УДК 519.2 DOI: 10.18698/2541-8009-2023-8-926

СРАВНЕНИЕ МОДЕЛЕЙ ПОЛИНОМИАЛЬНОГО ХАОСА И РЕГРЕССИОННОЙ НА ОСНОВЕ ПОЛИНОМОВ КОЛМОГОРОВА — ГАБОРА

Фам Куок Вьет fv20if04@student.bmstu.ru

МГТУ им. Н.Э. Баумана, Москва, Российская Федерация

Аннотация

Рассмотрено обобщенное разложение полиномиального хаоса применительно к задаче анализа регрессионных моделей. Коэффициенты полиномиального хаоса вычислены неинтрузивными методами: методом наименьших квадратов и методом эластичной сети. В качестве эталонной функции метода группового учета аргументов использованы полиномы Колмогорова — Габора. Сравнение методов проведено на функции Ишигами. Показано, что при большом интервале изменения значений случайных величин модели полиномиального хаоса обеспечивают лучший результат и нечувствительны к мультиколлинеарности. Продемонстрировано, что модели на основе полиномов Колмогорова — Габора дают нестабильной диапазон ошибок при более медленной скорости выполнения, но предпочтительны при большой размерности входных данных.

Ключевые слова

Полиномиальный хаос, схема Аски — Винера, эластичная сеть, неинтрузивная спектральная проекция, полиномиальная нейронная сеть, метод группового учета аргументов, полиномы Колмогорова — Габора, функция Ишигами

Поступила в редакцию 21.05.2023

Введение. Повседневные сценарии связаны с неопределенностями в различных областях: инвестиционные возможности, медицинские диагнозы, результаты спортивных игр, прогнозы погоды и т. д. Модели, применяемые для количественной оценки неопределенности, призваны оценить, насколько вероятны полученные результаты, если некоторые аспекты системы точно не известны. В работе изучаются две модели, которые относятся к классу полиномиальных регрессионных моделей: разложение полиномиального хаоса (ПХ) и регрессионная модель на основе полиномов Колмогорова — Габора.

Полиномиальный хаос — метод представления случайной величины в виде разложения по ортогональному базису. Преимущества модели обусловлены предварительной оценкой закона распределения входных данных, на основе которого можно построить ортогональные базисы пространства признаков.

Метод группового учета аргументов, или полиномиальная нейронная сеть — семейство индуктивных алгоритмов компьютерного математического моделирования многопараметрических наборов данных. Он отличается полностью автоматической структурой и параметрической оптимизацией моделей, которые минимизируют необходимость в предварительных знаниях о входных данных. В качестве базовых функций используется последовательность полиномов Колмогорова — Габора.

Разложение полиномиального хаоса. Пусть Ω — пространство элементарных событий и Θ — пространство функций, сопоставляющих элементам $\omega \in \Omega$. Функция $\theta : \omega \in \Omega \to \mathbb{R}$ — случайная величина. Пусть $\{\xi_i\}_{i=1}^{\infty}$ — бесконечное, но исчисляемое множество независимых нормированных гауссовских случайных величин. Определим:

$\widehat{\Gamma}_p$ — пространство всех многочленов степени меньше или равной p в $\{\xi_i\}_{i=1}^{\infty}$;

Γ_p — множество многочленов пространства $\widehat{\Gamma}_p$, ортогональных пространству $\widehat{\Gamma}_{p-1}$;

$\tilde{\Gamma}_p$ — пространство, порожденное Γ_p.

Тогда говорят, что $\tilde{\Gamma}_p$ — p-й гомогенный хаос, а Γ_p — полиномиальные хаосы порядка p.

