Kiefer, Wolfowitz: Stocastic Estimation Of The Maximum Of A Regression Function

1 Introduction

$Y \in \mathbb{R}$ is a random variable which depends on a parameter $x \in \mathbb{R}$. Let

$$M(x) = \mathbb{E}[Y(x)]$$

and assume $Var(Y(x)) \leq S$ $(\forall x \in \mathbb{R})$.

Close form of $M(x)$ and distribution of Y is not known. The only way to calculate $M(x)$ is to draw random samples from $H(y|x)$ which is the cdf of Y given x.

This paper presents the following procedure to maximize $M(x)$ via sampling $Y(x)$.

$$z_{n+1} = z_n + a_n \frac{y_{2n} - y_{2n-1}}{c_n}$$

where z_1 is an arbitrary number, y_{2n}, y_{2n-1} are sampled from $H(y|z_n + c_n), H(y|z_n - c_n)$ respectively. a_n, c_n are predetermined sequences of positive numbers.

It can be proved z_n converges to the optimal solution of maximizing $M(x)$ in probability under regularity conditions on $M(x), a_n, c_n$.

2 Regularity conditions

2.1 Regularity conditions on a_n, c_n

- $\lim_{n\to\infty} c_n = 0$.
- $\sum_{n=1}^{\infty} a_n = \infty$.
- $\sum_{n=1}^{\infty} a_n c_n < \infty$.
- $\sum_{n=1}^{\infty} \frac{a_n^2}{c_n^2} < \infty$.

2.2 Regularity conditions on $M(x)$

- $M(x)$ is maximized at $x = \theta$ and $M(x)$ strictly increasing when $x < \theta$, strictly decreasing when $x > \theta$.
- There exist $\beta > 0$ and $B > 0$ such that

$$|x' - \theta| + |x'' - \theta| < \beta \text{ implies } |M(x') - M(x'')| < B|x' - x''|$$

- There exist $\rho > 0$ and $R > 0$ such that

$$|x' - x''| < \rho \text{ implies } |M(x') - M(x'')| < R$$

- For every $\delta > 0$ there exists $\pi(\delta) > 0$ such that

$$|z - \theta| > \delta \text{ implies } \inf_{\frac{1}{2}\delta > \epsilon > 0} \frac{|M(z+\epsilon) - M(z-\epsilon)|}{\epsilon} > \pi(\delta)$$

3 Proof of convergence

Prove $z_n \xrightarrow{P} \theta$ when regularity conditions on $a_n, c_n, M(x)$ hold.

$$\begin{aligned}\mathbb{E}[(z_{n+1}-\theta)^2] &= \mathbb{E}[(z_n-\theta+a_n\frac{y_{2n}-y_{2n-1}}{c_n})^2] \\ &= \mathbb{E}[(z_n-\theta)^2]+2\frac{a_n}{c_n}\mathbb{E}[(z_n-\theta)(y_{2n}-y_{2n-1})]+\frac{a_n^2}{c_n^2}\mathbb{E}[(y_{2n}-y_{2n-1})^2] \\ &= \mathbb{E}[(z_1-\theta)^2]+2\sum_{j=1}^{n}\frac{a_j}{c_j}\mathbb{E}[(z_j-\theta)(y_{2j}-y_{2j-1})]+\sum_{j=1}^{n}\frac{a_j^2}{c_j^2}\mathbb{E}[(y_{2j}-y_{2j-1})^2]\end{aligned}$$

Check sign of $\mathbb{E}[(z_n-\theta)(y_{2n}-y_{2n-1})]$.

