

УДК 681.5.015

ББК 32.965.4

ВАРИАЦИОННЫЕ МЕТОДЫ ИДЕНТИФИКАЦИИ ЛИНЕЙНЫХ ДИНАМИЧЕСКИХ СИСТЕМ И ПРОБЛЕМА ЛОКАЛЬНЫХ ЭКСТРЕМУМОВ¹

Ломов А. А.²

(ФГБУН Институт математики им. С. Л. Соболева Сибирского
отделения РАН, Новосибирский государственный университет,
Новосибирск)

Рассмотрена проблема большого числа локальных экстремумов, возникающих при использовании «прямых» методов идентификации параметров линейных динамических систем по конечным выборкам наблюдений. Определен новый класс вариационных («непрямых») оценок параметров через свойство проективности слагаемых ядра целевой функции. Построены вариационные целевые функции с числом локальных экстремумов не больше числа элементов в матрицах системы. Получены условия состоятельности вариационных оценок в пределе большого числа наблюдений независимых траекторий конечной длины.

Ключевые слова: идентификация параметров, разностные уравнения, динамические системы.

1. Введение

Вычислительная сложность метода идентификации в значительной степени определяется экстремальными свойствами целевой функции. Известно, что при идентификации по конечным выборкам «прямыми» методами число экстремумов может быть

¹ Работа выполнена при финансовой поддержке РФФИ (грант № 10-01-00035) и при поддержке Сибирского отделения Российской академии наук (междисциплинарный проект № 80).

² Андрей Александрович Ломов, кандидат физико-математических наук, доцент (lomov@math.nsc.ru).

значительным и даже расти с ростом объема выборки наблюдений [26]. Поэтому заслуживают внимания методы идентификации, целевые функции у которых обладают «хорошими» экстремальными свойствами, возможно, за счет увеличения объема вычислений. В статье определен новый класс вариационных («непрямых») методов, интересный тем, что для ряда типичных представителей этого класса удается доказать конечность числа критических точек целевой функции независимо от объема выборки наблюдений.

Опишем структуру статьи. Во введении описываются исследуемые системы, даются основные понятия и обозначения. Затем формулируется проблема большого числа локальных экстремумов при идентификации «прямыми» методами по выборкам конечного объема. Решение проблемы может быть найдено при переходе к вариационной постановке задачи идентификации. В разделе 2 определяется новый класс «непрямых» (вариационных) оценок параметров через свойство проективности слагаемых ядра целевой функции. Показано, что в этот класс входят как известные в литературе оценки (ОР [42, 48], ОРМ [31], ВИ [9, 10], TLS [39], GTLS [49], STLS [45]), так и ряд новых предложенных в статье оценок — в частности, многомерные ОРМ и ОРС. Далее изучаются свойства вариационных оценок (раздел 3). Установлено, что многие задачи идентификации состоятельно решаются вариационными методами с целевыми функциями, число экстремумов которых не превосходит числа элементов в матричных коэффициентах уравнения системы независимо от объема L , N выборки измерений. В этом смысле вариационные постановки задач идентификации обладают «регуляризующим», «сглаживающим» эффектом, и поэтому могут рассматриваться как средство для решения проблемы большого числа локальных экстремумов.

1.1. ИССЛЕДУЕМЫЕ СИСТЕМЫ, ОЦЕНКИ ПАРАМЕТРОВ, ОБОЗНАЧЕНИЯ

Будем рассматривать линейные динамические системы

$$(1) \quad \alpha_p y[k+p] + \dots + \alpha_0 y[k] = \beta_p u[k+p] + \dots + \beta_0 u[k], \\ k \in \overline{1, N-p}, \quad N \geq p+1.$$

Здесь $\alpha_i = \alpha_{i,\theta} \in \mathbb{R}^{r \times r}$, $\beta_i = \beta_{i,\theta} \in \mathbb{R}^{r \times m}$ — матричные коэффициенты, которые зависят от параметра $\theta \in \Theta$, множество $\Theta \subseteq \mathbb{R}^v$ открытое. Параметр θ фиксирован и подлежит идентификации по множеству наблюдений

$$\{\check{y}[1], \check{u}[1], \dots, \check{y}[N], \check{u}[N]\},$$

$$\check{y}[k] = y[k] + \varepsilon_y[k], \quad \check{u}[k] = u[k] + \varepsilon_u[k],$$

где $y[k] \in \mathbb{R}^r$, $u[k] \in \mathbb{R}^m$ — переменные выхода, входа, $\varepsilon_y[k] \in \mathbb{R}^r$, $\varepsilon_u[k] \in \mathbb{R}^m$ — измерительные возмущения. Существенно, что старший матричный коэффициент α_p может быть вырожденным; тогда уравнение (1) описывает широкий класс многомерных (MIMO) систем с обратными связями и запаздываниями.

Определим вектор *процесса*

$$z \doteq (z[1]; \dots; z[N]) \in \mathbb{R}^{N(r+m)}, \quad z[k] \doteq \begin{pmatrix} y[k] \\ u[k] \end{pmatrix} \in \mathbb{R}^{r+m},$$

и вектор возмущений

$$\varepsilon \doteq (\varepsilon[1]; \dots; \varepsilon[N]) \in \mathbb{R}^{N(r+m)}, \quad \varepsilon[k] \doteq \begin{pmatrix} \varepsilon_y[k] \\ \varepsilon_u[k] \end{pmatrix} \in \mathbb{R}^{r+m}.$$

Здесь и далее, следуя [34], запятыми $(*, \dots, *)$ обозначаем вектор-строку, а точками с запятой $(*; \dots; *)$ — вектор-столбец. Аналогичные обозначения используются для клеточных матриц: $(A, B) \doteq (A \ B)$, $(A; B) \doteq \begin{pmatrix} A \\ B \end{pmatrix}$.

Пусть $z_{(i)} \in \mathbb{R}^{N(r+m)}$, $i \in \overline{1, L}$, — некоторый «полный» с точки зрения идентификации набор процессов в системе, $\varepsilon_{(i)}$

— случайные возмущения, ограниченные «естественными» предположениями о распределении (точные определения даны ниже). Нас будут интересовать следующие свойства оценки как функции наблюдений:

1. Единственность (по смыслу это свойство лучше назвать точность, или состоятельность в детерминированном смысле):

$$\hat{\theta}(z_{(1)}, \dots, z_{(L)}) = \theta.$$

2. Устойчивость:

$$\hat{\theta}(z_{(1)} + \varepsilon_{(1)}, \dots, z_{(L)} + \varepsilon_{(L)}) \xrightarrow[\substack{\|\varepsilon_{(i)}\| \rightarrow 0, \\ i \in \overline{1, L}}]{\rightarrow} \theta.$$

3. Состоятельность (в статистическом смысле):

$$\hat{\theta}(z_{(1)} + \varepsilon_{(1)}, \dots, z_{(L)} + \varepsilon_{(L)}) \xrightarrow{L \rightarrow \infty} \theta.$$

Если число процессов L ограничено, то состоятельность рассматривается в предельном случае больших длин процессов N :

$$\hat{\theta}(z + \varepsilon) \xrightarrow{N \rightarrow \infty} \theta, \quad z + \varepsilon = \begin{pmatrix} z^{[1]} + \varepsilon^{[1]} \\ \vdots \\ z^{[N]} + \varepsilon^{[N]} \end{pmatrix}.$$

Последнее предполагает, что система не изменяет своих параметров в течение длительного отрезка времени Nh , где h — время дискретизации.

В общем случае оценки $\hat{\theta}(\check{z}_{(1)}, \dots, \check{z}_{(L)})$ вычисляются не по явным формулам, а минимизацией той или иной целевой функции, зависящей от наблюдений $\check{z}_{(i)}$. Под методом идентификации будем понимать целевую функцию вместе с вычислительным алгоритмом минимизации.

1.2. ПРЯМЫЕ МЕТОДЫ ИДЕНТИФИКАЦИИ И ПРОБЛЕМА ЛОКАЛЬНЫХ ЭКСТРЕМУМОВ

С вычислительной точки зрения важны экстремальные свойства целевой функции — в частности, число критических точек

в зависимости от множества наблюдений. Этим в значительной степени определяется устойчивость алгоритмов идентификации.

В статье рассмотрим экстремальные свойства двух больших групп методов идентификации во временной области:

- «прямые» [28, раздел 5.3] (или «наивные» [38], «разомкнутые» [24]);
- «непрямые» (или «замкнутые» [24], «вариационные» [10]).

По-видимому, первым, кто попытался выделить эти две группы, был К.Б. Норкин [24], затем А.О. Егоршин [9] и Л. Глезер [38]. Данные ими определения были скорее интуитивными, мотивированными желанием найти «наиболее правильные»³ постановки задач идентификации. Методы первой группы («прямые») в вычислительном отношении более просты, но с ними связан ряд проблем, описанных ниже. Этим вызвано внимание к более сложным «непрямым» методам.

В статье далее делается попытка уточнить смысл понятия «правильной постановки» задачи идентификации с точки зрения экстремальных свойств целевых функций. В рамках этой программы предлагается (по-видимому, впервые) точное определение для второй группы («непрямых») методов, и на основании исследования экстремальных свойств устанавливаются частные случаи, в которых «непрямые» методы имеют преимущества по сравнению с методами первой группы.

Кратко опишем особенности «прямых» методов. Для этой группы, насколько известно, точного определения пока не дано. Имеется следующее описание, данное А. А. Красовским [28, раздел 5.3]: «Прямыми методами параметрической идентификации» называются те, в которых «неизвестные параметры ... модели определяются на основе того или иного способа (обычно приближенного) решения системы уравнений, получающихся путем [прямой] подстановки (выделено мной — А. Л.) в оператор модели последовательности значений входных и выходных величин

³ Приблизительная формулировка.

реального объекта». Согласно этому нестрогому определению, к «прямым» относятся методы, основанные на минимизации нормы «ошибки прогноза» (РЕМ, «prediction error methods») и их многочисленные модификации, включающие в себя рекуррентные методы наименьших квадратов по невязке уравнения и их варианты типа алгоритмов стохастической аппроксимации и алгоритмов Качмажа [11, 23, 25, 29, 28, 35, 52], а также методы типа инструментальных переменных [51, глава 8]. Прямые методы идентификации повсеместно употребляются при решении задач адаптивного управления [1, 32]. Для линейных моделей со структурой общего вида

$$(2) \quad A(s)\check{y}[k] = \frac{B(s)}{F(s)}\check{u}[k] + \frac{C(s)}{D(s)}\varepsilon[k]$$

(где s — оператор сдвига, и \check{y} , \check{u} — наблюдаемые переменные) целевые функции прямых методов строятся исходя из нормы оценки невязки уравнения $\frac{C(s)}{D(s)}\varepsilon[k]$ или ошибки прогноза $\frac{C(s)}{A(s)D(s)}\varepsilon[k]$ [23].

Заметим, что идея прямой подстановки измерений в уравнение системы восходит к ставшим уже классическими работам А.Н. Колмогорова [12] и Н. Винера [54]. В этой связи уместно назвать и работу К. Гаусса по методу наименьших квадратов [5].

В применении к идентификации динамических систем прямые методы при сравнительной простоте вычислительных алгоритмов имеют ограничения: 1) система должна быть управляемой; 2) длина N интервала наблюдения должна быть большой в сравнении с длительностью переходных процессов в системе [35].

