

dual simplex 法の基礎

February 3, 2019

1 双対性

dual simplex を説明するためには、当然 dual を説明する必要がある。dual については、数学的に色々面白い話があるが、ここではそれら厳密な話を全てすつとばして、気分的な紹介にとどめる¹。

以下の線形計画問題を考える。

$$\begin{aligned} \min_x \quad & c^T x \\ \text{s.t.} \quad & \sum_j A_{ij} x_j \geq b_i \ (i \in \{1, \dots, m\}), \\ & x \geq 0, x \in \mathbb{R}^n \end{aligned} \tag{1}$$

勿論、 $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ である。

ここで、ちょっとだらけたことを考える。何を考えるのかというと、「制約がきつすぎるので、制約を完全に満たすのではなく、そこそこ満たしていて良い感じの解がほしい」ということを考える。そのような際に有用な方法として、緩和がある。例えば (1) 式であれば、以下がそのような緩和になる。

$$\begin{aligned} \min_x \quad & c^T x - y^T (Ax - b) \\ \text{s.t.} \quad & x \geq 0, x \in \mathbb{R}^n \end{aligned} \tag{2}$$

ここに y は、 $y \geq 0$ なとりあえずパラメータである（あとで、このパラメータも最適化することを考えるがまずは説明のためにパラメータと思ってもらって OK）。

まずは、こいつが目的通りであることを見ておく。constraint として、 $Ax \geq b$ が入っていないので、(2) の解は当然のごとくこの制約を満たす必要はない。しかし、(2) において $Ax \geq b$ をやぶると大変なことになる。具体的には、 $y \geq 0$ であるため、この制約を破ると objective 最小化、という意味では不利であ

¹ というか、ちゃんとやろうとすると凸解析を一からやる必要があるので無理ゲーすぎる。

る．ということで (2) を解くと「制約を完全には満たしていないが，そこそこ満たした良い感じの解」が得られることが期待されるわけである．

さて，この緩和した問題であるが，以下の性質がある． x^* を (1) の最適解だとすると，当然 $Ax^* \geq b$, $x^* \geq 0$ に気をつければ，

$$c^T x^* \geq c^T x^* - y^T (Ax^* - b) \geq \min_{x, x \geq 0} (c^T x - y^T (Ax - b)) \quad (3)$$

という不等式を導くことができる．よって，(2) の最適解の objective は (1) の最適解の objective 以下である．この不等式自体は任意の $y \geq 0$ で成立するので，実は以下が成立する．

$$c^T x^* \geq \max_{y, y \geq 0} \min_{x, x \geq 0} (c^T x - y^T (Ax - b)) . \quad (4)$$

この式の右辺をもっと変形してみよう．まず objective であるが，これは $(c - A^T y)^T x + b^T y$ と変形することができる．これの $\min_{x, x \geq 0}$ を考えるが， x については $x \geq 0$ 以外の制約がないことを考えると

$$\min_{x, x \geq 0} (c - A^T y)^T x + b^T y = \begin{cases} b^T y & (c - A^T y \geq 0) \\ -\infty & (\exists i. s.t. (c - A^T y)_i < 0) \end{cases}$$

となる．続いて (4) 式の右辺の \max_y を考えていくが，上記で $c - A^T y < 0$ のケースでは $-\infty$ に飛ぶことを考えると \max では $c - A^T y \geq 0$ のケースのみを考えておけば十分である．よって (4) 式の右辺は

$$\begin{aligned} \max_y \quad & b^T y \\ \text{s.t.} \quad & A^T y \leq c \\ & y \geq 0, y \in \mathbb{R}^m \end{aligned} \quad (5)$$

と変形される．これが (1) に対する dual problem と言われる問題で，この導出からもわかるように，dual の最適解の objective は元問題 (primal problem と呼ばれる) の最適解の objective 以下の値をとる．

ここまでは (1) のように，不等式制約を考えてきたが，LP の基本形である等号制約な場合，つまり，

$$\begin{aligned} \min_x \quad & c^T x \\ \text{s.t.} \quad & Ax = b \\ & x \geq 0, x \in \mathbb{R}^n \end{aligned} \quad (6)$$

