УДК 681.5.015

ИССЛЕДОВАНИЕ МЕТОДОВ ПАРАМЕТРИЧЕСКОЙ ИДЕНТИФИКАЦИИ ПРИ НАЛИЧИИ НЕТИПИЧНЫХ ВОЗМУЩЕНИЙ ПО ЛОКАЛЬНЫМ РАЗЛОЖЕНИЯМ ЦЕЛЕВЫХ ФУНКЦИЙ

А.А. Ломов

Институт математики им. С.Л. Соболева СО РАН;

Новосибирский государственный университет

Россия, 630090, Новосибирск, пр. Академика Коптюга, 4; ул. Пирогова, 2

E-mail: lomov@math.nsc.ru; a.lomov@g.nsu.ru

А. В. Федосеев

Институт математики им. С.Л. Соболева СО РАН;

Новосибирский государственный университет

Россия, 630090, Новосибирск, пр. Академика Коптюга, 4; ул. Пирогова, 2

E-mail: alexeyfedoseev.nsk@gmail.com

Ключевые слова: линейные динамические системы, идентификация параметров, функции чувствительности, нетипичные возмущения.

Аннотация: В докладе изучается возможность сравнения методов идентификации в условиях нетипичных возмущений с использованием теории чувствительности. Для достижения цели получены локальные разложения целевых функций линейного МНК, метода инструментальных переменных в частотной области и вариационного метода (STLS, GTLS) в условиях смешанного шума — в наблюдениях и в невязке уравнения. Теоретические результаты проверяются вычислительным моделированием.

1. Введение

Субъективность посылок о характере шумов в реальных экспериментах приводит к определенного рода скепсису относительно применения асимптотической статистической теории, и возникают вопросы об устойчивости того или иного метода идентификации в случае нарушения априорных постулатов о свойствах возмущений [1]. Одним из общих подходов здесь является применение методов теории чувствительности [2], основанных на разложении целевой функции (правдоподобия) (ЦФ) вблизи точки глобального минимума в предположении малости шумов, как типичных, так и нетипичных для исследуемого метода идентификации [3,4]. Линеаризация градиента ЦФ позволяет описать локальный разброс оценок. Для исследования смещенности требуется вычисление членов более высокого порядка, чем линейные, что в общем случае оказывается весьма непростой задачей. Также остается открытым вопрос о величине возмущений, при которых локальные разложения адекватно отражают реальное поведение оценок.

В докладе исследуются границы применимости подхода, основанного на линеаризации градиента ЦФ, для количественного сравнения разброса оценок на примере сравнения трех методов идентификации параметров системы разностных уравнений при смешанных возмущениях (аддитивных и в невязке уравнения). Теоретические результаты проверяются вычислительным экспериментом.

2. Целевые функции

Рассматривается система разностных уравнений

$$
\begin{aligned}
x[k+1] &= A_\theta x[k] + B_\theta u[k] + w[k], \quad k = \overline{1, N-1}, \\
\check{z}[k] &\doteq \begin{bmatrix} \check{x}[k] \\ \check{u}[k] \end{bmatrix} = \begin{bmatrix} x[k] \\ u[k] \end{bmatrix} + \begin{bmatrix} \xi_x[k] \\ \xi_u[k] \end{bmatrix} \doteq z[k] + \xi[k],
\end{aligned} \tag{1}
$$

где $x[k] \in \mathbb{R}^n$, $u[k] \in \mathbb{R}^m$ — переменные, $A_\theta \in \mathbb{R}^{n\times n}$, $B_\theta \in \mathbb{R}^{n\times m}$ — матрицы, зависящие от фиксированного параметра $\theta \in \mathbb{R}^v$, подлежащего идентификации. Векторы $w[k]$ и $\xi[k]$ играют роль независимых случайных возмущений, $\check{z}[k] \in \mathbb{R}^{n+m}$ — наблюдаемые переменные. Величины $w[k], \xi[k], w[l], \xi[l]$ для $k \neq l$ статистически независимы и распределены с нулевыми математическими ожиданиеми и диагональными матрицами вторых моментов, $\xi \doteq \left[\xi[1]^{\mathrm{T}}, \ \ldots, \ \xi[N]^{\mathrm{T}}\right]^{\mathrm{T}}$,

$$
\mathbf{D}\, w[k] = \sigma_w^2 I_n, \quad \mathbf{D}\, \xi = I_N \otimes \operatorname{diag}\left[\sigma_1^2 \quad \ldots \quad \sigma_{n+m}^2\right] \doteq K. \tag{2}
$$

Для последующей линеаризации налагается ограничение на диаметры носителей распределений: $\operatorname{diam}\operatorname{supp}\mathrm{P}_{\xi,w} < \mathrm{const}\cdot\max\{\sigma_\xi, \sigma_w\}$.

