Linear Programming

1　主問題

主問題は以下のように定義される．

$$
\begin{aligned}
(\mathrm{P}):\quad & \text{Minimize } \boldsymbol{c}^\top \boldsymbol{x} \\
& \text{subject to } A\boldsymbol{x} = \boldsymbol{b} \\
& \qquad\qquad\quad \boldsymbol{x} \geq \boldsymbol{0}
\end{aligned} \tag{1}
$$

但し，最適化変数を $\boldsymbol{x} = \begin{bmatrix} x_1 \cdots x_n \end{bmatrix}^\top$ としている．制約が線形関数を用いて $h_l(x) \leq 0$ となるような場合，スラック変数の導入や変数変換によって上記の等式標準形まで変形することができる．

2　双対問題

主問題に対する Lagrange 双対問題は以下のようになる（Fenchel 双対問題でも同じ結果となる）．

$$
\begin{aligned}
(\mathrm{D}):\quad & \text{Maximize } \boldsymbol{b}^\top \boldsymbol{y} \\
& \text{subject to } A^\top \boldsymbol{y} + \boldsymbol{s} = \boldsymbol{c} \\
& \qquad\qquad\quad \boldsymbol{s} \geq \boldsymbol{0}
\end{aligned} \tag{2}
$$

最適化変数は $\boldsymbol{y} \in \mathbb{R}^m$ であり，$\boldsymbol{s} \in \mathbb{R}^n$ はスラック変数．（$\boldsymbol{y}$ 及び $\boldsymbol{s}$ はそれぞれ，主問題 (1) の等式制約と不等式制約に対する Lagrange 乗数に相当する）

3　双対定理

3.1　弱双対定理

主問題 (1) の任意の実行可能解 $\boldsymbol{x}$ と双対問題 (2) の任意の実行可能解 $\boldsymbol{y}$ に対して，不等式

$$
\boldsymbol{c}^\top \boldsymbol{x} \geq \boldsymbol{b}^\top \boldsymbol{y} \tag{3}
$$

が成り立つ．

3.2　強双対定理

主問題 (1) と双対問題 (2) がともに実行可能解をもつことと，両者に最適解が存在して最適解が一致することとは，同値である．

4　最適性条件

4.1　相補性条件

ベクトル $\boldsymbol{c} - A^\top \boldsymbol{y}$ の第 j 成分を $(\boldsymbol{c} - A^\top \boldsymbol{y})_j$ で表す．主問題 (1) の実行可能解 $\boldsymbol{x}$ と双対問題 (2) の実行可能解 $\boldsymbol{y}$ がそれぞれの問題の最適解であるための必要十分条件は，条件

$$
x_j(\boldsymbol{c} - A^\top \boldsymbol{y})_j = 0, \quad j = 1, ..., n \tag{4}
$$

が成り立つことである．

最適解であるための必要十分条件は，以下のように主問題と双対問題の組にすると見やすい．

4.2 最適性条件

$\boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{y} \in \mathbb{R}^m, \boldsymbol{s} \in \mathbb{R}^n$ がそれぞれ主問題 (1) と双対問題 (2) の最適解であるための必要十分条件は，これらが

$$\begin{aligned} &A\boldsymbol{x} = \boldsymbol{b} \\ &A^\top \boldsymbol{y} + \boldsymbol{s} = \boldsymbol{c} \\ &x_j s_j = 0, \quad j = 1, ..., n \\ &\boldsymbol{x} \geq \boldsymbol{0}, \quad \boldsymbol{s} \geq \boldsymbol{0} \end{aligned} \tag{5}$$

を満たすことである．

更に $X = \mathrm{diag}(\boldsymbol{x})$ とすれば，最適化条件は以下のように書き下せる．

$$\begin{aligned} &A\boldsymbol{x} = \boldsymbol{b} \\ &A^\top \boldsymbol{y} + \boldsymbol{s} = \boldsymbol{c} \\ &X\boldsymbol{s} = \boldsymbol{0} \\ &\boldsymbol{x} \geq \boldsymbol{0}, \quad \boldsymbol{s} \geq \boldsymbol{0} \end{aligned} \tag{6}$$

これを関数を定義して，以下のようにも書き下せる．

$$\begin{aligned} &h(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s}) = \boldsymbol{0} \\ &\boldsymbol{x} \geq \boldsymbol{0}, \quad \boldsymbol{s} \geq \boldsymbol{0} \end{aligned} \tag{7}$$