Разложение ПХ случайной величины $\theta(\omega)$ принимает вид

$$\theta(\omega) = a_0\Gamma_0 + \sum_{i_1=1}^{\infty} a_{i_1}\Gamma_1\left(\xi_{i_1}(\omega)\right) + \sum_{i_1=1}^{\infty}\sum_{i_2=1}^{\infty} a_{i_1 i_2}\Gamma_2\left(\xi_{i_1}(\omega),\ \xi_{i_2}(\omega)\right) + \ldots = \sum_{k=0}^{\infty}\theta_k\Psi_k\left(\xi(\omega)\right),$$

где θ_k — детерминированные коэффициенты; Ψ_k — случайные полиномы, ортогональные в L_2 пространстве [1]:

$$\Psi_i, \Psi_j = \int \Psi_i(\xi)\Psi_j(\xi)\, p(\xi)\, d\xi = \delta_{ij}\Psi_i, \Psi_j.$$

Пусть p обозначает порядок разложения ПХ. Конечномерное разложения ПХ случайной величины конечного порядка определяется формулой

$$\theta(\xi) \approx \sum_{k=0}^{P}\theta_k\Psi_k(\xi), \xi = \{\xi_1, \xi_2, \ldots, \xi_d\},$$

где базисный размер связан с d и p соотношением

$$P + 1 = \frac{(p+d)!}{p!d!}.$$

Сю и Карниадакис использовали схему Аски, чтобы обобщить разложение полиномиального хаоса Винера на обычные негауссовы меры. Это обобщение может быть полезно для улучшения сходимости разложения для негауссовских случайных величин [2].

Если данные факторы распределены по другому закону, то можно использовать нелинейное отображение, чтобы свести соответствующие многочлены к табличным (см. таблицу).

Соответствие типов полиномиального хаоса Винера — Аски и лежащих в их основе случайных величин

Вид распределения случайной величины	Плотность	Ортогональные полиномы	Ортогональный базис $\{\Psi_j\}$
Равномерное	$\dfrac{I_{[-1,1]}(x)}{2}$	Лежандра $P_k(x)$	$\dfrac{P_k(x)}{\sqrt{\dfrac{1}{2k+1}}}$
Нормальное	$\dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$	Эрмита $He_k(x)$	$\dfrac{He_k(x)}{\sqrt{k!}}$
Гамма	$x^\alpha e^{-x} I_{\mathbb{R}^+}(x)$	Лагерра $L_k^\alpha(x)$	$\dfrac{L_k^\alpha(x)}{\sqrt{\dfrac{\Gamma(k+\alpha+1)}{k!}}}$
Бета	$\dfrac{(1-x)^\alpha(1+x)^\beta}{B(a)B(b)} I_{[-1;1]}(x)$	Якоби $J_k^{\alpha,\beta}$	$\dfrac{J_k^{\alpha,\beta}(x)}{J_{\alpha,\beta,k}}$
Биномиальное	$\begin{pmatrix} n \\ k \end{pmatrix} p^k(1-p)^{n-k}$	Кравчука $K_k(x,p,n)$	$(-p)^k \begin{pmatrix} n \\ k \end{pmatrix} K_k\left(x,p,n\right)$
Пуассона	$\dfrac{\lambda^x}{x!}e^{-\lambda}$	Шарлье C_n^λ	$\sqrt{\dfrac{\omega_{C(x,\lambda)}}{\rho_{C(n,\lambda)}}} C_n^\lambda(x)$

Обозначая одномерные ортогональные полиномы из схемы Аски $\dfrac{(1-x)^\alpha(1+x)^\beta}{B(a)B(b)} I_{[-1;1]}(x)$ и предполагая, что случайные величины независимы, многомерный базис обобщенного полиномиального хаоса $\{\Psi_i\}$ можно построить на основе тензорных произведений соответствующих одномерных полиномов:

$$\Psi_i\left(\xi_1,\xi_2,\ldots,\xi_d\right) = \prod_{k=1}^{d} \varphi_{\alpha_k^i}(\xi_k), \quad \left|\alpha^i\right| = \sum_{k=1}^{d} \alpha_k^i \le p.$$

Определение коэффициентов разложения полиномиального хаоса. Для определения коэффициентов ПХ выделяют два класса методов.

1. Интрузивный — слабое решение математической модели со случайным вводом ищется с помощью галеркинской проекции уравнений модели на базе ПХ; требует модификации предварительно определенной модели. Метод проек-

ций Галеркина является интрузивным подходом для оценки желаемых коэффициентов ПХ. При задействовании данного метода генерируют систему детерминированных уравнений, в которой коэффициенты разложения полиномиального хаоса являются неизвестными. Затем полученную систему решают с помощью подходящих численных методов [3].

