УДК 519.2:53.05

Ф.Ф. Идрисов

ПРИБЛИЖЕННЫЕ АЛГОРИТМЫ ВЫДЕЛЕНИЯ ТРЕНДОВ В ЗАДАЧАХ ФИНАНСОВОЙ РАЗВЕДКИ. Часть 1. МОМЕНТЫ ПОЯВЛЕНИЯ ЭЛЕМЕНТОВ ФИНАНСОВОГО ПОТОКА ИЗВЕСТНЫ ТОЧНО

Введение

Во многих областях науки, техники, экономики и медицины наблюдаемые процессы нередко представляют собой временные ряды. Чаще всего наблюдения осуществляются через равные промежутки времени. Такие ряды принято называть эквидистантными. Их анализу посвящена обширная литература, например ставшие уже классическими монографии [1–3].

Однако на практике встречаются ситуации, когда моменты измерений, порождающие случайный процесс, являются случайными. Особенно часто это наблюдается в торговле и управлении запасами. Что касается банковских структур и страховых компаний, такие ситуации вполне типичны (например, появление вкладчиков в банке или наступление страховых случаев, связанных с ДТП, для страховых компаний). Все это приводит к необходимости разработки теории и алгоритмов анализа временных рядов при наблюдениях в случайные моменты времени (далее в целях удобства назовем их рандомизированными рядами).

Еще более актуальной такая проблема становится при анализе маскируемых финансовых потоков «грязных денег», осуществляемом органами финансовой разведки с использованием облачных технологий инструментальных средств Big Data.

В данной работе (идейно стыкующейся с [4]) рассматриваются следующие ситуации:

а) моменты появления элементов финансового потока известны точно;

б) моменты появления элементов финансового потока не известны.

Постановка задачи

Предположим, наблюдается рандомизированный ряд элементов финансового потока $x(t)$ на интервале [0, T], его значения наблюдаются в случайные моменты времени t_i. Будем считать, что на данном интервале измеренные значения элементов финансового потока $x_i = x(t_i)$ могут быть представлены в виде

$$x_i = x(t_i) = \sum_{k=1}^{s} \theta_k \varphi_k(t_i) + n_i, \tag{1}$$

где $\varphi_k(t)$ — постулируемые функции времени, θ_k — неизвестные коэффициенты (параметры) , а n_i — случайные добавки.

В дальнейшем будем считать, что $\varphi_k(t)$ — непрерывные ограниченные функции, так что все интегралы типа $\int_0^T \varphi_k^m(t)\varphi_l^n(t)dt$ существуют и конечны. Будем также считать, что эти функции имеют производные тех порядков, которые необходимы для вывода требуемых соотношений.

Что касается случайных добавок n_i, то будем считать, что они независимые одинаково распределенные случайные величины, у которых существуют начальные моменты всех порядков. Главное ограничение, которое используется ниже, будет заключаться в том, что все моменты нечетных порядков равны нулю, т.е. $\forall m\ M\{n_i^{2m+1}\} = 0$. Моменты четных порядков будем считать конечными, т.е. $M\{n_i^{2m}\} = \mu_{2m}$ и $M\{n_i^2\}$ обозначим как σ^2 (дисперсия n_i).

Что касается моментов измерений t_i, предполагается, что они образуют пуассоновский поток постоянной интенсивности λ.

Необходимо исследовать следующие варианты:

а) моменты измерений t_i известны точно;

б) величины t_i неизвестны, но известно, что $0 < t_i < t_2 < t_3 < \ldots < t_N < T$.

Задача выделения тренда финансового потока сводится к оценке параметров θ_k, $k = \overline{1, s}$, и исследованию этих оценок.

Упрощенный алгоритм решения задачи

Представим сжато классическую теорию метода наименьших квадратов (в целях удобства при сравнении с полученными результатами).

Итак, пусть измерения $\{t_i,\ i = \overline{1, N}\}$ произведены в отрезке $[0, T]$ и на всем этом отрезке x_i могут быть представлены в виде (1). Для построения оценок $\hat{\theta}_k$ параметров θ_k, определяющих тренд, обычно используется метод наименьших квадратов [5] и оценки находятся из условия

$$R = \sum_{i=1}^{N} \left(x_i - \sum_{k=1}^{s} \hat{\theta}_k \varphi_k(t_i) \right)^2 \Rightarrow \min_{\hat{\theta}_k}. \qquad (2)$$

Приравнивая к нулю производную от R по $\hat{\theta}_l$, можно получить

$$\sum_{i=1}^{N} \left(x_i - \sum_{k=1}^{s} \hat{\theta}_k \varphi_k(t_i) \right) \varphi_l(t_i) = 0,$$

что дает систему уравнений, определяющую оценку $\hat{\theta}_k$ параметров θ_k:

$$\sum_{k=1}^{s} \hat{\theta}_k \left(\sum_{i=1}^{N} \varphi_k(t_i)\varphi_l(t_i) \right) = \sum_{i=1}^{N} x_i \varphi_l(t_i). \qquad (3)$$

Записав систему (3) в матричном виде

$$(\varphi^T \varphi)\hat{\vec{\theta}} = \varphi^T \vec{x}, \qquad (4)$$

получаем явное выражение для оценок $\hat{\vec{\theta}}$ параметров $\vec{\theta}$:

$$\hat{\vec{\theta}} = (\varphi^T \varphi)^{(-1)} \varphi^T \vec{x}. \qquad (5)$$