Кроме того, И. И. Перельманом было показано [26], что для прямых методов при конечных N существует проблема большого числа локальных экстремумов целевой функции. Наличие локальных экстремумов отмечалось также в [33, 50]. Приведем простейший пример [26]:

$$y[k] = ay[k - 1] + u[k], \quad \check{y}[k] = y[k] + \varepsilon[k], \quad \check{u}[k] = u[k].$$

Уравнение для ошибки прогноза имеет вид

$$\begin{aligned}\varepsilon[k] &= \check{y}[k] - \check{u}[k] - a\check{y}[k-1] + a\varepsilon[k-1] \doteq \\ &\doteq \check{\omega}[k] + a\varepsilon[k-1],\end{aligned}$$

где $\check{\omega}[k]$ — измерения. Принимая в качестве целевой функции квадрат нормы ошибки прогноза на конце траектории, получаем

$$\varepsilon^2[N] = (a^{N-1}\varepsilon[1] + a^{N-2}\check{\omega}[1] + \dots + \check{\omega}[N])^2 \doteq J(a).$$

Локальные экстремумы определяются из уравнения $\frac{\partial J}{\partial a} = 0$, которое может иметь до $(2N - 3)$ действительных корней. Очевидно, число корней растет с ростом N , и все они должны рассматриваться как претенденты на точки глобального минимума. Этот факт, как пишет И. И. Перельман, «приводит к обоснованным сомнениям относительно практической разрешимости» поставленной простейшей задачи идентификации. Трудности только усугубляются, если рассмотреть системы с контурами обратной связи, т. е. системы с несколькими уравнениями.

В пределе $N \rightarrow \infty$ вероятность большого числа экстремальных точек стремится к нулю [3]. Однако этот факт не снимает отмеченных трудностей идентификации «прямыми» методами по конечному множеству наблюдений.

2. Вариационные (непрямые) оценки

Решение проблемы локальных экстремумов может быть найдено при переходе к вариационной постановке задачи идентификации. В одном частном случае это было показано М. Левиным [42] и М. Аоки, П. Ю [31], которые рассматривали системы (1) из одного уравнения ($r = 1$) с простой параметризацией. Для идентификации параметров они применяли метод ортогональной регрессии [48], целевая функция которого имеет вид отношения квадратичных форм (22), и как следствие, число экстремумов ограничивается сверху размерностью вектора параметров (см. раздел 2.1).

Термины «вариационный метод», «вариационная постановка задачи идентификации» были предложены А.О. Егоршиным [10], который сконструировал один из наиболее эффективных с точки зрения приложений «замкнутый» метод идентификации [9]. Поясним примером. Пусть дано наблюдение $\check{z} \in \mathbb{R}^{N(r+m)}$. Оценка вектора параметров системы (1) может быть получена минимизацией по θ функции

$$(3) \quad J(\theta) = J(\theta, \check{z}) = \|\check{z} - z_{\text{opt}}(\theta)\|^2,$$

где $z_{\text{opt}}(\theta)$ — проекция вектора \check{z} на множество всех процессов системы (1). Величина $J(\theta)$ играет роль квадрата расстояния от «точки» \check{z} до множества процессов (1). Функция $\hat{\theta}(\check{z}) = \arg \min_{\theta} J(\theta, \check{z})$ является примером вариационной (непрямой) оценки. Эффективные алгоритмы ее вычисления были предложены в [9, 10].

В этом разделе определим новый класс «вариационных»⁴ оценок. Он будет включать в себя большинство известных в литературе непрямых оценок: оценки ортогональной регрессии [48], модифицированные оценки М. Аоки, П. Ю [31], оценки по методу вариационной идентификации А.О. Егоршина [9, 10], оценки по модифицированному методу Прони [37, 47], оценки STLS [45] и ряд других, в том числе новых оценок. Общим свойством для всего класса будет *независимость следа матрицы ядра целевой функции от оцениваемого параметра*. В частности, это свойство позволяет доказывать состоятельность в пределе $L \rightarrow \infty$ по одной и той же схеме, предложенной в [15]. Одним из полезных побочных результатов будет также описание различных представителей класса вариационных методов в единых обозначениях (см. разделы 2.1–2.5).

Перейдем к определениям. Введем матричную запись урав-

⁴ *Название не является устоявшимся.*

нения (1):

$$(4) \quad G_\theta z = 0,$$

$$G_\theta = \begin{pmatrix} \gamma_0 & \gamma_1 & \dots & \gamma_p & & 0 \\ & \gamma_0 & \gamma_1 & \dots & \gamma_p & \\ & & \ddots & \ddots & & \ddots \\ 0 & & & \gamma_0 & \gamma_1 & \dots & \gamma_p \end{pmatrix}$$

$$\in \mathbb{R}^{(N-p)r \times N(r+m)},$$

$$\gamma_i \doteq \gamma_{i,\theta} = (\alpha_{i,\theta}, -\beta_{i,\theta}) \in \mathbb{R}^{r \times (r+m)}.$$

Тогда $z_{\text{opt}}(\theta) = (I - \Pi_\theta) \tilde{z}$, где $\Pi_\theta \doteq G_\theta^\top (G_\theta G_\theta^\top)^{-1} G_\theta$ есть матрица проектора ($\Pi_\theta^2 = \Pi_\theta$) на ортогональное дополнение к линейному многообразию решений системы (1). С учетом этого обозначения целевая функция (3) принимает вид

$$(5) \quad J(\theta) = \tilde{z}^\top \Pi_\theta \tilde{z}.$$

Матрицу Π_θ будем называть *ядром* целевой функции $J(\theta)$.

Пусть θ_* — некоторое неизвестное «истинное» значение параметра системы (1).

Определение 1. Вариационной оценкой $\hat{\theta}$ параметра θ_* по наблюдению $\tilde{z} = z_* + \varepsilon$, где z_* — истинный процесс, ε — возмущение, называем результат минимизации квадратичной целевой функции

$$(6) \quad J(\theta, \tilde{z}) \doteq \tilde{z}^\top (\Pi_{1,\theta} + \dots + \Pi_{M,\theta}) \tilde{z},$$

$$(7) \quad \hat{\theta} = \hat{\theta}(\tilde{z}) = \arg \min_{\theta} J(\theta, \tilde{z}),$$

где все матрицы $\Pi_j = \Pi_{j,\theta} \geq 0$ являются неотрицательно определенными, симметричными, проективными ($\Pi_j^2 = \Pi_j$) и такими, что верно соотношение $(G_{\theta_*} z = 0) \Rightarrow (\Pi_{j,\theta_*} z = 0)$. Целевую функцию $J(\theta, \tilde{z})$ (6) также называем вариационной.

Можно сказать, что целевые функции, ядра которых есть сумма проекторов, являются вариационными (или *проективными*).

Приведем еще ряд определений, расширяющих понятие вариационной оценки. Пусть возмущения $\varepsilon_{(1)}, \dots, \varepsilon_{(L)} \in \mathbb{R}^{N(r+m)}$

являются независимыми одинаково распределенными случайными векторами с нулевыми математическими ожиданиями и скалярной матрицей вторых моментов:

$$(8) \quad \mathbf{M} \varepsilon_{(i)} = 0, \quad \mathbf{M} \varepsilon_{(i)} \varepsilon_{(i)}^\top = \sigma^2 I.$$

Тогда имеет смысл следующее определение.

Определение 2. Вариационной оценкой $\hat{\theta}_L$ параметра θ_* по множеству наблюдений $\{\check{z}_{(i)} = z_{*(i)} + \varepsilon_{(i)}, i \in \overline{1, L}\}$, где $z_{*(i)}$ — истинные процессы, $\varepsilon_{(i)}$ — возмущения (8), называем результат минимизации квадратичной целевой функции

$$(9) \quad J_L(\theta, \check{z}_{(1)}, \dots, \check{z}_{(L)}) \doteq L^{-1} \sum_{i=1}^L \check{z}_{(i)}^\top (\Pi_{1,\theta} + \dots + \Pi_{M,\theta}) \check{z}_{(i)},$$

$$(10) \quad \hat{\theta}_L = \hat{\theta}(\check{z}_{(1)}, \dots, \check{z}_{(L)}) = \arg \min_{\theta} J_L(\theta, \check{z}_{(1)}, \dots, \check{z}_{(L)}),$$

где все матрицы $\Pi_j = \Pi_{j,\theta} \geq 0$ являются неотрицательно определенными, симметричными, проективными ($\Pi_j^2 = \Pi_j$) и такими, что верно соотношение $(G_{\theta_*} z = 0) \Rightarrow (\Pi_{j,\theta_*} z = 0)$.

Пусть теперь возмущения $\varepsilon_{(1)}, \dots, \varepsilon_{(L)} \in \mathbb{R}^{N(r+m)}$ являются независимыми одинаково распределенными случайными векторами со свойствами

$$(11) \quad \mathbf{M} \varepsilon_{(i)} = 0, \quad \mathbf{M} \varepsilon_{(i)} \varepsilon_{(i)}^\top = \Sigma \Sigma^\top,$$

где $\Sigma \in \mathbb{R}^{N(r+m) \times n}$ — матрица с линейно независимыми столбцами.

Определение 3. Вариационной оценкой $\hat{\theta}_L$ параметра θ_* по множеству наблюдений $\{\check{z}_{(i)} = z_{*(i)} + \varepsilon_{(i)}, i \in \overline{1, L}\}$, где $z_{*(i)}$ — истинные процессы, $\varepsilon_{(i)}$ — возмущения (11), называем результат минимизации целевой функции

$$(12) \quad J_{\Sigma,L}(\theta, \check{z}_{(1)}, \dots, \check{z}_{(L)}) \doteq \\ \doteq L^{-1} \sum_{i=1}^L \check{z}_{(i)}^\top (W_{1,\theta} + \dots + W_{M,\theta}) \check{z}_{(i)},$$

$$(13) \quad \hat{\theta}_L = \hat{\theta}(\check{z}_{(1)}, \dots, \check{z}_{(L)}) = \arg \min_{\theta} J_{\Sigma,L}(\theta, \check{z}_{(1)}, \dots, \check{z}_{(L)}),$$

где все матрицы $W_{j,\theta} \geq 0$ неотрицательно определены, симметричны и таковы, что верно соотношение $(G_{\theta_*} z = 0) \Rightarrow (W_{j,\theta_*} z = 0)$, а произведения $\Sigma^\top W_j \Sigma = \Pi_j$ являются проективными матрицами со свойством $\Pi_j^2 = \Pi_j$.

В следующем заключительном (наиболее общем) определении уже не будем опираться на проективность матриц, а возьмем за основу более общее свойство независимости следа матрицы от оцениваемого параметра. Полученные таким образом оценки уже не обязательно имеют ядро целевой функции в виде суммы проекторов, но доказательство их состоятельности осуществляется по той же схеме, что и для оценок из определений 2–3 (для оценок с целевой функцией (9) см. в [15]).

