A. V. Mayer, V. A. Simakhin

NONPARAMETRIC SENSORS FOR STOHASTIC STATIONARY PROCESS

In the article we consider an algorithm of nonparametric generator's building for stohastic stationary process. Dependency interval of stochastic process is determined with the help of nonparametric algorithms of prognosis.

Keywords: generator, nonparametric, prognosis, process, modeling.

УДК 519.224:330.46

А. А. Новоселов

ПОСТРОЕНИЕ МНОГОМЕРНЫХ ДИСКРЕТНЫХ РАСПРЕДЕЛЕНИЙ С ЗАДАННОЙ КОРРЕЛЯЦИОННОЙ СТРУКТУРОЙ

Рассмотрены методы воспроизведения многомерного дискретного распределения с заданной корреляционной структурой и маргинальными распределениями. Для воспроизведения используются смеси базовых распределений и решение некоторых оптимизационных задач.

Ключевые слова: дискретное распределение, корреляция, копула, смесь.

Пусть заданы нормальные распределения со средними значениями $\mu_1, \ldots, \mu_d$ и стандартными отклонениями $\sigma_1, \ldots, \sigma_d$. Для произвольной корреляционной матрицы R существует единственное многомерное нормальное распределение, обладающее такими маргинальными распределениями и корреляционной матрицей. Хорошо известный алгоритм воспроизведения соответствующего случайного вектора $X = (X_1,..., X_d)$ основан на факторизации ковариационной матрицы.

Обозначим

$$\Lambda = \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \sigma_d \end{pmatrix}$$

диагональную матрицу со стандартными отклонениями на диагонали, тогда $C = \Lambda R \Lambda$ является ковариационной матрицей распределения вектора X. Будучи неотрицательно определенной и симметричной, ковариационная матрица C может быть представлена в виде

$$C = A'A \tag{1}$$

с некоторой матрицей A, причем последняя определяется не одним образом. Примерами такого представления являются разложение Холецкого и ортогональное разложение.

При наличии разложения (1) вектор X воспроизводится из стандартного нормального случайного вектора Z по формуле

$$X = A'Z. \tag{2}$$

Действительно, для Z справедливо $\boldsymbol{E}ZZ' = I$, где I – единичная матрица соответствующего размера, поэтому $\boldsymbol{E}XX' = \boldsymbol{E}(A'ZZ'A) = A'(\boldsymbol{E}ZZ')A = A'A = C$, так что X обладает требуемой ковариационной структурой.

В случае, когда компоненты X имеют фиксированные дискретные распределения, аналогичный метод оказывается неприменимым. Во-первых, заданным маргинальным распределениям и ковариационной матрице соответствуют, вообще говоря, многие многомерные дискретные распределения. Может оказаться и так, что подходящее многомерное распределение не существует.

Во-вторых, алгоритм вращения (2) не сохраняет дискретную решетку значений, на которой задано распределение.

В работах [1; 2] анонсированы методы воспроизведения двумерного дискретного распределения с заданными маргинальными распределениями и корреляцией, основанные на смесях некоторых базовых распределений и минимизации уклонения от независимого распределения. В настоящей работе предлагается обоснование этих методов.

Описание двумерного дискретного распределения. Пусть размерность $d = 2$. Обозначим $K = \{1,\ldots,m\} \times \{1,\ldots,n\}$. Дискретное распределение вектора $X = (X_1, X_2)'$ задается на прямоугольной сетке значений $\{x_{11},\ldots, x_{1m}\} \times \{x_{21},\ldots, x_{2n}\}$ в виде $\boldsymbol{P}(X_1 = x_{1i}; X_2 = x_{2j}) = r_{ij}$, $(i, j) \in K$. Обозначим совместное распределение компонент вектора X:

$$r = \{r_{ij}, (i, j) \in K\}. \tag{3}$$

Здесь $P(X_1 = x_{1i}) = p_i$, $i = 1,\ldots,m$ и $P(X_2 = x_{2j}) = q_j$, $j = 1,\ldots,n$, так что векторы

$$p = (p_1,\ldots,p_m),\ q = (q_1,\ldots,q_n) \tag{4}$$

описывают маргинальные распределения компонент.

