УДК 517.962.27

© *А. О. Егоршин*

ОБ ОДНОЙ ВАРИАЦИОННОЙ ЗАДАЧЕ КУСОЧНО–ЛИНЕЙНОЙ ДИНАМИЧЕСКОЙ АППРОКСИМАЦИИ[1]

Изучаются свойства дискретной вариационной задачи динамической аппроксимации в комплексном евклидовом $(L+1)$-мерном пространстве E. Она обобщает известные задачи среднеквадратической полиномиальной аппроксимации функций, заданных своими отсчетами в конечном интервале. В рассматриваемой задаче аппроксимация последовательности $\mathbf{y} = \{y_i\}_0^L$ отсчетов функции $y(t) \in L^2[0,T]$, $T = Lh$ на сетке I_h осуществляется решениями однородных линейных дифференциальных или разностных уравнений заданного порядка n с постоянными, но, возможно, неизвестными коэффициентами. Тем самым показано, что в последнем случае задача аппроксимации включает в себя и задачу идентификации. Анализ ее особенностей — основная тема статьи. Ставится задача нахождения вектора коэффициентов α разностного уравнения $\sum_0^n \widehat{y}_{i+k}\alpha_i = 0$, где $k = \overline{0, L-n}$. Оптимизируются коэффициенты α и начальные условия переходного процесса $\widehat{\mathbf{y}}$ этого уравнения. Цель оптимизации — наилучшая аппроксимация исследуемого динамического процесса $\mathbf{y} \in E$. Критерий аппроксимации — минимум величины $\|\mathbf{y} - \widehat{\mathbf{y}}\|_E^2$. Показано, что изучаемая вариационная задача сводится к задачам проектирования в E вектора $\mathbf{y}$ на ядра разностных операторов с неизвестными коэффициентами $\alpha \in \omega \subset \mathcal{S} \subset E^{n+1}$. Здесь α — направление, $\mathcal{S}$ — сфера или гиперплоскость. Показана связь изучаемой задачи с задачами дискретизации и идентифицируемости. Тогда координаты вектора $\mathbf{y} \in E$ есть точное решение дифференциального уравнения на сетке I_h и $\mathbf{y} = \widehat{\mathbf{y}}$. Дано сравнение изучаемой задачи вариационной идентификации с алгебраическими методами идентификации. Показано, что ортогональные дополнения к ядрам разностных операторов всегда имеют теплицев базис. Это приводит к быстрым проекционным алгоритмам вычислений. Показано, что задача нахождения оптимального вектора $\widehat{\alpha}$ сводится к задаче безусловной минимизации функционала идентификации, зависящего от направления α в E^{n+1}. Предложена итерационная процедура его минимизации на сфере с широкой областью и высокой скоростью сходимости. Изучаемую вариационную задачу можно применять при математическом моделировании в управлении и научных исследованиях. При этом на конечных интервалах может использоваться, в частности, возможность кусочно-линейной динамической аппроксимации сложных динамических процессов разностными и дифференциальными уравнениями указанного типа.

Ключевые слова: вариационная идентификация, алгебраическая идентификация, кусочно–линейная динамическая аппроксимация, ортогональная регрессия, неградиентная оптимизация.

Введение

В статье рассмотрены некоторые вопросы построения по экспериментальным данным математических моделей динамических систем. Под построением модели понимается постановка и решение соответствующей математической обратной задачи обработки сигналов, а именно нахождение математического описания исследуемого динамического процесса или объекта. Обратные задачи моделирования, в отличие от прямых (*simulating*), называются задачами математического моделирования или идентификации.

Задачи математического моделирования реальных физических объектов имеют особенности, связанные с их конечной «памятью», с рассеиванием информации о «предыстории». Поэтому их уместно причислять к задачам математической физики. Математическими моделями операторов «вход–выход» реальных объектов должны быть необратимые (например, вполне непрерывные) операторы. Одним из способов регуляризации таких задач является параметризация, в частности использование моделей с конечным числом параметров.

[1]Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (проект № 10–01–00035) и Сибирского отделения Российской академии наук (междисциплинарный проект № 80).

Задачи *статистического* оценивания коэффициентов уравнений динамических систем называются также задачами идентификации. При этом стохастическими (случайными, имеющими соответствующие стохастические модели) элементами таких задач чаще всего считают измерительные ошибки в исходных данных. Однако в задачах математического моделирования не редки и так называемые структурные ошибки. Они связаны с тем, что используемая для идентификации математическая модель, как правило, проще (иногда существенно), чем исследуемый объект. Поэтому о задаче математического моделирования можно говорить и как о задаче аппроксимации сложного объекта более простой моделью.

Исследуется аппроксимационный подход к постановке задач математического моделирования сложных объектов. Используются модели, описываемые дифференциальными или разностными уравнениями определенного вида. Такой подход приводит к вариационным параметрическим обратным задачам [1, 2]. При этом оцениваются две группы параметров модели — оптимальные начальные условия аппроксимирующих динамических процессов и оптимальные коэффициенты их уравнений. Для краткости первая из этих задач называется задачей сглаживания, вторая — задачей идентификации. При использовании неоднородных уравнений в качестве моделей исследуемых процессов может решаться также и третья обратная задача — оценивание входного сигнала [3, 1]. Она здесь не рассматривается.

Предложенный подход назван *вариационной идентификацией* (ВИ) [1]. Используются модели, традиционные для теории систем автоматического управления. Имеются в виду модели, описываемые обыкновенными линейными, стационарными (с постоянными коэффициентами) дифференциальными или разностными уравнениями. Отметим, что решения большей части задач теории оптимального управления — «ровесницы» проблемы идентификации, появившейся в одно время (в 50-х гг. прошлого века) и для тех же целей (адаптивное управление сложными объектами аэрокосмической и атомной техники), — реализуемы на практике в основном для линейных моделей, используемых в общей задаче ВИ [4].

Аппроксимация такими моделями сложных динамических процессов на конечных интервалах называется их *кусочно–линейной динамической аппроксимацией.*

§ 1. Предварительные сведения

«Прототипом» рассматриваемой здесь вариационной задачи идентификации является классическая задача аппроксимации — задача сглаживания и восстановления функций с помощью степенных полиномов и метода наименьших квадратов (МНК) .

Задачу аппроксимации функции $y(t) \in L^2[I_T]$, $I_T = [0,T]$ полиномом степени $n-1$ сформулируем следующим образом. Найти в $L^2 = L^2[0,T]$ n раз дифференцируемую функцию $\widehat{y}(t)$, на которой минимален функционал

$$J_y(\widehat{y}(t)) = \|y(t) - \widehat{y}(t)\|^2_{L^2}\,, \quad \text{если} \quad \widehat{y}^{(n)}(t) = 0, \quad t \in [0,T] = I_T. \tag{1}$$

Пусть $E^{L+1} = E$. Обозначим через $\mathbf{y} = \{y_i\}_0^L$ $(L+1)$-вектор отсчетов функции $y(t)$ на h-сетке I_h интервала I_T. Здесь $T = Lh$. Такие векторы называются далее реализациями (или h-реализациями) на I_T. Так как $\widehat{y}^{(n)}(t) = 0 \to \boldsymbol{\Delta}^n \widehat{\mathbf{y}} = 0$, то дискретный аналог задачи (1) в E (или в $l^2[0,L]$) формулируется аналогичным образом. Требуется найти вектор $\widehat{\mathbf{y}} = \{\widehat{y}_i\}_0^L$ в E, на котором минимален функционал

$$J_y(\widehat{\mathbf{y}}) = \|\mathbf{y} - \widehat{\mathbf{y}}\|^2_E\,, \quad \text{если} \quad \boldsymbol{\Delta}^n \widehat{\mathbf{y}} = 0 \quad \text{или} \quad \widehat{\mathbf{y}} \in \ker \boldsymbol{\Delta}^n. \tag{2}$$

Здесь

$$\boldsymbol{\Delta}^n \mathbf{y} = \boldsymbol{\Delta}^n \{y_k\}_0^N = \{\Delta^n y_k\}_{k=0}^N\,, \quad \text{где} \quad \Delta^n y_k = \sum_{i=0}^{n} (-1)^{n-i} C_n^i s^i y_k = \sum_{i=0}^{n} (-1)^{n-i} C_n^i y_{k+i},$$

$N = L - n$, а s — оператор сдвига в E'. Мы рассматриваем координаты в E как функционалы. Поэтому $y_k = e_k^* \mathbf{y} = \langle \mathbf{y}, e_k \rangle$, где $e_k^* = |\delta_{kj}|_{j=0}^L = \langle \,\cdot\,, e_k \rangle$, $k = \overline{0,L}$ — орты из E'. Оператор сдвига s определяется действием на ортах в E'. Матрица оператора s в E' есть матрица сдвига

$$I^{1T} = \{\delta_{i+1,j}\}_0^L\,; \qquad s e_k^* = s \langle \,\cdot\,, e_k \rangle = e_{k+1}^* \longrightarrow e_k^* I^{1T} = \langle \,\cdot\,, I^1 e_k \rangle = \langle \,\cdot\,, e_{k+1} \rangle\,. \tag{3}$$

Здесь $k = \overline{0, L}$, e_k, — орты в E, а $I^1 = \{\delta_{i-1,j}\}_0^L$. Символы вида δ_{ij} и $\delta_{i+1,j}$ обозначают символы Кронекера. Запятая во втором случае используется для однозначности восприятия индекса.

Конструкция $\langle\,\cdot\,,\,\cdot\,\rangle = \langle\,\cdot\,,\,\cdot\,\rangle_E$ обозначает скалярное произведение в E, если индексом не указано иное. Ее аргументами могут быть и матрицы, и множества векторов. Тогда она обозначает соответствующее множество скалярных произведений. С помощью скалярного произведения определяем и функционалы. Если $\mathbf{x} \in E$, то $\langle\,\cdot\,, \mathbf{x}\rangle = \mathbf{x}^* = \mathbf{x}'$.

Задачи (1), (2) допускают обобщения в разных направлениях [1, 4]. Эти обобщения охватывают широкий круг приложений — от идентификации и математического моделирования динамических объектов различной природы до аппроксимации функций суммой комплексных экспонент и квазиполиномами, то есть решениями однородных дифференциальных и разностных уравнений.

Замечание 1. Задачей аппроксимации функций экспонентами интересовались многие математики. Например (см. [5]), ей была посвящена диссертация Л. Шварца.

