

УДК 519.2:53.05

Ф.Ф. Идрисов

РЕКУРРЕНТНЫЕ МОДЕЛИ
РАНДОМИЗИРОВАННЫХ ПРОЦЕССОВ
В ЗАДАЧАХ ВНЕШНЕЙ РАЗВЕДКИ. Часть 2.
УПРОЩЕННЫЕ МОДЕЛИ ВЫСШЕГО ПОРЯДКА*

Ключевые слова: разведанные, рандомизированный процесс, тренд, сплайн, дисперсия, ковариация оценок, рекуррентные оценки, имитационное моделирование.

Введение

Данная публикация является продолжением работы [1], где исследованы рекуррентные алгоритмы оценивания параметров простейших моделей (в виде сплайнов первого порядка). Вместе с тем особый интерес представляет некоторое усложнение данной задачи: построение упрощенных моделей в виде сплайнов второго и третьего порядков (в целях удобства принята сквозная нумерация формул для обеих частей).

Выделение тренда сплайном второго порядка дефекта 2

Основные идеи представлены в работе [1], что позволяет вести дальнейшее изложение в более кратком формате.

Как и ранее, будем считать, что интервал $[0, T]$ разбит на n подынтервалов длины $T_0 = \frac{T}{n}$ и на отрезке $[rT_0, (r+1)T_0]$, $r = \overline{0, n-1}$. Тогда тренд $x_r(t)$ представим в виде

$$x_r(t) = a_r \left(r + 1 - \frac{t}{T_0} \right)^2 + a_{r+1} \left(\frac{t}{T_0} - r \right)^2 + b_r \left(r + 1 - \frac{t}{T_0} \right) \left(\frac{t}{T_0} - r \right),$$

так что $x_r(rT_0) = a_r$, $x_r((r+1)T_0) = a_{r+1}$, а коэффициент b_r определяет недостающее слагаемое в полиноме второй степени. Таким образом, сплайн сшит по значению тренда на концах интервалов, но не сшит по значению производной. При $rT_0 \leq t_i < (r+1)T_0$ измеренное значение $x_i = x(t_i)$ представим в виде

$$x_i = x_r(t_i) + n_i.$$

А. Моменты изменений в наблюдаемом объекте известны точно. Следуя работе [1], для оценки коэффициентов a_r, b_r, a_{r+1} рассмотрим статистики вида

$$S_r = \frac{1}{\lambda T_0} \sum_{i \in M_r} \left(\alpha + \beta \frac{t_i - rT_0}{T_0} + \gamma \left(\frac{t_i - rT_0}{T_0} \right)^2 \right) x_i. \quad (48)$$

* Исследование выполнено за счет проекта 8.9562 2017/8.9.
© Ф.Ф. ИДРИСОВ, 2019

Вычисляя математическое ожидание оценок (48) теми же приемами, что и в [1] (см. соотношение (7)–(13)), получим

$$M\{S_r\} = \int_0^1 (\alpha + \beta x + \gamma x^2)(a_r(1-x)^2 + b_r x(1-x) + a_{r+1}x^2) dx =$$

$$= a_r \left(\frac{\alpha}{3} + \frac{\beta}{12} + \frac{\gamma}{30} \right) + b_r \left(\frac{\alpha}{6} + \frac{\beta}{12} + \frac{\gamma}{20} \right) + a_{r+1} \left(\frac{\alpha}{3} + \frac{\beta}{4} + \frac{\gamma}{5} \right). \quad (49)$$

Если в (49) взять $\alpha = 9$, $\beta = -35$, $\gamma = 30$, то получится $M\{S_r\} = a_r$; при $\alpha = -18$, $\beta = 120$, $\gamma = -120$ получаем $M\{S_r\} = b_r$; наконец, при $\alpha = 3$, $\beta = -24$, $\gamma = 30$ — $M\{S_r\} = a_{r+1}$.

Это позволяет предложить рекуррентный алгоритм оценивания параметров сплайна. Оцениваются коэффициенты a_r, b_r для $r = \overline{0, n-1}$ по формулам

$$\hat{a}_r = \frac{1}{\lambda T_0} \sum_{i \in M_r} \left(9 - 36 \frac{t_i - rT_0}{T_0} + 30 \left(\frac{t_i - rT_0}{T_0} \right)^2 \right) x_i,$$

$$\hat{b}_r = \frac{1}{\lambda T_0} \sum_{i \in M_r} \left(-18 + 120 \frac{t_i - rT_0}{T_0} - 120 \left(\frac{t_i - rT_0}{T_0} \right)^2 \right) x_i,$$

а коэффициент a_n оценивается отдельно

$$\hat{a}_n = \frac{1}{\lambda T_0} \sum_{i \in M_r} \left(3 - 24 \frac{t_i - rT_0}{T_0} + 30 \left(\frac{t_i - rT_0}{T_0} \right)^2 \right) x_i.$$

При таком методе оценивания будут коррелированы лишь пары \hat{a}_r, \hat{b}_r для $r = \overline{0, n-2}$ и лишь при $r = n-1$ будет коррелирована тройка $\hat{a}_{n-1}, \hat{b}_{n-1}, \hat{a}_n$.

