КОМПЬЮТЕРНАЯ МАТЕМАТИКА

УДК 519.21

А.С. САМОСЕНОК

АСИМПТОТИЧЕСКИЕ СВОЙСТВА ОЦЕНКИ МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ ДЛЯ ГИББСОВСКИХ ПОЛЕЙ

Рассматриваются условия, при которых возможна аппроксимация критериальной функции марковского процесса заданного гиббсовским распределением с единственной точкой минимума ее эмпирической оценкой. Доказываются теоремы о сходимости приближенных оценок, полученных методом максимального правдоподобия, как для случая конечного множества состояний марковского процесса, так и для случая произвольного множества.

Введение. Точное решение любой задачи возможно только при наличии строгих условий, но поскольку в реальном мире связи и отношения между объектами настолько сложны и многообразны, то в математическом аспекте значительный класс задач сводится к востановлению по экспериментальным данным некоторых неизвестных параметров объекта. Процесс выбора таких характеристик модели из заданного класса для наилучшего описания результатов и подразумевают под понятием оценивания.

В терминологии теории статистических решений эти задачи тесно связаны с асимптотическими свойствами оценок неизвестных параметров, а именно, состоятельностью, асимптотическим распределением, скоростью сходимости оценок и т. д. И если при этом предположение о состоятельности оценок было сделано без достаточных оснований, то практическое использование таких алгоритмов нецелесообразно. Кроме этого, естественно, что условия сходимости существенным образом зависят от функции критерия и вероятностных свойств случайных величин, от которых она зависит. В данной работе будем рассматривать модели, общий вид распределения которых известен, имеет форму гиббсовского и обладает марковским свойством.

Рассмотрим стохастическую модель, в которой семейство вероятностных распределений формирует марковское поле с диск-

ретным временем, зависящее от неизвестного

параметра v ($v \in \theta$), который может быть как одномерным, так и многомерным. Будем предполагать, что каждый элемент поля $s \in S$ с некоторой вероятностью может находиться в одном из своих состояний $x_i^s, i = 1, 2, ..., n$, принадлежащих некоторому множеству X. Марковость в нашем случае будет означать, что это значение зависит лишь от состояний, принадлежащих окрестности ∂s вершины s. Сами распределения имеют форму гиббсовских:

$$\Pi(x;v) = Z(v)^{-1} \exp\left(\sum_{i=1}^{m} H_i(x,v)\right),$$

где $H_i(x,v)$ некоторые действительные функции, которые назовем потенциалами, $v \in \theta$ – неизвестный параметр, $Z(v)$ – нормирующий множитель, равный сумме всех потенциалов системы.

Последовательность состояний такой системы локально независима в следующем смысле:

$$\prod_{s \in T} \Pi(x^s \mid x^{S \setminus s}; v) = \prod_{s \in T} \Pi(x^S \mid x^{\partial s}; v),$$

где T – некоторое конечное подмножество множества вершин S; x^s – состояние вершины s.

Задача состоит в оценивании значения неизвестного параметра v_0 или нахождении некоторой неизвестной функции от этого параметра $g(v_0)$ на основании результатов наблюдений $x \in X$. Зачастую такой функцией может быть математическое ожидание и, поскольку, на практике его вид неизвестен мы можем аппроксимировать его известными функциями $G^n(x;v)$.

В нашем случае эти функции имеют следующий вид:

$$g(v) = E(\ln \Pi(x;v)), \text{ с точкой максимума } v_0;$$

$$G^n(x^1, ... x^n; v) = \frac{1}{n} \sum_{i=1}^{n} \ln \Pi(x^i; v), \text{ с точкой максимума } v_n.$$

Логично, что вероятностные свойства рассматриваемых случайных величин в большой мере влияют на условия сходимости таких оценок к истинным параметрам. В работах [1, 2] эти подходы тщательно проанализированы, в них же приведен обзор основных результатов, полученных в этой области. В данной же статье для решения задачи оценивания неизвестного параметра рассматривается метод максимального правдоподобия и исследуются условия сходимости таких оценок к истинным значениям параметра.

