УДК 519.21

ОЦЕНИВАНИЕ МАТЕМАТИЧЕСКОГО ОЖИДАНИЯ И КОВАРИАЦИОННОЙ ФУНКЦИИ СТАЦИОНАРНОЙ СЛУЧАЙНОЙ ПОСЛЕДОВАТЕЛЬНОСТИ УСРЕДНЕНИЕМ ПО ВРЕМЕНИ И МНОЖЕСТВУ РЕАЛИЗАЦИЙ

В.С. МУХА, А.Ф. ТРОФИМОВИЧ

Белорусский государственный университет информатики и радиоэлектроники
П. Бровки,6, Минск, 220013, Беларусь

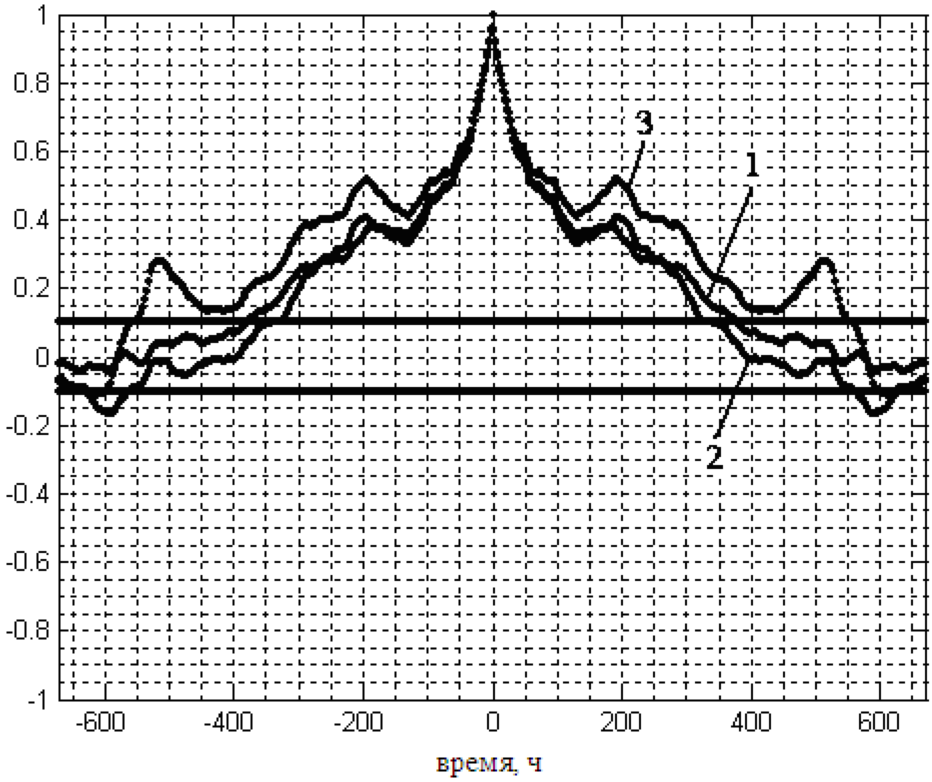
Поступила в редакцию 7 октября 2008

Предложены алгоритмы оценивания математического ожидания и ковариационной функции многомерно-матричной стационарной случайной последовательности путем усреднения как по времени, так и по множеству реализаций. Доказаны теоремы о несмещенности и состоятельности предложенных оценок.

Ключевые слова: стационарная случайная последовательность, оценка ковариационной функции, усреднение по времени и множеству реализаций.

