М.: Мир, 1980. 536с.

11. *Wilkinson J.* The algebraic aingenvalue aroblem. Oxford: Oxford Uni. Press, 1965. 350p.

УДК 550.3:519.2

А.Ф.Кушнир

ИДЕНТИФИКАЦИЯ ЛИНЕЙНЫХ ДИНАМИЧЕСКИХ СИСТЕМ КАК СТАТИСТИЧЕСКАЯ ЗАДАЧА С МЕШАЮЩИМИ ПАРАМЕТРАМИ

A.F.Kushnir

IDENTIFICATION OF LINEAR DYNAMIC SYSTEMS AS A STATISTICAL PROBLEM WITH INCIDENTAL PARAMETERS

The problem of linear multidimensional system parameter estimation is discussed. Estimation is based on observations of input and output obscured by noise. In geophysics such problems arise in analysis of seismic response of media and structures using experimental data. The lower bounds of estimator variance are found and optimal estimation procedures are designed. Stability of estimator quality are examined, when statistical noise properties are varied.

Статистическая постановка задачи идентификации

Широкий круг задач обработки экспериментальных данных с помощью ЭВМ сводится к проблеме оценки импульсной переходной характеристики (ИПХ) H_τ некоторой линейной системы, связывающей входные и выходные данные (дискретизированные с целью обработки на ЭВМ) уравнением типа свертки:

$$u_t=\sum_{\tau=-\infty}^{\infty} H_\tau x_{t-\tau}, \quad t \in \mathbb{Z}, \quad \sum_{\tau=-\infty}^{\infty} \| H_\tau\| < \infty , \tag{1}$$

где $x_\tau \in R^k$ - входной сигнал линейной системы; $u_t \in R^n$ - ее выходной сигнал. Измерение процессов x_t и u_t практически всегда осуществляется на фоне помех ξ_t и η_t, которые естественно описывать, как случайные временные ряды. Поэтому определение (n x m)-матричной функции H_τ, $\tau \in \mathbb{Z}$ по выборкам наблюдений

$$y_t= x_t+ \xi_t, \quad z_t= u_t+ \eta_t, \quad t \in \overline{1,N} \tag{2}$$

представляет собой. статистическую задачу.

В частных постановках применительно к конкретным вопросам обработки геофизических экспериментальных данных эта задача исследовалась в [1-4]. Теоретические подходы к ее решению методами асимптотической статистики изучались в [5-9]. В настоящей работе определение функции H_τ рассматривается как задача статистического оценивания при наличии мешающих параметров, т.е. в том аспекте, который не был подробно освещен в предыдущих публикациях. При этом используется более широкая параметрическая модель наблюдений, учитывающая отсутствие полной априорной информации о входном сигнале x_t.

Во многих геофизических задачах (рассмотренных, например, в [2-4,10]) можно на основе априорной информации о H_τ построить ее параметрическую модель, т.е. подобрать такой класс функций $\{H_\tau(\theta), \theta \in \Theta\subset R^P\}$, зависящих от конечномерного параметра θ, что неизвестная ИПХ H_τ хорошо аппроксимируется одним из элементов этого класса $H_\tau(\theta_0)$ при конкретном (априори неизвестном) значении параметра θ_0. В этой ситуации задача оценивания ИПХ линейной системы (1) по наблюдениям (2) сводится к проблеме оценивания параметра θ совместной функции распределения $dP(\mathbb{V}_N,\theta)$ наблюдений

$$\mathbb{V}_N= (v_1^T,\dots, v_N^T)^T, \qquad v_t= (y_t^T, z_t^T)^T. \tag{3}$$

Для построения оптимальных статистических оценок параметров θ необходимы сведения о вероятностных характеристиках случайных помех ξ_t и η_t. В большинстве ситуаций достаточно реалистической моделью временных рядов ξ_t и η_t является описание их как гауссовских стационарных процессов с нулевыми средними, причем часто существенным оказывается влияние корреляционных связей между компонентами случайных величин ξ_t и η_t, $t \in \overline{1,N}$ при различных t. Предположение, что компоненты ξ_t и η_t независимы при всех t означает неоправданное упрощение задачи, которое приводит к оценкам ИПХ, далеким от наилучших. Более уместным является предположение, что векторный временной ряд $\zeta_t= (\xi_t^T,\eta_t^T)^T$ есть регулярный гауссовский стационарный процесс с матричной спектральной плотностью

$$F_\zeta(\lambda) = \begin{bmatrix} F_{\xi\xi}(\lambda), & F_{\xi\eta}(\lambda) \\ F_{\eta\xi}(\lambda), & F_{\eta\eta}(\lambda) \end{bmatrix}. \tag{4}$$

Информация о $F_\zeta(\lambda)$ может быть почерпнута из наблюдений помех ξ_t,η_t (например в интервалы времени, когда x_t отсутствует). Возможна также адаптивная постановка задачи идентификации, когда $F_\zeta(\lambda)$ оценивается по наблюдениям v_t, $t \in \overline{1,N}$ совместно с оцениванием параметров θ ИПХ линейной системы.

Объединяя (1)-(3) в одно соотношение, получим статистическую модель наблюдений, которую ниже будем называть моделью параметрической идентификации линейной системы:

$$v_t = G_t(\theta) * x_t + \zeta_t; \quad v_t, \zeta_t \in R^M, \quad x_t \in R^k, \quad M=k+n, \quad t \in \overline{1,N}, \tag{5}$$

где $G_t(\theta) = \begin{bmatrix} I \\ H_t(\theta) \end{bmatrix}$, $\zeta_t = \begin{pmatrix} \xi_t \\ \eta_t \end{pmatrix}$, $a_t * b_t = \sum_{\tau=-\infty}^{+\infty} a_\tau b_{t-\tau}$ - свертка двух последовательностей; ζ_t - гауссовский регулярный стационарный временной ряд с матричной спектральной плоскостью $F_\zeta(\lambda)$; $\theta \in \Theta$; $\Theta \subset R^p$ - ограниченное множество.

Методика оценивания параметров θ по наблюдениям (5) существенно зависит от априорной информации о входном сигнале x_t, регистрируемом на фоне помех ξ_t: достаточно ли ее для построения параметрической модели x_t или нет. Мы рассмотрим оба эти случая. В первом будем исследовать достаточно общую модель:

$$x_t = s_t(\rho) + \mu_{t,\rho}, \tag{6}$$

где $s_t(\rho)$ - детерминированная последовательность, зависящая от конечномерного мешающего параметра $\rho \in Q$; $Q \subset R^q$ - ограниченное множество; $\mu_{t,\rho}$ - гауссовский стационарный временной ряд с нулевым средним и матричной спектральной плотностью $F_\mu(\lambda,\rho)$, также зависящей от мешающего параметра ρ. Представление (6) отражает тот факт, что во многих практических ситуациях входной сигнал x_t "в среднем" известен: $Ex_t = s_t(\rho)$ (хотя и с точностью до неопределенных параметров ρ). В то же время возможны его случайные отклонения от среднего: $\mu_{t,\rho} = x_t - s_t(\rho)$, которые удобно полагать гауссовским стационарным процессом со спектром, известным с точностью до мешающих параметров ρ. Часто $\mu_{t,\rho}$ действительно есть некоторая случайная помеха, воздействующая на вход линейной системы одновременно с полезным сигналом $s_t(\rho)$. Естественно считать, что $E\,\zeta_t \mu_{t,\rho} = 0$.

В тех случаях, когда невозможно построить для x_t конечно-параметрическую модель, приходится рассматривать x_t как бесконечномерный мешающий параметр в модели идентификации (5). В этой ситуации определенные оптимальные свойства статистических процедур идентификации доказаны лишь в предположении, что x_t - эргодическая последовательность в том смысле, что у нее существует временная автоковариационная функция

$$R_\tau = \lim_{N\to\infty} R_{\tau,N}, \quad \text{где} \quad R_{\tau,N} = \frac{1}{N}\sum_{t=1}^{N} x_t x^T_{t+|\tau|}, \tag{7}$$

и временная матричная энергетическая спектральная плотность

$$F_x(\lambda) = \lim_{N\to\infty} F_{x,N}(\lambda), \quad \text{где} \quad F_{x,N}(\lambda) = \sum_{\tau=-N}^{N} R_{\tau,N} e^{i\lambda\tau}. \tag{8}$$

В настоящей статье мы ограничимся анализом задачи параметрической идентификации линейных систем при известной спектральной матричной плотности помех $F_\zeta(\lambda)$. В этой ситуации представляет интерес построение *оптимального* алгоритма идентификации, обеспечивающего оценку параметра θ линейной системы, наилучшую с точки зрения выбранного критерия качества. Однако нахождение оптимальной оценки не исчерпывает проблемы статистической идентификации. Важнейшим свойством любого алгоритма обработки данных является его устойчивость к предположениям о статистических характеристиках помех (робастность). Если при отклонении этих характеристик от предполагаемых (заложенных в оптимальный алгоритм идентификации) качество алгоритма резко ухудшается, то его практическая значимость сомнительна.

