НОВЫЕ СРЕДСТВА
КИБЕРНЕТИКИ, ИНФОРМАТИКИ,
ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКИ
И СИСТЕМНОГО АНАЛИЗА

А.С. САМОСЁНОК

УДК 519.21

ИССЛЕДОВАНИЕ ЭМПИРИЧЕСКИХ ОЦЕНОК ПАРАМЕТРОВ ГИББСОВСКОГО РАСПРЕДЕЛЕНИЯ, ПОЛУЧЕННЫХ МЕТОДОМ МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ

Ключевые слова: *гиббсовское распределение, метод максимального правдоподобия, марковские случайные поля, стохастические процессы, оценка параметров.*

ПОСТАНОВКА ЗАДАЧИ

Рассмотрим стохастическую модель, в которой семейство вероятностных распределений формирует марковское поле с дискретным временем, зависящее от неизвестного параметра v ($v \in \theta$); он 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_s, принадлежащих некоторому множеству X. Марковость в данном случае означает, что эта величина зависит лишь от состояний элементов поля, принадлежащих окрестности ∂s вершины s. Каждой паре соседних элементов поля сопоставляется функция $H(x, v)$ — ее потенциал, зависящий от значений, принимаемых реализацией поля x на этих элементах [1]. Сумму потенциалов всех пар, в которые входит рассматриваемый элемент s, назовем энергетической функцией $U_s(x, v)$ по окрестности элемента s:

$$U_s(x, v) = \sum_{i=1}^{m} H_i(x, v).$$

Общий вид распределения вероятности наблюдения конкретного состояния известен с точностью до неизвестного параметра $v_0 \in \theta$ и имеет форму гиббсовского распределения

$$\Pi_s(x, v) = Z(v)^{-1} \exp(U_s(x, v)),$$

где $Z(v)$ — нормирующий множитель, равный сумме всех потенциалов системы.

Последовательность состояний такой системы локально независима в следующем смысле:

$$\prod_{s \in T} \Pi(x_s | x_{S \setminus s}; v) = \prod_{s \in T} \Pi(x_s | x_{\partial s}; v),$$

где T — некоторое конечное подмножество множества вершин S; x_s — состояние вершины s.

Рассмотрим последовательность независимых наблюдений $x_s^1, \ldots, x_s^n$, $x_s^i \in X$, $i = 1, 2, \ldots, n$, за состоянием некоторого элемента s марковского поля. Зада-

© А.С. Самосёнок, 2013

ча состоит в оценивании значения неизвестного параметра v_0 или нахождении некоторой неизвестной функции от этого параметра $F(v_0)$ на основании результатов наблюдений $x \in X$. Часто ввиду важности такой функцией является математическое ожидание. Поскольку на практике нередко его вид неизвестен, можно аппроксимировать математическое ожидание известными функциями, полученными на основании наблюдений.

В данной статье для решения задачи оценивания неизвестного параметра используется метод максимального правдоподобия и исследуются условия сходимости таких оценок к истинным значениям параметра.

Так как далее будем рассматривать наблюдения и производить расчеты относительно некоторого фиксированного элемента поля, а влияние его «соседей» ограничивается функцией $U_s(x, v)$, для удобства опустим нижний индекс при аргументе x, подразумевая под x^i состояние элемента поля s в i-м наблюдении.

Определим функцию максимального правдоподобия

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

Поскольку согласно методу максимального правдоподобия наблюдаемые результаты соответствуют параметру, при котором вероятность получения этих данных максимальна, 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 $L(x, v)$, в дальнейшем можем рассматривать следующую функцию:

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

Обозначим

$$f(x^i, v) = U(x^i; v) - \ln Z(v)$$

и аппроксимируем неизвестную функцию $F(v) = E(f(x, v))$ с точкой максимума v_0 функцией $F_n(x^1, \ldots, x^n; v) = \frac{1}{n}\sum_{i=1}^{n} f(x^i, v)$ с точкой максимума v_n.

