УДК 004.8, 004.93,159.95 Дата подачи статьи: 29.12.17
DOI: 10.15827/0236-235X.121.099-101 2018. Т. 31. № 1. С. 099–101

РЕГУЛЯРИЗОВАННЫЕ ОРТОГОНАЛЬНЫЕ МОДЕЛИ ВЕРОЯТНОСТНЫХ ХАРАКТЕРИСТИК С УСЛОВИЕМ ВЫПОЛНЕНИЯ ИХ ОСНОВНЫХ СВОЙСТВ

С.А. Прохоров [1], *д.т.н., профессор, зав. кафедрой, sp.prokhorov@gmail.com*
И.М. Куликовских [1], *к.т.н., доцент, kulikovskikh.i@gmail.com*

[1] *Самарский национальный исследовательский университет им. академика С.П. Королева, Московское шоссе, 34, г. Самара, 443086, Россия*

Представление модели вероятностной характеристики в виде разложения в ряд Фурье – один из наиболее эффективных способов понижения сложности модели. Широкое применение при решении задач идентификации, фильтрации и анализе динамических систем нашли ортогональные системы функций, позволяющие представить длинные временные последовательности в виде более короткого спектра разложения.

Для повышения устойчивости и достоверности ортогональной модели вероятностной характеристики используются различные подходы: оптимизация параметров ортогональной модели; применение устойчивых численных схем для вычисления коэффициентов разложения; построение взвешенных моделей, позволяющих учесть особенности поведения динамических систем, в частности, нестационарных систем и систем с запаздыванием, и т.д.

Особый интерес представляет метод, учитывающий необходимость выполнения ортогональной моделью основного свойства вероятностной характеристики. Он сводится к необходимости вычисления поправок для коэффициентов разложения, величина которых рассчитывается с учетом основных свойств вероятностных характеристик и базисных функций. Таким образом, построение совокупности ортогональных моделей требует различных реализаций в зависимости от постановки задачи и анализируемой динамической системы.

Данная работа обобщает предложенный ранее математический аппарат для реализации описанного метода, который позволяет следующее: предложить обобщенное описание поправочных коэффициентов для учета основных свойств функциональных характеристик в матричной форме; показать, что предложенные оценки аналогичны оценкам коэффициентов, регуляризованным по норме L_2; ввести ядровую функцию для описания основных свойств функциональных характеристик. Результатом работы является теоретическое обоснование устойчивости предложенных регуляризованных оценок.

Предложенный математический аппарат для описания регуляризованных ортогональных моделей позволяет свести оценку поправочных коэффициентов для различных вероятностных характеристик и базисных функций к единому алгоритму и, как следствие, существенно повысить вычислительную эффективность ее программной реализации.

Ключевые слова: *регуляризация, ядровые функции, ортогональные модели, вероятностные характеристики.*

Ортогональные модели вероятностных характеристик получены разложением в ряд по ортогональным базисам эмпирических зависимостей, таких как оценки плотностей и функций распределения вероятностей, корреляционных и спектральных функций и плотностей мощности [1–4] и т.п. Так как такой подход к построению моделей универсален, он является непараметрическим, что понижает интерпретируемость результирующих оценок. Один из способов повышения достоверности и устойчивости оценок вероятностных характеристик, полученных по ортогональной модели, – наложение дополнительного условия [1]. Данное условие формализуется исходя из необходимости выполнения основных свойств вероятностных характеристик, таких как значение в нуле и единице функции распределения, значение в нуле корреляционной функции, нормировка спектральной плотности мощности и т.д. В соответствии с этим условием оценки коэффициентов разложения пересчитываются с учетом поправочных коэффициентов.

В работах [1, 2] более детально изложены специфика предлагаемого подхода, а также аналитические соотношения поправочных коэффициентов для различных вероятностных характеристик при разложении в ряды Лагерра, Лежандра, Дирихле, Чебышева, включая обобщенные базисы Сонина–Лагерра и Якоби [2].

Однако построение совокупности ортогональных моделей требует различных алгоритмических и программных реализаций в зависимости от постановки задачи и анализируемой динамической системы [1–4]. Данная работа обобщает предложенный ранее математический аппарат для формализации описанного метода, предлагая описание поправочных коэффициентов для учета основных свойств функциональных характеристик в матричной форме. Показано, что предложенные оценки аналогичны оценкам коэффициентов, регуляризованным по норме L_2, что позволяет свести задачу вычисления поправочных коэффициентов к более общему представлению в форме задачи регуляризации исходных оценок коэффициентов. Кроме того, в работе дано теоретическое обоснование устойчивости полученных оценок через введение ядровой функции для описания основных свойств функциональных характеристик.

