МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

DOI: 10.24143/2072-9502-2021-3-126-133
УДК 681.5.015.3+001.18

МЕТОД ПОШАГОВОГО СГЛАЖИВАНИЯ ЭКСПЕРИМЕНТАЛЬНЫХ ЗАВИСИМОСТЕЙ ДЛЯ ЗАДАЧ КРАТКОСРОЧНОГО ПРОГНОЗИРОВАНИЯ

А. А. Ермаков, Т. К. Кириллова

*Иркутский государственный университет путей сообщения,
Иркутск, Российская Федерация*

В работе рассмотрен метод пошагового сглаживания как одного из возможных алгоритмов краткосрочного прогнозирования статистики равноточных измерений монотонных функций, представляющей собой значения определяющих параметров, которые оценивают динамику состояний сложных технических систем по наработке. Истинное значение самого контролируемого параметра считается неизвестным, а обрабатываемые значения измерений распределены нормально. Измерения подвергаются обработке методом пошагового сглаживания. В результате обработки образуется новая статистика, представляющая собой статистику прогнозов, каждое значение которой представляет собой полусумму самого измерения и так называемого частного прогноза. Доказывается, что полученные таким образом прогнозы имеют, во-первых, тот же закон распределения, что и закон распределения выборки равноточных измерений; во-вторых, тренд прогнозов должен быть тем же, что и тренд измерений, и соответствовать теоретическому тренду, т. е. истинным значениям монотонной функции; в-третьих, дисперсия полученной статистики должна быть не больше дисперсии исходной выборки. Делается вывод, что метод пошагового сглаживания может быть предложен для краткосрочного прогнозирования.

Ключевые слова: монотонная функция, измерения, прогноз, частный прогноз, тренд, дисперсия, метод пошагового сглаживания, временной интервал, оценка измерения, закон распределения.

Для цитирования: *Ермаков А. А., Кириллова Т. К.* Метод пошагового сглаживания экспериментальных зависимостей для задач краткосрочного прогнозирования. Серия: Управление, вычислительная техника и информатика. 2021. № 3. С. 126–133. DOI: 10.24143/2072-9502-2021-3-126-133.

Введение

Для эффективного обслуживания сложных технических систем (СТС) в процессе их использования, а также при исследовании поведения этих систем по наработке необходимо обрабатывать большие временные массивы значений выходных параметров. Обработка таких данных любыми методами [1, 2] часто предполагает предварительную обработку исходной статистики для снижения влияния случайных составляющих. Предлагаем рассмотреть один из возможных методов предварительной обработки статистических данных, а именно метод пошагового сглаживания (МПС). Суть метода заключается в вычислении прогнозной оценки измерения $\tilde{y}_i$ в виде полусуммы значений измерения в i-й момент времени y_i^* и значения частного прогноза $\tilde{y}_{(i-1)i}$ на i-й момент времени, полученного в предыдущий $(i-1)$ момент:

$$\tilde{y}_i = \frac{y_i^* + \tilde{y}_{(i-1)i}}{2}. \qquad (1)$$

Несомненным достоинством метода является то, что МПС обеспечивает обработку измерений, в отличие, например, от метода скользящего среднего [3], начиная с самого первого измерения. В дальнейшем используется вся предыстория исходного временного ряда. В результате обработки возникает новый случайный процесс с ожидаемой меньшей, в среднем, дисперсией.

Кроме того, с помощью МПС расчеты можно вести сразу после первого измерения, что упрощает вычисления, в отличие от других методов, где для первой и последующих оценок необходимо использовать группу последовательных измерений.

Результаты, полученные в [3], лишь только предполагают свойства метода, но не доказывают их. Настоящая работа направлена на строгое доказательство свойств МПС, которые бы позволили ему быть конкурентоспособным в ряду других методов, осуществляющих предварительное сглаживание.

