オペレーションズ・リサーチ II (2)

田中 俊二

shunji.tanaka@okayama-u.ac.jp

スケジュール

No.	内容
1	導入 (非線形最適化問題，ゲーム理論，多目的最適化問題)
2	非線形計画 1 (勾配，ヘッセ行列，凸性，最適性条件，ニュートン法)
3	非線形計画 2 (最急降下法，準ニュートン法，共役勾配法，信頼領域法)
4	非線形計画 3 (制約つき問題の最適性条件，KKT 条件，ペナルティ関数法，２次計画法，逐次 2 次計画法)
5	ゲーム理論 1 (種々のゲーム，標準形，純粋戦略，混合戦略，ナッシュ均衡)
6	ゲーム理論 2 (展開形ゲーム，繰り返しゲーム)
7	多目的最適化 (パレート最適性，重み付け法，ϵ 制約法，重み付きメトリック法)

制約なし非線形計画問題

制約なし非線形計画問題 (P0)

$$\begin{array}{ll} \text{min} & f(\boldsymbol{x}) \\ \text{s.t.} & \boldsymbol{x} \in \mathbb{R}^n \end{array}$$

仮定

$f(\boldsymbol{x})$ は C^1 **級**

C^n 級 (n 回連続的微分可能)

- n 回偏微分可能，かつすべての n 階偏導関数が連続
- C^∞ 級: 無限回偏微分可能

C^∞ 級関数の例

$f(\boldsymbol{x}) = e^{x_1+x_2}$ は,

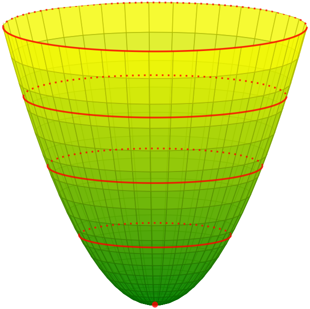
$$\frac{\partial f}{\partial x_1} = e^{x_1+x_2} = f(\boldsymbol{x}), \quad \frac{\partial f}{\partial x_2} = e^{x_1+x_2} = f(\boldsymbol{x})$$

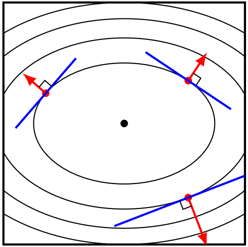
より，無限回偏微分可能

勾配 (復習)

勾配 (gradient)・勾配ベクトル (gradient vector)

$f(\boldsymbol{x})$ の偏微分係数を並べたベクトル $\nabla f(\boldsymbol{x}) = \begin{pmatrix} \frac{\partial}{\partial x_1} f(\boldsymbol{x}) \\ \vdots \\ \frac{\partial}{\partial x_n} f(\boldsymbol{x}) \end{pmatrix}$





等高線

- 等高線の拡大方向
- 等高線の接線に直交

局所的最適解 (極小点)

局所的最適解 (極小点)

$\boldsymbol{x}^{(0)}$ の近傍で $f(\boldsymbol{x}^*) \le f(\boldsymbol{x})$ が成り立つ $\Rightarrow$ $\boldsymbol{x}^*$ は (P0) の**局所的最適解**

局所的最適解の必要条件 (最適性の 1 次の必要条件)

$f(\boldsymbol{x})$ は C^1 級とする． $\boldsymbol{x}^*$ が (P0) の局所的最適解ならば $\nabla f(\boldsymbol{x}^*) = \boldsymbol{0}$

$\nabla f(\boldsymbol{x}^{(0)}) = \boldsymbol{0}$ を満たす点 $\boldsymbol{x}^{(0)}$ を $f(\boldsymbol{x})$ の**停留点 (臨界点)** という

必要条件なので，停留点が必ず局所的最適解になるわけではない．

$\nabla f(\boldsymbol{x}^{(0)}) = \boldsymbol{0}$ を満たす $\boldsymbol{x}^{(0)}$ が局所的最適解とならない例

$f(x_1, x_2) = -x_1^2 + x_2^2$ は $\nabla f(0,0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ を満たすが，任意の $\epsilon > 0$ に対して

$f(\epsilon, 0) < f(0,0) < f(0,\epsilon)$ となるので，原点 $(0,0)$ は局所的最適解ではない．

凸関数の場合 (十分条件)

$f(\boldsymbol{x})$ は C^1 級かつ凸関数であるとする．このとき，$f(\boldsymbol{x})$ の停留点は (P0) の大域的最適解．さらに，$f(\boldsymbol{x})$ が狭義凸関数ならば，大域的最適解は一意に定まる．

ヘッセ行列

ヘッセ行列 (Hessian matrix)

C^2 級関数 $f(\boldsymbol{x})$ の 2 階偏微分係数を並べた行列

$$\nabla^2 f(\boldsymbol{x}) = \begin{pmatrix} \frac{\partial^2}{\partial x_1^2} f(\boldsymbol{x}) & \frac{\partial^2}{\partial x_1 \partial x_2} f(\boldsymbol{x}) & \cdots & \frac{\partial^2}{\partial x_1 \partial x_n} f(\boldsymbol{x}) \\ \frac{\partial^2}{\partial x_2 \partial x_1} f(\boldsymbol{x}) & \frac{\partial^2}{\partial x_2^2} f(\boldsymbol{x}) & & \frac{\partial^2}{\partial x_2 \partial x_n} f(\boldsymbol{x}) \\ \vdots & & \ddots & \vdots \\ \frac{\partial^2}{\partial x_n \partial x_1} f(\boldsymbol{x}) & \frac{\partial^2}{\partial x_n \partial x_2} f(\boldsymbol{x}) & \cdots & \frac{\partial^2}{\partial x_n^2} f(\boldsymbol{x}) \end{pmatrix}$$

ただし， $\frac{\partial^2}{\partial x_i \partial x_j} f(\boldsymbol{x}) = \frac{\partial}{\partial x_i}\left(\frac{\partial}{\partial x_j} f(\boldsymbol{x})\right)$

C^2 級なら $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$ が成り立つので，ヘッセ行列は**対称行列**

テイラー展開とヘッセ行列

$\boldsymbol{x} = \boldsymbol{x}^{(0)}$ におけるテイラー展開 (1 次まで)

$$f(\boldsymbol{x}) = f(\boldsymbol{x}^{(0)}) + \nabla^\top f(\boldsymbol{x}^{(0)})(\boldsymbol{x} - \boldsymbol{x}^{(0)}) + R_2(\boldsymbol{x}; \boldsymbol{x}^{(0)})$$

$$R_2(\boldsymbol{x}; \boldsymbol{x}^{(0)}) = \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}^{(0)})^\top \nabla^2 f(\theta\boldsymbol{x} + (1-\theta)\boldsymbol{x}^{(0)})(\boldsymbol{x} - \boldsymbol{x}^{(0)}) = o(\|\boldsymbol{x} - \boldsymbol{x}^{(0)}\|)$$

