УДК 519.234.3

ФУНКЦИОНАЛЫ ВЛИЯНИЯ РОБАСТНЫХ ОЦЕНОК ПАРАМЕТРОВ АВТОРЕГРЕССИОННЫХ ПОЛЕЙ

В.Б. Горяинов

МГТУ им. Н.Э. Баумана, Москва
e-mail: vb-goryainov@mail.ru

Для оценок наименьших квадратов и наименьших модулей, М-оценок и обобщенных М-оценок коэффициентов авторегрессионных полей вычислены функционалы влияния и коэффициенты чувствительности к большой ошибке в наблюдениях.

Ключевые слова: пространственная авторегрессия, М-оценки, функционал влияния, коэффициент чувствительности к большой ошибке.

INFLUENCE FUNCTIONALS OF ROBUST ESTIMATIONS OF PARAMETERS OF AUTOREGRESSIVE FIELDS

V.B. Goryainov

Bauman Moscow State Technical University, Moscow
e-mail: vb-goryainov@mail.ru

Influence functionals and coefficients of sensitivity to a gross error in observations are calculated for estimations of least squares and least moduli, M-estimations and generalized M-estimations of autoregressive fields.

Keywords: spatial autoregression, M-estimation, influence functional, gross error sensitivity coefficient.

Введение. Рассмотрим стационарное поле X_{ij} на целочисленной прямоугольной решетке, описываемое разностным авторегрессионным уравнением

$$X_{ij} = a_{10}X_{i-1,j}+a_{01}X_{i,j-1}+a_{11}X_{i-1,j-1}+\varepsilon_{ij}, \quad i,j = 0, \pm1, \pm2, \ldots, \quad (1)$$

где $a = (a_{10}, a_{01}, a_{11})^T$ — (неслучайные) коэффициенты, а ε_{ij} – независимые одинаково распределенные случайные величины с нулевым математическим ожиданием $\mathsf{E}\varepsilon_{ij}$ и конечной дисперсией $\mathsf{D}\varepsilon_{ij}$. Такие поля описывают различные характеристики изображений (яркость, интенсивность, градации серого и т.д.) в теории распознавания образов и обработки изображений, одной из основных задач которой является фильтрация изображений на фоне шума посредством оценивания авторегрессионных коэффициентов $a = (a_{10}, a_{01}, a_{11})^T$ по наблюдениям X_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n$ [1].

Если обновляющее поле ε_{ij} является гауссовским, то наилучшими оценками параметра a будут оценки наименьших квадратов [2], определяемые как точка минимума функции

$$\mathcal{L}_{LS}(a) = \sum_{i=1}^{m}\sum_{j=1}^{n}\left(X_{ij} - a_{10}X_{i-1,j} - a_{01}X_{i,j-1} - a_{11}X_{i-1,j-1}\right)^2. \qquad (2)$$

На практике, однако, предположение о гауссовости ε_{ij} (а значит, и X_{ij}) обычно нарушается из-за грубых ошибок в измерении X_{ij} [3]. С помощью компьютерного моделирования было показано, что в этом случае оценки наименьших квадратов уступают в эффективности знаковым и ранговым оценкам, оценкам наименьших модулей, М-оценкам и обобщенным М-оценкам [4–8].

В данной работе вводятся теоретические характеристики устойчивости оценок к засорению наблюдений грубыми ошибками, которые вычисляются для М-оценок, обобщенных М-оценок, оценок наименьших модулей и оценок наименьших квадратов параметра a. Подтверждается вывод о предпочтительности обобщенных М-оценок при засорении наблюдений грубыми ошибками.

