

目次

第 3 章 数理計画法の基礎	3
3.1 問題設定	3
3.2 反復法	4
3.3 勾配法	5
3.4 ステップサイズの規準	12
3.5 Newton 法	21
3.6 拡大関数法	27
3.7 制約つき問題に対する勾配法	28
3.7.1 簡単なアルゴリズム	32
3.7.2 複雑なアルゴリズム	37
3.8 制約つき問題に対する Newton 法	43
3.8.1 簡単なアルゴリズム	46
3.8.2 複雑なアルゴリズム	53
3.9 第 3 章のまとめ	53
3.10 第 3 章の演習問題	54

第3章 数理計画法の基礎

第2章では有限次元ベクトル空間上の最適化問題について極小点であれば成り立つ条件 (極小点の必要条件) と最小点であることを保証する条件 (最小点の十分条件) をみてきた。そこでは極小点をみつける方法 (解法) については詳しく説明されなかった。本章では、解法に焦点をあてたい。最適化問題の解法は、**数理計画法**とよばれ、**オペレーションズ・リサーチ** (OR) とよばれる学問分野で盛んに研究が行われている。ここでは、最適化問題の解法に関する考え方や理論的に得られる結果を示しながら、アルゴリズムについて考えてみたい。ここで示される内容の多くは第7章の抽象的最適設計問題においても有効である。第7章では、関数空間を舞台にして同じアルゴリズムが成り立つことをみることになる。

3.1 問題設定

最適設計問題は、第1章でみてきたように、設計変数 $\xi \in \Xi$ と状態変数 $u \in U$ によって定義された評価関数 $f_0(\xi, u), \dots, f_m(\xi, u)$ に対する等式制約 (状態方程式) と不等式制約つき最適化問題であるとみなされた。第2章では、そのような問題は $x = (\xi, u) \in \Xi \times U$ を設計変数とする $f_0(\xi, u), \dots, f_m(\xi, u)$ で構成された最適化問題とみなされた。

本章では、第1章の定義に戻って、 ξ を設計変数とよび、 $f_0(\xi, u(\xi)), \dots, f_m(\xi, u(\xi))$ を $\tilde{f}_0(\xi), \dots, \tilde{f}_m(\xi)$ とみなすことにする。 ξ に対する $\tilde{f}_0(\xi), \dots, \tilde{f}_m(\xi)$ の微分は、2.8節でみてきたように、随伴変数法で求められる。さらに、 $\tilde{f}_0(\xi), \dots, \tilde{f}_m(\xi)$ は非線形関数であると仮定する。実際、第1章の最適設計問題 (問題 1.1.4) では、 $f_0(u)$ が u に対する線形関数であっても、等式制約関数 $h(a, u) = -K(a)u + p$ は (a, u) に対して非線形となるために、 $\tilde{f}_0(a)$ は非線形関数となった。

そこで、本章では、設計変数 $\xi \in \Xi$ を $x \in X = \mathbb{R}^d$ とかき、非線形関数 $\tilde{f}_0, \dots, \tilde{f}_m$ を f_0, \dots, f_m とかき、それらの x に対する勾配を g_0, \dots, g_m とかくことにして、等式制約を含まない次の問題を考えることにしよう。

問題 3.1.1 (非線形最適化問題) $X = \mathbb{R}^d$ とする. $f_0, \dots, f_m \in C^1(X; \mathbb{R})$ に対して,

$$\min_{x \in X} \{f_0(x) \mid f_1(x) \leq 0, \dots, f_m(x) \leq 0\}$$

を満たす x を求めよ. □

本章の構成は次のようである. 3.2 節で非線形最適化問題の解法に関する基本的な考え方である反復法の定義と収束性に関する定義を示す. その後, 3.3 節から 3.5 節までにおいて, 制約なし最適化問題に対する解法をみていくことにする. 不等式制約つき最適化問題 (問題 3.1.1) に対する解法は, 3.6 節以降においてみていくことにする.

3.2 反復法

非線形最適化問題の解法を考えたとき, 特別な前処理をせずに連立 1 次方程式を 1 回解くだけで最適解が得られるような方法はないと考えられる. 通常は, 次に見える反復法が基本となる.

定義 3.2.1 (反復法) 問題 3.1.1 に対して, 最小点ではない $x_0 \in X$ を選び, $y_g \in X$ を決めながら, $k \in \{0, 1, 2, \dots\}$ に対して,

$$x_{k+1} = x_k + y_g = x_k + \bar{\epsilon}_g \bar{y}_g \tag{3.2.1}$$

を求めていく方法を反復法という. ここで, y_g を探索ベクトル, その大きさ $\|y_g\|_X$ をステップサイズという. それに対して, \bar{y}_g は方向のみを与えるベクトルで, 本書では, 探索方向とよんで, 探索ベクトルとは区別することにする. \bar{y}_g の大きさは 1 である必要はないと仮定する. $\bar{\epsilon}_g$ はその大きさを調整するための正の定数とする. また, x_0 を初期点, $k \in \{0, 1, 2, \dots\}$ に対して x_k を試行点という. □

この定義から, 反復法を使ったアルゴリズムを考える場合は, 探索方向 \bar{y}_g を求める方法とステップサイズ $\|y_g\|_X$ を適切に決める方法を具体的に示す必要がある. 3.3 節以降で, それらの方法についてみていくことにする. なお, このような反復法とは別に, 直接法とよばれる最適化問題の数値解法が知られている. 直接法は, ある決められた有限回の手順で解を求めることができる方法の総称として用いられる. 直接法は, 主に線形最適化問題に対して想定される方法であることから, 本書では省略することにする.

のちの議論のために, 反復法の特徴や性能を表す用語を定義しておこう.

定義 3.2.2 (大域的収束性) 初期点が最小点から十分離れていても、最小点に収束する性質があるとき、反復法は**大域的収束性**をもつという。□

定義 3.2.3 (収束率) \mathbf{x} を極小点、 $\{\mathbf{x}_k\}_{k \in \mathbb{N}}$ を反復法でえられた点列とする。このとき、ある k_0 がとれて、任意の $k \geq k_0$ に対して、

$$\|\mathbf{x}_{k+1} - \mathbf{x}\|_X \leq r \|\mathbf{x}_k - \mathbf{x}\|_X^p$$

が成り立つような $p \in [1, \infty)$ がとれるとき、 p をそのアルゴリズムの**収束次数**という。ただし、 $p = 1$ のとき $r \in (0, 1)$ として、 $p > 1$ のとき r は正の定数とする。また、 r が 0 に収束する数列 $\{r_k\}_{k \in \mathbb{N}}$ に置き換えられるとき、**超 p 次収束**という。□

3.3 勾配法

最初に、探索方向 $\bar{\mathbf{y}}_g$ をみつける方法の一つとして、勾配法についてみてみよう。ここでは、 $i \in \{0, 1, \dots, m\}$ の中から一つの評価関数 f_i を選んで、それが降下する方向 $\bar{\mathbf{y}}_{gi}$ を求めることを考えよう。このような $\bar{\mathbf{y}}_{gi}$ を f_i の**降下方向**とよぶことにする。

問題 3.1.1 において、最小点が許容集合の内点にある (すべての不等式制約が無効となる) ならば、 f_0 の降下方向 $\bar{\mathbf{y}}_{g0}$ は式 (3.2.1) の探索方向 $\bar{\mathbf{y}}_g$ となる。また、いずれかの不等式制約条件が有効になった場合の探索方向 $\bar{\mathbf{y}}_g$ を求める際にも、3.7 節以降で示されるように、目的関数 f_0 と有効な制約関数 f_i の降下方向 $\bar{\mathbf{y}}_{g0}$ と $\bar{\mathbf{y}}_{gi}$ を使って、不等式制約を満たす探索方向 $\bar{\mathbf{y}}_g$ が求められる。そのときにも勾配法が使われる。

本書では、3.3 節から 3.5 節までは、制約なしの問題を考える。そこで、簡単のために、 f_i, \mathbf{g}_i および $\bar{\mathbf{y}}_{gi}$ をそれぞれ f, \mathbf{g} および $\bar{\mathbf{y}}_g$ のようにかくことにする。

まず、図 3.1 と 図 3.2 をみながら記号を定義しよう。 $k \in \{0, 1, 2, \dots\}$ に対して $\mathbf{x}_k \in X$ を試行点、 \mathbf{g} を \mathbf{x}_k における f の勾配で、 $\mathbf{0}_X$ ではないとする。次に、 \mathbf{g} を既知として、 f が減少する方向 $\bar{\mathbf{y}}_g \in X$ を求めることを考えよう。

ここで、 \mathbf{g} の意味について確認しておこう。 f を \mathbf{x}_k の周りで Taylor 展開すれば、

$$f(\mathbf{x}_k + \mathbf{y}) = f(\mathbf{x}_k) + \mathbf{g} \cdot \mathbf{y} + o(\|\mathbf{y}\|_X) \quad (3.3.1)$$

となる。ここで、第 4 章で示される Fréchet 微分の定義 (定義 4.5.4) を先取りすれば、 \mathbf{g} は X の双対空間 X' (定義 4.4.5) の要素で、 \mathbf{g} の大きさ (ノルム) は

$$\|\mathbf{g}\|_{X'} = \max_{\mathbf{y} \in X} \frac{|\langle \mathbf{g}, \mathbf{y} \rangle_{X' \times X}|}{\|\mathbf{y}\|_X} = \max_{\mathbf{y} \in X, \|\mathbf{y}\|_X = 1} |\langle \mathbf{g}, \mathbf{y} \rangle_{X' \times X}|$$

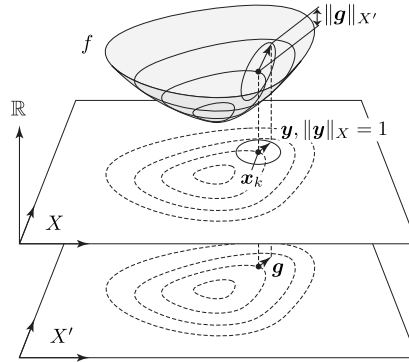
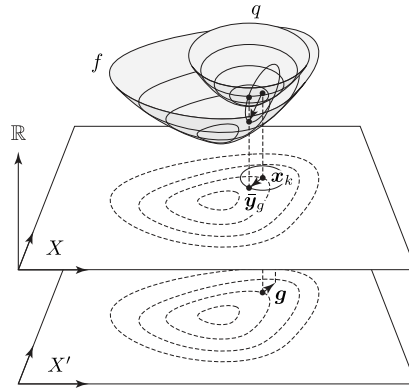
図 3.1: 勾配 g の定義

図 3.2: 勾配法

で定義される． $X = \mathbb{R}^d$ の場合は $X' = \mathbb{R}^d$ および双対積 $\langle g, y \rangle_{X' \times X} = g \cdot y$ となる．この定義によれば， $\|g\|_{X'}$ は， $\|y\|_X = 1$ なる任意の $y \in X$ に対する $|g \cdot y|$ の最大値を意味することになる．また， g の方向は f の等高線に対して垂直な方向を向いている．なぜならば，式 (3.3.1) において $x_k + y \in X$ を等高線上の点として， y を十分小さなベクトルにとったとき，

$$g \cdot y \approx f(x_k + y) - f(x_k) = 0$$

が成り立つためである．図 3.1 はこの関係を示している．

これらの関係より， g は f がもっとも増加する方向を向いていることになる．そこで，有限次元ベクトル空間では $X = X'$ が成り立つ (4.4.6 項) ことから，

$$\bar{y}_g = -g \tag{3.3.2}$$

のように $\bar{\mathbf{y}}_g \in X$ を選べば,

$$f(\mathbf{x}_k + \bar{\mathbf{y}}_g) - f(\mathbf{x}_k) = -\|\bar{\mathbf{y}}_g\|_X^2 + o(\|\bar{\mathbf{y}}_g\|_X)$$

となる. ここで, $\|\bar{\mathbf{y}}_g\|_X$ が十分小さければ, f が減少することになる.

この方法を一般化しよう. 降下方向を次の問題の解 $\bar{\mathbf{y}}_g \in X$ として求める方法を **勾配法** という.

問題 3.3.1 (勾配法) $X = \mathbb{R}^d$ とする. $\mathbf{A} \in \mathbb{R}^{d \times d}$ を正定値実対称行列 (定義 2.4.5) とする. $f \in C^1(X; \mathbb{R})$ に対して, 極小点ではない $\mathbf{x}_k \in X$ における f の勾配を $\mathbf{g}(\mathbf{x}_k) \in X' = \mathbb{R}^d$ とする. このとき, 任意の $\mathbf{y} \in X$ に対して

$$\bar{\mathbf{y}}_g \cdot (\mathbf{A}\mathbf{y}) = -\mathbf{g}(\mathbf{x}_k) \cdot \mathbf{y} \quad (3.3.3)$$

を満たす $\bar{\mathbf{y}}_g \in X$ を求めよ. \square

式 (3.3.3) は, 任意の $\mathbf{y} \in X$ との内積を用いた表現になっている. この方程式は,

$$\bar{\mathbf{y}}_g = -\mathbf{A}^{-1}\mathbf{g} \quad (3.3.4)$$

によって $\bar{\mathbf{y}}_g$ を求めることと同値である. わざわざ内積を用いたのは, 第 7 章で関数空間での勾配法を定義するときに, それが問題 3.3.1 の自然な拡張になるようにするためである. また, 式 (3.3.2) は \mathbf{A} に単位行列 \mathbf{I} を使った場合の勾配法であったことになる. 問題 3.3.1 の解 $\bar{\mathbf{y}}_g$ が f を減少させることを次の定理で確認しておこう.

定理 3.3.2 (勾配法) 問題 3.3.1 の解 $\bar{\mathbf{y}}_g$ は, f の \mathbf{x}_k における降下方向である. \square

証明 \mathbf{A} は正定値実対称行列なので, 任意の $\mathbf{y} \in X$ に対して

$$\mathbf{y} \cdot (\mathbf{A}\mathbf{y}) \geq \alpha \|\mathbf{y}\|_X^2, \quad \mathbf{A} = \mathbf{A}^\top$$

を満たす正定値 α が存在する. この関係と式 (3.3.3) を用いれば, 正の定数 $\bar{\epsilon}$ に対して,

$$\begin{aligned} f(\mathbf{x}_k + \bar{\epsilon}\bar{\mathbf{y}}_g) - f(\mathbf{x}_k) &= \bar{\epsilon}\mathbf{g} \cdot \bar{\mathbf{y}}_g + o(\bar{\epsilon}) = -\bar{\epsilon}\bar{\mathbf{y}}_g \cdot (\mathbf{A}\bar{\mathbf{y}}_g) + o(\bar{\epsilon}) \\ &\leq -\bar{\epsilon}\alpha \|\bar{\mathbf{y}}_g\|_X^2 + o(\bar{\epsilon}) \end{aligned}$$

が成り立つ. ここで, $\bar{\epsilon}$ が十分小さければ, f は減少することになる. \square

ここで, 降下方向の **降下角** を次のように定義しよう.

定義 3.3.3 (降下角) $\mathbf{x}_k \in X$ において, $\mathbf{g} \in X'$ を勾配, $\bar{\mathbf{y}}_g \in X$ を降下方向とする. このとき,

$$\cos \theta = -\frac{\langle \mathbf{g}, \bar{\mathbf{y}}_g \rangle_{X' \times X}}{\|\mathbf{g}\|_{X'} \|\bar{\mathbf{y}}_g\|_X}$$

で定義された $\theta \in [0, \pi]$ を \mathbf{x}_k における $\bar{\mathbf{y}}_g$ の降下角という. \square

勾配法 (問題 3.3.1) において, A を単位行列 I とおいたとき, \bar{y}_g の降下角 θ はゼロとなる. そこで, このときの反復法は**最急降下法**とよばれる. しかし, 最急降下法がいつも収束が速いとは限らない. そのことは, あとで示される共役勾配法 (問題 3.4.10) との比較で明らかになる.

さらに, 勾配法で得られる \bar{y}_g の幾何学的意味について考えてみよう. 問題 3.3.1 は,

$$q(\bar{y}_g) = \min_{y \in X} \left\{ q(y) = \frac{1}{2} y \cdot (Ay) + g \cdot y + f(x_k) \right\} \quad (3.3.5)$$

を満たす $\bar{y}_g \in X$ を求めることと同値である. 実際, 任意の $y \in X$ に対して $q'(\bar{y}_g)[y] = 0$ が成り立つことは, 式 (3.3.3) が成り立つことと同値である. 図 3.2 にそのときの関数 q を示す. ここで, q は楕円放物面で, その最小点が $x_k + \bar{y}_g$ になっている. \bar{y}_g の大きさは A の選び方に依存する. したがって, ステップサイズ $\|y_g\|_X = \|\bar{y}_g\|_X$ が ϵ_g になるようにしたければ, 次のような計算をおこなえばよいことになる. 正の定数 c_a を調整パラメータとして導入して, 式 (3.3.4) を

$$y_g = -(c_a A)^{-1} g \quad (3.3.6)$$

に変更する. ここで, c_a を大きくすれば y_g の大きさは小さくなることに注意する. そこで, ステップサイズ ϵ_g と初期点 x_0 における勾配法 (問題 3.3.1) の解 $\bar{y}_g = \bar{y}_g(x_0)$ が与えられたとき,

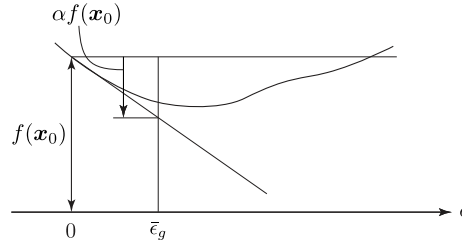
$$c_a = \frac{\|\bar{y}_g\|_X}{\epsilon_g} \quad (3.3.7)$$

とおくことにする. 最初のステップ ($k = 0$) において, c_a をこのように定めて, $k \in \mathbb{N}$ に対してもこの c_a を用いて式 (3.3.6) で探索ベクトルを求めていけば, ステップサイズ $\|y_{gk}\|_X$ は, しばらくは, およそ ϵ_g の大きさをもつことになる. また, c_a を最初のステップで求めた値に固定することによって, 収束に近づいたときにステップサイズは 0 に近づくことが期待される. このように c_a を導入したときには, 式 (3.3.5) は

$$q(y_g) = \min_{y \in X} \left\{ q(y) = \frac{1}{2} y \cdot (c_a A y) + g \cdot y + f(x_k) \right\} \quad (3.3.8)$$

に変更される.

c_a の決め方に関して, ステップサイズのイメージが明確で, ϵ_g として具体的な値が思い浮かべることができる場合には, 式 (3.3.7) を用いることができる. 例えば, 第9章で扱うような領域変動を考える場合には, 領域変動の大きさ (ステップサイズ) は領域変動によるひずみのノルム (ひずみを2乗して積分した値の平方根, あ

(追加) 目的関数の減少率 α

るいは最大ひずみ)などで定義される．そこで， ϵ_g にはひずみのノルムのイメージで，例えば 0.05 などと，値を設定することができる．しかしながら，そのようなイメージが得られない場合には， c_a を適切に決めるために別の方法を考えなければならない．

その方法の一つは，期待する目的関数の減少率を与えて， c_a を決める方法である¹． $k=0$ のときの評価関数 $f(\mathbf{x}_0)$ と勾配 $\mathbf{g}(\mathbf{x}_0)$ は与えられ，勾配法によって探索ベクトル $\bar{\mathbf{y}}_g$ も与えられたと仮定する．その上で，目的関数の減少率 $\alpha \in (0, 1)$ を与えて，

$$f(\mathbf{x}_0 + \bar{\epsilon}_g \bar{\mathbf{y}}_g) - f(\mathbf{x}_0) \approx \alpha f(\mathbf{x}_0) \approx \bar{\epsilon}_g \mathbf{g}(\mathbf{x}_0) \cdot \bar{\mathbf{y}}_g$$

を予測する．このとき， \approx を $=$ とみなせば， c_a は

$$c_a = \frac{1}{\bar{\epsilon}_g} = \frac{\mathbf{g}(\mathbf{x}_0) \cdot \bar{\mathbf{y}}_g}{\alpha f(\mathbf{x}_0)} \quad (3.3.9)$$

によって与えられる．

これまでの考察に基づいて，勾配法を使った簡単なアルゴリズムを示すことにしよう．なお，本章でアルゴリズムを記述するときには，最適設計問題を解くことを想定して，次のような表記をすることにする．「 $f(\mathbf{x}_k)$ を計算する」とかくところを「状態決定問題を解いて， $f(\mathbf{x}_k)$ を計算する」とかく．また，「 $\mathbf{g}(\mathbf{x}_k)$ を計算する」とかくところを「 f に対する随伴問題を解いて， $\mathbf{g}(\mathbf{x}_k)$ を計算する」とかく．その理由は，3.1節の冒頭で説明されたように，最適設計問題ではそのような計算になるためである．

これらの背景をふまえて，勾配法を使ったアルゴリズムの例を次に示す．最初にもっとも簡単な例を示そう．ステップサイズを調整するパラメータ c_a が，直接，決め打ちで与えられると考える．図 3.3 はその概要を示す．

¹Julius Fergy T. Rabago (私信)

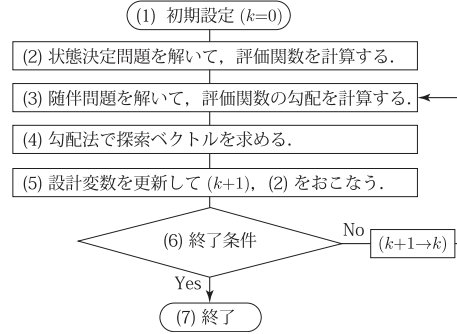


図 3.3: 勾配法のアルゴリズム

アルゴリズム 3.3.4 (勾配法) 問題 3.1.1 において f_0 を f とかき、すべての不等式制約は無効とする。

- (1) 初期点 \mathbf{x}_0 , 正定値対称行列 \mathbf{A} (特に指定がなければ \mathbf{I}), ステップサイズを調整する正の定数 c_a および収束判定に用いる正の定数 ϵ_0 を与える. $k = 0$ とおく.
- (2) 状態決定問題を解いて, $f(\mathbf{x}_k)$ を計算する.
- (3) f に対する随伴問題を解いて, $\mathbf{g}(\mathbf{x}_k)$ を計算する.
- (4) 式 (3.3.6) で \mathbf{y}_g を計算する.
- (5) $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{y}_g$ とおき, 状態決定問題を解いて, $f(\mathbf{x}_{k+1})$ を計算する.
- (6) 終了条件 $|f(\mathbf{x}_{k+1}) - f(\mathbf{x}_k)| \leq \epsilon_0$ を判定する.
 - 終了条件が満たされたとき, (7) に進む.
 - そうではないとき, $k+1$ を k に代入し, (3) に戻る.
- (7) 計算を終了する.

