計画数理学特論
～第13回：その他の最適化手法②
逐次2次計画法～

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

計画数理学特論
～第13回：その他の最適化手法②
逐次2次計画法～

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

今回の講義：逐次2次計画法

- 一般的な制約付きの最適化問題(

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x})$$
$$\text{subject to } \; g_i(\boldsymbol{x}) \le 0, i = 1,2,\dots,m$$
$$g_i(\boldsymbol{x}) = 0, i = m+1,\dots,p$$

- 議論をわかりやすくするため，まずは等式制約のみの最適化問題で議論する．($f(\boldsymbol{x})$，$g_i(\boldsymbol{x})$は2回連続微分可能)

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x})$$
$$\text{subject to } \; g_i(\boldsymbol{x}) = 0, i = 1,2,\dots,p$$

- この問題に対して，**凸2次計画問題を繰り返して解き**，解を更新しながら最適解を求めていく方法が**逐次2次計画法**

(参考)ニュートン法

・制約なし最適化問題：$\min_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x})$を直線探索による反復法で解く

(直線探索)

・現在の点：$\boldsymbol{x}_k \in \mathbb{R}^n$

・探索方向：$\boldsymbol{d}_k$ (ただし，$\nabla f(\boldsymbol{x}_k)^t \boldsymbol{d}_k < 0$ となるように選ぶと，$\boldsymbol{d}_k$は降下方向)

・(アルミホ条件やウルフ条件などから求めた)ステップ幅：$\alpha_k > 0$

・以上を用いて，$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k$ として次の点を設定する.

(参考)ニュートン法

・制約なし最適化問題：$\min_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x})$を直線探索による反復法で解く

(直線探索による反復法)

STEP0：初期点：$\boldsymbol{x}_0 \in \mathbb{R}^n$を設定し，$k \leftarrow 0$とする.

STEP1：停止条件が満たされるならば，$\boldsymbol{x}_k$を解として出力し，終了.

STEP2：降下方向として$\boldsymbol{d}_k$を設定.

STEP3：$\phi(\alpha) = f(\boldsymbol{x}_k + \alpha \boldsymbol{d}_k)$, $\alpha \geq 0$に対する直線探索により，ステップ幅α_kを計算.

STEP4：$\boldsymbol{x}_{k+1} \leftarrow \boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k$ と更新.

STEP5：$k \leftarrow k + 1$として，STEP1へ戻る.

(参考)ニュートン法

・制約なし最適化問題：$\min_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x})$を直線探索による反復法で解く

・最急降下法： $\boldsymbol{d}_k = -\nabla f(\boldsymbol{x}_k)$とした場合の反復法

・**ニュートン法**：ヘッセ行列$\nabla^2 f(\boldsymbol{x}_k)$が正定値行列と仮定して，

$$\boldsymbol{d}_k = -\left(\nabla^2 f(\boldsymbol{x}_k)\right)^{-1}\nabla f(\boldsymbol{x}_k), \alpha_k = 1 \text{とした場合の反復法}$$

(ニュートン法のメリット)

・局所最適解$\boldsymbol{x}^*$の近傍から反復法を始めた場合，局所最適解に2次収束する
(つまり，かなり高速に局所最適解を得ることができる)

→ このニュートン法を制約付き最適化問題へ応用する

1次の必要条件

- 等式制約のみの最適化問題

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x})$$
$$\text{subject to } \ g_i(\boldsymbol{x}) = 0, i = 1,2,\dots,p$$

- 1次の必要条件

$$\nabla f(\boldsymbol{x}^*) + \sum \lambda_i^* \nabla g_i(\boldsymbol{x}^*) = \boldsymbol{0}$$
$$g_i(\boldsymbol{x}^*) = 0, i = 1,\dots,p$$

