Федеральное государственное бюджетное образовательное
учреждение высшего профессионального образования

«РОССИЙСКАЯ АКАДЕМИЯ НАРОДНОГО ХОЗЯЙСТВА
И ГОСУДАРСТВЕННОЙ СЛУЖБЫ ПРИ ПРЕЗИДЕНТЕ РОССИЙСКОЙ
ФЕДЕРАЦИИ»

УДК
№ госрегистрации
Инв. №

«УТВЕРЖДАЮ»
Ректор Академии

__________________ В.А Мау

«____ » ______________ 2013г.

ОТЧЕТ
О НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ
по теме:

Структурные VAR и VECM

Научно-исследовательская работа выполнена в соответствии с Государственным заданием РАНХиГС при Президенте Российской Федерации на 2013 год

Москва
2014

Реферат

Содержание

Введение

Глава 1

Теория SVAR и SVECM

Innovation accounting и анализ импульсных откликов в контексте моделей векторых авторегрессий (VAR) начался с работ (Sims, 1980, 1981) в качестве альтернативного подхода к классическому макроэкономическому анализу. Главный элемент критики Симса по поводу последнего типа анализа заключается в том, что макроэкономические модели часто не основаны на надежных экономических теориях или имеющиеся теории не способны предложить полностью специфицированую модель. Если экономические теории не доступны для спецификации модели, должны применяться статистические инструменты. При таком подходе полагается достаточно свободная модель, которая не накладывает жесткие априорные ограничения на процесс порождения. Статистические инструменты используется после этого для определения возможных ограничений. VAR модели представляют собой класс свободных моделей, которые могут быть использованы при таком подходе. Конечно, чтобы интерпретировать эти модели, должны быть сделаны некоторые ограничительные предположения. В частности, упорядочение переменных может иметь важное значение для интерпретации импульсных откликов. (Sims, 1981) предполагает попробовать различное упорядочивание и исследовать чувствительность соответствующих ортогонализованных импульсных откликов и соответствующих выводов на упорядочения переменных.

Данный раздел частично основан на (Lütkepohl, 2005), добавляя новые работы.

1.1 Структурные VAR-модели

В эмпирических работах часто имеет интерес отклик одной переменной на импульс другой переменной в система, которая также включает дополнительные переменные. Таким образом, хочется исследовать импульсные отклики между двумя переменными в системе с более высокой размерностью. Конечно, если есть реакция одной переменной на импульс другой переменной, мы можем назвать последнюю причинной для первой. В этом разделе мы изучим этот тип причинности, прослеживая влияние экзогенного шока или инновации в одной переменной на некоторую или все другие переменные. Этот вид анализа поверхности отклика иногда называется *анализом мультипликаторов.*

1.1.1 Отклики на ошибки прогноза

Рассмотрим простой случай VAR(1) процесса:

$$y_t = v + A_1 y_{t-1} + u_t, \tag{1.1}$$

где $y_t = (y_{1t}, \dots, y_{Kt})'$ – случайный вектор размерности $(K \times 1)$, A_i – фиксированная матрциа коэффициентов размерности $(K \times K)$, $v = (v_1, \dots, v_K)'$ – фиксированный вектор констант размерности $(K \times 1)$, который допускает возможсость ненулевого среднего $E(y_t)$. Вектор $u_t = (u_{1t}, \dots, u_{Kt})'$ являестя K-мерным *процессом белого шума* или *инноваций*, то есть $E(u_t) = 0$, $E(u_t u_t') = \Sigma_u$ и $E(u_t u_s') = 0$ для $s \neq t$. Ковариационная матрица Σ_u предполагается невырожденной, если не предполагается обратное.

Процесс (1.1) имеет представление скользящего среднего (MA) при условии стабильности[1]:

$$y_t = \mu + \sum_{i=1}^{\infty} \Phi_i u_{t-i}, \quad t = 0, \pm 1, \pm 2, \dots, \tag{1.2}$$

где $\mu = (I_K - A_1)^{-1} v$, $\Phi_i = A_1^i$.

Предположим, что нас интересует влияние инновации в какой-то из переменных. Чтобы изолировать такое влияние, предположим, что все переменные равны своему среднему значению до момента времени $t = 0$, $y_t = \mu$ для $t < 0$, а какая-то из переменных (i-я переменная) увеличивается на одну единицу в период $t = 0$, то есть $u_{i,0} = 1$. Мы можем проследить, что случится с системой в течение периодов $t = 1, 2, \dots$, если больше не происходит никаких шоков, то есть $u_{j,0} = 0$ для всех $j \neq i$. Поскольку нас не интересует среднее системы, а только вариаци переменных около их средних, можно предположить, что все переменные имеют нулевое среднее и положить $v = 0$ в (1.1). Рассмотрим следующий простой пример трехмерной системы:

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \end{bmatrix} = \begin{bmatrix} 0.5 & 0 & 0 \\ 0.1 & 0.1 & 0.3 \\ 0 & 0.2 & 0.3 \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ y_{3,t-1} \end{bmatrix} + \begin{bmatrix} u_{1,t} \\ u_{2,t} \\ u_{3,t} \end{bmatrix} \tag{1.3}$$

Пусть в первой переменной в период $t = 0$ происходит единичный шок. Тогда можно получить, что

$$\begin{aligned} y_0 &= \begin{bmatrix} y_{1,0} \\ y_{2,0} \\ y_{3,0} \end{bmatrix} = \begin{bmatrix} u_{1,0} \\ u_{2,0} \\ u_{3,0} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \\ y_1 &= \begin{bmatrix} y_{1,1} \\ y_{2,1} \\ y_{3,1} \end{bmatrix} = A_1 y_0 = \begin{bmatrix} 0.5 \\ 0.1 \\ 0 \end{bmatrix}, \\ y_2 &= \begin{bmatrix} y_{1,2} \\ y_{2,2} \\ y_{3,2} \end{bmatrix} = A_1 y_1 = A_1^2 y_0 = \begin{bmatrix} 0.25 \\ 0.06 \\ 0.02 \end{bmatrix}. \end{aligned}$$

[1] Процесс VAR(1) является стабильным, если все собственные значения A_1 меньше единицы по модулю. Это эквивалентно условию $\det(I_K - A_1 z) \neq 0$ для $|z| \leq 1$.

Продолжая эту процедуру, оказывается, что $y_i = (y_{1,i}, y_{2,i}, y_{3,i})'$ является только первым столбцом матрицы A_1^i. Аналогично можно показать, что единичный шок во второй (третьей) переменной в $t = 0$ после i периодов приводит к вектору y_i, который является вторым (третьим) столбцом матрицы A_1^i. Таким образом, элементы A_1^i представляют эффект единичных шоков в переменных системы после i периодов. Следовательно, они называются импульсными откликами, или динамическими мультипликаторами. Также отметим, что поскольку $A_1^i = \Phi_i$ является i-й матрицей коэффициентов в MA-представлении (1.2) процесса VAR(1), матрицы коэффициентов MA содержат импульсные отклики системы.

Результат выше можно обобщить на случай процесса векторной авторегрессии порядка p, VAR(p):

$$y_t = v + A_1 y_{t-1} + \cdots + A_p y_{t-p} + u_t. \tag{1.4}$$

Любой VAR(p)-процесс можно переписать в виде VAR(1) как

$$Y_t = \mathbf{v} + \mathbf{A} Y_{t-1} + U_t, \tag{1.5}$$

где

$$\begin{aligned} Y_t &= \underset{(Kp\times 1)}{\begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{bmatrix}}, \quad \mathbf{v} = \underset{(Kp\times 1)}{\begin{bmatrix} v \\ 0 \\ \vdots \\ 0 \end{bmatrix}}, \\ \mathbf{A} &= \underset{(Kp\times Kp)}{\begin{bmatrix} A_1 & A_2 & \ldots & A_{p-1} & A_p \\ I_K & 0 & \ldots & 0 & 0 \\ 0 & I_K & & 0 & 0 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & I_K & 0 \end{bmatrix}}, \quad U_t = \underset{(Kp\times 1)}{\begin{bmatrix} u_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}}. \end{aligned}$$

Процесс Y_t являестя стабильным, если $\det(I_{K_p} - \mathbf{A}_1 z) \neq 0$ для $|z| \leq 1$. При этом условии Y_t имеет MA-представление:

$$Y_t = \boldsymbol{\mu} + \sum_{i=1}^{\infty} \mathbf{A}_i U_{t-i}. \tag{1.6}$$

Процесс y_t можно переписать в виде

$$y_t = JY_t = J\boldsymbol{\mu} + \sum_{i=1}^{\infty} J\mathbf{A}_i J' J U_{t-i} = \mu + \sum_{i=1}^{\infty} \Phi_i u_{t-i}, \tag{1.7}$$

где $J = [I_K \ : \ 0 \ : \ \cdots \ : \ 0]$ – матрица размерности $(K \times Kp)$, $\mu = J\boldsymbol{\mu}$, $\Phi_i = J\mathbf{A}_i J'$ и из-за особой структуры процесса белого шума U_t мы имеем $U_t = J'JU_t$ и $JU_t = u_t$. Матрица Φ_i иначе может быть записана как $\Phi_0 = I_K$, $\Phi_i = \sum_{j=1}^{i} \Phi_{i-j} A_j$, $i = 1, 2, \ldots$.

Относительно импульсных откликов, как и в случае VAR(1) предположим, что $y_t = 0$ для $t < 0$, $u_t = 0$ для $t > 0$ и $y_0 = u_0$ – K-мерный единичный вектор e_k, скажем, с единицей на k-ой координате и нулями в остальных случаях. Тогда $Y_0 = (e_k', 0, \dots, 0)'$ и $Y_i = \mathbf{A}^i Y_0$. Следовательно, импульсные отклики являются элементами верхнего левого $(K \times K)$ блока матрицы $\mathbf{A}^i$. Эта матрица, однако, является i-ой матрицей коэффициентов Φ_i MA-представление (1.7) для y_t. Другими словами, $\phi_{jk,i}$ – jk-ый элемент Φ_i, представляет реакцию j-ой переменной системы на единичный шок в переменной k, произошедший i периодов назад, если, конечно, эффект не загрязнен другими шоками в системе. Поскольку u_t является только прогнозом ошибки на один шаг вперед VAR процесса, шоки, рассмотренные здесь, можно рассматривать как ошибки прогноза, а импульсные отклики иногда называют*импульсными откливами ошибки прогноза*.

Если переменные имеют различные масштабы, иногда полезно рассмотреть инновации одного стандартного отклонения, а не единичные шоки. Однако это всего лишь вопрос масштабирования импульчных откликов, такое шкалирвоание может иногда давать лучшие картинки динамических соотношений, поскольку средний размер инноваций, происходящих в системе, зависит от их стандартного отклонения.

Импульсные отклики равны нулю, если одна из переменных не является причинной по Гренджеру для других переменных, взятых в качестве группы. Более точно, инновация в переменной k не имеет влияния на другие переменные, если она не причинна по Гренджеру для множества оставшихся переменных. В прикладных работах часто первостепенный интерес заклчается в том, имеет ли одна переменная влияние на конкретную другую переменную. То есть хотелось бы знать, для некоторого $k \neq j$ равно ли значение $\phi_{jk,i}$ нулю для $i = 1, 2, \dots$. Если $\phi_{jk,i}$ представляет фактические реакции переменной j на единичный шок в переменной k, мы можем назвать последнюю *непричинной* для j-ой переменной, если $\phi_{jk,i} = 0$ для $i = 1, 2, \dots$. Чтобы проверить последнее условие, нет необходимости вычислять бесконечно много матриц Φ_i. Достаточно проверить первые $p(K-1)$ матриц Φ_i, поскольку если для K-мерного стационарного и стабильного VAR(p) первые $p(K-1)$ откликов переменной j на импульс в переменной k нулевые, то все следующие отклики должны быть также нулевыми (см. (Lütkepohl, 2005, Proposition 2.4))

Иногда исследователя интересует накопленное влияние в течение нескольких периодов от шока одной переменной. Этот влияние может быть определено путем суммирования матриц коэффициентов MA-представления. Например, k-ый столбец в $\Psi_n = \sum_{i=0}^{n} \Phi_i$ содержит *накопленные отклики* в течение n периодов от единичного шока в k-ой переменной системы. Эти значения иногда называют n-ыми *промежуточными мультипликаторами* (interim multipliers). Полные накпленные влияния для всех будущих периодов могут быть получены, суммируя все матрицы коэффициентов MA. $\Psi_\infty = \sum_{i=0}^{\infty} \Phi_i$ инога называют матрицей *долгосрочных эффектов*, или *полным мультипликатором.* Поскольку MA-оператор $\Phi(z)$ является обратным VAR-оператора $A(z) = I_K - A_1 z - \dots - A_p z^p$, долгосрочные эффекты могут быть легко получены как

$$\Psi_\infty = \Phi(1) = (I_K - A_1 - \dots - A_p)^{-1}. \tag{1.8}$$

1.1.2 Отклики на ортогональные импульсы

Проблематичное предположение о данном виде анализа импульсных откликов заключается в том, что шок происходит только в одной переменной. Такое предположение может быть разумно, если шоки в различных переменных являюстя независимыми. Если это не так, можно утверждать, что ошибки состоят из всех влияний и переменных, которые не включены непосредственно во множество переменных y. Таким образом, в дополнении к силам, которые влияют на все переменные, могут быть силы, которые влияют, скажем, только на 1-у переменную. Если шок в первой переменной произошел из-за таких сил, то можно разумно интерпретировать коэффициенты Φ_i как динамические отклики. С другой стороны, корреляция ошибок может указывать на то, что шок в одной переменной вероятно будет сопровождаться шоком в другой переменной. В этом случае, полагая все другие остатки нулю, можно прийти к вводящим в заблуждение выводам о фактических динамических соотношениях между переменными. Все это являестя причиной, почему анализ импульсных откликов часто применяется в терминах следующего MA-представления:

$$y_t = \mu + \sum_{i=0}^{\infty} \Phi_i P P^{-1} u_{t-i} = \mu + \sum_{i=0}^{\infty} \Theta_i w_{t-i}, \tag{1.9}$$

где $\Sigma_u = PP'$, P – нижнетреугольная матрица с положительными диагональными элементами, $\Theta_i = \Phi_i P$, компоненты $w_t = (w_{1t}, \ldots, w_{Kt})'$ являются некоррелированными и имеют единичную дисперсию ($w_t = P^{-1}u_t$, $\Sigma_w = P^{-1}\Sigma_u(P^{-1})' = I_K$), $\Sigma_w = I_K$. Компоненту среднего можно исключить из этого уравнения, поскольку она не является важной для текущего анализа. В (1.9) разумно предположить, что изменение в одной компоненте w_t не имеет влияния на другие компоненты, поскольку эти компоненты ортогональны (некоррелированы). Кроме того, дисперсии этих компонент равны единице. Таким образом, единичная инновация является только инновацией размера одного стандартного отклонения. Элементы Θ_i интерпретируются как отклики системы на такие инновации. Более точно, jk-ый элемент Θ_i предполагаестя представляющим влияние на переменную j от единичной инновации в k-ой переменной, которая произошла i периодов назад.

Чтобы соотнести эти импульсные отклики с VAR-моделью, рассмотрим процесс VAR(p) с нулевым средним:

$$y_t = A_1 y_{t-1} + \cdots + A_p y_{t-p} + u_t. \tag{1.10}$$

Этот процесс может быть переписан таким способом, чтобы остатки в различных уравнениях были некоррелированы. Для этой цели мы выбираем разложение ковариационной матрицы белого шума $\Sigma_u = W\Sigma_\varepsilon W'$, где Σ_ε – диагональная матрциа с положительными диагональными элементами, а W – нижнетреугольная матрица с единичной диагональю. Это разложение получено из разложения Холецкого $\Sigma_u = PP'$, определяя диагональную матрицу D, которая имеет ту же самую главную диагональ, и специфицируя $W = PD^{-1}$ и $\Sigma_\varepsilon = DD'$.

Умножая слева (1.10) на $\mathsf{A} = W^{-1}$, получаем

$$\mathsf{A} y_t = A_1^* y_{t-1} + \cdots + A_p^* y_{t-p} + \varepsilon_t, \tag{1.11}$$

где $A_i^* = \mathsf{A} A_i$, $i = 1, \ldots, p$, и $\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{Kt})' = \mathsf{A} u_t$ имеет диагональную ковариационную матрицу

$$\Sigma_\varepsilon = E(\varepsilon_t \varepsilon_t') = \mathsf{A} E(u_t u_t') \mathsf{A}' = \mathsf{A} \Sigma_u \mathsf{A}'.$$

Добавление $(I_K - \mathsf{A})$ к обоим сторонам (1.11) дает

$$y_t = A_0^* y_t + A_1^* y_{t-1} + \cdots + A_p^* y_{t-p} + \varepsilon_t, \tag{1.12}$$

где $A_0^* = I_K - \mathsf{A}$. Поскольку W является нижнетреугольной с единичной диагональю, то же самое верно и для A. Следовательно,

$$A_0^* = I_K - \mathsf{A} = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ \beta_{21} & 0 & \dots & 0 & 0 \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ \beta_{K1} & \beta_{K2} & \dots & \beta_{K,K-1} & 0 \end{bmatrix}$$

является нижнетреугольной матрицей с нулевой диагональю, и, таким образом, в представлении (1.12) нашего VAR(p)-процесса первое уравнение не содержит текущих y-ов с правой стороны. Второе уравнение может содержать y_{1t} в правой стороне. В общем случае, k-е уравнение может содержать $y_{1t}, \dots, y_{k-1,t}$ и не содержать $y_{kt}, \dots, y_{K,t}$ в правой стороне. Таким образом, если (1.12) отражает фактическое поведение в системе, y_{st} не может иметь мгновенное влияние на y_{kt} для $k < s$. В эконометрической литературе такая система называется *рекурсивной моделью* (см. (**?**, Section 9.6)). Herman Wold has advocated эти модели, где исследователь должен специфицировать мгновенный (текущий?) "причинный" порядок переменных. Этот тип причинности иногда называют *Вольд-причинностью* (Wold-causality). Если мы проследим ε_{it}-инновации размера одной стандартной ошибки через систему (1.12), мы просто получим Θ импульсные отклики. Это может быть видно, решая систему (1.12) для y_t,

$$y_t = (I_K - A_0^*)^{-1} A_1^* y_{t-1} + \dots + (I_K - A_0^*)^{-1} A_p^* y_{t-p} + (I_K - A_0^*)^{-1} \varepsilon_t.$$

Замечая, что $(I_K - A_0^*)^{-1} = W = PD^{-1}$ показывает, что текущие эффекты шоков одного стандартного отклонения (ε_{it} размера одного стандартного отклонения) на систему представлены элементами $WD = P = \Theta_0$, поскольку диагональные элементы в D являются просто стандартными отклонениями компонент ε_t. Θ_i может быть получена, прослеживая эти эффекты через систему.

Заметим, что $\Theta_0 = P$ является нижнетреугольной, и некоторые элементы ниже диагонали будут ненулевыми, если Σ_u имеет ненулевые элементы вне диагонали. Если ковариационная матрица белого шума Σ_u содержит нули, некоторые компоненты $u_t = (u_{1t}, \dots, u_{Kt})'$ одновременно некоррелированы. Предположим, например, что u_{1t} некоррелированы с u_{it} для $i = 2, \dots, K$. В этом случае $\mathsf{A} = W^{-1}$ и, таким образом, A_0^* имеют блок нулей, так что y_1 не имеет мгновенного влияния на y_i, $i = 2, \dots, K$. Поскольку элементы $P = \Theta_0$ представляют немедленные отклики системы на единичные инновации, они иногда называются *мультипликаторами вклада* (impact multipliers).

Чтобы определить, нет ли отклика одной переменной на импульс в одной из других переменных, достаточно рассмотреть первые $pK - p$ коэффициентов откликов и немедленный эффект.

Тот факт, что Θ_0 – нижнетреугольная, показывает, что имеет значение упорядочивание переменных, то есть важно, какая из переменных называется y_1, какая – y_2 и так далее. Одна из проблем с этим типом анализа импульсных откликов заключается в том, что упорядочивание переменных не может быть определено статистическими методами, а должно быть специфицировано исследователем. Порядок должен быть таким, чтобы первая переменная являлась единственной, которая имеет потенциальное немедленное влияние на другие переменные. Вторая переменная может иметь влияние на оставшиеся $K-2$ компоненты y_t, но не на y_1, и так далее. Установление такого упорядочивания может быть достаточно сложным упражнением на практике. Выбор упорядочивания, причинное упорядочивание Вольда, может в значительной мере определять импульсные отклики и поэтому является критически важным для интерпретации системы. В настоящее время мы имеем дело только с известными системами. В этой ситуации, предположение, что порядок является известным, не может быть большим ограничением.

