

**ПРИБЛИЖЕННЫЕ МОДЕЛИ ФИНАНСОВОЙ
РАЗВЕДКИ ПРИ НЕТОЧНО ЗАДАННЫХ
МОМЕНТАХ ВРЕМЕНИ ОСУЩЕСТВЛЕНИЯ
ФИНАНСОВЫХ ТРАНЗАКЦИЙ***

Введение

Финансовая разведка в контексте возросших террористических угроз — очень важный инструмент антитеррора. В этих целях была создана FATF как межправительственная организация, вырабатывающая всеобщие стандарты противодействия отмыванию преступных доходов и финансированию терроризма, а также осуществляющая оценки соответствия национальных систем государств этим стандартам.

Имеющиеся в открытой печати работы, посвященные данной проблеме, нацелены либо на освещение организационных и методологических вопросов [1–3], либо анализируют проблему в теоретико-игровом аспекте, весьма удаленном от практических приложений [4, 5].

В данной работе группы в вероятностной постановке рассматриваются вопросы выявления и оценивания финансовых потоков, направляемых в преступные группы. Так, в работах [6, 7] рассматривались вопросы оценивания трендов финансовых потоков для случаев, когда моменты осуществления финансовых транзакций были заданы точно, либо были неизвестны. Но вместе с тем встречаются ситуации, когда моменты появления элементов финансового потока могут быть известны с некоторой погрешностью, т.е. известны не сами моменты времени финансовых транзакций t_i , а величины $\tau_i = t_i + \xi_i$, где ξ_i — независимые одинаково распределенные случайные величины с $M\{\xi_i\} = 0$ и дисперсией $D\{\xi_i\} = \sigma_0^2$. Для определенности рассмотрим случай, когда ξ_i — нормальные случайные величины.

Постановка задачи

Как и в предыдущих работах [6, 7], будем считать, что наблюдается поток финансовых транзакций на интервале времени $[0, T]$ в моменты времени t_i , образующих пуассоновский поток событий постоянной интенсивности λ и N — количество таких наблюдений. Допустим, что измеренные значения $x_i = x(t_i)$ представимы в виде

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

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

Всюду в дальнейшем будем полагать, что $\varphi_k(t)$ — непрерывные ограниченные функции. Оценки параметров будем искать по модифицированному методу наименьших квадратов [7] из условия

$$R = \sum_{i=1}^N \left(x_i - \sum_{k=1}^s \hat{\theta}_k \bar{\varphi}_k(\tau_i) \right)^2 \Rightarrow \min_{\hat{\theta}_k},$$

* Работа выполнена при поддержке Министерства образования и науки Российской Федерации. Проект 8 9562 2017/8,9.

© Ф.Ф. ИДРИСОВ, 2018

Международный научно-технический журнал
«Проблемы управления и информатики», 2018, № 6

где

$$\bar{\varphi}_k(\tau_i) = \int_{-\infty}^{\infty} \varphi_k(t_i) p(t_i | \tau_i) dt_i, \quad (1)$$

т.е. $\bar{\varphi}_k(\tau_i)$ — условное среднее функции $\varphi_k(t_i)$, если τ_i известно.

Заметим, что, вообще говоря, надо было бы искать условное среднее $\varphi_k(t_i)$ когда известны все значения $\tau_1, \tau_2, \dots, \tau_N$, но это настолько усложняет задачу, что делает результирующие формулы необозримыми.

Найдем $p(t_i | \tau_i)$. Как было показано в работе [7] (см. раздел «Упрощенный алгоритм решения задачи», вывод соотношений (7) и (8)), при $N \gg 1$ можно приближенно считать, что t_i является нормальной случайной величиной

$$M\{t_i | N\} = T \frac{i}{N+1}, \quad D\{t_i | N\} = T^2 \frac{i(N+1-i)}{(N+1)^2(N+2)}. \quad (2)$$

Величина τ_i при фиксированном t_i по предположениям о свойствах величин ξ_i также является нормальной случайной величиной с $M(\tau_i | t_i) = t_i$, $D(\tau_i | t_i) = \sigma_0^2$. В силу этого при $N \gg 1$ двумерная случайная величина (t_i, τ_i) — нормальная случайная величина с $M\{t_i | N\} = M\{\tau_i | N\}$ и ковариационной матрицей

$$\begin{bmatrix} t_i & D\{t_i | N\} \\ D\{t_i | N\} & \sigma_0^2 + D\{t_i | N\} \end{bmatrix}.$$

Поэтому согласно свойствам многомерных нормальных величин [8, 9] условная плотность вероятностей $p(t_i | \tau_i)$ момента i -го измерения при известном значении τ_i и числе измерений N будет также нормальной с параметрами

$$M\{t_i | \tau_i\} = T \frac{i}{N+1} + \frac{D\{t_i | N\}}{\sigma_0^2 + D\{t_i | N\}} \left(\tau_i - T \frac{i}{N+1} \right),$$

$$D\{t_i | \tau_i\} = \frac{\sigma_0^2 D\{t_i | N\}}{\sigma_0^2 + D\{t_i | N\}}. \quad (3)$$

Если $\sigma_0^2 \ll D\{t_i | N\}$, то приближенно

$$M\{t_i | \tau_i\} = \tau_i + \frac{\sigma_0^2}{D\{t_i | N\}} \cdot T \cdot \frac{i}{N+1},$$

$$D\{t_i | N\} \approx \sigma_0^2.$$

Отметим, что при $\sigma_0^2 \rightarrow \infty$ формулы (3) переходят в (2), т.е. в случай, когда о t_i известен лишь их порядок [7].

