

Прикладная эконометрика, 2018, т. 49, с. 115–142.

Applied Econometrics, 2018, v. 49, pp. 115–142.

Н. Г. Арефьев, Р. А. Хабибуллин¹

Байесовская идентификация структурных векторных авторегрессий

В статье предложен новый метод идентификации структурных векторных авторегрессий на основе байесовского усреднения моделей. В отличие от существующих алгоритмов байесовского усреднения структурных векторных авторегрессий, предложенный метод позволяет идентифицировать циклические модели при выполнении ряда условий. Поиск моделей в рамках данного подхода осуществляется только по множеству моделей, идентифицируемых на данных. Для того чтобы проиллюстрировать корректность и стабильность результатов алгоритма, а также проанализировать его чувствительность к значениям истинных параметров, числу наблюдений и значениям гиперпараметров априорных распределений, проведен симуляционный анализ моделей малой размерности.

Ключевые слова: структурная векторная авторегрессия; идентификация; байесовское усреднение моделей; байесовский выбор моделей.

JEL classification: C11; C32; C52.

1. Введение

Структурные векторные авторегрессии (structural vector autoregressions, SVAR) применяются для анализа воздействия макроэкономической политики на макроэкономические показатели. Для оценки SVAR требуется решить задачу идентификации: из оцененных ковариационных матриц нужно найти все причинно-следственные связи, породившие наблюдаемые в данных ковариации. Во многих работах такая задача решается наложением априорных ограничений на структурные связи между переменными на основании теоретических предположений. Однако часто либо теория не дает достаточного количества априорных ограничений, либо существует ряд противоречащих друг другу теорий, на основании которых невозможно выбрать единственный набор ограничений на рассматриваемые параметры.

В данной работе предлагается метод выбора набора ограничений на структурные параметры SVAR-моделей, основанный на методологии байесовского усреднения моделей. Идея подобных методов заключается в оценке апостериорных вероятностей ограничений, что дает критерий для сравнения наборов ограничений между собой. В данной работе также

¹ Арефьев Николай Геннадьевич — Национальный исследовательский университет «Высшая школа экономики», Москва; n.arefiev@gmail.com.

Хабибуллин Рамис Арсланович — Национальный исследовательский университет «Высшая школа экономики», Москва; rawirtschaft@gmail.com.

оценивается качество работы предложенного алгоритма на основании серии симуляционных экспериментов моделей малой размерности. Новизна работы заключается в том, что предложенный метод позволяет идентифицировать не только рекурсивные, но и циклические SVAR-модели.

Оценка апостериорных вероятностей SVAR-моделей является вычислительно затратной процедурой, и поэтому в данной работе используются численные методы для их аппроксимации. В рамках данной работы для байесовской оценки параметров SVAR-моделей применяется алгоритм Метрополиса–Гастингса (Baumeister, Hamilton, 2015). Для расчета апостериорных вероятностей моделей используется метод численного интегрирования (Chib, Jeliazkov, 2001).

Данная работа близка к исследованиям, посвященным байесовскому выбору и усреднению моделей. В наиболее близких работах (Ahelegbey et al., 2016; George et al., 2008) предложены алгоритмы Monte Carlo Markov Chain Model Composition (MC3) и стохастического поиска моделей, позволяющие идентифицировать ограничения SVAR-моделей, но только при условии треугольных (рекурсивных, ациклических) моделей.

Байесовские методы также позволяют накладывать более мягкие ограничения в качестве априорных распределений значений параметров. Такие априорные предположения о параметрах недостаточны для идентификации единственно возможной модели, но, тем не менее, могут использоваться для интерпретации распределений таких величин, как функции импульсного отклика. Авторы первой работы, посвященной байесовской оценке SVAR-моделей (Sims, Zha, 1998), рассматривали только рекурсивные модели, но уже в (Waggoner, Zha, 2003) был предложен алгоритм оценки параметров с помощью сэмплирования Гиббса для нерекурсивных моделей. В работах (Baumeister, Hamilton, 2015; Arias et al., 2014) используются знаковые ограничения: априори предполагается, что значение параметра может быть только положительным или только отрицательным. Помимо априорных предположений на значения параметров, в рамках байесовского подхода существует возможность использовать априорную информацию о функциях импульсного отклика, декомпозиции прогнозных дисперсий и прочих величин, зависящих от структурных параметров. Так, в работе (Plagborg-Møller, 2016) предлагается использовать априорные распределения на функции импульсного отклика и оценивать структурные векторные модели скользящего среднего. Использование такого рода априорных предположений является одним из возможных расширений предлагаемой в данной работе методологии.

В качестве альтернативного подхода выступают методы данно-ориентированной идентификации, позволяющие использовать информацию, содержащуюся в данных, для тестирования и выбора идентифицирующих ограничений модели. Во-первых, это методы, основанные на анализе значимости частных корреляций, представленные работами (Arefiev, 2014; Fragetta, Melina, 2013; Swanson, Granger, 1997; Moneta, 2008; Pearl, 2015). Предложенные в данных работах алгоритмы позволяют отвергнуть наборы ограничений, не удовлетворяющие наблюдаемым ковариациям, однако они подвержены проблеме оценивания при условии последовательно тестируемых гипотез. Предлагаемый в настоящей работе алгоритм не выбирает одну единственную модель, но позволяет «взвешивать» модели по вероятностям, что учитывает проблему неопределенности относительно моделей.

Во-вторых, это методы, основанные на идентификации с помощью условной гетероскедастичности и изменяющихся во времени параметрах. В работе (Lütkepohl, Velinov, 2016) представлен обзор таких методов. Однако предложенные в рамках этого подхода алгоритмы требуют дополнительных предположений о процессе, генерирующем данные: в первую

очередь о характере изменения параметров во времени (процесс с марковскими переключениями или процесс случайного блуждания параметров и т. д.), а также об изменении ограничений во времени. Эти предположения дают возможность тестировать ограничения на структурные параметры. В рамках данной работы рассматривается только случай постоянных во времени структурных параметров.

Настоящая статья организована следующим образом. В разделе 2 описана постановка задачи байесовского усреднения SVAR-моделей и приведены условия, требуемые для реализации предложенного метода. В разделе 3 описана процедура Метрополиса–Гастингса и метод расчета маргинальной функции плотности для него. Раздел 4 описывает основные проблемы и недостатки предложенной методологии. В разделе 5 представлены результаты анализа симуляций.

2. Постановка задачи и условия работы метода

В рамках данной работы предполагается структурная авторегрессия в следующей форме:

$$y_t' A = z_t' B + u_t' D^{-0.5},$$

где y_t — $(n \times 1)$ -вектор эндогенных переменных модели, z_t — $(k \times 1)$ -вектор предопределенных и экзогенных переменных, включающий в себя как лаговые значения эндогенных переменных, так и другие детерминированные и экзогенные переменные. Элементы матрицы A соответствуют структурным связям между эндогенными переменными, а элементы матрицы B — связям между эндогенными и экзогенными переменными. Матрица D^{-1} является диагональной, j -й диагональный элемент которой соответствует дисперсии структурного шока j -й эндогенной переменной.

При этом предполагается, что модель гауссовская. Это значит, что структурные шоки сгенерированы из стандартного нормального многомерного распределения $u_t \sim \mathcal{N}_n(0, I_n)$, причем $\mathbb{E}u_t u_t' = I_n$. Предположение об ортогональности структурных шоков мотивировано гипотезой о том, что отсутствуют каузальные связи между структурными шоками, а все каузальные связи между переменными уже явно учтены в модели.

Описанное выше уравнение для каждого периода времени $t = 1, \dots, T$ можно представить в виде матриц, состоящих из векторов-строк данных $Y = [y_1, \dots, y_T]'$, $Z = [z_1, \dots, z_T]'$ и ошибок $U = [u_1, \dots, u_T]'$:

$$YA = ZB + UD^{-0.5}. \quad (1)$$

Пусть $X = [Y, Z]$ — матрица данных, состоящая из векторов-строк эндогенных и экзогенных переменных. Далее функции плотности будут обозначаться как $p(\cdot | \Omega)$ при условии Ω . Обозначим также $\theta = [A, B, D]$ — все параметры структурной модели. «Моделью» $\mathcal{M}$ будет называться некоторый набор ограничений и априорных предположений о параметрах, позволяющий идентифицировать систему (1). Предполагается также, что эквивалентны следующие формы записи функции правдоподобия: $p(X | \theta; \mathcal{M}) = p(Y, Z | \theta; \mathcal{M})$, а также апостериорной функции плотности параметров: $p(\theta | Y, Z; \mathcal{M}) = p(\theta | X; \mathcal{M})$.

Предлагаемый метод основан на двух других, уже существующих методах. Во-первых, это теория тестируемой графической идентификации, описанная в работе (Arefiev, 2014),

в которой были выведены условия тестируемости ограничений SVAR-модели на данных. В рамках предлагаемого метода для точечных ограничений рассматриваются только те модели, которые являются потенциально тестируемыми на данных по методологии (Arefiev, 2014).

Во-вторых, это байесовский поиск и усреднение моделей, подробно описанные в работах (Gelfand, Dey, 1994; Madigan et al., 1995, 1996; Hoeting et al., 1999; Giudici, Green, 1999). В этом направлении разрабатываются методы сравнения моделей в рамках байесовского подхода.

Рассмотрим оба метода более подробно.

Критерий графической идентификации (Arefiev, 2014) основан на тестировании ограничений на трансформированную матрицу концентрации, которая может быть выражена

как $C = \begin{bmatrix} A \\ B \end{bmatrix} D [A' \ B']$. Значения элементов матриц A , B и D определяют значения элементов

матрицы C . Следовательно, ограничения на значения матриц A и B влияют на область возможных значений элементов матрицы C . Более того, существует такой набор ограничений на элементы матриц A и B , что некоторые элементы матрицы C окажутся нулевыми. Сеть концентрации отражает подобного рода ограничения.

При анализе данных недиагональные элементы матрицы C часто интерпретируются ненаправленным графом, который далее будет называться сетью концентрации.