Тем не менее поиск оптимальной оценки важен в том отношении, что он связан с построением теоретической границы качества алгоритмов идентификации, которая достигается на оптимальной процедуре. Характеристики любой другой (неоптимальной) процедуры, обладающей каким-либо желаемым свойством, например, устойчивостью к изменению спектральной плотности помех $F_\zeta(\lambda)$, естественно сравнивать с этой границей, чтобы понять, ценой какого проигрыша в качестве это свойство приобретается.

Связь статистической задачи идентификации со стандартными проблемами статистического анализа данных

Введем дискретное (конечное) преобразование Фурье (ДПФ)

$$A_j = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} a_t \exp\{i\lambda_j t\}, \qquad a_t = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} A_j \exp\{-i\lambda_j t\}, \qquad \lambda_j = \frac{2\pi j}{N}.$$

Применяя ДПФ к обеим частям равенства (5), получим

$$V_j = G(\lambda_j,\theta)X_j + \zeta_j + O_j\left(\frac{1}{\sqrt{N}}\right); \quad G(\lambda,\theta) = \sum_{t=-\infty}^{+\infty} G_t(\theta)\exp\{i\lambda t\}, \tag{9}$$

где $j \in \overline{1,N}$, V_j, X_j, ζ_j – ДПФ от v_t, x_t и ζ_t соответственно, $|[O_j(\frac{1}{\sqrt{N}})]/\sqrt{N}\,| < K$. Наличие члена $O_j(\frac{1}{\sqrt{N}})$ в (9) связано с тем, что теорема о свертке для ДПФ верна лишь асимптотически (при конечных N верна теорема о циклической свертке [11]). Заметим, что V_j, $G(\lambda_j,\theta)$, X_j и ζ_j – комплексные величины, причем ζ_j – комплексные гауссовские случайные векторы с моментами [12]:

$$E\,\zeta_j = 0, \quad E\,\zeta_j\zeta_l^* = \delta_{jl}F_\zeta(\lambda_j) + O_{jl}\left(\tfrac{1}{N}\right); \quad \delta_{jl} = \begin{cases} 1, & j=l \\ 0, & j\neq l. \end{cases} \tag{10}$$

Отбрасывая в (9) и (10) члены $o_j(\frac{1}{\sqrt{N}})$ и $o_{jl}(\frac{1}{N})$, равномерно по $j,l\in\overline{1,N}$ стремящиеся к нулю с ростом размера выборки, получаем, что при больших N векторы V_j приближенно удовлетворяют стандартной модели нелинейной регрессии [13]:

$$V_j = G(\lambda_j,\theta)X_j + \zeta_j\ ,\quad E\,\zeta_j = 0\ ,\quad E\,\zeta_j\zeta_l^* = \delta_{jl}F_\zeta(\lambda_j)\,. \tag{11}$$

Однако применение для оценивания параметра θ в модели наблюдений (5) алгоритмов и программ классического регрессионного анализа затруднено рядом обстоятельств. Первое имеет теоретический характер и связано с пренебрежением членами $o_j(1/\sqrt{N})$ и $o_{jl}(1/N)$ при выводе оптимальных алгоритмов оценивания θ или анализе качества оценок $\tilde{\theta}(\mathbb{V}_N)$ на основе замены модели наблюдений (5) на модель наблюдений (11). Второе связано с тем, что матрицы ковариаций векторов ζ_j различны для разных j, в то время как в классическом регрессионном анализе они, как правило, предполагаются одинаковыми. И, наконец, третье обстоятельство - это комплексность входящих в (9), (10) величин, что делает невозможным применение большинства стандартных программ регрессионного анализа.

Если в параметрической модели входного сигнала (6) $s_t(\rho)\not\equiv 0$, то соотношение (11) можно переписать в виде

$$V_j = S_j(\rho)G(\lambda_j,\theta) + \nu_{j\theta\rho}\ , \tag{12}$$

где $E\ \nu_{j\rho\theta}= 0,\quad E\ \nu_{j\rho\theta}\nu^*_{l\rho\theta} = \delta_{jl}\big[F_\zeta(\lambda_j)+G(\lambda_j,\theta)F_\mu(\lambda_j,\rho)G^*(\lambda_j,\theta)\big].$

Выражение (12) соответствует "F-модели" нелинейной регрессии, в которой информация о параметрах задачи (θ,ρ) заключена и в среднем значении, и в ковариациях наблюдений V_j. Класс ИРДЖИНА-оценок для этой модели исследован в [14]. Однако непосредственное использование результатов работы [14] для рассматриваемой задачи идентификации затруднено ввиду отмечавшихся выше ее особенностей.

При $s_t(\rho)\equiv 0$, т.е. чисто случайном входном сигнале, соотношение (11) имеет вид

$$V_j = G(\lambda_j,\theta)\mu_{j\rho}+ \zeta_j,\qquad E\ V_j= 0, \tag{13}$$

$$E\ V_jV_l^* = \delta_{jl}\big[F_\zeta(\lambda_j) + G\ (\lambda_j,\rho)F_\mu(\lambda_j,\rho)G^*(\lambda_j,\theta)\big],$$

т.е. проблема идентификации соответствует модели факторного анализа (модели "структурного соотношения" [15]), в которой информация о параметрах заключена только в ковариационной матрице наблюдений.

И наконец, когда x_t – "полностью" неизвестная последовательность, (11) представляет собой обобщение модели "функционального соотношения". Эта модель подробно исследована в работе [15], однако в более простой ситуации, когда ковариации ζ_j не зависят от j и $G(\lambda_j,\theta)\equiv G$, где G – неизвестная матрица параметров.

Из сказанного следует, что статистическая задача идентификации (1),(2), приведенная к форме модели наблюдений (5), хотя и имеет аналогии со стандартными постановками регрессионного и факторного анализа, тем не менее отличается от них рядом специфических особенностей и представляет собой самостоятельную проблему статистического анализа многомерных временных рядов, не сводимую непосредственно к классическим схемам многомерной статистики.

Статистические критерии качества алгоритмов параметрической идентификации

Задача построения оптимальных оценок параметра θ в моделях наблюдений (5),(6) и (5),(8) будет рассматриваться при определенных ограничениях, типа требований "гладкости" функций $G(\lambda,\theta)$, $F_\mu(\lambda,\rho)$, $F_\zeta(\lambda)$, $s_t(\rho)$ по λ,θ и ρ, а также "эргодичности" последовательностей $s_t(\rho)$ и x_t, $t\in\mathbb{Z}$. Сформулируем их, следуя работам [5-8].

Обозначим $\mathbb{C}([0,2\pi]\times\Theta)$, где Θ – компакт в R^p, пространство непрерывных функций $f(\lambda,\theta)$ с равномерной метрикой

$$\| f_1(\lambda,\theta) - f_2(\lambda,\theta) \| = \sup_{\substack{\lambda\in[0,2\pi]\\ \theta\in\Theta}} |f_1(\lambda,\theta) - f_2(\lambda,\theta)| ,$$

и будем говорить, что для $f(\lambda,\theta)$ существует производная по θ_k в $\mathbb{C}([0,2\pi)\times\Theta)$: $\dot f_k(\lambda,\theta) = \frac{\bar\partial}{\partial\theta_k} f(\lambda,\theta)$, если

$$\lim_{\mu\to 0} \sup_{\substack{\lambda\in[0,2\pi)\\ \theta\in\Theta}} |[f(\lambda,\theta+e_k\mu) - f(\lambda,\theta)]/\mu - \dot f_k(\lambda,\theta)| = 0,$$

где e_k – вектор, имеющий единицу на k-м месте и нули – на остальных. Напомним, что функция $f(\lambda)$ удовлетворяет условию Гельдера с показателем $\alpha\le 1$, если $|f(\lambda_1)-f(\lambda_2)| < c\,|\lambda_1-\lambda_2|^\alpha$ (при $\alpha = 1$ это условие Липшица).