□

アルゴリズム 3.3.4 では, 終了条件に $|f(\mathbf{x}_{k+1}) - f(\mathbf{x}_k)| \leq \epsilon_0$ が使われた. これ以外に, 設計変数の変動量に注意がむけられた $\|\mathbf{y}_g\|_X \leq \epsilon_0$ や k が上限値を超えない条件などが考えられる.

最初のステップサイズが指定された ϵ_g になるように c_a を決めたい (あるいは目的関数の減少率 α を用いる) 場合には, 次のようなアルゴリズムが考えられる. 図 3.4 はその概要を示す.

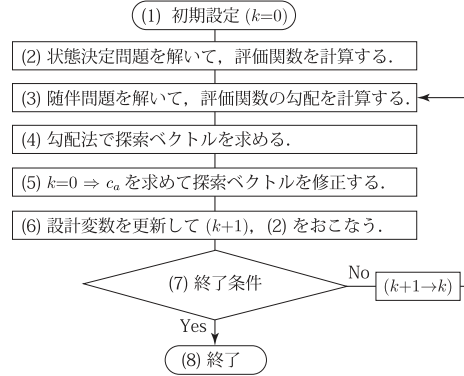


図 3.4: ステップサイズの初期値が与えられた勾配法のアルゴリズム

アルゴリズム 3.3.5 (初期ステップサイズ指定の勾配法) 問題 3.1.1 において f_0 を f とかき、すべての不等式制約は無効とする。

- (1) 初期点 \mathbf{x}_0 , 正定値対称行列 \mathbf{A} , 初期のステップサイズを与える正の定数 ϵ_g (あるいは目的関数の減少率 α) および収束判定に用いる正の定数 ϵ_0 を与える。
 $c_a = 1$ および $k = 0$ とおく。
- (2) 状態決定問題を解いて、 $f(\mathbf{x}_k)$ を計算する。
- (3) f に対する随伴問題を解いて、 $\mathbf{g}(\mathbf{x}_k)$ を計算する。
- (4) 式 (3.3.6) で \mathbf{y}_g を計算する。
- (5) $k = 0$ のとき、 $\mathbf{y}_g = \bar{\mathbf{y}}_g$ とおき、式 (3.3.7) (あるいは式 (3.3.9)) で c_a を求める。また、 $\bar{\mathbf{y}}_g/c_a$ を \mathbf{y}_g に代入する。
- (6) $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{y}_g$ とおき、状態決定問題を解いて、 $f(\mathbf{x}_{k+1})$ を計算する。
- (7) 終了条件 $|f(\mathbf{x}_{k+1}) - f(\mathbf{x}_k)| \leq \epsilon_0$ を判定する。
 - 終了条件が満たされたとき、(8) に進む。
 - そうではないとき、 $k+1$ を k に代入し、(3) に戻る。
- (8) 計算を終了する。

□

3.4 ステップサイズの規準

次に、ステップサイズ $\|\mathbf{y}_g\|_X$ を適切に決める方法について考えてみよう。

探索方向 $\bar{\mathbf{y}}_g$ が既知であれば、最適化問題の変数は式 (3.2.1) の $\bar{\epsilon}_g$ だけとなる。そこで、 $\bar{\epsilon}_g$ を設計変数においた次のような最適化問題を考えて、その解でステップサイズ $\|\bar{\epsilon}_g \bar{\mathbf{y}}_g\|_X$ を決定する方法が考えられる。この方法は厳密直線探索法とよばれる。

問題 3.4.1 (厳密直線探索法) $X = \mathbb{R}^d$ とする。 $f \in C^1(X; \mathbb{R})$, $\mathbf{x}_k \in X$ と $\bar{\mathbf{y}}_g \in X$ が与えられたとき、

$$\min_{\bar{\epsilon}_g \in (0, \infty)} f(\mathbf{x}_k + \bar{\epsilon}_g \bar{\mathbf{y}}_g)$$

を満たす $\bar{\epsilon}_g$ を求めよ。 □

問題 3.4.1 を解くアルゴリズムには、非線形方程式を解くためのアルゴリズムが使われる。たとえば次のような方法が考えられる。

- 二分法
- セカント法

二分法は、 f が減少から増加に転じるような $\bar{\epsilon}_g$ の区間をみつめて、その区間の中間点を求める操作を繰り返す。そのときの $\bar{\epsilon}_g$ の厳密解への収束次数は 1 となる。セカント法を用いる場合には、 f の $\bar{\epsilon}_g$ に対する勾配がゼロとなるような $\bar{\epsilon}_g$ を求める問題とみなす。その上で、あとで示される Newton-Raphson 法において f の勾配を差分におきかえたときの更新式を用いて極小点が求められる (演習問題 3.1 と 3.2)。この方法の収束次数は黄金比 $(1 + \sqrt{5})/2$ であることが知られている。

厳密直線探索法によってえられた \mathbf{x}_k の厳密解への収束性については、次のような結果が得られる。評価関数を 2 次関数と仮定する。探索方向は最急降下法によってえられていると仮定する。このとき、厳密直線探索法によって得られるステップサイズは次の問題の解となる。

問題 3.4.2 (2 次最適化問題の厳密直線探索法) $X = \mathbb{R}^d$ とする。 $B \in \mathbb{R}^{d \times d}$ を正定値実対称行列、 $\mathbf{b} \in X$ を既知ベクトルとして、評価関数を

$$f(\mathbf{x}) = \frac{1}{2} \mathbf{x} \cdot (B\mathbf{x}) + \mathbf{b} \cdot \mathbf{x} \tag{3.4.1}$$

とおく。このとき、 $\mathbf{x}_k \in X$ に対して最急降下法により $\bar{\mathbf{y}}_g \in X$ を求め、問題 3.4.1 を満たす $\bar{\epsilon}_g \in (0, \infty)$ を求めよ。 □

解答 $f(\mathbf{x}_k + \bar{\epsilon}_g \bar{\mathbf{y}}_g)$ を $\bar{f}(\bar{\epsilon}_g)$ とかけば,

$$\begin{aligned}\bar{f}(\bar{\epsilon}_g) &= \frac{1}{2} (\mathbf{x}_k + \bar{\epsilon}_g \bar{\mathbf{y}}_g) \cdot \{ \mathbf{B} (\mathbf{x}_k + \bar{\epsilon}_g \bar{\mathbf{y}}_g) \} + \mathbf{b} \cdot (\mathbf{x}_k + \bar{\epsilon}_g \bar{\mathbf{y}}_g) \\ &= \bar{\epsilon}_g^2 \frac{1}{2} \bar{\mathbf{y}}_g \cdot (\mathbf{B} \bar{\mathbf{y}}_g) + \bar{\epsilon}_g \bar{\mathbf{y}}_g \cdot \mathbf{g} + f(\mathbf{x}_k)\end{aligned}$$

とかける. ここで, $\mathbf{g} = \mathbf{B}\mathbf{x}_k + \mathbf{b}$ が 2 番目の等号で使われた. 厳密直線探索法では,

$$\frac{d\bar{f}}{d\bar{\epsilon}_g} = \bar{\epsilon}_g \bar{\mathbf{y}}_g \cdot (\mathbf{B} \bar{\mathbf{y}}_g) + \bar{\mathbf{y}}_g \cdot \mathbf{g} = 0$$

より,

$$\bar{\epsilon}_g = - \frac{\bar{\mathbf{y}}_g \cdot \mathbf{g}}{\bar{\mathbf{y}}_g \cdot (\mathbf{B} \bar{\mathbf{y}}_g)}$$

が得られる. さらに, $\bar{\mathbf{y}}_g$ が最急降下法の解ならば, $\bar{\mathbf{y}}_g = -\mathbf{g}$ より

$$\bar{\epsilon}_g = - \frac{\bar{\mathbf{y}}_g \cdot \mathbf{g}}{\bar{\mathbf{y}}_g \cdot (\mathbf{B} \bar{\mathbf{y}}_g)} = \frac{\mathbf{g} \cdot \mathbf{g}}{\mathbf{g} \cdot (\mathbf{B} \mathbf{g})} = \frac{\mathbf{g} \cdot \mathbf{g}}{\bar{\mathbf{y}}_g \cdot (\mathbf{B} \bar{\mathbf{y}}_g)} \quad (3.4.2)$$

が成り立つ. \square

このように厳密直線探索法により $\bar{\epsilon}_g$ を求めながら反復法を繰り返したときの収束性について, 次の結果が得られる.

定理 3.4.3 (厳密直線探索法を用いた最急降下法の収束性) 問題 3.4.2 の解 $\bar{\mathbf{y}}_g \in X$ と $\bar{\epsilon}_g$ を用いた反復法によって生成される点列 $\{\mathbf{x}_k\}_{k \in \mathbb{N}}$ は

$$\|\mathbf{x}_{k+1} - \mathbf{x}\|_B \leq \left| \frac{\lambda_d - \lambda_1}{\lambda_1 + \lambda_d} \right| \|\mathbf{x}_k - \mathbf{x}\|_B$$

を満たす. ただし, \mathbf{x} は極小点, λ_1, λ_d は \mathbf{B} の最小および最大の固有値, $\|\mathbf{x}\|_B = \sqrt{\mathbf{x} \cdot (\mathbf{B} \mathbf{x})}$ とする. \square

証明 問題 3.4.2 の目的関数は

$$f(\mathbf{x}) = \frac{1}{2} (\mathbf{x} + \mathbf{B}^{-1} \mathbf{b}) \cdot \{ \mathbf{B} (\mathbf{x} + \mathbf{B}^{-1} \mathbf{b}) \} - \frac{1}{2} \mathbf{b} \cdot (\mathbf{B}^{-1} \mathbf{b})$$

とかける. ここで, $\mathbf{x} + \mathbf{B}^{-1} \mathbf{b}$ を \mathbf{x} におきかえても $\mathbf{x}_{k+1} - \mathbf{x}$ の評価は変わらない. また, 上式の右辺第 2 項は \mathbf{x} の関数ではないので, それを省略しても $\mathbf{x}_{k+1} - \mathbf{x}$ の評価は変わらない. そこで,

$$\bar{f}(\mathbf{x}) = \frac{1}{2} \mathbf{x} \cdot (\mathbf{B} \mathbf{x})$$

の最小点を求める問題を考える. $\mathbf{g}_k = \mathbf{g}(\mathbf{x}_k) = \mathbf{B}\mathbf{x}_k$ とかくとき, 最急降下法より, $\bar{\mathbf{y}}_k = -\mathbf{g}_k$ が得られる. さらに, 厳密直線探索法の結果, 式 (3.4.2) より,

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \frac{\mathbf{g}_k \cdot \mathbf{g}_k}{\mathbf{g}_k \cdot (\mathbf{B} \mathbf{g}_k)} \bar{\mathbf{y}}_k$$

によって点列が生成される。このとき、

$$\begin{aligned}
 \bar{f}(\mathbf{x}_{k+1}) &= \frac{1}{2} \left\{ \mathbf{x}_k - \frac{\mathbf{g}_k \cdot \mathbf{g}_k}{\mathbf{g}_k \cdot (B\mathbf{g}_k)} \mathbf{g}_k \right\} \cdot \left\{ B \left(\mathbf{x}_k - \frac{\mathbf{g}_k \cdot \mathbf{g}_k}{\mathbf{g}_k \cdot B\mathbf{g}_k} \mathbf{g}_k \right) \right\} \\
 &= \frac{1}{2} \left[\mathbf{x}_k \cdot (B\mathbf{x}_k) - \frac{2(\mathbf{g}_k \cdot \mathbf{g}_k) \{\mathbf{g}_k \cdot (B\mathbf{x}_k)\} - (\mathbf{g}_k \cdot \mathbf{g}_k)^2}{\mathbf{g}_k \cdot (B\mathbf{g}_k)} \right] \\
 &= \frac{1}{2} \left\{ \mathbf{x}_k \cdot (B\mathbf{x}_k) - \frac{(\mathbf{g}_k \cdot \mathbf{g}_k)^2}{\mathbf{g}_k \cdot (B\mathbf{g}_k)} \right\} \\
 &= \frac{1}{2} \mathbf{x}_k \cdot (B\mathbf{x}_k) \left(1 - \frac{(\mathbf{g}_k \cdot \mathbf{g}_k)^2}{\{\mathbf{x}_k \cdot (B\mathbf{x}_k)\} \{\mathbf{g}_k \cdot (B\mathbf{g}_k)\}} \right) \\
 &= \left(1 - \frac{(\mathbf{g}_k \cdot \mathbf{g}_k)^2}{\{\mathbf{g}_k \cdot (B^{-1}\mathbf{g}_k)\} \{\mathbf{g}_k \cdot (B\mathbf{g}_k)\}} \right) f(\mathbf{x}_k)
 \end{aligned}$$

が成り立つ。ここで、任意の $\mathbf{y} \in X$ に対して、Kantorovich の不等式

$$\frac{4\lambda_1\lambda_d}{(\lambda_1 + \lambda_d)^2} \leq \frac{(\mathbf{y} \cdot \mathbf{y})^2}{\{\mathbf{y} \cdot (B^{-1}\mathbf{y})\} \{\mathbf{y} \cdot (B\mathbf{y})\}}$$

が成り立つことを用いれば、

$$\bar{f}(\mathbf{x}_{k+1}) \leq \left\{ 1 - \frac{4\lambda_1\lambda_d}{(\lambda_1 + \lambda_d)^2} \right\} \bar{f}(\mathbf{x}_k) = \left(\frac{\lambda_d - \lambda_1}{\lambda_1 + \lambda_d} \right)^2 \bar{f}(\mathbf{x}_k)$$

が得られる。 □

以上の結果をふまえて、厳密直線探索法の特徴について考えてみよう。厳密直線探索法では、ステップサイズのみを設計変数においた最小化問題を正確に解くことを要請している。最小点を正確に求めるためには、二分法やセカント法などのような繰り返し計算が必要となる。もしも、勾配を求める計算が容易ではなく、設計変数に対して評価関数を求める計算が容易な問題に対しては、二分法などのような勾配の計算が不要なアルゴリズムは有効であると考えられる。しかしながら、設計変数が探索方向に移動したあとでは勾配は変化し、勾配法から決定される探索方向も変化することになる。そのような状況にあっても古い探索方向を使って最小点を正確に求めることは、必ずしも得策ではないと考えられる。特に、セカント法のように $\bar{\epsilon}_g$ を更新したあとに勾配の再計算が必要となる (演習問題 3.2) アルゴリズムを用いる場合には、古い探索方向を使い続けるよりは、勾配法によって探索方向を更新したほうがより収束性が向上すると思われる。

そこで、次に、厳密であることにはこだわらずに、評価関数の非線形性と勾配の有効な範囲に注目した方法について考えてみることにしよう。以下に示される規準は、ステップサイズの上限と下限を与える。ステップサイズ $\|\bar{\epsilon}_g \bar{\mathbf{y}}_g\|_X$ の上限に関しては、次のような条件が知られている [1].

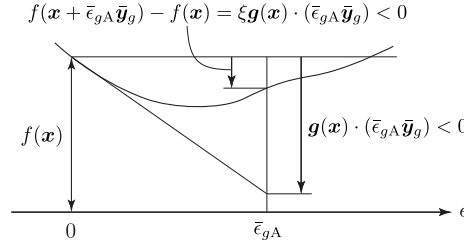


図 3.5: Armijo の規準

定義 3.4.4 (Armijo の規準) $g(x_k)$ を $f(x_k)$ の勾配, \bar{y}_g を探索方向, $\xi \in (0, 1)$ をステップサイズの上限を調整するパラメータとする. このとき, $\bar{\epsilon}_g > 0$ に対して

$$f(x_k + \bar{\epsilon}_g \bar{y}_g) - f(x_k) \leq \xi g(x_k) \cdot (\bar{\epsilon}_g \bar{y}_g) < 0 \quad (3.4.3)$$

が成り立つとき, $\bar{\epsilon}_g$ は [Armijo の規準](#)を満たすという. \square

Armijo の規準を満たす $\bar{\epsilon}_g$ の上限値を $\bar{\epsilon}_{gA}$ とかけば, 図 3.5 のような関係が成り立つ. 式 (3.4.3) の左辺は, x が x_k から $\bar{\epsilon}_g \bar{y}_g$ だけ変動したときに非線形関数 f が実際に減少した負の値を表している. それに対して右辺の $g(x_k) \cdot (\bar{\epsilon}_g \bar{y}_g)$ は, 勾配を使って f の減少を予測したときの負の値を表している. $\bar{\epsilon}_g$ が十分小さければ両者は一致するはずである. しかし, $\bar{\epsilon}_g$ がある大きさをもつときには異なってくる. $\xi \in (0, 1)$ はその違いを許容する割合を与えている. ξ を 1 に近づけることはその違いを許容しないことを表し, ξ を 0 に近づけることは違いを許容することを表す. したがって, Armijo の規準は, 勾配を使った予測値が f の実際の減少値からおおきく外れない程度にステップサイズを決めるための条件を与えていることになる.

また, ξ の目安として, 次の結果が参考になる. $f(x)$ が 2 次関数ならば, $\xi = 1/2$ のときの Armijo の規準の上限値を $\bar{\epsilon}_{gA}$ とするとき, $x_k + \bar{\epsilon}_{gA} \bar{y}_g$ が f の最小点となる (図 3.6). 実際, 式 (3.4.1) の 2 次関数に対して

$$f(x_k + \bar{\epsilon}_{gA} \bar{y}_g) - f(x_k) = g(x_k) \cdot (\bar{\epsilon}_{gA} \bar{y}_g) + \frac{1}{2} (\bar{\epsilon}_{gA} \bar{y}_g) \cdot \{B(\bar{\epsilon}_{gA} \bar{y}_g)\} \quad (3.4.4)$$

が成り立つ. ここで, $x_k + \bar{\epsilon}_{gA} \bar{y}_g$ が極小点ならば, $g(x_k)$ の Taylor 展開について

$$g(x_k + \bar{\epsilon}_{gA} \bar{y}_g) = g(x_k) + B(\bar{\epsilon}_{gA} \bar{y}_g) = \mathbf{0}_{X'} \quad (3.4.5)$$

が成り立つ. 式 (3.4.5) を式 (3.4.4) に代入すれば,

$$f(x_k + \bar{\epsilon}_{gA} \bar{y}_g) - f(x_k) = \frac{1}{2} g(x_k) \cdot (\bar{\epsilon}_{gA} \bar{y}_g)$$

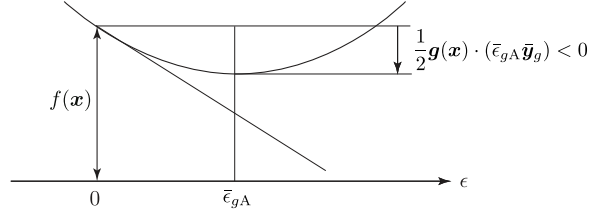


図 3.6: 2 次関数に対する Armijo の規準

が得られる.

一方, ステップサイズ $\|\bar{\epsilon}_g \bar{\mathbf{y}}_g\|_X$ の下限を与える条件として次のような条件が知られている [3]².

定義 3.4.5 (Wolfe の規準) $\mathbf{g}(\mathbf{x}_k)$ を $f(\mathbf{x}_k)$ の勾配, $\bar{\mathbf{y}}_g$ を探索方向, $\xi \in (0, 1)$ を Armijo の規準で使われたパラメータ, $\mu \in (0, 1)$ をステップサイズの下限を調整するパラメータで $0 < \xi < \mu < 1$ を満たすとする. このとき, $\bar{\epsilon}_g > 0$ に対して

$$\mu \mathbf{g}(\mathbf{x}_k) \cdot (\bar{\epsilon}_g \bar{\mathbf{y}}_g) \leq \mathbf{g}(\mathbf{x}_k + \bar{\epsilon}_g \bar{\mathbf{y}}_g) \cdot (\bar{\epsilon}_g \bar{\mathbf{y}}_g) < 0 \quad (3.4.6)$$

が成り立つとき, $\bar{\epsilon}_g$ は **Wolfe の規準**を満たすという. □

Wolfe の規準を満たす $\bar{\epsilon}_g$ の下限値を $\bar{\epsilon}_{gW}$ とかけば, 図 3.7 のような関係が成り立つ. 式 (3.4.6) の左辺における $\mathbf{g}(\mathbf{x}_k)$ は, \mathbf{x}_k のときの f の勾配を表している. 一方, 右辺の $\mathbf{g}(\mathbf{x}_k + \bar{\epsilon}_g \bar{\mathbf{y}}_g)$ は, 実際に \mathbf{x} を \mathbf{x}_k から $\mathbf{x}_k + \bar{\epsilon}_g \bar{\mathbf{y}}_g$ に移動したときに得られる f の勾配を表している. Wolfe の規準に対して, 次のことが成り立つ.

