

計画数理学特論 ～第6回：勾配法の基礎①～

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

本日の講義について

- 制約なしの非線形最適化問題で(局所)最適解を導出する方法を深掘する
- 反復法を用いながら, (局所)最適解に近づくように点列を更新していくが, その更新をどのように行えばよいかを考える
- まずは, 直線探索について諸々の条件を確認する
- その後, 特に最急降下法について練習問題を交え, 学習する.

(再掲)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**

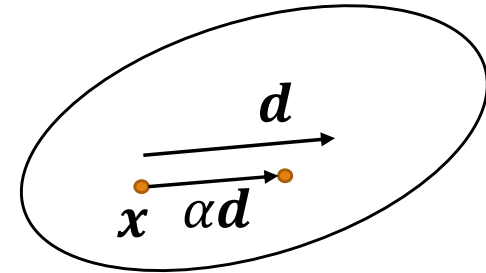
(反復法のステップ)

- STEP0 : 初期解 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 $\|x_{k+1} - x_k\| < \varepsilon$)を満たすなら, 結果を出力し終了.
- **STEP2 : 関数 $f(x)$, 実行可能領域 S , 生成される点列 x_0, x_1, \dots, x_k などの情報から, x_{k+1} を計算する.**
- STEP3 : $k \leftarrow k + 1$ として, STEP1へ戻る.

Q : 点列 x_0, x_1, \dots, x_k をどのように生成していくか?

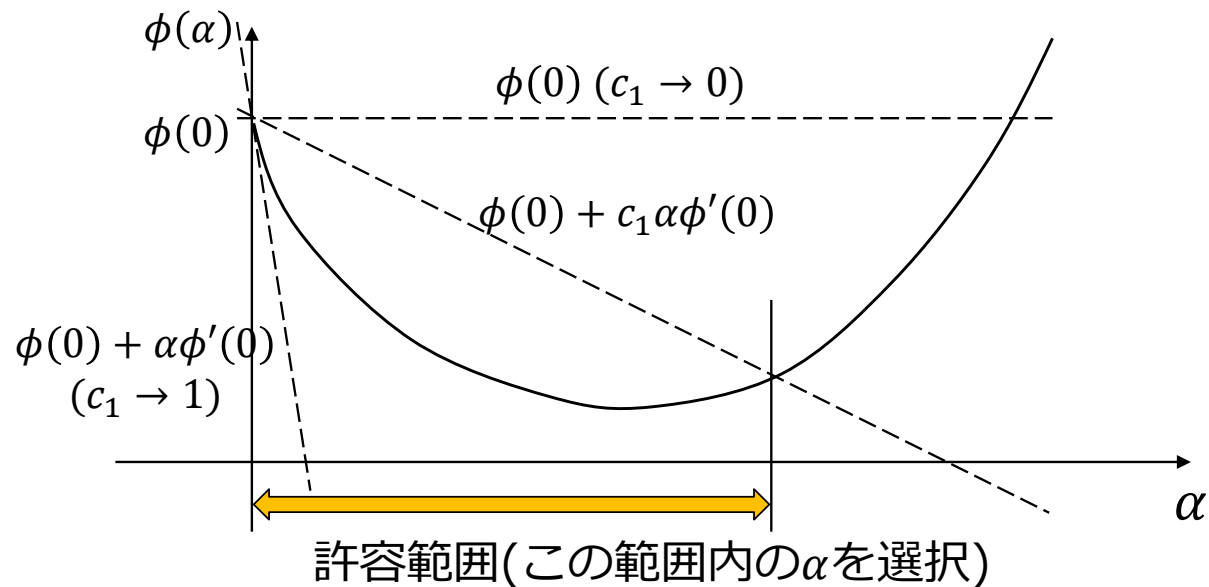
直線探索法

- 関数 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ として, 点 $x \in \mathbb{R}^n$ を d 方向に更新する状況を想定
- 1変数関数 $\phi(\alpha) =$ を設定(ただし, $\alpha > 0$)
- $\phi(0) = f(x)$: 現在の点 x における関数の値
- $\phi(0) > \phi(\alpha)$ となる $\alpha > 0$ があれば, 最小値の方向にもっていくことができる
→ できるだけ $\phi(\alpha)$ を小さくできる α を求める計算法: **直線探索法**
- 計算速度や効率性を考えると, $\phi(0) > \phi(\alpha)$ となる α が見つければOK
(もちろん $\phi(\alpha)$ の値が大きく減少すればよいが…)