Наша интерпретация ортогонализованных импульсных откликов основана на представлении (1.12), и импульсы рассматриваются как изменения в наблюдаемых переменных. Иногда более правдоподобно сосредоточить внимание на импульсы, которые не могут быть легко связанными с изменениями конкретной наблюдаемой переменной внутри системы. В этом случае может быть более логично основывать интерпретацию на MA-представлении (1.9), которое раскладывает переменные по их вкладу инноваций w_{kt}. Если эти инновации могут быть связаны с конкретным импульсом на систему, ортгонализованные импульсные отклики отражают реакции переменных таким возможно ненаблюдаемым инновациям. В этом случае конкретный импульс или шок в систему может иметь текущий эффект на несколько переменных, в то время как некоторый другой импульс может иметь только текущий эффект на одну конкретную переменную и может влиять на другие переменные только с запаздыванием. Раскладывая $\Sigma_u = PP'$ с некоторой нетреугольной матрицей P, также возможно, что шоки имеют текущее влияние на все наблюдаемые переменные в системе. В данном виде интерпретации нахождение матрицы разложения P и, следовательно, инноваций w_t, которые фактически могут быть связаны с интересующими шоками, часто является сложной частью анализа.

1.1.3 Идентификация структурных шоков

A-модель, B-модель и AB-модель

В данном подразделе мы рассмотрим A-модель, B-модель и AB-модель в терминологии (Lütkepohl, 2005). Рассмотрим модель (1.11). В отличие от MA-представления (1.9), где $\Theta_i = \Phi_i P$, рассмотри другое MA-представление:

$$y_t = \Theta_0 \varepsilon_t + \Theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots, \tag{1.13}$$

где $\Theta_j = \Phi_j \mathsf{A}^{-1}$ $(j = 0, 1, 2, \dots)$, а матрица A не специфицируется конкретным образом (рассмотренная ранее рекурсивная идентификация выбирается только если существвет некоторое теоретическое обоснование выбранного упорядочивания), а только должна обеспечивать диагональную ковариационную матрицу у ε_t. Теперь ϵ_t не обязательно должны иметь единичные дисперсии, в отличие от w_t в (1.9).

В рассматрвиаемой модели идентификационные ограничения накладываются на матрицу A, так чтобы $\varepsilon_t = \mathsf{A} u_t$ имел диагональную ковариационную матрицу. Эта модель в дальнейшем будем называть A-моделью. Учитывая то, как мы ввели связанные ограничения, вполне вероятно предположить, что A имеет единичную главную диагональ. В этом случае требуется $K(K-1)/2$ огарничений на внедиагональные элементы A, чтобы гарантировать точноидентифицируемые шоки ε_t и, следовательно, точноидентифицированные импульсные отклики.

Что касается ограничений на матрицу A понятно, что эти ограничения не могут быть произвольными. Записывая их в форме $C_{\mathsf{A}} \operatorname{vec}(\mathsf{A}) = c_{\mathsf{A}}$, где C_{A} – selection matrix размерности $(\frac{1}{2}K(K+1)\times K^2)$, и c_{A} – соответствующий фиксированный вектор размерности $(\frac{1}{2}K(K+1)\times 1)$, ограничения должны быть такими, чтобы система уравнений

$$\mathsf{A}^{-1}\Sigma_\varepsilon \mathsf{A}^{-1\prime} = \Sigma_u \quad \text{и} \quad C_{\mathsf{A}} \operatorname{vec}(\mathsf{A}) = c_{\mathsf{A}} \tag{1.14}$$

имела единственное решение, по крайней мере локально. Ясно, что система является нелинейной по A. Следовательно, мы можем только надеяться на локальную единственность, или идентификацию в целом. Необходимое и достаточное условие для (1.14), чтобы эта система имела единственное решение и, таким образом, для локальной идентификации структурных параметров, заключается в следующем. Пусть Σ_ε является положительноопределенной диагональной матрицей размерности $(K\times K)$, и пусть A является невырожденной матрицей размерности $(K\times K)$. Тогда для заданной симмутричной положительноопределенной $(K\times K)$-матрицы Σ_u, $(N\times K^2)$-матрицы C_{A} и фиксированного $(N\times 1)$-вектора c_{A} система уравнение в (1.14) имеет локально единственное решение для A и диагональных элементов Σ_ε, тогда и только тогда, когда

$$\operatorname{rk}\begin{bmatrix} -2\mathbf{D}_K^+(\Sigma_u\otimes \mathsf{A}^{-1}) & \mathbf{D}_K^+(\mathsf{A}^{-1}\otimes \mathsf{A}^{-1})\mathbf{D}_K \\ C_{\mathsf{A}} & 0 \\ 0 & C_\sigma \end{bmatrix} = K^2 + \frac{1}{2}K(K+1).$$

Здесь $\mathbf{D}_K$ – дуплицирующая матрица размерности $(K^2\times \frac{1}{2}K(K+1))$, $\mathbf{D}_K^+ = (\mathbf{D}_K'\mathbf{D}_K)^{-1}\mathbf{D}_K'$, C_σ – selection matrix размерности $(\frac{1}{2}K(K-1)\times\frac{1}{2}K(K+1))$, которая выбирает элементы $\operatorname{vech}(\Sigma_\varepsilon)$ ниже главной диагонали. Глобально единственное решение получается, если диагональные элементы в A ограничены, чтобы равняться 1.

Для практических целей является проблематичным, что это условие идентификации включает неизвестные параметры. Следовательно, строго говоря, это можно только проверить, когда истинные параметры известны. На практике неизвестные значения могут быть заменены на оценки, и улсовие можно проверить, используя оцененные матрицы, поскольку можно показать, что ранг соответствующей матрицы является или меньше $K^2 + \frac{1}{2}K(K+1)$ всюду в пространстве параметров, или ранговое условие выполняется почти всюду. В последнем случае это может потерпеть неудачу только на множестве нулевой меры Лебега. Таким образом, если рассматривается случайно извлеченный вектор из пространства параметров, он должен удовлетворяеть ранговому условию с вероятностью единица, если модель локлаьно идентифицируема. В любом случае, C_{A} должна иметь по крайней мере $K(K+1)/2$ строк, чтобы гарантировать идентификацию. Другими словами, наличие $K(K+1)/2$ ограничений являестя необходимым условием идентификации.

Хотя мы сформулировали ограничения на матрицу A в виде $C_{\mathsf{A}} \operatorname{vec}(\mathsf{A}) = c_{\mathsf{A}}$, можно записать их в виде

$$\operatorname{vec}(\mathsf{A}) = R_{\mathsf{A}}\gamma_{\mathsf{A}} + r_{\mathsf{A}},$$

где R_{A} и r_{A} – соответствующая фиксированная матрица и соответствующий вектор, соответственно, а γ_{A} – вектор неограниченных параметров.

Как правило, в анализе импульсных откликов смещается акцен от спецификации соотношений между наблюдаемыми переменными на интерпретацию неожидаемой части их изменений или шоков. Следовательно, это не редкость, чтобы идентифицировать структурные инновации

ε_t непосредственно из ошибок прогноза или остатков u_t в приведенной форме. Один из способов сделать это – думать об ошибках прогноза как о линейных функциях структурных инноваций. В этом случае у нас есть соотношения $u_t = \mathsf{B}\varepsilon_t$. Следовательно, $\Sigma_u = \mathsf{B}\Sigma_\varepsilon \mathsf{B}'$. Нормализуя дисперсии структурных инноваций на единицу, то есть предполагая $\varepsilon_t \sim (0, I_K)$, мы получаем

$$\Sigma_u = \mathsf{B}\mathsf{B}'. \tag{1.15}$$

Из-за симметричности ковариационной матрицы нам снова необходимо $K(K-1)/2$ дополнительных ограничений для идентификации всех K^2 элементов в B.

Текущая модель с

$$u_t = \mathsf{B}\varepsilon_t$$

и $\varepsilon \sim (0, I_K)$ будем называть B-моделью. Если существуют только нулевые ограничения, они могут быть записаны в виде

$$C_{\mathsf{B}} \operatorname{vec} \mathsf{B} = 0, \tag{1.16}$$

где C_{B} – selection matrix размерности $(N \times K^2)$. Необходимое и достаточное ранговое условие для локальной идентификации модели заключается в следующем. Пусть B является невырожденной $(K \times K)$-матрицей. Тогда для заданной симметричной положительно определенной $(K \times K)$-матрицы Σ_u и $(N \times K^2)$-матрицы C_{B} система уравнений в (1.15) и (1.16) имеет локально единственное решение, тогда и только тогда, когда

$$\operatorname{rk} \begin{bmatrix} 2\mathbf{D}_K^+(\mathsf{B} \otimes I_K) \\ C_{\mathsf{B}} \end{bmatrix} = K^2.$$

Необходимым условием для того, чтобы $((\frac{1}{2}K(K+1) + N) \times K^2)$-матрицы

$$\begin{bmatrix} 2\mathbf{D}_K^+(\mathsf{B} \otimes I_K) \\ C_{\mathsf{B}} \end{bmatrix}$$

имела ранг, равный K^2, является $N = \frac{1}{2}K(K-1)$. Другими словами, нам нужно $\frac{1}{2}K(K-1)$ ограничений для идентификации, как упомянуто ранее.

Легко видеть, что решение системы (1.15)/(1.16) не является глобально единственным, поскольку для любой матрицы B, удовлетворяющей уравнениям, $-\mathsf{B}$ будет также являться решением. Этот результат выполняется из-за того факта, что B входит в уравнения (1.15), будучи возведенной в квадрат. Фактически для любого решения B матрица $\mathsf{B}\Lambda$ будет также являться решением для любой диагональной матрицы Λ, у которой только элементы 1 и -1 на главной лиагонали. Очевидно, что если B такова, что (1.15) и (1.16) выполняются, $\Sigma_u = \mathsf{B}\Lambda\Lambda'\mathsf{B}'$ также выполняется, поскольку $\Lambda\Lambda' = I_K$. Кроме того,

$$C_{\mathsf{B}} \operatorname{vec}(\mathsf{B}\Lambda) = C_{\mathsf{B}}(\Lambda \otimes I_K) \operatorname{vec}(\mathsf{B}) = 0,$$

поскольку для каждого элемента $\mathsf{b}_{ij} = 0$ мы имеем $-\mathsf{b}_{ij} = 0$. Таким образом, каждый столбец в B можно заменить на столбец с противоположным знаком. Следовательно, ограничение в (1.16) идентифицирует B только локально в общем случае. Единственность может быть потенциально получена, фиксируя знаки диагональных элементов, однако. Знаки диагональных элементов в B определяют знаки шоков. Таким образом, если мы хотим изучить влияние положительного шока на определенную величину, в то время как соответствующий диагональный элемент в B отрицательный, мы можем просто поменять знаки всех элементов в соответствующем столбце в B, или, другими словами, мы можем просто поменять знаки всех текущих откликов на соответттствующий шок, чтобы найти желаемый результат.

Также стоит отметить, что ограничения могут быть выражены в альтернативном виде:

$$\operatorname{vec}(\mathsf{B}) = R_{\mathsf{B}}\gamma_{\mathsf{B}}, \tag{1.17}$$

где γ_{B} содержит все неограниченные коэффициенты в B, а R_{B} является фиксированной матрицей из нулей и единиц.

Также возможно рассмотреть оба типа ограничений одновременно. То есть мы можем рассмотреть так называемую AB-модель:

$$\mathsf{A}u_t = \mathsf{B}\varepsilon_t, \quad \varepsilon_t \sim (0, I_K). \tag{1.18}$$

В этом случае одновременные уравнения системы формулируются для ошибок в модели приведенной формы, а не для наблюдаемых переменных непосредственно. Тем самым модель составляет переход от спецификации прямых соотнощений для наблюдаемых переменных к формулированию соотношений для инноваций.

В этой модели мы получаем из (1.18), что $u_t = \mathsf{A}^{-1}\mathsf{B}\varepsilon_t$, и, следовательно, $\Sigma_u = \mathsf{A}^{-1}\mathsf{B}\mathsf{B}'\mathsf{A}^{-1\prime}$. Таким образом, у нас есть $K(K+1)/2$ уравнений вида

$$\operatorname{vech}(\Sigma_u) = \operatorname{vech}(\mathsf{A}^{-1}\mathsf{B}\mathsf{B}'\mathsf{A}^{-1\prime}), \tag{1.19}$$

в то время как две матрицы A и B имеют K^2 элементов каждая Таким образом, нам нужно дополнительно $2K^2 - \frac{1}{2}K(K+1)$ ограничений для идентификации всех $2K^2$ элементов в A и B по крайней мере локально. Даже если диагональные элементы в A полагаются равными единице, для идентификации нужно $2K^2 - K - \frac{1}{2}K(K+1)$ дополнительных ограничений. Следовательно, пожалуй, неудивительно, что большинство приложений рассматривает частные случаи с $\mathsf{A} = I_K$ (B-модель) или $\mathsf{B} = I_K$ (A-модель). Тем не менее, общая модель является полезной основой для анализа SVAR. Ограничения, как правило, являются нормализацией или нулевыми ограничениями, которые могут быть записаны в виде линейных уравнений вида

$$\operatorname{vec}(\mathsf{A}) = R_{\mathsf{A}}\gamma_{\mathsf{A}} + r_{\mathsf{A}}, \quad \text{и} \quad \operatorname{vec}(\mathsf{B}) = R_{\mathsf{B}}\gamma_{\mathsf{B}} + r_{\mathsf{B}}, \tag{1.20}$$

где R_{A} и R_{B} – соответствующие фиксированные матрицы, состоящие из нулей и единиц, γ_{A} и γ_{B} – векторы свободных параметров, r_{A} и r_{B} – векторы фиксированных параметров, которые допускают, например, нормализацию диагональных элементов в A. Хотя r_{B} обычно равен нулю, как в (1.17), мы представляем ограничения на B с общим вектором r_{B}, поскольку эта дополниетльная компонента не будет усложнять анализ.

Умножая два набора уравнений в (1.20) на ортогональные дополнения R_{A} и R_{B}, $R_{\mathsf{A}\perp}$ и $R_{\mathsf{B}\perp}$, соответственно, легко видеть,ч то их можно записать в виде

$$C_{\mathsf{A}} \operatorname{vec}(\mathsf{A}) = c_{\mathsf{A}} \quad \text{и} \qquad C_{\mathsf{B}} \operatorname{vec}(\mathsf{B}) = c_{\mathsf{B}}, \tag{1.21}$$

где $C_{\mathsf{A}} = R_{\mathsf{A}\perp}$, $C_{\mathsf{B}} = R_{\mathsf{B}\perp}$, $c_{\mathsf{A}} = R_{\mathsf{A}\perp}r_{\mathsf{A}}$ и $c_{\mathsf{B}} = R_{\mathsf{B}\perp}r_{\mathsf{B}}$. О матрицах C_{A} и C_{B} можно думать как о соответствующих selection matrices. Снова в общем случае ограничения будут гарантировать только локальную единственность для A и B из-за нелинейной природы полного множества уравнений. Ранговое условие для локальной идентификации заключается в следующем. Пусть A и B – невырожденные $(K \times K)$-матрицы. Тогда для заданной симметричной положительно определенной $(K \times K)$-матрицы Σ_u система уравнений (1.19) и (1.21) имеют локально единственное решение тогда и только тогда, когда

$$\mathrm{rk}\begin{bmatrix} -2\mathbf{D}_K^+(\Sigma_u \otimes \mathsf{A}^{-1}) & 2\mathbf{D}_K^+(\mathsf{A}^{-1}\mathsf{B} \otimes \mathsf{A}^{-1}) \\ C_{\mathsf{A}} & 0 \\ 0 & C_{\mathsf{B}} \end{bmatrix} = 2K^2.$$

Идентификация Бланшара-Кваха

Ясно, что не всегда легко найти подходящие и в целом приемлемые ограничения на матрицы A и B. Накладывание ограничений непосредственно на эти матрицы фактически не является необходимым для идентификации структурных инноваций и импульсных откликов. Другой тип ограничений обсуждался в (Blanchard and Quah, 1989). Авторы рассмотрели накопленные эффекты шоков на систему. В терминах структурных импульсных откликов (1.13) они сконцентрировали внимание на *полной матрицы вклада* (total impact matrix):

$$\Xi_\infty = \sum_{i=0}^{\infty} \Theta_i = (I_K - A_1 - \cdots - A_p)^{-1}\mathsf{A}^{-1}\mathsf{B}, \tag{1.22}$$

и они идентифицировали структурные инновации, полагая нулевые ограничения на эту матрицу. Другими словами, они предполагали, что некоторые шоки не имеют полный долгосрочный эффект.

1.1.4 Декомпозиция дисперсий ошибок прогноза

Если инновации, которые фактически приводят в движение систему могут быть идентифицированы, доступен дальнейший способ интерпретации VAr-модели. Предположим, что доступна рекурсивная идентификационная схема, так что может быть рассмотрено MA-представление с ортогонализованными инновациями белого шума. В контексте представления

$$y_t = \mu + \sum_{i=0}^{\infty} \Theta_i w_{t-i}, \tag{1.23}$$

с $\Sigma_w = I_K$ ошибка оптимального прогноза на h шагов равна

$$\begin{aligned} y_{t+h} - y_t(h) &= \sum_{i=0}^{h-1} \Phi_i u_{t+h-i} = \sum_{i=0}^{h-1} \Phi_i P P^{-1} u_{t+h-i} \\ &= \sum_{i=0}^{h-1} \Theta_i w_{t+h-i}. \end{aligned} \tag{1.24}$$

Обозначая mn-ый эелемент в Θ_i как $\theta_{mn,i}$, ошибка прогноза на h шагов j-ой компоненты y_t равна

$$\begin{aligned} y_{j,t+h} - y_{j,t}(h) &= \sum_{i=0}^{h-1} (\theta_{j1,i} w_{1,t+h-i} + \cdots + \theta_{jK,i} w_{K,t+h-i}) \\ &= \sum_{k=1}^{K} (\theta_{jk,0} w_{k,t+h} + \cdots + \theta_{jk,h-1} w_{k,t+1}). \end{aligned} \tag{1.25}$$

Таким образом, ошибка прогноза j-ой компоненты потенциально состоит из всех инноваций $w_{1t}, \ldots, w_{Kt}$. Конечно, некоторые из $\theta_{mn,i}$ могут быть нулевыми, так что некоторые компоненты могут не появляться в (1.25). Поскольку $w_{k,t}$ некоррелированы и имеют единичные дисперсии, MSE от $y_{j,t}(h)$ равна

$$E(y_{j,t+h} - y_{j,t}(h))^2 = \sum_{k=1}^{K} (\theta^2_{jk,0} + \cdots + \theta^2_{jk,h-1}).$$

Следовательно,

$$\theta^2_{jk,0} + \theta^2_{jk,1} + \cdots + \theta^2_{jk,h-1} = \sum_{i=1}^{h-1} (e_j' \Theta_i e_k)^2 \tag{1.26}$$

иногда интерпретируется как вклад инноваций в переменной k на ошибку прогноза дисперсии, или MSE прогноза на h шагов переменной j. Здесь e_k – k-ый столбец матрциы I_K. Деление (1.26) на

$$\text{MSE}[y_{j,t}(h)] = \sum_{i=0}^{h-1} \sum_{k=1}^{K} \theta^2_{jk,i}$$

дает

$$\omega_{jk,h} = \sum_{i=0}^{h-1} (e_j' \Theta_i e_k)^2 / \text{MSE}[y_{j,t}(h)], \tag{1.27}$$

что является пропорцией ошибки дисперсии прогноза переменной j на h шагов, которая приходится на долю инноваций w_{kt}. Если w_{kt} может быть связано с переменной k, $\omega_{jk,h}$ представляет пропорцию ошибки дисперсии прогноза переменной j на h шагов, которая приходится на долю инноваций в переменной k. Тем самым дисперсия ошибки прогноза раскладывается на компоненты, которые приходились на инновации в различных переменных системы. Из (1.24) видно, что матрица MSE прогноза на h шагов будет

$$\Sigma_y(h) = \text{MSE}[y_t(h)] = \sum_{i=0}^{h-1} \Theta_i \Theta_i' = \sum_{i=0}^{h-1} \Phi_i \Sigma_u \Phi_i'.$$

Диагональные элементы этой матрицы являются MSE от переменных y_{jt}, которые могут быть использованы в (1.27). Анализ декомпозиции дисперсий иногда называется *учетом инноваций* (innovation accounting).