Постановка задачи. М-оценки параметра a по наблюдениям X_{ij}, $i = 1,\ldots,m$, $j = 1,\ldots,n$, определяются [7] как точка минимума функции

$$\mathcal{L}_{M}(a) = \sum_{i=1}^{m}\sum_{j=1}^{n}\rho(X_{ij} - a_{10}X_{i-1,j} - a_{01}X_{i,j-1} - a_{11}X_{i-1,j-1}), \qquad (3)$$

где, например,

$$\rho_H(x) = \begin{cases} x^2, & \text{если } |x| \le k, \\ 2k|x| - k^2, & \text{если } |x| > k, \end{cases} \qquad (4)$$

— семейство функций Хьюбера [9, 10], $k > 0$, или

$$\rho_T(x) = \begin{cases} 1 - \left(1 - \left(\dfrac{x}{k}\right)^2\right)^3, & \text{если } |x| \le k, \\ 1, & \text{если } |x| > k, \end{cases}$$

— семейство функций, называемое бивесом Тьюки [9, 10], $k > 0$. Рекомендации по выбору $\rho(x)$ имеются в [9, 10]. В частности, если $\rho(x) = x^2$, то получаются оценки наименьших квадратов, а если $\rho(x) = |x|$ — оценки наименьших модулей.

Пусть плотность f случайных величин ε_{ij} является смесью двух гауссовских плотностей:

$$f(x) = (1-\gamma)\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{x^2}{2\sigma_1^2}} + \gamma\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{x^2}{2\sigma_2^2}}, \quad 0 < \gamma < 1, \qquad (5)$$

имитирующих появление с небольшой вероятностью γ среди ε_{ij}, $\mathsf{D}\varepsilon_{ij} = \sigma_1^2$, величин с аномально большой диспрсией σ_2^2, $\sigma_2^2 \gg \sigma_1^2$.

Тогда, например, для семейства функций Хьюбера (4) при $k \in (1{,}5; 2)$ М-оценки почти не уступают в эффективности оценкам наименьших квадратов при $\gamma = 0$ в (5) и почти так же эффективны, как оценки наименьших модулей при $\gamma \in (0{,}05; 0{,}2)$, когда эффективность оценок наименьших квадратов невысока (см. [7]).

Предположим теперь, что вместо поля X_{ij} наблюдается поле Y_{ij} вида

$$Y_{ij} = X_{ij} + \nu_{ij}\zeta_{ij}, \tag{6}$$

где ζ_{ij} — независимые одинаково распределенные случайные величины, а ν_{ij} — независимые бернуллиевские случайные величины, принимающие значения 1 и 0 с вероятностями δ и $1 - \delta$ соответственно, $0 \leq \delta \leq 1$. Предположим, что поля ε_{ij}, ν_{ij} и ζ_{ij} не зависят друг от друга. Модель (6) описывает загрязнение поля X_{ij} небольшой долей δ (обычно на практике $0 < \delta < 0{,}2$) случайных ошибок ζ_{ij}. Например, при измерении поля X_{ij} с вероятностью δ происходит сбой измерительной аппаратуры, и во время сбоя вместо X_{ij} наблюдается ζ_{ij}. В этом случае М-оценки теряют эффективность.

Дело в том, что если ρ — выпуклая дифференцируемая функция, то минимизация $\mathcal{L}_M(a)$ в (3) равносильна решению системы уравнений

$$L_M(a) = 0, \tag{7}$$

где

$$L_M(a) = \sum_{i=1}^{m}\sum_{j=1}^{n} \psi\left(X_{ij} - \tilde{X}_{ij}^T a\right)\tilde{X}_{ij}, \tag{8}$$

$\psi(x) = \rho'(x)$, $\tilde{X}_{ij} = (X_{i-1,j}, X_{i,j-1}, X_{i-1,j-1})^T$, а T — символ операции транспонирования. Потеря эффективности М-оценок при загрязнениях (6) связана с неограниченным множителем $\tilde{X}_{ij}$ в системе уравнений (7), влияние которого на решение этой системы может быть сколь угодно велико при замене X_{ij} на Y_{ij} вида (6) и достаточно больших значениях ζ_{ij} в (6).