Затронул эту тему и Р. Беллман. В частности, он обратил внимание на связь этой задачи с разностными уравнениями. Его тезисы 1960г. [6] и были посвящены использованию этого факта для определения числа экспонент в их сумме. □

Условимся о некоторых обозначениях для векторов и матриц. Конструкции вида $\{x_i\}_k^m$ будут обозначать вектор–столбец, $|x_j|_l^n$ или $|x_l, \dots, x_n|$ — вектор–строку, а $\{x_{ij}\}_{kl}^{mn}$ — матрицу (i — номера строк) множества элементов вида x. Если элементы x_i в $\{x_i\}_k^m$ — $(n-l+1)$-векторы–стро́ки, а x_j в $|x_j|_l^n$ или в $|x_l, \dots, x_n|$ — $(m-k+1)$-векторы–столбцы, то эти конструкции определяют матрицу $\{x_{ij}\}_{kl}^{mn}$. Надстрочный символ $\widehat{\ }$ отмечает отсчеты и функции модели.

Верхний индекс $()^*$ для скаляров будет обозначать комплексное сопряжение. Для векторов и матриц он будет обозначать двойную инволюцию: комплексное сопряжение и транспонирование. Специальные конструкции: 0_i — нулевой i-вектор–столбец, 0_i^T — i-вектор–строка, $0_{i,j}$ — $(i \times j)$-матрица. Необходимые примеры: $I_{(n)i}^T = |0_{n+1,i}, I_{n+1}, 0_{n+1,N-i}| \,:\, E \to E^{n+1}$, $i = \overline{0, N}$ и $I_{(N)i}^T = |0_{N+1,i} I_{N+1}, 0_{N+1,n-i}| \,:\, E \to E^{N+1}$, $i = \overline{0, n}$ — «вырезающие» матрицы.

§ 2. Обобщения степенной задачи аппроксимации

Ниже мы приводим некоторые обобщения задач степенной аппроксимации, сформулированных в предыдущем разделе. В них в качестве моделей функций или их отсчетов на h-сетке использовались дифференциальные и разностные уравнения $\widehat{y}^{(n)}(t) = 0$ и $\boldsymbol{\Delta}^n \widehat{\mathbf{y}} = 0$ соответственно. Мы обобщаем эти модели, ограничиваясь в этой статье также однородными и скалярными уравнениями. Мы осуществляем аппроксимацию функций переходными процессами таких моделей. Это есть, соответственно, квазиполиномы вида

$$\sum_{i=1}^{r} \sum_{j=0}^{k_i-1} t^j \exp(t\mu_i) \quad \text{или} \quad \sum_{i=1}^{r} \sum_{j=0}^{k_i-1} (hk)^j \nu_i^k, \quad \text{где} \quad \sum_{1}^{r} k_i = n, \quad \nu_i = \exp(h\mu_i).$$

Рассматриваются следующие обобщения моделей задач (1) и (2).

1. Линейные однородные дифференциальные (ДУ) или разностные (РУ) уравнения с постоянными, возможно, неизвестными коэффициентами.

$$\text{ДУ:} \qquad \sum_{i=0}^{n} \widehat{y}^{(i)}(t) a_i^* = 0, \quad a_n \neq 0, \quad t \in I_T = [0, T]; \tag{4}$$

$$\text{РУ:} \quad \sum_{i=0}^{n} \widehat{y}_{k+i} \alpha_i^* = 0 = \sum_{i=0}^{n} \left\langle \widehat{\mathbf{y}}, I^i e_k \alpha_i \right\rangle, \quad \alpha_n \neq 0 \quad k = \overline{0, N}. \tag{5}$$

2. Уравнения (то есть их коэффициенты и решения) могут быть комплексными.

Векторы коэффициентов a и α в ДУ (4) и РУ (5) представляются в виде

$$a^* = \left|a_0^*, \dots, a_{n-1}^*, a_n^*\right| = |\overline{a}^*, a_n^*| \quad \text{и} \quad \alpha^* = \left|\alpha_0^*, \dots, \alpha_{n-1}^*, \alpha_n^*\right| = |\overline{\alpha}^*, \alpha_n^*|\,.$$

На вектор коэффициентов могут быть наложены ограничения. Априорные ограничения на коэффициенты РУ (5) будем обозначать в виде $\alpha \in \omega$, где ω — множество допустимых значений оптимизируемого вектора коэффициентов α. Он задается, как видно из (5), с точностью до постоянного множителя, то есть отыскивается как направление. В отсутствие ограничений это сфера или гиперплоскость (например, $\alpha_n = 1$) в E^{n+1}.

Априорная информация о линейных ограничениях на коэффициенты уравнений может быть задана в виде уравнений линейных связей, например, $\alpha = \mathbb{T}\gamma$ или $\alpha = \overline{\mathbb{T}}\overline{\gamma} + \alpha_*$. В частности, параметры γ могут быть коэффициентами a ДУ (4). Тогда $\gamma = a$, а столбцы матрицы $\mathbb{T}$ есть векторы коэффициентов соответствующих конечных разностей [4].

Обобщения задач (1), (2) на неоднородные уравнения и на системы неоднородных уравнений (с векторными функциями и матричными коэффициентами) рассмотрены в [1, 4].

§ 3. Дискретизация и идентификация

3.1. Аналитическая и вариационная дискретизациия.

Определение 1. Назовем *полной* h-дискретизацией ДУ (4) в интервале I_T получение такого РУ (5), то есть его коэффициентов α, подпространство системы фундаментальных решений (СФР) которого совпадает с подпространством h-реализаций СФР ДУ (4) на сетке I_h интервала I_T. Система называется дискретизируемой, если решение задачи дискретизации — вектор коэффициентов РУ (5) α — единственно. □

Это определение означает, что с помощью СФР РУ (5) могут быть точно описаны h-реализации любого решения ДУ (4) на h-сетке I_h.

Пусть $\Phi = \Phi_h = \exp(\mathcal{A}h)$ — $(n \times n)$-матрица дискретной системы $\{\Phi, d\}$: $x_k = \Phi x_{k-1}$, $y_k = d^* x_k$, где $y(t)$ — так называемый «выход» системы. Эта система есть h-дискретный аналог дифференциальной системы $\{\mathcal{A}, d\}$: $\dot{x}(t) = \mathcal{A}x(t)$, $y(t) = d^* x(t)$, $t \in I_T$.

Пусть $m_k = \sum\limits_0^n y_{i+k}\alpha_i^* = \sum\limits_0^n d^*\Phi^{i+k}x_0\alpha_i^*$, $k = \overline{0, N}$. По теореме Гамильтона–Кэли $m_k \equiv 0$. Это равенство для $k = \overline{0, N}$ дает, во–первых, РУ(5). Во–вторых, если $\alpha_n = 1$, то при любом конечном k, в силу произвольности x_0, оно дает уравнения аналитической или *структурной* дискретизации $\overline{\alpha} = -d^*F^{-1} = -d^*\Phi^k\big(F\Phi^k\big)^{-1}$. Здесь значение k любое, а $F = \{d^*\Phi^{i-n}\}_0^{n-1}$ — матрица наблюдаемости решений ДУ на h-сетке I_h. Условия структурной дискретизируемости есть условия на матрицу $\mathcal{A}$ $(\operatorname{rank} \mathcal{A} = n)$.

Вариационный или *функциональный* способ дискретизации ДУ (4) основан на *идентификации* РУ (5), исходя из *точной* h-реализации $\mathbf{y}$ решения $y(t)$ ДУ (4). Условия вариационной дискретизируемости есть условия на свойства системы, требования к шагу сетки и ограничения на вектор $[y]_0$ начальных условий реализации $\mathbf{y}$.

Можно показать, что вектор начальных условий $[y]_0$ должен обеспечить в исходных данных $\mathbf{y}$ наличие h-реализаций всех r основных собственных движений (мод) $\varphi_{i/(k_i-1)}(t) = t^{k_i-1}\exp(\mu_i t)$ дифференциальной системы $\{\mathcal{A}, d\}$ на сетке I_h. Здесь $i = \overline{1, r}$, $j = \overline{0, k_i - 1}$, r — количество, а k_i — кратности $\left(\sum_1^r k_i = n\right)$ различных корней μ_i характеристического полинома $p\,(a) = \sum_0^n p^i a_i^* = 0$. Реализации этих мод на h-сетке есть векторы

$$\boldsymbol{\varphi}_{i/(k_i-1)} = \{(hk)^{k_i-1}\nu_i^k\}_{k=0}^L, \quad \text{где} \quad \nu_i = \exp(\mu_i h).$$

Теорема 1. *Пусть аналитическая h-дискретизация ДУ (4) — то есть РУ (5), полученное из теоремы Гамильтона–Кэли, как описано выше, — единственна, и пусть $\mathbf{y} = \{y_i\}_0^L$ — реализация решения $y(t)$, $t \in I_T$ ДУ (4) на сетке I_h. Чтобы решение $\widehat{\alpha}$ задачи минимизации функционала МНК J_y из (2) с условиями (5) было единственным, необходимо и достаточно, чтобы реализация $\mathbf{y}$ была полной, то есть содержала h-реализации $\boldsymbol{\varphi}_{i/(k_i-1)} = \{(hk)^{k_i-1}\exp(\mu_i hk)\}_{k=0}^L$ всех основных мод $\varphi_{i/(k_i-1)}(t) = t^{k_i-1}\exp(\mu_i t)$ ДУ (4) на h-сетке I_h.*

Следствие 1. (a) $\widehat{\mathbf{y}} = \mathbf{y}$ *и значение $\widehat{J}_{\widehat{\alpha}}$ минимума величины J_y из* (2) *равно нулю;*

(b) *значение $\widehat{\alpha}$ есть единственное решение задачи h-дискретизации ДУ* (4).

Следствие 2. *Пусть* $L_1 = L + n$ *и* h*-реализация* $\mathbf{y}(L_1)$ *решения* $y(t)$ *ДУ на интервале* $I_1 = [0, (L+n)h]$ *— полна. Тогда* n *ее подреализаций* $\mathbf{y}_i = I^T_{(L)i}\mathbf{y}(L_1) = \{y_{k+i}\}^L_{k=0}$, $i = \overline{0, n-1}$, *линейно независимы и* $((L+1) \times n)$*-матрица* $\overline{V} = \overline{V}_{(L)} = |\mathbf{y}_i|^{n-1}_0$ *имеет ранг* n.

Доказательства мы опускаем. Они не относятся к главной теме статьи.

Пусть $S(\cdot)$ — линейная оболочка аргументов, указанных в скобках $(\cdot)$. Если аргумент — матрица, то имеются в виду ее столбцы. Пусть $\mathbf{y}(L_1) = \{y_i\}^{L_1}_0$ — решение РУ (5) длины L_1 — есть реализация решения $y(t)$ ДУ (4) на h-сетке в интервале I_1. Пусть, наконец, $\overline{V}_{(L)} = \{\overline{v}^*_i\}^L_0 = |\,\mathbf{y}_j|^{n-1}_0$ — ганкелева $\big((L+1) \times n\big)$-матрица отсчетов этого решения. Ее строки есть n-выборки $\overline{v}^*_i = |\,y_{i+j}|^{n-1}_{j=0}$, а столбцы — $(L+1)$-подреализации $\mathbf{y}_j = I^T_{(L)j}\mathbf{y} = \{y_{j+i}\}^L_{i=0}$ из реализации $\mathbf{y}(L+n)$. Для матрицы $\overline{V}_{(L)}$ имеет место равенство $\operatorname{rank}\overline{V}_{(L)} = n$ (следствие 2), и поэтому она есть один из возможных базисов подпространства СФР РУ (5), и это подпространство — ядро разностного оператора D_α в E — совпадает с линейной оболочкой $S(\overline{V}_{(L)})$.