Для стационарного, стабильного, K-мерного VAR(p)-процесса y_t все пропорции дисперсии ошибки прогноза переменной j, которые приходятся на инновации в переменной k, будут равны нулю, если $\omega_{jk,h} = 0$ для $h = pK - p + 1$. В этом контексте стоит отметить соотношение между причинностью по Гренджеру и компонентами дисперсии ошибок прогноза. Для этой цели мы рассмотрим сначала двумерную систему $y_t = (z_t, x_t)'$. В такой системе, если z_t не являестя причинной по Гренджеру для x_t, пропорции дисперсий ошиок прогноза x_t, которые приходятся на инновации в z_t, могут все еще быть ненулевыми. Это свойство прямо следует из определения Θ_i. Непричинность по Гренджеру предполагае нулвые огарничения на Φ_i, что может исчезать в Θ_i, елси ковариационная матрица Σ_u не являестя диагональной. Другими словами, если Σ_u диагональная, так что нет текущей причинности между z_t и x_t, и если в дополнение z_t не является причинной по Гренджеру для x_t, нижний левый элемент в Θ_i будет нулевым. Следовательно, пропорция дисперсии ошибок прогноза x_t, которая приходится на инновации z_t, будет также нулевой.

В системах более высокого порядка предположим, что набор переменных z_t не являестя причиной по Гренджеру для оставшихся переменных x_t, и не существует текущей причинности между двумя наборами переменных. В этом случае пропорции прогноза MSE всех переменных x_t, которые приходятся на переменных z_t, будут нулевыми.

Важно понимать, однако, что причинность по Гренджеру и разложение дисперсии ошибок прогноза являются достаточно различными понятиями, поскольку являюстя различными поянтиями причинность по Гренджеру и текущая причинность. В то время как причинность по Гренджеру – единственно определяемое свойство двух подмножеств переменных в заданном процессе, декомпозиция дисперсии ошибок прогноза не являесят единственной, так как зависит от матриц θ_i, и, таким образом, от выбора преобразующей матрицы P. Следовательно, интерпретация разложения дисперсии ошибок прогноза подвержена влиянию аналогичных критических замечаний, как и интерпретации импульсных откликов. В дополнение применимы все критические вопросы, поднятые в контексте причинности по Гренжеру. То есть компоненты дисперсии ошибок прогноза являюстя условными по отношению к рассматриваемой системе. Они могут измениться, если система будет расширена дополнительными переменными или если переменные будут удалены из системы. Также ошибки измерения, сезонная коррекция и использование агрегирования могут загрязнить декомпозиции дисперсий ошибок прогноза.

1.1.5 Оценивание SVAR-модели

Предположим, что мы хотим оценить следующую SVAR-модель:

$$\mathsf{A}y_t = \mathsf{A}AY_{t-1} + \mathsf{B}\varepsilon_t, \tag{1.28}$$

где $y'_{t-1} = [y'_{t-1}, \dots, y'_{t-p}]$, $A = [A_1, ; A_p]$ и ϵ_t – гауссовский белый шум с ковариационной матрицей I_K, $\varepsilon_t \sim N(0, I_K)$. Предположение о нормальности сделано только для удобства получения оценок. асимптотические свойства оценок будут теми же самыми при более общих предположениях о распределении. Остатки в приведенной форме, соответствующие (1.28), будут иметь вид $u_t = \mathsf{A}^{-1}\mathsf{B}\varepsilon$.

Функция логарифмического правдоподобия для выборки $y_1, \dots, y_T$ равна

$$\begin{aligned}\ln l(A, \mathsf{A}, \mathsf{B}) &= -\frac{KT}{2}\ln 2\pi - \frac{T}{2}\ln|\mathsf{A}^{-1}\mathsf{B}\mathsf{B}'\mathsf{A}'^{-1}| \\ &\quad -\frac{1}{2}\mathrm{tr}\{(Y-AX)'[\mathsf{A}^{-1}\mathsf{B}\mathsf{B}'\mathsf{A}'^{-1}](Y-AX)\} \\ &= \text{constant} + \frac{T}{2}\ln|\mathsf{A}|^2 - \frac{T}{2}\ln|\mathsf{B}|^2 \\ &\quad \frac{1}{2}\mathrm{tr}\{\mathsf{A}'\mathsf{B}'^{-1}\mathsf{B}^{-1}\mathsf{A}(Y-AX)(Y-AX)'\}, \end{aligned} \tag{1.29}$$

где, как обычно, $Y = [y_1, \dots, y_T]$, $X = [Y_0, \dots, Y_{T-1}]$, и используются матричные правила $|\mathsf{A}^{-1}\mathsf{B}\mathsf{B}'\mathsf{A}'^{-1}| = |\mathsf{A}^{-1}|^2|\mathsf{B}|^2 = |\mathsf{A}'|^{-2}|\mathsf{B}|^2$ и $\mathrm{tr}(VW) = \mathrm{tr}(WV)$.

Предположим, что нет ограничений на параметры приведенной формы A. Тогда для любых заданных A и B функция логарифмического правдоподобия $\ln l(A, \mathsf{A}, \mathsf{B})$ максимизируется по A как $\hat{A} = YX'(X'X)^{-1}$. таким образом, заменяя A на $\hat{A}$ в (1.29), мы получаем концентрированную функцию логарифмического правдоподобия

$$\ln l_c(\mathsf{A}, \mathsf{B}) = \text{constant} + \frac{T}{2}\ln|\mathsf{A}|^2 - \frac{T}{2}\ln|\mathsf{B}|^2 - \frac{1}{2}\mathrm{tr}(\mathsf{A}'\mathsf{B}'^{-1}\mathsf{B}^{-1}\mathsf{A}\tilde{\Sigma}_u), \tag{1.30}$$

где $\tilde{\Sigma}_u = T^{-1}(Y - \hat{A}X)(Y - \hat{A}X)'$ – оцененная ковариационная матрица VAR-остатков (см. (**?**)).

Максимизация функции (1.29) по A и B относительно структурных ограничений (1.20) или (1.21) должна быть сделана при помощи численных методов, поскольку аналитическое решение обычно недоступно. Если ограничения задаются в виде (1.20), ограниченная максимизации концентрированной логарифмической функции правдоподобия достигается максимизацией относительно γ_{A} и γ_{B}.

Если эти параметры локально идентифицируемы, ML-оценки имеют стандартные асимптотические свойства:

$$\sqrt{T}\left(\begin{bmatrix}\tilde{\gamma}_{\mathsf{A}} \\ \tilde{\gamma}_{\mathsf{B}}\end{bmatrix} - \begin{bmatrix}\gamma_{\mathsf{A}} \\ \gamma_{\mathsf{B}}\end{bmatrix}\right) \xrightarrow{d} \left(0, \mathcal{I}_\alpha\begin{pmatrix}\gamma_{\mathsf{A}} \\ \gamma_{\mathsf{B}}\end{pmatrix}^{-1}\right),$$

где $\mathcal{I}_\alpha(\cdot)$ – асимпттическая информационная матрица. Она имеет вид

$$\mathcal{I}_\alpha\begin{pmatrix}\gamma_{\mathsf{A}} \\ \gamma_{\mathsf{B}}\end{pmatrix} = \begin{bmatrix}R'_{\mathsf{A}} & 0 \\ 0 & R'_{\mathsf{B}}\end{bmatrix}\mathcal{I}_\alpha\begin{pmatrix}\mathrm{vec}\,\mathsf{A} \\ \mathrm{vec}\,\mathsf{B}\end{pmatrix}\begin{bmatrix}R_{\mathsf{A}} & 0 \\ 0 & R_{\mathsf{B}}\end{bmatrix}$$

и

$$\mathcal{I}_\alpha\begin{pmatrix}\mathrm{vec}\,\mathsf{A} \\ \mathrm{vec}\,\mathsf{B}\end{pmatrix} = \begin{bmatrix}\mathsf{A}^{-1}\mathsf{B}\otimes\mathsf{B}'^{-1} \\ -(I_K\otimes\mathsf{B}'^{-1})\end{bmatrix}(I_{K^2} + \mathbf{K}_{KK})\big[(\mathsf{B}'\mathsf{A}'^{-1}\otimes\mathsf{B}^{-1}) \ : \ -(I_K\otimes\mathsf{B}'^{-1})\big] \tag{1.31}$$

Если γ_{A} и γ_{B} идентифицируемы, то же самое верно и для A и B. Оценивание этих матриц таково, что $\mathrm{vec}(\tilde{\mathsf{A}}) = R_{\mathsf{A}}\tilde{\gamma}_{\mathsf{A}} + r_{\mathsf{A}}$ и $\mathrm{vec}(\tilde{\mathsf{B}}) = R_{\mathsf{B}}\tilde{\gamma}_{\mathsf{B}} + r_{\mathsf{B}}$, соответственно, мы получаем следующее следствие:

$$\sqrt{T}\left(\begin{bmatrix}\mathrm{vec}\,\tilde{\mathsf{A}} \\ \mathrm{vec}\,\tilde{\mathsf{B}}\end{bmatrix} - \begin{bmatrix}\mathrm{vec}\,\mathsf{A} \\ \mathrm{vec}\,\mathsf{B}\end{bmatrix}\right) \xrightarrow{d} N(0, \Sigma_{\mathsf{AB}}),$$

где

$$\Sigma_{\mathsf{AB}} = \begin{bmatrix}R_{\mathsf{A}} & 0 \\ 0 & R_{\mathsf{B}}\end{bmatrix}\mathcal{I}_\alpha\begin{pmatrix}\gamma_{\mathsf{A}} \\ \gamma_{\mathsf{B}}\end{pmatrix}^{-1}\begin{bmatrix}R'_{\mathsf{A}} & 0 \\ 0 & R'_{\mathsf{B}}\end{bmatrix}.$$

Если накладываются только точноидетифицируемые огарничения на структурные параметры, у нас есть ML-оценка для Σ_u,

$$\tilde{\Sigma}_u = T^{-1}(Y - \hat{A}X)(Y - \hat{A}X)' = \tilde{\mathsf{A}}^{-1}\tilde{\mathsf{B}}\tilde{\mathsf{B}}'\tilde{\mathsf{A}}'^{-1}.$$

Если, однако, на A и/или B накладываются сверхидентифицирующие ограничения, соответствующая оценка для Σ_u,

$$\tilde{\Sigma}_u^r = \tilde{\mathsf{A}}^{-1}\tilde{\mathsf{B}}\tilde{\mathsf{B}}'\tilde{\mathsf{A}}'^{-1}, \tag{1.32}$$

будет отличаться от $\tilde{\Sigma}_u$. Фактически LR-статистику,

$$\lambda_{LR} = T(\ln|\tilde{\Sigma}_u^r| - \ln|\tilde{\Sigma}_u|), \tag{1.33}$$

можно использовать для проверки сверхидентифицирующих огарничений. при нулевой гипотезе, что ограничения верны, эта статистика имеет асимптотическое распределение χ^2 с числом степеней свободы, равной числу сверхидентифицирующих ограничений. Другими словами, число степеней свободы равно числу независимых ограничений, накладываемых на A и B минус $2K^2 - \frac{1}{2}K(K+1)$.

Рассмотрим теперь максимизацию функции (1.29). В (**?**) предлагается использовать скоринговый алгоритм, где i-ая итерация имеет вид

$$\begin{bmatrix}\tilde{\gamma}_{\mathsf{A}}\\ \tilde{\gamma}_{\mathsf{B}}\end{bmatrix}_{i+1} = \begin{bmatrix}\tilde{\gamma}_{\mathsf{A}}\\ \tilde{\gamma}_{\mathsf{B}}\end{bmatrix}_{i} + l\mathcal{I}\left(\begin{bmatrix}\tilde{\gamma}_{\mathsf{A}}\\ \tilde{\gamma}_{\mathsf{B}}\end{bmatrix}_{i}\right)^{-1}\mathbf{s}\left(\begin{bmatrix}\tilde{\gamma}_{\mathsf{A}}\\ \tilde{\gamma}_{\mathsf{B}}\end{bmatrix}_{i}\right), \tag{1.34}$$

где $\mathcal{I}(\cdot)$ обозначает информационную матрицу свободных параметров γ_{A} и γ_{B}, то есть в этом случае $\mathcal{I}(\cdot) = T\mathcal{I}_\alpha(\cdot)$, $\mathbf{s}(\cdot)$ – вектор вклада (score vector), l – длина шага.

Вектор вклада может быть получен, исопльзуя правила матричного и векторного дифференцирования. Прменяя цепное парвило для векторного дифференцирования, получим

$$\mathbf{s}\begin{pmatrix}\gamma_{\mathsf{A}}\\ \gamma_{\mathsf{B}}\end{pmatrix} = \frac{\partial \ln l}{\partial(\gamma_{\mathsf{A}}', \gamma_{\mathsf{B}}')'} = \begin{bmatrix}R_{\mathsf{A}}' & 0\\ 0 & R_{\mathsf{B}}'\end{bmatrix}\mathbf{s}\begin{pmatrix}\operatorname{vec}\mathsf{A}\\ \operatorname{vec}\mathsf{B}\end{pmatrix}, \tag{1.35}$$

и

$$\mathbf{s}\begin{pmatrix}\operatorname{vec}\mathsf{A}\\ \operatorname{vec}\mathsf{B}\end{pmatrix} = \frac{\partial \ln l}{\partial\begin{pmatrix}\operatorname{vec}\mathsf{A}\\ \operatorname{vec}\mathsf{B}\end{pmatrix}} = \begin{bmatrix}(I_K \otimes \mathsf{B}'^{-1})\\ -(\mathsf{B}^{-1}\mathsf{A} \otimes \mathsf{B}'^{-1})\end{bmatrix}\mathbf{s}(\operatorname{vec}[\mathsf{B}^{-1}\mathsf{A}])$$

с

$$\mathbf{s}(\operatorname{vec}[\mathsf{B}^{-1}\mathsf{A}]) = T\operatorname{vec}([\mathsf{B}^{-1}\mathsf{A}]'^{-1}) - T(\tilde{\Sigma}_u \otimes I_K)\operatorname{vec}(\mathsf{B}^{-1}\mathsf{A}).$$

На практике скоринговый алгоритм останавливается, если предварительно специфицированные критерии сходимости, такие как относительные изменения логарифмического правдоподобия и параметров, выполняются. Чтобы этот алгоритм работал, обращение информационной матрицы должно существовать, что гарантируется идентификацией параметров, по крайней мере в окрестности их истинных значений. В (Giannini, 1992) используется это свойство для получения альтернативных условий дял идентификации модели. Более точно, автор получил идентификационные условия из огоф акта, что, например, AB-модель локально идентифицируема тога и только тогда, когда матрица

$$\begin{bmatrix} \mathcal{I}_\alpha \begin{pmatrix} \operatorname{vec} \mathsf{A} \\ \operatorname{vec} \mathsf{B} \end{pmatrix} \\ \begin{bmatrix} C_{\mathsf{A}} & 0 \\ 0 & C_{\mathsf{B}} \end{bmatrix} \end{bmatrix} \tag{1.36}$$

имеет полный ранг по столбцам, когда $\mathcal{I}_\alpha(\cdot)$ оценена в истинных значениях параметров.

Данный подход легко обобщить на случай наличия детерминированных компонент. Кроме того, та же самая концентрированная функция правдоподобия получается, если есть ограничения на A. В этом случае теоретически мы можем просто заменить $\tilde{\Sigma}_u$ на ковариационную матрицу остатков ограниченной ML-оценки приведенной формы. Для практических целей, однако, точная ML-оценка обычно не используется, поскольку в общем случае ее вычисление требует методов нелинейной оптимизации. Вместо этого может быть использован доступный GLS, чьи асимптотические свойства совпадают с ML. В этом случае (1.30) не строго концентрированная функция логарифмического правдоподобия, но для простоты мы будем ссылаться на нее как на концентрированную функцию логарифмического правдоподобия.

Рассмотрим теперь оценвиание долгосрочных ограничений типа Бланшара-Кваха. Если матрциа полного вклада Ξ_∞ ограничена треугольной формой как в (Blanchard and Quah, 1989) и (Galí, 1999), оценивание становится особенно легким. Специфицируя $\mathsf{A} = I_K$, используя соотношение $\Xi_\infty = (I_K - A_1 - \cdots - A_p)^{-1}\mathsf{B}$ и замечая, что

$$\Xi_\infty \Xi_\infty' = (I_K - A_1 - \cdots - A_p)^{-1}\Sigma_u(I_K - A_1' - \cdots - A_p')^{-1},$$

матрица B может быть вычислена умножением разложения Холецкого для матрицы

$$(I_K - \hat{A}_1 - \cdots - \hat{A}_p)^{-1}\tilde{\Sigma}_u(I_K - \hat{A}_1' - \cdots - \hat{A}_p')^{-1}$$

на $(I_K - \hat{A}_1 - \cdots - \hat{A}_p)$.

Последняя процедура работает только если оператор VAR являестя стабильным, и процесс является стационарным, поскольку для интегрированных процессов не существует обращения $(I_K - \hat{A}_1 - \cdots - \hat{A}_p)$. С другой стороны, коинтегрированные переменные не создают проблемы для других методов оценивания SVAR-моделей.

1.1.6 Асимптотические распределения импульсных откликов и декомпозиции дисперсий ошибок прогноза

В данном подразделе приводятся асимптотические распределения оценок Φ_i и Θ_i в (1.7) и (1.9), предполагая их неизвестными и вычисленными посредством коэффициентов VAR и ковариационной матрицы ошибок. заметим, что нам не нужно существование MA-представления. Мы только предполагаем, что Φ_i получены из заданных коэффициентов $A_1, \ldots, A_p$ по рекурсивной формуле $\Phi_i = \sum_{j=1}^{i} \Phi_{i-j}A_j$, $i = 1, 2, \ldots$, начиная с $\Phi_0 = I_k$ и полагая $A_j = 0$ для $j > p$. Кроме того, Θ_i получены из $A_1, \ldots, A_p$ и Σ_u как $\Theta_i = \Phi_i P$, где выражение для P зависит от метода идентификации (либо $P = \mathsf{A}^{-1}$, либо $P = \mathsf{B}$, либо $P = \mathsf{A}^{-1}\mathsf{B}$). В дополнение даются распределения накопленных импульсных откликов

$$\begin{aligned}\Psi_n &= \sum_{i=0}^{n}\Phi_i, \quad \Psi_\infty = \sum_{i=0}^{\infty}\Phi_i = (I_K - A_1 - \cdots - A_p)^{-1}, \text{ если существует,}\\ \Xi_n &= \sum_{i=0}^{n}\Theta_i, \quad \Xi_\infty = \sum_{i=0}^{\infty}\Theta_i = (I_K - A_1 - \cdots - A_p)^{-1}P, \text{ если существует}\end{aligned}$$

и компонент дисперсии ошиок прогноза

$$\omega_{jk,h} = \sum_{i=0}^{h-1}(e_j'\Theta_i e_k)^2/\mathrm{MSE}_j(h).$$

Здесь e_k – k-ый столбец в I_K, и

$$\mathrm{MSE}_j(h) = \sum_{i=0}^{h-1} e_j'\Phi_i\Sigma_u\Phi_i' e_j$$

является j-ым диагональным элементом MSE-матрицы $\Sigma_y(h)$ прогноза на h шагов.

