

9 非線形計画法 (I) : 制約のない問題の最適化法

9.1 数学的準備

まず, いくつかの基本的な概念と数学的表記を整理しておく.

非線形計画問題の一般形

制約条件 $x \in V$ をみたす x の中で, 実数値関数 $f(x)$ を最小にする x を求めよ.

ベクトル x のノルム

$$\|x\| = \sqrt{x^T x}$$

点 x の ϵ 近傍

$$B(x, \epsilon) \equiv \{y \in E^n : \|y - x\| < \epsilon\}$$

許容解

$$x \in V \subseteq E^n$$

極小解 (極小点, 局小解), 極小値

許容解 x^* に対して, 正数 δ が存在して, 任意の $x \in V \cap B(x^*, \delta)$ に対して, $f(x^*) \leq f(x)$ が成立するとき, x^* を極小解または極小点, 局小解, $f(x^*)$ を極小値とよぶ.

最小解 (最小点), 最小値

任意の $x \in V$ に対して, $f(x^*) \leq f(x)$ が成立するとき, x^* を最小解または最小点, $f(x^*)$ を最小値とよぶ.

関数 $f(x)$ の勾配ベクトル

$$\nabla f(x) \equiv (\partial f / \partial x_1, \dots, \partial f / \partial x_n)^T$$

停留点 (臨界点)

$\nabla f(x) = 0$ を満足する点 x . 極小値, 鞍点, 極大値が該当する.

関数 $f(x)$ の Hesse 行列

$f(x)$ が 2 回微分可能なとき, 2 階偏微分係数 $\partial^2 f / \partial x_i \partial x_j$ を (i, j) 要素とする正方行列を Hesse 行列とよび $\nabla^2 f(x)$ と記す.

非負定値行列, 正定値行列

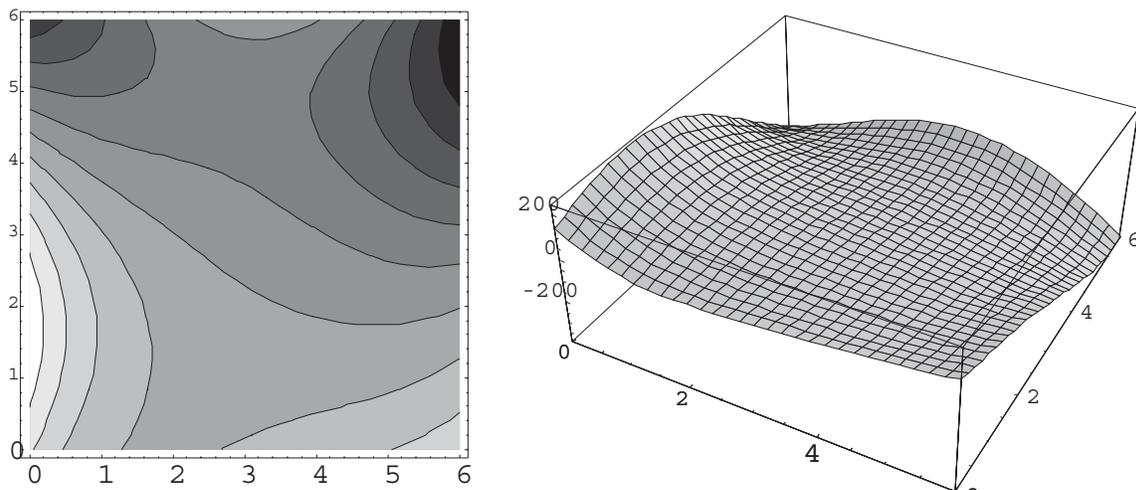
対称な n 次元正方行列 A が, 0 でない任意のベクトル $t \in R^n$ に対して, $t^T A t \geq 0$ のとき非負定値行列, $t^T A t > 0$ のとき正定値行列とよぶ.

9.2 制約なし最適化問題

非線形計画問題において, 制約が与えられないものを, 制約なし最適化問題 (unconstrained optimization problem) とよび, その最適解 x^* は以下の性質を満たす.

1 次の必要条件: x^* が極小解であるとき, $f \in C^1$ ならば $\nabla f(x^*) = 0$ である.

2 次の必要条件: x^* が極小解であるとき, $f \in C^2$ ならば $\nabla f(x^*) = 0$, かつ $\nabla^2 f(x^*)$ は非負定値である.



$$144 - 184x + 84x^2 - 15x^3 + x^4 + 92y - 61xy + 5x^2y - 10y^2 + 17xy^2 - 2x^2y^2 - 9y^3 + y^4$$

図 9.1: 多項式関数による等高線表示 (左) と立体的表示 (右)

2 次の十分条件 : $f \in C^2$ のとき x^* が $\nabla f(x^*) = 0$ を満たし $\nabla^2 f(x^*)$ が正定値ならば, x^* は制約なし非線形最適化問題の極小解 (孤立局所最適解) である .

