VỀ MỘT XẤP XỈ ƯỚC LƯỢNG BAYES TRONG MÔ HÌNH THỐNG KÊ PHI TUYẾN

Ung Ngọc Quang

Đại học tổng hợp Tp. HOCHIMINH

I. Mở đầu

Bài toán ước lượng tối ưu trong các mô hình thống kê tuyến tính và phi tuyến đã được khảo sát nhiều (xem [1,2]). Trong [3] đã chứng minh sự tồn tại ước lượng Bayes cho mô hình phi tuyến 1-chiều. Bài này nới rộng kết quả trên sang trường hợp nhiều chiều và đưa ra cách xấp xỉ ước lượng Bayes trong mô hình trên.

2. Về sự tồn tại ước lượng Bayes trong mô hình phi tuyến nhiều chiều

Trước hết ta đưa ra vài ký hiệu thường dùng

- $M(n \times q)$, $M(p \times r)$ - các không gian tất cả các ma trận cấp $n \times q$ và $p \times r$.
- $\mathcal{B}(n \times q)$, $\mathcal{B}(p \times r)$ - Các σ-đại số Borel trên các không gian $M(n \times q)$ và $M(p \times r)$.
- R^n, R^p - Các không gian Euclide n-chiều và p-chiều.

Xét mô hình thống kê có dạng sau

$$X = \varphi(\theta) + \varepsilon, \tag{1}$$

trong đó:

- X - phần tử quan trắc ngẫu nhiên có trị trong $M(n \times q)$.
- ε - phần tử sai ngẫu nhiên có trị trong $M(n \times q)$.
- θ - tham số chưa biết, $\theta \in \Theta$, với Θ là tập compact thuộc $M(p \times r)$.
- φ: hàm phi tuyến cho trước, $\varphi : \Theta \to M(n \times q)$.

Mô hình (1) được gọi là mô hình phi tuyến r-chiều với không gian tham compact. Nếu $q = 1$, $r = 1$ (tức là $M(n \times q) = R^n$, $M(p \times r) = r^p$) thì (1) được gọi là mô hình phi tuyến 1-chiều với không gian tham compact. Mô hình này được khảo sát trong [3].

Định nghiã 2.1. Hàm $h : M(n \times q), \mathcal{B}(n \times q)) \to M(p \times r), \mathcal{B}(p \times r))$ gọi là ước lượng của tham ẩn $\theta \in \Theta \subset M(p \times r)$ nếu h là hàm Borel đo được.

Hàm Borel đo được h gọi là bị chặn nếu

$$\sup_{x \in M(n \times q)} \|h(x)\|_{M(p \times r)} < +\infty.$$

Hàm Borel đo được h gọi là hầu hết bị chặn nếu tồn tại tập $B \in \mathcal{B}(n \times q), \mu(B) = 0$,

$$\sup_{x \in M(n \times q) - B} \|h(x)\|_{M(p \times r)} < +\infty,$$

trong đó μ là độ đo σ-hữu hạn trong không gian đo được $M(p \times r), \mathcal{B}(p \times r)$.

Ký hiệu $B(M(p \times r))$ là không gian tất cả các hàm đo được và bị chặn, $L^\infty(\mu, M(p \times r))$ là không gian tất cả hàm đo được và hầu hết bị chặn (thực chất là các lớp tương đương sau khi đã đồng nhất các hàm hầu hết bằng nhau). Dễ thấy chúng là các không gian Banach với chuẩn

$$\|h\|_{B(M(p \times r))} = \sup_{x \in M(n,q)} \|h(x)\|_{M(p \times r)}$$

$$\|h\|_\infty = \inf_{B, \mu(B)=0} \sup_{x \in M(n,q) - B} \|h(x)\|_{M(p \times r)}$$

và theo định nghiã chúng tạo thành các lớp ước lượng của tham ẩn $\theta \in \Theta \subset M(p \times r)$.

Định nghiã 2.2. Cho h là hàm được xác định bởi

$$H : M(n \times q) \times \Theta \to M(p \times r) \times \Theta, \quad H(x, \theta) = (h(x), \theta)$$

và cho hàm không âm

$$L : M(p \times r) \times \Theta \to \overline{R}^+ = [0, +\infty].$$

Khi ấy hàm hợp $L(h(.), .) = L.H : M(n \times q)\Theta \to \overline{R}^+$ được gọi là hàm tổn thất.

