

(23) 縮小勾配法の最小化法への適用
AN APPLICATION OF REDUCED GRADIENTS
TO UNCONSTRAINED MINIMIZATION METHODS

平田 恭久*
by Yasuhisa HIRATA

The author has already proposed an optimization method using selection techniques of active constraint surfaces. In this method an optimum point of the problem is found by minimization of the objective function that is calculated on the active constraints. In the Generalized Reduced Gradient method and the above method, reduced gradients of the Lagrangian function are used. If the Quasi-Newton techniques are used to modify search directions, then applications of the reduced gradient to these methods is difficult because a variable partition change results in a loss of the calculation of the direction vector.

Such difficulties can be avoided by an adaptation of the gradient that is calculated on the active constraint. In this paper a calculation method of this gradient is proposed, and some consideration in application of the Conjugate Gradient method and the Quasi-Newton method to the minimization method are stated.

Key Words : reduced gradient, minimization method,
quasi-Newton method, direction vector

1. まえがき

GRG等の縮小勾配を用いる最適化手法では、共役方向法、準Newton法(共役方向法等と略すことがある)による探索のとき、独立変数と従属変数の入れ換えが発生するとそれまでに蓄積した情報が使えなくなり、共役方向法等を新たに開始しなければならない。これに対し、独立変数と従属変数の区分を固定しておくことは実際上無理なので、この点が縮小勾配を用いる最適化手法に共役方向法等の最小化法を適用する場合の重大な障害になっている。筆者の開発した等式制約法も探索での最急勾配として縮小勾配を用いており、その値は活性な制約式とこれに対する制約変数(従属変数)と探索変数(独立変数)の区分(変数の区分と略すことがある)に依存している。

探索の開始点と最適解とでは活性な制約式が異なるのが普通なので、探索過程で活性な制約式が異なってくるのは止むを得ないことであるが、本研究では問題を簡単にするため次のように扱うことにした。①変数

* 群馬高専助教授 土木工学科 (〒371 群馬県前橋市鳥羽町580)

の区分が異なることと ②制約式が異なることを切り離して、活性な制約式は一定で変数の区分のみが異なってくる場合を以下の方法で解決することを試みた。探索過程で上記①はしばしば生ずるが、上記②が生ずる機会は比較的少ないので、上記①を解決することの効果は大きい。

活性な制約面上での最急勾配はある点 x_k についてただ一つ存在するので、縮小勾配($\nabla_s L$ と記す)を活性な制約面上での最急勾配(∇F と記す)に変換することにより、変数の区分が異なっても同じ最急勾配が得られるようにすれば、上述の重大な障害が取り除かれることになる。本研究ではLagrange未定乗数法を用いて ∇F 算出式を導き、この ∇F を共役方向法での共役方向算出式、準Newton法でのDFP公式に適用し、通常の縮小勾配と同様の探索が実行できることを確認した。

2. 最急勾配 ∇F の誘導

(1) ∇F を求める最適化問題

活性な制約面上での目的関数 f の最急勾配 ∇F を見つける問題を考えるが、ここでは探索方向 $dx = -\nabla F$ で定式化している。 dx は制約面上の条件を満足し、 $df = \nabla f^T dx$ が最小の方向なので、式(1)の最適化問題になる。¹⁾ 式(1)第2行の内容は式(2)であり、 dx が制約面上に存在するための条件である。式(1)第3行は正規化条件である。

式(1)の行列 A_s は E^T と I_s から、ベクトル dx は dx_m と dx_s から構成されている。 E は式(3)で表され、 g_m は m 個の活性な制約式である。²⁾ 式(1)は等号制約条件つき最適化問題なので、Lagrange未定乗数法が適用できる。式(1)についてLagrange関数は式(4)になり、極値の条件は式(5)になる。

(2) ∇F の解

式(5)を次のようにして解けば dx_s が得られる。式(5)第2行の dx を式(5)第1行に代入すると式(6)になり、式(6)を制約変数 x_m に関する部分と探索変数 x_s に関する部分に分けると式(7)になる。変数の個数を n とすると $n = m + s$ であり、 I は n 次元、 I_m は m 次元、 I_s は s 次元の単位行列である。式(7)より dx_s 、 μ_m は式(8)、式(9)になるが、 $w = 1$ はスケールファクタである。³⁾