Введение

В литературе известны алгоритмы оценивания математического ожидания и ковариационной функции скалярной стационарной случайной последовательности по одной достаточно длинной реализации [1]. Вместе с тем на практике встречаются случаи, когда вместо одной длинной реализации последовательности имеется ряд более коротких реализаций, полученных в различные промежутки времени и по ним требуется получить оценки математического ожидания и ковариационной функции. Можно, конечно, соединить ("сшить") отдельные реализации в одну и применить известные алгоритмы оценивания. Однако поскольку отдельная реализация не является продолжением какой-то другой реализации, то в местах "сшивания" реализаций мы получим разрывы первого рода, приводящие к ошибкам в оценках. Этот эффект отчетливо обнаруживается при расчете оценок ковариационных функций метеорологических процессов для отдельных периодов года. В течение одного месяца метеорологические процессы (температура воздуха, атмосферное давление и др.) можно предположить стационарными. Если для получения оценок использовать реализации за несколько лет, то они будут отделены одна от другой не менее чем на год и не будут естественным продолжением одна другой. На рисунке приведены оценки корреляционной (нормированной ковариационной) функции температуры воздуха в декабре по метеостанции Минск по семи реализациям температуры (за 7 лет) при неизвестном математическом ожидании по традиционному алгоритму [1]. Каждая реализация содержит 248 отсчетов. Кривая *1* соответствует случаю, когда месячные реализации сшивались в хронологическом порядке, кривая *2* — когда реализации "сшивались" в порядке, обратном хронологическому. Видно, что эти кривые совпадают в достаточно узкой области небольших значений времени и заметно отличаются друг от друга в области больших времен. Таким образом, различный порядок "сшивания" реализаций приводит к существенному различию оценок, что нельзя считать приемлемым. Возникает необходимость поиска других алгоритмов оценивания, адекватных имеющимся исходным данным. Случай оценивания по множеству реализаций скалярного стационарного случайного процесса рассматривался в работе [2]. Однако предложенные в [2] оценки относятся к процессу с непрерывным временем и не пригодны для использования цифровой (компьютерной) обработки. В данной работе предлагаются оценки для случайной последовательности, причем рассматривается случай многомерно-матричной случайной последовательности.



Оценки корреляционной функции для различных алгоритмов оценивания

Постановка задачи, выражения для оценок

Пусть имеется реализация $x(iT)$ p-мерно-матричной случайной последовательности $\gamma(iT)$, $i=\overline{1,n}$, T — период квантования по времени, ν_γ — математическое ожидание, $\mu_{\gamma\gamma}(kT)$, $k=0,\pm1,\pm2,\ldots$, — ковариационная функция этой последовательности, и требуется получить оценки указанных моментных функций этой последовательности. В качестве оценок в этом случае можно использовать традиционные статистики [1]

$$\hat{\nu}_\gamma = \frac{1}{n}\sum_{i=1}^{n} x(iT),$$

$$\hat{\mu}_{\gamma\gamma}(kT) = \frac{1}{n-k}\sum_{i=1}^{n-k} \overset{\circ}{x}(iT)\,\overset{\circ}{x}(iT+kT), \tag{1}$$

$$\bar{\mu}_{\gamma\gamma}(kT) = \frac{1}{n-k}\sum_{i=1}^{n-k} \overset{*}{x}(iT)\,\overset{*}{x}(iT+kT),\ k=\overline{0,n-1}, \tag{2}$$

где $\overset{\circ}{x}(iT)=x(iT)-\nu_\gamma$, $\overset{*}{x}(iT)=x(iT)-\hat{\nu}_\gamma$. Оценка $\hat{\mu}_{\gamma\gamma}(kT)$ используется, когда математическое ожидание ν_γ случайной последовательности известно, а оценка $\bar{\mu}_{\gamma\gamma}(kT)$ — когда оно неизвестно. Как отмечено выше, эти оценки не адекватны исходным данным в виде множества реализаций последовательности.

Пусть теперь имеется m независимых реализаций $x_s(iT)$ различной длины n_s, $s=\overline{1,m}$, p-мерно-матричной случайной последовательности $\gamma(iT)$, $i=1,2,\ldots$, и требуется по ним получить оценки математического ожидания ν_γ и ковариационной функции $\mu_{\gamma\gamma}(kT)$ этой последовательности. В этом случае целесообразно сначала получить оценки по отдельным реализациям, а затем усреднить их по всем реализациям. Для отдельной s-й реализации можно использовать традиционные оценки математического ожидания и ковариационной функции:

$$\hat{\nu}_\gamma = \frac{1}{n_s}\sum_{i=1}^{n_s} x_s(iT),$$

$$\hat{\mu}_{\gamma\gamma}(kT) = \frac{1}{n_s - k}\sum_{i=1}^{n_s-k} \overset{\circ}{x}_s(iT)\overset{\circ}{x}_s(iT+kT),$$

$$\bar{\mu}_{\gamma\gamma}(kT) = \frac{1}{n_s - k}\sum_{i=1}^{n_s-k} \overset{*}{x}_s(iT)\overset{*}{x}_s(iT+kT),\ k = \overline{0, n_s - 1},$$

