信阳师范学院学报(自然科学版) Journal of Xinyang Teachers College
第11卷 第4期 1998年10月 (Natural Science Edition) Vol 11 No. 4 Oct. 1998

“步幅常数”型随机算法收敛性分析

贾元强 杨宏志
(信阳师范学院 . 河南, 信阳 464000)

摘 要 研究了步幅为常数的随机算法: $H_{k+1}=H_k+\mu(P_k-F_kH_k)$, 得到了一弱收敛定理, 确定了收敛点与步幅 μ 之间的关系 .

关键词 收敛步幅; 弱收敛; 随机算法
分类号 $O232$

步幅 μ 为常数的随机算法:

$$H_{k+1}=H_k+\mu(P_k-F_k\cdot H_k) \tag{1}$$

在适应滤波及估计上有着广泛的应用, 因此对(1)式收敛性的研究具有深远的意义。本文就是对(1)式的收敛条件进行研究, 得到了一弱收敛定理。

记 $R=E(F_k)$, $P=E(P_k)$, H_* 满足 $R\cdot H_*=P$

定理 设 F_i 是 $N\times N$ 正定对称矩阵, P_i 是 $N\times 1$ 随机向量, 如果存在一正整数 M, 使得

$B_0)$ $M>\dfrac{N}{12}$,

$B_1)$ $\{(P_i,F_i): i\le k\}$ 与 $\{(P_j,F_j): j\ge k+M\}$ 对任意 k 是平稳且独立的;

$B_2)$ $E(\|F_k\|^{nM})<\infty$, $\forall n\le 12$;

$B_3)$ $E(|P_k|^{4+\epsilon})<\infty$,

则存在一对实正数 (μ_0,β) 使得由(1)定义的 H_k 满足

$$\lim_{k\to\infty}\sup E(|H_k-H_*|^2)\le\beta\mu,\ \forall\mu\le\mu_0.$$

定理对当 F_k, P_k 的适当矩存在时, 给出了迭代值的均方误差与迭代步幅的关系。定理的条件 B_0 至 B_3 具有可验证性及反映数据概率统计特性, 其实际意义可见一斑。

记 $V_k=H_k-H_*$, $Z_k=P_k-F_k\cdot H_*$.

$$U_{k,t}\overset{\triangle}{=\!=}(I-\mu F_t)\cdots(I-\mu F_{k+1}),\ t\ge k,$$

$$U_{k,t}=I,\ t=k \tag{2}$$

则 $E(Z_k)=0$, $V_{k+1}=(I-\mu F_k)V_k+\mu Z_k$.

由(2)可知 $V_k=U_{0,k}\cdot V_0+\mu\sum_{j=1}^{k}U_{j,k}Z_j$.

记 $W_k=U_{0,k}V_0$, $h_k=\mu\sum_{j=1}^{k}U_{j,k}Z_j$, 则

$$V_k=W_k+h_k$$

要证明定理, 先证如下两个引理

引理1 若(1)式所定义的 H_k 满足如下条件:

①对 $\forall k>0$, 序列 $\{F_i, i\le k\}$ 与 $\{F_j, j\ge k+M\}$ 是独立 .

② $E(\|F_k\|^{nM})<\infty$, $\forall n\le 12$, 则对 $\forall p>0$, 存在一正数对 (μ,L) 使得 $\forall\mu\le\mu_0$,

$$E(\|U_{r,r+p}\|^{12})$$

$$\le\left|1-12\mu E\left(\lambda_{min}\left(\sum_{i=k+1}^{r+p}F_i\right)\right)\right|+L\mu^2$$

证明 由 $U_{r,r+p}$ 的定义可知

$$\|U_{r,r+p}\|\le\left\|I-\mu\sum_{i=r+1}^{r+p}F_{i1}\right\|$$

$$+\mu^2\{\mu^{q-2}\|F_{i_1}\|\cdot\cdot\cdot\|F_{iq}\|\} \tag{3}$$

这里 $2\le q\le p$, $i,\cdots,i_q$ 是大于 r 小于等于 $r+p$ 的不同的整数。则

$$E(\|U_{r,r+p}\|^{12})\le E\left(\left\|I-\mu\sum_{t=r+1}^{r+p}F_i\right\|^{12}\right.$$