Средние значения

$$a_1 = EX_1 = \sum_{i=1}^{m} x_{1i} p_i \ , \quad a_2 = EX_2 = \sum_{j=1}^{n} x_{2j} q_j \ ,$$

$$a_{12} = E(X_1 X_2) = \sum_{i=1}^{m} \sum_{j=1}^{n} x_{1i} x_{2j} r_{ij} \ ,$$

стандартные отклонения

$$\sigma_1 = \sqrt{E(X_1 - EX_1)^2} \ , \quad \sigma_2 = \sqrt{E(X_2 - EX_2)^2}$$

и коэффициент корреляции

$$c = \frac{E(X_1 X_2) - EX_1 EX_2}{\sigma_1 \sigma_2} \tag{5}$$

вычисляются как обычно.

Если маргинальные распределения (4) известны, то выполняются следующие соотношения:

$$\sum_{i=1}^{m} r_{ij} = q_j, j = 1, \ldots, n \ ; \tag{6}$$

$$\sum_{j=1}^{n} r_{ij} = p_i, i = 1, \ldots, m \ . \tag{7}$$

Отметим, что среди $m + n$ уравнений (6), (7) имеется лишь $m + n - 1$ независимых, поскольку сумма компонент любого распределения равна 1.

Если же дополнительно известен коэффициент корреляции c компонент X, то справедливо и уравнение

$$\sum_{i=1}^{m} \sum_{j=1}^{n} x_{1i} x_{2j} r_{ij} = c\sigma_1 \sigma_2 + a_1 a_2 \ . \tag{8}$$

Обозначим $F(p, q)$ класс всех двумерных распределений (3) с маргинальными распределениями (4) (его еще называют классом Фреше), а $F_c(p, q)$ его подкласс распределений с корреляцией c. Известно [3], что среди всех распределений в $F(p, q)$ наименьшей корреляцией $c_{\min} = c_{\min}(p, q)$ обладает антикомонотонное распределение $R^{-}(p, q)$, а наибольшей корреляцией $c_{\max} = c_{\max}(p, q)$ обладает комонотонное распределение $R^{+}(p, q)$, причем $c_{\min} > -1$ и $c_{\max} < 1$. Класс Фреше представим в виде

$$F(p,q) = \bigcup_{c \in [c_{\min}, c_{\max}]} F_c(p,q) \ . \tag{9}$$

Обозначим еще $R^0(p, q)$ независимое распределение из класса Фреше $F(p, q)$.

Пример. Пусть на сетке значений

$$\{0, 1, 2\} \times \{0, 1\} \tag{10}$$

заданы маргинальные распределения

$$p = (1/3, 1/2, 1/6), q = (2/5, 3/5). \tag{11}$$

Основные характеристики маргинальных распределений равны:

$$a_1 = \frac{5}{6}, a_2 = \frac{3}{5}, \sigma_1 = \frac{\sqrt{17}}{6}, \sigma_2 = \frac{\sqrt{6}}{5}.$$

Двухпараметрическое представление класса Фреше имеет вид

$$R = \begin{pmatrix} u & v & 2/5 - u - v \\ 1/3 - u & 1/2 - v & u + v - 7/30 \end{pmatrix}. \tag{12}$$

Область допустимых значений параметров задается неравенствами

$$0 \le u \le 1/3, 0 \le v \le 1/2, 7/30 \le u + v \le 2/5 \tag{13}$$

и представлена на рисунке. Окружностями на нем отмечены антикомонотонное $R^{-}(p, q)$

$$R^{-}(p,q) = \begin{pmatrix} 0 & 7/30 & 1/6 \\ 1/3 & 4/15 & 0 \end{pmatrix}; \tag{14}$$