Постановка задачи

Поставим задачу построения ортогональной модели в матричной форме. Введем следующие обозначения. Пусть $\boldsymbol{f} \in \mathbf{R}^n$ – вероятностная характеристика, $\Phi \in \mathbf{R}^{n\times m}$ – система ортогональных функций, $\boldsymbol{\beta}$, $\boldsymbol{b}$, $\boldsymbol{c} \in \mathbf{R}^m$ – векторы стандартных, дополненных и поправочных коэффициентов Фурье соответственно. Тогда величина невязки [5–10] $\Delta \in \mathbf{R}$ может быть задана в виде

$$\Delta(\boldsymbol{\beta}) = (\boldsymbol{f} - \Phi\boldsymbol{\beta})^{\mathrm{T}}(\boldsymbol{f} - \Phi\boldsymbol{\beta}), \quad (1)$$

откуда $\partial\Delta(\boldsymbol{\beta})/\partial\boldsymbol{\beta} = -2\Phi^{\mathrm{T}}(\boldsymbol{f} - \Phi\boldsymbol{\beta})$.

Потребуем выполнения необходимого условия

$$\partial\Delta(\boldsymbol{\beta})/\partial\boldsymbol{\beta} = 0,\ -\Phi^{\mathrm{T}}\boldsymbol{f} + \Phi^{\mathrm{T}}\Phi\boldsymbol{\beta} = 0, \quad (2)$$

что позволяет представить оценки стандартных коэффициентов в виде

$$\boldsymbol{\beta} = (\Phi^{\mathrm{T}}\Phi)^{-1}\Phi^{\mathrm{T}}\boldsymbol{f}. \quad (3)$$

Вычисление оценок коэффициентов с дополнительным условием

Перепишем соотношение (1) с учетом дополнительного условия согласно [1]:

$$\Delta(\boldsymbol{b}) = (\boldsymbol{f} - \Phi\boldsymbol{b})^{\mathrm{T}}(\boldsymbol{f} - \Phi\boldsymbol{b}) + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{k}_{\phi}\lambda, \quad (4)$$

где $\boldsymbol{k}_{\phi} \in \mathbf{R}^m$ задает вектор значений, определяющий основные свойства системы ортогональных функций; $\lambda \in \mathbf{R}$ – нормирующий коэффициент, определяющий поправочные коэффициенты.

По аналогии можно записать:

$$\partial\Delta(\boldsymbol{b})/\partial\boldsymbol{b} = -2\Phi^{\mathrm{T}}(\boldsymbol{f} - \Phi\boldsymbol{b}) + \boldsymbol{k}_{\phi}\lambda,$$

$$\partial\Delta(\boldsymbol{b})/\partial\boldsymbol{b} = 0,$$

$$-2\Phi^{\mathrm{T}}\boldsymbol{f} + 2\Phi^{\mathrm{T}}\Phi\boldsymbol{b} + \boldsymbol{k}_{\phi}\lambda = 0.$$

С учетом нормального уравнения $\Phi^{\mathrm{T}}\Phi\boldsymbol{\beta} = \Phi^{\mathrm{T}}\boldsymbol{f}$, следующего из (2), получим $-2\Phi^{\mathrm{T}}\Phi\boldsymbol{\beta} + 2\Phi^{\mathrm{T}}\Phi\boldsymbol{b} + \boldsymbol{k}_{\phi}\lambda = 0$, откуда оценки коэффициентов, дополненных к (3), примут вид

$$\boldsymbol{b} = \boldsymbol{\beta} - (\Phi^{\mathrm{T}}\Phi)^{-1}\boldsymbol{k}_{\phi}\lambda/2. \quad (5)$$

Запишем основное свойство вероятностной характеристики в виде

$$\boldsymbol{k}_{\phi}^{\mathrm{T}}\boldsymbol{b} = \boldsymbol{k}_f, \quad (6)$$

где $\boldsymbol{k}_f \in \mathbf{R}$ задает величину, характеризующую основное свойство вероятностной характеристики.