- (1) ある $\bar{\epsilon}_g > 0$ に対して $\mathbf{g}(\mathbf{x}_k + \bar{\epsilon}_g \bar{\mathbf{y}}_g) \cdot (\bar{\epsilon}_g \bar{\mathbf{y}}_g) \leq \mathbf{g}(\mathbf{x}_k) \cdot (\bar{\epsilon}_g \bar{\mathbf{y}}_g) < 0$ が成り立つとき, 式 (3.4.6) が成り立つような $\bar{\epsilon}_g > 0$ は存在しない. この条件は, \mathbf{x} が $\bar{\mathbf{y}}_g$ の方向に移動したときに, f が減少するような負の勾配がさらに大きな負の勾配になることを表している. このようなときにはステップサイズの下限を設ける必要はないことを表している.
- (2) ある $\bar{\epsilon}_g > 0$ に対して $\mathbf{g}(\mathbf{x}_k) \cdot (\bar{\epsilon}_g \bar{\mathbf{y}}_g) < \mathbf{g}(\mathbf{x}_k + \bar{\epsilon}_g \bar{\mathbf{y}}_g) \cdot (\bar{\epsilon}_g \bar{\mathbf{y}}_g) < 0$ が成り立つとき, 式 (3.4.6) が成り立つような $\bar{\epsilon}_g > 0$ は存在する. この条件は, 上記 (1) とは反対に, \mathbf{x} が $\bar{\mathbf{y}}_g$ の方向に移動したときに, 勾配の大きさが減少することを表している. 式 (3.4.6) において μ を 1 よりも小さくしていけば, $\bar{\epsilon}_g$ の下限値 $\bar{\epsilon}_{gW}$ は大きくなっていく.

²文献では Wolfe の規準の中に Armijo の規準が含まれているが, 本書では $\bar{\epsilon}_g$ の下限を与える条件のみを Wolfe の規準とよぶことにする.

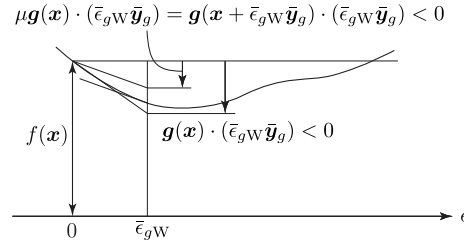


図 3.7: Wolfe の規準

そこで, Wolfe の規準は, 勾配の有効性が μ の割合まで失われる程度にステップサイズを大きくすることを要請する条件になっている.

また, $\xi < \mu$ の根拠は次のような関係に基づいている. f の $x_k + \bar{\epsilon}_g \bar{y}_g$ まわりにおける Taylor 展開を

$$f(x_k) = f(x_k + \bar{\epsilon}_g \bar{y}_g) - g(x_k + \bar{\epsilon}_g \bar{y}_g) \cdot (\bar{\epsilon}_g \bar{y}_g) + o(\bar{\epsilon}_g)$$

とかく. このとき, Wolfe の規準が満たされるならば,

$$\begin{aligned} \mu g(x_k) \cdot (\bar{\epsilon}_g \bar{y}_g) - o(\bar{\epsilon}_g) &\leq g(x_k + \bar{\epsilon}_g \bar{y}_g) \cdot (\bar{\epsilon}_g \bar{y}_g) - o(\bar{\epsilon}_g) \\ &= f(x_k + \bar{\epsilon}_g \bar{y}_g) - f(x_k) \end{aligned}$$

が成り立つ. 一方, Armijo の規準が満たされるならば,

$$f(x_k + \bar{\epsilon}_g \bar{y}_g) - f(x_k) \leq \xi g(x_k) \cdot (\bar{\epsilon}_g \bar{y}_g)$$

が成り立つ. そこで, 二つの規準を同時に満たすならば,

$$(\mu - \xi) g(x_k) \cdot (\bar{\epsilon}_g \bar{y}_g) \leq o(\bar{\epsilon}_g)$$

が成り立つ必要がある. ここで, $g(x_k) \cdot (\bar{\epsilon}_g \bar{y}_g) \leq 0$ であることから, 右辺が正, あるいは絶対値が十分に小さい場合には, $\xi < \mu$ のときにこの不等式は成り立つことになる.

Armijo の規準と Wolfe の規準が満たされるようにステップサイズが制御されたアルゴリズムの一例を次に示す. 図 3.8 はその概要を示す.

アルゴリズム 3.4.6 (Armijo の規準と Wolfe の規準) 問題 3.1.1 において f_0 を f とかき, すべての不等式制約は無効とする.

- (1) 初期点 x_0 , 正定値実対称行列 A , ステップサイズ ϵ_g , 収束判定値 ϵ_0 および Armijo の規準と Wolfe の規準で使われるパラメータ ξ と μ ($0 < \xi < \mu < 1$) を与える. $c_a = 1$ および $k = 0$ とおく.

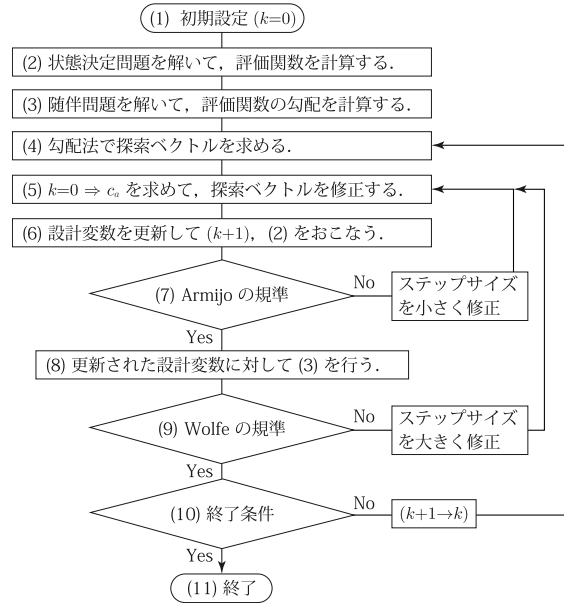


図 3.8: Armijo の規準と Wolfe の規準を用いた勾配法のアルゴリズム

- (2) 状態決定問題を解いて、 $f(\mathbf{x}_k)$ を計算する。
- (3) f に対する随伴問題を解いて、 $\mathbf{g}(\mathbf{x}_k)$ を計算する。
- (4) 式 (3.3.6) で \mathbf{y}_g を計算する。
- (5) $k = 0$ のとき、 $\mathbf{y}_g = \bar{\mathbf{y}}_g$ とおき、式 (3.3.7) で c_a を求める。また、 $\bar{\mathbf{y}}_g/c_a$ を \mathbf{y}_g に代入する。
- (6) $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{y}_g$ とおき、状態決定問題を解いて、 $f(\mathbf{x}_{k+1})$ を計算する。
- (7) Armijo の規準 (式 (3.4.3)) を判定する。
 - 満たされていれば、次に進む。
 - そうではないとき、 $\alpha > 1$ として c_a に αc_a を代入し、 \mathbf{y}_g に $\alpha \mathbf{y}_g$ を代入して、(5) に戻る。
- (8) $\mathbf{g}(\mathbf{x}_{k+1})$ を計算する。
- (9) Wolfe の規準 (式 (3.4.6)) を判定する。
 - 満たされていれば、次に進む。

- そうではないとき, $\beta \in (0, 1)$ として c_a に βc_a を代入し, \mathbf{y}_g に $\beta \mathbf{y}_g$ を代入して, (5) に戻る.

(10) 終了条件 $|f_0(\mathbf{x}_{k+1}) - f_0(\mathbf{x}_k)| \leq \epsilon_0$ を判定する.

- 終了条件が満たされたとき, (11) に進む.
- そうではないとき, $k+1$ を k に代入し, (4) に戻る.

(11) 計算を終了する.

□

Armijo の規準と Wolfe の規準が満たされるようにステップサイズが制御されたアルゴリズムによって得られる点列に対して, [大域的収束性](#)に関する次の結果が得られる.

定理 3.4.7 (大域的収束定理) $X = \mathbb{R}^d$ とする. 関数 $f : X \rightarrow \mathbb{R}$ は下界をもち, $\mathbf{x}_0 \in X$ の水準集合 $L = \{\mathbf{x} \in X \mid f(\mathbf{x}) \leq f(\mathbf{x}_0)\}$ の近傍で微分可能で, 勾配 \mathbf{g} は L の近傍で Lipschitz 連続 (定義 4.3.1) とする. \mathbf{x}_k における探索ベクトルを \mathbf{y}_{gk} とかいて, \mathbf{y}_{gk} は降下角 θ_k に対して $\cos \theta_k > 0$ を満たすとする. ステップサイズ $\|\bar{\epsilon}_g \bar{\mathbf{y}}_g\|_X$ は Armijo の規準と Wolfe の規準を満たすとする. このとき, 勾配法で生成される点列 $\{\mathbf{x}_k\}_{k \in \mathbb{N}}$ は

$$\sum_{k \in \mathbb{N}} \|\mathbf{g}(\mathbf{x}_k)\|_X^2 \cos^2 \theta_k < \infty \quad (3.4.7)$$

を満たす.

□

証明 Armijo の規準より, $\{\mathbf{x}_k\}_{k \in \mathbb{N}}$ は L の近傍に含まれる. Wolfe の規準より,

$$(\mu - 1) \mathbf{g}(\mathbf{x}_k) \cdot \mathbf{y}_g \leq (\mathbf{g}(\mathbf{x}_{k+1}) - \mathbf{g}(\mathbf{x}_k)) \cdot \mathbf{y}_g$$

が成り立つ. 一方, \mathbf{g} が Lipschitz 連続であることから, ある $\beta > 0$ に対して

$$(\mathbf{g}(\mathbf{x}_{k+1}) - \mathbf{g}(\mathbf{x}_k)) \cdot \mathbf{y}_g \leq \beta \|\mathbf{x}_{k+1} - \mathbf{x}_k\|_X \|\mathbf{y}_g\|_X = \bar{\epsilon}_g \beta \|\mathbf{y}_g\|_X^2$$

が成り立つ. これらの式より,

$$\bar{\epsilon}_g \geq \frac{(\mathbf{g}(\mathbf{x}_{k+1}) - \mathbf{g}(\mathbf{x}_k)) \cdot \mathbf{y}_g}{\beta \|\mathbf{y}_g\|_X^2} \geq \frac{(\mu - 1) \mathbf{g}(\mathbf{x}_k) \cdot \mathbf{y}_g}{\beta \|\mathbf{y}_g\|_X^2}$$

が得られる. この式を Armijo の規準に代入すれば,

$$f(\mathbf{x}_{k+1}) \leq f(\mathbf{x}_k) + \xi \bar{\epsilon}_g \mathbf{g}(\mathbf{x}_k) \cdot \mathbf{y}_g = f(\mathbf{x}_k) - \xi \frac{\mu - 1}{\beta} \left(\frac{\mathbf{g}(\mathbf{x}_k) \cdot \mathbf{y}_g}{\|\mathbf{y}_g\|_X} \right)^2$$

$$= f(\mathbf{x}_k) - \xi \frac{\mu - 1}{\beta} \|\mathbf{g}(\mathbf{x}_k)\|_{X'}^2 \cos^2 \theta_k$$

となる。したがって、

$$f(\mathbf{x}_{k+1}) \leq f(\mathbf{x}_k) - \xi \frac{\mu - 1}{\beta} \sum_{k \in \{0, \dots, m\}} \|\mathbf{g}(\mathbf{x}_k)\|_{X'}^2 \cos^2 \theta_k$$

が成り立つ。\$f\$ は下に有界であることから、定理の式が成り立つ。 \$\square\$

式 (3.4.7) は Zoutendijk 条件とよばれる。定理 3.4.7 の結果と、無限級数が収束するための必要条件 \$\lim_{k \rightarrow \infty} \|\mathbf{g}(\mathbf{x}_k)\|_{X'}^2 \cos^2 \theta_k = 0\$ を用いれば、次のような結果が得られる。

系 3.4.8 (大域的収束定理) 定理 3.4.7 の仮定に加えて、\$\mathbf{y}_g\$ は \$-\mathbf{g}(\mathbf{x}_k)\$ と直角に交わる方向に漸近することがないとき、すなわち \$\cos \theta_k > 0\$ のとき、

$$\lim_{k \rightarrow \infty} \mathbf{g}(\mathbf{x}_k) = \mathbf{0}_{X'}$$

が成り立つ。 \$\square\$

この結果は、問題設定が適切であれば、勾配法で探索方向を求め、Armijo の規準と Wolfe の規準が満たされるようにステップサイズをとっていけば、生成される点列 \$\{\mathbf{x}_k\}_{k \in \mathbb{N}}\$ は大域的収束性をもつことを示している。

この節の最後に、勾配法が拡張された方法として共役勾配法を紹介しておこう。まず、ベクトルどうしの共役を次のように定義する。

定義 3.4.9 (共役) \$B \in \mathbb{R}^{d \times d}\$ を正定値実対称行列とする。\$\mathbf{x} \in X\$ と \$\mathbf{y} \in X\$ に対して \$\mathbf{x} \cdot (B\mathbf{y}) = 0\$ のとき、\$\mathbf{x}\$ と \$\mathbf{y}\$ は共役であるという。 \$\square\$

問題 3.4.2 に対して、共役勾配法は次のように定義される。

問題 3.4.10 (共役勾配法) \$\mathbf{x}_0 \in X\$ に対して、探索方向 \$\bar{\mathbf{y}}_{g_0}\$ とステップサイズを調整するパラメータ \$\bar{\epsilon}_{g_0}\$ は、それぞれ最急降下法と厳密直線探索法により与えられているとする。\$k \in \mathbb{N}\$ に対して、\$\bar{\mathbf{y}}_{g_{k-1}}\$ を与えて、\$\bar{\mathbf{y}}_{g_{k-1}}\$ と共役になるように \$\bar{\mathbf{y}}_{g_k}\$ を求めよ。また、厳密直線探索法によってステップサイズを調整するパラメータ \$\bar{\epsilon}_{g_k}\$ を求めよ。 \$\square\$

図 3.9 は、\$X = \mathbb{R}^2\$ のときの共役勾配法によってえられた探索ベクトルを表している。\$\bar{\mathbf{y}}_{g_0}\$ と \$\bar{\mathbf{y}}_{g_1}\$ が共役になるように選ばれることによって、2次元ベクトル空間上の問題 3.4.2 であれば、探索ベクトルを 2 回求めるだけで最小点に到達できることになる。

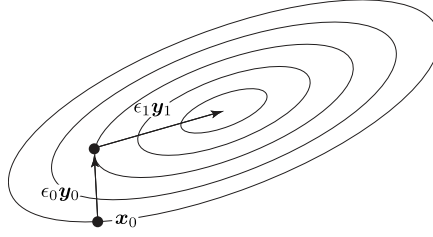


図 3.9: 共役勾配法で得られる探索ベクトル

共役勾配法の一例を示そう. $x_0 = \mathbf{0}_X$ とおく. 最急降下法により探索方向を $\bar{\mathbf{y}}_{g_0} = -\mathbf{g}_0 = -\mathbf{g}(x_0) = -\mathbf{B}x_0 - \mathbf{b} = -\mathbf{b}$ とおく. $k \in \{0, 1, 2, \dots\}$ に対して, $\bar{\mathbf{y}}_{g_k}$ と \mathbf{g}_k が与えられたとき, 式 (3.4.2) (厳密直線探索法) により,

$$\bar{c}_{gk} = \frac{\bar{\mathbf{y}}_{gk} \cdot \mathbf{g}_k}{\bar{\mathbf{y}}_{gk} \cdot (\mathbf{B}\bar{\mathbf{y}}_{gk})} \quad (3.4.8)$$

を求める. さらに, $k \in \mathbb{N}$ に対して,

$$\mathbf{x}_k = \mathbf{x}_{k-1} + \bar{c}_{gk-1} \bar{\mathbf{y}}_{gk-1}, \quad (3.4.9)$$

$$\mathbf{g}_k = \mathbf{g}_{k-1} + \bar{c}_{gk-1} \mathbf{B}\bar{\mathbf{y}}_{gk-1}, \quad (3.4.10)$$

$$\beta_k = \frac{\mathbf{g}_k \cdot \mathbf{g}_k}{\mathbf{g}_{k-1} \cdot \mathbf{g}_{k-1}}, \quad (3.4.11)$$

$$\bar{\mathbf{y}}_{gk} = -\mathbf{g}_k + \beta_k \bar{\mathbf{y}}_{gk-1} \quad (3.4.12)$$

によって点列を構成する. このとき, $\bar{\mathbf{y}}_{gk-1}$ と $\bar{\mathbf{y}}_{gk}$ は共役になる (演習問題 3.3).

式 (3.4.11) は [Fletcher-Reeves 公式](#) とよばれる. また, それと同値な

$$\beta_k = \frac{\mathbf{g}_k \cdot (\mathbf{g}_k - \mathbf{g}_{k-1})}{\mathbf{g}_{k-1} \cdot \mathbf{g}_{k-1}}$$

は [Polak-Ribiere 公式](#) とよばれる. これらの他にもいくつかの公式が知られている. これらの公式は, 2 次最適化問題 (問題 3.4.2) に対しては同値でも, そうではない非線形最適化問題に $\mathbf{g}_k = \mathbf{g}(x_k)$ を用いた場合には, 異なる結果となる.

3.5 Newton 法

勾配法では勾配 \mathbf{g} を用いて探索方向が求められた. ステップサイズは厳密直線探索法あるいは Armijo の規準や Wolfe の規準を満たすように決められた. ここでは, \mathbf{g} と Hesse 行列 \mathbf{H} を使って, 探索方向とステップサイズを同時に求める方法について

て考えてみよう．その方法は [Newton 法](#) とよばれる．Newton 法は，任意の $\mathbf{y} \in X$ に対して $\mathbf{g} \cdot \mathbf{y}$ の \mathbf{x}_k まわりの Taylor 展開

$$\mathbf{g}(\mathbf{x}_k + \mathbf{y}_g) \cdot \mathbf{y} = \mathbf{g}(\mathbf{x}_k) \cdot \mathbf{y} + \mathbf{y}_g \cdot (\mathbf{H}(\mathbf{x}_k) \mathbf{y}) + o(\|\mathbf{y}_g\|_X)$$

において， $o(\|\mathbf{y}_g\|_X)$ を無視し，

$$\mathbf{g}(\mathbf{x}_k + \mathbf{y}_g) \cdot \mathbf{y} = \mathbf{g}(\mathbf{x}_k) \cdot \mathbf{y} + \mathbf{y}_g \cdot (\mathbf{H}(\mathbf{x}_k) \mathbf{y}) = 0$$

とおくことにより探索ベクトル $\mathbf{y}_g \in X$ を求める方法である．すなわち，次のように定義される．

問題 3.5.1 (Newton 法) $X = \mathbb{R}^d$ とする． $f \in C^2(X; \mathbb{R})$ に対して，極小点ではない $\mathbf{x}_k \in X$ における f の勾配と Hesse 行列を $\mathbf{g}(\mathbf{x}_k)$ と $\mathbf{H}(\mathbf{x}_k)$ とする．このとき，任意の $\mathbf{y} \in X$ に対して

$$\mathbf{y}_g \cdot (\mathbf{H}(\mathbf{x}_k) \mathbf{y}) = -\mathbf{g}(\mathbf{x}_k) \cdot \mathbf{y} \quad (3.5.1)$$

が満たされるように $\mathbf{y}_g \in X$ を求めよ． \square

Newton 法は，勾配法 (定義 3.3.1) において使われた正定値実対称行列 \mathbf{A} を Hesse 行列におきかえて，さらに， $\bar{\epsilon}_g = 1$ とおいたときの勾配法になっている．Newton 法について次の結果が得られる．

定理 3.5.2 (Newton 法) f は極小点 \mathbf{x} の近傍で 2 階微分可能で，Hesse 行列 \mathbf{H} は Lipschitz 連続 (定義 4.3.1) であるとする．また， $\mathbf{H}(\mathbf{x})$ は正定値であるとする．このとき，極小点に十分近い点を \mathbf{x}_0 にして，Newton 法で生成された点列 $\{\mathbf{x}_k\}_{k \in \mathbb{N}}$ は \mathbf{x} に 2 次収束する． \square

証明 極小点を \mathbf{x} とする． \mathbf{x} の近傍で Hesse 行列 \mathbf{H} は Lipschitz 連続，かつ $\mathbf{H}(\mathbf{x})$ は正則であることから，ある $\beta > 0$ に対して

$$\|\mathbf{H}^{-1}(\mathbf{x})(\mathbf{H}(\mathbf{x}_k) - \mathbf{H}(\mathbf{x}))\|_{\mathbb{R}^d \times d} \leq \|\mathbf{H}^{-1}(\mathbf{x})\|_{\mathbb{R}^d \times d} \beta \|\mathbf{x}_k - \mathbf{x}\|_{\mathbb{R}^d} < \frac{1}{2} \quad (3.5.2)$$

が満たされるように極小点に十分近い点 \mathbf{x}_k をとることができる．ただし，

$$\|\mathbf{H}^{-1}(\mathbf{x})\|_{\mathbb{R}^d \times d} = \max_{\mathbf{y} \in \mathbb{R}^d, \|\mathbf{y}\|_{\mathbb{R}^d} = 1} \|\mathbf{H}^{-1}(\mathbf{x}) \mathbf{y}\|_{\mathbb{R}^d}$$

と定義する．この関係に対して [Banach の摂動定理](#) (たとえば，[4, p.240]) より，

$$\|\mathbf{H}^{-1}(\mathbf{x}_k)\|_{\mathbb{R}^d \times d} \leq \frac{\|\mathbf{H}^{-1}(\mathbf{x})\|_{\mathbb{R}^d \times d}}{1 - \|\mathbf{H}^{-1}(\mathbf{x})(\mathbf{H}(\mathbf{x}_k) - \mathbf{H}(\mathbf{x}))\|_{\mathbb{R}^d \times d}} < 2 \|\mathbf{H}^{-1}(\mathbf{x})\|_{\mathbb{R}^d \times d} \quad (3.5.3)$$