Определим обобщенные М-оценки параметра a как решение системы

$$L_{GM}(a) = 0, \tag{9}$$

где

$$L_{GM}(a) = \sum_{i=1}^{m}\sum_{j=1}^{n} \psi\left(X_{ij} - \tilde{X}_{ij}^T a\right) g_{ij}(\tilde{X}), \tag{10}$$

$g_{ij}(\tilde{X}) = (g(X_{i-1,j}), g(X_{i,j-1}), g(X_{i-1,j-1}))^T$, а g — некоторая функция. Выбрав в качестве g ограниченную функцию, можно ограничить влияние экстремальных значений ζ_{ij} на решение уравнения (9). Например, в качестве g можно взять половину производной

$$\frac{1}{2}\rho'_H(x) = \begin{cases} x, & \text{если } |x| \le k, \\ k, & \text{если } |x| > k, \end{cases}$$

ρ-функции Хьюбера (4). В этом случае, если $X_{i-1,j}$, $X_{i,j-1}$, $X_{i-1,j-1}$ невелики, то слагаемое $\psi(X_{ij} - \tilde{X}_{ij}^T a) g_{ij}(\tilde{X})$ в (10) совпадает со слагаемым $\psi(X_{ij} - \tilde{X}_{ij}^T a)\tilde{X}_{ij}$ в (8). В противном случае вектор $\psi(X_{ij} - \tilde{X}_{ij}^T a) g_{ij}(\tilde{X})$ будет “подрезан” и его длина никогда не превысит $k^2\sqrt{3}$.

В работах [6–8] чувствительность обобщенных М-оценок, М-оценок и оценок наименьших модулей к загрязнениям вида (6) исследовалась с помощью компьютерного моделирования. Показано, что с ростом δ в (6) разность между оценкой параметра a и самим параметром увеличивается, что может свидетельствовать о смещенности этих оценок. Возникает необходимость в количественной оценке этого смещения, например, как это было сделано для более простых моделей путем определения кривой чувствительности, кривой влияния и функции влияния [9, 10].

Определения и свойства функционала влияния и коэффициента чувствительности к большой ошибке. В работах [6–8] доказана состоятельность обобщенных М-оценок, М-оценок и оценок наименьших модулей при $\delta = 0$ в (6). Исследуем поведение этих оценок при $\delta > 0$.

Предположим, что $\hat{a}_{mn}$ — оценка параметра a и при $m, n \to \infty$ по вероятности $\hat{a}_{mn} \to a(\delta)$. Если $\delta \ne 0$, то оценка $\hat{a}_{mn}$, вообще говоря, перестает быть состоятельной, т.е. $a(\delta) \ne a^{(0)}$. Определим функционал влияния $IF(a(\delta), F_\zeta)$ оценки $\hat{a}_{mn}$ по формуле

$$IF(a(\delta), F_\zeta) = \left. \frac{d}{d\delta} a(\delta) \right|_{\delta=0};$$

$IF(a(\delta), F_\zeta)$ характеризует величину главного линейного члена в разложении асимптотического смещения

$$a(\delta) - a^{(0)} = IF\delta + o(\delta), \quad \delta \to 0,$$

и зависит от $a(\delta)$ и от функции распределения F_ζ случайной величины ζ_{11}.

Лучше других противостоять засорениям вида (6) наблюдений X_{ij} будут оценки с ограниченным $IF(a(\delta), F_\zeta)$. Обозначим через $\mathfrak{F}$ множество возможных функций распределения случайных величин ζ_{ij}. Назовем величину

$$GES(\mathfrak{F}, a(\delta)) = \sup_{F_\zeta \in \mathfrak{F}} |IF(a(\delta), F_\zeta)|$$

коэффициентом чувствительности оценки $\hat{a}_{mn}$ к большой ошибке. Оценку $\hat{a}_{mn}$ будем называть робастной на семействе распределений $\mathfrak{F}$, если $GES(\mathfrak{F}, a(\delta)) < \infty$.

Найдем функционал влияния и коэффициент чувствительности к большой ошибке для обобщенных М-оценок, М-оценок, оценок наименьших квадратов и оценок наименьших модулей.