При исследовании состоятельности такой оценки рассмотрим два случая: конечное фазовое пространство значений марковского поля; произвольное множество состояний марковской последовательности. Множество параметров θ в обоих случаях является компактным подмножеством $\mathfrak{R}$.

СОСТОЯТЕЛЬНОСТЬ ОЦЕНКИ МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ В СЛУЧАЕ КОНЕЧНОГО МНОЖЕСТВА СОСТОЯНИЙ МАРКОВСКОГО ПОЛЯ

Предварительно сформулируем утверждения, необходимые для доказательства состоятельности оценок максимального правдоподобия.

Теорема 1 [2]. Пусть $(\Omega, \mathfrak{F}, P)$ — вероятностное пространство, K — компактное подмножество некоторого банахова пространства с нормой $||\cdot||$; $\{\mathfrak{F}_n, n \in N\}$ — последовательность σ-алгебр, $\mathfrak{F}_n \subset \mathfrak{F}$, $\mathfrak{F}_{n_1} \subset \mathfrak{F}_{n_2}$, $n_1 < n_2$. Пусть также $\{Q_n(s) = Q_n(s, w) : (s, w) \in K \times \Omega,\ n \in N\}$ — 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 $Q_n(s, w)$, $w \in \Omega$, $\mathfrak{F}_n$-измерима;

3) для некоторого элемента $s_0 \in K$ и любого $s \in K$ имеем $Q_n(s, w) \to \Phi(s; s_0)$,

$n \to \infty$, 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

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

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 [1]. 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 имеем

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

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 $F(v)$. Тогда для любой точки максимума v_n функции $F_n(x^1, \ldots, x^n; v)$ справедливо $v_n \to v_0$, $n \to \infty$, по вероятности.

Доказательство. Для доказательства используем теорему 1; возьмем $Q_n = F_n$ и $\Im_n = \sigma\{x^i, \; i = \overline{1, n}\}$ и проверим выполнение условий 1–4 для рассматриваемого случая.

Условия 1 и 2 очевидны в силу выбора функции F_n.

Согласно закону больших чисел для марковских цепей

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

для каждого $v \in \theta$ и любого положительного δ, откуда следует условие 3.

Перейдем к доказательству выполнимости условия 4. Обозначим

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

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

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

Согласно теореме 2

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

Обозначив $c(\gamma) = 2E\{\psi(\cdot\,, \gamma)\}$, в силу того, что при любом $x \in X$ имеем $\psi(x, \gamma) \to 0$ при $\gamma \to 0$, на основании неравенства (1) запишем

$$P\left\{ \lim_{n \to \infty} \sup_{\|u - u_1\| < \gamma} |F_n(u) - F_n(u_1)| < c(\gamma) \right\} = 1, \text{ где } c(\gamma) \to 0 \text{ при } \gamma \to 0.$$

Таким образом, выполняется условие 4 теоремы 1. Следовательно, из теоремы 1 вытекает теорема 3. □

СИЛЬНАЯ СОСТОЯТЕЛЬНОСТЬ ОЦЕНКИ МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ

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

Пусть прошлое процесса X_t определяется σ-алгебрами вида $\mathfrak{F}_{-\infty}^{\tau}$, а будущее — вида $\mathfrak{F}_{\tau}^{\infty}$.

Определение 1. Стационарный процесс в узком смысле X_t называется удовлетворяющим условию сильного перемешивания, если

$$\sup_{A\in\mathfrak{F}_{-\infty}^{0},\, B\in\mathfrak{F}_{\tau}^{\infty}} |P(AB)-P(A)P(B)| = \alpha(\tau) \le \frac{c}{\tau^{1+\varepsilon}},\ \ \tau>0,\ \ \varepsilon>0,\ \ c>0\,.$$

Функцию $\alpha(\tau)$ назовем коэффициентом сильного перемешивания.