При формальном написании асимптотических распределений, мы используем следующие обозначения:

$$\begin{aligned}\boldsymbol{\alpha} &= \mathrm{vec}(A_1,\ldots,A_p) && (K^2p\times 1),\\ \mathbf{A} &= \begin{bmatrix} A_1 & A_2 & \ldots & A_{p-1} & A_p\\ I_K & 0 & \ldots & 0 & 0\\ 0 & I_K & & 0 & 0\\ \vdots & & \ddots & \vdots & \vdots\\ 0 & 0 & \ldots & I_K & 0\end{bmatrix} && (Kp\times Kp),\\ \boldsymbol{\sigma} &= \mathrm{vech}(\Sigma_u) && (\frac{1}{2}K(K+1)\times 1),\end{aligned}$$

а соответствующие оценки обозначаются с символом крышки. Оператор vec является оператором векторизации, а vech – оператор который собирает элементы ниже главной диагонали. Мы также используем коммутационную матрицу $\mathbf{K}_{mn}$, определенную таким образом, чтобы для любой $(m\times n)$ матрицы G размерности $\mathbf{K}_{mn}\,\mathrm{vec}(G) = \mathrm{vec}(G')$, дуплицирующую матрицу $\mathbf{D}_m$ размерности $(m^2\times\frac{1}{2}m(m+1))$, определенную таким образом, чтобы $\mathbf{D}_m\,\mathrm{vech}(F) = \mathrm{vec}(F)$ для любой симметричной $(m\times m)$ матрицы F и матрицу исключения $\mathbf{L}_m$ размерности $(\frac{1}{2}m(m+1)\times m^2)$, определенную таким образом, чтобы $\mathrm{vech}\,F = \mathbf{L}_m\,\mathrm{vec}(F)$ для любой $(m\times m)$ матрицы F. Кроме того, $J = [I_K \ : \ 0 \ : \ \cdots \ : \ 0]$ – матрица размерности $(K\times Kp)$. С этими обозначениями в (Lütkepohl, 1990) было установлено асимптотическое следующее поведение импульсных откликов.

Предположим, что

$$\sqrt{T}\begin{bmatrix}\hat{\boldsymbol{\alpha}}-\boldsymbol{\alpha}\\ \hat{\boldsymbol{\sigma}}-\boldsymbol{\sigma}\end{bmatrix} \xrightarrow{d} N\left(0, \begin{bmatrix}\Sigma_{\hat{\boldsymbol{\alpha}}} & 0\\ 0 & \Sigma_{\hat{\boldsymbol{\sigma}}}\end{bmatrix}\right), \tag{1.37}$$

где $\Sigma_{\hat{\boldsymbol{\alpha}}} = \Gamma_Y(0)^{-1} \otimes \Sigma_u$, $\Sigma_{\hat{\boldsymbol{\sigma}}} = 2\mathbf{D}_K^+(\Sigma_u \otimes \Sigma_u)\mathbf{D}_K^{+\prime}$, $\Gamma_Y(0) = E(Y_t^0 Y_t^{0\prime})$, $Y_t^0 = (y_t - \mu, \times, y_{t-p+1} - \mu)'$ – матрица размерности $(Kp \times 1)$, $\mathbf{D}_K^+ = (\mathbf{D}_K'\mathbf{D}_K)^{-1}\mathbf{D}_K'$ – псевдообратная матрица Мура-Пенроуза от дуплицирующей матрицы $\mathbf{D}_K$.

Тогда

1. $\sqrt{T}\,\mathrm{vec}(\hat{\Phi}_i - \Phi_i) \xrightarrow{d} N(0, G_i \Sigma_{\hat{\boldsymbol{\alpha}}} G_i')$, $i = 1, 2, \ldots$, где
$$G_i = \frac{\partial\,\mathrm{vec}(\Phi_i)}{\partial \boldsymbol{\alpha}'} = \sum_{m=1}^{i-1} J(\mathbf{A}')^{i-1-m} \otimes \Phi_m.$$

2. $\sqrt{T}\,\mathrm{vec}(\hat{\Psi}_i - \Psi_i) \xrightarrow{d} N(0, F_n \Sigma_{\hat{\boldsymbol{\alpha}}} F_n')$, $i = 1, 2, \ldots$, где $F_n = G_1 + \cdots + G_n$.

3. Если $(I_K - A_1 - \cdots - A_p)$ невырождена, $\sqrt{T}\,\mathrm{vec}(\hat{\Psi}_\infty - \Psi_\infty) \xrightarrow{d} N(0, F_\infty \Sigma_{\hat{\boldsymbol{\alpha}}} F_\infty')$, где $F_\infty = \underbrace{(\Psi_\infty', \ldots, \Psi_\infty')}_{p \text{ раз}} \otimes \Psi_\infty$.

4. $\sqrt{T}\,\mathrm{vec}(\hat{\Theta}_i - \Theta_i) \xrightarrow{d} N(0, C_i \Sigma_{\hat{\boldsymbol{\alpha}}} C_i' + \bar{C}_i \Sigma_{\hat{\boldsymbol{\alpha}}} \bar{C}_i')$, $i = 0, 1, 2, \ldots$, где
$$C_0 = 0,\ C_i = (P' \otimes I_K)G_i,\ i = 1, 2, \ldots,\ \bar{C}_i = (I_K \otimes \Phi_i)H,\ i = 0, 1, \ldots,$$
и
$$\begin{aligned} H &= \frac{\partial\,\mathrm{vec}(P)}{\partial \boldsymbol{\sigma}'} = \mathbf{L}_K'\{\mathbf{L}_K[(I_K \otimes P)\mathbf{K}_{KK} + (P \otimes I_K)]\mathbf{L}_K'\}^{-1} \\ &= \mathbf{L}_K'\{\mathbf{L}_K(I_{K^2} + \mathbf{K}_{KK})(P \otimes I_K)\mathbf{L}_K'\}^{-1} \end{aligned}$$

5. $\sqrt{T}\,\mathrm{vec}(\hat{\Xi}_n - \Xi_n) \xrightarrow{d} N(0, B_n \Sigma_{\hat{\boldsymbol{\alpha}}} B_n' + \bar{B}_n \Sigma_{\hat{\boldsymbol{\sigma}}} \bar{B}_n')$, где $B_n = (P' \otimes I_K)F_n$ и $\bar{B}_n = (I_K \otimes \Psi_n)H$.

6. Если $(I_K - A_1 - \cdots - A_p)$ невырождена, $\sqrt{T}\,\mathrm{vec}(\hat{\Xi}_\infty - \Xi_\infty) \xrightarrow{d} N(0, B_\infty \Sigma_{\hat{\boldsymbol{\alpha}}} B_\infty' + \bar{B}_\infty \Sigma_{\hat{\boldsymbol{\sigma}}} \bar{B}_\infty')$, где $B_\infty = (P' \otimes I_K)F_\infty$ и $\bar{B}_\infty = (I_K \otimes \Psi_\infty)H$.

7. $\sqrt{T}\,\mathrm{vec}(\hat{\omega}_{jk,h} - \omega_{jk,h}) \xrightarrow{d} N(0, d_{jk,h} \Sigma_{\hat{\boldsymbol{\alpha}}} d_{jk,h}' + \bar{d}_{jk,h} \Sigma_{\hat{\boldsymbol{\sigma}}} \bar{d}_{jk,h}')$, $j, k = 1, \ldots, K$, $h = 1, 2, \ldots$, где
$$\begin{aligned} d_{jk,h} = \frac{2}{\mathrm{MSE}_j(h)^2} \sum_{i=0}^{h-1} \big[&\mathrm{MSE}_j(h)(e_j'\Phi_i P e_k)(e_k'P' \otimes e_j')G_i \\ &- (e_j'\Phi_i P e_k)^2 \sum_{m=0}^{h-1} e_j'\Phi_m \Sigma_u \otimes e_j' G_m\big] \end{aligned}$$
с $G_0 = 0$ и
$$\begin{aligned} \bar{d}_{jk,h} = \sum_{i=0}^{h-1} \big[&2\mathrm{MSE}_j(h)(e_j'\Phi_i P e_k)(e_k' \otimes e_j'\Phi_i) \\ &- (e_j'\Phi_i P e_k)^2 \sum_{m=0}^{h-1} e_j'\Phi_m \otimes e_j'\Phi_m \mathbf{D}_K\big]/\mathrm{MSE}_j(h)^2. \end{aligned}$$

Далее приводим несколько замечаний.

Замечание 1. Некоторые матрицы частных производных могут быть нулевыми. Например, если оценивается VAR(1)-модель, хотя истинный порядок нулевой (то есть y_t – белый шум), то

$$G_2 = J\mathbf{A}' \otimes I_K + JI_K \otimes \Phi_1 = 0,$$

поскольку $\mathbf{A} = A_1 = 0$ и $\Phi_1 = A_1 = 0$. Следовательно, для $\sqrt{T}(\hat{\Phi}_2 - \Phi_2)$ получается вырожденное предельное распределение с нулевой ковариационной матрицей. Оценивание ковариационных матриц, полученных заменой неизвестных значений их обычными оценками, может быть проблематичным, когда асимптотическое распределение вырожденное. В этом случае обычные t-статистики и доверительные интервалы могут не быть подходящими и приводящими к неверным выводам.

Замечание 2. Процесс y_t не обязательно должен быть стабильным для всех установленных результатов. Решающим улсовием является асимптотическое поведение параметров в (1.37), однако асимптотическая нормальность также может быть получена для некоторых нестационарных и нестабильных процессов. Следовательно, некоторые части утверждения выше могут быть полезны и в нестационарном контексте.

Замечание 3. Блочно-диагональная структура ковариационной матрицы асимпотического распределения в (1.37) ни в коей мере не являесят необходимой для асимптотической нормальности импульсных откликов. Фактически асимптотические рапсределения в пунктах 1-3 остаются неизменными, даже если асимптотическая ковариационная матрица не является блочно-диагональной. С другой стороны, без блочно-диагональнйо структуры простая аддитивная структура асимптотических ковариационных матриц в пунктах 4-7 пропадает, хотя можно было бы обобщить эти результаты на случай общей асимптотической ковариационной матрицы в (1.37).

Замечание 4. Зная выражения для $\Sigma_{\hat{\boldsymbol{\alpha}}}$ и $\Sigma_{\hat{\boldsymbol{\sigma}}}$ в (1.37), можно получить некотоыре упрощения в ковариационных матрицах. Например, ковариационная матрица $G_i\Sigma_{\hat{\boldsymbol{\alpha}}}G_i'$ в пункте 1 становится равна

$$\begin{aligned}
G_i\Sigma_{\hat{\boldsymbol{\alpha}}}G_i' &= \\
&= \left[\sum_{m=0}^{i-1} J(\mathbf{A}')^{i-1-m} \otimes \Phi_m\right] (\Gamma_Y(0)^{-1} \otimes \Sigma_u) \left[\sum_{n=0}^{i-1} J(\mathbf{A}')^{i-1-n} \otimes \Phi_n\right]' \\
&= \sum_{m=0}^{i-1}\sum_{n=0}^{i-1} [J(\mathbf{A}')^{i-1-m}\Gamma_Y(0)^{-1}\mathbf{A}^{i-1-n}J'] \otimes (\Phi_m\Sigma_u\Phi_n'),
\end{aligned}$$

что вычислительно более удобно, поскольку все ключенные матрицы имеют относительно небольшой размер. преимущество общей формулировки в том, что она может использовать другие матрицы $\Sigma_{\hat{\boldsymbol{\alpha}}}$.

Замечание 5. На практике для использования асимптотических рапсределений для статистических выводов неизвестные значения в ковариационных матрицах могут быть заменены на их обычные оценки (LS, ML и др.) для случая стационарного и стабильного y_t (см., однако, Замечание 1).

Замечание 6. Суммируя компоненты ошибок дисперсии прогноза по k,

$$\sum_{k=1}^{K} \omega_{jk,h} = \sum_{k=1}^{K} \hat{\omega}_{jk,h} = 1$$

для каждого j и h. Эти ограничения не принимаются во внимание в получении асимптотических распределений в пункте 7. Легко проверить, однако, что для размерности $K = 1$ стандартные ошибки, полученные в пункте 7, нулевые, как они и должны быть, поскольку компоненты дисперсии ошибок прогноза равны 1 в этом случае. Проблема в данном контексте заключается в том, что асимптотическое рапсределение $\hat{\omega}_{jk,h}$ не может быть использовано обычным способом для тестов на значимость и построения доверительных интервалов, если $\omega_{jk,h} = 0$. В этом случае из определенеия $d_{jk,h}$ и $\bar{d}_{jk,h}$ дисперсия асимптотического распределения, как легко видно, будет нулевой, следовательно, оценивание этих значений путем замены неизвестных параметров их обычными оценками может привести к t-статистикам, не имеющим асимптотически стандартное нормальное распределение, и, следовательно, они не могут использоваться обычным способом для статистических выводом (см. Замечание 1). Такое положение дел вызывает сожаление с практической точки зрения, потому что проверка значимости компонентов дисперсии ошибок пргнозов представляет особый интерес на практике. Однако следует отметить, что

$$\omega_{jk,h} = 0 \iff \theta_{jk,i} = 0 \text{ для } i = 0, \ldots, h.$$

Может быть возможно тестировать последюю гипотезу.

Замечание 7. Совместные доверительные области и тестовые статистики для проверки гипотез, которые включают несколько коэффициентов откликов, могут быть полученыиз пунктов 1-7 обычным способом. Однако должно быть принято во внимание, что, например, элементы $\hat{\Phi}_i$ и $\hat{\Phi}_j$ не будут асимптотически независимыми. Если элементы из двух или более MA-матриц включены, должно быть определено совместное распределение всех матриц. Это распределение может быть легко получено из результатов, заданных в пунктах 1-7. Например, ковариационная матрица совместного асимптотического распределения $\mathrm{vec}(\hat{\Phi}_i, \hat{\Phi}_j)$ равно

$$\frac{\partial \,\mathrm{vec}(\hat{\Phi}_i, \hat{\Phi}_j)}{\partial \boldsymbol{\alpha}'} \Sigma_{\hat{\boldsymbol{\alpha}}} \frac{\partial \,\mathrm{vec}(\hat{\Phi}_i, \hat{\Phi}_j)'}{\partial \boldsymbol{\alpha}},$$

где

$$\frac{\partial \,\mathrm{vec}(\hat{\Phi}_i, \hat{\Phi}_j)}{\partial \boldsymbol{\alpha}'} = \begin{bmatrix} \dfrac{\partial \,\mathrm{vec}(\hat{\Phi}_i)}{\partial \boldsymbol{\alpha}'} \\ \dfrac{\partial \,\mathrm{vec}(\hat{\Phi}_j)}{\partial \boldsymbol{\alpha}'} = \end{bmatrix}$$

и так далее. Мы выбрали сформулировтаь утверждение для для индивидуальных матриц коэффициентов MA, поскольку тем самым все необходимые матрицы имеют относительно небольшие размеры и, следовательно, легко вычисляемы.

Замечание 8. Обозначая jk-ые элементы Φ_i и Θ_i как $\phi_{jk,i}$ и $\theta_{jk,i}$, соответственно, очевидно инетерсующими гипотезами для $j \neq k$ являются

$$H_0 : \phi_{jk,i} = 0 \text{ для } i = 1, 2, \ldots \tag{1.38}$$

и

$$H_0 : \theta_{jk,i} = 0 \text{ для } i = 0, 1, 2, \ldots, \tag{1.39}$$

поскольку они могут интерпретироваться как гипотезы непричинности переменной k для переменной j, то есть импульс в переменной k не вызывает како-либо отклик в переменной j. Как уже было замечено в Разделе 1.1.2, это эквивалентно тому, что

$$H_0 : \phi_{jk,i} = 0 \text{ для } i = 1, 2, \ldots, p(K-1) \tag{1.40}$$

и

$$H_0 : \theta_{jk,i} = 0 \text{ для } i = 0, 1, 2, \ldots, p(K-1). \tag{1.41}$$

Используя неравенство Бонферрони, тест для (1.40) с уровнем значимости по крайней мере $100\gamma\%$ получается, отвергая H_0, если

$$|\sqrt{T}\hat{\phi}_{jk,i}/\hat{\sigma}_{\phi_{jk}}(i)| > z_{(\gamma/2p(K-1))} \tag{1.42}$$

для по крайней мере одного $i \in \{1, 2, \ldots, p(K-1)\}$. Здесь $z_{(\gamma)}$ является верхней 100γ процентной точкой стандартного нормального распределения, а $\hat{\sigma}_{\phi_{jk}}(i)$ является оценкой асимптотического стандартного отклонения $\sigma_{\phi_{jk}}(i)$ для $\sqrt{T}\hat{\phi}_{jk,i}$. Чтобы получить асимптотическое стандартное нормальное распределение t-отношения $\sqrt{T}\hat{\phi}_{jk,i}/\hat{\sigma}_{\phi_{jk}}(i)$, дисперсия $\sigma^2_{\phi_{jk}}(i)$ должна быть ненулевой.

Тест (1.41) с уровнем значимости по крайней мере γ получается, отвергая H_0, если

$$|\sqrt{T}\hat{\theta}_{jk,i}/\hat{\sigma}_{\theta_{jk}}(i)| \begin{cases} > z_{(\gamma/2(pK-p+1))} & \text{для по крайней мере одного } i \in \{0, 1, 2, \ldots, p(K-1)\}, \text{ если } j > k \\ < z_{(\gamma/2(pK-p))} & \text{для по крайней мере одного } i \in \{0, 1, 2, \ldots, p(K-1)\}, \text{ если } j < k \end{cases} \tag{1.43}$$

Здесь $\hat{\sigma}_{\theta_{jk}}(i)$ является состоятельной оценкой стандартного отклонения асимптотического распределения $\sqrt{T}\hat{\theta}_{jk,i}$, которое предполагается ненулевым.

Тест, основанный на принципе Бонферрони, может иметь достаточно низкую мощность, посколькy фактический уровень значимости может быть намного меньше, чем заданная верхняя гарница. Следовательно, тест, основанный на χ^2- или F-статистиках, будет предпочтительнее. К сожалению, такие тесты не являются легкодоступными для нынешней ситуации. Для большего обсуждения см. также (Lütkepohl, 1990), а для различных подходов представления неопределенности в оцененных импульсных откликов см. (Sims and Zha, 1999).

Альтернативой использования асимптотических результатов, представленных выше, является методы бутстрепа. Хотя эти методы достаточно затратны с точки зрения вычислительного времени, они часто используются для оценивания свойств функций импульсных откликов (см., например, (Runkle, 1987) и (Kilian, 1998, 1999)), поскольку не требуют явного получения выражений асимптотических дисперсий.

В текущей ситуации существует несколько подходов для симуляций. Один из них предполагает конкретный вид распределения процесса белого шума, например, $u_t \sim N(0, \hat{\Sigma}_u)$ и генерирует большое число реализаций временного ряда, основанное на оценивании VAR-коэффициентов. По этим временным рядам затем оценивается новый набор коэффициентов, соответствующие импульсные отклики и компоненты дисперсии ошибок прогноза. Полученные этим способом эмпирические распределения могут быть использованы для исследования фактических распределений интересующих параметров.

Если не сделано никаких предположений о распределении инноваций, можно использовать методы бутстрепа путем создания нового набора остатков, извлеченных из оцененных остатков. Кратко рассмотрим основные положения данного метода. Сначала оценивается интересующая модель. Если оцененные остатки обозначаются как $\hat{u}_t$, вычисляются центрированные остатки $\hat{u}_1 - \bar{\hat{u}}, \ldots, \hat{u}_T - \bar{\hat{u}}$, а бутстреп=остатки, $u_1^*, \ldots, u_T^*$ вычисляются путем случайного извлечения с повторением из центрированных остатков. Затем u_t^* используются для вычисления бутстреп-временных рядов рекурсивно, начиная с заданных начальных значений $y_{-p+1}, \ldots, y_0$ для модели с p лагами. Модель переоценивается и интересующие параметры определяются на основе оценок параметров. Повторяя этот шаг много раз, мы получим эмпирическое бутстреп-распределение интересующих параметров. Из этого распределения могут быть получены необходимые квантили и, следовательно, доверительные интервалы.

Обозначим символами ϕ, $\hat{\phi}$ b $\hat{\phi}^*$, соответственно, коэффициент импульсного отклика, его оценку и соответствующую бутстреп-оценку. Обычно рассматрвиаются следующие доверительные интервалы (см., например, (Benkwitz, Lütkepohl and Wolters, 2001)):

- *Стандартный процентильный интервал*
  Самым употребимым на практике методом построения доверительных интервалов для импульсных откликов является интервал
  $$CI_S = [s^*_{\gamma/2}, s^*_{(1-\gamma/2)}],$$
  где $s^*_{\gamma/2}$ и $s^*_{(1-\gamma/2)}$ – $\gamma/2$- и $(1-\gamma/2)$-квантили, соответственно, эмпирического распределения $\hat{\phi}^*$. Интервал CI_S является процентильным доверительным интервалом, опсианным, например, в (Efron and Tibshirani, 1993).

