Основания Байесовских SVAR-моделей

Хабибуллин Рамис Арсланович

17 июня 2016 г.

Внимание! Данный файл является просто конспектом материала книг, статей и учебников с пояснениями и создан исключительно для лучшего понимания предмета. Все представленные ниже скрипты, написанные автором в программных пакетах `R` и `matlab`, доступны по ссылке: git link

1 Введение про байесовскую идентификацию.

Классическое решение проблемы идентификации заключается в использовании строгих априорных предположений о структуре модели на основании экономической теории. Однако, зачастую теория не даёт достаточного количества априорных ограничений, либо может существовать ряд противоречащих друг другу теорий, на основании которых невозможно построить единственную каузальную структуру экономики.

В качестве альтернативного подхода выступают методы данно-ориентированной идентификации, представленной работами (**???**), позволяющие использовать информацию, содержащуюся в данных для идентификации модели.

Первый подход использует только жёсткую априорную информацию о структуре модели. Второй подход использует только информмацию, содержащуюся в данных. В рамках данной работы была сделана попытка объединить оба подхода с помощью Байесовских методов. Сделано это с помощью методологии, реализованной в литературе по модельной неопределённости и Байесовском осреднении моделей.

Основной идеей Байесовской идентификации является не выбор единственной модели, а нахождение апостериорных вероятностей возможных моделей при условии данных $p(\mathcal{M}_s|\text{data})$. Это даёт два преимущества перед частотным методом данно-ориентированной идентификации (**?**).

Первое преимущество заключается в минимизации риска использования оценок различных моделей. С точки зрения политика, контролирующего некоторый инструмент политики, принципиальным является вопрос, как изменение инструмента политики повлияет на экономические переменные, такие как выпуск y и инфляция π. Представим центральный банк с квадратичной функцией потерь $L = (y-y^*)^2+(\pi-\pi^*)^2$, имеющего в распоряжении два инструмента политики: краткосрочную ставку процента R_{SR} (традиционная монетарная политика) и долгосрочную ставку процента R_{LR}(нетрадиционная монетарная политика). Для того, чтобы принять решение о выборе инструмента политики, он нанимает двух экспертов $\mathcal{E}_{\text{prior}}$ и $\mathcal{E}_{\text{data}}$. С целью расчёта эффекта шока инструмента политики каждый из них строит структурную векторную авторегрессию, одно из уравнений которой интерпретируется как совокупный спрос и отражает каузальное влияние долгосрочной и краткосрочной ставок процента на текущее значение выпуска.

Эксперт $\mathcal{E}_{\text{prior}}$ идентифицирует модель с помощью априорных предположений относительно каузальной структуры модели, причём он а приори предполагает, что экономика находится в состоянии ловушки ликвидности, и канал влияния ставки процента на выпуск перекрыт (коэффициент перед R_{SR} в уравнении совокупного спроса ограничен нулевым значением). Совокупность ограничений эксперта $\mathcal{E}_{\text{prior}}$ образуют модель $\mathcal{M}_{prior}$.

Эксперт $\mathcal{E}_{\text{data}}$ для решения своей задачи использует процедуру тестируемой идентификации. Представим, что на основании данного подхода он сформировал набор ограничений $\mathcal{M}_{data}$, одним из которых является ограничение на долгосрочную ставку процента в структурном уравнении совокупного спроса. При этом краткосрочную ставку процента в этом же уравнении он не ограничивает.

В итоге центральный банк получит две конкурирующих друг с другом модели $\mathcal{M}_{data}$ и $\mathcal{M}_{prior}$, которые дают противоречивые рекомендации. Проблема с моделью $\mathcal{M}_{\text{prior}}$ заключается в том, что она получена априорными предположениями, которые могут оказаться неверными. Проблема же с моделью $\mathcal{M}_{\text{data}}$ заключается в том, что она получена статистическим тестированием гипотез, и поэтому существует ненулевая вероятность того, что и эта модель неверна.

Таким образом, центральный банк рискует ошибиться при выборе политики, если доверится только одной из двух моделей, и поэтому для принятия решения ему нужно каким-то образом сравнить, насколько вероятна та или иная модель.

Байесовские методы позволяют примирить оба подхода и адекватно оценить вероятности моделей при условии данных, используя априорную информацию о вероятностях моделей и информацию, доступную из данных. Политик, зная отклик экономики на шоки при каждой из моделей и зная вероятности каждой из моделей, может сделать выбор на основании взвешенной по этим вероятностям функции потерь. В литературе, посвящённой робастному контролю, известным фактом является то, что в такой ситуации эффективным решением центрального банка будет выбор политики на основании взвешенной по вероятностям моделей при условии данных функции потерь.

Более того, Байесовский подход к идентификации позволяет решить и другую проблему. Метод тестируемой идентификации, предложенный (**?**), основан на процедуре последовательного тестирования гипотез, что может привести к проблемам построения доверительных интервалов для оценок параметров и функций импульсного отклика модели. Это один из основных выводов эконометрических работ, посвящённых pretest-оценкам. Однако, как правило, политика при выборе решения интересует, значим ли эффект от воздействия конкретного инструмента политики, что невозможно сделать без меры точности оценки. Однако, если доверительные интервалы построены неверно, может быть сделан неверный вывод о значимости эффекта политики. Более того, при условии выбора корректных априорных значений, доверительные интервалы оценённых Байесовскими методами параметров, как правило, более узкие, чем оценённые частотными методами, например, с помощью bootstrap репликаций (**?**). При этом как показано в литературе по Байесовским методам (например, в книге **?**) сами интервалы, построенные на основании Байесовских методов, имеют более корректную интерпретацию, нежели интервалы, оценённые частотными методами.

В данной работе собраны основные факты и утверждения с подробным доказательством, необходимые для анализа и интерпретации Байесовских SVAR-моделей.

2 Теоремы о распределениях

Утверждение 1. ***О совместном нормальном распределении.*** *Предположим, что* $\mathbf{x} = [\mathbf{y}', \mathbf{z}']'$, *где* $\mathbf{y}$ *— вектор размерности* $n \times 1$, *а* $\mathbf{z}$ *— вектор размерности* $k \times 1$. *При этом* $\mathbf{x} \sim \mathcal{N}_{kn}(\boldsymbol{\mu}_{\mathbf{X}}, \boldsymbol{\Sigma}_{\mathbf{xx}})$. *Пусть при этом можно сделать следующее разбиение параметров:*

$$\boldsymbol{\mu}_{\mathbf{x}} = \begin{bmatrix} \boldsymbol{\mu}_{\mathbf{y}} \\ \boldsymbol{\mu}_{\mathbf{z}} \end{bmatrix} \qquad \boldsymbol{\Sigma}_{\mathbf{xx}} = \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{yy}} & \boldsymbol{\Sigma}_{\mathbf{yz}} \\ \boldsymbol{\Sigma}_{\mathbf{zy}} & \boldsymbol{\Sigma}_{\mathbf{zz}} \end{bmatrix} \tag{1}$$

Следует помнить, что для любых $\mathbf{i}$ *и* $\mathbf{j} \neq \mathbf{i}$ *переменных* $\boldsymbol{\Sigma}_{\mathbf{ii}} = \boldsymbol{\Sigma}'_{\mathbf{ii}}$, *а* $\boldsymbol{\Sigma}'_{\mathbf{ij}} = \boldsymbol{\Sigma}_{\mathbf{ji}}$

Тогда маргинальные и условные случайные величины имеют следующие распределения:

$$\mathbf{y} \sim \mathcal{N}_n(\boldsymbol{\mu}_{\mathbf{y}}, \boldsymbol{\Sigma}_{\mathbf{yy}}) \tag{2}$$
$$\mathbf{z} \sim \mathcal{N}_k(\boldsymbol{\mu}_{\mathbf{z}}, \boldsymbol{\Sigma}_{\mathbf{zz}}) \tag{3}$$
$$\mathbf{y}|\mathbf{z} \sim \mathcal{N}_n(\boldsymbol{\mu}_{\mathbf{y}|\mathbf{z}} = \boldsymbol{\mu}_{\mathbf{y}} + \boldsymbol{\Sigma}_{\mathbf{yz}}\boldsymbol{\Sigma}_{\mathbf{zz}}^{-1}(\mathbf{z}-\boldsymbol{\mu}_{\mathbf{z}}), \boldsymbol{\Sigma}_{\mathbf{yy}|\mathbf{z}} = \boldsymbol{\Sigma}_{\mathbf{yy}} - \boldsymbol{\Sigma}_{\mathbf{yz}}\boldsymbol{\Sigma}_{\mathbf{zz}}^{-1}\boldsymbol{\Sigma}_{\mathbf{zy}}) \tag{4}$$
$$\mathbf{z}|\mathbf{y} \sim \mathcal{N}_k(\boldsymbol{\mu}_{\mathbf{z}|\mathbf{y}} = \boldsymbol{\mu}_{\mathbf{z}} + \boldsymbol{\Sigma}_{\mathbf{zy}}\boldsymbol{\Sigma}_{\mathbf{yy}}^{-1}(\mathbf{y}-\boldsymbol{\mu}_{\mathbf{y}}), \boldsymbol{\Sigma}_{\mathbf{zz}|\mathbf{y}} = \boldsymbol{\Sigma}_{\mathbf{zz}} - \boldsymbol{\Sigma}_{\mathbf{zy}}\boldsymbol{\Sigma}_{\mathbf{yy}}^{-1}\boldsymbol{\Sigma}_{\mathbf{yz}}) \tag{5}$$
$$\tag{6}$$

Доказательство. Совместная функция многомерного нормального распределения $\mathbf{x}$ выглядит следующим образом:

$$\begin{aligned} p(\mathbf{x}) = p(\mathbf{y},\mathbf{z}) = (2\pi)^{-\frac{nk}{2}}|\det(\boldsymbol{\Sigma}_{\mathbf{xx}})|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{x}})'\boldsymbol{\Sigma}_{\mathbf{xx}}^{-1}(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{x}})\right) = \\ = (2\pi)^{-\frac{nk}{2}}|\det(\boldsymbol{\Sigma}_{\mathbf{xx}})|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\left[(\mathbf{y}-\boldsymbol{\mu}_{\mathbf{y}})', (\mathbf{z}-\boldsymbol{\mu}_{\mathbf{z}})'\right]\begin{bmatrix}\boldsymbol{\Sigma}_{\mathbf{yy}} & \boldsymbol{\Sigma}_{\mathbf{yz}} \\ \boldsymbol{\Sigma}_{\mathbf{zy}} & \boldsymbol{\Sigma}_{\mathbf{zz}}\end{bmatrix}^{-1}\begin{bmatrix}(\mathbf{y}-\boldsymbol{\mu}_{\mathbf{y}}) \\ (\mathbf{z}-\boldsymbol{\mu}_{\mathbf{z}})\end{bmatrix}\right) \end{aligned} \tag{7}$$

Определим $\mathbf{S}(\mathbf{x})$ как:

$$\mathbf{S}(\mathbf{x}) = \left[(\mathbf{y}-\boldsymbol{\mu}_{\mathbf{y}})', (\mathbf{z}-\boldsymbol{\mu}_{\mathbf{z}})'\right]\begin{bmatrix}\boldsymbol{\Sigma}_{\mathbf{yy}} & \boldsymbol{\Sigma}_{\mathbf{yz}} \\ \boldsymbol{\Sigma}_{\mathbf{zy}} & \boldsymbol{\Sigma}_{\mathbf{zz}}\end{bmatrix}^{-1}\begin{bmatrix}(\mathbf{y}-\boldsymbol{\mu}_{\mathbf{y}}) \\ (\mathbf{z}-\boldsymbol{\mu}_{\mathbf{z}})\end{bmatrix} \tag{8}$$