であっても同じように dual の問題を考えることができる．導出は簡単で， $Ax = b$ を $Ax \geq b$ かつ $Ax \leq b$ と分けて計算すれば良く，結果は

$$\begin{aligned} \max_y \quad & b^T y \\ \text{s.t.} \quad & A^T y \leq c, y \in \mathbb{R}^m \end{aligned} \quad (7)$$

となる。

この節の最後に、LP における強双対性を示す。ここまでの話、特に、(4) 式から分かるように、dual の最適解は primal の最適解以下である。しかし、実はもっと強いことが言えて、これは以下ではなくイコールである。以下ではそれを確認する。

primal 側の最適な実行可能基底解を $x_B^* = B^{-1}b$, $x_N^* = 0$ とする。primal simplex の最適性条件から $c_N - N^T B^{-1T} c_B \geq 0$ である。これらの B, N を使って、dual の解らしきものとして、 $y = B^{-1T} c_B$ を考えてみよう。これは以下に見るように dual の constraint を満たしている。

$$A^T y = \begin{bmatrix} B^T \\ N^T \end{bmatrix} B^{-1T} c_B \leq \begin{pmatrix} c_B \\ c_N \end{pmatrix}.$$

よって、 $y = B^{-1T} c_B$ は dual の実行可能解である。ということで、primal 側の最適解を絡めた以下の不等式が成立する。

$$c^T x^* = c_B^T x_B^* = c_B^T B^{-1} b = (B^{-1T} c_B) b = b^T y \leq \max_{y, \text{dual constraint}} b^T y. \quad (8)$$

よって (4) 式の逆の不等式が言える。よって実はイコールが成立する。

2 dual simplex method

無事 dual を導入できたので、この章では、dual simplex 法について紹介する。まず、dual simplex を説明するために必要な用語 (simplex の時にも出てきた実行可能基底解だとかそういった話) をしたのちに、dual simplex method 御本尊を紹介する。

2.1 用語

simplex の時と同様に基底解などについてお話する。まず、イコール制約でないとか何とか面倒なので、(7) に slack を入れて

$$\begin{aligned} \max_y \quad & b^T y \\ \text{s.t.} \quad & A^T y + s = c, \\ & s \geq 0, y \in \mathbb{R}^m, s \in \mathbb{R}^n \end{aligned} \quad (9)$$

という形に変えて考える。

この問題の基底解というところからお話を始める。上記制約式は n 本あり、変数の数は合計で $n + m$ 個ある。よって $n + m$ 個の中から m 個の変数を 0 とおけば、基本的には $A^T y + s = c$ という方程式を解くことができる。

特に、ここでは、 s にある n 個の変数の中から m 個を選び、それが 0 となるような解を考える。そのような解を双対基底解と呼ぶ。勿論そのような双対基底解は $s \geq 0$ を満たすとは限らない。simplex の時と同様に双対基底解でかつこの制約も満たすものを双対実行可能基底解と呼ぶことにする。

と言葉で説明してもさっぱりなので、やっぱり数式を用いて述べておく。 s にある n 個の変数 index を m 個と $n - m$ 個のグループに分ける。前者を B_{index} 、後者を N_{index} と書くことにしよう。このとき、 A の列 (A^T の行というべきかも) を $B_{index} \rightarrow N_{index}$ の順番に並び替えて $A = [B|N]$ と分割する。同様に並び替えることで、 s も $s = [s_B|s_N]^T$ と並び替え、さらに c も $c = [c_B|c_N]^T$ と並び替える。すると制約式 $A^T y + s = c$ は

$$\left[\begin{array}{c} B^T \\ N^T \end{array} \right] y + \left[\begin{array}{c} s_B \\ s_N \end{array} \right] = \left[\begin{array}{c} c_B \\ c_N \end{array} \right]$$

となる。ここで、 s_B だけを右辺に持っていて左辺をまとめて書くと以下のよう書き直すことができる。

$$\left[\begin{array}{c|c} B^T & 0 \\ \hline N^T & I \end{array} \right] \left[\begin{array}{c} y \\ s_N \end{array} \right] = \left[\begin{array}{c} c_B \\ c_N \end{array} \right] - \left[\begin{array}{c} s_B \\ 0 \end{array} \right]$$