Định nghiã 2.3. Cho không gian tham compact $\Theta \subset M(p \times r)$ và σ-đại số Borel $\mathcal{B}(\Theta)$ trên Θ. Xác định một độ đo xác suất τ trên $(\Theta, \mathcal{B}(\Theta))$ và gọi τ là phân phối xác suất trên nghiệm của $\theta \in \Theta$. Biết rằng với phần tử quan trắc ngẫu nhiên X luân luân tồn tại phân phối xác suất có điều kiện chính quy p^X, thường được ký hiệu Q_θ (xem [4]). Giả sử $Q_\theta << \mu$ với mọi $\theta \in \Theta$. Khi ấy theo định nghiã Radon-Nykodym tồn tại hàm mật độ

$$f_\theta(x) = \frac{Q_\theta(dx)}{\mu(dx)}.$$

Hàm $\psi : B(M(p \times r)) \to \overline{R}^{+}$ (hoặc $\psi : L^{\infty}(\mu, M(p \times r)) \to \overline{R}^{+}$) được gọi là hàm mạo hiểm Bayes với phân phối xác suất tiền nghiệm τ nếu

$$\psi = \int_{\Theta} \int_{M(n\times q)} L(h(x), \theta) f_{\theta}(x) \mu(dx) \tau(d\tau).$$

Ước lượng $\hat{h} \in B(M(p \times r))$ (hoặc $\hat{h} \in L^{\infty}(\mu, M(p \times r))$ được gọi là ước lượng Bayes với phân phối xác suất tiền nghiệm τ nếu

$$\psi(\hat{h}) = \inf_{h \in B(M(p\times r))} \psi(h) \quad (\text{hoặc } \psi(\hat{h}) = \inf_{h \in L^{\infty}(\mu, M(p\times r))} \psi(h)).$$

Định lý 2.4. *Giả sử lớp ước lượng K của hàm ẩn $\theta \in \Theta \subset M(p \times r)$ thoả mãn các điều kiện:*

(i) $h(M(n \times q)) \subset \Theta$, $\forall h \in K$

(ii) $\forall \varepsilon > 0$, $\exists$ *phân hoạch hữu hạn* $\{E_i\}_{i=1}^{m} \subset M(n \times q)$ *và các điểm* $x_i \in E_i$, $i = 1, \ldots, m$ *sao cho*

$$\sup_{x \in E_i} \|h(x) - h(x_i)\|_{M(p\times r)} < \varepsilon, \quad \forall h \in K, \quad \forall i = 1, 2, \ldots, m,$$

(iii) tồn tại $C > 0$ sao cho

$$|L(y, theta) - L(y', \theta)| \le C\|y - y'\|_{M(p\times r)}, \quad \forall y,\ y' \in M(p \times r).$$

Khi ấy K là tập compact tương đối trong không gian Banach $B(M(p \times r))$ và trong lớp ước lượng $\overline{K}$ tồn tại ước lượng Bayes (với $\overline{K}$ là bao đóng của K).

Cách chứng minh định lý này tiến hành tương tự như trong [3] với những thay đổi thích hợp về chuẩn. Hơn nữa cũng có thể phát biểu và chứng minh định lý tồn tại ước lượng Bayes cho lớp ước lượng compact $\overline{K}$ thuộc không gian $L^{\infty}(\mu, M(p \times r))$ tương tự như 2.4.

3. Về một xấp xỉ ước lượng Bayes trong mô hình phi tuyến 1-chiều

Xét trường hợp $n = q = p = r = 1$, tức là X là biến số ngẫu nhiên. Giả sử tập trị của X là đóng và bị chặn thuộc R^1, ký hiệu là I. Giả thử không gian tham Θ là tập compact của R^1. Ký hiệu $B(I)$ là họ tất cả các ước lượng bị chặn của tham ẩn $\theta \in \Theta \subset R^1$. Ký hiệu $C(I)$ là họ tất cả các hàm liên tục xác định trên I. Hiển nhiên $B(I)$ và $C(I)$ là các không gian Banach và $C(I) \subset B(I)$.

Định lý 3.1. *Cho K là lớp ước lượng của hàm ẩn $\theta \in \Theta \subset R^1$ và thoả mãn các điều kiện trong 2.4. Giả sử hàm mật độ $f_{\theta}(x)$ bị chặn đều. Khi ấy xây dựng được một đa thức xấp xỉ ước lượng Bayes thuộc $\overline{K}$.*

Chứng minh. Vì K thoả mãn các điều kiện của định lý 2.4 nên trong $\overline{K}$ tồn tại ước lượng Bayes $\hat{h}$. ta sẽ tìm cách xấp xỉ $\hat{h}$ bằng một đa thức.