Матрица концентрации может быть расписана по правилам нахождения обратной матрицы в следующей форме:

$$\mathbf{C}_{\mathbf{xx}} = \boldsymbol{\Sigma}_{\mathbf{xx}}^{-1} = \begin{bmatrix}\mathbf{C}_{\mathbf{yy}} & \mathbf{C}_{\mathbf{yz}} \\ \mathbf{C}_{\mathbf{zy}} & \mathbf{C}_{\mathbf{zz}}\end{bmatrix} = \begin{bmatrix}(\boldsymbol{\Sigma}_{\mathbf{yy}} - \boldsymbol{\Sigma}_{\mathbf{yz}}\boldsymbol{\Sigma}_{\mathbf{zz}}^{-1}\boldsymbol{\Sigma}_{\mathbf{zy}})^{-1} & -(\boldsymbol{\Sigma}_{\mathbf{yy}} - \boldsymbol{\Sigma}_{\mathbf{yz}}\boldsymbol{\Sigma}_{\mathbf{zz}}^{-1}\boldsymbol{\Sigma}_{\mathbf{zy}})^{-1}\boldsymbol{\Sigma}_{\mathbf{yz}}\boldsymbol{\Sigma}_{\mathbf{zz}}^{-1} \\ -(\boldsymbol{\Sigma}_{\mathbf{zz}} - \boldsymbol{\Sigma}_{\mathbf{zy}}\boldsymbol{\Sigma}_{\mathbf{yy}}^{-1}\boldsymbol{\Sigma}_{\mathbf{yz}})^{-1}\boldsymbol{\Sigma}_{\mathbf{zy}}\boldsymbol{\Sigma}_{\mathbf{yy}}^{-1} & (\boldsymbol{\Sigma}_{\mathbf{zz}} - \boldsymbol{\Sigma}_{\mathbf{zy}}\boldsymbol{\Sigma}_{\mathbf{yy}}^{-1}\boldsymbol{\Sigma}_{\mathbf{yz}})^{-1}\end{bmatrix} \tag{9}$$

Более того, очевидно, что:

$$\mathbf{C}_{\mathbf{yz}} = -\boldsymbol{\Sigma}_{\mathbf{yy}}^{-1}\boldsymbol{\Sigma}_{\mathbf{yz}}\mathbf{C}_{\mathbf{zz}} = -\mathbf{C}_{\mathbf{yy}}\boldsymbol{\Sigma}_{\mathbf{zy}}\boldsymbol{\Sigma}_{\mathbf{zz}}^{-1} = \mathbf{C}_{\mathbf{zy}}' \tag{10}$$

А по правилам раскрытия обратной от суммы матриц можно получить:

$$\mathbf{C}_{\mathbf{zz}} = (\boldsymbol{\Sigma}_{\mathbf{zz}} - \boldsymbol{\Sigma}_{\mathbf{zy}}\boldsymbol{\Sigma}_{\mathbf{yy}}^{-1}\boldsymbol{\Sigma}_{\mathbf{yz}})^{-1} = \boldsymbol{\Sigma}_{\mathbf{zz}}^{-1} + \boldsymbol{\Sigma}_{\mathbf{zz}}^{-1}\boldsymbol{\Sigma}_{\mathbf{zy}}\mathbf{C}_{\mathbf{yy}}\boldsymbol{\Sigma}_{\mathbf{yz}}\boldsymbol{\Sigma}_{\mathbf{zz}}^{-1} \tag{11}$$
$$\mathbf{C}_{\mathbf{yy}} = (\boldsymbol{\Sigma}_{\mathbf{yy}} - \boldsymbol{\Sigma}_{\mathbf{yz}}\boldsymbol{\Sigma}_{\mathbf{zz}}^{-1}\boldsymbol{\Sigma}_{\mathbf{zy}})^{-1} = \boldsymbol{\Sigma}_{\mathbf{yy}}^{-1} + \boldsymbol{\Sigma}_{\mathbf{yy}}^{-1}\boldsymbol{\Sigma}_{\mathbf{yz}}\mathbf{C}_{\mathbf{zz}}\boldsymbol{\Sigma}_{\mathbf{zy}}\boldsymbol{\Sigma}_{\mathbf{yy}}^{-1} \tag{12}$$

Таким образом, можно расписать

$$\mathbf{S}(\mathbf{x}) = (\mathbf{y}-\boldsymbol{\mu}_{\mathbf{y}})'\mathbf{C}_{\mathbf{yy}}(\mathbf{y}-\boldsymbol{\mu}_{\mathbf{y}}) + 2(\mathbf{y}-\boldsymbol{\mu}_{\mathbf{y}})'\mathbf{C}_{\mathbf{yz}}(\mathbf{z}-\boldsymbol{\mu}_{\mathbf{z}}) + (\mathbf{z}-\boldsymbol{\mu}_{\mathbf{z}})'\mathbf{C}_{\mathbf{zz}}(\mathbf{z}-\boldsymbol{\mu}_{\mathbf{z}}) \tag{13}$$

Тогда в терминах $\mathbf{C}_{\mathbf{zz}}$ можно вывести следующее соотношение:

$$\begin{aligned} \mathbf{S}(\mathbf{x}) = (\mathbf{y}-\boldsymbol{\mu}_{\mathbf{y}})'(\boldsymbol{\Sigma}_{\mathbf{yy}}^{-1} + \boldsymbol{\Sigma}_{\mathbf{yy}}^{-1}\boldsymbol{\Sigma}_{\mathbf{yz}}\mathbf{C}_{\mathbf{zz}}\boldsymbol{\Sigma}_{\mathbf{zy}}\boldsymbol{\Sigma}_{\mathbf{yy}}^{-1})(\mathbf{y}-\boldsymbol{\mu}_{\mathbf{y}}) - \\ - 2(\mathbf{y}-\boldsymbol{\mu}_{\mathbf{y}})'(\boldsymbol{\Sigma}_{\mathbf{yy}}^{-1}\boldsymbol{\Sigma}_{\mathbf{yz}}\mathbf{C}_{\mathbf{zz}})(\mathbf{z}-\boldsymbol{\mu}_{\mathbf{z}}) + (\mathbf{z}-\boldsymbol{\mu}_{\mathbf{z}})'\mathbf{C}_{\mathbf{zz}}(\mathbf{z}-\boldsymbol{\mu}_{\mathbf{z}}) \end{aligned} \tag{14}$$

Обозначим

$$\mathbf{b} = \boldsymbol{\Sigma}_{\mathbf{zy}}\boldsymbol{\Sigma}_{\mathbf{yy}}^{-1}(\mathbf{y}-\boldsymbol{\mu}_{\mathbf{y}}) \tag{15}$$

$$\begin{aligned} (\mathbf{y}-\boldsymbol{\mu}_{\mathbf{y}})'\boldsymbol{\Sigma}_{\mathbf{yy}}^{-1}(\mathbf{y}-\boldsymbol{\mu}_{\mathbf{y}}) + (\mathbf{z}-\boldsymbol{\mu}_{\mathbf{z}})'\mathbf{C}_{\mathbf{zz}}(\mathbf{z}-\boldsymbol{\mu}_{\mathbf{z}}) + \mathbf{b}'\mathbf{C}_{\mathbf{zz}}\mathbf{b} - 2\mathbf{b}'\mathbf{C}_{\mathbf{zz}}(\mathbf{z}-\boldsymbol{\mu}_{\mathbf{z}}) = \\ = (\mathbf{y}-\boldsymbol{\mu}_{\mathbf{y}})'\boldsymbol{\Sigma}_{\mathbf{yy}}^{-1}(\mathbf{y}-\boldsymbol{\mu}_{\mathbf{y}}) + \left[\mathbf{z} - (\boldsymbol{\mu}_{\mathbf{z}} + \mathbf{b})\right]'\mathbf{C}_{\mathbf{zz}}\left[\mathbf{z} - (\boldsymbol{\mu}_{\mathbf{z}} + \mathbf{b})\right] \end{aligned} \tag{16}$$

В силу утверждения, верно следующее:

$$|\det(\mathbf{\Sigma_{xx}})| = |\det(\mathbf{\Sigma_{yy}})||\det(\mathbf{\Sigma_{zz}} - \mathbf{\Sigma_{zy}}\mathbf{\Sigma_{yy}^{-1}}\mathbf{\Sigma_{yz}})| = |\det(\mathbf{\Sigma_{yy}})||\det(\mathbf{C_{zz}})| \tag{17}$$

Следовательно, совместная функция распределения выглядит следующим образом:

$$p(\mathbf{x}) = (2\pi)^{-\frac{nk}{2}}|\det(\mathbf{\Sigma_{yy}})|^{-\frac{1}{2}}|\det(\mathbf{C_{zz}})|^{-\frac{1}{2}}\times \\ \times \exp\left\{-\frac{1}{2}\left((\mathbf{y}-\boldsymbol{\mu}_\mathbf{y})'\mathbf{\Sigma_{yy}^{-1}}(\mathbf{y}-\boldsymbol{\mu}_\mathbf{y}) + [\mathbf{z}-(\boldsymbol{\mu}_\mathbf{z}+\mathbf{b})]'\,\mathbf{C_{zz}}\,[\mathbf{z}-(\boldsymbol{\mu}_\mathbf{z}+\mathbf{b})]\right)\right\} \tag{18}$$

Для того, чтобы найти маргинальную функцию правдоподобия величины $\mathbf{y}$, необходимо найти следующий интеграл:

$$p(\mathbf{y}) = \int_\mathbf{z} p(\mathbf{x})\,\mathrm{d}\mathbf{z} = (2\pi)^{-\frac{n}{2}}|\det(\mathbf{\Sigma_{yy}})|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu}_\mathbf{y})'\mathbf{\Sigma_{yy}^{-1}}(\mathbf{y}-\boldsymbol{\mu}_\mathbf{y})\right\}\times \\ \times\int_\mathbf{z}(2\pi)^{-\frac{k}{2}}|\det(\mathbf{C_{zz}})|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\left([\mathbf{z}-(\boldsymbol{\mu}_\mathbf{z}+\mathbf{b})]'\,\mathbf{C_{zz}}\,[\mathbf{z}-(\boldsymbol{\mu}_\mathbf{z}+\mathbf{b})]\right)\right\}\mathrm{d}\mathbf{z} = \\ = (2\pi)^{-\frac{n}{2}}|\det(\mathbf{\Sigma_{yy}})|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu}_\mathbf{y})'\mathbf{\Sigma_{yy}^{-1}}(\mathbf{y}-\boldsymbol{\mu}_\mathbf{y})\right\}, \tag{19}$$

что есть плотность нормального распределения с параметрами математического ожидания $\boldsymbol{\mu}_\mathbf{y}$ и дисперсией $\mathbf{\Sigma_{yy}}$.

Для того, чтобы получить условную функцию распределения $\mathbf{z}|\mathbf{y}$, необходимо провести следующую операцию:

$$p(\mathbf{z}|\mathbf{y}) = \frac{p(\mathbf{x})}{p(\mathbf{y})} = (2\pi)^{-\frac{k}{2}}|\det(\mathbf{C_{zz}})|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\left([\mathbf{z}-(\boldsymbol{\mu}_\mathbf{z}+\mathbf{b})]'\,\mathbf{C_{zz}}\,[\mathbf{z}-(\boldsymbol{\mu}_\mathbf{z}+\mathbf{b})]\right)\right\}, \tag{20}$$

что является плотностью нормального распределения с параметром математического ожидания $\boldsymbol{\mu}_\mathbf{z} + \mathbf{\Sigma_{zy}}\mathbf{\Sigma_{yy}^{-1}}(\mathbf{y}-\boldsymbol{\mu}_\mathbf{y})$ и параметром дисперсии $\mathbf{\Sigma_{zz}} - \mathbf{\Sigma_{zy}}\mathbf{\Sigma_{yy}^{-1}}\mathbf{\Sigma_{yz}}$.