Решением задачи идентификации (оценкой вектора параметров θ) называем функцию наблюдений $\hat{\theta}(\check{z}) = \hat{\theta}(\check{z}[1], \ldots, \check{z}[N]) : \ \mathbb{R}^{N(n+m)} \to \mathbb{R}^v$ со свойством непрерывности $\lim_{\|\check{z}\|\to\|z\|} \hat{\theta}(\check{z}) = \theta$. В статье сравниваются оценки, получаемые тремя методами: 1) линейным методом наименьших квадратов (МНК); 2) методом инструментальных переменных в частотной области (МИП) [5] 3) вариационным методом (ВМ) [6] (за рубежом его разновидности известны как модифицированный метод Прони [7], метод STLS (Structured Total Least Squares) [8], метод GTLS (Global Total Least Squares) [9]). Каждая из рассматриваемых задач идентификации формулируется как минимизация некоторой целевой функции $\mathbb{R}^v \to \mathbb{R}$. Для оценок МНК и МИП известны явные решения в виде формул, для вариационного метода существуют итерационные процедуры с хорошими свойствами сходимости. Выбор методов идентификации в контексте статьи не играет решающей роли и призван проиллюстрировать методику сравнения. Известно, что оценки МНК являются статистически оптимальными при возмущениях в невязке уравнения, а оценки ВМ асимптотически оптимальны при возмущениях в наблюдениях решений. Одной из целей работы является изучение свойств оценок при возмущениях смешанного типа. Оценки МИП привлекают внимание своей вычислительной простотой и тем, что они используются зарубежными инженерами для идентификации коэффициентов уравнений летательных аппаратов [5], будучи в определенном смысле более устойчивыми, чем оценки МНК.

Система уравнений (1) записывается в матричном виде

$$(3)\qquad \underbrace{\begin{bmatrix} \gamma_0 & \gamma_1 & 0 & 0 & \dots & 0 \\ 0 & \gamma_0 & \gamma_1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & 0 & \gamma_0 & \gamma_1 \end{bmatrix}}_{G_\theta} \underbrace{\begin{bmatrix} z\,[1] \\ z\,[2] \\ \vdots \\ z\,[N] \end{bmatrix}}_{z} = \underbrace{\begin{bmatrix} w\,[1] \\ w\,[2] \\ \vdots \\ w\,[N-1] \end{bmatrix}}_{w}.$$

Здесь

$$(4)\qquad \gamma_0 \doteq -\begin{bmatrix} A_\theta & B_\theta \end{bmatrix}, \quad \gamma_1 \doteq \begin{bmatrix} I_n & 0 \end{bmatrix}, \quad \gamma_{0,1} \in \mathbb{R}^{n\times(n+m)}.$$

В матричной записи $G_\theta z = w$. Имеет место тождество $G_\theta z \equiv V\gamma$ при следующих обозначениях:

$$\gamma \doteq \operatorname{vect}\begin{bmatrix} \gamma_0 & \gamma_1 \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^{2n(n+m)}, \quad \operatorname{vect}\begin{bmatrix} a^{\mathrm{T}} & b^{\mathrm{T}} \\ c^{\mathrm{T}} & d^{\mathrm{T}} \end{bmatrix} \doteq \begin{bmatrix} a^{\mathrm{T}} & b^{\mathrm{T}} & c^{\mathrm{T}} & d^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}},$$

$$V \doteq \bar{V} \otimes I_n, \quad \bar{V} \doteq \begin{bmatrix} z\,[1]^{\mathrm{T}} & z\,[2]^{\mathrm{T}} \\ z\,[2]^{\mathrm{T}} & z\,[3]^{\mathrm{T}} \\ \vdots & \vdots \\ z\,[N-1]^{\mathrm{T}} & z\,[N]^{\mathrm{T}} \end{bmatrix} \in \mathbb{R}^{(N-1)\times 2(n+m)}.$$