Следуя тому, как была выведена формула (14) в [1], найдем дисперсии и ковариацию оценок:

$$D\{\hat{a}_r\} = \frac{\sigma^2}{T_0} \int_0^1 (9 - 36x + 30x^2)^2 dx +$$

$$+ \frac{1}{T_0} \int_0^1 (9 - 36x + 30x^2)^2 (a_r(1-x^2) + b_r x(1-x) + a_{r+1}x^2)^2 dx.$$

Отсюда

$$D\{\hat{a}_r\} = 9 \frac{\sigma^2}{T_0} + \frac{1}{70T_0} (406a_r^2 + 7b_r^2 + 34a_{r+1}^2 + 61a_r b_r + 14a_r a_{r+1} + 13b_r a_{r+1}).$$

Аналогично можно записать следующее соотношение:

$$D\{\hat{b}_r\} = 84 \frac{\sigma^2}{T_0} + \frac{1}{35T_0} (668a_r^2 + 98b_r^2 + 668a_{r+1}^2 + 254a_r b_r + 196a_r a_{r+1} + 254b_r a_{r+1}),$$

$$\text{cov}(\hat{a}_r, \hat{b}_r) = -18 \frac{\sigma^2}{T_0} - \frac{1}{35T_0} (308a_r^2 + 11b_r^2 + 104a_{r+1}^2 +$$

$$+ 41a_r b_r + 22a_r a_{r+1} + 35b_r a_{r+1}). \quad (50)$$

Несмещенные оценки величин (50) имеют вид

$$\begin{aligned}\hat{D}\{\hat{a}_r\} &= \frac{1}{(\lambda T_0)^2} \sum_{i \in M_r} \left(9 - 36 \frac{t_i - rT_0}{T_0} + 30 \left(\frac{t_i - rT_0}{T_0} \right)^2 \right)^2 x_i^2, \\ \hat{D}\{\hat{b}_r\} &= \frac{1}{(\lambda T_0)^2} \sum_{i \in M_r} \left(-18 + 120 \frac{t_i - rT_0}{T_0} - 120 \left(\frac{t_i - rT_0}{T_0} \right)^2 \right)^2 x_i^2, \\ \text{cov}(\hat{a}_r, \hat{b}_r) &= \frac{1}{(\lambda T_0)^2} \sum_{i \in M_r} \left(9 - 36 \frac{t_i - rT_0}{T_0} + 30 \left(\frac{t_i - rT_0}{T_0} \right)^2 \right) \times \\ &\quad \times \left(-18 + 120 \frac{t_i - rT_0}{T_0} - 120 \left(\frac{t_i - rT_0}{T_0} \right)^2 \right) x_i^2.\end{aligned}\quad (51)$$

Б. Моменты изменений в наблюдаемом объекте неизвестны. Разбивая значения N изменений в наблюдаемом объекте на n групп по N_0 измерений в каждой, так что $N = nN_0$, по аналогии с (20) для k -й группы представим x_{kN_0+j} в виде

$$\begin{aligned}x_{kN_0+j} &= a_{k+1} \left(\frac{t_{kN_0+j}}{T_0} - k \right)^2 + a_k \left(k + 1 - \frac{t_{kN_0+j}}{T_0} \right)^2 + \\ &\quad + b_k \left(\frac{t_{kN_0+j}}{T_0} - k \right) \left(k + 1 - \frac{t_{kN_0+j}}{T_0} \right).\end{aligned}\quad (52)$$

Тогда с учетом (19), пренебрегая слагаемыми, имеющими порядок малости $1/N$ для (52), по аналогии с (22) получим соотношение

$$M\{x_{kN_0+j}|N\} = a_{k+1} \left(\frac{j}{N_0} \right)^2 + a_k \left(1 - \frac{j}{N_0} \right)^2 + b_k \frac{j}{N_0} \left(1 - \frac{j}{N_0} \right).$$

Рассмотрим теперь статистику вида

$$S_k = \frac{1}{N_0} \sum_{j=1}^{N_0} \left(\alpha + \beta \frac{j}{N} + \gamma \left(\frac{j}{N} \right)^2 \right) x_{kN_0+j},$$

