数理計画法 第11回

第3章 非線形計画

3.2 制約なし最適化
3.3 制約つき最適化

担当： 塩浦昭義
(情報科学研究科 助教授)
shioura@dais.is.tohoku.ac.jp

期末試験と単位について

期末試験
- 日時：2月8日午後1時～2時半
- 試験範囲：
  - ネットワーク計画 (最大フロー、最小費用フロー)
  - 非線形計画 (制約なし最適化)
- 教科書、ノートなどの持ち込みは一切不可

単位について
単位に関する各種問合せ：2月16日 (金) 午後5時まで受付
それ以降は一切成績の修正はしません
期末試験等の得点の問合せについてはいつでも答えます

12／21の演習問題の解答例

問題２:関数 f_1, f_2 に対し，$x=(1, 1)$ におけるテイラー展開を求めなさい．
また，テイラー展開と元の関数の間の誤差が0．1以下となるxの範囲を計算しなさい．

$f_1(x_1, x_2) = x_1 + 2x_2$

$\nabla f_1(x_1, x_2) = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

f_1 の(1, 1)でのテイラー展開 $f_1(1,1) + \nabla f_1(1,1)^T \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = 3 + d_1 + 2d_2$

元の関数 $f_1(1+d_1, 1+d_2) = (1+d_1) + 2(1+d_2) = 3 + d_1 + 2d_2$

テイラー展開は元の関数に一致→誤差はない

$f_2(x_1, x_2) = x_1^2 + x_2^2 - 1 \qquad \nabla f_2(x_1, x_2) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}$

f_2 の(1, 1)でのテイラー展開 $f_2(1,1) + \nabla f_2(1,1)^T \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = 1 + 2d_1 + 2d_2$

元の関数 $f_2(1+d_1, 1+d_2) = (1+d_1)^2 + (1+d_2)^2 - 1 = 1 + 2d_1 + 2d_2 + d_1^2 + d_2^2$

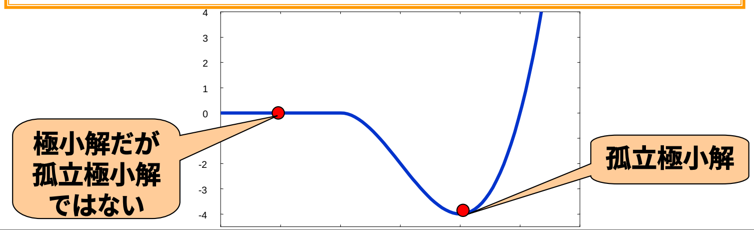
誤差は $d_1^2 + d_2^2 \Rightarrow (x_1-1)^2 + (x_2-1)^2 \le 0.1$ の範囲で誤差は0.1以下

復習：2次の最適性条件[p.99, 100]

ヘッセ行列を用いた最適性条件

定理 (2次の必要条件)：
x＊: 制約なし問題の極小解 ⇒ H f(x＊) は半正定値

定理 (2次の十分条件)：
x＊: 停留点, H f(x＊) は正定値 ⇒ x＊: 孤立極小解



復習：ニュートン法

ニュートン法のアイディア：

狭義2次凸関数の最適解は簡単に求められる！

一般の関数 ｆは狭義2次凸関数とは限らない

⇒ **ｆを2次のテイラー展開により近似**

$$f(\boldsymbol{x}+\boldsymbol{d}) \cong f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^T \boldsymbol{d} + \frac{1}{2}\boldsymbol{d}^T \mathrm{H}f(\boldsymbol{x})\boldsymbol{d}$$

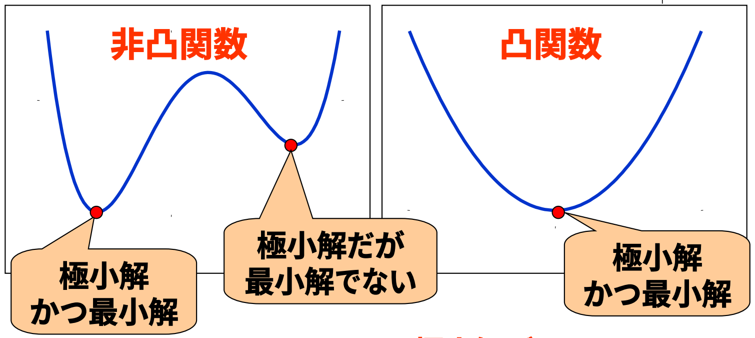
ヘッセ行列 H f(x) **が正定値のとき**
最適解は $d = -\mathrm{H}f(x)^{-1}\nabla f(x)$