Принимая во внимание (5), представим (6) в виде $\boldsymbol{k}_{\phi}^{\mathrm{T}}(\boldsymbol{\beta} - \lambda/2(\Phi^{\mathrm{T}}\Phi)^{-1}\boldsymbol{k}_{\phi}) = \boldsymbol{k}_f$, откуда

$$\lambda/2 = -(\boldsymbol{k}_{\phi}^{\mathrm{T}}(\Phi^{\mathrm{T}}\Phi)^{-1}\boldsymbol{k}_{\phi})^{-1}(\boldsymbol{k}_f - \boldsymbol{k}_{\phi}^{\mathrm{T}}\boldsymbol{\beta}). \quad (7)$$

С учетом (5) и (7) введем обозначение для дополненных оценок $\boldsymbol{b} = \boldsymbol{\beta} + \boldsymbol{c}$, откуда вектор оценок поправочных коэффициентов примет вид

$$\boldsymbol{c} = -(\Phi^{\mathrm{T}}\Phi)^{-1}\boldsymbol{k}_{\phi}\lambda/2 =$$

$$= (\Phi^{\mathrm{T}}\Phi)^{-1}\boldsymbol{k}_{\phi}(\boldsymbol{k}_{\phi}^{\mathrm{T}}(\Phi^{\mathrm{T}}\Phi)^{-1}\boldsymbol{k}_{\phi})^{-1}(\boldsymbol{k}_f - \boldsymbol{k}_{\phi}^{\mathrm{T}}\boldsymbol{\beta}).$$

Перепишем (1) в дополненных коэффициентах:

$$\Delta(\boldsymbol{b}) = (\boldsymbol{f} - \Phi\boldsymbol{b})^{\mathrm{T}}(\boldsymbol{f} - \Phi\boldsymbol{b}),$$

$$\Delta(\boldsymbol{b}) = \boldsymbol{f}^{\mathrm{T}}\boldsymbol{f} - 2\boldsymbol{f}^{\mathrm{T}}\Phi\boldsymbol{b} + \boldsymbol{b}^{\mathrm{T}}\Phi\Phi^{\mathrm{T}}\boldsymbol{b}.$$

С учетом того, что $\boldsymbol{b} = \boldsymbol{\beta} + \boldsymbol{c}$,

$$\Delta(\boldsymbol{b}) = \boldsymbol{f}^{\mathrm{T}}\boldsymbol{f} - 2\boldsymbol{f}^{\mathrm{T}}\Phi(\boldsymbol{\beta} + \boldsymbol{c}) + (\boldsymbol{\beta} + \boldsymbol{c})^{\mathrm{T}}\Phi\Phi^{\mathrm{T}}(\boldsymbol{\beta} + \boldsymbol{c}),$$

$$\Delta(\boldsymbol{b}) = \boldsymbol{f}^{\mathrm{T}}\boldsymbol{f} - 2\boldsymbol{\beta}^{\mathrm{T}}\Phi\Phi^{\mathrm{T}}\boldsymbol{\beta} - 2\boldsymbol{\beta}^{\mathrm{T}}\Phi\Phi^{\mathrm{T}}\boldsymbol{c} + \boldsymbol{\beta}^{\mathrm{T}}\Phi\Phi^{\mathrm{T}}\boldsymbol{\beta} + 2\boldsymbol{\beta}^{\mathrm{T}}\Phi\Phi^{\mathrm{T}}\boldsymbol{c} + \boldsymbol{c}^{\mathrm{T}}\Phi\Phi^{\mathrm{T}}\boldsymbol{c},$$

$$\Delta(\boldsymbol{b}) = \boldsymbol{f}^{\mathrm{T}}\boldsymbol{f} - \boldsymbol{\beta}^{\mathrm{T}}\Phi\Phi^{\mathrm{T}}\boldsymbol{\beta} + \boldsymbol{c}^{\mathrm{T}}\Phi\Phi^{\mathrm{T}}\boldsymbol{c}.$$

Покажем, что при условии $\boldsymbol{\beta}^{\mathrm{T}}\Phi\Phi^{\mathrm{T}}\boldsymbol{c} = 0$ полученная модель идентична регуляризованной с некоторой ядровой функцией, заданной в матричной форме как $K \in \mathbf{R}^{m\times m}$. При этом введенное ядро задает основные свойства вероятностных характеристик и системы базисных функций.