Пусть $\gamma = D\theta + d$, где матрица $D \in \mathbb{R}^{n(n+m)\times v}$ и столбец $d \in \mathbb{R}^{n(n+m)\times 1}$ заданы, $\operatorname{rank}\begin{bmatrix} D & d \end{bmatrix} = \max = v+1$. Тогда можно записать

$$G_\theta z = V\gamma = VD\theta + Vd \doteq W_1\theta + W_2 = \begin{bmatrix} W_1 & W_2 \end{bmatrix}\begin{bmatrix} \theta \\ 1 \end{bmatrix}.$$

Обозначим $\check{V} \doteq V(\check{z})$. Оценка МНК находится минимизацией по θ целевой функции

$$(5)\qquad J_{\text{МНК}}(\theta) \doteq \|G_\theta \check{z}\|^2 = \left(\check{W}_1\theta + \check{W}_2\right)^{\mathrm{T}}\left(\check{W}_1\theta + \check{W}_2\right) = \gamma^{\mathrm{T}}\check{V}^{\mathrm{T}}\check{V}\gamma,$$

$$\hat{\theta}_{\text{МНК}} = -\left(\check{W}_1^{\mathrm{T}}\check{W}_1\right)^{-1}\check{W}_1^{\mathrm{T}}\check{W}_2.$$

Оценка МИП [5] имеет вид $\hat{\theta}_{\text{МИП}} = -\left[\operatorname{Re}\left(\widetilde{Y}^*\widetilde{W}_1\right)\right]^{-1}\operatorname{Re}\left(\widetilde{Y}^*\widetilde{W}_2\right)$, где знак $*$ обозначает транспонирование и комплексное сопряжение, $\widetilde{W}_{1,2} \doteq \Phi F\check{W}_{1,2}$, $\Phi \doteq \bar{\Phi} \otimes I_n$, $F \doteq \bar{F} \otimes I_n$ — матрицы дискретного преобразования Фурье $\bar{\Phi} \in \mathbb{C}^{M\times(N-1)}$ и префильтрации $\bar{F} \in \mathbb{R}^{(N-1)\times(N-1)}$, $\widetilde{Y} \doteq \Phi W_1(z)$ — инструментальные переменные. Соответствующая целевая функция имеет вид

$$J_{\text{МИП}}(\theta) = \left(\widetilde{W}_1\theta + \widetilde{W}_2\right)^*\widetilde{Y}\widetilde{Y}^*\left(\widetilde{W}_1\theta + \widetilde{W}_2\right).$$

Целевая функция вариационного метода [6] строится из соображений наилучшей аппроксимации наблюдений поведением модели:

$$(6)\qquad J_{\text{ВМ}}(\theta) = (\check{z} - z_{\text{opt}}(\theta))^{\mathrm{T}} K^{\#}(\check{z} - z_{\text{opt}}(\theta)) \doteq \|\check{z} - z_{\text{opt}}(\theta)\|^2_{K^{\#}},$$

$$(7)\qquad z_{\text{opt}}(\theta) \doteq \arg\min_{z:G_\theta z=0}\|\check{z} - z\|^2_{K^{\#}} = \left[I - KG_\theta^{\mathrm{T}}C_\theta G_\theta\right]\check{z}, \quad C_\theta \doteq (G_\theta K G_\theta^{\mathrm{T}})^{-1}.$$

Здесь $K^{\#}$ — псевдообратная для K (2). Величина $z_{\text{opt}}(\theta)$ играет роль оценки процесса z для параметра θ. С учетом равенств (7) имеем $J_{\text{ВМ}}(\theta) = \gamma^{\mathrm{T}}\check{V}^{\mathrm{T}}C_\theta\check{V}\gamma$ (ср. (5)). Можно показать, что задача минимизации целевой функции (6) равносильна решению задач вида STLS [8] или GTLS [9]. Минимум целевой функции (6) устойчиво находится с помощью итерационных процедур с большим радиусом сходимости [6, 7].

3. Дисперсии оценок при малых возмущениях

Сравним дисперсии оценок МНК, МИП и ВМ в предположении малости возмущений в невязке и наблюдениях, применяя линеаризацию градиента целевой функции. Для упрощения формул наложим несущественное условие $\sigma_1 = \sigma_2 = \ldots = \sigma_{n+m} \doteq \sigma_\xi$.