Выводы для распределений $\mathbf{z}$ и $\mathbf{y}|\mathbf{z}$ производятся аналогично, если расписать совместную плотность в терминах матрицы $\mathbf{C_{yy}}$. □

Утверждение 2. ***Об условных и маргинальных нормальных распределениях.*** *Предположим, что $\mathbf{z} \sim \mathcal{N}_k(\boldsymbol{\mu}_\mathbf{z}, \mathbf{\Sigma_{zz}})$, а условное распределение $\mathbf{y}|\mathbf{z}$ может быть описано линейной моделью: $\mathbf{y}|\mathbf{z} \sim \mathcal{N}_n(\mathbf{a}+\mathbf{Bz}, \mathbf{\Sigma_{yy|z}})$ [1],где $\mathbf{a}$ — вектор размерности $n\times 1$, а $\mathbf{B}$ — матрица гиперпараметров размерности $n \times k$. Другими словами:*

$$\mathbf{y} = \mathbf{a} + \mathbf{Bz} + \boldsymbol{\varepsilon} \tag{21}$$

$$\mathbf{z} = \boldsymbol{\mu}_\mathbf{z} + \boldsymbol{\xi}, \tag{22}$$

где $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma_{yy|z}})$, $\boldsymbol{\xi} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma_{zz}})$, $\mathrm{cov}(\boldsymbol{\varepsilon}, \boldsymbol{\xi}) = \mathbb{E}(\boldsymbol{\varepsilon}\boldsymbol{\xi}') = \mathbf{0}$.

В этом случае совместная $\mathbf{x} = [\mathbf{y}', \mathbf{z}']'$ и маргинальная $\mathbf{y}$ случайные величины распределены нормально со следующими гиперпараметрами:

$$\mathbf{z} \sim \mathcal{N}(\boldsymbol{\mu}_\mathbf{z}, \mathbf{\Sigma_{zz}}) \tag{23}$$

$$\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}_\mathbf{x}, \mathbf{\Sigma_{xx}}) \qquad \boldsymbol{\mu}_\mathbf{x} = \begin{bmatrix} \mathbf{a} + \mathbf{B}\boldsymbol{\mu}_\mathbf{z} \\ \boldsymbol{\mu}_\mathbf{z} \end{bmatrix} \qquad \mathbf{\Sigma_{xx}} = \begin{bmatrix} \mathbf{B}\mathbf{\Sigma_{zz}}\mathbf{B}' + \mathbf{\Sigma_{yy|z}} & \mathbf{B}\mathbf{\Sigma_{zz}} \\ \mathbf{\Sigma_{zz}}\mathbf{B}' & \mathbf{\Sigma_{zz}} \end{bmatrix} \tag{24}$$

Если при этом $\mathbf{B} = \mathbf{\Sigma_{zy}}\mathbf{\Sigma_{zz}^{-1}}$, то:

$$\boldsymbol{\mu}_\mathbf{x} = \begin{bmatrix} \mathbf{a} + \mathbf{\Sigma_{zy}}\mathbf{\Sigma_{zz}^{-1}}\boldsymbol{\mu}_\mathbf{z} \\ \boldsymbol{\mu}_\mathbf{z} \end{bmatrix} \qquad \mathbf{\Sigma_{xx}} = \begin{bmatrix} \mathbf{\Sigma_{yz}}\mathbf{\Sigma_{zz}^{-1}}\mathbf{\Sigma_{zy}} + \mathbf{\Sigma_{yy|z}} & \mathbf{\Sigma_{yz}} \\ \mathbf{\Sigma_{zy}} & \mathbf{\Sigma_{zz}} \end{bmatrix} \tag{25}$$

[1] Например, стандартная линейная модель регрессии $\mathbf{y} = \mathbf{Z}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ может быть записана в векторной относительно $\mathbf{Z}$ форме: $\mathbf{y} = (\boldsymbol{\beta}' \otimes \mathbf{I}_T)\,\mathrm{vec}(\mathbf{Z}) + \boldsymbol{\varepsilon}$.

Доказательство.

$$\boldsymbol{\mu}_{\mathbf{y}} = \mathbb{E}_{\mathbf{y}}\boldsymbol{\mu}_{\mathbf{y}|\mathbf{z}} = \mathbb{E}_{\mathbf{y}}(\mathbf{y}) = \mathbf{a} + \mathbf{B}\boldsymbol{\mu}_{\mathbf{z}} \tag{26}$$

$$\boldsymbol{\Sigma}_{\mathbf{yy}} = \mathrm{var}(\mathbf{y}) = \mathbf{B}\boldsymbol{\Sigma}_{\mathbf{zz}}\mathbf{B}' + \boldsymbol{\Sigma}_{\mathbf{yy}|\mathbf{z}} \tag{27}$$

Таким образом, $\mathbf{y} \sim \mathcal{N}_n(\mathbf{a} + \mathbf{B}\boldsymbol{\mu}_{\mathbf{z}}, \mathbf{B}\boldsymbol{\Sigma}_{\mathbf{zz}}\mathbf{B}' + \boldsymbol{\Sigma}_{\mathbf{yy}|\mathbf{z}})$.

Для того, чтобы выразить совместную функцию плотности, следует найти математическое ожидание и ковариационную матрицу вектора $\mathbf{x} = [\mathbf{y}', \mathbf{z}']'$. Очевидно, что $\mathbb{E}_{\mathbf{x}} = [\boldsymbol{\mu}'_{\mathbf{y}}, \boldsymbol{\mu}_{\mathbf{z}}]'$. Ковариационную матрицу же можно найти следующим образом:

$$\begin{aligned}\boldsymbol{\Sigma}_{\mathbf{xx}} = \mathrm{var}\left(\begin{bmatrix}\mathbf{y}\\ \mathbf{z}\end{bmatrix}\right) = \mathrm{var}\left(\begin{bmatrix}\mathbf{a}\\ \boldsymbol{\mu}_{\mathbf{z}}\end{bmatrix} + \begin{bmatrix}\mathbf{Bz}+\boldsymbol{\varepsilon}\\ \boldsymbol{\xi}\end{bmatrix}\right) = \mathrm{var}\left(\begin{bmatrix}\mathbf{Bz}+\boldsymbol{\varepsilon}\\ \boldsymbol{\xi}\end{bmatrix}\right) = \\ = \mathbb{E}\left(\begin{bmatrix}\mathbf{Bz}+\boldsymbol{\varepsilon}\\ \boldsymbol{\xi}\end{bmatrix}\begin{bmatrix}\mathbf{z}'\mathbf{B}'+\boldsymbol{\varepsilon}' & \boldsymbol{\xi}'\end{bmatrix}\right) - \left(\mathbb{E}\begin{bmatrix}\mathbf{Bz}+\boldsymbol{\varepsilon}\\ \boldsymbol{\xi}\end{bmatrix}\right)\left(\mathbb{E}\begin{bmatrix}\mathbf{Bz}+\boldsymbol{\varepsilon}\\ \boldsymbol{\xi}\end{bmatrix}\right)'\end{aligned} \tag{28}$$

Распишем каждый член разности отдельно:

$$\begin{aligned}\mathbb{E}\left(\begin{bmatrix}\mathbf{Bz}+\boldsymbol{\varepsilon}\\ \boldsymbol{\xi}\end{bmatrix}\begin{bmatrix}\mathbf{z}'\mathbf{B}'+\boldsymbol{\varepsilon}' & \boldsymbol{\xi}'\end{bmatrix}\right) = \mathbb{E}\begin{bmatrix}\mathbf{Bzz}'\mathbf{B}' + \boldsymbol{\varepsilon}\mathbf{z}'\mathbf{B}' + \mathbf{Bz}\boldsymbol{\varepsilon}' + \boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' & \mathbf{Bz}\boldsymbol{\xi}' + \boldsymbol{\varepsilon}\boldsymbol{\xi}'\\ \boldsymbol{\xi}\mathbf{z}'\mathbf{B}' + \boldsymbol{\xi}\boldsymbol{\varepsilon}' & \boldsymbol{\xi}\boldsymbol{\xi}'\end{bmatrix} = \\ = \begin{bmatrix}\mathbf{B}\boldsymbol{\mu}_{\mathbf{z}}\boldsymbol{\mu}'_{\mathbf{z}}\mathbf{B}' + \mathbf{B}\boldsymbol{\Sigma}_{\mathbf{zz}}\mathbf{B}' + \boldsymbol{\Sigma}_{\mathbf{yy}|\mathbf{z}} & \mathbf{B}\boldsymbol{\Sigma}_{\mathbf{zz}}\\ \boldsymbol{\Sigma}_{\mathbf{zz}}\mathbf{B}' & \boldsymbol{\Sigma}_{\mathbf{zz}}\end{bmatrix}\end{aligned} \tag{29}$$

а также:

$$\mathbb{E}\begin{bmatrix}\mathbf{Bz}+\boldsymbol{\varepsilon}\\ \boldsymbol{\xi}\end{bmatrix}\mathbb{E}\begin{bmatrix}\mathbf{z}'\mathbf{B}'+\boldsymbol{\varepsilon}' & \boldsymbol{\xi}'\end{bmatrix} = \begin{bmatrix}\mathbf{B}\boldsymbol{\mu}_{\mathbf{z}}\\ \mathbf{0}\end{bmatrix}\begin{bmatrix}\boldsymbol{\mu}'_{\mathbf{z}}\mathbf{B}' & \mathbf{0}\end{bmatrix} = \begin{bmatrix}\mathbf{B}\boldsymbol{\mu}_{\mathbf{z}}\boldsymbol{\mu}'_{\mathbf{z}}\mathbf{B}' & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix} \tag{30}$$

Итого:

$$\boldsymbol{\Sigma}_{\mathbf{xx}} = \mathbb{E}\left(\begin{bmatrix}\mathbf{Bz}+\boldsymbol{\varepsilon}\\ \boldsymbol{\xi}\end{bmatrix}\begin{bmatrix}\mathbf{z}'\mathbf{B}'+\boldsymbol{\varepsilon}' & \boldsymbol{\xi}'\end{bmatrix}\right) - \mathbb{E}\begin{bmatrix}\mathbf{Bz}+\boldsymbol{\varepsilon}\\ \boldsymbol{\xi}\end{bmatrix}\mathbb{E}\begin{bmatrix}\mathbf{z}'\mathbf{B}'+\boldsymbol{\varepsilon}' & \boldsymbol{\xi}'\end{bmatrix} = \begin{bmatrix}\mathbf{B}\boldsymbol{\Sigma}_{\mathbf{zz}}\mathbf{B}' + \boldsymbol{\Sigma}_{\mathbf{yy}|\mathbf{z}} & \mathbf{B}\boldsymbol{\Sigma}_{\mathbf{zz}}\\ \boldsymbol{\Sigma}_{\mathbf{zz}}\mathbf{B}' & \boldsymbol{\Sigma}_{\mathbf{zz}}\end{bmatrix} \tag{31}$$

При этом, если подставить $\mathbf{B} = \boldsymbol{\Sigma}_{\mathbf{yz}}\boldsymbol{\Sigma}_{\mathbf{zz}}^{-1}$, то уравнение (31) может быть переписано как:

$$\boldsymbol{\Sigma}_{\mathbf{xx}} = \begin{bmatrix}\boldsymbol{\Sigma}_{\mathbf{yz}}\boldsymbol{\Sigma}_{\mathbf{zz}}^{-1}\boldsymbol{\Sigma}_{\mathbf{zy}} + \boldsymbol{\Sigma}_{\mathbf{yy}|\mathbf{z}} & \boldsymbol{\Sigma}_{\mathbf{yz}}\\ \boldsymbol{\Sigma}_{\mathbf{zy}} & \boldsymbol{\Sigma}_{\mathbf{zz}}\end{bmatrix} \tag{32}$$