$$+\mu^2E\{\mu^{k(q-2)}\|F_{i1}\|^k\cdots\|F_{iq}\|^k\cdot$$

$$\cdot\left\|I-\mu\sum_{i=r+1}^{r+p}F_i\right\|^{12-k}\} \tag{4}$$

这里 $1\le k\le 12$, 因为

$$\|I-\mu F_i\|\le 1+\mu\|F_i\| \tag{5}$$

所以(4)式后一项括号中的项由多项式

$$\sum\mu^mE(\|F_{i_1}\|^k\cdots\|F_{i_p}\|^k\cdot\|F_i\|^n) \tag{6}$$

◦ 收稿日期 1998—03—03

所控制。这个多项式的系数是矩

$$E(\|F_{i_1}\|^{p_1}\cdots\|F_{i_k}\|^{p_k}),\ p_1,\cdots,p_k\le 12 \quad (7)$$

这里所有整数 $i_1,\cdots,i_k$ 是互不相同的,且最高指数是 12。在(7)中,把下标 $i_1,\cdots,i_k$ 分成 M 组,每组 G_j 的内部各下标相差 M,则由 Holdier 不等式

$$E(\|F_{i_1}\|^{p_1}\cdots\|F_{i_k}\|^{p_k})$$
$$\le[\prod_{j=1}^{M}E(\prod_{i\in G_j}\|F_i\|^{p_iM})]^{\frac{1}{M}} \quad (8)$$

由 B_1 知

$$E(\|F_{i1}\|^{p_1}\cdots\|F_{i_k}\|^{p_k})$$
$$\le\prod_{j=1}^{M}\prod_{i\in G_j}[E(\|F_i\|)^{p_iM}]^{\frac{1}{M}} \quad (9)$$

由(9)及 B_2 知形如(7)的所有矩

$$E(\|F_{i_1}\|^{p_1}\cdots\|F_{i_k}\|^{p_k})<\infty,\ \forall p_1,\cdots,p_k\le 12 \quad (12)$$

因为 μ 是有界的,所以多项式(6)本身是有界的。故对一些正的固定的常数 L_1

$$E(\|U_{r,r+p}\|^{12})<E(\|I-\mu\sum_{i=r+1}^{r+p}F_i\|)+L_1\mu^2 \quad (11)$$

记 $y_r^p=\sum_{i=r+1}^{r+p}F_i$ (12)

ω 表示样本空间中任意一点,$p(d\omega)$ 表示概率测度,则有

$$E(\|I-\mu y_r^p\|^{12})=\int_{\Omega}\|I-\mu y_r^p\|^{12}p(dw)$$
$$\le\int_{\{\omega:\mu\|y_r^p\|\ge1\}}[1-\mu\lambda\min(y_r^p)]^{12}p(dw)$$
$$+\int_{\{\omega:\mu\|y_r^p\|\ge1\}}(2\mu\|y_r^p\|)^{12}p(dw) \quad (13)$$

则

$$E(\|I-\mu y_r^p\|^{12}\le E\{(1-\mu\lambda_{\min}(y_r^p))^{12}\}$$
$$+(2\mu)^{12}E(\|y_r^p\|^{12}) \quad (14)$$

把(10)的结果用于(12),可得对一些正数 L_2

$$E(\|y_r^p\|^k)\le L_2,\ \forall k\le 12 \quad (15)$$

故

$$E\{(1-\mu\lambda_{\min}(y_r^p)^{12})\}$$
$$\le|1-12\mu E(\lambda_{\min}(y_r^p))+C_{12}^2\mu^2E(\|y_r^p\|^{12})$$
$$+\cdots+\mu^2E(\|y_r^p\|^{12})| \quad (16)$$

由(14)及(16)知:存在 $L_3>0$,使得对 $\forall r,\ \forall\mu\le\mu_0$,

$$E(\|I-\mu y_r^p\|^{12})\le|1-12\mu E(\lambda_{\min}(y_r^p))|+L_3\mu^2 \quad (17)$$

选择 $L=L_1+L_3$ 即可得

$$E(\|U_{r,r+p}\|^{12})$$
$$\le|1-12\mu E(\lambda_{\min}(\sum_{i=r+1}^{r+p}F_i))|+L\mu^2$$ 证毕