Определим функцию максимального правдоподобия

$$L_n(x,v) = \Pi(x^1;v) \cdot \Pi(x^2;v) \cdot ... \cdot \Pi(x^n;v).$$

Поскольку согласно методу максимального правдоподобия наблюдаемые результаты соответствуют параметру, при котором вероятность получения этих данных максимальна, то оценкой оптимального параметра v_0 будет решение задачи максимизации функции максимального правдоподобия:

$$v_n = \arg\max L_n(x, v).$$

Так как функция $\ln L_n(x, v)$ при фиксированных x достигает максимума в той же точке, что и $L(x, v)$, то в дальнейшем будем рассматривать следующую функцию:

$$\ln L_n(\cdot, v) = \ln\left(\Pi(x^1; v) \cdot \ldots \cdot \Pi(x^n; v)\right) = \sum_{i=1}^{n} \ln \Pi(x^i; v).$$

При решении поставленной задачи будем рассматривать два случая: случай конечного фазового пространства значений марковского поля и случай произвольного множества состояний марковской последовательности. Множество параметров θ в обоих случаях является компактным подмножеством $\mathfrak{R}$.

1. Состоятельность оценки максимального правдоподобия в случае конечного множества состояний марковского поля. Предварительно сформулируем утверждения, которые необходимы для доказательства утверждений о состоятельности таких оценок.

Теорема 1 [3]. Пусть (Ω, U, P) – вероятностное пространство, K – компактное подмножество некоторого банахова пространства с нормой $\|\square\|$. Пусть $\{U_n, n \in N\}$ – последовательность σ-алгебр, $U_n \subset U$, $U_{n_1} \subset U_{n_2}$, $n_1 < n_2$. Пусть $\{Q_n(s) = Q_n(s, w) : (s, w) \in K \times \Omega, n \in N\}$ – семейство действительных функций, удовлетворяющее следующим условиям:

1) для фиксированных n, w функция $Q_n(s, w)$, $s \in K$ непрерывна;

2) для любых n, s функция $Q_n(s, w)$, $w \in \Omega$ U_n – измерима;

3) для некоторого элемента $s_0 \in K$ и любого $s \in K$ $Q_n(s, w) \to \Phi(s; s_0)$, $n \to \infty$ по вероятности, где $\Phi(s; s_0)$, $s \in K$ – действительная функция, непрерывная на K и удовлетворяющая условию $\Phi(s; s_0) < \Phi(s_0; s_0)$, $s \neq s_0$;

4) Существуют $\gamma_0 > 0$ и функция $c(\gamma) : \gamma > 0,\ c(\gamma) \to 0,\ \gamma \to 0,$ такие что при всех $s' \in K$ и $\gamma : 0 < \gamma < \gamma_0,$ имеем

$$\lim_{n \to \infty} P\left\{ \sup_{\|s - s'\| < \gamma} |G_n(s) - G_n(s')| < c(\gamma) \right\} = 1.$$

Тогда если $s_n = \arg\max_{s \in K} Q_n(s),$ то последовательность $\{s_n, n \geq 1\}$ сходится по вероятности к s_0:

для любого $\varepsilon > 0$

$$\lim_{n\to\infty} P\{\|s_n - s_0\| > \varepsilon\} = 0,$$

а последовательность $\{Q_n(s_n), n \geq 1\}$ сходится по вероятности к $\Phi(s_0; s_0)$.

Теорема 2 [4]. Пусть $\{\xi_i, i \in N\}$ – однородная цепь Маркова с примитивным марковским ядром *P* на конечном множестве *X*. Тогда для любого начального распределения ν и любой функции $f : X \to \Re$ по вероятности

$$\frac{1}{n}\sum_{i=1}^{n} f(\xi_i) \to E_\mu(f), n \to \infty,$$

где μ – инвариантное распределение марковского ядра *P*.

Для доказательства сходимости оценки максимального правдоподобия по вероятности докажем справедливость следующего утверждения.

Теорема 3. Пусть множество состояний *X* случайной последовательности является конечным, и v_0 – единственная точка максимума функции $g(v)$, тогда для любой точки максимума v_n функции $G^n(x^1,...x^n;v)$ справедливо $v_n \to v_0, n \to \infty$ по вероятности.

Доказательство. Для доказательства используем теорему 1. Возьмем $Q_n = G^n$ и $u_n = \sigma\{x^i, i = \overline{1,n}\}$ и проверим все 4 условия для нашего случая.