для которой

$$\begin{aligned}M\{S_k|N\} &= a_{k+1} \frac{1}{N_0} \sum_{j=1}^{N_0} \left(\alpha + \beta \frac{j}{N} + \gamma \left(\frac{j}{N_0} \right)^2 \right) \frac{j^2}{N_0^2} + \\ &\quad + b_k \frac{1}{N_0} \sum_{j=1}^{N_0} \left(\alpha + \beta \frac{j}{N_0} + \gamma \left(\frac{j}{N_0} \right)^2 \right) \frac{j}{N_0} \left(1 - \frac{j}{N_0} \right) + \\ &\quad + a_k \frac{1}{N_0} \sum_{j=1}^{N_0} \left(\alpha + \beta \frac{j}{N_0} + \gamma \left(\frac{j}{N_0} \right)^2 \right) \left(1 - \frac{j}{N_0} \right)^2.\end{aligned}\quad (53)$$

Заменяя в (53) приближенно суммы интегралами при больших N_0 , получим

$$M\{S_k|N\} = a_{k+1} \left(\frac{\alpha}{3} + \frac{\beta}{4} + \frac{\gamma}{5} \right) + b_k \left(\frac{\alpha}{6} + \frac{\beta}{12} + \frac{\gamma}{20} \right) + a_k \left(\frac{\alpha}{3} + \frac{\beta}{12} + \frac{\gamma}{30} \right).$$

Поэтому оценки \hat{a}_r, \hat{b}_r следует брать в виде

$$\begin{aligned}\hat{a}_r &= \frac{1}{N_0} \sum_{j=1}^{N_0} \left(9 - 36 \frac{j}{N_0} + 30 \left(\frac{j}{N_0} \right)^2 \right) x_{kN_0+j}, \\ \hat{b}_r &= \frac{1}{N_0} \sum_{j=1}^{N_0} \left(-18 + 120 \frac{j}{N_0} - 120 \left(\frac{j}{N_0} \right)^2 \right) x_{kN_0+j}\end{aligned}\quad (54)$$

для $r = \overline{0, n-1}$. Оценку \hat{a}_n параметра a_n следует брать в виде

$$\hat{a}_n = \frac{1}{N_0} \sum_{j=1}^{N_0} \left(3 - 24 \frac{j}{N_0} + 30 \left(\frac{j}{N_0} \right)^2 \right) x_{kN_0+j}. \quad (55)$$

Вычислим теперь асимптотические дисперсии и ковариации оценок (54) и (55). Как и ранее [1, см. (26)], представим t_{kN_0+j}/T_0 в виде

$$\frac{t_{kN_0+j}}{T_0} = k + \frac{j}{N_0} + \frac{\Delta t_{kN_0+j}}{T_0}.$$

Тогда

$$\begin{aligned}x_{kN_0+j} &= a_{k+1} \left(\frac{j}{N_0} + \frac{\Delta t_{kN_0+j}}{T_0} \right)^2 + a_k \left(1 - \frac{j}{N_0} - \frac{\Delta t_{kN_0+j}}{T_0} \right)^2 + \\ &+ b_k \left(\frac{j}{N_0} + \frac{\Delta t_{kN_0+j}}{T_0} \right) \left(1 - \frac{j}{N_0} - \frac{\Delta t_{kN_0+j}}{T_0} \right) + n_{kN_0+j}.\end{aligned}\quad (56)$$

После разложения выражения (56) в ряд Тейлора по $\Delta t_{kN_0+j}/T_0$, получим

$$\begin{aligned}x_{kN_0+j} &= a_{k+1} \left(\frac{j}{N_0} \right)^2 + b_k \frac{j}{N_0} \left(1 - \frac{j}{N_0} \right) + a_k \left(1 - \frac{j}{N_0} \right)^2 + \frac{\Delta t_{kN_0+j}}{T_0} \times \\ &\times \left[2a_{k+1} \frac{j}{N_0} - 2a_k \left(1 - \frac{j}{N_0} \right) + b_k \left(1 - 2 \frac{j}{N_0} \right) \right] + n_{kN_0+j},\end{aligned}$$

так что, например, \hat{a}_k можно представить в виде

$$\begin{aligned}\hat{a}_k &= a_k + \frac{1}{N_0} \sum_{j=1}^{N_0} \left(9 - 36 \frac{j}{N_0} + 30 \left(\frac{j}{N_0} \right)^2 \right) n_{kN_0+j} + \\ &+ \frac{\Delta t_{kN_0+j}}{T_0} \left(c_k \frac{j}{N_0} + d_k \left(1 - \frac{j}{N_0} \right) \right),\end{aligned}\quad (57)$$

где, для краткости, обозначено $c_k = 2a_{k+1} - b_k$, $d_k = b_k - 2a_k$.