 ISSN 0572-2691

Приведем свойства этих оценок. Соотношение (1) можно записать в виде

$$\vec{x} = \varphi\vec{\theta} + \vec{n}\,. \tag{6}$$

Но тогда с учетом (5) и (6) при фиксированных моментах измерений $\{t_i\}$

$$M\{\vec{x} \mid \{t_i\}\} = \varphi\vec{\theta}\,,$$

и поэтому

$$M\left\{\hat{\vec{\theta}} \mid \{t_i\}\right\} = (\varphi^T\varphi)^{(-1)}(\varphi^T\varphi)\vec{\theta} = \vec{\theta}\,,$$

т.е. оценки $\hat{\vec{\theta}}$ являются несмещенными независимо от моментов измерений $\{t_i\}$.

Выведем формулу для матрицы ковариаций этих оценок.

$$\hat{\vec{\theta}} - \vec{\theta} = (\varphi^T\varphi)^{(-1)}\varphi^T\vec{x} - (\varphi^T\varphi)^{(-1)}(\varphi^T\varphi)\vec{\theta} = (\varphi^T\varphi)^{(-1)}\varphi^T(\vec{x} - \varphi\vec{\theta}) = (\varphi^T\varphi)^{(-1)}\varphi^T\vec{n}\,.$$

Матрица ковариаций V оценок $\hat{\vec{\theta}}$ имеет вид

$$\mathrm{V} = M\left\{(\hat{\vec{\theta}} - \vec{\theta})\cdot(\hat{\vec{\theta}} - \vec{\theta})^T\right\} = (\varphi^T\varphi)^{(-1)}\varphi^T M\{\vec{n}\vec{n}^T\}\varphi(\varphi^T\varphi)^{(-1)}\,.$$

В силу сделанных предположений $M\{\vec{n}\vec{n}^T\} = \sigma^2\mathbf{E}_N$, где $\mathbf{E}_N$ — единичная матрица размерности $N \times N$. Отсюда

$$\mathrm{V} = \sigma^2(\varphi^T\varphi)^{(-1)}\,, \tag{7}$$

что позволяет находить матрицу ковариаций при известных моментах измерений, зная $\{t_i\}$.

Обычно величина σ^2 заранее неизвестна. Однако ее оценку легко построить, используя минимальное значение критерия R (2):

$$R_{\min} = \sum_{i=1}^{N}\left(x_i - \sum_{k=1}^{s}\hat{\theta}_k\varphi_k(t_i)\right)^2 = \left(\vec{x} - \varphi\hat{\vec{\theta}}\right)^T\left(x - \varphi\hat{\vec{\theta}}\right). \tag{8}$$

С учетом (7) и (8) найдем $M\{R_{\min}\big|\{t_i\}\}$. Подставляя вместо $\hat{\vec{\theta}}$ их явное выражение, получим

$$\vec{x} - \varphi\hat{\vec{\theta}} = \varphi\vec{\theta} + \vec{n} - \varphi(\varphi^T\varphi)^{(-1)}\varphi^T(\varphi\vec{\theta} + \vec{n}) = (\mathrm{E}_N - \varphi(\varphi^T\varphi)^{(-1)}\varphi^T)\vec{n}.$$

Поэтому

$$R_{\min} = \vec{n}^T(\mathrm{E}_N - \varphi(\varphi^T\varphi)^{(-1)}\varphi^T)^2\vec{n}\,.$$

Но легко получить, что

$$\left(\mathrm{E}_N - \varphi(\varphi^T\varphi)^{(-1)}\varphi^T\right)^2 = \mathrm{E}_N - \varphi(\varphi^T\varphi)^{(-1)}\varphi^T\,.$$

А так как $M\{n_i n_j\} = \sigma^2\delta_{ij}$, где δ_{ij} — символ Кронекера,

$$M\{R_{\min} \mid \{t_i\}\} = \sigma^2 \cdot \mathrm{Sp}(\mathrm{E}_N - \varphi(\varphi^T\varphi)^{(-1)}\varphi^T)\,.$$

Здесь $\mathrm{Sp}(\mathrm{E}_N) = N$ и

$$\mathrm{Sp}(\varphi(\varphi^T\varphi)^{(-1)}\varphi^T) = \mathrm{Sp}((\varphi^T\varphi)^{(-1)}(\varphi^T\varphi)) = \mathrm{Sp}(\mathrm{E}_s) = s\,,$$

так что

$$M\{R_{\min}\} = \sigma^2(N-s)\,. \qquad (9)$$

Отсюда в соответствии с (9) легко получается несмещенная оценка $\hat{\sigma}^2$ величины σ^2 и оценка $\hat{\mathrm{V}}$ матрицы вариаций V:

$$\hat{\sigma}^2 = R_{\min}/N-s\;, \qquad (10)$$

$$\hat{\mathrm{V}} = \left(R_{\min}(\varphi^T\varphi)^{(-1)} \right) \Big/ (N-s)\,. \qquad (11)$$

Знание $\hat{\mathrm{V}}$ (11) позволяет построить доверительные интервалы для параметров $\vec{\theta}$, по крайней мере для нормальных n_i.