При усреднении записанных оценок по множеству реализаций целесообразно их использовать с весами, пропорциональными количеству наблюдений, по которым они получены, т.е. с весами

$$\frac{n_s}{\sum_{i=1}^{m} n_i},\ \frac{n_s - k}{\sum_{i=1}^{m}(n_i - k)}$$

соответственно для математического ожидания и ковариационной функции. В результате получим следующие оценки:

$$\hat{\nu}_\gamma = \frac{1}{n}\sum_{s=1}^{m}\sum_{i=1}^{n_s} x_s(iT),\ n = \sum_{i=1}^{m} n_i, \tag{3}$$

$$\hat{\mu}_{\gamma\gamma}(kT) = \frac{1}{N}\sum_{s=1}^{m}\sum_{i=1}^{n_s-k} \overset{\circ}{x}_s(iT)\overset{\circ}{x}_s(iT+kT), \tag{4}$$

$$\bar{\mu}_{\gamma\gamma}(kT) = \frac{1}{N}\sum_{s=1}^{m}\sum_{i=1}^{n_s-k} \overset{*}{x}_s(iT)\overset{*}{x}_s(iT+kT),\ k < n_s,\ N = \sum_{i=1}^{m}(n_i - k). \tag{5}$$

Оценка $\hat{\mu}_{\gamma\gamma}(kT)$ (4) предназначена для случая, когда математическое ожидание случайной последовательности известно, а оценка $\bar{\mu}_{\gamma\gamma}(kT)$ (5) — для случая, когда оно неизвестно.

Среднее значение и дисперсия оценки математического ожидания

Исследуем свойства предложенных оценок (3)–(5), для чего найдем их математические ожидания и дисперсии (дисперсионные матрицы). Для среднего значения оценки $\hat{\nu}_\gamma$ (3) получим

$$E(\hat{\nu}_\gamma) = \frac{1}{n}\sum_{s=1}^{m}\sum_{i=1}^{n_s} E(x_s(iT)) = \frac{1}{n}\sum_{s=1}^{m}\sum_{i=1}^{n_s} \nu_\gamma = \frac{1}{n}\sum_{s=1}^{m} n_s \nu_\gamma = \nu_\gamma. \tag{6}$$

Это означает несмещенность оценки $\hat{\nu}_\gamma$.

Для дисперсии оценки $\hat{\nu}_\gamma$ будем иметь

$$D(\hat{\nu}_\gamma) = E(\hat{\nu}_\gamma - E(\hat{\nu}_\gamma)) = E\left(\left(\frac{1}{n}\sum_{s=1}^{m}\sum_{i=1}^{n_s} \overset{\circ}{x}_s(iT)\right)^2\right) = E\left(\left(\frac{1}{n}\sum_{s=1}^{m} z_s\right)^2\right),$$

где $z_s = \sum_{i=1}^{n_s} \overset{\circ}{x}_s(iT)$. Заметим, что в силу независимости реализаций $x_s(iT)$ случайные величины z_s также независимы по s. В таком случае получим

$$D(\hat{\nu}_\gamma) = \frac{1}{n^2}\sum_{s=1}^{m}\sum_{r=1}^{m} E(z_s z_r) = \frac{1}{n^2}\sum_{s=1}^{m} E(z_s^2) = \frac{1}{n^2}\sum_{s=1}^{m} E\left(\sum_{i=1}^{n_s}\sum_{j=1}^{n_s} \overset{\circ}{x}_s(iT)\overset{\circ}{x}_s(jT)\right) =$$

$$= \frac{1}{n^2}\sum_{s=1}^{m}\sum_{i=1}^{n_s}\sum_{j=1}^{n_s} \mu_{\gamma\gamma}((i-j)T) =$$

$$= \frac{1}{n^2}\sum_{s=1}^{m} \left[n_s\mu_{\gamma\gamma}(0) + 2(n_s-1)\mu_{\gamma\gamma}(T) + 2(n_s-2)\mu_{\gamma\gamma}(2T) + \ldots + 2\mu_{\gamma\gamma}((n_s-1)T) \right] =$$

$$= \frac{1}{n^2}\sum_{s=1}^{m} 2n_s \left(\frac{\mu_{\gamma\gamma}(0)}{2} + (1-\frac{1}{n_s})\mu_{\gamma\gamma}(T) + (1-\frac{2}{n_s})\mu_{\gamma\gamma}(2T) + \ldots + \frac{1}{n_s}\mu_{\gamma\gamma}((n_s-1)T) \right) =$$

$$= \frac{1}{n^2}\sum_{s=1}^{m} n_s \sum_{\lambda=-(n_s-1)}^{(n_s-1)} (1-\frac{|\lambda|}{n_s})\mu_{\gamma\gamma}(\lambda T) = \frac{1}{n^2}\sum_{s=1}^{m} n_s w_s \,, \qquad (7)$$