Следствие 3. *Пусть* $U = U_{(L)} = \{d^*\Phi^i\}^L_0$ *— есть* $\big((L+1)\times n\big)$*-матрица* h*-реализаций СФР ДУ (4) на сетке* I_h *интервала* I_T *и реализация* $\mathbf{y}(L_1)$ *является полной. Тогда*

(a) *подпространства* $S(\overline{V}_{(L)})$ *и* $S(U)$ *совпадают;*

(b) *вариационная дискретизация является полной;*

(c) *все квадратные блоки матрицы* $\overline{V}_{(L)}$ *невырождены. Это блоки* $\overline{V}_k = \big\{y_{i+j+k}\big\}^{n-1}_{i,j=0}$, *где* $k = \overline{0, L-n+1}$.

3.2. Матрицы Казорати. Отсчёты y_i, $i = \overline{0, L}$ решения ДУ (4) в условиях теоремы 1 известны точно. Поэтому коэффициенты $\overline{\alpha}^* = |\alpha^*_0, \ldots, \alpha^*_{n-1}|$ РУ (5) можно вычислить с помощью любого невырожденного $n \times (n+1)$-блока $V_k = |\overline{V}_k, \{y_{k+n+i}\}^{n-1}_0|$ ганкелевой $(N+1) \times (n+1)$-матрицы $V = V_{(N)} = |\mathbf{y}_i|^n_0$ отсчетов реализации $\mathbf{y} = \{y_i\}^L_0$. Здесь $\mathbf{y}_j = \{y_{j+i}\}^N_0$.

В условиях теоремы 1 и в соответствии с результатом (c) следствия 3 квадратные $(n \times n)$-блоки $\overline{V}_k = \{y_{k+i+j}\}^{n-1}_{i,j=0}$ ганкелевой $((N+1) \times n)$-матрицы $\overline{V} = \overline{V}_{(N)} = |\,\mathbf{y}_j|^{n-1}_0 = \{\overline{v}^*_i\}^N_0$ для всех значений $k = \overline{0, N-n+1}$ являются невырожденными.

Уточним, что в отличие от матрицы $\overline{V}_{(L)}$ в следствии 2 столбцы матрицы $\overline{V} = \overline{V}_{(N)}$ есть подреализации $\mathbf{y}_j = I^T_{(N)i}\mathbf{y}(L)$, $j = \overline{0, n-1}$ h-реализации $\mathbf{y}(L) = \mathbf{y} = \{y_i\}^L_0$ решения ДУ. Строками являются n-выборки $\overline{v}^*_i$, $i = \overline{0, N}$, состоящие из n подряд следующих отсчетов реализации $\mathbf{y}$. Итак,

$$\begin{gathered} V = |\mathbf{y}_j|^n_0 = \{v^*_i\}^N_0\,, \quad \text{где} \quad \mathbf{y}_i = I^T_{(N)i}\mathbf{y} = \{y_j\}^{N+i}_i \longrightarrow \\ V^* = |v_k|^N_0 = \big|(y)_{k(n+1)}\big|^N_0\,, \quad \text{где} \quad v_k = (y)_{k(n+1)} = \big\{y^*_j\big\}^{k+n}_k. \end{gathered} \tag{6}$$

Для $k = \overline{0, N-n+1}$ имеем: $V_k = \big\{y_{i+j+k}\big\}^{i=n-1,\, j=n}_{i,j=0} =$

$$\Big|\overline{V}_k, [y]_{k+n(n)}\Big| = \begin{vmatrix} y_k, & y_{k+1}, & \cdots, & y_{k+n-1}, & y_{k+n} \\ y_{k+1}, & y_{k+2}, & \cdots, & y_{k+n}, & y_{k+n+1} \\ y_{k+2}, & y_{k+3}, & \cdots, & y_{k+n+1}, & y_{k+n+2} \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ y_{k+n-1}, & y_{k+n}, & \cdots, & y_{k+2(n-1)}, & y_{k+2n-1}, \end{vmatrix} = \Big|[y]_{k(n)}, \overline{V}_{k+1}\Big|, \tag{7}$$

где $[y]_{k(n)} = \{y_j\}^{k+n-1}_k$. Ганкелева матрица V, ее $n \times (n+1)$-блоки V_k, $k = \overline{0, N-n+2}$ определяют $N-n+1$ систем из n уравнений для n-вектора $(\alpha) = \{\alpha^*_i\}^n_0 = \alpha^{*T}$. Они имеют вид

$$V_k \cdot (\alpha) = 0, \quad \text{или} \quad \overline{V}_k \cdot (\overline{\alpha}) = -\,[y]_{k+n(n)}, \quad \text{если} \quad \alpha_n = 1. \tag{8}$$

Обращаемые при решении систем (8) матрицы $\overline{V}_k$ из (7) названы матрицами Казорати [6].

3.3. Алгебраическая идентификация. Матрицы Казорати могут быть использованы (что предлагается иногда) для идентификации, то есть для вычисления коэффициентов РУ или оценки его порядка [6]. Такой линейный подход имеет существенный недостаток для практического применения из–за его высокой чувствительности к ошибкам в реальных отсчётах.

Во введении отмечались два вида ошибок в исходных данных $\mathbf{y}$, характерных для задач математического моделирования: измерительные ошибки и структурные. Наличие ошибок существенно при идентификации с помощью матриц выборок, так как они плохо обусловлены. Для линейных методов оценки оказываются смещенными и при случайных ошибках с нулевым средним, поскольку ими искажены обращаемые матрицы $\overline{V}_k$ систем уравнений (8).

При линейном подходе для «борьбы с ошибками» составляют переопределенную (при $L > 2n$) систему уравнений, объединив уравнения из (8) для всех $k = \overline{0, N-n+1}$. Матрица (6) такой системы линейных уравнений $V \cdot (\overline{\alpha}) = 0$ может быть представлена в виде $V = |\overline{V}, \overline{\mathbf{y}}|$. Здесь $\overline{\mathbf{y}} = \mathbf{y}_n$ — n-я подреализация, последний столбец матрицы V в (6). Теперь система алгебраических уравнений модели исследуемой последовательности $\mathbf{y}$ может быть записана в виде приближенных равенств

$$m = m(\overline{\alpha}) = V \cdot (\overline{\alpha}) \approx 0, \quad \alpha_n = 1, \longrightarrow \overline{V} \cdot (\overline{\alpha}) \approx -\overline{\mathbf{y}}, \tag{9}$$

где через $N+1$-вектор m обозначена невязка уравнения модели. Оценивая вектор $\overline{\alpha}$ по формулам линейного МНК (минимизация по $\overline{\alpha}$ величины $J_{ai} = \|m(\overline{\alpha})\|^2_{E^{N+1}}$), получим:

$$(\overline{\alpha}) = \overline{\alpha}^{*T} = \{\alpha_i^*\}_0^{n-1} = -\left(\overline{V}^*\overline{V}\right)^{-1}\overline{V}^*\overline{\mathbf{y}}. \tag{10}$$

Такой способ решения задачи оценивания коэффициентов РУ (5) также не удовлетворителен для многих приложений. В (9), как и в (8), искажена ошибками не только правая часть системы уравнений — $(N+1)$-вектор $\overline{\mathbf{y}}$, как это принимается в линейном МНК, — но и матрица системы — матрица отсчетов $\overline{V}$. Это приводит ко многим хорошо известным негативным последствиям. Причина их — обращаемая матрица $\overline{V}^*\overline{V}$ не только плохо обусловлена, но и искажена случайными и неопределенными ошибками в отсчетах, составляющих матрицу $\overline{V}$.

Различные методы составления и решения многообразных модификаций систем уравнений вида (9) как для ДУ, так и для РУ есть суть очень распространенного (из–за простоты формул (10)) алгебраического подхода к оценке коэффициентов ДУ и РУ [8, 9, 10].

§ 4. Вариационная идентификация

4.1. Модель. Пусть РУ, выбранное для аппроксимации заданной финитной последовательности (исходной реализации) $\mathbf{y} = \{y_i\}_0^L \in E^{L+1} = E$ отсчетов y_i, $i = \overline{0, L}$, имеет вид (5). Последовательность $\widehat{\mathbf{y}} = \{\widehat{y}_i\}_0^L$ в (5) называется сглаженной реализацией.

Уравнение (5) определяет $N+1$ соотношений для $(n+1)$-векторов $\widehat{v}_k^{*T} = [\widehat{y}]_{k(n+1)} = \{\widehat{y}_{k+j}\}_{j=0}^n$. Это «скользящие» выборки отсчетов из аппроксимирующей последовательности $\widehat{\mathbf{y}}$. Для $k = \overline{0, N}$ указанные $N+1$ соотношений могут быть записаны в виде

$$D_\alpha \widehat{\mathbf{y}} = 0 = \sum_{j=0}^n \alpha_j^* \widehat{y}_{k+j} \longrightarrow \alpha^* [\widehat{y}]_{k(n+1)} = 0 = \left\langle [\widehat{y}]_{k(n+1)}, \alpha \right\rangle_{E^{n+1}}. \tag{11}$$

Реализации $\widehat{\mathbf{y}}$ такие, что $D_\alpha \widehat{\mathbf{y}} = 0$, составляют ядро $\ker D_\alpha = \Psi_\alpha \subset E$ разностного оператора $D_\alpha = D : E \to E^{N+1}$. Эта характеристика подпространства Ψ записывается как $D_\alpha \Psi_\alpha = 0$. Оно называется подпространством модели. Ясно, что $\dim \Psi = L - N = n$.

Состояния модели (5) и (11) есть n-векторы $\widehat{\overline{v}}_k^{*T} = [\widehat{y}]_{k(n)} = \{\widehat{y}_{k+j}\}_{j=0}^{n-1}$, $k = \overline{0, N+1}$. При $k = 0$ имеем начальное состояние $[\widehat{y}]_0 = \{\widehat{y}_i\}_0^{n-1}$ модели (5) или (11).