Вершины такого графа — это случайные переменные структурной модели. Таким образом, первой переменной вектора X соответствует первая вершина графа, второй переменной — вторая вершина, и так далее. Ребра в таком графе определены следующим образом: ребро между вершинами i и j в сети концентрации отсутствует тогда и только тогда, когда для недиагонального элемента трансформированной матрицы концентрации выполняется ограничение $c_{ij} = 0$. Иначе ребро между вершинами i и j присутствует. При этом существует следующая связь между структурными параметрами и сетью концентрации: почти для всех значений параметров ребро от вершины x_i к вершине x_j присутствует в сети концентрации тогда и только тогда, когда в системе уравнений (1) присутствует хотя бы одно уравнение, в которое включены одновременно обе переменные x_i и x_j .

Таким образом, нулевые элементы матрицы концентрации дают информацию об ограничениях на структурные параметры SVAR-модели.

Перейдем к методологии байесовского усреднения моделей.

В данном направлении литературы предполагается, что в условиях неопределенности (относительно истинной) каждую модель можно задать с помощью некоторой дискретной случайной величины M , принимающей значение на множестве априорно заданных моделей $s = 1, \dots, S$. Далее реализация случайной величины $M = s$ будет означать, что истинной моделью является модель s .

В связи с этим исследователь может иметь некоторые априорные представления о вероятностях моделей $p(M = s)$ из некоторого априорно заданного множества моделей $s = 1, \dots, S$ (в литературе часто предполагается одинаковая априорная вероятность каждой модели: $p(M = s) = S^{-1} \ \forall s$). Предполагается, что набор данных X дает дополнительную информацию о вероятностях моделей, что дает критерий для их сравнения. Более того, это дает возможность подсчитать апостериорное распределение любой функции от структурных параметров при условии предполагаемого множества моделей с ненулевой априорной вероятностью:

$$p(g(\theta) | X) = \sum_{s=1}^S p(g(\theta) | X; \mathcal{M} = s) p(\mathcal{M} = s | X).$$

Апостериорная вероятность модели в общем виде может быть подсчитана по следующей формуле²:

$$p(\mathcal{M} = s | X) = \frac{p(Y | Z; \mathcal{M} = s) p(\mathcal{M} = s)}{\sum_{k=1}^S p(Y | Z; \mathcal{M} = k) p(\mathcal{M} = k)},$$

где $p(\mathcal{M} = s)$ — априорная вероятность модели s , а $p(Y | Z; \mathcal{M} = s)$ — маргинальная функция правдоподобия, которая рассчитывается как интеграл (по всем параметрам модели $\theta = [A, B]$) от совместной функции плотности параметров и данных для заданной модели:

$$p(Y | Z; \mathcal{M} = s) = \int_{\theta \in \Theta} p(Y, \theta | Z; \mathcal{M} = s) d\theta = \int_{\theta \in \Theta} p(Y | Z, \theta; \mathcal{M} = s) p(\mathcal{M} = s) d\theta. \quad (2)$$

Таким образом, задача расчета апостериорных вероятностей может быть разделена на две подзадачи: во-первых, оценка апостериорных распределений параметров SVAR-модели при условии данных, а во-вторых, расчет маргинальной функции правдоподобия и апостериорных вероятностей рассматриваемых моделей.

В рамках данной работы предлагается использовать процедуру Метрополиса–Гастингса для оценки апостериорных параметров. Однако оценка маргинальной функции правдоподобия в такой задаче затруднена. Во-первых, апостериорные параметры SVAR-модели не принадлежат ни к одному классу известных распределений. Во-вторых, расчет интеграла совместной плотности параметров и данных по распределению параметров в явном виде крайне сложен. По этой причине в данной работе используется численный метод оценки маргинальной функции правдоподобия, предложенный в (Chib, Jeliazkov, 2001).

В этом разделе предложенный метод будет описан более подробно. Для того чтобы предложенная процедура состоятельно оценивала апостериорные распределения параметров и апостериорные вероятности моделей при условии данных, необходимо выполнение ряда условий.

Условие 1. Число ограничений в истинной модели не меньше минимально необходимого для идентификации.

В рамках данной работы предполагается, что исследователь априори не знает ограничений истинной модели. Для того чтобы истинная модель имела более высокую апостериорную вероятность, чем другие рассматриваемые модели, необходимо, чтобы для множества допустимых значений параметров сеть концентрации отличалась от сети концентрации других рассматриваемых моделей. Таким образом, в рамках данного метода предполагается, что истинная модель содержит не меньшее число ограничений, чем того требует необходимое условие идентификации SVAR-моделей, т. е. не менее $0.5n(n-1)$ ограничений (Rothenberg, 1971). Более того, если это условие не выполняется, то вероятность любой тестируемой на данных модели стремится к нулю почти наверняка. Однако даже при выполнении этого условия в конечных выборках сохраняется риск неверного выбора моделей.

² Данная запись верна только в предположении о слабой экзогенности Z по отношению к Y (см. условие 5 ниже). Только в этом случае $p(X | \mathcal{M} = s) = p(Y | Z; \mathcal{M} = s) p(Z)$, где $p(Z)$ одинакова для любого s .

Условие 2. *Истинная модель содержит достаточно ограничений включения и исключения для полной идентификации модели.*

Несмотря на то что условие 1 предполагает выполнение необходимого условия идентификации, это не гарантирует выполнение достаточного рангового условия, предложенного в работе (Rubio-Ramirez et al., 2010). Однако предполагается, что ранговое условие также выполняется, что, с учетом условия 1, гарантирует идентифицируемость рассматриваемых моделей.

В данной работе рассматриваются только идентифицируемые модели по двум причинам. Во-первых, для заданного числа эндогенных и экзогенных переменных число неидентифицируемых моделей, эквивалентных на данных, сильно превышает число идентифицируемых моделей, потенциально отличимых на данных. Поэтому простым перебором невозможно рассмотреть все множество неидентифицируемых моделей. Во-вторых, если параметр неидентифицируем, его апостериорное распределение пропорционально априорному распределению. В случае использования априорного распределения параметра с большим значением дисперсии расчет интеграла в (2) с помощью методов Монте-Карло на марковских цепях (Monte Carlo Markov Chain, MCMC) может дать несостоятельную оценку маргинальной функции правдоподобия при условии конечного числа итераций.

Условие 3. *Априорные распределения параметров без ограничений являются собственными.*

Данная предпосылка исключает несобственные априорные распределения, для которых нарушаются некоторые из свойств функций плотности. Самым известным примером подобной функции плотности является абсолютно неинформативное распределение параметра $p(\theta) \propto 1 \quad \forall \theta$, интеграл от которого в пределах изменения параметра $(-\infty; \infty)$ не равен 1. В книге (Коор et al., 2008, р. 288) показано, что для подобных априорных распределений вывод апостериорных вероятностей моделей может быть некорректен.

Условие 4. *Истинная модель принадлежит классу рассматриваемых моделей, которым приписывается ненулевая априорная вероятность (M-closed problem).*

Так, в работе (Clarke et al., 2013) показано, что нарушение этого условия в задаче байесовского усреднения моделей приводит к несостоятельным оценкам апостериорных вероятностей моделей.

Истинное значение параметра должно принадлежать множеству значений, где он априори распределен. Например, если априори предполагается, что параметр принимает только положительные значения, а в истинной модели это не так, оценки параметров будут смещенными и несостоятельными.

Условие 5. *Переменные матрицы Z слабо экзогенны по отношению к Y .*

Согласно работе (Engle et al., 1983), слабая экзогенность Z характеризует следующую ситуацию. Представим, что совместная плотность данных X может быть описана набором параметров $\psi = \theta \times \lambda$ (здесь $\times$ — знак декартова произведения), где $\theta = \{A, B\}$ — набор параметров функции правдоподобия $Y|Z$, а λ — набор параметров безусловной функции плотности Z . В этом случае совместная функция плотности представима в виде $p(X|\psi) = p(Y, Z|\psi) = p(Y|Z; \theta)p(Z|\lambda)$. При этом не существует ограничений,

которые бы функционально связывали параметры θ и λ . Этот факт позволяет игнорировать распределение $p(Z|\lambda)$ при анализе условного распределения эндогенных переменных $p(Y|Z;\theta)$.

Представим, например, что векторы-строки z_i' матрицы Z распределены нормально с математическим ожиданием μ_Z и ковариационной матрицей Σ_{ZZ} : $z_i | \lambda \sim \mathcal{N}_k(\mu_Z, \Sigma_{ZZ})$. При этом векторы-строки y_i матрицы Y описываются SVAR процессом $y_i | z_i; \theta \sim \mathcal{N}_n(ZBA^{-1}, (A^{-1})'A^{-1})$. В этом случае параметрами маргинального распределения экзогенных переменных является набор $\lambda = \{\mu_Z, \Sigma_{ZZ}\}$, а набор параметров условного распределения эндогенных переменных есть $\theta = \{A, B\}$. Если θ и λ не связаны общими ограничениями, то Z является слабо экзогенной по отношению к Y . Более подробно слабая экзогенность описана в работе (Engle et al., 1983, Definitions 2.4, 2.5).

Если же, например, предположить, что $\mu_Z = f(A)$, то слабая экзогенность нарушится, и одного условного распределения $p(Y|Z;\theta)$ окажется недостаточно для оценки параметров θ .

Условие 6. Структурные шоки SVAR-модели распределены нормально.

Согласно работе (Baumeister, Hamilton, 2015), апостериорные оценки параметров асимптотически сходятся к истинным значениям параметров независимо от предположений о распределении. Несмотря на это, расчет маргинальной функции правдоподобия и, как следствие, апостериорных вероятностей моделей, требует предположений о функциональной форме распределения структурных шоков. Хотя методология оценки SVAR-моделей в случае нарушения предположения о нормальности разработана в статье (Hyyärinen et al., 2010), но расчет соответствующих апостериорных вероятностей моделей выходит за рамки настоящей работы.

3. Метод³

В данном разделе представлен общий алгоритм оценки параметров SVAR-модели и апостериорных вероятностей моделей. Для оценки параметров использовался алгоритм Метрополиса–Гастингса (Baumeister, Hamilton, 2015), а для расчета маргинальных функций правдоподобия — метод численного интегрирования, предложенный в статье (Chib, Jeliazkov, 2001).