Условия 1) и 2) очевидны в силу выбора функции G^n.

Согласно закону больших чисел для марковских цепей

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n}\ln\Pi(x^i;v) - E(\ln\Pi(\cdot;v))\right| \leq \delta\right) \to 1, n \to \infty$$

для каждого $v \in \theta$ и любого положительного δ. Откуда и следует условие 3).

Перейдем к доказательству выполнимости условия 4). Обозначим

$$\psi(x,\gamma) = \sup_{\substack{u,v\in\theta:\\ \|u-v\|<\gamma}} \left|\ln\Pi(x,u) - \ln\Pi(x,v)\right|, \gamma > 0, \ x \in X.$$

Для всех $n,\ u,\ u_1 \in \theta,\ \gamma > 0$

$$\sup_{\|u-u_1\|<\gamma}\left|G^n(u) - G^n(u_1)\right| \leq \frac{1}{n}\sum_{i=1}^{n}\psi(x^i,\gamma).$$

Согласно теореме 2

$$\frac{1}{n}\sum_{i=1}^{n}\psi\left(x^{i},\gamma\right)\to E_{\mu}\left(\psi(,\gamma)\right), n\to\infty \text{ по вероятности.}$$

При любом $x\in X$ $\psi(x,\gamma)\to 0$ при $\gamma\to 0$.

Обозначив $c(\gamma)=2E\{\psi(,\gamma)\}$ мы и получим, что $c(\gamma)\to 0, \gamma\to 0$, а значит, выполняется условие 4) теоремы 1. Следовательно, из теоремы 1 следует теорема 3. □

2. Сильная состоятельность оценки максимального правдоподобия

Для доказательства сильной состоятельности нам потребуются некоторые ранее доказанные факты и условия, которые необходимы для доказательства сходимости с вероятностью 1.

Пусть прошлое процесса X_t определяется σ-алгебрами вида $U_{-\infty}^{\tau}$, а будущее – вида U_{τ}^{∞}.

Определение. Стационарный процесс в узком смысле X_t называется удовлетворяющим условию сильного перемешивания если

$$\sup_{A\in U_{-\infty}^{0}, B\in U_{\tau}^{\infty}}\left|P(AB)-P(A)P(B)\right|=\alpha(\tau)\le\frac{C}{\tau^{1+\varepsilon}} \text{ и } \tau>0,\ \varepsilon>0,\ c>0.$$

Функцию $\alpha(\tau)$ будем называть коэффициентом сильного перемешивания, очевидно, что она монотонно не возрастает.

Определение. Если на σ-алгебре U_X существует конечная мера m с $m(U)>0$, целое число $v>1$ и число $\varepsilon>0$ такие, что вероятность перехода за v шагов $p^{(v)}(\xi,A)\le 1-\varepsilon$, при $m(A)\le\varepsilon$, то будем говорить, что выполняется гипотеза Деблина.

Если выполняется гипотеза Деблина и существует только один эргодический класс, причем он не содержит подклассов, тогда существует такое распределение $p(A)$ на U_X, что

$$\sup_{x,A}\left|p^{(n)}(x,A)-p(A)\right|\le c\rho^{n}, \tag{1}$$

где $c>0$, $0<\rho<1$ – константы (см. [4], глава 19).

Вернемся к нашему случаю, согласно [4] если справедливо соотношение (1) и начальное распределение $\Pi_0(x,v)$ стационарно, то стационарный процесс X_t удовлетворяет условию сильного перемешивания с коэффициентом $\alpha(n)\le 2c\rho^{n}$.

Теорема 4 [3]. Пусть (Ω,U,P) – вероятностное пространство, K – компактное подмножество некоторого банахова пространства с нормой $\|\cdot\|$.