2. Неинтрузивный — требует реализации только детерминированного кода для разных значений ввода. К неинтрузивным методам относятся: метод наименьших квадратов, метод эластичной сети, метод неинтрузивной спектральной проекции (NISP).

Метод наименьших квадратов (МНК). Пусть $\beta = \left(\beta_0, \ldots, \beta_P\right)^{\mathrm{T}}$ — вектор искомых коэффициентов ПХ в усеченном разложении выходных данных *y*; $\hat{\beta}$ — приближение коэффициентов β:

$$\hat{\beta} = \arg\min_{\beta} \sum_{i=1}^{n} \left(y^{(i)} - \sum_{k=0}^{P} \beta_k \Psi_k \left(\xi^{(i)} \right) \right)^2, \quad n > P + 1.$$

Эластичная сеть. Эластичная сеть (Elastic net) — пересмотренный алгоритм выбора переменной, основанный на модели Лассо (Least absolute shrinkage and selection operator) [4–6]. Регрессионная модель Лассо задается системой

$$\begin{cases} \hat{\beta} = \arg\min_{\beta} \left[\dfrac{1}{2N} \sum_{i=1}^{n} \left(y^{(i)} - \sum_{k=0}^{P} \beta_k \Psi_k \left(\xi^{(i)} \right) \right)^2 + \lambda \beta_1 \right]; \\ \beta_1 = \sum_{j=1}^{m} \left| a_j \right|, \end{cases}$$

где λ — штрафной параметр, имеющий смысл переменной сжатия.

Недостатком модели Лассо является невозможность обработки мультиколлинеарных данных, поэтому Zou et al. и Hanstie et al. предложили эластичную сеть для устранения этих недостатков:

$$\begin{cases} \hat{\beta} = \arg\min_{\beta} \left[\dfrac{1}{2N} \sum_{i=1}^{n} \left(y^{(i)} - \sum_{k=0}^{P} \beta_k \Psi_k \left(\xi^{(i)} \right) \right)^2 + \lambda_1 \beta_1 + \lambda_2 \beta_2 \right]; \\ \beta_1 = \sum_{j=1}^{m} \left| a_j \right|; \\ \beta_2 = \sum_{j=1}^{m} \left| a_j \right|^2 . \end{cases}$$

Эластичная сеть сжимает коэффициент регрессии дважды и приводит к увеличению смещения оценки. Видно, что в случае $\lambda_2 = 0$ модель становится регрессионной моделью Лассо, а в случае $\lambda_1 = 0$ — гребневой.

Неинтрузивная спектральная проекция. В методе неинтрузивной спектральной проекции используется ортогональность базы ПХ:

$$\beta_k = \frac{y(\xi), \Psi_k(\xi)}{\Psi_k(\xi), \Psi_k(\xi)},$$

где

$$f(\xi), g(\xi) = \int_{\Omega_d} f(\xi) g(\xi) p(\xi) d\xi.$$

Метод группового учета аргументов. Полиномиальная сеть высокого порядка, генерируемая алгоритмом метода группового учета аргументов (МГУА), по сути является многослойной нейронной сетью, которая была предложена советским ученым Ивахненко. Это эвристический и самоорганизованный метод моделирования. Самоорганизованная форма минимизирует необходимость в предыдущих знаниях. Модели не нужно указывать какие-либо начальные предположения, такие как количество нейронов и скрытых слоев, которые снижают субъективность и сложность моделирования.

Входные данные сначала делят на обучающий набор и набор тестирования. Данные образца обучения используются для оценки коэффициентов полинома Колмогорова — Габора (К-Г). Набор тестирования поставляется в сеть МГУА для проверки ошибок. Затем характеристические переменные обучающего набора попарно рекомбинируют. После рекомбинации каждая пара характеристических переменных является группой и обучается как нейрон сети. Затем выход каждого нейрона оценивают и проверяют по внешнему критерию. Устраняют нейроны, которые показали себя худшими, и сохраняют нейроны с хорошей производительностью в качестве следующего слоя. Весь процесс реструктуризации, обучения, тестирования и отбора снова повторяют на уровне нового слоя (рис. 1). Сеть МГУА работает до тех пор, пока ошибка прогноза нейронов больше перестанет уменьшаться [4].