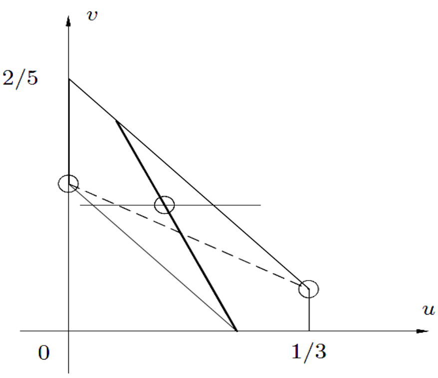
независимое $R^0(p, q)$

$$R^0(p,q) = \begin{pmatrix} 2/15 & 1/5 & 1/15 \\ 1/5 & 3/10 & 1/10 \end{pmatrix}; \tag{15}$$

комонотонное $R^{+}(p, q)$ распределение

$$R^{-}(p,q) = \begin{pmatrix} 1/3 & 1/15 & 0 \\ 0 & 13/30 & 1/6 \end{pmatrix}. \tag{16}$$

.



Допустимая область в плоскости параметров (u, v):
—— – некоррелированные компоненты; O (слева направо) – антикомонотонное, независимое и комонотонное распределения; —— – решения задачи 1

Класс $F_c(p, q)$ описывается в параметрическом представлении следующим уравнением:

$$v = \frac{7}{15} + c\frac{\sqrt{102}}{30} - 2u \ . \tag{17}$$

Некоррелированные распределения лежат на отрезке прямой, описываемом выражением

$$u + v = \frac{7}{15}, \frac{1}{15} \le u \le \frac{7}{30} \ .$$

Коэффициент корреляции в данном примере заключен в интервале

$$c \in \left[-\frac{7}{\sqrt{102}}; \frac{8}{\sqrt{102}} \right]. \tag{18}$$

Критерий близости к независимому распределению. Теперь сформулируем две задачи выбора из всего множества совместных распределений, удовлетворяющих условиям (6)–(8), единственного распределения, наименее уклоняющегося от независимого распределения в смысле критерия

$$f(r) = \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n}\left(r_{ij} - p_i q_j\right)^2 \to \min_r . \qquad (19)$$

Задача 1. В первой задаче целевая функция (19) минимизируется при ограничениях (6)–(8).

Функция Лагранжа этой задачи имеет вид

$$L(r,\lambda,\mu,\nu) = \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n}\left(r_{ij} - p_i q_j\right)^2 + \sum_{j=1}^{n}\lambda_j\left(\sum_{i=1}^{m} r_{ij} - q_j\right) +$$
$$+\sum_{i=1}^{m}\mu_i\left(\sum_{j=1}^{n} r_{ij} - p_i\right) + \nu\left(\sum_{i=1}^{m}\sum_{j=1}^{n} x_{1i}x_{2j}r_{ij} - c\sigma_1\sigma_2 - a_1 a_2\right).$$

Дифференцируя функцию Лагранжа по всем переменным и приравнивая производные к нулю, получаем уравнение

$$\partial L / \partial r_{kl} = (r_{kl} - p_k q_l) + \lambda_l + \mu_k + \nu x_{1k}x_{2l} = \\ = 0, k = 1,\ldots m, l = 1,\ldots n, \qquad (20)$$

а также выражения (6)–(8). Отметим, что ввиду соотношения $\sum_{i=1}^{m} p_i = \sum_{j=1}^{n} q_j = 1$ в системе (6)–(8) одно из уравнений, например первое, является следствием остальных, и его можно отбросить вместе с соответствующим множителем Лагранжа λ_1. Для решения полученной системы уравнений выразим элементы матрицы r из уравнения (20) в виде функции множителей Лагранжа λ, μ, ν:

$$r_{ij} = p_i q_j - \left(\lambda_j + \mu_i + \nu x_{1i}x_{2j}\right) = 0, (i,j) \in K . \qquad (21)$$