Отсюда, переходя в (57) сразу от сумм вида $\frac{1}{N_0} \sum_{j=1}^{N_0}$ к аппроксимирующим

их интегралам, получим

$$D\{\hat{a}_k\} = \frac{\sigma^2}{N_0} \int_0^1 (9 - 36x + 30x^2) dx + \frac{1}{N_0} \int_0^1 \int_0^1 (9 - 36x + 30x^2)(9 - 36y + 30y^2) \times \\ \times \left[\min(k+x, k+y) - \frac{(k+x)(k+y)}{n} \right] \cdot (c_k x + d_k(1-x))(c_k y + d_k(1-y)) dx dy.$$

Вычисления позволяют получить

$$D\{\hat{a}_k\} = 9 \frac{\sigma^2}{N_0} + \frac{1}{N_0} \left[d_k^2 k \left(1 - \frac{k}{n}\right) + \frac{1}{70} (c_k^2 + c_k d_k + 4d_k^2) \right].$$

Аналогично можно записать

$$D\{\hat{b}_k\} = 84 \frac{\sigma^2}{N_0} + \frac{1}{N_0} \left[(kc_k + (k+1)d_k) \left(1 - \frac{kc_k + (k+1)d_k}{n}\right) + \right. \\ \left. + \frac{1}{35} (32c_k^2 + 32d_k^2 + 61c_k d_k) \right],$$

$$\text{cov}(\hat{a}_k, \hat{b}_k) = -18 \frac{\sigma^2}{N_0} + \frac{1}{N_0} \left[d_k (kc_k + (k+1)d_k) \left(1 - \frac{k}{n}\right) - \right. \\ \left. - \frac{1}{70} (7c_k^2 + 69d_k^2 + 4c_k d_k) \right].$$

Оценка сплайном второго порядка тренда произвольного вида

Рассмотрим теперь вопрос о том, какова будет величина неустранимой ошибки, если сплайном второго порядка выравняется тренд произвольного вида $f(t)$. В этом случае, при $N \rightarrow \infty$, по аналогии с тем, как это сделано в [1], получим

$$a_k = \frac{1}{T_0} \int_{kT_0}^{(k+1)T_0} \left(9 - 36 \frac{t - kT_0}{T_0} + 30 \left(\frac{t - kT_0}{T_0} \right)^2 \right) f(t) dt = \\ = \int_0^1 (9 - 36x + 30x^2) f(xT_0 + kT_0) dx, \quad (58) \\ b_k = \int_0^1 (-18 + 120x - 120x^2) f(xT_0 + kT_0) dx.$$

Среднеквадратичная погрешность оценки тренда $f(t)$ сплайном второго порядка равна

$$\varepsilon^2 = \frac{1}{n} \sum_{k=0}^{n-1} \int_0^1 [f(yT_0 + kT_0) - a_k(1-y)^2 - b_k y(1-y) - a_{k+1} y^2]^2 dy,$$

где коэффициенты a_k и b_k даются формулами (58).

Рассчитаем величину ε для $f(t) = e^{-\alpha t}$. Тогда

$$a_k = a_0 e^{-\alpha k}, \quad b_k = b_0 e^{-\alpha k},$$

где $a = \alpha T_0$, и в результате получим

$$a_0 = \int_0^1 (9 - 36x + x^2) e^{-\alpha x} dx = \frac{3}{a^3} [3a^2 - 12a + 20 - (a^2 + 8a + 20)e^{-a}],$$

$$b_0 = \int_0^1 (-18 + 120x - 120x^2)e^{-ax} dx = \frac{6}{a^3} [(40 + 20a + 3a^2)e^{-a} - 3a^2 + 20a - 40].$$

В данном случае, как и ранее,

$$\varepsilon^2 = \varphi(a) \frac{1 - e^{-2an}}{n(1 - e^{-2a})},$$

где

$$\varphi(a) = \int_0^1 [e^{-ay}y - a_0e^{-a}y^2 - b_0y(1-y) - a_0(1-y)^2]^2 dy.$$

Явное выражение для $\varphi(a)$ не приводится исключительно из-за его громоздкости; разложение $\varphi(a)$ в ряд Тейлора по a имеет вид

$$\varphi(a) = \frac{11}{168000}a^6 - \frac{61}{504000}a^7 + \frac{8147}{63504000}a^8 - \frac{6341}{63504000}a^9 \pm \dots,$$

так что при малых a приближенно

$$\varepsilon \approx \sqrt{\varphi(a)} \approx \frac{a^3}{40} \cdot \sqrt{\frac{11}{105}} \approx 0,008a^3. \quad (59)$$

Сравнивая (59) с выражением (47), видим, что если при сплайне первого порядка ε пропорциональна a^2 , при сплайне второго порядка ε пропорциональна уже a^3 .