Пусть $\{U_n, n \in N\}$ – последовательность δ-алгебр, $U_n \subset U$, $U_{n_1} \subset U_{n_2}$, $n_1 < n_2$. Пусть $\{Q_n(s) = Q_n(s,w) : (s,w) \in K \times \Omega, n \in N\}$ – семейство действительных функций, удовлетворяющее следующим условиям:

1) для фиксированных n,w функция $Q_n(s,w)$, $s \in K$ непрерывна;

2) для любых n, s функция $Q_n(s,w)$, $w \in \Omega$ U_n – измерима;

3) для некоторого элемента $s_0 \in K$ для любого $s \in K$ $Q_n(s,w) \to \Phi(s;s_0)$, $n \to \infty$ с вероятностью 1, где $\Phi(s;s_0)$, $s \in K$ – действительная функция, непрерывная на K и удовлетворяющая условию $\Phi(s;s_0) < (s_0;s_0)$, $s \neq s_0$;

4) существуют $\gamma_0 > 0$ и функция $c(\gamma)$, $\gamma > 0$, $c(\gamma) \to 0$, $\gamma \to 0$, такие что при всех $s' \in K$ и $\gamma : 0 < \gamma < \gamma_0$, имеем

$$P\{\overline{\lim_{n\to\infty}} \sup_{\|s-s'\|<\gamma} |Q_n(s) - Q_n(s')| < c(\gamma)\} = 1.$$

Для всех n, w элемент $s_n = s_n(w)$ определяется соотношением $Q_n(s_n) = \max_{s\in K} Q_n(s)$.

Такой элемент можно выбрать U_n измеримым по w. Тогда

$$P\{s_n \to s_0, n \to \infty\} = 1,$$

а последовательность $\{Q_n(s_n), n \geq 1\}$ сходится по вероятности к $\Phi(s_0;s_0)$.

Для доказательства сильной состоятельности оценки максимального правдоподобия (с вероятностью 1) покажем, что справедлива следующая теорема.

Теорема 5. Если функция $g(v)$ имеет единственную точку максимума v_0, последовательность x^n удовлетворяет условию Деблина с $\rho^n \leq \frac{c}{n^{1+\varepsilon}}$, и $E\{\sup_{v\in\theta} \ln \Pi(x,v)\} < \infty$ тогда для любой точки максимума v_n функции $G^n(x^1,...x^n;v)$, справедливо $P\{v_n \to v_0, n \to \infty\} = 1$.

Доказательство. Для доказательства применим теорему 4, возьмем $Q_n = G^n$ и $u_n = \sigma\{x^n, i = \overline{1,n}\}$ и проверим все 4 условия для нашего случая. Очевидно, что условия 1), 2) теоремы 4 выполнены в силу выбора функции G^n.

Рассмотрим условие 3) теоремы 4. Так как последовательность x^n удовлетворяет условию равномерного сильного перемешивания с коэффициентом – показательной функцией мы можем выбрать ρ таким образом, что

$$\left|E \ln \Pi(x^i,v) \ln \Pi(x^j,v) - E \ln \Pi(x^i,v) E \ln \Pi(x^j,v)\right| \leq \frac{c}{1+|i-j|^{1+\varepsilon'}}, \ \varepsilon' > 0.$$

Обозначим $\eta_n(v)=G^n(v)-EG^n(v)$ и оценим $E\eta^2{}_n(v)$:

$$E\eta^2{}_n(v)=E\left\{\frac{1}{n}\sum_{i=1}^{n}\ln\Pi(x^i,v)-E\left(\frac{1}{n}\sum_{i=1}^{n}\ln\Pi(x^i,v)\right)\right\}^2=$$

$$=E\left\{\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left[\ln\Pi(x^i,v)-E\ln\Pi(x^i,v)\right]\left[\ln\Pi(x^j,v)-E\ln\Pi(x^j,v)\right]\right\}=$$

$$=\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}Ey_iy_j,$$

где $y_i=\ln\Pi(x^i,v)-E\ln\Pi(x^i,v)$.

Учитывая условие сильного перемешивания, имеем

$$Ey_iy_j\le\frac{c}{1+\left|i-j\right|^{1+\varepsilon'}},\ \varepsilon'>0.$$

Поэтому $\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}Ey_iy_j\le\frac{c}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{c}{1+\left|i-j\right|^{1+\varepsilon'}}\le\frac{c}{n}$.

Пусть $n=m^2$. Тогда согласно лемме Бореля – Кантелли

$$P\{\lim_{m\to\infty}\eta_{m^2}=0\}=1.$$

Обозначим $\zeta_m=\sup\limits_{m^2\le n\le(m+1)^2}\left|\eta_n-\eta_{m^2}\right|$.