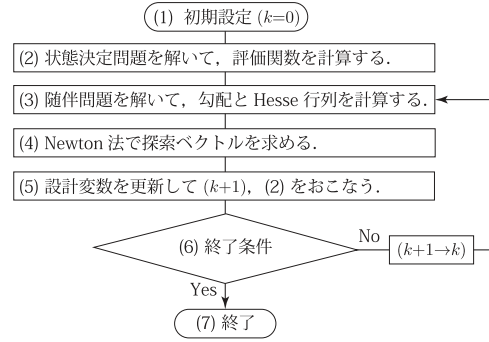


図 3.10: Newton 法のアルゴリズム

が成り立つ。このとき, $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{y}_g$, 式 (3.5.1) および $\mathbf{g}(\mathbf{x}) = \mathbf{0}_{\mathbb{R}^d}$ より,

$$\begin{aligned} \mathbf{x}_{k+1} - \mathbf{x} &= \mathbf{x}_k - \mathbf{H}^{-1}(\mathbf{x}_k) \mathbf{g}(\mathbf{x}_k) - \mathbf{x} + \mathbf{H}^{-1}(\mathbf{x}_k) \mathbf{g}(\mathbf{x}) \\ &= -\mathbf{H}^{-1}(\mathbf{x}_k) \{\mathbf{g}(\mathbf{x}_k) - \mathbf{g}(\mathbf{x}) - \mathbf{H}(\mathbf{x}_k)(\mathbf{x}_k - \mathbf{x})\} \end{aligned} \quad (3.5.4)$$

が成り立つ。一方,

$$\begin{aligned} &\|\mathbf{g}(\mathbf{x}_k) - \mathbf{g}(\mathbf{x}) - \mathbf{H}(\mathbf{x}_k)(\mathbf{x}_k - \mathbf{x})\|_{\mathbb{R}^d} \\ &= \left\| \int_0^1 (\mathbf{H}(\mathbf{x} + t(\mathbf{x}_k - \mathbf{x})) - \mathbf{H}(\mathbf{x}_k))(\mathbf{x}_k - \mathbf{x}) dt \right\|_{\mathbb{R}^d} \\ &\leq \|\mathbf{x}_k - \mathbf{x}\|_{\mathbb{R}^d} \int_0^1 \|\mathbf{H}(\mathbf{x} + t(\mathbf{x}_k - \mathbf{x})) - \mathbf{H}(\mathbf{x}_k)\|_{\mathbb{R}^d \times d} dt \\ &\leq \|\mathbf{x}_k - \mathbf{x}\|_{\mathbb{R}^d} \int_0^1 \beta \|\mathbf{x}_k - \mathbf{x}\|_{\mathbb{R}^d} (1-t) dt \\ &= \frac{1}{2} \beta \|\mathbf{x}_k - \mathbf{x}\|_{\mathbb{R}^d}^2 \end{aligned}$$

が成り立つ。したがって, 式 (3.5.2), 式 (3.5.3) および式 (3.5.4) より

$$\|\mathbf{x}_{k+1} - \mathbf{x}\|_{\mathbb{R}^d} \leq \frac{1}{2} \|\mathbf{H}^{-1}(\mathbf{x}_k)\|_{\mathbb{R}^d \times d} \beta \|\mathbf{x}_k - \mathbf{x}\|_{\mathbb{R}^d}^2 < \frac{1}{2} \|\mathbf{x}_k - \mathbf{x}\|_{\mathbb{R}^d} \quad (3.5.5)$$

が成り立つ。式 (3.5.5) の左辺と最右辺の関係により, 極小点 \mathbf{x} への収束が確認される。また, 式 (3.5.5) の左辺と第 1 右辺の関係により, 2 次収束が示された。□

Newton 法を使ったアルゴリズムの例を次に示す。図 3.10 はその概要を示す。

アルゴリズム 3.5.3 (Newton 法) 問題 3.1.1 において f_0 を f とかき, すべての不等式制約は無効とする。

(1) 初期値 \mathbf{x}_0 と収束判定値 ϵ_0 を定める。 $k = 0$ とおく。

(2) 状態決定問題を解いて, $f(\mathbf{x}_k)$ を計算する。

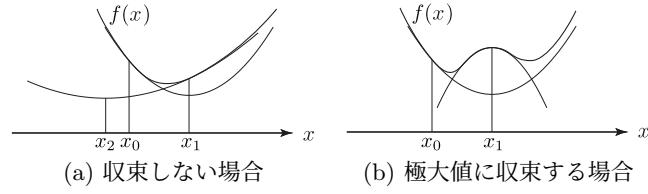


図 3.11: Newton 法

- (3) f に対する随伴問題を解いて, $\mathbf{g}(\mathbf{x}_k)$ および $\mathbf{H}(\mathbf{x}_k)$ を計算する.
- (4) 式 (3.5.1) で \mathbf{y}_g を計算する.
- (5) $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{y}_g$ とおき, 状態決定問題を解いて, $f(\mathbf{x}_{k+1})$ を計算する.
- (6) 終了条件 $|f_0(\mathbf{x}_{k+1}) - f_0(\mathbf{x}_k)| \leq \epsilon_0$ を判定する.
 - 終了条件が満たされたとき, (7) に進む.
 - そうではないとき, $k+1$ を k に代入し, (3) に戻る.
- (7) 計算を終了する.

□

Newton 法について, 次のことを指摘しておこう.

注意 3.5.4 (Newton 法) Newton 法は次の性質をもつ.

- (1) Newton 法は Hesse 行列が必要となる. Hesse 行列の計算量は, この行列が密な場合には, 設計変数の 2 乗に比例する. しかし, Hesse 行列が対角行列となる場合には, 設計変数に比例する程度である. 実際, 第 1 章の問題 1.1.4 では Hesse 行列は対角行列であった.
- (2) Newton 法は 2 次収束する (定理 3.5.2).
- (3) Hesse 行列が正定値ではない場合, Newton 法では収束しないことがある. また, 不定値の場合, 極大点に収束することがある (図 3.11).
- (4) Hesse 行列 \mathbf{H} が特異行列のとき, あるいは特異行列でなくても行列の条件数 (最大固有値の最小固有値に対する比) が大きいとき, 逆行列の計算が困難になる.

□

このようにみえてくると、勾配法 (問題 3.3.1) は、Newton 法 (問題 3.5.1) において Hesse 行列を正定値実対称行列におきかえた方法になっている。そこで、正定値実対称行列を Hesse 行列に漸近するように更新していくことで、Newton 法に近い性能をもつ勾配法が研究されている。それらは準 Newton 法とよばれる。代表的な更新公式として次のようなものが知られている。詳細は数理計画法の教科書を参照されたい。

- Davidon-Fletcher-Powell 法
- Broyden-Fletcher-Goldfarb-Shanno 法
- Broyden 法

また、Newton 法の原理は、非線形方程式の解を求めるときにも使われる。そのような場合には、Newton-Raphson 法ともよばれる。Newton-Raphson 法は、あとで示されるアルゴリズム 3.7.6 の中でも使われる。そこで、その説明をここでしておこう。Newton-Raphson 法が適用される問題は次のような問題である。

問題 3.5.5 (非線形方程式) $X = \mathbb{R}^d$ とする。 $f \in C^1(\mathbb{R}^d; \mathbb{R}^d)$ に対して、

$$f(x) = \mathbf{0}_{\mathbb{R}^d} \quad (3.5.6)$$

が満たされるように $x \in \mathbb{R}^d$ を求めよ。 \square

問題 3.5.5 の試行点を $k \in \{0, 1, 2, \dots\}$ に対して x_k とかくことにする。 x_k における f とその勾配 $G = (\partial f_i / \partial x_j)_{(i,j) \in \{1, \dots, d\}^2}$ が計算可能であるとする。このとき、 $x_{k+1} = x_k + y_g$ が x となるような y_g を求めることを考える。 f の x_k まわりの Taylor 展開は

$$f(x_k + y_g) = f(x_k) + G(x_k) y_g + o(\|y_g\|_{\mathbb{R}^d})$$

とかける。ここで、 $o(\|y_g\|_{\mathbb{R}^d})$ を無視すれば、

$$f(x_k + y_g) = f(x_k) + G(x_k) y_g = \mathbf{0}_{\mathbb{R}^d} \quad (3.5.7)$$

となる。式 (3.5.7) より、次の問題を考える。

問題 3.5.6 (Newton-Raphson 法) 問題 3.5.5 に対して、試行点 x_k における関数値 $f(x_k)$ とその勾配 $G(x_k)$ が与えられたとする。探索ベクトルを

$$y_g = -G^{-1}(x_k) f(x_k) \quad (3.5.8)$$

によって求めよ。 \square

Newton-Raphson 法は、問題 3.5.6 の解 \mathbf{y}_g を用いて、 $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{y}_g$ により問題 3.5.5 の解 \mathbf{x} に収束する点列 $\{\mathbf{x}_k\}_{k \in \mathbb{N}}$ を求める方法である。Newton 法 (問題 3.5.1) の \mathbf{g} と \mathbf{H} をそれぞれ \mathbf{f} と \mathbf{G} におきかえれば、Newton-Raphson 法 (問題 3.5.6) に一致することが理解されよう。

この節の最後に、1.1.6 節の最後に説明した Lagrange 乗数法による評価関数の 2 階微分を用いた場合の Newton 法について考えておこう。すなわち、評価関数の 2 階微分が式 (1.1.49) のような Hesse 勾配によって得られていることを仮定する。この場合、 $\mathbf{H}_0 \mathbf{b}_1$ が $\mathbf{g}_{H0}(\mathbf{a}, \mathbf{b}_1)$ として得られることから、 $\mathbf{g}_{H0}(\mathbf{a}, \mathbf{b}_1)$ を勾配とみなしたときの勾配法を考えればよい。しかし、 $\mathbf{g}_{H0}(\mathbf{a}, \mathbf{b}_1)$ を求めるために、評価関数の 1 階微分に対する随伴問題 (問題 1.1.6) を解く必要がある。さらに、その随伴問題を解くためには、 \mathbf{b}_1 が既知である必要がある。これらのことを考慮すれば、解くべき問題は次のようになる。

問題 3.5.7 (Hesse 勾配を用いた Newton 法) $X = \mathbb{R}^d$ とする。 $\mathbf{A} \in \mathbb{R}^{d \times d}$ と c_a を正定値実対称行列と正の定数とする。 $f \in C^2(X; \mathbb{R})$ に対して、極小点ではない $\mathbf{x}_k \in X$ における f の勾配 $\mathbf{g}(\mathbf{x}_k)$ 、探索ベクトル $\bar{\mathbf{y}}_g$ および Hesse 勾配 $\mathbf{g}_H(\mathbf{x}_k, \bar{\mathbf{y}}_g)$ が与えられたとする。このとき、任意の $\mathbf{y} \in X$ に対して

$$\mathbf{y}_g \cdot (c_a \mathbf{A}(\mathbf{x}_k) \mathbf{y}) = -(\mathbf{g}(\mathbf{x}_k) + \mathbf{g}_H(\mathbf{x}_k, \bar{\mathbf{y}}_g)) \cdot \mathbf{y} \quad (3.5.9)$$

が満たされるように $\mathbf{y}_g \in X$ を求めよ。 \square

問題 3.5.7 の解 \mathbf{y}_g は、 $c_a \mathbf{A}(\mathbf{x}_k) = \mathbf{I}$ および $\bar{\mathbf{y}}_g = \mathbf{y}_g$ のとき、Newton 法の解と一致する。問題 3.5.7 のを用いたアルゴリズムは、例えば、次のようになる。図 3.12 はその概要を示す。

アルゴリズム 3.5.8 (Hesse 勾配を用いた Newton 法) 問題 3.1.1 において f_0 を f とかき、すべての不等式制約は無効とする。

- (1) 初期値 \mathbf{x}_0 と収束判定値 ϵ_0 を定める。 $k = 0$ とおく。
- (2) 状態決定問題を解いて、 $f(\mathbf{x}_k)$ を計算する。
- (3) f に対する随伴問題を解いて、 $\mathbf{g}(\mathbf{x}_k)$ を計算する。
- (4) 勾配法 (式 (3.3.6)) で \mathbf{y}_g を計算する。
- (5) f' に対する随伴問題を解いて、Hesse 勾配 $\mathbf{g}_H(\mathbf{x}_k, \mathbf{y}_g)$ を計算する。
- (6) $\mathbf{g}_H(\mathbf{x}_k, \mathbf{y}_g)$ を用いた Newton 法 (式 (3.5.9)) で \mathbf{y}_g を計算する。

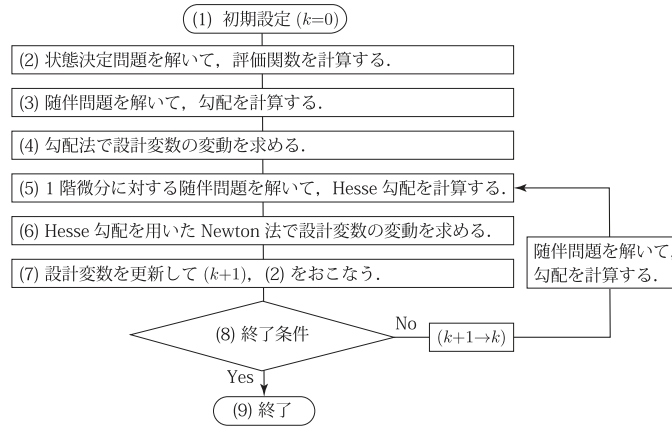


図 3.12: Hesse 勾配を用いた Newton 法のアルゴリズム

(7) $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{y}_g$ とおき、状態決定問題を解いて、 $f(\mathbf{x}_{k+1})$ を計算する。

(8) 終了条件 $|f_0(\mathbf{x}_{k+1}) - f_0(\mathbf{x}_k)| \leq \epsilon_0$ を判定する。

- 終了条件が満たされたとき、(7) に進む。
- そうではないとき、 $k+1$ を k に代入し、 f に対する随伴問題を解いて、 $\mathbf{g}(\mathbf{x}_k)$ を計算し、(5) に戻る。

(9) 計算を終了する。

□

3.6 拡大関数法

3.3 節から 3.5 節までは制約なし最適化問題に対する数値解法についてみてきた。この節以降では問題 3.1.1 に戻って、極小点で不等式制約が有効となる場合について考えよう。まず、本節では、制約関数に重みを表す定数をかけて目的関数に加えることで、制約なし問題におきかえる方法について考えてみることにする。このような方法は、**拡大関数法**とよばれる。拡大関数の最小値を求める方法には、3.3 節から 3.5 節までに示された制約なし最適化問題の解法が使われる。

拡大関数法の一つは、すべての不等式制約が満たされている点 (内点) を初期点として、許容集合の外に出ないように工夫された拡大関数を用いる方法である。問題 3.1.1 に対して、次のようにして得られる点列 $\{\mathbf{x}_k\}_{k \in \mathbb{N}}$ の収束点により解を求める方法を**障壁法**あるいは**内点法**という。

問題 3.6.1 (障壁法, 内点法) $\{\rho_k\}_{k \in \mathbb{N}}$ を正値をとる単調減少数列とする. $\mathbf{x}_0 \in X$ は $f_1(\mathbf{x}_0) < 0, \dots, f_m(\mathbf{x}_0) < 0$ が満たされるように与えられているとする. $k \in \mathbb{N}$ に対して, ρ_k と試行点 \mathbf{x}_{k-1} を与えて,

$$\min_{\mathbf{y} \in X} \left\{ \hat{f}_k(\mathbf{x}_k, \rho_k) = f_0(\mathbf{x}_k) - \rho_k \sum_{i \in \{1, \dots, m\}} \log(-f_i(\mathbf{x}_k)) \right\}$$

を満たす $\mathbf{x}_k = \mathbf{x}_{k-1} + \mathbf{y}$ を求めよ. \square

拡大関数法のもう一つの方法は, 不等式制約条件が満たされていない点を初期点に選び, 試行点を求めるごとに許容集合内に収まるように追い込んでいくように工夫された拡大関数を用いる方法である. 問題 3.1.1 に対して, 次のようにして得られる点列 $\{\mathbf{x}_k\}_{k \in \mathbb{N}}$ の収束点により解を求める方法を **罰金法**あるいは**外点法**という.

問題 3.6.2 (罰金法, 外点法) $\{\rho_k\}_{k \in \mathbb{N}}$ を正値をとるある単調増加数列とする. $\mathbf{x}_0 \in X$ は与えられているとする. $k \in \mathbb{N}$ に対して, ρ_k と試行点 \mathbf{x}_{k-1} を与えて,

$$\min_{\mathbf{y} \in X} \left\{ \hat{f}_k(\mathbf{x}_k, \rho_k) = f_0(\mathbf{x}_k) + \rho_k \sum_{i \in \{1, \dots, m\}} \max\{0, f_i(\mathbf{x}_k)\} \right\}$$

を満たす $\mathbf{x}_k = \mathbf{x}_{k-1} + \mathbf{y}$ を求めよ. \square

以上の定義から, 拡大関数法は, 原理が明快であることから, 利用しやすいと思われる. しかし, 拡大関数法を使用するためには, 問題に応じて単調減少数列あるいは単調増加数列 $\{\rho_k\}_{k \in \mathbb{N}}$ を適切に選ぶ必要がある. 本書では, その選択方法の詳細には触れないことにする.

3.7 制約つき問題に対する勾配法

次に, 3.7 節と 3.8 節において, **KKT 条件**を用いる方法について考えてみよう. ここで示されるアルゴリズムは第 7 章以降で使われる. 最初に, **制約つき問題に対する勾配法**について考えよう.

問題 3.1.1 に対して, 不等式制約が満たされる許容集合を

$$S = \{\mathbf{x} \in X \mid f_1(\mathbf{x}) \leq 0, \dots, f_m(\mathbf{x}) \leq 0\} \quad (3.7.1)$$

とかくことにする. また, $\mathbf{x} \in S$ に対して,

$$I_A(\mathbf{x}) = \{i \in \{1, \dots, m\} \mid f_i(\mathbf{x}) \geq 0\} = \{i_1, \dots, i_{|I_A(\mathbf{x})|}\} \quad (3.7.2)$$

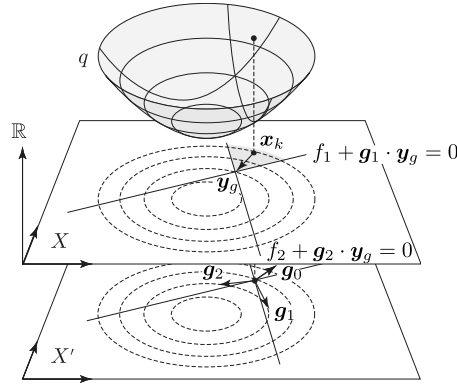


図 3.13: 制約つき問題に対する勾配法

を有効な制約に対する添え字の集合とする．混乱がない場合には $I_A(x_k)$ を I_A とかくことにする．

勾配法は, $k \in \{0, 1, 2, \dots\}$ に対して, 試行点 $x_k \in S$ のまわりで, 評価関数 f が式 (3.3.8) のような近似 2 次関数 $q(y)$ で近似されたときの Newton 法になっていた．そこで, 問題 3.1.1 に対して, 目的関数を式 (3.3.8) のように仮定して, 不等式制約が勾配を用いた 1 次関数で近似された次のような問題を考えることにしよう．

問題 3.7.1 (制約つき問題に対する勾配法) 問題 3.1.1 の試行点 $x_k \in S$ において $f_0(x_k), f_{i_1}(x_k) = 0, \dots, f_{i_{|I_A|}}(x_k) = 0, g_0(x_k), g_{i_1}(x_k), \dots, g_{i_{|I_A|}}(x_k)$ を既知とする．また, $A \in \mathbb{R}^{d \times d}$ を正定値実対称行列, c_a を正の定数とする．このとき,

$$q(y_g) = \min_{y \in X} \left\{ q(y) = \frac{1}{2} y \cdot (c_a A y) + g_0(x_k) \cdot y + f_0(x_k) \mid \begin{array}{l} f_i(x_k) + g_i(x_k) \cdot y \leq 0 \text{ for } i \in I_A(x_k) \end{array} \right\}$$

を満たす $x_{k+1} = x_k + y_g$ を求めよ．

□

問題 3.7.1 は, 2.2 節の最適化問題の分類によれば, 2 次最適化問題に分類される．さらに, A は正定値実対称行列であることから凸最適化問題となる．したがって, この問題に対する KKT 条件を満たす y_g は問題 3.7.1 の最小点となる (図 3.13)．その y_g をみつける方法について考えてみよう．

問題 3.7.1 の Lagrange 関数を,

$$\mathcal{L}_Q(y, \lambda_{k+1}) = q(y) + \sum_{i \in I_A(x_k)} \lambda_{i,k+1} (f_i(x_k) + g_i(x_k) \cdot y) \quad (3.7.3)$$

とおく. 問題 3.7.1 の最小点 \mathbf{y}_g における KKT 条件は

$$c_a \mathbf{A} \mathbf{y}_g + \mathbf{g}_0(\mathbf{x}_k) + \sum_{i \in I_A(\mathbf{x}_k)} \lambda_{i,k+1} \mathbf{g}_i(\mathbf{x}_k) = \mathbf{0}_{X'}, \quad (3.7.4)$$

$$f_i(\mathbf{x}_k) + \mathbf{g}_i(\mathbf{x}_k) \cdot \mathbf{y}_g \leq 0 \quad \text{for } i \in I_A(\mathbf{x}_k), \quad (3.7.5)$$

$$\lambda_{i,k+1} (f_i(\mathbf{x}_k) + \mathbf{g}_i(\mathbf{x}_k) \cdot \mathbf{y}_g) = 0 \quad \text{for } i \in I_A(\mathbf{x}_k), \quad (3.7.6)$$

$$\lambda_{i,k+1} \geq 0 \quad \text{for } i \in I_A(\mathbf{x}_k) \quad (3.7.7)$$

となる. すべての $i \in I_A(\mathbf{x}_k)$ に対して, 不等式制約が有効であると仮定すれば, 式 (3.7.4) と式 (3.7.5) は

$$\begin{pmatrix} c_a \mathbf{A} & \mathbf{G}^\top \\ \mathbf{G} & \mathbf{0}_{\mathbb{R}^{|I_A|} \times |I_A|} \end{pmatrix} \begin{pmatrix} \mathbf{y}_g \\ \boldsymbol{\lambda}_{k+1} \end{pmatrix} = - \begin{pmatrix} \mathbf{g}_0 \\ (f_i)_{i \in I_A} \end{pmatrix} \quad (3.7.8)$$

となる. ただし,

$$\mathbf{G}^\top = \begin{pmatrix} \mathbf{g}_{i_1}(\mathbf{x}_k) & \cdots & \mathbf{g}_{i_{|I_A(\mathbf{x}_k)|}}(\mathbf{x}_k) \end{pmatrix}$$

とする. ここで, $\mathbf{g}_{i_1}, \dots, \mathbf{g}_{i_{|I_A|}}$ が 1 次独立であれば, 式 (3.7.8) は $(\mathbf{y}_g, \boldsymbol{\lambda}_{k+1})$ について可解となる. この連立 1 次方程式の解に対して,

$$I_I(\mathbf{x}_k) = \{i \in I_A(\mathbf{x}_k) \mid \lambda_{i,k+1} < 0\} \quad (3.7.9)$$

を無効な制約条件の集合とおき, $I_I(\mathbf{x}_k) \neq \emptyset$ のときには $I_A(\mathbf{x}_k) \setminus I_I(\mathbf{x}_k)$ を $I_A(\mathbf{x}_k)$ とおきなおし, 式 (3.7.8) を再度解くことにする. このようにしてえられた $(\mathbf{y}_g, \boldsymbol{\lambda}_{k+1}) \in X \times \mathbb{R}^{|I_A|}$ は, 式 (3.7.4) から式 (3.7.7) を満たすことになる. このように, 有効な制約のみを残しながら繰り返す方法は有効制約法とよばれる (たとえば, [2, Section 10.10.6, p. 447]).

