# ПРИМЕНЕНИЕ МЕТОДОВ ПАРАМЕТРИЧЕСКОЙ ИДЕНТИФИКАЦИИ В КОИНТЕГРАЦИОННОМ АНАЛИЗЕ.

Народицкая Н.А., Подладчиков В.Н.
Институт прикладного системного анализа, УНК "ИПСА".

В последнее время большой интерес вызывает решение задач анализа и прогнозирования нестационарных финансовых рядов. В мировой литературе для обработки нестационарных временных рядов рекомендуется применять аппарат коинтеграционного анализа [1,2], который был разработан в ответ на растущую потребность в анализе соотношений между группами экономических показателей-переменных, чтобы получать концептуально и эмпирически более значимое измерение этих взаимосвязей в свете нестационарности отдельных временных рядов. Одной из основных задач коинтеграционного анализа является оценка константы коинтеграции, которая определяет характер долгосрочной зависимости между исследуемыми рядами. Для решения этой задачи наибольшее применение нашел метод наименьших квадратов (МНК). Но применение МНК требует выполнения гипотез, лежащих в его основе, нарушение которых приводит к смещенности оценок. В работе, впервые в коинтеграционном анализе, предложено применение методов идентификации статистических характеристик действующих шумов модели для повышения точности оценки константы коинтеграции и разработаны новые методы и приемы, позволяющие получить несмещенные оценки константы коинтеграции, которые рекомендуется применять в условиях высокой смещенности оценок МНК.

В работе предлагается использовать приближенную оценку смещения МНК, выведенную для одномерного случая. Общий случай оценки смещения МНК для случайных регрессоров описан в работе Себера[3].

Рассмотрим временные ряды с единичными корнями, которые наблюдаются в присутствии случайного шума

$$y(k) = \mu_y(k) + \delta_y(k) \qquad (1)$$

$$x(k) = \mu_x(k) + \delta_x(k) \qquad (2)$$

где $E[\delta_y(k)] = E[\delta_x(k)] = 0$, $E[\delta_y(k), \delta_x(k)] = 0$,

$$E[\delta_y(k), \delta_y(k-j)] = \begin{cases} \sigma_y^2, j = k \\ 0, j \neq k \end{cases}, \quad E[\delta_x(k), \delta_x(k-j)] = \begin{cases} \sigma_x^2, j = k \\ 0, j \neq k \end{cases}.$$

и имеют стохастические тренды,

$$\mu_x(k+1) = \mu_x(k) + \varepsilon_x(k) \tag{3}$$

$$\mu_y(k+1) = \mu_y(k) + \varepsilon_y(k) \tag{4}$$

связанные соотношением(5):

$$\mu_y(k+1) = \alpha + \lambda \mu_x(k), \tag{5}$$

Тогда между ними существует коинтеграционное соотношение, описываемое следующими параметрами

$$\hat{\alpha}(k) = \bar{y}(k) - \hat{\lambda}(k)\bar{x}(k) \tag{6}$$

$$\hat{\lambda}(k) = \frac{\sum_{t=1}^{T}(x(k) - \bar{x}(k))(y(k) - \bar{y}(k))}{\sum_{t=1}^{T}(x(k) - \bar{x}(k))^2} \tag{7}$$

где

$$\bar{y}(k) = \frac{1}{T}\sum_{t=1}^{T} y(k), \ \bar{x}(k) = \frac{1}{T}\sum_{t=1}^{T} x(k)$$

Тогда остаток

$$\hat{u}(k) = y(k) - \hat{\lambda}\, x(k) - \hat{\alpha} = (\bar{y}(k) - y(k)) - \hat{\lambda}(\bar{x}(k) - x(k))$$

$$= (\mu_y(k) + \delta_y(k) - \bar{\mu}_y - \bar{\delta}_y) - \hat{\lambda}(\mu_x(k) + \delta_x(k) - \bar{\mu}_x - \bar{\delta}_x)$$

где $\bar{\mu}_y = \frac{1}{T}\sum_{t=1}^{T}\mu_y(k)$, $\bar{\mu}_x = \frac{1}{T}\sum_{t=1}^{T}\mu_x(k)$,

$$\bar{\delta}_y = \frac{1}{T}\sum_{t=1}^{T}\delta_y(k), \bar{\delta}_x = \frac{1}{T}\sum_{t=1}^{T}\delta_x(k)$$