Обозначим для $i, j = 0, \pm 1, \pm 2, \ldots$

$\tilde{Y}_{ij}{=}(Y_{i-1,j}, Y_{i,j-1}, Y_{i-1,j-1})^T$, $g_{ij}(\tilde{Y}){=}(g(Y_{i-1,j}), g(Y_{i,j-1}), g(Y_{i-1,j-1}))^T$,

$$L_{mn}(a) = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\psi\left(Y_{ij} - \tilde{Y}_{ij}^T a\right) g_{ij}(\tilde{Y}), \tag{11}$$

$$L(\delta, a) = \mathsf{E}\Big[\psi\left(Y_{11} - \tilde{Y}_{11}^T a\right) g_{11}(\tilde{Y})\Big]. \tag{12}$$

Обозначим через $B = \mathsf{E}[\tilde{X}_{11} g_{11}(\tilde{X})^T]$ взаимную ковариационную матрицу векторов $\tilde{X}_{11}$ и $g_{11}(\tilde{X})$:

$$B = \begin{pmatrix} r_{0,0} & r_{1,-1} & r_{0,-1} \\ r_{-1,1} & r_{0,0} & r_{-1,0} \\ r_{0,1} & r_{1,0} & r_{0,0} \end{pmatrix},$$

где

$$r_{\alpha-p,\beta-q} = \mathsf{E}[X_{\alpha\beta} g(X_{pq})], \quad (\alpha, \beta) \in \mathcal{I}, \quad (p, q) \in \mathcal{I},$$
$$\mathcal{I} = \{(1,0), (0,1), (0,0)\}.$$

Теорема. *Пусть $\hat{a}_{mn}$ — решение уравнения $L_{mn}(a) = 0$, $\hat{a}_{mn} \to a(\delta)$ и $L(\delta, a(\delta)) = 0$ для любых достаточно малых δ,*

$$\mathsf{E}|\psi(Y_{11} - \tilde{Y}_{11}^T a) g_{11}(\tilde{Y})|^2 < \infty, \tag{13}$$

$$\mathsf{E}[\psi'(\varepsilon_{11})] \neq 0, \tag{14}$$

существуют и непрерывны $\dfrac{\partial L(\delta, a)}{\partial \delta}$, $\dfrac{\partial L(\delta, a)}{\partial a}$ *в некоторой окрестности* $(0, a^{(0)})$.

Тогда

$$\begin{aligned} &IF(a(\delta), F_\zeta) = \\ &\quad = \frac{1}{\mathsf{E}[\psi'(\varepsilon_{11})]} B^{-1}\Big(\mathsf{E}[\psi(\varepsilon_{11} + \zeta_{11})]\mathsf{E}[(g(X_{01}), g(X_{10}), g(X_{00}))^T] + \\ &\qquad + \mathsf{E}[\psi(\varepsilon_{11} - a_{01}^{(0)}\zeta_{01})(g(X_{01} + \zeta_{01}), g(X_{10}), g(X_{00}))^T] + \\ &\qquad + \mathsf{E}[\psi(\varepsilon_{11} - a_{10}^{(0)}\zeta_{10})(g(X_{01}), g(X_{10} + \zeta_{10}), g(X_{00}))^T] + \\ &\qquad + \mathsf{E}[\psi(\varepsilon_{11} - a_{00}^{(0)}\zeta_{00})(g(X_{01}), g(X_{10}), g(X_{00} + \zeta_{00}))^T]\Big). \end{aligned} \tag{15}$$

Замечание. Так как случайные поля X_{ij}, ν_{ij}, ζ_{ij} предполагаются стационарными, то поле $Z_{ij} = \psi\left(Y_{ij} - \tilde{Y}_{ij}^T a\right) g_{ij}(\tilde{Y})$ как измеримая функция от X_{ij}, ν_{ij}, ζ_{ij} также будет стационарным полем [11, с. 170, 182]. Поэтому по закону больших чисел для стационарных функций [11, с. 181] при $m, n \to \infty$

$$L_{mn}(a) \to L(\delta, a)$$

и при разумном выборе функций ψ и g оценка $\hat{a}_{mn}$ как решение уравнения $L_{mn}(a) = 0$ стремится к $a(\delta)$ — решению уравнения $L(\delta, a(\delta)) = 0$.