Результаты имитационного моделирования показали, что при измерениях в случайные моменты времени для получения хороших оценок требуются выборки достаточного большого объема — порядка 100 и выше. Но при этом начинают сказываться ошибки округления при вычислениях, особенно при обращении матрицы $(\varphi^T\varphi)$ для $s \geq 3$.

Поэтому ниже предлагается и исследуется упрощенный алгоритм оценки параметров $\vec{\theta}$, не требующий обращения матрицы $(\varphi^T\varphi)$. Кроме того, по аналогии с этим алгоритмом можно строить другие алгоритмы в более сложных случаях.

Элементы матрицы $(\varphi^T\varphi)$ имеют вид:

$$\sum_{i=1}^{N} \varphi_k(t_i)\varphi_l(t_i)\,. \qquad (12)$$

Введем обозначение

$$\varphi_{kl} = \frac{1}{T}\int_0^T \varphi_k(u)\varphi_l(u)du\,. \qquad (13)$$

Представим (12) в виде

$$\sum_{i=1}^{N} \varphi_k(t_i)\varphi_l(t_i) = \lambda T \cdot \frac{1}{\lambda T}\sum_{i=1}^{N} \varphi_k(t_i)\varphi_l(t_i).$$

Здесь последний сомножитель в силу сходимости почти наверное можно заменить на φ_{kl}. Тогда с учетом (13) при больших объемах выборки можно приближенно считать, что

$$\begin{aligned} &(\varphi^T\varphi) = \lambda T \cdot \Phi, \quad \Phi = \|\varphi_{kl}\|, \\ &(\varphi^T\varphi)^{(-1)} = \frac{1}{\lambda T}\Phi^{(-1)}. \end{aligned} \qquad (14)$$

Отсюда оценки $\hat{\vec{\theta}}$ параметров $\vec{\theta}$ с учетом условий (14) примут вид

$$\hat{\vec{\theta}} = \Phi^{(-1)} \cdot \frac{1}{\lambda T}(\varphi^T\vec{x}) \qquad (15)$$

и элементы вектор-столбца $\frac{1}{\lambda T}(\varphi^T\vec{x})$ суть статистики вида

$$\frac{1}{\lambda T}\sum_{i=1}^{N} \varphi_k(t_i)x_i, \quad k = \overline{1, s}\,. \qquad (16)$$

ISSN 0572-2691

Заметим, что $\Phi^{(-1)}$ — числовая матрица.
Рассмотрим частный случай, когда

$$\varphi_k(t) = (t/T)^k\ ,\ k = 0,\ 1,\ ...,\ s, \tag{17}$$

т.е. $\varphi_k(t)$ — степенная функция. Тогда

$$\varphi_{kl} = \frac{1}{T}\int_0^T \left(\frac{t}{T}\right)^k \left(\frac{t}{T}\right)^l dt = \frac{1}{k+l+1}$$

и матрица Φ имеет вид

$$\Phi = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{s+1} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \cdots & \frac{1}{s+2} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \cdots & \frac{1}{s+3} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{1}{s+1} & \frac{1}{s+2} & \frac{1}{s+3} & \cdots & \frac{1}{2s+1} \end{bmatrix}. \tag{18}$$

Приведем для матрицы (18) несколько матриц $\Phi^{(-1)}$, полученных на компьютере с помощью программы MathCad 6.0:

$$s = 1\ \ \Phi^{(-1)} = \begin{bmatrix} 4 & -6 \\ -6 & 12 \end{bmatrix}$$

$$s = 2\ \ \Phi^{(-1)} = \begin{bmatrix} 9 & -36 & 30 \\ -36 & 192 & -180 \\ 30 & -180 & 180 \end{bmatrix}$$

$$s = 3\ \ \Phi^{(-1)} = \begin{bmatrix} 16 & -120 & 240 & -140 \\ -120 & 1200 & -2700 & 1680 \\ 240 & -2700 & 6480 & -4200 \\ -140 & 1680 & -4200 & 2800 \end{bmatrix}$$

$$s = 4\ \ \Phi^{(-1)} = \begin{bmatrix} 25 & -300 & 1050 & -1400 & 630 \\ -300 & 4800 & -18900 & 26880 & -12600 \\ 1050 & -18900 & 79380 & -117600 & 56700 \\ -1400 & 26880 & -117600 & 179200 & -88200 \\ 630 & -12600 & 56700 & -88200 & 44100 \end{bmatrix}.$$

Очевидно, что с ростом s элементы матрицы $\Phi^{(-1)}$ очень быстро растут, и это приводит к неустойчивости оценок из-за влияния ошибок округления, в связи с чем использовать полиномы степени выше 4 не рекомендуется.

Анализ свойств упрощенного алгоритма оценки параметров при известных моментах осуществления финансовых транзакций

Исследуем свойства предложенных упрощенных оценок, считая, что поток моментов осуществления финансовых транзакций является пуассоновским с постоянной интенсивностью λ.