ニュートン方向

⇒ x＋ d**は ｆの最適解のより良い近似解と期待できる**

凸関数 [p.93]

最小化しやすい関数の形は？



最小解でない極小解がある
➔最小化が難しい

極小解が一つ
➔最小化しやすい

凸関数の定義 [p.94]

定義：関数 f は凸関数
⇔ 任意のベクトル x, y
および任意の 0≦t≦1 に対して
(1 – t) f(x) + t f(y) ≧ f((1 – t) x + t y)

注：教科書の定義と異なります

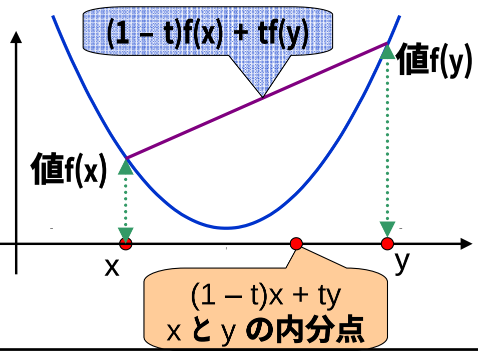
定義：関数 f は狭義凸関数
⇔ 任意の異なるベクトル x, y
および任意の 0＜t＜1 に対して
(1 – t) f(x) + t f(y) ＞ f((1 – t) x + t y)

凸関数の定義（続き） [p.94]

凸関数 ⇔ (1 – t) f(x) + t f(y) ≧ f((1 – t) x + t y)

狭義凸関数 ⇔ (1 – t) f(x) + t f(y) ＞ f((1 – t) x + t y)

狭義凸関数の例



凸関数の定義（続き） [p.94]

凸関数 ⇔ $(1-t)\,f(x) + t\,f(y) \geqq f((1-t)\,x + t\,y)$

狭義凸関数 ⇔ $(1-t)\,f(x) + t\,f(y) > f((1-t)\,x + t\,y)$

狭義凸関数の例

2次関数 $f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T V \boldsymbol{x} + \boldsymbol{c}^T \boldsymbol{x} + c_0$

（V: n ×n 行列, c: n次元ベクトル, c_0: 定数）

V ：正定値行列 ⇒ **狭義凸関数**

例えば

$$f(x_1, x_2) = \frac{1}{2}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^T \begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

凸関数の定義（続き） [p.94]

凸関数 ⇔ $(1-t)\,f(x) + t\,f(y) \geqq f((1-t)\,x + t\,y)$

狭義凸関数 ⇔ $(1-t)\,f(x) + t\,f(y) > f((1-t)\,x + t\,y)$

狭義凸関数の例

2次関数 $f(x) = a\,x^2$ $(a > 0)$ **は狭義凸関数**

（証明） 任意の異なる x, y **と** $0 < t < 1$ **に対して、**

$(1-t)\,ax^2 + t\,a\,y^2 - a\,[(1-t)\,x + t\,y]^2$

$= (1-t)\,ax^2 + t\,a\,y^2 - a\,(1-t)^2\,x^2 - a\,t^2\,y^2 - 2a\,(1-t)\,t\,x\,y$

$= (t-t^2)\,ax^2 + (t-t^2)\,a\,y^2 - 2a\,(t-t^2)\,x\,y$

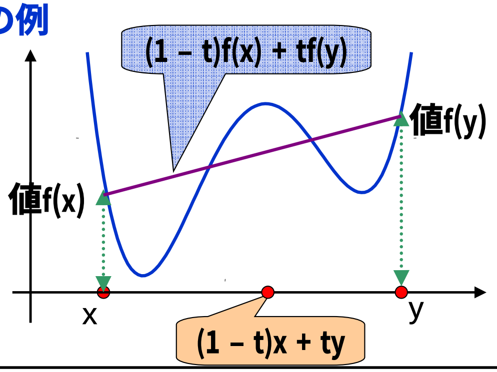
$= (t-t^2)\,a\,(x-y)^2$

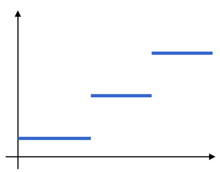
> 0 　　　　（$0 < t < 1$, $x \neq y$ より）

凸関数の定義（続き） [p.94]

凸関数 ⇔ $(1-t)\,f(x) + t\,f(y) \geqq f((1-t)\,x + t\,y)$

非凸関数の例





凸関数の最適解の性質 [p.101]