Определение 2. Если на σ-алгебре $\mathfrak{F}_X$ существует конечная мера m с $m(\mathfrak{F})>0$, целое число $v>1$ и число $\varepsilon>0$ такие, что вероятность перехода $p^{(v)}(\xi, A)$ за v шагов удовлетворяет условию $p^{(v)}(\xi, A) \le 1-\varepsilon$ при $m(A)\le\varepsilon$, то будем считать, что выполняется гипотеза Деблина.

Если выполняется гипотеза Деблина и существует только один эргодический класс, причем он не содержит подклассов, то существует такое распределение $p(A)$ на $\mathfrak{F}_X$, что

$$\sup_{x,A} |\, p^{(n)}(x, A) - p(A)| \le c\rho^{n}, \tag{2}$$

где $c>0$, $0<\rho<1$ — константы [3].

Вернемся к рассматриваемому случаю. Согласно [3], если справедливо соотношение (2) и начальное распределение $\Pi_0(x, v)$ стационарно, стационарный процесс X_t удовлетворяет условию сильного перемешивания с коэффициентом $\alpha(n)\le 2c\rho^{n}$.

Теорема 4 [4]. Пусть $(\Omega, \mathfrak{F}, P)$ — вероятностное пространство, K — компактное подмножество некоторого банахова пространства с нормой $||\cdot||$; $\{\mathfrak{F}_n, n\in N\}$ — последовательность δ-алгебр, $\mathfrak{F}_n \subset \mathfrak{F}$, $\mathfrak{F}_{n_1} \subset \mathfrak{F}_{n_2}$, $n_1 < n_2$. Пусть также $\{Q_n(s) = Q_n(s, w) : (s, w) \in K\times\Omega,\, n\in N\}$ — 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 $Q_n(s, w)$, $w\in\Omega$, $\mathfrak{F}_n$-измерима;

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

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

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

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 $\mathfrak{F}_n$-измеримым по w. Тогда $P\{s_n\to s_0, n\to\infty\} = 1$, а последовательность $\{Q_n(s_n), n\ge 1\}$ сходится по вероятности к $\Phi(s_0; s_0)$.

При наличии сходимости последовательности $s_n \to s_0$ теорему можно сформулировать иначе и именно в таком виде использовать при доказательстве асимптотической нормальности оценки неизвестного параметра.

Следствие 1. Пусть семейство действительных функций $\{Q_n(s) = Q_n(s, w):$ $(s, w) \in K \times \Omega,\ n \in N\}$ удовлетворяет следующим условиям:

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

2) для любых n, s функция $Q_n(s, w)$, $w \in \Omega$, $\mathfrak{F}_n$-измерима;

3) существует такой элемент $s_0 \in K$, что

$$P\{\lim_{n\to\infty} || s_n - s_0 || = 0\} = 1;$$

4) для любого $s \in K$ имеем $Q_n(s) \to \Phi(s)$ при $n \to \infty$ с вероятностью 1, где $\Phi(s)$ — действительная функция, непрерывная на K.

Тогда последовательность $\{Q_n(s_n), n \geq 1\}$ сходится к $\Phi(s_0)$ с вероятностью 1:

$$P\{\lim_{n\to\infty} Q_n(s_n) = \Phi(s_0)\} = 1.$$

Для доказательства сильной состоятельности оценки максимального правдоподобия покажем, что справедлива следующая теорема.

Теорема 5. Если функция $F(v)$ имеет единственную точку максимума v_0, последовательность x^n удовлетворяет условию Деблина при $\rho^n \leq \dfrac{c}{n^{1+\varepsilon}}$ и $E\{\sup\limits_{v\in\theta} f(x, v)\} < \infty$, то для любой точки максимума v_n функции $F_n(x^1, \ldots, x^n; v)$ справедливо соотношение $P\{v_n \to v_0, n \to \infty\} = 1$.