Векторы состояния $[\widehat{y}]_{k(n)}$ модели при $k = \overline{0, N}$ определяют ее рекуррентное представление, если $\alpha^* = |\overline{\alpha}^*, 1|$. Оно предназначено для последовательного вычисления N компонент реализации $\widehat{\mathbf{y}}$ с начального условия $[\widehat{y}]_0$ по формуле

$$\widehat{y}_{k+n} = -\sum_{j=0}^{n-1} \alpha_j^* \widehat{y}_{k+j} = -\overline{\alpha}^* [\widehat{y}]_{k(n)} = -\left\langle [\widehat{y}]_{k(n)}, \overline{\alpha} \right\rangle_{E^n}, \quad k = \overline{0, N}. \tag{12}$$

4.2. Постановка задачи. Исследуется следующая задача ортогонального проектирования. Минимизируем величину

$$J = \sum_{0}^{N+n} \|y_i - \widehat{y}_i\|^2 = \|\mathbf{y} - \widehat{\mathbf{y}}\|^2\,, \quad \text{если} \tag{13}$$

$$D_\alpha \widehat{\mathbf{y}} = 0 = \left\{ \sum_{0}^{n} \left\langle \widehat{\mathbf{y}}, \alpha_i I^i e_k \right\rangle \right\}_{k=0}^{N}, \qquad \text{или} \qquad \widehat{\mathbf{y}} \in \Psi_\alpha = \Psi = \ker D_\alpha. \tag{14}$$

Подчеркнем, что (11), (12), (14) — это эквивалентные записи совокупности $N+1$ условий (5) минимизации функционала МНК (13) и J_y из (2).

Лемма 1. *Число* $L+1$ *отсчетов в реализации* $\mathbf{y}$, *необходимое для однозначного решения задачи оценивания вектора* $\overline{\alpha}^* = |\alpha_0^*, \ldots, \alpha_{n-1}^*|$ *коэффициентов уравнения* (12), *должно быть таким, что* $L+1 \geqslant k_{min} = 2n \to L \geqslant 2n-1$.

Д о к а з а т е л ь с т в о. Число $k_{min} = 2n$ легко подсчитать из определений (7) матриц систем уравнений (8) в п. **4.2**. Оно равно числу отсчетов, необходимому для формирования одной из систем уравнений $V_k \cdot (\alpha) = 0$ (8). Заметим, что оно равно числу параметров задачи (13), (14). Это n начальных условий и n коэффициентов $\overline{\alpha}$. □

4.3. Сдвиговые матрицы и невязка уравнения. Введем две специальные матрицы:

$$A = \left|\eta_i\right|_0^N, \ \ \eta_i = I^1 \eta_{i-1}, \ \ \eta_0 = \left|\alpha^*, 0_N^T\right|^*, \ \ i = \overline{1, N} \qquad \text{и} \qquad V = |\mathbf{y}_i|_0^n = \left\{\, (y)^*_{k(n+1)} \,\right\}_0^N.$$

Первая — ленточная теплицева $(N+1) \times (L+1)$-матрица A. Матрицу $A = A(\alpha) = A_N$ мы называем *матрицей скользящего вектора* α (МСВ α). Будем обозначать также $\widehat{A} = A(\widehat{\alpha})$. Можно показать [7], что A^* есть матрица обобщенного разностного оператора $D_\alpha : E \to E^{N+1}$ РУ (5) в естественных базисах $\left\{\delta_{ij}\right\}$ пространств обобщенных отсчетов E и невязок обобщенных разностных уравнений E^{N+1}.

Вторая — ганкелева $(N+1) \times (n+1)$-матрица выборок $V = V_{(N)} = V(\mathbf{y})$ из (6). Столбцы матрицы $V^* = V^*(\mathbf{y})$ есть $(n+1)$-выборки $v_k = (\widehat{y})_{k(n+1)} = \{y_j^*\}_k^{k+n}$ из реализации $\mathbf{y} = \{y_i\}_0^L$. Будем обозначать также $\widehat{V} = V(\widehat{\mathbf{y}})$. Итак,

$$A^* = \begin{vmatrix} \alpha_0^* & \ldots & \alpha_n^* & & & & \\ & \ddots & \vdots & \ddots & & \mathbf{0} & \\ & & \alpha_0^* & \ddots & \alpha_n^* & & \\ & \mathbf{0} & & \ddots & \vdots & \ddots & \\ & & & & \alpha_0^* & \ldots & \alpha_n^* \end{vmatrix}; \qquad V = \begin{vmatrix} y_0, & y_1, & \ldots, & y_{n-1}, & y_n \\ y_1, & y_2, & \ldots, & y_n, & y_{n+1} \\ y_2, & y_3, & \ldots, & y_{n+1}, & y_{n+2} \\ y_3, & y_4, & \ldots, & y_{n+2}, & y_{n+3} \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ y_N, & y_{N+1}, & \ldots, & y_{L-1}, & y_L \end{vmatrix}. \tag{15}$$

Лемма 2. *Заменим в комплексном уравнении* (5) *сглаженную реализацию* $\widehat{\mathbf{y}}$ *на исходную* $\mathbf{y}$. *Для невязки* $m = m(\alpha) \in E^{N+1}$ (9) *имеет место следующее тождество* [3, 7]*:*

$$m = A^* \mathbf{y} \equiv V \cdot (\alpha) = V \alpha^{*T} \ \longrightarrow \ m = \langle \mathbf{y}, A \rangle \equiv \left\langle \alpha^{*T}, V^* \right\rangle = \langle (\alpha), V^* \rangle_{E^{n+1}}.$$

Следствие 1. *Пусть* $(\alpha) = \{\alpha_i^*\}_0^n = \alpha^{*T}$. *Условия* (14) *минимизации функционала* (13) *могут быть записаны двумя способами для комплексных скалярных компонент.*

$$A^* \widehat{\mathbf{y}} \equiv \widehat{V} \cdot (\alpha) = \widehat{V} \alpha^{*T} = 0 \ \longrightarrow \ \langle \widehat{\mathbf{y}}, A \rangle \equiv \left\langle \alpha^{*T}, \widehat{V}^* \right\rangle = \left\langle\, (\alpha)\,, \widehat{V}^* \right\rangle_{E^{n+1}} = 0.$$

Следствие 2. *В скалярном действительном случае тождество для невязки и условия минимизации задачи* (13), (14) *приобретают более простой вид.*

$$m = A^T \mathbf{y} \equiv V\alpha \ \longrightarrow \ m = \langle \mathbf{y}, A \rangle \equiv \left\langle \alpha, V^T \right\rangle_{E^{n+1}},$$

$$A^T \widehat{\mathbf{y}} \equiv \widehat{V} \alpha = 0 \ \longrightarrow \ \langle \widehat{\mathbf{y}}, A \rangle \equiv \left\langle \alpha, \widehat{V}^T \right\rangle_{E^{n+1}} = 0. \tag{16}$$

§ 5. Сглаживание и фильтрация

5.1. Формулы сглаживания. Обозначим через $S = S_N = S_\alpha = S\big(A(\alpha)\big)$ ортогональное дополнение $E \ominus \Psi_\alpha$ к подпространству модели Ψ_α. Проекция на это подпространство вычисляется по формулам ортогонального проектирования в евклидовом пространстве. Обозначим через C матрицу Грама системы векторов A.

Теорема 2. *Сглаженная реализация $\widehat{\mathbf{y}}_\alpha$ и перпендикуляр $\Delta\mathbf{y}_\alpha = \mathbf{y} - \widehat{\mathbf{y}}_\alpha$ имеют вид*

$$\widehat{\mathbf{y}}_\alpha = \Pi\mathbf{y}, \quad \Pi = \Pi_\alpha = \Pi_N = I - P, \quad P = P_\alpha = P_N, \quad \Delta\mathbf{y}_\alpha = P\mathbf{y}, \tag{17}$$

где Π есть проектор на Ψ_α, а $P = P(A)$ — проектор на $S_\alpha = S(A) = E \ominus \Psi_\alpha$:

$$P = P\,(A) = A\,\langle A, A\rangle^{-1}\,\langle\,\cdot\,, A\rangle = AC^{-1}A^*, \quad C = A^*A = \langle A, A\rangle = C_\alpha. \tag{18}$$

Формулы (17), (18) есть беспоисковое решение задачи минимизации (13) по локальным условиям реализации $\widehat{\mathbf{y}}$. Это следует из того, что из формул (17), (18) достаточно вычислить лишь, например, первые n отсчетов $[\widehat{y}]_0$. Остальные отсчеты реализации можно вычислить по рекуррентному уравнению модели (12). Для этого при больших N важны рекуррентные способы вычисления проекторов и проекций.

Отметим факты, лежащие в основе таких вычислений.

Теорема 3. (a) *Ортогональное дополнение S в E к ядру стационарного разностного оператора $D_\alpha : E \to E^{N+1}$ всегда имеет теплицев базис скользящего вектора.*

(b) *Его матрица* (15) $A = A(\alpha) = A_N = |\eta_i|_0^N$ *есть МСВ α — вектора коэффициентов оператора D_α. Она определяется начальным вектором $\eta_0 = |\alpha^*, 0_N^T|^*$ и степенями оператора сдвига* (3) $I^1 = \{\delta_{i-1,j}\}_0^L$ *так, что $\eta_i = I^i\eta_0$, $i = \overline{0, L}$.*

5.2. Рекуррентные уравнения для проекции. Введем понятие частично сглаженной реализации $\widehat{\mathbf{y}}_k = \Pi_k\mathbf{y}$. Проектор $\Pi_k = I - P_k$, где $P_k = P(S_k)$ изменяет только первые $\overline{0, k+n}$ отсчетов из реализации $\mathbf{y}$. Поэтому в частично сглаженной реализации — $(L+1)$- векторе $\widehat{\mathbf{y}}_k$ — последние $N-k$ отсчетов равны исходным, а первые $k+n$ отсчетов — «сглажены» проектором Π_k. Эти факты следуют из теоремы 3.

Прогнозами состояний $[\widehat{y}]_{k+1(n)}$ модели (12) на тактах $k = \overline{-1, N}$ мы называем n-векторы

$$[\widehat{y}]_{k+1/k} = [\widehat{y}]_{k+1(n)/k} = \{\widehat{y}_{k+i+1/k}\}_{i=0}^{n-1} = I^T_{(n)k+1}\widehat{\mathbf{y}}_k. \tag{19}$$

Векторы (19) есть последние n из $k+n$ сглаженных отсчетов в частично сглаженной реализации $\widehat{\mathbf{y}}_k$. Оценки (19) определяют прогнозы $\widehat{y}_{k+1+n/k}$ модели по формуле (12) на последний сглаживаемый $(k+n+1)$-й отсчет следующей, $(k+1)$-й частично сглаженной реализации $\widehat{\mathbf{y}}_{k+1}$.

Ошибка этого прогноза равна

$$\pi_{k+1} = \langle\widehat{\mathbf{y}}_k, \eta_{k+1}\rangle = y_{k+1+n} - \widehat{y}_{k+1+n/k}, \qquad k = \overline{-1, N-1}.$$

Она называется также процессом обновления.

