbol{y}^T\boldsymbol{A}\boldsymbol{d}
$$

に対して

$$
\begin{array}{l}
\nabla_{\boldsymbol{d}} L = \boldsymbol{c} - \mu\boldsymbol{X}_k^{-1}\boldsymbol{e} + \mu\boldsymbol{X}_k^{-2}\boldsymbol{d} - \boldsymbol{A}^T\boldsymbol{y} = \boldsymbol{0} \\
\nabla_{\boldsymbol{y}} L = \boldsymbol{A}\boldsymbol{d} = \boldsymbol{0}
\end{array}
\tag{9}
$$

をみたす．これより

$$
\begin{array}{l}
\boldsymbol{y} = (\boldsymbol{A}\boldsymbol{X}_k^2\boldsymbol{A}^T)^{-1}\boldsymbol{A}\boldsymbol{X}_k^2(\boldsymbol{c}-\mu\boldsymbol{X}_k^{-1}\boldsymbol{e}) \\
\boldsymbol{d} = \frac{1}{\mu}\boldsymbol{X}_k^2(\boldsymbol{A}^T\boldsymbol{y} + \mu\boldsymbol{X}_k^{-1}\boldsymbol{e} - \boldsymbol{c})
\end{array}
\tag{10}
$$

となる (演習問題)．ここで,

$$
\begin{array}{l}
\boldsymbol{x}' = \boldsymbol{x}^k + \boldsymbol{d} \\
\boldsymbol{z}' = \boldsymbol{c} - \boldsymbol{A}^T\boldsymbol{y} = \mu\boldsymbol{X}_k^{-1}\boldsymbol{e} - \mu\boldsymbol{X}_k^{-2}\boldsymbol{d}
\end{array}
\tag{11}
$$

とおけば，$\boldsymbol{A}\boldsymbol{x}' = \boldsymbol{b}$，$\boldsymbol{A}^T\boldsymbol{y} + \boldsymbol{z}' = \boldsymbol{c}$ を満たす．以下では，パラメータ $\beta \in (0,1)$ と $\gamma \in (0,1)$ を適当に選び，$\mu = \gamma\mu_k$ とすれば，$\boldsymbol{x}'$ が不等式 $\boldsymbol{x}' > \boldsymbol{0}$ を満たし主問題 (1) の実行可能解となり，$\boldsymbol{z}'$ が不等式 $\boldsymbol{z}' > \boldsymbol{0}$ を満たし，$(\boldsymbol{y},\boldsymbol{z}')$ が双対問題 (2) の実行可能解となり，さらに

$$
\|\boldsymbol{X}'\boldsymbol{z}' - \mu\boldsymbol{e}\| \leq \beta\mu
$$

を満たすこと，すなわち $\boldsymbol{x}' \in N(\mu,\beta)$ となることを示す.

補題 1.4 上のように $\boldsymbol{d}$ を定義すれば,

$$\|\boldsymbol{X}_k^{-1}\boldsymbol{d}\| \leq \frac{1}{\mu}\|\boldsymbol{X}_k\boldsymbol{z}^k - \mu\boldsymbol{e}\|$$

が成立する.

証明 仮定より，$\boldsymbol{x}^k \in N(\mu_k, \beta)$ であるから，ある $(\boldsymbol{y}^k, \boldsymbol{z}^k)$ が存在し,

$$\|\boldsymbol{X}_k\boldsymbol{z}^k - \mu_k\boldsymbol{e}\| \leq \beta\mu_k,\ \boldsymbol{A}^T\boldsymbol{y}^k + \boldsymbol{z}^k = \boldsymbol{c},\ \boldsymbol{z}^k > \boldsymbol{0}$$

となる．これより，$\boldsymbol{A}\boldsymbol{d} = \boldsymbol{0}$ を使うと，関係式

$$\begin{aligned}
\|\boldsymbol{X}_k^{-1}\boldsymbol{d}\|^2 &= \boldsymbol{d}^T\boldsymbol{X}_k^{-2}\boldsymbol{d} \\
&= \frac{1}{\mu}\boldsymbol{d}^T(\boldsymbol{A}^T\boldsymbol{y} + \mu\boldsymbol{X}_k^{-1}\boldsymbol{e} - \boldsymbol{c}) \\
&= \frac{1}{\mu}\boldsymbol{d}^T(\mu\boldsymbol{X}_k^{-1}\boldsymbol{e} - \boldsymbol{A}^T\boldsymbol{y}^k - \boldsymbol{z}^k) \\
&= \frac{1}{\mu}\boldsymbol{d}^T\boldsymbol{X}_k^{-1}(\mu\boldsymbol{e} - \boldsymbol{X}_k\boldsymbol{z}^k) \\
&\leq \frac{1}{\mu}\|\boldsymbol{X}_k^{-1}\boldsymbol{d}\|\|\boldsymbol{X}_k\boldsymbol{z}^k - \mu\boldsymbol{e}\|
\end{aligned}$$

が得られる．この両辺を $\|\boldsymbol{X}_k^{-1}\boldsymbol{d}\|$ で割ることにより，補題の結果を得る． ▮

上の補題の右辺にあるノルムの大きさを評価すると

$$\begin{aligned}
\|\boldsymbol{X}_k\boldsymbol{z}^k - \mu\boldsymbol{e}\| &\leq \|\boldsymbol{X}_k\boldsymbol{z}^k - \mu_k\boldsymbol{e}\| + |\mu_k - \mu|\|\boldsymbol{e}\| \\
&\leq \beta\mu_k + |\mu_k - \mu|\sqrt{n}
\end{aligned}$$

となる．したがって，次の結果が得られる.