反復法による点列生成

- 現在の点： $\left(\boldsymbol{x}^{(k)}, \boldsymbol{\lambda}^{(k)}\right)$
- $\left(\boldsymbol{x}^{(k+1)}, \boldsymbol{\lambda}^{(k+1)}\right) = \left(\boldsymbol{x}^{(k)} + \Delta\boldsymbol{x}, +\Delta\boldsymbol{s}, \boldsymbol{\lambda}^{(k)} + \Delta\boldsymbol{\lambda}\right)$

$$\rightarrow \begin{cases} \nabla f\left(\boldsymbol{x}^{(k)} + \Delta\boldsymbol{x}\right) = \nabla f\left(\boldsymbol{x}^{(k)}\right) + \nabla^2 f\left(\boldsymbol{x}^{(k)}\right)\Delta\boldsymbol{x} \\ g_i\left(\boldsymbol{x}^{(k)} + \Delta\boldsymbol{x}\right) = g_i\left(\boldsymbol{x}^{(k)}\right) + \nabla g_i\left(\boldsymbol{x}^{(k)}\right)\Delta\boldsymbol{x} \\ \nabla g_i\left(\boldsymbol{x}^{(k)} + \Delta\boldsymbol{x}\right) = \nabla g_i\left(\boldsymbol{x}^{(k)}\right) + \nabla^2 g_i\left(\boldsymbol{x}^{(k)}\right)\Delta\boldsymbol{x} \end{cases}$$

- これらの式を先ほどの1次の必要条件で$\boldsymbol{x}^* = \boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \Delta\boldsymbol{x}$, $\boldsymbol{\lambda}^* = \boldsymbol{\lambda}^{(k+1)} = \boldsymbol{\lambda}^{(k)} + \Delta\boldsymbol{\lambda}$として代入すると…

1次の必要条件

- 等式制約のみの最適化問題

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x})$$
$$\text{subject to } \; g_i(\boldsymbol{x}) = 0, i = 1,2,\dots,p$$

- 1次の必要条件

$$\nabla f\left(\boldsymbol{x}^{(k)}\right) + \left(\nabla^2 f\left(\boldsymbol{x}^{(k)}\right) + \sum \lambda_i^{(k)} \nabla^2 g_i\left(\boldsymbol{x}^{(k)}\right)\right)\Delta\boldsymbol{x} + \sum\left(\lambda_i^{(k)}\nabla g_i\left(\boldsymbol{x}^{(k)}\right)\right)$$
$$= \nabla f\left(\boldsymbol{x}^{(k)}\right) + \nabla_{\boldsymbol{xx}}^2 L\left(\boldsymbol{x}^{(k)}, \boldsymbol{\lambda}^{(k)}\right)\Delta\boldsymbol{x} + \sum\left(\lambda_i^{(k)}\nabla g_i\left(\boldsymbol{x}^{(k)}\right)\right) = \boldsymbol{0}$$
$$g_i\left(\boldsymbol{x}^{(k)}\right) + \nabla g_i\left(\boldsymbol{x}^{(k)}\right)\Delta\boldsymbol{x} = 0, i = 1,2,\dots,p$$

2次計画問題の導入

・等式制約のみの最適化問題

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x})$$
$$\text{subject to}\ \ g_i(\boldsymbol{x}) = 0, i = 1,2, \dots, p$$

に対して，以下の2次計画問題を考える(先ほどのテイラー展開の$\Delta\boldsymbol{x}$を$\boldsymbol{d}$として代入したものとみなせる(ヘッセ行列の部分だけ異なる))

$$\min_{\boldsymbol{d}\in\mathbb{R}^n}\ \ \frac{1}{2}\boldsymbol{d}^t\nabla^2_{\boldsymbol{xx}}L\left(\boldsymbol{x}^{(k)},\boldsymbol{\lambda}^{(k)}\right)\boldsymbol{d} + \nabla f\left(\boldsymbol{x}^{(k)}\right)^t\boldsymbol{d}$$
$$\text{subject to}\ \ g_i\left(\boldsymbol{x}^{(k)}\right) + \nabla g_i\left(\boldsymbol{x}^{(k)}\right)^t\boldsymbol{d} = 0, i = 1,2, \dots, p$$