На рис. 2 для того же примера, что и на рис. 1, приведено выделение тренда в виде сплайна второго порядка дефекта 2.

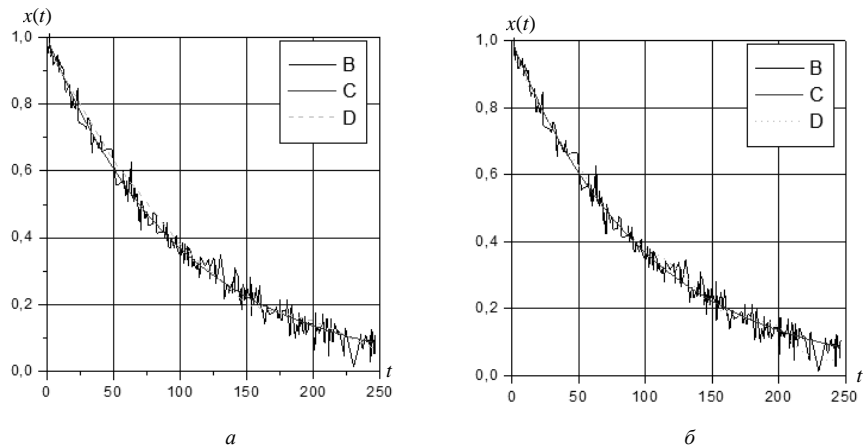


Рис. 1

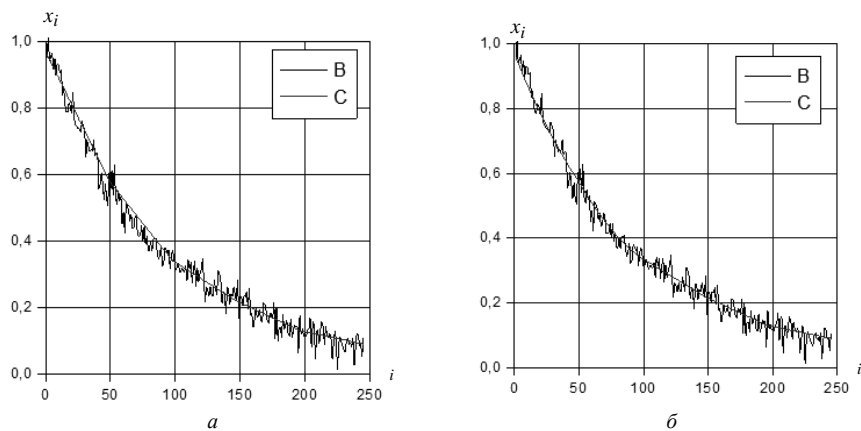


Рис. 2

Выделение тренда сплайном третьего порядка дефекта 2

Описанные выше сплайны имеют тот недостаток, что в точках сшивания у них наблюдается излом кривой, т.е. они не шиты по производной.

Этого дефекта можно было бы избежать при использовании сплайна второго порядка дефекта 1. Однако, как показали исследования [2], рекуррентное построение такого сплайна не приводит к хорошим результатам, так как начинается «раскачивание» производной на правом конце отрезка, а это приводит к тому, что выделенный сплайн мало похож на истинный тренд. С этим можно бороться, однако все равно ошибки выделения тренда остаются большими.

Более удовлетворительные результаты получаются при использовании сплайнов третьего порядка дефекта 2. Ниже приводятся основные формулы для них с минимумом пояснений.

В случае, когда моменты измерений известны точно, на k -м интервале $[kT_0, (k+1)T_0]$ сплайна выделяющий тренд лучше брать в форме

$$f(x) = a_k(1-x)^2(1+2x) + b_k x(1-x)^2 - b_{k+1}x^2(1-x) + a_{k+1}x^2(3-2x),$$

где $x = (t - kT_0)/T_0$. В этом случае

$$\begin{aligned} f(0) &= a_k; \quad f(1) = a_{k+1}; \\ f'(0) &= b_k; \quad f'(1) = b_{k+1}, \end{aligned}$$

т.е. коэффициенты a_k и b_k имеют ясный геометрический смысл.