ただし，$0 < \theta < 1$

$\boldsymbol{x} = \boldsymbol{x}^{(0)}$ におけるテイラー展開 (2 次まで)

$$f(\boldsymbol{x}) = f(\boldsymbol{x}^{(0)}) + \nabla^\top f(\boldsymbol{x}^{(0)})(\boldsymbol{x} - \boldsymbol{x}^{(0)}) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}^{(0)})^\top \nabla^2 f(\boldsymbol{x}^{(0)})(\boldsymbol{x} - \boldsymbol{x}^{(0)}) + R_3(\boldsymbol{x}; \boldsymbol{x}^{(0)})$$

$$R_3(\boldsymbol{x}; \boldsymbol{x}^{(0)}) = o(\|\boldsymbol{x} - \boldsymbol{x}^{(0)}\|^2)$$

ランダウ (Landau) の漸近記法

$f(\boldsymbol{x}) = o(\|\boldsymbol{x} - \boldsymbol{x}^{(0)}\|^k)$ の意味： $\displaystyle\lim_{\boldsymbol{x}\to\boldsymbol{x}^{(0)}} \frac{|f(\boldsymbol{x})|}{\|\boldsymbol{x} - \boldsymbol{x}^{(0)}\|^k} = 0$ $\quad \left(\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x}^t\boldsymbol{x}} = \sqrt{\sum_{i=1}^{n} x_i^2} \right)$

$\boldsymbol{x}$ が $\boldsymbol{x}^{(0)}$ に近づくとき，$f(\boldsymbol{x})$ は $\|\boldsymbol{x} - \boldsymbol{x}^{(0)}\|^k$ より早く 0 に収束

$o(x)$ は「リトルオー」「スモールオー」と読む．$O(x)$ (ビッグオー) もよく使われる

テイラー展開とヘッセ行列

$\boldsymbol{x} = \boldsymbol{x}^{(0)}$ におけるテイラー展開 (1 次まで)

$$f(\boldsymbol{x}) = f(\boldsymbol{x}^{(0)}) + \nabla^\top f(\boldsymbol{x}^{(0)})(\boldsymbol{x} - \boldsymbol{x}^{(0)}) + R_2(\boldsymbol{x}; \boldsymbol{x}^{(0)})$$

$$R_2(\boldsymbol{x}; \boldsymbol{x}^{(0)}) = \frac{1}{2}\underbrace{(\boldsymbol{x} - \boldsymbol{x}^{(0)})^\top \nabla^2 f(\theta \boldsymbol{x} + (1-\theta)\boldsymbol{x}^{(0)})(\boldsymbol{x} - \boldsymbol{x}^{(0)})}_{\boxed{\text{2 次形式}}} = o(\|\boldsymbol{x} - \boldsymbol{x}^{(0)}\|)$$

ただし，$0 < \theta < 1$

$\boldsymbol{x} = \boldsymbol{x}^{(0)}$ におけるテイラー展開 (2 次まで)

$$f(\boldsymbol{x}) = f(\boldsymbol{x}^{(0)}) + \nabla^\top f(\boldsymbol{x}^{(0)})(\boldsymbol{x} - \boldsymbol{x}^{(0)}) + \frac{1}{2}\underbrace{(\boldsymbol{x} - \boldsymbol{x}^{(0)})^\top \nabla^2 f(\boldsymbol{x}^{(0)})(\boldsymbol{x} - \boldsymbol{x}^{(0)})}_{\boxed{\text{2 次形式}}} + R_3(\boldsymbol{x}; \boldsymbol{x}^{(0)})$$

$R_3(\boldsymbol{x}; \boldsymbol{x}^{(0)}) = o(\|\boldsymbol{x} - \boldsymbol{x}^{(0)}\|^2)$

ランダウ (Landau) の漸近記法

$f(\boldsymbol{x}) = o(\|\boldsymbol{x} - \boldsymbol{x}^{(0)}\|^k)$ の意味： $\displaystyle\lim_{\boldsymbol{x} \to \boldsymbol{x}^{(0)}} \frac{|f(\boldsymbol{x})|}{\|\boldsymbol{x} - \boldsymbol{x}^{(0)}\|^k} = 0$ $\quad\left(\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x}^t \boldsymbol{x}} = \sqrt{\sum_{i=1}^n x_i^2} \right)$

$\boldsymbol{x}$ が $\boldsymbol{x}^{(0)}$ に近づくとき，$f(\boldsymbol{x})$ は $\|\boldsymbol{x} - \boldsymbol{x}^{(0)}\|^k$ より早く 0 に収束

$o(x)$ は「リトルオー」「スモールオー」と読む．$O(x)$ (ビッグオー) もよく使われる

2 次形式・正定・負定

2 次形式 (quadratic form)

n 次元ベクトル $\boldsymbol{x}$ と n 次 (実) 対称行列 A について,

$$Q(\boldsymbol{x}) = \boldsymbol{x}^\top A \boldsymbol{x} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$$

n 次対称行列 A が

- **正定** (poisitive definite): 任意の $\boldsymbol{x} \neq \boldsymbol{0}$ に対して $Q(\boldsymbol{x}) > 0$
- **負定** (negative definite): 任意の $\boldsymbol{x} \neq \boldsymbol{0}$ に対して $Q(\boldsymbol{x}) < 0$
- **半正定**・**準正定** (positive semidefinite): 任意の $\boldsymbol{x}$ に対して $Q(\boldsymbol{x}) \geq 0$
- **半負定**・**準負定** (negative semidefinite): 任意の $\boldsymbol{x}$ に対して $Q(\boldsymbol{x}) \leq 0$

正定 (負定) 行列は $A > O$ ($A < O$), 準正定 (準負定) 行列は $A \geq O$ ($A \leq O$) などと表す.

正定性・負定性と固有値の関係

- $A > O$ ($A < O$) $\Leftrightarrow$ A の固有値がすべて**正** (**負**)
- $A \geq O$ ($A \leq O$) $\Leftrightarrow$ A の固有値がすべて**非負** (**非正**)

対称行列の固有値はすべて実数であることに注意.

ヘッセ行列と最適解の関係

仮定

$f(\boldsymbol{x})$ は C^2 級

$\boldsymbol{x} = \boldsymbol{x}^{(0)}$ におけるテイラー展開

$$f(\boldsymbol{x}) = f(\boldsymbol{x}^{(0)}) + \nabla^\top f(\boldsymbol{x}^{(0)})(\boldsymbol{x} - \boldsymbol{x}^{(0)}) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}^{(0)})^\top \nabla^2 f(\boldsymbol{x}^{(0)})(\boldsymbol{x} - \boldsymbol{x}^{(0)}) + o(\|\boldsymbol{x} - \boldsymbol{x}^{(0)}\|^2)$$

$\boldsymbol{x}^{(0)}$ が停留点なら $f(\boldsymbol{x}) = f(\boldsymbol{x}^{(0)}) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}^{(0)})^\top \nabla^2 f(\boldsymbol{x}^{(0)})(\boldsymbol{x} - \boldsymbol{x}^{(0)}) + o(\|\boldsymbol{x} - \boldsymbol{x}^{(0)}\|^2)$

停留点 $\boldsymbol{x}^{(0)}$ の近傍での $\boldsymbol{x}$ の振る舞い

$o(\|\boldsymbol{x} - \boldsymbol{x}^{(0)}\|^2)$ の項は十分小さいので，

2 次形式 $(\boldsymbol{x} - \boldsymbol{x}^{(0)})^\top \nabla^2 f(\boldsymbol{x}^{(0)})(\boldsymbol{x} - \boldsymbol{x}^{(0)})$ が正： $f(\boldsymbol{x}) > f(\boldsymbol{x}^{(0)})$

