計画数理学特論
～第1回：ガイダンス・序論～

担当：蓮池隆　(経営システム工学科)
e-mail: thasuike@waseda.jp

この講義全体について

- オペレーションズ・リサーチ，統計解析・機械学習などのデータ解析の多くの場面で登場する『**連続最適化**』，特に『**非線形最適化**』に焦点を当てて解説
- **連続最適化**：最適化問題における決定変数が連続値を取る
  (cf. 離散最適化：決定変数が0 or 1や整数値のみの場合)
- **非線形最適化**：最適化問題の定式化に非線形項が含まれる

例：n個のデータ対$\left(x_j, y_j\right), (j = 1,2, \ldots, n)$を利用し，最小二乗法により回帰直線

この講義全体について

・オペレーションズ・リサーチ，統計解析・機械学習などのデータ解析の多くの場面で登場する『**連続最適化**』，特に『**非線形最適化**』に焦点を当てて解説

・**連続最適化**：最適化問題における決定変数が連続値を取る
(cf. 離散最適化：決定変数が0 or 1や整数値のみの場合)

・**非線形最適化**：最適化問題の定式化に非線形項が含まれる

例：m箇所の供給地(各供給地の供給可能量：$a_i, (i = 1,2, \dots, m)$)からn箇所の需要地(各需要地の需要量：$b_j, (j = 1,2, \dots, n)$)に商品を運びたい

講義の主な内容

・連続変数を持つ非線形最適化問題の中でも，『**制約付き最適化問題**』を中心に，「なぜその手法を用いるのか？」「なぜその手法が開発されたのか」を解説．(できる限り後藤先生の授業と被らないようにします)

・基本事項(第1回)：最適化問題の数学的記法，微積分の基礎
反復法と収束速度

・凸解析(第2回，第3回)

・最適性条件(第4回)

・勾配法とその応用(第5回，第6回)

(前半は，制約有り無しに関わらず解析に必要な数理的手法について学習)

講義の主な内容

・連続変数を持つ非線形最適化問題の中でも，『**制約付き最適化問題**』を中心に，「なぜその手法を用いるのか？」「なぜその手法が開発されたのか」を解説．(できる限り後藤先生の授業と被らないようにします)

・制約付き最適化：等式制約(第7回)，不等式制約(第8回)
・主問題に対する最適化法(第9回・第10回)
・ラグランジュ関数を用いた最適化法(第11回・第12回)
・上界最小化アルゴリズムとEMアルゴリズム(第13回)
・その他，最終演習課題(第14回，第15回)
(後半は，制約付き最適化問題に関わる様々なケースを学習)

テキスト(持ってなくてもOK)

- 金森，鈴木，竹内，佐藤著
  『機械学習のための連続最適化』　講談社
- 統計学習プロフェッショナルシリーズの1冊
- ORの非線形最適化に関する内容も一通り網羅
- 最後の方は，そこそこ最新の機械学習手法で利用される最適化手法(スパース学習，多様体上の最適化)などが掲載されている

最適化問題の記載方法

- 本講義で主に利用する最適化問題の記載方法
  (決定変数：$\boldsymbol{x} = (x_1, x_2, \dots, x_n)^t \in \mathbb{R}^n$ (n次元実数ベクトル))

$$\boxed{\begin{array}{l} \text{Minimize } \; f(\boldsymbol{x}) \\ \text{subject to } \; \boldsymbol{x} \in S \end{array}}$$

- $f(\boldsymbol{x})$：目的関数，$f\colon \mathbb{R}^n \to \mathbb{R}$ (n次元空間から1次元空間への写像)
  (今のところ，微分可能であるかどうかは問わない)
- S：実行可能領域．この領域内に含まれる$\boldsymbol{x}$が実行可能解
- $f(\boldsymbol{x})$とSを問題に合わせて変更することにより，実社会の様々な問題が最適化問題として表現できる

最適化問題の記載方法

(参考)

・決定変数$\boldsymbol{x}$が
- ・$\boldsymbol{x} \in \mathbb{R}^n$：連続最適化問題
- ・$\boldsymbol{x} \in \mathbb{Z}^n$：離散最適化問題(特に，$\boldsymbol{x} \in \{0, 1\}^n$なら0-1整数計画問題)

