Некоторые свойства неоднородных марковских последовательностей с периодическими матрицами переходных вероятностей

Огородников В.А., Савельев Л.Я., Каргаполова Н.А
Институт вычислительной математики и математической геофизики СО РАН, Новосибирск;
Новосибирский государственный университет, Новосибирск;
ova@osmf.sscc.ru

Введение. При исследовании реальных временных рядов, например, метеорологических (балл облачности, количество осадков температура и т.д.) возникает необходимость моделирования случайных рядов с конечным числом состояний. Такие модели могут быть использованы для моделирования любых метеорологических рядов, но при условии, что рассматриваются не все допустимые значения исследуемой величины, а только некоторые их градации. Возможны несколько подходов к моделированию дискретных рядов с заданными вероятностными свойствами. Достаточно часто применяется метод, основанный на использовании марковских моделей различной степени связности [2,3]. Другой распространенный метод моделирования основан на пороговом преобразовании специально подобранного гауссовского процесса [3]. При исследоваении реальных процессов важную роль играют процессы с периодическими свойствами. В работе [1] исследованы различные классы периодических случайных процессов, а также рассмотрены области их применения.

В данной работе рассматриваются марковские неоднородные последовательности с двумя состояниями, у которых периодически изменяются матрицы переходных вероятностей, и исследуются их свойства. Последовательности такого типа могут быть использованы, например, для моделирования индикаторов осадков, выходов температуры воздуха за заданный уровень и т.д. с учетом суточного хода. Для иллюстрации использованы данные многолетних наблюдений на станции Астрахань.

1. Неоднородная марковская цепь с периодически изменяющимися матрицами переходных вероятностей

a1. Определение процесса $\xi(k)$. Рассмотрим двоичную неоднородную марковскую последовательность ξ случайных переменных $\xi(k)$, $k \geq 0$ с множеством значений $C = \{1, 0\}$, начальным вектором A и переходными вероятностями Q, R

$$A = (a_1,\ a_0) = (a,\ 1-a),$$
$$Q = \begin{pmatrix} q_{11} & q_{10} \\ q_{01} & q_{00} \end{pmatrix} = \begin{pmatrix} p & 1-p \\ 1-q & q \end{pmatrix},\quad R = \begin{pmatrix} p_{11} & p_{10} \\ p_{01} & p_{00} \end{pmatrix} = \begin{pmatrix} r & 1-r \\ 1-s & s \end{pmatrix},$$

$$\Pr[\xi[0] = \alpha] = a_\alpha,$$
$$\Pr[\xi[2i+1] = \beta\,|\xi[2i] = \alpha] = q_{\alpha\beta},\quad \Pr[\xi[2i+2] = \beta\,|\xi[2i+1] = \alpha] = r_{\alpha\beta},$$
$$(i \geq 0,\ \ \alpha = 0, 1,\ \ \beta = 0, 1).$$

Переходные матрицы Q, R применяются последовательно, начиная с Q. Благодаря стохастичности они определяются четырьмя независимыми параметрами p,q,r,$s \in [0, 1]$.

Общая матрица вероятностей перехода $p_{\alpha\beta}[k] = \Pr[\xi[k+1] = \beta\,|\xi[k] = \alpha]$ имеет вид

$$P[k] = (1-\theta[k])Q + \theta[k]R = \begin{pmatrix} p_{11}[k] & p_{10}[k] \\ p_{01}[k] & p_{00}[k] \end{pmatrix} = \begin{pmatrix} p[k] & 1-p[k] \\ 1-q[k] & q[k] \end{pmatrix},$$

где $\theta[2i-1] = 0, \quad \theta[2i] = 1, \quad i \geq 1$ и $p_{\alpha\beta}[k] = (1-\theta[k])q_{\alpha\beta} + \theta[k]r_{\alpha\beta}$.

а2. Распределение $\xi(k)$. Верны следующие равенства для произведений матриц $P[k]$:

$$\prod_{k=1}^{2m} P[k] = (QR)^m, \quad \prod_{k=1}^{2m+1} P[k] = (QR)^m Q, \quad m \geq 1.$$

Положим $S = QR$. Тогда

$$S = QR = \begin{pmatrix} 1-s+pt & s-pt \\ r-qt & 1-r+qt \end{pmatrix} = \begin{pmatrix} 1-s & s \\ r & 1-r \end{pmatrix} + t\begin{pmatrix} p & -p \\ -q & q \end{pmatrix},$$

где $t = r+s-1$. Будут использоваться также обозначения

$$d = Det[s] = Det[Q]Det[R] = (p+q-1)(r+s-1),$$
$$b = (r-qt)/(1-d), \; d \neq 1.$$