Определение 4. *Обобщенной вариационной оценкой $\hat{\theta}_L$ параметра θ_* по множеству наблюдений $\{\check{z}_{(i)} = z_{*(i)} + \varepsilon_{(i)}, i \in \overline{1, L}\}$, где $z_{*(i)}$ — истинные процессы, $\varepsilon_{(i)}$ — возмущения (11), называем результат минимизации целевой функции*

$$J_{\Sigma, L}(\theta, \check{z}_{(1)}, \dots, \check{z}_{(L)}) \doteq L^{-1} \sum_{i=1}^L \check{z}_{(i)}^{\top} U_{\theta} \check{z}_{(i)},$$

$$\hat{\theta}_L = \hat{\theta}(\check{z}_{(1)}, \dots, \check{z}_{(L)}) = \arg \min_{\theta} J_{\Sigma, L}(\theta, \check{z}_{(1)}, \dots, \check{z}_{(L)}),$$

где матрица $U_{\theta} \geq 0$ неотрицательно определена, симметрична и такова, что верно соотношение $(G_{\theta_*} z = 0) \Rightarrow (U_{\theta_*} z = 0)$, а след произведения $\Sigma^{\top} U_{\theta} \Sigma$ не зависит от θ :

$$\text{Sp } \Sigma^{\top} U_{\theta} \Sigma \equiv \text{const}(\theta).$$

В итоге, определяющим свойством для введенного в этом разделе класса вариационных оценок является независимость следа матрицы ядра целевой функции от оцениваемого параметра. Заметим, что все известные нам примеры вариационных оценок ограничены оценками из определений 2–3, т. е. с проективными ядрами.

Определения 2 и 3 по существу не вносят новых идей в понятие вариационной оценки по сравнению с определением 1. Чтобы продемонстрировать конструктивность данных определений, покажем, что ряд известных в литературе оценок являются частными случаями вариационных из определения 1 ($L = 1$), а также построим новые оценки.

Введем обозначения. Из матриц $\gamma_i \in \mathbb{R}^{r \times (r+m)}$ (4) составим матрицу

$$(14) \quad \gamma_\theta \doteq (\gamma_0 \ \gamma_1 \ \dots \ \gamma_p) \in \mathbb{R}^{r \times (r+m)(p+1)}.$$

Пусть $\text{vect } \gamma$ – вектор, полученный последовательным выстраиванием транспонированных строк матрицы γ . Будем предполагать, что матрица γ_θ (14) зависит от вектора параметров θ аффинным образом:

$$(15) \quad \text{vect } \gamma_\theta = d + D\theta \doteq D\vartheta,$$

где матрица $D \in \mathbb{R}^{r(r+m)(p+1) \times v}$ и вектор $d \in \mathbb{R}^{r(r+m)(p+1)}$ заданы, столбцы составной матрицы $D \doteq (d, D)$ линейно независимы, $\vartheta \doteq \begin{pmatrix} 1 \\ \theta \end{pmatrix}$.

Система уравнений (4) всегда может быть записана в виде

$$(16) \quad G_\theta z \equiv V \text{vect } \gamma_\theta \equiv VD\vartheta = 0,$$

где знак « \equiv » обозначает равенство для всех $z \in \mathbb{R}^{N(r+m)}$, $V = V(z)$ – клеточно-ганкелевая матрица из элементов вектора z :

$$(17) \quad V(z) \doteq \begin{pmatrix} I_r \otimes v_{(1)}^\top \\ \vdots \\ I_r \otimes v_{(N-p)}^\top \end{pmatrix}, \quad v_{(i)}^\top \doteq (z[i]^\top \dots z[i+p]^\top).$$

Символ \otimes обозначает кронекерово произведение матриц, I_r – единичная матрица $r \times r$.

Дополнительно предположим:

(А) для всех $\theta \in \Theta$ каноническая форма многочленной матрицы

$$\gamma_\theta(s) \doteq \gamma_{p,\theta} s^p + \gamma_{p-1,\theta} s^{p-1} + \dots + \gamma_{0,\theta}$$

не имеет тождественно нулевых строк; это условие при $N \geq 2(p+1)$, $p \geq 0$ равносильно линейной независимости строк числовой матрицы G_θ (4) [14, предложение 4.1];

(Б) степень определителя многочленной матрицы знаменателя $\alpha_\theta(s)$ не зависит от параметра θ :

$$\deg \det \alpha_\theta(s) \equiv \text{const},$$

$$\alpha_\theta(s) \doteq \alpha_{p,\theta} s^p + \alpha_{p-1,\theta} s^{p-1} + \dots + \alpha_{0,\theta};$$

(В) степень каждой строки многочленной матрицы знаменателя $\alpha_\theta(s)$ не меньше степени соответствующей строки многочленной матрицы числителя

$$\beta_\theta(s) \doteq \beta_{p,\theta} s^p + \beta_{p-1,\theta} s^{p-1} + \dots + \beta_{0,\theta}$$

(условие причинности); степенью строки многочленной матрицы называется наибольшая степень среди образующих эту строку многочленов.

Заметим, что на систему (4) не налагаются условия управляемости и устойчивости. Последнее возможно благодаря конечности N .

Ввиду равенства $G_\theta z = 0$ условие $\Pi_{j,\theta} z = 0$ из определения 1 позволяет выразить проекторы $\Pi_{j,\theta}$ через линейные комбинации строк матрицы G_θ :

$$(18) \quad \begin{aligned} \Pi_{j,\theta} &= \Phi_{j,\theta}^\top \left(\Phi_{j,\theta} \Phi_{j,\theta}^\top \right)^{-1} \Phi_{j,\theta} \doteq \Phi_{j,\theta}^\top C_{j,\theta} \Phi_{j,\theta}, \\ \Phi_{j,\theta} &= P_j G_\theta, \quad P_j \in \mathbb{R}^{n_j \times (N-p)r}, \end{aligned}$$

где строки матриц P_j (составленные из коэффициентов линейных комбинаций) линейно независимы. Последнее вместе с условием (А) гарантирует строгую положительную определенность матриц $\Phi_{j,\theta} \Phi_{j,\theta}^\top$ и $C_{j,\theta} \doteq \left(\Phi_{j,\theta} \Phi_{j,\theta}^\top \right)^{-1}$.

Проективная целевая функция (6) с учетом параметризации (15) и тождества (16) принимает вид

$$(19) \quad \begin{aligned} J(\theta, \tilde{z}) &\doteq \tilde{z}^\top (\Pi_{1,\theta} + \dots + \Pi_{M,\theta}) \tilde{z} = \\ &= \vartheta^\top D^\top V(\tilde{z})^\top \left(\sum_{j=1}^M P_j^\top C_{j,\theta} P_j \right) V(\tilde{z}) D \vartheta. \end{aligned}$$

Аналогичное выражение получается для целевой функции $J_{\Sigma,L}$ (12): по лемме 2 (из приложения)

$$W_{j,\theta} = G_\theta^\top P_j^\top \underbrace{\left(P_j G_\theta \Sigma \Sigma^\top G_\theta^\top P_j^\top \right)^{-1}}_{C_{\Sigma,j,\theta}} P_j G_\theta,$$

$$\begin{aligned}
 &= \begin{pmatrix} \pi_\theta & & & 0 \\ & \pi_\theta & & \\ & & \ddots & \\ 0 & & & \pi_\theta \end{pmatrix} = \begin{pmatrix} \frac{\gamma_\theta^\top \gamma_\theta}{\gamma_\theta \gamma_\theta^\top} & & & 0 \\ & \frac{\gamma_\theta^\top \gamma_\theta}{\gamma_\theta \gamma_\theta^\top} & & \\ & & \ddots & \\ 0 & & & \frac{\gamma_\theta^\top \gamma_\theta}{\gamma_\theta \gamma_\theta^\top} \end{pmatrix} = \\
 &= G_{\text{OR},\theta}^\top \left(G_{\text{OR},\theta} G_{\text{OR},\theta}^\top \right)^{-1} G_{\text{OR},\theta} \doteq \Pi_\theta.
 \end{aligned}$$

Несложно убедиться, что каждое слагаемое в сумме (21) является проектором ввиду равенств $\pi_\theta^2 = \left(\frac{\gamma_\theta^\top \gamma_\theta}{\gamma_\theta \gamma_\theta^\top} \right)^2 = \pi_\theta$. Также проектором является вся сумма (21): $\Pi_\theta^2 = \Pi_\theta$. Поэтому оценка М. Левина принадлежит классу вариационных.

В оценке М. Левина каждая матрица P_j из (19) является строкой из нулей и одной единицы на месте с номером $(j-1)(p+1)+1$, $j \in \overline{1, M}$, $M = \frac{N}{p+1}$:

$$P_j = \left(\underbrace{0 \dots 0}_{(j-1)(p+1)} \quad 1 \quad 0 \dots 0 \right).$$

$N-p$

Запись (19) принимает вид:

$$\begin{aligned}
 J(\theta, \check{z}) &= \frac{\vartheta^\top D^\top V(\check{z})^\top \left(\sum_{j=1}^M P_j^\top P_j \right) V(\check{z}) D \vartheta}{\gamma_\theta^\top \gamma_\theta} = \\
 (22) \quad &= \frac{\vartheta^\top D^\top V_{\text{OR}}(\check{z})^\top V_{\text{OR}}(\check{z}) D \vartheta}{\vartheta^\top D^\top D \vartheta},
 \end{aligned}$$

где матрица $V_{\text{OR}}(\check{z})$ получена прореживанием ганкелевой матрицы $V(\check{z})$ (17) — удаляются строки, содержащие элементы, имеющиеся в других строках.

Устойчивость оценок с целевой функцией вида (22) исследовали Г. Голуб и Ч. Ван Лоан [39], назвав их оценками TLS (Total Least Squares estimates). Тожественное совпадение оценок TLS Г. Голуба и Ч. Ван Лоана с классическими оценками ортогональной регрессии [48] было отмечено в [36].

2.2. МОДИФИЦИРОВАННЫЕ ОЦЕНКИ ОРТОГОНАЛЬНОЙ РЕГРЕССИИ (ОРМ)

М. Аоки и П. Ю предложили модифицированный вариант (ОРМ) оценки М. Левина, имеющий локально меньший разброс [31]. В наших обозначениях ядро целевой функции модифицированной оценки принимает вид суммы проекторов:

$$\begin{aligned}
 (23) \quad & \Pi_{1,\theta} + \dots + \Pi_{N-p,\theta} = \\
 & = \begin{pmatrix} \frac{\gamma_\theta^\top \gamma_\theta}{\gamma_\theta \gamma_\theta^\top} & & & 0 \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{pmatrix} + \begin{pmatrix} 0 & & & 0 \\ & \frac{\gamma_\theta^\top \gamma_\theta}{\gamma_\theta \gamma_\theta^\top} & & \\ & & \ddots & \\ & & & 0 \end{pmatrix} + \dots \\
 & \dots + \begin{pmatrix} 0 & & & 0 \\ & 0 & & \\ & & \ddots & \\ & & & \frac{\gamma_\theta^\top \gamma_\theta}{\gamma_\theta \gamma_\theta^\top} \end{pmatrix}.
 \end{aligned}$$

Поэтому модифицированная оценка принадлежит классу вариационных. Нули на диагонали есть числа, а клетки $\frac{\gamma_\theta^\top \gamma_\theta}{\gamma_\theta \gamma_\theta^\top}$ имеют размерность $(p+1) \times (p+1)$. Данная сумма уже не является проектором. Число слагаемых $N-p$ здесь больше, чем в методе ОР.

Каждая матрица P_j , $j \in \overline{1, N-p}$ в оценке ОРМ является j -й строкой единичной матрицы I_{N-p} :

$$(24) \quad P_j = \underbrace{(0 \dots 0 \ 1 \ 0 \dots 0)}_{N-p}.$$

Выражение (19) принимает вид

$$(25) \quad J(\theta, \check{z}) = \frac{\vartheta^\top D^\top V(\check{z})^\top V(\check{z}) D \vartheta}{\vartheta^\top D^\top D \vartheta}.$$

Введенные обозначения позволяют достаточно просто построить оценку ОРМ для случая нескольких уравнений $r > 1$.