Доказательство теоремы. Из условия теоремы следует, что по теореме о неявной функции в некоторой окрестности точки $(0, a^{(0)})$ уравнение $L(\delta, a) = 0$ определяет однозначную дифференцируемую функцию $a = a(\delta)$ и

$$\left.\frac{da}{d\delta}\right|_{\delta=0} = -\frac{\left.\dfrac{\partial L(\delta, a)}{\partial \delta}\right|_{\substack{\delta=0\\ a=a^{(0)}}}}{\left.\dfrac{\partial L(\delta, a)}{\partial a}\right|_{\substack{\delta=0\\ a=a^{(0)}}}}.$$

Определим полную группу событий:

$$H_{ijkl} = \{\nu_{11} = i, \nu_{01} = j, \nu_{10} = k, \nu_{00} = l\}, \quad i, j, k, l = 0, 1.$$

По формуле полного математического ожидания [12, с. 90, 230]

$$L(\delta, a) = \sum_{i,j,k,l=0}^{1} \mathsf{E}\left[\psi\left(Y_{11} - \tilde{Y}_{11}^T a\right) g_{11}(\tilde{Y}) | H_{ijkl}\right] \mathsf{P}(Y_{ijkl}) =$$

$$= \sum_{i,j,k,l=0}^{1} \delta^{i+j+k+l}(1-\delta)^{4-i-j-k-l} \mathsf{E}\left[\psi\left(Y_{11} - \tilde{Y}_{11}^T a\right) g_{11}(\tilde{Y}) | H_{ijkl}\right],$$

где $\mathsf{E}\left[\psi\left(Y_{11} - \tilde{Y}_{11}^T a\right) g_{11}(\tilde{Y}) | H_{ijkl}\right]$ не зависит от δ. Поэтому

$$\left.\frac{\partial L(\delta, a)}{\partial \delta}\right|_{\substack{\delta=0\\ a=a^{(0)}}} =$$

$$= \mathsf{E}\left[\psi\left(Y_{11} - \tilde{Y}_{11}^T a^{(0)}\right) g_{11}(\tilde{Y}) | H_{1000}\right] +$$

$$+ \mathsf{E}\left[\psi\left(Y_{11} - \tilde{Y}_{11}^T a^{(0)}\right) g_{11}(\tilde{Y}) | H_{0100}\right] +$$

$$+ \mathsf{E}\left[\psi\left(Y_{11} - \tilde{Y}_{11}^T a^{(0)}\right) g_{11}(\tilde{Y}) | H_{0010}\right] +$$

$$+\,\mathsf{E}\left[\psi\left(Y_{11}-\tilde{Y}_{11}^{T}a^{(0)}\right)g_{11}(\tilde{Y})|H_{0001}\right]+$$

$$+\,\mathsf{E}\left[\psi\left(Y_{11}-\tilde{Y}_{11}^{T}a^{(0)}\right)g_{11}(\tilde{Y})|H_{0000}\right];$$

$$\mathsf{E}\left[\psi\left(Y_{11}-\tilde{Y}_{11}^{T}a^{(0)}\right)g_{11}(\tilde{Y})|H_{1000}\right]=$$
$$=\mathsf{E}\left[\psi\left(X_{11}+\zeta_{11}-a_{10}^{(0)}X_{01}-a_{01}^{(0)}X_{10}-a_{11}^{(0)}X_{00}\right)\right]\times$$
$$\times\,\mathsf{E}[(g(X_{01}),g(X_{10}),g(X_{00}))^{T}]=$$
$$=\mathsf{E}[\psi(\varepsilon_{11}+\zeta_{11})]\mathsf{E}[(g(X_{01}),g(X_{10}),g(X_{00}))^{T}].$$