Lấy bất kỳ $h \in \overline{K}$. Theo định lý Lusin (xem [5]), tồn tại hàm liên tục $g \in C(I)$ sao cho với bất kỳ $\varepsilon > 0$ cho trước, ta có:

$$\mu(\{x \in I: \ h(x) \neq g(x)\}) < \frac{\varepsilon}{C.C'C''},$$

trong đó μ là độ đo Lebegue trên R^1. Theo giả thiết $\exists C', C$ sao cho $|h(x)| \leq C'$, $|g(x)| \leq C'$. Đặt $A = \{h(x) \neq g(x)\}$. Xét hệ thức (và nhớ rằng theo giả thiết $\exists C'': \ |f_\theta(x)| \leq C''$) :

$$\begin{aligned} |\psi(h) - \psi(g)| &\leq \int_\Theta \int_I |L(h(x), \theta) - L(g(x), \theta)| f_\theta(x)\mu(dx)\tau(d\theta) \\ &\leq \int_\Theta \int_{I \setminus A} C|h(x) - g(x)| f_\theta(x)\mu(dx)\tau(d\theta) + \int_\Theta \int_A C|h(x) - g(x)| f_\theta(x)\mu(dx)\tau(d\theta) \ < \frac{\varepsilon}{2}. \end{aligned}$$

Mặt khác, với $\varepsilon > 0$ như trên, theo định lý xấp xỉ Weierstrass (xem [6]) tồn tại một đa thức $P_{n(\varepsilon),a}$, bậc $n(\varepsilon)$, hệ số $a = (a_0, a_1, ..., a_n) \in R^{n+1}$, thuộc $C(I)$ sao cho

$$||g - P_{n(\varepsilon),a}||_{C(I)} \ < \ \frac{\varepsilon}{2C.C''}.$$

Do đó

$$|\psi(g) - \psi(P_{n(\varepsilon),a}| \leq \int_\Theta \int_I C|g(x) - P_{n(\varepsilon),a}|_{C(I)} f_\theta(x)\mu(dx)\tau(d\theta) < \frac{CC''\varepsilon}{2CC''} = \frac{\varepsilon}{2}.$$

Suy ra

$$|\psi(g) - \psi(P_{n(\varepsilon),a}| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Theo xây dựng trên, ta có đa thức bậc $n(\varepsilon)$, hệ số $a \in R^{n(\varepsilon)+1}$. Ta sẽ cố định $n(\varepsilon) = n$. Theo định nghiã cuả phiếm hàm ψ thì $\psi(P_{n,a})$ chỉ phụ thuộc vào hệ số $a \in R^{n+1}$. Vậy với mọi $a \in R^{n+1}$ tồn tại duy nhất một số $\psi(P_{n,a})$. Nên, tồn tại một hàm số nhiều biến $F: \ R^{n+1} \to R^1$ sao cho $F(a) = \psi(P_{n,a})$.

Chú ý rằng với mỗi $h \in \overline{K}$ và với $\varepsilon > 0$ cho trước, tồn tại một số $n = n(\varepsilon)$ như trên và do đó tồn tại không gian R^{n+1}. Nên, với $h' \in \overline{K}$, $h' \neq h$, sẽ có một không gian $R^{n'+1}$ khác. Vậy, hàm F xác định trên các không gian khác nhau. Về mặt lý thuyết lấy supremum cuả tất cả các $n(\varepsilon)$ này vì nó có thể bằng $+\infty$. Tuy nhiên trên quan điểm thức hành (vì đây là bài toán xấp xỉ), với mọi $h \in \overline{K}$ ta có thể chọn được một số $n(\varepsilon)$ đủ lớn sao cho hàm F xác định trên cùng một không gian $R^{n(\varepsilon)+1}$ và thoả $F(a) = \psi(P_{n,a})$.