Алгоритм нахождения оценок при пуассоновском потоке финансовых транзакций

Теперь имеется возможность вычислить и $\bar{\varphi}_k(\tau_i)$, так как $p(t_i | \tau_i)$, входящая в (1), известна. При некоторых конкретных видах $\varphi_k(t)$ этот интеграл достаточно просто вычисляется. Если, например, $\varphi_k(t) = (t/T)^k$, то

$$\bar{\varphi}_0(\tau_i) = 1,$$

$$\bar{\varphi}_1(\tau_i) = \frac{i}{N+1} + \frac{D\{t_i|N\}}{\sigma_0^2 + D\{t_i|N\}} \left(\frac{\tau_i}{T} - \frac{i}{N+1} \right),$$

$$\bar{\varphi}_2(\tau_i) = \left[\frac{i}{N+1} + \frac{D\{t_i|N\}}{\sigma_0^2 + D\{t_i|N\}} \left(\frac{\tau_i}{T} - \frac{i}{N+1} \right) \right]^2 + \frac{\sigma_0^2 D\{t_i|N\}}{T^2(\sigma_0^2 + D\{t_i|N\})},$$

$$\begin{aligned} \bar{\varphi}_3(\tau_i) &= \left[\frac{i}{N+1} + \frac{D\{t_i|N\}}{\sigma_0^2 + D\{t_i|N\}} \left(\frac{\tau_i}{T} - \frac{i}{N+1} \right) \right]^3 + \\ &+ 3 \left[\frac{i}{N+1} + \frac{D\{t_i|N\}}{\sigma_0^2 + D\{t_i|N\}} \left(\frac{\tau_i}{T} - \frac{i}{N+1} \right) \right] \cdot \frac{\sigma_0^2 D\{t_i|N\}}{T^2(\sigma_0^2 + D\{t_i|N\})} \end{aligned}$$

и так далее. Сложность этих формул быстро растет с ростом степени полинома.

Зная величины $\bar{\varphi}_k(\tau_i)$, можно получить и явный вид оценок $\hat{\theta}_k$ параметров θ_k .

Приравнявая к нулю $\partial R / \partial \hat{\theta}_s$, оценки находим из системы уравнений

$$\sum_{k=1}^s \hat{\theta}_k \sum_{i=1}^N \bar{\varphi}_k(\tau_i) \bar{\varphi}_l(\tau_i) = \sum_{i=1}^N \bar{\varphi}_l(\tau_i) x_i.$$

Переходя к матричным обозначениям

$$\bar{\varphi} = \begin{bmatrix} \bar{\varphi}_1(\tau_1) & \bar{\varphi}_2(\tau_1) & \dots & \bar{\varphi}_s(\tau_1) \\ \bar{\varphi}_1(\tau_2) & \bar{\varphi}_2(\tau_2) & \dots & \bar{\varphi}_s(\tau_2) \\ \dots & \dots & \dots & \dots \\ \bar{\varphi}_1(\tau_N) & \bar{\varphi}_2(\tau_N) & \dots & \bar{\varphi}_s(\tau_N) \end{bmatrix},$$

можем записать

$$\hat{\theta} = (\bar{\varphi}^T \bar{\varphi})^{(-1)} \bar{\varphi}^T \bar{x}, \quad (4)$$

что и дает явные выражения для оценок неизвестных параметров.

Вводя матрицу

$$\bar{\Phi} = \frac{1}{N} (\bar{\varphi}^T \bar{\varphi}),$$

запишем (4) в виде

$$\hat{\theta} = \bar{\Phi}^{(-1)} \cdot \frac{1}{N} \bar{\varphi}^T \bar{x},$$

а элементы вектор-столбца $\frac{1}{N} \bar{\varphi}^T \bar{x}$ имеют вид

$$\frac{1}{N} \sum_{i=1}^N \varphi_k(\tau_i) x_i.$$

Свойства оценок параметров тренда финансовых транзакций

Как было сказано ранее, $\bar{x} = \varphi \bar{\theta} + \bar{n}$, где φ — матрица $\|\varphi_k(t_i)\|$. Усредняя при фиксированных τ_i , получаем

$$M\{\bar{x}|\{\tau_i\}\} = M\{\varphi(t_i)|\tau_i\} \bar{\theta} + M\{\bar{n}\} = \bar{\varphi} \bar{\theta},$$

поэтому

$$M\{\hat{\theta}\} = (\bar{\varphi}^T \bar{\varphi})^{-1} (\bar{\varphi}^T \bar{\varphi}) \bar{\theta} = \bar{\theta},$$

т.е. оценки являются несмещенными.

Найдем теперь ковариационную матрицу оценок $\hat{\theta}$. Имеем

$$\hat{\theta} - \bar{\theta} = (\bar{\varphi}^T \bar{\varphi})^{-1} \bar{\varphi}^T (\varphi \bar{\theta} + \bar{n}) - \bar{\theta} = (\bar{\varphi}^T \bar{\varphi})^{-1} \bar{\varphi}^T [(\varphi - \bar{\varphi}) \bar{\theta} + \bar{n}].$$

Отсюда

$$(\hat{\theta} - \bar{\theta})^T = [\bar{\theta}^T (\varphi - \bar{\varphi})^T + \bar{n}^T] \bar{\varphi} (\bar{\varphi}^T \bar{\varphi})^{(-1)},$$

поэтому матрица ковариаций V оценок $\hat{\theta}$ при фиксированных τ_i имеет вид (усреднение производится при фиксированных τ_i)

$$\begin{aligned} \mathbf{V} = M\{(\hat{\theta} - \bar{\theta})(\hat{\theta} - \bar{\theta})^T\} &= (\bar{\varphi}^T \bar{\varphi})^{-1} \bar{\varphi} M\{\bar{n}\bar{n}^T\} \times \\ &\times \bar{\varphi} (\bar{\varphi}^T \bar{\varphi})^{(-1)} + (\bar{\varphi}^T \bar{\varphi})^{(-1)} \bar{\varphi}^T M\{(\varphi - \bar{\varphi}) \bar{\theta} \bar{\theta}^T \cdot (\varphi - \bar{\varphi})^T\} \bar{\varphi} (\bar{\varphi}^T \bar{\varphi})^{(-1)}, \end{aligned} \quad (5)$$

где учтено, что \bar{n} и t_i независимы и $M\{\bar{n}\} = 0$. А так как $M\{\bar{n}\bar{n}^T\} = \sigma^2 E_N$, то первое слагаемое в (5) равно $\sigma^2 (\bar{\varphi}^T \bar{\varphi})^{(-1)}$.