Так как $\vec{x} = \varphi\vec{\theta} + \vec{n}$,

$$M\left\{\frac{1}{\lambda T}\varphi^T \vec{x} | \{t_i\}\right\} = \frac{1}{\lambda T}(\varphi^T \varphi)\vec{\theta},$$

так что

$$M\left\{\hat{\vec{\theta}} \,|\, \{t_i\}\right\} = \Phi^{(-1)} \cdot \frac{1}{\lambda T}(\varphi^T \varphi)\vec{\theta}, \tag{19}$$

т.е. в упрощенных оценках (19) исчезает свойство несмещенности оценок при любых наборах моментов измерений $\{t_i\}$. Свойство несмещенности сохраняется лишь в среднем, после усреднения по моментам измерений $\{t_i\}$:

$$M\left\{\hat{\vec{\theta}}\right\} = \Phi^{(-1)} M\left\{\frac{1}{\lambda T}(\varphi^T \varphi)\right\}\vec{\theta} = \Phi^{(-1)}\Phi\vec{\theta} = \vec{\theta}. \tag{20}$$

Далее замена $(\varphi^T\varphi)^{(-1)}$ на $\Phi^{(-1)}$ ведет к тому, что дисперсия оценок $\hat{\vec{\theta}}$ параметров $\vec{\theta}$ возрастает. Выведем поэтому матрицу ковариаций для оценок $\hat{\vec{\theta}}$ параметров $\vec{\theta}$.

Заметим, что

$$M\left\{\frac{1}{\lambda T}\sum_{i=1}^{N}\varphi_k(t_i)x_i\right\} = M\left\{\frac{1}{\lambda T}\sum_{r=1}^{s}\theta_r\sum_{i=1}^{N}\varphi_k(t_i)\varphi_r(t_i) + \frac{1}{\lambda T}\sum_{i=1}^{N}\varphi_k(t_i)n_i\right\} =$$
$$= \frac{1}{\lambda T}\sum_{r=1}^{s}\theta_r \cdot \lambda\int_0^T \varphi_k(u)\varphi_r(u)du = \sum_{r=1}^{s}\varphi_{kr}\theta_r,$$

так что математическое ожидание вектора $\left[\frac{1}{\lambda T}\sum_{i=1}^{N}\varphi_k(t_i)x_i\right]$ равно. Поэтому

$$\hat{\vec{\theta}} - \vec{\theta} = \Phi^{(-1)}\left[\frac{1}{\lambda T}(\varphi^T\vec{x}) - \Phi\vec{\theta}\right]$$

и

$$\mathrm{V} = M\left\{\left(\hat{\vec{\theta}} - \vec{\theta}\right)\left(\hat{\vec{\theta}} - \vec{\theta}\right)^T\right\} = \Phi^{(-1)} M\left\{\left[\frac{1}{\lambda T}(\varphi^T\vec{x} - \Phi\theta)\right]\cdot\left[\frac{1}{\lambda T}(\varphi^T\vec{x}) - \Phi\vec{\theta}\right]^T\right\}\Phi^{(-1)}.$$

А так как $\vec{x} = \varphi\vec{\theta} + \vec{n}$,

$$\frac{1}{\lambda T}(\varphi^T\vec{x}) - \Phi\vec{\theta} = \frac{1}{\lambda T}(\varphi^T\vec{x})\vec{\theta} + \frac{1}{\lambda T}\varphi^T\vec{n} - \Phi\vec{\theta} = \left(\frac{1}{\lambda T}(\varphi^T\varphi) - \Phi\right)\vec{\theta} + \frac{1}{\lambda T}\varphi^T\vec{n}.$$

Поэтому

$$M\left\{\left[\frac{1}{\lambda T}(\varphi^T\vec{x}) - \Phi\vec{\theta}\right]\cdot\left[\frac{1}{\lambda T}(\varphi^T\vec{x}) - \Phi\vec{\theta}\right]^T\right\} =$$
$$= M\left\{\left[\left(\frac{1}{\lambda T}(\varphi^T\varphi) - \Phi\right)\vec{\theta} + \frac{1}{\lambda T}\varphi^T\vec{n}\right]\cdot\left[\left(\frac{1}{\lambda T}(\varphi^T\varphi) - \Phi\right)\vec{\theta} + \frac{1}{\lambda T}\varphi^T\vec{n}\right]^T\right\}. \tag{21}$$

Усредняя по $\vec{n}$ с учетом того, что $M\{\vec{n}\} = 0$ и $M\{\vec{n}\vec{n}^T\} = \sigma^2 \cdot \mathrm{E}_N$, получим, что при фиксированных $\{t_i\}$ правая часть (21) равна

 ISSN 0572-2691

$$\frac{\sigma^2}{(\lambda T)^2} M\{\varphi^T\varphi\} + M\left\{\left(\frac{1}{\lambda T}(\varphi^T\varphi) - \Phi\right)\vec{\theta}\vec{\theta}^T\left(\frac{1}{\lambda T}(\varphi^T\varphi) - \Phi\right)\right\}, \qquad (22)$$

где осталось усреднение лишь по $\{t_i\}$.