一方, 式 (3.7.8) を直接解く代わりに, $i \in I_A(\mathbf{x}_k)$ ごとに f_i に対して勾配法が使われた結果を用いる方法も考えられる. その方法は次のようである. $\mathbf{g}_0, \mathbf{g}_{i_1}, \dots, \mathbf{g}_{i_{|I_A|}}$ を用いて, 個別に勾配法を適用する. すなわち,

$$\mathbf{y}_{gi} = -(c_a \mathbf{A})^{-1} \mathbf{g}_i \quad (3.7.10)$$

が満たされるように $\mathbf{y}_{g0}, \mathbf{y}_{gi_1}, \dots, \mathbf{y}_{gi_{|I_A|}}$ を求める. ここで, Lagrange 乗数 $\boldsymbol{\lambda}_{k+1} \in \mathbb{R}^{|I_A|}$ を未知数として,

$$\mathbf{y}_g = \mathbf{y}_g(\boldsymbol{\lambda}_{k+1}) = \mathbf{y}_{g0} + \sum_{i \in I_A(\mathbf{x}_k)} \lambda_{i,k+1} \mathbf{y}_{gi} \quad (3.7.11)$$

とおく. このとき, \mathbf{y}_g は式 (3.7.8) の 1 行目を満たすことになる. 一方, 式 (3.7.8) の 2 行目は

$$\begin{pmatrix} \mathbf{g}_{i_1} \cdot \mathbf{y}_{gi_1} & \cdots & \mathbf{g}_{i_1} \cdot \mathbf{y}_{gi_{|I_A|}} \\ \vdots & \ddots & \vdots \\ \mathbf{g}_{i_{|I_A|}} \cdot \mathbf{y}_{gi_1} & \cdots & \mathbf{g}_{i_{|I_A|}} \cdot \mathbf{y}_{gi_{|I_A|}} \end{pmatrix} \begin{pmatrix} \lambda_{i_1 k+1} \\ \vdots \\ \lambda_{i_{|I_A|} k+1} \end{pmatrix} \\ = - \begin{pmatrix} f_{i_1} + \mathbf{g}_{i_1} \cdot \mathbf{y}_{g0} \\ \vdots \\ f_{i_{|I_A|}} + \mathbf{g}_{i_{|I_A|}} \cdot \mathbf{y}_{g0} \end{pmatrix}$$

となる. この式を

$$(\mathbf{g}_i \cdot \mathbf{y}_{gj})_{(i,j) \in I_A^2} (\lambda_{j k+1})_{j \in I_A} = - (f_i + \mathbf{g}_i \cdot \mathbf{y}_{g0})_{i \in I_A} \quad (3.7.12)$$

とかくことにする. ここでも, $\mathbf{g}_{i_1}, \dots, \mathbf{g}_{i_{|I_A|}}$ が 1 次独立ならば, λ_{k+1} は式 (3.7.12) により一意に決定されることになる. この連立 1 次方程式の解に対して, 有効制約法を適用する. すなわち, 式 (3.7.9) の $I_1(\mathbf{x}_k)$ が \emptyset ではないときには $I_A(\mathbf{x}_k) \setminus I_1(\mathbf{x}_k)$ を $I_A(\mathbf{x}_k)$ とおきなおし, 式 (3.7.8) を再度解くことにする. このようにしてえられた $(\mathbf{y}_g, \lambda_{k+1}) \in X \times \mathbb{R}^{|I_A|}$ は, 式 (3.7.4) から式 (3.7.7) を満たすことになる.

また, 式 (3.7.12) において, 有効な制約関数 $f_{i_1}, \dots, f_{i_{|I_A|}}$ の値がすべて 0 であれば, すべての $\mathbf{y}_{g0}, \mathbf{y}_{gi_1}, \dots, \mathbf{y}_{gi_{|I_A|}}$ に任意の定数をかけても λ_{k+1} は変化しない. このことは, ステップサイズ $\|\mathbf{y}_g\|_X$ が適切に設定されていなくても, λ_{k+1} が求められることを意味する. この関係は, 3.7.2 項においてステップサイズの初期値が与えられた大きさになるように c_a を決定するときに使われる.

以上をまとめると, 制約つき問題に対する勾配法は, 式 (3.7.8) により探索ベクトル \mathbf{y}_g と Lagrange 乗数 λ_{k+1} を直接解くか, あるいは $i \in I_A(\mathbf{x}_k)$ ごとに式 (3.7.10) より \mathbf{y}_{gi} を解いて, それらを用いて式 (3.7.12) より λ_{k+1} を求め, それらを式 (3.7.11) に代入することにより \mathbf{y}_g を求めるかのいずれかの方法により, \mathbf{y}_g を更新していく反復法であるということができる.

具体的なアルゴリズムを示す前に, いくつかの状況について考えておこう. 一つは, 問題 3.1.1 の不等式制約が等式制約 $f_i(\mathbf{x}) = 0$ におきかえられた状況である. 実は, 不等式制約が有効な場合には等式制約と同じ扱いになることから, この状況はいつも起こりえることになる. このような等式制約は, 二つの不等式制約 $f_i(\mathbf{x}) \leq 0$ および $-f_i(\mathbf{x}) \leq 0$ に置き換えられる. しかし, これら二つの不等式制約が非線形の場合, それらを厳密に満たすように \mathbf{x} を決めることは一般には困難である. そこで, 正の定数 ϵ_i を定めて, $|f_i(\mathbf{x})| \leq \epsilon_i$ のように緩和する必要がある. あとで示されるアルゴリズムでは不等式制約のみが仮定されているので, ϵ_i を用いて制約を緩

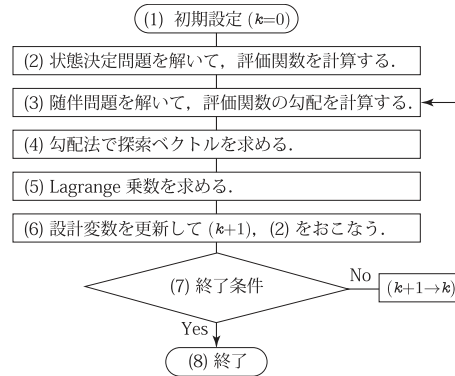


図 3.14: パラメータ調整なし制約つき問題に対する勾配法のアルゴリズム

和する必要はないように思われる。しかしながら、不等式制約が有効な場合は等式制約 $f_i(\mathbf{x}) = 0$ と同じ意味となり、正の定数 ϵ_i によるチェックが必要となる。

もう一つは、のちに示されるアルゴリズムでは、初期点 \mathbf{x}_0 においてすべての不等式制約が満たされている状況を仮定する。このような状況が満たされていない場合には、次のステップを前処理として行うことで、すべての不等式制約を満たす $\mathbf{x}_0 \in S$ をみつけることができる。みつからない場合には問題設定を見直す必要がある。

- (0) 目的関数 $f_0 = 0$ および $\mathbf{g}_0 = \mathbf{0}_{\mathbb{R}^d}$ において、すべての不等式制約が満たされるまで、あとのアルゴリズムで示される既定のステップを繰り返す。

3.7.1 簡単なアルゴリズム

上でみてきたことをふまえて、まずは、簡単なアルゴリズムについて考えてみよう。本項では、ステップサイズを調整するパラメータ c_a を決め打ちで与え、設計変数の更新後に不等式制約のチェックを行わないときのアルゴリズムの一例を示す。図 3.14 にその流れ図を示す。

アルゴリズム 3.7.2 (パラメータ調整なし制約つき問題に対する勾配法) 問題 3.1.1 の極小点を次のようにして求める。

- (1) 初期点 \mathbf{x}_0 を $f_1(\mathbf{x}_0) \leq 0, \dots, f_m(\mathbf{x}_0) \leq 0$ が満たされるように定める。式 (3.7.10) の正定値行列 \mathbf{A} とステップサイズを調整する正の定数 c_a , f_0 の収束判定に用いる正の定数 ϵ_0 および f_1, \dots, f_m の許容範囲を与える正の定数 $\epsilon_1, \dots, \epsilon_m$ を定める。また、 $k = 0$ とおく。

- (2) 状態決定問題を解いて, \mathbf{x}_k において $f_0(\mathbf{x}_k), f_1(\mathbf{x}_k), \dots, f_m(\mathbf{x}_k)$ を計算する. また,

$$I_A(\mathbf{x}_k) = \{i \in \{1, \dots, m\} \mid f_i(\mathbf{x}_k) \geq -\epsilon_i\}$$

とおく.

- (3) $f_0, f_{i_1}, \dots, f_{i_{|I_A|}}$ に対する随伴問題を解いて, \mathbf{x}_k において $\mathbf{g}_0, \mathbf{g}_{i_1}, \dots, \mathbf{g}_{i_{|I_A|}}$ を計算する.

- (4) 式 (3.7.10) で $\mathbf{y}_{g_0}, \mathbf{y}_{g_{i_1}}, \dots, \mathbf{y}_{g_{i_{|I_A|}}}$ を計算する.

- (5) 式 (3.7.12) で λ_{k+1} を求める. 式 (3.7.9) の $I_1(\mathbf{x}_k)$ が \emptyset ではないときには, $I_A(\mathbf{x}_k) \setminus I_1(\mathbf{x}_k)$ を $I_A(\mathbf{x}_k)$ とおきなおし, 式 (3.7.12) を再度解く.

- (6) 式 (3.7.11) で \mathbf{y}_g を求め, $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{y}_g$ とおき, 状態決定問題を解いて, $f_0(\mathbf{x}_{k+1}), f_1(\mathbf{x}_{k+1}), \dots, f_m(\mathbf{x}_{k+1})$ を計算する. また,

$$I_A(\mathbf{x}_{k+1}) = \{i \in \{1, \dots, m\} \mid f_i(\mathbf{x}_{k+1}) \geq -\epsilon_i\}$$

とおく.

- (7) 終了条件 $|f_0(\mathbf{x}_{k+1}) - f_0(\mathbf{x}_k)| \leq \epsilon_0$ を判定する.

- 終了条件が満たされたとき, (8) に進む.
- そうではないとき, $k+1$ を k に代入し, (3) に戻る.

- (8) 計算を終了する.

□

アルゴリズム 3.7.2 を使って, 第 1 章の例題 1.1.7 (平均コンプライアンス最小化問題の数値例) に対する試行点を求めてみよう.

例題 3.7.3 (平均コンプライアンス最小化問題の数値例) 例題 1.1.7 に対して, 初期点を $\mathbf{a}_{(0)} = (1/2, 1/2)^\top$ とおき, アルゴリズム 3.7.2 を使って $k \in \{0, 1\}$ のときの試行点を求めよ. ただし, 必要となる行列や数値は適当に定めよ. □

解答 平均コンプライアンス $\tilde{f}_0(\mathbf{a})$ と体積に対する制約関数 $f_1(\mathbf{a})$ は, それぞれ

$$\tilde{f}_0(\mathbf{a}) = \frac{4}{a_1} + \frac{1}{a_2}, \quad (3.7.13)$$

$$f_1(\mathbf{a}) = a_1 + a_2 - 1 \quad (3.7.14)$$

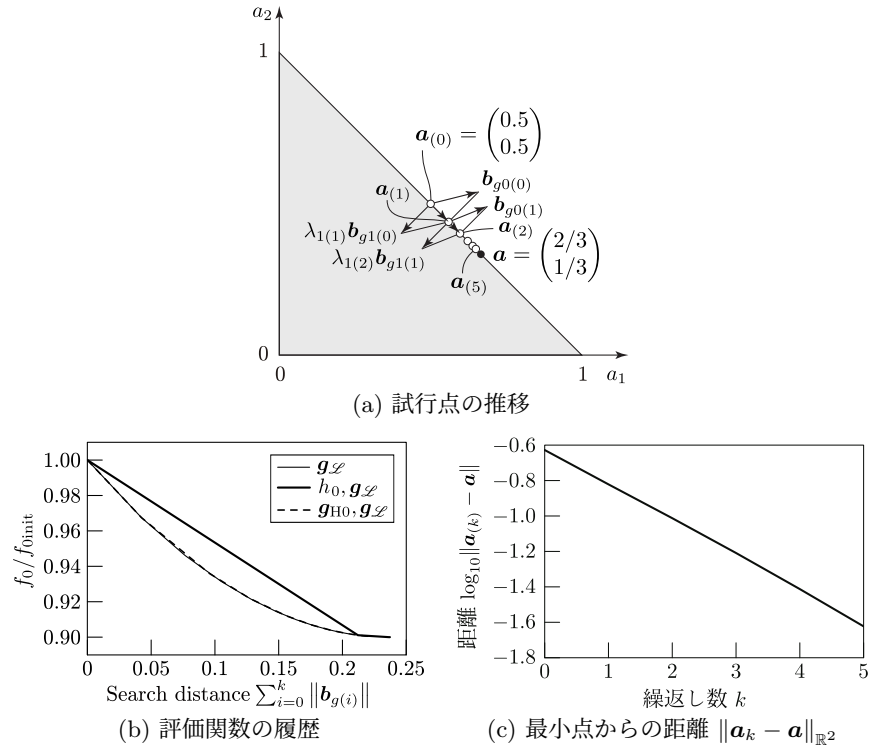


図 3.15: パラメータ調整なし制約つき問題に対する勾配法による平均コンプライアンス最小化問題の数値例

で与えられる．また，それらの断面積微分は

$$\mathbf{g}_0 = - \begin{pmatrix} \frac{4}{a_1^2} \\ \frac{1}{a_2^2} \end{pmatrix}, \quad (3.7.15)$$

$$\mathbf{g}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (3.7.16)$$

となる．アルゴリズム 3.7.2 に沿って，数値を求めていく．ただし，ステップ数 k のときの設計変数を $\mathbf{a}_{(k)}$ とかくことにする． $\mathbf{b}_{g0(k)}$, $\mathbf{b}_{g1(k)}$ および $\lambda_{1(k)}$ も同様とする．

- (1) 初期点 $\mathbf{a}_{(0)} = (1/2, 1/2)^\top$ では $f_1(\mathbf{a}_{(0)}) = 0$ は満たされる．式 (3.7.10) の正定値行列を $\mathbf{A} = \mathbf{I}$ ，ステップサイズを調整する正の定数を $c_a = 100$ (あとで示される計算より，ステップサイズ $\|\mathbf{b}_{g(0)}\|_{\mathbb{R}^2} = 0.0848528$ となった)， $\epsilon_0 = 10^{-3} \tilde{f}_0(\mathbf{a}_{(0)})$ ， $\epsilon_1 = 10^{-3}$ とおく．また， $k = 0$ とおく．
- (2) 式 (3.7.13) と式 (3.7.14) より， $\tilde{f}_0(\mathbf{a}_{(0)}) = 10$ と $f_1(\mathbf{a}_{(0)}) = 0$ が得られる．また， $I_A(\mathbf{a}_{(0)}) = \{1\}$ とおく．
- (3) 式 (3.7.15) と式 (3.7.16) より， $\mathbf{g}_{0(0)} = -(16, 4)^\top$ ， $\mathbf{g}_{1(0)} = (1, 1)^\top$ が得られる．
- (4) 式 (3.7.10) より， $\mathbf{b}_{g0(0)} = (0.16, 0.04)^\top$ ， $\mathbf{b}_{g1(0)} = -(0.01, 0.01)^\top$ が得られる．
- (5) 式 (3.7.12) で $\lambda_{1(1)} = 10$ が得られる．
- (6) 式 (3.7.11) で $\mathbf{b}_{g(0)} = (0.06, -0.06)^\top$ を得て， $\mathbf{a}_{(1)} = \mathbf{a}_{(0)} + \mathbf{b}_{g(0)} = (0.56, 0.44)^\top$ とおき， $\tilde{f}_0(\mathbf{a}_{(1)}) = 9.41558$ ， $f_1(\mathbf{a}_{(1)}) = 0$ が得られる．また， $I_A(\mathbf{a}_{(1)}) = \{1\}$ とおく．
- (7) $|\tilde{f}_0(\mathbf{a}_{(1)}) - \tilde{f}_0(\mathbf{a}_{(0)})| = 0.584416 \geq \epsilon_0 = 0.01$ より，終了条件は満たされないと判定し，1 を k に代入し，(3) に戻る．
- (3) 式 (3.7.15) と式 (3.7.16) より， $\mathbf{g}_{0(1)} = -(12.7551, 5.16529)^\top$ ， $\mathbf{g}_{1(1)} = (1, 1)^\top$ が得られる．
- (4) 式 (3.7.10) より， $\mathbf{b}_{g0(1)} = (0.127551, 0.0516529)^\top$ ， $\mathbf{b}_{g1(1)} = -(0.01, 0.01)^\top$ が得られる．
- (5) 式 (3.7.12) で $\lambda_{1(2)} = 8.9602$ が得られる．
- (6) 式 (3.7.11) で $\mathbf{b}_{g(1)} = (0.0379491, -0.0379491)^\top$ を得て， $\mathbf{a}_{(2)} = \mathbf{a}_{(1)} + \mathbf{b}_{g(1)} = (0.597949, 0.402051)^\top$ とおき， $\tilde{f}_0(\mathbf{a}_{(2)}) = 9.17678$ ， $f_1(\mathbf{a}_{(2)}) = 0$ が得られる．また， $I_A(\mathbf{a}_{(1)}) = \{1\}$ とおく．
- (7) $|\tilde{f}_0(\mathbf{a}_{(2)}) - \tilde{f}_0(\mathbf{a}_{(1)})| = 0.238804 \geq \epsilon_0 = 0.01$ より，終了条件は満たされないと判定し，2 を k に代入し，(3) に戻る．

図 3.15 にこれらとその後の結果を示す．ただし， $f_{0 \text{ init}}$ は $k = 0$ のときの f_0 の値を示す．

□

次に，例題 1.1.7 の目的関数と制約関数をいれかえてみよう．

例題 3.7.4 (体積最小化問題の数値例) 目的関数と制約関数をそれぞれ

$$f_0(\mathbf{a}) = a_1 + a_2, \quad (3.7.17)$$

$$\tilde{f}_1(\mathbf{a}) = \frac{4}{a_1} + \frac{1}{a_2} - 9 \quad (3.7.18)$$

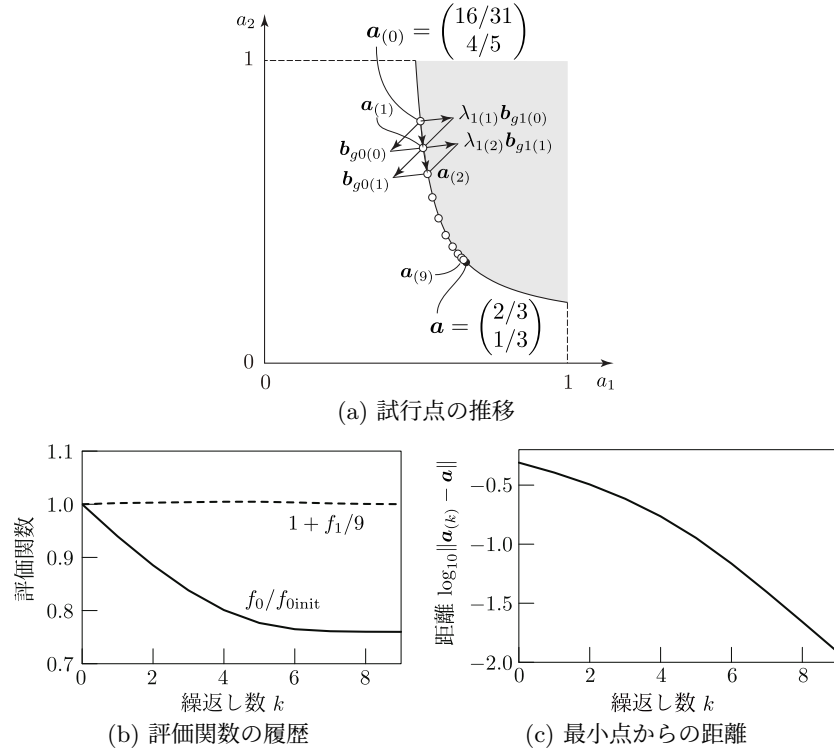


図 3.16: パラメータ調整なし制約つき問題に対する勾配法による体積最小化問題の数値例

とおく. このとき, $\tilde{f}_1(\mathbf{a}) \leq 0$ を満たす制約のもとで $f_0(\mathbf{a})$ を最小化する問題 (平均コンプライアンス制約つき体積最小化問題) に対して, 初期点を $\mathbf{a}_{(0)} = (16/31, 4/5)^\top \approx (0.516, 0.8)^\top$ とおき, アルゴリズム 3.7.2 を使って $k \in \{0, 1\}$ のときの試行点を求めよ. ただし, 必要となる行列や数値は適当に定めよ. \square

解答 評価関数 $f_0(\mathbf{a})$ と $\tilde{f}_1(\mathbf{a})$ の断面積微分はそれぞれ

$$\mathbf{g}_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad (3.7.19)$$

$$\mathbf{g}_1 = - \begin{pmatrix} \frac{4}{a_1^2} \\ \frac{1}{a_2^2} \end{pmatrix} \quad (3.7.20)$$

となる. アルゴリズム 3.7.2 に沿って, 数値を求めていく. ここでも, ステップ数 k のときの設計変数を $\mathbf{a}_{(k)}$ とかくことにする. $\mathbf{b}_{g0(k)}$, $\mathbf{b}_{g1(k)}$ および $\lambda_{1(k)}$ も同様とする.