Такие многомерные оценки в литературе ранее не встречались. Ядро целевой функции:

$$\begin{aligned} & \Pi_{1,\theta} + \dots + \Pi_{N-p,\theta} = \\ & = \begin{pmatrix} \pi_\theta & & 0 \\ & 0 & \\ & & \ddots \\ 0 & & & 0 \end{pmatrix} + \begin{pmatrix} 0 & & 0 \\ & \pi_\theta & \\ & & \ddots \\ 0 & & & 0 \end{pmatrix} + \dots \\ & \dots + \begin{pmatrix} 0 & & 0 \\ & 0 & \\ & & \ddots \\ 0 & & & \pi_\theta \end{pmatrix}, \end{aligned}$$

$$\pi_\theta \doteq \gamma_\theta^\top \left(\gamma_\theta \gamma_\theta^\top \right)^{-1} \gamma_\theta \in \mathbb{R}^{(r+m)(p+1) \times (r+m)(p+1)},$$

$$0 \in \mathbb{R}^{(r+m) \times (r+m)}, \quad \Pi_{j,\theta} \in \mathbb{R}^{(r+m)(N-p) \times (r+m)(N-p)}.$$

Матрицы P_j (24) клеточные:

$$(26) \quad P_j = \underbrace{\left(\underbrace{0 \dots 0}_{(j-1)} \ 1 \ 0 \dots 0 \right)}_{N-p} \otimes I_{r+m}.$$

Целевая функция (19) записывается следующим образом:

$$\begin{aligned} (27) \quad J(\theta, \tilde{z}) &= \vartheta^\top \mathbf{D}^\top \left[\sum_{j=1}^{N-p} V_j(\tilde{z})^\top \left(\gamma_\theta \gamma_\theta^\top \right)^{-1} V_j(\tilde{z}) \right] \mathbf{D} \vartheta = \\ &= \vartheta^\top \mathbf{D}^\top V(\tilde{z})^\top C_{\mathbf{M},\theta} V(\tilde{z}) \mathbf{D} \vartheta, \\ C_{\mathbf{M},\theta} &\doteq I_{N-p} \otimes \left(\gamma_\theta \gamma_\theta^\top \right)^{-1}. \end{aligned}$$

2.3. ОЦЕНКИ ПО МЕТОДУ ВАРИАЦИОННОЙ ИДЕНТИФИКАЦИИ (ВИ) И ПО МЕТОДУ GTLS

Более сложным в вычислительном отношении является метод вариационной идентификации А.О. Егоршина [9]; его многомерный вариант при $r > 1$ изложен в [10]. С точки зрения

разброса оценок при случайных возмущениях этот метод лучше ОР и ОРМ [19]. Целевая функция метода ВИ при $L = 1$ есть (9) с ядром

$$\Pi_{1,\theta} = \dots = \Pi_{M,\theta} \doteq \Pi_\theta = G_\theta^\top \left(G_\theta G_\theta^\top \right)^{-1} G_\theta.$$

Формально $M = 1$,

$$(28) \quad \begin{aligned} P_1 &= I, \\ C_{1,\theta} &\doteq C_\theta = \left(G_\theta G_\theta^\top \right)^{-1}. \end{aligned}$$

Тогда (19) принимает вид

$$(29) \quad J(\theta, \tilde{z}) = \vartheta^\top D^\top V(\tilde{z})^\top C_\theta V(\tilde{z}) D \vartheta.$$

Следует упомянуть также метод идентификации GTLS (Global Total Least Squares), предложенный позже Б. Роордой и К. Хейджем [49]. По целевой функции этот метод совпадает с методом ВИ, поэтому также принадлежит классу вариационных.

2.4. ОЦЕНКИ STLS

Оценки STLS (Structured TLS) были предложены Б. Де Муром [45] как модификация оценок ОР (TLS), учитывающая специальную структуру матрицы наблюдений $V(z)$ (17). В случае $L = 1$, $M = 1$ эта структура клеточно-ганкелевая. Оценки ОР, модифицированные ОР, оценки метода ВИ являются частными случаями STLS. Нас будет интересовать вопрос, совпадает ли класс оценок STLS с классом вариационных оценок из определения 3.

Задача STLS может быть сформулирована следующим образом [40]. Пусть матрица $V(z)$ аффинно зависит от элементов вектора $z \in \mathbb{R}^n$:

$$(30) \quad \begin{aligned} z &= (z_1; \dots; z_n), \\ V(z) &= V_0 + V_1 z_1 + \dots + V_n z_n, \end{aligned}$$

где матрицы V_0, \dots, V_n заранее известны и фиксированы. Пусть $\tilde{z} \in \mathbb{R}^n$ есть вектор наблюдений с матрицей ковариации RR^\top .

Найти минимум

$$(31) \quad \min_{z \in \mathbb{R}^n, \theta \in \mathbb{R}^v} \|R^{-1}(\tilde{z} - z)\|^2 \quad \text{при условии} \quad V(z)g(\theta) = 0,$$

где $g(\theta)$ — матрица, аффинно зависящая от $\theta \in \mathbb{R}^v$:

$$(32) \quad g(\theta) = g_0 + g_1\theta_1 + \dots + g_v\theta_v.$$

С начала 1990-х гг. задачи STLS активно исследуются за рубежом. Широкий круг приложений обсуждался в [45]. В [41] установлена тождественность между оценками STLS и оценками по методу CTLS (Constrained TLS) Т.Абатзоглу и Дж.Менделя [30]. Что касается свойств оценок STLS, то авторами [40] доказана состоятельность многомерных оценок STLS ($g(\theta) \in \mathbb{R}^{r \times v}$) в предельном случае $N \rightarrow \infty$ для клеточно-ганкелевых матриц $V(z)$. Ниже мы установим состоятельность вариационных оценок в случае $L \rightarrow \infty$, и доказав равносильность классов STLS и вариационных, распространим тем самым новое утверждение о состоятельности и на оценки STLS.

Как было отмечено, при $L = 1$, $M = 1$ вариационная оценка (7) является оценкой STLS. Покажем, что в общем случае это отношение сохраняется, более того, имеет место совпадение обоих классов оценок. Этот новый результат устанавливает связь между свойствами оценок вариационных и STLS, исследованных независимо разными авторами [7, 9, 15, 22, 31, 37, 40, 45, 46].

Теорема 1. *Оценка STLS (30)–(32) является вариационной оценкой (13). И наоборот, вариационная оценка (13) всегда может быть получена как решение некоторой задачи STLS (30)–(32).*

Доказательство. См. приложение.

Как следствие этой теоремы, оценки STLS допускают новые проективные формы записи (12), (20) для целевой функции (31), и все полученные ниже результаты о свойствах вариационных оценок — условия единственности, устойчивости, состоятельности при $L \rightarrow \infty$ и конечных N (теоремы 2–6 далее) — распространяются и на оценки STLS. Необходимые переобозначения приведены в приложении в доказательстве теоремы 1.

2.5. СКАЛЯРНЫЕ ОЦЕНКИ ОРТОГОНАЛЬНОЙ РЕГРЕССИИ ДЛЯ СИСТЕМ ИЗ НЕСКОЛЬКИХ УРАВНЕНИЙ (ОРС)

Введем еще один тип вариационных оценок параметров системы (1) из нескольких уравнений ($r > 1$), ранее неизвестный в литературе. По свойствам целевой функции эти оценки наиболее близки к оценкам наиболее простого скалярного ($r = 1$) метода ОРМ. Назовем эти оценки ОРС, обозначим θ_S . Идея состоит в том, чтобы по одному измерению \tilde{z} идентифицировать параметры уравнения с матрицей γ_θ (14) из r строк так, как бы идентифицировалась каждая из этих строк по отдельности скалярным методом ОРМ из раздела 2.2. Существенно, что в оценках ОРС учитывается зависимость коэффициентов разных уравнений от общего параметра θ .

Введем обозначения. В случае $r > 1$ матрица $V(z)$ (17) перестановкой строк приводится к виду

$$V(z) \sim \begin{pmatrix} \bar{V} & & 0 \\ & \ddots & \\ 0 & & \bar{V} \end{pmatrix} = I_r \otimes \bar{V},$$

$$(33) \quad \text{где } \bar{V} \doteq \begin{pmatrix} v_{(1)}^\top \\ \vdots \\ v_{(N-p)}^\top \end{pmatrix} \in \mathbb{R}^{(N-p) \times (p+1)(r+m)}$$

совпадает с матрицей $V(z)$ (17) с $r = 1$. Обозначим $(\gamma_0^i \gamma_1^i \dots \gamma_p^i) \doteq \gamma_\theta^i$ i -ю строку матрицы γ_θ (14). Тогда соотношение (15) запишется в виде

$$\begin{aligned} \text{vect } \gamma_\theta &= \begin{pmatrix} \gamma_\theta^{1\top} \\ \vdots \\ \gamma_\theta^{r\top} \end{pmatrix} = \begin{pmatrix} d_1 + D_1\theta \\ \vdots \\ d_r + D_r\theta \end{pmatrix} = \\ &= \begin{pmatrix} d_1 & D_1 \\ \vdots & \vdots \\ d_r & D_r \end{pmatrix} \begin{pmatrix} 1 \\ \theta \end{pmatrix} \doteq \begin{pmatrix} D_1 \\ \vdots \\ D_r \end{pmatrix} \vartheta \doteq D\vartheta. \end{aligned}$$

Оценки ОРС определяются соотношением

$$(34) \quad \theta_S = \arg \min_{\theta} J_S(\theta, \tilde{z}),$$

$$J_S(\theta, \tilde{z}) \doteq \frac{\vartheta^T D_1^T \bar{V}(\tilde{z})^T \bar{V}(\tilde{z}) D_1 \vartheta}{\vartheta^T D_1^T D_1 \vartheta} + \dots + \frac{\vartheta^T D_r^T \bar{V}(\tilde{z})^T \bar{V}(\tilde{z}) D_r \vartheta}{\vartheta^T D_r^T D_r \vartheta}.$$

Несложно установить равенство

$$(35) \quad J_S(\theta, \tilde{z}) = \vartheta^T D^T V(\tilde{z})^T C_{S,\theta} V(\tilde{z}) D \vartheta,$$

$$C_{S,\theta} \doteq I_{N-p} \otimes \begin{pmatrix} \gamma_{\theta}^1 \gamma_{\theta}^{1T} & & 0 \\ & \ddots & \\ 0 & & \gamma_{\theta}^r \gamma_{\theta}^{rT} \end{pmatrix}^{-1}.$$

Выражение (35) получается из (27) занулением недиагональных элементов матрицы $\gamma_{\theta}^1 \gamma_{\theta}^{1T}$.

Матрицы P_j остаются теми же, что и в методе ОРМ (26).

Убедимся, что оценки ОРС принадлежат классу вариационных. Действительно, ввиду соотношений (23), (25), (34) ядро целевой функции $J_S(\theta, \tilde{z})$ есть сумма проекторов

$$\Pi_{1,\theta}^{(1)} + \dots + \Pi_{N-p,\theta}^{(1)} + \dots + \Pi_{1,\theta}^{(r)} + \dots + \Pi_{N-p,\theta}^{(r)},$$

где проекторы $\Pi_{j,\theta}^{(i)}$ получаются из проекторов $\Pi_{j,\theta}$ (23) заменой γ_{θ} на строку γ_{θ}^i . Далее см. определение 2.

3. Свойства вариационных оценок

Установим условия единственности вариационных оценок (7), (13). Воспользуемся представлениями (18), (19), (20) для целевых функций (6), (12).