Вычисление регуляризованных оценок коэффициентов

Представим соотношение (4) как (1) с регуляризацией по норме L_2 [7–10]:

$$\Delta(\boldsymbol{\beta}) = (\boldsymbol{f} - \Phi\boldsymbol{\beta})^{\mathrm{T}}(\boldsymbol{f} - \Phi\boldsymbol{\beta}) + \boldsymbol{\beta}^{\mathrm{T}}K^{-1}\boldsymbol{\beta},$$

откуда

$$\partial\Delta(\boldsymbol{\beta})/\partial\boldsymbol{\beta} = -2\Phi^{\mathrm{T}}(\boldsymbol{f} - \Phi\boldsymbol{\beta}) + 2K^{-1}\boldsymbol{\beta},$$

$$\partial\Delta(\boldsymbol{\beta})/\partial\boldsymbol{\beta} = 0,$$

$$-2\Phi^{\mathrm{T}}(\boldsymbol{f} - \Phi\boldsymbol{\beta}) + 2K^{-1}\boldsymbol{\beta} = 0,$$

$$-\Phi^{\mathrm{T}}\boldsymbol{f} + \Phi^{\mathrm{T}}\Phi\boldsymbol{\beta} + K^{-1}\boldsymbol{\beta} = 0,$$

$$\boldsymbol{\beta}(\Phi^{\mathrm{T}}\Phi + K^{-1}) = \Phi^{\mathrm{T}}\boldsymbol{f},$$

$$\boldsymbol{\beta} = (\Phi^{\mathrm{T}}\Phi + K^{-1})^{-1}\Phi^{\mathrm{T}}\boldsymbol{f}. \quad (8)$$

Так как двойная операция обращения матриц может значительно повысить вычислительные затраты, преобразуем (8) к виду $\Phi\boldsymbol{\beta} = \Phi(\Phi^{\mathrm{T}}\Phi + K^{-1})^{-1}\Phi^{\mathrm{T}}\boldsymbol{f}$.

Введя $P = \Phi K\Phi^{\mathrm{T}}$, получим $\Phi\boldsymbol{\beta} = P(P + I_n)^{-1}\boldsymbol{f}$, где $I_n \in \mathbf{R}^{n\times n}$.

Тогда оценки коэффициентов примут вид $\boldsymbol{\beta} = K\Phi^{\mathrm{T}}(\Phi K\Phi^{\mathrm{T}} + I_n)^{-1}\boldsymbol{f}$, что идентично регуляризованным оценкам для взвешенной системы базисных функций.

Выводы

В данной работе рассмотрен предложенный ранее метод повышения достоверности и устойчивости ортогональных моделей вероятностных характеристик в терминах L_2-регуляризации. Такая трактовка позволила решить следующие задачи: предложить обобщенное описание поправочных коэффициентов для учета основных свойств функциональных характеристик в матричной форме; показать, что предложенные оценки аналогичны оценкам коэффициентов, регуляризованным по норме L_2; ввести ядровую функцию для описания основных свойств функциональных характеристик.

Результатом данной работы является теоретическое обоснование устойчивости предложенных регуляризованных оценок. Практическая значимость полученного результата заключается в возможности оценивания поправочных коэффициентов для различных вероятностных характеристик и базисных функций с помощью одного алгоритма, что позволит предложить более эффективную программную реализацию.

Работа выполнена при поддержке грантов Президента РФ № МК-6218.2018.9 и Минобрнауки РФ № 074-U01.

Литература

1. Прохоров С.А. Аппроксимативный анализ случайных процессов. Самара: Изд-во СамНЦ РАН, 2001. 380 с.

2. Прохоров С.А., Куликовских И.М. Ортогональные модели корреляционно-спектральных характеристик случайных процессов: Лабораторный практикум. Самара: Изд-во СамНЦ РАН, 2008. 301 с.
3. Wahlberg B. System identification using Laguerre models. IEEE Trans. Automatic Control, 1991, vol. AC-36, pp. 551–562.
4. Van den Hof P.M., Heuberger P.S., Bokor J. System identification with generalized orthonormal basis functions. Automatica, 1995, vol. 31, no. 12, pp. 1821–1834.
5. Agresti A. Foundations of linear and generalized linear models. Wiley series in Probability and Statistics, 2015, 472 p.
6. Hastie T., Tibshirani R., Friedman J. The elements of statistical learning: Data mining, inference, and prediction. Springer series in Statistics, 2013, 745 p.
7. Pilonetto G., Nicolao G.D. A new kernel-based approach for linear system identification. Automatica, 2010, vol. 46, no. 1, pp. 81–93.
8. Chen T., Ljung L. Implementation of algorithms for tuning parameters in regularized least squares problems in system identification. Automatica, 2013, vol. 49, no. 7, pp. 2213–2220.
9. Pillonetto G., Dinuzzo F., Chen T., De Nicolao G., Ljung L. Kernel methods in system identification, machine learning and function estimation: A servey. Automatica, 2014, vol. 50, no. 3, pp. 657–682.
10. Aronszajn N. Theory of reproducing kernels. Transactions of American Mathematical Society, 1950, pp. 337–404.