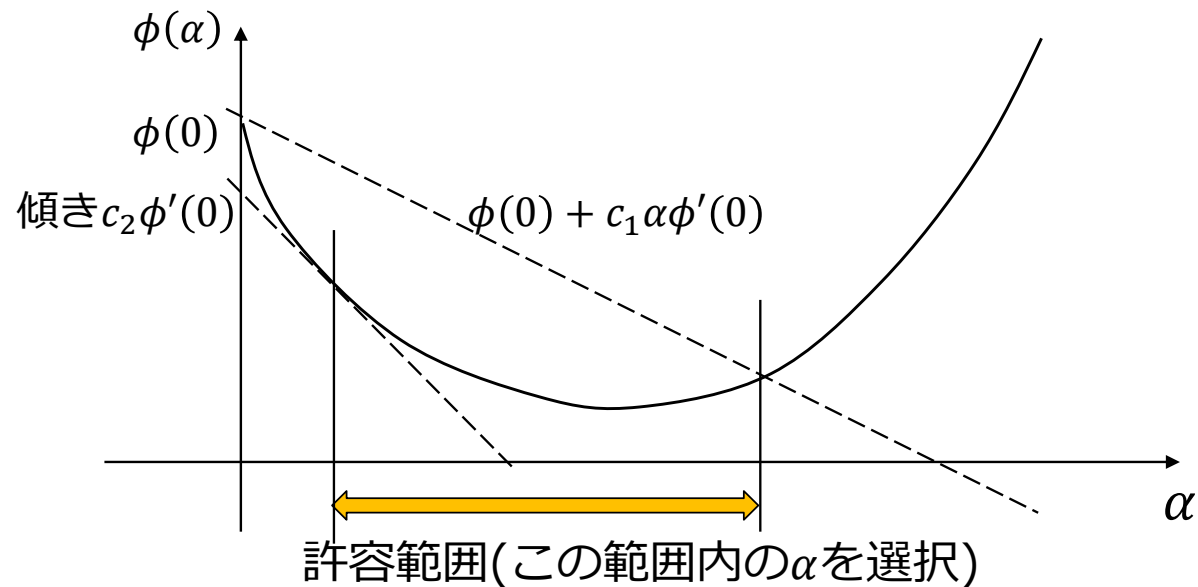
アルミホ条件(Armijo condition)

- $c_1 \in (0,1)$ として, $\phi'(\alpha)$ が $\phi(\alpha)$ の1回微分として
$$\phi(\alpha) \leq \phi(0) + c_1 \alpha \phi'(0)$$
を満たす $\alpha > 0$ を選択すれば, $\phi(0) > \phi(\alpha)$ となる.



ウルフ条件(Wolfe condition)

- $0 < c_1 < c_2 < 1$ として, $\phi'(\alpha)$ が $\phi(\alpha)$ の1回微分として
$$\phi(\alpha) \leq \phi(0) + c_1\alpha\phi'(0), \phi'(\alpha) \geq c_2\phi'(0)$$
を満たす $\alpha > 0$ を選択する.



アルミホ条件を満たす α の求め方

・アルミホ条件下を満たす α の求め方は？ → **バックトラッキング法**

STEP0 : アルミホ条件の定数 $c_1 \in (0,1)$, 縮小率 $\rho \in (0,1)$, 初期値 $\alpha > 0$ を設定

STEP1 : α がアルミホ条件を満たすなら, α をステップ幅として出力.

アルゴリズム終了

STEP2 : $\alpha \leftarrow \rho\alpha$ として, STEP1に戻る.

つまりは, アルミホ条件を満たすまで, α の値を縮小していくシンプルな方法

直線探索を用いる反復法

- 現在の点 : $\mathbf{x}_k \in \mathbb{R}^n$
- 探索方向 : \mathbf{d}_k (ただし, $\mathbf{d}_k^T \mathbf{g}_k < 0$ となるように選ぶ)
- (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, $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k$ として次の点を設定する.

- この時, $\mathbf{d}_k^T \mathbf{g}_k < 0$ となる \mathbf{d}_k を選んでいることで,
$$f(\mathbf{x}_k + \alpha_k \mathbf{d}_k) - f(\mathbf{x}_k) < 0$$
が成り立つ. (よって, \mathbf{d}_k を**降下方向**とも呼ぶ)

直線探索を用いる反復法

(反復法の流れ)

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

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

STEP2 : 降下方向として \mathbf{d}_k を設定.