Теорема 2. *Необходимым и достаточным условием единственности вариационной оценки (7) параметра θ системы (4) по невозмущенному наблюдению $\tilde{z} = z$, $G_{\theta} z = 0$, является строгая положительная определенность матрицы*

$$Q_M \doteq D^T \left(\sum_{j=1}^M V_j(z)^T V_j(z) \right) D > 0,$$

где $V_j(z) \doteq P_j V(z)$, и матрицы P_j определены в (18).

Доказательство. Единственность равносильна соотношению

$$\forall \Delta\theta \in \mathbb{R}^v \quad J(\theta + \Delta\theta, z) > J(\theta, z).$$

Из (19) следует равенство

$$\begin{aligned} J(\theta + \Delta\theta, z) - J(\theta, z) &= \\ &= \Delta\theta^\top D^\top \left(\sum_{j=1}^M V_j(z)^\top C_{j,\theta+\Delta\theta} V_j(z) \right) D \Delta\theta. \end{aligned}$$

Здесь учтено равенство $V_j(z) (D\theta + d) = 0$. Имеет место строгая положительная определенность матриц $C_{j,\theta+\Delta\theta} > 0$. Поэтому неравенство $J(\theta + \Delta\theta, z) > J(\theta, z)$ имеет место тогда и только тогда, когда строго положительно определена матрица

$$D^\top \left(\sum_{j=1}^M V_j(z)^\top V_j(z) \right) D > 0.$$

Теорема доказана.

Теорема 3. *Необходимым и достаточным условием единственности вариационной оценки (13) параметра θ системы (4) по набору невозмущенных решений $\{\check{z}_{(1)} = z_{(1)}, \dots, \check{z}_{(L)} = z_{(L)}\}$, $G_\theta z_{(i)} = 0$, является строгая положительная определенность матрицы*

$$(36) \quad Q_{(L),M} \doteq L^{-1} \sum_{i=1}^L D^\top \left(\sum_{j=1}^M V_j(z_{(i)})^\top V_j(z_{(i)}) \right) D > 0.$$

Доказательство этой теоремы опирается на соотношение (20) и аналогично доказательству теоремы 2.

3.1. УСЛОВИЯ КОРРЕКТНОСТИ ПАРАМЕТРИЗАЦИИ

Выясним, как зависит выполнение условия единственности (36) от устройства матрицы D . Согласно (15), матрица D используется для задания параметризации $\theta \leftrightarrow G_\theta$.

Определение 5 [43]. Пусть $Z = \{z_{(1)}, \dots, z_{(L)}\}$ — множество наблюдений $z_{(i)}$ таких, что $G_\theta z_{(i)} = 0$. отождествим множество Z с матрицей $(z_{(1)} \dots z_{(L)})$, тогда $G_\theta Z = 0$. Параметризация $\theta \leftrightarrow G_\theta$ называется корректной на множестве наблюдений Z , если для каждого значения параметра $\xi \in \Theta \setminus \{\theta\}$ имеет место неравенство $G_\xi Z \neq 0$.

Дадим обобщение этого определения на случай набора матриц (18) $\Phi_{j,\theta} \doteq P_j G_\theta$, $j \in \overline{1, M}$. Введем обозначения

$$(37) \quad \Phi_\theta \doteq \begin{pmatrix} \Phi_{1,\theta} \\ \vdots \\ \Phi_{M,\theta} \end{pmatrix} = \underbrace{\begin{pmatrix} P_1 \\ \vdots \\ P_M \end{pmatrix}}_{\tilde{P}} G_\theta \doteq \tilde{P} G_\theta.$$

Определение 6. Параметризация $\theta \leftrightarrow \Phi_\theta$ называется корректной на множестве наблюдений Z , если для каждого значения параметра $\xi \in \Theta \setminus \{\theta\}$ имеет место неравенство $\Phi_\xi Z \doteq \tilde{P} G_\xi Z \neq 0$.

Предложение 1. Условие единственности (36) выполнено тогда и только тогда, когда параметризация $\theta \leftrightarrow \Phi_\theta$ корректна на множестве наблюдений $Z = \{z_{(1)}, \dots, z_{(L)}\}$.

Доказательство. Согласно определению (16), неравенство $\Phi_\xi Z \neq 0$ равносильно неравенству

$$\underbrace{\begin{pmatrix} \tilde{P}V(z_{(1)}) \\ \vdots \\ \tilde{P}V(z_{(L)}) \end{pmatrix}}_W (D\xi + d) \doteq W (D\xi + d) \neq 0.$$

Учитывая соотношения $\Phi_\theta Z = 0$, $W (D\theta + d) = 0$, получаем:

$$W (D\xi + d) - W (D\theta + d) = WD (\xi - \theta) \neq 0.$$

Ввиду произвольности ξ последнее означает, что столбцы матрицы WD линейно независимы. Следовательно, строго положи-

тельно определена матрица

$$Q_{(L),M} \doteq L^{-1} \sum_{i=1}^L D^\top \left(\sum_{j=1}^M V_j(z_{(i)})^\top V_j(z_{(i)}) \right) D = \\ = L^{-1} D^\top W^\top W D > 0.$$

Обратное тоже верно. Предложение доказано.

Определение 7. Пусть $Z_\theta \doteq \{z : G_\theta z = 0\}$ — линейное многообразие всех решений системы (4). И пусть H_θ — матрица, столбцы которой образуют фундаментальную систему решений (4) (базис Z_θ). Параметризация $\theta \leftrightarrow G_\theta$ (или $\theta \leftrightarrow \Phi_\theta$) называется корректной, если она корректна на множестве $Z = H_\theta$.

Заметим, что параметризация $\theta \leftrightarrow G_\theta$ полностью определяется системой (4), а параметризация $\theta \leftrightarrow \Phi_\theta$ зависит еще и от метода идентификации (ОР, ОРМ, ВИ, ОРС, ...), т.е. от набора матриц P_j (37).

Из корректности параметризации $\theta \leftrightarrow \Phi_\theta$ следует корректность параметризации $\theta \leftrightarrow G_\theta$. Обратное, вообще говоря, неверно. Установим условия, при которых корректность $\theta \leftrightarrow G_\theta$ равносильна корректности $\theta \leftrightarrow \Phi_\theta$. В этом случае условия единственности оценок не будут зависеть от метода идентификации (от набора матриц P_j).

Определение 8 [15]. Множество наблюдений $Z = \{z_{(1)}, \dots, z_{(L)}\} \subset Z_\theta$ называем полным, если оно содержит базис многообразия Z_θ .

Покажем, что для оценок ОРМ, ВИ, ОРС условия корректности параметризаций $\theta \leftrightarrow G_\theta$ и $\theta \leftrightarrow \Phi_\theta$ совпадают. Другими словами, имеет место следующая теорема.

Теорема 4. При полном множестве наблюдений $Z = \{z_{(1)}, \dots, z_{(L)}\} \subset Z_\theta$ единственность оценок ОРМ, ВИ, ОРС имеет место тогда и только тогда, когда корректна параметризация $\theta \leftrightarrow G_\theta$.

Доказательство. Согласно предложению 1, единственность любой оценки вида (13) равносильна корректности параметриза-

ции $\theta \leftrightarrow \Phi_\theta$, т. е. когда для всех значений параметров $\xi \in \Theta \setminus \{\theta\}$ имеет место неравенство $\Phi_\xi H_\theta \neq 0$. Заметим, что для каждой из оценок ОРМ, ВИ, ОРС соответствующие матрицы $\tilde{P} = \begin{pmatrix} P_1 \\ \vdots \\ P_M \end{pmatrix}$ (26), (28) все имеют линейно независимые столбцы. Поэтому соотношения $\Phi_\xi H_\theta = \tilde{P} G_\xi H_\theta \neq 0$ и $G_\xi H_\theta \neq 0$ для этого случая равносильны. Теорема доказана.

Важно, что условия корректности параметризации $\theta \leftrightarrow G_\theta$ допускают равносильную формулировку в виде ограничений на ранги специальных подматриц из элементов «малой» матрицы γ_θ (14) [16, 17, 18].

Определение 9 [17]. Если все ненулевые строки матрицы D образуют единичную подматрицу, то параметризация называется простой (или правильной).

В литературе это наиболее часто встречающийся тип параметризаций.

Замечание 1. Ввиду того, что свойство максимальности ранга устойчиво по отношению к малым изменениям элементов матрицы, для простых параметризаций корректность оказывается в пространстве параметров свойством общего положения (этот факт следует из теоремы 4 статьи [17]).

К этому замечанию мы вернемся ниже при обсуждении состоятельности.

3.2. УСТОЙЧИВОСТЬ

Устойчивость вариационной оценки $\hat{\theta}(\tilde{z})$ (7) прямо следует из результатов статьи [21]. Сформулируем утверждение.

Теорема 5. Пусть выполнено условие единственности из теоремы 2, $\epsilon \doteq \|\tilde{z} - z\|$, $\Delta\theta \doteq \hat{\theta}(\tilde{z}) - \hat{\theta}(z) = \hat{\theta}(\tilde{z}) - \theta$,

$$(38) \quad R_0 \doteq D^\top \left(\sum_{j=1}^M V_j(z)^\top C_{j,\theta} V_j(z) \right) D$$

и $\lambda_{\min}(R_0)$ — наименьшее собственное число. Тогда

$$\|\Delta\theta\| \leq \frac{1}{\sqrt{\lambda_{\min}(R_0)}} \epsilon + O(\epsilon^2).$$

Верхняя граница для величины остаточного члена $O(\epsilon^2)$ также получена в [21]. Она имеет сложный вид и здесь не приводится.

Для оценки (13) утверждение об устойчивости аналогично теореме 5. Ввиду соотношения (20) нужно наложить условие единственности из теоремы 3 и заменить матрицу R_0 (38) на матрицу

$$(39) \quad R_{\Sigma,0} \doteq L^{-1} \sum_{i=1}^L D^\top \left(\sum_{j=1}^M V_j(z_{(i)})^\top C_{\Sigma,j,\theta} V_j(z_{(i)}) \right) D.$$

На основании теоремы 5 можно предложить количественные критерии идентифицируемости параметров уравнения (1), основанные на вычислении минимального собственного числа матрицы R_0 (38) (или $R_{\Sigma,0}$ (39)). На практике матрица R_0 может быть заменена оценкой \hat{R}_0 , получаемой заменой неизвестных истинных значений θ, z на оценки $\hat{\theta}, z_{\text{opt}}$ [22].

3.3. СОСТОЯТЕЛЬНОСТЬ

Прежде изучения свойств оценки на конечных выборках следует убедиться в ее состоятельности. В статье [40] была установлена состоятельность оценок STLS в пределе $N \rightarrow \infty$ по наблюдению одного процесса ($L = 1$). Здесь рассмотрим предельный случай $L \rightarrow \infty$ наблюдений большого числа процессов конечных длин N с независимыми начальными условиями (подлежащими оцениванию совместно с параметрами уравнения). В этом случае оценка $\hat{\theta}(\tilde{z})$ (7) переходит в оценку $\hat{\theta}(\tilde{z}_{(1)}, \dots, \tilde{z}_{(L)}) \doteq \hat{\theta}_L$ (10) или (13), и сильная состоятельность означает сходимость $\hat{\theta}_L \rightarrow \theta$ п. н. в пределе $L \rightarrow \infty$, где $\hat{\theta}_L$ — одна из точек локального минимума целевой функции. Сформулируем утверждение.