Для оценки коэффициентов a_k, b_k следует использовать статистики

$$S_k = \frac{1}{\lambda T_0} \sum_{i \in M_k} \left(\alpha + \beta \frac{t_i - kT_0}{T_0} + \gamma \left(\frac{t_i - kT_0}{T_0} \right)^2 + \delta \left(\frac{t_i - kT_0}{T_0} \right)^3 \right) x_i,$$

математическое ожидание которых равно

$$\begin{aligned} M\{S_k\} &= a_k \left(\frac{1}{2} \alpha + \frac{3}{20} \beta + \frac{1}{15} \gamma + \frac{1}{28} \delta \right) + b_k \left(\frac{1}{12} \alpha + \frac{1}{30} \beta + \frac{1}{60} \gamma + \frac{1}{105} \delta \right) - \\ &- b_{k+1} \left(\frac{1}{12} \alpha + \frac{1}{20} \beta + \frac{1}{30} \gamma + \frac{1}{42} \delta \right) + a_{k+1} \left(\frac{1}{2} \alpha + \frac{7}{20} \beta + \frac{4}{15} \gamma + \frac{3}{14} \delta \right). \end{aligned}$$

Так как

$$\begin{bmatrix} \frac{1}{2} & \frac{2}{30} & \frac{1}{15} & \frac{1}{28} \\ \frac{1}{12} & \frac{1}{30} & \frac{1}{60} & \frac{1}{105} \\ \frac{1}{12} & \frac{1}{20} & \frac{1}{30} & \frac{1}{42} \\ \frac{1}{2} & \frac{7}{20} & \frac{4}{15} & \frac{3}{14} \end{bmatrix}^{(-1)} = \begin{bmatrix} 16 & -120 & 60 & -4 \\ -120 & 1200 & -840 & 60 \\ 240 & -2400 & 2340 & -180 \\ -140 & 1680 & -1680 & 140 \end{bmatrix},$$

этим определяется выбор коэффициентов $\alpha, \beta, \gamma, \delta$. При $\alpha = 16, \beta = -120, \gamma = 240, \delta = -240$ получается $M\{S_k\} = a_k$; при $\alpha = -120, \beta = 1200, \gamma = -2400, \delta = 1680$ — $M\{S_k\} = b_k$. Это позволяет оценивать коэффициенты a_k, b_k для $k = \overline{0, n-1}$,

и лишь коэффициенты a_n, b_n оцениваются по коэффициентам, стоящим в двух последних столбцах обратной матрицы.

Не будем выписывать явных выражений для дисперсий и ковариации оценок \hat{a}_k, \hat{b}_k — их легко получить с помощью программы MathCad. Скажем только, что для пуассоновского потока моментов измерений их несмещенные оценки имеют вид

$$\begin{aligned} \hat{D}\{\hat{a}_k\} &= \frac{1}{(\lambda T_0)^2} \sum_{i \in M_k} \left(16 - 120 \frac{t_i - kT_0}{T_0} + 240 \left(\frac{t_i - kT_0}{T_0} \right)^2 - 140 \left(\frac{t_i - kT_0}{T_0} \right)^3 \right)^2 x_i^2, \\ \hat{D}\{\hat{b}_k\} &= \frac{1}{(\lambda T_0)^2} \sum_{i \in M_k} \left(-120 + 1200 \frac{t_i - kT_0}{T_0} - 2400 \left(\frac{t_i - kT_0}{T_0} \right)^2 + 1680 \left(\frac{t_i - kT_0}{T_0} \right)^3 \right)^2 x_i^2, \\ \text{cov}(\hat{a}_k, \hat{b}_k) &= \frac{1}{(\lambda T_0)^2} \sum_{i \in M_k} \left(16 - 120 \frac{t_i - kT_0}{T_0} + 240 \left(\frac{t_i - kT_0}{T_0} \right)^2 - 140 \left(\frac{t_i - kT_0}{T_0} \right)^3 \right) \times \\ &\quad \times \left(-120 + 1200 \frac{t_i - kT_0}{T_0} - 2400 \left(\frac{t_i - kT_0}{T_0} \right)^2 + 1680 \left(\frac{t_i - kT_0}{T_0} \right)^3 \right) x_i^2. \end{aligned}$$

В случае, когда относительно моментов измерений известен лишь их порядок, оценки \hat{a}_k, \hat{b}_k коэффициентов a_k, b_k имеют вид

$$\frac{1}{N_0} \sum_{j=1}^{N_0} \left(\alpha + \beta \frac{j}{N_0} + \gamma \left(\frac{j}{N_0} \right)^2 + \delta \left(\frac{j}{N_0} \right)^3 \right) x_{kN_0+j}$$

с теми же наборами значений $\alpha, \beta, \gamma, \delta$, что и выше. Так можно оценить a_k, b_k для $k = \overline{0, n-1}$, и только a_n, b_n требуют отдельных оценок.

