局所線形回帰を用いた条件付き密度推定の改良

46222 福地 崇史
指導教員　駒木 文保 助教授

2006 年 2 月 7 日

1　はじめに

密度関数の推定の良さを表す値として平均積分 2 乗誤差と呼ばれるものがある．平均積分 2 乗誤差は小さいほうが好まれ，真の値と推定値とのバイアスおよび分散が小さいほど平均積分 2 乗誤差は小さくなる．Fan et al. [2] により局所線形回帰を用いた条件付き密度の推定が提案された．この手法の漸近的なバイアスはバンド幅に関して 2 次のオーダーになることが知られている．本論文では局所線形回帰を用いた条件付き密度の推定に改良を加えて，漸近的なバイアスを 4 次のオーダーに下げる手法を提案する．

2　カーネル関数

本論文におけるカーネル関数 $W(u)$ は簡単のため以下の性質を持つものを仮定する．

- $W(u) \geq 0$
- $\int_{-\infty}^{\infty} W(u)\mathrm{d}u = 1$
- $W(u)$ は 0 に関して対称.

また以下の記法を定義しておく.

- $W_{h_1}(u) = h_1^{-1} W(u/h_1)$
- $\omega_j = \int u^j W(u)\mathrm{d}u$

上記のバンド幅 h_1 はどのくらいの範囲のデータを採用するかを決めるパラメータである.

3　局所線形回帰

互いに独立な n 個のペアのデータ $(X_1, Y_1), \ldots, (X_n, Y_n)$ に対して $Y = m(X) + \epsilon$ で与えられるモデルを考える．$m(x) = E(Y|X = x)$ は未知である．局所線形回帰では,

$$S(a(x), b(x)) = \sum_{i=1}^{n} \{Y_i - a(x) - b(x)(X_i - x)\}^2 W_{h_1}(X_i - x)$$

を最小にする $\hat{a}(x)$ を推定値 $\hat{m}(x)$ とする.

4　条件付き密度の推定

互いに独立な n 個のペアのデータ $(X_1, Y_1), \ldots, (X_n, Y_n)$ に対して，$X = x$ を与えたときの，$Y = y$ の条件付き密度 $g(y|x)$ を推定することを考える．Fan et al. [2] は，$g(y|x)$ を推定する問題に対して，$h_2 \to 0$ のとき

$$\mathrm{E}[K_{h_2}(Y - y)|X = x] \approx g(y|x) \qquad (1)$$

となることに注目した．ただし $K(u)$ はカーネル関数とする．(1) 式の左辺について考える．この式は y, h_2 を固定すると，一般の回帰問題に帰着できる．そこで局所線形回帰を用いて，この回帰問題を解くと $\beta(x, y) = (\beta_0(x, y), \beta_1(x, y))$ として

$$S(\beta(x, y)) = \sum_{i=1}^{n} \{K_{h_2}(Y_i - y) - \beta_0(x, y) - \beta_1(x, y)(X_i - x)\}^2 W_{h_1}(X_i - x)$$

を最小化する $\hat{\beta} = (\hat{\beta}_0(x, y), \hat{\beta}_1(x, y))$ の $\hat{\beta}_0(x, y)$ が推定値 $\hat{g}(y|x)$ となる．バイアスを計算すると

$$\begin{aligned}\mathrm{E}[\hat{g}(y|x)|X_1, \ldots, X_n] = g(y|x) &+ \frac{1}{2}\omega_2 \frac{\partial^2 g(y|x)}{\partial x^2} h_1^2 \\ &+ \frac{1}{2}\kappa_2 \frac{\partial^2 g(y|x)}{\partial y^2} h_2^2 + \mathrm{o_p}(h_1^2) + \mathrm{o}(h_2^2)\end{aligned}$$

となる．ただし，$\kappa_j = \int u^j K(u)\mathrm{d}u$ とする．

5　重み付き平均を用いた改良

Fan et al. [2] による手法のオーダーが 2 次であること，y, h_2 を固定すると一般の回帰問題に帰着できることを述べた．そこで Choi and Hall [1] による手法を今回の条件付き密度の推定の場合に応用して，バイアスを求めた．