где

$$w_s = \sum_{\lambda=-(n_s-1)}^{(n_s-1)} (1-\frac{|\lambda|}{n_s})\mu_{\gamma\gamma}(\lambda T),$$

а также

$$tr\, D(\hat{\nu}_\gamma) = \frac{1}{n^2}\sum_{s=1}^{m} n_s tr\, w_s \,,$$

где tr означает след матрицы. Поскольку

$$\left|\sum_{s=1}^{m} n_s tr\, w_s\right| \le \sum_{s=1}^{m} n_s \left|tr\, w_s\right| \le \sum_{s=1}^{m} n_s \left|tr\, w_{\max}\right| = n\left|tr\, w_{\max}\right|,$$

где $\left|tr\, w_{\max}\right|$ — максимальное из чисел $\left|tr\, w_1\right|, \left|tr\, w_2\right|, \ldots, \left|tr\, w_m\right|$, то

$$tr\, D(\hat{\nu}_\gamma) \le \frac{1}{n}\left|tr\, w_{\max}\right|.$$

Если выполняется условие

$$\sum_{\lambda=-\infty}^{\infty} tr\, \mu_{\gamma\gamma}(\lambda T) < \infty, \qquad (8)$$

то $\left|tr\, w_{\max}\right| < \infty$, и мы получим, что

$$\lim_{n\to\infty} tr\, D(\hat{\nu}_\gamma) = 0. \qquad (9)$$

Это означает состоятельность оценки $\hat{\nu}_\gamma$ (см. работу [3], результаты которой, относящиеся к векторным случайным последовательностям, обобщаются на многомерно-матричные случайные последовательности).

Таким образом, доказана следующая теорема.

Теорема 1. *Пусть* ν_γ *— математическое ожидание,* $\mu_{\gamma\gamma}(kT)$, $k = 0, \pm 1, \pm 2, \ldots$, *— ковариационная функция* p*-мерно-матричной последовательности* $\gamma(iT)$, $i = 1, 2, \ldots$, *и* $x_s(iT)$, $s = \overline{1, m}$, *—* m *реализаций различной длины* $n_1, n_2, \ldots, n_m$ *этой последовательности. Тогда математическое ожидание и дисперсионная матрица оценки* (3) *определяются выражениями* (6) *и* (7). *Если выполняется условие* (8)*, то имеет место сходимость* (9).

Таким образом, оценка вида (3) математического ожидания стационарной случайной последовательности несмещенная и при условии (8) состоятельная. Важно отметить, что состоятельность рассматриваемой оценки можно обеспечить увеличением, как числа реализаций m, так и длин отдельных реализаций $n_1, n_2, \ldots, n_m$.

Среднее значение и дисперсия оценки ковариационной функции

Для среднего значения оценки $\hat{\mu}_{\gamma\gamma}(kT)$ (4) получим

$$E(\hat{\mu}_{\gamma\gamma}(kT)) = \frac{1}{N} \sum_{s=1}^{m} \sum_{i=1}^{n_s - k} E\left(\overset{\circ}{x}_s(iT) \overset{\circ}{x}_s(iT + kT) \right) = \frac{1}{N} \sum_{s=1}^{m} \sum_{i=1}^{n_s - k} \mu_{\gamma\gamma}(kT) = \mu_{\gamma\gamma}(kT). \tag{10}$$

Это означает, что оценка $\hat{\mu}_{\gamma\gamma}(kT)$ несмещенная.

Найдем теперь дисперсию этой оценки:

$$D(\hat{\mu}_{\gamma\gamma}(kT)) = E\ (\hat{\mu}_{\gamma\gamma}(kT) - \mu_{\gamma\gamma}(kT))^2\ =$$

$$= E\left(\left(\frac{1}{N} \sum_{s=1}^{m} \sum_{i=1}^{n_s - k} \overset{\circ}{x}_s(iT) \overset{\circ}{x}_s(iT + kT) - \mu_{\gamma\gamma}(kT) \right)^2 \right) =$$

$$= \frac{1}{N^2} E\left(\left(\sum_{s=1}^{m} y_s(kT) \right)^2 \right),$$

где

$$y_s(kT) = \sum_{i=1}^{n_s - k} \overset{\circ}{x}_s(iT) \overset{\circ}{x}_s(iT + kT) - \mu_{\gamma\gamma}(kT).$$