По индукции легко показать, что верно следующее выражение для m-той степени матрицы S

$$S^m = \begin{pmatrix} b+(1-b)d^m & 1-b-(1-b)d^m \\ b-bd^m & 1-b+bd^m \end{pmatrix}.$$

Распределение

$$(\Pr[\xi(2m)=1], \; \Pr[\xi(2m)=0]) = AS^m = (p_{2m}, 1-p_{2m})$$

в четный момент $r = 2m, \; m \geq 1$ имеет вид:

$$p_{2m} = \Pr[\xi(2m)=1] = b + (a-b)d^m,$$
$$1-p_{2m} = \Pr[\xi(2m)=0] = 1-b-(a-b)d^m,$$

а распределение

$$(\Pr[\xi(2m+1)=1], \; \Pr[\xi(2m+1)=0]) = AS^m Q = (p_{2m+1}, 1-p_{2m+1})$$

в нечетный момент $r = 2m+1, \; m \geq 1$ имеет вид:

$$p_{2m+1} = \Pr[\xi[2m+1]=1] = 1-q+bu+(a-b)ud^m,$$
$$1-p_{2m+1} = \Pr[\xi[2m+1]=0] = q-bu-(a-b)ud^m,$$

где $u = p+q-1$.
а3. Пределы $\Pr[\xi(2m)=1], \; \Pr[\xi(2m)=1]$. Если $|d| < 1$, то четное и нечетное предельные распределения имеют вид

$$f_\infty = \lim_{m\to\infty} \Pr[\xi[2m]=1] = b, \;\; g_\infty = \lim_{m\to\infty} \Pr[\xi[2m+1]=1] = 1-q+bu$$

а4. Распределение длительностей 1-серий. Будем рассматривать стационарную последовательность. Распределение длительностей 1-серий, начинающихся в четные моменты времени, равно

$$P(L_1 = k) = \frac{P(\xi[2t-1] = 0, \xi[2t] = 1, \dots, \xi[2t+k-1] = 1, \xi[2t+k] = 0)}{P(\xi[2t-1] = 0, \xi[2t] = 1)}, k = 1,\ 2,\ \dots$$

Тогда

$$P(L_1 = 2m) = p^m r^{m-1}(1-r), \quad k = 2m, \quad m = 1, 2, \dots,$$
$$P(L_1 = 2m-1) = p^{m-1} r^{m-1}(1-p), \quad k = 2m-1, \quad m = 1, 2, \dots.$$

Если серии начинаются с нечетного момента времени, то

$$P(L_1 = 2m) = p^{m-1} r^m(1-p), \quad k = 2m, \quad m = 1, 2, \dots,$$
$$P(L_1 = 2m-1) = p^{m-1} r^{m-1}(1-r), \quad k = 2m-1, \quad m = 1, 2, \dots.$$

Среднее значение и дисперсия длительности серии, в случае, когда серия начинается с элемента с четным номером, имеют вид

$$ML_1 = \sum_{m=1}^{\infty} (2m)\, p^m r^{m-1}\,(1-r) + \sum_{m=1}^{\infty} (2m-1)\, p^{m-1} r^{m-1}\,(1-p) = \tfrac{1+p}{(1-pr)},$$
$$DL_1 = ML_1^2 - (ML_1)^2 = \tfrac{-p^2+2p+4pr-1}{(1-pr)^2} + \tfrac{(1-p)}{(1-pr)}.$$

Если серия начинается с нечетного места, то

$$ML_1 = \frac{1+r}{(1-pr)}, \;\; DL_1 = \frac{-r^2+2r+4pr-1}{(1-pr)^2} + \frac{(1-r)}{(1-pr)}.$$

а5. Корреляционная функция. Корреляционная функция в стационарном случае выражается соотношением

$$corr(\xi_t, \xi_{t+h}) = \frac{P(\xi_t = 1, \xi_{t+h} = 1) - P(\xi_t = 1)P(\xi_{t+h} = 1)}{\sqrt{P(\xi_t = 1) - P^2(\xi_t = 1)}\sqrt{P(\xi_{t+h} = 1) - P^2(\xi_{t+h} = 1)}}.$$

При $m \geq 0, \;\; k \geq 1$ необходимо рассмотреть 4 случая.

$$t = 2m, \;\; h = 2k-1: \qquad corr\left(\xi_{2m}, \xi_{2m+2k-1}\right) = \tfrac{\sqrt{b(1-b)}u}{\sqrt{(1-q+bu)(q-bu)}} d^{k-1},$$
$$t = 2m, \;\; h = 2k: \qquad corr\left(\xi_{2m}, \xi_{2m+2k}\right) = \tfrac{b(1-b)d^k}{b(1-b)} = d^k,$$
$$t = 2m+1, \; h = 2k-1: \quad corr\left(\xi_{2m+1}, \xi_{2m+1+(2k-1)}\right) = \tfrac{\sqrt{(1-q+bu)(r-b)}}{\sqrt{(q-bu)}\sqrt{b(1-b)}} d^{k-1},$$
$$t = 2m+1, \;\; h = 2k: \qquad corr\left(\xi_{2m+1}, \xi_{2m+1+2k}\right) = \tfrac{(r-b)u}{(q-bu)} d^{k-1}.$$

2. Распределения длительностей выхода температуры воздуха за заданный уровень.