Но $M\{(\varphi^T\varphi)\} = \lambda T \cdot \Phi$, так что первое слагаемое в (22) равно $\sigma^2\Phi/\lambda T$. Во втором слагаемом усредняемое выражение образует матрицу с элементами

$$\left(\frac{1}{\lambda T}\sum_{r=1}^{s}\theta_r\sum_{i=1}^{N}\varphi_k(t_i)\,\varphi_r(t_i) - \sum_{r=1}^{s}\theta_r\varphi_{kr}\right)\cdot\left(\frac{1}{\lambda T}\sum_{p=1}^{s}\theta_p\sum_{j=1}^{N}\varphi_l(t_j)\varphi_p(t_j) - \sum_{p=1}^{s}\theta_p\varphi_{lp}\right). \qquad (23)$$

Усреднение этого выражения по моментам измерений $\{t_i\}$ дает

$$\frac{1}{(\lambda T)^2}\sum_{r=1}^{s}\sum_{p=1}^{s}\theta_r\theta_p M\left\{\sum_{i=1}^{N}\sum_{j=1}^{N}\varphi_k(t_i)\varphi_r(t_i)\varphi_l(t_j)\varphi_p(t_j)\right\} - \sum_{r=1}^{s}\sum_{p=1}^{s}\theta_r\theta_p\varphi_{kr}\varphi_{lp}.$$

Но

$$M\left\{\frac{1}{(\lambda T)^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\varphi_k(t_i)\varphi_r(t_i)\varphi_l(t_j)\varphi_p(t_j)\right\} = \varphi_{kr}\varphi_{lp} + \frac{1}{\lambda T}\varphi_{krlp},$$

где

$$\varphi_{krlp} = \frac{1}{T}\int_0^T \varphi_k(u)\varphi_r(u)\varphi_l(u)\varphi_p(u)du, \qquad (24)$$

поэтому математическое ожидание (23) с учетом (24) равно

$$\frac{1}{\lambda T}\sum_{r,\,p=1}^{s}\theta_r\theta_p\varphi_{krlp}. \qquad (25)$$

Объединяя все вместе, получим матрицу вариаций оценок $\hat{\vec{\theta}}$, имеющую вид

$$\mathrm{V} = \frac{1}{\lambda T}\Phi^{(-1)}\left[\sigma^2\Phi + \left\|\sum_{r,\,p=1}^{s}\theta_r\theta_p\varphi_{krlp}\right\|\right]\Phi^{(-1)}. \qquad (26)$$

Второе дополнительное слагаемое в (26) дает нам добавочную дисперсию оценок, возникающую из-за замены матрицы $(\varphi^{\mathrm{T}}\varphi)^{(-1)}$ на $\Phi^{(-1)}$.

Заметим, однако, что при $\lambda\mathrm{T}\to\infty$ матрица $\mathbf{V}\to 0$, так что получаемые оценки сходятся в среднеквадратичном смысле к истинным значениям параметров $\vec{\theta}$. Относительное увеличение дисперсий оценок, по сравнению с оценками метода наименьших квадратов, определяется матрицей с элементами

$$\frac{1}{\sigma^2}\left\|\sum_{r,\,p=1}^{s}\theta_r\theta_p\varphi_{krlp}\right\|$$

и зависит от значений параметров $\vec{\theta}$ и функций $\varphi_k(t)$.

Рассмотрим теперь более подробно асимптотические свойства упрощенных оценок при ограничениях на величины n_i и функции времени $f(t)$.

Теорема 1. Пусть $x_i = f(t_i) + n_i$. Тогда статистики вида

$$S = \frac{1}{\lambda T} \sum_{i=1}^{N} \varphi(t_i) x_i \tag{27}$$

при $\lambda \to \infty$ (для случая $\lambda_n = \lambda_0 \cdot n$) сходятся почти наверное к величине

$$\frac{1}{T} \int_0^T \varphi(t) f(t) dt\,. \tag{28}$$

Доказательство. Действительно, используя представление для x_i, запишем S в виде

$$S = \frac{1}{\lambda T} \sum_{i=1}^{N} \varphi(t_i) f(t_i) + \frac{1}{\lambda T} \sum_{i=1}^{N} \varphi(t_i) n_i\,. \tag{29}$$

Что касается первого слагаемого, то нетрудно показать, что при $\lambda \to \infty$ оно сходится почти наверное к величине (28). Также нетрудно показать, что и второе слагаемое

$$S_1 = \frac{1}{\lambda T} \sum_{i=1}^{N} \varphi(t_i) n_i$$

при $\lambda \to \infty$ почти наверное сходится к нулю.

Объединяя все вместе, получим, что при $\lambda \to \infty$

$$S \xrightarrow{\text{п.н.}} \frac{1}{T} \int_0^T f(t) \varphi(t) dt\,. \tag{30}$$

Теорема доказана.

Так как все упрощенные оценки представляют собой статистики вида (27), то отсюда следует, что при $\lambda \to \infty$ $\hat{\vec{\theta}} \xrightarrow{\text{п.н.}} \vec{\theta}$, т.е. оценки являются сильно состоятельными.

Теорема 2. Пусть $x_i = f(t_i) + n_i$. Тогда статистики вида

$$S = \sqrt{\lambda} \left(\frac{1}{\lambda} \sum_{i=1}^{N} \varphi(t_i) x_i - \int_0^T f(u) \varphi(u) du \right) \tag{31}$$

при $\lambda \to \infty$ сходятся по распределению к нормальной случайной величине с нулевым математическим ожиданием и дисперсией

$$\sigma^2 \int_0^T \varphi^2(u) du + \int_0^T f^2(u) \varphi^2(u) du\,. \tag{32}$$

Доказательство. Рассмотрим статистику S_t вида

$$S_t = \sqrt{\lambda} \left(\frac{1}{\lambda} \sum_{\{i: t_i > t\}} \varphi(t_i) x_i - \int_t^T f(u) \varphi(u) du \right), \tag{33}$$

так что $S = S_0$, и пусть $g(\omega, t) = M\{e^{i\omega S_t}\}$ — характеристическая функция статистики S_t.