- *Процентильный интервал Холла*
  В (Hall, 1992) была представлена обычная бутстреп-аналогия, утверждающая что распределение $(\hat{\phi} - \phi)$ приближенное равно распределению $(\hat{\phi}^* - \hat{\phi})$на больших выборках. Из этого результата вытекает, что можно получить интервал
  $$CI_H = [\hat{\phi} - t^*_{(1-\gamma/2)}, \hat{\phi} - t^*_{\gamma/2}],$$
  где $t^*_{(1-\gamma/2)}$ и $t^*_{\gamma/2}$ – $(1-\gamma/2)$- и $\gamma/2$-квантили, соответственно, эмпирического распределения $(\hat{\phi}^* - \hat{\phi})$.

- *Стюдентизированный интервал Холла*
  В некоторых ситуациях может быть предпочтительнее использовать стьюдентизированную статистику $(\hat{\phi}-\phi)/\widehat{\mathrm{var}}(\hat{\phi})^{1/2}$ для построения доверительного интервала. В этом случае используются бутстреп-квантили $t^{**}_{\gamma/2}$ и $t^{**}_{1-\gamma/2}$ из распределения $(\hat{\phi}^* - \hat{\phi})/\widehat{\mathrm{var}}(\hat{\phi}^*)^{1/2}$ для построения интервала
  $$CI_{SH} = [\hat{\phi} - t^{**}_{(1-\gamma/2)}\widehat{\mathrm{var}}(\hat{\phi}^*)^{1/2}, \hat{\phi} - t^{**}_{\gamma/2}\widehat{\mathrm{var}}(\hat{\phi}^*)^{1/2}].$$
  В этом подходе дисперсии оцениваются через бутстреп внутри каждого бутстреп-повторения.

Нужно подчеркнуть, что бутстреп не решает проблемы вырожденного асимптотического распределения импульсных откликов и получающихся потенциально неверных статистических выводов. Если асимптотическое распределение вырожденное, бутстреп может давать осмысленные доверительные интервалы. Более детальное обсуждение этой проблемы см. в (Benkwitz, Lütkepohl and Neumann, 2000), где также рассматриваются методы для коррекции асимптотических выводов. Одно из возможных решений – рассмотреть бутстреп-процедуры, которые адаптируются к типу вырожденности предельного распределения. Например, можно использовтаь subsampling для оценивания скорости сходимости оценок параметров в дополнение к параметрам модели. Эти и другие методы имеют, однако недостатки в эмпирическихз приложениях. Или они не очень практичны для процессов реалистичной размерности и авторегрессионного порядка, или они не применяются хорошо в выборках типичного объема. Второе возможное решение – удалить все точки, где может происходить вырожденность асимптотического распределения до проведения анализа импульсных откликов. В данном случае это предложение сводится к определению всех нулевых коэффициентов в первом этапе анализа и применению полученных нулевых ограничений на следующем этапе, где в результате subset model оценивается и используется для вычисления импульсных откликов. Третья возможность обойти проблему – рассмотреть VAR-процесс бесконечного порядка, допускающий увеличение порядка, когда становится достыпно большее количество информации (другими словами, порядок модели предполагается расходящимся к бесконечности с соответствующей скоростью при увеличении объема выборки).

Другая проблема посроения доверительных интервалов была указана в (Lütkepohl, 2013b). Известно, что шоки идентифицированы с точностью до знака, и эта особенность может привести к вводящим в заблуждение доверительным интервалам для импульсных откликов, если для их построения используются методы симуляций, в частности, бутстреп. Рассмотрим следующий иллюстративный пример. Пусть b является i-ым столбцом матрицы B, то есть b представляет вектор текущих эффектов i-го шока. Должно быть всегда сделано фиксирование знака для элемента, который явно отличается от нуля, поскольку тогда обеспечивание определенного знака не является ограничительным для других элементов. Если фиксирование знака сделано через элемент, который близок к нулю, это может привести к искаженным симулированным доверительным интервалам других параметров в b. Предположм, что $b = (b_1, b_2)'$ – просто двумерный вектор, и фиксирование знака сделано через b_2, хотя истинное знаечние $b_2 = 0$. Дополнительно предположим, что фактическое распределение оценки $\hat{b}$ есть $N(b, \Sigma_b)$. Тогда $(1-\gamma)$-доверительный интервал для b_1 должен быть

$$\hat{b}_1 - c_{1-\gamma/2}\sigma_{11}, \hat{b}_1 + c_{1-\gamma/2}\sigma_{11},$$

где c_η – η-квантиль стандартного нормального распределения, а σ_{11}^2 – верхний левый элемент в Σ_b. Однако, если мы строим доверительный интервал следующим симуляционным методом, может быть получен достаточно отличный, более широкий доверительный интервал. Пусть $b^{(n)}$ для $n = 1, \ldots, N$ следующий:

1. Извлечем $\tilde{b} = (\tilde{b}_1, \tilde{b}_2)'$ из распределения $N(b, \Sigma_b)$.

2. Положим
$$b^{(n)} = \begin{cases} \tilde{b}, & \text{если } \tilde{b}_2 \geq 0, \\ -\tilde{b}, & \text{если } \tilde{b}_2 < 0. \end{cases}$$

Выбирая доверительный интервал для b_1 из эмпирического распределения как

$$[b_1^{(N\gamma/2)}, b_1^{(N(1-\gamma/2))}] \tag{1.44}$$

может очевидно привести к более широкому доверительному интервалу с не очень большой вероятностью включения, поскольку знак $\hat{b}_1$ меняется с веоятностью около 0.5, если истинное значение b_1 ясно и значимо положительно. Фактическая вероятность покрытия и длина интервала (1.44) зависит, конечно, от истинных значений параметра. Проблема заключаестя в том, что b_2 будет давать положительные и отрицательные знаечния в симуляциях. Всякий раз, когда получается отрицательное значение, оно будет конвертировано в положительное значение. В то же время b_1 будет также отражается в другую сторону от начала координат. Фактически пара $(b_1, b_2)'$ из части пространства параметров, связанной с различными локальными максимумами функции правдоподобия, выбрана этим способом. Когда b_1 всегда положительно в области пространства параметров, соответствующей первому оптимуму функции правдоподобия, это может привести к очень широкому доверительному интервалу для b_1, поскольку включены значения b_1 из двух различных локальных оптимумов функции правдоподобия, даже если область вариации b_1 мала в окрестности одного локального оптимума. Если b_1 была использована для фиксирования знака, ситуация будет достаточно отличной. Эти рассмотрения предлагают, что хорошей стратегией являестя исключение фиксирвоания знака, основанного на неточно оцененных impact effects. Другими словами, можно захотеть проверить t-отношения оцененных элементов в B и рассмотреть переменные для фиксирвоания знака, которые имеют impact effects с большими (по абсолютному значению) t-статистиками. В (Lütkepohl, 2013b) рассматриваются два примера с применением этой методологии.

Доверительные дапазоны

Обычно доверительные диапазоны строятся относительно индивидуальных коэффициентов импульсных откликов. Если индивидуальные доверительные интервалы для заданного уровня значимости строятся около коэффициентов импульсных откликов для каждого горизонта отклика отдельно, то нет гарантии, что общий уровень покрытия для всех импульсных откликов одной переменной будет соответствовать предварительно специфицированному уровню доверия. Другими словами, вероятность диапазона, содержащего истинную функцию импульсного отклика конкретной переменной будет, в общем случае, не равна $1-\gamma$, если доверительный диапазон строится как объединение индивидуальных $(1-\gamma) \times 100\%$-доверительных интервалов, то есть игнорируется совместное рапсределение. Такие доверительные диапазоны называют *наивными* диапазонами. Следовательно, желательно построить доверительные диапазоны с общей предварительно специфицированной вероятностью покрытия. Соответствующие методы для этого рассматриваются в (Lütkepohl, Staszewska-Bystrova and Winker, 2013), ограничиваясь классическим оцениванием (хотя соответствующие Байесовские доверительные диапазоны уже были рассмотрены, например, в (Sims and Zha, 1999)).

Многие из методов, рассматрвиаемые в (Lütkepohl, Staszewska-Bystrova and Winker, 2013), основаны на стандартном остаточном бутстрепе. Также вместо обычных OLS-оценок VAR-коэффициентов используются скорректированные на смещение оценки, поскольку они оказались лучше работающими в этом контексте (см. (Kilian, 1998, 2001)). Рассматриваются следующие процедуры:

1. *Наивный диапазон (naive)*
   Как уже было отмечено, наивный доверительный диапазон задается собранием доверительных интервалов для индивидуальных коэффициентов импульсных откликов. Пусть $s^*_{\gamma/2}$ и $s^*_{1-\gamma/2}$ обозначают $\gamma/2$ и $1-\gamma/2$ бутстреп-распределения оцененного коэффициента импульсного отклика на горизонте h ($\phi_{ij,h}$), где $h = 0, \dots, H$. Тогда $(1-\gamma) \times 100\%$-доверительный интервал для $\phi_{ij,h}$ просто задается как
   $$s^*_{\gamma/2}, s^*_{1-\gamma/2}$$
   для всех горизонтов $h = 0, \dots, H$. Хотя метод не гарантирует совместную вероятность покрытия $1-\gamma$, он часто используется в прикладных работах возможно на том основании, что доверительные диапазоны призваны просто служить указанием неопределенности выборки, связанной с оцененными импульсными откликами.

2. *Традиционный диапазон Бонферрони (B)*
   В отличие от наивных диапазонов, диапазоны Бонферрони принимают во внимание стохастическую зависимость в оцененных коэффициентах импульсных откликах путем увеличения величины диапазонов, так чтобы гарантировать предварительно специфицированную нижнюю границу уровня доверия по крайней мере асимптотически. То есть они в общем случае будут консервативными. Формально диапазоны Бонферрони строятся по доверительным интервалам для индивидуальных коэффициентов импульсных откликов. Номинальный уровень доверия каждого интервала задается как $(1-\gamma/L) \times 100\%$, где $L = H + 1$ если эффект вклада не ограничивается нулем, и $L = H$, если эффект вклада ограничен нулем. Для каждого горизонта h ($h = 0, \dots, H$) $(1-\gamma/L) \times 100\%$-доверительный интервал для $\phi_{ij,h}$ задается как
   $$s^*_{\gamma/2L}, s^*_{1-\gamma/2},$$
   где $s^*_{\gamma/2L}$ и $s^*_{1-\gamma/2}$ являются $\gamma/2L$ и $1-\gamma/2L$ квантили бутстреп-распределения оценки $\phi_{ij,h}$.

3. *Скорректированный диапазон Бонферрони (B-adj)*
   Поскольку традиционные диапазоны Бонферрони являются консервативными по построению, в (Lütkepohl, Staszewska-Bystrova and Winker, 2013) предлагается коррекция, которая уменьшает диапазоны, но все еще соблюдает общий уровень покрытия относительно бутстреп импульсных откликов. В дальнейшем число бутстреп повторений обозначается как b. Сначала диапазон Бонферрони строится как в пункте 2. Этот диапазон покрывает и полные, и неполные бутстреп функции откликов. Бутстреп импульсные отклики полностью внутри диапазоны идентифицируются и сохраняются, и новый интервал получается как граница этих функций. Получившийся в результате диапазон является более узким, или по крайней мере не шире, чем традиционный диапазон Бонферрони. На последнем шаге число бутстреп импульсных откликов, покрытых диапазоном, вычисляется и обозначается как nb. Если $nb > (1-\gamma)b$, применяется последовательная процедура, направленная на устранение $nb - (1-\gamma)b$ бутстреп импульсных откликов. На каждом шаге бутстреп импульсные отклики идентифицируются, чтобы предоставлять по крайней мере одну точку на текущих диапазоных. Существует самое большее $2(H+1)$ таким функций. Функция, которая наибольшим образом способствует размеру текущего диапазона (измеренного как сумма ширин индивидуальных интервалов) отвергается. Процедура останав-

ливается после того, как устраняется $nb - (1-\gamma)b$ функций. Диапазон получается как граница оставшихся $(1-\gamma)b$ бутстреп импульсных откликов.

Эта процедура называется скорректирвоанном методом Бонферрони и сокращается как B-adj в дальнейшем. Она имеет асимптотически корректный уровень покрытия, всякий раз когда бутстреп предлагает доверительное множество с асимптотически точным уровнем покрытия, поскольку данный метод сохраняет $(1-\gamma)b$ бутстреп выборок и исопльзует внешнюю границу доверительного множества.

4. *Диапазон Шеффе (Scheffé)*
Так называемый диапазон Шеффе был преложен в (Jordà, 2009). Он основан на асимптотическом распределении импульсных откликов и вычисляется, используя нисходящую процедуру (step-down procedure), предложенную в (Jordà and Marcellino, 2010) в контексте построения совместных диапазонов прогноза. Для описания этого метода более формально мы исопльзуем следующее обозначение: $\phi_{ij} = (\phi_{ij,0}, \ldots, \phi_{ij,H})$ является $(H+1)$-мерным вектором откликов переменной i на j-ый шок, $\phi_{ij,1,\ldots,H} = (\phi_{ij,1}, \ldots, \phi_{ij,H})$ обозначает соответствующий вектор, когда эффект вклада ограничен нулем и, следовательно, не оценивается. Оцененную асимптотическую ковариационную матрицу для $\hat{\phi}_{ij}$, $\hat{\Omega}_{ij}/T$ можно разложить как $\hat{\phi}_{ij}$, $\hat{\Omega}_{ij}/T = PP'$, где P – нижнетреугольная матрциа разложения Холецкого. $c^2_\gamma(h+1)$ являеся $(1-\gamma)$-квантилем рапсределения χ^2 с $h+1$ степенью свободы, а $\left[\sqrt{c^2_\gamma(h+1)/(h+1)}\right]$ – $((H+1)\times 1)$-вектор, $(h+1)$-ый элемент которого равен $\sqrt{c^2_\gamma(h+1)/(h+1)}$ (для $h = 0, 1, \ldots, H$). Используя эти обозначения, диапазон задается как

$$\hat{\phi}_{ij} \pm P\left[\sqrt{\frac{c^2_\gamma(h+1)}{h+1}}\right], \tag{1.45}$$

если эффект вклада шока неограничен, и как

$$\hat{\phi}_{ij,1,\ldots,H} \pm P^*\left[\sqrt{\frac{c^2_\gamma(h)}{h}}\right], \tag{1.46}$$

если эффект вклада шока равен нулю по построению. Здесь P^* являестя нижнетреугольной матрицей, полученной из разложения Холецкого матрицы, полученной из $\hat{\Omega}_{ij}/T$ путем удаления первого столбца и первой строки, а $\left[\sqrt{c^2_\gamma(h)/h}\right]$ – $(H\times 1)$-вектор с h-ым элементом, равным $\sqrt{c^2_\gamma(h)/h}$ (для $h = 1, 2, \ldots, H$).

Заметим, что диапазон Шеффе основан на оценках наименьших квадратов VAR-коэффициентов, как изначально предлагалось, а не на скорректированных на смещение оценках. Также для больших H (например, $H = 10$) может возникнуть проблема получения разложения Холецкого, и в (Lütkepohl, Staszewska-Bystrova and Winker, 2013) предлагается заменять все собствнные знаечния оцененной ковариационной матрицы, коорые меьше или раны 0, на значение 10^{-6}.

5. *Диапазон соседней траектории (neighbouring path band, NP)*
   Метод соседней траектории (NP), предложенная в (Staszewska, 2007), является процедурой, основнаной на b бутстреп траекториях откликов $\hat{\phi}_{ij}^{c*}$. В последовательном алгоритме удаляется γb бутстреп траекторий. Граница оставшихся $(1-\gamma)b$ траекторий дает доверительный диапазон. While eliminating individual paths, метод направлен на достижения узкого доверительного диапазона (все еще содержащего $((1-\gamma)b$ бутстреп траекторий). Следовательно, траектории кандидаты для удаления на каждом шаге являются теми, которые содержат по крайней мере одну точку вне границы других бутстреп траекторий. Из группы этих траекторий кандидатов алгоритм выбирает одну самый дальний (в смысле евклидова расстояния) от импульсного отклика, полученного из исходных данных. После устранения этой крайней траектории, шаг удаления повторяется до тех пор, пока не будут исключены γb траекторий.

6. *Диапазон Wolf & Wunderli (Wolf & Wunderli band, WW)*
   В (Wolf and Wunderli, 2012) были предложены диапазоны для multiple horizon forecasts, и этот метод был адаптирован в (Lütkepohl, Staszewska-Bystrova and Winker, 2013) для построения диапазонов импульсных откликов. Диапазон основан на b бутстреп тракториях откликов $\hat{\phi}_{ij}^{c*}$ и вычисляется как

$$\hat{\phi}_{ij}^{c} \pm \left[d_{|\cdot|,1-\gamma}^{\max,c} \sqrt{\hat{\omega}_{h+1,h+1}^{c}} \right],$$

   где $\left[d_{|\cdot|,1-\gamma}^{\max,c} \sqrt{\hat{\omega}_{h+1,h+1}^{c}} \right]$ – вектор размерностью $((H+1)\times 1)$, $(h+1)$-ый элемент которого (для $h=0,1,\ldots,H$) задается как $d_{|\cdot|,1-\gamma}^{\max,c} \sqrt{\hat{\omega}_{h+1,h+1}^{c}}$. Здесь $\hat{\omega}_{h+1,h+1}^{c}$ является $(h+1)$-ым элементом главной диагонали $\hat{\Omega}_{ij}^{c}$, а $d_{|\cdot|,1-\gamma}^{\max,c}$ вычисляеся следующим образом:

   - В отдельной итерации бутстреп-метода вычисляются

$$\hat{s}_{ij}^{c*}(h) = \frac{\hat{\phi}_{ij,h}^{c*} - \hat{\phi}_{ij,h}^{c}}{\sqrt{\hat{\omega}_{h+1,h+1}^{c*}}}, \; h=0,\ldots,H.$$

     Вычисляется величина $\max_{|\cdot|}^{c*} \equiv \max(|\hat{S}^{c*}(H)|)$, где $\hat{S}^{c*}(H) = (\hat{s}_{ij}^{c*}(0),\ldots,\hat{s}_{ij}^{c*}(H))'$.
   - После применения b итераций бутстреп метода и получения $\max_{|\cdot|,1}^{c*},\ldots,\max_{|\cdot|,b}^{c*}$ вычисляется $d_{|\cdot|,1-\gamma}^{\max,c}$, обозначающая эмпирическую квантиль уровня $(1-\gamma)$ статистик $\max_{|\cdot|,1}^{c*},\ldots,\max_{|\cdot|,b}^{c*}$.

Исходя из результатов симуляци, авторы приходят к слеудующим общим выводам:

1. Наивный и Шеффе методы могут иметь очень плохие уровни покрытия, намного меньшие номинального уровня. Это верно в частности для метода Шеффе. Эти два метода не могут быть рекомендованы, если желателен хороший уровень покрытия. Конечно, метод Шеффе должен быть dismissed для построения доверительных диапазонов для импульсных откликов.

2. Метод WW может давать чрезвычайно широкие доверительные диапазоны по сравнению с методами Бонферрони и NP, особенно когда объем выборки мал. Таким образом, этот метод должен быть использован с осторожностью, в частности для относительно малых выборок.

3. Традиционный метод Бонферони, как парвило, производит несколько более широкие диапазоны, чем методы B-adj и NP, особенно если propagation horizon is large.

4. Методы B-adj и NP примняются хорошо в терминах уровня покрытия и ширины диапазона для большинства сценариев. Исключением являетcя процессы с единичным корнем, когда объем выборки мал.

Таким образом, эксперименты Монте-Карло подтверждают, что метод B-adj достаточно конкурентен, и в некоторых отношениях даже лучший. В то время как он имеет хорошее покрытие, он также приводит к достаточно малым диапазонам. Метод NP является ближайшим конкурентом, однако, до сих пор ему не хватает теоретической базы.