Вычисления прогнозной оценки измерения параметров

Полученная в результате предварительного сглаживания выборка является исходным материалом для анализа поведения исследуемой системы и для краткосрочного прогнозирования [1]. Прогнозирование будет корректным, если такая выборка отвечает следующим требованиям:

1. Закон распределения выборки прогнозов $\tilde{y}$ относительно тренда должен соответствовать закону распределения выборки равноточных измерений y^*.

2. Тренд выборки прогнозов $\tilde{y}$ должен быть тем же, что и тренд выборки измерений y^* и соответствовать теоретическому тренду $y(t)$.

3. Дисперсия полученной выборки должна быть не больше дисперсии исходной выборки.

Таким образом, необходимо доказать:

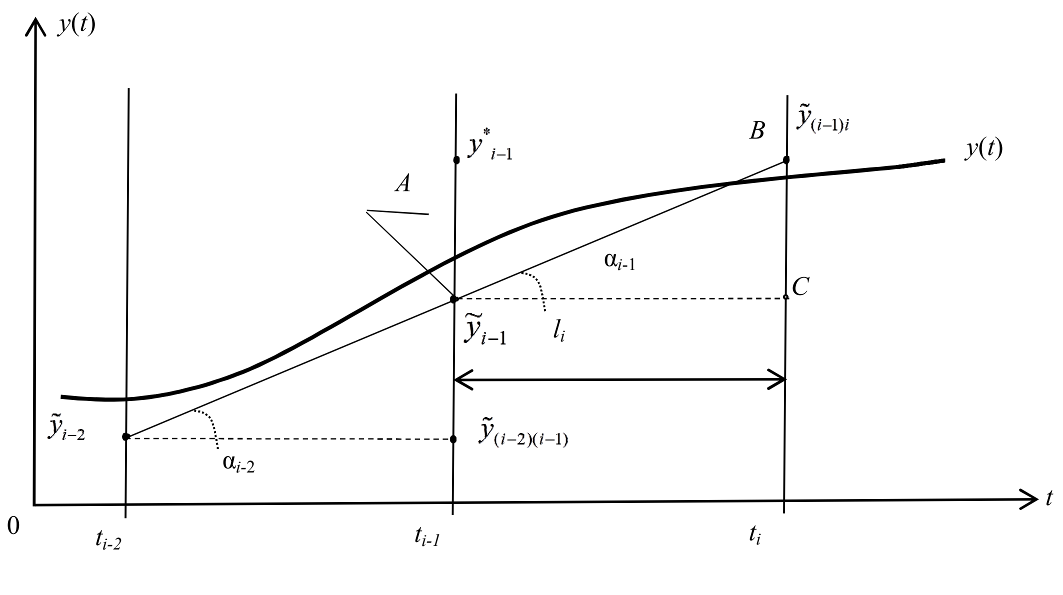
– что закон распределения $F(\tilde{y})$ идентичен закону распределения $F(y^*)$;

– равенство математических ожиданий: $M_{\tilde{y}}(t) = M_{y^*}(t) = y(t)$;

– соотношение дисперсий: $D_{y^*} \geq D_{\tilde{y}}$.

Закон распределения оценок. Пусть характер изменения наблюдаемого параметра СТС $y(t)$ представляет собой монотонную функцию [4, 5].

В выражении (1) не определено значение частного прогноза $\tilde{y}_{(i-1)i}$. Для его получения необходимо рассмотреть три следующих друг за другом сечения времени: t_{i-2}, t_{i-1}, t_i. Пусть для моментов времени t_{i-2}, t_{i-1} найдены соответствующие оценки прогноза $\tilde{y}_{i-2}$ и $\tilde{y}_{i-1}$. Отрезок линии прогноза, лежащий между точками $(t_{i-2}, \tilde{y}_{i-2})$ и $(t_{i-1}, \tilde{y}_{i-1})$, наклонен к оси абсцисс под углом (рис.).



Частный прогноз $\tilde{y}_{(i-1)i}$

Для монотонных функций вполне естественно предположить, что тенденция угловых характеристик линии прогноза функции $y(t)$ на временном интервале (t_{i-1}, t_{i-2}) сохранится и на интервале (t_i, t_{i-1}). При продолжении отрезка этой линии прогноза до сечения t_i образуется точка их пересечения B, ордината которой и будет значением частного прогноза $\tilde{y}_{(i-1)i}$ возможного значения $y(t_i)$.