Теорема 6. Пусть выполнено предположение о возмущениях (11). Если (и только если) выполнено условие

$$(40) \quad Q_{(\infty),M} \doteq \lim_{L \rightarrow \infty} Q_{(L),M} > 0,$$

где матрица $Q_{(L),M}$ определена в (36), то вариационная оценка $\hat{\theta}_L$ (13) сильно состоятельна по $L \rightarrow \infty$, где под $\hat{\theta}_L$ понимается одна из точек локального минимума целевой функции (12).

Доказательство близко по идее к доказательству теоремы 1 из [15] и полностью будет опубликовано в другой статье.

В ряде случаев оказывается возможным заменить целевую функцию на более простую. Необходимым условием является сохранение состоятельности оценок.

Определение 10. При возмущениях (11) две оценки вида (13) с различающимися наборами матриц

$$\mathcal{W}_1 \doteq \{W_{1_1}, \dots, W_{M_1}\}, \quad \mathcal{W}_2 \doteq \{W_{1_2}, \dots, W_{M_2}\}$$

называются равносильными по состоятельности, если условия (40) для соответствующих наборов \mathcal{W}_1 и \mathcal{W}_2 равносильны. При отсутствии возмущений оценки вида (13) равносильны по состоятельности в детерминированном смысле, если для них при любых одинаковых множествах невозмущенных наблюдений $\{\check{z}_{(1)} = z_{(1)}, \dots, \check{z}_{(L)} = z_{(L)}\}$ равносильны условия (36).

Заметим, что из равносильности по состоятельности в детерминированном смысле всегда следует равносильность по состоятельности.

Равносильные по состоятельности оценки могут отличаться асимптотическими свойствами.

Теорема 7. Вариационные оценки ВИ, ОРМ, ОРС (с целевыми функциями соответственно (29), (27) и (35)) равносильны по состоятельности в детерминированном смысле.

Доказательство. Достаточно заметить, что перечисленные оценки отличаются только матрицами C_θ , $C_{m,\theta}$, $C_{s,\theta}$ в ядрах целевых функций (29), (27) и (35). Эти матрицы не участвуют в формулировке условий единственности (36).

3.4. СОСТОЯТЕЛЬНОСТЬ И ИДЕНТИФИЦИРУЕМОСТЬ КАК СВОЙСТВА НЕОБЩЕГО ПОЛОЖЕНИЯ

Из теорем 4, 6 следует, что при «естественных» предположениях о распределении процессов $z_{*(i)}$ и возмущений $\varepsilon_{(i)}$ для состоятельности оценок ОРМ, ВИ, ОРС необходимо и достаточно

выполнить условия корректности параметризации $\theta \leftrightarrow G_\theta$. Поэтому, согласно замечанию 1, для простых параметризаций состоятельность оказывается свойством общего положения. Подобный факт был установлен ранее П. Стойкой, который исследовал прямые оценки, получаемые методами типа инструментальных переменных (МИП), и ввел понятие *состоятельности в общем положении* (*generic consistency*) [51, с. 266]. Оно означает отсутствие состоятельности в точках множества меры ноль в пространстве параметров. Причиной является вырождение матрицы ковариации наблюдений [51, с. 272]. В таких точках бесконечно растет и разброс оценок МИП [51, с. 273].

Заметим, что если использовать *количественные* показатели разброса оценок (например, через детерминант или сингулярные числа матрицы ковариации), то понятие состоятельности в общем положении теряет смысл. Например, пусть считается неприемлемым разброс оценок больше заданного уровня E , тогда отсутствие состоятельности (идентифицируемости) будет происходить уже на множестве меры > 0 (эта мера будет зависеть от E).

Нарушение корректности параметризации также происходит на множестве меры ноль (см. замечание 1). Введение количественных показателей идентифицируемости (раздел 3.2) делает множество неидентифицируемых систем множеством меры > 0 , и идентифицируемость перестает быть свойством общего положения.

3.5. МИНИМИЗАЦИЯ ПРОЕКТИВНЫХ ЦЕЛЕВЫХ ФУНКЦИЙ

Минимизация $J(\theta, \tilde{z})$ (22) по θ осуществляется итерациями с обратной матрицей:

$$(41) \quad \vartheta_{(k)} \doteq \begin{pmatrix} 1 \\ \theta_{(k)} \end{pmatrix}, \quad \begin{cases} \tau \doteq A^{-1} B \vartheta_{(k)}, \\ \vartheta_{(k+1)} = \frac{1}{(10 \dots 0) \tau} \tau, \end{cases}$$

$$A \doteq D^\top V_{\text{OR}}(\tilde{z})^\top V_{\text{OR}}(\tilde{z}) D, \quad B \doteq D^\top D.$$

Запись $\frac{1}{(10 \dots 0) \tau}$ означает нормирование вектора τ с приведением первого элемента к единице. Этот алгоритм является мо-

дификацией стандартного способа поиска минимального (соответствующего минимальному собственному числу) собственного вектора симметричной матрицы [8, с. 421].

Целевая функция метода ВИ (29) имеет сложный характер изоповерхностей [6, 13]. Применение универсальных алгоритмов минимизации типа Ньютона крайне затруднено из-за их малого радиуса сходимости [44]. Тем не менее, в [9, 10] был найден эффективный вычислительный метод, основанный на модификации алгоритма (41):

$$(42) \quad \vartheta_{(k)} \doteq \begin{pmatrix} 1 \\ \theta_{(k)} \end{pmatrix}, \quad \begin{cases} \tau \doteq A_{(k)}^{-1} B \vartheta_{(k)}, \\ \vartheta_{(k+1)} = \frac{1}{(10\dots 0)\tau} \tau, \end{cases}$$

$$A_{(k)} \doteq D^T V(\tilde{z})^T C_{\theta_{(k)}} V(\tilde{z}) D, \quad B \doteq D^T D.$$

Итерации типа (42) в частном случае однородных систем (1) независимо от А. О. Егоршина использовал М. Осборн [46] и позже переоткрыл Б. Де Мур [45], применяя их для решения задачи STLS (см. раздел 2.4). Особенностью итераций Егоршина—Осборна является высокая скорость сходимости в 2–4 итерации в малую окрестность глобального минимума при слабой зависимости от начального приближения. Примеры расчетов приведены в [15, 20, 22]. Локальную сходимость итерационной процедуры (42) исследовал В. Г. Демиденко [7].

В методе ОРМ минимизация целевой функции (27) может быть осуществлена итерациями (42) с заменой $C_{\theta} = (G_{\theta} G_{\theta}^T)^{-1}$ на $C_{m,\theta}$ (27) (результаты экспериментальной проверки этого факта здесь не приводим).

3.6. ЧИСЛО ЭКСТРЕМУМОВ ЦЕЛЕВОЙ ФУНКЦИИ ОРС

Для вариационной идентификации параметров систем (1) оценки ОРС оказываются одними из наиболее простых и в то же время состоятельными (теорема 7), с хорошими экстремальными свойствами.

Теорема 8. Число локальных экстремумов целевой функции ОРС $J_S(\theta, \tilde{z})$ (34) не превосходит $r(p+1)(r+m)$.

Доказательство. Из выражения (34) получаем уравнение на критические точки функции $J_S(\vartheta) \doteq J_S(\theta, \xi)$:

$$\frac{\partial J_S(\vartheta)}{\partial \vartheta} = 2\vartheta^\top \left[\sum_{i=1}^r \frac{D_i^\top \bar{V}^\top \bar{V} D_i}{(\vartheta^\top D_i^\top D_i \vartheta)} - \frac{(\vartheta^\top D_i^\top \bar{V}^\top \bar{V} D_i \vartheta)}{(\vartheta^\top D_i^\top D_i \vartheta)^2} D_i^\top D_i \right] = 0.$$

Обозначим

$$\begin{aligned} a_i(\vartheta) &\doteq \vartheta^\top D_i^\top \bar{V}^\top \bar{V} D_i \vartheta, & A_i &\doteq D_i^\top \bar{V}^\top \bar{V} D_i, \\ b_i(\vartheta) &\doteq \vartheta^\top D_i^\top D_i \vartheta, & B_i &\doteq D_i^\top D_i, \\ J_i(\vartheta) &\doteq \frac{a_i(\vartheta)}{b_i(\vartheta)}. \end{aligned}$$

Уравнение на критические точки запишем в виде

$$\left[\sum_{i=1}^r \frac{1}{b_i(\vartheta)} D_i^\top (\bar{V}^\top \bar{V} - J_i(\vartheta) I) D_i \right] \vartheta = 0.$$

В матричной записи

$$\begin{aligned} (43) \quad & \left(D_1^\top \quad \dots \quad D_r^\top \right) \times \\ & \times \begin{pmatrix} \frac{(\bar{V}^\top \bar{V} - J_1(\vartheta) I)}{b_1(\vartheta)} & & 0 \\ & \ddots & \\ 0 & & \frac{(\bar{V}^\top \bar{V} - J_r(\vartheta) I)}{b_r(\vartheta)} \end{pmatrix} \begin{pmatrix} D_1 \\ \vdots \\ D_r \end{pmatrix} \vartheta \doteq \\ & \doteq D^\top \begin{pmatrix} W_1 & & 0 \\ & \ddots & \\ 0 & & W_r \end{pmatrix} D \vartheta \doteq D^\top W D \vartheta = 0. \end{aligned}$$

Несложно проверить тождество $\forall \vartheta \quad \vartheta^\top D^\top W D \vartheta = 0$. Оно показывает, что производная $\frac{\partial J_S(\vartheta)}{\partial \vartheta}$ перпендикулярна направлению ϑ , т. е. $J_S(\vartheta)$ не зависит от нормы ϑ .

Как следует из принятых обозначений, матрица W симметричная. Ввиду линейной независимости столбцов матрицы D условие (43) при $\vartheta \neq 0$ означает, что матрица W вырождена. Отсюда следует необходимое условие существования экстремальной точки:

$$\exists i \in \overline{1, r} \quad \det \left(\bar{V}^\top \bar{V} - J_i(\vartheta) I \right) = 0.$$

Следовательно, число экстремальных точек не превосходит числа rt , где $t = (p + 1)(r + m)$ есть размерность матрицы $\bar{V}^T \bar{V}$ (см. (33)). Теорема доказана.

Будем говорить, что разные уравнения системы (1) параметризованы аффинно и независимо, если матрица D из условия (15) имеет клеточно-диагональный вид

$$(44) \quad D = \begin{pmatrix} D_{11} & & 0 \\ & \ddots & \\ 0 & & D_{rr} \end{pmatrix},$$

$$D_{ii} \in \mathbb{R}^{(p+1)(r+m) \times v_i}, \quad \sum_{i=1}^r v_i = v + 1,$$

а вектор параметров $\vartheta = (1; \theta)$ распадается на r подвекторов $\vartheta^i = (1; \theta^i)$, $i \in \overline{1, r}$, так что каждая из строк γ^i в матрице γ_θ (14) системы (1), (4) зависит только от i -го подвектора параметров θ^i .