- (1) 初期点 $\mathbf{a}_{(0)} = (16/31, 4/5)^\top$ のとき $\tilde{f}_1(\mathbf{a}_{(0)}) = 0$ は満たされる. 式 (3.7.10) の正定値行列を $\mathbf{A} = \mathbf{I}$, ステップサイズを調整する正の定数を $c_a = 10$ (あとで示され

- る計算より, ステップサイズ $\|\mathbf{b}_{g(0)}\|_{\mathbb{R}^2} = 0.089113$ となった), $\epsilon_0 = 10^{-3} f_0(\mathbf{a}_{(0)})$, $\epsilon_1 = 9 \times 10^{-3}$ とおく. また, $k = 0$ とおく.
- (2) 式 (3.7.17) と式 (3.7.18) より, $f_0(\mathbf{a}_{(0)}) = 1.31613$ と $\tilde{f}_1(\mathbf{a}_{(0)}) = 0$ が得られる. また, $I_A(\mathbf{a}_{(0)}) = \{1\}$ とおく.
- (3) 式 (3.7.15) と式 (3.7.16) より, $\mathbf{g}_{0(0)} = (1, 1)^\top$, $\mathbf{g}_{1(0)} = -(15.0156, 1.5625)^\top$ が得られる.
- (4) 式 (3.7.10) より, $\mathbf{b}_{g0(0)} = -(0.1, 0.1)^\top$, $\mathbf{b}_{g1(0)} = (1.50156, 0.15625)^\top$ が得られる.
- (5) 式 (3.7.12) で $\lambda_{1(1)} = 0.0727397$ が得られる.
- (6) 式 (3.7.11) で $\mathbf{b}_{g(0)} = (0.00922315, -0.0886344)^\top$ を得て, $\mathbf{a}_{(1)} = \mathbf{a}_{(0)} + \mathbf{b}_{g(0)} = (0.525352, 0.711366)^\top$ とおき, $f_0(\mathbf{a}_{(1)}) = 1.23672$, $\tilde{f}_1(\mathbf{a}_{(1)}) = 0.019687$ が得られる. また, $I_A(\mathbf{a}_{(1)}) = \{1\}$ とおく.
- (7) $|f_0(\mathbf{a}_{(1)}) - f_0(\mathbf{a}_{(0)})| = 0.0794113 \geq \epsilon_0 = 0.00131613$ より, 終了条件は満たされないと判定し, 1 を k に代入し, (3) に戻る.
- (3) 式 (3.7.15) と式 (3.7.16) より, $\mathbf{g}_{0(1)} = (1, 1)^\top$, $\mathbf{g}_{1(1)} = -(14.493, 1.97612)^\top$ が得られる.
- (4) 式 (3.7.10) より, $\mathbf{b}_{g0(1)} = -(0.1, 0.1)^\top$, $\mathbf{b}_{g1(1)} = (1.4493, 0.197612)^\top$ が得られる.
- (5) 式 (3.7.12) で $\lambda_{1(2)} = 0.0778958$ が得られる.
- (6) 式 (3.7.11) で $\mathbf{b}_{g(1)} = (0.0128945, -0.0846068)^\top$ を得て, $\mathbf{a}_{(2)} = \mathbf{a}_{(1)} + \mathbf{b}_{g(1)} = (0.538247, 0.626759)^\top$ とおき, $f_0(\mathbf{a}_{(2)}) = 1.16501$, $\tilde{f}_1(\mathbf{a}_{(2)}) = 0.0270467$ が得られる. また, $I_A(\mathbf{a}_{(1)}) = \{1\}$ とおく.
- (7) $|f_0(\mathbf{a}_{(2)}) - f_0(\mathbf{a}_{(1)})| = 0.0717123 \geq \epsilon_0 = 0.00131613$ より, 終了条件は満たされないと判定し, 2 を k に代入し, (3) に戻る.

図 3.16 にこれらとそのあとの結果を示す. □

例題 3.7.3 では, $f_1(\mathbf{a}_{(1)}) = f_1(\mathbf{a}_{(2)}) = 0$ のように, 制約は常に満たされていた. しかし, 例題 3.7.4 では, $\tilde{f}_1(\mathbf{a}_{(1)}) = 0.019687$ および $\tilde{f}_1(\mathbf{a}_{(2)}) = 0.0459682$ のように, 制約は満たされておらず, 反復を繰り返すごとにその超過量は拡大した. このようなことを防ぐ方法については, 次の項で考えることにしよう.

3.7.2 複雑なアルゴリズム

次に, 簡単なアルゴリズム (アルゴリズム 3.7.2) に次のような機能を追加することを考えよう.

- (i) ステップサイズの初期値 ϵ_g (あるいは目的関数の減少率 α) を与えて, $\|\mathbf{y}_g\|_X = \epsilon_g$ となるように c_a を決定する機能
- (ii) 設計変数が \mathbf{x}_{k+1} に更新されたとき, $i \in I_A(\mathbf{x}_{k+1})$ に対して, $|f_i(\mathbf{x}_{k+1})| \leq \epsilon_i$ と $\lambda_{i\ k+1} \geq 0$ が満たされるように $\boldsymbol{\lambda}_{k+1} = (\lambda_{i\ k+1})_{i \in I_A(\mathbf{x}_{k+1})}$ を修正する機能

(iii) 目的関数 f_0 の収束判定値 ϵ_0 に対して, 制約関数 f_1, \dots, f_m の許容値 $\epsilon_1, \dots, \epsilon_m$ を適切化する機能

(iv) 大域的収束性が保証されるようにステップサイズ $\|\mathbf{y}_g\|_X$ を調整する機能

上記 (i) は, アルゴリズム 3.3.5 と同様に, 式 (3.3.7) で c_a を求めることで解決される. あとで示されるアルゴリズム 3.7.6 では, ステップ (6) に取り入れられている.

また, 上記 (ii) については次のような方法が考えられる. アルゴリズム 3.7.2 のステップ (5) で計算された λ_{k+1} は, 制約つき問題に対する勾配法 (問題 3.7.1) の KKT 条件は満たしていても, 有効な不等式制約関数の非線形性により, \mathbf{x}_{k+1} において指定された許容範囲で満たしているとは限らない.

\mathbf{x}_{k+1} で不等式制約が指定された許容範囲で満たすためには, λ_{k+1} が修正されて, それにより式 (3.7.11) で \mathbf{y}_g が修正され, その結果, \mathbf{x}_{k+1} は $\mathbf{x}_k + \mathbf{y}_g$ に更新される必要がある. ここでは, $\lambda_{k+1} = \lambda_{k+1l}$ とおき, $l \in \{0, 1, 2, \dots\}$ に対して λ_{k+1l} を与えて λ_{k+1l+1} を求める計算を繰り返すことを考える. そこでは, 非線形方程式の解法である Newton-Raphson 法 (問題 3.5.6) が使われる.

その方法は次のようである. $i \in I_A(\mathbf{x}_{k+1})$ に対して $f_i(\mathbf{x}_k + \mathbf{y}_g(\lambda_{k+1l}))$ を $\bar{f}_i(\lambda_{k+1l})$ とかくことにする. Newton-Raphson 法の説明では, $k \in \{0, 1, 2, \dots\}$ に対して式 (3.5.7) の関数 $\mathbf{f}(\mathbf{x}_k + \mathbf{y}_g) = \mathbf{0}_{\mathbb{R}^d}$ を考えた. ここでは, $l \in \{0, 1, 2, \dots\}$ に対して $(\bar{f}_i(\lambda_{k+1l} + \delta\lambda))_{i \in I_A} = \mathbf{0}_{\mathbb{R}^{|I_A|}}$ を考える. また, $\mathbf{y}_g(\lambda_{k+1l})$ は式 (3.7.11) で定義された λ_{k+1l} の 1 次関数であることを考慮して, 式 (3.5.7) の $\mathbf{G}(\mathbf{x}_k)$ の代わりに $(\mathbf{g}_i(\lambda_{k+1l}) \cdot \mathbf{y}_{gj}(\lambda_{k+1l}))_{(i,j) \in I_A^2}$ を考える. このとき, 式 (3.5.8) に代わって,

$$\begin{aligned} \delta\lambda &= (\delta\lambda_j)_{j \in I_A} \\ &= -(\mathbf{g}_i(\lambda_{k+1l}) \cdot \mathbf{y}_{gj}(\lambda_{k+1l}))_{(i,j) \in I_A^2}^{-1} (f_i(\lambda_{k+1l}))_{i \in I_A} \end{aligned} \quad (3.7.21)$$

が得られる. 式 (3.7.21) の $\delta\lambda$ を用いて, λ_{k+1} は $\lambda_{k+1l+1} = \lambda_{k+1l} + \delta\lambda$ に更新されることになる. さらに, 式 (3.7.11) により \mathbf{y}_g は $\mathbf{y}_g(\lambda_{k+1l+1})$ に修正される. その結果, $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{y}_g$ は $\mathbf{x}_{k+1l+1} = \mathbf{x}_k + \mathbf{y}_g(\lambda_{k+1l+1})$ に更新される. この更新は, あとで示されるアルゴリズム 3.7.6 のステップ (11) の中で使われている.

例題 3.7.4 に対して, λ_{k+1} を修正する方法を試してみよう.

例題 3.7.5 (体積最小化問題の数値例) 例題 3.7.4 では, $\tilde{f}_1(\mathbf{a}_{(1)}) = 0.019687$ であった. 式 (3.7.21) を用いて, $\lambda_{1(1)} = 0.0727397$ を修正し, $\tilde{f}_1(\mathbf{a}_{(1)[l]}) \leq 10^{-4}$ が成り立つような試行点 $\mathbf{a}_{(1)[l]}$ を求めよ. ただし, ステップ数 k において修正回数 l のときの λ_1 を $\lambda_{1(k+1)[l]}$ とかくことにする. \square

解答 式 (3.7.21) を例題 3.7.4 に適用すれば,

$$\delta\lambda_1 = -\frac{\tilde{f}_1(\mathbf{a}_{(1)[l]})}{\mathbf{g}_{1(0)[l]} \cdot \mathbf{b}_{g1(0)[l]}} \quad (3.7.22)$$

となる. $l = 0$ のとき, $\mathbf{a}_{(1)[0]} = \mathbf{a}_{(1)}$, $\mathbf{g}_{1(0)[0]} = \mathbf{g}_{1(0)}$ および $\mathbf{b}_{g1(0)[0]} = \mathbf{b}_{g1(0)}$ より, $\delta\lambda_1 = 0.000863807$ が得られる. そこで, λ_1 を

$$\lambda_{1(1)[1]} = \lambda_1 + \delta\lambda_1 = 0.0736035$$

に更新する. この $\lambda_{1(1)[1]}$ を用いて, 式 (3.7.11) で $\mathbf{b}_{g(0)}$ を再計算すれば,

$$\mathbf{b}_{g(0)[1]} = (0.0105202, -0.0884995)^\top$$

となる. この探索ベクトルを用いて設計変数を更新すれば,

$$\mathbf{a}_{(1)[1]} = \mathbf{a}_{(1)} + \mathbf{b}_{g(0)[1]} = (0.526649, 0.711501)^\top$$

となる. このとき, $\tilde{f}_1(\mathbf{a}_{(1)[1]}) = 0.00066837 > 10^{-4}$ となる. 許容制約条件は満たされない. そこで, $l = 1$ において, 上記のステップを繰り返す. 式 (3.7.22) を適用し,

$$\delta\lambda_1 = -\frac{\tilde{f}_1(\mathbf{a}_{(1)[1]})}{\mathbf{g}_{1(0)[1]} \cdot \mathbf{b}_{g1(0)[1]}} = 0.000029326$$

が得られる. そこで, λ_1 を

$$\lambda_{1(1)[2]} = \lambda_{1(1)[1]} + \delta\lambda_1 = 0.0736328$$

に更新する. この $\lambda_{1(1)[2]}$ を用いて, 式 (3.7.11) で $\mathbf{b}_{g(0)}$ を再計算すれば,

$$\mathbf{b}_{g(0)[2]} = (0.0105202, -0.0884995)^\top$$

となる. この探索ベクトルを用いて設計変数を更新すれば,

$$\mathbf{a}_{(1)[2]} = \mathbf{a}_{(1)} + \mathbf{b}_{g(0)[2]} = (0.526693, 0.711505)^\top$$

となる. このとき, $\tilde{f}_1(\mathbf{a}_{(1)[2]}) = 0.0000243138 \leq 10^{-4}$ となる. □

さらに, 上記 (iii) については, 次のような方法が考えられる. 原問題 (問題 3.1.1) の Lagrange 関数は

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = f_0(\mathbf{x}) + \sum_{i \in I_A(\mathbf{x})} \lambda_i f_i(\mathbf{x}) \quad (3.7.23)$$

で与えられる. $i \in I_A(\mathbf{x}_k)$ に対して, $|f_i(\mathbf{x}_k)| \leq \epsilon_i$ が満たされているとき, $\mathcal{L}(\mathbf{x}_k, \boldsymbol{\lambda}_k) \approx f_0(\mathbf{x}_k)$ が成り立つためには

$$\epsilon_0 \gg \sum_{i \in I_A(\mathbf{x}_k)} \lambda_{ki} \epsilon_i \quad (3.7.24)$$

である必要がある。そこで、この条件が満たされるようにするために、 σ を 1 よりも十分小さい正の定数として、制約許容値の規準を、すべての $i \in I_A(\mathbf{x}_k)$ に対して、

$$\epsilon_i \leq \frac{\sigma \epsilon_0}{|I_A(\mathbf{x}_k)| \lambda_{ki}} \quad (3.7.25)$$

が成り立つことであるとする。この条件が満たされない i がある場合には、 ϵ_i に $\sigma \epsilon_0 / (|I_A(\mathbf{x}_k)| \lambda_{ki})$ よりも小さな値を代入することで規準が満たされるようにする。このような制約許容値の規準は、このあとで示されるアルゴリズム 3.7.6 のステップ (12) で使われている。

一方、上記 (iv) については、Lagrange 関数に対して Armijo の規準と Wolfe の規準が満たされるようにステップサイズ $\|\mathbf{y}_g\|_X$ (すなわち c_a) を決定する方法が考えられる。大域的収束性を保証する根拠となった定理 3.4.7 では、制約なしの問題を対象にしていた。ここでは、上記 (ii) と (iii) により、 \mathbf{x}_k と $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{y}_g$ において Lagrange 乗数と不等式制約に対する KKT 条件が満たされていると仮定する (あとで示されるアルゴリズム 3.7.6 では、ステップサイズが調整されたあとで Lagrange 乗数と不等式制約に対する KKT 条件が満たされるようになっている)。このとき、Lagrange 関数は f_0 と一致し、制約なし問題に対する Armijo の規準と Wolfe の規準を適用することが可能となる。実際に、それらの規準を示しておこう。原問題 (問題 3.1.1) の Lagrange 関数を式 (3.7.23) とおく。 $f_0(\mathbf{x}_k), f_{i_1}(\mathbf{x}_k), \dots, f_{i_{|I_A|}}(\mathbf{x}_k)$ の勾配を $\mathbf{g}_0(\mathbf{x}_k), \mathbf{g}_{i_1}(\mathbf{x}_k), \dots, \mathbf{g}_{i_{|I_A|}}(\mathbf{x}_k)$ とかき、 $f_0(\mathbf{x}_k + \mathbf{y}_g), f_{i_1}(\mathbf{x}_k + \mathbf{y}_g), \dots, f_{i_{|I_A|}}(\mathbf{x}_k + \mathbf{y}_g)$ の勾配を $\mathbf{g}_0(\mathbf{x}_k + \mathbf{y}_g), \mathbf{g}_{i_1}(\mathbf{x}_k + \mathbf{y}_g), \dots, \mathbf{g}_{i_{|I_A|}}(\mathbf{x}_k + \mathbf{y}_g)$ とかくことにする。このとき、Armijo の規準は $\xi \in (0, 1)$ に対して

$$\begin{aligned} & \mathcal{L}(\mathbf{x}_k + \mathbf{y}_g, \boldsymbol{\lambda}_{k+1}) - \mathcal{L}(\mathbf{x}_k, \boldsymbol{\lambda}_k) \\ & \leq \xi \left(\mathbf{g}_0(\mathbf{x}_k) + \sum_{i \in I_A(\mathbf{x}_k)} \lambda_{ki} \mathbf{g}_i(\mathbf{x}_k) \right) \cdot \mathbf{y}_g \end{aligned} \quad (3.7.26)$$

となる。また、Wolfe の規準は μ ($0 < \xi < \mu < 1$) に対して

$$\begin{aligned} & \mu \left(\mathbf{g}_0(\mathbf{x}_k) + \sum_{i \in I_A(\mathbf{x}_k)} \lambda_{ki} \mathbf{g}_i(\mathbf{x}_k) \right) \cdot \mathbf{y}_g \\ & \leq \left(\mathbf{g}_0(\mathbf{x}_k + \mathbf{y}_g) + \sum_{i \in I_A(\mathbf{x}_{k+1})} \lambda_{i, k+1} \mathbf{g}_i(\mathbf{x}_k + \mathbf{y}_g) \right) \cdot \mathbf{y}_g \end{aligned} \quad (3.7.27)$$

で与えられる。これらの規準は、このあとで示されるアルゴリズム 3.7.6 のステップ (8) と (10) で使われている。

以上のような方法を組み込んだアルゴリズムの一例を次に示す。図 3.17 にその流れ図を示す。

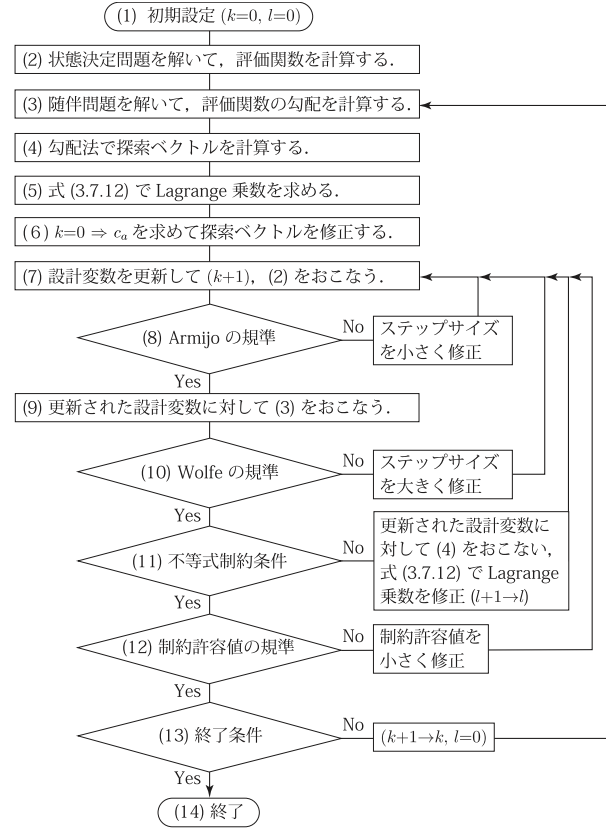


図 3.17: パラメータ調整あり制約つき問題に対する勾配法のアルゴリズム

アルゴリズム 3.7.6 (パラメータ調整あり制約つき問題に対する勾配法) 問題 3.1.1 の極小点を次のようにして求める。

- (1) 初期点 \mathbf{x}_0 を $f_1(\mathbf{x}_0) \leq 0, \dots, f_m(\mathbf{x}_0) \leq 0$ が満たされるように定める。式 (3.7.10) の正定値行列 \mathbf{A} と初期ステップサイズ ϵ_g (あるいは目的関数の減少率 α)、 f_0 の収束判定値 ϵ_0 、 f_1, \dots, f_m の許容範囲の初期値 $\epsilon_1, \dots, \epsilon_m$ 、Armijo と Wolfe の規準値 ξ と μ ($0 < \xi < \mu < 1$) および制約許容値の規準値 σ ($\ll 1$) を定める。また、 $c_a = 1$ 、 $k = 0$ および $l = 0$ とおく。
- (2) 状態決定問題を解いて、 \mathbf{x}_k において $f_0(\mathbf{x}_k), f_1(\mathbf{x}_k), \dots, f_m(\mathbf{x}_k)$ を計算する。また、

$$I_A(\mathbf{x}_k) = \{i \in \{1, \dots, m\} \mid f_i(\mathbf{x}_k) \geq -\epsilon_i\}$$

とおく。

- (3) $f_0, f_{i_1}, \dots, f_{i_{|I_A|}}$ に対する随伴問題を解いて, \mathbf{x}_k において $\mathbf{g}_0, \mathbf{g}_{i_1}, \dots, \mathbf{g}_{i_{|I_A|}}$ を計算する.
- (4) 式 (3.7.10) で $\mathbf{y}_{g_0}, \mathbf{y}_{g_{i_1}}, \dots, \mathbf{y}_{g_{i_{|I_A|}}}$ を計算する.
- (5) 式 (3.7.12) で $\lambda_{k+1} = \lambda_{k+1l}$ を求める. 式 (3.7.9) の $I_1(\mathbf{x}_k)$ が \emptyset ではないときには, $I_A(\mathbf{x}_k) \setminus I_1(\mathbf{x}_k)$ を $I_A(\mathbf{x}_k)$ とおきなおし, 式 (3.7.12) を再度解く.
- (6) 式 (3.7.11) で \mathbf{y}_g を求める. $k=0$ のとき, $\mathbf{y}_g = \bar{\mathbf{y}}_g$ とおき, 式 (3.3.7) (あるいは式 (3.3.9)) で c_a を求める. また, $i \in I_A(\mathbf{x}_k)$ に対して $\bar{\mathbf{y}}_{gi}/c_a$ を \mathbf{y}_{gi} に代入する.
- (7) $\mathbf{x}_{k+1l} = \mathbf{x}_k + \mathbf{y}_g(\lambda_{k+1l})$ とおき, 状態決定問題を解いて, $f_0(\mathbf{x}_{k+1l}), f_1(\mathbf{x}_{k+1l}), \dots, f_m(\mathbf{x}_{k+1l})$ を計算する. また,

$$I_A(\mathbf{x}_{k+1}) = \{i \in \{1, \dots, m\} \mid f_i(\mathbf{x}_{k+1l}) \geq -\epsilon_i\}$$

とおく.

