非線形多次元確率近似法の停止則

Yuichiro Honda
Morita lab. M1

2016/12/05

論文紹介

非線形多次元確率近似法の停止則　第 57 回自動制御連合会　2014 年 11 月 10 日-12 日

動機

- 未知の非線形方程式を雑音付き観測値から反復解放で解く（確率近似法）
- 有限回の反復で停止して得られる解候補が，真の解にどの程度近いか評価したい
- 従来研究では，各反復での計算の事後情報を利用していて事前に反復回数を見積もることができない，反復回数が十分に大きいと仮定しているなどの問題点があった
- 本論文は，1 変数の場合の議論を多変数へと拡張したものであることに注意する

定式化

標準的な確率近似法

$$x_{k+1} = x_k - a_k y_k \tag{1}$$

により，未知の非線形方程式

$$f(x) = 0 \tag{2}$$

の解を求める問題を考える（$f : \mathbb{R}^n \mapsto \mathbb{R}^n$）．ただし，$k \in \mathbb{N}$ は反復回数，

$$y_k = f(x_k) + \xi_k \tag{3}$$

は雑音付き観測値，a_k はステップサイズ（後述）である．観測雑音 $\xi_k \in \mathbb{R}^n$ は独立で同じ分布に従う確率変数ベクトルであり，平均 $\mathrm{E}[\xi_k] = 0$，有界な分散 $\mathrm{Var}[\xi_k] = \Sigma \geq 0$ を持つとする．

仮定

仮定 1

関数 f は微分可能であり，任意の x について，

$$\| \int_0^1 I - \frac{\nabla f(x^* + t(x - x^*))}{\zeta} dt \| \leq 1 - \eta \tag{4}$$

を満足する $\zeta \in (0, \infty)$, $\eta \in (0, 1)$ が存在する．

この ζ, η に対して，ステップサイズ a_k を

$$a_k = \frac{1}{\eta\zeta(k_0 + k)}$$

と定める．

主結果

定理 1

任意に与えた定数 $\alpha \in (0, \infty)$，$\beta \in (0, \infty)$ をもとに，$\bar{k} \in \mathbb{N}, k_0 \in \mathbb{N}$ が

$$\bar{k} \geq \max\{\tau_1, \tau_2\}, \qquad k_0 \geq \frac{1}{\eta} \tag{5}$$

$$\tau_1 := \frac{k_0 + 1}{\sqrt{\alpha}} - k_0, \qquad \tau_2 := \frac{4 \operatorname{tr} \Sigma}{\beta \eta^2 \zeta^2} - k_0 \tag{6}$$

を満足するように選ぶ．このとき，任意の初期解候補 $x_1 \in \mathbb{R}^n$ に対して，アルゴリズム (1) が生成する $\bar{k}$ 番目以降の解候補 x_k，$k > \bar{k}$ は

$$\mathrm{E}[\|x_k - x^*\|^2] \leq \alpha \|x_1 - x^*\|^2 + \beta \tag{7}$$

を満たす．ただし，k_0 は安定化定数とよばれる．

定理 1 の意義

定理 1(7) (再掲)

$$\mathrm{E}[\|x_k - x^*\|^2] \le \alpha\|x_1 - x^*\|^2 + \beta$$

- 事後情報に依存していない．
- 適切な反復回数 $\bar{k}$ を与えれば，解の候補 x_k と x^* の距離の二乗の期待値が指定したパラメータ以下に抑えられる．
- 特に，$\mathrm{Var}[\|x_k - x^*\|] \le \mathrm{E}[\|x_k - x^*\|^2] \le \alpha\|x_1 - x^*\|^2 + \beta$ であるから，$\bar{k}$ 番目以降の解候補の分散を抑えられる．
- 定理 1 の式 (7) を用いて，x_k と x^* の距離をより直接的に評価することができる．

より直接的な解候補の評価

補足 1

定理 1 が成り立つとき，任意の $\gamma \in (0,1)$ に対し，$\bar{k}$ 番目以降の解候補 x_k は

$$\mathrm{P}\left[\|x_k - x^*\| \leq \left(1 + \frac{1}{\sqrt{\gamma}}\right)\left(\sqrt{\alpha}\|x_k - x^*\| + \sqrt{\beta}\right)\right] \geq 1 - \gamma$$