Будем говорить, что для моделей наблюдений (5),(6) и (5),(8) справедливы ограничения B, если

BI. Детерминированная составляющая $s_t(\rho)$, $t\in\mathbb{Z}$ в модели (6) входного сигнала x_t равномерно по $\rho\in Q$ обладает следующими свойствами:

1) $s_t(\rho)$ имеет временную энергетическую спектральную плотность $T_s(\lambda,\rho)$, $\lambda\in[0,2\pi)$, определяемую формулами (7),(8).

2) $s_t(\rho)$ не имеет "слишком больших выбросов", т.е.

$$\max_{1<t<N} |s_t(\rho)|^2 < c\, N^{\beta}, \quad \beta\in [0,1).$$

3) $s_t(\rho)$ при каждом $t\in \mathbb{Z}$ имеет частные производные

$$\dot{s}_{tk}(\rho) = \frac{\overline{\partial}}{\partial\rho_k} s_t(\rho), \qquad k \in \overline{1,r},$$

причем последовательности $s_{tk}(\rho)$, $t\in\mathbb{Z}$, удовлетворяют ограничениям BI.1 и BI.2 и следующему условию гладкости по ρ:

$$\lim_{N\to\infty} (1/N) \sum_{t=1}^{N} |s_{tk}(\rho_N+r_N/\sqrt{N})-s_{tk}(\rho_N)|^2 = 0$$

для любых последовательностей ρ_N, $r_N\in Q$. Это условие выполняется, если существуют вторые производные $\ddot{s}_{tkl}(\rho) = \frac{\overline{\partial}}{\partial\rho} \dot{s}_{tk}(\rho)$, удовлетворяющие условиям BI.1 и BI.2.

BII. Частотная характеристика линейной системы $G(\lambda,\theta)$ при всех $\theta \in \Theta$ и $\lambda \in [0,2\pi)$ удовлетворяет следующим условиям:

1) существуют (односторонние) производные $\frac{\partial^2}{\partial\lambda^2} G(\lambda,\theta)$;

2) матрицы $G(\lambda,\theta)$ дважды дифференцируемы по θ_k, $k\in \overline{1,p}$ в $\mathbb{C}([0,2\pi) \times \Theta)$, т.е. существуют матрицы

$$\dot{G}_k(\lambda,\theta) = \frac{\overline{\partial}}{\partial\theta_k} G(\lambda,\theta) \quad \text{и} \quad \ddot{G}_{kl}(\lambda,\theta) = \frac{\overline{\partial}}{\partial\theta_l} \dot{G}_k(\lambda,\theta).$$

3) элементы матриц $\dot{G}_k(\lambda,\theta)$ и $\ddot{G}_{kl}(\lambda,\theta)$ как функции $\lambda\in [0,2\pi)$ равномерно по $\theta\in\Theta$ удовлетворяют условию Гельдера с $\alpha \ge 1/2 + \beta/2$ и $\alpha \ge 1/2$ соответственно.

BIII. Матричная энергетическая спектральная плотность $F_{\mu}(\lambda,\rho)$ случайной составляющей входного сигнала $\mu_{t\rho}$ при всех $\rho \in Q$ и $\lambda \in [0,2\pi)$ удовлетворяет условиям BII.

BIV. Матричная энергетическая спектральная плотность помех $F_{\zeta}(\lambda)$ невырождена и

1) $$\inf_{\lambda\in[0,2\pi)} \det F_{\zeta}(\lambda) = \varepsilon > 0.$$

2) ее элементы удовлетворяют условию Липшица.

BV. В случае, когда входной сигнал линейной системы x_t неизвестен, т.е.не подчиняется модели (6), он должен удовлетворять ограничениям BI,1 и BI,2, а его временная матричная спектральная плотность $T_x(\lambda)$ - соответствовать условиям BIV.

Важнейшую роль при построении оптимальных оценок $\tilde{\theta}_N = \tilde{\theta}(\mathbb{V}_N)$ играет критерий качества. Мы будем использовать как меру точности матрицу среднеквадратического рассеяния оценки относительно истинного значения параметра θ:

$$U_{\tilde{\theta}_N} = E\,(\tilde{\theta}_N - \theta)\,(\tilde{\theta}_N - \theta)^T. \qquad (14)$$

Естественно применять на практике только состоятельные оценки, такие, что

$$U_{\tilde{\theta}_N} \to 0 \quad \text{при} \quad N \to \infty. \qquad (15)$$

Можно показать, что при сделанных выше предположениях о статистических характеристиках сигналов и помех в модели наблюдений (5) существует целый класс $\sqrt{N}$-состоятельных оценок, для которых

$$U_{\tilde{\theta}_N} = \frac{C_\theta}{N} + o\,\left(\frac{1}{N}\right), \qquad (16)$$

где C_θ - неотрицательно определенная матрица:

$$C_\theta = \lim_{N\to\infty} N\,E\,(\tilde{\theta}_N - \theta)\,(\tilde{\theta}_N - \theta)^T. \qquad (17)$$

Будем называть C_θ асимптотической ковариацией оценки $\tilde{\theta}(\mathbb{Y}_N)$.

Если x_t удовлетворяет модели (6), то распределение наблюдений (5) зависит от конечномерного мешающего параметра ρ, и матрица $U_{\tilde{\theta}_N}$ есть функция истинных значений информативного параметра θ и мешающего параметра ρ : $U_{\tilde{\theta}_N} = U_{\tilde{\theta}_N}(\theta,\rho)$. Если x_t - произвольная последовательность, обладающая свойством (7), то $U_{\tilde{\theta}_N} = U_{\tilde{\theta}_N}(\theta,\{x_t\})$. Мы будем сохранять обозначение $U_{\tilde{\theta}_N}(\theta,\rho)$, рассматривая $\rho=\{x_t\}$ как "бесконечномерный" вектор. Класс оценок $\tilde{\theta}(\mathbb{Y}_N)$ параметра θ, удовлетворяющих условиям (14)-(17) равномерно по $\theta \in \Theta$, $\rho \in Q$, будем обозначать $\mathbb{K}$.

Для сравнения различных оценок между собой удобнее иметь *числовую* меру их точности. Для этого будем использовать асимптотический риск оценки $\tilde{\theta}(\mathbb{Y}_N)$:

$$R_{\tilde{\theta}_N}(\theta,\rho) = r\,(NU_{\tilde{\theta}_N}(\theta,\rho)), \qquad (18)$$

где $r(C)$ - функция риска из класса $\mathbb{R}$ неотрицательных непрерывных выпуклых функций от матрицы, имеющих степенной рост при $|C| \to \infty$. Например, часто используются такие функции риска, как: $r_1(C)=\mathrm{tr}C$ - сумма асимптотических среднеквадратических отклонений компонент оценки $\tilde{\theta}_{N,l}$, $l\in\overline{1,p}$ от истинных значений θ_l; $r_2(C) = a^TC\,a$ - асимптотическая дисперсия оценки какой-либо важной линейной комбинации $\upsilon = a^T\theta$ параметров θ_1; $r_3(\tilde{C}) = \det C$ - объем эллипсоида асимптотического рассеяния оценки $\tilde{\theta}_N$, и т.п. В терминах функции асимптотического риска $r(NU_{\tilde{\theta}}(\theta,\rho))$ можно формулировать и решать проблему построения наилучшей - асимптотически эффективной - оценки $\hat{\theta}(\mathbb{V}_N)$ параметров линейной системы в классе $\mathbb{K}$ оценок, удовлетворяющих условиям (14) - (17).

Асимптотически эффективные оценки для конечно-параметрической модели входного сигнала

Пусть входной сигнал x_t удовлетворяет конечно-параметрической модели (6) и выполняются ограничения BI-BIV. В этом случае из результатов [5 - 7,15] следует, что асимптотический риск произвольной оценки $\tilde{\theta}(\mathbb{V}_N)$ из класса $\mathbb{K}$ при всех $\delta>0$ удовлетворяет неравенству

$$\lim_{N\to\infty} \sup_{\substack{|\theta-\theta_0|<\delta \\ |\rho-\rho_0|<\delta}} r(NU_{\tilde{\theta}_N}(\theta,\rho)) \ge r(L^{-1}(\theta_0,\rho_0)), \tag{19}$$

где $L(\theta,\rho) = \Gamma_{\theta\theta}(\theta,\rho) - \Gamma_{\theta\rho}(\theta,\rho)\Gamma^{-1}_{\rho\rho}(\theta,\rho)\Gamma_{\rho\theta}(\theta,\rho)$.