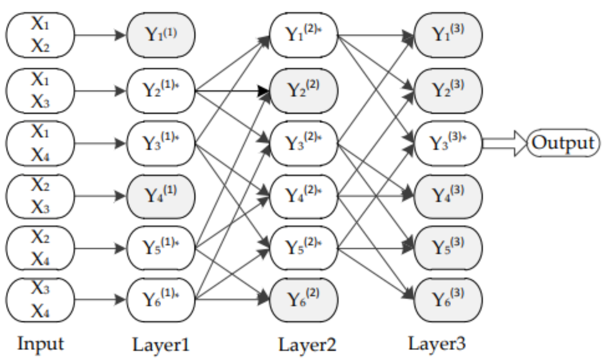

Рис. 1. Сетевая архитектура метода группового учета аргументов

Связь между входом и выходом каждого нейрона может быть представлена полиномом К-Г:

$$\hat{y}_n = f(x_1, x_2, \ldots, x_m) = \sum_{i=1}^{m} a_i x_{in} + \sum_{i=1}^{m}\sum_{j=1}^{m} a_{ij} x_{in} x_{jn} + \sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k=1}^{m} a_{ijk} x_{in} x_{jn} x_{kn} + \ldots$$

Обозначим ω — значение внешнего критерия. Чем меньше ω, тем лучше нейрон и лучше эффект подгонки полинома К-Г. Более бедные нейроны отвергаются, а превосходные нейроны сохраняются и становятся входом следующего слоя. Между тем минимальное значение $\omega_{\min}$ этого слоя записано. Когда $\omega_{\min}$ этого уровня больше не уменьшается по сравнению с предыдущим уровнем, это указывает на то, что ошибка прогнозирования сети больше не уменьшается; затем сеть перестает расширяться и выводит результат предыдущего уровня:

$$\omega = \frac{\sum_{n=1}^{P} (\hat{y}_n - y_n)^2}{\sum_{n=1}^{P} (y_n)^2},$$

где P — размер набора тестирования.

Функция Ишигами. Проиллюстрируем использование разложения ПХ и МГУА в моделировании выходных данных функции Ишигами. Благодаря высокой нелинейности и немонотонности функция Ишигами широко используется для оценки методологий анализа неопределенности и чувствительности. Рассмотрим сначала случай, когда входные данные $\boldsymbol{x}$ находятся в двумерном пространстве:

$$\begin{cases} f(\boldsymbol{x}) = \sin x_1 + a \sin^2 x_2 + b x_3^4 \sin x_1; \\ x_1, x_2 \sim R[-\pi, \pi]; \\ x_3 = 1;\ a = 7;\ b = 0{,}1. \end{cases}$$

В качестве примера рассмотрим модели: МГУА на основе эластичной сети с квадратичной и кубической эталонными функциями; модель полиномиального хаоса степеней 2, 3 и 6. Будем сравнивать эффективность моделей на основе ошибок, возвращаемых в тестовых наборах, а также по времени работы при однопоточной программе. На рис. 2–5 приведены результаты работы программы на 100 различных случайных наборах, каждый из которых состоит из 1000 входных данных, разделенных на обучающий и тестовый в соотношении 8 : 2.

Модели реализованы в программе, написанной на языке Python, которая запускается на компьютере со следующей конфигурацией: процессор Intel(R) Core i5-9600KF; оперативная память RAM 16Gb; видеокарта NVIDIA GeForce RTX 2060.