2 次形式 $(\boldsymbol{x} - \boldsymbol{x}^{(0)})^\top \nabla^2 f(\boldsymbol{x}^{(0)})(\boldsymbol{x} - \boldsymbol{x}^{(0)})$ が負： $f(\boldsymbol{x}) < f(\boldsymbol{x}^{(0)})$

$$\Downarrow$$

ヘッセ行列 $\nabla^2 f(\boldsymbol{x}^{(0)})$ が正定なら，$\boldsymbol{x} \neq \boldsymbol{x}^{(0)}$ のとき $(\boldsymbol{x} - \boldsymbol{x}^{(0)})^\top \nabla^2 f(\boldsymbol{x}^{(0)})(\boldsymbol{x} - \boldsymbol{x}^{(0)}) > 0$ なので，**停留点 $\boldsymbol{x}^{(0)}$ は局所最適解**

ヘッセ行列と最適解の関係 (続き)

ヘッセ行列と局所的最適解 (2 次の最適性条件)

$f(\boldsymbol{x})$ は C^2 級とする．

2 次の必要条件　$\boldsymbol{x}^*$ が (P0) の局所的最適解なら †，$\nabla^2 f(\boldsymbol{x}^*) \succeq O$

2 次の十分条件　$\nabla f(\boldsymbol{x}^*) = \boldsymbol{0}$ かつ $\nabla^2 f(\boldsymbol{x}^*) \succ O$ なら，$\boldsymbol{x}^*$ は (P0) の局所的最適解

† $\boldsymbol{x}^*$ が (P0) の局所的最適解なら，1 次の必要条件より $\nabla(\boldsymbol{x}^*) = \boldsymbol{0}$ ($\boldsymbol{x}^*$ は停留点)

ヘッセ行列と凸性の関係

C^2 級関数 $f(\boldsymbol{x})$ が凸 (狭義凸) $\Leftrightarrow$ 任意の $\boldsymbol{x}$ に対して $\nabla^2 f(\boldsymbol{x}) \succeq O$ ($\nabla^2 f(\boldsymbol{x}) \succ O$)

ヘッセ行列と大域的最適解

任意の $\boldsymbol{x}$ に対して $\nabla^2 f(\boldsymbol{x}) \succeq O$ なら，$f(\boldsymbol{x})$ は凸関数なので，$f(\boldsymbol{x})$ の停留点は (P0) の大域的最適解

$\nabla^2 f(\boldsymbol{x}^*) \succeq O$ の場合，$\boldsymbol{x}^*$ が (P0) の局所的最適解になるとは限らないことに注意

参考：$\nabla^2 f(\boldsymbol{x}^*) \geq O$ の停留点 $\boldsymbol{x}^*$ が局所的最適解とならない例

$f(x_1, x_2) = \sin x_1 + \sin x_2 + \sin(x_1 + x_2)$ の停留点は，

$$\nabla f(x_1, x_2) = \begin{pmatrix} \frac{\partial}{\partial x_1} f(x_1, x_2) \\ \frac{\partial}{\partial x_2} f(x_1, x_2) \end{pmatrix} = \begin{pmatrix} \cos x_1 + \cos(x_1 + x_2) \\ \cos x_2 + \cos(x_1 + x_2) \end{pmatrix}$$

より，$0 \leq x_1, x_2 < 2\pi$ の範囲では $\left(\frac{\pi}{3}, \frac{\pi}{3}\right), (\pi, \pi), \left(\frac{5\pi}{3}, \frac{5\pi}{3}\right)$ の 3 点．ヘッセ行列は，

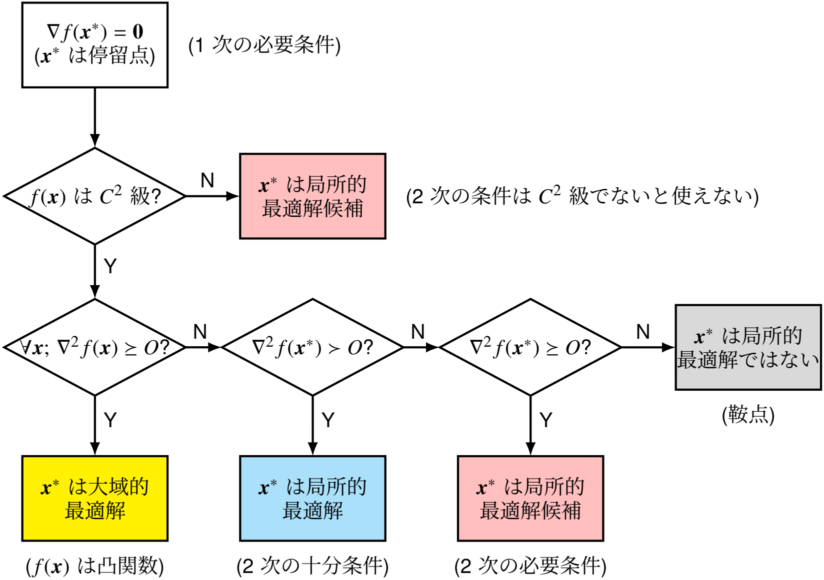
$$\nabla^2 f(x_1, x_2) = \begin{pmatrix} \frac{\partial^2}{\partial x_1^2} f(x_1, x_2) & \frac{\partial^2}{\partial x_1 \partial x_2} f(x_1, x_2) \\ \frac{\partial^2}{\partial x_2 x_1^2} f(x_1, x_2) & \frac{\partial^2}{\partial x_2^2} f(x_1, x_2) \end{pmatrix}$$
$$= \begin{pmatrix} -\sin x_1 - \sin(x_1 + x_2) & -\sin(x_1 + x_2) \\ -\sin(x_1 + x_2) & -\sin x_2 - \sin(x_1 + x_2) \end{pmatrix}$$

したがって，点 (π, π) におけるヘッセ行列は $\nabla^2 f(\pi, \pi) = O$ であり半正定．(π, π) における目的関数値は $f(\pi, \pi) = 0$ であり，

$$f(\pi - \delta, \pi - \delta) = 2\sin(\pi - \delta) + \sin(2\pi - 2\delta) = 2\sin(\pi - \delta)(1 + \cos(\pi - \delta))$$
$$f(\pi + \delta, \pi + \delta) = 2\sin(\pi + \delta) + \sin(2\pi + 2\delta) = 2\sin(\pi + \delta)(1 + \cos(\pi + \delta))$$

より，点 (π, π) の近傍に $f(\pi - \delta, \pi - \delta) > 0$ の点 $(\pi - \delta, \pi - \delta)$ と $f(\pi + \delta, \pi + \delta) < 0$ の点 $(\pi + \delta, \pi + \delta)$ が存在する $(\delta > 0)$ ので，局所的最適解ではない．
一方，$\left(\frac{5\pi}{3}, \frac{5\pi}{3}\right)$ は，ヘッセ行列が $\nabla^2 f\left(\frac{5\pi}{3}, \frac{5\pi}{3}\right) = \begin{pmatrix} \sqrt{3} & \sqrt{3}/2 \\ \sqrt{3}/2 & \sqrt{3} \end{pmatrix}$ より正定なので，局所的最適解．