まず Choi and Hall [1] による手法を応用したときの推定法を述べる．$(\hat{\beta}_0(x,y), \hat{\beta}_1(x,y))$ を用いて次の $\hat{m}(u|x,y)$ を導入する．

$$\hat{m}(u|x,y) = \hat{\beta}_0(x,y) + \hat{\beta}_1(x,y)(u-x)$$

次に任意の $\lambda > 0$ に対して

$$\begin{aligned} l &= l(\lambda) \\ &= \{(1+2\lambda)\omega_2/(2\lambda)\}^{\frac{1}{2}} \end{aligned}$$

を導入し

$$\begin{aligned} \tilde{g}(y|x) = \{&\lambda\hat{m}(x|x-lh_1,y) + \hat{m}(x|x,y) \\ &+\lambda\hat{m}(x|x+lh_1,y)\}/(2\lambda+1) \end{aligned} \tag{2}$$

とするとバイアスは,

$$\begin{aligned} \mathrm{E}\{\tilde{g}(y|x)|X_1,\ldots,X_n\} &= g(y|x) + A(x,y)h_1^4 \\ &+ B(x,y)h_2^2 + \mathrm{o_p}(h_1^4) + \mathrm{O}(h_2^4) \end{aligned}$$

となる．ただし

$$\begin{aligned} A(x,y) = \{16f(x)\}^{-1}\Big[&2(\omega_2^2-\omega_4)\Big\{2f''(x)\frac{\partial^2 g(y|x)}{\partial x^2} \\ &+4f'(x)\frac{\partial^3 g(y|x)}{\partial x^3} + f(x)\frac{\partial^4 g(y|x)}{\partial x^4}\Big\} \\ &-\lambda^{-1}\omega_2^2 f(x)\frac{\partial^2 g(y|x)}{\partial x^2}\Big], \\ B(x,y) = \frac{1}{2}\kappa_2&\frac{\partial^2 g(y|x)}{\partial y^2} \end{aligned}$$

とする．

6　幾何平均を用いた改良

さらにオーダーを下げるために改良を加える．バンド幅 h_1, h_2 とし（2）式により得られる推定値を $\tilde{g}_{h_1,h_2}(y|x)$ と書くこととする．このとき $c > 0$ なる任意の c に対して

$$c_1 = c, c_2 = c^2, \alpha = \frac{c^4}{c^4-1}$$

とし

$$\tilde{g}_{\mathrm{new}}(y|x) = (\tilde{g}_{h_1,h_2}(y|x))^{\alpha}(\tilde{g}_{c_1h_1,c_2h_2}(y|x))^{1-\alpha}$$

を推定値とする．（ただし積分して 1 にならないため，基準化が必要.）このときのバイアスは

$$\mathrm{E}[\tilde{g}_{\mathrm{new}}(y|x)|X_1,\ldots,X_n] = g(y|x) + \mathrm{o_p}(h_1^4) + \mathrm{O}(h_2^4)$$

となる．

7　まとめと今後の課題

局所線形回帰を用いた条件付き密度の推定を

1. 重み付き平均
2. 幾何平均

という改良を加えて，漸近的なバイアスをバンド幅の 4 次のオーダーに下げた．

今後の課題として，バンド幅の選択という問題が考えられる．一般の局所線形回帰でのバンド幅の選択法はクロスバリデーションやブートストラップなど，よりデータを反映させた手法がある．よって今回の手法にも，そういったバンド幅の選び方が考えられる．

参考文献

[1] Choi, E. and Hall, P., On bias reduction in local linear smoothing. *Biometrika* **85,** 2, pp. 333-345, 1998.

[2] Fan, J., Yao, Q. and Tong, H., Esimation of conditional densities and sensitivity measures in nonlinear dynamical systems. *Biometrika* **83,** 1, pp. 189-206, 1996.

[3] Hyndoman, R. J. and Yao, Q., Nonparametric estimation and symmetry testes for conditional density funcutions. *Nonparametric Statistics* **14,** 3, pp. 259-278, 2002.

[4] Terrell, G. R. and Scott, D. W., On improving convergence rates for nonnegative kernel density estimators. *Annals of Statistics* **8,** pp. 1160-1163, 1980.