Из рисунка видно, что

$$\tilde{y}_{(i-1)i} = \tilde{y}_{i-1} + l_i \mathrm{tg}\alpha_{i-1}, \tag{2}$$

где α_{i-1} – угол между осью абсцисс и продолжением линии прогноза до сечения времени t_i; l_i – проекция гипотенузы AB ΔABC на ось абсцисс.

Отсюда выражение (1) примет вид

$$\tilde{y}_i = \frac{y_i^* + \tilde{y}_{i-1} + l_i \mathrm{tg}\alpha_{i-1}}{2}. \tag{3}$$

Из (2) значение $\mathrm{tg}\alpha_{i-1}$ имеет вид $\mathrm{tg}\alpha_{i-1} = \dfrac{\tilde{y}_{i-1} - \tilde{y}_{i-2}}{l_i}$. Подстановка этого значения в (3) преобразует последнее следующим образом:

$$\tilde{y}_i = \frac{y_i^* + \tilde{y}_{i-1} + (\tilde{y}_{i-1} - \tilde{y}_{i-2})}{2}. \tag{4}$$

Таким образом, частный прогноз $\tilde{y}_{(i-1)i}$ определяется величинами оценки прогноза $\tilde{y}_{i-1}$ и измерением y_i^*. Это предполагает, что угловые характеристики линии прогноза на временном интервале (t_{i-2}, t_{-1i}) сохранятся неизменными до сечения t_i. Следовательно, через точки $(t_{i-2}, \tilde{y}_{i-2}), (t_{i-1}, \tilde{y}_{i-1}), (t_i, \tilde{y}_{(i-1)i})$ проводится одна прямая. В результате измерений и описанных расчетов для любого $i = k, k > 1$ возникает два временных ряда: ряд измерений $y^* = (y_0^*, y_1^*, ..., y_k^*)$ и ряд оценок прогнозов $\tilde{y} = (\tilde{y}_0, \tilde{y}_1, ..., \tilde{y}_k)$.

В момент времени $i = 0$ произведено лишь одно измерение, результатом которого является величина y_0^*. Поскольку до начального момента времени никакой предыстории не существовало, а кроме измерения никакой иной информации нет, то для $i = 0$ угол $\alpha = 0$, а величина формального частного прогноза принимается равной измерению $\tilde{y}_{(0-1)0} = y_0^*$. Поэтому из (3) можно сделать следующее допущение: $\tilde{y}_0 = y_0^*$.

Рассмотрим несколько последовательных моментов времени, начиная с $i = 0$ до $i = k$, и проведем в (4) необходимые преобразования. Получим общее выражение для любой прогнозной оценки $\tilde{y}_k$ при $k > 1$:

$$\tilde{y}_k = \frac{\tilde{y}_0 + \sum_{i=1}^{k} (2^{(i-1)} y_i^* + \sum_{i=1}^{k} \left([2^{(i-1)} (\tilde{y}_{i-1} - \tilde{y}_{i-2})]\right)}{2^k}. \tag{5}$$

Действительно, при $k = 0$ нулевой член равен $\tilde{y}_0 = y_0^*$, при $k = 1$ – значению $\tilde{y}_1 = \dfrac{y_1^* + 2\tilde{y}_0}{2}$ и т. д.

В выражении (5) все составляющие определены. Оно является удобным для программирования. Кроме того, из него не следует никаких ограничений на величину интервала $\Delta t_i = t_i - t_{i-1}$. В этом случае естественно предположить, что $\Delta t_1 = \Delta t_2 = ... = \Delta t_n$.

