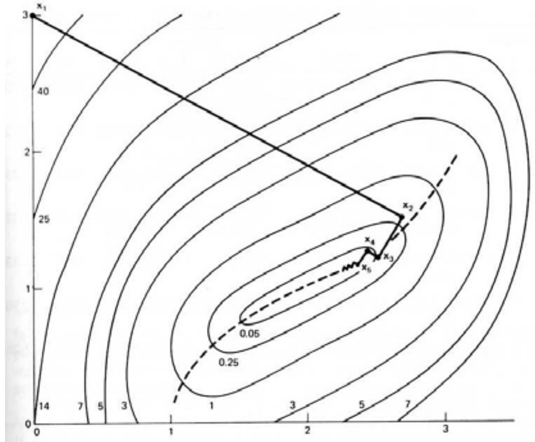
数理計画法 第10回

第3章　非線形計画

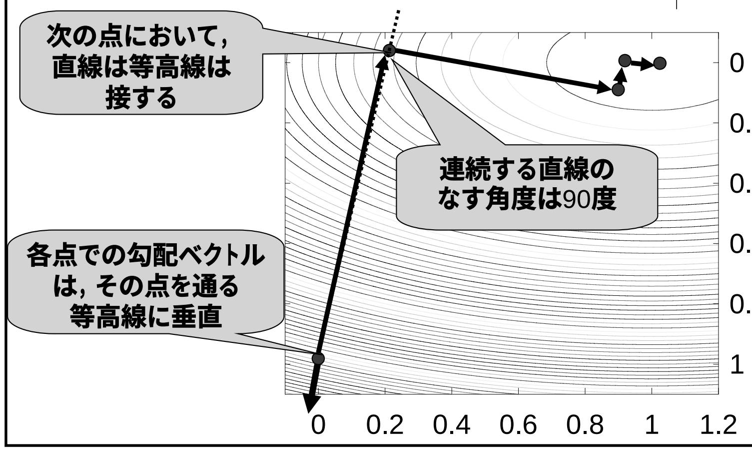
3.2　制約なし最適化
3.3　制約つき最適化

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

前回のレポートについて：最急降下法の問題



前回のレポートについて：最急降下法の実行例のポイント



復習：扱う問題

制約なし最適化問題

入力：目的関数 f(x) のみ

最小化　f(x)　条件　なし　($x \in R^n$)

どうやって最適解を求めるか？

関数 f の勾配ベクトル、ヘッセ行列を利用

復習：正定値性、半正定値性[p.99]

定義：対称な正方行列 A は半正定値
⇔ 任意のベクトル y に対して $y^T A y \geqq 0$

定義：対称な正方行列 A は正定値
⇔ 任意の非零ベクトル y に対して $y^T A y > 0$

※ A **が対称な**2×2**行列のとき、**
A**は半正定値** ⇔ $a_{11} \geqq 0$, $a_{22} \geqq 0$, $a_{11}a_{22} - a_{12}^2 \geqq 0$

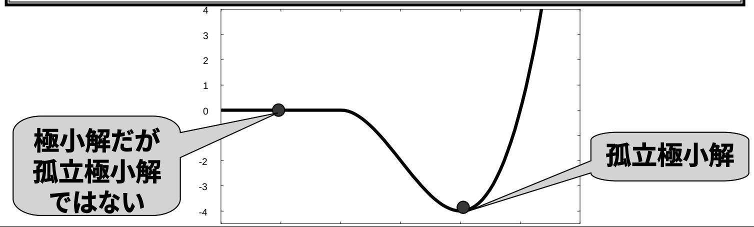
$\begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix}$**正定値** $\begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$**半正定値** $\begin{pmatrix} 2 & -2 \\ -2 & 0 \end{pmatrix}$**半正定値ではない**

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** ⇒ Hf(x*) **は半正定値**

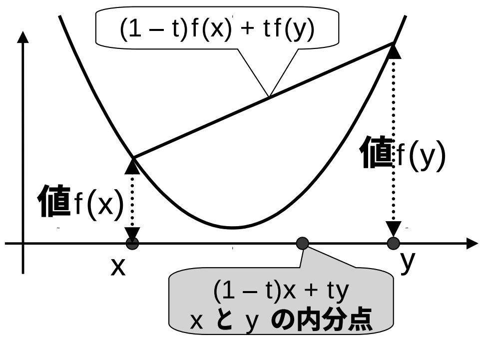
定理（2次の十分条件）：
x*: **停留点**, Hf(x*) **は正定値** ⇒ x*: **孤立極小解**



復習：凸関数 [p.94]

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

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



定理：
f: **凸関数**
x*: f **の極小解**
⇩
x***は最適解**

制約なし問題の解法2：ニュートン法 [p.105]

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！**

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T V \boldsymbol{x} + \boldsymbol{c}^T \boldsymbol{x} + c_0 \quad \text{(V は正定値行列)}$$

$$\nabla f(\boldsymbol{x}) = V\boldsymbol{x} + \boldsymbol{c} \qquad \mathrm{H}\, f(\boldsymbol{x}) = V$$