$$\begin{aligned}\mathbb{E}[(z_n-\theta)(y_{2n}-y_{2n-1})] = \mathbb{E}[(z_n-\theta)\mathbb{E}[(y_{2n}-y_{2n-1})|z_n]] &= \mathbb{E}[(z_n-\theta)(M(z_n+c_n)-M(z_n-c_n))] \\ &= \mathbb{E}[\frac{z_n-\theta}{2c_n}\cdot 2c_n\cdot(M(z_n+c_n)-M(z_n-c_n))]\end{aligned}$$

$$|z_n-\theta|\geq c_n \Rightarrow (z_n-\theta)(M(z_n+c_n)-M(z_n-c_n))\leq 0$$

Separate sign of $\mathbb{E}[(z_n-\theta)(y_{2n}-y_{2n-1})]$. Denote $U_n(x)=(x-\theta)(M(x+c_n)-M(x-c_n))$, $U_n^+(x)=\frac{1}{2}(U_n(x)+|U_n(x)|)$, $U_n^-(x)=\frac{1}{2}(U_n(x)-|U_n(x)|)$.

$$\mathbb{E}[(z_n-\theta)(y_{2n}-y_{2n-1})]=\mathbb{E}[U_n(z_n)]=\mathbb{E}[U_n^+(z_n)+U_n^-(z_n)]$$

Therefore

$$\begin{aligned}\mathbb{E}[(z_{n+1}-\theta)^2] &= \mathbb{E}[(z_1-\theta)^2]+2\sum_{j=1}^{n}\frac{a_j}{c_j}\mathbb{E}[U_j^+(z_j)]+2\sum_{j=1}^{n}\frac{a_j}{c_j}\mathbb{E}[U_j^-(z_j)]+\sum_{j=1}^{n}\frac{a_j^2}{c_j^2}\mathbb{E}[(y_{2j}-y_{2j-1})^2] \\ &\triangleq \mathbb{E}[(z_1-\theta)^2]+2\sum_{j=1}^{n}\frac{a_j}{c_j}P_j+2\sum_{j=1}^{n}\frac{a_j}{c_j}N_j+\sum_{j=1}^{n}\frac{a_j^2}{c_j^2}e_j\end{aligned} \tag{3.1}$$

$\sum_{j=1}^{n}\frac{a_j}{c_j}P_j$ and $\sum_{j=1}^{n}\frac{a_j^2}{c_j^2}e_j$ converge $\Rightarrow \lim_{n\to\infty}\mathbb{E}[(z_{n+1}-\theta)^2]$ exist.

Firstly, prove convergence of $\sum_{j=1}^{n}\frac{a_j}{c_j}P_j$.

$$U_n^+(z_n)>0\Leftrightarrow U_n(z_n)>0\Rightarrow |z_n-\theta|<c_n$$

When n is large enough,

$$|z_n+c_n-\theta|+|z_n-c_n-\theta|=2c_n<\beta$$
$$\Downarrow$$
$$|M(z_n+c_n)-M(z_n-c_n)|<2Bc_n$$
$$\Downarrow$$
$$\frac{a_n}{c_n}P_n\leq 2Ba_nc_n$$

Therefore $\sum_{j=1}^{n}\frac{a_j}{c_j}P_j$ converge.

Secondly, prove convergence of $\sum_{j=1}^{n}\frac{a_j^2}{c_j^2}e_j$.

$$e_n=\mathbb{E}[(y_{2n}-y_{2n-1})^2]=\mathbb{E}\{\mathbb{E}[(y_{2n}-y_{2n-1})^2|z_n]\}$$

$$\begin{aligned}\mathbb{E}[(y_{2n}-y_{2n-1})^2|z_n] &= \mathbb{E}[\{y_{2n}-M(z_n+c_n)-(y_{2n-1}-M(z_n-c_n))+M(z_n+c_n)-M(z_n-c_n)\}^2|z_n] \\ &= \mathbb{E}[(y_{2n}-M(z_n+c_n))^2|z_n]+\mathbb{E}[(y_{2n-1}-M(z_n-c_n))^2|z_n]+[M(z_n+c_n)-M(z_n-c_n)]^2 \\ &\leq 2S+R^2 (\text{when n is large enough})\end{aligned}$$

Therefore

$$\frac{a_n^2}{c_n^2}e_n\leq (2S+R^2)\frac{a_n^2}{c_n^2}$$

which implies $\sum_{j=1}^{n}\frac{a_j^2}{c_j^2}e_j$ converge.

