

(1) 活性制約面と最適解の関係

RELATION BETWEEN A CHANGE OF ACTIVE CONSTRAINTS AND AN OPTIMUM SOLUTION

平田 恭久*
Yasuhisa HIRATA

*群馬高専教授 環境都市工学科 (〒371-0845 群馬県前橋市鳥羽町580)

If an active set of constraints changes in a search process, an optimum solution which is based on the active set also changes. The author took up the optimum problem of a quadratic function with linear inequality constraints, and studied the above change of the optimum solution. The author makes clear the transition from the former optimum solution to the latter optimum solution. Non-linear constraints can be approximately expressed by linear constraints. The above transition also is present between the optimum solution of the non-linear constraints and the linear approximation constraints. In this paper an equation that expresses the relation between the two optimum solutions is proposed, and some considerations in the above transition are stated.

Key Words: optimum solution, non-linear constraint, linear approximation

1. まえがき

非線形最適化手法の解法に関する主要な問題点は、①不等号制約式から活性制約式の選択、②効率の良い最小化、の2点である。上記①について著者は、等式制約法と称する活性制約式とその従属変数を選択する方法を考案し、これにより解決済みである。上記②については、準Newton法や共役方向法を用いて効率の良い探索方向を算出することが考えられるが、一般に非線形最適化問題では探索途中で活性制約面が変化してくる。これに伴い活性制約面上の目的関数が変化する(別の関数になる)ため、以前の探索過程で蓄積してきたデータが使用できなくなり、準Newton法や共役方向法を再出発させるという不便がある。よって、活性制約面の変化を前提にした最小化法が必要になってくる。

最適構造設計問題は、①目的関数は構造物の体積などの簡単な関数、②制約条件は応力度、たわみなどの解析に労力のかかる関数、で構成されるが普通である。探索方向の算出には、①制約面上の目的関数の1階微分である最急勾配方向、②制約面上の目的関数の2階微分であるHesse行列の利用、の二つがある。最急勾配方向をそのまま使用する最急降下法は効率的でなく、最急勾配方向の効率的利用である準Newton法や共役方向法は活性制約面の変化に対して不便である。Hesse行列の利用では制約条件式の2階微分に伴う計算量の増加がある。一般に最適化問題を解くときには制約条件式の1階微分は最低限必要なので、2階微分が簡単に得られる目的関数の

Hesse行列と制約条件式の1階微分を組み合わせた方法が浮上してくる。これはある点 x で線形近似した活性制約面について目的関数の最小点を求める問題であり、元の非線形最適化問題はこれを反復して解くことになるので、等式制約付き逐次2次計画法の形になる。

本論文では等式制約付き逐次2次計画法を解くのが主眼ではなく、非線形制約面での最適解の探索に関する種々の問題点に考察を加えていき、非線形最適化問題での効率的最小化の基礎を固めることに焦点を当てており、以下の点に着目することにした。探索途中での活性制約面の変化としては、①新たな制約式に抵触して活性制約式が以前とは異なったものになる、②非線形性の強い制約式による活性制約面の傾きの急激な変化、の2種類があるが、いずれの場合も当初に考えていた最適解の推定位置が大幅に移動するので、活性制約面と最適解の関係について知見が必要になる。Hesse行列を用いた探索方向の算出について非線形制約面では、①活性制約面の傾きが場所により変化する。②活性制約面上に乗っているために変数の拘束がある、という点で線形制約面とは異なった状態にあるので、これらの関係を把握しておく必要がある。

以上の事柄を出発点として本論文の構成は次のようになっている。第2章「活性制約面と最適解」では、縮小勾配を用いたHesse行列による勾配差法から、①異なった線形制約面、②非線形制約面と線形近似制約面、の二つの場合について最適解の移動量を表す式を誘導してい

る。第3章「Hesse行列と探索方向」では、Hesse行列とそれより得られる探索方向について線形項と非線形項を分類し、線形項のみから得られる探索方向は変数区分に依存しないが、非線形項を含む場合の探索方向は変数区分により異なった値を持つことを述べている。第4章「最適解の移動量と数値例」では、探索例でのHesse行列と探索方向を示し、最適解の移動量を表す式の探索例での適合度を調べているが、縮小勾配だけでなく全変数での最急勾配についてもHesse行列とその探索方向について触れている。第5章「線形近似制約面最小化への準備」では、探索例での種々の探索方法を比較し、変数区分と等式制約法との関係に触れ、最適解の移動量と線形近似制約面最小化との結び付きについて考察している。第6章は「まとめ」である。

2. 活性制約面と最適解

2.1 縮小勾配を用いたHesse行列

不等式制約付き非線形最適化問題は式(1)で表され、活性制約式 g_m とその従属変数 x_m を等式制約法を適用して選択すると式(2)になる。式(2)についてのLagrange関数は式(3)であり、式(3)の微分を独立変数 x_s と x_m に分けて表すと式(4)になる。式(4)第2行 $=0$ と置くと式(5) λ_m が得られ、 λ_m を式(4)第1行に代入すると式(6)になり、これが縮小勾配であるが、制約面の傾きを表す式(7) E を用いて書き直してある。 f の2階微分である。Hesse行列 G は式(8)であり、制約面上に存在する条件より dx_m と dx_s の関係は式(9)になる。

$$\begin{aligned} \min \quad & f(x) \\ \text{subject} \quad & g(x) \leq 0 \end{aligned} \quad \dots\dots\dots (1)$$

$$\begin{aligned} \min \quad & f(x) \\ \text{subject} \quad & g_m(x) = 0 \end{aligned} \quad \dots\dots\dots (2)$$

$$L(x, \lambda_m) = f(x) + g_m(x)^T \lambda_m \quad \dots\dots (3)$$

$$\left. \begin{aligned} \partial L / \partial x_s &= \nabla_s L = \nabla_s f + \nabla_s g_m^T \lambda_m \\ \partial L / \partial x_m &= \nabla_m L = \nabla_m f + \nabla_m g_m^T \lambda_m \end{aligned} \right\} \dots\dots (4)$$

$$\lambda_m = -(\nabla_m g_m^T)^{-1} \nabla_m f \quad \dots\dots (5)$$

$$\left. \begin{aligned} \nabla_s L &= \nabla_s f - \nabla_s g_m^T (\nabla_m g_m^T)^{-1} \nabla_m f \\ &= \nabla_s f + E \cdot \nabla_m f \end{aligned} \right\} \dots\dots (6)$$

$$E = -\nabla_s g_m^T (\nabla_m g_m^T)^{-1} \quad \dots\dots (7)$$

$$G = \begin{bmatrix} G_{ss} & G_{sm} \\ G_{ms} & G_{mm} \end{bmatrix} \quad \dots\dots (8)$$