因 F_i 为正定矩阵,故存在 δ_0,对 $\forall p>0$,

$$E(\lambda_{\min}(y_r^p))\ge\delta_0>0 \quad (18)$$

由(18)式及引理 1 的条件选取 $P=M$,则可得存在一正数对 (δ,μ_2),$\delta\mu_0<1$,对 $\forall r,\ \forall\mu\le\mu_0$

$$E(\|U_{r,r+p}\|^{12})\le1-\delta\mu \quad (19)$$

引理 2 若假定 B_0- B_3 成立,则

$$\mu^{-1}E(|h_k|^2)$$
$$\le\mu[G_1(\mu)+G_2(\mu)+G_3(\mu)+G_4(\mu)] \quad (20)$$

这里 $G_1(\mu)=a_1[1-r(\mu)^{q_1}]^{-1}+b_1$;

$G_2(\mu)=\mu a_2[1-r(\mu)^{q_2}]^{-2}$;

$+\mu b_2[1-r(\mu)^{q_2}]^{-1}+\mu C_2$;

$G_3(\mu)=a_3[1-r(\mu)^{q_3}]^{-1}$;

$G_4(\mu)=a_4,\ r(\mu)=(1-\delta\mu)^{\frac{1}{3M}}<1$ (21)

其中 $a_1,a_2,a_3,a_4,b_1,b_2,c_2,q_1,q_2,q_3$ 是给定的常数。

证明 记

$S\triangleq\{(j,k),\ j=1,2,\cdots,t-1,k=2,3,\cdots,t,k>j\}$;

$S_2\triangleq\{(j,k)\in S: k\ge j+M,\ j\le t-2M\}$;

$S_3\triangleq\{(j,k)\in S: k<j+M,\ j\le t-2M\}$;

$S_4\triangleq\{(j,k)\in S: j>t-2M\}$ (22)

由 $h_t\triangleq\mu\sum_{j=1}^{t}U_{j,t}Z_j$ 知

$$E(|h_t|^2)=\mu^2\sum_{j=1}^{t}E(|U_{j,t}Z_j|^2)$$
$$+2\mu^2\sum_{(j,k)\in S}E(|Z_j^TU_{j,t}^TU_{k,t}Z_k| \quad (23)$$

T 表示转置,由 Schwary 不等式有

$$\sum_{j=1}^{t}E(|U_{j,t}Z_j|^2)\le\sum_{j=1}^{t}\{E(\|U_{j,t}\|^4)E(|Z_j|^4\}^{\frac{1}{2}} \quad (24)$$

由 B_2,B_3 知存在常数 B,使得 $E(|Z_j|^4)\le B<\infty$ (25)

由(23),(24),式知

$$\sum_{j=1}^{t}E(|U_{j,t}Z_j|^2\le\sum_{j=1}^{t}\{r(\mu)^{t-j}B\}^{\frac{1}{2}}$$
$$\le[2M+(1-r(\mu)^{\frac{1}{2}})^{-1}]\sqrt{3}\triangleq G_1(\mu) \quad (26)$$

现考虑(23)式的第二部分

记 $\Sigma^2\underset{(j,k)\in S_2}{\triangleq}E(Z_j^TU_{j,t}^TU_{R,t}Z_k)$;

$\Sigma^3\underset{(j,k)\in S_3}{\triangleq}E(Z_j^TU_{k,t}^TU_{k,t}Z_k)$;

$\Sigma^4\underset{(j,k)\in S_4}{\triangleq}E(Z_j^TU_{j,t}^TU_{k,t}Z_k)$;

下证 Σ^2 的有界性

$$E(Z_j^T U_{j,t}^T U_{k,t} Z_k) = E(Z_j^T U_{j,j+M-1}^T U_{j+M-1,t}^T U_{k,t} Z_k) \tag{28}$$

设 $U_{j,j+M-1}^T = I + y_j$ (29)

因在 S_2 中 $k \geq j+M$, 由 B_1 的独立性及(28)知

$$E(Z_j^T U_{j,t}^T U_{k,t} Z_k) = E(Z_j^T) \cdot E(U_{j+M-1,t}^T U_{k,t} Z_k) + E(Z_j^T Y_j U_{j,j+m-1}^T U_{k,t} Z_k) \tag{30}$$

而由 $E(Z_k) = 0$ 知

$$E(Z_j^T U_{k,t}^T Z_k) = E(Z_j^T y_j U_{j+M-1,t}^T U_{k,t} \cdot Z_k) \tag{31}$$