In conclusion, $\lim_{n\to\infty}\mathbb{E}[(z_{n+1}-\theta)^2]$ exist.

Next show there exist a subsequnce $z_{n_j}(n_1 < n_2 < n_3 < \cdots)$ of z_n, $z_{n_j} \xrightarrow{P} \theta(j \to \infty)$. Let

$$K_n = |\frac{M(z_n + c_n) - M(z_n - c_n)}{c_n}|$$

Then

$$\mathbb{E}[K_n|z_n - \theta|] = \frac{P_n - N_n}{c_n}$$

Due to the convergence of $\sum_{j=1}^{n} a_j \frac{P_j - N_j}{c_j}$ and divergence of $\sum_{j=1}^{n} a_j$,

$$\liminf_{n\to\infty} \frac{P_n - N_n}{c_n} = 0$$

By the defination of limit inferior, denote $H_n = \inf_{k>n} \frac{P_k - N_k}{c_k}$, then $\lim_{n\to\infty} H_n = \liminf_{n\to\infty} \frac{P_n - N_n}{c_n}$.

Obviously, $\lim_{n\to\infty} H_n \geq 0$ or $\lim_{n\to\infty} H_n = \infty$.

Assume $\lim_{n\to\infty} H_n = \mu > 0$

$$\text{take } \epsilon = \frac{\mu}{2}, \text{then } \exists N, \text{ when } n \geq N, \ H_n \geq \mu - \epsilon = \frac{\mu}{2}$$

which is

$$\frac{P_n - N_n}{c_n} > \frac{\mu}{2} (\forall n > N)$$

then

$$\sum_{j=N}^{\infty} a_j \frac{P_j - N_j}{c_j} \geq \frac{\mu}{2} \sum_{j=N}^{\infty} a_n$$

diverge, which is a contradiction.

Therefore there exists a sequence of positive integers $n_1 < n_2 < n_3 < \cdots$ such that

$$\lim_{j\to\infty} E\left[K_{n_j} \left|z_{n_j} - \theta\right|\right] = 0$$

Moreover, $z_{n_j} \xrightarrow{P} \theta(j \to \infty)$.

Assume $z_{n_j} \overset{P}{\nrightarrow} \theta$. Then

$$\exists \epsilon, \eta > \quad \mathbb{P}(|z_{n_j} - \theta| > \eta) \geq \epsilon$$

Let j be large enough, $c_{n_j} \leq \frac{\eta}{2}$. By regularity conditions on $M(x)$,

$$K_{n_j} \geq \inf_{\frac{\eta}{2} > \epsilon > 0} \frac{|M(z_{n_j} + \epsilon) - M(z_{n_j} - \epsilon)|}{\epsilon} > \pi(\eta) \text{ when } |z_{n_j} - \theta| > \eta \text{ hold.}$$

$$E\left[K_{n_j} \left|z_{n_j} - \theta\right|\right] \geq E\left[K_{n_j} \left|z_{n_j} - \theta\right| I_{\{|z_{n_j} - \theta| > \eta\}}\right] > \eta\pi(\eta)\mathbb{E}[I_{\{|z_{n_j} - \theta| > \eta\}}] \geq \eta\pi(\eta)\epsilon > 0$$

Which is a contradiction.