Аналогично

$$\mathsf{E}\left[\psi\left(Y_{11}-\tilde{Y}_{11}^{T}a^{(0)}\right)g_{11}(\tilde{Y})|H_{0100}\right]=$$
$$=\mathsf{E}[\psi(\varepsilon_{11}-a_{01}^{(0)}\zeta_{01})(g(X_{01}+\zeta_{01}),g(X_{10}),g(X_{00}))^{T}],$$

$$\mathsf{E}\left[\psi\left(Y_{11}-\tilde{Y}_{11}^{T}a^{(0)}\right)g_{11}(\tilde{Y})|H_{0010}\right]=$$
$$=\mathsf{E}[\psi(\varepsilon_{11}-a_{10}^{(0)}\zeta_{10})(g(X_{01}),g(X_{10}+\zeta_{10}),g(X_{00}))^{T}],$$

$$\mathsf{E}\left[\psi\left(Y_{11}-\tilde{Y}_{11}^{T}a^{(0)}\right)g_{11}(\tilde{Y})|H_{0001}\right]=$$
$$=\mathsf{E}[\psi(\varepsilon_{11}-a_{00}^{(0)}\zeta_{00})(g(X_{01}),g(X_{10}),g(X_{00}+\zeta_{00}))^{T}],$$

$$\mathsf{E}\left[\psi\left(Y_{11}-\tilde{Y}_{11}^{T}a^{(0)}\right)g_{11}(\tilde{Y})|H_{0000}\right]=$$
$$=\mathsf{E}[\psi(\varepsilon_{11})]\mathsf{E}[(g(X_{01}),g(X_{10}),g(X_{00}))^{T}]=0.$$

Далее

$$\left.\frac{\partial L(\delta,a)}{\partial a}\right|_{\substack{\delta=0\\ a=a^{(0)}}}=$$
$$=\frac{\partial}{\partial a}\left(\mathsf{E}\left[\psi\left(Y_{11}-\tilde{Y}_{11}^{T}a\right)g_{11}(\tilde{Y})|H_{0000}\right]\right)\bigg|_{\substack{\delta=0\\ a=a^{(0)}}}=-\mathsf{E}[\psi'(\varepsilon_{11})]B.$$

Поэтому

$$\left.\frac{da}{d\delta}\right|_{\delta=0}=\frac{1}{\mathsf{E}[\psi'(\varepsilon_{11})]}B^{-1}\Big(\mathsf{E}[\psi(\varepsilon_{11}+\zeta_{11})]\mathsf{E}[(g(X_{01}),g(X_{10}),g(X_{00}))^{T}]+$$

$$+\,\mathsf{E}[\psi(\varepsilon_{11}-a_{01}^{(0)}\zeta_{01})(g(X_{01}+\zeta_{01}),g(X_{10}),g(X_{00}))^{T}]+$$

$$+ \mathsf{E}[\psi(\varepsilon_{11} - a_{10}^{(0)}\zeta_{10})(g(X_{01}), g(X_{10} + \zeta_{10}), g(X_{00}))^T] +$$

$$+ \mathsf{E}[\psi(\varepsilon_{11} - a_{00}^{(0)}\zeta_{00})(g(X_{01}), g(X_{10}), g(X_{00} + \zeta_{00}))^T]\Big).$$

Теорема доказана.

Из теоремы следует, что если функции ψ и g ограничены, то функционал влияния обобщенных М-оценок ограничен и коэффициент чувствительности к большой ошибке будет конечным.

Для обычных М-оценок $g(x) = x$ и поэтому $\mathsf{E}[g(X_{ij})] = \mathsf{E}[X_{ij}] = 0$, а B — ковариационная матрица вектора (X_{01}, X_{10}, X_{00}), следовательно,

$$IF(a(\delta), F_\zeta) = \\ = \frac{1}{\mathsf{E}[\psi'(\varepsilon_{11})]} B^{-1}\Big(\mathsf{E}[\psi(\varepsilon_{11} - a_{01}^{(0)}\zeta_{01})\zeta_{01}], \mathsf{E}[\psi(\varepsilon_{11} - \\ - a_{10}^{(0)}\zeta_{10})\zeta_{10}], \mathsf{E}[\psi(\varepsilon_{11} - a_{00}^{(0)}\zeta_{00})\zeta_{00}]\Big)^T. \quad (16)$$