式(8)の $\nabla_s L$ は式(10)で表されるが、 $\nabla_s L$ は本来の最適化問題でのLagrange関数の探索変数 x_s についての微分である。活性な制約面を得るための掃き出し計算で E 、 $\nabla_s L$ は求まっているので、式(8)から dx_s が計算できる。

式(8)の dx_s は探索方向 dx の s 成分なので、 ∇F の s 成分 $\nabla_s F$ は式(11)第1行のように dx_s の符号を変えたものになる。 ∇F の m 成分 $\nabla_m F$ は式(11)第2行になり、 $\nabla_s F$ と $\nabla_m F$ を合わせたものが ∇F である。式(8)では $(I_s + EE^T)$ の逆行列を解いているが、 $\nabla_s L$ が存在するなら ∇F が存在する筈なので、 $(I_s + EE^T)^{-1}$ の計算は可能である。

$$\left. \begin{array}{l} \min \nabla f^T dx \\ dx \\ \text{subject to} \\ A_s dx_s - dx_m = 0 \\ dx^T dx - 1 = 0 \end{array} \right\} \dots\dots (1)$$

$$\left. \begin{array}{l} E^T dx_s - dx_m = 0 \\ I_s dx_s - dx_s = 0 \end{array} \right\} \dots\dots (2)$$

$$E = -\nabla_s g_m^T (\nabla_m g_m^T)^{-1} \dots (3)$$

$$\left. \begin{array}{l} L(dx, \mu, w) = \nabla f^T dx \\ + \mu^T (A_s dx_s - dx_m) \\ + w (dx^T dx - 1) \end{array} \right\} (4)$$

$$\left. \begin{array}{l} \partial L / \partial dx = \nabla f^T + \mu^T (A_s - I) \\ + 2 w dx^T = 0 \\ \partial L / \partial \mu = A_s dx_s - dx_m = 0 \\ \partial L / \partial w = dx^T dx - 1 = 0 \end{array} \right\} (5)$$

$$\nabla f + (A_s - I)^T \mu_m + 2 w A_s dx_s = 0 \quad (6)$$

$$\left. \begin{array}{l} \nabla_m f - I_m \mu_m + 2 w E^T dx_s = 0 : m \text{ 個} \\ \nabla_s f + E \mu_m + 2 w I_s dx_s = 0 : s \text{ 個} \end{array} \right\} (7)$$

$$dx_s = -(I_s + EE^T)^{-1} (\nabla_s f + E \nabla_m f) / 2 w = -(I_s + EE^T)^{-1} \nabla_s L / 2 w \dots (8)$$

$$\mu_m = (I_s + EE^T)^{-1} (\nabla_m f - E^T \nabla_s f) \dots (9)$$

$$\nabla_s L = \nabla_s f + E \nabla_m f \dots\dots\dots (10)$$

$$\left. \begin{array}{l} \nabla_s F = -dx_s \\ \nabla_m F = E^T \nabla_s F \end{array} \right\} \dots\dots\dots (11)$$

3. 縮小勾配 $\nabla_s L$ と最急勾配 ∇F の関係

(1) $\nabla_s L$ と ∇F の図示

目的関数 f 、制約式 g が線形の場合について $-\nabla_s L$ と $-\nabla F$ を図示したのが図-1である。 $\nabla_s L$ は活性な制約面上での f を s 次元に射影したときの最急勾配であり、 $\nabla_s L$ を n 次元に拡張するときの $\nabla_m L$ は式(12)になる。 $\nabla_s L$ と $\nabla_m L$ を合わせたものが ∇L であり、探索方向 $-\nabla L$ は制約面上の条件を満足している。

図-1では $x_2 - x_3$ 平面が s 次元であり、 x_1 が制約変数 x_m になる。 ∇F は n 次元での" g 上で f 一定"に直交しており、 $\nabla_s L$ は s 次元に射影した" g 上で f 一定"に直交している。これより $\nabla_s L$ は s 次元に射影したときの最急勾配であることが分かる。 ∇F は x_k 点での g 上の最急勾配なので唯一であるが、 $\nabla_s L$ はどの平面に射影(x_s の選択)したかにより異なってくるので、組み合わせは $n C_s$ 個存在する。上述のように探索過程で x_s の選択を固定しておくことは困難なので、どの x_s を選択しても唯一の ∇F が計算できればこの問題点は解決される。