Для вычисления второго слагаемого надо найти $M\{(\varphi_k(t_i) - \bar{\varphi}_k(\tau_i))(\varphi_l(t_j) - \bar{\varphi}_l(\tau_j))\}$. Для этого рассмотрим два момента измерения: t_i и t_j . Как показано выше, средние значения, дисперсии и ковариации для них равны:

$$\begin{aligned} M\{t_i|N\} &= T \cdot \frac{i}{N+1}; \quad M\{t_j|N\} = T \cdot \frac{j}{N+1}; \\ D\{t_i|N\} &= T^2 \cdot \frac{i(N+1-i)}{(N+1)^2(N+2)}; \quad D\{t_j|N\} = T^2 \cdot \frac{j(N+1-j)}{(N+1)^2(N+2)}; \\ \text{cov}(t_i, t_j) &= \frac{T^2}{(N+1)^2(N+2)} [(N+1) \cdot \min(i, j) - i \cdot j]. \end{aligned}$$

Кроме того, при $N \gg 1$ величины t_i и t_j асимптотически совместно нормальны.

Поскольку $\tau_i = t_i + \xi_i$; $\tau_j = t_j + \xi_j$ и $M\{\xi_i\} = M\{\xi_j\} = 0$, то $M\{\tau_i|N\} = M\{t_i|N\}$; $M\{\tau_j|N\} = M\{t_j|N\}$. Далее, величины ξ_i и ξ_j нормальны, независимы между собой, а также независимы от t_i и t_j , и их дисперсия равна σ_0^2 . Поэтому ковариационная матрица величин t_i, t_j, τ_i и τ_j имеет вид

$$\begin{array}{cccc} & t_i & t_j & \tau_i & \tau_j \\ \begin{array}{c} t_i \\ t_j \\ \tau_i \\ \tau_j \end{array} & \begin{bmatrix} D\{t_i|N\} & \text{cov}(t_i, t_j) & D\{t_i|N\} & \text{cov}(t_i, t_j) \\ \text{cov}(t_i, t_j) & D\{t_j|N\} & \text{cov}(t_i, t_j) & D\{t_j|N\} \\ D\{t_i|N\} & \text{cov}(t_i, t_j) & \sigma_0^2 + D\{t_i|N\} & \text{cov}(t_i, t_j) \\ \text{cov}(t_i, t_j) & D\{t_j|N\} & \text{cov}(t_i, t_j) & \sigma_0^2 + D\{t_j|N\} \end{bmatrix} \end{array}$$

Поскольку эти величины совместно нормальны, условная плотность вероятностей $p(t_i, t_j | \tau_i, \tau_j)$ также является нормальной с математическими ожиданиями

$$\begin{aligned} M\{t_i | \tau_i\} &= T \cdot \frac{i}{N+1} + \frac{D\{t_i | N\}}{\sigma_0^2 + D\{t_i | N\}} \left(\tau_i - T \frac{i}{N+1} \right), \\ M\{t_j | \tau_j\} &= T \cdot \frac{j}{N+1} + \frac{D\{t_j | N\}}{\sigma_0^2 + D\{t_j | N\}} \left(\tau_j - T \frac{j}{N+1} \right) \end{aligned} \quad (6)$$

и условной ковариационной матрицей $\text{cov}(t_i, t_j | \tau_i, \tau_j)$

$$\begin{aligned} & \begin{bmatrix} D\{t_i | N\} & \text{cov}(t_i, t_j) \\ \text{cov}(t_i, t_j) & D\{t_j | N\} \end{bmatrix}^{-1} \begin{bmatrix} D\{t_i | N\} & \text{cov}(t_i, t_j) \\ \text{cov}(t_i, t_j) & D\{t_j | N\} \end{bmatrix} \times \\ & \times \begin{bmatrix} \sigma_0^2 + D\{t_i | N\} & \text{cov}(t_i, t_j) \\ \text{cov}(t_i, t_j) & \sigma_0^2 + D\{t_j | N\} \end{bmatrix}^{(-1)} \begin{bmatrix} D\{t_i | N\} & \text{cov}(t_i, t_j) \\ \text{cov}(t_i, t_j) & D\{t_j | N\} \end{bmatrix}. \end{aligned}$$

Можно выписать явный вид элементов этой матрицы, но он весьма громоздкий. Приближенное выражение для $M\{(\varphi_k(t_i) - \bar{\varphi}_k(\tau_i))(\varphi_l(t_j) - \bar{\varphi}_l(\tau_j))\}$ можно найти, разлагая $(\varphi_k(t_i) - \bar{\varphi}_k(\tau_i))$ в ряд Тейлора около точки $M\{t_i | \tau_i\}$:

$$(\varphi_k(t_i) - \bar{\varphi}_k(\tau_i)) = \varphi'_k(M\{t_i | \tau_i\})(t_i - M\{t_i | \tau_i\}) + \dots$$

Тогда

$$\begin{aligned} & M\{(\varphi_k(t_i) - \bar{\varphi}_k(\tau_i))(\varphi_l(t_j) - \bar{\varphi}_l(\tau_j))\} = \\ & = \varphi'_k(M\{t_i | \tau_i\})\varphi'_l(M\{t_j | \tau_j\}) \cdot \text{cov}(t_i, t_j | \tau_i, \tau_j) \end{aligned}$$

и поэтому для элементов матрицы \mathbf{R}

$$\mathbf{R} = \bar{\varphi}^T M\{(\varphi - \bar{\varphi})\bar{\theta}\bar{\theta}^T \cdot (\varphi - \bar{\varphi})^T\} \bar{\varphi}$$

получим выражение

$$\begin{aligned} R_{mn} &= \sum_{k=1}^s \sum_{l=1}^s \theta_k \theta_l \cdot \sum_{i,j=1}^N \varphi_m(M\{t_i | \tau_i\}) \cdot \varphi'_k(M\{t_i | \tau_i\}) \times \\ & \times \varphi_n(M\{t_j | \tau_j\}) \cdot \varphi'_l(M\{t_j | \tau_j\}) \cdot \text{cov}(t_i, t_j | \tau_i, \tau_j), \end{aligned}$$