Метод расчета маргинальных функций правдоподобия исходит из того, что совместную функцию плотности параметров и данных для заданной модели s можно записать двумя способами:

$$p(Y, \theta | Z; \mathcal{M}=s) = p(Y | Z, \theta; \mathcal{M}=s)p(\theta | \mathcal{M}=s) = p(\theta | X; \mathcal{M}=s)p(Y | Z; \mathcal{M}=s).$$

Из приведенной формулы можно выразить логарифм маргинальной функции плотности для какого-то конкретного значения параметра θ^* (например, апостериорной оценки параметров):

$$\ln p(Y | Z; \mathcal{M}=s) = \ln p(Y | Z, \theta^*; \mathcal{M}=s) + \ln(p(\theta^* | \mathcal{M}=s)) - \ln p(\theta^* | X; \mathcal{M}=s). \quad (3)$$

³ Все программные скрипты могут быть найдены по ссылке: https://github.com/rakhbay/Bayesian_Identification.

Здесь $p(Y|Z, \theta^*; \mathcal{M}=s)$ — это функция правдоподобия, а $p(\theta^* | \mathcal{M}=s)$ — функция плотности априорного распределения параметров модели. Обе функции известны и могут быть рассчитаны в явном виде. Основные проблемы возникают при расчете апостериорной функции плотности $p(\theta^* | X; \mathcal{M}=s)$, ибо, как правило, при применении алгоритмов МСМС данная функция в общем виде неизвестна (либо не определена ее интеграционная константа). Идея метода (Chib, Jeliazkov, 2001) заключается в расчете $p(\theta^* | X; \mathcal{M}=s)$ на основании выборок, полученных процедурами МСМС.

Помимо упомянутых методов расчета маргинальных функций правдоподобия, в литературе существует ряд других методов численного интегрирования, однако метод (Chib, Jeliazkov, 2001) дает наиболее устойчивую оценку. Чтобы показать это, для сравнения были рассчитаны маргинальные функции правдоподобия с помощью алгоритма Annealed importance sampling (Neal, 2001) и метода гармонического среднего (Raftery et al., 2007).

В следующих подразделах будут подробно разобраны алгоритмы генерации выборок параметров и расчета апостериорных вероятностей моделей.

3.1. Процедура Метрополиса–Гастингса

В данной работе предлагается использовать для оценки апостериорных распределений параметров SVAR-моделей модифицированную процедуру Метрополиса–Гастингса. Для SVAR-модели в форме (1) предполагаются следующие априорные распределения:

$$d_j | A; \mathcal{M}=s \sim \text{Gamma}(\kappa_j^{\text{prior}}, \tau_j^{\text{prior}}), \quad b_j | A, D; \mathcal{M}=s \sim \mathcal{N}(0, d_j^{-1} V_{j.s}^{\text{prior}}), \\ a_j | \mathcal{M}=s \sim \mathcal{N}_k(0, S_{j.s}^{\text{prior}}).$$

В (Baumeister, Hamilton, 2015) выведены апостериорные плотности в следующем виде:

$$b_j | A, D; X; \mathcal{M}=s \sim \mathcal{N}(m_{j.s}^{\text{post}}, d_j^{-1} V_{j.s}^{\text{post}}), \quad V_{j.s}^{\text{post}} = \left((V_{j.s}^{\text{prior}})^{-1} + Z'Z \right)^{-1}, \quad m_{j.s}^{\text{post}} = V_{j.s}^{\text{post}} Z'Y a_j, \\ d_j | A; \mathcal{M}=s \sim \text{Gamma}(\kappa_j^{\text{post}}, \tau_{j.s}^{\text{post}}), \\ \kappa_j^{\text{post}} = 0.5T + \kappa_j^{\text{prior}}, \quad \tau_{j.s}^{\text{post}} = \tau_j^{\text{prior}} + 0.5Ta_j' \Omega_{j.s} a_j, \quad \Omega_{j.s} = T^{-1}(Y'Y - Z'Y V_{j.s}^{\text{post}} Y'Z).$$

Для известных значений матрицы A параметры матриц D и B могут быть сгенерированы из представленных выше распределений.

В работе (Rubio-Ramirez et al., 2010) показано, что с помощью ограничений на матрицу параметров при экзогенных переменных B можно идентифицировать отдельный шок или, в некоторых случаях, систему в целом. Однако в рамках алгоритма Метрополиса–Гастингса параметры A генерируются независимо от параметров B из маргинальной плотности $p(A)$. В связи с этим оценка параметров матрицы A не учитывает априорные предположения о параметрах матрицы B , и поэтому с помощью ограничений на эти параметры невозможно идентифицировать SVAR-модель с помощью данного метода. Это существенно сокращает множество идентифицируемых моделей.

Для каждого шага g процедуры МСМС параметры матрицы A в данной работе предлагается генерировать из условного распределения $p(A^{(g)} | B^{(g-1)}, D^{(g-1)}, X; \mathcal{M}=i)$ с помощью

процедуры Метрополиса–Гастингса со случайным блужданием. Каждый параметр генерируется как процесс случайного блуждания с ошибками, распределенными нормально с нулевым математическим ожиданием и дисперсией σ_{prop}^2 :

$$a_{ij}^{prop} = a_{ij}^{(g-1)} + v; \quad v \sim \mathcal{N}(0, \sigma_{prop}^2).$$

На данном шаге марковской цепи параметру присваивается значение $a_{ij}^{(g)} = a_{ij}^{prop}$, и оно принимается с вероятностью

$$\alpha(a_{ij}^{(g)} \rightarrow a_{ij}^{prop} | A_{(-i),(-j)}^{(g-1)}, D^{(g-1)}, B^{(g-1)}, X; \mathcal{M} = s) = \\ = \min \left\{ 1, \frac{p(Y | a_{ij}^{prop}, A_{(-i),(-j)}^{(g-1)}, D^{(g-1)}, B^{(g-1)}, Z) p(a_{ij}^{prop} | \mathcal{M} = s)}{p(Y | a_{ij}^{(g-1)}, A_{(-i),(-j)}^{(g-1)}, D^{(g-1)}, B^{(g-1)}, Z) p(a_{ij}^{(g-1)} | \mathcal{M} = s)} \right\}.$$

С вероятностью $1 - \alpha(a_{ij}^{(g)} \rightarrow a_{ij}^{prop} | A_{(-i),(-j)}^{(g-1)}, D^{(g-1)}, B^{(g-1)}, X; \mathcal{M} = s)$ присваивается значение с предыдущего шага марковской цепи: $a_{ij}^{(g)} = a_{ij}^{(g-1)}$. В статье (Baumeister, Hamilton, 2015) авторы предлагают генерировать параметры A из маргинального распределения $p(A | \mathcal{M} = i)$. Однако в этом случае возникает проблема с моделями, идентифицированными с помощью ограничений на матрицу B , описанная выше. Более того, авторы предлагают генерировать все параметры A одновременно. Однако, как показано в работе (Chib, Greenberg, 1995), такой генератор случайных чисел в среднем требует большее число итераций для сходимости марковской цепи к апостериорному распределению, чем блочный алгоритм Метрополиса–Гастингса, предлагаемый в настоящей работе.

3.2. Расчет маргинальной функции правдоподобия на основе процедуры Метрополиса–Гастингса

В работе (Chib, Jeliazkov, 2001) представлено расширение описанного выше метода для случая генерации параметров с помощью алгоритма Метрополиса–Гастингса, при этом сложность алгоритма растет с увеличением числа блоков в алгоритме.

Применительно к задаче расчета маргинальных функций правдоподобия для SVAR-модели необходимо найти следующую плотность распределения в точке апостериорных значений параметров θ^* :

$$\ln p(\theta^* | X; \mathcal{M} = s) = \ln p(B^* | A^*, D^*, X; \mathcal{M} = s) + \ln p(D^* | A^*, X; \mathcal{M} = s) + \ln p(A^* | X; \mathcal{M} = s).$$

Функциональные формы первых двух членов данной суммы известны и могут быть подсчитаны напрямую, последний же член неизвестен, однако он может быть записан с помощью цепной формулы

$$\ln p(A^* | X; \mathcal{M} = s) = \sum_{k=1}^{n(n-1)} \ln p(a_k^* | \{a_r^*, r < k\}, X; \mathcal{M} = s).$$

Предположим, что требуется рассчитать $p(a_k^* | \{a_r^*, r < k\}, X; \mathcal{M} = s)$. В работе (Chib, Jeliazkov, 2001) предлагается строить оценку данной плотности с помощью следующего алгоритма.

1. Зафиксировать $\{a_r^*, r < k\}$ и сгенерировать значения параметров из распределения $p(\{a_r^{(q)}\}_{r \geq k}, B^{(q)}, D^{(q)} | \{a_r^*\}_{r < k}, X; \mathcal{M} = s)$ алгоритмом Метрополиса–Гастингса ($q = 1, \dots, Q$, а Q — число симуляций марковской цепи).

2. Сгенерировать также значения параметров из распределения

$$p(\{a_r^{(f)}, r > k\}, B^{(f)}, D^{(f)} | \{a_r^*, r < k\}, a_k^*, X; \mathcal{M} = s),$$

где $f = 1, \dots, F$, а F — число симуляций марковской цепи. После этого для каждого шага алгоритма дополнительно сгенерировать с помощью случайного блуждания параметр $a_k^{(f)} = a_k^* + v$, где $v \sim \mathcal{N}(0, \sigma_{prop}^2)$.

3. Оценка искомой функции плотности параметров может быть подсчитана следующим образом:

$$p(a_k^* | \{a_r^*, r < k\}, X; \mathcal{M} = s) = \frac{Q^{-1} \sum_{l=1}^Q \alpha(a_{ij}^{(q)} \rightarrow a_{ij}^* | A_{(-i), (-j)}^{(q)}, D^{(q)}, B^{(q)}, X; \mathcal{M} = s) \pi(a_{ij}^{(q)} \rightarrow a_{ij}^*)}{F^{-1} \sum_{f=1}^F \alpha(a_{ij}^* \rightarrow a_{ij}^{(f)} | A_{(-i), (-j)}^{(f)}, D^{(f)}, B^{(f)}, X; \mathcal{M} = s)}.$$

Здесь $\alpha(x \rightarrow y | A_{(-i), (-j)}^{(f)}, D^{(f)}, B^{(f)}, X)$ — вероятность перехода марковской цепи из состояния x в состояние y при условии $A_{(-i), (-j)}^{(f)}, D^{(f)}, B^{(f)}; X$. $\pi(x \rightarrow y)$ — вспомогательная плотность Метрополиса–Гастингса (proposal density), которую в нашем случае можно рассчитывать. В рамках данной работы использовалась плотность нормального распределения с математическим ожиданием x и дисперсией σ_{prop}^2 , подсчитанная в точке x .

