＜第6回＞
最　適　化

統計数理研究所

伊藤　聡

sito@ism.ac.jp

最 適 化

概 要

数理モデルあるいは統計モデルを用いた意思決定を行う際に必要となる数理最適化手法の基礎について学びます

1. 最適化とは
2. 無制約最適化とニュートン法
3. 制約つき最適化—線形計画を中心に
4. 離散最適化
5. ソフトウェアパッケージ

参考文献

1. 最適化とは

- 最適化問題

  minimize $f(x)$ ← 目的関数

  subject to $g(x) \leq 0,\ h(x) = 0,\ x \in S$ ← 制約条件

  - 決定変数，制約集合（制約領域, 許容領域, 実行可能領域）
  - 許容解（実行可能解），最適解（大域的, 局所的）
  - 多目的最適化（パレート最適解→選好最適化）
- 連続最適化
  - 線形計画（LP），2次計画（QP），非線形計画（NLP），確率計画
  - 半正定値計画（SDP），2次錐計画（SOCP）
- 離散最適化（組合せ最適化）
  - 整数計画（IP），混合整数計画（MIP），0-1整数計画（0-1IP）
  - 最短路，ネットワーク設計，施設配置，配送計画，スケジューリング

2. 無制約最適化とニュートン法

- 制約なし最適化問題

$$\min\ f(x) \qquad （ただし\ x \in R^n）$$

f が凸かつ微分可能ならば、非線形方程式 $\nabla f(x) = 0$ を解くことと等価

- ニュートン法

$x = x_c$ における局所2次モデル

$$m_c(x) = f(x_c) + \nabla f(x_c)^T (x - x_c) + \frac{1}{2}(x - x_c)^T \nabla^2 f(x_c)(x - x_c)$$

$\nabla^2 f(x_c) > 0$ ならば、凸2次関数 m_c の最小解 x_+ は

$$0 = \nabla m_c(x) = \nabla f(x_c) + \nabla^2 f(x_c)(x_+ - x_c)$$

すなわち

$$x_+ = x_c - \nabla^2 f(x_c)^{-1} \nabla f(x_c)$$

で与えられる

$$\nabla^2 f(x_c)\, s = -\nabla f(x_c)$$

$$x_+ = x_c + s$$

- ニュートン法
  - 収束が速いが局所的
    - 十分に滑らかな関数に対して、局所的最適解の十分近くから始めればその解に**2次収束**する
  - 大域的収束性を持たせるために
    - **直線探索（Line search）**… s 方向へ厳密に探索するのではなく、目的関数値の減少量 $f(x_c) - f(x_c + \alpha s)$ を制御する。Armijo 則や Wolfe 則など。
    - **信頼領域法（Trust region method）**… 信頼領域 $T(\Delta) := \{ |x - x_c| \le \Delta \}$ 上で f の局所2次モデル m_c を最小化する点を求める。
  - 計算を簡単にするために
    - 最適解から離れているときはニュートン方程式 $\nabla^2 f(x_c)\, s = -\nabla f(x_c)$ を正確に解く必要がない（→反復法により残差を制御する**近似ニュートン法**）
    - $s = x_+ - x_c$ と $y = \nabla f(x_+) - \nabla f(x_c)$ の情報からヘッセ行列の近似行列を作り上げる（→**準ニュートン法（quasi-Newton method）**, BFGS更新公式）

3. 制約つき最適化—線形計画を中心に

- 等式制約最適化問題

$$\min \ f(x)$$
$$\text{subj. to} \ \ h(x) = 0$$

  - **変分問題**として18世紀から数学者を魅了してきた
  - 最適解が満たすべき必要条件の一つは **Lagrange 乗数則**として知られる

- 不等式制約最適化問題

$$\min \ f(x)$$
$$\text{subj. to} \ \ g(x) \leq 0$$

  - 比較的新しく、20世紀中盤より**非線形計画**（**Nonlinear Programming, NLP**）として研究されるようになった
  - 最適解が満たすべき必要条件には、**Karush–Kuhn–Tucker**（**KKT**）**条件**がある

- 線形関数のみからなる場合は**線形計画**（**Linear Programming, LP**）と呼ばれる

- 線形計画の最適性条件と双対理論
  - 線形計画問題の標準形

主問題（Primal problem）

(P) $\min_{x} \; c^{\mathrm{T}}x \quad \text{subj. to} \;\; Ax = b, \;\; x \geq 0$ （where $x \in \mathrm{R}^{n}, A \in \mathrm{R}^{m \times n}, b \in \mathrm{R}^{m}$）

  - 最適性条件（KKT条件）

Lagrange 関数を $L(x, y, z) \;:=\; c^{\mathrm{T}}x + y^{\mathrm{T}}(b - Ax) - z^{\mathrm{T}}x$ とおくと

$$\begin{cases} Ax = b, \;\; x \geq 0 & \text{主制約（Primal feasibility）} \\ A^{\mathrm{T}}y + z = c, \;\; z \geq 0 & \text{双対制約（Dual feasibility）} \\ x^{\mathrm{T}}z = 0 & \text{相補性（Complimentarity）} \end{cases}$$