由定义(29)式知

$$\| y_j \| \leq \Sigma\{\mu^p \| F_{i1} \| \cdots \| F_{ip} \| \text{项}\} \tag{32}$$

$1 \leq p < M-1$, $i_1, \cdots, i_p$ 是互不相同的整数。

由引理 1 中的(10)式知存在一常数 c_1, 使得

$$E(\| y_j \|^4) \leq c_1^4 \mu^4, \forall \mu \leq \mu_0 \tag{33}$$

对(31)式利用 Holdier 不等式知

$$E(Z_j^T U_{j,t}^T U_{k,t} Z_k) = E(Z_j^T y_j U_{j+m-1,t}^T U_{k,t} Z_k) \leq \{E(|Z_j|^4 |Z_k|^4) \cdot E(\| y_j \|^4) \cdot E(\| U_{j+H-1,t} \|^4 \cdot E(\| U_{k,t} \|^4)\}^{\frac{1}{4}} \tag{34}$$

因 Z_j, Z_k 在 S_2 中是独立的, 由(33), (34)及(25)得

$$|E(Z_j^T U_{j,t}^T \cdot U_{k,t} Z_k| \leq \sqrt{B}\, \mu c_1 \{E(\| U_{j+M-1,t} \|^4) \cdot E(\| U_{k,t} \|^4\}^{\frac{1}{4}} \tag{35}$$

由(35)式知

$$|\Sigma^2| \leq \frac{1}{2} G_2(\mu) \tag{36}$$

$$G_2(\mu) = \mu \cdot 2c_1 \sqrt{B}\{M_2 + (1 - r(\mu)^{\frac{1}{4}})^{-2} + 2M(1 - r(\mu)^{\frac{1}{4}-1}\} \tag{37}$$

下证 Σ^3 的有界性, 由 Holdier 不等式

$$|\Sigma^3| \leq \sum_{(j,k)\in S_3} \{E(|Z_j|^4 \cdot E(|Z_k|^4) \cdot E(\| U_{j,t} \|^4) \cdot E(\| U_{k\cdot t} \|^4)\}^{\frac{1}{4}} \tag{38}$$

由(25)式得

$$|\Sigma^3| \leq \sum_{(j,k)\in S_3} \sqrt{B}\,(r(\mu))^{t-j} \tag{39}$$

再求和知

$$|\Sigma^3| \leq \frac{1}{2} G_3(\mu) \tag{40}$$

$$G_3(\mu) \overset{\triangle}{=\!=} 2M \sqrt{B}\,(1 - r(\mu)^{\frac{1}{4}})^{-1} \tag{41}$$

再证 Σ^4 的有界性

由 Holdier 不等式

$$|\Sigma^4| \leq \sum_{(j,k)\in S_4} \sqrt{3} = 2M\sqrt{B} \overset{\triangle}{=\!=} \frac{1}{2} G_4(\mu) \tag{42}$$

从而 $\left|\sum_{(j,k)\in S} E(Z_j^T U_{j,k}^T U_{k,t} Z_k)\right| \leq \frac{1}{2}(G_2(\mu) + G_3(\mu) + G_4(\mu))$, 即

$$\mu^{-1} E(|h_i|^2) \leq \mu[G_1(\mu) + G_2(\mu) + G_3(\mu) + G_4(\mu)] \tag{43}$$

引理 2 得证

定理的证明

第一步, 证 $\lim_{k\to\infty} E(|W_k|^2) = 0$

现把 $[j, t]$ 分成三部分

$\Gamma_1 = \{j+1, \cdots, j+M\} \cup \{j+2M+1, \cdots, j+3M\} \cup \cdots;$

$\Gamma_2 = \{j+M+1, \cdots, j+2M\} \cup \{j+3M+1, \cdots, j+4M\} \cup \cdots;$

$\Gamma_3 = \{j+2nM+1, j+2nM+2, \cdots, t\}$, 这里 $n = [\frac{t-j}{2M}]$ (44)