4. Проблемы методологии

У класса методов, основанных на критерии апостериорных вероятностей моделей, существует ряд внутренних проблем, не решенных на настоящий момент. Эти проблемы подробно описаны в (Gelman, Shalizi, 2013).

Во-первых, оценка вероятностей модели при заданных данных подвержена парадоксу Lindley, подробно описанному в работе (Shafer, 1982). Суть парадокса заключается в том, что при достаточно малом значении гиперпараметра дисперсии априорного распределения вероятность модели с ограничениями может оказаться высокой, в то время как в рамках частотного подхода гипотеза о равенстве параметра нулю не отвергается.

Парадокс может быть описан на следующем упрощенном примере. Пусть функция правдоподобия зависит только от одного параметра θ , априорное математическое ожидание которого равно нулю, и рассматриваются две модели. В модели без ограничений (unrestricted, UR) параметр имеет априорную дисперсию, значительно превышающую оценку дисперсии параметра, полученную методом максимального правдоподобия. В модели с ограничениями (restricted, R) априорная дисперсия примерно равна нулю. Если дисперсия примет нулевое значение, априорная функция плотности параметра равна $p(\theta = 0 | \mathcal{M} = R) = 1$ и $p(\theta | \mathcal{M} = R) = 0, \forall \theta \neq 0$. Тогда маргинальная функция правдоподобия такой модели равна значению функции правдоподобия в точке $\theta = 0$.

Значение же маргинальной функции правдоподобия модели без ограничений UR примерно равно нулю, когда априорная дисперсия стремится к бесконечности, ибо с ростом

дисперсии величина $p(\theta | \mathcal{M} = UR)$ для каждого значения θ стремится к 0, а значит, и интеграл $\int p(y | \theta; \mathcal{M} = UR)p(\theta | \mathcal{M} = UR)d\theta$ стремится к 0 для заданной функции правдоподобия $p(y | \theta; \mathcal{M} = UR)$.

Таким образом, если исследователь выберет максимально неинформативные априорные значения для модели без ограничений и максимально ограничит возможное распределение параметра для модели с ограничениями, значение маргинальной функции правдоподобия последней может оказаться больше по построению. Тогда апостериорная вероятность модели с ограничениями окажется значительно ниже апостериорной вероятности модели без ограничений, независимо от значений функции правдоподобия.

Во-вторых, помимо байесовского риска, существует проблема индуктивного выбора множества проверяемых моделей. Если истинная модель не принадлежит множеству проверяемых моделей (априори это невозможно состоятельно определить), исследователь сталкивается с риском неопределенности по Найту. Таким образом, исследователь вынужден выбрать как можно более широкое множество проверяемых моделей перед началом процедуры оценивания и проверки гипотез. С другой стороны, с ростом числа эндогенных и экзогенных переменных число возможных моделей растет экспоненциально, что вынуждает ограничивать множество рассматриваемых моделей. Это увеличивает риск ошибки выбора моделей.

В-третьих, Gelman, Shalizi (2013) утверждают, что процедура расчета апостериорных вероятностей моделей является индуктивной по своей природе и противоречит научному методу фальсификации и верификации научных теорий. Так, в работе (Porreger, 1956) показано, что невозможно на основе индуктивного перебора всех вероятных с точки зрения исследователя моделей получить состоятельную верификацию той или иной теории. В качестве альтернативы эти авторы предлагают метод апостериорного P -значения. Однако в настоящий момент этот метод является слабо разработанным, и вопрос о его реализации для SVAR-моделей выходит за рамки данной работы.

В-четвертых, как показано в комментарии Neal (1999) к работе (Chib, 1995), используемый метод расчета маргинальной функции правдоподобия может дать ошибочный результат в случае, когда рассматриваются неидентифицируемые модели с многомодальными апостериорными значениями параметров. По этой причине в данной работе рассматриваются только идентифицируемые модели со стандартными гауссовскими априорными предположениями.

5. Анализ симуляций

В рамках анализа симуляций в данном подразделе предполагается самая простая структура структурной векторной авторегрессии с двумя эндогенными и двумя экзогенными переменными, которые формируются как лаги эндогенных переменных.

Прежде всего, были сгенерированы векторы Z и ε размерности $1 \times T$, $T = 300$ как независимые реализации стандартных нормальных случайных величин $\mathcal{N}(0, 1)$.

Каждую модель можно выразить с помощью матриц ограничений на структурные параметры уравнения (1). Обозначим через Λ матрицу ограничений на параметры A ($\Lambda_{ij} = 0$, если $a_{ij} = 0$ для предполагаемой модели s , и $\Lambda_{ij} = 1$ иначе), а через Θ — матрицу ограничений на параметры B . Тогда истинные матрицы параметров могут быть выражены с помощью операции поэлементного умножения, обозначаемой символом $\circ$. Например, для модели 5,

представленной в Приложении 2, матрицы ограничений и матрицы параметров выглядит следующим образом:

$$\Lambda = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}; \quad \Theta = \begin{bmatrix} 0 & 1 \end{bmatrix} \Leftrightarrow A = \Lambda \circ \tilde{A} = \begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{bmatrix}; \quad B = \Theta \circ \tilde{B} = \begin{bmatrix} 0 & b_{12} \end{bmatrix}.$$

Для каждой модели и априори заданных матриц неограниченных структурных параметров $\tilde{A}$ и $\tilde{B}$ были сгенерированы эндогенные переменные Y :

$$Y(\Lambda \circ \tilde{A}) = Z(\Theta \circ \tilde{B}) + UD^{-0.5} \Leftrightarrow Y\left(\Lambda \circ \begin{bmatrix} 1 & 5 \\ 3 & 1 \end{bmatrix}\right) = Z(\Theta \circ \begin{bmatrix} 3 & 5 \end{bmatrix}) + UD^{-0.5}. \quad (4)$$

Столь большие значения параметров истинной модели $\tilde{A}$ и $\tilde{B}$ были выбраны для того, чтобы в матрице концентрации не возникало близких к нулю значений в тех элементах, в которых ограничения не предполагают нулевых значений. Это сделано для того, чтобы модели могли быть различимы на данных. Для построения оценки байесовских структурных векторных авторегрессий были выбраны следующие гиперпараметры априорных распределений структурных параметров SVAR-модели без ограничений:

$$d_j \sim \text{Gamma}(0.5, 2c) \Leftrightarrow \pm d_j^{0.5} \sim \mathcal{N}(0, c), \quad j = 1, 2, \\ b_j \sim \mathcal{N}(0, c), \quad j = 1, 2; \quad a_{ij} \sim \mathcal{N}(0, c), \quad i, j = 1, 2,$$

где c — параметр дисперсии. Для модели без ограничений этот параметр был выбран равным 200, а для модели с ограничениями — 0.02. Такой выбор параметров был сделан с целью использования как можно меньшей априорной информации при построении структурной модели. Коэффициент c — мера дисперсии априорного распределения генерируемых параметров. В силу парадокса Lindley в целях байесовского усреднения моделей выбор значений гиперпараметра дисперсии, много больших оценки дисперсии методом максимального правдоподобия $c \gg \sigma_{MLE}^2$, может привести к выбору моделей с максимальным числом ограничений при оценке апостериорных вероятностей моделей. Но слишком маленькие априорные дисперсии приведут к тому, что, во-первых, оценка параметров будет сильно смещена в сторону априорного математического ожидания параметра, и выводы из оценок окажутся неверными. Во-вторых, параметры с небольшими дисперсиями могут оказаться численно неотличимыми от ограниченных параметров. По этой причине были выбраны промежуточные априорные значения дисперсий параметров.

В силу размерности эндогенных и экзогенных переменных необходимо проанализировать результаты метода для $2^{0.5n(n+1)+k} = 2^4 = 16$ моделей. Случай модели с отсутствием ограничений на структурные параметры был исключен из анализа. Все структурные модели, рассмотренные в анализе симуляций, выписаны в Приложении 2. Априорные вероятности всех моделей предполагаются одинаковыми. Несмотря на условие 1 об априори большом числе ограничений в истинной модели, были проанализированы максимально неинформативные априорные предположения.

В силу условия 2 об идентифицируемости истинной модели, из анализа была исключена модель 12. В силу предпосылки о возможности тестирования истинной модели на данных

Таблица 1. Средние (по 1000 выборкам данных) значения апостериорных вероятностей моделей, подсчитанные методом (Chib, Jeliazkov, 2001).

Тестируемая модель	Истинная модель						
	3	5	6	7	10	11	15
1	0.0001 (3.3)	0 (2.2)	0 (8.2)	0 (2.4)	0 (5.8)	0 (5.0)	0 (2.0)
2	0.0001 (3.3)	0 (3.7)	0 (3.4)	0 (2.2)	0 (3.4)	0 (3.8)	0 (2.0)
3	0.7801 (3.2)	0 (2.7)	0 (9.1)	0 (2.4)	0 (5.0)	0 (3.9)	0 (2.2)
4	0.2196 (1.5)	0 (1.3)	0 (7.8)	0 (1.3)	0 (1.4)	0 (3.0)	0 (1.4)
5	0 (4.7)	0.998 (0.0)	0.02 (0.0)	0.0001 (0.0)	0 (3.5)	0 (4.6)	0 (0.0)
6	0 (2.3)	0.002 (1.3)	0.9742 (1.1)	0.0001 (1.3)	0 (3.3)	0 (4.4)	0 (1.0)
7	0 (2.1)	0 (1.5)	0 (2.1)	0.9127 (1.3)	0 (3.3)	0 (4.1)	0 (1.1)
8	0.0001 (1.5)	0 (2.6)	0 (2.2)	0 (1.3)	0 (7.8)	0 (2.8)	0 (1.3)
9	0 (1.5)	0 (1.6)	0 (8.2)	0 (1.4)	0.0092 (3.1)	0.0007 (1.3)	0 (1.4)
10	0 (2.1)	0 (1.5)	0 (2.1)	0 (1.3)	0.9899 (3.3)	0.0009 (4.4)	0 (1.0)
11	0 (2.3)	0 (1.8)	0 (1.8)	0 (1.3)	0 (2.7)	0.9984 (4.1)	0 (1.0)
12	0 (4.7)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
13	0 (0.6)	0 (0.4)	0.0025 (0.3)	0 (0.4)	0.0004 (0.4)	0 (0.0)	0.0001 (1.2)
14	0 (0.4)	0 (0.4)	0.0033 (0.3)	0 (0.4)	0.0006 (0.4)	0 (0.1)	0.0001 (0.4)
15	0 (0.7)	0 (0.7)	0 (0.7)	0.0872 (0.4)	0 (0.7)	0 (0.1)	0.9999 (0.5)

Примечание. Серым фоном отмечены апостериорные вероятности, превышающие значение 0.1. В скобках указаны дисперсии апостериорных вероятностей, умноженные на 1000.