これらの最適性条件によって, f がある条件を満足する場合には, $f(x)$ の極小解を探す, という問題が, $\nabla f(x) = 0$ というベクトル方程式を解く, という問題に置き換わったことになる . さて, それでは, $\nabla f(x) = 0$ を満足する x はどうやったら求まるのであろうか . 残念ながら $\nabla f(x) = 0$ という方程式もそうそう解析的には解くことができない . そこで, 通常は反復法 (iterative method) によって解を求めることになる .

9.3 反復法の概要

反復法は逐次的近似解法ともよばれ, 初期近似解 x_0 から出発して, 次に挙げる漸化式

$$x_{k+1} = x_k + \alpha_k d_k \quad k = 0, 1, \dots$$

によって, 次々に近似解を生成して最適解へと接近していく方法である . ここで, d_k は探索方向を表す方向ベクトル (direction vector), α_k は探索方向のステップ幅 (step size) を表す正の数である . 目的関数が $f(x_{k+1}) < f(x_k)$ を満足するように, 探索方向とステップ幅を定めて移動していくので, 降下法 (descent method) とよばれる . 降下法の中で, 探索方向を決定する時に, $\nabla f(x_k)$ を利用するものを特に勾配法 (gradient method) と呼ぶことがある . 図 9.1 に多項式関数の作る曲面を等高線表示したものと立体的表示をしたものを示す . 多項式のパラメータを少し変更するだけで, この曲面の様子は大きく変化する . 降下法を考える上では,

- 反復することでちゃんと最適解に到達できるのか? (大域的収束性 : global convergence property)
- 到達するまでの反復の回数は? (収束の速さ : speed of convergence)

が重要なポイントとなる .

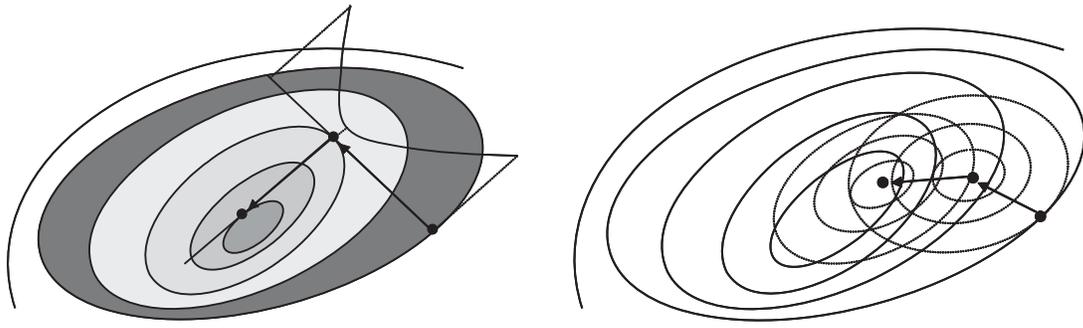


図 9.2: 最急降下法のイメージ (左) とニュートン法のイメージ (右)

ここでは典型的な降下法¹ について見ていることにする。

9.4 最急降下法

最急降下法 (steepest descent method) とは, 降下の方向ベクトルとして $\mathbf{d}_k = -(\nabla f(\mathbf{x}))$ をとる方法である。すなわち, 各反復において点 \mathbf{x}_k の下り勾配の方向に降下していく。ステップ幅の決定は, 降下の方向が決まっていると, 点 \mathbf{x}_k を起点として \mathbf{d}_k 方向での最小化問題となる。すなわち, $\phi(\alpha) = f(\mathbf{x}_k + \alpha \mathbf{d}_k)$ を $\alpha > 0$ の範囲で見つけるという問題が出てくる。これを直線探索とよび, 厳密に最適解を求めるもの (厳密直線探索), おおよその最適解でやめるもの (非厳密直線探索) など, 色々な方法がある。代表的な方法としては, 黄金分割 (golden section) を使った最小解を含む区間の縮小がある。

最急降下法によって最適解に収束していく様子 of 概念図を図 9.2(左) に示す。いまの点でもっとも傾斜している方向 (下り勾配方向: $-(\nabla f(\mathbf{x}_k))$) に降りていき, その方向において最も低くなった点で, 次の下り勾配方向 (先ほどの降下方向とは直交する) へと切り替えて進んでいく。2 次関数を例にとった場合に, 楕円が円に近い場合には, まっすぐに最適解へと向かうが, 扁平になると, ジグザグな経路で解へと収束していく。また, 最急降下法においては, 初期解の位置によって決定される直交する 2 方向 (Stiefel's cage) を交互に選択しながら収束していくため, 解空間の探索効率は高くない。