□

Утверждение 3. ***О совместном Вишарт-нормальном распределении.*** *Предположим, что набор случайных величин $(\boldsymbol{\Phi}, \boldsymbol{\Omega})$ имеют обратный Вишарт-нормальное распределение со следующими параметрами: $(\boldsymbol{\Phi}, \boldsymbol{\Omega}) \sim \mathcal{IW}_n\mathcal{MN}_{kn}(\nu, \boldsymbol{\Psi}_{\boldsymbol{\Omega}}, \mathbf{M}_{\boldsymbol{\Phi}}, \mathbf{V}_{\boldsymbol{\Phi}})$. Здесь предполагается, что $\boldsymbol{\Omega}$ — положительно определённая матрица размерности $n \times n$, а $\boldsymbol{\Phi}$ — некоторая случайная матрица размерности $k \times n$. Верны следующие соотношения:*

$$\boldsymbol{\Phi}, \boldsymbol{\Omega}^{-1} \sim \mathcal{W}_n\mathcal{MN}_{kn}(\nu, \boldsymbol{\Psi}_{\boldsymbol{\Omega}}^{-1}, \mathbf{M}_{\boldsymbol{\Phi}}, \mathbf{V}_{\boldsymbol{\Phi}}) \tag{33}$$

$$\boldsymbol{\Omega} \sim \mathcal{IW}_n(\nu, \boldsymbol{\Psi}_{\boldsymbol{\Omega}}) \tag{34}$$

$$\boldsymbol{\Omega}^{-1} \sim \mathcal{W}_n(\nu, \boldsymbol{\Psi}_{\boldsymbol{\Omega}}^{-1}) \tag{35}$$

$$\boldsymbol{\Phi}|\boldsymbol{\Omega} = \boldsymbol{\Phi}|\boldsymbol{\Omega}^{-1} \sim \mathcal{MN}_{kn}(\mathbf{M}_{\boldsymbol{\Phi}}, \mathbf{V}_{\boldsymbol{\Phi}}) \tag{36}$$

$$\mathrm{vec}\,(\boldsymbol{\Phi})\,|\boldsymbol{\Omega} = \mathrm{vec}\,(\boldsymbol{\Phi})\,|\boldsymbol{\Omega}^{-1} \sim \mathcal{N}_{kn}(\mathbf{M}_{\boldsymbol{\Phi}}, \boldsymbol{\Omega} \otimes \mathbf{V}_{\boldsymbol{\Phi}}) \tag{37}$$

$$\mathrm{vec}\,(\boldsymbol{\Phi})\,, \boldsymbol{\Omega} \sim \mathcal{IW}_n\mathcal{N}_{kn}(\nu, \boldsymbol{\Psi}_{\boldsymbol{\Omega}}, \mathbf{M}_{\boldsymbol{\Phi}}, \boldsymbol{\Omega} \otimes \mathbf{V}_{\boldsymbol{\Phi}}) \tag{38}$$

$$\boldsymbol{\Phi} \sim \mathcal{M}t_{kn}(\nu, \mathbf{M}_{\boldsymbol{\Phi}}, \boldsymbol{\Psi}_{\boldsymbol{\Omega}}, \mathbf{V}_{\boldsymbol{\Phi}}) \tag{39}$$

Доказательство. Для начала выразим распределение случайной величины $\mathbf{\Phi}, \mathbf{\Omega}$, с которой мы будем работать:

$$p(\mathbf{\Phi}, \mathbf{\Omega}) = k \left[|\det(\mathbf{\Omega})|^{-\frac{\nu+n+1}{2}} \exp\left(-\frac{1}{2}\operatorname{tr}\left(\mathbf{\Psi_\Omega}\mathbf{\Omega}^{-1}\right)\right)\right] \times \\ \times \left[|\det(\mathbf{\Omega})|^{-\frac{k}{2}} \exp\left(-\frac{1}{2}\operatorname{tr}\left(\mathbf{\Omega}^{-1}\left(\mathbf{\Phi}-\mathbf{M_\Phi}\right)'\mathbf{V_\Phi}^{-1}\left(\mathbf{\Phi}-\mathbf{M_\Phi}\right)\right)\right)\right] \quad (40)$$

Где $k = \left[(2)^{-\nu n/2}|\det(\mathbf{\Psi_\Omega})|^{\nu/2}\Gamma_n^{-1}(\frac{\nu}{2})\right]\left[(2\pi)^{-nk/2}|\det(\mathbf{V_\Phi})|^{-n/2}\right]$.

Для того, чтобы получить распределение $\mathbf{\Phi}, \mathbf{\Omega}^{-1}$, нужно умножить (78) на Якобиан замены $\mathbf{\Omega} \to \mathbf{\Omega}^{-1}$. При этом и $\mathbf{\Omega}$, и $\mathbf{\Omega}^{-1}$ являются симметричными матрицами, а следовательно, в силу уравнения (238), $|\operatorname{vech}(\mathbf{\Omega})/\operatorname{vech}(\mathbf{\Omega}^{-1})| = |\det(\mathbf{\Omega})|^{(n+1)}$:

$$p(\mathbf{\Phi}, \mathbf{\Omega}^{-1}) = |\det(\mathbf{\Omega})|^{(n+1)} p(\mathbf{\Phi}, \mathbf{\Omega}) = k \left[|\det(\mathbf{\Omega}^{-1})|^{\frac{\nu-(n+1)}{2}} \exp\left(-\frac{1}{2}\operatorname{tr}\left(\mathbf{\Psi_\Omega}\mathbf{\Omega}^{-1}\right)\right)\right] \times \\ \times \left[|\det(\mathbf{\Omega}^{-1})|^{\frac{k}{2}} \exp\left(-\frac{1}{2}\operatorname{tr}\left(\mathbf{\Omega}^{-1}\left(\mathbf{\Phi}-\mathbf{M_\Phi}\right)'\mathbf{V_\Phi}^{-1}\left(\mathbf{\Phi}-\mathbf{M_\Phi}\right)\right)\right)\right], \quad (41)$$

что является случайной величиной, распределённой по Вишарт-нормальному закону распределения со следующими гиперпарамтрами: $\mathbf{\Phi}, \mathbf{\Omega}^{-1} \sim \mathcal{W}_n\mathcal{MN}_{kn}(\nu, \mathbf{\Psi_\Omega}^{-1}, \mathbf{M_\Phi}, \mathbf{V_\Phi})$.

Теперь получим маргинальную функцию плотности матрицы $\mathbf{\Omega}$. Для этого необходимо найти следующий интеграл:

$$p(\mathbf{\Omega}) = \int_{\mathbf{\Phi}\in S(\mathbf{\Phi})} p(\mathbf{\Omega}, \mathbf{\Phi})\,\mathrm{d}\mathbf{\Phi} = \left[(2)^{-\frac{\nu n}{2}}|\det(\mathbf{\Psi_\Omega})|^{\frac{\nu}{2}}\Gamma_n^{-1}\left(\frac{\nu}{2}\right)\right]\left[|\det(\mathbf{\Omega})|^{-\frac{\nu+n+1}{2}}\exp\left(-\frac{1}{2}\operatorname{tr}\left(\mathbf{\Psi_\Omega}\mathbf{\Omega}^{-1}\right)\right)\right] \times \\ \times \int_{\mathbf{\Phi}\in S(\mathbf{\Phi})}\left[(2\pi)^{-\frac{nk}{2}}|\det(\mathbf{V_\Phi})|^{-\frac{n}{2}}\right]\left[|\det(\mathbf{\Omega})|^{-\frac{k}{2}}\exp\left(-\frac{1}{2}\operatorname{tr}\left(\mathbf{\Omega}^{-1}\left(\mathbf{\Phi}-\mathbf{M_\Phi}\right)'\mathbf{V_\Phi}^{-1}\left(\mathbf{\Phi}-\mathbf{M_\Phi}\right)\right)\right)\right]\mathrm{d}\mathbf{\Phi} = \\ = \left[(2)^{-\frac{\nu n}{2}}|\det(\mathbf{\Psi_\Omega})|^{\frac{\nu}{2}}\Gamma_n^{-1}\left(\frac{\nu}{2}\right)\right]\left[|\det(\mathbf{\Omega})|^{-\frac{\nu+n+1}{2}}\exp\left(-\frac{1}{2}\operatorname{tr}\left(\mathbf{\Psi_\Omega}\mathbf{\Omega}^{-1}\right)\right)\right], \quad (42)$$

что есть случайная величина, распределённая по закону распределения обрантого Вишарта: $\mathbf{\Omega} \sim \mathcal{IW}_n(\nu, \mathbf{\Psi_\Omega})$. Аналогично можно показать, что $\mathbf{\Omega}^{-1} \sim \mathcal{W}_n(\nu, \mathbf{\Psi_\Omega}^{-1})$.

Для того, чтобы выразить условную функцию плотности $\mathbf{\Phi}|\mathbf{\Omega}$, необходимо совместную функцию плотности разделить на маргинальную плотность для $\mathbf{\Omega}$:

$$p(\mathbf{\Phi}|\mathbf{\Omega}) = \frac{p(\mathbf{\Phi}, \mathbf{\Omega})}{p(\mathbf{\Omega})} = \left[(2\pi)^{-\frac{nk}{2}}|\det(\mathbf{V_\Phi})|^{-\frac{n}{2}}\right]\left[|\det(\mathbf{\Omega})|^{-\frac{k}{2}}\exp\left(-\frac{1}{2}\operatorname{tr}\left(\mathbf{\Omega}^{-1}\left(\mathbf{\Phi}-\mathbf{M_\Phi}\right)'\mathbf{V_\Phi}^{-1}\left(\mathbf{\Phi}-\mathbf{M_\Phi}\right)\right)\right)\right], \quad (43)$$

что является плотностью матричного нормального распределения, то есть $\mathbf{\Phi}|\mathbf{\Omega} \sim \mathcal{MN}_{kn}(\mathbf{M_\Phi}, \mathbf{V_\Phi})$. Очевидно, что также и $\mathbf{\Phi}|\mathbf{\Omega}^{-1} \sim \mathcal{MN}_{kn}(\mathbf{M_\Phi}, \mathbf{V_\Phi})$

При этом можно показать, что если $\mathbf{\Phi}|\mathbf{\Omega}$ является матричной нормальной величиной, то векторизация матрицы $\mathbf{\Phi}$ имеет многомерное нормальное распределение со следующими параметрами: $\operatorname{vec}(\mathbf{\Phi})|\mathbf{\Omega} \sim \mathcal{N}_{kn}(\mathbf{M_\Phi}, \mathbf{\Omega} \otimes \mathbf{V_\Phi})$. Для этого рассмотрим разложение Холецкого матриц $\mathbf{\Omega} = \mathbf{LL}'$ и $\mathbf{V_\Phi}^{-1} = \mathbf{PP}'$. Тогда можно рассмотреть член, который стоит под экспонентой условной функции распределения (здесь используются теоремы 1 и 2):

$$\begin{aligned}&\operatorname{tr}\left(\mathbf{L}\mathbf{L}'(\boldsymbol{\Phi}-\mathbf{M}_{\boldsymbol{\Phi}})'\mathbf{V}_{\boldsymbol{\Phi}}^{-1}(\boldsymbol{\Phi}-\mathbf{M}_{\boldsymbol{\Phi}})\right)=\operatorname{tr}\left(\mathbf{L}'(\boldsymbol{\Phi}-\mathbf{M}_{\boldsymbol{\Phi}})'\mathbf{P}\mathbf{P}'(\boldsymbol{\Phi}-\mathbf{M}_{\boldsymbol{\Phi}})\mathbf{L}\right)=\\&=\operatorname{tr}\left((\mathbf{P}'(\boldsymbol{\Phi}-\mathbf{M}_{\boldsymbol{\Phi}})\mathbf{L})'(\mathbf{P}'(\boldsymbol{\Phi}-\mathbf{M}_{\boldsymbol{\Phi}})\mathbf{L})\right)=((\mathbf{L}'\otimes\mathbf{P}')\operatorname{vec}(\boldsymbol{\Phi}-\mathbf{M}_{\boldsymbol{\Phi}}))'((\mathbf{L}'\otimes\mathbf{P}')\operatorname{vec}(\boldsymbol{\Phi}-\mathbf{M}_{\boldsymbol{\Phi}}))=\\&=(\operatorname{vec}(\boldsymbol{\Phi}-\mathbf{M}_{\boldsymbol{\Phi}}))'\underbrace{(\mathbf{L}\otimes\mathbf{P})(\mathbf{L}'\otimes\mathbf{P}')}_{(\mathbf{L}\mathbf{L}'\otimes\mathbf{P}\mathbf{P}')}(\operatorname{vec}(\boldsymbol{\Phi}-\mathbf{M}_{\boldsymbol{\Phi}}))=(\operatorname{vec}(\boldsymbol{\Phi}-\mathbf{M}_{\boldsymbol{\Phi}}))'\left(\boldsymbol{\Omega}^{-1}\otimes\mathbf{V}_{\boldsymbol{\Phi}}^{-1}\right)(\operatorname{vec}(\boldsymbol{\Phi}-\mathbf{M}_{\boldsymbol{\Phi}})),\end{aligned}\tag{44}$$