по критерию (Arefiev, 2014), множество сгенерированных моделей включает в себя модели (3, 5, 6, 7, 10, 11, 15).

Для каждой из сгенерированных моделей предполагалось, что исследователь не знает истинную модель, описывающую процесс генерации данных. Следовательно, для каждой модели перебирались все возможные 15 моделей (кроме модели без ограничений), для каждой из них были сгенерированы данные X , подсчитаны маргинальная функция правдоподобия и апостериорные вероятности моделей по формуле (3).

Для каждой истинной модели эксперимент повторялся 1000 раз, с генерированием данных на каждой итерации заново. Для оценки параметров каждой модели использовалось 10 000 итераций марковской цепи со значением burn-in в 500 итераций.

В таблице 1 представлены средние оценки апостериорных вероятностей для 1000 выборок искусственно сгенерированных данных.

Как видно из табл. 1, для всех идентифицируемых на данных моделей апостериорные вероятности истинных моделей близки к 1, а вероятности остальных близки к 0. При этом дисперсии апостериорных вероятностей близки к нулю.

Таким образом, предлагаемый метод позволяет идентифицировать те модели, которые идентифицированы на данных по критерию работы (Arefiev, 2014).

5.1. Анализ чувствительности метода

В данном разделе будет проанализирована чувствительность метода к истинным значениям параметров, априорным предположениям о параметрах и числу наблюдений.

Для того чтобы проверить чувствительность алгоритма к значениям параметров, был реализован следующий ряд экспериментов. Была рассмотрена модель 4, в которой ограничен только параметр $b_{11} : b_{11} = 0$ (см. Приложение 2). Данная модель сравнивалась с моделью 6, в которой, помимо этого параметра, ограничено также значение параметра $a_{21} : a_{21} = 0$. При этом из Приложения 2 видно, что обе модели не являются эквивалентными на данных по критерию из работы (Arefiev, 2014). Однако естественно предположить, что если исследователь оценивает модель без ограничений 4, но при этом истинное значение параметра a_{21} численно близко к 0, то апостериорная вероятность модели 4 должна стремиться к 0, а вероятность модели 6 — к 1.

Были сгенерированы 1000 выборок из истинной модели 4 для каждого значения параметра a_{21} в интервале $[0, 0.5]$. Для каждой репликации была рассчитана апостериорная вероятность моделей 4 и 6. При этом в силу того, что модель 4 эквивалентна на данных нескольким другим моделям, использовались априорные вероятности моделей 4 и 6, равные 0.5, и нулевые априорные вероятности для остальных моделей.

Результаты экспериментов представлены на рис. 1.

Из графика видно, что для значений $a_{21} < 0.1$ при условии рассматриваемой параметризации апостериорная вероятность модели с ограничениями 6 принимает значения, близкие к 1, с 95%-ным доверительным интервалом в пределах $[0.85, 1]$. При этом для значений параметра $a_{21} > 0.47$ апостериорная вероятность модели 6 близка к нулю и лежит в интервале $[0, 0.2]$. Для значений параметра в интервале $[0.1, 0.47]$ апостериорные вероятности обеих моделей могут меняться от 0 до 1, это говорит о том, что в данном интервале истинную модель невозможно идентифицировать на основании предложенного метода — значение апостериорной вероятности модели без ограничений 4 может быть как выше, так и ниже, чем в модели с ограничениями 6.

Помимо малых значений структурных параметров, интересно проанализировать влияние гиперпараметра дисперсии априорного распределения структурных параметров. Из-за парадокса Lindley можно предположить следующую связь с гиперпараметром дисперсии. Пусть сравниваются две модели: одна с ограничениями (R), в которой априорный параметр дисперсии стремится к нулю, а другая — без ограничений (UR), в которой дисперсия стремится к бесконечности. Если для модели без ограничений параметр априорной дисперсии значительно превышает оценку дисперсии параметра методом максимального правдоподобия, апостериорная вероятность модели с ограничениями будет больше, чем у модели без ограничений, независимо от функции правдоподобия.

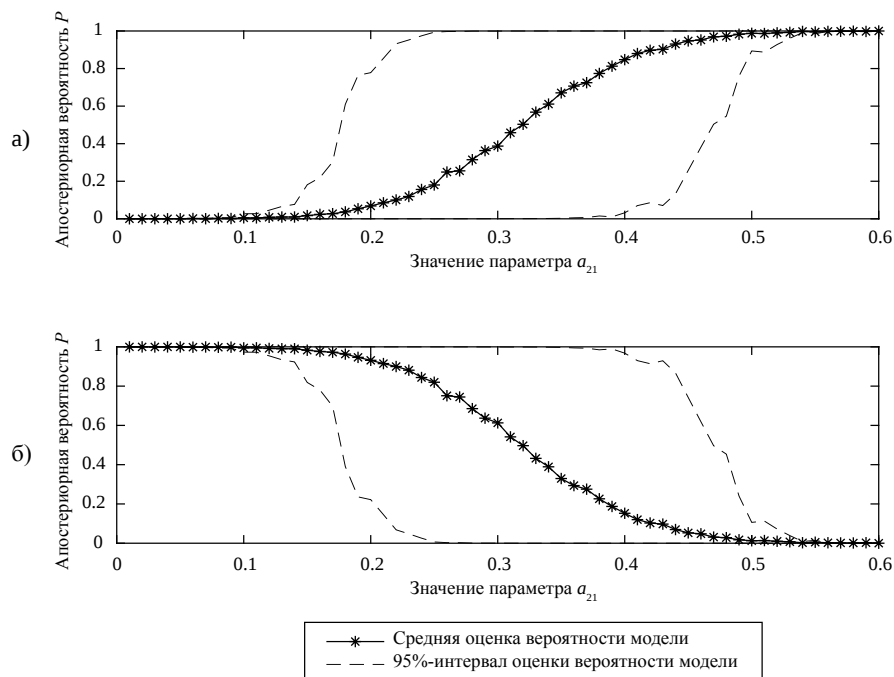


Рис. 1. Чувствительность апостериорных вероятностей:

а) модель 4 без ограничений на параметр a_{21} ;

б) модель 6 с ограничениями на параметр a_{21}

Для того чтобы проверить данный эффект, рассматривалась модель 14, в которой только параметр a_{12} не имеет ограничений. В качестве модели без ограничений рассматривалась ситуация, в которой гиперпараметр дисперсии априорного распределения a_{12} предполагается равным 200, а для модели с ограничениями — параметр c принимает значения от 10^{-16} до 10^{-14} .

Результаты экспериментов представлены на рис. 2. Из графика видно, что при достаточно низких значениях гиперпараметра дисперсии параметра c апостериорная вероятность модели с ограничениями превышает апостериорную вероятность модели без ограничений. При этом истинная модель — без ограничений.

Таким образом, априорные предположения о распределении параметров без ограничений могут повлиять на результат метода, и потому при реализации алгоритма байесовского усреднения моделей следует всегда проводить анализ чувствительности результатов к выбору гиперпараметров априорного распределения.

Был также проведен анализ чувствительности метода к числу наблюдений в выборке. Рисунок 3 иллюстрирует результаты следующего эксперимента. Были сгенерированы по 1000 выборок модели 10 с числом наблюдений T от 20 до 600. Затем были подсчитаны апостериорные вероятности модели 10 при априорном предположении о равной вероятности каждой из 15 моделей. Как видно из рисунка, в условиях малой выборки (в случае нашего эксперимента для $T < 200$) апостериорные вероятности модели имеют большую дисперсию: она может оказаться как близкой к 1, так и близкой к 0 с 95%-ной вероятностью. Однако при увеличении размера выборки апостериорная вероятность стремится к 1, а ее дисперсия — к 0.

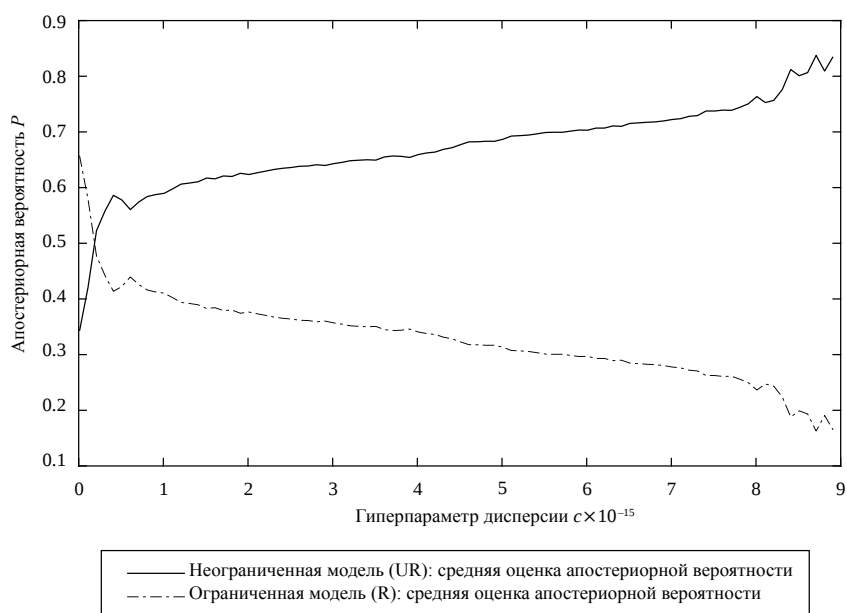


Рис. 2. Чувствительность апостериорных вероятностей модели без ограничений 14 и модели с ограничениями 15 к гиперпараметру априорного распределения c

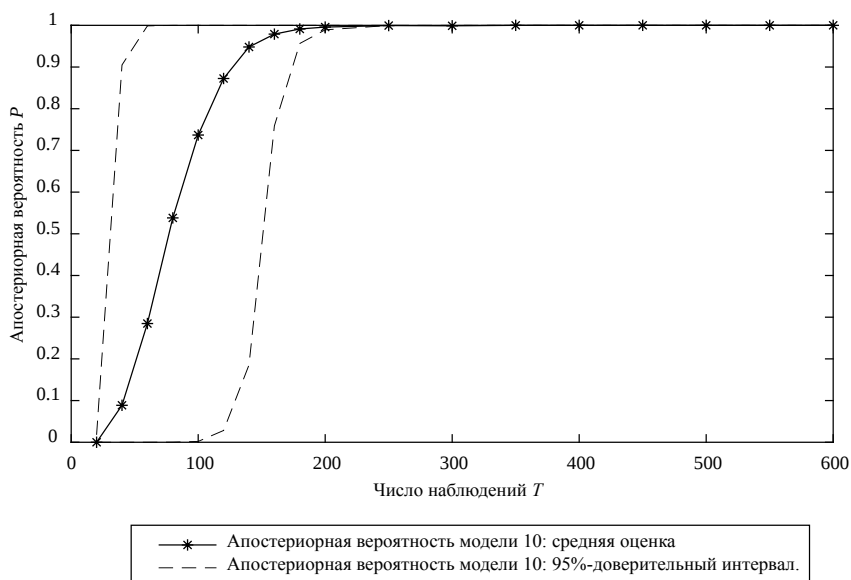


Рис. 3. Чувствительность апостериорных вероятностей к числу наблюдений в выборке. Истинная модель — 10

Таким образом, в условиях малого числа наблюдений предложенный в данной статье метод не дает возможности адекватно оценить апостериорные вероятности моделей.