停留点は $x^* = -V^{-1}c$, **ヘッセ行列は** V **（正定値）**

⇨ 2**次の十分条件より** x* **は最適解**

定理：x*: **停留点**, Hf(x*) **は正定値** ⇒ x*: **孤立極小解**

制約なし問題の解法2：ニュートン法 [p.105]

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 Hf(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**

ニュートン法のアルゴリズム [p.106]

現在の点 x を繰り返しニュートン方向へ移動、最適解に近づける

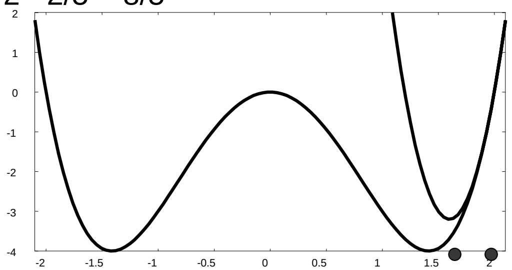
入力：関数 f とその勾配ベクトル∇f, ヘッセ行列Hf
初期点 x^0

ステップ0： k =0 **とする**
ステップ1： x^k **が最適解に十分近ければ終了**
ステップ2：ニュートン方向 $-\mathrm{H}f(x^k)^{-1}\nabla f(x^k)$ **を計算**
ステップ3： $x^{k+1}=x^k - \mathrm{H}f(x^k)^{-1}\nabla f(x^k)$ **とおく**
ステップ4： k=k+1**として、ステップ1に戻る**

ニュートン法の例 [p.106]

例1：一変数関数 $f(x) = x^4 - 4x^2$

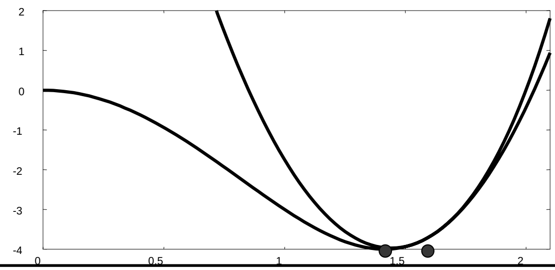
初期点 x = 2 **において f のテイラー近似を求める**
⇒ $f(2+d) \quad 0 + 16d + (40/2)d^2$
d = -2/5 **のとき最小**
⇒ **次の点は** x = 2 - 2/5 = 8/5



ニュートン法の例 [p.106]

例1(続き)：一変数関数 $f(x) = x^4 - 4x^2$

点 x = 8/5 **において f のテイラー近似を求める**
⇒ $f(8/5+d) \quad -3.69 + 3.58d + 11.36d^2$
d = -0.11 **のとき最小**
⇒ **次の点は** x = 1.6 – 0.11 = 1.49



ニュートン法の特徴 [p.107]

長所：

- 最急降下法より反復回数が少ない
  - ■ 狭義2次凸関数に対しては一反復で終了
- 直線探索が不要

短所：

- ヘッセ行列の逆行列の計算が必要
  - ■ ヘッセ行列の計算ができないと破綻
  - ■ ヘッセ行列が正則でないと破綻
- ヘッセ行列が正定値でない場合には
  目的関数値が増加する可能性あり

ニュートン法の問題点 [p.107]

■ ヘッセ行列が正則でないと破綻

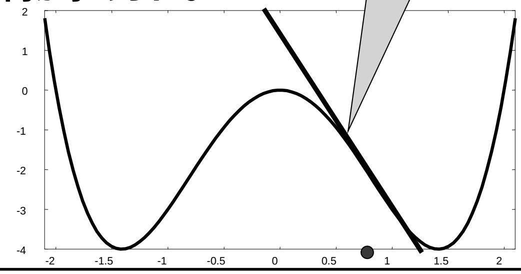
例1 (続き)：一変数関数 $f(x) = x^4 - 4x^2$

初期点 $x = \sqrt{2/3}$ のとき

⇒ ヘッセ行列は $Hf(x) = 0$ (正則でない)

⇒ ニュートン方向が求められない

f を2次近似すると直線になる



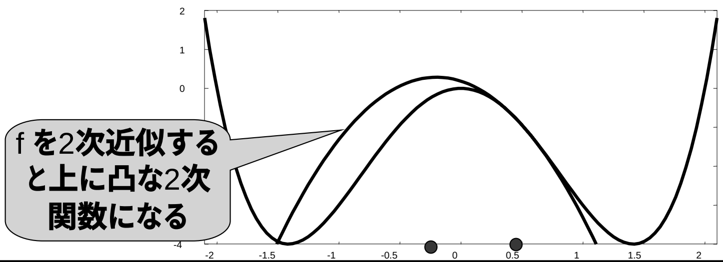
ニュートン法の問題点 [p.107]

- ヘッセ行列が正定値でない場合には
  目的関数値が増加する可能性あり

初期点 $x = 1/2$ のとき

⇒ ヘッセ行列は $Hf(x) = -5$ (正定値でない)