При этом известно, что $(\boldsymbol{\Omega}\otimes\mathbf{V}_{\boldsymbol{\Phi}})^{-1}=\boldsymbol{\Omega}^{-1}\otimes\mathbf{V}_{\boldsymbol{\Phi}}^{-1}$ и $|\det(\boldsymbol{\Omega})|^{-k/2}||\det(\mathbf{V}_{\boldsymbol{\Phi}})|^{-n/2}=|\det(\boldsymbol{\Omega}^{-1}\otimes\mathbf{V}_{\boldsymbol{\Phi}}^{-1})|^{\frac{1}{2}}=|\det(\boldsymbol{\Omega}\otimes\mathbf{V}_{\boldsymbol{\Phi}})|^{-\frac{1}{2}}$. Тогда можно переписать функцию плотности (43) в следующем виде (поскольку $\boldsymbol{\Phi}$ и $\operatorname{vec}(\boldsymbol{\Phi})$ отличаются только структурой расположения элементов, Якобиан замены $=1$):

$$p(\operatorname{vec}(\boldsymbol{\Phi})|\boldsymbol{\Omega})=(2\pi)^{-\frac{nk}{2}}|\det(\boldsymbol{\Omega}\otimes\mathbf{V}_{\boldsymbol{\Phi}})|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}(\operatorname{vec}(\boldsymbol{\Phi}-\mathbf{M}_{\boldsymbol{\Phi}}))'\left(\boldsymbol{\Omega}^{-1}\otimes\mathbf{V}_{\boldsymbol{\Phi}}^{-1}\right)(\operatorname{vec}(\boldsymbol{\Phi}-\mathbf{M}_{\boldsymbol{\Phi}}))\right),\tag{45}$$

что является плотностью многомерного нормального распределения: $\operatorname{vec}(\boldsymbol{\Phi})|\boldsymbol{\Omega}\sim\mathcal{N}_{kn}(\mathbf{M}_{\boldsymbol{\Phi}},\boldsymbol{\Omega}\otimes\mathbf{V}_{\boldsymbol{\Phi}})$

Для того, чтобы найти маргинальную функцию распределения для $\boldsymbol{\Phi}$, необходимо найти следующий интеграл (удобнее рассмотреть совместную функцию плотности $(\boldsymbol{\Phi},\boldsymbol{\Omega}^{-1})$ и проинтегрировать по положительно определённой матрице $\boldsymbol{\Omega}^{-1}>0$):

$$\begin{aligned}&\int_{\boldsymbol{\Omega}^{-1}>0}p(\boldsymbol{\Phi},\boldsymbol{\Omega}^{-1})\,\mathrm{d}\boldsymbol{\Omega}=\int_{\boldsymbol{\Omega}^{-1}>0}\left[(2)^{-\frac{\nu n}{2}}|\det(\boldsymbol{\Psi}_{\boldsymbol{\Omega}})|^{\frac{\nu}{2}}\Gamma_n^{-1}\left(\frac{\nu}{2}\right)\right]\left[|\det(\boldsymbol{\Omega}^{-1})|^{\frac{\nu-(n+1)}{2}}\exp\left(-\frac{1}{2}\operatorname{tr}\left(\boldsymbol{\Psi}_{\boldsymbol{\Omega}}\boldsymbol{\Omega}^{-1}\right)\right)\right]\times\\&\times\left[(2\pi)^{-\frac{nk}{2}}|\det(\mathbf{V}_{\boldsymbol{\Phi}})|^{-\frac{n}{2}}|\det(\boldsymbol{\Omega})|^{-\frac{k}{2}}\exp\left(-\frac{1}{2}\operatorname{tr}\left(\boldsymbol{\Omega}^{-1}(\boldsymbol{\Phi}-\mathbf{M}_{\boldsymbol{\Phi}})'\mathbf{V}_{\boldsymbol{\Phi}}^{-1}(\boldsymbol{\Phi}-\mathbf{M}_{\boldsymbol{\Phi}})\right)\right)\right]\mathrm{d}\boldsymbol{\Omega}^{-1}=\\&=\left[\underbrace{(2)^{-\frac{\nu n}{2}}|\det(\boldsymbol{\Psi}_{\boldsymbol{\Omega}})|^{\frac{\nu}{2}}\Gamma_n^{-1}\left(\frac{\nu}{2}\right)(2\pi)^{-\frac{nk}{2}}|\det(\mathbf{V}_{\boldsymbol{\Phi}})|^{-\frac{n}{2}}}_{D}\right]\times\\&\times\int_{\boldsymbol{\Omega}^{-1}>0}\left|\det(\boldsymbol{\Omega}^{-1})\right|^{\frac{\nu+k-(n+1)}{2}}\exp\left\{-\frac{1}{2}\operatorname{tr}\left[\boldsymbol{\Omega}^{-1}\left(\underbrace{\boldsymbol{\Psi}_{\boldsymbol{\Omega}}+(\boldsymbol{\Phi}-\mathbf{M}_{\boldsymbol{\Phi}})'\mathbf{V}_{\boldsymbol{\Phi}}^{-1}(\boldsymbol{\Phi}-\mathbf{M}_{\boldsymbol{\Phi}})}_{\mathbf{S}}\right)\right]\right\}\mathrm{d}\boldsymbol{\Omega}^{-1}=\\&=D|\det(\mathbf{S})|^{-(\frac{\nu+k-(n+1)}{2}+\frac{(n+1)}{2})}(2)^{n(\frac{\nu+k-(n+1)}{2}+\frac{n+1}{2})}\Gamma_n\left(\frac{\nu+k-(n+1)}{2}+\frac{n+1}{2}\right)=\\&=(\pi)^{-\frac{nk}{2}}\Gamma_n\left(\frac{\nu+k}{2}\right)\Gamma_n^{-1}\left(\frac{\nu}{2}\right)|\det(\mathbf{V}_{\boldsymbol{\Phi}})|^{-\frac{n}{2}}|\det(\boldsymbol{\Psi}_{\boldsymbol{\Omega}})|^{\frac{\nu}{2}}\left|\det\left(\boldsymbol{\Psi}_{\boldsymbol{\Omega}}+(\boldsymbol{\Phi}-\mathbf{M}_{\boldsymbol{\Phi}})'\mathbf{V}_{\boldsymbol{\Phi}}^{-1}(\boldsymbol{\Phi}-\mathbf{M}_{\boldsymbol{\Phi}})\right)\right|^{-\frac{\nu+k}{2}}=\\&=(\pi)^{-\frac{nk}{2}}\Gamma_n\left(\frac{\nu+k}{2}\right)\Gamma_n^{-1}\left(\frac{\nu}{2}\right)|\det(\mathbf{V}_{\boldsymbol{\Phi}})|^{-\frac{n}{2}}|\det(\boldsymbol{\Psi}_{\boldsymbol{\Omega}})|^{-\frac{k}{2}}\left|\det\left(\mathbf{I}_n+\boldsymbol{\Psi}_{\boldsymbol{\Omega}}^{-1}(\boldsymbol{\Phi}-\mathbf{M}_{\boldsymbol{\Phi}})'\mathbf{V}_{\boldsymbol{\Phi}}^{-1}(\boldsymbol{\Phi}-\mathbf{M}_{\boldsymbol{\Phi}})\right)\right|^{-\frac{\nu+k}{2}},\end{aligned}\tag{46}$$

что является матричным t-распределением с параметрами $\boldsymbol{\Phi}\sim\mathcal{M}t_{kn}(\nu,\mathbf{M}_{\boldsymbol{\Phi}},\boldsymbol{\Psi}_{\boldsymbol{\Omega}},\mathbf{V}_{\boldsymbol{\Phi}})$.

Можно показать, что в векторном виде $\operatorname{vec}(\boldsymbol{\Phi})$ распределено как многомерное t-распределение:

$$\begin{aligned}&p(\operatorname{vec}(\boldsymbol{\Phi}))=(\pi)^{-\frac{nk}{2}}\Gamma_n\left(\frac{\nu+k}{2}\right)\Gamma_n^{-1}\left(\frac{\nu}{2}\right)|\det(\boldsymbol{\Psi}_{\boldsymbol{\Omega}}^{-1}\otimes\mathbf{V}_{\boldsymbol{\Phi}}^{-1})|^{-\frac{1}{2}}\times\\&\times\left(1+\operatorname{vec}(\boldsymbol{\Phi}-\mathbf{M}_{\boldsymbol{\Phi}})'\left(\boldsymbol{\Psi}_{\boldsymbol{\Omega}}^{-1}\otimes\mathbf{V}_{\boldsymbol{\Phi}}^{-1}\right)\operatorname{vec}(\boldsymbol{\Phi}-\mathbf{M}_{\boldsymbol{\Phi}})\right)^{-\frac{\nu+k}{2}}\end{aligned}\tag{47}$$

□

Утверждение 4. *О переходе к Вишарт-нормальному распределению. Если* $\boldsymbol{\Omega}^{-1} \sim \mathcal{W}_n(\nu, \boldsymbol{\Psi}_{\boldsymbol{\Omega}})$ *и* $\boldsymbol{\Phi}|\boldsymbol{\Omega} \sim \mathcal{MN}_{kn}(\mathbf{M}_{\boldsymbol{\Phi}}, \boldsymbol{\Omega}, \mathbf{V}_{\boldsymbol{\Phi}})$*, то совместная функция распределения принадлежит семейству Вишарт-нормальных распределений:* $\boldsymbol{\Phi}, \boldsymbol{\Omega}^{-1} \sim \mathcal{W}_n\mathcal{MN}_{kn}(\nu, \boldsymbol{\Psi}_{\boldsymbol{\Omega}}, \mathbf{M}_{\boldsymbol{\Phi}}, \mathbf{V}_{\boldsymbol{\Phi}})$*. Соответственно, если* $\boldsymbol{\Omega} \sim \mathcal{IW}_n(\nu, \boldsymbol{\Psi}_{\boldsymbol{\Omega}})$ *и* $\boldsymbol{\Phi}|\boldsymbol{\Omega} \sim \mathcal{MN}_{kn}(\mathbf{M}_{\boldsymbol{\Phi}}, \boldsymbol{\Omega}, \mathbf{V}_{\boldsymbol{\Phi}})$*, то* $\boldsymbol{\Phi}, \boldsymbol{\Omega} \sim \mathcal{IW}_n\mathcal{MN}_{kn}(\nu, \boldsymbol{\Psi}_{\boldsymbol{\Omega}}, \mathbf{M}_{\boldsymbol{\Phi}}, \mathbf{V}_{\boldsymbol{\Phi}})$*.*