Представляя t_{kN_0+j}/T_0 в виде

$$\frac{t_{kN_0+j}}{T_0} = k + \frac{j}{N_0} + \frac{\Delta t_{kN_0+j}}{T_0},$$

легко получить

$$\begin{aligned} x_{kN_0+j} &= a_k \left(1 - \frac{j}{N_0} \right)^2 \left(1 + 2 \frac{j}{N_0} \right) + b_k \frac{j}{N_0} \left(1 - \frac{j}{N_0} \right)^2 - \\ &- b_{k+1} \left(\frac{j}{N_0} \right)^2 \left(1 - \frac{j}{N_0} \right) + a_{k+1} \left(\frac{j}{N_0} \right)^2 \left(3 - 2 \frac{j}{N_0} \right) + \frac{\Delta t_{kN_0+j}}{T_0} \varphi \left(\frac{j}{N_0} \right) + n_{kN_0+j}, \end{aligned}$$

где

$$\varphi(x) = -6a_k x(1-x) + b_k(1-x)(1-3x) - b_{k+1}x(2-3x) + 6a_{k+1}x(1-x).$$

Это позволяет вычислять дисперсии оценок \hat{a}_k, \hat{b}_k и их ковариацию

$$D\{\hat{a}_k\} = 16 \frac{\sigma^2}{N_0} + \frac{1}{N_0} \int_0^1 \int_0^1 (16 - 120x + 240x^2 - 140x^3)(16 - 120y + 240y^2 - 140y^3) \times$$

$$\begin{aligned} & \times \varphi(x)\varphi(y) \left[\min(k+x, k+y) - \frac{(k+x)(k+y)}{n} \right] dx dy, \\ D\{\hat{b}_k\} &= 1200 \frac{\sigma^2}{N_0} + \frac{1}{N_0} \int_0^1 \int_0^1 (-120 + 1200x - 2400x^2 + 1680x^3) \times \\ & \times (-120 + 1200y - 2400y^2 + 1680y^3) \varphi(x)\varphi(y) \left[\min(k+x, k+y) - \frac{(k+x)(k+y)}{n} \right] dx dy, \\ \text{cov}(\hat{a}_k, \hat{b}_k) &= -120 \frac{\sigma^2}{N_0} + \frac{1}{N_0} \int_0^1 \int_0^1 (16 - 120x + 240x^2 - 140x^3) (-120 + \\ & + 1200y - 2400y^2 + 1680y^3) \varphi(x)\varphi(y) \left[\min(k+x, k+y) - \frac{(k+x)(k+y)}{n} \right] dx dy. \end{aligned}$$

Явные выражения для этих величин легко получить с помощью программы MathCad.

Следует отметить, однако, что из-за того, что коэффициенты α , β , γ и δ очень велики, начинают сильно сказываться случайные ошибки. Это видно из выражений для дисперсий оценок — например, в $D\{\hat{b}_k\}$ перед $\frac{\sigma^2}{N_0}$ стоит коэффициент 1200. Поэтому такими сплайнами можно пользоваться лишь при очень больших N_0 — порядка 300 и более, что редко встречается на практике.

Имитационное моделирование

Опишем теперь кратко результаты имитационного моделирования предлагаемых сплайновых оценок.

На рис. 1, *a*, *б* показано выделение тренда для реализации с $\lambda = 1$ и $T = 250$. В этой реализации оказалось 245 наблюдений.

Тренд брался в виде $f(t) = e^{-0,01t}$, величины n_i — как нормальные случайные величины с нулевым математическим ожиданием и среднеквадратичным отклонением, равным 0,03.

На рис. 1, *a* показано выделение тренда сплайном первого порядка, когда моменты изменений наблюдаемого объекта считались точно известными. Параметр T_0 брался равным 50, так что весь интервал наблюдений разбит на пять отрезков. На рис. 1, *a* кривая *B* — это величины x_i , соединенные отрезками прямых, кривая *C* — истинный тренд, кривая *D* — тренд в виде сплайна первого порядка. На рис. 1, *б* для этой же реализации выделение тренда проведено в виде сплайна второго порядка. Как видно из рис. 1, *б*, выделение тренда достаточно хорошее. Хотя доверительные интервалы не приведены, во всех случаях истинный тренд лежит внутри 95 % доверительного интервала, построенного к оценке тренда в виде сплайна.

На рис. 2, *a*, *б* для той же реализации приведено выделение тренда в виде сплайна первого и второго порядков, когда моменты изменений считались неизвестными. Здесь по горизонтальной оси отложен номер изменения, по вертикальной — сами изменения. Истинный тренд не приведен, так как не совсем понятно, как его строить — ведь моменты измерений неизвестны. Приведен лишь выде-

ленний тренд в виде сплайна первого (рис. 2, а) и второго (рис. 2, б) порядков. И здесь выделение тренда достаточно хорошее.

В общем случае выделение тренда сплайнами первого и второго порядков достаточно хорошее, когда величина N_0 или λT_0 больше 50. Сплайны третьего порядка требуют больших значений этих величин.

Заключение

Итак, рассмотрена задача выделения трендов рандомизированных наблюдений в виде сплайнов. Описаны следующие типы сплайнов:

- сплайны первого порядка (кусочно-ломаные прямые);
- сплайны второго порядка дефекта 2;
- сплайны третьего порядка дефекта 2.