Для $m^2\le n\le(m+1)^2$ справедливо следующее неравенство

$$\left|\eta_n\right|\le\left|\eta_{m^2}\right|+\sup_{m^2\le n\le(m+1)^2}\left|\eta_n-\eta_{m^2}\right|.$$

Для ζ_m запишем

$$E(\zeta_m)^2=E\sup_{m^2\le n\le(m+1)^2}\left|\eta_n-\eta_{m^2}\right|\le\sup_{m^2\le n\le(m+1)^2}\left|\frac{1}{m^2}\sum_{i=1}^{n}\sum_{j=1}^{n}Ey_iy_j\right|\le$$

$$\le\frac{1}{m^2}\sum_{i=m^2+1}^{(m+1)^2}\sum_{j=m^2+1}^{(m+1)^2}E\left|y_iy_j\right|\le c\left[\frac{(m+1)^2-m^2-1}{m^2}\right]^2=2\frac{c}{m^2}.$$

Откуда согласно лемме Бореля – Кантелли $P\{\lim\limits_{m\to\infty}\zeta_n=0\}=1$, тогда $P\{\lim\limits_{m\to\infty}\eta_n=0\}=1$, и следовательно $P\left\{\lim\limits_{n\to\infty}G_n(v)=E\ln\Pi(x,v)\right\}=1$. А значит и условие 3) теоремы 4 выполнено.

Перейдем к условию 4). Обозначим

$$\psi(x,\gamma)=\sup_{\substack{u,v\in\theta:\\ \|u-v\|<\gamma}}\left|\Pi(x,u)-\Pi(x,v)\right|,\gamma>0,\ \ x\in X.$$

Для всех $n,\ u,\ u_1\in\theta$, $\gamma>0$

$$\sup_{\|u-u_1\|<\gamma}\left|G^n(u)-G^n(u_1)\right|\le\frac{1}{n}\sum_{i=1}^{n}\psi(x^i,\gamma).$$

Учитывая закон больших чисел $\frac{1}{n}\sum_{i=1}^{n}\psi(x^i,\gamma) \to E_\mu\left(\psi(,\gamma)\right), n \to \infty$ с вероятностью 1.

При любом $x \in X$ $\psi(x,\gamma) \to 0$ при $\gamma \to 0$. Обозначив $c(\gamma) = 2E\{\psi(,\gamma)\}$ получаем, что $c(\gamma) \to 0, \gamma \to 0,$ а значит, выполняется условие 4) теоремы 1. Следовательно из теоремы 1 следует теорема 3.

О.С. Самосьонок

АСИМПТОТИЧНІ ВЛАСТИВОСТІ ОЦІНКИ МАКСИМАЛЬНОЇ ПРАВДОПОДІБНОСТІ ДЛЯ ГІБСОВСЬКИХ ПОЛІВ

Розглядаються умови, за яких можлива апроксимація критеріальної функції марківського процессу заданого гібсовським розподілом з єдиною точкою мінімуму її емпіричною оцінкою. Доводяться теореми про збіжність наближених оцінок, отриманих методом максимальної правдоподібності, як для випадку кінечної множини станів марківського процесу, так і для випадку довільної множини.

A.S. Samosonok

ASYMPTOTIC PROPERTIES OF MAXIMUM LIKELIHOOD ESTIMATOR FOR GIBBS FIELDS

The article focuses on asymptotic consistency of maximum likelihood estimators applied to Markov fields with Gibbs distribution. Theorems on the approximate estimate convergence are proved. The estimates are obtained with the maximum likelihood method for cases of finite set of states of Markov process and for arbitrary set.

1. *Gerhard Winkler*. Analysis, Random Fields and Dynamic Monte Carlo Methods. – Springer, 1995.
2. *Xavier Guyon.* Random Fields on Network. – Springer-Verlag. 1995.
3. *Дороговцев А.Я.* Теория оценок параметров случайных процессов. – К.: Вища школа, 1982.
4. *Ибрагимов И.А., Линник Ю.В*. Независимые и стационарно связанные величины. – М.: Наука, 1980.

Получено 13.12.2011

Об авторе:

Самосенок Александр Сергеевич,
младший научный сотрудник Института кибернетики имени В.М. Глушкова НАН Украины.