由 $U_{k,t}$ 的定义知

$$U_{j,t} = U_{m,t} \cdot U_{r,m} \cdot U_{j,r},\ j \leq r \leq m \leq t \tag{45}$$

$$E(\| U_{j,t} \|^4 = E[\prod_{i=0}^{n-1}(\| U_{j+2M,j+(2i+1)M} \|^4 \cdot \| U_{j+(2i+1)M,j+(2i+2)M} \|^4) \cdot \| U_{j+2nM,t} \|^4] \tag{46}$$

由 Holdier 不等式得

$$E(\| U_{j,t} \|^4) \leq \{E(\prod_{i=0}^{n-1} \| U_{j+2M,j+(2i+1)M} \|^{12}\}^{\frac{1}{3}} \cdot \{E(\| U_{j+2nM,t} \|^{12})\}^{\frac{1}{3}} \cdot \{E(\prod_{i=0}^{n-1} \| U_{j+(2i+1)M,j+(2i+2)M} \|^{12})\}^{\frac{1}{3}} \tag{47}$$

由假定 B_1 的 $M-$ 无关性及 Γ_2、Γ_2 的结构知

$$E(\| U_{j,t} \|^4) \leq \{\prod_{i=0}^{n=1} E(\| U_{j+2M,j+(2i+1)M} \|^{12})\}^{\frac{1}{3}} \cdot \{E \| U_{j+2nM,t} \|^{12})\}^{\frac{1}{3}} \cdot \{\prod_{i=0}^{n=1} E(\| U_{j+(2i+1)M,j+(2i+2)M} \|^{12})\}^{\frac{1}{3}} \tag{48}$$

由(19)式知

$$E(\| U_{j,t} \|^4) \leq \{\prod_{i=0}^{n-1}(1 - \delta\mu)\}^{\frac{1}{3}}(1 - \delta u)^{\frac{1}{3}} \cdot \{\prod_{i=0}^{n-1}(1 - \delta\mu)\}^{\frac{1}{3}} = (1 - \delta\mu)^{\frac{1}{3}(2n+1)} \tag{49}$$

由 Schwarz 不等式

$$E(|W_k|^2) \leq \{E(\| U_{0,k} \|^4)\}^{\frac{1}{2}} \cdot \{E(|V_0|^4)\}^{\frac{1}{2}} \leq (r(\mu)^k)^{\frac{1}{2}} \cdot \{E(|V_0|^4)\}^{\frac{1}{2}}$$

而 $r(\mu) = (1 - \delta\mu)^{\frac{1}{3M}} < 1$, 所以 $\lim_{k\to\infty} E(|W_k|^2) = 0$.

第二步, 存在一正数 β^1 使得

$$\lim_{k\to\infty} \mu^{-1} E(|h_k|^2 \leq \beta^1,$$

由(19)式知

$$\lim_{k\to\infty} \mu[1 - r(\mu)^q]^{-1} < \infty, \forall q > 0$$

则

$$\mu^{-1} E(|h_k|^2 \leq \mu[G_1(\mu) + G_2(\mu)$$

$+ G_3(\mu) + G_4(\mu)] \leq \beta^1$,

从而 $E(|V_k|^2) \leq 2E(|W_k|^2) + 2E(|h_k|^2)$, 则

$\lim\limits_{k\to\infty} \sup E(|v_k|^2)$

$\leq 2\lim\limits_{k\to\infty} \sup E(|h_k|^2)$

$\leq 2\beta^1\mu \overset{\triangle}{=\!=} \beta\mu$

定理得证。

参 考 文 献

1 Kushner H J and Shwartz. *Convergece of Stochastic Approximations with State Dependent Noise Under Weak Condtions.* C D C IEEE, 1982, 517～521.

2 Farden D C. *Stochastic Approximation with Correlated Date.* IEEE Trans Inform Theo, 1981, (2): 105～113.

3 MA D C, Makoski A M and Shwartz A. *Stockastic Approximations for Finite-state Markov Chains.* Stoch, Proce and their Appli, 1990, (35): 27～45.

Studies on the Convergence of Stochastic Algorithm for the Constant Step Size

Jia Yuanqiang Yang Hongzhi

(Dept of Math, Xinyang Teachers College. Xinyang, Henan; China 464000)

Abstract The stochastic algorthm for the constant step size $H_{k+1} = H_k + \mu(p_k - F_k H_k)$ is discussed and a theorem of weak convergence is proved.

Key words Step size; Weak convergence; Stochastic Algorithm

责任编辑 郭红建