$$dx_m = E^T dx_s \quad \dots\dots (9)$$

$$\left. \begin{aligned} d\nabla_s L / dx_s &= d\nabla_s f / dx_s + d(E \cdot \nabla_m f) / dx_s \\ &= \partial \nabla_s f / \partial x_s + \partial \nabla_s f / \partial x_m \cdot dx_m / dx_s \\ &\quad + \partial (E \cdot \nabla_m f) / \partial x_s + \partial (E \cdot \nabla_m f) / \partial x_m \cdot dx_m / dx_s \\ &= G_{ss} + G_{sm} \cdot E^T + \nabla_s (E \cdot \nabla_m f) + \nabla_m (E \cdot \nabla_m f) E^T \\ &= G_{ss} + G_{sm} \cdot E^T + (\nabla_s E \cdot \nabla_m f) + E \cdot G_{ms} \\ &\quad + (\nabla_m E \cdot \nabla_m f) E^T + E \cdot G_{mm} \cdot E^T = G_{sL} \end{aligned} \right\} \dots\dots (10)$$

$$dx_s = -G_{sL}^{-1} \nabla_s L \quad \dots\dots (11)$$

$$G_{sL} \cdot dx_s = \nabla_s L^* - \nabla_s L = -\nabla_s L \quad \dots\dots (12)$$

式(6) $\nabla_s L$ を x_s で微分して縮小勾配を用いたHesse行列 G_{sL} を導くと式(10)になる。独立変数 x_s が s 個とすると G_{sL} は $s \times s$ の行列である。式(10)で $\nabla_s E$ ($\nabla_m E$ も同様)は3次元の配列であり、 $(\nabla_s E \cdot \nabla_m f)$ は行列になるが、制約面の傾き E の変化を表しており、線形制約面ではこの非線形項は0になる。式(10) G_{sL} と式(6) $\nabla_s L$ を用いてある点 x での探索方向 dx_s は式(11)で得られるが、逆行列のままでは扱いにくいので、勾配差式の式(12)を考察での基本式として用いる。式(12)で $\nabla_s L^* = 0$ は最適解 x^* での縮小勾配であり、式(12)は x^* と x との勾配差が $-\nabla_s L$ であることを示している。

2.2 異なった線形制約面

制約面の傾き E_1, E_2 の線形制約面が点 x で交差しており、その最適解が x_1^*, x_2^* の場合を図-1に示す。¹⁾線形制約式は式(13)になり、式(14) D を用いて x_m と x_s の関係が式(15)で表されるが、式(15)は非線形制約面では成立しない。式(13)では定数項 $=b$ としているが、定数項 $=0$ での最適解との差を求めると式(16)になり、定数項による最適解の移動量を表している。制約面の傾きの差に基づく勾配差は式(17)になり、これより最適解の移動量 dx^* は式(18)、式(19)で表される。式(18)、式(19)と対になるのが式(20)、式(21)である。

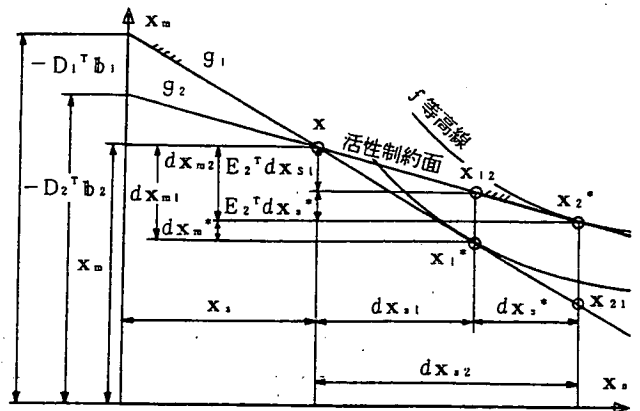


図-1 線形制約面 g_1, g_2 と最適解 x_1^*, x_2^* の関係

$$g_m(x) = \nabla g_m \cdot x = b \quad \dots\dots (13)$$

$$D = -(\nabla_m g_m^T)^{-1} \quad \dots\dots (14)$$

$$x_m = E^T x_s - D^T b \quad \dots\dots (15)$$

$$x_0^* - x_b^* = -G_{sL}^{-1} (G_{sm} + E \cdot G_{mm}) D^T b \quad \dots\dots (16)$$

$$G_{sL2} \cdot dx_{s2} - G_{sL1} \cdot dx_{s1} = -(E_2 - E_1) \nabla_m f \quad (17)$$

$$\begin{aligned} (G_{ss} + E_2 \cdot G_{ms}) dx_{s1}^* + (G_{sm} + E_2 \cdot G_{mm}) dx_m^* \\ = -(E_2 - E_1) \nabla_m f^* \end{aligned} \quad \dots\dots (18)$$

$$G_{sL2} \cdot dx_{s1}^* = -\nabla_{sL12} \quad \dots\dots (19)$$

$$\begin{aligned} (G_{ss} + E_1 \cdot G_{ms}) dx_{s2}^* + (G_{sm} + E_1 \cdot G_{mm}) dx_m^* \\ = -(E_2 - E_1) \nabla_m f^* \end{aligned} \quad \dots\dots (20)$$

$$G_{sL1} \cdot dx_{s2}^* = \nabla_{sL21} \quad \dots\dots (21)$$

2.3 非線形制約面と線形近似制約面

非線形制約面の最適解を x_2^* 、ある点 x での線形近似

制約面上の最適解 x_1^* とし、線形近似制約面の傾き E_1 、 x_2^* での非線形制約面の傾きを E_2 とする。上記2.2では、点 x での縮小勾配は E_1 上が ∇_{sL1} 、 E_2 上が ∇_{sL2} となっていたが、図-2の点 x では二つの制約面の傾きは E_1 なので、縮小勾配は共に ∇_{sL1} である。

点 x での情報で勾配差式を導くとすると、 G_{sL1} と G_{sL2} での制約面の傾きは共に E_1 であるが、 G_{sL2} には式(10)での非線形項が含まれており、 G_{sL1} には含まれていない。勾配差式を導くと式(22)が得られ、 $G_{sL1} \cdot dx_{s1}$ は式(23)として表すこともできる。 G_{sL2} の非線形項に dx_{s2} を掛けると式(24)のようになる。式(22)から式(23)を引くと勾配差での非線形項 $dE \cdot \nabla_{mF}$ は式(25)としても表すことができる。ここで導いている式には非線形項が含まれているので、近似解である。