Пусть $J = J(\theta, z)$ — целевая функция, и $\theta(z) \doteq \arg\min_\tau J(\tau, z)$ — оценка параметра θ по наблюдению z. В пределе $\sigma_w, \sigma_\xi \to 0$ выполним разложение градиента J'_θ в ряд Тейлора относительно точки $\theta(z)$:

$$J'_\theta(\theta + \Delta\theta, z + \Delta z) = J'_\theta(\theta, z) + J''_{\theta\theta}\Delta\theta + J''_{\theta z}\Delta z + O(\sigma^2_{w,\xi}). \tag{8}$$

Верны равенства $J'_\theta(\theta, z) = 0$, $J'_\theta(\theta + \Delta\theta, z + \Delta z) = 0$. Тогда из (8) с точностью до слагаемых второго порядка малости получаем уравнение

$$\Delta\theta = -\left(J''_{\theta\theta}\right)^{-1} J''_{\theta z}\Delta z \doteq -\left(J''_{\theta\theta}\right)^{-1} \Delta_z J'^{\mathrm{T}}_\theta. \tag{9}$$

Имеет место следующее утверждение [11].

Теорема. *Пусть Δz в (9) — отклонение точного решения системы (1), вызванное малыми шумами ξ, w. Тогда в пределе $\sigma_\xi, \sigma_w \to 0$ имеют место следующие соотношения для отклонений оценок:*

$$\mathbf{D}\,\Delta\theta_{ВМ} \;\to\; \sigma^2_\xi \left(W_1^{\mathrm{T}} C W_1\right)^{-1} + \sigma^2_w \left(W_1^{\mathrm{T}} C W_1\right)^{-1} W_1^{\mathrm{T}} C^2 W_1 \left(W_1^{\mathrm{T}} C W_1\right)^{-1}, \tag{10}$$

$$\mathbf{D}\,\Delta\theta_{МНК} \;\to\; \sigma^2_\xi \left(W_1^{\mathrm{T}} W_1\right)^{-1} W_1^{\mathrm{T}} G G^{\mathrm{T}} W_1 \left(W_1^{\mathrm{T}} W_1\right)^{-1} + \sigma^2_w \left(W_1^{\mathrm{T}} W_1\right)^{-1}, \tag{11}$$

$$\mathbf{D}\,\Delta\theta_{МИП} \;\to\; \sigma^2_\xi \Psi^{\mathrm{T}}\Psi + \sigma^2_w \Psi^{\mathrm{T}} H H^{\mathrm{T}} \Psi, \tag{12}$$

$$\Psi \doteq \mathrm{Re}\left(\widetilde{G}^* \widetilde{Y} \widetilde{Y}^* \widetilde{W}_1\right) \left[\mathrm{Re}\left(\widetilde{W}_1^* \widetilde{Y} \widetilde{Y}^* \widetilde{W}_1\right)\right]^{-1},$$

$$\widetilde{G} \doteq \left(I_M \otimes \begin{bmatrix}\gamma_0 & \gamma_1\end{bmatrix}\right)\left(\bar{\Phi}\bar{F} \otimes I_{2(n+m)}\right) E, \quad E \doteq \frac{\partial \operatorname{vect} \bar{V}}{\partial z},$$

$$H \doteq \frac{\partial z}{\partial w} = \begin{bmatrix}0 & G_1^{-\mathrm{T}} & 0\end{bmatrix}^{\mathrm{T}}, \quad G_1 \doteq \begin{bmatrix} I_n & & 0 \\ -A_\theta & I_n & \\ 0 & \ddots & \ddots \end{bmatrix}.$$

Интерес представляют величины диагональных элементов матриц дисперсий (10)-(12) при различных соотношениях σ_w и σ_ξ. Были проведены расчеты для системы (3), (4) второго порядка ($n = 2$, $m = 1$, $v = 6$, $N = 250$) при значениях $\sigma_\xi \in [0, 0.1\sigma_z] = [0, 0.4]$, $\sigma_z \doteq \frac{\|z\|}{\sqrt{N}} \simeq 4.0$, с вычислением $\sigma_w = 0.15\,(0.4 - \sigma_\xi)$. Из формул (10)-(12) следует, что оценки МНК имеют преимущество при возмущениях в невязке ($\sigma_w \gg \sigma_\xi$), а оценки ВМ становятся предпочтительными при увеличении аддитивных возмущений в наблюдениях ($\sigma_w \ll \sigma_\xi$), что находится в согласии с известными свойствами оценок МНК и ВМ: первые статистически оптимальны при возмущениях в невязке уравнения, а вторые — при возмущениях в наблюдениях переменных.