Software & Systems Received 29.12.17
DOI: 10.15827/0236-235X.121.099-101 2018, vol. 31, no. 1, pp. 099–101

REGULARIZED ORTHOGONAL MODELS OF PROBABILISTIC CHARACTERISTICS WITH THE CONDITION OF ITS BASIC PROPERTIES EXECUTION

S.A. Prokhorov [1]*, Dr.Sc. (Engineering), Professor, Head of Chair, sp.prokhorov@gmail.com*
I.M. Kulikovskikh [1]*, Ph.D. (Engineering), Associate Professor, kulikovskikh.i@gmail.com*

[1] *Samara National Research University, Moskovskoe Highway 34, Samara, 443086, Russian Federation*

Abstract. The representation of a probabilistic characteristic model as Fourier series is one of the most effective methods for model complexity reduction. The orthogonal function systems are widely used to solve the problems of identification, filtering, and analysis of dynamic systems. This success is often connected to the ability of presenting long time series as a more compact model.

There are different methods to guarantee stability and validity of such probabilistic characteristic orthogonal model. They include: optimizing orthogonal model parameters, using numerically stable schemes for computing Fourier coefficients, designing weighted models that allow taking into consideration the aspects of dynamic system behavior, in particular, non-stationary systems, delayed system and so on.

Nevertheless, the primary focus of the paper is on the method that takes into account the requirement of executing basic probabilistic characteristics property by an orthogonal model. This method is based on the need of computing the corrections for Fourier coefficients, which are determined with regard to the basic properties of both probabilistic characteristics and basic functions. As a result, a set of orthogonal models requires different implementations depending on a problem statement and a dynamic system.

This paper generalizes the previously proposed mathematical models to implement the method that allows the following: to propose a general description of modifying factors in order to take into account basic properties of functional characteristics in a matrix form; to show that the proposed estimates are identical to the estimates regularized with L_2-norm; to introduce a kernel function for describing basic properties of functional characteristics. The findings of the present study prove the stability of the proposed regularized estimates.

The generalized mathematical models reduce the method of computing the corrected estimates for various functional characteristics to a unique algorithm. Therefore, this significantly improves the computational efficiency of further software implementation.

Keywords: regularization, kernel functions, orthogonal models, probabilistic characteristics.

Acknowledgements. *This work has been supported by grants of the Prezident of the Russian Federation no. МК-6218.2018.9 and the Ministry of Education and Science of the Russian Federation no. 074-U01.*

References

1. Prokhorov S.A. *Approksimativny analiz sluchaynykh protsessov* [Approximative Analysis of Stochastic Processes]. 2nd ed., Samara, SSC RAS Publ., 2001, 380 p.
2. Prokhorov S.A., Kulikovskikh I.M. *Ortogonalnye modeli korrelyatsionno-spektralnykh kharakteristik sluchaynykh protsessov: laboratorny praktikum* [Orthogonal models of correlation-spectral characterisitcs. assignments book]. Samara, SSC RAS Publ., 2008, 301 p.
3. Wahlberg B. System identification using Laguerre models. *IEEE Trans. Automatic Control*. 1991, vol. AC-36, pp. 551–562.
4. Van den Hof P.M., Heuberger P.S., Bokor J. System identification with generalized orthonormal basis functions. *Automatica*. 1995, vol. 31, no. 12, pp. 1821–1834.
5. Agresti A. *Foundations of linear and generalized linear models*. Wiley Series in Probability and Statistics, 2015, 472 p.
6. Hastie T., Tibshirani R., Friedman J. *The elements of statistical learning: data mining, inference, and prediction*. 2nd ed., Springer Series in Statistics, 2013, 745 p.
7. Pilonetto G., Nicolao G.D. A new kernel-based approach for linear system identification. *Automatica*. 2010, vol. 46, no. 1, pp. 81–93.
8. Chen T., Ljung L. Implementation of algorithms for tuning parameters in regularized least squares problems in system identification. *Automatica*. 2013, vol. 49, no. 7, pp. 2213–2220.
9. Pillonetto G., Dinuzzo F., Chen T., De Nicolao G., Ljung L. Kernel methods in system identification, machine learning and function estimation: A servey. *Automatica*. 2014, vol. 50, no. 3, pp. 657–682.
10. Aronszajn N. Theory of reproducing kernels. *Trans. of American Mathematical Society*. 1950, pp. 337–404.