В силу независимости по s величин $y_s(kT)$ получим

$$D(\hat{\mu}_{\gamma\gamma}(kT)) = \frac{1}{N^2} \sum_{s=1}^{m} \sum_{r=1}^{m} E\ y_s(kT) y_r(kT)\ = \frac{1}{N^2} \sum_{s=1}^{m} E\ y_s^2(kT)\ .$$

Найдем теперь $E\{y_s^2(kT)\}$:

$$E\ y_s^2(kT)\ =$$

$$= \sum_{i=1}^{n_s - k} \sum_{j=1}^{n_s - k} E\left((\overset{\circ}{x}_s(iT) \overset{\circ}{x}_s(iT + kT) - \mu_{\gamma\gamma}(kT))(\overset{\circ}{x}_s(jT) \overset{\circ}{x}_s(jT + kT) - \mu_{\gamma\gamma}(kT)) \right) =$$

$$= \sum_{i=1}^{n_s - k} \sum_{j=1}^{n_s - k} E\left(\overset{\circ}{x}_s(iT) \overset{\circ}{x}_s(iT + kT) \overset{\circ}{x}_s(jT) \overset{\circ}{x}_s(jT + kT) \right) + \sum_{i=1}^{n_s - k} \sum_{j=1}^{n_s - k} E\ \mu_{\gamma\gamma}^2(kT)\ -$$

$$- \sum_{i=1}^{n_s - k} \sum_{j=1}^{n_s - k} E\left(\overset{\circ}{x}_s(iT) \overset{\circ}{x}_s(iT + kT) \mu_{\gamma\gamma}(kT) \right) -$$

$$-\sum_{i=1}^{n_s-k}\sum_{j=1}^{n_s-k} E\left(\mu_{\gamma\gamma}(kT)\overset{\circ}{x}_s(jT)\overset{\circ}{x}_s(jT+kT)\right)=$$

$$=\sum_{i=1}^{n_s-k}\sum_{j=1}^{n_s-k}\mu_{\gamma\gamma\gamma\gamma}(kT,(j-i)T,(j+k-i)T)-\sum_{i=1}^{n_s-k}\sum_{j=1}^{n_s-k}\mu_{\gamma\gamma}^2(kT)=$$

$$=\sum_{\lambda=-(n_s-k-1)}^{n_s-k-1}(n_s-k)\left(1-\frac{|\lambda|}{n_s-k}\right)\ \mu_{\gamma\gamma\gamma\gamma}(kT,\lambda T,(k+\lambda)T)-\mu_{\gamma\gamma}^2(kT)\ ,$$

где $\mu_{\gamma\gamma\gamma\gamma}(k_1T,k_2T,k_3T)=E\left(\overset{\circ}{\gamma}(iT)\overset{\circ}{\gamma}(iT+k_1T)\overset{\circ}{\gamma}(iT+k_2T)\overset{\circ}{\gamma}(iT+k_3T)\right)$ — центральный момент четвертого порядка. Вводя в рассмотрение семиинвариант четвертого порядка, допускающий представление в виде [1]

$$\kappa_{\gamma\gamma\gamma\gamma}(k_1T,k_2T,k_3T)=\mu_{\gamma\gamma\gamma\gamma}(k_1T,k_2T,k_3T)-\mu_{\gamma\gamma}(k_1T)\mu_{\gamma\gamma}((k_2-k_3)T)-$$

$$-\mu_{\gamma\gamma}(k_2T)\mu_{\gamma\gamma}((k_1-k_3)T)-\mu_{\gamma\gamma}(k_3T)\mu_{\gamma\gamma}((k_1-k_2)T),$$

получим

$$E\ \ y_s^2(kT)\ \ =\sum_{\lambda=-(n_s-k-1)}^{n_s-k-1}(n_s-k)\left(1-\frac{|\lambda|}{n_s-k}\right)\times$$

$$\times\ \kappa_{\gamma\gamma\gamma\gamma}(kT,\lambda T,(k+\lambda)T)+\mu_{\gamma\gamma}^2(\lambda T)+\mu_{\gamma\gamma}((k+\lambda)T)\mu_{\gamma\gamma}((k-\lambda)T)\ \ .$$

В результате для $D(\widehat{\mu}_{\gamma\gamma}(kT))$ будем иметь выражение

$$D(\widehat{\mu}_{\gamma\gamma}(kT))=\frac{1}{N^2}\sum_{s=1}^{m}(w_{1,s}+w_{2,s}+w_{3,s}),\ N=\sum_{i=1}^{m}(n_i-k), \tag{11}$$