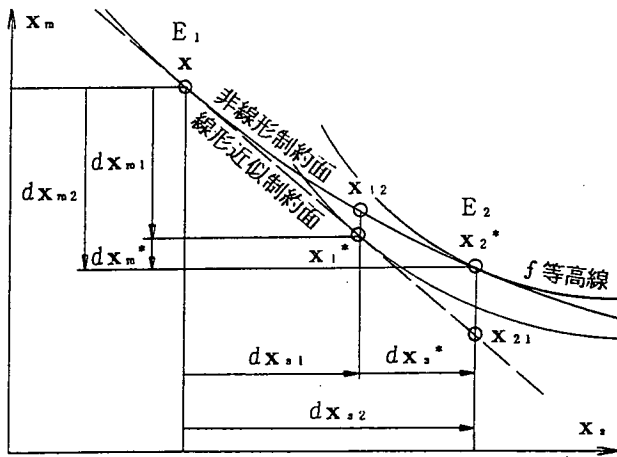


図-2 非線形制約面と点 x での線形近似制約面

$$\begin{aligned} G_{sL2} \cdot dx_{s2} - G_{sL1} \cdot dx_{s1} \\ = G_{sL1} \cdot dx_{s2}^* + \{(\nabla_s E \cdot \nabla_{mF}) + (\nabla_s E \cdot \nabla_{mF}) E_1^T\} dx_{s2} = G_{sL1} \cdot dx_{s2}^* + dE \cdot \nabla_{mF} \\ = \nabla_{sL2}^* - \nabla_{sL1} - (\nabla_{sL1}^* - \nabla_{sL1}) = \nabla_{sL2}^* = 0 \end{aligned} \quad (22)$$

$$\begin{aligned} G_{sL1} \cdot dx_{s2}^* = \nabla_{sL21} - \nabla_{sL1}^* = \nabla_{sL21} \quad (23) \\ \{(\nabla_s E \cdot \nabla_{mF}) + (\nabla_m E \cdot \nabla_{mF}) E_1^T\} dx_{s2} \\ = dE \cdot \nabla_{mF} = (E_2 - E_1) \nabla_{mF} \end{aligned} \quad (24)$$

$$dE \cdot \nabla_{mF} = \nabla_{sL2}^* - \nabla_{sL21} = -\nabla_{sL21} \quad (25)$$

3. Hesse行列と探索方向

3.1 Hesse 行列 G_{sL} と探索方向 dx_s の内訳

Hesse行列 G_{sL} とそれより得られる探索方向 dx_s を図-2を参照して線形項と非線形項に分けると式(26)になる。点 x での縮小勾配 ∇_{sL1} を用いた勾配差式は ∇_{sL2}^* について式(27)が、 ∇_{sL1}^* について式(28)が成立する。両者は共に右边が $-\nabla_{sL1}$ になっており、これを利用すると式(27)、式(28)から式(29)が得られる。式(29)の $G_{sL2} - G_{sL1}$ は非線形項すなわち E の変化を表しており、 dx_{s2} を掛けることにより $dE \cdot \nabla_{mF}$ となるが、この関係は式(23)、式(24)からも得られる。

縮小勾配 ∇_{sL} は独立変数 x_s の次元で算出しており、従属変数 x_m をどのように選択するかにより ∇_{sL} は異なった値になる。線形項のみで算出した dx_{s1} とこのときの従属変数の変化 dx_{m1} を合わせた dx_{s1} は、変数区分をどのように選択していても同じ値を持つという特性がある。非線形項も含んで算出した dx_{s2} とこのときの従属変数の変化 dx_{m2} を合わせた dx_{s2} は、変数区分に依存した値となり、変数区分により dx_{s2} の絶対値が大きく変化する。このため、 x_m の選択に注意が必要になる。

$$\begin{aligned} G_{sL1} \cdot dx_{s1} & : \text{線形項} \\ G_{sL2} - G_{sL1}, dx_{s2}^* = dx_{s2} - dx_{s1} & : \text{非線形項} \end{aligned} \quad (26)$$

$$G_{sL2} \cdot dx_{s2} = (G_{sL1} + G_{sL2} - G_{sL1})(dx_{s1} + dx_{s2}^*) = G_{sL1} \cdot dx_{s1} + G_{sL1} \cdot dx_{s2}^* + (G_{sL2} - G_{sL1}) \cdot dx_{s2} = \nabla_{sL2}^* - \nabla_{sL1} = -\nabla_{sL1} \quad (27)$$

$$G_{sL1} \cdot dx_{s1} = \nabla_{sL1}^* - \nabla_{sL1} = -\nabla_{sL1} \quad (28)$$

$$G_{sL1} \cdot dx_{s2}^* = -(G_{sL2} - G_{sL1}) dx_{s2} = -dE \cdot \nabla_{mF} = -(E_2 - E_1) \nabla_{mF} \quad (29)$$

3.2 勾配差式と変数交換

dx_{s1} と dx_{s2} では変数区分の与える影響が異なっている理由を説明するため、等式制約法で用いている変数交換を取り上げる。²⁾ 変数交換とはある X_{s1} とある X_{m1}' とを交換することであり、これにより、変数区分が変化する。 X_1, X_2, X_3 のどれか一つが従属変数 x_m になるような3変数の例題を用いて、勾配差式 $d\nabla_{sL}$ での変数交換を説明する。図-3は変数交換しない場合で $x_m = X_1$ または X_2 または X_3 の3通りが示してあり、空白の