制約なし問題の最適解フローチャート



制約なし凸 2 次計画問題

制約なし凸 2 次計画問題 (convex quadratic programming) の一般形

Q は n 次**半正定**行列，$\boldsymbol{q}$ は n 次元ベクトルとして，

$$\begin{aligned} \min \quad & f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^\top Q\boldsymbol{x} + \boldsymbol{q}^\top \boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{x} \in \mathbb{R}^n \end{aligned}$$

$Q = O, \boldsymbol{q} = \boldsymbol{0}$ の場合は考えないものとする．

$f(\boldsymbol{x})$ のヘッセ行列は

$$\begin{aligned} \nabla f(\boldsymbol{x}) &= Q\boldsymbol{x} + \boldsymbol{q} \\ \nabla^2 f(\boldsymbol{x}) &= Q \end{aligned}$$

より半正定．$\Rightarrow f(\boldsymbol{x})$ は**凸関数**

制約なし凸 2 次計画問題の大域的最適解

- Q が正定：$f(\boldsymbol{x})$ は狭義凸なので，停留点が大域的最適解．$\boldsymbol{x}^* = -Q^{-1}\boldsymbol{q}$
- Q が正定ではない (半正定)
  - $Q\boldsymbol{x} = -\boldsymbol{q}$ が解を持つ: $Q\boldsymbol{x} = -\boldsymbol{q}$ のすべての解
  - $Q\boldsymbol{x} = -\boldsymbol{q}$ が解を持たない: 単調増加もしくは単調減少であり，非有界

参考: 2 次形式の偏微分

$Q(\boldsymbol{x}) = \boldsymbol{x}^\top A\boldsymbol{x}$ の偏微分

$$Q(\boldsymbol{x}) = \boldsymbol{x}^\top A\boldsymbol{x} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j$$

より，$Q(\boldsymbol{x})$ の x_k の項を抜き出すと，

$$\begin{aligned} &a_{1k}x_1x_k + \cdots + a_{kk}x_kx_k + \cdots + a_{nk}x_nx_k \\ &\quad + a_{k1}x_kx_1 + \cdots + a_{k,k-1}x_kx_{k-1} + a_{k,k+1}x_kx_{k+1} \cdots a_{kn}x_kx_n \end{aligned}$$

したがって，

$$\begin{aligned} \frac{\partial}{\partial x_k}Q(\boldsymbol{x}) &= a_{1k}x_1 + \cdots + 2a_{kk}x_k + \cdots + a_{nk}x_n + a_{k1}x_1 + \cdots + a_{k,k-1}x_{k-1} + a_{k,k+1}x_{k+1} \cdots + a_{kn}x_n \\ &= \sum_{i=1}^{n} a_{ik}x_k + \sum_{j=1}^{n} a_{kj}x_j \end{aligned}$$

A は対称行列なので，$a_{kj} = a_{jk}$ に注意すれば，

$$\frac{\partial}{\partial x_k}Q(\boldsymbol{x}) = 2\sum_{i=1}^{n} a_{ik}x_k$$

よって，

$$\nabla Q(\boldsymbol{x}) = 2A\boldsymbol{x}$$

制約なし凸 2 次計画問題の例

例 1

$f(x_1, x_2) = 2x_1^2 + 2x_1x_2 + x_2^2 - x_1$ の大域的最適解

行列で書き直す

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^\top Q\boldsymbol{x} + \boldsymbol{q}^\top \boldsymbol{x}$$

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad Q = \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix}, \quad \boldsymbol{q} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

Q は正定行列なので凸 2 次計画問題．大域的最適解 $\boldsymbol{x}^*$ は，$\nabla f(\boldsymbol{x}^*) = \boldsymbol{0}$ を解いて

$$Q\boldsymbol{x}^* = -\boldsymbol{q}$$

$$\boldsymbol{x}^* = -Q^{-1}\boldsymbol{q} = \begin{pmatrix} 1/2 \\ -1/2 \end{pmatrix}$$

最適値は $f\left(\frac{1}{2}, -\frac{1}{2}\right) = -\frac{1}{4}$

参考: 2 次対称行列の正定性の判定法

- $\det Q > 0$ のとき，$q_{11} > 0$ なら正定，$q_{11} < 0$ なら負定
- $\det Q = 0$ のとき，$q_{11} + q_{22} \geq 0$ なら半正定，$q_{11} + q_{22} \leq 0$ なら半負定
- $\det Q < 0$ のとき，正定でも負定でもない

制約なし凸 2 次計画問題の例 (その 2)

例 2

$f(x_1, x_2) = x_1^2 + 2x_1x_2 + x_2^2 - x_1 - x_2$ の大域的最適解

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^\top Q\boldsymbol{x} + \boldsymbol{q}^\top\boldsymbol{x}$$

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad Q = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}, \quad \boldsymbol{q} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

Q は半正定行列，大域的最適解 $\boldsymbol{x}^*$ は連立 1 次方程式 $Q\boldsymbol{x} = -\boldsymbol{q}$ の解 $\boldsymbol{x}^* = \begin{pmatrix} x_1 \\ 1/2 - x_1 \end{pmatrix}$ (x_1 任意)．最適値は $-\dfrac{1}{4}$.

例 3

$f(x_1, x_2) = x_1^2 + 2x_1x_2 + x_2^2 - x_1$ の大域的最適解

$$Q = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}, \quad \boldsymbol{q} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

$Q\boldsymbol{x}^* = -\boldsymbol{q}$ は解を持たないので，非有界．実際，$f(x_1, 1 - x_1) = 1 - x_1$ より，

$$\lim_{x_1 \to \infty} f(x_1, 1 - x_1) = -\infty$$

練習問題：制約なし凸 2 次計画問題

練習問題

$f(x_1, x_2) = x_1^2 + 3x_1x_2 + \frac{5}{2}x_2^2 + x_1 + 2x_2$ の大域的最適解

行列で書き直す

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^\top Q\boldsymbol{x} + \boldsymbol{q}^\top \boldsymbol{x}$$

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad Q = \begin{pmatrix} \quad & \quad \\ \quad & \quad \end{pmatrix}, \quad \boldsymbol{q} = \begin{pmatrix} \quad \\ \quad \end{pmatrix}$$

Q は　　行列なので凸 2 次計画問題．大域的最適解 $\boldsymbol{x}^*$ は,

最適値は

練習問題：制約なし凸 2 次計画問題

練習問題

$f(x_1, x_2) = x_1^2 + 3x_1x_2 + \frac{5}{2}x_2^2 + x_1 + 2x_2$ の大域的最適解

行列で書き直す

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^\top Q\boldsymbol{x} + \boldsymbol{q}^\top \boldsymbol{x}$$

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad Q = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}, \quad \boldsymbol{q} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

Q は**正定**行列なので凸 2 次計画問題．大域的最適解 $\boldsymbol{x}^*$ は,

最適値は $f(1, -1) = -\frac{1}{2}$.