2次計画問題の1次の必要条件

・2次計画問題

$$\min_{\boldsymbol{d}\in\mathbb{R}^n} \quad \frac{1}{2}\boldsymbol{d}^t\nabla^2_{\boldsymbol{xx}}L\left(\boldsymbol{x}^{(k)},\boldsymbol{\lambda}^{(k)}\right)\boldsymbol{d}+\nabla f\left(\boldsymbol{x}^{(k)}\right)^t\boldsymbol{d}$$

$$\text{subject to} \quad g_i\left(\boldsymbol{x}^{(k)}\right)+\nabla g_i\left(\boldsymbol{x}^{(k)}\right)^t\boldsymbol{d}=0, i=1,2,\dots,p$$

・1次の必要条件

$$\nabla^2_{\boldsymbol{xx}}L\left(\boldsymbol{x}^{(k)},\boldsymbol{\lambda}^{(k)}\right)\boldsymbol{d}+\nabla f\left(\boldsymbol{x}^{(k)}\right)+\sum\left(v_i\nabla g_i\left(\boldsymbol{x}^{(k)}\right)\right)=\boldsymbol{0}$$

$$g_i\left(\boldsymbol{x}^{(k)}\right)+\nabla g_i\left(\boldsymbol{x}^{(k)}\right)^t\boldsymbol{d}=0, i=1,2,\dots,p$$

→ $\boldsymbol{d}=\Delta\boldsymbol{x}, v_i=\lambda_i^{(k)}$とおけば，元々の1次の必要条件と同じ.
つまり，上記の2次計画問題を解けば元の最適化問題を解くことになる.

不等式制約の場合の2次計画問題

・不等式制約付きの最適化問題

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x})$$
$$\text{subject to}\ \ g_i(\boldsymbol{x}) \leq 0, i = 1,2,\dots,p$$

・先ほどと同様に2次計画問題を導入

$$\min_{\boldsymbol{d}\in\mathbb{R}^n}\ \ \frac{1}{2}\boldsymbol{d}^t\nabla^2_{\boldsymbol{xx}}L\left(\boldsymbol{x}^{(k)},\boldsymbol{\lambda}^{(k)}\right)\boldsymbol{d} + \nabla f\left(\boldsymbol{x}^{(k)}\right)^t\boldsymbol{d}$$
$$\text{subject to}\ \ g_i\left(\boldsymbol{x}^{(k)}\right) + \nabla g_i\left(\boldsymbol{x}^{(k)}\right)^t\boldsymbol{d} \leq 0, i = 1,2,\dots,p$$

→ $\nabla^2_{\boldsymbol{xx}}L\left(\boldsymbol{x}^{(k)},\boldsymbol{\lambda}^{(k)}\right)$が正定値行列であれば，内点法などで効率的に解ける

正定値行列でない場合：$\nabla^2_{\boldsymbol{xx}}L\left(\boldsymbol{x}^{(k)},\boldsymbol{\lambda}^{(k)}\right)$を正定値対称行列$B_k$で置き換える

KKT条件

- B_kで置き換えた2次計画問題(＝凸2次計画問題)を導入

$$\min_{\boldsymbol{d}\in\mathbb{R}^n} \quad \frac{1}{2}\boldsymbol{d}^t B_k \boldsymbol{d} + \nabla f\left(\boldsymbol{x}^{(k)}\right)^t \boldsymbol{d}$$

$$\text{subject to} \quad g_i\left(\boldsymbol{x}^{(k)}\right) + \nabla g_i\left(\boldsymbol{x}^{(k)}\right)^t \boldsymbol{d} \leq 0, i = 1,2,\dots,p$$