Теорема 4. (a) *Частично сглаженная реализация $\widehat{\mathbf{y}}_{k+1} = \Pi_{k+1}\mathbf{y}$ описывается уравнением*

$$\widehat{\mathbf{y}}_{k+1} = \widehat{\mathbf{y}}_k - f_{k+1}a_{k+1}\pi_{k+1}, \qquad \widehat{\mathbf{y}}_{-1} = \mathbf{y}, \tag{20}$$

где $\pi_{k+1} = \langle\widehat{\mathbf{y}}_k, \eta_{k+1}\rangle = y_{k+1+n} - \widehat{y}_{k+1+n/k}$ — процесс обновления, $k = \overline{-1, N-1}$, а $\Pi_{-1} = I$.

(b) *Прогнозы $\widehat{y}_{k+1+n/k}$ на отсчеты y_{k+1+n} осуществляются с помощью рекуррентного уравнения модели* (12)*. Они выполняются по формуле* (12) *на основе оценки $[\widehat{y}]_{k+1/k}$* (19) *состояния $[\widehat{y}]_{k+1(n)}$. Эта оценка содержится в частично сглаженной реализации $\widehat{\mathbf{y}}_k$ и есть результат предыдущего, k-го такта уравнения* (20)*. Прогноз $[\widehat{y}]_{k+1/k}$ состояния $[\widehat{y}]_{k+1(n)}$ модели есть последние n сглаженных отсчетов* (19) *частично сглаженной реализации $\widehat{\mathbf{y}}_k$.*

(c) *Векторы $f_{k+1} \in E$ в уравнении* (20) *для проекции $\widehat{\mathbf{y}}_{k+1}$ есть ортогонализирующие векторы прямого процесса ортогонализации Грама–Шмидта. Они определяются формулой $f_{k+1} = \Pi_k\eta_{k+1}$ этого известного процесса. Числа $a_{k+1} = \|f_{k+1}\|^{-2}$ есть обратные квадраты норм этих векторов. Здесь также $k = \overline{-1, N-1}$.*

5.3. Уравнения двусторонней (встречной) ортогонализации.

Теорема 5 (см. [7]). (a) *Процесс ортогонализации базиса A скользящего вектора определяется ортогонализирующими векторами встречных (прямого и возвратного) процессов ортогонализации Грама–Шмидта. Это, соответственно, векторы*

$$f_{k+1} = \Pi_k \eta_{k+1}, \qquad \widetilde{f}_{k+1} = \Pi_{1,k+1}\eta_0 \quad k = \overline{-1, N-1}.$$

Они подчиняются уравнениям двусторонней ортогонализации. Так названы перекрестные разностные уравнения

$$f_{k+1} = I^1 f_k - \widetilde{f}_k \theta^*_{k+1}, \qquad \widetilde{f}_{k+1} = \widetilde{f}_k - I^1 f_k \theta_{k+1},$$

$$k = \overline{0, N-1}, \qquad \Pi_{-1} = I, \qquad f_0 = \widetilde{f}_0 = \eta_0.$$

(b) *Скалярная функция $\theta_k = \theta(k)$ вычисляется по формулам*

$$\theta_{k+1} = a_k \mu_{k+1}, \qquad \mu_{k+1} = \left\langle \widetilde{f}_k, I^1 f_k \right\rangle, \quad \text{где} \quad a_k = \|f_k\|^{-2} = \|\widetilde{f}_k\|^{-2}.$$

(c) *Положительные скалярные множители a_k описываются нелинейными уравнениями*

$$a_{k+1} = \left(I - \theta_{k+1}\theta^*_{k+1}\right)^{-1} a_k, \qquad a_0 = \|\eta_0\|^{-2} = \|\alpha\|^{-2}_{E^{n+1}}.$$

(d) *Если вектор α коэффициентов РУ* (11) *имеет хотя бы одну не нулевую компоненту, то есть матрица A имеет полный ранг, равный $N+1$, то все обращаемые числа положительны:* $1 - |\theta_{k+1}|^2 > 0 \quad \forall \quad k = \overline{0, N-1}$.

Уравнения теорем 4, 5 содержат и уравнения фильтрации (фильтры) для состояний. Это указанный в п. (c) теоремы 4 минимальный набор n компонент в уравнениях (20). Им соответствуют достаточные n компонент ортогонализирующих векторов в уравнениях теоремы 5.

Следствие 1. *Для вычисления функций μ_k и a_k уравнений теоремы* 5, *а также прогнозов π_k теоремы* 4 *необходимо и достаточно вычисления n последних значимых компонент*

$$[f\,]_k = \{f_{ik}\}_{k+1}^{k+n} \qquad \text{и} \qquad \left[\widetilde{f}\,\right]_k = \left\{\widetilde{f}_{ik}\right\}_{k+1}^{k+n}, \qquad k = \overline{0, N},$$

ортогонализирующих $L+1$-векторов $f_k = \left\{f_{ik}\right\}_0^L$ и $\widetilde{f}_k = \left\{\widetilde{f}_{ik}\right\}_0^L$ в теореме 5. *Указанные компоненты ортогонализирующих векторов есть n граничных компонент этих векторов перед $N-k$ последними нулевыми их компонентами.*

Д о к а з а т е л ь с т в о. Оно следует из формулы для перекрестных множителей $\mu_{k+1} = \left\langle \widetilde{f}_k, I^1 f_k \right\rangle = \left\langle \widetilde{f}_k, \eta_{k+1} \right\rangle$, из определения векторов η_k в теореме 3, а также из рекуррентного уравнения модели в виде (12). С помощью него и вычисляются прогнозы $\widehat{y}_{k+1+n/k}$.

§ 6. Оптимизация уравнения модели

6.1. Уравнения для длины перпендикуляра. Введем множество

$$\Omega = \Omega_\omega = \left\{\Psi_\alpha \subset E \,:\, \langle \Psi_\alpha, A_\alpha \rangle = 0, \quad \alpha \in \omega \subset E^{n+1}\right\}$$

подпространств

$$\Psi = \Psi_\alpha = \ker D_\alpha = \{\widehat{\mathbf{y}} \in E \,:\, \langle \widehat{\mathbf{y}}, A_\alpha \rangle = 0, \quad \alpha \in \omega\} \in \Omega,$$

определяемых векторами α из некоторого допустимого их множества $\omega \subset E^{n+1}$.

Введем также множество

$$\widehat{\Omega} = \widehat{\Omega}_\omega = \widehat{\Omega}_\omega(\mathbf{y}) = \{\widehat{\mathbf{y}}_\alpha \in E \,:\, \widehat{\mathbf{y}}_\alpha = \Pi_\alpha \mathbf{y} \in \Psi_\alpha, \ \alpha \in \omega\} \subset \Omega$$

ортогональных проекций $\widehat{\mathbf{y}}_\alpha = \Pi_\alpha \mathbf{y}$ вектора исходных данных $\mathbf{y}$ на подпространства Ψ_α.

Теорема 2 приводит к следующим результатам для этих множеств, зависящих от вектора α и области его допустимых значений. Множество $\widehat{\Omega}$ зависит и от исходной реализации $\mathbf{y}$.

Теорема 6. (a) *Полное (оптимизация и по* $\widehat{\mathbf{y}}$, *и по* α*) решение задачи* (13), (14) *дается ближайшим (по расстоянию, определенному в* (13)*) к исходной реализации* $\mathbf{y}$ *элементом* $\widehat{\mathbf{y}}_{\widehat{\alpha}}$ *в множестве* Ω *подпространств* Ψ_α, $\alpha \in \omega$.

(b) *Этот элемент* $\widehat{\mathbf{y}}_{\widehat{\alpha}}$ *принадлежит множеству* $\widehat{\Omega}$ *проекций* $\widehat{\mathbf{y}} \in \widehat{\Omega}_\omega(\mathbf{y}) \subset \Omega_\omega$. *Он есть ортогональная проекция вектора исходных данных* $\mathbf{y}$ *на ближайшее к нему (минимизирующее длину перпендикуляра* $\Delta\mathbf{y} = \mathbf{y} - \widehat{\mathbf{y}}_\alpha$*) подпространство* $\Psi_{\widehat{\alpha}} = \Psi_{\widehat{\alpha}}(\mathbf{y}) \in \Omega$ *модели* (14).

(c) *Параметры* $\widehat{\alpha}$ *подпространства* $\Psi_{\widehat{\alpha}}$ *есть решение поставленной задачи идентификации — минимизации функционала* (13) *по параметрам* α *уравнения* (14).

Точку $\widehat{\mathbf{y}}_{\widehat{\alpha}}$ называем проекцией вектора $\mathbf{y}$ на множества Ω и $\widehat{\Omega}$. Проекцией вектора на произвольное множество понимаем точку множества, ближайшую к этому вектору. Эта точка определяет расстояние от данного вектора до множества.

Если множество — подпространство, то расстояние до него есть длина перпендикуляра.

Следствие 1. *Решение задачи идентификации единственно тогда и только тогда, когда*

(a) *или единственна проекция реализации* $\mathbf{y}$ *на множество* Ω;

(b) *или единственно ближайшее к реализации* $\mathbf{y}$ *подпространство* $\Psi_{\widehat{\alpha}} = \Psi_{\widehat{\alpha}}(\mathbf{y}) \in \Omega$;

(c) *или единственна ближайшая к реализации* $\mathbf{y}$ *точка* $\widehat{\mathbf{y}}_{\widehat{\alpha}}$ *в множестве* $\widehat{\Omega} = \widehat{\Omega}_\omega = \widehat{\Omega}(\mathbf{y})$.

Доказательство. Проекция $\widehat{\mathbf{y}}_\alpha$ реализации $\mathbf{y}$ на подпространство Ψ_α единственна. Длина перпендикуляра $\Delta\mathbf{y}_\alpha = \mathbf{y} - \widehat{\mathbf{y}}_\alpha$ минимальна среди всех векторов $\widehat{\mathbf{y}} \in \Psi_\alpha$. Отсюда следует, что для единственности задачи идентификации необходима и достаточна единственность подпространства $\Psi_{\widehat{\alpha}} \in \Omega$, для которого длина перпендикуляра $\Delta\mathbf{y}_{\widehat{\alpha}} = \mathbf{y} - \widehat{\mathbf{y}}_{\widehat{\alpha}}$ минимальна. Тогда и проекция вектора $\mathbf{y}$ на множество Ω подпространств Ψ_α единственна. Единственна тогда и точка в множестве $\widehat{\Omega}$, ближайшая к вектору $\mathbf{y}$. □

Определение 2. *Функционалом идентификации* $\widehat{J} = \widehat{J}_\alpha$ называется квадрат расстояния от реализации $\mathbf{y}$ до точек $\widehat{\mathbf{y}}_\alpha$ множества $\widehat{\Omega}_\omega(\mathbf{y})$. □

Эти точки, как указано в теореме 6, есть проекции вектора $\mathbf{y}$ на подпространства $\Psi_\alpha \in \Omega_\omega \supset \widehat{\Omega}_\omega(\mathbf{y})$. Подпространство Ψ_α определяется *только* вектором коэффициентов α.