Для численной проверки утверждений теоремы использовали $L = 50$ независимых реализаций нормально распределенных шумов $w\,[k]$ и $\xi\,[k]$ с нулевыми математическим ожиданием и матрицами вторых моментов $\mathbf{M}\,\xi[k]\xi[k]^{\mathrm{T}} = \sigma^2_\xi I_3$, $\mathbf{M}\,w[k]w[k]^{\mathrm{T}} = \sigma^2_w I_2$, $\sigma_\xi \in [0, 0.4]$, $\sigma_w = 0.15\,(0.4 - \sigma_\xi)$. Для относительных величин $\overline{\sigma}_{w,\xi} \doteq \sigma_{w,\xi}/\sigma_z$ при $\sigma_z \simeq 4.0$ диапазоны изменений были соответственно равны $\overline{\sigma}_\xi \in [0, 10\%]$, $\overline{\sigma}_w \in [0, 1.5\%]$. Начальные условия выбирались нулевыми или случайными с распределением $\mathrm{N}(0, 100 I_2)$. Результаты идентификации по данным с нулевыми или случайными н. у. качественно не различаются.

По результатам экспериментов можно сделать вывод, что имеет место хорошее соответствие между эмпирическими значениями дисперсий оценок параметров и теоретическими значениями, полученными по линейной модели (9), при относительном уровне случайных возмущений до 5–10%. При этом эмпирические смещения оценок МНК значительны, что принципиально не может быть предсказано линейной моделью.

4. Заключение

Теоретическое сравнение методов идентификации путем линеаризации градиентов целевых функций имеет ряд ограничений, как принципиальных (невозможность изучить смещение оценок), так и связанных с недостаточным развитием теории, в частности, с чрезвычайно сильными предположениями о малости возмущений. Тем не менее, линеаризованные модели зависимости отклонений оценок от возмущений демонстрируют хорошее совпадение с результатами вычислительного эксперимента при практически значимых уровнях возмущений. Это дает основания использовать локальный подход в качестве инструмента исследования свойств оценок на конечных выборках. Качественные характеристики разброса оценок трех методов идентификации в условиях смешанных возмущений уверенно предсказываются линейной теорией.

Работа выполнена при поддержке Российского фонда фундаментальных исследований (19-01-00754).

Список литературы

1. Lemmerling P., Moor De B. Misfit Versus Latency // Automatica. 2001. Vol. 37. P. 2057-2067.
2. Измаилов А.Ф. Чувствительность в оптимизации. М.: Физматлит, 2006.
3. Ломов А.А. Сравнение методов оценивания параметров линейных динамических систем по измерениям коротких участков переходных процессов // Автоматика и телемеханика. 2005. № 3. С. 39-48.
4. Kruglov I., Mishulina O., Bakirov M. Quantile based decision making rule of the neural networks committee for ill-posed approximation problems // Neurocomputing. 2012. Vol. 96. P. 74-82.
5. Klein V., Morelli E.A. Aircraft System Identification. Theory and Practice / AIAA Education Series. Reston: AIAA, 2006.
6. Егоршин А.О. Метод наименьших квадратов и «быстрые» алгоритмы в вариационных задачах идентификации и фильтрации (метод ВИ) // Автометрия. 1988. № 1. С. 30-42.
7. Osborne M. R., Smyth G. K. A modified Prony algorithm for fitting functions defined by difference equations // SIAM J. Sci. Statist. Comput. 1991. Vol. 12. P. 362-382.
8. Moor De B. Structured total least squares and L_2 approximation problems // Linear Algebra Appl. 1993. Vol. 188-189. P. 163-207.
9. Roorda B., Heij C. Global total least squares modelling of multivariable time series // IEEE Trans. on Aut. Con. 1995. Vol. AC-40. P. 50-63.
10. Ломов А.А. О количественных априорных показателях идентифицируемости коэффициентов линейных динамических систем // Изв. РАН ТСУ. 2011. № 1. С. 3-15.
11. Ломов А.А., Федосеев А.В. Сравнение методов параметрической идентификации линейных динамических систем в условиях смешанных возмущений // Сибирский журнал чистой и прикладной математики. 2018. Т. 18. № 3. С. 45-59.