

# 計画数理学特論

## ～第11回：拡張ラグランジュ関数法～

---

担当：蓮池隆（経営システム工学科）

e-mail: [thasuike@waseda.jp](mailto:thasuike@waseda.jp)

# 前回の講義について

---

- 一般的な制約付きの最適化問題

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} & f(\mathbf{x}) \\ \text{subject to} & \mathbf{x} \in S \end{aligned}$$

- ペナルティ関数法

実行可能領域外の場合(=制約条件を違反した場合)にペナルティをかける

$q(c, \mathbf{x}) = f(\mathbf{x}) + cP(\mathbf{x})$ で,  $c \rightarrow \infty$ に近づけると最適解に近づく

- バリア関数法

実行可能領域内において, 境界に近づくにつれて目的関数を改悪する

$r(c, \mathbf{x}) = f(\mathbf{x}) + \frac{1}{c}B(\mathbf{x})$ で,  $c \rightarrow \infty$ に近づけると最適解に近づく

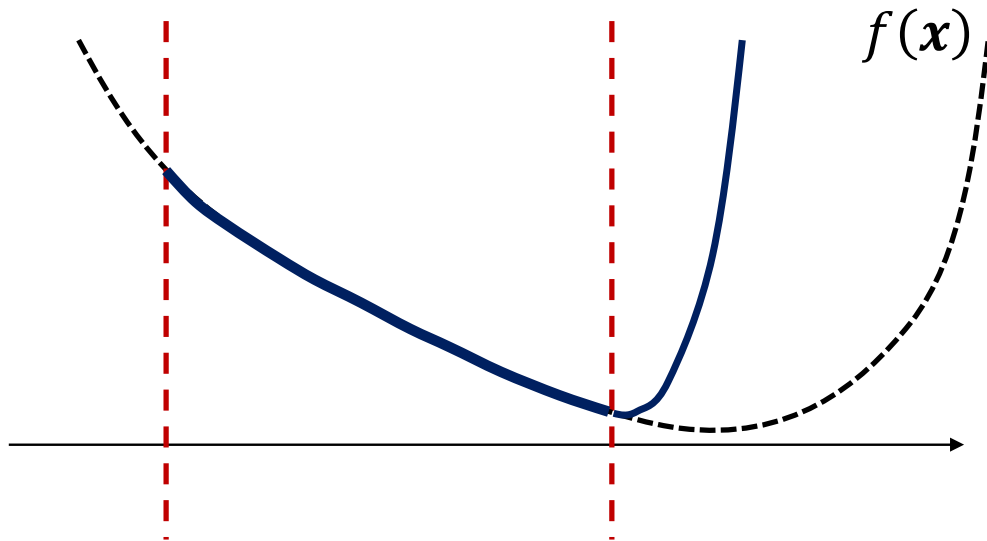
→ いずれも,  $c$ が大きくなると, 境界付近で最適化計算が不安定になる

# 前回の講義について

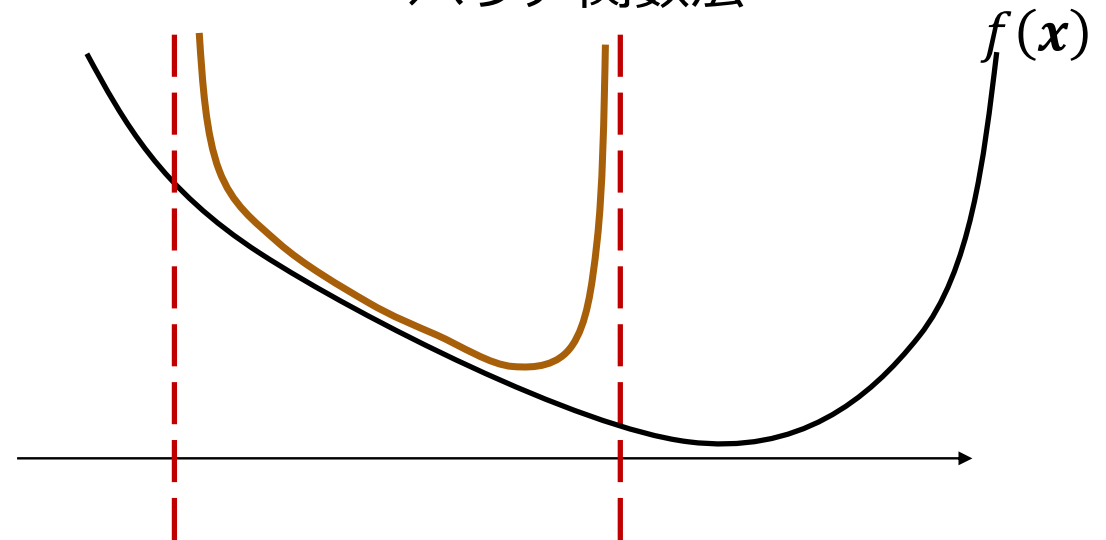
- 一般的な制約付きの最適化問題

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} & f(\mathbf{x}) \\ \text{subject to} & \mathbf{x} \in S \end{aligned}$$

ペナルティ関数法



バリア関数法



# 今回の講義について

---

- 一般的な制約付きの最適化問題

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & \mathbf{x} \in S \end{aligned}$$

- 拡張ラグランジュ関数を設定

→ まずは等式制約付き最適化問題を考える

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & g_i(\mathbf{x}) = 0, i = 1, 2, \dots, p \end{aligned}$$

→ 拡張ラグランジュ関数(ただし,  $\rho > 0$ )

$$L_\rho(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}) + \mathbf{y}^t \mathbf{g}(\mathbf{x}) + \frac{\rho}{2} \|\mathbf{g}(\mathbf{x})\|^2 = f(\mathbf{x}) + \sum_i y_i g_i(\mathbf{x}) + \frac{\rho}{2} \sum_i g_i(\mathbf{x})^2$$