Обозначим $\tilde{\mu}_y(k) = \mu_y(k) - \bar{\mu}_y$, $\tilde{\mu}_x(k) = \mu_x(k) - \bar{\mu}_x$,

$$\tilde{\delta}_y(k) = \delta_y(k) - \bar{\delta}_y, \tilde{\delta}_x(k) = \delta_x(k) - \bar{\delta}_x$$

Тогда остаток $\hat{u}(k)$ можно представить в виде

$$\hat{u}(k) = \tilde{\mu}_y(k) + \tilde{\delta}_y(k) - \hat{\lambda}(k)(\tilde{\mu}_x(k) + \tilde{\delta}_x(k))$$

Учитывая, что $\tilde{\mu}_y(k) = \lambda\, \tilde{\mu}_x(k)$, получим

$$\hat{u}(k) = \lambda\, \tilde{\mu}_x(k) + \tilde{\delta}_y(k) - \hat{\lambda}(k)(\tilde{\mu}_x(k) + \tilde{\delta}_x(k))$$

$$\hat{u}(k) = (\lambda - \hat{\lambda}(k))\, \tilde{\mu}_x(k) + \tilde{\delta}_y(k) - \hat{\lambda}(k)\, \tilde{\delta}_x(k)$$

При применении МНК для оценивания константы коинтеграции минимизируется сумма квадратов ошибок. Следовательно, процедура оценивания такова, что минимизируется дисперсия ошибки.

Дисперсия остатка $\hat{u}(k)$ равна

$$\sigma^2_{\hat{u}(k)} = (\lambda - \hat{\lambda}(k))^2 \sigma^2_{\tilde{\mu}_x(k)} + \sigma^2_{\tilde{\delta}_y(k)} + \hat{\lambda}^2(k)\sigma^2_{\tilde{\delta}_x(k)} \quad ,$$

где $\sigma^2_{\mu_x}$ дисперсия тренда $\mu_x(t)$

Если бы значения дисперсий $\sigma^2_{\mu_x}, \sigma^2_{\delta_x}, \sigma^2_{\delta_y}$ были априорно известны, то оценку МНК $\hat{\lambda}(k)$ константы $\lambda$ можно было бы получить из уравнения

$$\frac{d\sigma^2_{u(k)}}{d\,\hat{\lambda}(k)} = 0 \quad \text{или} \quad 2(\lambda - \hat{\lambda}(k)\sigma^2_{\mu_x} + 2\,\hat{\lambda}(k)\sigma^2_{\delta_x} = 0$$

Откуда смещение можно оценить по формуле

$$C\lambda = \frac{\sigma^2_{\mu_x}}{\sigma^2_{\mu_x} + \sigma^2_{\delta_x}}\lambda \tag{8}$$

Выражение (8) удобно использовать для качественного анализа оценок МНК константы коинтеграции $\lambda$.

При $\sigma^2_{\delta_x} > 0$ , коэффициент С < 1.Отсюда следует, что сама идея минимизации дисперсии ошибки оценки $\lambda$ при наблюдаемых в присутствии шумов значений ряда x(k) обуславливает смещение оценки $\lambda$ в сторону уменьшения модуля его значения.

С ростом дисперсии $\sigma^2_{\delta_x}$ ошибок наблюдений коэффициент С убывает и, следовательно, смещение возрастает. Если дисперсия $\sigma^2_{\mu_x}$ тренда $\mu_x(k)$ увеличивается, то возрастает коэффициент С, а, следовательно, смещение уменьшается. При одновременном увеличении $\sigma^2_{\delta_x}$ и $\sigma^2_{\mu_x}$ в одно и тоже число раз смещение не изменится.

Так как тренд $\mu_x(k)$ изменяется в соответствии с моделью случайного блуждания(3), то $\sigma^2_{\mu_x} = T\sigma^2_{\varepsilon_x}$, где $\sigma^2_{\varepsilon_x}$ дисперсия шума $\varepsilon_x(t)$. При $T \to \infty$ дисперсия $\sigma^2_{\mu_x} \to \infty$ и коэффициент $C \to 1$. Следовательно, асимптотическое возрастание дисперсии трендов временных рядов с единичным корнем обуславливает асимптотическую несмещенность оценок МНК константы коинтеграции.