5 主双対内点法

バリアパラメータ ν を導入して，以下のような最適性条件を考える．

$$\begin{aligned} &A\boldsymbol{x} = \boldsymbol{b} \\ &A^\top \boldsymbol{y} + \boldsymbol{s} = \boldsymbol{c} \\ &X\boldsymbol{s} = \nu\boldsymbol{1} \\ &\boldsymbol{x} > \boldsymbol{0}, \quad \boldsymbol{s} > \boldsymbol{0} \end{aligned} \tag{8}$$

内点実行可能解（$\boldsymbol{x} > \boldsymbol{0}, \boldsymbol{s} > \boldsymbol{0}$ を満たす実行可能解）が存在するならば，各 $\nu > 0$ に対して最適解が一意に存在する．その解 $(\boldsymbol{x}(\nu), \boldsymbol{y}(\nu), \boldsymbol{s}(\nu))$ に対して，$\nu > 0$ を変化させたときの軌跡 $\{(\boldsymbol{x}(\nu), \boldsymbol{y}(\nu), \boldsymbol{s}(\nu)) | \nu > 0\}$ を主双対中心曲線と呼ぶ．主双対中心曲線は主問題と双対問題の実行可能領域の内部を通る滑らかな曲線であり，$\nu \searrow 0$ で最適解に収束する．

関数を用いて表せば，バリア関数を導入した場合の最適性条件は以下のように書ける．

$$\begin{aligned} &h(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s}) - \begin{bmatrix} 0 \\ 0 \\ \nu\boldsymbol{1} \end{bmatrix} = \boldsymbol{0} \\ &\boldsymbol{x} > \boldsymbol{0}, \quad \boldsymbol{s} > \boldsymbol{0} \end{aligned} \tag{9}$$

5.1 初期点

初期点としては内点 $\boldsymbol{x}_0 > \boldsymbol{0}, \boldsymbol{s}_0 > \boldsymbol{0}$ を満たすような初期点 $(\boldsymbol{x}_0, \boldsymbol{y}_0, \boldsymbol{s}_0)$ を用いる．（線形計画の？）主双対内点法（パス追従）においては初期点を内点実行可能解に選ぶ必要はない．

5.2 更新則

探索方向 $(\Delta\boldsymbol{x}, \Delta\boldsymbol{y}, \Delta\boldsymbol{s})$ とステップ幅 α_k を用いて以下のように更新する．

$$\begin{bmatrix} \boldsymbol{x}_{k+1} \\ \boldsymbol{y}_{k+1} \\ \boldsymbol{s}_{k+1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}_k \\ \boldsymbol{x}_k \\ \boldsymbol{x}_k \end{bmatrix} + \alpha_k \begin{bmatrix} \Delta\boldsymbol{x}_k \\ \Delta\boldsymbol{y}_k \\ \Delta\boldsymbol{s}_k \end{bmatrix} \tag{10}$$

ステップ幅 α については主問題と双対問題で別々のステップ幅を選んでもよい．つまり，

$$\begin{bmatrix} \boldsymbol{x}_{k+1} \\ \boldsymbol{y}_{k+1} \\ \boldsymbol{s}_{k+1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}_k \\ \boldsymbol{x}_k \\ \boldsymbol{x}_k \end{bmatrix} + \begin{bmatrix} \alpha_k^P \Delta \boldsymbol{x}_k \\ \alpha_k^D \Delta \boldsymbol{y}_k \\ \alpha_k^D \Delta \boldsymbol{s}_k \end{bmatrix} \tag{11}$$

としても良い．

5.3 バリアパラメータの更新

双対ギャップ一定の集合はバリアパラメータ ν を用いて，

$$F_{PD}(n\nu) = \{(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s}) \mid \boldsymbol{c}^\top \boldsymbol{x} - \boldsymbol{b}^\top \boldsymbol{y} = n\nu,\ A\boldsymbol{x} = \boldsymbol{b},\ A^\top \boldsymbol{y} + \boldsymbol{s} = \boldsymbol{c},\ \boldsymbol{x} \geq \boldsymbol{0},\ \boldsymbol{s} \geq \boldsymbol{0}\} \tag{12}$$