	$G_{sL1} \cdot dx_{s1}^*$	$G_{sL1} \cdot dx_{s2}^*$	$(G_{sL2} - G_{sL1}) dx_{s2}$	$-\nabla_{sL1}$
(X_1)				
X_2	$(d\nabla_{2F} + e_2 \cdot d\nabla_{1F})_{s1}$	$(d\nabla_{2F} + e_2 \cdot d\nabla_{1F})_{s2}^*$	$de_2 \cdot \nabla_{1F}$	$-\nabla_{2F} - e_2 \cdot \nabla_{1F}$
X_3	$(d\nabla_{3F} + e_3 \cdot d\nabla_{1F})_{s1}$	$(d\nabla_{3F} + e_3 \cdot d\nabla_{1F})_{s2}^*$	$de_3 \cdot \nabla_{1F}$	$-\nabla_{3F} - e_3 \cdot \nabla_{1F}$
X_1	$(d\nabla_{1F} + 1/e_2 \cdot d\nabla_{2F})_{s1}$	$(d\nabla_{1F} + 1/e_2 \cdot d\nabla_{2F})_{s2}^*$	$d(1/e_2) \nabla_{2F}$	$-\nabla_{1F} - 1/e_2 \cdot \nabla_{2F}$
(X_2)				
X_3	$(d\nabla_{3F} - e_3/e_2 \cdot d\nabla_{2F})_{s1}$	$(d\nabla_{3F} - e_3/e_2 \cdot d\nabla_{2F})_{s2}^*$	$d(-e_3/e_2 \cdot de_2) \nabla_{2F}$	$-\nabla_{3F} + e_3/e_2 \cdot \nabla_{2F}$
X_1	$(d\nabla_{1F} + 1/e_3 \cdot d\nabla_{3F})_{s1}$	$(d\nabla_{1F} + 1/e_3 \cdot d\nabla_{3F})_{s2}^*$	$d(1/e_3) \nabla_{3F}$	$-\nabla_{1F} - 1/e_3 \cdot \nabla_{3F}$
X_2	$(d\nabla_{2F} - e_2/e_3 \cdot d\nabla_{3F})_{s1}$	$(d\nabla_{2F} - e_2/e_3 \cdot d\nabla_{3F})_{s2}^*$	$d(-e_2/e_3) \nabla_{3F}$	$-\nabla_{2F} + e_2/e_3 \cdot \nabla_{3F}$
(X_3)				

図-3 変数交換なし ($x_m = X_1$ または X_2 または X_3)

	$G_{sL1} \cdot d\mathbf{x}_s^*$	$G_{sL1} \cdot d\mathbf{x}_s^*$	$(G_{sL2} - G_{sL1})d\mathbf{x}_{s2}$	$-\nabla_{sL1}$
(X ₁)				
X ₂	$(d\nabla_{2F} + e_2 \cdot d\nabla_{1F})_{s1}$	$(d\nabla_{2F} + e_2 \cdot d\nabla_{1F})_{s1}^*$	$de_2 \cdot \nabla_{1F}$	$-\nabla_{2F} - e_2 \cdot \nabla_{1F}$
X ₃	$(d\nabla_{3F} + e_3 \cdot d\nabla_{1F})_{s1}$	$(d\nabla_{3F} + e_3 \cdot d\nabla_{1F})_{s1}^*$	$de_3 \cdot \nabla_{1F}$	$-\nabla_{3F} - e_3 \cdot \nabla_{1F}$
(X ₂)				
X ₁	$(d\nabla_{1F} + 1/e_2 \cdot d\nabla_{2F})_{s1}$	$(d\nabla_{1F} + 1/e_2 \cdot d\nabla_{2F})_{s1}^*$	$1/e_2 \cdot de_2 \cdot \nabla_{1F}$	$-\nabla_{1F} - 1/e_2 \cdot \nabla_{2F}$
X ₃	$(d\nabla_{3F} - e_3/e_2 \cdot d\nabla_{2F})_{s1}$	$(d\nabla_{3F} - e_3/e_2 \cdot d\nabla_{2F})_{s1}^*$	$(de_3 - e_3/e_2 \cdot de_2) \nabla_{1F}$	$-\nabla_{3F} + e_3/e_2 \cdot \nabla_{2F}$
(X ₃)				
X ₂	$(d\nabla_{2F} - e_2/e_3 \cdot d\nabla_{3F})_{s1}$	$(d\nabla_{2F} - e_2/e_3 \cdot d\nabla_{3F})_{s1}^*$	$(de_2 - e_2/e_3 \cdot de_3) \nabla_{1F}$	$-\nabla_{2F} + e_2/e_3 \cdot \nabla_{3F}$
X ₁	$(d\nabla_{1F} + 1/e_3 \cdot d\nabla_{3F})_{s1}$	$(d\nabla_{1F} + 1/e_3 \cdot d\nabla_{3F})_{s1}^*$	$1/e_3 \cdot de_3 \cdot \nabla_{1F}$	$-\nabla_{1F} - 1/e_3 \cdot \nabla_{3F}$

図-4 $\mathbf{x}_m = X_1$ から X_2 または X_3 へ変数交換

変数行は $\mathbf{x}_m$ 行((X₁)のように表示)である。 $d\nabla_{1F}$ 等は ∇_{1F} の勾配差, e_2 等は $\mathbf{x}_m = X_1$ のときのEの X_2 行要素、 $de_2 \cdot \nabla_{1F}$ 等は $dE \cdot \nabla_{mF}$ の X_2 行要素である。式(27), 式(28), 式(29)の関係が図-3に盛り込まれている。

図-4は $\mathbf{x}_m = X_1$ または X_2 または X_3 へ変数交換した場合であり、記号等の説明は図-3と同じである。図-5は $\mathbf{x}_m = X_1$ から $\mathbf{x}_m = X_2$ へ変数交換することにより式(7)E, 式(14)Dが変化する状況を示しているが、変数交換に伴い変数行の位置が移動している。これと同じ操作が図-4の勾配差式にも適用されている。

(X ₁)	d_1	0	0
X ₂	e_2	1	0
X ₃	e_3	0	1

↓ $\mathbf{x}_m = X_1$ から X_2 へ変数交換

(X ₂)	$d_1 - d_1 \cdot 1/e_2 = 0$	$1 - d_1 \cdot 1/e_2$	1
X ₁	$e_2/e_2 = 1$	$1/e_2$	0
X ₃	$e_3 - e_3 \cdot 1/e_2 = 0$	$0 - e_3 \cdot 1/e_2$	1

図-5 E, (e_2, e_3), D (d_1)についての変数交換

3.3 変数交換に伴う $d\mathbf{x}_1$ と $d\mathbf{x}_2$ の相違点

図-3と図-4を対比して変数交換の影響を考えてみる。 $-\nabla_{sL1}$ と線形項のみの $G_{sL1} \cdot d\mathbf{x}_{s1}$ では変数交換しても、変数交換しない場合と同じ式になっているが、非線形項を含む $(G_{sL2} - G_{sL1})d\mathbf{x}_{s2}$ では変数交換により、変数交換しない場合とは異なった式になっている。これが重要な相違点であり、これに基づいて以下の説明がなされる。

$d\mathbf{x}_{s1}$ が変数区分に関係なく同じ値を持つ理由は、① $d\nabla_{1F}$ は変数交換に伴って変数行・変数区分を変えているが $d\nabla_{1F}$ の値は変化していない(数値例からも確認できる)、② $d\nabla_{1F} = \nabla_{1F} \cdot dX_1$ で ∇_{1F} はEに無関係である、③よって $d\nabla_{1F}$ が変化しないことは $d\nabla_{1F} = \nabla_{1F} \cdot dX_1$ より変数交換に伴って dX_1 が変化しないことを意味している、④変数交換した場合としない場合では値が同じかつ式が同じであることは dX_1 が変数区分に関係なく同