где $M\{t_i | \tau_i\}$ и $M\{t_j | \tau_j\}$ выписаны в (6). Явное выражение для R_{mn} очень громоздко, но на ЭВМ может быть вычислено. Сама матрица ковариаций оценок $\hat{\theta}$ приобретает вид

$$\mathbf{V} = (\bar{\varphi}^T \bar{\varphi})^{-1} [\sigma^2 (\bar{\varphi}^T \bar{\varphi}) + \mathbf{R}] (\bar{\varphi}^T \bar{\varphi})^{-1}.$$

Аналогично тому, как это сделано выше, для построения оценки $\hat{\mathbf{V}}$ матрицы ковариаций \mathbf{V} рассмотрим статистику

$$S = \sum_{i=1}^N \varphi_m(\tau_i) \varphi_n(\tau_i) x_i^2.$$

Находя математическое ожидание этой статистики при фиксированных τ_i , получим

$$M\{S | \tau_i\} = \sigma^2 \sum_{i=1}^N \varphi_m(\tau_i) \varphi_n(\tau_i) + \sum_{k=1}^s \sum_{l=1}^s \theta_k \theta_l \cdot \sum_{i=1}^N \bar{\varphi}_m(\tau_i) \cdot \bar{\varphi}_n(\tau_i) \cdot \bar{\varphi}_k(\tau_i) \cdot \bar{\varphi}_l(\tau_i).$$

Первое слагаемое этого выражения образует матрицу $\sigma^2(\bar{\varphi}^T \bar{\varphi})$. Поэтому оценки элементов матрицы $\mathbf{Q} = \sigma^2(\bar{\varphi}^T \bar{\varphi}) + \mathbf{R}$ следует брать в виде

$$\hat{Q}_{mn} = \sum_{i=1}^N \varphi_m(\tau_i) \varphi_n(\tau_i) x_i^2 + \sum_{k=1}^s \sum_{l=1}^s \hat{\theta}_k \hat{\theta}_l \cdot \left[\sum_{i,j=1}^N \varphi_m(M\{t_i|\tau_i\}) \times \right. \\ \times \varphi'_k(M\{t_i|\tau_i\}) \varphi_n(M\{t_i|\tau_i\}) \varphi'_l(M\{t_i|\tau_i\}) \text{cov}(t_i, t_j | \tau_i, \tau_j) - \\ \left. - \sum_{i=1}^N \bar{\varphi}_m(\tau_i) \cdot \bar{\varphi}_n(\tau_i) \cdot \bar{\varphi}_k(\tau_i) \cdot \bar{\varphi}_l(\tau_i) \right],$$

что и позволяет оценить матрицу вариаций \mathbf{V} .

Упрощенные оценки параметров тренда финансовых транзакций

Как видно из предыдущего изложения, предлагаемые оценки $\hat{\theta}_k$ параметров θ_k достаточно сложны в вычислительном отношении и при больших объемах выборки могут возникать дополнительные погрешности, обусловленные ошибками округления, которые особенно сказываются при обращении матрицы $(\bar{\varphi}^T \bar{\varphi})$. Поэтому приведем и исследуем упрощенные оценки $\hat{\theta}_k$ параметров θ_k .

Для построения упрощенных оценок рассмотрим статистику вида

$$S_m = \frac{1}{\lambda T} \sum_{i=1}^N \varphi_m(\tau_i) x_i. \quad (7)$$

Усредняя по n_i при фиксированном t_i , получим

$$M\{S_m | \{t_i\}\} = \frac{1}{\lambda T} \sum_{i=1}^N \sum_{k=1}^s \varphi_m(t_i + \xi_i) \varphi_k(t_i) \theta_k.$$

Усредняя теперь по t_i и ξ_i с учетом того, что ξ_i не зависит ни от t_i , ни от n_i , получим

$$M\{S_m\} = \sum_{k=1}^s \frac{1}{T} \int_0^T du \int_{-\infty}^{\infty} d\xi \varphi_m(u + \xi) \varphi_k(u) \theta_k.$$

Обозначим $\tilde{\Phi}$ матрицу с элементами

$$\tilde{\Phi}_{mk} = \frac{1}{T} \int_0^T du \int_{-\infty}^{\infty} \varphi_m(u + \xi) \varphi_k(u) d\xi.$$

Тогда

$$M\{S_m\} = \sum_{k=1}^s \tilde{\Phi}_{mk} \theta_k$$

или в матричном виде $M\{\bar{S}\} = \tilde{\Phi} \bar{\theta}$, где $\bar{S} = [S_1, S_2, \dots, S_s]^T$. Исходя из этого соотношения, естественно брать оценки $\hat{\theta}$ неизвестных параметров $\bar{\theta}$ в виде

$$\hat{\theta} = \tilde{\Phi}^{(-1)} \bar{S} = \tilde{\Phi}^{(-1)} \cdot \left\| \frac{1}{\lambda T} \sum_{i=1}^N \varphi_m(t_i) x_i \right\|, \quad (8)$$

где \bar{S} — вектор-столбец с элементами (7).

Прежде чем исследовать свойства этих оценок, приведем пример матриц $\tilde{\Phi}$ и $\tilde{\Phi}^{(-1)}$. Пусть $\varphi_k(t) = (t/T)^k$, $k = 0, 1, 2, 3$, а ξ_i — нормальные случайные величины с $M\{\xi_i\} = 0$ и $D\{\xi_i\} = \sigma_0^2$. Обозначив $s_0 = \sigma_0^2/T^2$, приведем явный вид элементов Φ_{mk} (таблица).