Переходя к моменту времени $t + \Delta t$, получим, что с вероятностью $1 - \lambda \cdot \Delta t + o(\Delta t)$ на интервале $[t, t + \Delta t]$ измерений сделано не будет. В случае

$$S_t = S_{t+\Delta t} - \varphi(t) f(t) \Delta t \sqrt{\lambda} + o(\Delta t)$$

с вероятностью $\lambda \cdot \Delta t + o(\Delta t)$ на интервале $[t, t+\Delta t]$ будет сделано измерение. В этом случае

$$S_t = S_{t+\Delta t} + \frac{1}{\sqrt{\lambda}} \varphi(t) x(t) - \varphi(t) f(t) \sqrt{\lambda} \Delta t + o(\Delta t)\,.$$

Поэтому для $g(\omega,\ t)$ получаем соотношение

$$g(\omega,\, t) = (1 - \lambda \Delta t)\, g\,(\omega,\, t + \Delta t)\, \mathrm{e}^{-i\omega\sqrt{\lambda} f(t)\varphi(t)\Delta t} +$$

$$+ \lambda \Delta t\, g(\omega, t + \Delta t) M \left\{ \mathrm{e}^{i\omega \frac{1}{\sqrt{\lambda}} \varphi(t) f(t) + i\omega \frac{1}{\sqrt{\lambda}} \varphi(t) n + O(\Delta t)} \right\} + o(\Delta t). \quad (34)$$

Разлагая в ряды Тейлора

$$g(\omega,\, t + \Delta t) = g(\omega,\, t) + \frac{\partial g(\omega,\, t)}{\partial t} \Delta t + o(\Delta t),$$

$$\mathrm{e}^{-i\omega\sqrt{\lambda} f(t)\varphi(t)\Delta t} = 1 - i\omega f(t)\varphi(t)\sqrt{\lambda}\Delta t + o(\Delta t),$$

$$\mathrm{e}^{\frac{i\omega}{\sqrt{\lambda}}(\varphi(t) f(t) + \varphi(t) n + O(\Delta t))} =$$

$$= 1 + \frac{i\omega}{\sqrt{\lambda}} (\varphi(t) f(t) + \varphi(t) n) - \frac{\omega^2}{2\lambda} (\varphi(t) f(t) + \varphi(t) n)^2 + O(\frac{1}{\lambda^{3/2}},\, \Delta t),$$

подставляя эти разложения в (34) и усредняя по величине n, получим

$$g(\omega,\, t) = g(\omega,\, t) + \Delta t \cdot \{-\lambda g(\omega,\, t) + \frac{\partial g(\omega,\, t)}{\partial t} - i\sqrt{\lambda} f(t)\varphi(t) g(\omega,\, t) + \lambda g(\omega,\, t) +$$

$$+ i\omega\sqrt{\lambda} f(t)\varphi(t) g(\omega,\, t) - \frac{\omega^2}{2} g(\omega,\, t)[\sigma^2 \varphi^2(t) + \varphi^2(t) f^2(t)\ ]\} + o(\Delta t).$$

Сокращая $g(\omega, t)$, деля на Δt и переходя сначала к пределу $\Delta t \to 0$, а затем $\lambda \to \infty$, получим уравнение, определяющее $g(\omega, t)$:

$$\frac{\partial g(\omega,t)}{\partial t} = g(\omega,t) \cdot \frac{\omega^2}{2} [\sigma^2 \varphi^2(t) + f^2(t) \varphi^2(t)]\,. \quad (35)$$

Решение (35) имеет вид

$$g(\omega,\, t) = \exp\left(-\frac{\omega^2}{2} \left(\sigma^2 \int_t^T \varphi^2(u) du + \int_t^T f^2(u) \varphi^2(u) du \right) \right). \quad (36)$$

В частности, при $t = 0$

$$g(\omega,\, 0) = M\{\mathrm{e}^{i\omega S}\} = \exp\left(-\frac{\omega^2}{2} \left(\sigma^2 \int_0^T \varphi^2(u) du + \int_0^T f^2(u) \varphi^2(u) du \right) \right), \quad (37)$$

что соответствует нормальной случайной величине с нулевым математическим ожиданием и дисперсией

$$\sigma^2 \int_0^T \varphi^2(u) du + \int_0^T f^2(u) \varphi^2(u) du.$$

Теорема доказана.

Отметим, что аналогично доказывается асимптотическая нормальность статистик вида

$$S_k = \sqrt{\lambda}\left(\frac{1}{\lambda}\sum_{i=1}^{N}\varphi_k(t_i)x_i - \int_0^T f(u)\varphi_k(u)du\right), \tag{38}$$

только выкладки становятся более громоздкими.