定理： f: 凸関数, x^*: f の極小解
⇒ x^*は制約なし問題の最適解

証明： x^* は極小解
⇒ ある ε > 0が存在して、
任意の x に対し $\|x - x^*\| < \varepsilon$ ならば $f(x) \geqq f(x^*)$

$f(y) < f(x^*)$ なる y が存在すると仮定
f は凸関数
⇒ $0 < t < 1$ なる任意の t に対して
$f((1-t)\,y + t\,x^*) \leqq (1-t)\,f(y) + t\,f(x^*) < f(x^*)$
t を1に近づけると
$(1-t)\,y + t\,x^*$ と x^* の距離 < ε （矛盾）

制約つき最適化問題

入力: 目的関数 $f(x)$

制約を表す関数 $g_i(x)$ $(i = 1, 2, \ldots, m)$

最小化 $f(x)$

条件 $g_i(x) \leqq 0$ $(i = 1, 2, \ldots, m)$

極小解: x^* の付近だけに注目したとき、x^* は最小

ある $\delta > 0$ が存在して、$\|x - x^*\| \leqq \delta$ を満たすすべての**許容解** x に対して $f(x) \geqq f(x^*)$

最適解ならば極小解

制約つき最適化問題の最適性条件

カルーシュ・キューン・タッカー (KKT) 条件

ある λ_i $(i = 1, \ldots, m)$ が存在して以下の式を満たす

$$\nabla f(\boldsymbol{x}) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\boldsymbol{x}) = \boldsymbol{0}$$

$$\lambda_i\, g_i(\boldsymbol{x}) = 0 \quad (i = 1, \ldots, m)$$

$\boldsymbol{x}$は許容解 $\quad g_i(\boldsymbol{x}) \le 0 \quad (i = 1, \ldots, m)$

$$\lambda_i \ge 0 \quad (i = 1, \ldots, m)$$

定理 (制約つき最適化問題の最適性条件 (必要条件)):
ある仮定の下で

$\tilde{\mathbf{x}}$:極小解 $\Rightarrow$ $\tilde{\mathbf{x}}$はKKT条件を満たす

制約つき最適化問題の最適性条件

カルーシュ・キューン・タッカー (KKT) 条件

ある λ_i $(i = 1, \ldots, m)$ が存在して以下の式を満たす

$$\nabla f(\boldsymbol{x}) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\boldsymbol{x}) = \boldsymbol{0}$$

$$\lambda_i\, g_i(\boldsymbol{x}) = 0 \quad (i = 1, \ldots, m)$$

$$g_i(\boldsymbol{x}) \le 0 \quad (i = 1, \ldots, m)$$

$$\lambda_i \ge 0 \quad (i = 1, \ldots, m)$$

定理 (制約つき最適化問題の最適性条件 (十分条件)):
f, g_i **は凸関数と仮定**

$\tilde{\mathbf{x}}$はKKT条件を満たす $\Rightarrow$ $\tilde{\mathbf{x}}$:最適解

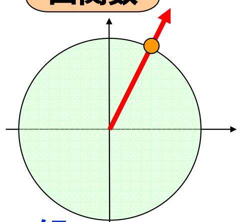
制約つき最適化問題の最適性条件

例 **最小化** $f(x_1, x_2) = -x_1 - 2x_2$

条件 $g(x_1, x_2) = x_1^2 + x_2^2 - 6 \le 0$

共に凸関数

$$\nabla f(x_1, x_2) = \begin{pmatrix} -1 \\ -2 \end{pmatrix} \quad \nabla g(x_1, x_2) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}$$



KKT条件

$$\begin{bmatrix} -1 \\ -2 \end{bmatrix} + \lambda \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} = 0$$

$$\lambda(x_1^2 + x_2^2 - 6) = 0$$

$$x_1^2 + x_2^2 - 6 \le 0$$

$$\lambda \ge 0$$

KKT条件の解

$$(x_1, x_2) = \left(\sqrt{\frac{6}{5}}, \sqrt{\frac{24}{5}}\right)$$

$$\lambda = \frac{1}{x_2}$$

最適解

LPに対するKKT条件

KKT条件は，LPの最適性条件を一般化したもの

LPの標準形

最小化 $\sum_{j=1}^{n} c_j x_j$

条件 $\sum_{j=1}^{n} a_{ij} x_j \geq b_i \ (i=1,\ldots,m)$

$x_j \geq 0 \ (j=1,\ldots,n)$

不等式を書き換え ⇒

最小化 $\sum_{j=1}^{n} c_j x_j$

条件 $-\sum_{j=1}^{n} a_{ij} x_j + b_i \leq (i=1,\ldots,m)$

$-x_j \leq 0 \ (j=1,\ldots,n)$

LPのKKT条件は次のようになる

$$\nabla f(\boldsymbol{x}) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\boldsymbol{x}) = \boldsymbol{0} \ \Rightarrow \ c_j - \sum_{i=1}^{m} y_i a_{ij} - z_j = 0 \ (j=1,\ldots,n)$$