- (8) $\lambda_{k+1} = \lambda_{k+1l}$ とおき, Armijo の規準 (式 (3.7.26)) を判定する.
- 満たされていれば, 次に進む.
 - そうではないとき, $\alpha > 1$ として αc_a を c_a に代入し, $\mathbf{y}_{g_0}/c_a, \mathbf{y}_{g_{i_1}}/c_a, \dots, \mathbf{y}_{g_{i_{|I_A|}}}/c_a$ を $\mathbf{y}_{g_0}, \mathbf{y}_{g_{i_1}}, \dots, \mathbf{y}_{g_{i_{|I_A|}}}$ に代入し, (7) に戻る.
- (9) \mathbf{x}_{k+1} において $\mathbf{g}_0, \mathbf{g}_{i_1}, \dots, \mathbf{g}_{i_{|I_A|}}$ を計算する.
- (10) $\lambda_{k+1} = \lambda_{k+1l}$ とおき, Wolfe の規準 (式 (3.7.27)) を判定する.
- 満たされていれば, 次に進む.
 - そうではないとき, $\beta \in (0, 1)$ として βc_a を c_a に代入し, $\beta \mathbf{y}_{g_0}, \beta \mathbf{y}_{g_{i_1}}, \dots, \beta \mathbf{y}_{g_{i_{|I_A|}}}$ を $\mathbf{y}_{g_0}, \mathbf{y}_{g_{i_1}}, \dots, \mathbf{y}_{g_{i_{|I_A|}}}$ に代入し, (7) に戻る.
- (11) すべての $i \in I_A(\mathbf{x}_{k+1})$ に対して, $|f_i(\mathbf{x}_{k+1})| \leq \epsilon_i$ を判定する.
- 満たされていれば, 次に進む.
 - そうではないとき, \mathbf{x}_{k+1} において $\mathbf{g}_0, \mathbf{g}_{i_1}, \dots, \mathbf{g}_{i_{|I_A|}}$ を計算し, 式 (3.7.10) で $\mathbf{y}_{g_0}, \mathbf{y}_{g_{i_1}}, \dots, \mathbf{y}_{g_{i_{|I_A|}}}$ を計算し, 式 (3.7.21) で $\delta \lambda$ を求め, $\lambda_{k+1l+1} = \lambda_{k+1l} + \delta \lambda$ とおく. $l+1$ を l に代入して (7) に戻る.
- (12) すべての $i \in I_A(\mathbf{x}_{k+1})$ に対して, 制約許容値の規準 (式 (3.7.25)) を判定する.
- 満たされていれば, 次に進む.

- そうではないとき, 満たされない i に対して, $\beta \in (0, 1)$ として $\beta \sigma \epsilon_0 / (|I_A(\mathbf{x}_{k+1})| \lambda_{i, k+1})$ を ϵ_i に代入し, (7) に戻る.

(13) 終了条件 $|f_0(\mathbf{x}_{k+1}) - f_0(\mathbf{x}_k)| \leq \epsilon_0$ を判定する.

- 終了条件が満たされたとき, (14) に進む.
- そうではないとき, $k+1$ を k に代入し, $l=0$ とおき, (3) に戻る.

(14) 計算を終了する.

□

3.8 制約つき問題に対する Newton 法

評価関数の \mathbf{x} の変動に対する Hesse 行列が得られる場合には, 勾配法を Newton 法に変更することができる. その方法を制約つき問題に対する Newton 法とよぶことにする. ここでは, f_0, f_1, \dots, f_m の Hesse 行列を $\mathbf{H}_0, \mathbf{H}_1, \dots, \mathbf{H}_m$ とかくことにする. また, 混乱がないときには $I_A(\mathbf{x}_k)$ を I_A とかくことにする.

制約つき問題に対する勾配法 (問題 3.7.1) に対して定義された式 (3.7.3) の Lagrange 関数 \mathcal{L}_Q を, ここでは,

$$\begin{aligned} \mathcal{L}_Q(\mathbf{y}, \boldsymbol{\lambda}_{k+1}) &= \frac{1}{2} \mathbf{y} \cdot (\mathbf{H}_0(\mathbf{x}_k) \mathbf{y}) + \mathbf{g}_0(\mathbf{x}_k) \cdot \mathbf{y} + f_0(\mathbf{x}_k) \\ &\quad + \sum_{i \in I_A(\mathbf{x}_k)} \left\{ \lambda_{i, k+1} (f_i + \mathbf{g}_i(\mathbf{x}_k) \cdot \mathbf{y}) + \lambda_{ik} \frac{1}{2} \mathbf{y} \cdot (\mathbf{H}_i(\mathbf{x}_k) \mathbf{y}) \right\} \end{aligned} \quad (3.8.1)$$

におきかえる. ここで, $\boldsymbol{\lambda}_k = (\lambda_{ik})_i$ は前のステップでえられた Lagrange 乗数とする. $k=0$ のときには, 制約つき問題に対する勾配法で使われた方法で決定されると仮定する. この \mathcal{L}_Q は, 次の問題の Lagrange 関数となっている.

問題 3.8.1 (制約つき問題に対する Newton 法) 問題 3.1.1 の試行点 $\mathbf{x}_k \in X$ において, Lagrange 乗数 $\boldsymbol{\lambda}_k \in \mathbb{R}^{|I_A|}$ は式 (3.7.5) から式 (3.7.7) (ただし, $k+1$ を k とみなす) を満たすとする. また, $f_0(\mathbf{x}_k), f_{i_1}(\mathbf{x}_k) = 0, \dots, f_{i_{|I_A|}}(\mathbf{x}_k) = 0$ および $\mathbf{g}_0(\mathbf{x}_k), \mathbf{g}_{i_1}(\mathbf{x}_k), \dots, \mathbf{g}_{i_{|I_A|}}(\mathbf{x}_k)$ と $\mathbf{H}_0(\mathbf{x}_k), \mathbf{H}_{i_1}(\mathbf{x}_k), \dots, \mathbf{H}_{i_{|I_A|}}(\mathbf{x}_k)$ を既知として,

$$\mathbf{H}_{\mathcal{L}}(\mathbf{x}_k) = \mathbf{H}_0(\mathbf{x}_k) + \sum_{i \in I_A(\mathbf{x}_k)} \lambda_{ik} \mathbf{H}_i(\mathbf{x}_k) \quad (3.8.2)$$

とおく. このとき,

$$q(\mathbf{y}_g) = \min_{\mathbf{y} \in X} \left\{ q(\mathbf{y}) = \frac{1}{2} \mathbf{y} \cdot (\mathbf{H}_{\mathcal{L}}(\mathbf{x}_k) \mathbf{y}) + \mathbf{g}_0(\mathbf{x}_k) \cdot \mathbf{y} + f_0(\mathbf{x}_k) \mid \right. \\ \left. f_i(\mathbf{x}_k) + \mathbf{g}_i(\mathbf{x}_k) \cdot \mathbf{y} \leq 0 \text{ for } i \in I_A(\mathbf{x}_k) \right\}$$

を満たす $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{y}_g$ を求めよ. \square

問題 3.8.1 は, 2 次最適化問題に分類される. $\mathbf{H}_{\mathcal{L}}(\mathbf{x}_k)$ は正定値実対称行列とは限らないが, そのような性質が満たされたときは, 問題 3.8.1 は凸最適化問題となる. ここでも, この問題に対する KKT 条件を用いて, \mathbf{y}_g をみつける方法について考えてみよう.

問題 3.8.1 の最小点 \mathbf{y}_g における KKT 条件は,

$$\mathbf{H}_{\mathcal{L}}(\mathbf{x}_k) \mathbf{y}_g + \mathbf{g}_0(\mathbf{x}_k) + \sum_{i \in I_A(\mathbf{x}_k)} \lambda_{i,k+1} \mathbf{g}_i(\mathbf{x}_k) = \mathbf{0}_{X'}, \quad (3.8.3)$$

$$f_i(\mathbf{x}_{k+1}) = f_i(\mathbf{x}_k) + \mathbf{g}_i(\mathbf{x}_k) \cdot \mathbf{y}_g \leq 0 \quad \text{for } i \in I_A(\mathbf{x}_k) \quad (3.8.4)$$

$$\lambda_{i,k+1} (f_i(\mathbf{x}_k) + \mathbf{g}_i(\mathbf{x}_k) \cdot \mathbf{y}_g) = 0 \quad \text{for } i \in I_A(\mathbf{x}_k), \quad (3.8.5)$$

$$\lambda_{i,k+1} \geq 0 \quad \text{for } i \in I_A(\mathbf{x}_k) \quad (3.8.6)$$

となる. これらを満たす $(\mathbf{y}_g, \boldsymbol{\lambda}_{k+1}) \in X \times \mathbb{R}^{|I_A|}$ は次のようにして求められる. すべての $i \in I_A(\mathbf{x}_k)$ に対して, 不等式制約が有効であると仮定すれば, 式 (3.8.3) と式 (3.8.4) は

$$\begin{pmatrix} \mathbf{H}_{\mathcal{L}} & \mathbf{G}^\top \\ \mathbf{G} & \mathbf{0}_{\mathbb{R}^{|I_A|} \times |I_A|} \end{pmatrix} \begin{pmatrix} \mathbf{y}_g \\ \boldsymbol{\lambda}_{k+1} \end{pmatrix} = - \begin{pmatrix} \mathbf{g}_0 \\ (f_i)_{i \in I_A} \end{pmatrix} \quad (3.8.7)$$

となる. ただし,

$$\mathbf{G}^\top = (\mathbf{g}_{i_1}, \dots, \mathbf{g}_{i_{|I_A|}}).$$

とする. $\mathbf{g}_{i_1}, \dots, \mathbf{g}_{i_{|I_A|}}$ が 1 次独立で, かつ $\mathbf{H}_{\mathcal{L}}$ が正則ならば, 式 (3.8.7) は $(\mathbf{y}_g, \boldsymbol{\lambda}_{k+1})$ について可解となる. この連立 1 次方程式の解に対して,

$$I_I(\mathbf{x}_k) = \{i \in I_A(\mathbf{x}_k) \mid \lambda_{i,k+1} < 0\} \quad (3.8.8)$$

を無効な制約条件の集合とおき, $I_I(\mathbf{x}_k) \neq \emptyset$ のときには $I_A(\mathbf{x}_k) \setminus I_I(\mathbf{x}_k)$ を $I_A(\mathbf{x}_k)$ とおきなおし, 式 (3.8.7) を再度解くことにする. このようにしてえられた $(\mathbf{y}_g, \boldsymbol{\lambda}_{k+1}) \in X \times \mathbb{R}^{|I_A|}$ は, 式 (3.8.3) から式 (3.8.6) を満たすことになる.

また、制約つき問題に対する勾配法 (3.7 節) においても示されたように、式 (3.8.7) の連立 1 次方程式を直接解く代わりに、次のような方法も考えられる。 $i \in I_A(x_k)$ ごとに g_i に対して

$$y_{gi} = -H_{\mathcal{L}}^{-1} g_i \quad (3.8.9)$$

が満たされるように $y_{g0}, y_{gi_1}, \dots, y_{gi_{|I_A|}}$ を求める。さらに、式 (3.7.12) で λ_{k+1} を求める。ただし、 $I_I(x_k)$ が \emptyset でないときには $I_A(x_k) \setminus I_I(x_k)$ を $I_A(x_k)$ とおきなおし、式 (3.7.8) を再度解くことにする。それらの結果を用いれば、 y_g は式 (3.7.11) によって求められることになる。

制約つき問題に対する勾配法とこの方法の違いは $c_a A$ が $H_{\mathcal{L}}$ におきかえられただけである。しかし、この方法でえられた探索ベクトル y_g は、評価関数の Hesse 行列を用いていることから、注意 3.5.4 であげた Newton 法の性質が成り立つことが期待される。ただし、次のことに注意する必要がある。

注意 3.8.2 (制約つき問題に対する Newton 法) 問題 3.8.1 の評価関数 q には、制約関数の Hesse 行列が前のステップの Lagrange 乗数が乗じられて加えられている。その結果、制約条件には Hesse 行列が使われていない。このことから、制約関数の非線形性がつよく、KKT 条件を満たす Lagrange 乗数がおおきく変化する問題に対しては、ステップサイズを小さくしなければ、収束しない場合がある。 \square

さらに、評価関数の 2 階微分が Hesse 勾配によって得られる場合の Newton 法は次のようになる。問題 3.8.1 は次の問題に置き換えられる。

問題 3.8.3 (Hesse 勾配を用いた Newton 法) $X = \mathbb{R}^d$ とする。 $A \in \mathbb{R}^{d \times d}$ と c_a を正定値実対称行列と正の定数とする。問題 3.1.1 の試行点 $x_k \in X$ において、Lagrange 乗数 $\lambda_k \in \mathbb{R}^{|I_A|}$ は式 (3.7.5) から式 (3.7.7) (ただし、 $k+1$ を k とみなす) を満たすとする。また、 $f_0(x_k), f_{i_1}(x_k) = 0, \dots, f_{i_{|I_A|}}(x_k) = 0$ および勾配 $g_0(x_k), g_{i_1}(x_k), \dots, g_{i_{|I_A|}}(x_k)$, 探索ベクトル \bar{y}_g および Hesse 勾配 $g_{H0}(x_k, \bar{y}_g), g_{Hi_1}(x_k, \bar{y}_g), \dots, g_{Hi_{|I_A|}}(x_k, \bar{y}_g)$ を既知として、

$$g_{H\mathcal{L}}(x_k, \bar{y}_g) = g_{H0}(x_k, \bar{y}_g) + \sum_{i \in I_A(x_k)} \lambda_{ik} g_{Hi}(x_k, \bar{y}_g) \quad (3.8.10)$$

とおく。このとき、

$$q(y_g) = \min_{y \in X} \left\{ q(y) = \frac{1}{2} y \cdot (c_a A y) + (g_0(x_k) + g_{H\mathcal{L}}(x_k, \bar{y}_g)) \cdot y \right. \\ \left. + f_0(x_k) \mid f_i(x_k) + g_i(x_k) \cdot y \leq 0 \text{ for } i \in I_A(x_k) \right\}$$

を満たす $x_{k+1} = x_k + y_g$ を求めよ。 \square

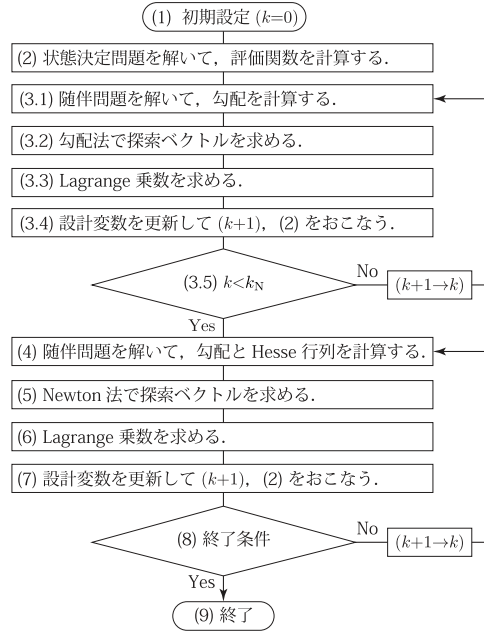


図 3.18: 制約つき問題に対する Newton 法のアルゴリズム

問題 3.8.3 の解を求めるために、評価関数ごとに求めた探索ベクトルを求める方法を採用する場合には、式 (3.8.9) を

$$\mathbf{y}_{gi} = -(\mathbf{c}_a \mathbf{A})^{-1} (\mathbf{g}_i + \mathbf{g}_{Hi}) \quad (3.8.11)$$

におきかえることで Newton 法と同じアルゴリズムが利用可能である。

3.8.1 簡単なアルゴリズム

上でみてきたことをふまえて、制約つき問題に対する Newton 法の考え方による簡単なアルゴリズムについて考えてみよう。図 3.18 にその流れ図を示す。ここでも、 $i \in I_A(\mathbf{x}_k)$ ごとに \mathbf{g}_i を用いて式 (3.8.9) で \mathbf{y}_{gi} を求める方法を用いることにする。

アルゴリズム 3.8.4 (制約つき問題に対する Newton 法) 問題 3.1.1 の極小点を次のようにして求める。

- (1) 初期点 \mathbf{x}_0 を $f_1(\mathbf{x}_0) \leq 0, \dots, f_m(\mathbf{x}_0) \leq 0$ が満たされるように定める。 f_0 の収束判定に用いる収束判定に用いる正の定数 ϵ_0 および f_1, \dots, f_m の許容範囲を与える正の定数 $\epsilon_1, \dots, \epsilon_m$ を定める。また、Newton 法を開始するまでの繰り返し数 k_N を定める。 $k = 0$ とおく。

(2) 状態決定問題を解いて, $f_0(\mathbf{x}_k), f_1(\mathbf{x}_k), \dots, f_m(\mathbf{x}_k)$ を計算する. また,

$$I_A(\mathbf{x}_k) = \{i \in \{1, \dots, m\} \mid f_i(\mathbf{x}_k) \geq -\epsilon_i\}$$

とおく.

(3) $k < k_N$ のとき次をおこなう.

(3.1) $f_0, f_{i_1}, \dots, f_{i_{|I_A|}}$ に対する随伴問題を解いて, \mathbf{x}_k において $\mathbf{g}_0, \mathbf{g}_{i_1}, \dots, \mathbf{g}_{i_{|I_A|}}$ を計算する.

(3.2) 式 (3.7.10) で $\mathbf{y}_{g_0}, \mathbf{y}_{g_{i_1}}, \dots, \mathbf{y}_{g_{i_{|I_A|}}}$ を計算する.

(3.3) 式 (3.7.12) で λ_{k+1} を求める. 式 (3.7.9) の $I_I(\mathbf{x}_k)$ が \emptyset ではないときには, $I_A(\mathbf{x}_k) \setminus I_I(\mathbf{x}_k)$ を $I_A(\mathbf{x}_k)$ とおきなおし, 式 (3.7.12) を再度解く.

(3.4) 式 (3.7.11) で \mathbf{y}_g を求め, $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{y}_g$ とおき, 状態決定問題を解いて, $f_0(\mathbf{x}_{k+1}), f_1(\mathbf{x}_{k+1}), \dots, f_m(\mathbf{x}_{k+1})$ を計算する. また,

$$I_A(\mathbf{x}_{k+1}) = \{i \in \{1, \dots, m\} \mid f_i(\mathbf{x}_{k+1}) \geq -\epsilon_i\}$$

とおく.

(3.5) $k+1$ を k に代入する. $k < k_N$ のとき (3.1) に戻る. そうでないとき, (4) に進む.

(4) $f_0, f_{i_1}, \dots, f_{i_{|I_A|}}$ に対する随伴問題を解いて, \mathbf{x}_k において, $\mathbf{g}_0, \mathbf{g}_{i_1}, \dots, \mathbf{g}_{i_{|I_A|}}$ および $\mathbf{H}_0, \mathbf{H}_{i_1}, \dots, \mathbf{H}_{i_{|I_A|}}$ を計算する.

(5) 式 (3.8.9) で $\mathbf{y}_{g_0}, \mathbf{y}_{g_{i_1}}, \dots, \mathbf{y}_{g_{i_{|I_A|}}}$ を計算する.

(6) 式 (3.7.12) で λ_{k+1} を求める. 式 (3.7.9) の $I_I(\mathbf{x}_k)$ が \emptyset ではないときには, $I_A(\mathbf{x}_k) \setminus I_I(\mathbf{x}_k)$ を $I_A(\mathbf{x}_k)$ とおきなおし, 式 (3.7.12) を再度解く.

(7) 式 (3.7.11) で \mathbf{y}_g を求め, $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{y}_g$ とおき, 状態決定問題を解いて, $f_0(\mathbf{x}_{k+1}), f_1(\mathbf{x}_{k+1}), \dots, f_m(\mathbf{x}_{k+1})$ を計算する. また,

$$I_A(\mathbf{x}_{k+1}) = \{i \in \{1, \dots, m\} \mid f_i(\mathbf{x}_{k+1}) \geq -\epsilon_i\}$$

とおく.

(8) 終了条件 $|f_0(\mathbf{x}_{k+1}) - f_0(\mathbf{x}_k)| \leq \epsilon_0$ を判定する.

- 終了条件が満たされたとき, (9) に進む.
- そうでないとき, $k+1$ を k に代入し, (4) に戻る.