Доказательство. Применим теорему 4; возьмем $Q_n = F_n$, $\mathfrak{F}_n = \sigma\{x^n,\ i = \overline{1, n}\}$ и проверим выполнимость условий 1–4 для рассматриваемого случая. Очевидно, что условия 1 и 2 выполнены в силу выбора функции F_n.

Рассмотрим условие 3 теоремы 4. Так как последовательность x^n удовлетворяет условию равномерного сильного перемешивания с коэффициентом — показательной функцией, выберем ρ таким образом, что

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

Обозначим $\eta_n(v) = F_n(v) - EF_n(v)$ и оценим $E\eta_n^2(v)$:

$$E\eta_n^2(v) = E\left\{ \frac{1}{n} \sum_{i=1}^n f(x^i, v) - E\left(\frac{1}{n} \sum_{i=1}^n f(x^i, v) \right) \right\}^2 =$$

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

где $y_i = f(x^i, v) - Ef(x^i, v)$.

Учитывая условие сильного перемешивания, имеем

$$Ey_i y_j \leq \frac{c}{1 + |i-j|^{1+\varepsilon'}}, \qquad \varepsilon' > 0,$$

поэтому

$$\frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n Ey_i y_j \leq \frac{c}{n^2} \sum_{i=1}^n \sum_{j=1}^n \frac{c}{1 + |i-j|^{1+\varepsilon'}} \leq \frac{c}{n}.$$

Пусть $n = m^2$. Тогда согласно лемме Бореля–Кантелли $P\{\lim_{m\to\infty} \eta_{m^2} = 0\} = 1$.

Обозначим $\zeta_m = \sup_{m^2 \le n \le (m+1)^2} |\eta_n - \eta_{m^2}|$. Для $m^2 \le n \le (m+1)^2$ справедливо следующее неравенство:

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

Для ζ_m запишем

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

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

Отсюда согласно лемме Бореля–Кантелли $P\{\lim_{m\to\infty} \zeta_n = 0\} = 1$. Тогда $P\{\lim_{m\to\infty} \eta_n = 0\} = 1$ и, следовательно, $P\{\lim_{n\to\infty} F_n(v) = Ef(x,v)\} = 1$. А значит, и условие 3 теоремы 4 выполнено.

Перейдем к условию 4. Обозначим

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

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

$$\sup_{\|u-u_1\|<\gamma} |F_n(u) - F_n(u_1)| \le \frac{1}{n} \sum_{i=1}^{n} \psi(x^i, \gamma). \tag{3}$$

Учитывая закон больших чисел, получаем $\frac{1}{n} \sum_{i=1}^{n} \psi(x^i, \gamma) \to E(\psi(\,\cdot\,, \gamma)), \ n \to \infty$, с вероятностью 1.

Обозначив $c(\gamma) = 2E\{\psi(\,\cdot\,, \gamma)\}$, в силу того, что для любого $x \in X$ имеем $\psi(x,\gamma) \to 0$ при $\gamma \to 0$, на основании неравенства (3) запишем

$$P\left\{ \lim_{n\to\infty} \sup_{\|u-u_1\|<\gamma} |F_n(u) - F_n(u_1)| < c(\gamma) \right\} = 1, \text{ где } c(\gamma) \to 0 \text{ при } \gamma \to 0.$$

Таким образом, выполняется условие 4 теоремы 4. Следовательно, теперь из теоремы 4 вытекает теорема 5. □

АСИМПТОТИЧЕСКАЯ НОРМАЛЬНОСТЬ

Исследуем условия, при которых оценка неизвестного параметра, полученная методом максимального правдоподобия для гиббсовского поля, является асимптотически нормальной. Сформулируем некоторые утверждения и рассмотрим некоторые условия, касающиеся асимптотического поведения разности $v_n - v_0$ при большом количестве наблюдений ($n \to \infty$).