練習問題：制約なし凸 2 次計画問題

練習問題

$f(x_1, x_2) = x_1^2 + 3x_1x_2 + \frac{5}{2}x_2^2 + x_1 + 2x_2$ の大域的最適解

行列で書き直す

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^\top Q\boldsymbol{x} + \boldsymbol{q}^\top \boldsymbol{x}$$

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad Q = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}, \quad \boldsymbol{q} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

Q は**正定**行列なので凸 2 次計画問題．大域的最適解 $\boldsymbol{x}^*$ は，$\nabla f(\boldsymbol{x}^*) = \boldsymbol{0}$ を解いて

$$\begin{aligned} Q\boldsymbol{x}^* &= -\boldsymbol{q} \\ \boldsymbol{x}^* &= -Q^{-1}\boldsymbol{q} \\ &= -\begin{pmatrix} 5 & -3 \\ -3 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \end{aligned}$$

最適値は $f(1, -1) = -\frac{1}{2}$.

ニュートン法

ニュートン法 (Newton's method)

- 方程式を数値的に解くための反復解法
- 1 次の最適性条件 $\nabla f(\boldsymbol{x}) = \boldsymbol{0}$ に適用
- 目的関数 $f(\boldsymbol{x})$ は C^2 級を仮定

基本方針

1. 第 k 反復の解を $\boldsymbol{x}^{(k)}$ として，テイラー展開を使って $f(\boldsymbol{x})$ を 2 次近似 $\Rightarrow \tilde{f}_k(\boldsymbol{x})$ とおく

$$f(\boldsymbol{x}) = f(\boldsymbol{x}^{(k)}) + \nabla^\top f(\boldsymbol{x}^{(k)})(\boldsymbol{x} - \boldsymbol{x}^{(k)}) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}^{(k)})^\top \nabla^2 f(\boldsymbol{x}^{(k)})(\boldsymbol{x} - \boldsymbol{x}^{(k)}) + o(\|\boldsymbol{x} - \boldsymbol{x}^{(k)}\|^2)$$

$$\tilde{f}_k(\boldsymbol{x}) = f(\boldsymbol{x}^{(k)}) + \nabla^\top f(\boldsymbol{x}^{(k)})(\boldsymbol{x} - \boldsymbol{x}^{(k)}) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}^{(k)})^\top \nabla^2 f(\boldsymbol{x}^{(k)})(\boldsymbol{x} - \boldsymbol{x}^{(k)}) \tag{*}$$

2. $\nabla \tilde{f}_k(\boldsymbol{x}) = \boldsymbol{0}$ となる $\boldsymbol{x}$ を求めて $\boldsymbol{x}^{(k+1)}$ とする．(*) の両辺を $\boldsymbol{x}$ で偏微分して

$$\nabla \tilde{f}_k(\boldsymbol{x}) = \nabla f(\boldsymbol{x}^{(k)}) + \nabla^2 f(\boldsymbol{x}^{(k)})(\boldsymbol{x} - \boldsymbol{x}^{(k)})$$

(左辺)= $\boldsymbol{0}$ とおいて

$$\nabla f(\boldsymbol{x}^{(k)}) + \nabla^2 f(\boldsymbol{x}^{(k)})(\boldsymbol{x} - \boldsymbol{x}^{(k)}) = \boldsymbol{0}$$

$\nabla^2 f(\boldsymbol{x}^{(k)})$ が正則なら

$$\boldsymbol{x} = \boldsymbol{x}^{(k)} - \nabla^2 f(\boldsymbol{x}^{(k)})^{-1} \nabla f(\boldsymbol{x}^{(k)})$$

$\boldsymbol{d}^{(k)} = -\nabla^2 f(\boldsymbol{x}^{(k)})^{-1} \nabla f(\boldsymbol{x}^{(k)})$：ニュートン方向

ニュートン法 (続き)

ニュートン法のアルゴリズム

1. **(初期化)** 適当に初期解 $\boldsymbol{x}^{(0)}$ を決める．$k := 0$ とする
2. **(終了判定)** 終了条件を満たしているなら $\boldsymbol{x}^{(k)}$ を解として出力し終了
3. **(解の更新)** ニュートン方向 $\boldsymbol{d}^{(k)} := -\nabla^2 f(\boldsymbol{x}^{(k)})^{-1} \nabla f(\boldsymbol{x}^{(k)})$ を用いて，解を $\boldsymbol{x}^{(k+1)} := \boldsymbol{x}^{(k)} + \boldsymbol{d}^{(k)}$ により更新
4. **(次の反復へ)** $k := k + 1$ として 2 へ

終了条件

- $\|\nabla f(\boldsymbol{x}^{(k)})\| \leq \epsilon$
- $|f(\boldsymbol{x}^{(k+1)}) - f(\boldsymbol{x}^{(k)})| \leq \epsilon$
- $\|\boldsymbol{d}^{(k)}\| \leq \epsilon$
- etc.

$\boldsymbol{d}^{(k)}$ を計算する際は，連立 1 次方程式 $\nabla^2 f(\boldsymbol{x}^{(k)})\boldsymbol{d}^{(k)} = -\nabla f(\boldsymbol{x}^{(k)})$ を解けばよい (逆行列を求めなくてよい)

ニュートン法の適用例

例題

$$\begin{aligned} \text{min} \quad & f(x_1, x_2) = (x_1^2 + x_2^2 - 4)^2 + 8x_1^2x_2^2 \\ \text{s.t.} \quad & x_1,\ x_2 \in \mathbb{R} \end{aligned}$$

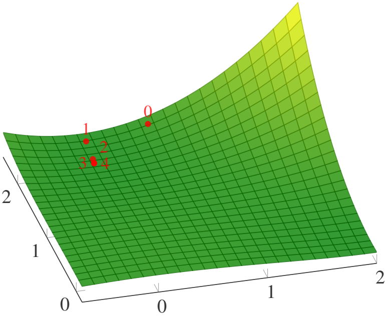
$$\nabla f(x_1, x_2) = 4\begin{pmatrix} x_1(x_1^2 + 5x_2^2 - 4) \\ x_2(5x_1^2 + x_2^2 - 4) \end{pmatrix}, \quad \nabla^2 f(x_1, x_2) = 4\begin{pmatrix} 3x_1^2 + 5x_2^2 - 4 & 10x_1x_2 \\ 10x_1x_2 & 5x_1^2 + 3x_2^2 - 4 \end{pmatrix}$$

1. ここでは初期解は $\boldsymbol{x}^{(0)} = \begin{pmatrix} 0.6000 \\ 2.4000 \end{pmatrix}$ とする
2. $\nabla f(0.6000, 2.4000) = \begin{pmatrix} 60.3840 \\ 34.1760 \end{pmatrix}$, $\nabla^2 f(0.6000, 2.4000) = \begin{pmatrix} 103.5200 & 57.6000 \\ 57.6000 & 60.320 \end{pmatrix}$ より，

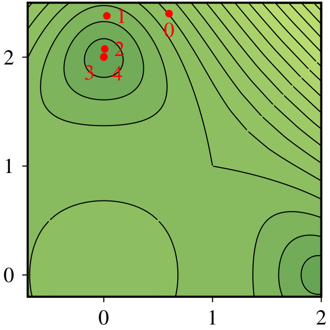
ニュートン方向は $\boldsymbol{d}^{(0)} = \begin{pmatrix} -0.5719 \\ -0.0204 \end{pmatrix}$. したがって，$\boldsymbol{x}^{(1)} = \boldsymbol{x}^{(0)} + \boldsymbol{d}^{(0)} = \begin{pmatrix} 0.0281 \\ 2.3796 \end{pmatrix}$.