5.2. Альтернативные методы расчета маргинальных функций правдоподобия

Помимо метода (Chib, Jeliazkov, 2001), маргинальные функции правдоподобия были рассчитаны с помощью методов, активно используемых в машинном обучении для решения подобной задачи для байесовских сетей. Это алгоритмы гармонического среднего (Raftery et al., 2007) и Annealed importance sampling (Neal, 2001). Однако все перечисленные выше методы дают недостоверные оценки апостериорных вероятностей моделей. В Приложении 3 дано краткое описание алгоритмов, а таблицы 2 и 3 показывают результат симуляций для двух альтернативных методов. Серой заливкой в них отмечены апостериорные вероятности, превышающие 0.1.

Из таблиц видно, что оба метода дают смещенную оценку апостериорных вероятностей моделей. По этой причине авторами предлагается использовать метод (Chib, Jeliazkov, 2001), дающий более стабильные оценки апостериорных вероятностей, хотя и более сложный с вычислительной точки зрения.

5.3. Сравнение предложенного метода с другими

В данном подразделе обсуждается вопрос о том, как предложенный в настоящей работе метод соотносится с другими, представленными в литературе (см. George et al., 2008; Ahelegbey et al., 2016). Этот вопрос может быть разделен на два более конкретных вопроса. Во-первых, как время компьютерных вычислений, необходимое для оценки модели с использованием предложенного метода, соотносится со временем вычислений, необходимым для оценки традиционными методами? Во-вторых, какой метод дает наиболее достоверные результаты?

Предложенный метод проигрывает другим методам по времени вычислений, т. к. он основан на точном решении, а не на аппроксимации методом стохастического поиска или процедуре МСЗ. Поскольку рассматриваемая задача по сложности относится к классу NP-полных (Арефьев, 2014), использование точного решения сильно ограничивает размерность модели, которая может быть оценена за приемлемое время. Так, модель, включающая в себя три переменные с ограничениями на матрицу текущих эффектов и на матрицу эффектов первого лага, может быть оценена на персональном компьютере за несколько часов без использования каких-либо специальных приемов программирования. Модель, включающая в себя 4–5 переменных, может быть решена за то же время на персональном компьютере при использовании методов программного ускорения, таких как параллельные вычисления на графическом сопроцессоре (Aldrich et al., 2011). Модель, включающая 6 или 7 переменных, потребует использования суперкомпьютера. Естественным следующим шагом развития предложенного в работе подхода является включение в алгоритм какого-либо аппроксимативного метода поиска истинной модели, что позволит расширить применимость подхода до моделей, включающих несколько десятков переменных.

Чтобы понять, какой из методов дает наиболее достоверные результаты, можно рассмотреть два случая. В первом заранее не известно, является ли истинная модель рекурсивной или циклической. Так как представленные в литературе (George et al., 2008; Ahelegbey et al., 2016) методы предполагают априори, что истинная модель является рекурсивной, в случае циклической истинной модели эти методы достоверно выберут ложную модель. В рассмотренном же выше примере предложенный в данной работе метод с высокой апостериорной вероятностью выбирает истинную модель. Таким образом, этот метод позволяет оценивать некоторый класс циклических моделей, что невозможно сделать другими существующими методами.

Таблица 2. Средние (по 1000 репликациям данных) значения апостериорных вероятностей моделей, подсчитанные методом гармонического среднего

Тестируемая модель	Истинная модель						
	3	5	6	7	10	11	15
1	0.2558	0.3678	0.0992	0.1751	0.0151	0.209	0.0279
2	0.2873	0.0379	0.2798	0.1619	0.3409	0.0002	0.0186
3	0.4569	0	0	0.3716	0	0.2778	0.0461
4	0	0.0605	0.143	0.0033	0	0	0
5	0	0.5335	0.0049	0.067	0	0	0.0589
6	0	0.0003	0.473	0.0626	0	0	0.0462
7	0	0	0	0.1585	0	0	0.108
8	0	0	0	0	0.2818	0.0019	0
9	0	0	0	0	0.0029	0.1299	0.0501
10	0	0	0	0	0.3593	0.3687	0.0497
11	0	0	0	0	0	0.0125	0.1072
12	0	0	0	0	0	0	0.0013
13	0	0	0	0	0	0	0.1418
14	0	0	0	0	0	0	0.0981
15	0	0	0	0	0	0	0.2461

Таблица 3. Средние (по 1000 репликациям данных) значения апостериорных вероятностей моделей, подсчитанные методом Annealed importance sampling

Тестируемая модель	Истинная модель						
	3	5	6	7	10	11	15
1	0	0	0	0.0089	0	0	0
2	0	0.0001	0.1475	0.004	0.0788	0.0132	0.0004
3	1	0	0	0.0385	0	0.0845	0.0033
4	0	0.7683	0.5392	0.1464	0.2523	0.1577	0.0337
5	0	0.2317	0	0.0642	0	0	0.002
6	0	0	0.3133	0.0613	0	0	0.0023
7	0	0	0	0.6767	0	0	0.0168
8	0	0	0	0	0	0.0065	0.566
9	0	0	0	0	0	0	0
10	0	0	0	0	0.6689	0.0368	0.0035
11	0	0	0	0	0	0.7012	0.023
12	0	0	0	0	0	0	0.2086
13	0	0	0	0	0	0	0.0149
14	0	0	0	0	0	0	0.0169
15	0	0	0	0	0	0	0.1083

Во втором случае заранее известно, что истинная модель является рекурсивной. В примерах, рассмотренных авторами и в литературе, все методы в этом случае выбирают истинную модель с высокой апостериорной вероятностью. Однако эти примеры не показывают, насколько надежным является тот или иной метод в общем случае. Ответ на этот вопрос зависит от множества факторов, включая число наблюдений, значение априорной дисперсии параметров, значение истинных параметров и т. д., и требует отдельного исследования. Однако вряд ли ответ на этот вопрос важен с практической точки зрения, т. к. сложно представить ситуацию, в которой заранее известно, что истинная модель рекурсивна, но при этом неизвестен рекурсивный порядок уравнений. Если же рекурсивный порядок уравнений заранее известен, то все рассматриваемые в данной работе методы являются избыточными.

Таким образом, предложенный метод проигрывает имеющимся в литературе методам по скорости вычислений, однако дает более достоверные оценки в случае, если заранее неизвестно, является ли истинная модель циклической или рекурсивной.

6. Заключение

В работе предложен метод байесовского усреднения SVAR-моделей на основе оценивания параметров методом Метрополиса–Гастингса (Baumeister, Hamilton, 2015) и оценки маргинальной функции правдоподобия методом (Chib, Jeliazkov, 2001). В отличие от существующих алгоритмов, предложенный метод позволяет идентифицировать циклические модели наряду с рекурсивными, хотя расчет апостериорных вероятностей моделей методами работ (George et al., 2008) и (Ahelegbey et al., 2016) можно сделать с меньшими затратами операционного времени. В качестве расширения исследования предполагается рассмотреть аппроксимативные методы поиска моделей, что даст возможность применять метод для моделей большей размерности.

На основании ряда симуляционных экспериментов было показано, что метод в среднем (для этих экспериментов) дает наибольшую оценку апостериорной вероятности истинным моделям при выполнении ряда требований, описанных в разделе 2. Симуляционные эксперименты также показали, что использование численного метода расчета маргинальной функции правдоподобия (Chib, Jeliazkov, 2001) дает более стабильные оценки, чем методы гармонического среднего и Annealed importance sampling. Предложенный метод был также проверен на чувствительность к априорным гиперпараметрам структурных параметров, значениям истинных параметров и числу наблюдений.

Благодарности. Исследование осуществлено в рамках Программы фундаментальных исследований НИУ ВШЭ. Авторы выражают благодарность Б. Б. Демешеву, Д. В. Левандо, О. А. Малаховской, С. Э. Пекарскому, А. А. Пересецкому, А. Д. Слестникову, Д. Фантаццини, А. М. Ярину и двум анонимным рецензентам за ценные комментарии к данной работе.

Список литературы

Ahelegbey D. F., Billio M., Casarin R. (2016). Bayesian graphical models for structural vector autoregressive processes. *Journal of Applied Econometrics*, 31 (2), 357–386.

Aldrich E. M., Fernández-Villaverde J., Gallant A. R., Rubio-Ramírez J. F. (2011). Tapping the supercomputer under your desk: Solving dynamic equilibrium models with graphics processors. *Journal of Economic Dynamics and Control*, 35 (3), 386–393.

Arefiev N. (2014). Structural models with testable identification. *Higher School of Economics Research Paper No. WP BRP 79/EC/2014*.

Arias J., Rubio-Ramirez J. F., Waggoner D. F. (2014). Inference based on SVAR identified with sign and zero restrictions: Theory and applications. *International Finance Discussion Papers Number 1100*.

Baumeister C., Hamilton J. D. (2015). Sign restrictions, structural vector autoregressions, and useful prior information. *Econometrica*, 83 (5), 1963–1999.