Видно, что ограниченность функции ψ уже не является достаточной для конечности $GES(\mathfrak{F}, a(\delta))$. Если ψ ограничена, но

$$\sup_{C\in\mathbb{R}} \Big|\mathsf{E}[\psi(\varepsilon_{11} - a_{11}^{(0)}C)C\Big| = \infty,$$

то $IF(a(\delta), F_\zeta)$ может быть сколь угодно большим для такой ψ, например, если ошибки $\zeta_{ij} = C$ являются неслучайными, т.е. $\zeta_{ij} = C$, где $C \in \mathbb{R}$ — произвольная постоянная. Поэтому если класс распределений $\mathfrak{F}$ содержит всевозможные δ-функции (функции распределения постоянных C, $C \in \mathbb{R}$), то коэффициент чувствительности к большой ошибке для обычных М-оценок может быть неограниченным и М-оценки в этом случае будут неробастными, вообще говоря, даже для ограниченной функции $\psi(x)$.

Обозначив плотности случайных величин ε_{ij} и ζ_{ij} через $f_\varepsilon(x)$ и $f_\zeta(x)$, получим, что

$$\mathsf{E}[\psi(\varepsilon_{11} - a_{ij}^{(0)}\zeta_{ij})\zeta_{ij}] = \int_{-\infty}^{\infty} \psi(x - a_{ij}^{(0)}y) y f_\varepsilon(x) f_\zeta(y)\, dxdy.$$

Отсюда следует, что для конечности коэффициента чувствительности к большой ошибке достаточно условия

$$\sup_{(x,y)\in\mathbb{R}^2} \Big|\psi(x - a_{ij}^{(0)}y)y\Big| < \infty.$$

Вычислим функционал влияния для двух частных случаев М-оценок — оценок наименьших модулей и оценок наименьших квадратов.

Оценка наименьших модулей — точка минимума функции

$$\mathcal{L}_{LD}(a) = \sum_{i=1}^{m}\sum_{j=1}^{n} |X_{ij} - a_{10}X_{i-1,j} - a_{01}X_{i,j-1} - a_{11}X_{i-1,j-1}|$$

или, что равносильно, решение уравнения $L_{LD}(a) = 0$, где

$$L_{LD}(a) = \sum_{i=1}^{m}\sum_{j=1}^{n} \psi\left(X_{ij} - \tilde{X}_{ij}^T a\right) g_{ij}(\tilde{X}).$$

Оценка наименьших модулей — частный случай М-оценки при $\rho(x) = |x|$ или $\psi(x) = \operatorname{sign}(x)$, где

$$\operatorname{sign}(x) = \begin{cases} 1, & \text{если } x \geq 0, \\ -1, & \text{если } x < 0. \end{cases}$$

Отметим, что

$$\operatorname{sign}(x) = 1 - 2I(\{x < 0\}),$$

где $I(A)$ — индикаторная функция множества A. Обозначим через $\mathfrak{A}_\zeta$ σ-алгебру, порожденную случайной величиной ζ_{ij}. Воспользовавшись формулой полного математического ожидания, получим

$$\mathsf{E}[\zeta_{ij}\operatorname{sign}(\varepsilon_{11} - a_{ij}^{(0)}\zeta_{ij})] = \mathsf{E}\left(\mathsf{E}[\zeta_{ij}\operatorname{sign}(\varepsilon_{11} - a_{ij}^{(0)}\zeta_{ij})|\mathfrak{A}_\zeta]\right) = \\ = \mathsf{E}\left(\zeta_{ij}\mathsf{E}[\operatorname{sign}(\varepsilon_{11} - a_{ij}^{(0)}\zeta_{ij})|\mathfrak{A}_\zeta]\right) = \mathsf{E}\left(\zeta_{ij}\mathsf{E}[1 - 2F_\varepsilon(a_{ij}^{(0)}\zeta_{ij})]\right).$$

Отметим, что $\mathsf{E}[\psi'(\varepsilon_{11})] = 2f_\varepsilon(0)$. Поэтому формула (16) превращается в

$$IF(a(\delta), F_\zeta) = \frac{1}{2f_\varepsilon(0)} B^{-1}(e_{01}, e_{10}, e_{00})^T, \tag{17}$$

где $e_{ij} = \mathsf{E}(\zeta_{ij}\mathsf{E}[1 - 2F_\varepsilon(a_{ij}^{(0)}\zeta_{ij})]$, а F_ε — функция распределения случайных величин ε_{ij}.