1.1.7 Другие возможности идентификации структурных шоков

Идентификация через сдвиги в волатильности

Одна из возможностей идентификации структурных шоков использует предположение о том, что существует по крайней мере один сдвиг в волатильности остатков, и, следовательно, остатки в модели будут гетероскедастичными. Этот подход был исопльзован в (Rigobon, 2003), (Rigobon and Sack, 2003) и (Lanne and Lütkepohl, 2008*a*,*b*). Рассмотрение в контексте обобщенной модели было предложено в (Lütkepohl, 2013*a*). Предположим, что существует единственный сдвиг в волатильности переменных в течение периода выборки. Следовательно, предположим, что у нас есть выборка объема T и сдвиг в волатильности шоков происходит, скажем, в период T_B:

$$\Sigma_u = \begin{cases} \Sigma_1 & \text{для } t = 1, \dots, T_B - 1, \\ \Sigma_2 & \text{для } t = T_B, \dots, T. \end{cases} \tag{1.47}$$

Ковариационные матрицы Σ_1 и Σ_2 можно диагонализировать одновременно, то есть существует $(K \times K)$-матрица W и диагональная матрица $\Psi = \text{diag}(\psi_1, \dots, \psi_K)$ с положительными диагональными элементами ψ_i, $i = 1, \dots, K$, такие что $\Sigma_1 = WW'$ и $\Sigma_2 = W\Psi W'$. Здесь диагональные элементы Ψ отражают сдвиги в дисперсиях шоков после возможно произошедшего изменения в волатильности. Фактически сдвигв волатильности произошел, если значения ψ_i отличаются от единицы. В (Lanne and Lütkepohl, 2008*b*) было показано, что W является единственной за исключением перемены знака, если все ψ_i различны и упорядочены некоторым способом (см. также (Lanne, Lütkepohl and Maciejowska, 2010)). Таким образом, если мы выбираем $\mathsf{B} = W$, мы получаем единственность шоков ε_t (за исключением перемены знака) без необходимости делать дополнительные идентификационные предположения. Структурные шоки имеют тогда единичную ковариационную матрицу в первом режиме и Ψ во втором режиме, то есть

$$\Sigma_\varepsilon = \begin{cases} I_K & \text{для } t = 1, \dots, T_B - 1, \\ \Psi & \text{для } t = T_B, \dots, T. \end{cases} \tag{1.48}$$

Следовательно, они ортогональны обоим режимам. Это является достаточным для идетификации шоков.

Рассмотрим иллюстративный пример, предложенный в (Lütkepohl, 2013*a*). пусть у нас есть двумерная система, такая что

$$u_t = \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

Для этого случая соотношения $\Sigma_1 = WW'$ и $\Sigma_2 = W\Psi W'$ соответствуют

$$\begin{bmatrix} \sigma^2_{1,1} & \sigma_{12,1} \\ \sigma_{12,1} & \sigma^2_{2,1} \end{bmatrix} = \begin{bmatrix} b^2_{11} + b^2_{12} & b_{11}b_{21} + b_{12}b_{22} \\ b_{11}b_{21} + b_{12}b_{22} & b^2_{21} + b^2_{22} \end{bmatrix}$$

и

$$\begin{bmatrix} \sigma^2_{1,2} & \sigma_{12,2} \\ \sigma_{12,2} & \sigma^2_{2,2} \end{bmatrix} = \begin{bmatrix} \lambda_1 b^2_{11} + \lambda_2 b^2_{12} & \lambda_1 b_{11}b_{21} + \lambda_2 b_{12}b_{22} \\ \lambda_1 b_{11}b_{21} + \lambda_2 b_{12}b_{22} & \lambda_1 b^2_{21} + \lambda_2 b^2_{22} \end{bmatrix}.$$

Таким образом, мы получаем шесть соотношений:

$$\begin{aligned} \sigma^2_{1,1} &= b^2_{11} + b^2_{12}, \\ \sigma_{12,1} &= b_{11}b_{21} + b_{12}b_{22}, \\ \sigma^2_{2,1} &= b^2_{21} + b^2_{22}, \\ \sigma^2_{1,2} &= \lambda_1 b^2_{11} + \lambda_2 b^2_{12}, \\ \sigma_{12,2} &= \lambda_1 b_{11}b_{21} + \lambda_2 b_{12}b_{22}, \\ \sigma^2_{2,2} &= \lambda_1 b^2_{21} + \lambda_2 b^2_{22}, \end{aligned}$$

которые мы можем решить относительно шести структурных параметров b_{11}, b_{12}, b_{11}, b_{22}, λ_1, λ_2. Здесь последние два параметра являюстя диагональными элементами матрицы Ψ. Решение единственное (до знака), если λ_1 и λ_2 различны и упорядочены некоторым конкретным образом, например, $\lambda_1 < \lambda_2$. Заметим, что дисперсии структурных шоков нормализованы на единицу в первой части выборки. Следовательно, λ_i указывает на изменение дисперсии при переключении с первого на второй режим. Другими словами, они интерпретируются как относительные диспесии во втором режиме выборки. Требование единственности только означает, что сдвиг в дисперсии не является однородным среди обеих переменных. Фактически только дисперсия одной переменной должна измениться. Этого уже достаточно для единственности решения.

Требование, что все все ψ_i различны, удовлетворяется, если сдвиги в волатильности не пропорциональны во всех переменных. Даже если волатильность в одном из шоков не меняется вообще, то есть одно из ψ_i может быть единицей, значения ψ_i могут, конечно, быть различными, и это является существенным требованием для единственности B. Любые другие ограничения на B будут сверхидентифицирующими. Так как сдвиг в дисперсии являестя тестируемым предположением, мы не должны полагаться исключительно на информацию из экономической теории или других ресурсов, чтобы гарантировать идентификацию структурных шоков. Вместо этого мы можем использовать информацию из данных и применить статистические процедуры для получения идентификации.

Для получения единственности W и, следовательно, B, мы можем, например, упорядочить ψ_i от наименьшего к наибольшему. Если ограничения накладываются на B, они могут быть только совместимы с одним конкретным порядком в ψ_i и не обязательно с порядком согласно размеру. Следовательно, в оцеривании с ограничениями на B мы не накладываем какое-нибудь конкретное упорядочивание на ψ_i, но позволяем данным решать, какое являестя лучшим. Это не являестя проблемой, поскольку локальная идентификация гарантируется любым упорядочиванием ψ_i, при условии что все они различны.

Факт о том, что всегда возможно поменять знаки всех элементов в единственном столбце матрицы $W = \mathsf{B}$ без влияния на правдоподобие, не являестя проблемой в данном контексте, поскольку для выполнения асимптотической теории нам нужна только локльная идентификация, которая гарантируется несмотря на перемены знака. Для практических целей изменение всех знаков в столбце в W просто означает, что рассматривается отрицательный шок, если шоки изначально положительны, и наоборот.

Также возможно учитывать возможно больше одного сдвига в волатильности. Если существует $n+1$ различных режимов, и ковариации в этих режимах равны $WW', W\Psi_1 W', \dots, W\Psi_n W'$, где все Ψ_i являются диагональными матрциами, единственность W (до знака) гарантируется, например, если все диагональные элементы только в одной из Ψ_i различны, и м можем снова выбирать $\mathsf{B} = W$ (см. также (Lütkepohl, 2013a, Section 3.1)). Проблемы, возможные в данной постановке, заключаются в следующем: короткие режимы, что может привести к плохо интерпретируемым оценкам, а также неизвестные даты переключения этих режимов (см. (Ehrmann, Fratzscher and Rigobon, 2011), где производится предварительное тестирвоание дат сдвигов).

Оценивается модель посредством метода максимального правдоподобия. Определим оценки ковариационных матриц в двух режимах как

$$\tilde{\Sigma}_1 = \frac{1}{T_B - 1}\sum_{t=1}^{T_B-1} \hat{u}_t \hat{u}_t' \quad \text{и} \quad \tilde{\Sigma}_2 = \frac{1}{T - T_B + 1}\sum_{t=T_B}^{T} \hat{u}_t \hat{u}_t',$$

где $\hat{u}_t$ – остатки от оценивания VAR(p)-модели. Заменяя VAR-параметры в гауссовской функции правдоподобия на их OLS-оценки, мы получаем концентрированную функцию логарифма правдоподобия вида

$$\log L_H = \frac{T_B - 1}{2}\left(\log\det(WW') + \operatorname{tr}\left\{\tilde{\Sigma}_1 (WW')^{-1}\right\}\right) \\ - \frac{T - T_B + 1}{2}\left(\log\det(W\Psi W') + \operatorname{tr}\left\{\tilde{\Sigma}_1 (W\Psi W')^{-1}\right\}\right). \quad (1.49)$$

Однако полученные оценки не являются ML-оценками, поскольку OLS-оценки VAR-коэффициентов не являются ML-оценками в данном случае, посколкьу они не принимают во внимание гетероскедастичнсть ошибок. Аналогично LR-тест на сверхидетифицирующие ограничения не являестя LR-тестом в обычном смысле, поскольку он основан на максимизации псевдо концентрированной функции правдоподобия. Однако LR-тест все еще имеет обычные асимптотические свойства стандартного LR-теста, и все еще можно использовать распределение χ^2 для этого теста.

Вторая возможность идентификации через изменения в волатильности – предположить не гетероскедастичные ошибки, а условно гетероскедастичные. В (Normandin and Phaneuf, 2004) предлагается использовать многомерную GARCH-модель для учета изменения в волатильности. пусть ошибки в приведенной форме порождаются многомерной GARCH (MGARCH) процессом вида

$$\Sigma_{u,t|t-1} = E(u_t u_t' | u_{t-1}, \dots) = W \Sigma_{\varepsilon,t|t-1} W', \tag{1.50}$$

где $\Sigma_{\varepsilon,t|t-1} = \mathrm{diag}(\sigma^2_{1,t|t-1}, \dots, \sigma^2_{K,t|t-1})$ является диагональной матрицей с

$$\sigma^2_{k,t|t-1} = \gamma_{k0} + \sum_{j=1}^{q} \gamma_{kj} \varepsilon^2_{t-j} + \sum_{j=1}^{s} g_{kj} \sigma^2_{k,t-j|t-j-1}, \; k = 1, \dots, K. \tag{1.51}$$

Другими словами, предполагается, что структурные шоки ε_t instantaneously некоррелированными и имеют диагональную MGARCH(q, s)-структуру (см. (van der Weide, 2002) и (**?**)). Для простоты предполагается MGARCH(1,1)-модель.

В (Sentana and Fiorentini, 2001) предлагается общие результаты, которые влекут идентификацию структурных шоков (за исключением перестановок и изменений знаков), если все, кроме одной компоненты фактически условно гетероскедастичны и $\Gamma'\Gamma$ обратима, где $(\sigma^2_{k,10}, \dots, \sigma^2_{k,T|T-1})$ – k-ая строка в Γ'. Другими словами, в случае GARCH(1,1) у нас есть самое большее один $k \in \{1, \dots, K\}$, в котором $\gamma_{k1} = g_{k1} = 0$. Обратимость $\Gamma'\Gamma$ гарантирует, что условная гетероскедастичность не обусловлена небольшим количеством процессов как в факторной модели, где единственный процесс может определять развитие большого множества переменных. Конечно, это требование соответствует условию, что должна быть существенная неоднородность в измнениях режимов волатильности, когда существует только конечное число режимов волатильности. При условии, что идентификационные условию выполняются, структура (1.50) гарантирует, что если мы выберем $\mathsf{B} = W$, структурные шоки будут единственными, а импульсные отклики, включающие эффекты воздействия (impact effects), являются инвариантными по времени.

Для оценивания параметров в (Normandin and Phaneuf, 2004) предлагается двухшаговая процедура, в которой на первом шаге оценивается VAR(p)-модель, а затем путем максимизации соответствующей функции правдоподобия получаются оценки параметров GARCH и структурных параметров. Другими словами, матрица B и параметры GARCH оцениваются путем максимизации

$$\log L_{GARCH}(W, \boldsymbol{\gamma}) = -\frac{1}{2} \sum_{t=1}^{T} [\log \det(W \Sigma_{\varepsilon,t|t-1} W') + \hat{u}_t' (W \Sigma_{\varepsilon,t|t-1} W')^{-1} \hat{u}_t], \tag{1.52}$$

где $\hat{u}_t$ – остатки от оценивания VAR(p) на первом шаге, а $\boldsymbol{\gamma}$ – вектор GARCH-параметров. В (Bouakez and Normandin, 2010) было замечено, что полная оптимизация (1.52) может быть сложной в системах высокой размерности. В этом случае могут быть необходимы дополнительные упрощения. например, можно разделить систему и ввести дополнительные шаги в оценивание, в частности, если допустимы соответствующие ограничения на B.

В принципе, условия идентификации можно проверитьс татистическими методами. В частности, можно тестировать, являются ли параметры GARCH в (1.51) статистически значимыми. Кроме того, можно проверить невырожденность матрицы $\Gamma'\Gamma$.

Третья возможность идентификации через изменения в волатильности предлагает конечное число режимов волатильсноти, которые порождаются эндогенно через марковский процесс. Данный подход (в контексте идентификации структурных шоков) был предложен в (Lanne, Lütkepohl and Maciejowska, 2010). Более полная методология для использования этого в модели SVAR была представлена в (Herwartz and Lütkepohl, 2011). В этом подходе предполагается, что распределение ошибок u_t зависит от дискретного марковского процесса s_t $(t = 0, \pm 1, \pm 2, \dots)$ с M состояниями. Другими словами, $s_t \in \{1, \dots, M\}$. Переходные вероятности между состояниями равны

$$p_{ij} = \Pr(s_t = j | s_{t-1} = i), \quad i, j = 1, \dots, M.$$

В (Lanne *et al.*, 2010) предполагается, что условное распределение u_t при заданном состоянии s_t является нормальным,

$$u_t | s_t \sim N(0, \Sigma_{s_t}).$$

Это предположение сделано для удобства и может быть ослаблено. Также VAR-коэффициенты не зависят от марковского процесса s_t и предполагаются инвариантными по времени. Хотя существует только конечное количество состояний волатильности, модель может смешивать эти состояния, назначая вероятности строго между нулем и единицей в любой определенный период t. Таким образом, она может охватить постепенный переход из одного состояния в другое, и можно утверждать, что она может также порождать континуум состояний. Они сосредоточены на конечным количестве состояний, что может упростить интерпретацию. Ковариационные матрицы состояний $\Sigma_1, \dots, \Sigma_M$ используются для идентификации шоков, то есть они раскладываются аналогичным способом, как и в предыдущих моделях.

При оценивании существуют некоторые дополнительные проблемы: число состояний волатильности должно быть определено, параметры полной модели должны быть оценены, диагональные элементы матрицы Ψ_m, $m = 1, \dots, M$, должны быть сравнимы, и возможно идентифицирующие ограничения из обычного анализа должны быть тестированы внутри текущей структуры. При предположении оь условном нормальном распределении при заданных состояниях логарифмическая функция правдоподобия имеет вид

$$\log l_{MS}(W, \Psi_1, \dots, \Psi_M, P) = \sum_{t=1}^{T} \log \left(\sum_{m=1}^{M} \Pr(s_t = m | Y_{t-1}) f(y_t | s_t = m, Y_{t-1}) \right) \tag{1.53}$$

где P – матрица переходных вероятностей, $Y_{t-1} = (y'_{t-1}, \dots, y'_{t-p})'$ и

$$f(y_t | s_t = m, Y_{t-1}) = (2\pi)^{-K/2} \det(\Sigma_m)^{-1/2} \exp\left\{ -\frac{1}{2} u'_t \Sigma_m^{-1} u_t \right\}.$$

Для максимизации в (Herwartz and Lütkepohl, 2011) предлагается соответствующий EM-алгоритм. К сожалению, этот алгоритм надежно работает только для относительно небольших систем с малым числом переменных K, малым числом состояний волатильности M и с умеренным числом лагов p.

Статистическая проверка параметров матриц Ψ_i может быть выполнена, используя стандартные процедуры. Например, раевнство некоторых параметров можно проверить стандарными тестами Вальда или LR. Однако последние тесты имеют недостаток в текущей ситуации, поскольку включают в себя две сложные процедуры максимизации функций правдоподобия. Следовательно, тесты Вальда может быть предпочтительнее с точки зрения вычислительной точки зрения. К сожалению, изветсно, что они имеют плохие свойства на конечных выборках в похожих ситуациях и, следовательно, могут не быть полностью удовлетворительными. В любом случае, в прицнипе, эти тесты могут использоваться для проверки условий идентифиации. Кроме того, если шоки могут идентифицироваться через структуру волатильности, другие ограничения можно также проверить обычными тестами. Для выбора между моделями с различным количеством режимов можно пользоваться стандартными критериями (см. (Psaradakis and Spagnolo, 2006)).

Когда структурная модель с идентифицируемыми шоками специфицирвоана, оан может исопльзоваться для структурного анализа. Импульсные отклики являются стандартной техников для этих целей. В классическом ML-анализе доверительные интервалы для импульсных откликов обычно порождаются методами бутстрепа. Ясно, что такие методы проблематичны, когда единственное оценивание сложно. Для уменьшения вычислительных проблем в (Herwartz and Lütkepohl, 2011) предлагается процедура fixed design wild bootstrap. такая процедура имеет смысл в текущей модели, поскольку она может позаботиться об условной гетероскедастичности (см. (Goncalves and Kilian, 2004)). Бутстреп-выборки строятся условно на ML-оценки, исопльзуя

$$y_t^* = \hat{v} + \hat{A}_1 y_{t-1} + \cdots + \hat{A}_p y_{t-p} + u_t^*.$$

Здесь переменные в правой стороне для заданного t всегда являются оригинальными лагированными переменными, а ошибки $u_t^* = \eta_t \hat{u}_t$, где η_t – бинарная случайная величина, принимающая значения 1 или -1 с равной вероятностью. Процедура сохраняет гетероскедастичность и картину одновременной зависимости данных. В (Herwartz and Lütkepohl, 2011) предлагаются бутстреп-оценки параметров θ^* для $\theta = \mathrm{vec}[v, A_1, \ldots, A_p]$ и W^* для W, условные на начлаьные оценки переходных вероятностей и Ψ_i, чтобы ограничить вычислительную нагрузку. Вычисление бутстреп-импульсных откликов требует нелинейной оптимизации логарифмической функции правдоподобия даже с этими упрощениями, что вычислительно затратно.

SVAR с ненормальными ошибками

В (Lane and Lütkepohl, 2010) предлагается подход, в котором предположения о распределении ошибок могут помочь в идентификации шоков. Рассмотрим следующую приведенную форму VAR(p):

$$A(L) y_t = u_t, \tag{1.54}$$

где предполагается, что n-мерный вектор ошибок u_t являестя смесью двух серийно независимых нормальных случайных векторов, таких что

$$u_t = \begin{cases} e_{1t} \sim N(0, \Sigma_1) & \text{с вероятностью } \gamma, \\ e_{2t} \sim N(0, \Sigma_2) & \text{с вероятностью } 1-\gamma. \end{cases} \tag{1.55}$$

Предплагается, что $\Sigma_1 \neq \Sigma_2$, поскольку в противном случае γ не будет идентифицирована. Заметим, что $u_t \sim N(0, \gamma\Sigma_1 + (1-\gamma)\Sigma_2)$. Рассмотрим так называемую B-модель, то есть где $u_t = \mathsf{B}\varepsilon_t$ (хотя анализ может быть обобщен на случай AB-модели). Существует альтернативный способ выбрать единственную матрицу B. Существуют диагональная матрица $\Psi = \mathrm{diag}(\psi_1, \ldots, \psi_n)$, $\psi_i > 0$, $1 = 1, \ldots, n$ и $(n \times n)$-матрица W, такие что $\Sigma_1 = WW'$ и $\Sigma_2 = W\Psi W'$, и W являестя единственной, за исключением случая изменения всех знаков в столбце, при условии что все ψ_i различны. Поэтому мы можем перепараметризовать Σ_u как

$$\Sigma_u = W(\gamma I_K + (1-\gamma)\Psi)W'.$$

В этом случае мы находим (локально) единственную матрицу $\mathsf{B} = W(\gamma I_K + (1-\gamma)\Psi)^{1/2}$. Если смешанное нормальное распределение представляет данные из двух различных режимов, такой выбор B приводит к ортогональности шоков к обоим режимам, в который они входят, $N(0, \Sigma_1)$ или $N(0, \Sigma_2)$, и, конечно, ε_t имеет единичную ковариационную матрицу. Локальная идентификация – это все, на что мы можем надеяться в моделировании SVAR, поскольку все знаки в столбце B могут всегда измениться без изменения произведения $\mathsf{B}\mathsf{B}'$.