と表せる．バリアパラメータの更新がないと，双対ギャップ $n\nu$ が一定のパスを移動することになり，解析的中心に近づくが（パスに近づくが）目的関数である双対ギャップは小さくならない．双対ギャップの尺度 ρ_k を

$$\rho_k = \frac{\boldsymbol{x}_k^\top \boldsymbol{s}_k}{n} \tag{13}$$

として定義し，$\gamma \in (0, 1]$ を用いて，

$$\nu_{k+1} = \gamma \rho_k \tag{14}$$

として更新する．

5.4 探索方向

探索方向は Newton 方向を用いる．k 回目の反復において，$\nu = \nu_{k+1}$ とした最適性条件 (9) の解がわかればよいので（次のパスを目指す），点 $(\boldsymbol{x}_k, \boldsymbol{y}_k, \boldsymbol{s}_k)$ における，最適性条件 (9) の Newton 方向を考えればよい．非線形方程式 $f(\boldsymbol{u}) = \boldsymbol{0}$ の点 $\boldsymbol{u}'$ における Newton 方向 $\Delta\boldsymbol{u}$ は $\nabla f(\boldsymbol{u}')\Delta\boldsymbol{u} = -f(\boldsymbol{u}')$ の解であるから，最適性条件 (9) の Newton 方向は

$$\nabla h(\boldsymbol{x}_k, \boldsymbol{y}_k, \boldsymbol{s}_k) \begin{bmatrix} \Delta \boldsymbol{x}_k \\ \Delta \boldsymbol{y}_k \\ \Delta \boldsymbol{s}_k \end{bmatrix} = -\left(h(\boldsymbol{x}_k, \boldsymbol{y}_k, \boldsymbol{s}_k) - \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \\ \nu_{k+1}\boldsymbol{1} \end{bmatrix} \right) \tag{15}$$

の解である．実際に計算すると，

$$\begin{bmatrix} A & O & O \\ O & A^\top & I \\ S_k & O & X_k \end{bmatrix} \begin{bmatrix} \Delta \boldsymbol{x}_k \\ \Delta \boldsymbol{y}_k \\ \Delta \boldsymbol{s}_k \end{bmatrix} = -\begin{bmatrix} \boldsymbol{r}_k^b \\ \boldsymbol{r}_k^c \\ X_k \boldsymbol{s}_k - \nu_{k+1}\boldsymbol{1} \end{bmatrix} \tag{16}$$

となる．但し

$$S_k = \mathrm{diag}(\boldsymbol{s}_k) \tag{17}$$

$$\boldsymbol{r}_k^b = A\boldsymbol{x}_k - \boldsymbol{b} \tag{18}$$

$$\boldsymbol{r}_k^c = A^\top \boldsymbol{y}_k + \boldsymbol{s}_k - \boldsymbol{c} \tag{19}$$

である．ここでは $\boldsymbol{x}_{k+1} > \boldsymbol{0}, \boldsymbol{s}_{k+1} > \boldsymbol{0}$ については考慮していない．これについてはステップ幅で考慮することになる．

Newton 方程式については，上記のように最適性条件 (9) に対して（非線形方程式の）Newton 方程式を適応して Newton 方向を求めているとしてもよいし，（恐らく）目的関数を双対ギャップとして，その Newton 方程式を考えても良い．すなわち，主問題と双対問題をまとめて，双対ギャップ $\rho(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s})$ を最小化するとみる．その際，最適性条件 (9) は一次最適性条件 (KKT 条件) であると考えられ，不等式制約を無視すれば，

$$\nabla \rho(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s}) = h(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s}) - \begin{bmatrix} 0 \\ 0 \\ \nu\boldsymbol{1} \end{bmatrix} = \boldsymbol{0} \tag{20}$$

である．この ρ に対して Newton 法を適応したいので，ρ の二次近似から得られる条件（二次近似 ρ^Q の一次最適性条件 $\nabla\rho^Q = \mathbf{0}$）として Newton 方程式が得られ，

$$\nabla^2\rho\boldsymbol{d} = -\nabla\rho \tag{21}$$

となる．これを ρ の一次最適性条件 (20) を用いて変形すれば，

$$\nabla h(\boldsymbol{x}_k, \boldsymbol{y}_k, \boldsymbol{s}_k)\begin{bmatrix}\Delta\boldsymbol{x}_k\\ \Delta\boldsymbol{y}_k\\ \Delta\boldsymbol{s}_k\end{bmatrix} = -\left(h(\boldsymbol{x}_k, \boldsymbol{y}_k, \boldsymbol{s}_k) - \begin{bmatrix}\mathbf{0}\\ \mathbf{0}\\ \nu_{k+1}\mathbf{1}\end{bmatrix}\right) \tag{22}$$