Với quy ước n như trên, đặt

$$A_{\varepsilon,h} = \{a \in R^{n+1}: \ |\psi(h) - F(a)| < \varepsilon\}, \ A_\varepsilon = \cup_{h \in K} A_{\varepsilon,h}.$$

Giả sử F thoả các điều kiện đạt cực tiểu trên A_ε (để ý rằng chỉ cần giả thiết F liên tục thì A_ε là tập mở thuộc R^{n+1}). Khi ấy tồn tại $a \in A_\varepsilon$ sao cho

$$F(a^*) = \inf_{a \in A_\varepsilon} F(a).$$

Gọi $\hat{h}$ là ước lượng Bayes thuộc $\overline{K}$, tức là $\psi(\hat{h}) = \inf_{h \in \overline{K}} \psi(h)$. Với $\hat{h}$ này, theo lập luận trên, sẽ tồn tại đa thức $P_{n,\hat{a}}$ với hệ số $\hat{a} \in R^{n+1}$ sao cho

$$|F(\hat{a}) - \psi(\hat{h})| < \varepsilon. \tag{2}$$

Do đó $\hat{a} \in A_\varepsilon$. Suy ra

$$|F(a^*) - \psi(\hat{h})| < 3\varepsilon. \tag{3}$$

Mặt khác, ta có đồng thời

$$F(a) - \psi(h) < \varepsilon$$
$$\psi(h) - \psi(h^*) < \varepsilon$$
$$F(h^*) - F(a^*) < \varepsilon$$

Nên

$$F(a^*) - F(\hat{a}) > -3\varepsilon. \tag{4}$$

Từ (3) và (4) suy ra

$$|F(a^*) - F(\hat{a})| < 3\varepsilon. \tag{5}$$

Từ (2) và (5) ta được

$$|F(a^*) - \psi(\hat{h})| < 4\varepsilon.$$

Từ các hệ số $a^* \in R^{n+1}$, ta thiết lập được một đa thức P_{n,a^*}, bậc n, hệ số a^*. Với đa thức này, ta có $\psi(P_{n,a^*}) = F(a^*)$, theo định nghĩa của F. Vậy ta có

$$|\psi(\hat{h}) - \psi(P_{n,a^*})| < 4\varepsilon.$$

Điều này có nghĩa ta có thể lấy đa thức (mà ta sẽ gọi là đa thức cực tiểu) làm xấp xỉ của các ước lượng Bayes $\hat{h}$ và định lý chứng minh xong.

Nhận xét. Giá trị a^* là cực tiểu toàn cục của hàm nhiều biến F. Tuy nhiên trong thực tế, rất khó tìm được cực tiểu toàn cục. Vì vậy, như thường lệ, ta sẽ chỉ tìm cực tiểu địa phương a^* và dùng a^* này làm hệ số của đa thức cực tiểu để xấp xỉ ước lượng Bayes $\hat{h}$. Thậm chí, ta có thể chỉ lấy a^* là giá trị của điểm dừng và dùng chúng làm hệ số của đa thức cực tiểu.

Trên đây ta đã làm bài toán xấp xỉ cho trường hợp X là biến số ngẫu nhiên. Có thể nới rộng bài toán này sang trường hợp X là véc tơ ngẫu nhiên hay X là ma trận ngẫu nhiên. Khi ấy ta phải sử dụng định lý Lusin và định lý xấp xỉ Weierstrass trong trường hợp không gian nhiều chiều. Cách chứng minh sẽ không khó hơn, tuy nhiên cồng kềnh hơn.

Tài liệu tham khảo

1. Rao C. R., *Linear Statistical Infrence and Its Applications*, John Wiley 1973.
2. Rumak K. M., *Statische Methoden der Modellbildung,* Bandl I, Academie-Verlage, 1977.
3. Ung Ngọc Quang, *Về sự tồn tại ước lượng Bayes trong mô hình thống kê với không gian tham compact.* Tạp chí toán học, T. XVIII, 1990, 1-8.
4. Giklman I. and Skorokhod, A., *The theory of stachastic process,* T. I, Moscow, Nauka, 1971 (in Russian).
5. Rudin W., *Real and Complex Analysis,* Tata McGraw - Hill, Publ. Company, New Delhi, 1976.
6. Rudin W., *Principles of Mathematical Analysis,* McGraw - Hill Book Company, New York 1964.

Abstract

An approach to the Bayes estimates in nonlinear statiscal models

In this note, we consider the problem of finding an approach to the Bayesian estimates in the nonlinear statistical models $X = \varphi(\theta) + \varepsilon$, *where* X *is a matric of observations,* φ *is a known nonlinear function, and* $\theta \in \Theta$, Θ *is a compact subset of* R^1.

Khoa toán, Đại học Tổng hợp Tp. HOCHIMINH
227 Nguyễn Văn Cừ, Q. 5, Tp. HCM