Элементы матриц $\Gamma_{\alpha\beta}$, $\alpha,\beta\in\{\theta,\rho\}$ определяются формулами

$$\Gamma^{(kl)}_{\alpha\beta}(\theta,\rho) = \Phi^{(kl)}_{\alpha\beta}(\theta,\rho) + \Psi^{(kl)}_{\alpha\beta}(\theta,\rho),\ k,l\in\overline{1,q_\alpha},\ q_\alpha=p,\ q_\beta=q;$$

$$\Phi^{(kl)}_{\alpha\beta} = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{tr}F^{-1}(\lambda)\dot{Q}^{\alpha\beta}_{kl}(\lambda)d\lambda;$$

$$\Psi^{(kl)}_{\alpha\beta} = \frac{1}{8\pi}\int_0^{2\pi} \mathrm{tr}F^{-1}(\lambda)\dot{F}^{\alpha}_{k}(\lambda)F^{-1}(\lambda)\dot{F}_l^{\beta}(\lambda)F^{-1}(\lambda)d\lambda;$$

$$F(\lambda) = F_\zeta(\lambda)+G(\lambda,\rho)F_\mu(\lambda,\rho)G^*(\lambda,\rho);\quad \dot{F}^{\theta}_k = \frac{\bar{\partial}}{\partial\theta_k}F,\quad \dot{F}^{\rho}_l = \frac{\bar{\partial}}{\partial\rho_l}F; \tag{20}$$

$$\dot{Q}^{\theta\theta}_{kl} = \dot{G}_k(\lambda,\theta)T_{ss}(\lambda,\rho)\dot{G}^*_l(\lambda,\theta);\quad \dot{G}_k = \frac{\bar{\partial}}{\partial\theta_k}G;\quad k,l\in\overline{1,p};$$

$$\dot{Q}^{\rho\rho}_{kl} = G(\lambda,\theta)T_{\dot{s}_k\dot{s}_l}(\lambda,\rho)G^*(\lambda,\theta);\quad k,l\in\overline{1,q};$$

$$\dot{Q}^{\theta\rho}_{kl} = \dot{G}_k(\lambda,\theta)T_{s\dot{s}_l}(\lambda,\rho)G^*(\lambda,\theta);\quad k\in\overline{1,p},\quad l\in\overline{1,q};$$

T_{ss} - временная энергетическая спектральная плотность (ВЭСП) детерминированной составляющей $s_t(\rho)$ сигнала; $T_{\dot{s}_k \dot{s}_l}$ – взаимная ВЭСП $\dot{s}_{tl}(\rho)$ и $\dot{s}_{tk}(\rho)$, где $\dot{s}_{tk}(\rho) = \frac{\partial}{\partial \rho_k} s_t(\rho)$; $T_{s\dot{s}_k}$ – взаимная ВЭСП $s_t(\rho)$ и $\dot{s}_{tk}(\rho)$. Отметим, что блочная матрица

$$\Gamma = \begin{bmatrix} \Gamma_{\theta\theta}\Gamma_{\theta\rho} \\ \Gamma_{\rho\theta}\Gamma_{\rho\rho} \end{bmatrix}$$

есть матрица Фишера распределения наблюдений V_j модели (11) по отношению к объединенному вектору параметров задачи $\theta = (\theta^T, \rho^T)^T$.

Оценку $\hat{\theta}(\mathbb{V}_N)$ называют асимптотически эффективной (АЭ) в классе оценок $\mathbb{K}$, если при $\delta \to 0$ для нее в соотношении (19) при любых функциях потерь $r(C) \in \mathbb{R}$ достигается знак равенства.

Из [5 - 7,15] следует, что АЭ-оценка $\hat{\theta}(\mathbb{V}_N)$ может быть найдена (совместно с АЭ-оценкой $\hat{\rho}(\mathbb{V}_N)$ - мешающего параметра) как решение следующей системы, вообще говоря, нелинейных уравнений:

$$\Delta(\mathbb{V}_N, \theta, \rho) = 0, \qquad \theta \in \Theta, \qquad \rho \in Q, \tag{21}$$

где вектор-функция $\Delta(\mathbb{V}_N, \theta, \rho) = (\Delta_1^\theta, \dots, \Delta_p^\theta, \Delta_1^\rho, \dots, \Delta_q^\rho)$ выражается следующими формулами:

$$\begin{gathered} \Delta_k^\alpha(\mathbb{V}_N, \theta, \rho) = \varphi_k^\alpha(\mathbb{V}_N, \theta, \rho) + \psi_k^a(\mathbb{V}_N, \theta, \rho), \\ k \in \overline{1, q_\alpha}, \quad \alpha \in \{\theta, \rho\}, \quad q_\alpha = p, \quad q_\beta = q. \end{gathered} \tag{22}$$

где $\quad \varphi_k^\alpha(\mathbb{V}_N, \theta, \rho) = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} [(V_j - U_j)^* F_j^{-1} \dot{U}_{kj}^\alpha]$;

$$\psi_k^\alpha = \frac{1}{2\sqrt{N}} \sum_{j=1}^{N} [(V_j - U_j)^* F_j^{-1} \dot{F}_{kj}^\alpha F_j^{-1} (V_j - U_j) - \mathrm{tr} \dot{F}_{kj} F_j^{-1}];$$

$$U_j = G(\lambda_j, \theta) S_j(\rho); \quad F_j = F(\lambda_j, \theta, \rho);$$

$$\dot{U}_{kj}^\theta = \frac{\partial}{\partial \theta_k} U_j; \quad \dot{U}_{kj}^\rho = \frac{\partial}{\partial \rho_k} U_j; \quad \lambda_j = \frac{2\pi j}{N};$$

V_j, S_j – ДПФ от v_t и $s_t(\rho)$ соответственно.

При ограничениях BI-BIV АЭ-оценка оказывается асимптотически нормальной с параметрами $(\theta, N^{-1}L^{-1}(\theta_0, \rho_0))$, т.е. ее асимптотическая ковариация равна

$$C_{\hat{\theta}}(\theta_0,\rho_0) = L^{-1}(\theta_0,\rho_0). \tag{23}$$

Система уравнений (21) в случае нелинейной зависимости частотной характеристики $G(\lambda,\theta)$ от параметра θ должна решаться итерационными процедурами. В частности, в [6,7] доказана сходимость следующей процедуры (при условии, что начальное приближение достаточно близко к истинному значению параметра):

$$\hat{\theta}^{(n+1)} = \hat{\theta}^{(n)} + \frac{1}{\sqrt{N}}\,\Gamma^{-1}(\hat{\theta}^{(n)})\;\Delta\,(\mathbb{V}_N,\hat{\theta}^{(n)}_N),\quad n = 0,1,\ldots \tag{24}$$

где $\quad \hat{\theta}^{(n)} = (\hat{\theta}^{(n)T},\hat{\rho}^{(n)T})^T, \qquad \Gamma(\theta) = \begin{bmatrix} \Gamma_{\theta\theta} & \Gamma_{\theta\rho} \\ \Gamma_{\rho\theta} & \Gamma_{\rho\rho} \end{bmatrix},$

блоки $\Gamma_{\alpha\beta}(\theta)$ выражаются формулами (20).

Рассмотрим подробнее два частных случая конечно-параметрической модели входного сигнала (6): а) x_t – квазидетерминированный процесс: $x_t = s_t(\rho)$, $(\mu_t \equiv 0)$, и б) x_t – регулярный случайный стационарный гауссовский процесс: $x_t = \mu_{t,\rho}$, $(s_t \equiv 0)$.

а) **Квазидетерминированный входной сигнал.** В этом случае система уравнений (21) имеет вид

$$\Delta_k^\alpha(\mathbb{V}_N,\theta,\rho) = \varphi_k^\alpha(\mathbb{V}_N,\theta,\rho) = 0,\quad \alpha\in\{\theta,\rho\},\ k\in\overline{1,q_\alpha},\ q_\theta = p,\ q_\rho = q, \tag{25}$$

а матрица Γ в итерационной процедуре (24) равна

$$\Gamma = \begin{bmatrix} \Phi_{\theta\theta} & \Phi_{\theta\rho} \\ \Phi_{\rho\theta} & \Phi_{\rho\rho} \end{bmatrix}.$$

Отметим, что матричная функция $F(\lambda,\theta,\rho)$, фигурирующая в выражениях (20) и (22) в данном случае равна $F_\zeta(\lambda)$, т.е. не зависит от θ,ρ.