じであることを意味する、のように説明できる。

$d\mathbf{x}_{s2}$ が変数区分により異なった値を持つ理由は、①変数交換に伴って dX_1 が変化しないことは $d\mathbf{x}_{s1}$ と同じで最初の $\mathbf{x}_m$ 選択で定まる、②式(29)より図-4の $(G_{sL2} - G_{sL1})d\mathbf{x}_{s2} = -G_{sL1} \cdot d\mathbf{x}_s^*$ であり $G_{sL1} \cdot d\mathbf{x}_s^*$ の $d\nabla_{sL1}$ は $G_{sL1} \cdot d\mathbf{x}_{s1}$ の $d\nabla_{sL1}$ と同じ式である(添字に示すように $d\mathbf{x}_{s1}$ と $d\mathbf{x}_s^*$ の違いがあるが)、③変数交換した場合としない場合とは $(G_{sL2} - G_{sL1})d\mathbf{x}_{s2}$ の式が異なる(値が異なることは数値例からも確認できる)、④ $(G_{sL2} - G_{sL1})d\mathbf{x}_{s2}$ と同じ絶対値の $G_{sL1} \cdot d\mathbf{x}_s^*$ が上記②と $G_{sL1} \cdot d\mathbf{x}_{s1}$ の性質より変数交換後も同じ式である、⑤変数交換した場合としない場合では値が異なりかつ式が同じであることは dX_1 が異なることを意味する、⑥以上より図-4の dX_1 が図-3の dX_s とは異なっている。最初の $\mathbf{x}_m$ 選択に特有な値である、のように説明できる。

4. 最適解の移動量と数値例

4.1 数値例と $\mathbf{x}_m$ 選択の影響

上記3.2, 3.3で述べたことを数値的に確認するため、式(30)の非線形等式制約付き2次関数を探索例として用いる。探索例の最適解は式(31), 点 $\mathbf{x}$ の座標は式(32)であるが、この探索例は非線形性の強い制約式である。縮小勾配を用いたHesse行列 G_{sL2} (非線形項を含む)、 G_{sL1} (線形項のみ)と $G_{sL2} - G_{sL1}$ (非線形項の $\nabla E \cdot \nabla_{mF}$)の数値例を表-1に示すが、 ∇E は式(7)Eを微分して求めている。 $\mathbf{x}_m$ 選択による影響を調べるため $\mathbf{x}_m = X_1$, または X_2 または X_3 の3ケースであるが、 G_{sL1} は独立変数 $\mathbf{x}_s$ に係わるものなので、 $\mathbf{x}_m$ 行は空白行になる。

表-1の G_{sL2} , G_{sL1} により式(11)で $d\mathbf{x}_{s2}$ (非線形項を含む), $d\mathbf{x}_{s1}$ (線形項のみ)を算出したが、 $d\mathbf{x}_s^* = d\mathbf{x}_{s2} - d\mathbf{x}_{s1}$ と $d\mathbf{x}_s$ を算出するときの点 $\mathbf{x}$ の縮小勾配 ∇_{sL1} と合わせて表-2に示す。 $d\mathbf{x}_1$ の $\mathbf{x}_m$ 行は $d\mathbf{x}_{m1} = E_1^T d\mathbf{x}_{s1}$ で求めたものであるが、どの X_1 を $\mathbf{x}_m$ として選択しても同じ $d\mathbf{x}_1$ になっている。 $d\mathbf{x}_2$ については $\mathbf{x}_m$ 選択の影響が大きく、絶対値にかなりの差がある。表-2に真の最適解への探索方向である最適 $d\mathbf{x}$ が付記してあり、

表-1 Hesse行列 G_{sL2} (非線形項を含む), G_{sL1} (線形項のみ)

	G_{sL2}		G_{sL1}		$G_{sL2} - G_{sL1}$	
(X_1)						
X_2	33.100	-301.03	7.6477	-76.991	25.452	-224.04
X_3	-301.03	2740.0	-76.991	1562.7	-224.04	1177.2
X_1	41.098	-207.44	10.940	-46.676	30.158	-160.77
(X_2)						
X_3	-207.44	1185.8	-46.676	986.27	-160.77	199.49
X_1	12.228	-1.9291	9.7096	2.0467	2.5186	-3.9757
X_2	-1.9291	4.6412	2.0467	4.2809	-3.9757	0.36026
(X_3)						

表-2 探索方向 dx_{s2} (非線形項を含む), dx_{s1} (線形項のみ)

	dx_2	dx_1	dx^*	∇_{sL1}	最適 dx	$dx_{2(m)}$	∇_{sL2}
(X_1)	-21.666	-0.07046			-0.1	-2.7496	
X_2	-38.736	-0.09181	-38.644	-0.19164	-0.175	-38.736	-168.50
X_3	-4.2598	-0.01160	-4.2482	11.063	-0.01	-4.2598	-153.30
X_1	-0.34441	-0.07046	-0.27395	0.22929	-0.1	-0.34441	-0.50957
(X_2)	-0.60687	-0.09181			-0.175	0.12023	
X_3	-0.06713	-0.01160	-0.05553	8.1543	-0.01	-0.06713	-9.6628
X_1	-0.09586	-0.07046	-0.02540	0.87203	-0.1	-0.09586	0.00614
X_2	-0.15560	-0.09181	-0.06379	0.53722	-0.175	-0.15560	0.06780
(X_3)	-0.01780	-0.01160			-0.01	-0.00994	

表-3 制約面の傾き E_1 , E_2 と最適解の移動量を表す $G_{sL1} \cdot dx_{s1}^*$

	E_1	E_2	$E_2 - E_1$	$G_{sL1} \cdot dx_{s1}^*$	∇_{sL21}	$-(E_{21} - E_1) \cdot \nabla_{mF21}$
(X_1)						
X_2	-0.83582	0.004805	0.84063	31.536	31.734	29.585
X_3	12.687	0.61476	-12.072	-3663.6	-3663.6	-3415.7
X_1	-1.1964	2.1283	3.3248	-0.40526	-0.40526	2.7236
(X_2)						
X_3	15.179	-7.8110	-22.990	-41.978	-41.978	-35.628
X_1	0.07882	0.01239	-0.06643	-0.37717	-0.37717	-0.41034
X_2	0.06588	0.003051	-0.06283	-0.32506	-0.32506	-0.35364
(X_3)						