9.5 ニュートン法

最急降下法は点 \mathbf{x}_k での f の一階微分係数を用いていたが, f が 2 回連続微分可能であれば, より効率的に降下方向を決定することができる。その代表的な方法がニュートン法 (Newton's method) である。ニュートン法は降下方向 \mathbf{d}_k を

$$\nabla^2 f(\mathbf{x}_k) \mathbf{d}_k = -\nabla f(\mathbf{x}_k)$$

という連立一次方程式の解として求めてやり, ステップ幅を $\alpha_k \equiv 1$ として次のステップの近似解を生成する。

すなわち, 目的関数 f をある点 \mathbf{x}_k のまわりで局所的に 2 次関数近似

$$f(\mathbf{x}) \simeq f(\mathbf{x}_k) + \nabla f(\mathbf{x}_k)(\mathbf{x} - \mathbf{x}_k) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_k)^T \nabla^2 f(\mathbf{x}_k)(\mathbf{x} - \mathbf{x}_k)$$

とし, 右辺の 2 次関数の最小点を次の点 \mathbf{x}_{k+1} とする。

ニュートン法は二階微分係数を用いることにより, 非常によい解への収束を与えるが, ニュートン法が適用できる場合は限定される。すなわち, Hesse 行列 $\nabla^2 f(\mathbf{x}_k)$ が解析的に求まらない場合や, 正定値

¹ただし, 最急降下法と Newton 法は最適化法としては多くの場合実用的ではない。

でない場合には適用できない。また、適用できる場合においても、 $\nabla^2 f(\mathbf{x}_k)$ の計算に手間がかかる場合には適切ではない。

9.6 準ニュートン法

ニュートン法においては、 $\nabla^2 f(\mathbf{x}_k)$ の計算が問題となっていた。そこで、直接に $\nabla^2 f(\mathbf{x}_k)$ を計算せずに、 $\{\mathbf{x}_k\}$ 、 $\{\nabla f(\mathbf{x}_k)\}$ を使って、 $\nabla^2 f(\mathbf{x}_k)$ の近似行列の系列 $\{H_k\}$ を構成してやることを考える。こうしたアプローチを総称して準ニュートン法 (quasi-Newton method) と呼ぶ。

今、2 つのベクトルを

$$\mathbf{s}_k \equiv \mathbf{x}_{k+1} - \mathbf{x}_k, \quad \mathbf{y}_k \equiv \nabla f(\mathbf{x}_{k+1}) - \nabla f(\mathbf{x}_k)$$

と定義する。すると近似的には、 $\nabla^2 f(\mathbf{x}_k) \mathbf{s}_k \simeq \mathbf{y}_k$ が成り立つ。そこで、 \mathbf{s}_k と \mathbf{y}_k の組が得られるたびに、

$$H_{k+1} \mathbf{s}_k = \mathbf{y}_k$$

を満足するように、近似行列 H_k を順次、更新していく。 H_k の決め方は連立一次方程式が不定であるため、無数に考えられる。その中で最も優れているとされている決定方法は、C.G.Broyden, R. Fletcher, D. Goldfarb, D.F. Shanno による以下に示す BFGS 公式である。

$$H_{k+1} = H_k - \frac{H_k \mathbf{s}_k \mathbf{s}_k^T H_k}{\mathbf{s}_k^T H_k \mathbf{s}_k} + \frac{\mathbf{y}_k \mathbf{y}_k^T}{\mathbf{s}_k^T \mathbf{y}_k}$$

準ニュートン法は制約なし最適化問題において最もよく用いられる方法であるが、問題の次元 (変数の数) が多くなると、行列 H_k に大きな記憶領域が必要となってくる。

9.7 共役方向法

問題が大規模になり、準ニュートン法の適用が困難な場合に効果を発揮するのが、共役方向法 (conjugate direction method) である。この方法の探索方向のベクトルは、反復公式

$$\mathbf{d}_k = -\nabla f(\mathbf{x}_k) + \beta_k \mathbf{d}_{k-1}$$

によって生成される。ただし、

$$\beta_k = \frac{(\nabla f(\mathbf{x}_k))^T (\nabla f(\mathbf{x}_k) - \nabla f(\mathbf{x}_{k-1}))}{\|\nabla f(\mathbf{x}_{k-1})\|^2}, \quad k \geq 1$$

$$\beta_0 = 1$$

とする。これは、2 つの探索方向のベクトルの間に、 $\mathbf{d}_k^T \nabla^2 f(\mathbf{x}_k) \mathbf{d}_{k-1} = 0$ という共役関係が成り立つ。この方法では準ニュートン法のように行列を記憶しておく必要がないため、大規模な問題に適用することが可能である。

【参考図書】

- 今野浩，山下浩著：非線形計画法，日科技連出版社，ISBN4-8171-5306-7
- 藤田宏，今野浩，田邊國士：岩波講座応用数学 9 [方法 7] 最適化法，岩波書店