Заметим что идентификация достигается путем использования только специфической структуры ковариационной матрицы. Такая структура также получается, конечно, если смешаны любые два распределения с нулевыми средними и ковариационными матрицами Σ_1 и Σ_2. Предположение о смешанном нормальном рапсределении сделано просто для удобства, так как это делает легким оценивание.

Данный подход позволяет использовать чисто статистический подход, чтобы не делать предположение о том, что существуют нефундаментальные шоки. Интерпретация ошибок как смеси двух нормлаьных рапсределений может быть поянтна, если предположить, что ненормальность распределения исходит из-за, например, выбросов, то есть два режима отличаются друг от друга волатильностью. Наличие нефундаментальных шоков в данном подходе можно отдельно проверять, например, LR-тестом.

Предположим, что у нас есть m шоков, которые идентифицируются непосредственно накладывая огарничения на матрицу B и предположим, что, без потери общности, что они устанавливаются в последних m позициях в ε_t. Тогдя для точной идентификации всех шоков в разложении $\Sigma_2 = W\Psi W'$ с $\Psi = \mathrm{diag}(\psi_1, \ldots, \psi_n)$ только первые $n-m$ значений ψ_i должны быть различными. если некоторые из m значений ψ_i также различны, шоки фактически сверхидентифицируемы. В этом случае можнотестировать ограничения до тех пор, пока шоки остаются идентифицируемы при нулевой гипотезе.

Для оценивания параметров данной модели строится функция правдоподобия, в которой плотности являются просто линейной (по вероятностям) комбинацией соответствующих плотностей для каждого из режимов. Более конкретно, мы оцениваем параметры γ, Ψ и W путем макимизации псевдоконцентрированной функции правдоподобия

$$L_{MN}(W, \Psi, \gamma) = \prod_{t=1}^{T} f_{t-1}(y_t), \tag{1.56}$$

где

$$\begin{aligned} f_{t-1}(y_t) = \gamma \det(W)^{-1} \exp\left\{-\frac{1}{2}(A(L)y_t)'(WW')^{-1}(A(L)y_t)\right\} \\ + (1-\gamma)\det(\Psi)^{-1/2}\det(W)^{-1}\exp\left\{-\frac{1}{2}(A(L)y_t)'(W\Psi W')^{-1}(A(L)y_t)\right\} \end{aligned} \tag{1.57}$$

Данный подход также использовался в (Lanne and Lütkepohl, 2008a), где предлагается на первом шаге оценить VAR(p)-модель, получить остатки (поскольку оценки удут состоятельными), а затем полученные остатки $\hat{u}_t$ подставить в функцию правдоподобия (1.56) вместо $A(L)y_t$.

В целом, подход, использующий ненормальные ошибки, аналогичен предыдущему подходу, использующему информацию о сдвиге в волатильности, только в последнем подходе режимы изменяются в некоторый фиксированный момент времени, в то время как в подходе с ненормальными ошибками режимы регулируются случайным механизмом.

1.1.8 Критика анализа импульсных откликов

Помимо спецификации соответствующих импульсов в системе, существует некоторое количество других проблем, которые делают сложным интерпретацию импульсных откликов. главное ограничение нашей системы заключается в ее потенциальной неполности. Хотя в реальной экономической системн почти все всегда зависит от всего, мы обычно работаем в VAR-системе небольшой размерности. Все эффекты пропущенных переменных предполагаются содержащимися в инновациях. Если важная переменная пропущена в системе, это может привести к существенным искажениям импульсных откликов и делает их бесполезными для структурных интерпретаций. Однако, система может все еще быть полезной для предсказания.

Чтобы видеть связанные с этим проблемы более ясно, рассмотрим систему y_t, которая разивается на векторы z_t и x_t . Если рассматрвиаются только переменные z_t, а переменные x_t пропущены, мы получаем систему

$$\begin{aligned} z_t &= \mu_1 + \sum_{i=0}^{\infty} \Phi_{11,i} u_{1,t-i} + \sum_{i=0}^{\infty} \Phi_{12,i} u_{2,t-i} \\ &= \mu_1 + \sum_{i=0}^{\infty} F_i v_{t-i}. \end{aligned} \tag{1.58}$$

Фактические реакции компонент z_t на инновации u_{1t} могут быть заданы матрциами $\Phi_{11,i}$. С другой стороны, F_i или соответствующие ортогонализованные “импульсные отклики” вероятно будут интерпретироваться как импульсные отклики, если исследователь не понимает, что пропущены важные переменные. Здесь F_i будет равна $\Phi_{11,i}$, если и тлько если x_t не является причинуой по Гренджеру для z_t.

Дальнейшие проблемы, связанные с интерпретацией MA-коэффициентов как динамических мультипликаторов или импульсных откликов, получаются из-за ошибок измерения и использования сезонно скорректированных или временно и/или одновремено агрегированных переменных. Детальное изложение пробелмы агрегирования дано в (Lütkepohl, 1987). Эти проблемы существенно ограничивают интерпретируемость MA-коэффициентов VAR-системы как импульсных откликов.

1.2 Структурные VECM-модели

Интегрированные и коинтегрированные системы нужно интерпретировтаь с осторожностью. В коинтегрированной системе принято считать, что коинтегрирующий вектор $\beta' y_t$ (который, напомним, должен быть стационарной линейной комбинацией) обычно представляет долгосрочное равновесное соотношение между переменными. Предположим, что существует только одно такое соотношение, скажем

$$\beta_1 y_{1t} + \cdots + \beta_K y_{Kt} = 0,$$

или, если $\beta_1 \neq 0$,

$$y_{1t} = -\frac{\beta_2}{\beta_1} y_{2t} - \cdots - \frac{\beta_K}{\beta_1} y_{Kt}.$$

Очень хотелось бы утверждать, что долгосрочный эффект единичного увеличения в y_2 будет изменять y_1 на величину β_2/β_1. Это, однако, игнорирует все другие соотношения между переменными в VAR(p) или соответствующей VECM. Елинчиная инновация в некоторый момент времени в y_2 может влиять на различные другие переменные, которые также имеют влияние на y_1. Следовательно, долгосрочный эффект инноваций в y_2 на y_1 может достаточно отличаться от $-\beta_2/\beta_1$. Импульсные отклики могут давать лучшую картину соотношений между переменными.

В Разделе 1.1 было показано, что импульсные отклики стационарного и стабильного процесса VAR(p) являются коэффициентами конкретных MA-представлений. Нестабильный, интегрированный или коинтегрированный процесс VAR(p) не обладает обоснованными MA-представлениями, представленными в Разделе 1.1. Несмотря на это матрицы Φ_i и Θ_i могут быть вычеслены, как и ранее, и элементы матрицы Φ_i – $\phi_{jk,i}$ – могут редставлять импульсные отклики, как и в случае стабильности. Но в стабильных процессах отклики затухают с течением времени ($i \to \infty$), однако, это свойство не выполняется для нестабильных систем, в которых влияние импульса в некоторый момент времени не затухает асимптотически. Аналогично можно вычислять и накопленные импульсные отклики ортогонализованным остаткам и декомпозиции дисперсии ошибок прогноза, кроме полных "долгосрочных эффектов", или полных мультипликаторов, Ψ_∞ и Ξ_∞, поскольку они могут не быть конечными (например, стабильность требуется для ограничений Бланшара-Кваха, поскольку матрица $(I_K - A_1 - \cdots - A_p)$ является вырожденной для коинтегрированных процессов). Другими словами, мы можем даже специфицировать и оценить VECM в приведенной форме, преобразовать эту модель в модель VAR в уровнях, а затем использовать ее как основу для AB-анализа, как обсуждалось в Разделе 1.1.3. Однако существуют некоторые недостатки этго подхода.

В (Phillips, 1998) были получены следующие результаты. Пусть процесс y_t не является стабильным, то есть имеет корни вне или на единичной окружности. Тогда если горизонт для импульсных откликов предполагается фиксированным, то функции импульсных откликов остаются состоятельными и асимптотически нормальными, как и в случае стабильных VAR. Однако асимптотическая ковариационная матрица оцененных коэффициентов VAR будет вырожденной, поскольку только стационарные компоненты системы имеют $\sqrt{T}$-асимптотику. Однако, если задвать горизон прогноза как $i/T \to 0$ при $T, i \to \infty$ или как $i = fT$, где $f > 0$ – фиксированная доля от выборки, то при наличиие единичных корней или почти единичных корней (near unit roots) в VAR импульсные отклики, оцененные неограниченной OLS-регрессией, будут несостоятельными. В частности, пределы оцененных откликов будут некоторыми случайными величинами, а не истинными импульсными откликами. Это может показаться удивительным, поскольку наличие единичных корней или почти единичных корней ускоряет сходимость оценок коэффициентов в OLS-регрессии, и на основе этого можно ожидать, что импульсные отклики будут сходиться ыстрее в некоторых направлениях. Причина несостоятельности заключается в том, что истинные импульсные отклики больше не затухают при увеличении горизонта, то есть элементы Θ_i в (1.13) не стремятся к нулю при $i \to \infty$, но несут эффекты единичных корней с их бесконечностью. Однако единичные корни, оцененные с ошибкой, и эффекты оценивания сохраняется в приделе при $n \to \infty$, когда рассматриваются долговременные импульсы Θ_i, где i – некоторая доля (f) от объема выборки (T). В отличие от этого, когда система стабильна, элементы Θ_i стремятся к нулю при $i \to \infty$, и оцененные ошибки не оказывают влияния в пределе. То есть результат Филлипса охватывает случай стабильных VAR.

Если мы оценивиаем импульсные отклики в коинтегрированной VAR-модели, то они будут состоятельными на краткосрочных и долгосрочных горизонтах, если ранг коинтеграции выбирается состоятельной процедурой. На долгосрочных горизонтах щоки будут иметь инерционный эффект на систему бесконечно в будущее. Именно эти инерционные эффекты состоятельно оценвиаются регрессией с пониженным рангом (reduced rank regression, RRR). Если в системе присутствуют почти единичные корни, регерссия с пониженным рангом ошибочно принимает близкие к единице корни как единичные корни при росте выборки. То же самое остается верным, когда исопльзуются информационные критерии для выбора ранга коинтеграции. Таким образом, долговременные импульсные отклики несостоятельно оцениваются, когда существую почти единичные корни. Предельное расgределение этих оцененных импульсов, однако, все еще нормально. Это распределение обладает нежелатиельным свойством, которое заключается в том, что асимптотические дисперсии расходятся при $i \to \infty$. Это довольно озадачивает в связи с тем, что истинные матрицы импульсных откликов сходятся к фиксированным матрицам при $i \to \infty$, и мпульсные отклики, оцененные в регрессии с пониженным рангом, являюстя состоятельными. В (Arai and Yamamoto, 2000) предлагается альтернативное представление асимптотического рапсределения RRR-оценок импульсных откликов со сходящимися асимптотическими дисперсиями для коинтегрированных VAR-систем. Получение близко следует (Phillips, 1998), за исключением того факта, что авторы явно используют тот факт, что $n = K - r$ единичных корней не оценевается RRR, где K – размерность системы, а r – ранг коинтеграции.

Также Филлипс показал, что оцененная матрица дисперсии ошибки прогноза из оцененной неограниченной VAR с некоторыми корнями, близкими или равными единице, ведет себя как случайная матрица, умноженная на горизонт прогноза h, а не как постоянная матрица, умноженная на h. Таким образом, оцененные матрицы дисперсий ошибок прогноза в долгосрочном плане в неограниченной VAR являются несостоятельными. Тот же самый результат следует для соответствующих оценок разложений дисперсий ошибок прогноза. Кроме того, предсказание по неограниченной VAR-регрессии не являесят асимптотически оптимлаьным в том смысле, что предсказания не сходятся к оптимлаьным предикторам, по крайней мере на длиных горизонтах прогноза. Этот результат являестя полностью противоположным стационарной VAR-модели, где ошибки от оценивания коэффициентов не влияют асимптотически, и разница между доступным и оптимлаьным прогнозами стремится к нулю при $T \to \infty$.

Если оценивается регрессия в пониженным рангом, то разложение дисперсий ошибок прогноза для фиксированного горизонта имеет ту же самую предельную теорию, как и неограниченная VAR. Но когда горизонт прогноза стремится к бесконечности с ростом выборки, регрессия с пониженным рангом продолжает быть состоятельной, в отличие от неограниченной VAR, которая несостоятельна и имеет случайные пределы. Прогнозы в этом случае дял корректно специфицированной модели будут асимптотически эквивалентными оптимальному предиктору. Однако, когда в системе есть почти единичные корни вместо единичных корней, прогнозы по регрессии с пониженным рангом не будут больше асимптотически эквивалентными оптимальному предиктору.

Мы предполагаем, что все переменные по крайней мере I(1), и что процесс порождения данных может быть представлен как VECM с рангом коинтеграции r в виде

$$\Delta y_t = \boldsymbol{\alpha}\boldsymbol{\beta}' y_{t-1} + \boldsymbol{\Gamma}_1 \Delta y_{t-1} + \cdots + \boldsymbol{\Gamma}_{p-1} \Delta y_{t-p+1} + u_t, \tag{1.59}$$

где все символы имеют свое стандартное значение. Другими словами, y_t является K-мерным вектором наблюдаемых переменных, $\boldsymbol{\alpha}$ – матрица корректирующих коэффициентов размерности $(K \times r)$, $\boldsymbol{\beta}$ – матрица коинтегрирующих векторов размерности $(K \times r)$, $\boldsymbol{\Gamma}_j$ – матрица краткосрочных коэффициентов для $j = 1, \ldots, p-1$ размерности $K \times K$, и $u_t \sim (0, \Sigma_u)$ – вектор ошибок.

Процесс (1.59) по Теореме предствления Гренджера имеет следующее MA-представление типа Converиджа-Нельсона

$$y_t = \Xi \sum_{i=1}^{t} u_i + \sum_{j=0}^{\infty} \Xi_j^* u_{t-j} + y_0^* = C(L)u_t, \tag{1.60}$$

где Ξ_j^* абсолютно суммируемы, так что бесконечная сумма определена, компонента y_0^* содержит начальные значения, а $C(L)$ – некоторый полином. Абсолютная суммируемость Ξ_j^* подразумевает, что эти матрицы сходятся к нулю при $j \to \infty$. Таким образом, долгосрочные эффекты шоков охватываются компонентой общих трендов $\Xi \sum_{i=1}^{t} u_i$. Матрица

$$\Xi = \boldsymbol{\beta}_\perp \left[\boldsymbol{\alpha}'_\perp \left(I_K - \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_i \right) \boldsymbol{\beta}_\perp \right] \boldsymbol{\alpha}'_\perp$$

имеет ранг $K-r$. Таким образом, существует $K-r$ общих трендов, и если инновации, embodied in the u_t, могут быть восстановлены, самое большее r из них могут иметь преходящее влияние, только потому что матрица Ξ или ее невырожденное преобразование не могут иметь более r нулевых столбцов. Таким образом, зная ранг коинтеграции системы, мы уже знаем максимальное число преходящих шоков.

В этом контексте в центре внимания, как правило, остатки и, следовательно, для того чтобы определить структурные инновации обычно используется анализ B-модели. Другими словами, мы ищем матрицу B, такую что

$$u_t = \mathsf{B}\varepsilon_t \quad \text{с} \quad \varepsilon_t \sim (0, I_K).$$

Подставление этого соотношения в компоненту общих трендов дает $\Xi\mathsf{B} \sum_{i=1}^{t} \varepsilon_i$. Следовательно, долгосрочные эффекты структурных инноваций задаются матрицей $\Xi\mathsf{B}$. Посколкьу структурные инновации представляют регулярный случайный вектор с невырожденной ковариационной матрицей, матрица B также невырождена. Вспомните, что $\Sigma_u = \mathsf{B}\mathsf{B}'$. Таким образом, $\mathrm{rk}(\Xi\mathsf{B}) = K - r$ и может быть может быть не более r нулевых столбцов в этой матрице. Другими словами, r из структурных инноваций могут иметь преходящее влияние и $K-r$ из них могут иметь перманентное влияние. если есть r преходящих шоков, мы можем ограничить r столбцов в $\Xi\mathsf{B}$, полагая их нулевыми. Поскольку матрица имеет пониенный ранг $K-r$, каждый нулевой столбец выступает только за $K-r$ независимых ограничений. Таким образом, r преходящих шоков представляет только $r(K-r)$ независимых ограничений. Все еще полезно заметить, что ограничения могут быть наложены на основе нашего знания о ранге коинтеграции системы, который может быть определен статистическими способами. Однако, дальнейшие теоретические рассмотрения требуют накладывания дополнительных ограничений.

Для локальной точной идентифицируемости структурных инноваций в B-модели нам нужно в общей сложности $K(K-1)/2$ ограничений. Предполагая, что существует только r шоков с преходящими эффектами, мы уже имеем $r(K-r)$ ограничений из коинтеграционной структуры модели. Остается $\frac{1}{2}K(K-1)-r(K-r)$ дополнительных ограничений для точной идентифицируемости структурных инноваций. Фактически $r(r-1)/2$ дополнительных одновременных ограничений необходимо для распутывания преходящих шоков и $(K-r)((K-r)-1)/2$ ограничений для идентификации перманентных шоков (см., например, (King, Plosser, Stock and Watson, 1991) и (Gonzalo and Ng, 2001)). Затем у нас есть в общей сложности $\frac{1}{2}r(r-1)+\frac{1}{2}(K-r)((K-r)-1)=\frac{1}{2}K(K-1)-r(K-r)$ ограничений, как и требуется. Таким образом, недостаточно наложить произвольные ограничения на B или $\Xi\mathsf{B}$, но мы должны выбрать их, чтобы идентифицировать преходящие и перманентные шоки по крайней мере локально. Фактически преходящие шоки могут быть только идентифицируемы через ограничения непосредственно на B, поскольку они соответствуют нулевым столбцам в $\Xi\mathsf{B}$. Таким образом, должны ыть наложены на B непосредственно $r(r-1)/2$ ограничений. Как правило, ограничения имеют вид

$$C_{\Xi\mathsf{B}}\operatorname{vec}(\Xi\mathsf{B})=c_l, \quad \text{или} \quad C_l\operatorname{vec}(\mathsf{B})=c_l, \quad \text{и} \quad C_s\operatorname{vec}(\mathsf{B})=c_s, \tag{1.61}$$

где $C_l=C_{\Xi\mathsf{B}}(I_K\otimes\Xi)$ являестя матрицей долгосрочных ограничений, то есть $C_{\Xi\mathsf{B}}$ – подходящая selection matrix , такая что $C_{\Xi\mathsf{B}}\operatorname{vec}(\Xi\mathsf{B})$, а C_s специфицирует краткосрочные или текущие ограничения, ограничивая элементы в B непосредственно. Здесь c_l и c_s – векторы соответствующих размерностей. В прикладных работах они обычно являюстя нулевыми векторами. Другими словами, нулевые ограничения специфицированы в (1.61) для $\Xi\mathsf{B}$ и B.

Как обсуждалось для стационарного случая в Разделе 1.1.3, матрциа B будет только локально идентифицируема. в частности, в общем случае мы можем поменять знаки столбцов в B, чтобы найти другую обоснованную матрицу. Формальное необходимое и достаточное условие для локальной идентифицируемости заключается в следующем. Предположим, что B невырожденная $(K\times K)$-матрица. Тогда множество уравнений

$$\Sigma_u=\mathsf{B}\mathsf{B}', \quad C_l\operatorname{vec}(\mathsf{B})=c_l \quad \text{и} \quad C_s\operatorname{vec}(\mathsf{B})=c_s$$

с C_l, c_l, C_s и c_s определенных как в (1.61) имеют локально единственное решение для B тогда и только тогда, когда

$$\operatorname{rk}\begin{bmatrix}2\mathbf{D}_K^+(\mathsf{B}\otimes I_K)\\ C_l\\ C_s\end{bmatrix}=K^2.$$

Далее более детально рассмотрим подходы разделения перманентных и преходящих шоков. В (Gonzalo and Ng, 2001) (см. также (Yang, 1998)) определяется $(K\times 1)$-вектор ϵ_t, первые $K-r$ элементов которого являются перманентными шоками, а оставшиеся r – преходящими шоками, как $\epsilon_t=Gu_t$, где

$$G=\begin{bmatrix}\alpha'_\perp\\ \beta'\end{bmatrix}$$

и $\Sigma_\epsilon=G\Sigma_u G'$. Предполагается, что $(p\times p)$-матрица G имеет полный ранг. Долгосрочная матрица вклада перманентных-преходящих шоков ϵ_t равна $\Xi G^{-1}=C(1)G^{-1}$. Легко показать, что последние r вектор-столбцов этой долгосрочной матрицы вклада нулевые, замечая что

$$G^{-1}=[\boldsymbol{\beta}_\perp(\boldsymbol{\alpha}'_\perp\boldsymbol{\beta}_\perp)^{-1}\ \boldsymbol{\alpha}(\boldsymbol{\beta}'\boldsymbol{\alpha})^{-1}],$$

как и требуется для перманентного-преходящего разложения. Далее применяется дополнительное преобразование для ортогонализации перманентных-преходящих шоков: $\varepsilon_t=H^{-1}\epsilon_t=$

$H^{-1}G\varepsilon_t$, где H – единственная нижнетреугольная матрица, такая что $HH' = G\Sigma_u G'$. Другими словами, $\mathsf{B} = G^{-1}H$, и ортогонализованная матрциа вклада перманентны-преходящих шоков будет равна $\Xi\mathsf{B} = \Xi G^{-1}H$. P-T разложение сохраняется, поскольку умножение ΞG^{-1} на нижнетреугольную матрицы приводит к матрице, все последние r вектор-столбцов которой нулевые.