以下，同様に解を更新すると，$\boldsymbol{x}^{(2)} = \begin{pmatrix} 0.0084 \\ 2.0754 \end{pmatrix}$, $\boldsymbol{x}^{(3)} = \begin{pmatrix} 0.0007 \\ 2.0040 \end{pmatrix}$, $\boldsymbol{x}^{(4)} = \begin{pmatrix} 0.0000 \\ 2.0000 \end{pmatrix}$.

ニュートン法の適用例 (続き)



$f(x_1, x_2)$ のグラフ



等高線

$$\boldsymbol{x}^{(0)} = \begin{pmatrix} 0.6000 \\ 2.4000 \end{pmatrix}, \quad \boldsymbol{x}^{(1)} = \begin{pmatrix} 0.0281 \\ 2.3796 \end{pmatrix}, \quad \boldsymbol{x}^{(2)} = \begin{pmatrix} 0.0084 \\ 2.0754 \end{pmatrix}, \quad \boldsymbol{x}^{(3)} = \begin{pmatrix} 0.0007 \\ 2.0040 \end{pmatrix}, \quad \boldsymbol{x}^{(4)} = \begin{pmatrix} 0.0000 \\ 2.0000 \end{pmatrix}$$

$$f(x_1^{(0)}, x_2^{(0)}) = 21.083, \quad f(x_1^{(1)}, x_2^{(1)}) = 2.8017, \quad f(x_1^{(2)}, x_2^{(2)}) = 0.0968, \quad f(x_1^{(3)}, x_2^{(3)}) = 0.0003, \quad f(x_1^{(4)}, x_2^{(4)}) = 0.0000$$

(局所) 最適解 $(\pm 2, 0), (0, \pm 2)$ の 1 つである $(0, 2)$ に収束

ニュートン法の練習問題

例題その 2

$$\begin{aligned} \text{min} \quad & f(x_1, x_2) = (x_1^2 - 1)^2 + (x_2^2 - 1)^2 \\ \text{s.t.} \quad & x_1,\ x_2 \in \mathbb{R} \end{aligned}$$

$\nabla f(x_1, x_2) =$, $\nabla^2 f(x_1, x_2) =$

1. 初期解を $\boldsymbol{x}^{(0)} = \begin{pmatrix} 1.5000 \\ 2.0000 \end{pmatrix}$ とする
2. $\nabla f(1.5000, 2.0000) =$, $\nabla^2 f(1.5000, 2.0000) =$ より，

ニュートン方向は $\boldsymbol{d}^{(0)} =$ ．したがって，$\boldsymbol{x}^{(1)} =$

以下，同様に解を更新すると，

ニュートン法の練習問題

例題その 2

$$\min \quad f(x_1, x_2) = (x_1^2 - 1)^2 + (x_2^2 - 1)^2$$
$$\text{s.t.} \quad x_1,\ x_2 \in \mathbb{R}$$

$$\nabla f(x_1, x_2) = 4\begin{pmatrix} x_1(x_1^2 - 1) \\ x_2(x_2^2 - 1) \end{pmatrix}, \quad \nabla^2 f(x_1, x_2) = 4\begin{pmatrix} 3x_1^2 - 1 & 0 \\ 0 & 3x_2^2 - 1 \end{pmatrix}$$

1. 初期解を $\boldsymbol{x}^{(0)} = \begin{pmatrix} 1.5000 \\ 2.0000 \end{pmatrix}$ とする
2. $\nabla f(1.5000, 2.0000) =$, $\nabla^2 f(1.5000, 2.0000) =$ より，

ニュートン方向は $\boldsymbol{d}^{(0)} =$ ．したがって，$\boldsymbol{x}^{(1)} =$

以下，同様に解を更新すると，

ニュートン法の練習問題

例題その 2

$$\begin{aligned} \text{min} \quad & f(x_1, x_2) = (x_1^2 - 1)^2 + (x_2^2 - 1)^2 \\ \text{s.t.} \quad & x_1,\ x_2 \in \mathbb{R} \end{aligned}$$

$$\nabla f(x_1, x_2) = 4\begin{pmatrix} x_1(x_1^2 - 1) \\ x_2(x_2^2 - 1) \end{pmatrix}, \quad \nabla^2 f(x_1, x_2) = 4\begin{pmatrix} 3x_1^2 - 1 & 0 \\ 0 & 3x_2^2 - 1 \end{pmatrix}$$

1. 初期解を $\boldsymbol{x}^{(0)} = \begin{pmatrix} 1.5000 \\ 2.0000 \end{pmatrix}$ とする
2. $\nabla f(1.5000, 2.0000) = \begin{pmatrix} 7.5000 \\ 24.0000 \end{pmatrix}$, $\nabla^2 f(1.5000, 2.0000) = \begin{pmatrix} 23.0000 & 0 \\ 0 & 44.0000 \end{pmatrix}$ より,

ニュートン方向は $\boldsymbol{d}^{(0)} =$. したがって, $\boldsymbol{x}^{(1)} =$

以下，同様に解を更新すると,

ニュートン法の練習問題

例題その 2

$$\min \quad f(x_1, x_2) = (x_1^2 - 1)^2 + (x_2^2 - 1)^2$$
$$\text{s.t.} \quad x_1,\ x_2 \in \mathbb{R}$$

$$\nabla f(x_1, x_2) = 4\begin{pmatrix} x_1(x_1^2 - 1) \\ x_2(x_2^2 - 1) \end{pmatrix}, \quad \nabla^2 f(x_1, x_2) = 4\begin{pmatrix} 3x_1^2 - 1 & 0 \\ 0 & 3x_2^2 - 1 \end{pmatrix}$$

1. 初期解を $\boldsymbol{x}^{(0)} = \begin{pmatrix} 1.5000 \\ 2.0000 \end{pmatrix}$ とする
2. $\nabla f(1.5000, 2.0000) = \begin{pmatrix} 7.5000 \\ 24.0000 \end{pmatrix}$, $\nabla^2 f(1.5000, 2.0000) = \begin{pmatrix} 23.0000 & 0 \\ 0 & 44.0000 \end{pmatrix}$ より，

ニュートン方向は $\boldsymbol{d}^{(0)} = \begin{pmatrix} -0.3261 \\ -0.5455 \end{pmatrix}$．したがって，$\boldsymbol{x}^{(1)} =$

以下，同様に解を更新すると，

ニュートン法の練習問題

例題その 2

$$\begin{aligned} \text{min} \quad & f(x_1, x_2) = (x_1^2 - 1)^2 + (x_2^2 - 1)^2 \\ \text{s.t.} \quad & x_1,\ x_2 \in \mathbb{R} \end{aligned}$$

$$\nabla f(x_1, x_2) = 4\begin{pmatrix} x_1(x_1^2 - 1) \\ x_2(x_2^2 - 1) \end{pmatrix}, \quad \nabla^2 f(x_1, x_2) = 4\begin{pmatrix} 3x_1^2 - 1 & 0 \\ 0 & 3x_2^2 - 1 \end{pmatrix}$$