Рис. 2. Ошибка при моделировании функции Ишигами в двумерном пространстве

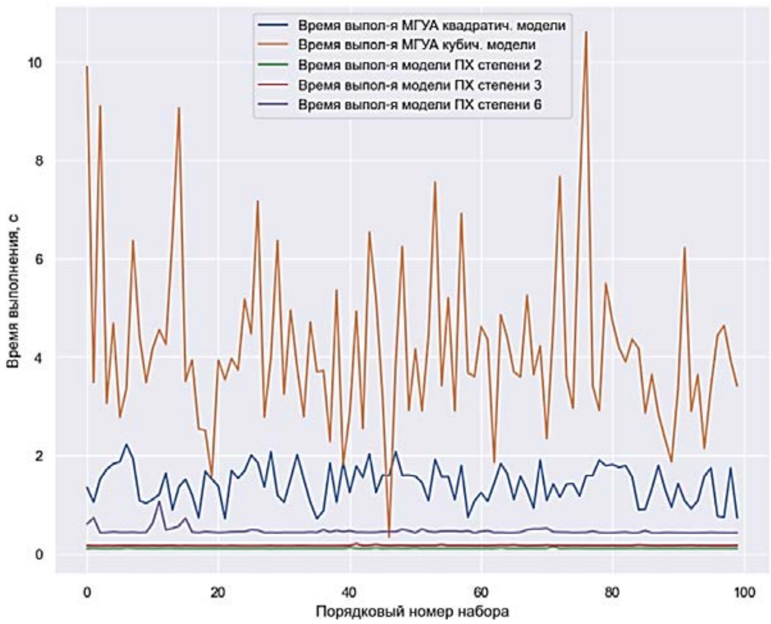

Рис. 3. Время выполнения моделей при моделировании функции Ишигами в двумерном пространстве

Видно, что в двумерном пространстве ошибки двух моделей МГУА и ПХ одного порядка имеют явные различия. С одной стороны, модель МГУА, в которой используется квадратичную эталонную функцию, самая устойчива и имеет наименьшую погрешность среди моделей второго и третьего порядка. Впрочем, эта ошибка достаточно велика в сравнении с моделью ПХ степени 6. С другой стороны, хотя время работы модели ПХ степени 6 значительно больше, чем у моделей первого и второго порядка, оно все же намного меньше по сравнению с большими и нестабильными результатами, полученными на основе двух моделей МГУА.

Рассмотрим теперь полный случай функции Ишигами в трехмерном пространстве:

$$\begin{cases} f(\boldsymbol{x}) = \sin x_1 + a\sin^2 x_2 + bx_3^4 \sin x_1; \\ x_1, x_2, x_3 \sim R[-\pi, \pi]; \\ a = 7;\ b = 0{,}1. \end{cases}$$

Видно, что в этом случае погрешности обеих моделей одного порядка не сильно различаются. Амплитуда ошибки велика, а ее среднее значение для каждой модели находится в интервале [0,2; 0,3]. Однако это не относится к модели ПХ степени 6, ошибка которой остается стабильно низкой. Время выполнения значительно увеличилось, но при этом скорость выполнения моделей ПХ степени 2 и 3 все еще по-прежнему почти мгновенна, скорость модели ПХ степени 6 увеличилась до среднего времени выполнения квадратичной модели МГУА. Это происходит потому, что, когда переменных становится больше, число базовых полиномов также увеличивается вместе с числом коэффициентов, которые необходимо оценить.