Доказательство. Для доказательства необходимо просто вывести совместную функцию плотности:

$$
\begin{aligned}
p(\boldsymbol{\Phi}, \boldsymbol{\Omega}^{-1}) = p(\boldsymbol{\Omega}^{-1})p(\boldsymbol{\Phi}|\boldsymbol{\Omega}^{-1}) = k \left[|\det(\boldsymbol{\Omega}^{-1})|^{\frac{\nu-(n+1)}{2}} \exp\left(-\frac{1}{2} \operatorname{tr}\left(\boldsymbol{\Psi}_{\boldsymbol{\Omega}} \boldsymbol{\Omega}^{-1} \right) \right) \right] \times \\
\times \left[|\det(\boldsymbol{\Omega})|^{-\frac{k}{2}} \exp\left(-\frac{1}{2} \operatorname{tr}\left(\boldsymbol{\Omega}^{-1} \left(\boldsymbol{\Phi} - \mathbf{M}_{\boldsymbol{\Phi}} \right)' \mathbf{V}_{\boldsymbol{\Phi}}^{-1} \left(\boldsymbol{\Phi} - \mathbf{M}_{\boldsymbol{\Phi}} \right) \right) \right) \right]
\end{aligned} \tag{48}
$$

Где $k = \left[(2)^{-\nu n/2} |\boldsymbol{\Psi}_{\boldsymbol{\Omega}}|^{\nu/2} \Gamma_n^{-1}(\frac{\nu}{2}) \right] \left[(2\pi)^{-nk/2} |\det(\mathbf{V}_{\boldsymbol{\Phi}})|^{-n/2} \right]$. Аналогично выражается $p(\boldsymbol{\Phi}, \boldsymbol{\Omega})$. □

3 Байесовские методы для VAR-моделей

Рассмотрим стандартную векторную авторегрессию p-порядка, описывающую динамику n переменных:

$$
\mathbf{y}_t' = \mathbf{y}_{t-1}' \boldsymbol{\Phi}_1 + \ldots + \mathbf{y}_{t-p}' \boldsymbol{\Phi}_p + \mathbf{d}_t' \boldsymbol{\Psi} + \mathbf{u}_t', \tag{49}
$$

где $u_t \sim \text{i.i.d.} \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega})$, $\mathbf{y}_t$ — вектор зависимой переменной размерности $(n \times 1)$, $\mathbf{d}_t$ — вектор детерминированных переменных размерности $(k_d \times 1)$. Эти переменные могут включать в себя константу, тренд, сезонные дамми и прочие дамми-переменные.

Тогда можно записать динамику системы $\mathbf{Y}_{1:T}$ при условии начальных значений $\mathbf{Y}_{(1-p):0}$ в матричной форме:

$$
\underbrace{\begin{bmatrix} \mathbf{y}_1' \\ \mathbf{y}_2' \\ \ldots \\ \mathbf{y}_T' \end{bmatrix}}_{\substack{\mathbf{Y} \\ (T \times n)}} = \underbrace{\begin{bmatrix} \mathbf{y}_0' & \mathbf{y}_{-1}' & \ldots & \mathbf{y}_{1-p}' & \mathbf{d}_1' \\ \mathbf{y}_1' & \mathbf{y}_0' & \ldots & \mathbf{y}_{2-p}' & \mathbf{d}_2' \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \mathbf{y}_{T-1}' & \mathbf{y}_{T-2}' & \ldots & \mathbf{y}_{T-p}' & \mathbf{d}_T' \end{bmatrix}}_{\substack{\mathbf{Z} \\ (T \times k)}} \underbrace{\begin{bmatrix} \boldsymbol{\Phi}_1 \\ \boldsymbol{\Phi}_2 \\ \ldots \\ \boldsymbol{\Phi}_p \\ \boldsymbol{\Psi} \end{bmatrix}}_{\substack{\boldsymbol{\Phi} \\ (k \times n)}} + \underbrace{\begin{bmatrix} \mathbf{u}_1' \\ \mathbf{u}_2' \\ \ldots \\ \mathbf{u}_T' \end{bmatrix}}_{\substack{\mathbf{u} \\ (T \times n)}} \tag{50}
$$

$$
\mathbf{Y} = \mathbf{Z}\boldsymbol{\Phi} + \mathbf{u} \tag{51}
$$

Для удобства обозначим число экзогенных переменных k и матрицу всех доступных данных $\mathbf{X} = [\mathbf{Y}, \mathbf{Z}]$.

Функция правдоподобия модели 49 может быть записана с помощью строк матриц, входящих в уравнение 51:

$$
\begin{aligned}
&p(\mathbf{Y}|\mathbf{Z}, \boldsymbol{\Phi}, \boldsymbol{\Omega}) \propto \\
&\propto |\det(\boldsymbol{\Omega})|^{-\frac{T}{2}} \exp\left(-\frac{1}{2} \sum_{t=1}^{T} (\mathbf{y}_t' - \sum_{j=1}^{p} \mathbf{y}_{t-j}' \boldsymbol{\Phi}_j - \mathbf{d}_t' \boldsymbol{\Psi}) \boldsymbol{\Omega}^{-1} (\mathbf{y}_t' - \sum_{j=1}^{p} \mathbf{y}_{t-j}' \boldsymbol{\Phi}_j - \mathbf{d}_t' \boldsymbol{\Psi})' \right) \propto \\
&\propto |\det(\boldsymbol{\Omega})|^{-\frac{T}{2}} \exp\left(-\frac{1}{2} \sum_{t=1}^{T} (\mathbf{Y}_{t,\bullet} - \mathbf{Z}_{t,\bullet} \boldsymbol{\Phi}) \boldsymbol{\Omega}^{-1} (\mathbf{Y}_{t,\bullet} - \mathbf{Z}_{t,\bullet} \boldsymbol{\Phi})' \right)
\end{aligned}
$$

В литературе для удобства дальнейших операций с функцией правдоподобия обычно записывают член под экспонентой с помощью следа матрицы в виде: $\operatorname{tr}\left\{ (\mathbf{Y} - \mathbf{X}\boldsymbol{\Phi}) \boldsymbol{\Omega}^{-1} (\mathbf{Y} - \mathbf{X}\boldsymbol{\Phi})' \right\}$. Чтобы понять, правомерна ли подобная замена, рассмотрим произведение трех матриц:

$$
\underset{(T \times n)}{\mathbf{M}} \; \mathbf{H} \underset{(n \times T)}{\mathbf{M}'} = \begin{bmatrix} \mathbf{M}_{1,\bullet} \\ \mathbf{M}_{2,\bullet} \\ \ldots \\ \mathbf{M}_{T,\bullet} \end{bmatrix} \mathbf{H} \begin{bmatrix} \mathbf{M}_{1,\bullet}', & \mathbf{M}_{2,\bullet}', & \ldots, & \mathbf{M}_{T,\bullet}' \end{bmatrix} = \begin{bmatrix} \mathbf{M}_{1,\bullet}\mathbf{H}\mathbf{M}_{1,\bullet}' & \mathbf{M}_{1,\bullet}\mathbf{H}\mathbf{M}_{2,\bullet}' & \ldots & \mathbf{M}_{1,\bullet}\mathbf{H}\mathbf{M}_{T,\bullet}' \\ \mathbf{M}_{2,\bullet}\mathbf{H}\mathbf{M}_{1,\bullet}' & \mathbf{M}_{2,\bullet}\mathbf{H}\mathbf{M}_{2,\bullet}' & \ldots & \mathbf{M}_{2,\bullet}\mathbf{H}\mathbf{M}_{T,\bullet}' \\ \vdots & \ldots & \ddots & \vdots \\ \mathbf{M}_{T,\bullet}\mathbf{H}\mathbf{M}_{1,\bullet}' & \mathbf{M}_{T,\bullet}\mathbf{H}\mathbf{M}_{2,\bullet}' & \ldots & \mathbf{M}_{T,\bullet}\mathbf{H}\mathbf{M}_{T,\bullet}' \end{bmatrix}
$$

Из последней записи видно, что $\operatorname{tr}(\mathbf{MHM}') = \sum_{t=1}^{T} \mathbf{M}_{t,\bullet}\mathbf{H}\mathbf{M}'_{t,\bullet}$, что и требовалось доказать.

Таким образом, с учётом вышеназванной замены, получаем следующую функцию правдоподобия:

$$p(\mathbf{Y}|\mathbf{Z},\mathbf{\Phi},\mathbf{\Omega}) \propto |\det(\mathbf{\Omega})|^{-\frac{T}{2}} \exp\left(-\frac{1}{2}\operatorname{tr}\left\{(\mathbf{Y}-\mathbf{Z\Phi})\mathbf{\Omega}^{-1}(\mathbf{Y}-\mathbf{Z\Phi})'\right\}\right) \propto$$
$$\propto |\det(\mathbf{\Omega})|^{-\frac{T}{2}} \exp\left(-\frac{1}{2}\operatorname{tr}\left\{\mathbf{\Omega}^{-1}(\mathbf{Y}-\mathbf{Z\Phi})'(\mathbf{Y}-\mathbf{Z\Phi})\right\}\right) \quad (52)$$

Утверждение 5. *Об естественных сопряжённых распределениях параметров VAR-моделей. Для VAR-модели, что может быть записана в виде уравнения* (51)*, естественным сопряжённым априорным распределением параметров может считаться семейство обратного Вишарт-нормальных распределений. Другими словами:*

$$\mathbf{\Omega},\mathbf{\Phi}|\mathbf{X} \sim \mathcal{IW}_n\mathcal{MN}_{kn}(\overline{\nu} = T-k, \overline{\mathbf{\Psi}} = \widehat{\mathbf{S}}, \overline{\mathbf{\Phi}} = \widehat{\mathbf{\Phi}}, \mathbf{\Sigma} = \mathbf{\Omega}, \overline{\mathbf{V}}_{\mathbf{\Phi}} = (\mathbf{Z}'\mathbf{Z})^{-1}) \quad (53)$$
$$\mathbf{\Omega}|\mathbf{X} \sim \mathcal{IW}_n(\overline{\nu} = T-k, \overline{\mathbf{\Psi}} = \widehat{\mathbf{S}}) \quad (54)$$
$$\mathbf{\Omega}^{-1}|\mathbf{X} \sim \mathcal{W}_n(\overline{\nu} = T-k, \overline{\mathbf{\Psi}} = \widehat{\mathbf{S}}^{-1}) \quad (55)$$
$$\mathbf{\Phi}|\mathbf{\Omega},\mathbf{X} \sim \mathcal{MN}_{kn}(\overline{\mathbf{\Phi}} = \widehat{\mathbf{\Phi}}, \mathbf{\Sigma} = \mathbf{\Omega}, \overline{\mathbf{V}}_{\mathbf{\Phi}} = (\mathbf{Z}'\mathbf{Z})^{-1}) \quad (56)$$
$$\mathbf{\Phi}|\mathbf{X} \sim \mathcal{M}t_{kn}(\overline{\nu} = T-k, \mathbf{M} = \widehat{\mathbf{\Phi}}, \mathbf{P} = \mathbf{\Omega}, \mathbf{Q} = (\mathbf{Z}'\mathbf{Z})^{-1}) \quad (57)$$

Тогда в терминах векторизации параметров $\operatorname{vec}(\mathbf{\Phi})$*:*

$$\mathbf{\Omega},\operatorname{vec}(\mathbf{\Phi})|\mathbf{X} \sim \mathcal{IW}_n\mathcal{N}_{kn}(\overline{\nu} = T-k, \overline{\mathbf{\Psi}} = \widehat{\mathbf{S}}, \overline{\boldsymbol{\varphi}} = \widehat{\boldsymbol{\varphi}}, \overline{\mathbf{V}}_{\boldsymbol{\varphi}} = \mathbf{\Omega}\otimes(\mathbf{Z}'\mathbf{Z})^{-1}) \quad (58)$$
$$\mathbf{\Omega}|\mathbf{X} \sim \mathcal{IW}_n(\overline{\nu} = T-k, \overline{\mathbf{\Psi}} = \widehat{\mathbf{S}}) \quad (59)$$
$$\operatorname{vec}(\mathbf{\Phi})|\mathbf{\Omega},\mathbf{X} \sim \mathcal{N}_{kn}(\overline{\boldsymbol{\varphi}} = \widehat{\boldsymbol{\varphi}}, \overline{\mathbf{V}}_{\boldsymbol{\varphi}} = \mathbf{\Omega}\otimes(\mathbf{Z}'\mathbf{Z})^{-1}) \quad (60)$$
$$\operatorname{vec}(\mathbf{\Phi})|\mathbf{X} \sim t_{kn}(\overline{\nu} = T-k, \boldsymbol{\mu} = \widehat{\boldsymbol{\varphi}}, \mathbf{V} = \mathbf{\Omega}\otimes(\mathbf{Z}'\mathbf{Z})^{-1}) \quad (61)$$

Доказательство. Для того, чтобы выразить естественные сопряжённые распределения параметров VAR-модели, можно выразить апостериорные распределения для случая Jeffreys priors.