Пусть выполняются условия, гарантирующие существование состоятельной оценки неизвестного параметра:

1) параметрическое множество θ — замкнутое подмножество $\mathfrak{R}^l$, $l \ge 1$;

2) функция $f(x,v) = \ln \Pi(x;v)$ непрерывна по v, $v \in \theta$.

Теорема 6 [5]. Пусть $\xi_i \in \Re^p$, $i \geq 1$, — последовательность независимых одинаково распределенных случайных величин, $E\{\xi_i\} = a$, $D\{\xi_i\} = \sigma^2 > 0$. Тогда

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (\xi_i - a) \to N(0, E\{(\xi_i - a)^2\}) \text{ по распределению.}$$

Теорема 7 [6]. Пусть $\{\xi_i, i \in N\}$, $\xi_i \in \Re^k$, — последовательность случайных величин такая, что ξ_i сходится по распределению к некоторой случайной величине ξ, $\{\eta_i, i \in N\}$, $\eta_i \in \Re^l$, — последовательность случайных величин, которая сходится по вероятности к величине c.

Тогда если φ — непрерывная функция $\Re^{k+l} \to \Re^m$, то $\varphi(\xi_i, \eta_i)$ сходится по распределению к $\varphi(\xi, c)$ при $i \to \infty$.

Теорема 8. Пусть выполняются следующие условия:

1) $E\{f(x, v)\} < \infty$;

2) v_0 — внутренняя точка θ;

3) $f(x^i, v)$ дважды непрерывно дифференцируемая в окрестности O_{v_0} точки v_0 для любого $x^i \in X$;

4) $E\{\max\limits_{v \in O_{v_0}} ||\vec{\nabla} f(x, v)||\} < \infty$, где

$$\vec{\nabla} f(x, v) = \left(\frac{\partial f(x, v)}{\partial v_j} \right)_{j=1}^{l}, \quad v = (v_j)_{j=1}^{l};$$

5) $E\{\max\limits_{v \in O_{v_0}} ||\Phi(x, v)||\} < \infty$, где

$$\Phi(x; v) = \left(\frac{\partial^2 f(x, v)}{\partial v_j \partial v_k} \right)_{j,k=1}^{l};$$

6) $\det A_0 \neq 0$, где $A_0 = E\{\Phi(x, v_0)\}$;

7) $E\{||\vec{\nabla} f(x, v_0)||^2\} < \infty$ и $\det C \neq 0$, где

$$C = E\{\vec{\nabla} f(x, v_0)(\vec{\nabla} f(x, v_0))'\}.$$

Тогда $\sqrt{n}(v_n - v_0)$ сходится по распределению к $N(0, (A_0)^{-1} C (A_0)^{-1})$, $n \to \infty$.

Пусть также выполняются следующие условия:

8) $E\{(f(x, v_0))^2\} < \infty$;

9) $\sigma^2 = E\{(f(x, v_0) - E\{f(x, v_0)\})^2\} > 0$.

Тогда $\sqrt{n}(F_n(v_n) - F(v_0))$ сходится по распределению к $N(0, \sigma^2)$, $n \to \infty$.

Доказательство. Учитывая состоятельность оценки v_n и условие 3, применим формулу Тейлора для каждой компоненты вектор-функции $\vec{\nabla}(F_n(v))$ в окрестности точки v_0:

$$\vec{\nabla} F_n(v) = \vec{\nabla} F_n(v_0) + B_0(v - v_0) + o(v),$$

где

$$\vec{\nabla} F_n(v) = \left(\frac{\partial F_n(v)}{\partial v_j} \right)_{j=1}^{l} = \frac{1}{n} \sum_{i=1}^{n} \vec{\nabla} f(x^i, v),$$

$$B_0 = \left(\frac{\partial^2 F_n(v_0)}{\partial v_j \partial v_k} \right)_{j,k=1}^{l} = \left(\frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2 f(x^i, v_0)}{\partial v_j \partial v_k} \right)_{j,k=1}^{l}.$$