Теорема 7. (a) *Функционал идентификации* $\widehat{J}$ *есть значения функционала* J (13) *на проекциях* $\widehat{\mathbf{y}}_\alpha$ (18) *реализации* $\mathbf{y}$ *на подпространства модели* $\Psi_\alpha = S_\alpha^\perp$, $\forall\, \alpha \in \omega$. *Эти значения*

$$\widehat{J}_\alpha = \rho^2(\alpha) = \rho_N^2 = \langle \mathbf{y}, P(A)\mathbf{y} \rangle = \langle A, \mathbf{y} \rangle \langle A, A \rangle^{-1} \langle \mathbf{y}, A \rangle$$

представляют собой квадраты длины перпендикуляров $\Delta\mathbf{y}_\alpha = \mathbf{y} - \widehat{\mathbf{y}}_\alpha$ *на подпространства* Ψ_α.

(b) *Значения* ρ_k^2 *функционала* J (13) *на последовательности частично сглаженных реализаций* $\widehat{\mathbf{y}}_k$, $k = \overline{0, N}$ *описываются уравнениями теоремы* 4 *для проекций и уравнениями встречной ортогонализации теоремы* 5 *для теплицевых базисов* $A_k = |\eta_i|_0^k$ *этих подпространств.*

(c) *Для* ρ_k^2 *имеем уравнение*

$$\rho_{k+1}^2 = \rho_k^2 + \pi_{k+1}^* a_{k+1} \pi_{k+1}, \quad \text{где} \quad \rho_{(-1)}^2 = 0, \quad k = \overline{-1, N-1},$$

а функция π_k *вычисляется с помощью уравнений утверждения* (b) *этой теоремы.*

Доказательство. Из теоремы 6 следует, что задача идентификации сводится к задаче поиска подпространства $\Psi_{\widehat{\alpha}} \in \Omega$, ближайшего к исходной реализации $\mathbf{y}$. Утверждение (a) следует из теоремы 6 и из формул (17), (18) теоремы 2. Утверждение (b) следует из теоремы 3 и формул (17), (18), если в них положить $A = A_k$. Формула п. (c) следует из формулы теоремы 4. Уравнения для последовательных (по k) вычислений выводятся из уравнений сглаживания теоремы 4 и уравнений встречной ортогонализации теоремы 5. □

Следствие 1. *Значение функционала МНК* (13) *на проекции* $\widehat{\mathbf{y}}_\alpha$ *есть сумма взвешенных норм ошибок прогнозов* π_k *процесса обновления* $\boldsymbol{\pi} = \{\pi_k\}_0^N$: $\rho_N^2 = \sum_{k=0}^{N} \pi_k^* a_k \pi_k$.

6.2. Безусловная оптимизация. Теоремами 6, 7 показано, что задача минимизации функционала J (13) по параметрам α условий (14) должна решаться минимизацией значений функционала (13) на проекциях $\widehat{\mathbf{y}}_\alpha$. Эти значения есть квадраты ρ^2 длин $\rho = \rho(\alpha) = \|\Delta\mathbf{y}_\alpha\|$ перпендикуляров $\Delta\mathbf{y}_\alpha = \mathbf{y} - \widehat{\mathbf{y}}_\alpha$.

Теорема 8. *Определенный на множестве $\widehat{\Omega} = \widehat{\Omega}_\omega(\mathbf{y})$ проекций $\widehat{\mathbf{y}}_\alpha = \Pi_\alpha\mathbf{y}$ функционал идентификации*

$$\widehat{J}_\alpha = \rho_N^2(\alpha) = \|P_\alpha\mathbf{y}\|^2 = \mathbf{y}^*A(A^*A)^{-1}A^*\mathbf{y} = m^*C_\alpha^{-1}m \tag{21}$$

зависит только от исходной реализации $\mathbf{y}$ *и коэффициентов* α *РУ* (14).

Следствие 1. *Функционал идентификации инвариантен к длине* $\|\alpha\|$ *вектора* α.

Следствие 2. *Функционал идентификации может быть приведен к виду нелинейной «квазиквадратичной» формы*

$$\widehat{J}_\alpha = \rho_\alpha^2 = \alpha^T Q_\alpha \alpha^{*T} = (\alpha)^* Q_\alpha(\alpha), \quad \text{где} \quad Q_\alpha = V^*C_\alpha^{-1}V. \tag{22}$$

Матрица $Q = Q_\alpha$ *неконстантна по* α *на сфере* $\mathcal{S}(c) = \{\alpha \in E^{n+1} \,:\, \|\alpha\| = const = c\}$.

Доказательство. Утверждение теоремы 8 есть следствие теорем 6, 7 и формул (18). Утверждение следствия 1 очевидно из формулы (21). Следствие 2 — результат применения тождества леммы 2 в формуле (21). □

Формулы теоремы 2 описывают аналитический этап минимизации функционала аппроксимации (13). Он осуществляется по параметрам (n локальным условиям) реализаций $\widehat{\mathbf{y}}_\alpha \;\forall\, \alpha \in \omega$. Нахождение оптимального вектора α из ω есть поисковый этап минимизации функционала (13). Он осуществляется безусловной минимизацией функционала идентификации (22).

Теорема 9. *Вектор–функция* $\widehat{\mathbf{y}}_\alpha = \Pi_\alpha\mathbf{y} \,:\, \omega \to \widehat{\Omega}$, $\alpha \in \omega$ *устанавливает на реализации* $\mathbf{y}$ *взаимно однозначное соответствие между множеством* ω *допустимых значений вектора коэффициентов* α *уравнения* (11), (12) *и множеством* $\widehat{\Omega} = \widehat{\Omega}_\omega(\mathbf{y}) \subset \Omega \subset E$ *проекций* $\widehat{\mathbf{y}}_\alpha$ *этой реализации на подпространства моделей* $\Psi_\alpha \subset \Omega \subset E$.

Доказательство. Каждая точка множества $\widehat{\Omega} = \widehat{\Omega}_\omega$ есть единственная точка $\widehat{\mathbf{y}}_\alpha = \Pi_\alpha\widehat{\mathbf{y}}$ из подпространства Ψ_α, $\forall\, \alpha \in \omega$, составляющих множество подпространств Ω_ω. Эти точки есть основания перпендикуляров $\Delta\mathbf{y} = \mathbf{y} - \widehat{\mathbf{y}}_\alpha \in \Psi_\alpha^\perp = S(A_\alpha)$.

Обратное тоже верно. Из определения МСВ $A(\alpha)$ (15) следует, что двум различным (не нулевым и с точностью до постоянного множителя) векторам α (то есть направлениям α) из ω соответствуют различные размерности n подпространства $\Psi_\alpha = S_\alpha^\perp = S^\perp\big(A(\alpha)\big)$. Им соответствуют различные проекции $\widehat{\mathbf{y}}_\alpha$ из $\Psi_\alpha \subset \Omega$ и различные точки из $\widehat{\Omega}$.

Таким образом, отображение $\widehat{\mathbf{y}}_\alpha = \Pi_\alpha\mathbf{y} \,:\, \omega \to \widehat{\Omega} = \widehat{\Omega}_\omega$ взаимно однозначно. □

Следствие 1. *Следующие две задачи оптимизации эквивалентны:*

$$\widehat{\mathbf{y}}_{\widehat{\alpha}} = \arg\min_{\widehat{\mathbf{y}} \in \widehat{\Omega}_\omega} J \iff \widehat{\alpha} = \arg\min_{\alpha \in \omega} \widehat{J}_\alpha.$$

Доказательство. Элементы множества $\widehat{\Omega} = \widehat{\Omega}_\omega$ зависят только от искомого вектора $\alpha \in \omega$, поэтому в этом и только в этом множестве осуществляется поиск наилучшей аппроксимации $\widehat{\mathbf{y}}_{\widehat{\alpha}}$ исходной реализации $\mathbf{y}$. □

Следствие 2. *Вариационная задача аппроксимации — условной минимизации функционала* (13) *— сводится к задаче безусловной минимизации функционала идентификации* $\widehat{J}$ (22).

6.3. Геометрическая интерпретация. Обозначим через $\mathcal{S}_r^{L+1}(c)$ сферу в $E = E^{L+1}$ с центром r и радиусом c. Определим:

$$\mathcal{S} = \mathcal{S}\left(\mathbf{y}/2\right) = \mathcal{S}_{\|\mathbf{y}/2\|}^{L+1}\left(\mathbf{y}/2\right) = \left\{\mathbf{x} \in E \,:\, \big\|\mathbf{x} - \mathbf{y}/2\big\| = \|\mathbf{y}\|/2\right\}.$$

Теорема 10. *Множество* $\widehat{\Omega} = \widehat{\Omega}_\omega$ *есть множество точек, максимально удаленных от начала координат в пересечениях* $\mathbf{s} = \mathbf{s}_\alpha = \Psi_\alpha \cap \mathcal{S}$ *подпространств* Ψ_α, $\alpha \in \omega$ *со сферой* $\mathcal{S} \subset E = E^{L+1}$ *радиуса* $\|\mathbf{y}\|/2$ *и с центром в точке* $\mathbf{y}/2$.

Доказательство. Все точки $\mathbf{x} \in \mathcal{S} = \mathcal{S}(\mathbf{y}/2)$ сферы $\mathcal{S}$, в том числе и принадлежащие пересечению $\mathbf{s} = \Psi \cap \mathcal{S}$, имеют два свойства:

(a) $\langle \mathbf{x}, \mathbf{y} - \mathbf{x} \rangle = 0$ — ортогональность направлений из $\mathbf{x}$ на нулевую точку $\mathbf{0}$ и точку $\mathbf{y}$;

(b) $\|\mathbf{y} - \mathbf{x}\|^2 + \|\mathbf{x}\|^2 = \|\mathbf{y}\|^2$ — теорема Пифагора.