Метод пошагового сглаживания (5) предлагает вместо исходной выборки y_i^* сглаженную выборку $\tilde{y} = (\tilde{y}_0, \tilde{y}_1, ..., \tilde{y}_k)$ при условии, что трендом исходной выборки является монотонная функция. В случае немонотонного тренда пользоваться (5) нецелесообразно, т. к. этот метод будет «опаздывать» в отражении перегибов тренда. Действительно, на каждом очередном i-м наблюдении частный прогноз $\tilde{y}_{(i-1)i}$ поддерживает угловые тенденции прогноза $\tilde{y}_{i-1}$, полученного на предыдущем $(i-1)$-м наблюдении. Наличие такого ограничения можно отнести к недостаткам метода.

Перепишем выражение (5) в следующем виде:

$$\tilde{y}_k = \frac{1}{2^k}\tilde{y}_0 + \sum_{i=1}^{k}\frac{2^{(i-1)}}{2^k}y_i^* + \sum_{i=1}^{k}\frac{2^{(i-1)}}{2^k}(\tilde{y}_{(i-1)} - \tilde{y}_{(i-2)}),$$

и, произведя суммирование и группирование, получим

$$\tilde{y}_k = \sum_{i=1}^{k}a_i y_i^* - \sum_{i=1}^{k-1}b_i\tilde{y}_i, \tag{6}$$

где $a_i = \dfrac{2^{(i-1)}}{2^k}$ при $i = 0, 1, ..., k-1$; $b_i = \dfrac{2^i}{2^k}$ при $i = 0, 1, ..., k-2$.

При постановке задачи МПС определялось, что значения выборки y^* являются результатами равноточных измерений монотонной функции $y(t)$. Каждое измерение y_i^* представляет собой элемент случайной выборки со средним $y(t)$. Полученный в результате равноточных измерений, временной ряд $y_0^*, y_1^*, ..., y_k^*$ является рядом величин, которые выборочно характеризуют случайный процесс $y^*(t)$. В силу ранее сказанного $y^*(t)$ является некоррелированным случайным процессом с постоянной дисперсией и трендом в виде монотонной функции $y(t)$.

Из (6) видно, что для любого k оценка $\tilde{y}_k$ – это линейная комбинация k случайных величин $\tilde{y}_i$, полученных ранее. В каждом i-м сечении нормального процесса $y^*(t)$ случайные величины y_i^* распределены одинаково. Следовательно, для t_0 случайная величина $\tilde{y}_0$ распределена так же, т. к. $\tilde{y}_0 = y_0^*$. Далее, для $\tilde{y}_1$ имеем

$$\tilde{y}_1 = a_1 y_1^* - b_0\tilde{y}_0.$$

Но $\tilde{y}_0^*$, в силу равенства $\tilde{y}_0 = y_0^*$, распределена так же, как и y_0^*. Иначе говоря, $\tilde{y}_1$ – это линейная функция от двух одинаково распределенных случайных величин.

Аналогично можно показать, что такой же закон распределения имеет случайная величина $\tilde{y}_2$. Действительно, в выражении

$$\tilde{y}_2 = a_1 y_1^* + a_2 y_2^* - b_0\tilde{y}_0 - b_1\tilde{y}_1$$

все случайные аргументы распределены одинаково. Отсюда следует, что такое же распределение имеет случайная величина $\tilde{y}_2$.

Рассуждая аналогичным образом, можно заключить, что каждый член выборки оценок получается путем линейного преобразования соответствующего члена выборки измерений. Это означает, что закон распределения выборки оценок идентичен закону распределения выборки измерений.

Тренд оценок. Можно предположить, что математическое ожидание случайной величины $\tilde{y}_i$ смещено относительно математического ожидания случайной величины y_i^* на некоторое значение смещения ρ_i:

$$\rho_i = m_{\tilde{y}_i} - m_{y_i}, \tag{7}$$

где ρ_i – случайное число.

Рассмотрим некоторый ограниченный интервал времени от t_0 до t_n. При этом выберем такое количество наблюдений n, для которого, практически, можно считать, что $n \to \infty$. Тогда интервал между соседними наблюдениями будет $\Delta t = \dfrac{t_n - t_0}{n}$. В силу малости этих интервалов можно, без нарушения строгости рассуждений, допустить, что $\Delta t \to 0$.