を満たす．

チェビシェフの不等式（の特殊な場合）

確率変数 X とその平均 μ，分散 σ^2 に対し，

$$\forall k > 0; \quad P[X - \mu \geq k\sigma] \leq \frac{1}{k^2}$$

イェンゼンの不等式（の特殊な場合）

$$\mathrm{E}[X^2] \geq (E[X])^2$$

補足 1 の証明

$\|x_k - x^*\|$ は確率変数であるから，チェビシェフの不等式で $X = \|x_k - x^*\|$，$k = \frac{1}{\sqrt{\gamma}}$ とすると，

$$\mathrm{P}\left[\|x_k - x^*\| - \mathrm{E}[\|x_k - x^*\|] \geq \frac{1}{\sqrt{\gamma}}\sqrt{\mathrm{Var}[\|x_k - x^*\|]}\right] \leq \gamma$$

すなわち，

$$\mathrm{P}\left[\|x_k - x^*\| \leq \mathrm{E}[\|x_k - x^*\|] + \frac{1}{\sqrt{\gamma}}\sqrt{\mathrm{Var}[\|x_k - x^*\|]}\right] \geq 1 - \gamma \tag{8}$$

ここで，分散と期待値の関係より，

$$\sqrt{\mathrm{Var}[\|x_k - x^*\|^2]} = \sqrt{\mathrm{E}[\|x_k - x^*\|^2] - (\mathrm{E}[\|x_k - x^*\|])^2} \leq \sqrt{\mathrm{E}[\|x_k - x^*\|^2]} \tag{9}$$

また，イェンゼンの不等式より，$(\mathrm{E}[\|x_k - x^*\|])^2 \leq \mathrm{E}[\|x_k - x^*\|^2]$ であるので，

$$\mathrm{E}[\|x_k - x^*\|] \leq \sqrt{\mathrm{E}[\|x_k - x^*\|^2]} \tag{10}$$

補足 1 の証明 (2)

(9), (10) より，(8) は

$$P\left[\|x_k - x^*\| \leq \left(1 + \frac{1}{\sqrt{\gamma}}\right) \sqrt{E[\|x_k - x^*\|^2]}\right] \geq 1 - \gamma$$

となる．いま，定理 1 の式 (7) より，$E[\|x_k - x^*\|^2] \leq \alpha\|x_k - x^*\|^2 + \beta$ であったから，

$$P\left[\|x_k - x^*\| \leq \left(1 + \frac{1}{\sqrt{\gamma}}\right) \sqrt{\alpha\|x_k - x^*\|^2 + \beta}\right] \geq 1 - \gamma$$

$x, y \geq 0$ に対して，$\sqrt{x + y} \leq \sqrt{x} + \sqrt{y}$ であるため，

$$P\left[\|x_k - x^*\| \leq \left(1 + \frac{1}{\sqrt{\gamma}}\right)\left(\sqrt{\alpha}\|x_k - x^*\| + \sqrt{\beta}\right)\right] \geq 1 - \gamma$$

補足 1 の式が導かれた．

定理 1 の証明

補題 1

$$\mathrm{E}[\|\tilde{x}_k\|^2] \le \phi(k,1)\|\tilde{x}_1\|^2 + \sum_{j=1}^{k-1} \frac{\phi(k, j+1)}{\eta^2 \zeta^2 (k_0 + j)^2} \operatorname{tr} \Sigma \tag{11}$$

が成り立つ．ただし，

$$\phi(i,j) = \begin{cases} \displaystyle\prod_{l=j}^{i-1} \left(1 - \frac{1}{k_0 + l}\right)^2 & (i > j) \\ 1 & (i \le j) \end{cases}$$

である．

定理 1 の証明 (2)

補題 2

$$\mathrm{E}[\|\tilde{x}_k\|^2] \leq \left(\frac{k_0+1}{k_0+k}\right)^2 \|\tilde{x}_1\|^2 + \frac{4\,\mathrm{tr}\,\Sigma}{\eta^2\zeta^2(k_0+k)} \tag{12}$$

補題 2 と定理 1 の (5)，(6) より，$k \geq \bar{k}$ ならば，

$$\begin{aligned}
\mathrm{E}[\|\tilde{x}_k\|^2] &\leq \left(\frac{k_0+1}{k_0+k}\right)^2 \|\tilde{x}_1\|^2 + \frac{4\,\mathrm{tr}\,\Sigma}{\eta^2\zeta^2(k_0+k)} \\
&\leq \left(\frac{k_0+1}{k_0+k}\right)^2 \|\tilde{x}_1\|^2 + \frac{4\,\mathrm{tr}\,\Sigma}{\eta^2\zeta^2(k_0+k)} \\
&= \alpha\|\tilde{x}_1\|^2 + \beta
\end{aligned}$$

- 与えたパラメータに対して適当な反復回数を選ぶと，アルゴリズム停止時に得られた解候補と真の解との距離が確率的に評価できることを見た．
- α，β を小さく選べば，精度保証に必要な反復回数は多くなるものの，解候補と真の解の距離の二乗平均値，分散が α，β の式で抑えられることを確認した．