Таблица

$k \backslash m$	0	1	2	3	4
0	1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{5}$
1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{5}$	$\frac{1}{6}$
2	$\frac{1}{3} + s_0$	$\frac{1}{4} + \frac{s_0}{2}$	$\frac{1}{5} + \frac{s_0}{3}$	$\frac{1}{6} + \frac{s_0}{4}$	$\frac{1}{7} + \frac{s_0}{5}$
3	$\frac{1}{4} + \frac{3s_0}{2}$	$\frac{1}{5} + s_0$	$\frac{1}{6} + \frac{3s_0}{4}$	$\frac{1}{7} + \frac{3s_0}{5}$	$\frac{1}{8} + \frac{s_0}{2}$
4	$\frac{1}{5} + 2s_0 + 3s_0^2$	$\frac{1}{6} + \frac{3s_0}{2} + \frac{3s_0^2}{2}$	$\frac{1}{7} + \frac{6s_0}{5} + s_0^2$	$\frac{1}{8} + s_0 + \frac{3s_0^2}{4}$	$\frac{1}{9} + \frac{6s_0}{7} + \frac{3s_0^2}{5}$

Приведем матрицы $\tilde{\Phi}^{(-1)}$ для некоторых значений s_0 , полученные с помощью программы MATHCAD v.6.0:

$$s = 2 \quad \tilde{\Phi}^{(-1)} = \begin{bmatrix} 4 & -6 \\ -6 & 12 \end{bmatrix},$$

$(k=0,1)$

$$s = 3 \quad \tilde{\Phi}^{(-1)} = \begin{bmatrix} 9 - 30s_0 & -36 & 30 \\ -36 + 180s_0 & 192 & -180 \\ 30 - 180s_0 & -180 & 180 \end{bmatrix},$$

$(k=0,1,2)$

$$s = 4 \quad \tilde{\Phi}^{(-1)} = \begin{bmatrix} -16 + 240s_0 & -120 + 420s_0 & 240 & -140 \\ -120 + 2700s_0 & 1200 - 5040s_0 & -2700 & 1680 \\ 240 - 6480s_0 & -2700 + 12600s_0 & 6480 & -4200 \\ -140 + 4200s_0 & 1680 - 8400s_0 & -4200 & 2800 \end{bmatrix},$$

$(k=0,1,2,3)$

$$s = 5 \quad \tilde{\Phi}^{(-1)} = \begin{bmatrix} 25 - 1050s_0 + 1890s_0^2 & -300 + 4200s_0 \\ -300 + 18900s_0 - 37800s_0^2 & 4800 - 80640s_0 \\ 1050 - 79380 + 170100s_0^2 & -18900 + 352800s_0 \\ -1400 + 117600s_0 - 264600s_0^2 & 26880 - 537600s_0 \\ 630 - 56700s_0 + 132300s_0^2 & -12600 + 264600s_0 \end{bmatrix},$$

$(k=0,1,2,3,4)$

$$\begin{bmatrix} 1050 - 3780s_0 & -1400 & 630 \\ 118900 + 75600s_0 & 26880 & -12600 \\ 79380 - 340200s_0 & -117600 & 56700 \\ -117600 + 529200s_0 & 179200 & -88200 \\ 56700 - 264600s_0 & -88200 & 44100 \end{bmatrix}.$$

Исследование свойств упрощенных оценок

Исследуем теперь свойства предлагаемых упрощенных оценок. Для краткости формул введем обозначение

$$\overline{\varphi}_k^r(t) = \int_{-\infty}^{\infty} \varphi_k^r(t + \xi) p(\xi) d\xi.$$

Прежде всего отметим, что предлагаемые оценки (8) несмещенные, поскольку именно из этого требования они и строились.

Для общности рассмотрим свойства статистик вида

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

где x_i может быть представлено $x_i = x(t_i) + n_i$. Тогда легко получить, что

$$M\{S\} = \bar{S} = \frac{1}{T} \int_0^T f(u) \overline{\varphi}(u) du. \quad (9)$$

Дальнейшее изложение оформим в виде теорем.

Теорема 1. При $\lambda \rightarrow \infty$ S сходится к \bar{S} в среднеквадратичном смысле.

Доказательство. Имеем

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

Усредняя по n_i при фиксированных $\{t_i\}$ и $\{\xi_i\}$, получаем

$$M_n\{S^2\} = \frac{1}{(\lambda T)^2} \left[\sigma^2 \sum_{i=1}^N \varphi^2(t_i + \xi_i) + \sum_{i=1}^N \sum_{j=1}^N \varphi(t_i + \xi_i) \varphi(t_j + \xi_j) f(t_i) f(t_j) \right].$$

Усредняя по t_i и ξ_i , получаем

$$M\{S^2\} = \frac{\sigma^2}{\lambda T} \cdot \frac{1}{T} \int_0^T \overline{\varphi^2}(t) dt + \frac{1}{\lambda T} \cdot \frac{1}{T} \int_0^T f^2(t) \overline{\varphi^2}(t) dt + \left(\frac{1}{T} \int_0^T f(t) \overline{\varphi}(t) dt \right)^2,$$

отсюда

$$D\{S\} = \frac{1}{\lambda T} \left[\sigma^2 \frac{1}{T} \int_0^T \overline{\varphi^2}(t) dt + \frac{1}{T} \int_0^T f^2(t) \overline{\varphi^2}(t) dt \right]. \quad (10)$$

Из выражения (9) видно, что $D\{S\}$ убывает как $\frac{1}{\lambda T}$ и при $\lambda \rightarrow \infty$ $D\{S\} \rightarrow 0$.

Это и означает, что $S \xrightarrow{rms} \bar{S}$ при $\lambda \rightarrow \infty$.

Теорема доказана.

Из этого следует, что при $\lambda \rightarrow \infty$ оценки $\hat{\theta}$ в виде (10) сходятся к истинным значениям параметров $\bar{\theta}$, по крайней мере, в среднеквадратичном смысле.

Теорема 2. При $\lambda \rightarrow \infty$ S сходятся к \bar{S} почти наверное (п.н.).

Доказательство. Представим S в виде $S = S_1 + S_2$, где $S_1 = \frac{1}{\lambda T} \times \sum_{i=1}^N \varphi(t_i + \xi_i) n_i$, $S_2 = \frac{1}{\lambda T} \sum_{i=1}^N \varphi(t_i + \xi_i) f(t_i)$.