$$\lambda_i \, g_i(\boldsymbol{x}) = 0 \ \Rightarrow \ y_i \Big(-\sum_{j=1}^{n} a_{ij} \boldsymbol{x}_j + b_i\Big) = 0 \ (i=1,\ldots,m)$$

$$z_j \boldsymbol{x}_j = 0 \ (j=1,\ldots,n)$$

LPに対するKKT条件

LPに対するKKT条件は

- $c_j - \sum_{i=1}^{m} y_i a_{ij} - z_j = 0 \ (j=1,\ldots,n)$
  $\Rightarrow \ z_j = c_j - \sum_{i=1}^{m} y_i a_{ij} \ (j=1,\ldots,n)$
- $y_i \Big(-\sum_{j=1}^{n} a_{ij} \boldsymbol{x}_j + b_i\Big) = 0 \ (i=1,\ldots,m), \quad z_j \boldsymbol{x}_j = 0 \ (j=1,\ldots,n)$
- $-\sum_{j=1}^{n} a_{ij} \boldsymbol{x}_j + b_i \leq 0 \ (i=1,\ldots,m), \quad -\boldsymbol{x}_j \leq 0 \ (j=1,\ldots,n)$
  $\Rightarrow \ \sum_{j=1}^{n} a_{ij} \boldsymbol{x}_j \geq b_i \ (i=1,\ldots,m), \quad \boldsymbol{x}_j \geq 0 \ (j=1,\ldots,n)$
- $y_i \geq 0 \ (i=1,\ldots,m), \quad z_j \geq 0 \ (j=1,\ldots,n)$

LPに対するKKT条件

z_j を消去し，整理すると，

$$\left.\begin{aligned} & y_i \Big(-\sum_{j=1}^{n} a_{ij} \boldsymbol{x}_j + b_i\Big) = 0 \ (i=1,\ldots,m) \\ & \Big(c_j - \sum_{i=1}^{m} y_i a_{ij}\Big) \boldsymbol{x}_j = 0 \ (j=1,\ldots,n) \end{aligned}\right\}$$

相補性条件

$\sum_{j=1}^{n} a_{ij} \boldsymbol{x}_j \geq b_i \ (i=1,\ldots,m), \quad \boldsymbol{x}_j \geq 0 \ (j=1,\ldots,n)$ **主実行可能性**

$y_i \geq 0 \ (i=1,\ldots,m), \quad c_j - \sum_{i=1}^{m} y_i a_{ij} \geq 0 \ (j=1,\ldots,n)$ **双対実行可能性**

LPの最適性条件に一致

等式制約のみの場合のKKT条件

等式制約のみの問題

最小化 f(x)

条件 $g_i(x) = 0$ (i = 1, 2, …, m)

等式を書き換え ⇒ $g_i(x) \leqq 0$, $g_i(x) \geqq 0$ (i = 1, 2, …, m)

KKT条件は次のようになる

$$\nabla f(\boldsymbol{x}) + \sum_{i=1}^{m} (\lambda_i^+ - \lambda_i^-) \nabla g_i(\boldsymbol{x}) = \boldsymbol{0}$$

$$\lambda_i^+ g_i(\boldsymbol{x}) = 0, \ \lambda_i^- g_i(\boldsymbol{x}) = 0 \ (i=1,\ldots,m)$$

$$g_i(\boldsymbol{x}) = 0 \quad (i=1,\ldots,m)$$

$$\lambda_i^+ \geq 0, \quad \lambda_i^- \geq 0 \quad (i=1,\ldots,m)$$

⇒

$$\nabla f(\boldsymbol{x}) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\boldsymbol{x}) = \boldsymbol{0}$$

$$g_i(\boldsymbol{x}) = 0 \ (i=1,\ldots,m)$$

等式だけの条件へ

ラグランジュの未定乗数法

- **等式制約のみ**の問題に対する解法
- ＫＫＴ条件（連立方程式）を満たす解を求める

$$\nabla f(\boldsymbol{x}) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\boldsymbol{x}) = \boldsymbol{0}$$