Таким образом, оценка $\hat{\lambda}(k)$ константы коинтеграции $\lambda$ смещенная. Смещение оценки пропорционально истинному значению $\lambda$, увеличивается с ростом дисперсии шума наблюдений и уменьшается с ростом дисперсии временного ряда.

Несмотря на смещения оценок, МНК остается самым распространенным и наиболее используемым методом для оценки константы коинтеграции. Поэтому в работе ставилась задача определить зону высоких смещений МНК, при неизвестных параметрах шумов. Для решения этой проблемы были выбраны широко известные методы статистической идентификации, обладающие свойством сходимости. Эти методы описаны в работах Мехра[4], Подладчикова[5], Борна-Тейпли (алгоритмы, основаны на анализе невязок субоптимального фильтра) и Андерсона (алгоритм использует мультипликативную функцию от измерений).

Если в результате применения методов параметрической идентификации к модели (1)-(4) было установлено, что мы находимся в зоне высоких смещений оценок МНК, то для повышения точности оценки предлагается использовать следующие методы и приемы:

1. Оценка константы коинтеграции на основе фильтрации рядов.

Если априорно известно, или рассчитано с помощью методов идентификации, что интенсивность шума измерения выше интенсивности шума состояния, т. е. $\sigma^2_{\varepsilon_x} < \sigma^2_{\delta_x} \quad \sigma^2_{\varepsilon_y} < \sigma^2_{\delta_y}$, то целесообразно отфильтровать данные с помощью дискретного фильтра Калмана и, используя полученные оценки, рассчитать константу коинтеграции на основе МНК

$$\hat{\mu}_y(k) = \lambda * \hat{\mu}_x(k) + \nu(k) \qquad (9)$$

2. Оценка константы коинтеграции на основе обратной регрессии.

Если $\sigma^2_{\varepsilon_x} < \sigma^2_{\delta_x}, \sigma^2_{\varepsilon_y} > \sigma^2_{\delta_y}$, то предлагается использовать обратную регрессию для построения коинтеграционного соотношения

$$x(k) = \frac{1}{\lambda} * y(k) + w(k) \qquad (10)$$

3. Оценка константы коинтеграции на основе соотношения.

$$\mathrm{E}\left[\mathrm{y(k)y^T(k-1)}\right] = \lambda * \mathrm{E}\left[\mathrm{y(k)x^T(k-1)}\right]$$

Справедливость данного соотношения легко показать исходя из способа построения ряда y(k). Оценку константы коинтеграции предлагается производить по формуле

$$\lambda(n) = \left[\sum_{k=3}^{n} y(k)y(k-1)\right]\left[\sum_{k=3}^{n} y(k)x(k-1)\right]^{-1} \qquad (11)$$

Проведем сравнительный анализ указанных методов, на основе имитационного моделирования. Статистические характеристики шумов состояния и измерения [таблица 1] были выбраны таким образом, что бы исследовать поведения методов в зоне высоких смещений оценки МНК.

Таблица 1.

| Параметры | $\lambda$ | $\sigma^2_{\varepsilon_x}$ | $\sigma^2_{\delta_x}$ | $\sigma^2_{\varepsilon_y}=\lambda$ | $\sigma^2_{\delta_y}$ |
|---|---|---|---|---|---|
| Значения | 4 | 1 | 25 | 16 | 9 |

На графике рис. 1 представлены зависимости среднеквадратичной ошибки оценки константы коинтеграции от времени, полученные с помощью описанных выше методов.