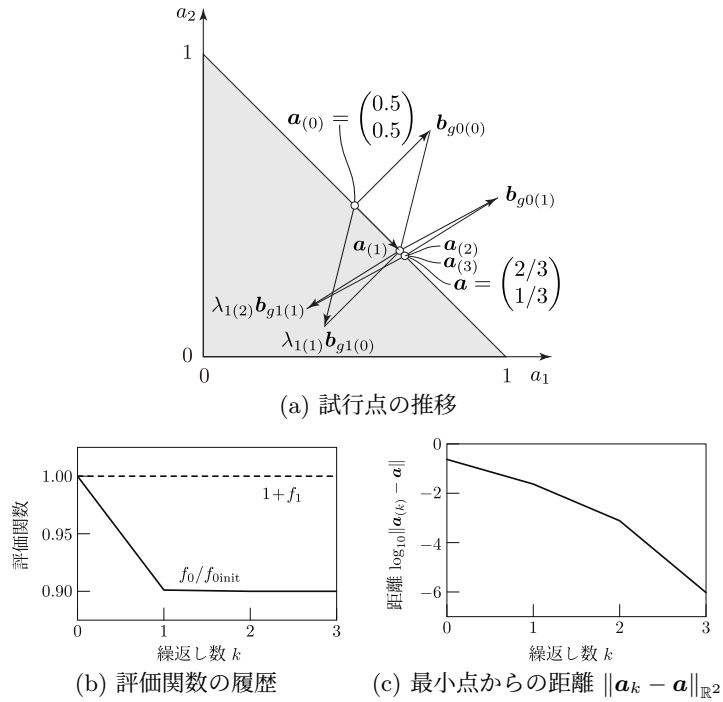


図 3.19: Newton 法による平均コンプライアンス最小化問題の数値例

(9) 計算を終了する。

□

アルゴリズム 3.8.4 を使って、第 1 章の例題 1.1.7 に対する試行点を求めてみよう。

例題 3.8.5 (平均コンプライアンス最小化問題の数値例) 例題 1.1.7 に対して、初期点を $\mathbf{a}_{(0)} = (1/2, 1/2)^\top$ とおき、アルゴリズム 3.8.4 を使って $k \in \{0, 1\}$ のときの試行点を求めよ。ただし、 k_N は可能な限り小さくとり、必要となる数値は適当に定めよ。

□

解答 平均コンプライアンス $\tilde{f}_0(\mathbf{a})$ と体積に対する制約関数 $f_1(\mathbf{a})$ は、それぞれ式 (3.7.13) と式 (3.7.14) で与えられる。また、それらの断面積微分は式 (3.7.15) と式 (3.7.16) で与えられる。 $\tilde{f}_0(\mathbf{a})$ の Hesse 行列は、

$$\mathbf{H}_0 = \begin{pmatrix} 8/a_1^3 & 0 \\ 0 & 2/a_2^3 \end{pmatrix} \quad (3.8.12)$$

となる．また， $\mathbf{H}_1 = \mathbf{0}_{\mathbb{R}^{2 \times 2}}$ より，式 (3.8.2) において λ_1 は不要となり， $k_N = 0$ とおくことができる．アルゴリズム 3.8.4 に沿って，数値を求めていく．ここでも，ステップ数 k のときの設計変数を $\mathbf{a}_{(k)}$ とかくことにする． $\mathbf{b}_{g0(k)}$ ， $\mathbf{b}_{g1(k)}$ および $\lambda_{1(k)}$ も同様とする．

- (1) 初期点 $\mathbf{a}_{(0)} = (1/2, 1/2)^\top$ では $f_1(\mathbf{a}_{(0)}) = 0$ は満たされる． $\epsilon_0 = 10^{-3}\tilde{f}_0(\mathbf{a}_{(0)})$ ， $\epsilon_1 = 10^{-6}$ とおく．また， $k = 0$ とおく．
- (2) 式 (3.7.13) と式 (3.7.14) より， $\tilde{f}_0(\mathbf{a}_{(0)}) = 10$ と $f_1(\mathbf{a}_{(0)}) = 0$ が得られる．また， $I_A(\mathbf{a}_{(0)}) = \{1\}$ とおく．
- (3) $k = k_N$ なので，先に進む．
- (4) 式 (3.7.15) と式 (3.7.16) より， $\mathbf{g}_{0(0)} = -(16, 4)^\top$ ， $\mathbf{g}_{1(0)} = (1, 1)^\top$ が得られる．
- (5) 式 (3.8.9) より， $\mathbf{b}_{g0(0)} = (1/4, 1/4)^\top$ ， $\mathbf{b}_{g1(0)} = -(1/64, 1/16)^\top$ が得られる．
- (6) 式 (3.7.12) で $\lambda_{1(1)} = 6.4$ が得られる．
- (7) 式 (3.7.11) で $\mathbf{b}_{g(0)} = (0.15, -0.15)^\top$ を得て， $\mathbf{a}_{(1)} = \mathbf{a}_{(0)} + \mathbf{b}_{g(0)} = (0.65, 0.35)^\top$ とおき， $\tilde{f}_0(\mathbf{a}_{(1)}) = 9.01099$ ， $f_1(\mathbf{a}_{(1)}) = 0$ が得られる．また， $I_A(\mathbf{a}_{(1)}) = \{1\}$ とおく．
- (8) $|\tilde{f}_0(\mathbf{a}_{(1)}) - \tilde{f}_0(\mathbf{a}_{(0)})| = 0.989011 \geq \epsilon_0 = 0.01$ より，終了条件は満たされないと判定し，1 を k に代入し，(4) に戻る．
- (4) 式 (3.7.15) と式 (3.7.16) より， $\mathbf{g}_{0(1)} = -(9.46746, 8.16327)^\top$ ， $\mathbf{g}_{1(1)} = (1, 1)^\top$ が得られる．
- (5) 式 (3.8.9) より， $\mathbf{b}_{g0(1)} = (0.325, 0.175)^\top$ ， $\mathbf{b}_{g1(1)} = -(0.0343281, 0.0214375)^\top$ が得られる．
- (6) 式 (3.7.12) で $\lambda_{1(2)} = 8.9661$ が得られる．
- (7) 式 (3.7.11) で $\mathbf{b}_{g(1)} = (0.0172107, -0.0172107)^\top$ を得て， $\mathbf{a}_{(2)} = \mathbf{a}_{(1)} + \mathbf{b}_{g(1)} = (0.667211, 0.332789)^\top$ とおき， $\tilde{f}_0(\mathbf{a}_{(2)}) = 9.00001$ ， $f_1(\mathbf{a}_{(2)}) = 1.11022 \times 10^{-16}$ が得られる．また， $I_A(\mathbf{a}_{(1)}) = \{1\}$ とおく．
- (8) $|\tilde{f}_0(\mathbf{a}_{(2)}) - \tilde{f}_0(\mathbf{a}_{(1)})| = 0.010977 \geq \epsilon_0 = 0.01$ より，終了条件は満たされないと判定し，2 を k に代入し，(4) に戻る．

図 3.19 にこれらとそのあとの結果を示す． $k = 3$ のとき， $|\tilde{f}_0(\mathbf{a}_{(3)}) - \tilde{f}_0(\mathbf{a}_{(2)})| = 0.0000119968 \leq \epsilon_0 = 0.01$ より，終了条件は満たされた．□

さらに，評価関数の 2 階微分が Hesse 勾配によって得られる場合のアルゴリズムは，アルゴリズム 3.8.4 のステップ (4) と (5) を次のように修正したものとなる．

- (4) \mathbf{x}_k において， $f_0, f_{i_1}, \dots, f_{i_{|I_A|}}$ に対する随伴問題を解いて， $\mathbf{g}_0, \mathbf{g}_{i_1}, \dots, \mathbf{g}_{i_{|I_A|}}$ を計算する．また， $f'_0, f'_{i_1}, \dots, f'_{i_{|I_A|}}$ に対する随伴問題を解いて， $\mathbf{g}_{H0}, \mathbf{g}_{Hi_1}, \dots, \mathbf{g}_{Hi_{|I_A|}}$ を計算する．
- (5) 式 (3.8.11) で $\mathbf{y}_{g0}, \mathbf{y}_{gi_1}, \dots, \mathbf{y}_{gi_{|I_A|}}$ を計算する．

図 3.20 と図 3.21 は，例題 1.1.7 に対して初期点を $\mathbf{a}_{(0)} = (1/2, 1/2)^\top$ とおいたときの f_0 の Hesse 勾配 \mathbf{g}_{H0} を用いた Newton 法によって得られた結果を勾配法

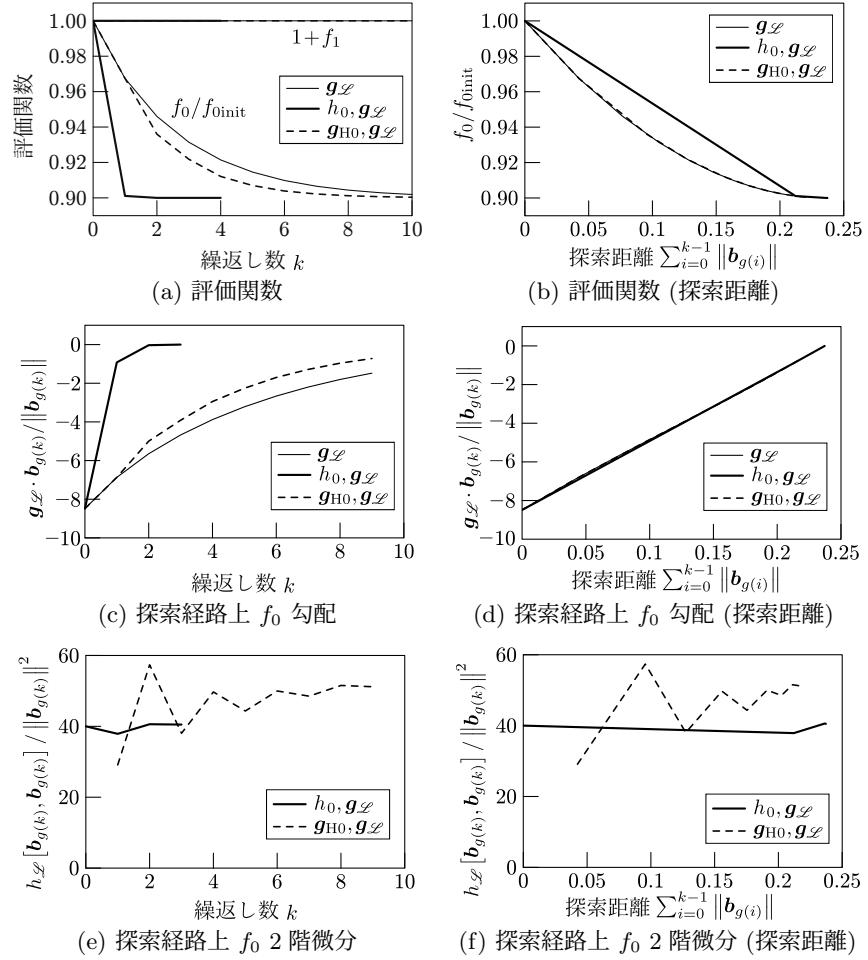


図 3.20: 平均コンプライアンス最小化に対する数値例: 評価関数と探索経路に沿った評価関数の勾配と2階微分 (g_L : H^1 勾配法, h_0, g_L : H^1 Newton 法, g_{H0}, g_L : Hesse 勾配を用いた H^1 Newton 法)

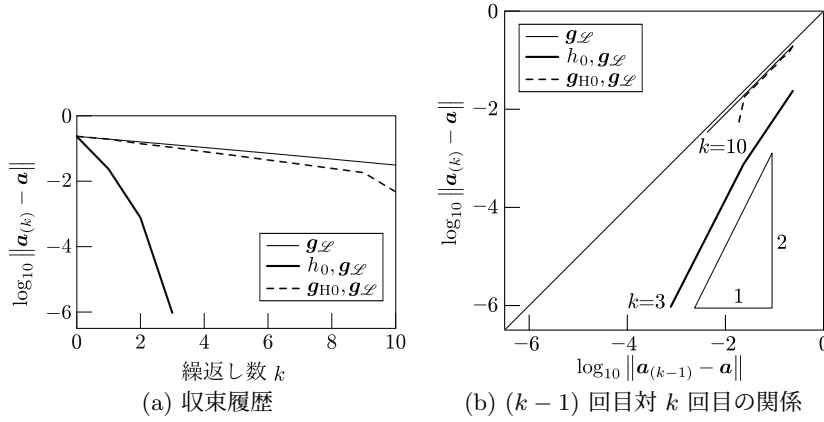


図 3.21: 平均コンプライアンス最小化に対する数値例: 最小点からの距離 $\|a_k - a\|_{\mathbb{R}^2}$ ($g_{\mathcal{L}}$: 勾配法, $h_0, g_{\mathcal{L}}$: Newton 法, $g_{H0}, g_{\mathcal{L}}$: Hesse 勾配を用いた Newton 法)

(例題 3.7.3) と Newton 法 (例題 3.8.5) の結果と合わせて示している. 勾配法では, $A = I$, $c_a = 200$ とおいた. Hesse 勾配を用いた Newton 法では, $A = I$ とおき, $k = 0$ の勾配法では $c_a = 200$ を用いて, $k \geq k_N = 1$ のときに $c_a = 100$ に変更した.

図 3.20 (a) は, 初期形状のときの f_0 を表す $f_{0\text{init}}$ と初期形状の体積を表す $c_1 = 1$ で標準化された評価関数 $f_0/f_{0\text{init}}$ と $1+f_1$ の繰返し数 k に対する変化を示している. 図 3.20 (b) は, それらの評価関数を, $X = \mathbb{R}^2$ 上の探索距離 $\sum_{i=0}^{k-1} \|b_{g(i)}\|_X$ に対して示している. また, 図 3.20 (c) と (d) は, 探索経路に沿って観測したときの f_0 の勾配 (すなわち Lagrange 関数 $\mathcal{L} = \mathcal{L}_0 + \lambda_1 f_1$ の勾配) $g_{\mathcal{L}} \cdot b_{g(k)} / \|b_{g(k)}\|_X$ の変化をそれぞれ繰返し数と探索距離に対して示している. さらに, 図 3.20 (e) と (f) は, 探索経路に沿って観測したときの f_0 の 2 階微分 $h_{\mathcal{L}}[b_{g(k)}, b_{g(k)}] / \|b_{g(k)}\|_X^2$ (Hesse 勾配を用いた Newton 法の場合は $(g_{H0} \cdot b_{g(k)} + \lambda_1 h_1[b_{g(k)}, b_{g(k)}]) / \|b_{g(k)}\|_X^2 = g_{H0} \cdot b_{g(k)} / \|b_{g(k)}\|_X^2$) の変化をそれぞれ繰返し数と探索距離に対して示している.

図 3.20 より, 繰返し数に対するグラフでは収束速度の違いにより異なった曲線になるが, 探索距離に対するグラフはほぼ同一の曲線になることが観察される. その理由は, 図 3.15 (a) や図 3.19 (a) に示されたように, 探索経路は同一であったためである. このようなグラフは, 第 8 章と第 9 章の数値例においても示される. そのときは, 探索経路は図にかけないが, このようなイメージであることを想像してほしい.

また, 図 3.21 (a) には, 3つの方法で得られた最小点からの距離 $\|a^{(k)} - a\|_X$ を繰返し数 k に対して示している. この図から, Hesse 勾配を用いた Newton 法の結果は, 収束次数が 1 次以上であることが確認される. さらに, 図 3.21 (b) は, 繰返

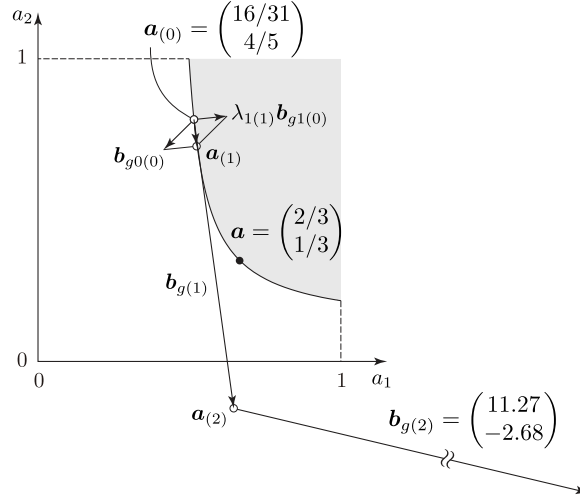


図 3.22: Newton 法による体積最小化問題の数値例

し数 $(k-1)$ 回目に対して k 回目の距離 $\|\mathbf{a}_k - \mathbf{a}\|_X$ をプロットしている．このグラフの勾配は収束次数を表す．なぜならば， $\|\mathbf{a}_{(k)} - \mathbf{a}\|_X = r \|\mathbf{a}_{(k-1)} - \mathbf{a}\|_X^p$ の関係を仮定したとき，

$$\log_{10} \|\mathbf{a}_{(k)} - \mathbf{a}\|_X = p \log_{10} \|\mathbf{a}_{(k-1)} - \mathbf{a}\|_X + \log_{10} r \quad (3.8.13)$$

が成り立つためである．この関係から，グラフの勾配が次数 p に対応し，対角線からのずれが $\log_{10} r$ に対応する．これらのグラフより，勾配法と Newton 法の収束次数はそれぞれ 1 次と 2 次になっていることが確認される．

このように，平均コンプライアンス最小化問題に対しては Newton 法が有効に機能することが確かめられた．

一方，例題 3.7.4 のような体積最小化問題に対して Newton 法をそのまま適用すれば，収束しない結果におちいる．図 3.22 に，アルゴリズム 3.8.4 によってえられる試行点の推移を示す．ただし， $k=0$ のときの Lagrange 乗数 $\lambda_{1(0)}$ には，例題 3.7.4 の $\lambda_{1(1)}$ が使われた．この問題では， f_0 の Hesse 行列は $\mathbf{0}_{\mathbb{R}^{2 \times 2}}$ で， f_1 の Hesse 行列が正定値となっている．すなわち，注意 3.8.2 で指摘された条件が成り立っている．このような場合にも Newton 法が使えるようにするためには，ステップサイズを調整する必要がある．

3.8.2 複雑なアルゴリズム

図 3.22 のような場合が起こりうることを考えると、ステップサイズを調整する機能をはじめ、3.7.2 節で説明された機能を追加する必要がある。それらの機能とアルゴリズムの関係は 3.7.2 項に示されている。それらをもう一度説明することは省略する。

また、Hesse 行列が正定値ではないような場合には、正定値行列を加えることで正定値化する方法や、負の固有値をもつ固有モードの成分のみをとりだして正定値化する方法などが知られている。さらに、正定値ではないときには勾配法を適用して、収束がちかづいてきてから Newton 法にきりかえる方法なども考えられる。

3.9 第3章のまとめ

第3章では、有限次元ベクトル空間上の非線形最適化問題に対して、極小点を求める方法についてみてきた。要点は以下のようである。

- (1) 非線形最適化問題の解法には基本的に反復法が使われる。反復法は、初期点を与えて、探索ベクトル (探索方向とステップサイズ) を適切に決めながら試行点を更新していく方法である (3.2 節)。
- (2) 制約なし最適化問題に対して、探索方向を決める代表的な方法は勾配法である。勾配法は、評価関数の設計変数に対する勾配を既知として、勾配とは逆向きに探索方向を決める方法である (3.3 節)。
- (3) 制約なし最適化問題に対して、ステップサイズの適切性を判定する規準として Armijo の規準と Wolfe の規準が知られている。それらの規準が満たされるようにステップサイズが決定された反復法は、大域的収束性をもつ (3.4 節)。
- (4) 制約なし最適化問題に対して、評価関数の勾配と Hesse 行列が計算可能なとき、Newton 法を用いれば探索方向とステップサイズを同時に決めることができる。Newton 法でえられる試行点は2次収束する。しかし、Hesse 行列の計算にコストがかかる (3.5 節)。
- (5) 不等式制約つき最適化問題の解法の一つとして、拡大関数法が知られている。拡大関数法は、制約関数に重みを表す定数をかけて目的関数に加えることで制約なし問題におきかえる方法である。しかし、それらの方法を用いるためには、単調数列を問題ごとに適切に選ぶ必要がある (3.6 節)。

- (6) 不等式制約つき最適化問題の解法には、KKT 条件を用いた解法が考えられる。すべての評価関数の勾配が計算可能であれば、制約つき問題に対する勾配法が使われる。この方法では、評価関数ごとの勾配とそれらを用いた勾配法によってえられる探索ベクトルで構成される行列を使って Lagrange 乗数が決定される。この関係は、実用的なアルゴリズムを考える際に有効に利用される (3.7 節)。
- (7) 不等式制約つき最適化問題において、評価関数の Hesse 行列が計算可能であれば、制約つき問題に対する Newton 法が使われる。この方法では、制約つき問題に対する勾配法で使われた正定値対称行列を Hesse 行列におきかえるだけで、勾配法と同様のアルゴリズムが使われる。この方法が有効に機能するためには、制約関数の非線形性がよわいことが条件になる (3.8 節)。

3.10 第 3 章の演習問題

- 3.1 非線形関数 $f \in C^2(\mathbb{R}; \mathbb{R})$ に対して、勾配 $g(x) = 0$ を満たす $x \in \mathbb{R}$ を求める問題を考える。 $k \in \{0, 1, 2, \dots\}$ に対して、 $x_k \in \mathbb{R}$ と $g(x_k)$ が与えられたとき、Newton-Raphson 法 (問題 3.5.6) によって x_{k+1} を求める式を示せ。また、 $g(x_k)$ を差分

$$\frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}$$

におきかえたときに、 x_{k+1} を求める式を示せ。この式はセカント法の公式を示す。

- 3.2 問題 3.4.1 (厳密直線探索法) を解くアルゴリズムとしてセカント法を用いることを考える。 $l \in \{0, 1, 2, \dots\}$ に対して $\bar{\epsilon}_{gl}$ が与えられたとき、 $\bar{\epsilon}_{gl+1}$ を求める式を示せ。
- 3.3 共役勾配法 (問題 3.4.10) の一例として示された式 (3.4.8) から式 (3.4.12) で計算される探索方向 \bar{y}_{gk+1} は \bar{y}_{gk} と共役になることを確かめよ。
- 3.4 演習問題 1.6 の評価関数 $f(\mathbf{a})$ (4 面体の体積) が最小となる設計変数 \mathbf{a} (2 辺の長さ) を求める問題を考える。設計変数の初期値 $\mathbf{a}_0 = (a_{01}, a_{02})^\top$ が与えられたとき、Newton 法で探索ベクトル \mathbf{b} を求めよ。

参考文献

- [1] Armijo, L. Minimization of functions having Lipschitz-continuous first partial derivatives. *Pacific J. Math.*, Vol. 16, pp. 1–3, 1966.
- [2] Murty, K. G. *Linear complementarity, linear and nonlinear programming*. Heldermann Verlag, Berlin, 1988.
- [3] Wolfe, P. Convergence conditions for ascent methods. *SIAM Review*, Vol. 11, pp. 226–235, 1969.
- [4] 矢部博. 工学基礎: 最適化とその応用. 数理工学社, 東京, 2006.