Запишем (4), с учетом выражения (7), в следующем виде:

$$2(x_i + \rho_i + \tilde{\Delta}_i) = (x_i + \Delta_i) + 2(x_{i-1} + \rho_{i-1} + \tilde{\Delta}_{i-1}) - (x_{i-2} + \rho_{i-2} + \tilde{\Delta}_{i-2}), \qquad (8)$$

где $\widetilde{\Delta}_i$ – случайное отклонение $\widetilde{y}_i$ от своего математического ожидания $m_{\tilde{y}_i}$; Δ_i – отклонение величины y_i^* от математического ожидания $m_{y_i^*}$.

Поскольку при любом i величина y_i^* является измерением, на основании центральной предельной теоремы [6] и опыта исследования поведения СТС [7, 8] она распределена нормально со средним $m_{y_i^*} = y_i$. Такое же распределение имеет и отклонение Δ_i с нулевым средним. Относительно величин $\tilde{y}_i$ и $\tilde{\Delta}_i$ можно утверждать то же самое, т. к. любое $\tilde{y}_i$ можно представить в виде комбинации измерений, что следует из (1) и рис. Найдем математическое ожидание выражения (8):

$$2M\left[y_i + \rho_i + \tilde{\Delta}_i\right] = M\left[y_i + \Delta_i\right] + 2M\left[y_{x-1} + \rho_{i-1} + \tilde{\Delta}_{i-1}\right] - M\left[y_{i-2} + \rho_{i-2} + \tilde{\Delta}_{i-2}\right],$$

или

$$\rho_i - \rho_{i-1} + \frac{1}{2}\rho_{i-2} = -\frac{1}{2}y_i + y_{i-1} - \frac{1}{2}\rho_{i-2}. \qquad (9)$$

Так как функция $y(t)$ непрерывна на интервале от t_0 до t_n, предел правой части (9) при выбранных условиях $n \to \infty; \Delta t \to 0$ будет равен

$$\lim_{\Delta t \to 0}(-\frac{1}{2}y_i + y_{i-1} - \frac{1}{2}y_{i-2}) = \lim_{\Delta t \to 0}(-\frac{1}{2}y(t_i) + y(t_i - \Delta t) - \frac{1}{2}y(t_i - 2\Delta t)) = 0.$$

Тогда, окончательно, выражение (9) примет вид

$$\rho_i - \rho_{i-1} + \frac{1}{2}\rho_{i-2} = 0.$$

Это разностное уравнение. Запишем характеристическое уравнение $\lambda^2 - \lambda + \frac{1}{2}\lambda = 0$.

Решением этого уравнения будет $\rho_i = C_1(\sqrt{\frac{1}{2}})^i \cos(i\frac{\pi}{4}) + C_2(\sqrt{\frac{1}{2}})^i \sin(i\frac{\pi}{4})$.

Так как постоянные C_1 и C_2, определяемые начальными условиями $t_0 = 0;\ t_1 = \Delta t,$ конечны, то

$$\lim_{i \to \infty} \rho_i = 0.$$

Иначе говоря, смещение ρ_i с увеличением асимптотически стремится к нулю.

Приведенные рассуждения дают основание переписать выражение (7) в следующем виде:

$$\rho_i = m_{\tilde{y}_i} - m_{y_i^*} = 0,$$

или

$$m_{\tilde{y}_i} = m_{y_i^*}.$$

Иначе говоря, элементы $y_0, y_1, ..., y_n$ тренда $y(t)$ в предельном случае являются средними соответствующих случайных нормально распределенных величин. Тогда для случайной функции $\tilde{y}(t)$, равно как и для случайной функции $\tilde{y}(t)$, трендом служит неслучайная функция $y(t)$:

$$M_{\tilde{y}}(t) = M_{y^*}(t) = y(t).$$

Таким образом, тренд оценок идентичен тренду измерений.