Значительный интерес в рассматриваемом случае представляет ситуация, когда помехи ξ_t и η_t на входе и выходе линейной системы порождаются одним источником, так что они когерентны и матрица

$$F_\zeta = \begin{bmatrix} F_{\xi\xi} & F_{\xi\eta} \\ F_{\eta\xi} & F_{\eta\eta} \end{bmatrix}$$

вырождена. При этом непосредственно применять формулы (20)-(24) с $F(\lambda,\theta,\rho) = F_\zeta(\lambda)$ невозможно. В [7] показано, что оптимальный алгоритм идентификации в этой ситуации сводится к решению нелинейной системы уравнений:

$$\varphi_k^\alpha(\mathbb{V}_N,\theta,\rho) = \frac{1}{\sqrt{N}}\sum_{j=1}^{N}(V_j - U_j)^* B_j \dot{U}_{kj}^\alpha = 0;\ \alpha \in \{\theta,\rho\},\ k \in \overline{1,q_\alpha}, \quad (26)$$

где $B = [I - Q(Q^*Q)^{-1}Q^*]$; $Q = [q_1(\lambda),\dots,q_r(\lambda)]$; r - ранг спектральной матрицы помех $F_\zeta(\lambda)$; $q_l(\lambda)$; $l \in \overline{1.r}$ - собственные векторы этой матрицы, соответствующие ненулевым собственным значениям. Интересным свойством оценки $\theta_N^* = (\theta_N^*, \rho_N^*)^T$ - решения системы уравнений (26) при когерентных помехах ξ_t, η_t, является ее своеобразная "суперэффективность". В отличие от регулярного случая, при когерентных помехах для матрицы среднеквадратического отклонения оценки θ_N^* вместо (16) справедливо соотношение [7]:

$$U_{\theta_N^*} = O\left(\frac{1}{N^\nu}\right),$$

где $\nu = 1 + \min[1-\beta, 2\alpha-(1+\beta)]$, $0<\beta<1$, $1/2+\beta/2 < \alpha < 1$;

α и β - константы, фигурирующие в условиях BI и BII. Следовательно, $\theta^*(\mathbb{V}_N)$ стремится к точному значению параметров θ быстрее обычного. Это объясняется "взаимной компенсацией помех" при вычислении функционала $\varphi_K^\alpha(\mathbb{V}_N,\theta,\rho)$, содержащего матрицу B, поскольку B есть проектор на ортогональное дополнение к подпространству $\mathbb{L} \subset R^M$, в котором действуют когерентные помехи ζ_t.

б) **Случайный регулярный входной сигнал.** В противоположном случае, когда $x_t = \mu_{t\rho}$, т.е. входной сигнал - "чисто" случайный, уравнения для АЭ-оценки имеют вид

$$\Delta_k^\alpha(\mathbb{V}_N,\theta,\rho) = \psi_k^\alpha(\mathbb{V}_N,\theta,\rho) = 0, \quad \alpha \in \{\theta, q\}, \quad k \in \overline{1,q_\alpha}, \quad (27)$$

где ψ_k^α выражаются формулами (22), в которых $U_j = 0$. Матрица Γ в итерационной процедуре (24) в этом случае равна

$$\Gamma = \begin{bmatrix} \Psi_{\theta\theta}\Psi_{\theta\rho} \\ \Psi_{\rho\theta}\Psi_{\rho\rho} \end{bmatrix}.$$

При случайном входном сигнале $x_t = \mu_{t\rho}$ и когерентных помехах ξ_t и η_t матрица $F(\lambda,\theta,\rho)$ в (20), вообще говоря, невырождена, и уже невозможно построить "суперэффективные" оценки. Это объясняется тем, что модель сигнала x_t в виде случайного регулярного стационарного процесса содержит гораздо меньше априорной информации, чем модель в виде квазидетерминированного процесса $s_t(\rho)$ (допускающего безошибочный прогноз по наблюдениям конечной длительности).

Оптимальные алгоритмы идентификации при неизвестном эргодическом входном сигнале

Рассмотрим теперь случай, когда о входном сигнале x_t известно лишь то, что он имеет временную матричную энергетическую спектральную плотность $T_x(\lambda)$ (8) (которая также априори неизвестна). При этом согласно (1)-(5) любая конечная выборка наблюдений $\mathbb{V}_N$ формально зависит от бесконечномерного мешающего параметра - последовательности x_t, $t \in \mathbb{Z}$, и может показаться, что $\mathbb{V}_N$ не содержит достаточно информации для состоятельного оценивания параметра θ линейной системы, т.е. построение оценок $\tilde{\theta}(\mathbb{V}_N)$, удовлетворяющих условиям (14)-(17) невозможно. Однако $\| H_\tau \| \to 0$ при $\tau \to \pm \infty$, так что, по существу, можно в свертке (1) ограничиться конечным числом L членов. Следовательно, число "существенных" мешающих параметров в (5) равно $k(N+L)$, в то время как общее число наблюдений в выборке $\mathbb{V}_N$ равно MN, где M - размерность векторов v_t. Поскольку $M > k$, и $MN-k(N+L) \to \infty$ при $N \to \infty$, то общее число наблюдений асимптотически неограниченно превышает число "существенных" мешающих параметров, и, в принципе, возможно состоятельное оценивание информативных параметров θ линейной системы.

Важным примером $\sqrt{N}$-состоятельной и асимптотически нормальной оценки параметра θ при неизвестном входном сигнале x_t является оценка $\bar{\theta}(\mathbb{V}_N)$ по методу максимума правдоподобия. Формально она строится для модели (11) наблюдений V_j $j \in \overline{1,N}$ в спектральной области и находится совместной максимизацией функции правдоподобия этих наблюдений по параметрам θ и X_j. В результате оценка $\bar{\theta}(\mathbb{V}_N)$ является решением следующей системы нелинейных уравнений [4]:

$$\delta_k(\mathbb{V}_N, \theta) = \frac{\partial}{\partial \theta_k} \sum_{j=1}^{N} |(I - A_j(\theta)) W_j|^2 = 0, \quad k \in \overline{1,p}, \tag{28}$$

где $A_j(\theta) = D_j(\theta)[D_j(\theta) D_j^*(\theta)]^{-1} D_j^*(\theta)$, $W_j = F_\xi^{-1/2}(\lambda_j)\, V_j$,

$$D_j(\theta) = G(\lambda_j, \theta) F_\xi^{-1/2}(\lambda_j).$$

Можно показать [4,7,10], что $\bar{\theta}(\mathbb{V}_N)$ является $\sqrt{N}$-состоятельной и асимптотически нормальной оценкой параметра θ также и для модели наблюдений (5), (8), если x_t удовлетворяет условиям BV, а $G_t(\theta)$, $F(\lambda)$ - условиям BII, BIV. При этом ее асимптотическая ковариация равна

$$\mathbb{C}_{\bar{\theta}_N}(\theta) = K^{-1}(\theta) + K^{-1}(\theta)\Pi(\theta)\, K^{-1}(\theta), \tag{29}$$

где матрицы $K(\theta)$ и $\Pi(\theta)$ определяются формулами

$$K^{(kl)}(\theta) = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{tr}\dot{D}_k^*(\lambda,\theta)[I-A(\lambda,\theta)]\dot{D}_l(\lambda,\theta)T_x(\lambda)d\lambda,$$

$$\Pi^{(kl)}(\theta) = \frac{1}{\pi}\int_0^{2\pi} \mathrm{tr}\dot{A}_k^*(\lambda,\theta)\dot{A}_l(\lambda,\theta)d\lambda,$$

$$\dot{D}_k = \frac{\partial}{\partial\theta_k}D(\lambda,\theta), \quad \dot{A}_k = \frac{\partial}{\partial\theta_k}A(\lambda,\theta), \quad k,l \in \overline{1,p};$$

$T_x(\lambda)$ - временная матричная спектральная плотность последовательности x_t.