ここで 最初の $n \times n$ 行列の逆行列は

$$\left[\begin{array}{c|c} B^T & 0 \\ \hline N^T & I \end{array} \right]^{-1} = \left[\begin{array}{c|c} B^{-1T} & 0 \\ \hline -N^T B^{-1T} & I \end{array} \right]$$

なので、両辺にかけてやれば、

$$\left[\begin{array}{c} y \\ s_N \end{array} \right] = \left[\begin{array}{c} B^{-1T} c_B \\ c_N - N^T B^{-1T} c_B \end{array} \right] - \left[\begin{array}{c} B^{-1T} s_B \\ -N^T B^{-1T} s_B \end{array} \right]$$

となる。よって、双対基底解 $s_B = 0$ が双対実行可能基底解となる条件 $s_N \geq 0$ より

$$\rho \geq 0, \tag{10}$$

where

$$\rho = c_N - N^T B^{-1T} c_B \tag{11}$$

となる²。

²これは simplex 法の optimality condition と一致している。

2.2 dual simplex

dual simplex について紹介する．ここまでに紹介した用語を使うと，simplex 同様，dual simplex も input を双対実行可能基底解とし，output を最適な双対実行可能基底解を出力する algorithm である．以下でそれを見ていく．

2.2.1 最適性条件と pricing

まずは最適性条件を確認する．ある双対実行可能基底解とそれに関する B_{index}, N_{index} が与えられたとする．このとき，その解の周りで $[s_B|0]^T$ を動かすとする．その場合，dual problem (7) は (10) 式を用いて

$$\begin{aligned} \max_{s_B} \quad & b^T B^{-1T} c_B - b^T B^{-1T} s_B \\ \text{s.t.} \quad & \rho + N^T B^{-1T} w_B \geq 0, s_B \geq 0 \end{aligned} \quad (12)$$

となる．よって， $s_B \geq 0$ なので，maximize であることに注意をすれば，optimality condition は

$$b^T B^{-1T} \geq 0 \Leftrightarrow B^{-1}b \geq 0 \quad (13)$$

となる³．

逆に言えば， $(B^{-1T}b)_i < 0$ となる i が存在する場合，条件の範囲内で，objective を増加させることが可能である．以下では simplex 同様，そのような i の中から一つだけ選び，それを k として， k に関する local な問題を考えていく．このような k の選択を dual pricing とか dual pricing rule とか呼ぶ⁴．

2.2.2 ratio test

pricing で k が一つ選ばれたとする．そうすると s_B の k 成分だけ値を持った以下の local な問題を定義することができる． $(s_B)_k = \xi$ と書けば

$$\begin{aligned} \max_{\xi} \quad & b^T B^{-1T} c_B + |(B^{-1}b)_k| \xi \\ \text{s.t.} \quad & \rho_j + \sum_{\ell} (N^T)_{j\ell} (B^{-1T})_{\ell k} \xi \geq 0 \quad (j \in N_{index}) \\ & \xi \geq 0 \end{aligned} \quad (14)$$

となる．この local な問題であるが，simplex の時と同様，簡単に解くことが可能である．というのも，objective を見ると ξ が大きければ大きいほど良いの

³これは primal feasibility と同じ．

⁴ k の選び方については，simplex 同様，様々なルール，つまり pricing rule がある．

で、条件の範囲内で ξ を大きくすれば良いからである。ではどれだけ大きくして良いのか、というと、

$$\alpha_j = \sum_{\ell} (N^T)_{j\ell} (B^{-1T})_{\ell k} \quad (15)$$

と定義してやれば、条件より以下の θ まで大きくすることが可能であることが分かる。

$$j_{min} = \arg \min_j \left[\frac{\rho_j}{|\alpha_j|} \mid \alpha_j < 0 \right], \theta = \frac{\rho_{j_{min}}}{|\alpha_{j_{min}}|}. \quad (16)$$

このような j_{min} の選択を dual ratio test と呼ばれる。ちなみに、この ratio test の際、もし $\alpha \geq 0$ であった場合、いくらでも ξ を大きくすることができる。よって、この場合は unbounded である。