Для того, чтобы вывести Jeffreys prior для $\mathbf{\Phi},\mathbf{\Omega}$, определим $\boldsymbol{\varphi} = \operatorname{vec}(\mathbf{\Phi})$, $\boldsymbol{\omega} = \operatorname{vech}(\mathbf{\Omega})$, а также $\boldsymbol{\theta} = [\boldsymbol{\varphi}', \boldsymbol{\omega}]'$. Тогда по определению Jeffreys prior выглядит следующим образом:

$$p(\mathbf{\Phi},\mathbf{\Omega}) \propto |\det(\mathcal{I}(\boldsymbol{\theta}))|^{\frac{1}{2}} = \left|\det\left[\mathbb{E}\left(\frac{\partial^2 \ln(\mathcal{L}(\mathbf{\Phi},\mathbf{\Omega}))}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right)\right]\right|^{\frac{1}{2}} \quad (62)$$

При этом:

$$\ln(\mathcal{L}(\mathbf{\Phi},\mathbf{\Omega})) = const - \frac{T}{2}ln(|\det(\mathbf{\Omega})|) - \frac{1}{2}\left(\operatorname{tr}(\mathbf{\Omega}^{-1}\mathbf{Y}'\mathbf{Y}) + \operatorname{tr}(\mathbf{Z\Phi\Omega}^{-1}\mathbf{\Phi}'\mathbf{Z}') - 2\operatorname{tr}(\mathbf{Z}'\mathbf{Y\Omega}^{-1}\mathbf{\Phi}')\right) = \quad (63)$$
$$const - \frac{T}{2}ln(|\det(\mathbf{\Omega})|) - \frac{1}{2}\left(\operatorname{tr}(\mathbf{\Omega}^{-1}\mathbf{Y}'\mathbf{Y}) + \operatorname{tr}(\mathbf{\Phi}'\mathbf{Z}'\mathbf{Z\Phi\Omega}^{-1}) - 2\operatorname{tr}(\mathbf{\Omega}^{-1}\mathbf{Y}'\mathbf{Z\Phi})\right) \quad (64)$$
$$\frac{\partial\ln(\mathcal{L}(\mathbf{\Phi},\mathbf{\Omega}))}{\partial\operatorname{vec}(\mathbf{\Phi})} = \operatorname{vec}\left(-\mathbf{Z}'\mathbf{Z\Phi\Omega}^{-1} + \mathbf{Z}'\mathbf{Y\Omega}^{-1}\right) = (\mathbf{I}_n\otimes\mathbf{Z}'\mathbf{Y})\operatorname{vec}(\mathbf{\Omega}^{-1}) - (\mathbf{\Omega}^{-1}\otimes\mathbf{Z}'\mathbf{Z})\operatorname{vec}(\mathbf{\Phi}) \quad (65)$$
$$\frac{\partial\ln(\mathcal{L}(\mathbf{\Phi},\mathbf{\Omega}))}{\partial\operatorname{vech}(\mathbf{\Omega})} = \operatorname{vech}\left(\mathbf{\Omega}^{-1} - \frac{1}{2}\mathbf{Y}'\mathbf{Y} - \frac{1}{2}\mathbf{\Omega}^{-1}\mathbf{\Phi}'\mathbf{Z}'\mathbf{Z\Phi\Omega}^{-1} + \mathbf{\Omega}^{-1}\mathbf{\Phi}'\mathbf{ZY}'\mathbf{\Omega}^{-1}\right) \quad (66)$$
$$\frac{\partial^2\ln(\mathcal{L}(\mathbf{\Phi},\mathbf{\Omega}))}{\partial\operatorname{vec}(\mathbf{\Phi})\partial\operatorname{vec}(\mathbf{\Phi})'} = \mathbf{\Omega}^{-1}\otimes\mathbf{Z}'\mathbf{Z} \quad (67)$$
$$(68)$$

Я сдаюсь. В статьях пишут, что Jeffreys prior — это:

$$p(\mathbf{\Phi}|\mathbf{\Omega}) \propto 1 \tag{69}$$
$$p(\mathbf{\Omega}) \propto |\det(\mathbf{\Omega})|^{-\frac{n+1}{2}} \tag{70}$$
$$\tag{71}$$

Следовательно,

$$p(\mathbf{\Phi}, \mathbf{\Omega}) = p(\mathbf{\Omega})p(\mathbf{\Phi}|\mathbf{\Omega}) \propto |\det(\mathbf{\Omega})|^{-\frac{n+1}{2}} \tag{72}$$

Вывод апостериорных плотностей разделим на два этапа.

Во-первых, выразим функцию правдоподобия в удобной для Байесовского анализа форме. Во-вторых, выразим сами апостериорные плотности.

Этап первый. Для того, чтобы его реализовать, необходимо расписать функцию правдоподобия в терминах OLS-оценки параметров VAR: $\widehat{\mathbf{\Phi}} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}$. Для этого рассмотрим выражение под экспонентой в функции правдоподобия:

$$\begin{aligned}
&\operatorname{tr}\left\{\mathbf{\Omega}^{-1}(\mathbf{Y}-\mathbf{Z\Phi})'(\mathbf{Y}-\mathbf{Z\Phi})\right\} = \operatorname{tr}\left\{\mathbf{\Omega}^{-1}(\mathbf{Y}-\mathbf{Z}\widehat{\mathbf{\Phi}}+\mathbf{Z}\widehat{\mathbf{\Phi}}-\mathbf{Z\Phi})'(\mathbf{Y}-\mathbf{Z}\widehat{\mathbf{\Phi}}+\mathbf{Z}\widehat{\mathbf{\Phi}}-\mathbf{Z\Phi})\right\} = \\
&= \operatorname{tr}\left\{\mathbf{\Omega}^{-1}(\mathbf{Y}-\mathbf{Z}\widehat{\mathbf{\Phi}})'(\mathbf{Y}-\mathbf{Z}\widehat{\mathbf{\Phi}})\right\} + \operatorname{tr}\left\{\mathbf{\Omega}^{-1}(\widehat{\mathbf{\Phi}}-\mathbf{\Phi})'\mathbf{Z}'(\mathbf{Y}-\mathbf{Z}\widehat{\mathbf{\Phi}})\right\} + \operatorname{tr}\left\{\mathbf{\Omega}^{-1}(\mathbf{Y}-\mathbf{Z}\widehat{\mathbf{\Phi}})'\mathbf{Z}(\widehat{\mathbf{\Phi}}-\mathbf{\Phi})\right\} + \\
&+ \operatorname{tr}\left\{\mathbf{\Omega}^{-1}(\widehat{\mathbf{\Phi}}-\mathbf{\Phi})'\mathbf{Z}'\mathbf{Z}(\widehat{\mathbf{\Phi}}-\mathbf{\Phi})\right\} = \operatorname{tr}\left\{\mathbf{\Omega}^{-1}(\mathbf{Y}-\mathbf{Z}\widehat{\mathbf{\Phi}})'(\mathbf{Y}-\mathbf{Z}\widehat{\mathbf{\Phi}})\right\} + \operatorname{tr}\left\{\mathbf{\Omega}^{-1}(\widehat{\mathbf{\Phi}}-\mathbf{\Phi})'\mathbf{Z}'\mathbf{Z}(\widehat{\mathbf{\Phi}}-\mathbf{\Phi})\right\} + \\
&+ \operatorname{tr}\left\{\mathbf{\Omega}^{-1}(\widehat{\mathbf{\Phi}}-\mathbf{\Phi})'\underbrace{(\mathbf{Z}'\mathbf{Y}-\underbrace{\mathbf{Z}'\mathbf{Z}\widehat{\mathbf{\Phi}}}_{\mathbf{X}'\mathbf{Y}}+\underbrace{\mathbf{Z}'(\mathbf{Y}-\mathbf{Z}\widehat{\mathbf{\Phi}})}_{\mathbf{Z}'\mathbf{Y}-\mathbf{Z}'\mathbf{Y}})}_{\mathbf{0}}\mathbf{Z}'(\widehat{\mathbf{\Phi}}-\mathbf{\Phi})\right\}
\end{aligned} \tag{73}$$

Таким образом,

$$\operatorname{tr}\left\{\mathbf{\Omega}^{-1}(\mathbf{Y}-\mathbf{Z\Phi})'(\mathbf{Y}-\mathbf{Z\Phi})\right\} = \operatorname{tr}\left\{\mathbf{\Omega}^{-1}(\mathbf{Y}-\mathbf{Z}\widehat{\mathbf{\Phi}})'(\mathbf{Y}-\mathbf{Z}\widehat{\mathbf{\Phi}})\right\} + \operatorname{tr}\left\{\mathbf{\Omega}^{-1}(\widehat{\mathbf{\Phi}}-\mathbf{\Phi})'\mathbf{Z}'\mathbf{Z}(\widehat{\mathbf{\Phi}}-\mathbf{\Phi})\right\} \tag{74}$$

Следовательно, функция правдоподобия принимает следующий вид:

$$\begin{aligned}
&p(\mathbf{Y}|\mathbf{\Phi},\mathbf{\Omega},\mathbf{Z}) = \\
&= (2\pi)^{-\frac{nT}{2}}|\det(\mathbf{\Omega})|^{-\frac{T}{2}} \exp\left\{-\frac{1}{2}\operatorname{tr}\left(\mathbf{\Omega}^{-1}(\mathbf{Y}-\mathbf{Z}\widehat{\mathbf{\Phi}})'(\mathbf{Y}-\mathbf{Z}\widehat{\mathbf{\Phi}})\right)\right\} \exp\left\{-\frac{1}{2}\operatorname{tr}\left(\mathbf{\Omega}^{-1}(\mathbf{\Phi}-\widehat{\mathbf{\Phi}})'\mathbf{Z}'\mathbf{Z}(\mathbf{\Phi}-\widehat{\mathbf{\Phi}})\right)\right\}
\end{aligned} \tag{75}$$