$$g_i(\boldsymbol{x}) = 0 \ (i = 1, \ldots, m)$$

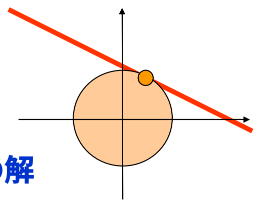
ラグランジュの未定乗数法

- **等式制約のみ**の問題に対する解法
- ＫＫＴ条件（連立方程式）を満たす解を求める

凸関数

例　最小化　$f(x_1, x_2) = x_1^2 + x_2^2$

　　条件　$g(x_1, x_2) = x_1 + 2x_2 - 3 = 0$



KKT条件

$$\begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} + \lambda \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 0$$

$$x_1 + 2x_2 - 3 = 0$$

KKT条件の解

$$(x_1, x_2) = \left(\frac{3}{5}, \frac{6}{5}\right), \ \lambda = -\frac{6}{5}$$

最適解

制約つき問題の一般的な解法

- ペナルティ関数法
- バリア関数法

後で説明

- 内点法　　ＫＫＴ条件を満たす解を反復計算により求める
- 逐次２次計画法　　２次関数の問題に繰り返し近似して解く

などなど

■ いずれも，極小解（の近似解）を求めることが目的
■ 目的関数および制約が凸関数のときは最適解（の近似解）が得られる

ペナルティ関数法

制約つき問題を制約なし問題へ変換して解く

最小化　$f(x)$　条件　$g_i(x) \leqq 0$ (i = 1, …, m)

ペナルティ関数の導入

P(x): x が制約を満たすとき ＝0
　　　満たさないとき ＞0

なるべく制約を満たしてほしい

(ペナルティ関数)

例えば，$P(x) = \Sigma_i \max\{g_i(x), 0\}$

最小化　$f(x) + \mu P(x)$　条件　なし　　$\mu > 0$は定数

ペナルティ問題

最急降下法，ニュートン法などを利用して解ける

ペナルティ関数法

ペナルティ問題

最小化　$f(x)+\mu P(x)$　**条件**　なし　　$\mu>0$は定数

μが十分大きい

⇒ペナルティ問題の最適解は$P(x)$　0を満たす

⇒x はもとの問題の制約をほぼ満たす

⇒x はもとの問題の最適解に近い

ペナルティ関数法の流れ

1. 適当にμの値を定める．
2. ペナルティ問題の最適解x_μを求める．
3. x_μが最適解に近い⇒終了
4. μを大きい値に変更，2に戻る．

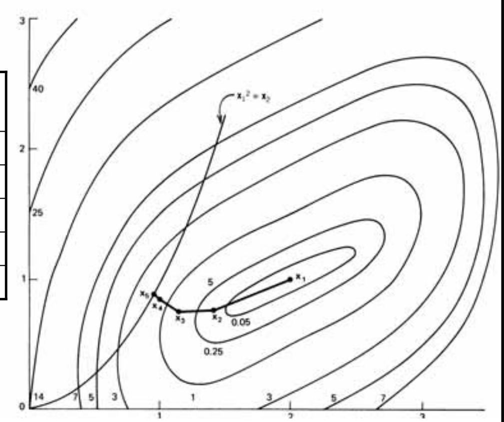
x_μは最初は制約を満たさない ⇒ 徐々に満たすようになる

ペナルティ関数法の実行例

最小化　$(x_1-2)^4+(x_1-2x_2)^2$

条件　$x_1^2-x_2=0$

μ	ペナルティ問題の最適解
0.1	(1.453, 0.7608)
1.0	(1.168, 0.7407)
10.0	(0.9906, 0.8425)
100.0	(0.9507, 0.8875)
1000.0	(0.9461, 0.8934)



バリア関数法

制約つき問題を制約なし問題へ変換して解く

最小化　$f(x)$　**条件**　$g_i(x)\leqq 0$ $(i=1,\ldots,m)$

バリア関数の導入

$B(x)$: 制約を満たすx に対して定義される

制約を必ず満たしてほしい

$g_i(x)\to 0$のとき$B(x)\to\infty$

(バリア関数)

例えば，$B(x)=-\Sigma_i\, 1/g_i(x)$

最小化　$f(x)+\mu B(x)$　**条件**　なし　　$\mu>0$は定数

バリア問題

最適解は，もとの問題の制約を必ず満たす

バリア関数法の実行例

最小化　$(x_1-2)^4+(x_1-2x_2)^2$

条件　$x_1^2-x_2\leqq 0$

μ	バリア問題の最適解
10.0	(0.7079, 1.5315)
1.0	(0.8282, 1.1098)
0.1	(0.8989, 0.9638)
0.01	(0.9294, 0.9162)
0.001	(0.9403, 0.9011)
0.0001	(0.94389, 0.89635)