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

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

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

ゾーテンダイク条件

- 関数 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ は下に有界で連続微分可能
- 初期点 \mathbf{x}_0 から定まるレベル集合 $\{\mathbf{x} \in \mathbb{R}^n \mid f(\mathbf{x}) \leq f(\mathbf{x}_0)\}$ を含む集合上で, f が γ 平滑.
- 直線探索にウルフ条件を用い, 探索方向 \mathbf{d}_k と勾配 $\nabla f(\mathbf{x}_k)$ の間の角度を

$$\cos\theta_k = -\frac{\nabla f(\mathbf{x}_k)^t \mathbf{d}_k}{\|\nabla f(\mathbf{x}_k)\| \|\mathbf{d}_k\|}$$

と定めるとき,

$$\sum_{k=0}^{\infty} (\cos\theta_k)^2 \|\nabla f(\mathbf{x}_k)\|^2 < \infty$$

が成り立つ.

ゾーテンダイク条件が言いたいこと

- ゾーテンダイク条件から, $(\cos\theta_k)^2 \|\nabla f(\mathbf{x}_k)\|^2$ の無限級数和が発散しないため,
 $(\cos\theta_k)^2 \|\nabla f(\mathbf{x}_k)\|^2 \rightarrow 0, k \rightarrow \infty$

となる.

- 降下方向 \mathbf{d}_k が勾配方向 $\nabla f(\mathbf{x}_k)$ と直交しないように,
$$\liminf_{k \rightarrow \infty} (\cos\theta_k)^2 > 0$$

と設定すると, $\nabla f(\mathbf{x}_k) \rightarrow 0, k \rightarrow \infty$ が成り立つ.

→ 初期点 \mathbf{x}_0 によらず, 勾配が0となる点に収束する(**大域的収束**)

座標降下法

- 目的関数 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ を各座標軸に沿って最適化していく方法
- ベクトル $e_i = (0, 0, \dots, 0, 1, 0, \dots, 0)^t$ (i 番目の要素だけ1, 他は0のベクトル)
- 探索方向: $\pm e_1, \pm e_2, \dots, \pm e_n$ の中から選択

(探索方向の選び方)

- ランダム
- 巡回的に $1, 2, \dots, n, 1, 2, \dots, n$ とする
- 勾配の要素の絶対値が最大となる座標軸

座標降下法のアルゴリズム

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

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

STEP2 : 降下方向として \mathbf{d}_k を, $\pm \mathbf{e}_1, \pm \mathbf{e}_2, \dots, \pm \mathbf{e}_n$ の中から選択する.

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

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

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

反復法で, STEP 2 の内容が座標降下法に合った内容に変わっただけ

最急降下法

- 勾配ベクトルが(できれば簡単に)計算可能な場合, 探索方向を降下方向として

とした場合の反復法. (勾配と逆方向を取ることで, 最も減る方向になる)

- ゴーテンダイク条件での $\cos\theta_k$ は

$$\cos\theta_k = -\frac{\nabla f(\mathbf{x}_k)^t \mathbf{d}_k}{\|\nabla f(\mathbf{x}_k)\| \|\mathbf{d}_k\|} = 1$$

となり, $\nabla f(\mathbf{x}_k) \rightarrow \mathbf{0}, k \rightarrow \infty$ が成り立つことから, 大域的収束となる.

最急降下法のアルゴリズム

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

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

STEP2 :

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

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

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

反復法で, STEP 2 の内容が最急降下法に合った内容に変わっただけ

最急降下法の例題

(例題) 関数 $f(x, y) = x^2 + y^2 + xy - 2x$ の最小化問題に対する局所最適解を初期点を $(1, 1)$ として最急降下法で求める(途中過程を示す).

(解答)

今回のまとめ

- 制約なしの非線形最適化問題で(局所)最適解を導出する方法(反復法)を用いながら, (局所)最適解に近づくように点列を更新していくが, その更新をどのように行えばよいかを考察
- ポイント①: 直線探索でステップ幅をどう設定すればよいか?
→ アルミホ条件, ウルフ条件
- ポイント②: 探索方向をどう設定すればよいか?
→ 座標降下法, 最急降下法