1. 初期解を $\boldsymbol{x}^{(0)} = \begin{pmatrix} 1.5000 \\ 2.0000 \end{pmatrix}$ とする
2. $\nabla f(1.5000, 2.0000) = \begin{pmatrix} 7.5000 \\ 24.0000 \end{pmatrix}$, $\nabla^2 f(1.5000, 2.0000) = \begin{pmatrix} 23.0000 & 0 \\ 0 & 44.0000 \end{pmatrix}$ より，

ニュートン方向は $\boldsymbol{d}^{(0)} = \begin{pmatrix} -0.3261 \\ -0.5455 \end{pmatrix}$．したがって，$\boldsymbol{x}^{(1)} = \boldsymbol{x}^{(0)} + \boldsymbol{d}^{(0)} = \begin{pmatrix} 1.1739 \\ 1.4545 \end{pmatrix}$

以下，同様に解を更新すると，

ニュートン法の練習問題

例題その 2

$$\begin{aligned} \min \quad & f(x_1, x_2) = (x_1^2 - 1)^2 + (x_2^2 - 1)^2 \\ \text{s.t.} \quad & x_1,\ x_2 \in \mathbb{R} \end{aligned}$$

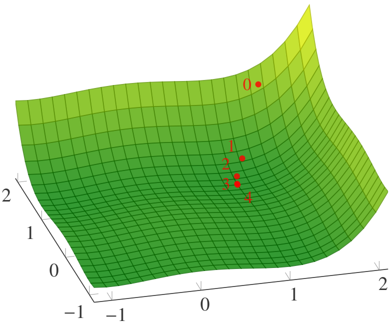
$$\nabla f(x_1, x_2) = 4\begin{pmatrix} x_1(x_1^2 - 1) \\ x_2(x_2^2 - 1) \end{pmatrix}, \quad \nabla^2 f(x_1, x_2) = 4\begin{pmatrix} 3x_1^2 - 1 & 0 \\ 0 & 3x_2^2 - 1 \end{pmatrix}$$

1. 初期解を $\boldsymbol{x}^{(0)} = \begin{pmatrix} 1.5000 \\ 2.0000 \end{pmatrix}$ とする
2. $\nabla f(1.5000, 2.0000) = \begin{pmatrix} 7.5000 \\ 24.0000 \end{pmatrix}$, $\nabla^2 f(1.5000, 2.0000) = \begin{pmatrix} 23.0000 & 0 \\ 0 & 44.0000 \end{pmatrix}$ より，

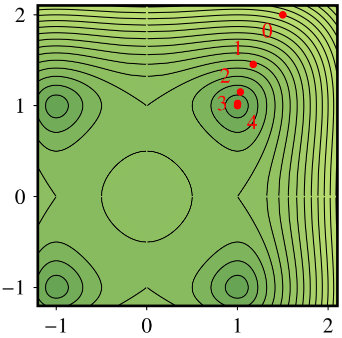
ニュートン方向は $\boldsymbol{d}^{(0)} = \begin{pmatrix} -0.3261 \\ -0.5455 \end{pmatrix}$． したがって，$\boldsymbol{x}^{(1)} = \boldsymbol{x}^{(0)} + \boldsymbol{d}^{(0)} = \begin{pmatrix} 1.1739 \\ 1.4545 \end{pmatrix}$

以下，同様に解を更新すると， $\boldsymbol{x}^{(2)} = \begin{pmatrix} 1.0323 \\ 1.1510 \end{pmatrix}$, $\boldsymbol{x}^{(3)} = \begin{pmatrix} 1.0015 \\ 1.0253 \end{pmatrix}$, $\boldsymbol{x}^{(4)} = \begin{pmatrix} 1.0000 \\ 1.0009 \end{pmatrix}$.

ニュートン法の練習問題 (続き)



$f(x_1, x_2)$ のグラフ



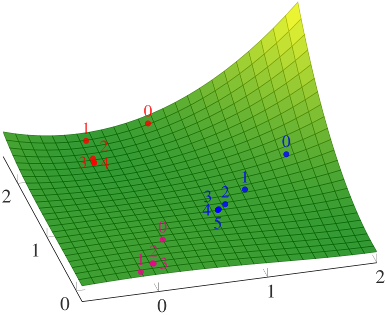
等高線

$$\boldsymbol{x}^{(0)} = \begin{pmatrix} 1.5000 \\ 2.0000 \end{pmatrix}, \quad \boldsymbol{x}^{(1)} = \begin{pmatrix} 1.1739 \\ 1.4545 \end{pmatrix}, \quad \boldsymbol{x}^{(2)} = \begin{pmatrix} 1.0323 \\ 1.1510 \end{pmatrix}, \quad \boldsymbol{x}^{(3)} = \begin{pmatrix} 1.0015 \\ 1.0253 \end{pmatrix}, \quad \boldsymbol{x}^{(4)} = \begin{pmatrix} 1.0000 \\ 1.0009 \end{pmatrix}$$

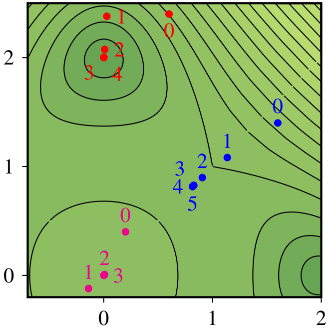
$$f(x_1^{(0)}, x_2^{(0)}) = 10.562, \quad f(x_1^{(1)}, x_2^{(1)}) = 1.3877, \quad f(x_1^{(2)}, x_2^{(2)}) = 0.1099, \quad f(x_1^{(3)}, x_2^{(3)}) = 0.0026, \quad f(x_1^{(4)}, x_2^{(4)}) = 0.0000$$

(局所) 最適解 $(\pm 1, \pm 1)$ (複号自由) の 1 つである $(1, 1)$ に収束

ニュートン法の適用例 (局所最適解に収束しない例)



$f(x_1, x_2)$ のグラフ



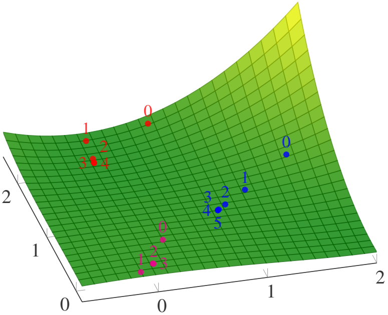
等高線

$$\boldsymbol{x}^{(0)} = \begin{pmatrix} 1.6000 \\ 1.4000 \end{pmatrix},\quad \boldsymbol{x}^{(1)} = \begin{pmatrix} 1.1357 \\ 1.0822 \end{pmatrix},\quad \boldsymbol{x}^{(2)} = \begin{pmatrix} 0.9061 \\ 0.8985 \end{pmatrix},\quad \boldsymbol{x}^{(3)} = \begin{pmatrix} 0.8275 \\ 0.8272 \end{pmatrix},\quad \boldsymbol{x}^{(4)} = \begin{pmatrix} 0.8167 \\ 0.8167 \end{pmatrix},\quad \boldsymbol{x}^{(5)} = \begin{pmatrix} 0.8165 \\ 0.8165 \end{pmatrix}$$