Однако МП-оценка $\bar{\theta}(V_N)$, вообще говоря, не является асимптотически наилучшей в задаче идентификации (5) при неизвестном эргодическом входном сигнале x_t. Для построения асимптотически эффективной оценки параметра θ необходимо найти нижнюю границу риска типа (19) для произвольных оценок из класса $\mathbb{K}$. Обобщая результаты работы [15], можно при ограничениях BII, BIV и BV получить следующее неравенство:

$$\begin{aligned} &\lim_{N\to\infty} \sup_{\substack{|\theta-\theta_0|<\delta \\ (\{x_t\}:T_{Nx}(\lambda)\in S_\delta)}} r[E_{\theta,\{x_t\}}N(\tilde{\theta}_N-\theta)(\tilde{\theta}_N-\theta)^T] \ge \\ &\ge \lim_{N\to\infty} \sup_{\substack{|\theta-\theta_0|<\delta \\ (T_x(\lambda)\in S_\delta)}} r[M_{\theta,T_x(\lambda)}N(\tilde{\theta}_N-\theta)(\tilde{\theta}_N-\theta)^T], \end{aligned} \tag{30}$$

где $E_{\theta,\{x_t\}}$ - усреднение по гауссовской мере в пространстве наблюдений v_t, $t\in\overline{1,N}$, зависящей от всех неизвестных параметров θ, $\{x_t\}$ модели (5); $M_{\theta,T(\lambda)}$ - усреднение по гауссовской мере в том же пространстве, в предположении, что x_t - регулярный гауссовский случайный процесс с матричной спектральной плотностью $T_x(\lambda)$; $S_\delta = \{T_x(\lambda): \| T_x(\lambda)-T_x^0(\lambda) \| <\delta \}$ - сфера в пространстве матричных функций от $\lambda\in [0,2\pi]$ с метрикой

$$\| T_1(\lambda)-T_2(\lambda) \| = \sup_{\lambda\in[0,2\pi]} | T_1(\lambda)-T_2(\lambda) |.$$

$|T(\lambda)|$ - евклидова норма матрицы $T(\lambda)$; $T_{Nx}(\lambda)$, $T_x^0(\lambda)$ определяются формулами (7), (8) для последовательностей $\{x_t\}$ и $\{x_t^0\}$, соответственно θ_0, $\{x_t^0\}$ - действующие значения информативного и мешающего параметров.

Неравенство (30) можно интерпретировать следующим образом: дополнительная информация о том, что x_t - гауссовский стационарный процесс, может лишь уменьшить статистический разброс оценки $\tilde{\theta}(\mathbb{V}_N)\in\mathbb{K}$. В настоящее время автору в деталях неизвестно, как строить АЭ-оценку в классе $\mathbb{K}$, для которой при ограничениях BII, BIV и BV в неравенстве (30) при $\delta\to 0$ достигался бы знак равенства. Однако, если сузить множество возможных временных спектральных плотностей сигнала x_t, такую оценку можно найти. Введем дополнительное ограничение BVI:

BVI. 1) Временная матричная спектральная плотность сигнала x_t принадлежит известному конечно-параметрическому множеству комплекснозначных эрмитовых матричных функций: $T_x(\lambda)=T(\lambda,\rho)$, $\rho\in Q\subset R^q$, где Q - ограниченное множество.

2) Функции $T(\lambda,\rho)$ удовлетворяют условиям BII.

При ограничениях BVI неравенство (30) может быть записано в виде

$$\lim_{N\to\infty}\sup_{\substack{|\theta-\theta_0|<\delta\\ (\{x_t\}:T_{NX}(\lambda)\in S_\delta(\rho_0))}} r[E_{\theta,\{x_t\}}N(\tilde{\theta}_N-\theta)(\tilde{\theta}_N-\theta)^T]\geq$$

$$\geq\lim_{N\to\infty}\sup_{\substack{|\theta-\theta_0|<\delta\\ |\rho-\rho_0|<\delta}} r[M_{\theta,\rho}N(\tilde{\theta}_N-\theta)(\tilde{\theta}_N-\theta)^T]\geq r[L^{-1}\theta_0,\rho_0)], \quad (31)$$

где $S_\delta(\rho_0)=\{T(\lambda):\ \|T(\lambda)-T(\lambda,\rho_0)\|<\delta\ \}$; $M_{\theta,\rho}$- усреднение по гауссовской мере наблюдений v_t, $t\in\overline{1,N}$, в предположении, что x_t - гауссовский стационарный процесс со спектральной плотностью $T(\lambda,\rho)$, $L(\theta,\rho)$ определяется формулами (19) и (20) при $\Gamma_{\alpha\beta}(\theta,\rho)=\Psi_{\alpha\beta}(\theta,\rho)$ и $F_\mu(\lambda,\rho)=T(\lambda,\rho)$. Второе неравенство в (31) есть следствие неравенства (19), так как предположение, что x_t- гауссовский случайный процесс с параметрически заданной спектральной плотностью, - частный случай модели наблюдений (5),(6).

Рассмотрим АЭ-оценку $\hat{\theta}(\mathbb{V}_N)$ параметра $\theta=(\theta^T\rho^T)^T$ для модели (5) при чисто случайном входном сигнале $x_t=\mu_{t,\rho}$ и $F_\mu(\lambda,\rho)=T(\lambda,r)$. Как сказано выше, она является решением системы уравнений (27). Наше утверждение заключается в том, что $\hat{\theta}(\mathbb{V}_N)$ - часть оценки $\hat{\theta}(\mathbb{V}_N)=(\hat{\theta}(\mathbb{V}_N)^T,\ \hat{\rho}(\mathbb{V}_N)^T)^T$, является АЭ-оценкой параметра θ в задаче идентификации (5), (8) при неизвестном входном сигнале x_t, удовлетворяющем ограничениям BV, BVI.

Чтобы показать это, вычислим асимптотическую ковариационную матрицу оценки $\hat{\theta}(\mathbb{V}_N)$ в предположении, что x_t- детерминированный временной ряд, подчиняющийся ограничениям BV, BVI. Эта матрица может быть найдена по следующим формулам [7, теорема 1]:

$$\lim_{N\to\infty} E_{\theta,\{x_t\}} N(\hat\theta_N - \theta)(\hat\theta_N - \theta)^T = W^{-1}(\theta)U(\theta)W^{*-1}(\theta), \qquad (32)$$

где $\quad U(\theta) = \lim_{N\to\infty} E_{\theta,\{x_t(\rho)\}} \psi^T(\mathbb{V}_N,\theta)\psi(\mathbb{V}_N,\theta),$

векторная статистика $\psi(\mathbb{V}_N,\theta)$ состоит из компонент $\psi_k^\alpha(\mathbb{V}_N,\theta)$, $\alpha\in\{\theta,\rho\}$, $k\in\overline{1,q_\alpha}$, $q_\theta=p$, $q_\rho=q$, которые определяются формулами (22).

$$W(\theta) = \lim_{N\to\infty} E_{\theta,\{x_t(\rho)\}} (1/\sqrt{N})\; D(\mathbb{V}_N,\theta),$$

блочная матрица $D = \begin{bmatrix} D_{\theta\theta} & D_{\theta\rho} \\ D_{\rho\theta} & D_{\rho\rho} \end{bmatrix}$ имеет следующие элементы:

$$D_{\alpha\beta}^{(jk)}(\mathbb{V}_N,\theta) = \frac{\partial}{\partial\beta_j}\,\psi_k^\alpha(\mathbb{V}_N,\theta),\; \alpha,\beta\in\{\theta,\rho\},\; j,k\in\overline{1,q_\alpha}.$$

Рассмотрим интегральные суммы вида

$$(1/N)\sum_{j=1}^{N} A(\lambda_j,\theta)X_jX_j^* = (1/2\pi)\int_0^{2\pi} A(\lambda,\theta)dK_{xN}(\lambda,\rho), \qquad (33)$$

где $A(\lambda,\theta)$ - матричные функции, удовлетворяющие условиям BII: X_j, $j\in\overline{1,N}$, - ДПФ от x_t, $t\in\overline{1,N}$:

$$K_{xN}(\lambda,\rho) = (1/N)\sum_{j=1}^{\nu(\lambda)} X_jX_j^* \quad -$$

оценка матричного временного кумулятивного спектра сигнала $\{x_t\}$ по конечной выборке x_t, $t\in\overline{1,N}$. $\nu(\lambda)=[\frac{\lambda N}{2\pi}]$, $[a]$ - целая часть a. Из ограничений BV на сигнал x_t вытекает, что интегральные суммы (33) сходятся равномерно по $\theta\in\Theta$ и $\rho\in Q$ к интегралам:

$$\lim_{N\to\infty}(1/N)\sum_{j=1}^{N} A(\lambda_j,\theta)X_jX_j^* = \int_0^{2\pi} A(\lambda)dK_x(\lambda,\rho) = \int_0^{2\pi} A(\lambda,\theta)T_x(\lambda,\rho)d\lambda, \quad (34)$$

где $T_x(\lambda,\rho) = \frac{\partial}{\partial\lambda} K_x(\lambda,\rho)$ определяется формулами (7),(8). С учетом соотношений (33),(34) непосредственные вычисления матриц (32) приводят к следующему результату:

$$W(\theta) = U(\theta) = \begin{bmatrix} \Psi_{\theta\theta} & \Psi_{\theta\rho} \\ \Psi_{\rho\theta} & \Psi_{\rho\rho} \end{bmatrix},$$

из которого немедленно следует, что

$$\lim E_{\theta,\{x_t(\rho)\}} N(\tilde{\theta}_N - \theta)(\theta_N - \theta)^T = L^{-1}(\theta,\rho),$$

и для любой функции потерь $r(C)$ в обоих неравенствах (31) при $\delta \to 0$ достигается знак равенства.