(2) 正規化 df

活性な制約面上を探索するときの f の微小変化 df は式(3)になる。最急降下法として式(3)の $d x_s$ に $-\nabla_s L$ を用いると式(4)になり、 $-\nabla_s F$ を用いると式(5)になる。 $\nabla_s L^T \nabla_s L \geq 0$ 、 $\nabla F^T \nabla F \geq 0$ なので、 $df(\nabla_s L) \leq 0$ 、 $df(\nabla_s F) \leq 0$ であり、最急降下法の探索方向に $-\nabla_s L$ または $-\nabla_s F$ を用いれば常に f を減少させ得る。

大きさを比較するための正規化 df は式(6)、式(7)になる。制約変数 x_m の選択により E 、 $\nabla_s L$ は異なったものになるので、 $df(\nabla_s L)$ は選択での組み合わせの数だけ存在する。これに対し、 $df(\nabla_s F)$ は最小の df であり、どの E 、 $\nabla_s L$ を用いても同じ ∇F (すなわち同じ d)が得られる。

(3) 最小の正規化 df

∇F は n 次元での最急勾配なので、 $-\nabla F$ 方向に探索すれば最も f が減少する筈である。すなわち式(8)に示すように正規化 $df(\nabla_s F)$ は最小であり、 $df \leq 0$ より

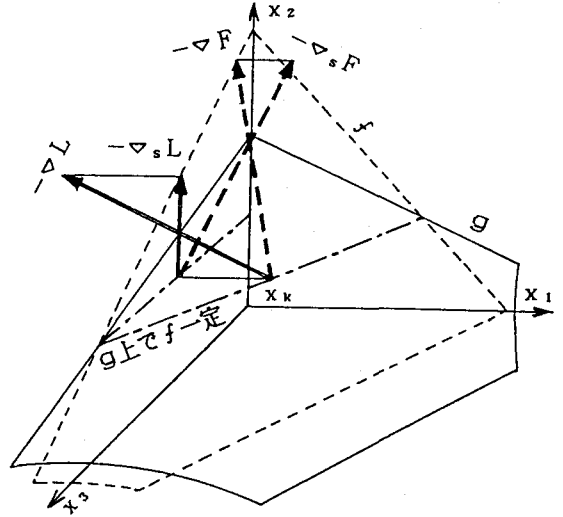


図-1 $-\nabla_s L$ と $-\nabla F$

$$\nabla_m L = E^T \nabla_s L \quad \dots\dots (12)$$

$$\begin{aligned} df &= \nabla f^T dx \\ &= \nabla_s f^T dx_s + \nabla_m f^T dx_m \\ &= \nabla_s L^T dx_s \quad \dots\dots (13) \end{aligned}$$

$$df(\nabla_s L) = -\nabla_s L^T \nabla_s L \quad \dots (14)$$

$$\begin{aligned} df(\nabla_s F) &= -\nabla_s L^T \nabla_s F \\ &= -2 w \nabla F^T \nabla F \quad \dots (15) \end{aligned}$$

$$\begin{aligned} \text{正規化} df(\nabla_s L) &= -\nabla_s L^T \nabla_s L / \sqrt{\nabla_s L^T \nabla_s L} \quad \dots (16) \end{aligned}$$

$$\begin{aligned} \text{正規化} df(\nabla_s F) &= -2 w \nabla F^T \nabla F / \sqrt{\nabla F^T \nabla F} \quad \dots (17) \end{aligned}$$

$$\text{正規化} df(\nabla_s F) \leq \text{正規化} df(\nabla_s L) \quad \dots (18)$$

$$\begin{aligned} \{\text{正規化} df(\nabla_s L)\}^2 \nabla L^T \nabla L &= (\nabla_s L^T \nabla_s L)^2 = (2 \nabla F^T \nabla L)^2 \\ &= (\alpha^T b)^2 \quad \dots (19) \end{aligned}$$

$$\begin{aligned} \{\text{正規化} df(\nabla_s F)\}^2 \nabla L^T \nabla L &= (2 \nabla F^T 2 \nabla F) (\nabla L^T \nabla L) = \alpha^2 b^2 \\ &\quad \dots (20) \end{aligned}$$

$$\alpha^2 b^2 \geq (\alpha^T b)^2 \quad \dots (21)$$

$$\begin{aligned} \{\text{正規化} df(\nabla_s F)\}^2 \nabla L^T \nabla L &\geq \{\text{正規化} df(\nabla_s L)\}^2 \nabla L^T \nabla L \\ &\quad \dots (22) \end{aligned}$$