・実行可能領域 S が
- ・$S = \{\boldsymbol{x} | A\boldsymbol{x} \leq \boldsymbol{b}\}$のように線形関数で書けて，かつ閉空間＝凸多面体
- ・通常は，$S = \{\boldsymbol{x} | g_k(\boldsymbol{x}) = 0, (k = 1,2, \dots)\}$や$S = \{\boldsymbol{x} | g_k(\boldsymbol{x}) \leq 0, (k = 1,2, \dots)\}$のように関数$g_k(\boldsymbol{x})$を用いて表現する

ベクトルのノルム(≒距離)

- **ノルム**：ベクトルの "長さ" の概念の一般化であり、ベクトル空間に対して「距離」を与えるための数学の道具
- 非線形最適化でよく利用するノルム：l_2ノルム(≒ユークリッド距離)
  → $\boldsymbol{x} \in \mathbb{R}^n$に対して,
- その他のノルム
  - l_1ノルム(≒マンハッタン距離)：
  - p次ノルム：
  - 無限大ノルム：

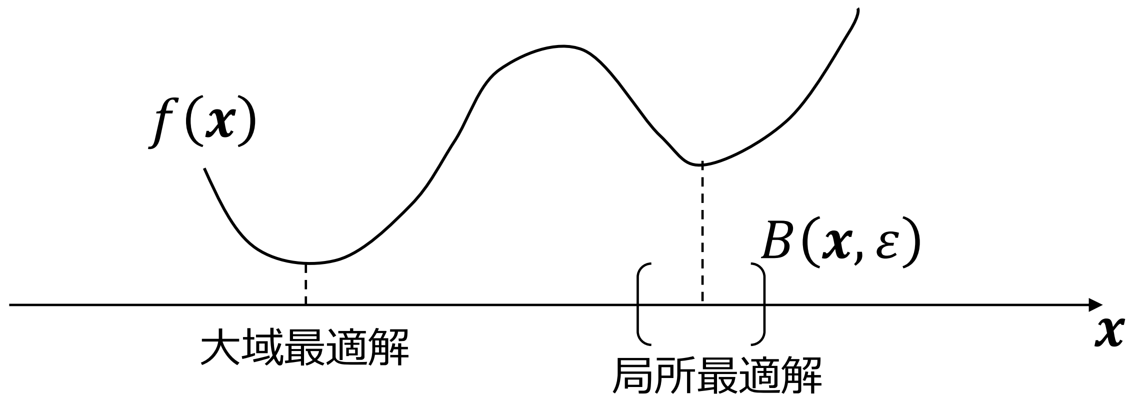
近傍と最適解の考え方

・ε近傍：点$\boldsymbol{x} \in \mathbb{R}^n$，$\varepsilon > 0$に対して，$B(\boldsymbol{x}, \varepsilon) =$

(大域最適解と局所最適解)　最適解$\boldsymbol{x}^*$として

・大域最適解：$\boldsymbol{x}^* \in S$が任意の$\boldsymbol{x} \in S$に対して，$f(\boldsymbol{x}^*) \leq f(\boldsymbol{x})$

・局所最適解：$\boldsymbol{x}^* \in S$が任意の$\boldsymbol{x} \in B(\boldsymbol{x}, \varepsilon) \cap S$に対して，$f(\boldsymbol{x}^*) \leq f(\boldsymbol{x})$

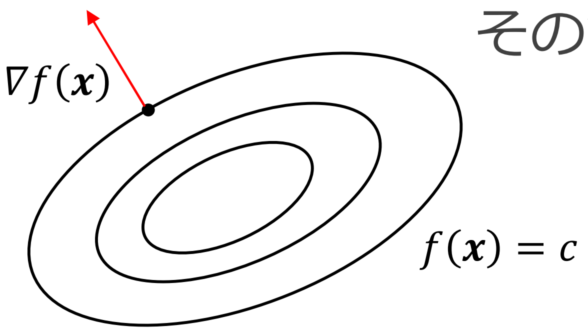


微積分の基礎

本講義で主に用いる微積分

・関数の勾配(ベクトル)$\nabla f(\boldsymbol{x}) = \left(\frac{\partial f}{\partial x_1}(\boldsymbol{x}), \frac{\partial f}{\partial x_2}(\boldsymbol{x}), \ldots, \frac{\partial f}{\partial x_n}(\boldsymbol{x})\right)^t \in \mathbb{R}^n$