$dx_{m2} = E_1^T dx_{s2}$ で定めた dx_{m2} を合わせた dx_2 と比較すると、 $x_m = X_3$ のケースが絶対値としては最も良い値になっている。 dx_1 は比較的良好で最適 dx に比べて絶対値が若干小さめになっている。

$$\begin{aligned} \min f(x) &= 5X_1^2 + 2X_2^2 + 2X_3^3 + 2X_1X_2 + 2X_2X_3 \\ &\quad - 2X_1X_3 - 6X_3 \\ \text{subject to } g(x) &= X_1^2 - 3X_1X_2 + 2X_2X_3 - X_1X_3 \\ &\quad + 4X_3 = 0 \end{aligned} \quad (30)$$

$$x^{*T} = [0, 0, 0] \text{ のとき } f^* = 0 \quad \dots\dots (31)$$

$$x^T = [0.1, 0.175, 0.01] \text{ のとき } f = 0.08795 \quad \dots(32)$$

4.2 最適解の移動量

表-2 で示すように dx_2 は x_m 選択により差が生じているが、これに最も関係しているのが制約面の傾き E である。表-3 に点 x での E_1 と dx_{s2} より求めた x_2^* での E_2 を示す。最も良好な探索方向である $x_m = x_3$ のケースでは E_1 の絶対値は小さい。

最適解の移動量 dx_{s1}^* を表す式の一つに式(23)があり、左辺の $G_{sL1} \cdot dx_{s1}^*$ と右辺の ∇_{sL21} とが良い精度で一致しているのは、線形近似制約面の E_1 で計算しているためである。式(29)では $G_{sL1} \cdot dx_{s1}^* = -dE \cdot \nabla_{mF}$ と表しているが、右辺に非線形項が含まれているため、数値例での適合度は悪くなっている。 $-dE \cdot \nabla_{mF}$ の中でどのケースにも比較的合致しているのは、 x_2^* と x_3 が

同じで線形近似制約面上の点 x_{21} での $-(E_{21} - E_1) \nabla_{mF21}$ である。これらの状況が表-3に示されている。目的関数 f のHesse行列 G は制約面に関係なく、探索例では f が2次関数なので G は定数になり、式(33)の勾配差式についての数値例では右辺と左辺は正確に一致する。

$$G \cdot dx = \nabla_{F2} - \nabla_{F1} \quad \dots\dots (33)$$

4.3 全変数での最急勾配 ∇F

非線形制約面では縮小勾配 ∇_{sL} を微分した G_{sL} より得られる dx_s は、 x_m の選択により異なった値を持つことから、 ∇_{sL} ではなく全変数での最急勾配 ∇F を微分したHesse行列 G_F を取り上げることにした。³⁾ ∇_{sL} では制約式について $g_m = 0$ を満足するための x_m が選択されているので、差分キザミ Δx_j や探索方向 dx_s に伴う $g_m = 0$ をどの変数で満足させるかは自動的に決まってくるが、 ∇F では $g_m = 0$ の満足をもどのようにするかを決めておく必要がある。

∇F は ∇_{sF} と ∇_{mF} に別れて式(34)で表されるが、 I_s は $s \times s$ の単位行列であり、 E を含む行列の逆行列が ∇_{sF} の算出に必要である。このため ∇F を微分して G_F を求めるのは困難なので、差分で G_F を算出することにした。式(30)の探索例について差分キザミ Δx_j に対して x_j 以外の任意の変数で $g_m = 0$ を満足させると、任意の変数の指定により差分が異なった値を持つことになる。これでは G_F を特定することができないので、 $\nabla_{m(i)} g_m$

方向 (Δx_j に対し x_j 以外の変数に関する g_m の勾配ベクトル方向)に x_j 以外の変数全部で分担して $g_m = 0$ を満足させることにした。

$$\left. \begin{aligned} \nabla_{sF} &= (I_s + E \cdot E^T)^{-1} \nabla_{sL} \\ \nabla_{mF} &= E^T \nabla_{sF} \end{aligned} \right\} \quad \dots\dots (34)$$

表-4は ∇F の差分によるHesse行列 G_F である。差分キザミ Δx_j で説明すると $g_m = 0$ を満足させる変数として、 x_2 または x_3 を指定している場合と x_2, x_3 で分担する場合が示してある。 x_m 分担での値は x_m 指定での値の中間になっている。 ∇_{sL} の差分により G_{sL} を求めるとHesse行列での対称性が得られること(表-1の G_{sL} は微分で求めたものであるが対称性がある)を確認しているが、表-4の G_F では対称性が得られていない。これは G_F の差分計算に問題があることを意味しており、原因は $g_m = 0$ の満足にあると考えられるが、制約面上での差分計算では $g_m = 0$ が達成できないと非線形項が表現できなくなる。 x_m 指定では変数は任意に指定しているので、 G_F は $2^3 = 8$ 通りの組み合わせができる。

5. 線形近似制約面最小化への準備

5.1 探索方向の比較

探索方向の善し悪しを判断する基準として、①正規化 dx と式(35)正規化 df 、② $x + dx$ での f と dx 絶対値、③ dx 方向 line search 最小点での f 、が考えられる。

⁴⁾ 通常の探索では上記③が標準であるが、ここではline search の代わりに最適 dx (真の最適解 $-x$) を基準にした式(36)を用いている。 dx についても $g_m = 0$ をどのように満足させるかが問題で、これにより f, df が変化してくる。上記③相当では $x + \text{式(36)}dx_s$ で $x_m = x_3$ として $g_m = 0$ を満足させており、 $f_m(\text{line})$ で表示してある。上記①、②相当では dx のままで $g_m = 0$ は満足させておらず、 $f(dx)$ 、正規化 df で表示してある。

$$\text{正規化 } df = \nabla f \cdot \text{正規化 } dx \quad \dots\dots (35)$$

$$\left. \begin{aligned} \text{line search 相当 } dx_s \\ = dx_s \cdot | \text{最適 } dx_s | / | dx_s | \end{aligned} \right\} \quad \dots\dots (36)$$

x 点でのこれらのデータを表-5、表-6に示す。縮小勾配 ∇_{sL} を用いて非線形項を含む G_{sL2} より算出した

表-4 最急勾配 ∇F の差分によるHesse行列 G_F

		$\nabla_1(\nabla F)$	$\nabla_2(\nabla F)$	$\nabla_3(\nabla F)$
x_m 指定	x_m	x_2	x_3	x_1
	x_1	14.650	-2.0511	155.48
	x_2	-7.4943	4.6015	-26.191
	x_3	-0.03370	0.69784	10.275
	x_m	x_3	x_1	x_2
	x_1	12.212	-12.132	-31.132
	x_2	-2.0563	6.2912	6.9889
	x_3	0.80891	0.00913	10.586
x_m 分担	x_m	x_2, x_3	x_3, x_1	x_1, x_2
	x_1	12.227	-2.1145	78.255
	x_2	-2.0803	4.6121	13.352
	x_3	0.80520	0.69353	10.453

