I-382　等式制約法での縮小勾配の改良

群馬高専　正員　平田恭久

１．はじめに

等式制約法では探索のときの最急勾配として*Lagrange*関数の探索変数についての微分を用いているが、これは縮小勾配であり、その値は活性な制約式とこれに対する制約変数と探索変数の区分に依存している。しかしながら、活性な制約式に対する制約変数の選択を常に固定しておくのは非常に困難なので、この点が探索方向の改良に共役方向法，準*Newton*法を適用するときの重大な障害になっている。このため、変数の区分に無関係な最急勾配が必要であり、ここではこの点の解決を試みた。

２．最急勾配∇Fの誘導

勾配射影法で探索方向を見つける問題にならって、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。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)についての*Lagrange*関数は式(3)になり、極値の条件は式(4)になる。式(4)第２行のdxを式(4)第１行に代入すると式(5)になり、式(5)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(6)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(6)よりdx_s，μ_mは式(7)，式(8)になるが、w＝１はスケールファクタである。式(7)の$\nabla_s L$は本来の最適化問題での*Lagrange*関数の探索変数x_s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(7)からdx_sが計算できる。

$$\min_{dx}\ \nabla f^T dx \quad \text{subject to} \quad A_s dx_s - dx = 0,\quad dx^T dx - 1 = 0 \qquad (1)$$

$$E = -\nabla_s g_m^T(\nabla_m g_m^T)^{-1},\quad dx_m = E^T dx_s,\quad dx_s = dx_s \qquad (2)$$

$$L(dx,\mu,w) = \nabla f^T dx + \mu^T(A_s dx_s - dx) + w(dx^T dx - 1) \qquad (3)$$

$$\begin{aligned}\partial L/\partial dx &= \nabla f^T + \mu^T(A_s - I) + 2w\,dx^T = 0\\ \partial L/\partial \mu &= A_s dx_s - dx = 0\\ \partial L/\partial w &= dx^T dx - 1 = 0\end{aligned} \qquad \cdots\cdots (4)$$

$$\nabla f + (A_s - I)^T\mu_m + 2wA_s dx_s = 0 \qquad (5)$$

$$\begin{aligned}&\nabla_m f - I_m\mu_m + 2wE^T dx_s = 0 \ :\ \text{m個}\\ &\nabla_s f + E\mu_m + 2wI_s dx_s = 0 \ :\ \text{s個}\end{aligned} \qquad (6)$$

$$\begin{aligned}dx_s &= -(EE^T + I_s)^{-1}(\nabla_s f + E\nabla_m f)/2w\\ &= -(EE^T + I_s)^{-1}\nabla_s L/2w\end{aligned} \qquad \cdots (7)$$

$$\mu_m = (EE^T + I_s)^{-1}(\nabla_m f - E^T\nabla_s f) \qquad \cdots (8)$$

３．∇Fと∇Lの関係

式(7)のdx_sは探索方向dxのｓ成分なので、最急勾配のｓ成分$\nabla_s F$は式(9)第１行のようにdx_sの符号を変えたものになる。最急勾配のm成分$\nabla_m F$は式(9)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。$\nabla_s L$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(10)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∇Lは制約面上の条件を満足している。活性な制約面上を探索するときのfの微小変化dfは式(11)になる。最急降下法として式(11)のdx_sに$-\nabla_s L$を用いると式(12)になり、$-\nabla_s F$を用いると式(13)になる。$\nabla_s L^T\nabla_s L \geqq 0$，$\nabla F^T\nabla F \geqq 0$なので、$df(\nabla_s L) \leqq 0$，$df(\nabla_s F) \leqq 0$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(14)，式(15)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(

$$\nabla_s F = -dx_s,\quad \nabla_m F = E^T\nabla_s F \qquad (9)$$

$$\nabla_m L = E^T\nabla_s L \qquad (10)$$

$$\begin{aligned}df &= \nabla f\,dx\\ &= \nabla_s f^T dx_s + \nabla_m f^T dx_m\\ &= \nabla_s L^T dx_s\end{aligned} \qquad \cdots (11)$$

$$df(\nabla_s L) = -\nabla_s L^T\nabla_s L \qquad (12)$$

$$\begin{aligned}df(\nabla_s F) &= -\nabla_s L^T\nabla_s F\\ &= -2w\nabla F^T\nabla F\end{aligned} \qquad (13)$$

$$\text{正規化}\,df(\nabla_s L) = -\nabla_s L^T\nabla_s L/\sqrt{\nabla L^T\nabla L} \qquad (14)$$