Подставив полученное выражение в уравнения (6)–(8), имеем:

$$-m\lambda_j - \sum_{i=1}^{m}\mu_i - \nu x_{2j}\sum_{i=1}^{m} x_{1i} = 0, j = 2,\ldots n ; \qquad (22)$$

$$-\sum_{j=1}^{n}\lambda_j - n\mu_i - \nu x_{1i}\sum_{j=1}^{n} x_{2j} = 0, i = 1,\ldots m ; \qquad (23)$$

$$\sum_{i=1}^{m}\sum_{j=1}^{n} x_{1i}x_{2j}\left(p_i q_j - \left(\lambda_j + \mu_i + \nu x_{1i}x_{2j}\right)\right) = c\sigma_1\sigma_2 + a_1 a_2 .$$

Обозначив $s_1 = \sum_{i=1}^{m} x_{1i}, s_2 = \sum_{j=1}^{n} x_{2j}, s_{12} = \sum_{i=1}^{m} x_{1i}^2 \sum_{j=1}^{n} x_{2j}^2$, преобразуем последнее уравнение к виду

$$-s_1\sum_{j=1}^{n}\lambda_j x_{2j} - s_2\sum_{i=1}^{m}\mu_i x_{1i} - \nu s_{12} = c\sigma_1\sigma_2 . \qquad (24)$$

Обозначим I_n единичную матрицу размера $n\times n$, а J_{mn} – прямоугольную матрицу размера $m \times n$, все элементы которой равны 1. Далее обозначим $\bar{x}_2 = (x_{22},\ldots,x_{2n})', \bar{x}_1 = (x_{11},\ldots,x_{1m})'$. В системе уравнений (6)–(8), как уже отмечалось, первое уравнение является следствием остальных, поэтому его можно отбросить. Кроме того, в полученной системе $m + n$ уравнений относительно $m + n + 1$ неизвестных одну из неизвестных можно выбрать произвольным образом. Мы будем полагать $\lambda_1 = 0$, что соответствует отбрасыванию первого столбца в матрице системы. После этих операций система уравнений запишется в виде

$$A\gamma = b , \qquad (25)$$

где неизвестные множители Лагранжа обозначены

$$\gamma = (\lambda_2,\ldots,\lambda_n,\mu_1,\ldots,\mu_n,\nu)', \qquad (26)$$

а матрица системы A и вектор правых частей имеют вид

$$A = \begin{pmatrix} -mI_{n-1} & -J_{n-1,m} & -s_1\bar{x}_2 \\ -J_{m,n-1} & -nI_m & -s_2\bar{x}_1 \\ -s_1\bar{x}_2' & -s_2\bar{x}_1' & -s_{12} \end{pmatrix}, \quad b = \left(0,\ldots,0,c\sigma_1\sigma_2\right) .$$

Квадратная система линейных уравнений (25) имеет единственное решение (26), добавляя в которое значение $\lambda_1 = 0$, из (21) вычисляем искомое распределение.

Задача 2. Целевая функция (19) минимизируется при ограничениях (6)–(8) и

$$r_{ij} \geq 0, (i,j) \in K . \qquad (27)$$

Аналитическое решение этой задачи в общем виде недоступно, однако численные методы позволяют эффективно решать ее. Приведем решения для рассмотренного примера.

Максимальному значению $c = c_{\max} = 8/\sqrt{102}$ соответствует решение

$$r = \begin{pmatrix} 1/3 & 1/15 & 0 \\ 0 & 13/30 & 1/6 \end{pmatrix},$$

которое, как нетрудно заметить, представляет собой комонотонное распределение (16).

Минимальному значению $c = c_{\min} = -7/\sqrt{102}$ соответствует решение

$$r = \begin{pmatrix} 0 & 7/30 & 1/6 \\ 1/3 & 4/15 & 0 \end{pmatrix},$$

представляющее собой антикомонотонное распределение (14).

Результаты работы позволяют эффективно строить дискретные вероятностные модели с заданной корреляционной структурой. Предложенная методика без труда обобщается на многомерный случай $d > 2$.

Библиографические ссылки

1. Новоселов А. А. Воспроизведение дискретных распределений с заданной ковариационной структурой // Материалы II Межрегион. конф. / КГТЭИ. Красноярск, 2009. С. 229–234.