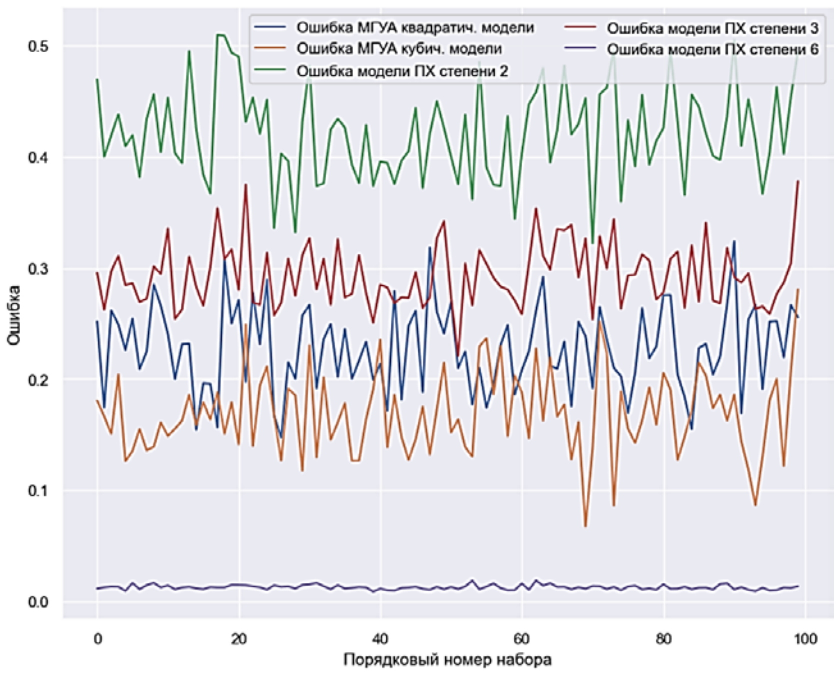

Рис. 4. Ошибка моделей при моделировании функции Ишигами в трехмерном пространстве

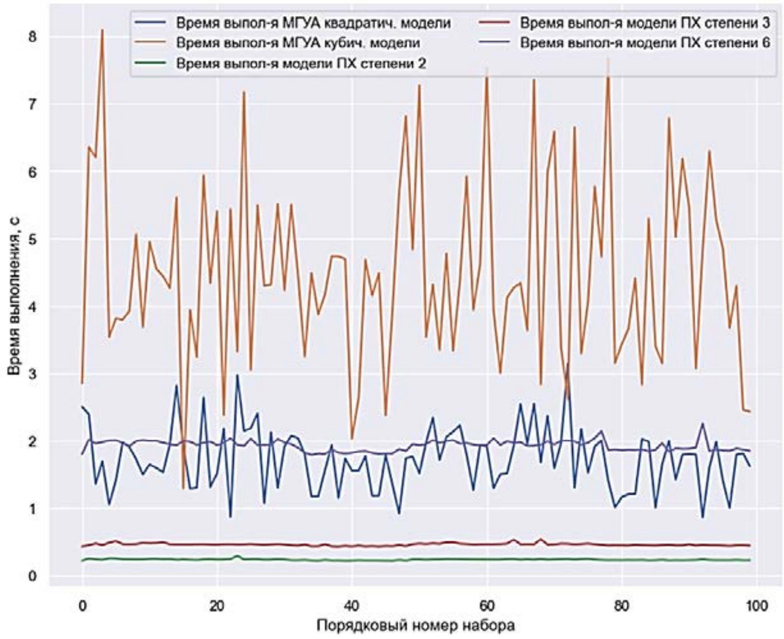

Рис. 5. Время выполнения моделей при моделировании функции Ишигами в трехмерном пространстве

Вывод. Проведенная работа по сравнению эффективности моделей в регрессионной задаче показывает, что при большом интервале изменения значения случайных величин модели полиномиального хаоса обеспечивают лучший результат. Модели на основе полинома Комогорова — Габора дают нестабильной диапазон ошибки при большем времени работы, но предпочтительны при большой размерности входных данных.

Литература

[1] Crestaux T., Le Maitre O., Martinez J.-M. Polynomial chaos expansion for sensitivity analysis. *Reliability Engineering System Safety*, 2009, vol. 94 (7), pp. 1161–1172. http://doi.org/10.1016/j.ress.2008.10.008

[2] Xiu D., Em Karniadakis G. The Wiener — Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput*., 2002, vol. 24, no. 2, pp. 619–644. http://doi.org/10.1137/S1064827501387826

[3] Pettersson M.P., Iaccarino G., Nordstrom J. *Polynomial chaos methods for hyperbolic partial differential equations. Numerical Techniques for Fluid Dynamics Problems in the Presence of Uncertainties*. Switzerland, Springer International Publishing, 2015.

[4] Liu W., Dou Z., Wang W. et. al. Shortterm load forecasting based on elastic net improved GMDH and difference degree weighting optimization. *Applied Sciences*, 2018, vol. 8 (9). http://doi.org/10.3390/app8091603

[5] Стрижов В.В., Крымова К.А. *Методы выбора регрессионных моделей*. Москва, Вычислительный центр им. А.А. Дородницына РАН, 2010, 60 с.