# 拡張ラグランジュ関数法

- 等式制約付き最適化問題のラグランジュ関数における1次の必要条件

$$\begin{aligned}\nabla f(\mathbf{x}^*) + \sum_{i=1}^p y_i^* \nabla g_i(\mathbf{x}^*) &= \mathbf{0} \\ g_i(\mathbf{x}^*) &= 0, i = 1, 2, \dots, p\end{aligned}$$

- 通常のラグランジュ関数と同様に, 以下の問題を考える

$$\begin{aligned}\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) + \frac{\rho}{2} \|\mathbf{g}(\mathbf{x})\|^2 \\ \text{subject to } g_i(\mathbf{x}) &= 0, i = 1, 2, \dots, p\end{aligned}$$

→ この問題のラグランジュ関数が  $L_\rho(\mathbf{x}, \mathbf{y})$  の1次の必要条件

$$\nabla_{\mathbf{x}} L_\rho(\mathbf{x}, \mathbf{y}) = \nabla f(\mathbf{x}) + \sum_i y_i \nabla g_i(\mathbf{x}) + \rho \sum_i g_i(\mathbf{x}) \nabla g_i(\mathbf{x})$$

# 拡張ラグランジュ関数法

- ラグランジュ関数が  $L_\rho(\mathbf{x}, \mathbf{y})$  の1次の必要条件

$$\nabla_{\mathbf{x}} L_\rho(\mathbf{x}, \mathbf{y}) = \nabla f(\mathbf{x}) + \sum_i y_i \nabla g_i(\mathbf{x}) + \rho \sum_i g_i(\mathbf{x}) \nabla g_i(\mathbf{x})$$

→ 局所最適解  $\mathbf{x} = \mathbf{x}^*$  とすると, 制約条件から  $g_i(\mathbf{x}^*) = 0$  となるので,

$$\nabla_{\mathbf{x}} L_\rho(\mathbf{x}, \mathbf{y}^*)|_{\mathbf{x}=\mathbf{x}^*} = \nabla f(\mathbf{x}^*) + \sum_i y_i^* \nabla g_i(\mathbf{x}^*) = \mathbf{0}$$

- つまり, 適切な  $\mathbf{y}^*$  が得られていれば, 拡張ラグランジュ関数を最適化することで, 1次の必要条件を満たす  $\mathbf{x}^*$  が見つかる