В (Fisher and Huh, 2007) доказываестя следующее утверждение. Предположим, что существует $K - r$ слабо экзогенных переменных относительно коинтегрирующего вектора, и они упорядочены таким образом, чтобы быть первыми в векторе y_t. Тогда ортогонализованные шоки, полученные непосредственно из VECM накладыванием рекурсивной структуры на одновременные вклады ошибок составляют P-T разложение, где перманентные шоки идентичны P-T разложению (Gonzalo and Ng, 2001). В дополнение, когда существует только одно коинтеграционное соотношение ($r = 1$), ортогонализованные шоки, полученные путем накладывания одновременной рекурсивной структуры на ошибки VECM, идентичны P-T разложению (Gonzalo and Ng, 2001). Также, если существует $0 < j < K - r$ слабо экзогенных переменных отосительно коинтеграционного вектора, и они упорядочены быть первыми в y_t, то ортогонализованные шоки, полученные путем накладывания одновременной рекурсивной структуры на ошибки VECM не составляют P-T разложение; первые j из этих шоков идентичны первым j перманентным шокам в P-T разложении (Gonzalo and Ng, 2001) (см. также (Fisher and Huh, 1999) где доказывается, что если у нас есть $0 < j < K - r$ слабо экзогенных переменных, то r структурных шоков с преходящими эффектами не имеет одновременного вклада на эти j слабо экзогенные переменные).

В (Lütkepohl, 2008) утверждается, что дополнительные сверхидентифицирующие ограничения нужно накладывать с осторожность, поскольку некоторые из них невозможны из-за того, что они приводят к вырожденной ковариационной матрице остатков, котоаря обычно исключается из предположения, а также неправдоподобна с теоретической точки зрения. Другими словами, некоторые сверхидентифицирующие огарничения мгут не быть возможными, посоклкьу они будут вне допустимого постранства параметров. Поэтому соответствующие t-статистики не могут интерпретироваться обычным способом. Эта проблема влияет также и на импульсные отклики. Автор приводит следующее утверждение. Предположим, что в ECM матрица $\boldsymbol{\alpha}_\perp$ размерности $(K \times (K - r))$ такова, что все множества $K - r$ строк являются невырожденными $((K-r)\times(K-r))$-матрциами, и существует $r^* \leq r$ преходящих шоков. Тогда число допустимых нулевых ограничений, размещенных на столбце в B, связанных с преходящим шоком, не может быть более $r-1$. Например, если ранг равен 1 (и один преходящий шок)), в столбце в матрице B не может быть ни одногоо ограничения. Если ранг равен 2 (и два преходящих шока), то может быть самое большее одно нулевой ограничение в каждом из двух столбцов в B, соответствующих преходящим шокам. Условие, что матрица $\boldsymbol{\alpha}_\perp$ такова, что все множества $K - r$ строк являются невырожденными $((K-r)\times(K-r))$-матрциами, будет всегда удовлетворяться, если α оценивается без ограничений. Однако при наличии слабо экзогенных переменных полученное утверждение не будет выполняться, и преходящие шоки не будут иметь одновременный эффект на слабо экзогенные переменные (см. (Fisher and Huh, 1999)).

Аналогичное свойство будет наблюдаться для ограничений на матрицу $\Xi\mathsf{B}$. Предположим, что $((K-r)\times(K-r))$-субматрицы $\boldsymbol{\beta}_\perp$ невырождены, и есть r преходящих шоков. Тогда число допустимых нулевых ограничений размещенных на столбце в $\Xi\mathsf{B}$, связанных с перманентным шоком, не может быть более $K - r - 1$.

В (Lucke, 2010) было получено обобщение подхода (Lütkepohl, 2008), которое выполняется независимо от того, являются ли некоторые структурные шоки преходящими: при структурном разложении $K-r$ столбцов матрицы $\Xi\mathsf{B}$ допускают самое большее $K-r-1$ нулевых ограничений каждое; оставшиеся r столбцов допускают вплоть до K нулевых ограничений каждое, из которых не более $K-r$ линейно независимы. Также предположим, что максимально допустимое число нулевых ограничений накладывается на $\Xi\mathsf{B}$. Тогда структурная идентификация для (1.59) требует дополнительных нулевых ограничений на B тогда и только тогда, когда $2r > K$.

Относительно дополнительных ограничений на матрицу B в (Lucke, 2010) доказывается следующий результат. Предположим, что в структурном разложении (1.59) существует $z \leq K$ нулевых ограничений, накладываемых на определенные столбцы матрицы $\Xi\mathsf{B}$. Пусть $v = \min\{z, K-r\}$. Если из $z < K-r$ следует, что все множества $(v \times v)$-субматриц в Ξ невырожденные, и если из $z > K-r$ следует, что все множества $(v \times v)$-субматриц в $\boldsymbol{\alpha}_\perp$ невырождены, то самое большее $K-v-1$ нулевых огарничений допускается на соответствующий столбец в B.

Оценивание и построение доверительных интервалов

Предположим теперь, что ограничения для VECM заданы в форме линейных ограничений на $\Xi\mathsf{B}$ и B, как в (1.61). Для вычисления оценок параметров мы можем заменить Ξ на ее ML-оценку приведенной формы,

$$\tilde{\Xi} = \tilde{\boldsymbol{\beta}}_\perp \left[\tilde{\boldsymbol{\alpha}}'_\perp \left(I_K - \sum_{i=1}^{p-1} \tilde{\boldsymbol{\Gamma}}_i \right) \tilde{\boldsymbol{\beta}}_\perp \right] \tilde{\boldsymbol{\alpha}}'_\perp,$$

где $\tilde{\boldsymbol{\Gamma}}_i$ – ML-оценки параметров $\boldsymbol{\Gamma}_i$, а $\tilde{\boldsymbol{\alpha}}_\perp$ и $\tilde{\boldsymbol{\beta}}_\perp$ – ортогональные дополнения ML-оценок $\tilde{\boldsymbol{\alpha}}$ и $\tilde{\boldsymbol{\beta}}$, соответственно. Ограниченная ML-оценка для B может быть получена, полагая $\mathsf{A} = I_K$ и оптимизируя концентрированную функцию логарифмического правдоподобия (1.30) с заменой C_l на

$$\tilde{C}_l = C_{\Xi\mathsf{B}}(I_K \otimes \tilde{\Xi})$$

(см. (Vlaar, 2004)). Хотя эта процедура приводит к множеству стохастических ограничений, с численной точки зрения у нас есть стандартная оптимизационная задача с ограничениями, которую можно решить, исопльзуя подход Лагранжа, поскольку $\tilde{\Xi}$ фиксирована в вычислении оценки для B. Из-за того факта, что для точно идентифицированной структурной модели максимум логарифма правдоподобия тот же самый, что и для приведенной формы, сравнение значений логарифма правдоподобия может служить проверкой надлежащей сходимости алгоритма оптимизации, используемого для структурного оценивания.

Свойства ML-оценки для B аналогично случаю SVAR-моделей. Другими словами, $\tilde{\mathsf{B}}$ является состоятельной и асимптотически нормлаьной при стандартных условиях,

$$\sqrt{T}\,\mathrm{vec}(\tilde{\mathsf{B}} - \mathsf{B}) \xrightarrow{d} N(0, \Sigma_{\mathsf{B}}).$$

асимптотическое распределение являестя вырожденным из-за ограничений, которые наложены на B. Таким образом, хотя t-отношения могут быть использованы для оценки значимости индивидуальных параметров, F-тесты, основанные на принципе Вальда, не будут в общем случае обоснованы и должны быть интерпретированы с осторожностью. Выражения для ковариационных матриц асимптотических распределений в терминах параметров модели могут быть получены обычным способом, используя соответствующие информационные матрицы (см. (Vlaar, 2004)). Для практических целей распространенным способом является использование методов бутстрепа для статистических выводов в данном контексте.

В принципе, можно использовать аналогичный подход, если есть сверхидентифицирующие ограничения на B. В этом случае $\tilde{\mathsf{B}}\tilde{\mathsf{B}}'$ не будет равно ковариационной матрицы оценки приведенной формы Σ_u, однако. Оценка B будет все еще состоятельной и асимптотически нормальной при более общих условиях, а также LR-статистика, заданная в (1.33) может использоваться для проверки правильности сверхидентифицирующих ограничений. Она будет иметь обычное асимптотическое распределение χ^2 с числом степеней свободы, равному числу сверхидентифицирующих ограничений.

Свойства доверительных интервалов для импульсных откликов SVECM на конечных выборках были анализированы в (Brüggemann, 2006). На основе симуляций Монте-Карло автор показал, что нет большой разницы между асимптотическими интервалами и процентильными бутстреп-интервалами Холла в SVECM с долгосрочными ограничениями. В отличие от этого стандартный процентильный бутстреп-интервал может быть менее подходящим в прикладных работах, так как он менее информативен о знаке лежащей в основе функции импульсного отклика. Сравнения доверительные интервалы для откликов на перманентные шоки, автор обнаружил, тчо асимптотические и процентильные бутстреп-интервалы Холла имеют похожую норму покрытия и указывают на правильный знак лежащего в основе отклика одинаково часто. Следовательно, они допускают одинаковую интерпертацию. Однако бутстреп-интервалы обычно асимммметричны и намного более широкие. В отличие от этого стандатные бутстреп-интервалы включают нулевую линия чаще, и, следовательно, указывают корректный знак реже. Для откликов на преходящие шоки доверительные интервалы имеют гораздо больше общего. Однако автор не рекомендует исопльзовать стьюдентизированные интервалы Холла из-за их вычислительных требований.

SVECM с ненормальными ошибками

Аналогично разделу 1.1.7 рассмотрим аналогичную идентификацию для VECM. Предположим, что накладываюстя только ограничения, связанные с разделением перманентных и преходящих шоков в матрице $\Xi\mathsf{B}$, то есть последние r столбцов в этой матрице нулевые и связаны с преходящими шоками. Тогда шоки будут идентифицирвоаны, если u_t имеет смешанное нормлаьное распределение как в (1.55) с ковариационными матрицами $\Sigma_1 = WW'$ и $\Sigma_2 = W\Psi W'$, где $\psi_1, \ldots, \psi_{n-r}$ различны, а также $\psi_{n-r+1}, \ldots, \psi_n$ различны.

Заметим, что в VECM с рангом коинтеграции $r < K$ все структурные шоки могут быть, в прицнипе, перманентными, поскольку матрица $\Xi\mathsf{B}$ может не иметь нулевых столбцов, даже если она и имеет пониженный ранг. Рассмотрение r шоков как преходящих являестя предположением, которое в идеали не основано исключительно на ранге коинтеграции. Поэтому, если u_t является fully general смешанным нормальным, как в (1.55), нулевые ограничения на $\Xi\mathsf{B}$ фактически являются сверхидентифицирующими, которые можно тестировать, например, LR-тестом. Другими словами, можно тестировать число перманентных и преходящих шоков, в том числе и последовательной процедурой, начиная с $H_0 : r^* = r$ и заканчивая $H_0 : r^* = 1$. Тестовая последовательность останавливается и число преходящих шоков выбирается тогда, когда одна из нулевых гипотез не отвергается.

Для оценивания такой модели на первом шаге используется регрессия пониженного ранга Йохансена (?) для оценвиания коинтеграционных соотношений. Эти оценки эквивалентны ML для гауссовской модели, но являюстя квази-ML при смешанном нормальном распределении. На втором шаге мы можем тогда использовать условие на оцененные коинтеграционные соотношения и максимизировтаь функцию правдоподобия относительно других параметров.

1.3 Система динамических одновременных уравнений

1.4 Байесовские SVAR и SVECM

Глава 2

Обзор литературы по эмпирическим приложениям SVAR

Глава 3

Эмпирика по российским данным

Заключение

Литература

Arai, Y. and Yamamoto, T. (2000). Alternative representation for asymptotic distributions of impulse responses in cointegrated VAR systems. *Economics Letters*, **67**, 261–271.

Benkwitz, A., Lütkepohl, H., and Neumann, M. (2000). roblems related to bootstrapping impulse responses of autoregressive processes. *Econometric Reviews*, **19**, 69–103.

Benkwitz, A., Lütkepohl, H., and Wolters, J. (2001). Comparison of bootstrap confidence intervals for impulse responses of German monetary systems. *Macroeconomic Dynamics*, **5**, 81–100.

Blanchard, O. and Quah, D. (1989). The dynamic effects of aggregate demand and supply disturbances. *American Economic Review*, **79**, 655–673.

Bouakez, H. and Normandin, M. (2010). Fluctuations in the foreign exchange market: How important are monetary policy shocks? *Journal of International Economics*, **81**, 139–153.

Brüggemann, R. (2006). Unpublished Manuscript.

Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. Chapman & Hall, New York.

Ehrmann, M., Fratzscher, M., and Rigobon, R. (2011). Stocks, bonds, money markets and exchange rates: Measuring international
nancial transmission. *Journal of Applied Econometrics*, **26**, 948–974.

Fisher, L. A. and Huh, H. (1999). Weak exogeneity and long-run and contemporaneous identifying restrictions in VEC models. *Economics Letters*, **63**, 159–165.

Fisher, L. A. and Huh, H. (2007). Permanent-Transitory Decompositions Under Weak Exogeneity. *Econometric Theory*, **23**, 183–189.

Galí, J. (1999). Technology, employment, and the business cycle: Do technology shocks explain aggregate fluctuations? *American Economic Review*, **89**, 249–271.

Giannini, C. (1992). *Topics in Structural VAR Econometrics*. Springer, Heidelberg.

Goncalves, S. and Kilian, L. (2004). Bootstrapping autoregressions with conditional heteroskedasticity of unknown form. *Journal of Econometrics*, **123**, 89–120.

Gonzalo, J. and Ng, S. (2001). A systematic framework for analyzing the dynamic effects of permanent and transitory shocks. *Journal of Economic Dynamics and Control*, **25**, 1527–1546.

Hall, P. (1992). *The Bootstrap and Edgeworth Expansion*. Springer-Verlag, New York.

Herwartz, H. and Lütkepohl, H. (2011). Structural vector autoregressions with markov switching: Combining conventional with statistical identification of shocks. Unpublished manuscript.

Jordà, O. (2009). imultaneous con
dence regions for impulse responses. *The Review of Economics and Statistics*, **91**, 629–647.

Jordà, O. and Marcellino, M. (2010). Path forecast evaluation. *Journal of Applied Econometrics*, **25**, 635–662.

Kilian, L. (1998). Small-sample confidence intervals for impulse response functions. *Review of Economics and Statistics*, **80**, 218–230.

Kilian, L. (1999). Finite-sample properties of percentile and percentile-t bootstrap confidence intervals for impulse responses. *Review of Economics and Statistics*, **81**, 652–660.

Kilian, L. (2001). Impulse response analysis in vector autoregressions with unknown lag order. *Journal of Forecasting*, **20**, 161–179.

King, R. G., Plosser, C. I., Stock, J. H., and Watson, M. W. (1991). Stochastic trends and economic fluctuations. *American Economic Review*, **81**, 819–840.

Lane, M. and Lütkepohl, H. (2010). Structural Vector Autoregressions With Nonnormal Residuals. *Journal of Business & Economic Statistics*, **28**, 159–168.

Lanne, M. and Lütkepohl, H. (2008*a*). A Statistical Comparison of Alternative Identification Schemes for Monetary Policy Shocks. Unpublished Manuscript.

Lanne, M. and Lütkepohl, H. (2008*b*). Identifying Monetary Policy Shocks via Changes in Volatility. *Journal of Money, Credit & Banking*, **40**, 1131–1149.

Lanne, M., Lütkepohl, H., and Maciejowska, K. (2010). Structural vector autoregressions with Markov switching. *Journal of Economic Dynamics and Control*, **34**, 121–131.

Lucke, B. (2010). Identification and overidentification in SVECMs. *Economics Letters*, **108**, 318–321.

Lütkepohl, H. (1987). *Forecasting Aggregated Vector ARMA Processes*. Springer, Berlin.

Lütkepohl, H. (1990). Asymptotic distributions of impulse response functions and forecast error variance decompositions of vector autoregressive models. *Review of Economics and Statistics*, **72**, 116–125.

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer, Berlin.

Lütkepohl, H. (2008). Problems related to over-identifying restrictions for structural vector error correction models. *Economics Letters*, **99**, 512–515.

Lütkepohl, H. (2013*a*). Identifying Structural Vector Autoregressions Via Changes in Volatility, *in* T. B. Fomby, L. Kilian, A. Murphy (eds.). *VAR Models in Macroeconomics – New Developments and Applications: Essays in Honor of Christopher A. Sims (Advances in Econometrics, Volume 32)*, Emerald Group Publishing Limited, 169–203.

Lütkepohl, H. (2013*b*). Reducing confidence bands for simulated impulse responses. *Statistical Papers*, **54**, 1131–1145.

Lütkepohl, H., Staszewska-Bystrova, A., and Winker, P. (2013). Comparison of Methods for Constructing Joint Confidence Bands for Impulse Response Functions. Unpublished Manuscript.

Normandin, M. and Phaneuf, L. (2004). Monetary policy shocks: Testing identification conditions under time-varying conditional volatility. *Journal of Monetary Economics*, **51**, 1217–1243.

Phillips, P. C. B. (1998). Impulse response and forecast error variance asymptotics in nonstationary VARs. *Journal of Econometrics*, **83**, 21–56.

Psaradakis, Z. and Spagnolo, N. (2006). Joint determination of the state dimension and autoregressive order for models with Markov regime switching. *Journal of Time Series Analysis*, **27**, 753–766.

Rigobon, R. (2003). Identification Through Heteroskedasticity. *The Review of Economics and Statistics*, **85**, 777–792.

Rigobon, R. and Sack, B. (2003). Measuring The Reaction Of Monetary Policy To The Stock Market. *The Quarterly Journal of Economics*, **118**, 639–669.

Runkle, D. E. (1987). Vector autoregressions and reality. *Journal of Business & Economic Statistics*, **5**, 437–442.

Sentana, E. and Fiorentini, G. (2001). Identification, estimation and testing of conditionally heteroskedastic factor models. *Journal of Econometrics*, **102**, 143–164.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, **48**, 1–48.

Sims, C. A. (1981). An autoregressive index model for the U.S. 1948-1975, *in* Kmenta, J. & Ramsey, J. B. (eds.). *Large-Scale Macro-Econometric Models*, North–Holland, Amsterdam, 283–327.

Sims, C. A. and Zha, T. (1999). Error bands for impulse responses. *Econometrica*, **67**, 1113–1155.

Staszewska, A. (2007). Representing uncertainty about impulse response paths: The use of heuristic optimization methods. *Computational Statistics & Data Analysis*, **52**, 121–132.

van der Weide, R. (2002). O-GARCH: A multivariate generalized orthogonal GARCH model. *Journal of Applied Econometrics*, **17**, 549–564.

Vlaar, P. J. G. (2004). On the asymptotic distribution of impulse response functions with long-run restrictions. *Econometric Theory*, **20**, 891–903.

Wolf, M. and Wunderli, D. (2012). Bootstrap joint prediction regions. Unpublished Manuscript.

Yang, M. (1998). On identifying permanent and transitory shocks in VAR models. *Economics Letters*, **58**, 171–175.