Дисперсия оценок. Для определения дисперсии оценки $\widetilde{y}_i$ воспользуемся выражением (4). Его можно представить в виде

$$\tilde{y}_i = \frac{y_i^* + 2\tilde{y}_{i-1} - \tilde{x}_{i-2}}{2} = \tilde{y}_{i-1} - \frac{1}{2}\tilde{y}_{i-2} + \frac{1}{2}y_i^*. \tag{10}$$

Аналогично запишем выражение для $\tilde{y}_{i-1}$:

$$\tilde{y}_{i-1} = \tilde{y}_{i-2} - \frac{1}{2}x\tilde{y}_{i-3} + y_{i-1}^* .$$

Подставим это выражение в (10):

$$\tilde{y}_i = (\tilde{y}_{i-2} - \frac{1}{2}\tilde{y}_{i-3} + \frac{1}{2}y_{i-1}^*) - \frac{1}{2}\tilde{y}_{i-2} + \frac{1}{2}y_i^* = \frac{1}{2}\tilde{y}_{i-2} - \frac{1}{2}\tilde{y}_{i-3} + \frac{1}{2}y_i^* + \frac{1}{2}y_{i-1}^*$$

и далее, применяя метод последовательных итераций, окончательно:

$$\tilde{y}_i = \frac{1}{2}(\tilde{y}_{i-3} - \frac{1}{2}\tilde{y}_{i-4} + \frac{1}{2}{}^{*}_{i-2}) - \frac{1}{2}\tilde{y}_{i-3} + \frac{1}{2}y_i^* + \frac{1}{2}y_{i-1}^* = -\frac{1}{4}\tilde{y}_{i-4} + \frac{1}{2}y_i^* + \frac{1}{2}y_{i-1}^* + \frac{1}{4}y_{i-2}^*.$$

Используя ранее принятые обозначения, найдем дисперсию последнего выражения:

$$D\left[\tilde{y}_i\right] = D\left[-\frac{1}{2}\tilde{y}_{i-4} + \frac{1}{2}y_i^* + \frac{1}{2}y_{i-1}^* + \frac{1}{4}y_{i-2}^*\right] =$$

$$= D\left[-\frac{1}{4}\tilde{y}_{i-4} + \frac{1}{2}y_i + \frac{1}{2}\Delta_i + \frac{1}{2}y_{i-1} + \Delta_{i-1} + \frac{1}{4}y_{i-2} + \frac{1}{4}\Delta_{i-2}\right].$$

Величины ${}_i y, y_{i-1}, y_{i-2}$ неслучайны, поэтому их дисперсия равна нулю. Случайные величины $\Delta_i, \Delta_{i-1}, \Delta_{i-2}$ независимы и, кроме того, независимы пары случайных величин $\tilde{y}_{i-4}, \Delta_{i-2}$; $\tilde{y}_{i-4}, \Delta_{i-1}$, $\tilde{y}_{i-4}, \Delta_i$, т. к. они не принимают участие в формировании друг друга. Следовательно, для коваріаций можно записать:

$$K\left[\tilde{y}_{i-4}, \Delta_{i-2}\right] = K\left[\tilde{y}_{i-4}, \Delta_{i-1}\right] = K\left[\tilde{y}_{i-4}, \Delta_i\right] = 0 .$$

Отсюда

$$D\left[\tilde{y}_i\right] = \frac{1}{16}D\left[\tilde{y}_{i-4}\right] + \frac{1}{4}D\left[\Delta_i\right] + \frac{1}{4}D\left[\Delta_{i-1}\right] + \frac{1}{16}D\left[\Delta_{i-2}\right],$$

но

$$D\left[\Delta_i\right] = D\left[\Delta_{i-1}\right] = D\left[\Delta_{i-2}\right] = D[y^*],$$

тогда

$$D\left[\tilde{y}_i\right] = \frac{1}{16}D\left[\tilde{y}_{i-4}\right] + \frac{9}{16}D[y^*]. \tag{11}$$

Для проверки того, что $D[\tilde{y}] \le D[y^*]$, рассмотрим выражение (11) для $i = 4$:

$$D\left[\tilde{y}_4\right] = \frac{1}{16}D\left[\tilde{y}_0\right] + \frac{9}{16}D[y^*] .$$