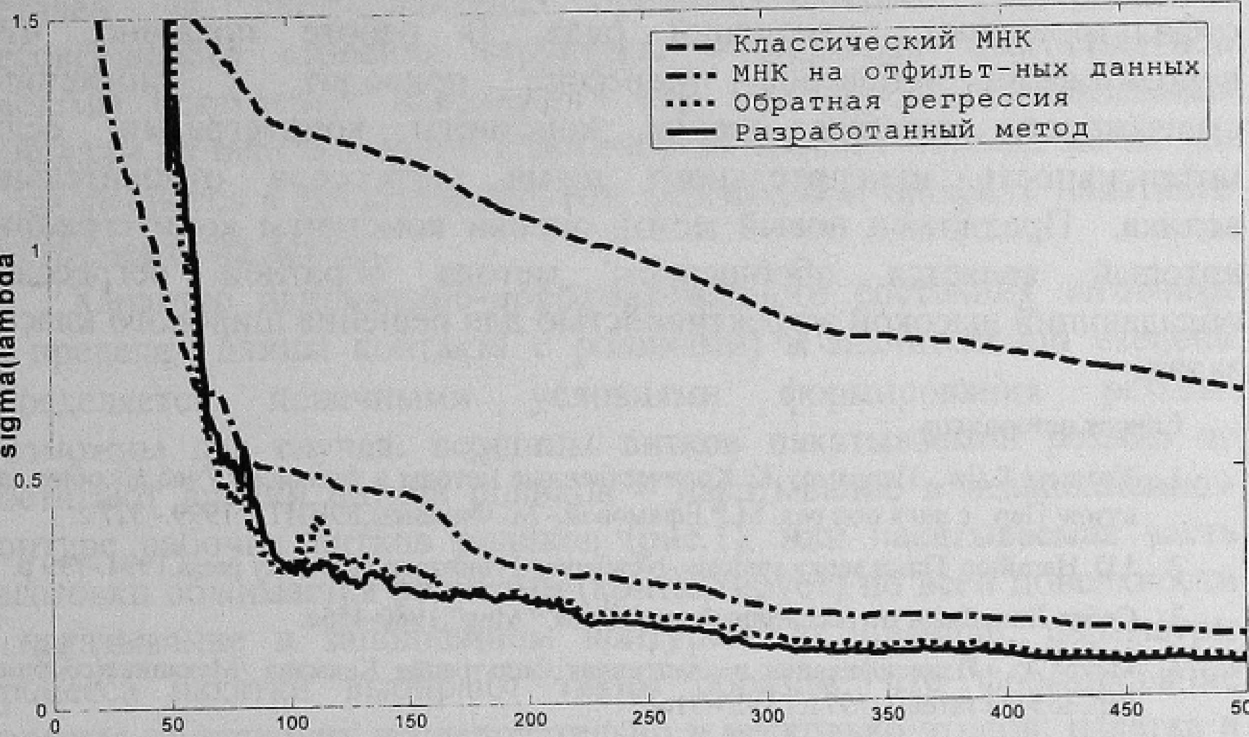


Рисунок 1 - Среднеквадратичная ошибка оценки константы коинтеграции

Из графика видно, что классический МНК дает существенное смещение оценки константы коинтеграции (на 150 точке погрешность составляет 30 %), таким образом мы находимся в зоне высоких смещений оценок, полученных на основе МНК. Применение МНК к отфильтрованным рядам, ведет к улучшению оценки более чем в 2 раза (150 точка- 11,5%). Метод обратной регрессии и разработанный метод практически не дают смещения (150 точка- 6,5%), и являются предпочтительными в данных условиях.

Выводы.

Таким образом, в работе, впервые в коинтеграционном анализе, предложено применение методов идентификации статистических характеристик шумов модели, для повышения эффективности оценки константы коинтеграции, позволяющие увеличить точность оценки, особенно в зонах высоких смещений оценок МНК.

На предварительном этапе предлагается проанализировать исследуемые ряды с помощью методов параметрической идентификации. При высоких интенсивностях обоих шумов целесообразно оценивать константу коинтеграции на основе отфильтрованных значений ряда. В работе показано, что использование обратной регрессии приводит к снижению смещённости ошибки оценки константы коинтеграции, если интенсивность измерительного шума регрессора относительно велика. Предложен новый метод оценки константы коинтеграции, который является обобщением метода обратной регрессии, обладающий высокой эффективностью для решения широкого класса задач.

Список источников

1. Уотшем Т.Дж., Паррамоу К. Количественные методы в финансах:Учеб.пособие для вузов/ Пер. с англ.под ред. М.Р.Ефимовой.- М.:Финансы,ЮНИТИ,1999.- 527 с.
2. J.D. Hamilton Times series analusis.-NewJersey:Princtonun university press,1994.-799 p.
3. Себер Линейный регрессионный анализ.-М.:"Мир",1980-456с.
4. Мехра Р. –Идентификация и адаптивная фильтрация Калмана /Механика(сборник переводов статей ).-1971,Т3.-53-51с.
5. Згуровський М.З.,Підладчиков В.М. Аналітичні методи калмановської фільтрации для систем с апріорною неозначеністю.-К.:"Наукова думка",1995.-282стр.