正規化 $df(\nabla_s F))^2$ は最大である。式(19), 式(20)のように $\nabla L^T \nabla L \geq 0$ を乗じたもので比較するが、 $2\nabla F = \alpha$, $\nabla L = \beta$ とくと式(21)のSchwarzの不等式が適用できる。⁴⁾

これより式(22)が得られ、両辺を $\nabla L^T \nabla L \geq 0$ で割っても不等号の向きは同じなので、(正規化 $df(\nabla_s F))^2$ 最大すなわち正規化 $df(\nabla_s F)$ 最小の式(18)が得られる。式(18)はどの $-\nabla_s L$ 方向(∇L 方向)よりも $-\nabla_s F$ 方向($-\nabla F$ 方向)の方が f の減少が大きい、すなわち ∇F が n 次元での最急勾配であることを示している。

4. $\nabla_s L$, ∇F の直交性

(1) 線形制約面での直交性

E は式(3)で表されるので、活性な制約式 g_m が線形なら E は定数であり、 x_k には依存しなくなる。ここでは議論を簡単にするため線形制約面を前提にする。ラインサーチの探索方向とラインサーチ終了点での最急勾配は直交しており、共役方向法、準Newton法では式の誘導のときにこの関係を利用している。⁵⁾ よって、これらの方法を適用する場合には、ラインサーチ探索方向とラインサーチ終了点の最急勾配との直交性を確かめておく必要がある。

dx_{k-1} はラインサーチ開始点($k-1$)での探索方向(n 次元)、 ∇F_k はラインサーチ終了点(k)での最急勾配(n 次元)とする。 dx_{k-1} は制約面上にあり、 ∇F_k は制約面上での最急勾配なので、ラインサーチ終了点では直交している筈である。 dx_{k-1} と ∇F_k の内積に式(23), (24), (25)を適用すると式(26)が得られる。これに対し、 dx_{k-1} と ∇L_k の内積は式(27)になり、一般に $dx_{k-1}^T \nabla L_k \neq 0$ で直交性は成立しない。

次に s 次元については、 $\nabla_s L_k$ は制約面上の f を s 次元に射影したときの最急勾配なので、 dx_{k-1} の s 成分である dx_{sk-1} とは式(28)のように直交している。これに対し、式(28)より明らかなように一般に $dx_{sk-1}^T \nabla_s F_k \neq 0$ であり、 dx_{sk-1} と $\nabla_s F_k$ との直交性は成立しない。図-1で" g 上で f 一定"の線上に dx_{k-1} が乗っていることから、上述の直交性は理解できる。

(2) 非線形制約面の場合

非線形制約面の場合に活性な制約面に沿って dx 方向に探索するとき、 dx_s は一定であるが dx_m は変化してくる。すなわち、 $dx_m = E^T dx_s$ において E が変化し、ラインサーチ開始点では $E = E_{k-1}$ であったものが、終了点では $E = E_k$ になる。この関係を式(28)に示すが、 dx_{k-1}' は終了点での探索方向である。

図-1を非線形制約面であるとして、 dx_{k-1} と F_k , dx_{sk-1} と $\nabla_s L_k$ の関係を示したが図-2である。終了点の x_k では ∇F_k は dx_{k-1}' と直交しているが、 dx_{k-1}' とは m 成分の異なる dx_{k-1} に対して $dx_{k-1}^T \nabla$

$$dx^T = [dx_s^T, dx_m^T E] \quad \dots (23)$$

$$\nabla F^T = [\nabla_s F^T, \nabla_m F^T E] \quad \dots (24)$$

$$\nabla_s L = 2(I_s + EE^T)\nabla_s F \quad \dots (25)$$

$$dx_{k-1}^T \nabla F_k = dx_{sk-1}^T (I_s + EE^T) \nabla_s F_k \\ = dx_{sk-1}^T \nabla_s L_k / 2 = 0 \quad \dots (26)$$

$$dx_{k-1}^T \nabla L_k = dx_{sk-1}^T (I_s + EE^T) \nabla_s L_k \quad \dots (27)$$

$$dx_{k-1}^T = [dx_{sk-1}^T, dx_{mk-1}^T E_{k-1}] \\ dx_{k-1}^T = [dx_{sk-1}^T, dx_{mk-1}^T E_k] \quad \dots (28)$$

$$dx_{k-1}^T \nabla F_k = [dx_{sk-1}^T, dx_{mk-1}^T E_k] \begin{bmatrix} \nabla_s F_k \\ E_k^T \nabla_s F_k \end{bmatrix} \\ = dx_{k-1}^T (I_s + E_k E_k^T) \nabla_s F_k \\ = dx_{sk-1}^T \nabla_s L_k / 2 = 0 \quad (29)$$