上述したように、 $\xi = \theta$ のときが local な問題の最適解である。ではこの最適解では、 s がどうなっているのかというと、まず、 $s_j^{new} = 0 (j \in B_{index} \setminus \{k\})$, $s_k^{new} = \theta$ である。さらに、 $s_N = \rho + N^T B^{-1T} s_B$ という関係から $s_j^{new} = \rho_j + \alpha_j \theta (j \in N_{index})$ となる。特に、 $s_{j_{min}}^{new} = 0$ となる。よって、この local な最適化問題の答えは、 $B_{index}^{new} = (B_{index} \setminus \{k\}) \cup \{j_{min}\}$, $N_{index}^{new} = (N_{index} \setminus \{j_{min}\}) \cup \{k\}$ とする双対実行可能基底解である。よって、この新しい双対実行可能基底解をベースとして、pricing を行えば、新しく local な問題も立てることができる、といった具合に iterative に objective を上げていくことができる。勿論、このような local search でうまくいく理由は今回の問題が linear であるためである。

2.2.3 algorithm

以上をまとめると、以下のような algorithm になる。

3 dual simplex の特徴

dual simplex には simplex 法には無い良い特徴がある。Bound Flipping Ratio Test など ratio test も工夫できるといった特徴もあるが、ここでは、warm start についてお話をする。これは、「LP を解いた後に、不等式制約を追加した場合に、古い問題の最適解から新しい問題の探索をできる」という性質である。これは simplex 法にはない特徴である。というのも、simplex 法を使った場合、新しい制約を加えた段階で古い問題の最適解は primal の意味の実行可能性が保証できず、結局二段階 simplex を一から解く必要が生じるためである。ところが、以下に見るように古い問題の最適解は新しい問題の双対実行可能基底解になっている。そのため dual simplex の場合はこのような warm start が可能である。

Require: 双対実行可能基底解

Ensure: 最適解

現在の双対実行可能規定解 \leftarrow input 双対実行可能基底解

loop

現在の双対実行可能基底解に基づいて $B^{-1}b$ を計算する.

if $B^{-1}b > 0$ **then**

return 現在の双対実行可能基底解

end if

– pricing –

なんらかのルールで $B^{-1}b$ が negative な index から k を選ぶ.

– ratio test –

(11) と (15) を計算

(16) によって $B_{index} \rightarrow N_{index}$ となる変数 j_{min} を選択

– pivot –

k, j_{min} を基に現在の双対実行可能基底解を更新

end loop

primal で

$$\begin{aligned} \min_x \quad & c^T x \\ \text{s.t.} \quad & Ax = b, x \geq 0, x \in \mathbb{R}^n \end{aligned} \quad (17)$$

と書かれる問題を考える. 勿論この dual の問題は

$$\begin{aligned} \max_y \quad & b^T y \\ \text{s.t.} \quad & A^T y + s = c, y \in \mathbb{R}^m, s \geq 0 \end{aligned} \quad (18)$$

である. この解を y^{old*}, s^{old*} と書くことにし, この基底 index 集合, 非基底 index 集合を $B_{index}^{old}, N_{index}^{old}$ と書くことにする. 勿論 $s_j^{old*} = 0 (j \in B_{index}^{old})$ であるし, (10) 式を思えば, $y^{old*} = B^{old-1T} c_{B^{old}}, s_j^{old*} = \rho^{old} = c_{N^{old}} - N^{oldT} B^{old-1T} c_{B^{old}} \geq 0 (j \in N_{index}^{old})$ である.

このとき, 今回の問題が解き終わった後に, primal 側に不等式制約 $a^T x \leq b'$ を加えた問題を考える. つまり,

$$\begin{aligned} \min_x \quad & c^T x + 0 \times w \\ \text{s.t.} \quad & Ax = b \\ & a^T x + w = b' \\ & x \geq 0, w \geq 0, x \in \mathbb{R}^n, w \in \mathbb{R} \end{aligned} \quad (19)$$

という問題を考える．ここに， w は slack である．ということで，イコール制約に直すことまで考えると，制約が一つ，変数が増えていることだけ述べておく．上記のままでも良いが，書くのが面倒なので，

$$\bar{x} = [x, w]^T, \bar{c} = [c, 0]^T, \bar{b} = [b, b']^T, \bar{A} = \left[\begin{array}{c|c} A & 0 \\ \hline a^T & 1 \end{array} \right]$$