すなわち同じdf)が得られる。dfについて式(16)が成立することは数値的に確かめられる。

$$\text{正規化}df(\nabla_s F) = -2w\nabla F^T\nabla F / \sqrt{\nabla F^T\nabla F} \quad (15)$$

$$\text{正規化}df(\nabla_s F) \leq \text{正規化}df(\nabla_s L) \quad \cdots (16)$$

４．▽Ｆ，▽Ｌの直交性

$$dx^T\nabla F = dx_s^T(EE^T + I_s)\nabla_s F = dx_s^T\nabla_s L / 2w = 0 \quad (17)$$

$$dx^T\nabla L = dx_s^T(EE^T + I_s)\nabla_s L \quad (18)$$

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事実を利用している。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はラインサーチ探索方向(ｎ次元)，▽Ｆはラインサーチ終了点での最急勾配(ｎ次元)とする。dxは制約面上にあり、▽Ｆ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り、dxと▽Ｆの内積は式(17)になる。これに対し、dxと▽Ｌの内積は式(18)になり、一般に$dx^T\nabla L \neq 0$で直交性は成立しない。$\nabla_s L$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のｓ成分であるdx_sとは直交している。これに対し、式(17)より、一般に$dx_s^T\nabla_s F \neq 0$であり、dx_sと$\nabla_s F$との直交性は成立しない。▽Ｆ，$\nabla_s L$は直交性が確かめられたので、いずれも共役方向法，準Newton法の適用のとき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。$\nabla_s L$のような縮小勾配には制約変数と探索変数の区分が固定していないと、共役方向法，準Newton法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非常に困難である。この点が共役方向法，準Newton法適用での最大のネックになっていたが、▽Ｆを最急勾配に採用することによりこの問題点が解決できた。

５．共役方向法，準Ｎｅｗｔｏｎ法の適用

$$dx_k(\nabla F) = -\nabla F_k + b_k(\nabla F)dx_{k-1} \quad (19)$$

$$dx_k(\nabla_s L) = -\nabla L_k + b_k(\nabla_s L)dx_{k-1} \quad (20)$$

$$b_k(\nabla F) = \nabla F_k^T\nabla F_k / (\nabla F_{k-1}^T\nabla F_{k-1}) = \frac{\nabla_s F_k^T(E_kE_k^T + I_s)\ \nabla_s F_k}{\nabla_s F_{k-1}^T(E_{k-1}E_{k-1}^T + I_s)\ \nabla_s F_{k-1}} \quad (21)$$

$$b_k(\nabla_s L) = \nabla_s L_k^T\nabla_s L_k / (\nabla_s L_{k-1}^T\nabla_s L_{k-1}) = \frac{\nabla_s F_k^T(E_kE_k^T + I_s)^2\nabla_s F_k}{\nabla_s F_{k-1}^T(E_{k-1}E_{k-1}^T + I_s)^2\nabla_s F_{k-1}} \quad (22)$$

探索方向の改良に共役方向法，準Newton法を適用する場合を共役方向法を例にして説明する。最急勾配としては直交性が確かめられた▽Ｆ，$\nabla_s L$を用いる。共役方向法でのｋ点の探索方向dx_kは▽Ｆを用いると式(19)，$\nabla_s L$を用いると式(20)になる。式(19)の$b_k(\nabla F)$は式(21)，式(21)の$b_k(\nabla_s L)$は式(22)になる。式(19)，式(20)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(21)，式(22)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。ｎ＝３，ｓ＝２でfが２次式，制約条件が線形の簡単な例題について、共役方向法，準Newton法(ＤＦＰ公式)を適用した結果を述べる。ＤＦＰ公式についても共役方向法と同様に最急勾配法に▽Ｆ，$\nabla_s L$を用いるか、Hesse行列の逆行列の近似行列H_kの次元は▽Ｆの場合はｎ次元，$\nabla_s L$の場合はｓ次元になる。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．まとめ

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活性な制約面上のfについての最急勾配▽Ｆを採用することによりこの問題点を解決できた。また、▽Ｆを最急降下法に用いると正規化df最小なので、最も効率の良い探索方向になる。