となる．

5.5 終了条件

制約と相補性を十分な精度で満たす点に達したとき終了する．ε を十分に小さい定数として，

$$\rho_k < \varepsilon, \quad \|A\boldsymbol{x}_k - \boldsymbol{b}\| < \varepsilon, \quad \|A^\top\boldsymbol{y}_k + \boldsymbol{s}_k - \boldsymbol{c}\| < \varepsilon \tag{23}$$

を満たす点 $(\boldsymbol{x}_k, \boldsymbol{y}_k, \boldsymbol{s}_k)$ が得られれば，最適解として出力し，終了する．

6 直線探索

主双対内点法における直線探索では，$\boldsymbol{x}_{k+1} > 0, \boldsymbol{s}_{k+1} > 0$ を満たすようにステップ幅 α を決めるため，先に α の上限 $\alpha_{\max}$ について考えなけらばならない．このような $\alpha_{\max}$ は以下となる．

$$\begin{aligned}\alpha^P_{\max} &= \max\{\alpha \in (0,1] \mid \boldsymbol{x}_k + \alpha\Delta\boldsymbol{x}_k \geq \varepsilon\boldsymbol{x}_k\}\\ \alpha^D_{\max} &= \max\{\alpha \in (0,1] \mid \boldsymbol{s}_k + \alpha\Delta\boldsymbol{s}_k \geq \varepsilon\boldsymbol{s}_k\}\end{aligned} \tag{24}$$

$\boldsymbol{x}_k > \mathbf{0}, \alpha > 0$ であるから，i 番目について $\Delta x^i_k > 0$ ならば，$x^i_{k+1} > 0$ が満たされるため，α について特別な上限値を定める必要はなく，$\alpha_{\max} = 1$ とすればよい．逆に，$\Delta x^i_k < 0$ ならば，α が大きいと，$x^i_{k+1} < 0$ となる可能性が生じるため，$\alpha_{\max}$ を定める．従って，

$$\begin{aligned}\alpha^P_{\max} &= \min\left\{\min_{\Delta x^i_k<0}\left(-\frac{x^i_k}{\Delta x^i_k}\right), 1\right\}\\ \alpha^D_{\max} &= \min\left\{\min_{\Delta s^i_k<0}\left(-\frac{s^i_k}{\Delta s^i_k}\right), 1\right\}\end{aligned} \tag{25}$$

として選べばよい．主問題と双対問題で同じ α を用いる場合は，単に

$$\alpha_{\max} = \min\left\{\min_{\Delta x^i_k<0}\left(-\frac{x^i_k}{\Delta x^i_k}\right), \min_{\Delta s^i_k<0}\left(-\frac{s^i_k}{\Delta s^i_k}\right), 1\right\} \tag{26}$$

とすればよい．

適切な大きさの α を選ぶには Armijo の条件とバックトラック法を用いる．Armijo の条件はある評価関数 $f(\boldsymbol{u}_k)$，探索方向 $\boldsymbol{d}_k$ に対して，評価関数が十分に減少していればその α をステップ幅として決定する手法である．すなわち，ある定数 $c_1 \in (0,1]$ に対して，

$$f(\boldsymbol{u}_k + \alpha\boldsymbol{d}_k) \leq f(\boldsymbol{u}_k) + c_1\langle\nabla f(\boldsymbol{u}_k), \boldsymbol{d}_k\rangle\alpha \tag{27}$$

が満たされていれば終了とする．

実際の評価関数には，例えば以下のような制約に対する正則化を施したメリット関数 ϕ を用いる．

$$\phi_{(\nu,\rho)}(\boldsymbol{x}) = f(\boldsymbol{x}) + \rho\sum_{l=1}^{k}|h_l(\boldsymbol{x})| - \nu\sum_{j=1}^{n}\log x_j \tag{28}$$

$\nu > 0$ は主双対内点法で用いたバリアパラメータ（最終的に $\nu \searrow 0$），$\rho > 0$ はペナルティ関数のペナルティパラメータで徐々に大きくしていく（$\rho \to \infty$ で最適解に収束．実際のプログラムでは発散しないように上限を設けておく．紛らわしいが，この ρ は前述の双対ギャップ ρ とは関係ない別の定数.）.