Очевидно, что $M\{S_1\} = 0$. Далее

$$M\{S_1^4\} = \frac{1}{(\lambda T)^4} M \left\{ \sum_{i,j,k,l=1}^N \varphi(t_i + \xi_i) \varphi(t_j + \xi_j) \varphi(t_k + \xi_k) \varphi(t_l + \xi_l) n_i n_j n_k n_l \right\} =$$

$$= \frac{m_4}{(\lambda T)^3} \cdot \frac{1}{T} \int_0^T \bar{\varphi}^4(t) dt + \frac{3\sigma^4}{(\lambda T)^2} \left(\frac{1}{T} \int_0^T \bar{\varphi}^2(t) dt \right)^2,$$

где $m_4 = M\{n_i^4\}$. Так как $M\{S_1^4\}$ убывает не медленнее, чем $\frac{1}{\lambda^2}$, то отсюда, как и в предыдущих теоремах, следует, что при $\lambda \rightarrow \infty$ $S_1 \xrightarrow{\text{i.f.}} 0$, по крайней мере, для последовательности $\lambda_n = n\lambda_0$.

Введем далее обозначение $A_k = \frac{1}{T} \int_0^T f^k(t) \bar{\varphi}^k(t) dt$. Тогда легко получить, что

$$M\{S_2\} = A_1; \quad M\{S_2^2\} = A_1^2 + \frac{A_2}{\lambda T}; \quad M\{S_2^3\} = A_1^3 + \frac{3}{\lambda T} A_2 A_1 + \frac{1}{(\lambda T)^2} A_3;$$

$$M\{S_2^4\} = A_1^4 + \frac{6}{\lambda T} A_2 A_1^2 + \frac{4}{(\lambda T)^2} A_3 A_1 + \frac{3}{(\lambda T)^2} A_2^2 + \frac{A_4}{(\lambda T)^3},$$

отсюда следует, что

$$M\{(S_2 - A_1)^4\} = \frac{3}{(\lambda T)^2} A_2^2 + \frac{A_4}{(\lambda T)^3},$$

т.е. $M\{(S_2 - A_1)^4\}$ убывает также не медленнее, чем $\frac{1}{\lambda^2}$ и поэтому при $\lambda \rightarrow \infty$ $S_2 \xrightarrow{\text{i.f.}} A_1 = M\{S\}$.

Объединяя эти результаты, получаем, что при $\lambda \rightarrow \infty$ $S \xrightarrow{\text{п.н.}} \bar{S}$, что и требовалось доказать.

Отсюда следует, что упрощенные оценки (8) при $\lambda \rightarrow \infty$ сходятся к истинным значениям параметров θ почти наверное.

Теорема 3. При $\lambda \rightarrow \infty$ статистика

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

сходится по распределению к нормальной случайной величине с нулевым математическим ожиданием и дисперсией $\sigma^2 \int_0^T \bar{\varphi}^2(u) du + \int_0^T f^2(u) \bar{\varphi}^2(u) du$.

Доказательство. Доказательство этой теоремы аналогично доказательству теоремы 2. Рассмотрим статистику S_t вида

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

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

Дальнейшие рассуждения близки к рассуждениям теоремы 2 [6], отличаясь от них лишь необходимостью усреднения по величинам ξ_i . Так, вместо (9) будем иметь

$$g(\omega, t) = g(\omega, t + \Delta t) \left[(1 - \lambda \Delta t) \cdot \exp(-i\omega \sqrt{\lambda} f(t) \bar{\varphi}(t) \Delta t) + \lambda \Delta t M_{n, \xi} \left\{ \exp \left(i\omega \frac{1}{\sqrt{\lambda}} \varphi(t + \xi) f(t) + i\omega \frac{1}{\sqrt{\lambda}} \varphi(t + \xi) n + O(\Delta t) \right) \right\} \right] + o(\Delta t).$$

После разложения экспонент в ряд Тейлора и усреднения по величинам n и ξ , получим

$$g(\omega, t) = g(\omega, t) + \Delta t \left[-\lambda g(\omega, t) + \frac{\partial g(\omega, t)}{\partial t} - i\sqrt{\lambda} f(t) \bar{\varphi}(t) \cdot g(\omega, t) + \lambda g(\omega, t) + i\omega \sqrt{\lambda} f(t) \bar{\varphi}(t) g(\omega, t) - \frac{\omega^2}{2} g(\omega, t) (\sigma^2 \bar{\varphi}^2(t) + f^2(t) \bar{\varphi}^2(t)) \right] + o(\Delta t)$$

и после сокращения $g(\omega, t)$, деления на Δt и предельного перехода $\Delta t \rightarrow 0$ будем иметь уравнение, аналогичное уравнению, полученному в [6]:

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

Дальнейшие рассуждения буквально повторяют рассуждения теоремы 2, приведенной в [6], и поэтому здесь не приводятся.

Аналогично можно доказать, что совокупность статистик S , отличающихся одна от другой функциями $\varphi(\tau_i)$, при $\lambda \rightarrow \infty$ асимптотически совместно нормальные.

Отсюда следует, что при $\lambda \rightarrow \infty$ построенные нами упрощенные оценки асимптотически нормальные.

Найдем теперь явное выражение для матрицы ковариаций упрощенных оценок. Для упрощения получаемых формул введем обозначение

$$\bar{\varphi}_{mm}(t) = \int_{-\infty}^{\infty} \varphi_m(t + \xi) \varphi_n(t + \xi) p(\xi) d\xi.$$

Имеем

$$\hat{\bar{\theta}} = \tilde{\Phi}^{(-1)} \cdot \left\| \frac{1}{\lambda T} \sum_{i=1}^N \varphi_m(\tau_i) x_i \right\|.$$

Так как $\bar{\theta} = \tilde{\Phi}^{(-1)} \cdot \tilde{\Phi} \cdot \theta$, то

$$\hat{\bar{\theta}} - \bar{\theta} = \tilde{\Phi}^{(-1)} \cdot \left\| \frac{1}{\lambda T} \sum_{i=1}^N \varphi_m(\tau_i) x_i - \sum_{k=1}^s \hat{\varphi}_{mk} \theta_k \right\|$$

и поэтому

$$\mathbf{V} = M\{(\hat{\bar{\theta}} - \bar{\theta})(\hat{\bar{\theta}} - \bar{\theta})^T\} = \frac{1}{\lambda T} \tilde{\Phi}^{(-1)} \cdot \mathbf{B} \cdot \tilde{\Phi},$$