В силу (b) и минимальности длины перпендикуляра, из всех точек $\mathbf{x}$ указанного пересечения $\mathbf{s}$ (а это сфера) ортогональной, в силу свойства (a), проекцией $\mathbf{y}$ на $\widehat{\mathbf{y}}$ является вектор $\mathbf{x}$ максимальной длины. Это точка диаметрально противоположна нулевой точке $\mathbf{0}$ в E. Последняя принадлежит и сфере $\mathbf{s}$. Отсюда следует утверждение теоремы для всех $\Psi_\alpha \in \Omega$. □

Следствие 1. *Имеет место эквивалентность следующих двух множеств, равных* $\widehat{\Omega}$:

$$\Big\{\mathbf{x} \in \Omega \,:\, \langle \mathbf{y} - \mathbf{x}, \Psi_\alpha \rangle = 0,\ \Psi_\alpha \in \Omega \Big\} = \widehat{\Omega};$$
$$\Big\{\mathbf{x} \in \mathcal{S} \,:\, \mathbf{x} = \arg\max \|\mathbf{z}\| \text{ при } \ \mathbf{z} \in \mathcal{S} \cap \Psi_\alpha,\ \Psi_\alpha \in \Omega \Big\} = \widehat{\Omega}.$$

Следствие 2. *Эквивалентны следующие два определения ортогональной проекции* $\widehat{\mathbf{y}} = P(\Psi)\mathbf{y}$ *некоторого вектора* $\mathbf{y} \in E$ *на некоторое подпространство* $\Psi \subset E$:

$$\widehat{\mathbf{y}} = \arg\max \|\mathbf{z}\| \quad \textit{при} \quad \mathbf{z} \in \mathcal{S}(\mathbf{y}/2) \cap \Psi \quad \Longleftrightarrow \quad \widehat{\mathbf{y}} = \arg\min \|\mathbf{y} - \mathbf{z}\| \quad \textit{при} \quad \mathbf{z} \in \Psi.$$

Пример в E^3, иллюстрирующий некоторые результаты раздела **6**, приведен в работах [4, 11].

§ 7. Функционал идентификации

7.1. Проблема поиска экстремума. Вопросы о типе экстремумов функционала идентификации, об их количестве, о наличии глобального экстремума остаются открытыми. Причина — сложность исследований этого функционала как аналитическими, так и экспериментальными методами. Выражения для вектора градиента и матрицы вторых производных (гессиана) функционала идентификации $\widehat{J}_\alpha$ (21), (22) получены автором в [4, 12]. Из этих формул следует, что они сложны как для вычислений, так и для анализа. Из приведенных формул для первых и вторых производных (их доказательства и выводы даны в [12]) можно видеть, что указанные производные непрерывны и по α и по $\mathbf{y}$. Следовательно, можно утверждать, что имеет место непрерывная зависимость решения $\widehat{\alpha}$ задачи идентификации от исходных данных $\mathbf{y}$.

Эксперименты с итерациями Ньютона–Раффсона, градиентного спуска показали, что они сходятся только при начальных условиях для коэффициентов, близких к решению. Настолько близких, что применение такого рода итераций не представляет практического интереса.

Экспериментальные исследования автора этой статьи, численные расчеты автора работы [13], а также результаты экспериментов В.Г. Демиденко (выполненных им для работы [14]), сообщенные автору, показали, что причина этого — очень острый характер глобального экстремума и сильная овражность в большой окрестности его расположения.

Автором найден неградиентный способ минимизации функционала (22). Экспериментально показано, что с его помощью осуществляется быстрое попадание итераций в малую окрестность его глобального экстремума. Часто она оказывается достаточной для практики.

Далее функции и коэффициенты уравнений считаются действительными.

7.2. О вариационной и орторегрессионной идентификации. Обращает на себя внимание внешнее сходство задачи минимизации функционала идентификации (22) — задачи ВИ (вариационной идентификации) — с известной задачей поиска минимума отношения квадратичных форм. Простейшая задача такого рода означает минимизацию функционала

$$r_\alpha^2 = \widehat{J}_{or} = \alpha^T Q \alpha, \qquad Q(\alpha) = V^T V \Big/ \|\alpha\|^2 = Q_{or}. \tag{23}$$

Сходство функционалов (22) и (23) не случайно. Одна из задач «тотального» МНК сводится к решению системы уравнений вида (9), матрица которой $V = V_{ist} + \Xi$ искажена ошибками Ξ. Здесь V_{ist} — истинная матрица системы. Пусть существует единственное значение α_{ist} такое, что $V_{ist}\alpha_{ist} = 0$. Задачу оценивания вектора α на фоне статистических шумов Ξ в приведенной линейной зависимости называют задачей ортогональной регрессии (ОР) [11].

Простейшая задача ОР в наших обозначениях ставится так: минимизировать

$$J_{or} = \left\|V - \widehat{V}\right\|^2, \quad \text{если} \quad \widehat{V}\alpha = 0, \tag{24}$$

где $\|V\|^2 = \mathrm{Sp}(V^T V)$. Функционал (23) есть значение функционала (24) на проекциях $\widehat{V}_\alpha$. Следовательно, функционал (23) есть функционал идентификации задачи ОР (24). Отметим еще близкие свойства задач (24) и (13,14).

Во-первых, как и в задаче ВИ (13,14), условная минимизация (24) сводится к безусловной минимизации. В задаче ВИ это функционал (22) $\widehat{J} = \widehat{J}_{vi}$, в задаче ОР — функционал (23) $\widehat{J}_{or}$. Во-вторых, имеется сходство условий минимизации в задачах ВИ (13,16) и ОР (24). В задаче ВИ это $A^*\widehat{\mathbf{y}} \equiv \widehat{V}\alpha = 0$, в задаче ОР это $\widehat{V}\alpha = 0$. В-третьих, функционал идентификации ОР (23) также инвариантен к длине вектора α. Так что имеет место и внутреннее (учет ошибок в матрице системы уравнений), и внешнее сходство задач оценивания в ВИ и ОР. Наконец, проектирование в задачах ВИ и ОР осуществляется на n-подпространства (размерности n).

Однако постановка (24) есть $N+1$ задач проектирования $\widehat{v}_i = (I - \alpha\alpha^T \big/ \|\alpha\|^2)v_i$ отдельных строк v_i^*, $i = \overline{0, N}$ матрицы V на n-подпространства $\alpha^\perp$ в E^{n+1}. Это следует из того, что функционал идентификации (23) задачи ОР (24) инвариантен к любым перестановкам строк матрицы V. Задача ВИ есть проектирование (17) $\widehat{\mathbf{y}} = \Pi\mathbf{y}$ *одной* реализации $\mathbf{y} \in E = E^{L+1}$ на n-подпространство $S^\perp(A_\alpha)$, где $\big((L+1)\times(N+1)\big)$-матрица A_α есть МСВ α. Реализация $\mathbf{y}$ определяет матрицу V в соответствии с тождеством $A^*\mathbf{y} = V\alpha$ (16).

Это тождество определяет взаимозависимость компонент матрицы $V(\mathbf{y})$ в задаче ВИ. В результате если в задаче ОР оценивается $(N+1)\times n$ компонент матрицы $\widehat{V}$, то в задаче ВИ — только n компонент реализации $\widehat{\mathbf{y}}$. Они определяют матрицу $\widehat{V}$ благодаря тождеству (16).

§ 8. Неградиентная оптимизация на сфере

8.1. Линейная минимальная спектральная задача. Минимум отношения квадратичных форм (23) — $\alpha^T Q\alpha \big/ \|\alpha\|^2$ — достигается на собственном векторе (СВ) минимального характеристического числа (МХЧ) матрицы $Q_{or} = V^T V \big/ \|\alpha\|^2$. Матрица Q_{or} функционала идентификации $\widehat{J}_{or} = \alpha^T Q_{or}\alpha$ (23) задачи ОР (24) постоянна на сфере $\mathcal{S}_c^{n+1}(0)$, $c > 0$. Таким образом, мы приходим к задаче поиска минимальной оси эллипсоида. Это фигура может быть описана как множество в E^{n+1}. Например,

$$\Phi_{or} = \Big\{x \in E^{n+1} \,:\, x = \alpha\cdot\left(\alpha^T Q_{or}\alpha\right), \qquad Q_{or} = V^T V \Big/ \|\alpha\|^2, \quad \|\alpha\| = 1\Big\}. \tag{25}$$

Минимум в методе ОР глобален, если МХЧ не кратное. Если допустимое множество ω для α есть вся сфера $\mathcal{S}_c^{n+1}(0)$, то отыскивается единственная минимальная ось $\widehat{\alpha}$ эллипсоида (25), то есть отыскивается одна из двух ближайших к началу координат точек $\widehat{\alpha}$ или $-\widehat{\alpha}$ множества (25). Эта задача поиска СВ МХЧ и названа минимальной спектральной задачей.

Итерации с обратной матрицей Q для минимизации функционала (23) в задаче ОР (24)

$$\widetilde{\alpha}_{[k+1]} = Q_{or}^{-1}\alpha_{[k]}, \qquad \alpha_{[k+1]} = \widetilde{\alpha}_{[k+1]} \Big/ \|\widetilde{\alpha}_{[k+1]}\| \tag{26}$$

начинаются с некоторого вектора $\alpha_{[0]}$. Основное к нему требование — ненулевая проекция на СВ матрицы Q_{or}, соответствующий ее МХЧ. Итерации (26) сходятся (к СВ МХЧ) тем быстрее, чем острее экстремум, то есть чем изолированнее МХЧ матрицы Q, чем дальше оно отстоит от следующего, большего по величине.

Как показывают эксперименты, скорость сходимости итераций вида (26) в малую окрестность СВ МХЧ определяется уровнем ошибок: чем он меньше, тем изолированнее МХЧ, острее экстремум, выше скорость сходимости. Этот факт характерен и для нелинейной минимальной спектральной задачи и аналогичных итераций в задаче ВИ.

8.2. Нелинейная минимальная спектральная задача. В задаче ВИ мы ищем точку, ближайшую к началу координат, на более сложной фигуре. Пусть $c = 1$. Тогда это множество

$$\Phi_{vi} = \left\{x \in E^{n+1} \,:\, x = \alpha \cdot \left(\alpha^T Q_{vi}\alpha\right), \qquad Q_{vi} = V^T \left(A_\alpha^T A_\alpha\right)^{-1} V, \quad \|\alpha\| = 1\right\}. \tag{27}$$

Оказывается, что эффективной (по скорости и величине области сходимости) итерационной процедурой (ИП) нахождения этой точки (глобального минимума функционала идентификации (22)) является ИП, подобная процедуре (26). Здесь это итерации

$$Q_{vi[k]} = Q_{vi}(\alpha_{[k]}), \qquad \widetilde{\alpha}_{[k+1]} = Q_{vi[k]}^{-1}\alpha_{[k]}, \qquad \alpha_{[k+1]} = \widetilde{\alpha}_{[k+1]}\Big/\|\widetilde{\alpha}_{[k+1]}\|, \tag{28}$$

начиная с некоторого значения $\alpha_{[0]}$ [1, 3, 4, 11, 12, 14].