# ヘッセ行列の計算

---

- 2次(ヘッセ行列)の計算

$$\begin{aligned}\nabla_{xx}L_\rho(\mathbf{x}, \mathbf{y}) &= \nabla^2 f(\mathbf{x}) + \sum_i y_i \nabla^2 g_i(\mathbf{x}) + \rho \sum_i \nabla g_i(\mathbf{x}) \nabla g_i(\mathbf{x})^t + \rho \sum_i g_i(\mathbf{x}) \nabla^2 g_i(\mathbf{x}) \\ &= \nabla^2 f(\mathbf{x}) + \sum_i (y_i + \rho g_i(\mathbf{x})) \nabla^2 g_i(\mathbf{x}) + \rho \sum_i \nabla g_i(\mathbf{x}) \nabla g_i(\mathbf{x})^t \\ &\rightarrow \nabla_{xx}L(\mathbf{x}^*, \mathbf{y}^*) + \rho \sum_i \nabla g_i(\mathbf{x}^*) \nabla g_i(\mathbf{x}^*)^t\end{aligned}$$

- $\nabla g_i(\mathbf{x}^*)^t \mathbf{d} = 0$ を満たす任意のベクトル $\mathbf{d}$ に対して,

$$\mathbf{d}^t \nabla_{xx}L_\rho(\mathbf{x}^*, \mathbf{y}^*) \mathbf{d} = \mathbf{d}^t \nabla_{xx}L(\mathbf{x}^*, \mathbf{y}^*) \mathbf{d} > 0$$

となる。(任意の $\mathbf{d}$ でも,  $\rho$ の値を十分大きくとれば同様に成り立つ)

$\rightarrow \nabla_{xx}L_\rho(\mathbf{x}^*, \mathbf{y}^*)$ が正定値行列であることがわかる

# 拡張ラグランジュ関数法

- ラグランジュ関数が $L_\rho(\mathbf{x}, \mathbf{y})$ の1次の必要条件

$$\nabla_{\mathbf{x}} L_\rho(\mathbf{x}, \mathbf{y}) = \nabla f(\mathbf{x}) + \sum_i y_i \nabla g_i(\mathbf{x}) + \rho \sum_i g_i(\mathbf{x}) \nabla g_i(\mathbf{x})$$

→ 局所最適解 $\mathbf{x} = \mathbf{x}^*$ とすると, 制約条件から $g_i(\mathbf{x}^*) = 0$ となるので,

$$\nabla_{\mathbf{x}} L_\rho(\mathbf{x}, \mathbf{y}^*)|_{\mathbf{x}=\mathbf{x}^*} = \nabla f(\mathbf{x}^*) + \sum_i y_i^* \nabla g_i(\mathbf{x}^*) = \mathbf{0}$$

- つまり, 適切な $\mathbf{y}^*$ が得られていれば, 拡張ラグランジュ関数を最適化することで, 1次の必要条件を満たす $\mathbf{x}^*$ が見つかる  
→ **この組 $(\mathbf{x}^*, \mathbf{y}^*)$ をうまく見つけることが重要**

# アルゴリズム構築

- 拡張ラグランジュ関数  $L_\rho(\mathbf{x}, \mathbf{y})$  において, ラグランジュ乗数  $\mathbf{y}^{(k)}$  が定められているとき最適化問題を解く  $\rightarrow$  解を  $\mathbf{x}^{(k+1)}$  とすると,

$$\begin{aligned} & \nabla_{\mathbf{x}} L_\rho(\mathbf{x}^{(k+1)}, \mathbf{y}^{(k)}) \\ &= \nabla f(\mathbf{x}^{(k+1)}) + \sum_i y_i^{(k)} \nabla g_i(\mathbf{x}^{(k+1)}) + \rho \sum_i g_i(\mathbf{x}^{(k+1)}) \nabla g_i(\mathbf{x}^{(k+1)}) = \mathbf{0} \end{aligned}$$

- 次に, ラグランジュ乗数を更新する:  $\mathbf{y}^{(k+1)} = \mathbf{y}^{(k)} + \rho \mathbf{g}(\mathbf{x}^{(k+1)})$