- KKT条件

$$B_k\boldsymbol{d} + \nabla f\left(\boldsymbol{x}^{(k)}\right) + \sum\left(v_i \nabla g_i\left(\boldsymbol{x}^{(k)}\right)\right) = \boldsymbol{0}$$

$$g_i\left(\boldsymbol{x}^{(k)}\right) + \nabla g_i\left(\boldsymbol{x}^{(k)}\right)\boldsymbol{d} \leq 0, i = 1,2,\dots,p$$

$$v_i \geq 0, \quad v_i\left(g_i\left(\boldsymbol{x}^{(k)}\right) + \nabla g_i\left(\boldsymbol{x}^{(k)}\right)^t \boldsymbol{d}\right) = 0$$

KKT条件の分析

・KKT条件

$$B_k \boldsymbol{d} + \nabla f\left(\boldsymbol{x}^{(k)}\right) + \sum \left(v_i \nabla g_i\left(\boldsymbol{x}^{(k)}\right)\right) = \boldsymbol{0}$$

$$g_i\left(\boldsymbol{x}^{(k)}\right) + \nabla g_i\left(\boldsymbol{x}^{(k)}\right)\boldsymbol{d} \leq 0, i = 1,2, \ldots, p$$

$$v_i \geq 0, \quad v_i \left(g_i\left(\boldsymbol{x}^{(k)}\right) + \nabla g_i\left(\boldsymbol{x}^{(k)}\right)^t \boldsymbol{d}\right) = 0$$

・$\boldsymbol{d} = \boldsymbol{0}$ならば，$\boldsymbol{x}^{(k)}$は元の不等式制約付き最適化問題のKKT条件となる

・$\boldsymbol{d} \neq \boldsymbol{0}$ならば，以下の関数(メリット関数)に対して，直線探索を適用する

$$f(\boldsymbol{x}) + \rho \sum max\{g_i(\boldsymbol{x}), 0\}$$

KKT条件の分析

・KKT条件

$$B_k \boldsymbol{d} + \nabla f\left(\boldsymbol{x}^{(k)}\right) + \sum \left(v_i \nabla g_i\left(\boldsymbol{x}^{(k)}\right)\right) = \boldsymbol{0}$$

$$g_i\left(\boldsymbol{x}^{(k)}\right) + \nabla g_i\left(\boldsymbol{x}^{(k)}\right)\boldsymbol{d} \leq 0, i = 1,2, \dots, p$$

$$v_i \geq 0, \quad v_i\left(g_i\left(\boldsymbol{x}^{(k)}\right) + \nabla g_i\left(\boldsymbol{x}^{(k)}\right)^t \boldsymbol{d}\right) = 0$$

・正定値対称行列B_kの更新

→ パウエルの修正BFGS公式などを適用(詳しくは，各自で調べておくこと)

逐次2次計画法のアルゴリズム

STEP1：初期点$\boldsymbol{x}^{(0)}$および初期の近似行列B_0を定め，$k = 0$とする.

STEP2： 凸2次計画問題を解いて，探索方向$\boldsymbol{d}$とラグランジュ乗数$\boldsymbol{\lambda}^{(k+1)}$を求める.

STEP3：ベクトル$\boldsymbol{d}$の大きさ$\|\boldsymbol{d}\|$が十分に小さい場合，アルゴリズムを終了.

STEP4：メリット関数に対して直線探索によりステップ幅α_kを求める.

$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha_k \boldsymbol{d}$とする.

STEP5：近似行列B_kを更新し，新たな近似行列B_{k+1}を生成する.

$k = k + 1$としてSTEP2へ戻る.

今回のまとめ

- 逐次2次計画法について，アルゴリズムも含めて解説
- 前回の内点法と同様に，逐次2次計画法も非常に有用な手法であり，適用範囲もかなり広い
- これまでの手法も含めて，基本的に非線形のままでは解きにくいため，テイラー展開による近似，直線探索，反復法をうまく組み合わせた手法が構築されている