С помощью рассмотренных марковских последовательностей построена численная стохастическая модель индикаторов выхода температуры воздуха за заданный уровень c. По построенной модели были получены соответствующие распределения длительностей для различных месяцев и уровней. На Рис. 1,2 приведены графики распределений $P(L_1 = k)$ для марта и декабря, полученные по модели и реальным данным (кривые 1, 3). Из рисунков видно, что рассмотренная выше марковская модель хорошо описывает затухающий колебательный характер распределений с периодом колебаний равным одним суткам. Однако, по величине эти вероятности значительно различаются, поэтому наряду с односвязной моделью была рассмотрена также численная двусвязная марковская модель с периодическими матрицами переходных вероятностей. Расчеты показали, что двусвязная модель также хорошо описывает колебательный характер распределений и существенно точнее воспроизводит значения реальных вероятностей (кривая 2 на Рис. 1,2).

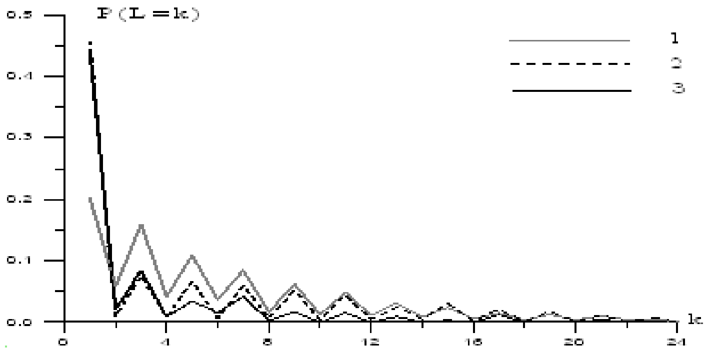

Рис. 1: Распределение длительности $P(L_1 = k)$. 1 – односвязная, 2 – двусвязная марковские модели, 3 – реальные данные. Март, $c = -2\ ^0C$.

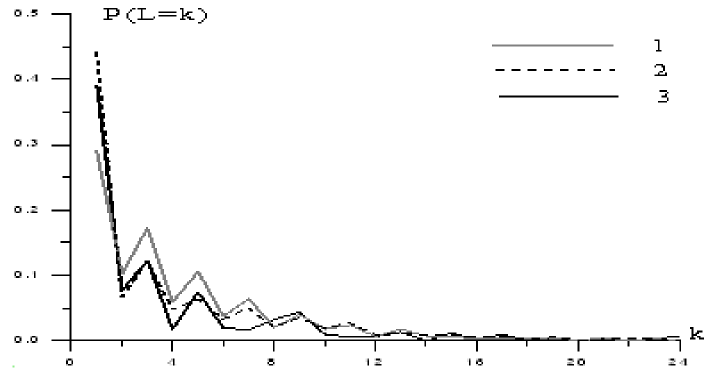

Рис. 2: Распределение длительности $P(L_1 = k)$. 1 – односвязная, 2 – двусвязная марковские модели, 3 – реальные данные. Декабрь, $c = -1\ ^0C$.

Следует отметить, что рассмотренные модели дают приемлемые результаты не для всех возможных уровней c. Это можно объяснить тем, что реальные процессы не всегда обладают марковскими свойствами, особенно в условиях учета суточного хода.

Работа выполнена при финансовой поддержке РФФИ (грант No 11-01-00641-а).

ЛИТЕРАТУРА

1. *Драган Я. П., Рожков В. А., Яворский И. Н.* Методы вероятностного анализа ритмики океанологических процессов. Л.: Гидрометеоиздат, 1987.

2. *Савельев Л. Я., Балакин С. В.* Совместные распределения чисел значенийи серий в троичных марковских последовательностях. Дискретная математика, том , вып. ,2004, т.16, No 3. С. 43-62.

3. *V. A.Ogorodnikov and S. M.Prigarin* Numerical Modelling of Random Processes and Fields: Algorithms and Applications. VSP, Utrecht, the Netherlands, 1996.