- $\nabla_{\mathbf{x}} L(\mathbf{x}^{(k+1)}, \mathbf{y}^{(k+1)})$ 

$$\begin{aligned} &= \nabla f(\mathbf{x}^{(k+1)}) + \sum_i y_i^{(k+1)} \nabla g_i(\mathbf{x}^{(k+1)}) \\ &= \nabla f(\mathbf{x}^{(k+1)}) + \sum_i \left( \mathbf{y}^{(k)} + \rho g_i(\mathbf{x}^{(k+1)}) \right) \nabla g_i(\mathbf{x}^{(k+1)}) = \nabla_{\mathbf{x}} L_\rho(\mathbf{x}^{(k+1)}, \mathbf{y}^{(k)}) = \mathbf{0} \end{aligned}$$

$\rightarrow$  つまり  $(\mathbf{x}^{(k+1)}, \mathbf{y}^{(k+1)})$  は1次の最適性条件を満たしている

# アルゴリズム

---

- STEP1 : 初期点 $(\mathbf{x}^{(0)}, \mathbf{y}^{(0)})$ およびパラメータの初期値 $\rho_0$ を定め,  $k = 0$ とする.
- STEP2 :  $\rho_k \sum_i g_i(\mathbf{x}^{(k)})^2$  が十分に小さければ, アルゴリズム終了
- STEP3 :  $\mathbf{x}^{(k)}$ を初期点として, 拡張ラグランジュ関数 $L_\rho(\mathbf{x}, \mathbf{y}^{(k)})$ を最小化する制約なし最適化問題を解き, 新たな点 $\mathbf{x}^{(k+1)}$ を求める.
- STEP4 :  $\mathbf{y}^{(k+1)} = \mathbf{y}^{(k)} + \rho_k \mathbf{g}(\mathbf{x}^{(k+1)})$ とする. また,  $\rho_{k+1} > \rho_k$ を満たすパラメータの値 $\rho_{k+1}$ を定め,  $k = k + 1$ としてSTEP2に戻る.

# 例を通してアルゴリズムを考える

---

(例題)

$$\begin{aligned} \min f(\mathbf{x}) &= (x_1 - 2)^4 + (x_1 - 2x_2)^2 \\ \text{subject to } g(\mathbf{x}) &= x_1^2 - x_2 = 0 \end{aligned}$$

(アルゴリズム)

# 例を通してアルゴリズムを考える

---

(例題)

$$\begin{aligned} \min f(\mathbf{x}) &= (x_1 - 2)^4 + (x_1 - 2x_2)^2 \\ \text{subject to } g(\mathbf{x}) &= x_1^2 - x_2 = 0 \end{aligned}$$

(アルゴリズム)

# 不等式制約の場合

---

- 不等式制約付き最適化問題

$$\begin{aligned} & \min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) \\ & \text{subject to } g_i(\mathbf{x}) \leq 0, i = 1, 2, \dots, m \end{aligned}$$

→ 不等式制約に  $s_i^2$  のようにスラック変数を導入すると

$$g_i(\mathbf{x}) \leq 0 \rightarrow g_i(\mathbf{x}) + s_i^2 = 0$$

- このスラック変数を用いて拡張ラグランジュ関数を最適化することにより、最適解を得ることができる

# 近接点アルゴリズムとの関係性

---

## 近接点アルゴリズムの概略

STEP1 : 初期解 $\mathbf{x}^{(0)}$ を設定し,  $k \leftarrow 0$ とする.

STEP2 : 停止条件

STEP3 :  $\mathbf{x}^{(k+1)} \leftarrow \operatorname{argmin} \left\{ f(\mathbf{x}) + \frac{\eta_k}{2} \|\mathbf{x} - \mathbf{x}^{(k)}\|^2 \right\}$ で更新. ( $\eta_k$ は正のステップ幅)

STEP4 :  $k \leftarrow k + 1$ として, STEP2へ戻る.

- 近接点アルゴリズムと拡張ラグランジュ関数法は双対関係にあることが知られている.

# 今回のまとめ

---

- 拡張ラグランジュ関数法について, アルゴリズムも含めて解説
- 拡張ラグランジュ関数法は, ペナルティ関数法の欠点を補った方法であり, 様々な場面で適用が可能