  - 双対理論

双対問題（Dual problem）

(D) $\max_{y, z} \; b^{\mathrm{T}}y \quad \text{subj. to} \;\; A^{\mathrm{T}}y + z = c, \;\; z \geq 0$ （where $y \in \mathrm{R}^{m}, z \in \mathrm{R}^{n}$）

- (P), (D) の任意の許容解 (x, y, z) に対して $c^{\mathrm{T}}x \geq b^{\mathrm{T}}y$ 【弱双対性】
- (P), (D) の最適解 (x^*, y^*, z^*) に対して $c^{\mathrm{T}}x^* = b^{\mathrm{T}}y^*$ 【強双対性】

- 線形計画法:単体法と内点法
  - 凸多面体上の線形最適化
  - **単体法（Simplex method）**
    - 多面体の頂点（実行可能基底解に相当）を辿り、有限回の反復で最適解に到達する。1947年 Dantzig によって提案された。多項式時間解法ではないが十分に実用的。
  - **内点法（Interior point method)**
    - 最適解に収束する点列を制約内部に生成する多項式時間解法であり、主双対法（primal–dual method）や予測子修正子法（predictor-corrector method）が代表的。大規模な問題に対して単体法よりも優れる。
  - KKT 条件との関係

$Ax = b,\ \ x \geq 0$ 【主制約条件】

$A^{\mathrm{T}}y + z = c,\ \ z \geq 0$ 【双対制約条件】

$x^{\mathrm{T}}z = 0$ 【相補性条件】

- 一般の非線形計画問題に対する解法としては**逐次2次計画法**（**Sequential Quadratic Programming**）←準ニュートン法の拡張
- **半正定値計画（Semi-definite programming, SDP）**
  - 実対称行列の空間における線形計画
  - 主問題と双対問題

$$
(\mathrm{P})\begin{cases}\displaystyle\min_{X}\ C \bullet X \\ \text{subj. to}\ \ A_i \bullet X = b_i \\ \qquad\qquad X \geq 0\end{cases}
\qquad
(\mathrm{D})\begin{cases}\displaystyle\max_{y,Z}\ b^{\mathrm{T}} y \\ \text{subj. to}\ \displaystyle\sum_{i=1}^{m} y_i A_i + Z = C \\ \qquad\qquad Z \geq 0\end{cases}
$$

where $X, Z, A_i, C \in S^n$（n 次実対称行列の空間）, $y, b \in \mathrm{R}^m$

  - 豊かな表現力を持ち、制御工学・信号処理・機械学習など多くの分野に応用されている【参考文献3】
- **２次錐計画（Second-order cone programming, SOCP）**

4. 離散最適化

- 組合せ最適化問題の難しさ
  - 30都市の巡回セールスマン問題の巡回路を列挙するのに100TFlopsの計算機でさえ2.5兆年以上（→組合せ爆発）
  - 現実問題の多くは**NP困難（NP-hard）**、すなわち厳密な最適解を求めるのに最悪の場合（入力サイズに対して）指数時間を要するだろうと考えられている
- 厳密解法と近似解法
  - 最悪の場合に指数時間をかけても**厳密解**を求めるか、あるいは現実的な時間でよい実行可能解すなわち**近似解**を求めるか
- 汎用ソルバーと専用ソルバー
  - （混合）整数計画問題に定式化して既存の**汎用ソルバー**を使うか、あるいは問題特有の構造をうまく利用したメタヒューリスティクス（発見的解法）による**専用ソルバー**を開発するか【参考文献4・5】

5. ソフトウェアパッケージ

- 商用ソルバーと非商用ソルバー
  - 商用ソルバー
    - Gurobi Optimizer (Gurobi Optimization), ILOG CPLEX (IBM), MATLAB Optimization Toolbox (MathWorks), NUOPT（NTTデータ数理システム）など
  - 非商用ソルバー
    - SCIP (ZIB) など
- Decision Tree for Optimization Software（H.D. Mittelmann）
- モデルの記述
  - GUI
  - CUI：LP形式, MPS形式, モデリング言語（AIMMS, AMPL, GAMS など）
  - API：C, C++, Java, MATLAB, Python など

参考文献

1. 矢部博『工学基礎 最適化とその応用』（数理工学社, 2006年）
2. S. Boyd and L. Vandenberghe『Convex Optimization』（Cambridge University Press, 2004年）
3. オペレーションズ・リサーチ Vol.55, No.7（2010）
特集「半正定値計画に対するソルバーと応用例」
4. オペレーションズ・リサーチ Vol.57, No.4（2012）
特集「はじめよう整数計画」
5. オペレーションズ・リサーチ Vol.58, No.12（2013）
特集「はじめようメタヒューリスティックス」
6. H.D. Mittelmann, Decision Tree for Optimization Software, http://plato.asu.edu/guide.html