Но в соответствии с допущением $\tilde{y}_0 = y_0^*$

$$D\left[\tilde{y}_0\right] = D\left[y_0^*\right] = D\left[y^*\right],$$

следовательно,

$$D\left[\tilde{y}_4\right] = \frac{1}{16}D\left[y^*\right] + \frac{9}{16}D[y^*] = \frac{5}{8}D[y^*] < D[y^*] .$$

Продолжая подобные преобразования для $i = 2, 3$, получим

$$D[\tilde{y}_3] = \frac{9}{16} D\left[y^*\right] < D[y^*]; \quad D[\tilde{y}_3] = \frac{3}{4} D\left[y^*\right] < D[y^*].$$

Аналогичные результаты дает и вычисление (11) для $i = 5$: $D[\tilde{y}_5] = \frac{41}{64} D\left[y^*\right] < D[y^*]$.

Отсюда можно сделать вывод, что для любых i условие $D[\tilde{y}] \le D[y^*]$ выполняется. Таким образом, для всех измерений, кроме нулевого и первого, дисперсия оценок меньше дисперсии измерений. Несоответствие порядка величин дисперсий нулевого и первого измерений порядку остальных можно объяснить принятыми допущениями. Следовательно, исходные предпосылки относительно дисперсии оценок, полученных МПС, верны.

Кроме того, из (11) и приведенных ранее значений дисперсии следует, что начиная с $i = 2$ дисперсия будет стремиться к некоторому пределу, т. к. как первое слагаемое в правой части выражения с увеличением i будет меньше ощущать влияние допущений.

Заключение

На основании проведенных в настоящей работе доказательств можно утверждать, что все поставленные задачи выполнены, а именно:

– закон распределения прогнозов, определяемых с помощью метода пошагового сглаживания, $F(\tilde{y})$ идентичен закону распределения измерений $F(y^*)$;

– функции математических ожиданий прогнозов и измерений равны между собой и равны теоретической монотонной функции $y(t)$;

– дисперсия полученных прогнозов не превышает дисперсию измерений.

Следовательно, метод пошагового сглаживания корректен и может быть предложен для краткосрочного прогнозирования.

СПИСОК ЛИТЕРАТУРЫ

1. *Афанасьев В. Н., Юзбашев Н. Н.* Анализ временных рядов и прогнозирование. М.: Финансы и статистика, 2001. 288 с.
2. *Рабочая* книга по прогнозированию / под ред. И. В. Бестужева-Лады. М.: Мысль, 1982. 430 с.
3. *Ермаков А. А., Михайлов Д. И.* Основы пошагового сглаживания при обработке статистических данных // Современные технологии, системный анализ, моделирование. Иркутск: ИрГУПС, 2014. № 2 (42). С. 97–103.
4. *Бородачев Н. А.* Основы теории точности производства. М.-Л.: Изд-во АН СССР, 1950. 256 с.
5. *Новицкий П. В., Зограф И. А.* Оценка погрешностей результатов измерений. Л.: Энергоатомиздат, 1985. 247 с.
6. *Вентцель Е. С.* Теория вероятностей. М.: Высш. шк., 2006. 576 с.
7. *Дедков В. К., Северцев Н. А.* Основные вопросы эксплуатации сложных систем. М.: Высш. шк., 1976. 406 с.
8. *Калашников В. В.* Сложные системы и методы их анализа. М.: Знание, 1980. 62 с.

Статья поступила в редакцию 19.05.2021

ИНФОРМАЦИЯ ОБ АВТОРАХ

Анатолий Анатольевич Ермаков – канд. техн. наук, доцент; профессор кафедры информационных систем и защиты информации; Иркутский государственный университет путей сообщения; Россия, 664074, Иркутск; ermak@irgups.ru.

Татьяна Климентьевна Кириллова – канд. экон. наук, доцент; зав. кафедрой информационных систем и защиты информации; Иркутский государственный университет путей сообщения; Россия, 664074, Иркутск; kirillova_tk@irgups.ru.