где

$$w_{1,s}=\sum_{\lambda=-(n_s-k-1)}^{n_s-k-1}(n_s-k)\left(1-\frac{|\lambda|}{n_s-k}\right)\kappa_{\gamma\gamma\gamma\gamma}(kT,\lambda T,(k+\lambda)T),$$

$$w_{2,s}=\sum_{\lambda=-(n_s-k-1)}^{n_s-k-1}(n_s-k)\left(1-\frac{|\lambda|}{n_s-k}\right)\mu_{\gamma\gamma}^2(\lambda T),$$

$$w_{3,s}=\sum_{\lambda=-(n_s-k-1)}^{n_s-k-1}(n_s-k)\left(1-\frac{|\lambda|}{n_s-k}\right)\mu_{\gamma\gamma}((k+\lambda)T)\mu_{\gamma\gamma}((k-\lambda)T).$$

Аналогично тому, как это было сделано в предыдущем разделе, можно показать, что

$$tr\,D(\widehat{\mu}_{\gamma\gamma}(kT))\le\frac{1}{N}(|tr\,w_{1,\max}|+|tr\,w_{2,\max}|+|tr\,w_{3,\max}|),$$

где $\left|tr\,w_{i,\max}\right|$ — максимальное из чисел $\left|tr\,w_{i,1}\right|,\left|tr\,w_{i,2}\right|,...,\left|tr\,w_{i,m}\right|$, $i=\overline{1,3}$. Если

$$\sum_{\lambda=-\infty}^{\infty}tr\,\kappa_{\gamma\gamma\gamma\gamma}(kT,\lambda T,(k+\lambda)T)<\infty, \tag{12}$$

$$\sum_{\lambda-\infty}^{\infty}\mu_{\gamma\gamma}^{2}(\lambda T)<\infty, \quad (13)$$

то $\left|tr\, w_{i,\max}\right|<\infty$, $i=\overline{1,3}$, и мы получим

$$\lim_{N\to\infty} tr\, D(\hat{\mu}_{\gamma\gamma}(kT))=0. \quad (14)$$

Последнее выражение означает состоятельность оценки $\hat{\mu}_{\gamma\gamma}(kT)$.

Также доказана следующая теорема.

Теорема 2. *Пусть* ν_{γ} *— математическое ожидание,* $\mu_{\gamma\gamma}(kT)$, $k=0,\pm1,\pm2,\dots$, *— ковариационная функция* p*-мерно-матричной последовательности* $\gamma(iT)$, $i=1,2,\dots$, *и* $x_s(iT)$, $s=\overline{1,m}$, *—* m *реализаций различной длины* $n_1,n_2,\dots,n_m$ *этой последовательности. Тогда математическое ожидание и дисперсионная матрица оценки* (4) *определяются выражениями* (10), (11). *Если выполняются соотношения* (12), (13), *то имеет место сходимость* (14).

Таким образом, оценка вида (4) ковариационной функции стационарной случайной последовательности $\gamma(iT)$ несмещенная и при условиях (12), (13) — состоятельная.

Анализ свойств оценки (5) ковариационной функции при неизвестном математическом ожидании сложнее рассмотренного анализа оценки (4). Однако, следуя соотношению между свойствами оценок (1), (2) традиционного алгоритма оценивания, имеющемуся в [1], можно предположить, что оценка $\bar{\mu}_{\gamma\gamma}(kT)$ (5) будет асимптотически несмещенной и состоятельной.

По предложенному алгоритму (5) была рассчитана оценка корреляционной функции температуры воздуха в декабре по метеостанции Минск по данным, описанным во введении (кривая *3* на рисунке). Данная оценка не изменяется при изменении порядка следования отдельных реализаций, что является подтверждением ее адекватности исходным данным.

ESTIMATION OF EXPECTATION AND COVARIANCE FUNCTION OF STATIONARY CASUAL SEQUENCE BY TIME AND SET OF IMPLEMENTATIONS AVERAGING

V.S. MUKHA, A.F. TROFIMOVICH

Abstract

Algorithms of estimation of expectation and covariance function of multidimensional-matrix stationary casual sequence are offered by time and set of implementations averaging. Theorems of unbiasedness and competence of the estimations are proved.

Литература

1. *Андерсон Т.* Статистический анализ временных рядов. М., 1976.
2. *Лившиц Н.А., Пугачев В.Н.* Вероятностный анализ систем автоматического управления. М., 1963. Ч. 1.
3. *Муха В.С.* Электромагнитные волны и электронные системы. 2008. № 4. С. 4–8.