[6] Стрижов В.В. *Порождение и выбор моделей в задачах регрессии и классификации*. Дис. … канд. техн. наук. Москва, 2014, 299 с.

Фам Куок Вьет — студент кафедры «Вычислительная математика и математическая физика», МГТУ им. Н.Э. Баумана, Москва, Российская Федерация.

Научный руководитель — Облакова Татьяна Васильевна, кандидат физико-математических наук, доцент кафедры «Вычислительная математика и математическая физика», МГТУ им. Н.Э. Баумана, Москва, Российская Федерация.

Ссылку на эту статью просим оформлять следующим образом:

Фам Куок Вьет. Сравнение моделей полиномиального хаоса и регрессионной на основе полиномов Колмогорова — Габора. *Политехнический молодежный журнал*, 2023, № 08 (85). http://dx.doi.org/10.18698/2541-8009-2023-8-926

POLYNOMIAL CHAOS AND REGRESSION MODELS COMPARISON BASED ON THE KOLMOGOROV — GABOR POLYNOMIALS

Pham Quoc Viet fv20if04@student.bmstu.ru

Bauman Moscow State Technical University, Moscow, Russian Federation

Abstract

The paper considers the polynomial chaos generalized expansion applied to the regression model analysis problem. The polynomial chaos coefficients were calculated by the non-intrusive methods, including the least squares and the elastic network methods. Kolmogorov — Gabor polynomials were used as a reference function in the method of the arguments group accounting. Methods were compared with the Ishigami function. It is shown that at the wide range of variation in the random variables values, the polynomial chaos models are providing the best result and stay insensitive to the multicollinearity. The paper demonstrates that models based on the Kolmogorov — Gabor polynomials are providing unstable error range at the slower execution speed, but are preferable at the large input data dimensions.

Keywords

Polynomial chaos, Askey — Wiener scheme, elastic network, non-intrusive spectral projection, polynomial neural network, method of the arguments group accounting, Kolmogorov — Gabor polynomials, Ishigami function

Received 21.05.2023

References

[1] Crestaux T., Le Maitre O., Martinez J.-M. Polynomial chaos expansion for sensitivity analysis. *Reliability Engineering System Safety*, 2009, vol. 94 (7), pp. 1161–1172. http://doi.org/10.1016/j.ress.2008.10.008

[2] Xiu D., Em Karniadakis G. The Wiener — Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.*, 2002, vol. 24, no. 2, pp. 619–644. http://doi.org/10.1137/S1064827501387826

[3] Pettersson M.P., Iaccarino G., Nordstrom J. *Polynomial chaos methods for hyperbolic partial differential equations. Numerical Techniques for Fluid Dynamics Problems in the Presence of Uncertainties*. Switzerland, Springer International Publishing, 2015.

[4] Liu W., Dou Z., Wang W. et. al. Shortterm load forecasting based on elastic net improved GMDH and difference degree weighting optimization. *Applied Sciences*, 2018, vol. 8 (9). http://doi.org/10.3390/app8091603

[5] Strizhov V.V., Krymova K.A. *Metody vybora regressionnykh modeley* [Regression model selection methods]. Moscow, Vychislitel'nyy tsentr im. A.A. Dorodnitsyna RAN Publ., 2010, 60 p. (In Russ.).

[6] Strizhov V.V. *Porozhdenie i vybor modeley v zadachakh regressii i klassifikatsii* [Generation and choice of models in problems of regression and classification]. Ph. D. Diss. Moscow, 2014, 299 p. (In Russ.).

Pham Quoc Viet — Student, Department of Computational Mathematics and Mathematical Physics, Bauman Moscow State Technical University, Moscow, Russian Federation.

Scientific advisor — Oblakova T.V., Ph. D. (Eng.), Associate Professor, Department of Computational Mathematics and Mathematical Physics, Bauman Moscow State Technical University, Moscow, Russian Federation.

Please cite this article in English as:

Pham Quoc Viet. Polynomial chaos and regression models comparison based on the Kolmogorov — Gabor polynomials. *Politekhnicheskiy molodezhnyy zhurnal*, 2023, no. 08 (85). (In Russ.). http://dx.doi.org/10.18698/2541-8009-2023-8-926