⇒ ニュートン方向に進むと関数値が増加する



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 $\boldsymbol{\lambda}_i$ (i = 1, ..., m)**が存在して以下の式を満たす**

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

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

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

$$\lambda_i \ge 0 \qquad (i = 1,\dots,m)$$

定理(制約つき最適化問題の最適性条件)：
ある仮定の下で
$\boldsymbol{x}$**:制約つき問題の極小解** ⇒ $\boldsymbol{x}$**はKKT条件を満たす**

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 $\sum_{j=1}^{n} a_{ij} x_j \ge b_i \; (i = 1,\dots,m)$

$x_j \ge 0 \; (j = 1,\dots,n)$

不等式を書き換え

最小化 $\sum_{j=1}^{n} c_j x_j$

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

$-x_j \le 0 \; (j = 1,\dots,n)$

LPのKKT条件は次のようになる

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

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

LPに対するKKT条件

LPに対するKKT条件は

- $c_j - \sum_{i=1}^{m} y_i a_{ij} - z_j = 0 \; (j = 1,\dots,n)$
  $\Rightarrow z_j = c_j - \sum_{i=1}^{m} y_i a_{ij} \; (j = 1,\dots,n)$
- $y_i\Big(-\sum_{j=1}^{n} a_{ij} x_j + b_i\Big) = 0 \; (i = 1,\dots,m), \quad z_j x_j = 0 \; (j = 1,\dots,n)$
- $-\sum_{j=1}^{n} a_{ij} x_j + b_i \le 0 \; (i = 1,\dots,m), \quad -x_j \le 0 \; (j = 1,\dots,n)$
  $\Rightarrow \sum_{j=1}^{n} a_{ij} x_j \ge b_i \; (i = 1,\dots,m), \quad x_j \ge 0 \; (j = 1,\dots,n)$
- $y_i \ge 0 \; (i = 1,\dots,m), \quad z_j \ge 0 \; (j = 1,\dots,n)$

LPに対するKKT条件

z_j **を消去し，整理すると，**

$$\left\{\begin{array}{l} y_i\Big(-\sum_{j=1}^{n} a_{ij} x_j + b_i\Big) = 0 \; (i = 1,\dots,m) \\ \Big(c_j - \sum_{i=1}^{m} y_i a_{ij}\Big) x_j = 0 \; (j = 1,\dots,n) \end{array}\right\}$$ **相補性条件**

$\sum_{j=1}^{n} a_{ij} x_j \ge b_i \; (i = 1,\dots,m), \quad x_j \ge 0 \; (j = 1,\dots,n)$ **主実行可能性**

$y_i \ge 0 \; (i = 1,\dots,m), \quad c_j - \sum_{i=1}^{m} y_i a_{ij} \ge 0 \; (j = 1,\dots,n)$ **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

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

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

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

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

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

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

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

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

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

例：

最小化　$x^2+y^2+z^2$

条件　6x+3y+4z=61

KKT条件は次のとおり

$$\begin{bmatrix} 2x \\ 2y \\ 2z \end{bmatrix} + \lambda \begin{bmatrix} 6 \\ 3 \\ 4 \end{bmatrix} = \boldsymbol{0}, \quad 6x + 3y + 4z = 61$$

解は(x,y,z) = (6,3, 4)

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
(ペナルティ関数)

なるべく制約を満たしてほしい

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

最小化　f(x) + μP(x)　条件　なし　　μ＞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) + μP(x)　**条件**　なし

μ＞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**

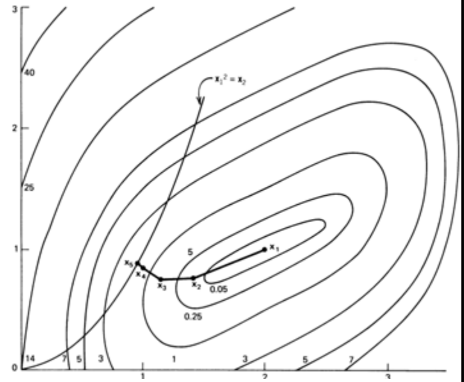
ペナルティ関数法の実行例

最小化　$(x_1-2)^4+(x_1-x_2)^2$

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)

ほぼ最適解



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, …, 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) + μB(x)　**条件**　なし

μ＞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

レポート問題（締め切り：2月2日）

(1) 対称な2×2**行列** A **に対し、次の関係を証明せよ。**

A**は半正定値 ⇔** $a_{11} \geqq 0$, $a_{22} \geqq 0$, $a_{11}a_{22} - a_{12}^2 \geqq 0$

(ヒント：教科書の問題3.7**の答えを参考にせよ)**

(2) 関数 $f(x) = x^3 + 6x^2$ **に対して**

(a) 初期点を x = 2 **としてニュートン法を適用せよ。**

(b) 初期点を x = 1 **としてニュートン法を適用せよ。**

それぞれ、反復は2回行うこと。

(3) 教科書　練習問題3．22