Найдем функционал влияния для оценки наименьших квадратов. Оценка наименьших квадратов является точкой минимума (3) с $\rho(x) = x^2$ или, что эквивалентно, решением (9) с $\psi(x) = x$, $g(x) = x$. Подставляя в (15) $\psi(x) = x$, $g(x) = x$, учитывая независимость ε_{11} от ζ_{ij}, а также, что $\mathsf{E}[\varepsilon_{ij}] = 0$, $\mathsf{E}[\psi'(\varepsilon_{11})] = 1$, получаем

$$IF(a(\delta), F_\zeta) = \mathsf{E}[\zeta_{00}^2] B^{-1}\left(a_{01}^{(0)}, a_{10}^{(0)}, a_{00}^{(0)}\right)^T, \tag{18}$$

где B — ковариационная матрица вектора (X_{01}, X_{10}, X_{00}).

Сравнение (17) и (18) показывает, что оценки наименьших модулей предпочтительнее оценок наименьших квадратов, поскольку $IF(a(\delta), F_\zeta)$ в (17) линейно зависит от ζ_{ij}, а в (18) — квадратично.

Заключение. Определены такие инфинитезимальные характеристики робастности оценок коэффициентов авторегрессионного поля, как функционал влияния и коэффициент чувствительности к большой ошибке. Эти характеристики вычислены для обычных и обобщенных М-оценок, в частности для оценок наименьших модулей и наименьших квадратов. Сделан вывод о предпочтительности обобщенных М-оценок в условиях искажения наблюдений авторегрессионного поля грубыми ошибками.

СПИСОК ЛИТЕРАТУРЫ

1. Olivier A., Olivier C. Choice of a 2-D causal autoregressive texture model using information criteria // Pattern Recognition Letters. – 2003. – Vol. 24. No. 9–10. – P. 1191–1201.
2. Tjostheim D. Statistical spatial series modelling // Advances in Applied Probability. – 1978. – Vol. 10. No. 1. – P. 130–154.
3. Kashyap R., Eom K. Robust image techniques with and image restoration application // IEEE Trans. on Acoustics, Speech, and Signal Processing. – 1988. – Vol. 36. No. 8. – P. 1313Ц-1325.
4. Горяинов В. Б., Горяинова Е. Р. Непараметрическая идентификация пространственной модели авторегрессии в условиях априорной стохастической неопределенности // Автоматика и телемеханика. – 2010. – № 2. – С. 31–41.
5. Горяинов В. Б. Идентификация пространственной авторегрессии ранговыми методами // Автоматика и телемеханика. – 2011. – № 5. – С. 82–95.
6. Горяинов В. Б. Оценки наименьших модулей коэффициентов пространственной авторегрессии. // Изв. РАН. Теория и системы управления. – 2011. № 4. – С. 58–65.
7. Горяинов В. Б. М-оценки коэффициентов пространственной авторегрессии // Автоматика и телемеханика. – 2012. – № 8. – С. 119–129.
8. Горяинов В. Б. Обобщенные М-оценки коэффициентов авторегрессионного поля // Автоматика и телемеханика. – 2012. – № 10. – С. 42–51.
9. Хьюбер П. Дж. Робастность в статистике. – М.: Мир, 1984.
10. Хампель Ф., Рончетти Э., Рауссеу П., Штаэль В. Робастность в статистике. Подход на основе функций влияния. – М.: Мир, 1989.
11. Stout W. F. Almost sure convergence. – New York: Academic Press, 1974.
12. Ширяев А. Н. Вероятность. – М.: Наука., 1980.

Статья поступила в редакцию 2.03.2012