Теорема 9. Пусть разные уравнения системы (1) параметризованы аффинно и независимо. Тогда число экстремальных точек целевой функции ОРС (34) не превосходит $r(v + 1)$, а оценка ОРС ϑ_S , минимизирующая (34), имеет вид $\vartheta_S = (\vartheta_S^1; \dots; \vartheta_S^r)$, где каждый из подвекторов ϑ_S^i является решением скалярной ($r = 1$) задачи ОРМ вида (25):

$$(45) \quad \vartheta_S^i = \arg \min_{\theta} J_i(\vartheta),$$

$$J_i(\vartheta) \doteq \frac{\vartheta^T D_{ii}^T \bar{V}^T \bar{V} D_{ii} \vartheta}{\vartheta^T D_{ii}^T D_{ii} \vartheta}.$$

Доказательство. В случае независимой параметризации уравнений из выражения (35) ввиду (44) и (33) получаем

$$J_S(\vartheta) = J_S(\vartheta^1, \dots, \vartheta^r) =$$

$$= \frac{\vartheta^1 T D_{11}^T \bar{V}^T \bar{V} D_{11} \vartheta^1}{\vartheta^1 T D_{11}^T D_{11} \vartheta^1} + \dots + \frac{\vartheta^r T D_{rr}^T \bar{V}^T \bar{V} D_{rr} \vartheta^r}{\vartheta^r T D_{rr}^T D_{rr} \vartheta^r}.$$

Целевая функция (45) имеет ровно $v + 1$ критических точек, где $v + 1$ — размерность вектора ϑ , поскольку все такие точки являются собственными векторами регулярного пучка матриц $D_{ii}^T (\bar{V}^T \bar{V} - \mu I) D_{ii}$ [4, гл.Х, п.6]. Отсюда сразу следует утверждение теоремы.

4. Заключение

После работ А.Н. Колмогорова и Н. Винера [12, 54] в области приложений методов идентификации возобладали подход, основанный на вычислении корреляционных функций наблюдаемых сигналов. Это направление со временем стало магистральным [1, 23, 29, 32, 35, 51], во многом благодаря ранним работам К. Острема с коллегами [25]. Несмотря на простоту, корреляционные («прямые») методы идентификации на конечных выборках наблюдений нередко приводили к трудностям в интерпретации результатов ввиду большого числа локальных экстремумов у целевой функции.

Известная работа Я. Виллемса [2] послужила толчком для развития направления, основанного более на идеях аппроксимации наблюдений, чем на вычислении коэффициентов корреляций. Эту группу методов мы называем вариационными («непрямыми»). Первые существенные результаты в этой области были получены значительно раньше А.О. Егоршиным [9, 10] и М. Осборном [46]. Развитие этого направления связано с именами Г. Голуба, Ч. Ван Лоана [39], Б. Де Мура [45], С. Ван Хуффель, Дж. Вандевалле [53] и мн. др. Непрямые методы идентификации более сложны с точки зрения числа арифметических операций.

В данной статье впервые предложено точное определение для всего класса вариационных методов и рассмотрены их экстремальные свойства. Показано, что в ряде существенных случаев увеличение числа арифметических операций приводит не к усложнению процедуры идентификации, а в конечном счете, наоборот, к ее упрощению, ввиду значительного уменьшения числа локальных экстремумов, и как следствие, улучшения свойств сходимости итерационных процедур.

Автор благодарит рецензентов за конструктивные замечания, способствовавшие улучшению текста статьи.

Приложение. Доказательство теоремы 1

Первое утверждение теоремы сформулируем более конкретно. Пусть θ_S, z_S — решение задачи (30)–(32). Тогда можно указать целевую функцию $J_S(\theta, \check{x}) = \check{x}^\top W_{S,\theta} \check{x}$, $\check{x} \doteq (1; \check{z})$, с ядром таким, что произведение $\Sigma^\top W_{S,\theta} \Sigma = \Pi_{S,\theta}$, $\Sigma = \begin{pmatrix} 0 & 0 \\ 0 & R \end{pmatrix}$, есть проективная матрица $\Pi_{S,\theta}^2 = \Pi_{S,\theta}$, и при этом $\theta_S = \arg \min_{\theta} J_S(\theta, \check{z})$. Докажем это утверждение.

Условие $V(z)g(\theta) = 0$ равносильно уравнению

$$[I \otimes g(\theta)] \text{vect } V(z) = 0.$$

Ввиду соотношения (30) имеем $\text{vect } V(z) = f + Fz$, где вектор f и матрица F вычисляются по матрицам V_0, \dots, V_n . Также из (32) следует представление

$$I \otimes g(\theta) = \Gamma_{\theta}, \quad \Gamma_{\theta} = \Gamma_0 + \Gamma_1 \theta_1 + \dots + \Gamma_v \theta_v,$$

где матрицы $\Gamma_0, \dots, \Gamma_v$ вычисляются по матрицам g_0, \dots, g_v .

Задачу (31) запишем в виде

$$\min_{z \in \mathbb{R}^n, \theta \in \mathbb{R}^v} \|R^{-1}(\check{z} - z)\|^2 \quad \text{при условии} \quad \Gamma_{\theta}(f + Fz) = 0.$$

Определим векторы $x(z) \doteq \begin{pmatrix} 1 \\ z \end{pmatrix} \in \mathbb{R}^{n+1}$, $\check{x} \doteq x(\check{z})$ и матрицы $\Sigma \doteq \begin{pmatrix} 0 & 0 \\ 0 & R \end{pmatrix}$, $G_{\theta} \doteq (\Gamma_{\theta} f \quad \Gamma_{\theta} F)$. Тогда $G_{\theta} x = 0$, $\|R^{-1}(\check{z} - z)\| = \|\Sigma^+ (\check{x} - x)\|$, где $\Sigma^+ = \begin{pmatrix} 0 & 0 \\ 0 & R^{-1} \end{pmatrix}$ — обобщенная обратная Мура–Пенроуза [27, 1.b.5.(IV)]. В новых обозначениях задача (31) принимает вид

$$\min_{x \in \mathbb{R}^{n+1}, \theta \in \mathbb{R}^v} \|\Sigma^+ (\check{x} - x)\|^2$$

при условиях

$$G_{\theta} x = 0, \quad (1, 0, \dots, 0)x = 1.$$

Минимизация по x приводит к безусловной задаче минимизации

$$\min_{\theta \in \mathbb{R}^v} \check{x}^\top \underbrace{G_\theta^\top \left(G_\theta \Sigma \Sigma^\top G_\theta^\top \right)^{-1} G_\theta}_{W_{S,\theta}} \check{x},$$

$$x_{\text{opt}} = \left[I - \Sigma \Sigma^\top G_\theta^\top \left(G_\theta \Sigma \Sigma^\top G_\theta^\top \right)^{-1} G_\theta \right] \check{x}.$$

Матрица $\Sigma^\top W_{S,\theta} \Sigma$ является проектором. Поэтому оценка STLS

$$\hat{\theta} = \arg \min_{\theta} \check{x}^\top W_{S,\theta} \check{x}$$

есть вариационная согласно определению 3. Первая часть теоремы доказана.

Прежде доказательства второй части установим три леммы.

Лемма 1. Пусть $FXF^\top \doteq \Pi = \Pi^\top \geq 0$ — неотрицательно определенная симметричная матрица со свойством проектора: $\Pi^2 = \Pi$. И пусть матрица X неособенная, и столбцы матрицы F линейно независимы. Тогда $X = (F^\top F)^{-1}$.

Доказательство. Из равенства $(FXF^\top)(FXF^\top) = FXF^\top$ следует $F^\top(FXF^\top)(FXF^\top)F = F^\top FXF^\top F$, откуда сразу получаем $XF^\top FX = X$. Учитывая неособенность X , приходим к соотношению $F^\top FX = I$, т. е. $X = (F^\top F)^{-1}$.

Лемма 2. Пусть $W^\top = W \geq 0$ — неотрицательно определенная симметричная матрица со свойством $(Gz = 0) \Rightarrow (Wz = 0)$, где G — заданная матрица с линейно независимыми строками. И пусть для заданной матрицы Σ с линейно независимыми столбцами выполняется соотношение $\Sigma^\top W \Sigma \doteq \Pi$, $\Pi^2 = \Pi$. Тогда $W = G^\top P^\top (PG\Sigma\Sigma^\top G^\top P^\top)^{-1} PG$, где P — некоторая матрица с линейно независимыми строками.

Доказательство. Из симметричности W и свойства $(Gz = 0) \Rightarrow (Wz = 0)$ следует $W = G^\top P^\top XPG$, где X — некоторая неособенная матрица. По условию,

$$\left(\Sigma^\top W \Sigma \right) \left(\Sigma^\top W \Sigma \right) = \Sigma^\top W \Sigma.$$

После переобозначения $PG\Sigma \doteq F^\top$ получаем равенство $(F XF^\top)(F XF^\top) = F XF^\top$. По лемме 1 $X = (F^\top F)^{-1}$, тогда

$$W = G^\top P^\top \left(PG\Sigma\Sigma^\top G^\top P^\top \right)^{-1} PG.$$

Лемма доказана.

Лемма 3. *Задача минимизации*

$$J = \min_{z \in \mathbb{R}^n} \|\Sigma^{-1}(\check{z} - z)\|^2 \quad \text{при условии} \quad Hz = 0$$

имеет единственное решение

$$J = \check{z}^\top H^\top \left(H\Sigma\Sigma^\top H^\top \right)^{-1} H\check{z},$$

$$z_{\text{opt}} = \left[I - \Sigma\Sigma^\top H^\top \left(H\Sigma\Sigma^\top H^\top \right)^{-1} H \right] \check{z}.$$

Доказательство. Введем функцию Лагранжа

$$J^*(z, \lambda) = \|\Sigma^{-1}(\check{z} - z)\|^2 + \lambda^\top Hz.$$

Необходимое условие минимума $J_\lambda^{*'} = 0$, $J_z^{*'} = 0$:

$$Hz = 0, \quad H^\top \lambda - 2\Sigma^{-\top} \Sigma^{-1}(\check{z} - z) = 0.$$

Левым умножением последнего уравнения на $H\Sigma\Sigma^\top$ получаем

$$\lambda = 2 \left(H\Sigma\Sigma^\top H^\top \right)^{-1} H\check{z}.$$

Подставляя λ в условие минимума, вычисляем

$$z_{\text{opt}} = \left[I - \Sigma\Sigma^\top H^\top \left(H\Sigma\Sigma^\top H^\top \right)^{-1} H \right] \check{z}.$$

Подстановка z_{opt} в J приводит к выражению

$$J = \check{z}^\top H^\top \left(H\Sigma\Sigma^\top H^\top \right)^{-1} H\check{z}.$$

Лемма доказана.

Докажем вторую часть теоремы. Для упрощения обозначений положим $L = 1$, $M = 1$. В случае $L > 1$, $M > 1$ ход рассуждений остается тем же.

Пусть (12), т. е. $J(\theta, \check{z}) = \check{z}^\top W_\theta \check{z}$. Матрица W_θ удовлетворяет условиям леммы 2, поэтому имеет место выражение

$$W_\theta = G_\theta^\top P^\top \left(P G_\theta \Sigma \Sigma^\top G_\theta^\top P^\top \right)^{-1} P G_\theta$$

для некоторой матрицы P с линейно независимыми строками. Обозначим $H_\theta \doteq P G_\theta$, тогда

$$J(\theta, \check{z}) = \check{z}^\top H_\theta^\top \left(H_\theta \Sigma \Sigma^\top H_\theta^\top \right)^{-1} H_\theta \check{z}.$$

По лемме 3

$$J(\theta, \check{z}) = \min_{z \in \mathbb{R}^n} \|\Sigma^{-1}(\check{z} - z)\|^2 \quad \text{при условии} \quad H_\theta z = 0.$$

Используя определения (16), (17), всегда можно построить матрицу $V(z)$ с аффинной параметризацией (30) такую, что верно соотношение $H_\theta z \equiv V(z) \text{vect } \gamma_\theta \equiv V(z) D \vartheta = 0$. Осталось заметить, что для матрицы (из одного столбца)

$$g(\theta) \doteq \text{vect } \gamma_\theta = D \vartheta = d + D \theta$$

выполнено условие (32). Мы получили, что оценка (13) есть оценка STLS (30)–(32). Теорема доказана.