Из этой теоремы следует, что величины $\sqrt{\lambda T}(\hat{\theta}_k - \theta_k)$ при $\lambda \to \infty$ являются асимптотическими, совместно нормальными величинами с нулевым математическим ожиданием и ковариационной матрицей

$$\mathrm{V}_0 = \Phi^{(-1)}\left[\sigma^2\Phi + \left\|\sum_{r,\,p=1}^{s}\theta_r\theta_p\varphi_{krlp}\right\|\right]\Phi^{(-1)}. \tag{39}$$

Для того чтобы иметь возможность строить доверительные интервалы для неизвестных параметров θ_k хотя бы в асимптотике $\lambda T \to \infty$, надо иметь оценку $\hat{\mathrm{V}}_0$ матрицы V_0. Построим такую оценку для матрицы V_0 (39)

$$\mathrm{R} = \sigma^2\Phi + \left\|\sum_{r,\,p=1}^{s}\theta_r\theta_p\varphi_{krlp}\right\| \tag{40}$$

с элементами

$$R_{kl} = \sigma^2\varphi_{kl} + \sum_{r,\,p=1}^{s}\theta_r\theta_p\varphi_{krlp}. \tag{41}$$

В качестве оценок $\hat{R}_{kl}$ величины R_{kl} (41) рассмотрим статистики

$$\hat{R}_{kl} = \frac{1}{\lambda T}\sum_{i=1}^{N}\varphi_k(t_i)\varphi_l(t_i)x_i^2. \tag{42}$$

Исследуем свойства этих оценок. Так как $x_i = n_i + \sum_{r=1}^{s}\theta_r\varphi_r(t_i)$,

$$\hat{R}_{kl} = \frac{1}{\lambda T}\sum_{i=1}^{N}\varphi_k(t_i)\varphi_l(t_i)\left[n_i^2 + 2n_i\sum_{r=1}^{s}\theta_r\varphi_r(t_i) + \sum_{r,\,p=1}^{s}\theta_r\theta_p\varphi_r(t_i)\varphi_p(t_i)\right].$$

Усредняя сначала по n_i, а затем по t_i, получим

$$M\{\hat{R}_{kl}\} = \sigma^2\varphi_{kl} + \sum_{r=1}^{s}\sum_{p=1}^{s}\theta_r\theta_p\varphi_{krlp},$$

так что $\hat{R}_{kl}$ является несмещенной оценкой величины R_{kl}.

Рассмотрим теперь асимптотические свойства этих оценок при $\lambda \to \infty$. В целях упрощения выкладок и большей общности рассмотрим статистики вида

$$S = \frac{1}{\lambda T}\sum_{i=1}^{N}\varphi(t_i)x_i^2, \tag{43}$$

считая, что $x_i = f(t_i) + n_i$.

Теорема 3. При $\lambda \to \infty$ статистика (43)

$$S = \frac{1}{\lambda T}\sum_{i=1}^{N}\varphi(t_i)x_i^2$$

сходится в среднеквадратичном к величине

$$\sigma^2 \cdot \frac{1}{T}\int_0^T \varphi(t)dt + \frac{1}{T}\int_0^T \varphi(t)f^2(t)dt. \tag{44}$$

Доказательство. Вычисляя $M\{S\}$, получим

$$M\{S\}=\frac{1}{\lambda T}M\left\{\sum_{i=1}^{N}\varphi(t_i)[n_i^2+2n_i f(t_i)+f^2(t_i)]\right\}=\sigma^2\cdot\frac{1}{T}\int_0^T\varphi(t)dt+\frac{1}{T}\int_0^T\varphi(t)f^2(t)dt.$$

Для вычисления дисперсии величины S найдем

$$M\{S^2\}=\frac{1}{(\lambda T)^2}M\left\{\sum_{i=1}^{N}\sum_{j=1}^{N}\varphi(t_i)\varphi(t_j)[n_i^2+2n_i f(t_i)+f^2(t_i)]\cdot[n_j^2+2n_j f(t_j)+f^2(t_j)]\right\}.$$

Выделяя отдельно слагаемые с $i=j$ и усредняя по n_i, получим

$$M\{S^2\}=\frac{1}{(\lambda T)^2}M_t\left\{\sum_{i=1}^{N}\varphi^2(t_i)[m_4+6\sigma^2 f^2(t_i)+f^4(t_i)]+\right.$$
$$\left.+\sum_{i\neq j}\varphi(t_i)\varphi(t_j)[\sigma^2+f^2(t_i)][\sigma^2+f^2(t_j)]\right\}.$$

Усредняя по моментам измерений $\{t_i\}$, получаем

$$M\{S^2\}=M\{S\}^2+\frac{1}{\lambda T}\left[m_4\cdot\frac{1}{T}\int_0^T\varphi^2(u)du+6\sigma^2\frac{1}{T}\int_0^T\varphi^2(u)f^2(u)du+\frac{1}{T}\int_0^T f^4(u)du\right],$$

отсюда

$$D\{S\}=\frac{1}{\lambda T}\left[m_4\cdot\frac{1}{T}\int_0^T\varphi^2(u)du+6\sigma^2\frac{1}{T}\int_0^T\varphi^2(u)f^2(u)du+\frac{1}{T}\int_0^T f^4(u)du\right].\qquad(45)$$

Мы видим, что $D\{S\}$ убывает как $\frac{1}{\lambda T}$ и при $\lambda T\to\infty$ $D\{S\}\to 0.$ Отсюда следует сходимость S к $M\{S\}$ в среднеквадратичном смысле при $\lambda\to\infty$.

Теорема доказана.