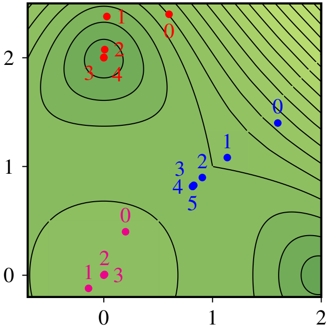
$$f(\boldsymbol{x}^{(0)}) = 40.4112 \quad f(\boldsymbol{x}^{(1)}) = 14.4545,\ f(\boldsymbol{x}^{(2)}) = 10.9273,\ f(\boldsymbol{x}^{(3)}) = 10.6705,\ f(\boldsymbol{x}^{(4)}) = 10.6667,\ f(\boldsymbol{x}^{(5)}) = 10.6667$$

$\left(\frac{\sqrt{6}}{3}, \frac{\sqrt{6}}{3}\right) \simeq (0.8165, 0.8165)$ に収束． $\nabla f\left(\frac{\sqrt{6}}{3}, \frac{\sqrt{6}}{3}\right) = \mathbf{0}$ だが，局所最適解ではない

ニュートン法の適用例 (局所最適解に収束しない例)



$f(x_1, x_2)$ のグラフ



等高線

$$\boldsymbol{x}^{(0)} = \begin{pmatrix} 0.2000 \\ 0.4000 \end{pmatrix}, \quad \boldsymbol{x}^{(1)} = \begin{pmatrix} -0.1404 \\ -0.1206 \end{pmatrix}, \quad \boldsymbol{x}^{(2)} = \begin{pmatrix} 0.0070321 \\ 0.0073798 \end{pmatrix}, \quad \boldsymbol{x}^{(3)} = \begin{pmatrix} 0.0000011316 \\ 0.0000011315 \end{pmatrix}$$

$$f(\boldsymbol{x}^{(0)}) = 14.4912, \quad f(\boldsymbol{x}^{(1)}) = 15.7294, \quad f(\boldsymbol{x}^{(2)}) = 15.9992, \quad f(\boldsymbol{x}^{(3)}) = 16.0000$$

極大点 $(0, 0)$ に収束． $\nabla f(0, 0) = \mathbf{0}$

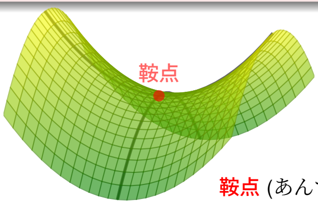
ニュートン法の長所・短所

ニュートン法の長所

- 収束が速い．局所的最適解の十分近くからスタートすれば 2 次収束 (**局所的収束**)
- 調整パラメータがなく，適用が容易

ニュートン法の短所

- $\nabla f(\boldsymbol{x}) = \boldsymbol{0}$ を満たす停留点を探す．$f(\boldsymbol{x})$ が凸でない場合は (停留点)=(局所的最適解) の保証なし (極大点や鞍点 (saddle point) なども見つかる)
- ヘッセ行列 $\nabla^2 f(\boldsymbol{x})$ が正則でないと，ニュートン方向を計算できない
- 1 回の反復計算に時間がかかる．決定変数の数が増えると大変
- 一般に**大域的収束**は保証されない



鞍点 (あんてん)：ある方向では極大点，別の方向では極小点

大域的収束と局所的収束

大域的収束 (global convergence)

任意の初期点から出発しても，いずれかの停留点が見つかる

- 数学的な定義は，点列 $\{\boldsymbol{x}^{(k)}\}$ の任意の集積点が停留点であること．したがって，探索点の列が発散する場合も含む

  例：数列 $\{(-1)^k\}$ は発散するが，集積点 $-1, 1$ を持つ
- 「大域的」は，初期点に依存しないという意味．大域的最適解と紛らわしいので注意

局所的収束 (global convergence)

停留点の近傍の初期点から出発すれば，停留点に収束する

局所的収束における収束の「速さ」も重要

局所的収束性の種類

点列 $\{\boldsymbol{x}^{(k)}\}$ は $\boldsymbol{x}^*$ に収束するとする

1 次収束 (linear convergence)

ある自然数 $N \in \mathbb{N}$ および実数 $0 \le r < 1$ が存在して，任意の $k \ge N$ に対して

$$\frac{\|\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^*\|}{\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^*\|} \le r$$

超 1 次収束 (superlinear convergence)

$$\lim_{k\to\infty} \frac{\|\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^*\|}{\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^*\|} = 0$$

p 次収束 ($p > 1$)

ある自然数 $N \in \mathbb{N}$ および実数 $r \ge 0$ が存在して，任意の $n \ge N$ に対して

$$\frac{\|\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^*\|}{\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^*\|^p} \le r$$

(速) p (> 1) 次収束 $\Rightarrow$ 超 1 次収束 $\Rightarrow$ 1 次収束 (遅)

ニュートン法の局所的収束性

ニュートン法の局所的収束性

$f(\boldsymbol{x})$ は C^2 級とする．もし停留点 $\boldsymbol{x}^*$ の近傍で $\nabla^2 f(\boldsymbol{x}^*)$ が正則なら，$\boldsymbol{x}^*$ の近傍に初期点 $\boldsymbol{x}^{(0)}$ を選んだニュートン法は $\boldsymbol{x}^*$ に**超 1 次収束**する．さらに，もし $\nabla^2 f(\boldsymbol{x})$ が $\boldsymbol{x}^*$ の近傍でリプシッツ連続 † なら，**2 次収束** (quadratic convergence).

† ある $L > 0$ が存在して，近傍内の任意の点 $\boldsymbol{x}$ において $\|\nabla^2 f(\boldsymbol{x}) - \nabla^2 f(\boldsymbol{x}^*)\| < L\|\boldsymbol{x} - \boldsymbol{x}^*\|$. ただし，行列 A のノルムは $\|A\| = \sup_{\|\boldsymbol{x}\|=1} \|A\boldsymbol{x}\|$ で定義される．

色々ややこしい前提条件はつくものの，ニュートン法は収束が速い!

$\delta^{(k)} = \|\boldsymbol{x}^{(k)} - \boldsymbol{x}^*\|$ の収束速度の比較 ($r = 0.8,\ r^{(k)} = 1/k$)

	1 次収束 $\delta^{(k+1)} = r\delta^{(k)}$	超 1 次収束 $\delta^{(k+1)} = r^{(k)}\delta^{(k)}$	2 次収束 $\delta^{(k+1)} = r(\delta^{(k)})^2$
$\delta^{(0)}$	1.0×10^{0}	1.0×10^{0}	1.0×10^{0}
$\delta^{(1)}$	8.0×10^{-1}	5.0×10^{-1}	8.0×10^{-1}
$\delta^{(2)}$	6.4×10^{-1}	1.7×10^{-1}	5.1×10^{-1}
$\delta^{(3)}$	5.1×10^{-1}	4.2×10^{-2}	2.1×10^{-1}
$\delta^{(4)}$	4.1×10^{-1}	8.3×10^{-3}	3.5×10^{-2}
$\delta^{(5)}$	3.3×10^{-1}	1.4×10^{-3}	9.9×10^{-4}
$\delta^{(6)}$	2.6×10^{-1}	2.0×10^{-4}	7.8×10^{-7}