今回のような線形計画問題の場合，メリット関数は以下のようになる．

$$\phi_{(\nu,\rho)}(\boldsymbol{x}) = \boldsymbol{c}^\top \boldsymbol{x} + \rho\|A\boldsymbol{x} - \boldsymbol{b}\| - \nu \sum_{j=1}^{n} \log x_j \tag{29}$$

この関数の勾配 $\nabla\phi_{(\nu,\rho)}(\boldsymbol{x})$ は解析的に求まり，

$$\nabla\phi_{(\nu,\rho)}(\boldsymbol{x}) = \boldsymbol{c} + \rho\frac{A^\top(A\boldsymbol{x} - \boldsymbol{b})}{\|A\boldsymbol{x} - \boldsymbol{b}\|} - \nu \begin{bmatrix} \frac{1}{x_1} \\ \vdots \\ \frac{1}{x_n} \end{bmatrix} \tag{30}$$

となる．右辺第二項の計算については，$e = \|A\boldsymbol{x} - \boldsymbol{b}\|$ とすると，

$$\nabla e^2 = 2e\nabla e \tag{31}$$

である．ここで

$$\begin{aligned} e^2 &= \|A\boldsymbol{x} - \boldsymbol{b}\|^2 &(32)\\ &= (A\boldsymbol{x} - \boldsymbol{b})^\top(A\boldsymbol{x} - \boldsymbol{b}) &(33)\\ &= \boldsymbol{x}^\top A^\top A\boldsymbol{x} - 2\boldsymbol{b}^\top A\boldsymbol{x} + \boldsymbol{b}^\top\boldsymbol{b} &(34)\end{aligned}$$

より，

$$\nabla e^2 = 2A^\top A\boldsymbol{x} - 2A^\top\boldsymbol{b} = 2A^\top(A\boldsymbol{x} - \boldsymbol{b}) \tag{35}$$

なので，

$$\nabla e = \frac{1}{2e}\nabla e^2 = \frac{A^\top(A\boldsymbol{x} - \boldsymbol{b})}{\|A\boldsymbol{x} - \boldsymbol{b}\|} \tag{36}$$

として求められる．

バックトラック法では，$\alpha = \alpha_{\max}$ から始め，Armijo の条件を満たすか判断し，満たす場合は終了，満たさない場合は $c_2 \in (0,1)$ を用いて $\alpha \leftarrow c_2\alpha$ として更新，再度 Armijo の条件を満たすかを繰り返す．（プログラムではこれに加えて α が十分に小さくなった場合でも終了判定を出している）.

7 アルゴリズム

7.1 主双対内点法

Step 1: 初期点として，$\boldsymbol{x}_0 > \boldsymbol{0}, \boldsymbol{s}_0 > \boldsymbol{0}$ を満たす点 $(\boldsymbol{x}_0, \boldsymbol{y}_0, \boldsymbol{s}_0)$ を選ぶ.

Step 2: 終了条件が満たされて入れば，$(\boldsymbol{x}_k, \boldsymbol{y}_k, \boldsymbol{s}_k)$ を最適解として終了.

Step 3: 双対ギャップ ρ_k を計算し，$\nu_{k+1} = \gamma\rho_k$ を計算する.

Step 4: ν_{k+1} を用いて，Newton 方程式から探索方向 $\begin{bmatrix}\Delta\boldsymbol{x}_k & \Delta\boldsymbol{y}_k & \Delta\boldsymbol{s}_k\end{bmatrix}^\top$ を決定する.

Step 5: 直線探索によってステップ幅 α を決定する.

Step 6: 更新則によって $(\boldsymbol{x}_{k+1}, \boldsymbol{y}_{k+1}, \boldsymbol{s}_{k+1})$ を求めて，Step 2 へ.

7.2 直線探索

Step 1: $\alpha_{\max}$ を計算し，初期値 $\alpha = \alpha_{\max}$ とする.（ρ, ν についても確認する．）

Step 2: Armijo の条件が満たされていれば $\alpha_k \leftarrow \alpha$ を解として終了.

Step 3:（十分小さい定数 ε_2 に対して $\alpha < \varepsilon_2$ であれば，$\alpha_k \leftarrow \alpha$ を解として終了.）

Step 4: $\alpha \leftarrow c_2\alpha$ として Step 2 へ.