где \mathbf{B} — матрица

$$\mathbf{B} = \lambda T \cdot M \left\{ \left\| \frac{1}{\lambda T} \sum_{i=1}^N \varphi_m(\tau_i) x_i - \sum_{k=1}^s \theta_k \tilde{\varphi}_{mk} \right\| \cdot \left\| \frac{1}{\lambda T} \sum_{j=1}^N \varphi_n(\tau_j) x_j - \sum_{l=1}^s \theta_l \tilde{\varphi}_{nl} \right\|^T \right\}.$$

Найдем явный вид элементов B_{mn} матрицы \mathbf{B} . Так как

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

то

$$B_{mn} = \lambda T \cdot M \left\{ \left[\sum_{k=1}^s \theta_k \left(\frac{1}{\lambda T} \sum_{i=1}^N \varphi_m(\tau_i) \varphi_k(t_i) - \tilde{\Phi}_{mk} \right) + \frac{1}{\lambda T} \sum_{i=1}^N \varphi_m(\tau_i) n_i \right] \times \right. \\ \left. \times \left[\sum_{l=1}^s \theta_l \left(\frac{1}{\lambda T} \sum_{j=1}^N \varphi_n(\tau_j) \varphi_l(t_j) - \tilde{\Phi}_{nl} \right) + \frac{1}{\lambda T} \sum_{j=1}^N \varphi_n(\tau_j) n_j \right] \right\}.$$

После усреднения по n_i , получаем

$$B_{mn} = \lambda T \cdot M \left\{ \frac{\sigma^2}{(\lambda T)^2} \sum_{i=1}^N \varphi_m(\tau_i) \varphi_n(\tau_i) \right\} + \lambda T \sum_{k=1}^s \sum_{l=1}^s \theta_k \theta_l \times \\ \times M \left\{ \left(\frac{1}{\lambda T} \sum_{i=1}^N \varphi_m(\tau_i) \varphi_k(t_i) - \tilde{\Phi}_{mk} \right) \cdot \left(\frac{1}{\lambda T} \sum_{j=1}^N \varphi_n(\tau_j) \varphi_l(t_j) - \tilde{\Phi}_{nl} \right) \right\}.$$

Усредняя теперь сначала по ξ_i , а затем по t_i , получаем

$$B_{mn} = \sigma^2 \frac{1}{T} \int_0^T \bar{\varphi}_{mn}(t) dt + \sum_{k=1}^s \sum_{l=1}^s \theta_k \theta_l \cdot \frac{1}{T} \int_0^T \bar{\varphi}_{mn}(t) \varphi_k(t) \varphi_l(t) dt,$$

что и дает явный вид элементов матрицы \mathbf{B} , а с ним и явный вид матрицы ковариаций

$$\mathbf{V} = \frac{1}{\lambda T} \tilde{\Phi}^{(-1)} \cdot \mathbf{B} \cdot \tilde{\Phi}.$$

Как обычно, в силу того, что σ^2 и $\bar{\theta}$ нам неизвестны, надо заменить \mathbf{B} ее оценкой $\hat{\mathbf{B}}$. В качестве оценок \hat{B}_{mn} элементов матрицы B_{mn} можно брать статистики вида

$$\hat{B}_{mn} = \frac{1}{\lambda T} \sum_{i=1}^N \varphi_m(\tau_i) \varphi_n(\tau_i) \cdot x_i^2. \quad (11)$$

Действительно, подставляя выражение для x_i , получаем

$$\hat{B}_{mn} = \frac{1}{\lambda T} \sum_{i=1}^N \varphi_m(\tau_i) \varphi_n(\tau_i) \cdot \left[\sum_{k=1}^s \theta_k \varphi_k(t_i) + n_i \right]^2,$$

отсюда после усреднения по n_i , ξ_i и t_i имеем

$$M\{\hat{B}_{mn}\} = \sigma^2 \frac{1}{T} \int_0^T \bar{\varphi}_{mn}(u) du + \sum_{k=1}^s \sum_{l=1}^s \theta_k \theta_l \cdot \frac{1}{T} \int_0^T \varphi_{mn}(u) \varphi_k(u) \varphi_l(u) du,$$

т.е. $M\{\hat{B}_{mn}\} = B_{mn}$ и (11) — несмещенная оценка величин B_{mn} .

Совершенно аналогично доказательству теоремы 3 и 4 (см. [6]) покажем, что при $\lambda \rightarrow \infty$ величины \hat{B}_{mn} сходятся к B_{mn} в среднеквадратичном смысле и почти

наверное. Доказательство этих утверждений не приводится, поскольку не содержит принципиально новых моментов и отличается от доказательств теорем 3 и 4 [6] лишь дополнительным усреднением по величинам ξ_i . Поэтому в качестве оценки \hat{V} матрицы ковариаций V можно использовать матрицу

$$\hat{V} = \frac{1}{\lambda T} \tilde{\Phi}^{(-1)} \cdot \hat{R} \cdot \tilde{\Phi}^{(-1)} = \frac{1}{(\lambda T)^2} \tilde{\Phi}^{(-1)} \left\| \sum_{i=1}^N \varphi_m(t_i) \varphi_n(t_i) \cdot x_i^2 \right\| \tilde{\Phi}^{(-1)}.$$

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

Имитационное моделирование

В данном разделе продолжается описание результатов, полученных в работах [6, 7] и, следовательно, постановочно стыкуется с ними (рис. 1–3). Взята та же реализация с $T=100, \lambda=1$, так что $\lambda T=100$. На моменты измерений t_i «набрасывались» ошибки ξ_i , поэтому известными считались $\tau_i = t_i + \xi_i$, где ξ_i были независимыми нормальными величинами с $M\{\xi_i\} = 0$ и $D\{\xi_i\} = \sigma_0^2$. Кривая B — данные моделирования x_i , соединенные отрезками прямых, прямая C — истинный тренд, прямая D — тренд, выделенный по точному алгоритму, прямая E — тренд, выделенный по приближенному алгоритму. Истинный тренд