Поскольку v_n является точкой минимума функции $F_n(v)$, значит, $\vec{\nabla} F_n(x, v_n) = \vec{0}$. В таком случае при $v = v_n$ левая часть равенства равна нулю, тогда

$$-\vec{\nabla} F_n(v_0) = B_0(v_n - v_0) + o(v).$$

Так как согласно определению

$$\vec{\nabla} F_n(v_0) = \left(\frac{\partial F_n(v_0)}{\partial v_j} \right)_{j=1}^{l} = \frac{1}{n} \sum_{i=1}^{n} \vec{\nabla} f(x^i, v_0),$$

запишем

$$-\frac{1}{n} \sum_{i=1}^{n} \vec{\nabla} f(x^i, v_0) \cong B_0(v_n - v_0),$$

или

$$\sqrt{n}(v_n - v_0) \cong -(B_0)^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \vec{\nabla} f(x^i, v_0). \tag{4}$$

Рассмотрим более детально каждый множитель правой части этого равенства.

Зафиксируем j, k. Согласно закону больших чисел с учетом условия 5 для каждого элемента матрицы B_0 справедливо равенство

$$P\left\{ \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2 f(x^i, v_0)}{\partial v_j \partial v_k} = E \frac{\partial^2 f(x, v_0)}{\partial v_j \partial v_k} \right\} = 1,$$

или

$$P\{ \lim_{n\to\infty} B_0 \to A_0 \} = 1.$$

Перейдем ко второму множителю. Применив теорему Лебега о предельном переходе, в силу условий 3 и 4 можем записать

$$\vec{\nabla} E\{f(x, v)\} = \left(\frac{\partial E\{f(x, v)\}}{\partial v_j} \right)_{j=1}^{l} = E\{\vec{\nabla} f(x, v)\}.$$

Следовательно, с учетом того, что v_0 — оптимальное значение параметра, справедливо $E\{\vec{\nabla} f(x, v_0)\} = 0$, а значит, выполняется равенство

$$-\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \vec{\nabla} f(x, v_0) = -\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (\vec{\nabla} f(x, v_0) - E_v\{\vec{\nabla} f(x, v_0)\}).$$

Правая часть этого равенства согласно теореме 6 и условию 7 асимптотически нормальна с нулевым математическим ожиданием и дисперсией $E\{\vec{\nabla} f(x, v_0)(\vec{\nabla} f(x, v_0))'\}$, т.е.

$$-\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \vec{\nabla} f(x, v_0) \to N(0, C), \ n \to \infty.$$

На основании изложенного, равенства (4) и теоремы 7 можно сделать вывод о справедливости утверждения в первой части теоремы, а именно

$$\sqrt{n}(v_n - v_0) \xrightarrow{D} N(0, (A_0)^{-1} C (A_0)^{-1}) .$$

Перейдем к доказательству второй части теоремы. Применим формулу Тейлора непосредственно к функции $F_n(x, v_n)$ в окрестности точки v_0:

$$F_n(x, v_n) - F_n(x, v_0) = (\vec{\nabla} F_n(x, v_0))'(v_n - v_0) +$$

$$+\frac{1}{2}(v_n - v_0)' \left(\frac{\partial^2 F_n(x, v)}{\partial v_j \, \partial v_k} \right)_{j,k=1}^{l} (v_n - v_0) ,$$

где $v = v_0 + \lambda(v_n - v_0)$, $0 < \lambda < 1$.