Таким образом, при параметрически заданной временной энергетической спектральной плотности неизвестного входного сигнала x_t асимптотически эффективная оценка параметра линейной системы θ получается в результате решения той же системы уравнений, что и в случае, когда дополнительно известно, что x_t – гауссовский регулярный случайный процесс. Этот факт свидетельствует, что часто используемое на практике искусственное предположение о гауссовском характере процессов, реальное распределение которых неизвестно, не всегда привносит в задачу неоправданную дополнительную информацию. Часто это предположение может рассматриваться лишь как формальный метод построения процедур обработки данных, которые обладают рядом оптимальных свойств и "за пределами" предположений о гауссовости.

Устойчивость асимптотически эффективных оценок к изменению спектральной плотности помех

Рассмотренные выше АЭ-оценки параметров наблюдений, удовлетворяющих моделям (5),(6) и (5),(8), построены в предположении, что помехи ξ_t и η_t, маскирующие входной и выходной сигналы линейной системы, а также случайная составляющая $\mu_{t\rho}$ входного сигнала (которая может рассматриваться как дополнительная помеха) – регулярные гауссовские процессы с *известными* матричными спектральными плотностями $F_\xi(\lambda)$, $F_\eta(\lambda)$, $F_{\xi\eta}(\lambda)$ и $F_\mu(\lambda,\rho)$. Эти функции "заложены" в алгоритмы вычисления АЭ-оценок. На практике, однако, приходится решать задачу идентификации в условиях "априорной неопределенности" относительно распределения помех. При этом важную роль играет робастность АЭ-оценок. Если предполагаемые характеристики помех отличаются от реальных асимптотический риск АЭ-оценок будет, как правило, выше теоретического нижнего предела, определяемого неравенствами (19) и (31). Тем не менее, если АЭ-оценка при изменившихся характеристиках помех продолжает оставаться $\sqrt{N}$-состоятельной, а точнее - оценкой из класса $\mathbb{K}$, ее можно назвать "устойчивой к распределению помех". К сожалению, рассмотренные выше АЭ-оценки параметров θ линейной системы, как для модели наблюдений (5),(6), так и для модели (5),(8), а также МП- оценка $\bar{\theta}(\mathbb{V}_N)$ (решение системы уравнений (28)) этим свойством устойчивости в общем случае не обладают. Чтобы показать это, рассмотрим, в чем кроется "механизм" состоятельности АЭ-оценки $\hat{\theta}(\mathbb{V}_N)$ параметров $\theta=(\theta^T,\rho^T)^T$ модели (5), (6). Как показано в [6], при выполнении условий BI-BIV случайная функция $\Delta(\mathbb{V}_N,\theta)$, $\theta\in \mathbb{U} = \Theta\times Q$ с компонентами, определяемыми формулами (22), сходится по мере в пространстве $\mathbb{C}(\mathbb{U})$ непрерывных функций над $\mathbb{U}$ с равномерной метрикой к детерминированной функции $t(\theta_0,\theta)$:

$$P_{N,\theta_0} - \lim_{N\to\infty} \sup_{\theta\in\mathbb{U}} \left|(1/\sqrt{N})\Delta(\mathbb{V}_N,\theta) - t(\theta_0,\theta)\right| = 0, \tag{35}$$

где P_{N,θ_0} - мера наблюдений v_t, $t\in \overline{1,N}$, определяемая спектральными плотностями помех $F_\zeta(\lambda)$ и $F_\mu(\lambda,\rho_0)$ и частотной характеристикой $G(\theta_0)$. При этом необходимым для состоятельности АЭ-оценки $\hat{\theta}(\mathbb{V}_N)$ является условие

$$t(\theta_0,\theta_0) = 0 \text{ для всех } \theta_0\in \mathbb{U}. \tag{36}$$

При ограничениях BI-BIV из (35) следует, что

$$t(\theta_0,\theta) = \lim_{N\to\infty} E_{\theta_0} (1/\sqrt{N})\Delta(\mathbb{V}_N,\theta). \tag{37}$$

В [7] показано, что условия типа (45), (46) являются необходимыми условиями $\sqrt{N}$-состоятельности *любой* оценки в виде корня уравнения $\delta(\mathbb{V}_N,\theta)=0$, где $\delta(\mathbb{V}_N,\theta)$ удовлетворяет условиям BII. Отсюда вытекает, что необходимое условие устойчивости АЭ-оценки к спектральным плотностям помех $F_\zeta(\lambda)$, $F_\mu(\lambda,\rho)$ имеет вид:

$$\tilde{t}(\theta_0,\theta_0) = 0, \quad \theta \in \mathbb{U}, \tag{38}$$

где $\tilde{t}(\theta_0,\theta) = \lim_{N\to\infty} \tilde{t}_N(\theta_0,\theta)$, $\qquad \tilde{t}_N(\theta_0,\theta) = (1/\sqrt{N})\, \tilde{E}_{\theta_0} \Delta(\mathbb{V}_N,\theta)$.

$\tilde{E}_{\theta_0}$ - математическое ожидание по распределению помех ζ_t, $\mu_{t,\rho}$ со спектральными плотностями $\tilde{F}_\zeta(\lambda)$ и $\tilde{F}_\mu(\lambda,\rho_0)$, отличающимися от предполагавшихся априори $F_\zeta(\lambda)$ и $F_\mu(\lambda,\rho_0)$. Непосредственное вычисление функции $\tilde{t}(\theta_0,\theta)$ для оценки $\hat{\theta}(\mathbb{V}_N)$ - корня системы уравнений (22) - приводит к следующей формуле для ее компонент:

$$\tilde{t}^{\alpha}_{Nk}(\theta_0,\theta) = (1/N)\sum_{j=1}^{N}(U_j(\theta_0)-U_j(\theta))^* F_j^{-1}(\theta)[\dot{U}^{\alpha}_{kj}(\theta)+U_j(\theta_0)-U(\theta)] +$$

$$+ (1/N)\sum_{j=1}^{N}\mathrm{tr}F_j^{-1}(\theta)\dot{F}^{\alpha}_{kj}(\theta)[F_j^{-1}(\theta)\tilde{F}_j(\theta_0)-I] + o(1), \qquad (39)$$

где $k\in\overline{1,q_\alpha}$, $\alpha\in\{\theta,\rho\}$, $q_\theta=p$, $q_\rho=q$. $U_j(\theta)$, $F_j(\theta)$, $U^{\alpha}_{kj}(\theta)$, $\dot{F}^{\alpha}_{kj}(\theta)$ и $\tilde{F}(\lambda,\theta_0)$ выражаются формулами (20) и (22), в которых используются $F_\zeta(\lambda)$, $F_\mu(\lambda,\rho)$ и $\tilde{F}_\zeta(\lambda)$, $\tilde{F}_\mu(\lambda,\rho)$ соответственно. Аналогично, для МП-оценки $\bar{\theta}(\mathbb{V}_N)$ параметра линейной системы при неизвестном входном сигнале (корня системы уравнений (28)) получаем, что $\tilde{\tau}_N(\theta_0,\theta)=(1/\sqrt{N})\tilde{E}_{\theta_0}\delta(\mathbb{V}_N,\theta)$ имеет следующие компоненты:

$$\tilde{\tau}_{Nk}(\theta_0,\theta) = (1/N)\sum_{j=1}^{N}\mathrm{tr}D_j^*(\theta_0)\dot{A}_{kj}(\theta)D_j(\theta_0)X_jX_j^* -$$

$$- (1/N)\sum_{j=1}^{N}\mathrm{tr}\dot{A}_{kj}(\theta)F_{\zeta j}^{-1/2}\tilde{F}_{\zeta j}F_{\zeta j}^{-1/2} + o(1), \quad k\in\overline{1,p}, \qquad (40)$$

где $D_j(\theta)$, $A_j(\theta)$ и $A_{kj}(\theta)$ выражаются формулами (28).