表-5 探索方向の比較(最適 dx, dx_2, dx_1)

		最適 dx	$dx_{2 \times 1}$	$dx_{2 \times 2}$	$dx_{2 \times 3}$	dx_1
dx	x_1	-0.1	-21.666	-0.34441	-0.09586	-0.07046
	x_2	-0.175	-38.736	-0.60687	-0.15560	-0.09181
	x_3	-0.01	-4.2598	-0.06713	-0.01781	-0.01160
$ dx $		0.20180	44.587	0.70101	0.18362	0.11631
正規化	x_1	-0.49553	-0.48592	-0.49131	-0.52206	-0.60580
	x_2	-0.86718	-0.86876	-0.86570	-0.84738	-0.78934
	x_3	-0.04955	-0.09554	-0.09576	-0.09698	-0.09975
正規化 df		-1.1690	-0.89046	-0.89352	-0.91048	-0.95233
$f_m(\text{line})$		0.0	2.21×10^{-5}	2.28×10^{-6}	2.84×10^{-4}	4.61×10^{-3}
$f(dx)$		0.0	7168.6	1.2536	0.04773	0.03257

dx_2 は、 x_m 選択により dx_2 絶対値は大きく異なっているが、正規化 dx_2 は正規化最適 dx に近いものになっている。線形項のみの $G_{s,1,1}$ より算出した dx_1 は、 dx_1 絶対値が最適 dx 絶対値の約1/2であり、 dx_2 に次ぐ探索方向である。全変数での最急勾配 ∇F を用いて非線形項を含む G_r より算出した dx_r は、 dx_2 、 dx_1 より若干悪くなっているが、これは差分による G_r 算出に問題があると考えられる。 dx_r 指定1は x_m 指定の中で最も良いケース、 dx_r 指定2は x_m 指定の中で最も悪いケースである。 dx_r では $g_m > 0$ (抵触側)になっており、 dx_r 指定1で $f(dx) < 0$ と真の最適解での $f = 0$ より小さくなっているのはこのためである。最急勾配方向 $dx = -\nabla_r$ はHesse行列を用いていないため探索方向としては劣っているが、縮小勾配での $dx_s = -\nabla_{s,1}$ よりは優れており、 $-\nabla_r$ 、 $-\nabla_{s,1}$ は最急降下法として用いられる。 dx_1' は最適解の移動量を表す式(29)を参考にして dx_1 の改良を試みたものである。

5.2 x_m 選択と等式制約法

全変数での最急勾配 ∇F の差分より求めたHesse行列 G_r には、非線形性を考慮した差分計算に問題があり、良好な探索方向 dx_r が得られないので、縮小勾配 $\nabla_{s,1}$ の微分または差分より求めた $G_{s,1}$ を使用した方が良いと考えられる。非線形項を含む $G_{s,1,2}$ を使用する場合には、探索方向としては良好であっても x_m の選択により dx_2 の絶対値が大きく異なってくるので、この点に関する配慮が必要である。 dx_2 絶対値が最適 dx 絶対値に近ければline searchでのstep幅調整が楽になり、場合によってはline searchを省略することも可能になる。

数値例では dx_2 絶対値が小さいケースで E の絶対値が小さくなっているが、図-2の非線形制約面と表-3の制約面の傾き E_1 、 E_2 とを合わせて考えてみる。 E の絶対値が小さければ線形近似制約面での x_1^* と非線形制約面での x_2^* とが近接してくると推定されるので、点 x での E_1 の絶対値を小さくするような x_m の選択が浮上してくる。

$$\left. \begin{aligned} G_{s,1} \cdot dx_s^* &= -(E_2 - E_1) \nabla_{m,r} \\ &= -(\nabla_s g_{m,2}^T \lambda_{m,2} - \nabla_s g_{m,1}^T \lambda_{m,1}) \end{aligned} \right\} \dots (37)$$

$$\left. \begin{aligned} \lambda_{i,j} &> 0 \text{ かつ } \min \lambda_{i,j} \text{ となる} \\ j \text{ 行を } x_m \text{ 行として選択する} \end{aligned} \right\} \dots (38)$$

式(29)右辺の E_1 、 E_2 に式(5) λ_m と式(7) E を合わせて適用すると式(37)が導ける。式(37)左辺の $dx_s^* = x_2^* - x_1^*$ は x_2^* 、 x_1^* の近接度合を表しており、式(37)右辺より $\lambda_{m,2}$ 、 $\lambda_{m,1}$ の絶対値がこれに関係しているので、 λ_m の絶対値を小さくするような x_m 選択を考えてみる。等式制約法での x_m 行選択規則は式(38)のようになっている。このように等式制約法での x_m 行選択規則を用いていけば、 λ_m の絶対値が小さい $\Rightarrow E$ の絶対値が小さい $\Rightarrow dx_2$ の絶対値が小さい、のようになり、 $dx_{s,2}$ 算出での x_m 選択が自動的に解決される。⁵⁾

等式制約法は活性制約面とその従属変数の選択に特色があるが、縮小勾配を用いた非線形項を含むHesse行列 $G_{s,1,2}$ より算出する探索方向についても、絶対値の小さい dx_2 すなわち最適 dx の絶対値に近い dx_2 が得られる、という利点があることが明らかになってきた。

5.3 線形近似制約面最小化の構想

等式制約法での x_m 行選択規則を用いけば、非線形項を含むHesse行列 $G_{s,1,2}$ より良好な探索方向 dx_2 が得られるが、 E についての微分または差分が必要になってくる。線形項のみの $G_{s,1,1}$ より得られる dx_1 は dx_2 より探索方向は若干悪くなっているが、図-2より明らかなように dx_2 より絶対値が小さく(数値例では最適 dx 絶対値の約1/2)、step幅調整の面でも安定感がある。目的関数 f については2階微分で制約式 g については1階微分のHesse行列 $G_{s,1,1}$ より算出した dx_1 を取り上げ、 dx_1 を用いて探索する方法が計算量の面から有効と考えられる。

線形近似制約面から非線形制約面への最適解の移動量 dx_s^* を表す式(29)で dx_s^* 近似することができれば、 $dx_{s,2} = dx_{s,1} + dx_s^*$ より dx_1 を改良することが可能になる。図-2を参照して dx_1 が既知とすると $g_m = 0$ 上では点 $x_{1,2}$ になり、 $E_{1,2}$ 、 $\nabla f_{1,2}$ を求めると式(39)の最適