Как и в (26), здесь применена нормировка итерируемого вектора коэффициентов к единичной длине. Это более надежно, чем нормировать итерируемый вектор к единичному значению одной из компонент, например компоненты α_n. В случае малости значений выбранной для такой нормировки компоненты вектора α (в процессе итераций или вблизи решения) это приведет к большим вычислительным ошибкам. Если МХЧ равно нулю (это значит, что модель (11) точна для реализации $\mathbf{y} = \widehat{\mathbf{y}}$, этот случай предусмотрен в теореме 1), то итерации (26) и (28) осуществляются с матрицей $Q_\varepsilon = Q + \varepsilon I_{n+1}$, где $\varepsilon > 0$ достаточно мало.

§ 9. Заключение

Итерации (28) сходятся в малую окрестность ближайшей к началу координат точки $\widehat{\alpha}$ фигуры (27) достаточно быстро. Если этого не происходит, то это означает, что длина $L + 1$ реализации $\mathbf{y}$ недостаточна для того уровня измерительных и/или структурных ошибок, который в ней содержится. В этом случае получение устойчивого решения маловероятно. Требуется новая информация: либо дополнительные отсчеты в реализации, либо регуляризация с дополнительной априорной информацией о параметрах α (см., например, п. **3** в разделе **2**).

Вопрос о сходимости итераций (28) в малую окрестность точки $\widehat{\alpha}$ множества (27) может быть иногда решен и априори, визуально. Если экспериментатор «видит» в искаженной ошибками реализации $\mathbf{y}$ интересующий его переходной процесс (ПП), то, вероятнее всего, он будет найден. Коэффициенты $\widehat{\alpha}$ его РУ будут найдены итерациями (28).

Если оценка коэффициентов РУ из (28) получена, оптимальный ПП может быть вычислен двумя способами. Первый — полностью по формулам теорем 2, 4, 5. Этот способ не рационален. Предпочтительнее исходить из следствия 1 теоремы 5. Достаточно использовать уравнения теорем 2, 4, 5 только для начальных условий оптимального аппроксимирующего ПП. Затем вычислить этот процесс с помощью рекуррентного уравнения модели (12).

Итерации (28) могут дать приемлемое решение и за один шаг, особенно если идентификация наблюдаемого объекта осуществляется в режиме мониторинга или слежения. В подобных случаях первое приближение $\alpha_{[0]}$ в (28) может быть задано не очень удаленным от решения $\widehat{\alpha}$ и благодаря полученным в статье быстрым алгоритмам, основанным на уравнениях встречной ортогонализации теоремы 5, процесс идентификации может быть реализован в реальном времени. Описанный процесс важен и для некоторых задач автоматического управления, в частности, для задач построения адаптивных и самонастраивающихся систем с идентификатором в обратной связи.

Автор признателен коллегам: А.А. Ломову за интересную ссылку [6] и В.Г. Демиденко за возможность ознакомления с его экспериментальными результатами.

СПИСОК ЛИТЕРАТУРЫ

1. Егоршин А.О. Метод наименьших квадратов и быстрые алгоритмы в вариационных задачах идентификации и фильтрации (метод ВИ) // Автометрия. 1988. № 1. С. 30–42.
2. Аоки М. Введение в методы оптимизации. М.: Наука, 1977.
3. Егоршин А.О. Вычислительные замкнутые методы идентификации линейных объектов // Оптимальные и самонастраивающиеся системы: сб. / Институт автоматики и электрометрии СО АН СССР. Новосибирск, 1971. С 40–53.
4. Егоршин А.О. Оптимизация параметров стационарных моделей в унитарном пространстве // Автоматика и телемеханика. 2004. № 12. С. 29–48.
5. Кутателадзе С.С. Сергей Соболев и Лоран Шварц: две судьбы, две славы // Препринт Института математики им. С.Л. Соболева СО РАН. № 121. 2003.
6. Bellman R. On separation of exponentials // Bollatino Della Unione Matematica. 1960. Ser. III. Vol. 15. № 1. P. 38–39.
7. Егоршин А.О. Об одной вариационной задаче сглаживания // Вестник Удмуртского университета. Математика. Механика. Компьютерные науки. 2011. Вып. 4. С. 9–22.
8. Эйкхофф П. Основы идентификации систем управления. М.: Мир, 1975.
9. Льюнг Л. Идентификация систем. М.: Наука, 1991.
10. Острем К.Ю. Введение в стохастическую теорию управления. М.: Мир, 1973.
11. Егоршин А.О. Об одном способе оценки коэффициентов моделирующих уравнений для последовательностей // Сибирский журнал индустриальной математики. 2000. Т. 3. № 2. С. 78–96.
12. Егоршин А.О. Об отслеживании параметров экстремума в вариационной задаче идентификации // Вестник НГУ. Серия: Математика. Механика. Информатика. 2011. Т. 11. Вып. 1. С. 95–114.
13. Костин В.И. О точках экстремума одной функции // Управляемые системы: сб. / Институт математики СО АН СССР. Новосибирск, 1984. Т. 24. С. 35–42.
14. Демиденко В.Г. Оценки устойчивости в задаче идентификации коэффициентов линейных разностных уравнений // Неклассические уравнения математической физики: сб. / Институт математики СО РАН. Новосибирск, 2010. С. 62–81.

Поступила в редакцию 20.04.2012

Егоршин Алексей Олегович, к. ф.-м. н., старший научный сотрудник, Институт математики имени С. Л. Соболева, СО РАН, 630090, Россия, г. Новосибирск, 90, просп. акад. В. А.Коптюга, 4.
E-mail: egorshin@math.nsc.ru

A. O. Egorshin

On one variational problem of the piecewise–linear dynamical approximation

Keywords: variational identification, algebraic identification, piecewise–linear dynamical approximation, orthogonal regression, non-gradient optimization.

Mathematical Subject Classifications: 65F25; 15A03

Some properties of the discrete variational problem of the dynamic approximation in the complex Euclidean $(L+1)$-dimensional space are studied here. It generalizes familiar problems of the mean square polynomial approximation of the functions given on the finite interval in accordance with their references. In the problem under consideration sequence approximation $\mathbf{y} = \{y_i\}_0^L$ of the references of the function $y(t) \in L^2[0,T]$, $T = Lh$ on the lattice I_h is achieved by solving homogeneous linear differential equations or difference equations of the given order n with constant but possibly unknown coefficients. Thus, it is shown that in the latter case the approximation problem also includes the identification problem. The analysis of its properties is the main subject of the article. The problem is set to find vector of coefficients α of difference equation $\sum_0^n \widehat{y}_{i+k}\alpha_i = 0$, where $k = \overline{0, L-n}$. Coefficients α and initial conditions of the transient process $\widehat{\mathbf{y}}$ of this equation are optimized. The optimization purpose is to achieve the best approximation of the dynamic process $\mathbf{y} \in E$ being considered here. The approximation criterion is a minimum of the quantity $\|\mathbf{y} - \widehat{\mathbf{y}}\|_E^2$. The variational problem under study is shown to be reduced to the problem of projecting vector $\mathbf{y}$ in E on the kernels of the difference operators with unknown coefficients $\alpha \in \omega \subset \mathcal{S} \subset E^{n+1}$, where α is a direction, S is a sphere or a hyperplane. The problem under study is shown to be related to the problems of the discretization and identifiability. In this case vector coordinates $\mathbf{y} \in E$ is an exact solution of differential equation on the lattice I_h and $\mathbf{y} = \widehat{\mathbf{y}}$. The problem of the variational identification is compared with algebraic methods of identification. The orthogonal complement to the kernels of the difference operators are shown to always have Toeplitz basis. This results in fast projecting algorithms of computation. The problem of finding optimal vector $\widehat{\alpha}$ is shown to be reduced to the problem of the absolute minimization of the identification functional depending on the direction α in E^{n+1}. The iterative procedure of its minimization on a sphere with wide domain and high speed of convergence is presented here. The variational problem considered here can be applied in mathematical modeling for control problem and research purposes. On the finite intervals, for example, it is possible to use piecewise-linear dynamic approximations of the complex dynamic processes with difference and differential equations of the specified type.

REFERENCES

1. Egorshin A.O. Least square method and fast algorithms of identification and filtration (VI-method), *Avtometriya*, 1988, no. 1, pp. 30–42.

2. Aoki M. *Introduction to optimization techniques*, Los Angeles: University of California, 1971, 335 p.

3. Egorshin A.O. Numerical closed methods of linear objects identification, *Optimal'lnye i samonastraivayushchiesya sistemy: sbornik statei* (Optimal and self-adjusting systems: Transactions), Institute of Automation and Electrometry, Siberian Branch of the Academy of Sciences of USSR, Novosibirsk, 1971, pp. 40–53.

4. Egorshin A.O. Optimization of parameters of stationary models in a unitary space, *Automation and Remote Control*, 2004, vol. 65, no. 12, pp. 1885–1903.

5. Kutateladze S.S. *Sergey Sobolev and Loran Schwarz: two fates, two glories*, Preprint of the Sobolev Institute of Mathematics, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, 2003, no. 121.

6. Bellman R. On separation of exponentials, *Bollatino Della Unione Matematica*, 1960, ser. III, vol. 15, no. 1, pp. 38–39.

7. Egorshin A.O. On one variational smoothing problem, *Vestn. Udmurt. Univ. Mat. Mekh. Komp'yut. Nauki*, 2011, no. 4, pp. 9–22.

8. Eykhoff P. *System identification, parameter and state estimation*, London, New York [etc.]: Wiley–Interscience, 1974, 555 p.

9. Ljung L. *System identification: theory for the user*, Englewood Cliffs, New Jersey: Prentice-Hall, 1987, 519 p.

10. Astrom K.J. *Introduction to stochastic control theory*, N.Y., London: Academic Press, 1970, 299 p.

11. Egorshin A.O. On one estimation method of modeling equation coefficients for sequences, *Sib. Zh. Ind. Mat.*, 2000, vol. 3, no. 2, pp. 78–96.

12. Egorshin A.O. On extremum parameters tracing in variational identification problem, *Vestn. Novosib. Gos. Univ., Ser. Mat. Mekh. Inform.*, 2011, vol. 11, no. 3, pp. 95–114.

13. Kostin V.I. On extremum points of one function, *Upravl'aemye sistemy: sbornik trudov* (Control systems: Transactions), Sobolev Institute of Mathematics, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, 1984, vol. 24, pp. 35–42.

14. Demidenko V.G. Stability estimations in identification problem of the linear difference equation coefficients, *Neklassicheskie uravneniya matematicheskoi fiziki: sbornik statei* (Nonclassical equation of mathematical physic: Transactions), Sobolev Institute of Mathematics, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, 2010, pp. 62–81.

Received 20.04.2012

Egorshin Aleksei Olegovich, Candidate of Physics and Mathematics, Senior Researcher, Sobolev Institute of Mathematics, Siberian Branch of the Russian Academy of Sciences, pr. Acad. Koptyug, 4, Novosibirsk, 630090, Russia.
E-mail: egorshin@math.nsc.ru