Из (39) легко получить, что при $F_\zeta(\lambda)\not\equiv\tilde{F}_\zeta(\lambda)$ условие (38) выполняется, только когда $\mu_{t\rho}\equiv 0$, т.е. входной сигнал линейной системы x_t - квазидетерминированный процесс. Лишь в этом случае АЭ-оценка параметров линейной системы устойчива к спектрам помех. Учет случайной составляющей $\mu_{t\rho}$ сигнала x_t при синтезе процедуры оценивания приводит к неустойчивой АЭ-оценке, т.е. оптимизация качества алгоритма оценивания "покупается" ценой потери устойчивости.

АЭ-оценка при неизвестном эргодическом входном сигнале (т.е. для модели наблюдений (5), (8)) при условиях BV, BVI совпадает с АЭ-оценкой при гауссовском регулярном входном сигнале (т.е. для модели наблюдений (5),(6) с $x_t=\mu_{t\rho}$ и $F_\mu(\lambda,\rho)=T(\lambda,\rho)$). Поэтому согласно (38),(39) эта оценка также неустойчива к спектральной плотности помех $F_\zeta(\lambda)$ и параметрической модели временной спектральной плотности $T(\lambda,\rho)$ входного сигнала.

Анализ устойчивости МП-оценки $\bar{\theta}(\mathbb{V}_N)$ на основе выражений (38), (40) не столь очевиден, как для АЭ-оценок. Покажем сначала, что при $\tilde{F}_\zeta(\lambda)\equiv F_\zeta(\lambda)$ выполнено условие (38). Вычислив производные $\dot{A}_{kj}(\theta)$, получим

$$\begin{aligned} &D_{j0}\dot{A}_{kj}= D^*_{j0}(I-D_{j0}(D^*_{j0}D_{j0})^{-1}D^*_{j0})\dot{G}_{kj0}(G^*_{j0}G_{j0})^{-1}G^*_{j0} = 0, \\ &\mathrm{tr}[\dot{A}_{kj}\theta F_{\zeta j}^{-1/2}\tilde{F}_{\zeta j}F_{\zeta j}^{-1/2}] = \frac{\partial}{\partial\theta_k}\mathrm{tr}[A_j(\theta)F_{\zeta j}^{-1/2}\tilde{F}_{\zeta j}F_{\zeta j}^{-1/2}] = 0, \end{aligned} \qquad (41)$$

где $D_{j0} = D_j(\theta_0)$, $G_{j0} = G_j(\theta_0)$. Последнее равенство в (41) справедливо потому, что $tr[AI] = trG(G^*G)^{-1}G = tr(G^*G)^{-1}G^*G \equiv k$. Из второго выражения в (41) нетрудно вывести, что при $\tilde{F}_\zeta(\lambda) \neq F_\zeta(\lambda)$, вообще говоря, $\tilde{\tau}_{Nk}(\theta_0,\theta_0) \neq 0$. То есть МП-оценка $\bar{\theta}(\mathbb{Y}_N)$ также не является устойчивой к спектральной плотности помех.

Основной вывод из приведенного выше анализа устойчивости состоит в следующем. При наличии случайной составляющей $\mu_{t\rho}$ входного сигнала или при неизвестном эргодическом входном сигнале x_t использование АЭ-оценок вряд ли разумно, поскольку предполагает наличие точной априорной информации об энергетических спектральных плотностях помех и входного сигнала. Если такая информация отсутствует, более целесообразная стратегия состоит в использовании оценок, наилучших в классах устойчивых процедур. Структура таких классов зависит от того, содержит ли входной сигнал x_t квазидетерминированную составляющую $s_t(\rho)$. Результаты, касающиеся построения классов устойчивых оценок и нахождения в них наилучших процедур, предполагается изложить во второй части настоящей работы.

Литература

1. *Кушнир А.Ф., Пинский В.И., Рукавишникова Т.А.* Модельные исследования оценок параметров линейных систем и диспергирующих сред // Логические и вычислительные методы в сейсмологии. М.: Наука, 1984. С.124-140. (Вычисл. сейсмология; Вып. 17).

2. ***Kushnir A.F., Pinsky V.I.*** Optimal compensation of natural electromagnetic field background in electric exploration work // Proceedings of the 5th International Mathmatical Geophysics Seminar held at the Free University of Berlin, Feb. 4-7, 1987. Braunschweig/Wiesbaden: Frier Vieweg & Sohn. P.269-285.

3. ***Kushnir A.F., Levshin A.L., Lokshtanov D.E.*** Determination of a regional velosity structure from surface wave seismograms, recorded at a set of stations // Proceeding of the 6th International Mathmatical Geophysics Seminar held at the Free University of Berlin, Feb. 3-6, 1988. Braunschweig/Wiesbaden: Frier Vieg & Sohn.P.269-285.

4. *Кушнир А.Ф., Локштанов Д.Е.* Статистический подход в задаче определения среды по поверхностным волнам, наблюдаемым на локальной группе станций // Проблемы сейсмологической информатики. М.: Наука, 1988. С.124-136. (Вычисл. сейсмология; Вып.21).

5. *Кушнир А.Ф., Пинский В.И.* Асимптотически эффективное оценивание параметров линейных систем. I. Локальная асимптотическая нормальность // Математические методы в сейсмологии и геодинамике М.: Наука, 1986. С.101-118. (Вычисл. сейсмология; Вып.19).

6. *Кушнир А.Ф.* Асимптотически эффективное оценивание параметров линейных систем. II. Упрощенные асимптотически эффективные оценки // Численное моделирование и анализ геофизических процессов М.:Наука, 1987. С.148-166. (Вычисл.сейсмология; Вып.20).

7. *Кушнир А.Ф.* Алгоритмы идентификации линейных систем при корре-

лированных помехах на входе и выходе // Пробл. передачи информ. 1986. Т.23, вып.2. С.61-74.

8. *Кушнир А.Ф.* Упрощенные асимптотически оптимальные оценки // Тез. докл. IV Междунар. конф. по теории вероятностей и математической статистике. Вильнюс, 1985. Т.II, С.94-96.

9. *Кушнир А.Ф., Лапшин В.М.* Параметрические методы анализа многомерных временных рядов. М.: Наука, 1986. 243с.

10. *Кушнир А.Ф.* Параметрические методы статистического анализа геофизических временных рядов: Дис. ... докт. физ.-мат. наук. 01.04.12. М.: ИФЗ АН СССР, 1989. 310с.

11. *Рабинер Л., Гоулд Б.* Теория и применение цифровой обработки сигналов. М.: Мир, 1978. 848с.

12. *Бриллинджер Д.* Временные ряды. Обработка данных и теория. М.: Мир, 1980. 536с.

13. *Андерсон Т.* Статистический анализ временных рядов. М.: Мир, 1976. 755с.

14. *Малютов М.Б.* Нижняя граница для средней длительности последовательно планируемого эксперимента // Изв. вузов. Математика. 1983. N 11. С.19-41.

15. *Nussbaum M.* An asymptotic minimax risk for estimation of a linear functional relationship//J. Multivar. Anal. 1984. Vol.14, N 3. P.300-314.

16. *Ибрагимов И.А., Хасьминский Р.З.* Асимптотическая теория оценивания. М.: Наука, 1979. 527с.

УДК 550.34

М.Н.Жижин

СИНТАКСИЧЕСКИЙ АНАЛИЗ ЗАПИСЕЙ СИЛЬНЫХ ДВИЖЕНИЙ.
II. СЛУЧАЙ БЕСКОНЕЧНОГО АЛФАВИТА

M.N.Zhizhin

SYNTACTIC ANALYSIS OF STRONG MOTION DATA.
II. THE CASE OF INFINITE GRAMMARS

In this work the Syntactic Pattern Recognition Scheme is constructed for the purpose of classification of multichannel seismic records in case of infinite cardinality of grammars, including the tools for verification of the classification reliability. It is applied for classification of three-component strong motion records from different geological regions.

Введение

Синтаксическое распознавание предполагает представление образов в виде предложений, составленных из символов некоторого алфавита с помощью набора грамматических правил (грамматики). Различные грамматики порождают образы-предложения различных классов. Обучение