METHOD OF STEPWISE SMOOTHING EXPERIMENTAL DEPENDENCES FOR PROBLEMS OF SHORT-TERM FORECASTING

A. A. Ermakov, T. K. Kirillova

Irkutsk State Transport University,
Irkutsk, Russian Federation

Abstract. The article considers the correspondence of the step-by-step smoothing method as one of the possible algorithms for short-term forecasting of statistics of equal-current measurements of monotone functions, which represent the values of the determining parameters that evaluate the dynamics of the states of complex technical systems based on the operating time. The true value of the monitored parameter is considered unknown, and the processed measurement values are distributed normally. The measurements are processed by step-by-step smoothing. As a result of processing, a new statistic is formed, which is a forecast statistic, each value of which is a half-sum of the measurement itself and the so-called private forecast. First, the forecasts obtained in this way prove to have the same distribution law as the distribution law of a sample of equally accurate measurements. Second, the forecast trend should be the same as the measurement trend and correspond to the theoretical trend, that is, the true values of the monotone function. Third, the variance of the obtained statistics should not exceed the variance of the original sample. It is inferred that the method of step-by-step smoothing method can be proposed for short-term forecasting.

Key words: monotonic function, measurements, forecast, private forecast, trend, variance, stepwise smoothing method, time interval, measurement estimation, distribution law.

For citation: Ermakov A. A., Kirillova T. K. Method of stepwise smoothing experimental dependences for problems of short-term forecasting. *Vestnik of Astrakhan State Technical University. Series: Management, Computer Science and Informatics*. 2021;3:126-133. (In Russ.) DOI: 10.24143/2072-9502-2021-3-126-133.

REFERENCES

1. Afanas'ev V. N., Iuzbashev N. N. *Analiz vremennykh riadov i prognozirovanie* [Analysis of time series and forecasting]. Moscow, Finansy i statistika Publ., 2001. 288 p.
2. *Rabochaia kniga po prognozirovaniiu* [Workbook on forecasting]. Pod redaktsiei I. V. Bestuzheva-Lady. Moscow, Mysl' Publ., 1982. 430 p.
3. Ermakov A. A., Mikhailov D. I. Osnovy poshagovogo sglazhivaniia pri obrabotke statisticheskikh dannykh [Fundamentals of step-by-step smoothing in statistical data processing]. *Sovremennye tekhnologii, sistemnyi analiz, modelirovanie*. Irkutsk, IrGUPS, 2014. N. 2 (42). Pp. 97-103.
4. Borodachev N. A. *Osnovy teorii tochnosti proizvodstva* [Principles of theory of production accuracy]. Moscow, Leningrad, Izd-vo AN SSSR, 1950. 256 p.
5. Novitskii P. V., Zograf I. A. *Otsenka pogreshnostei rezul'tatov izmerenii* [Assessing errors of measurement results]. Leningrad, Energoatomizdat, 1985. 247 p.
6. Venttsel' E. S. *Teoriia veroiatnostei* [Probability theory]. Moscow, Vysshaia shkola Publ., 2006. 576 p.
7. Dedkov V. K., Severtsev N. A. *Osnovnye voprosy ekspluatatsii slozhnykh sistem* [Basic questions of operation of complex systems]. Moscow, Vysshaia shkola Publ., 1976. 406 p.
8. Kalashnikov V. V. *Slozhnye sistemy i metody ikh analiza* [Complex systems and methods of their analysis]. Moscow, Znanie Publ., 1980. 62 p.

The article submitted to the editors 19.05.2021

INFORMATION ABOUT THE AUTHORS

Anatoly A. Ermakov – Candidate of Technical Sciences, Assistant Professor; Professor of the Department of Information Systems and Information Security; Irkutsk State Transport University; Russia, 664074, Irkutsk; ermak@irgups.ru.

Tatyana K. Kirillova – Candidate of Economic Sciences, Assistant Professor; Head of the Department of Information Systems and Information Security; Irkutsk State Transport University; Russia, 664074, Irkutsk; kirillova_tk@irgups.ru.