Finally use the defination of converge in probability to prove $z_n \xrightarrow{P} \theta$ which is to prove

$$\forall \epsilon\ ,\eta > 0,\ \exists N,\ \forall n > N\ \mathbb{P}[|z_n - \theta| > \eta] \leq \epsilon$$

$$\begin{aligned}\mathbb{P}[|z_n - \theta| > \eta] &= \mathbb{P}[|z_n - \theta| > \eta||z_{N_0} - \theta| < s]\mathbb{P}[|z_{N_0} - \theta| < s] + \mathbb{P}[|z_n - \theta| > \eta||z_{N_0} - \theta| \geq s]\mathbb{P}[|z_{N_0} - \theta| \geq s] \\ &\leq \mathbb{P}[|z_n - \theta| > \eta||z_{N_0} - \theta| < s] + \mathbb{P}[|z_{N_0} - \theta| \geq s]\end{aligned}$$

Firstly, let s be a positive number which is undetermined. Due to $z_{n_j} \xrightarrow{P} \theta$, choose N_0 which relies on s, ϵ to make $\mathbb{P}[|z_{N_0} - \theta| \geq s] \leq \frac{\epsilon}{2}$. Note that N_0 is not determined totally in this step.

Next apply markov inequality to connect $\mathbb{P}[|z_n - \theta| > \eta||z_{N_0} - \theta| < s]$ with second moment which has been explored previously.

$$\mathbb{P}[|z_n - \theta| > \eta||z_{N_0} - \theta| < s] \leq \frac{\mathbb{E}[(z_n - \theta)^2||z_{N_0} - \theta| < s]}{\eta^2}$$

$$\begin{aligned}\mathbb{E}[(z_n - \theta)^2|z_{N_0} = z] &= \mathbb{E}[(z_{n-1} - \theta + \frac{a_{n-1}(y_{2n-2} - y_{2n-3})}{c_{n-1}})^2|z_{N_0} = z] \\ &= \mathbb{E}[(z_{n-1} - \theta)^2|z_{N_0} = z] + 2\frac{a_{n-1}}{c_{n-1}}\mathbb{E}[(z_{n-1} - \theta)(y_{2n-2} - y_{2n-3})|z_{N_0} = z] \\ &\quad + \frac{a_{n-1}^2}{c_{n-1}^2}\mathbb{E}[(y_{2n-2} - y_{2n-3})^2|z_{N_0} = z] \\ &= (z - \theta)^2 + 2\sum_{j=N_0}^{n-1} \frac{a_j}{c_j}\mathbb{E}[(z_j - \theta)(y_{2j} - y_{2j-1})|z_{N_0} = z] + \sum_{j=N_0}^{n-1} \frac{a_j^2}{c_j^2}\mathbb{E}[(y_{2j} - y_{2j-1})^2|z_{N_0} = z] \\ &= (z - \theta)^2 + 2\sum_{j=N_0}^{n-1} \frac{a_j}{c_j}\mathbb{E}[U_j(z_j)|z_{N_0} = z] + \sum_{j=N_0}^{n-1} \frac{a_j^2}{c_j^2}\mathbb{E}[(y_{2j} - y_{2j-1})^2|z_{N_0} = z] \\ &\leq (z - \theta)^2 + 4B\sum_{j=N_0}^{\infty} a_j c_j + (2S + R^2)\sum_{j=N_0}^{\infty} \frac{a_j^2}{c_j^2}(\text{when } N_0 \text{ is large enough}) \\ &< (z - \theta)^2 + s(\text{Let } N_0 \text{ be larger to make } 4B\sum_{j=N_0}^{\infty} a_j c_j < \frac{s}{2}, (2S + R^2)\sum_{j=N_0}^{\infty} \frac{a_j^2}{c_j^2} < \frac{s}{2})\end{aligned}$$

Therefore

$$\mathbb{E}[(z_n - \theta)^2||z_{N_0} - \theta| < s] = \mathbb{E}[\mathbb{E}[(z_n - \theta)^2|z_{N_0}]||z_{N_0} - \theta| < s] \leq \mathbb{E}[(z_{N_0} - \theta)^2 + s||z_{N_0} - \theta| < s] < s^2 + s$$

Choose s to make $\frac{s^2+s}{\eta^2} \leq \frac{\epsilon}{2}$.

Combine with $\mathbb{P}[|z_{N_0} - \theta| \geq s] \leq \frac{\epsilon}{2}$, then proof is done.