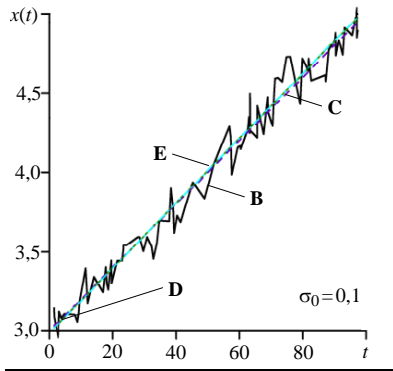


Рис. 1

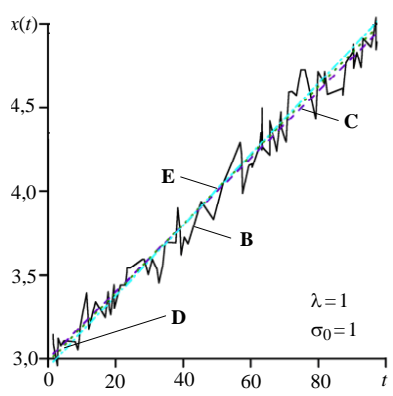


Рис. 2

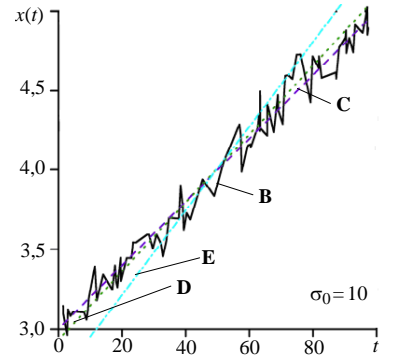


Рис. 3

$$x(t) = 3 + 2 \frac{t}{T}.$$

Рис. 1 соответствует $\sigma_0 = 0,1$. Как видно, выделение тренда здесь очень хорошее по обоим алгоритмам. Хотя доверительные интервалы и не нарисованы, но истинное значение тренда лежит в 95 % доверительном интервале.

Рис. 2 соответствует $\sigma_0 = 1$. Здесь совпадение хуже, но истинный тренд все еще находится в 95 % доверительном интервале.

Наконец, рис. 3 соответствует $\sigma_0 = 10$, когда ошибки в измерениях моментов времени очень велики. Но и в этом случае точный алгоритм достаточно хорошо выделяет тренд, тогда как приближенный алгоритм дает оценку тренда, весьма далекую от истинной.

Насколько можно судить по результатам имитационного моделирования, упрощенный алгоритм вполне удовлетворительно выделяет тренд в виде полинома первого порядка при $\lambda \sigma_0 \leq 1$. Что касается точного алгоритма, он достаточно хорошо выделяет тренд для всех значений $\sigma_0 \in [0, +\infty)$. Это связано с тем, что при $\sigma_0 \rightarrow \infty$ он переходит в алгоритм с полностью неизвестными t_i , когда известен лишь их порядок, и он так же хорошо выделяет линейные тренды.

Что касается выделения трендов в виде полиномов более высоких порядков, то о них верно все, сказанное в работах [6, 7]. Асимптотическая нормальность оценок линейного тренда достигается при λT порядка 100, причем эта граница увеличивается с ростом степени полинома. И поэтому не рекомендуется выделять тренды выше третьего порядка при небольших ($\lambda T < 1000$) объемах выборки, так как при этом возникают большие ошибки.

Ф.Ф. Ідрісов

НАБЛИЖЕНІ МОДЕЛІ ФІНАНСОВОЇ РОЗВІДКИ ПРИ НЕТОЧНО ЗАДАНИХ МОМЕНТАХ ЧАСУ ЗДІЙСНЕННЯ ФІНАНСОВИХ ТРАНЗАКЦІЙ

Розглянуто імовірнісну постановку питань виявлення та оцінювання фінансових потоків, що надсилаються в кримінальні і терористичні угруповання. Припускається, що моменти реалізації фінансових транзакцій відомі неточно. Досліджено питання статистичної коректності розглянутих моделей. Наведено результати імітаційного моделювання.

F.F. Idrisov

APPROXIMATE MODELS OF FINANCIAL INTELLIGENCE AT INACCURATELY DEFINED TIME INSTANTS FOR PERFORMING FINANCIAL TRANSACTIONS

The probabilistic formulation of the issues of detecting and evaluating financial flows directed to criminal and terrorist groups is discussed. It is assumed that the moments of the implementation of financial transactions are not known precisely. The questions of statistical correctness of the models presented are investigated. The results of simulation modeling are given.

1. Мелкумян К.С. ФАТФ в противодействии финансированию терроризма // Вестник МГИМО. — 2014. — 34, № 1. — С. 88–96.
2. Предотвращение отмывания денег и финансирования терроризма. Практическое руководство для банковских специалистов / П.-Л. Шатен, Дж. Макдауэл, С. Муссе, П.А. Шотт, Эмиль ван дер Дус де Вильбуа. — М. : Альбина Паблишер, 2011. — 316 с.
3. Кириленко В.С. Возможности противодействия отмыванию доходов, полученных преступным путем // Science Time. — 2014. — № 6. — С. 102–107.
4. Резников А.В. Модель финансового мониторинга, основанная на знаниях // Экономические науки. 3. Финансовые отношения. — 2015. — № 7. — С. 84–92.
5. Светлов В.А. Конфликт: модели, решения, менеджмент. 4-е изд. — СПб. : Питер, 2005. — 540 с.
6. Ідрісов Ф.Ф. Приближенные алгоритмы выделения трендов в задачах финансовой разведки. Часть 1. Моменты появления элементов финансового потока известны точно // Международный научно-технический журнал «Проблемы управления и информатики». — 2017. — № 6. — С. 7–18.
7. Ідрісов Ф.Ф. Приближенные алгоритмы выделения трендов в задачах финансовой разведки. Часть 2. Моменты появления элементов финансового потока неизвестны // Там же. — 2018. — № 1. — С. 146–155.
8. Андерсон Т. Введение в многомерный статистический анализ. — М. : Физматгиздат, 1963. — 500 с.
9. Leneman O.A.Z., Lewis J.P. Random sampling of random processes: mean square comparison of various interpolators // IEEE Trans. Automat. Control. — 1966. — 11. — P. 396–403.

Получено 18.12.2017
После доработки 12.07.2018