・勾配ベクトルの特徴：$f(\boldsymbol{x})$の等高線($f(\boldsymbol{x}) = c$として描かれる曲線)と
その等高線上にある点における勾配ベクトルは直交する



・ヘッセ行列$\nabla^2 f(\boldsymbol{x})$：$(i, j)$成分が$\frac{\partial f}{\partial x_i \partial x_j}(\boldsymbol{x}), (i, j = 1,2, \ldots, n)$

微積分の基礎

本講義で主に用いる微積分

・テイラー展開：$f(\boldsymbol{x})$が連続微分可能であるとき，$\boldsymbol{x}, \boldsymbol{\delta} \in \mathbb{R}^n$に対して

・1次のテイラー展開：$f(\boldsymbol{x}+\boldsymbol{\delta}) = f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^t \boldsymbol{\delta}$ +($\boldsymbol{\delta}$が2次以上の項)

・2次のテイラー展開：$f(\boldsymbol{x}+\boldsymbol{\delta}) = f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^t \boldsymbol{\delta} + \boldsymbol{\delta}^t \nabla^2 f(\boldsymbol{x})^t \boldsymbol{\delta}$
+($\boldsymbol{\delta}$が3次以上の項)

$\boldsymbol{\delta} \to \boldsymbol{0}$なら大きな次数の項は無視できる(無視できるくらい小さい)

・テイラー展開の利用方法：局所的に(＝ある近傍において)任意の連続関数 $f(\boldsymbol{x})$を1次関数や2次関数として近似可能

反復法

本講義で最適解を求める際に用いる基本アルゴリズム：**反復法**

(反復法のステップ)

・STEP0：初期解$\boldsymbol{x}_0$を設定．$k \leftarrow 0$とする

・STEP1：停止条件(例：微小の定数εに対して$\|\boldsymbol{x}_{k+1} - \boldsymbol{x}_k\| < \varepsilon$)を満たすなら，
結果を出力し終了.

・**STEP2：**

・STEP3：$k \leftarrow k + 1$として，STEP1へ戻る.

非線形最適化に用いられる手法の多くが，この反復法のパターンに帰着される
(STEP2の計算部分が異なるだけのものが多い)

収束速度

- 反復法などを用いる場合，どのくらいの速さで最適解(＝局所最適解) $\boldsymbol{x}^*$に収束するのか(＝収束速度)を知りたい
  (仮定：$\lim_{k\to\infty}\|\boldsymbol{x}_k-\boldsymbol{x}^*\|=0$, つまり点列$\{\boldsymbol{x}_k\}$が$\boldsymbol{x}^*$に収束する)

- 1次収束：$\|\boldsymbol{x}_{k+1}-\boldsymbol{x}^*\|\leq c\|\boldsymbol{x}_k-\boldsymbol{x}^*\|$となる$c, (0<c<1)$が存在する.
- 超1次収束：点列$\{c_k\}$で$\lim\limits_{k\to\infty}c_k=0$に収束する点列が存在するとき,
  $\|\boldsymbol{x}_{k+1}-\boldsymbol{x}^*\|\leq c_k\|\boldsymbol{x}_k-\boldsymbol{x}^*\|$となる場合
- 2次収束： $\|\boldsymbol{x}_{k+1}-\boldsymbol{x}^*\|\leq c\|\boldsymbol{x}_k-\boldsymbol{x}^*\|^2$となる$c, (0<c<1)$が存在する.

- 収束のスピード的には，2次収束が早く，1次収束が遅い

今日の講義のまとめ

- この講義で今後解説していく内容の中で，今のうちに知っておいてほしいことを中心に学習
- 特に，反復法は制約有り無しに関わらず，非線形最適化問題を解く上での基本となる内容であることから，理解をしておくこと

- 次回：凸解析の基礎(凸関数・凸集合，主に微分可能な凸関数に関する解析)