Теорема 4. При $\lambda\to\infty$

$$S=\frac{1}{\lambda T}\sum_{i=1}^{N}\varphi(t_i)x_i^2\xrightarrow{\text{п.н.}}M\{S\}=\sigma^2\cdot\frac{1}{T}\int_0^T\varphi(t)dt+\frac{1}{T}\int_0^T\varphi(t)f^2(t)dt\,.$$

Доказательство. Имеем

$$S=\frac{1}{\lambda T}\sum_{i=1}^{N}\varphi(t_i)n_i^2+\frac{2}{\lambda T}\sum_{i=1}^{N}f(t_i)\varphi(t_i)n_i+\frac{1}{\lambda T}\sum_{i=1}^{N}f^2(t_i)\varphi(t_i)\,.$$

Нетрудно видеть, что при $\lambda\to\infty$

$$\frac{1}{\lambda T}\sum_{i=1}^{N}f^2(t_i)\varphi(t_i)\xrightarrow{\text{п.н.}}\frac{1}{T}\int_0^T f^2(u)\varphi(u)du$$

и

$$\frac{1}{\lambda T}\sum_{i=1}^{N}f(t_i)\varphi(t_i)n_i\xrightarrow{\text{п.н.}}0\,.$$

И наконец, оценим сходимость первого слагаемого. Для этого рассмотрим статистику

$$S_1=\frac{1}{\lambda T}\sum_{i=1}^{N}\varphi(t_i)(n_i^2-\sigma^2)\,.$$

Усредним ее по n_i, тогда нетрудно показать, что при $\lambda\to\infty$ $S_1\xrightarrow{\text{п.н.}}0.$

Итак, имеем

$$\frac{1}{\lambda T}\sum_{i=1}^{N}\varphi(t_i)n_i^2-\sigma^2\frac{1}{\lambda T}\sum_{i=1}^{N}\varphi(t_i)\xrightarrow{\text{п.н.}}0,\quad\frac{1}{\lambda T}\sum_{i=1}^{N}\varphi(t_i)-\frac{1}{T}\int_0^T\varphi(u)du\xrightarrow{\text{п.н.}}0.$$

Поэтому при $\lambda \to \infty$

$$\frac{1}{\lambda T}\sum_{i=1}^{N} \varphi(t_i) n_i^2 \xrightarrow{\text{п.н.}} \sigma^2 \cdot \frac{1}{T}\int_0^T \varphi(u)du.$$

Объединяя все сходимости почти наверное, получим доказываемую теорему.

Теоремы 3 и 4 обосновывают возможность замены R_{kl} на $\hat{R}_{kl}$, по крайней мере при больших λT. Поэтому в качестве оценки $\hat{\text{V}}$ матрицы ковариаций V можно использовать матрицу

$$\hat{\text{V}} = \frac{1}{\lambda T}\Phi^{(-1)}\hat{\text{R}}\Phi^{(-1)} = \frac{1}{(\lambda T)^2}\cdot\Phi^{(-1)}\cdot\left\|\sum_{i=1}^{N}\varphi_k(t_i)\varphi_l(t_i)x_i^2\right\|\cdot\Phi^{(-1)}. \qquad (46)$$

Вместе с асимптотической нормальностью оценок эта формула позволяет строить доверительные интервалы для неизвестных параметров при $\lambda T >> 1$.

В целях проверки работоспособности исследованных алгоритмов, а также приложимости нормальной аппроксимации для оценок проведено их имитационное моделирование. Результаты имитационного моделирования данной части совместно с результатами имитационного моделирования второй части, как и общее заключение, приведены в конце работы.

Ф.Ф. Ідрісов

НАБЛИЖЕНІ АЛГОРИТМИ ВИДІЛЕННЯ ТРЕНДІВ В ЗАДАЧАХ ФІНАНСОВОЇ РОЗВІДКИ. Частина 1. МОМЕНТИ ПОЯВИ ЕЛЕМЕНТІВ ФІНАНСОВОГО ПОТОКУ ВІДОМІ ТОЧНО

Розглянуто наближені алгоритми оцінювання трендів в задачах фінансової розвідки. Досліджено випадок, коли моменти появи елементів фінансового потоку відомі точно. Вивчено статистичні характеристики отриманих оцінок.

F.F. Idrisov

APPROXIMATIVE ALGORITHMS OF ESTIMATING TRENDS FOR FINANCIAL INTELLIGENCE TASKS. Part I. THE MOMENTS WHEN THE ELEMENTS OF THE FINANCIAL FLOW APPEAR ARE KNOWN EXACTLY

Approximate algorithms for estimating trends in financial intelligence tasks are considered. The case when the moments of the appearance of the elements of the financial flow are known exactly. The statistical characteristics of the obtained estimates are studied.

1. *Андерсон Т.* Статистический анализ временных рядов. — М. :Мир,1976. — 755 с.
2. *Бокс Дж.*, *Дженкинс Г.* Анализ временных рядов. Прогноз и управление. — М. : Мир, 1974. — Вып. 1 — 406 с.; Вып. 2 — 197 с.
3. *Бриллинджер Д.Р.* Временные ряды. Обработка данных и теория. — М. : Мир, 1980. — 536 с.
4. *Идрисов Ф.Ф.* Оценка параметров динамических систем при случайных пропусках измерений // Международный научно-технический журнал «Проблемы управления и информатики». — 2016. — № 4. — С. 18–26
5. *Рао С.Р.* Линейные статистические методы и их приложения. — М. : Наука, 1968. — 547 с.

Получено 18.04.2017