定理 1.5 パラメータ $\beta = 0.5$ と $\gamma = 1 - \frac{1}{10\sqrt{n}}$ とし，$\mu = \gamma\mu_k$ とすれば，不等式

$$\|\boldsymbol{X}_k^{-1}\boldsymbol{d}\| \leq \frac{2}{3}$$

が成立する．これより，$\boldsymbol{x}'$ は，不等式 $\boldsymbol{x}' > \boldsymbol{0}$ を満たし，主問題 (1) の実行可能内点となり，$\boldsymbol{z}'$ は，不等式 $\boldsymbol{z}' > \boldsymbol{0}$ を満たし，双対問題 (2) の実行可能内点となる.

証明 $\beta = 0.5$ と $\gamma = 1 - \frac{1}{10\sqrt{n}}$ とし，$\mu = \gamma\mu_k$ とすれば，$\frac{\mu}{\mu_k} \geq \frac{9}{10}$ となるので，補題 1.4 と定理の上の不等式から

$$\|\boldsymbol{X}_k^{-1}\boldsymbol{d}\| \leq \frac{1}{\mu}(\beta\mu_k + |\mu_k - \mu|\sqrt{n})$$

$$
\begin{aligned}
&\le \frac{\mu_k}{\mu}\left(\frac{1}{2}+\frac{1}{10}\right)\\
&\le \frac{2}{3}
\end{aligned}
$$

となる．これより

$$\boldsymbol{x}' = \boldsymbol{x}_k + \boldsymbol{d} = \boldsymbol{X}_k(\boldsymbol{e} + \boldsymbol{X}_k^{-1}\boldsymbol{d}) > \boldsymbol{0}$$

となり

$$\boldsymbol{z}' = \mu\boldsymbol{X}_k^{-1}\boldsymbol{e} - \mu\boldsymbol{X}_k^{-2}\boldsymbol{d} = \mu\boldsymbol{X}_k^{-1}(\boldsymbol{e} - \boldsymbol{X}_k^{-1}\boldsymbol{d}) > \boldsymbol{0}$$

となる．また，$\boldsymbol{A}\boldsymbol{x}' = \boldsymbol{b}$ と $\boldsymbol{A}^T\boldsymbol{y} + \boldsymbol{z}' = \boldsymbol{c}$ は，定義 (11) より成り立つ． ▮

定理 1.6 パラメータ $\beta = 0.5$ と $\gamma = 1 - \frac{1}{10\sqrt{n}}$ とし，$\mu = \gamma\mu_k$ とすれば，$\boldsymbol{x}'$ と $\boldsymbol{z}'$ が

$$\|\boldsymbol{X}'\boldsymbol{z}' - \mu\boldsymbol{e}\| \le \beta\mu$$

を満たす.

証明 $\boldsymbol{D} = \mathrm{diag}(\boldsymbol{d})$ とすれば，定義 (11) より

$$
\begin{aligned}
\frac{1}{\mu}\|\boldsymbol{X}'\boldsymbol{z}' - \mu\boldsymbol{e}\| &= \frac{1}{\mu}\|(\boldsymbol{X}_k + \boldsymbol{D})(\mu\boldsymbol{X}_k^{-1}\boldsymbol{e} - \mu\boldsymbol{X}_k^{-2}\boldsymbol{d}) - \mu\boldsymbol{e}\|\\
&= \|\boldsymbol{e} - \boldsymbol{X}_k^{-1}\boldsymbol{d} + \boldsymbol{D}\boldsymbol{X}_k^{-1}\boldsymbol{e} - \boldsymbol{D}\boldsymbol{X}_k^{-2}\boldsymbol{d} - \boldsymbol{e}\|\\
&= \|\boldsymbol{D}\boldsymbol{X}_k^{-2}\boldsymbol{d}\|\\
&\le \|\boldsymbol{D}\boldsymbol{X}_k^{-2}\boldsymbol{d}\|_1\\
&= \boldsymbol{e}^T\boldsymbol{D}\boldsymbol{X}_k^{-2}\boldsymbol{d}\\
&= \|\boldsymbol{X}_k^{-1}\boldsymbol{d}\|^2\\
&\le \left(\frac{2}{3}\right)^2 \le \beta
\end{aligned}
$$

となる． ▮

上の定理 1.5 と 1.6 より，$\beta = 0.5$ と $\gamma = 1 - \frac{1}{10\sqrt{n}}$ とし，$\mu = \gamma\mu_k$ とすれば，アルゴリズム 1.3 により $\boldsymbol{x}^k \in N(\mu_k, \beta)$ を満たす点列 $\{\boldsymbol{x}^k\}$ が求められ，

$$\mu_k = \left(1 - \frac{1}{10\sqrt{n}}\right)^k \mu^0$$

が成立する．したがって，理論的には，$O(\sqrt{n}L)$ 反復で最適解を得ることができる．

パス追跡法について，より詳しくは Gonzaga [1, 2] を参照することを薦める．

演習問題 1.7 式 (9) の解が式 (10) となることを示せ．

参考文献

[1] Gonzaga, C. C.: “Large Step Path-Following Methods for Linear Programming, Part I: Barrier Function Method”, *SIAM Journal on Optimization* 1 (1991) 268–279.

[2] Gonzaga C. C.: “Path Following Methods for Linear Programming”, *SIAM Review* **34** (1992) 167–227.