表-6 探索方向の比較(dx_r 指定、 dx_r 分担、 $-\nabla F$ 、 dx_1')

		dx_r 指定1	dx_r 指定2	dx_r 分担	$-\nabla F$	dx_1'
dx	x_1	-0.08017	-12.803	-8.8907	-0.86391	-0.07772
	x_2	-0.25635	-15.295	-7.5247	-0.53043	-0.15051
	x_3	-0.00691	-0.03731	1.1743	-0.10304	-0.00672
$ dx $		0.26868	19.946	11.707	1.0190	0.16953
正規化	x_1	-0.29840	-0.64189	-0.75946	-0.84782	-0.45847
	x_2	-0.95409	-0.76679	-0.64277	-0.52055	-0.88783
	x_3	0.02572	0.00187	0.10031	-0.10112	-0.03963
正規化 df		-1.4241	-1.5700	-2.1842	-1.0190	-1.1963
$f_m(\text{line})$		1.29×10^{-2}	7.81×10^{-3}	2.79×10^{-2}	5.28×10^{-2}	4.96×10^{-4}
$f(dx)$		-0.09233	1647.7	622.75	4.2131	-0.01488

解移動量の近似式が得られる。式(29)の E_2 を E_{12} で ∇_{mf12} を ∇_{mf12} で置き換えたものが式(39)であり、理論的裏付けはまだなされていないが、数値例では比較的良好な結果が得られている。式(40)、式(41)を用いて算出した $d\mathbf{x}_1'$ についての探索方向の比較を表-6に付記してある。正規化 $d\mathbf{x}_1'$ は正規化最適 $d\mathbf{x}$ に近づいており、 $d\mathbf{x}_{2 \times 3}$ とはほぼ同程度になっている。 $d\mathbf{x}_1$ の改良がなされたことは明らかであるが、これは最適解の移動量を表す式の応用例である。

$$G_{sL1} \cdot d\mathbf{x}_s^* = -(E_{12} - E_1) \nabla_{mf12} \quad \cdots (39)$$

$$d\mathbf{x}_s^* = -G_{sL1}^{-1}(E_{12} - E_1) \nabla_{mf12} \quad \cdots (40)$$

$$\left. \begin{aligned} d\mathbf{x}_{s1}' &= d\mathbf{x}_{s1} + d\mathbf{x}_s^* \\ d\mathbf{x}_{m1}' &= E_{12}^T d\mathbf{x}_{s1}' \end{aligned} \right\} \cdots \cdots (41)$$

線形近似制約面最小化には種々の方法が考えられるが、大別すると、①探索方向 $d\mathbf{x}_1$ のままで反復する、② $d\mathbf{x}_1$ を改良して反復する、になる。最適解の移動量 $d\mathbf{x}_s^*$ について良好な近似が得られるなら、制約式 g についての1階微分だけで2階微分を用いた場合に近い探索方向が得られることになる。現段階では $d\mathbf{x}_s^*$ の近似について十分な検討が行われていないが、改良の余地は残っていると考えられる。

6. まとめ

文献では線形制約面を基にして探索方向を扱っている場合が多く、非線形制約面での探索方向については明確ではない部分も存在しているので、考察を行ってみることにした。上記2.～5.の考察より下記(1)～(7)の結果を得た。

(1) ①異なった線形制約面、②非線形制約面と線形近似制約面、の二つのケースについて勾配差式から最適解の移動量を表す式を誘導した。

(2) 縮小勾配 ∇_{sL} を微分して求めたHesse行列 G_{sL} について、制約面の傾き E の変化に伴う非線形項と線形項との関係を明らかにした。

(3) 線形項のみの G_{sL1} より算出される $d\mathbf{x}_1$ は従属変数 $\mathbf{x}_m$ の選択に関係なく一定の値を持つが、非線形項を含む G_{sL2} より算出される $d\mathbf{x}_2$ は $\mathbf{x}_m$ の選択により異なった値を持つことを、等式制約法での変数交換を用いて説明した。

(4) 全変数での最急勾配 ∇F には E の逆行列が含まれ、微分でHesse行列 G_F を算出するのは困難であるが、差分でHesse行列を算出するときに $g_m = 0$ の満足をどのようにするかに問題があり、良好な G_F が得られにくい。

(5) 探索例を用いて下記の探索方向について数値的な比較を行った。①非線形項を含む $d\mathbf{x}_2$ は $\mathbf{x}_m$ の選択により絶対値が大きく異なってくるが、正規化 $d\mathbf{x}_2$ は正規化最適 $d\mathbf{x}$ に近い探索方向である。②線形項のみの $d\mathbf{x}_1$ は探索方向としては $d\mathbf{x}_2$ より若干悪くなるが、 $\mathbf{x}_m$ 選択に関係なく一定の値を持ち、絶対値が $d\mathbf{x}_2$ より小さく安定した探索方向である。③ G_F より算出した $d\mathbf{x}_F$ は ∇F の差分に問題があるため、計算量が多い割には良好な探索方向になっていない。④最急勾配方向 $-\nabla F$ はHesse行列を用いていないため、上記①～③に比べ探索方向としては劣っている。

(6) 上記(3)で述べたように $d\mathbf{x}_2$ では $\mathbf{x}_m$ 選択により絶対値が大きく動くが、 $d\mathbf{x}_2$ 絶対値を小さくするような $\mathbf{x}_m$ 選択と等式制約法での $\mathbf{x}_m$ 行選択規則とが関連していることを明らかにした。

(7) 線形近似制約面最小化の構想を出したが、まだ理論的裏付けがなされていない。最適解の移動量を近似することにより $d\mathbf{x}_1$ を改良することができれば、 g_m についての1階微分だけで2階微分を用いた探索方向に近いものが得られる。

参考文献

- 1) 平田恭久：線形制約付き2次関数の最適解，土木学会第52回年次学術講演会講演概要集，I-A157
- 2) 平田恭久：制約式交換に伴う行列内容，土木学会第51回年次学術講演会講演概要集，I-A466
- 3) 平田恭久：縮小勾配法の最小化法への適用，構造工学論文集，Vol. 38A, pp. 413～pp. 420, 1992年 3月
- 4) 平田恭久：制約付き最小化法への縮小勾配の適用：構造工学論文集，Vol. 39A, pp. 485～pp. 492, 1993年 3月
- 5) 平田恭久：活性な制約式集合を用いる最適化手法と等式制約法の比較，構造工学論文集，Vol. 40A, pp. 549～pp. 558, 1994年 3月