2. Новоселов А. А. Дискретные распределения с заданной корреляцией, наименее отклоняющиеся от независимого // Тр. XIII Междунар. конф. по эвентологической математике и смежным вопросам / КГТЭИ ; СФУ. Красноярск, 2009. С. 126–131.

3. Nelsen R. B. An Introduction To Copulas. Springer, 1998.

A. A. Novosyolov

CONSTRUCTION OF MULTIDIMENSIONAL DISCRETE DISTRIBUTIONS WITH PREASSIGNED CORRELATION STRUCTURE

The paper is devoted to methods of construction of multidimensional discrete distributions with preassigned correlation structure and marginal distributions. The methods are based on mixture of the basic distributions and on decision of some optimization problems.

Keywords: discrete distribution, correlation, copula, mixture.

УДК 62-50:519.224

А. А. Новоселов

ПАРАМЕТРИЗАЦИЯ МОДЕЛЕЙ УПРАВЛЯЕМЫХ СИСТЕМ

Описывается применение метода ортогональных рядов для построения моделей управляемых систем параметрического вида в условиях непараметрической неопределенности. Ключевым элементом метода является выбор длины ортогонального ряда по данным наблюдений, т. е. определение параметрической структуры модели. Продемонстрировано применение метода к оцениванию плотности распределения и функции регрессии. Предложены пути обобщения оценок на многомерный случай.

Ключевые слова: плотность распределения, функция регрессии, ортогональный ряд, непараметрическая оценка.

В классической теории управления [1] обычно используются параметрические модели, которые строятся в условиях параметрической неопределенности. Здесь предполагается известным факт принадлежности модели к некоторому заранее определенному конечномерному классу моделей. Процедура идентификации (оценивания) модели сводится при этом к оцениванию по данным наблюдений некоторого конечного набора параметров.

Столь полную априорную информацию можно считать доступной далеко не всегда, поэтому параллельно развивается и непараметрическая теория управления [2; 3], в которой априорный класс моделей предполагается бесконечномерным; при этом говорят о непараметрической неопределенности. Малое количество априорной информации приходится компенсировать при этом большим количеством статистических наблюдений.

Основными строительными блоками для создания математических моделей систем управления по наблюдениям являются оценки плотности распределения и функции регрессии [1–3]. Непараметрическая оценка плотности распределения ядерного типа была впервые рассмотрена в [4; 5], а соответствующая непараметрическая оценка функции регрессии – в работах [6; 7].

Классические параметрические оценки этих функций гораздо проще для человеческого восприятия и лучше приспособлены для анализа, чем непараметрические оценки. Поэтому постоянно предпринимаются попытки построения оценок параметрического вида в условиях непараметрической неопределенности. Это направление можно назвать параметризацией моделей.

Одной из ветвей этого направления является использование оценок в виде отрезков ортогональных рядов, называемых еще проекционными оценками ввиду прозрачной геометрической аналогии. Оценка такого типа для плотности распределения была предложена Н. Н. Ченцовым в [8], а оценка для функции регрессии – в [9].

Проекционные оценки были подвергнуты тщательному изучению. Так, в статьях [10–13] исследовались вопросы сходимости оценок в различных смыслах, в публикациях [14; 15] проводилось сравнение различных непараметрических оценок. В [16–19] исследовались проекционные оценки в конкретных базисах. Во всех этих исследованиях длина отрезка ортогонального ряда выбиралась произвольно. В работах [20–24] предложен некий принцип оптимального выбора длины отрезка ортогонального ряда в проекционных оценках, исследованы свойства получаемых оптимальных проекционных оценок. В настоящей работе приведен обзор результатов [20–24], и представлены соображения по поводу обобщения оптимальных проекционных оценок на многомерный случай.

Проекционное приближение. Пусть L_r^2 – гильбертово пространство функций f, заданных на подмножестве U вещественной оси и интегрируемых с квадратом с весом r в смысле

$$\int_U f^2(x)r(x)dx < \infty,$$