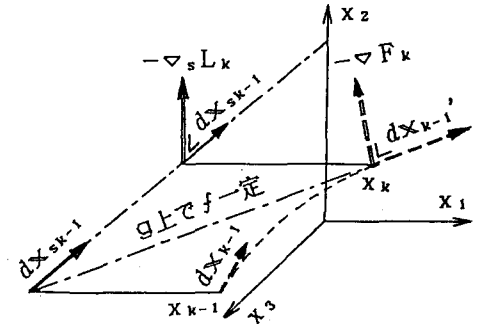


図-2 非線形制約面での直交性

$F_k \neq 0$ になる。これに対し、 $d\mathbf{x}_s$ は s 成分のみなので、 $\mathbf{x}_k$ 点でも $d\mathbf{x}_{sk-1}$ であり、常に直交性は保たれている。上記の関係を式で表したのが式29である。

5. 共役方向法，準Newton法の適用

(1) 適用に当たって

上記4. で $\nabla_s L$ ， ∇F の直交性が確かめられたので、いずれも共役方向法，準Newton法での最急勾配として使用可能であるが、両者の大きな相違点は $\nabla_s L$ が制約変数の選択に依存しているのに対し、 ∇F は全く依存していないことである。 $\nabla_s L$ のような縮小勾配では変数の区分が固定していないと共役方向法等の適用が不可能であるが、活性な制約面を得るときに変数の区分を固定するのは困難である。これに対し、 ∇F を最急勾配に採用すれば上記の問題点は解決されたことになる。

しかしながら、 ∇F には非線形制約面での直交性が保てないことが問題点として残されているが、ここでは ∇F による探索 (∇F 探索) が可能であることを確かめるため、線形制約面の場合について共役方向法等を適用してみた。比較のため $\nabla_s L$ による探索 ($\nabla_s L$ 探索) も行った。共役方向

法等は各探索方向についてラインサーチを行い、ラインサーチ最小点を求めていることから、式29の直交性は式30で表され、式30は共役方向法等の基礎となっている。⁶⁾

(2) 共役方向法

共役方向法での k 点の探索方向 $d\mathbf{x}_k$ は ∇F を用いると式30， $\nabla_s L$ を用いると式32になる。式30の $b_k(\nabla F)$ は式33に、式32の $b_k(\nabla_s L)$ は式34になる。 ∇F 探索と $\nabla_s L$ 探索とでは勾配の値が異なるため、途中の探索経路は異なってくるが、最適解では一致する筈である。

式30，式32で探索するとき、最適解に到達するまでのラインサーチ反復回数がどのようになるかを考えてみる。 ∇F は n 次元であるが、このうち $\nabla_s F$ のみが独立で $\nabla_m F$ は従属である。 $\nabla_s L$ は s 次元でこれに従属の $\nabla_m L$ を加えたのが n 次元の ∇L である。式30，式34はいずれも s 次元の $\nabla_s F$ で表されており、 f が2次式の場合は s 回の反復で最適解に到達すると予想される。これは、本来の f は n 次元であっても、制約面上の f または s 次元に射影された制約面上の f は s 次元に縮小していることから明らかである。

(3) 準Newton法

準Newton法のDFP公式に ∇F を用いると式35になり、 $\nabla_s L$ を用いると式35の $d\nabla F_k$ を $d\nabla_s L_k$ に、 $d\mathbf{x}_k$ を $d\mathbf{x}_{sk}$ に、 $\mathbf{l}$ を $\mathbf{l}_s$ に置き換えたものになる。⁷⁾ H_k はHesse行列の逆行列の近似行列であり、 ∇F 探索では n 次元、 $\nabla_s L$ 探索では s 次元になる。 $H_k(\nabla_s L)$ は s 次元の探索空間に対し、 s 次元の行列になっているので、DFP公式の標準的な使い方である。 $H_k(\nabla F)$ は s 次元の探索空間に対し n 次元の行列なので、 H_k の内容は標準型とは異なったものになる。