Обозначим

$$B_v = \left(\frac{\partial^2 F_n(x, v)}{\partial v_j \, \partial v_k} \right)_{j,k=1}^{l} = \left(\frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2 f(x^i, v)}{\partial v_j \, \partial v_k} \right)_{j,k=1}^{l}$$

и выразим $F_n(x, v_n)$ в виде

$$F_n(x, v_n) = \left(\frac{1}{n} \sum_{i=1}^{n} \vec{\nabla} f(x^i, v_0) \right)' (v_n - v_0) + \frac{1}{2}(v_n - v_0)' B_v (v_n - v_0) + \frac{1}{n} \sum_{i=1}^{n} f(x^i, v_0) .$$

Следовательно,

$$\sqrt{n}(F_n(x, v_n) - F(v_0)) = \frac{1}{\sqrt{n}} \left(\sum_{i=1}^{n} \vec{\nabla} f(x^i, v_0) \right)' (v_n - v_0) +$$

$$+\frac{\sqrt{n}}{2}(v_n - v_0)' B_v (v_n - v_0) + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (f(x^i, v_0) - E\{f(x, v_0)\}) .$$

Оценим каждое слагаемое правой части равенства.

Так как последовательность v_n сходится с вероятностью 1 к v_0, можем говорить о сходимости по вероятности к нулю первого слагаемого:

$$\frac{1}{\sqrt{n}} \left(\sum_{i=1}^{n} \vec{\nabla} f(x^i, v_0) \right)' (v_n - v_0) \to 0, \ n \to \infty .$$

Обозначим

$$\psi_n^{jk}(v) = \frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2 f(x^i, v)}{\partial v_j \, \partial v_k}$$

и применим следствие 1 теоремы 4 к последовательности функций $Q_n = \psi_n^{jk}(v)$. В силу такого выбора функции Q_n следует справедливость условия 1 (непрерывности) и условия 2 (измеримости) для функции $\psi_n^{jk}(v)$.

Условие 3 выполняется, так как v_n — состоятельная оценка параметра, т.е.

$$P\{\lim_{n \to \infty} || v_n - v_0 || = 0\} = 1 .$$

Зафиксируем j, k. Согласно закону больших чисел с учетом условия 5 теоремы 8 справедливо равенство

$$P\{\lim_{n \to \infty} \psi_n^{jk}(v) = \psi^{jk}(v)\} = 1 ,$$

где $\psi^{jk}(v)$ — элементы матрицы $A = E\{\Phi(x, v)\} = (\psi^{jk}(v))_{j,k=1}^{l}$, что, по сути, и является условием 4 следствия 1 .

Таким образом, все условия следствия 1 теоремы 4 справедливы. Значит,

$$P\{\psi_n^{jk}(v) = \psi^{jk}(v_0), n \to \infty\} = 1,$$

а следовательно,

$$P\{\lim_{n\to\infty} B_v \to A_0, n \to \infty\} = 1.$$

В таком случае, поскольку последовательность сходится с вероятностью 1 к v_0,

$$\frac{\sqrt{n}}{2}(v_n - v_0)' B_u (v_n - v_0) \to 0, \ n \to \infty .$$

В результате условия 8 и 9 позволяют записать

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (f(x^i, v_0) - E\{f(x, v_0)\}) \xrightarrow{D} N(0, \sigma^2).$$

Для завершения доказательства применим теорему 2 и таким образом получим справедливость второй части теоремы 3.

СПИСОК ЛИТЕРАТУРЫ

1. Winkler G. Analysis, random fields and dynamic Monte Carlo methods. — Berlin: Springer, 1995. — 325 p.
2. Дороговцев А.Я. Теория оценок параметров случайных процессов. — Київ: Вища шк., 1982. — 80 с.
3. Ибрагимов И.А., Линник Ю.В. Независимые и стационарно связанные величины. — М.: Наука, 1965. — 524 с.
4. Knopov P.S., Kasitskaya E.J. Properties of empirical estimates in stochastic optimization and identification problems // Ann. Oper. Res. — 1995. — **56**. — P. 225–239.
5. Королюк В.С., Портенко Н.И., Скороход А.В., Турбин А.Ф. Справочник по теории вероятности и математической статистике. — М.: Наука, 1985. — 640 с.
6. Schmetterer L. Introduction to mathematical statistics. — Berlin: Springer-Verlag, 1974. — 502 p.

Поступила 04.04.2012