を用意すれば，primal 側の問題は

$$\begin{aligned} \min_x \quad & \bar{c}^T \bar{x} \\ \text{s.t.} \quad & \bar{A} \bar{x} = \bar{b}, \bar{x} \geq 0, \bar{x} \in \mathbb{R}^{n+1} \end{aligned} \quad (20)$$

とかけし，dual 側は

$$\begin{aligned} \max_{\bar{y}} \quad & \bar{b}^T \bar{y} \\ \text{s.t.} \quad & \bar{A}^T \bar{y} + \bar{s} = \bar{c}, \bar{y} \in \mathbb{R}^{m+1}, \bar{s} \geq 0 \end{aligned} \quad (21)$$

と書ける．この書き方に習って，以下では，新しく追加した制約の index を $m+1$ と書き，新しく追加した slack 変数に対応する index を $n+1$ と書くことにする．

少々前置きが長くなってしまったが，ここまで準備できれば，あとは計算をするだけで以下を示すことができる．「 $B_{index}^{new} = B_{index}^{old} \cup \{n+1\}$ かつ $N_{index}^{new} = N_{index}^{old}$ に対応する基底解は双対実行可能基底解である．もう少し具体的には $\bar{y} = [y^{old*}, 0]^T$, $s_{B^{new}} = 0$, $s_{N^{new}} = s_{N^{old}}$ は双対実行可能である．」つまり，新しく加わった slack を基底に入れば，双対実行可能性は保持される，という訳である．

証明は超簡単．(10) 式を使ってただ計算するだけ． B_{index}^{new} , N_{index}^{new} の選び方から， \bar{A} を分割すると，

$$B^{new} = \left[\begin{array}{c|c} B^{old} & 0 \\ \hline a_{B^{old}}^T & 1 \end{array} \right], N^{new} = \left[\begin{array}{c} N^{old} \\ \hline a_{N^{old}}^T \end{array} \right]$$

となるが，(10) 式に必要なものを見ると

$$B^{new-1T} = \left[\begin{array}{c|c} B^{old-1T} & -B^{old-1T} a_{B^{old}} \\ \hline 0 & 1 \end{array} \right], N^{newT} = [N^{oldT} | a_{N^{old}}]$$

となっている．これをもとに B_{index}^{new} , N_{index}^{new} に対応する基底解 (つまり， $\bar{s}_{B^{new}} = 0$ な解) を (10) 式をベースに計算する．まず y であるが， $\bar{c}_{B^{new}} = [c_{B^{old}}, 0]^T$ であることに注意すると，

$$\bar{y} = B^{new-1T} \bar{c}_{B^{new}} = [B^{-1T} c_{B^{old}}, 0]^T = [y^{old*}, 0]^T \quad (22)$$

となる。そして、 $c_{N^{new}} = c_{N^{old}}$ に注意すれば、

$$\begin{aligned}
\bar{s}_{N^{new}} &= c_{N^{new}} - N^{newT} B^{new-1T} c_{B^{new}} \\
&= c_{N^{old}} - \left[N^{oldT} | a_{N^{old}} \right] \left[\frac{B^{old-1T} c_{B^{old}}}{0} \right] \\
&= c_{N^{old}} - N^{oldT} B^{old-1T} c_{B^{old}} \\
&= s_{N^{old}}^{old*} \geq 0
\end{aligned} \tag{23}$$

となっているので、 B_{index}^{new} , N_{index}^{new} に対応する基底解は双対実行可能であり、 $\bar{y} = [y^{old*}, 0]^T$, $\bar{s}_{B^{new}} = 0$, $\bar{s}_{N^{new}} = s_{N^{old}}^{old*}$ がその双対実行可能基底解である。

このように、dual simplex 法は、新しい制約を加えても、古い問題の最適解が新しい問題の双対実行可能基底解になっているので、古い問題の最適解を input として新しい問題を解くことが可能である。このような解き方を warm start と呼ばれ、これは simplex 法ではできず dual simplex 特有の方法だと言える。

4 まとめ

ここでは dual simplex を紹介した。この algorithm は、最後にも述べたように、制約式の追加して繰り返して解く際に非常に有用な algorithm となっている。そういった特徴があるため、例えば branch and bound では利用されている。また、そのほかの場面でもこのような特徴は活かせる可能性があるので、結構重要な algorithm だと言える。