$n = 3$ ， $s = 2$ で f が2次式、制約式が線形の簡単な例題を採用したが、この例題では制約式を代入した

$$d\mathbf{x}_{sj}^T \nabla_s L_k = 0, (j < k) \quad \dots\dots (30)$$

$$d\mathbf{x}_k(\nabla F) = -\nabla F_k + b_k(\nabla F) d\mathbf{x}_{k-1} \quad (31)$$

$$d\mathbf{x}_k(\nabla_s L) = -\nabla L_k + b_k(\nabla_s L) d\mathbf{x}_{k-1} \quad (32)$$

$$\begin{aligned} b_k(\nabla F) &= \nabla F_k^T \nabla F_k / (\nabla F_{k-1}^T \nabla F_{k-1}) \\ &= \frac{\nabla_s F_k^T (I_s + E_k E_k^T) \nabla_s F_k}{\nabla_s F_{k-1}^T (I_s + E_{k-1} E_{k-1}^T) \nabla_s F_{k-1}} \quad (33) \end{aligned}$$

$$\begin{aligned} b_k(\nabla_s L) &= \nabla_s L_k^T \nabla_s L_k / (\nabla_s L_{k-1}^T \nabla_s L_{k-1}) \\ &= \frac{\nabla_s F_k^T (I_s + E_k E_k^T)^2 \nabla_s F_k}{\nabla_s F_{k-1}^T (I_s + E_{k-1} E_{k-1}^T)^2 \nabla_s F_{k-1}} \quad (34) \end{aligned}$$

$$\left. \begin{aligned} H_{k+1} &= H_k + \frac{d\mathbf{x}_k d\mathbf{x}_k^T}{d\mathbf{x}_k^T d\nabla F_k} - \frac{H_k d\nabla F_k d\nabla F_k^T H_k}{d\nabla F_k^T H_k d\nabla F_k} \\ \text{ただし} \quad H_0 &= I \\ d\mathbf{x}_k &= \mathbf{x}_{k+1} - \mathbf{x}_k \\ d\nabla F_k &= \nabla F_{k+1} - \nabla F_k \end{aligned} \right\} \quad \dots\dots (35)$$

f (Lagrange関数に相当)が2次式なので、 s 回のラインサーチで正確に最適解に到達する筈である。共役方向法, DFP公式のそれぞれに ∇F または $\nabla_s L$ を用いたが、いずれのケースでも s 回のラインサーチで最適解に到達した。特にDFP公式に ∇F を用いると近似行列 H_k は n 次元であるにもかかわらず、 s 回のラインサーチで最適解に到達することが確かめられた。

6. 結論

(1) 制約変数と探索変数の区分が固定できないことが $\nabla_s L$ のような縮小勾配を用いた探索に共役方向法, 準Newton法を適用するときの重大な障害になっていたが、変数区分に依存しない最急勾配 ∇F を採用することによりこの問題点を解決できた。

(2) 線形制約面を持つ f が2次式の例題に共役方向法, 準Newton法を適用して、 ∇F による探索では s 回のラインサーチにより最適解に到達することが確かめられた。

(3) $-\nabla F$ が活性な制約面上での正規化 df 最小の方向であることが確かめられたが、 $-\nabla F$ を最急降下法に用いると最も効率の良い探索方向になる。

(4) 活性な制約面上を進む探索方法では、非線形制約面の場合に $\nabla_s L$ では直交性が成立するが、 ∇F では直交性が成立しないので、この点の処理をどのようにしたらよいかが問題点として残されている。

(5) s 次元の探索空間に対し、DFP公式での近似行列 H_k は n 次元なので、 H_k の内容は標準型とは異なっており、 H_k の内容についての吟味が必要である。

参考文献

- 1) 志水清孝：システム最適化理論, コロナ社, PP.75 ~PP.82, 1976
- 2) 平田恭久：等式制約法における最小化法について, 土木学会第44回年次学術講演会講演概要集, I - 151
- 3) 平田恭久：等式制約法での縮小勾配の改良, 土木学会第46回年次学術講演会講演概要集, I - 382
- 4) 吉田洋一・高橋健人：線形代数学, 培風館, PP.110, 1961
- 5) Peressini, A.L., Sullivan, F.E., Uhl, Jr., J.J. : The Mathematics of Nonlinear Programming, Springer-Verlag, PP.112 ~PP.128, 1988
- 6) L.C.W.ディクソン (松原正一訳)：非線形最適化計算法, 培風館, PP.61 ~PP.71, 1974
- 7) Ralph, W.P. : Optimization for Engineering Systems, Van Nostrand Reinhold, PP.209~PP.222, 1986