Литература

1. БУНИЧ А. Л., БАХТАДЗЕ Н. Н. *Синтез и применение дискретных систем управления с идентификатором*. – М.: Наука, 2003. – 232 с.
2. ВИЛЛЕМС Я. *От временного ряда к линейной системе // Теория систем. Математические методы и моделирование*. – М.: Мир, 1989. – С. 8–191.

3. ВОРЧИК Б.Г. *Единственность оценок максимального правдоподобия параметров стохастических систем (проблема локальных экстремумов)* // Автоматика и телемеханика. – 1984. – № 6. – С. 47–55.
4. ГАНТМАХЕР Ф.Р. *Теория матриц*. – М.: Наука, 1966. – 576 с.
5. ГАУСС К. Ф. *Избранные геодезические сочинения*. – Т. 1. М.: Геодезиздат, 1957. – С. 59–87.
6. ДЕМИДЕНКО В.Г. *Разработка и программная реализация метода оценки параметров линейных моделей генных сетей*. – Магистерская диссертация. Новосибирск: Новосибирский гос. университет, 2007.
7. ДЕМИДЕНКО В.Г. *Восстановление коэффициентов систем линейных разностных уравнений* // Вестник НГУ (Серия: математика, механика, информатика). – 2010. – Т. 10. Вып. 2. – С. 45–53.
8. ДЕМИДОВИЧ Б. П., МАРОН И. А. *Основы вычислительной математики*. – М.: Наука, 1966. – 664 с.
9. ЕГОРШИН А. О. *Вычислительные замкнутые методы идентификации линейных объектов* // Оптимальные и самонастраивающиеся системы. – Новосибирск, 1971. – С. 40–53.
10. ЕГОРШИН А. О. *Метод наименьших квадратов и «быстрые» алгоритмы в вариационных задачах идентификации и фильтрации (метод ВИ)* // Автотметрия. – 1988. – № 1. – С. 30–42.
11. КАШЬЯП Р. Л., РАО А. Р. *Построение динамических стохастических моделей по экспериментальным данным*. – М.: Наука, 1983. – 384 с.
12. КОЛМОГОРОВ А. Н. *Интерполирование и экстраполирование стационарных случайных последовательностей* // Изв. АН СССР. Сер. мат. – 1941. – Т. 5. – С. 3–14. (В кн.: Колмогоров А. Н. *Теория вероятностей и математическая статистика*. – М.: Наука, 1986. – С. 255–263.)
13. КОСТИН В. И. *О точках экстремума одной функции* //

- Управляемые системы. – Новосибирск: ИМ СО АН СССР, 1984. Т. 24. – С. 35–42.
14. ЛОМОВ А. А. *Минимальные описания стационарных линейных моделей* // Труды Института математики СО РАН. Модели и методы оптимизации. – Новосибирск: ИМ СО РАН, 1994. Т. 28, – С. 91–117.
 15. ЛОМОВ А. А. *Идентификация линейных динамических систем по коротким участкам переходных процессов при аддитивных измерительных возмущениях* // Известия РАН ТСУ. – 1997. – № 3. – С. 20–26.
 16. ЛОМОВ А. А. *Параметрическая идентифицируемость линейных стохастических систем по наблюдениям коротких отрезков траекторий* // Известия РАН ТСУ. – 2002. – № 2. – С. 53–58.
 17. ЛОМОВ А. А. *Условия различимости стационарных линейных систем* // Дифференц. уравнения. – 2003. – Т. 39. № 2. – С. 261–266.
 18. ЛОМОВ А. А. *О различимости стационарных линейных систем с коэффициентами, зависящими от параметра* // Сибирский журнал индустриальной математики. – 2003. – Т. 6. № 4(16). – С. 60–66.
 19. ЛОМОВ А. А. *Сравнение методов оценивания параметров линейных динамических систем по измерениям коротких участков переходных процессов* // Автоматика и телемеханика. – 2005. – № 3. – С. 39–47.
 20. ЛОМОВ А. А. *Оценка трендов и идентификация динамики временных рядов на коротких интервалах наблюдения* // Известия РАН ТСУ. – 2009. – № 1. – С. 25–37.
 21. ЛОМОВ А. А. *О локальной устойчивости в задаче идентификации коэффициентов линейного разностного уравнения* // Вестник НГУ. Серия: Математика, механика, информатика. – 2010. – Т. 10, вып. 4. – С. 81–103.
 22. ЛОМОВ А. А. *О количественных априорных показателях идентифицируемости коэффициентов линейных динамических систем* // Известия РАН ТСУ. – 2011. – № 1. – С. 3–

- 15.
23. ЛЬЮНГ Л. *Идентификация систем*. – М.: Наука, 1991. – 432 с.
24. НОРКИН К.Б. *Поисковые методы настройки управляемых моделей в задачах настройки параметров объектов* // Автоматика и телемеханика. – 1968. – № 11. – С. 61–67.
25. ОСТРЕМ К., БОЛИН Т. *Цифровая идентификация линейных динамических систем на основе данных о нормальном режиме работы* // Теория самонастраивающихся систем управления: Труды II Международного конгресса IFAC. М.: Наука, 1969. – С. 99–116.
26. ПЕРЕЛЬМАН И. И. *Методы состоятельного оценивания параметров линейных динамических объектов и проблематичность их реализации на конечных выборках* // Автоматика и телемеханика. – 1981. – № 3. – С. 49–55.
27. РАО С. Р. *Линейные статистические методы и их применение*. – М.: Наука, 1968. – 548 с.
28. *Справочник по теории автоматического управления* / Под. ред. А. А. Красовского. – М.: Наука, 1987. – 712 с.
29. ЦЫПКИН Я. З. *Основы информационной теории идентификации*. – М.: Наука, 1984. – 336 с.
30. АВАТЗОГЛУ Т. J., MENDEL J. M. *Constrained Total Least Squares* // Proc. 1987 IEEE ICASSP (Dallas). 1987. – P. 1485–1488.
31. АОКИ М., YUE P. C. *On A Priori Error Estimates of Some Identification Methods* // IEEE Trans. on Automat. Control. V. AC-15. 1970. – P. 541–548.
32. ÅSTROM K. J., WITTENMARK B. *Adaptive Control*. – 2nd ed. Mineola, New York: Dover Publ. Inc. 2008. – 580 p.
33. BOHLIN T. *On the Problem of Ambiguities in Maximum Likelihood Identification* // Automatica. – 1971. – V. 7. – P. 137–146.
34. BROCKETT R. *Finite Dimensional Linear Systems*. – New York: Wiley, 1970. – 129 p.
35. CHEN H.-F. *Recursive Identification of EIV ARMA Processes*

- // Proceedings of the 17th World Congress IFAC. Seoul, Korea, July 6–11, 2008. – P. 1366–1371.
36. CHENG C.-L., VAN NESS J. W. *Statistical Regression with Measurement Error* // Kendall's Library of Statistics, 6, Arnold, – London. 1999. – 262 p.
37. DUPUIS P., SELS T., DRIESEN T., BELMANS R. *Exponential Parameters Measurement Using a Modified Prony Method* // Proc. Instrumentation and Measurement Technology Conference. Como, Italy, 18–20 May. 2004. – P. 1590–1594.
38. GLESER L. J. *Improvements of the Naive Approach to Estimation in Nonlinear Errors-in-Variables Regression Models* // Contemporary Mathematics. V. 112. Statistical Analysis of Measurement Error Models and Applications. Providence, Rhode Island: AMS, 1990. – P. 99–114.
39. GOLUB G. H., VAN LOAN C. F. *An Analysis of the Total Least Squares Problem* // SIAM J. Numer. Anal. – 1980. – V. 17, – P. 883–893.
40. KUKUSH A., MARKOVSKY I., VAN HUFFEL S. *Consistency of the Structured Total Least Squares Estimator in a Multivariate Errors-in-variables Model* // Journal of Statistical Planning and Inference. – 2005. – V. 133. No. 2. – P. 315–358.
41. LEMMERLING PH., MOOR DE B., VANHUFFEL S. *On the Equivalence of Constrained Total Least Squares and Structured Total Least Squares* // IEEE Trans. on Signal Processing. – 1996. V. 44. No. 11. – P. 2908–2911.
42. LEVIN M. J. *Estimation of a System Pulse Transfer Function in the Presence of Noise* // IEEE Trans. on Automatic Control. – 1964. – V. AC-9. – P. 229–235.
43. LOMOV A. A. *Correct Parametrizations of Linear Models* // Siberian Advances in Mathematics. – 1994. V. 4. – P. 95–113.
44. MAINE R. E., ILIFF K. W. *Formulation and Implementation of a Practical Algorithm for Parameter Estimation with Process and Measurement Noise* // SIAM Journal on Applied

- Mathematics. – 1981. V. 41. No. 3. – P. 558–579.
45. MOOR DE B. *Structured Total Least Squares and L_2 Approximation Problems* // Linear Algebra Appl. – 1993. V. 188–189. – P. 163–207.
 46. OSBORNE M.R. *A Class of Nonlinear Regression Problems* // Data Representation / Eds. R. S. Anderssen and M. R. Osborne. St. Lucia: University of Queensland Press, 1970. – P. 94–101.
 47. OSBORNE M.R., SMYTH G.K. *A modified Prony algorithm for fitting functions defined by difference equations* // SIAM J. Sci. Statist. Comput. – 1991. – V. 12. – P. 362–382.
 48. PEARSON K. *On Lines and Planes of Closest Fit to Systems of Points in Space* // Phil. Mag. – 1901. – VI. No. 2. – P. 559–572.
 49. ROORDA B., HEIJ C. *Global Total Least Squares Modelling of Multivariable Time Series* // IEEE Trans. on Automatic Control. – 1995. – V. AC-40. – P. 50–63.
 50. SÖDERSTRÖM T. *On the Uniqueness of Maximum Likelihood Identification* // Automatica. – 1975. – V. 11. – P. 193–197.
 51. SÖDERSTRÖM T., STOICAP. *System Identification*. – London: Prentice-Hall, 1989. – 613 p.
 52. SÖDERSTRÖM T. *Errors-in-variables Methods in System Identification* // 14th IFAC Symposium on System Identification, Newcastle, Australia, 2006. – P. 1–19.
 53. VAN HUFFEL S., VANDEWALLE J. *The total least squares problem*. – SIAM, Philadelphia, 1991. – 300 p.
 54. WIENER N. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications*. NDRC Report to the Services 370, February 1, 1942. – Cambridge: MIT Press, 1949. – 176 p.

VARIATIONAL IDENTIFICATION METHODS FOR LINEAR DYNAMIC SYSTEMS AND THE LOCAL EXTREMA PROBLEM

Andrei Lomov, Sobolev Institute of Mathematics of the Siberian Branch of RAS, Novosibirsk State University, Novosibirsk;
Cand.Sc., assistant professor (lomov@math.nsc.ru).

Abstract: The problem of a large number of local extrema is considered. This problem arises when using «direct» methods to identify parameters of linear dynamical systems with finite-sample observations. A new class of variational («indirect») parameter estimators is defined by the projectivity property of matrix kernels in the objective function. The variational objective functions are constructed having the number of local extrema not greater than the number of elements in system matrices. We obtain conditions for consistency of variational estimates in the limit of large number of observations of independent finite-length trajectories.

Keywords: parameter identification, difference equations, dynamic systems.

*Статья представлена к публикации
членом редакционной коллегии Н. Н. Бахтадзе*