Во всех случаях рассмотрены две ситуации:

- моменты измерений известны точно;
- моменты измерений неизвестны, известен лишь их порядок.

Поток моментов измерений предполагался пуассоновским.

В данной работе для всех случаев

- построены оценки параметров сплайна;
- найдены дисперсии и ковариации оценок этих параметров;
- построены оценки дисперсий и ковариаций;
- оценена неустранимая погрешность при выделении тренда произвольного вида сплайнами.

Имитационное моделирование показало, что наиболее хорошо тренды выделяются сплайнами первого или второго порядков. Сплайны третьего порядка неустойчивы к случайным ошибкам, и их применение возможно лишь при очень большом объеме выборки.

Ф.Ф. Ідрісов

РЕКУРЕНТНІ МОДЕЛІ РАНДОМІЗОВАНИХ ПРОЦЕСІВ У ЗАДАЧАХ ЗОВНІШНЬОЇ РОЗВІДКИ. Частина 2. СПРОЦЕНІ МОДЕЛІ ВИЩОГО ПОРЯДКУ

Зовнішня розвідка в усі часи залишалася сферою особливої уваги багатьох держав. Її значення особливо зросло в даний час у зв'язку з гостросюжетним характером їх відносин. З'явилися нові форми конфліктів: мережеві та гібридні війни, масштабні терористичні акти, ускладнився простір їх проведення — до традиційних морського, сухопутного та повітряного просторів додалася нова сфера — кіберпростір. Розробка інструментів зовнішньої розвідки — процес тривалий та дорогий, вимагає постійного внесення доповнень і змін відповідно до подій на міжнародній арені. Однак використання для цих цілей теорії випадкових процесів наштовхується на одну вельми непросту проблему. Справа в тому, що класична теорія має на увазі обробку даних в умовах еквідистантних спостережень, в той час як в реальності розвіддані є потоки даних, одержуваних у випадкові моменти часу. Більш того, може надходити випадкове число даних у випадкові моменти часу i , нарешті, може бути і така ситуація, коли розвідданих настільки мало, що неможливо зробити будь-які висновки щодо прийняття рішення. Для таких умов, що дуже часто зустрічаються, в роботі запропоновано сплайнові моделі різного ступеня складності, що дозволяють синтезувати адаптивні алгоритми при точно відомих та невідомих моментах змін стану об'єкта, що розвідують.

Як модель потоку подій використовується пуассоновський процес. Коректність алгоритмів забезпечується математичною строгістю наведених міркувань, наведено результати імітаційного моделювання.

Ключові слова: розвідані, рандомізований процес, тренд, сплайн, дисперсія, коваріація оцінок, рекурентні оцінки, імітаційне моделювання.

F.F. Idrisov

RECURRENT MODELS OF RANDOMIZED PROCESSES IN FOREIGN INTELLIGENCE TASKS. PART II. SIMPLIFIED HIGH ORDER MODELS

Foreign intelligence at all times has always remained the focus of attention of many states. Its significance has particularly increased at the present time in connection with the acute nature of their relationship. New forms of conflicts have appeared: network and hybrid wars, large-scale terrorist acts, the space for their conduct has become more complicated — cyberspace has been added to the traditional maritime, land and air spaces. The development of foreign intelligence tools is a long and expensive process, requiring constant additions and changes in accordance with events in the international arena. However, the use of the theory of random processes for these purposes encounters one very difficult problem. The fact is that the classical theory implies the processing of data under conditions of equidistant observations, while in reality intelligence is a stream of data obtained at random times. Moreover, a random amount of data may arrive at random times and, finally, there may be a situation where intelligence is so small that it is impossible to draw any conclusions on the decision. For such very often encountered conditions, spline models of varying degrees of complexity are proposed in the work, which allow synthesizing adaptive algorithms with precisely known and unknown moments of changes in the state of the reconnaissance object. The Poisson process is used as an event flow model. The correctness of the algorithms is ensured by the mathematical rigor of the above reasoning; the results of simulation are presented.

Keywords: intelligence, randomized process, trend, spline, variance, covariance of estimates, recurrence estimates, simulation.

1. Идрисов Ф.Ф. Рекуррентные модели рандомизированных процессов в задачах внешней разведки. Часть 1. Упрощенные модели низшего порядка. *Международный научно-технический журнал «Проблемы управления и информатики»*. 2019. № 5. С. 137–149.
2. Лифшиц К.И. Сглаживание экспериментальных данных сплайнами. Томск : Изд-во Том. ун-та, 1991. 180 с.

*Получено 29.03.2019
После доработки 21.11.2019*