Bishop C. M. (2006). *Pattern recognition and machine learning*. Springer.

Chib S. (1995). Marginal likelihood from the Gibbs output. *Journal of the American Statistical Association*, 90 (432), 1313–1321.

Chib S., Greenberg E. (1995). Understanding the Metropolis–Hastings algorithm. *The American statistician*, 49 (4), 327–335.

Chib S., Jeliazkov I. (2001). Marginal likelihood from the Metropolis–Hastings output. *Journal of the American Statistical Association*, 96 (453), 270–281.

Clarke J. L., Clarke B., Yu C.-W. (2013). Prediction in M-complete problems with limited sample size. *Bayesian Analysis*, 8 (3), 647–690.

Engle R. F., Hendry D. F., Richard J. F. (1983). Exogeneity. *Econometrica*, 51 (2), 277–304.

Fragetta M., Melina G. (2013). Identification of monetary policy in SVAR models: A data-oriented perspective. *Empirical Economics*, 45 (2), 831–844.

Gelfand A. E., Dey D. K. (1994). Bayesian model choice: Asymptotics and exact calculations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 56 (3), 501–514.

Gelman A., Shalizi C. R. (2013). Philosophy and the practice of Bayesian statistics. *British Journal of Mathematical and Statistical Psychology*, 66 (1), 8–38.

Gelman A., Meng X. L., Stern H. (1996). Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica*, 6 (1), 733–760.

George E. I., Sun D., Ni S. (2008). Bayesian stochastic search for VAR model restrictions. *Journal of Econometrics*, 142 (1), 553–580.

Giudici P., Green A. P. (1999). Decomposable graphical Gaussian model determination. *Biometrika*, 86 (4), 785–801.

Hoeting J. A., Madigan D., Raftery A. E., Volinsky C. T. (1999). Bayesian model averaging: A tutorial. *Statistical science*, 14 (4), 382–401.

Hyvärinen A., Zhang K., Shimizu S., Hoyer P. O. (2010). Estimation of a structural vector autoregression model using non-Gaussianity. *Journal of Machine Learning Research*, 11 (5), 1709–1731.

Koop G., Poirier D. J., Tobias J. L. (2008). *Bayesian econometric methods*. Cambridge University Press.

Korobilis D. (2017). *Forecasting with many predictors using message passing algorithms*. <https://ssrn.com/abstract=2977838>.

Lütkepohl H., Velinov A. (2016). Structural vector autoregressions: Checking identifying long-run restrictions via heteroskedasticity. *Journal of Economic Surveys*, 30 (2), 377–392.

Madigan D., Andersson S. A., Perlman, M. D., Volinsky C. T. (1996). Bayesian model averaging and model selection for Markov equivalence classes of acyclic digraphs. *Communications in Statistics — Theory and Methods*, 25 (11), 2493–2519.

Madigan D., York J., Allard D. (1995). Bayesian graphical models for discrete data. *International Statistical Review/Revue Internationale de Statistique*, 63 (2), 215–232.

Moneta A. (2008). Graphical causal models and VARs: An empirical assessment of the real business cycles hypothesis. *Empirical Economics*, 35 (2), 275–300.

Neal R. M. (1999). Erroneous Results in marginal likelihood from the Gibbs output. <https://www.cs.toronto.edu/~radford/ftp/chib-letter.pdf>.

Neal R. M. (2001). Annealed importance sampling. *Statistics and computing*, 11 (2), 125–139.

Pearl J. (2015). Trygve Haavelmo and the emergence of causal calculus. *Econometric Theory*, 31 (01), 152–179.

Plagborg-Møller M. (2016). Bayesian inference on structural impulse response functions. https://scholar.harvard.edu/files/plagborg/files/irf_bayes.pdf.

Popper K. R. (1956). Three views concerning human knowledge. *Contemporary British Philosophy*, 387, 357–388.

Raftery A. E., Newton M. A., Satagopan J. M., Krivitsky P. N. (2007). Estimating the integrated likelihood via posterior simulation using the harmonic mean identity. *Bayesian Statistics*, 8, 1–45.

Rothenberg T. J. (1971). Identification in parametric models. *Econometrica*, 39 (3), 577–591.

Rubio-Ramirez J. F., Waggoner D. F., Zha T. (2010). Structural vector autoregressions: Theory of identification and algorithms for inference. *The Review of Economic Studies*, 77 (2), 665–696.

Shafer G. (1982). Lindley's paradox. *Journal of the American Statistical Association*, 77 (378), 325–334.

Sims C. A., Zha T. (1998). Bayesian methods for dynamic multivariate models. *International Economic Review*, 39 (4), 949–968.

Swanson N. R., Granger C. W. (1997). Impulse response functions based on a causal approach to residual orthogonalization in vector autoregressions. *Journal of the American Statistical Association*, 92 (437), 357–367.

Waggoner D. F., Zha T. (2003). A Gibbs sampler for structural vector autoregressions. *Journal of Economic Dynamics and Control*, 28 (2), 349–366.

Поступила в редакцию 28.07.2017;
принята в печать 08.11.2017.

Приложение 1

Таблица обозначений, используемых в данной работе

Обозначение	Описание	Формула
$x \sim \text{Gamma}(\alpha, \beta)$	Гамма-распределение	Плотность вероятности: $p(x) = \frac{\beta^{-\alpha}}{\Gamma(\alpha)} x^{\alpha-1} \exp\left(-\frac{x}{\beta}\right)$
$x \sim \mathcal{N}_n(\mu, \Sigma)$	Многомерное нормальное распределение вектора случайных величин x размерности n с математическим ожиданием μ и ковариационной матрицей Σ	Плотность вероятности: $p(x) = (2\pi)^{-0.5n} \det(\Sigma)^{-0.5} \times \exp(-0.5(x - \mu)' \Sigma^{-1} (x - \mu))$
$x \sim \mathcal{N}(\mu, \sigma^2)$	Одномерное нормальное распределение случайной величины x с математическим ожиданием μ и дисперсией σ^2	Плотность вероятности: $p(x) = (2\pi\sigma^2)^{-0.5} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right)$

Окончание Прил. 1

Обозначение	Описание	Формула
$A \circ B$	Позлементное умножение матриц. Каждый (i, j) -й элемент итоговой матрицы является произведением (i, j) -х элементов исходных матриц	$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \circ \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} =$ $= \begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} \\ a_{21}b_{21} & a_{22}b_{22} \end{bmatrix}$
$A_{(-i),(-j)}$	Все элементы матрицы A , кроме (i, j) -го элемента	$A_{(-i),(-j)} = \{a_{kl}\}; k, l = 1, \dots, n; l \neq k$
$p(\bullet \Omega)$	Функция плотности некоторой случайной величины $\bullet$ при условии Ω	
$\pi(x \rightarrow y \Omega)$	Proposal density: вспомогательная функция плотности y на шаге g алгоритма Метрополиса–Гастингса при условии нахождения цепи на шаге $(g-1)$ в значении x , а также при условии Ω	Для предложенного в данной работе алгоритма Метрополиса–Гастингса со случайным блужданием $\pi(x \rightarrow y) = \pi(y \rightarrow x) \propto$ $\propto \exp\{-0.5\sigma_{prop}^{-2}(x-y)^2\}$
$\alpha(x \rightarrow y \Omega)$	Вероятность перехода марковской цепи из состояния x на шаге $(g-1)$ в состояние y на шаге g при условии Ω	Для алгоритма Метрополиса–Гастингса со случайным блужданием $\alpha(x \rightarrow y \Omega) = \min\left(1, \frac{p(y \Omega)}{p(x \Omega)}\right)$
$\xrightarrow{a.s.}$	Сходимость почти наверное	
T	Число наблюдений в выборке	
n	Число эндогенных переменных	
k	Число экзогенных переменных	
A, a_j, a_{ij}	Матрица размерности $(n \times n)$ связей эндогенных переменных между собой, j -й столбец данной матрицы (соответствующий j -му уравнению), (i, j) -й элемент данной матрицы	
B, b_j, b_{ij}	Матрица размерности $(k \times n)$ связей эндогенных переменных с экзогенными, j -й столбец данной матрицы (соответствующий j -му уравнению), (i, j) -й элемент данной матрицы	
D, d_j	Диагональная матрица размерности $(n \times n)$ дисперсий структурных шоков, j -й диагональный элемент данной матрицы	
C, c_{ij}	Матрица концентрации размерности $(n+k) \times (n+k)$	$C = \begin{bmatrix} A \\ B \end{bmatrix} D \begin{bmatrix} A' & B' \end{bmatrix}$

Приложение 2

Модели, используемые в анализе симуляций

Идентифицируемость — свойство модели, состоящее в возможности однозначной оценки структурных параметров модели.

Тестируемость — свойство модели, заключающееся в возможности отличить ее от других на данных.

Модель	Каузальная диаграмма	Сеть концентрации	Идентифицируемость
1			Идентифицируема Нетестируема
2			Идентифицируема Нетестируема
3			Идентифицируема Тестируема
4			Идентифицируема Нетестируема
5			Идентифицируема Тестируема
6			Идентифицируема Нетестируема
7			Идентифицируема Тестируема
8			Идентифицируема Нетестируема

Окончание Прил. 2

Модель	Каузальная диаграмма	Сеть концентрации	Идентифицируемость
9			Идентифицируема Нетестируема
10			Идентифицируема Тестируема
11			Идентифицируема Тестируема
12			Неидентифицируема
13			Идентифицируема Нетестируема
14			Идентифицируема Нетестируема
15			Идентифицируема Тестируема

Приложение 3

Альтернативные методы оценки маргинальных функций правдоподобия

Метод гармонического среднего

Данный метод был предложен в работе (Raftery et al., 2007). Предположим, что $\theta^{(g)}$ — реализации из алгоритма Метрополиса–Гастингса, где g изменяется от 1 до G . Суть метода заключается в использовании следующего результата:

$$G \left(\sum_{g=1}^G \frac{1}{p(Y|Z, \theta^{(g)}; \mathcal{M}=s)} \right)^{-1} \xrightarrow{a.s.} p(Y|Z; \mathcal{M}=s),$$

где $p(Y|Z, \theta^{(g)}; \mathcal{M}=s)$ — известная функция правдоподобия. Алгоритм является вычислительно наиболее простым, но он, как показано в данной работе, нестабилен.

Метод Annealed importance sampling

Данный метод был предложен в работе (Neal, 2001). Алгоритм основан на одновременном сэмплировании ряда марковских цепей, индексируемых с помощью индексов $k=1, \dots, K$. Процедура выглядит следующим образом.

1. Выбрать последовательность чисел $0 < \beta_{K-1} < \dots < \beta_2 < \beta_1 < 1$, где β_k — параметр для k -й марковской цепи.
2. Параллельно сгенерировать G реализаций параметров $\theta_k^{(g)}$, $g=1, \dots, G$, для каждой из K марковских цепей алгоритмом Метрополиса–Гастингса:

2.1. $\theta_1^{(g)}$ из априорного распределения $f_1(\theta^{(g)}) = p(\theta^{(g)} | \mathcal{M}=s)$.

2.2. $\theta_k^{(g)}$ из распределения $f_k(\theta^{(g)}) = p(\theta^{(g)} | \mathcal{M}=s) (p(Y|Z, \theta^{(g)}; \mathcal{M}=s))^{1-\beta_{k-1}}$, $k=2, \dots, K$ с помощью процедуры Метрополиса–Гастингса со случайным блужданием, используя в качестве начального значения $\theta_{k-1}^{(g)}$. В качестве вспомогательного распределения используется нормальное распределение:

$\theta_k^{prop} = \theta_{k-1}^{(g)} + \varepsilon$, $\varepsilon \sim \mathcal{N}(0, \sigma_{prop}^2)$. Предложенное значение $\theta_k^{(g)} = \theta_k^{prop}$ принимается с вероятностью $\min(f_k(\theta_k^{prop})/f_k(\theta_{k-1}^{(g)}), 1)$.

- 2.3. Рассчитать два набора значений функций плотности распределения марковских цепей:

- по сэмплированным значениям параметров «текущей» марковской цепи k :

$$f_1(\theta_1^{(g)}), f_2(\theta_2^{(g)}), \dots, f_{K-2}(\theta_{K-2}^{(g)}), f_{K-1}(\theta_{K-1}^{(g)});$$

- по сэмплированным значениям параметров «предыдущей» марковской цепи $k-1$ (чем больше k , тем меньше β_k):

$$f_2(\theta_1^{(g)}), f_3(\theta_2^{(g)}), \dots, f_{K-1}(\theta_{K-2}^{(g)}), f_K(\theta_{K-1}^{(g)}),$$

где $f_k(\theta) = p(\theta | \mathcal{M}=s) p(Y|Z, \theta^{(g)}; \mathcal{M}=s)$.

3. Для каждой выборки (g) рассчитать веса в следующем виде:

$$w^{(g)} = \frac{f_2(\theta_1^{(g)})}{f_1(\theta_1^{(g)})} \cdot \frac{f_3(\theta_2^{(g)})}{f_2(\theta_2^{(g)})} \cdot \dots \cdot \frac{f_{K-1}(\theta_{K-2}^{(g)})}{f_{K-2}(\theta_{K-2}^{(g)})} \cdot \frac{f_K(\theta_{K-1}^{(g)})}{f_{K-1}(\theta_{K-1}^{(g)})},$$

где используются полученные в пункте 2.3 значения функций плотности распределения.

4. Рассчитать оценку маргинальной функции правдоподобия в следующем виде:

$$\hat{p}(Y|Z; \mathcal{M}=s) = G^{-1} \sum_{g=1}^G w^{(g)}.$$

Arefiev N., Khabibullin R. Bayesian identification of structural vector autoregression models. *Applied Econometrics*, 2018, v. 49, pp. 115–142.

Nikolay Arefiev

National Research University Higher School of Economics (NRU HSE), Moscow, Russian Federation;
n.arefiev@gmail.com

Ramis Khabibullin

National Research University Higher School of Economics (NRU HSE), Moscow, Russian Federation;
rawirtschaft@gmail.com

Bayesian identification of structural vector autoregression models

We propose a new method of Bayesian identification of a structural vector autoregression based on the Bayesian model averaging. As compared to the literature on Bayesian SVAR averaging, the proposed algorithm can identify not only recursive, but also cyclical models given that some conditions specified in the paper hold. Bayesian model selection is made within the set of distinguishable on data models. We use simulations to assess the performance of the algorithm. We also check sensitivity of the proposed algorithm with respect to true parameter values, number of observations, and with respect to the parameters of prior distribution.

Keywords: SVAR; identification; Bayesian model averaging; Bayesian model selection.

JEL classification: C11; C32; C52.

References

Ahelegbey D. F., Billio M., Casarin R. (2016). Bayesian graphical models for structural vector autoregressive processes. *Journal of Applied Econometrics*, 31 (2), 357–386.

Aldrich E. M., Fernández-Villaverde J., Gallant A. R., Rubio-Ramírez J. F. (2011). Tapping the supercomputer under your desk: Solving dynamic equilibrium models with graphics processors. *Journal of Economic Dynamics and Control*, 35 (3), 386–393.

Arefiev N. (2014). Structural models with testable identification. *Higher School of Economics Research Paper No. WP BRP 79/EC/2014*.

- Arias J., Rubio-Ramirez J. F., Waggoner D. F. (2014). Inference based on SVAR identified with sign and zero restrictions: Theory and applications. *International Finance Discussion Papers Number 1100*.
- Baumeister C., Hamilton J. D. (2015). Sign restrictions, structural vector autoregressions, and useful prior information. *Econometrica*, 83 (5), 1963–1999.
- Bishop C. M. (2006). *Pattern recognition and machine learning*. Springer.
- Chib S. (1995). Marginal likelihood from the Gibbs output. *Journal of the American Statistical Association*, 90 (432), 1313–1321.
- Chib S., Greenberg E. (1995). Understanding the Metropolis–Hastings algorithm. *The American statistician*, 49 (4), 327–335.
- Chib S., Jeliazkov I. (2001). Marginal likelihood from the Metropolis–Hastings output. *Journal of the American Statistical Association*, 96 (453), 270–281.
- Clarke J. L., Clarke B., Yu C.-W. (2013). Prediction in M-complete problems with limited sample size. *Bayesian Analysis*, 8 (3), 647–690.
- Engle R. F., Hendry D. F., Richard J. F. (1983). Exogeneity. *Econometrica*, 51 (2), 277–304.
- Fragetta M., Melina G. (2013). Identification of monetary policy in SVAR models: A data-oriented perspective. *Empirical Economics*, 45 (2), 831–844.
- Gelfand A. E., Dey D. K. (1994). Bayesian model choice: Asymptotics and exact calculations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 56 (3), 501–514.
- Gelman A., Shalizi C. R. (2013). Philosophy and the practice of Bayesian statistics. *British Journal of Mathematical and Statistical Psychology*, 66 (1), 8–38.
- Gelman A., Meng X. L., Stern H. (1996). Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica*, 6 (1), 733–760.
- George E. I., Sun D., Ni S. (2008). Bayesian stochastic search for VAR model restrictions. *Journal of Econometrics*, 142 (1), 553–580.
- Giudici P., Green A. P. (1999). Decomposable graphical Gaussian model determination. *Biometrika*, 86 (4), 785–801.
- Hoeting J. A., Madigan D., Raftery A. E., Volinsky C. T. (1999). Bayesian model averaging: A tutorial. *Statistical science*, 14 (4), 382–401.
- Hyvärinen A., Zhang K., Shimizu S., Hoyer P. O. (2010). Estimation of a structural vector autoregression model using non-Gaussianity. *Journal of Machine Learning Research*, 11 (5), 1709–1731.
- Koop G., Poirier D. J., Tobias J. L. (2008). *Bayesian econometric methods*. Cambridge University Press.
- Korobilis D. (2017). *Forecasting with many predictors using message passing algorithms*. <https://ssrn.com/abstract=2977838>.
- Lütkepohl H., Velinov A. (2016). Structural vector autoregressions: Checking identifying long-run restrictions via heteroskedasticity. *Journal of Economic Surveys*, 30 (2), 377–392.
- Madigan D., Andersson S. A., Perlman, M. D., Volinsky C. T. (1996). Bayesian model averaging and model selection for Markov equivalence classes of acyclic digraphs. *Communications in Statistics — Theory and Methods*, 25 (11), 2493–2519.
- Madigan D., York J., Allard D. (1995). Bayesian graphical models for discrete data. *International Statistical Review/Revue Internationale de Statistique*, 63 (2), 215–232.
- Moneta A. (2008). Graphical causal models and VARs: An empirical assessment of the real business cycles hypothesis. *Empirical Economics*, 35 (2), 275–300.

- Neal R. M. (1999). Erroneous Results in marginal likelihood from the Gibbs output. <https://www.cs.toronto.edu/~radford/ftp/chib-letter.pdf>.
- Neal R. M. (2001). Annealed importance sampling. *Statistics and computing*, 11 (2), 125–139.
- Pearl J. (2015). Trygve Haavelmo and the emergence of causal calculus. *Econometric Theory*, 31 (01), 152–179.
- Plagborg-Møller M. (2016). Bayesian inference on structural impulse response functions. https://scholar.harvard.edu/files/plagborg/files/irf_bayes.pdf.
- Popper K. R. (1956). Three views concerning human knowledge. *Contemporary British Philosophy*, 387, 357–388.
- Raftery A. E., Newton M. A., Satagopan J. M., Krivitsky P. N. (2007). Estimating the integrated likelihood via posterior simulation using the harmonic mean identity. *Bayesian Statistics*, 8, 1–45.
- Rothenberg T. J. (1971). Identification in parametric models. *Econometrica*, 39 (3), 577–591.
- Rubio-Ramirez J. F., Waggoner D. F., Zha T. (2010). Structural vector autoregressions: Theory of identification and algorithms for inference. *The Review of Economic Studies*, 77 (2), 665–696.
- Shafer G. (1982). Lindley's paradox. *Journal of the American Statistical Association*, 77 (378), 325–334.
- Sims C. A., Zha T. (1998). Bayesian methods for dynamic multivariate models. *International Economic Review*, 39 (4), 949–968.
- Swanson N. R., Granger C. W. (1997). Impulse response functions based on a causal approach to residual orthogonalization in vector autoregressions. *Journal of the American Statistical Association*, 92 (437), 357–367.
- Waggoner D. F., Zha T. (2003). A Gibbs sampler for structural vector autoregressions. *Journal of Economic Dynamics and Control*, 28 (2), 349–366.

Received 28.07.2017; accepted 08.11.2017.