Обозначим $\widehat{\mathbf{S}} = (\mathbf{Y}-\mathbf{Z}\widehat{\mathbf{\Phi}})'(\mathbf{Y}-\mathbf{Z}\widehat{\mathbf{\Phi}})$. Тогда функция правдоподобия записывается в виде:

$$p(\mathbf{Y}|\mathbf{\Phi},\mathbf{\Omega},\mathbf{Z}) = (2\pi)^{-\frac{nT}{2}}|\det(\mathbf{\Omega})|^{-\frac{T}{2}} \exp\left\{-\frac{1}{2}\operatorname{tr}\left(\mathbf{\Omega}^{-1}\widehat{\mathbf{S}}\right)\right\} \exp\left\{-\frac{1}{2}\operatorname{tr}\left(\mathbf{\Omega}^{-1}(\mathbf{\Phi}-\widehat{\mathbf{\Phi}})'\mathbf{Z}'\mathbf{Z}(\mathbf{\Phi}-\widehat{\mathbf{\Phi}})\right)\right\} \tag{76}$$

Помимо данного представления возможно представление в векторной форме параметров $\text{vec}(\mathbf{\Phi})$. Для этого необходимо выразить подынтегральное выражение в функции правдоподобия (76). Распишем разложение Холецкого матрицы концентрации $\mathbf{\Omega}^{-1} = \mathbf{LL}'$ и рассмотрим следующее выражение (для вывода использовались теоремы 1 и 2):

$$\operatorname{tr}\left(\mathbf{L}\mathbf{L}'(\mathbf{\Phi}-\widehat{\mathbf{\Phi}})'\mathbf{Z}'\mathbf{Z}(\mathbf{\Phi}-\widehat{\mathbf{\Phi}})\right)=\operatorname{tr}\left(\mathbf{L}'(\mathbf{\Phi}-\widehat{\mathbf{\Phi}})'\mathbf{Z}'\mathbf{Z}(\mathbf{\Phi}-\widehat{\mathbf{\Phi}})\mathbf{L}\right)=\operatorname{tr}\left(\left(\mathbf{Z}(\mathbf{\Phi}-\widehat{\mathbf{\Phi}})\mathbf{L}\right)'\left(\mathbf{Z}(\mathbf{\Phi}-\widehat{\mathbf{\Phi}})\mathbf{L}\right)\right)=$$
$$=\operatorname{vec}\left(\mathbf{Z}(\mathbf{\Phi}-\widehat{\mathbf{\Phi}})\mathbf{L}\right)'\operatorname{vec}\left(\mathbf{Z}(\mathbf{\Phi}-\widehat{\mathbf{\Phi}})\mathbf{L}\right)=\left((\mathbf{L}'\otimes\mathbf{Z})\operatorname{vec}(\mathbf{\Phi}-\widehat{\mathbf{\Phi}})\right)'\left((\mathbf{L}'\otimes\mathbf{Z})\operatorname{vec}(\mathbf{\Phi}-\widehat{\mathbf{\Phi}})\right)=$$
$$=\left(\operatorname{vec}(\mathbf{\Phi}-\widehat{\mathbf{\Phi}})\right)'\underbrace{(\mathbf{L}\otimes\mathbf{Z}')(\mathbf{L}'\otimes\mathbf{Z})}_{(\mathbf{L}\mathbf{L}'\otimes\mathbf{Z}'\mathbf{Z})}\left(\operatorname{vec}(\mathbf{\Phi}-\widehat{\mathbf{\Phi}})\right)=\left(\operatorname{vec}(\mathbf{\Phi}-\widehat{\mathbf{\Phi}})\right)'\left(\mathbf{\Omega}^{-1}\otimes\mathbf{Z}'\mathbf{Z}\right)\left(\operatorname{vec}(\mathbf{\Phi}-\widehat{\mathbf{\Phi}})\right),$$

где оценку $\widehat{\mathbf{\Phi}}$ можно переписать в следующем виде: $\operatorname{vec}(\widehat{\mathbf{\Phi}})=\operatorname{vec}\left((\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}\right)=\operatorname{vec}\left(((\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}')\,\mathbf{Y}\mathbf{I}_n\right)=(\mathbf{I}_n\otimes(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}')\operatorname{vec}(\mathbf{Y})$. Для удобства дальнейших рассуждений обозначим $\boldsymbol{\varphi}=\operatorname{vec}(\mathbf{\Phi})$. В этом случае функция правдоподобия, выраженная в терминах векторизованных параметров, выглядит следующим образом:

$$p(\mathbf{Y}|\operatorname{vec}(\mathbf{\Phi}),\mathbf{\Omega},\mathbf{Z})\propto\left[|\det(\mathbf{\Omega})|^{-\frac{T}{2}}\exp\left(-\frac{1}{2}\operatorname{tr}\left(\mathbf{\Omega}^{-1}\widehat{\mathbf{S}}\right)\right)\right]\times$$
$$\times\left[\left|\det(\mathbf{\Omega}^{-1}\otimes(\mathbf{Z}'\mathbf{Z}))\right|^{\frac{1}{2}}\exp\left(-\frac{1}{2}\operatorname{tr}\left((\operatorname{vec}(\mathbf{\Phi})-\widehat{\boldsymbol{\varphi}})'\left(\mathbf{\Omega}^{-1}\otimes\mathbf{Z}'\mathbf{Z}\right)(\operatorname{vec}(\mathbf{\Phi})-\widehat{\boldsymbol{\varphi}})\right)\right)\right]\left|\det(\mathbf{\Omega}^{-1}\otimes(\mathbf{Z}'\mathbf{Z}))\right|^{-\frac{1}{2}},\tag{77}$$

где $\widehat{\mathbf{S}}=(\mathbf{Y}-\mathbf{Z}\widehat{\mathbf{\Phi}})'(\mathbf{Y}-\mathbf{Z}\widehat{\mathbf{\Phi}})$ и $\widehat{\boldsymbol{\varphi}}=(\mathbf{I}_n\otimes(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}')\operatorname{vec}(\mathbf{Y})$. При этом из (228) видно, что $|\det(\mathbf{\Omega}^{-1}\otimes(\mathbf{Z}'\mathbf{Z}))|^{0.5}=|\det(\mathbf{\Omega})|^{-0.5k}|\det(\mathbf{Z}'\mathbf{Z})|^{0.5n}$. Тогда выражение (77) превращается в следующее уравнение:

$$p(\mathbf{Y}|\operatorname{vec}(\mathbf{\Phi}),\mathbf{\Omega},\mathbf{Z})\propto\left[|\det(\mathbf{\Omega})|^{-\frac{T-k}{2}}\exp\left(-\frac{1}{2}\operatorname{tr}\left(\mathbf{\Omega}^{-1}\widehat{\mathbf{S}}\right)\right)\right]\times$$
$$\times\left[|\det(\mathbf{\Omega})|^{-\frac{k}{2}}\exp\left(-\frac{1}{2}\operatorname{tr}\left((\operatorname{vec}(\mathbf{\Phi})-\widehat{\boldsymbol{\varphi}})'\left(\mathbf{\Omega}^{-1}\otimes\mathbf{Z}'\mathbf{Z}\right)(\operatorname{vec}(\mathbf{\Phi})-\widehat{\boldsymbol{\varphi}})\right)\right)\right]\tag{78}$$

Перейдём ко второму этапу. На втором этапе необходимо выразить апостериорную функцию распределения параметров. Для параметров $\mathbf{\Phi}$ в матричной форме апостериорная совместная функция плотности выглядит следующим образом:

$$p(\mathbf{\Phi},\mathbf{\Omega}|\mathbf{X})\propto p(\mathbf{\Phi},\mathbf{\Omega})p(\mathbf{Y}|\mathbf{\Phi},\mathbf{\Omega},\mathbf{Z})\propto|\mathbf{\Omega}|^{-\frac{n+1}{2}}p(\mathbf{Y}|\mathbf{\Phi},\mathbf{\Omega}^{-1},\mathbf{Z})\propto$$
$$\propto\left[|\det(\mathbf{\Omega})|^{-\frac{T-k+n+1}{2}}\exp\left\{-\frac{1}{2}\operatorname{tr}\left(\mathbf{\Omega}^{-1}\widehat{\mathbf{S}}\right)\right\}\right]\left[|\det(\mathbf{\Omega})|^{-\frac{k}{2}}\exp\left\{-\frac{1}{2}\operatorname{tr}\left(\mathbf{\Omega}^{-1}(\mathbf{\Phi}-\widehat{\mathbf{\Phi}})'\mathbf{Z}'\mathbf{Z}(\mathbf{\Phi}-\widehat{\mathbf{\Phi}})\right)\right\}\right],\tag{79}$$

При этом эта функция плотности принадлежит семейству Вишарт-нормальных распределений. Таким образом, в силу утверждения, верно следующее:

$$\mathbf{\Omega},\mathbf{\Phi}|\mathbf{X}\sim\mathcal{IW}_n\mathcal{MN}_{kn}(\overline{\nu}=T-k,\overline{\mathbf{\Psi}}=\widehat{\mathbf{S}},\overline{\mathbf{\Phi}}=\widehat{\mathbf{\Phi}},\mathbf{\Sigma}=\mathbf{\Omega},\overline{\mathbf{V}}_{\mathbf{\Phi}}=(\mathbf{Z}'\mathbf{Z})^{-1})\tag{80}$$
$$\mathbf{\Omega}|\mathbf{X}\sim\mathcal{IW}_n(\overline{\nu}=T-k,\overline{\mathbf{\Psi}}=\widehat{\mathbf{S}})\tag{81}$$
$$\mathbf{\Omega}^{-1}|\mathbf{X}\sim\mathcal{W}_n(\overline{\nu}=T-k,\overline{\mathbf{\Psi}}=\widehat{\mathbf{S}}^{-1})\tag{82}$$
$$\mathbf{\Phi}|\mathbf{\Omega},\mathbf{X}\sim\mathcal{MN}_{kn}(\overline{\mathbf{\Phi}}=\widehat{\mathbf{\Phi}},\mathbf{\Sigma}=\mathbf{\Omega},\overline{\mathbf{V}}_{\mathbf{\Phi}}=(\mathbf{Z}'\mathbf{Z})^{-1})\tag{83}$$
$$\mathbf{\Phi}|\mathbf{X}\sim\mathcal{M}t_{kn}(\overline{\nu}=T-k,\mathbf{M}=\widehat{\mathbf{\Phi}},\mathbf{P}=\mathbf{\Omega},\mathbf{Q}=(\mathbf{Z}'\mathbf{Z})^{-1})\tag{84}$$

Для параметров $\operatorname{vec}(\mathbf{\Phi})$ в векторизованной форме апостериорная совместная функция плотности выглядит следующим образом:

$$p(\mathrm{vec}(\mathbf{\Phi}),\mathbf{\Omega}^{-1}|\mathbf{X}) \propto \left[|\det(\mathbf{\Omega})|^{-\frac{T-k+n+1}{2}}\exp\left(-\frac{1}{2}\operatorname{tr}\left(\mathbf{\Omega}^{-1}\widehat{\mathbf{S}}\right)\right)\right]\times$$
$$\times\left[|\det(\mathbf{\Omega})|^{-\frac{k}{2}}\exp\left(-\frac{1}{2}\operatorname{tr}\left((\mathrm{vec}(\mathbf{\Phi})-\widehat{\boldsymbol{\varphi}})'\left(\mathbf{\Omega}^{-1}\otimes\mathbf{Z}'\mathbf{Z}\right)(\mathrm{vec}(\mathbf{\Phi})-\widehat{\boldsymbol{\varphi}})\right)\right)\right] \quad (85)$$

Таким образом, в силу утверждения, верно следующее:

$$\mathbf{\Omega},\mathrm{vec}(\mathbf{\Phi})|\mathbf{X}\sim\mathcal{IW}_n\mathcal{N}_{kn}(\overline{\nu}=T-k,\overline{\mathbf{\Psi}}=\widehat{\mathbf{S}},\overline{\mathbf{\Phi}}=\widehat{\mathbf{\Phi}},\overline{\mathbf{V}}_{\boldsymbol{\varphi}}=\mathbf{\Omega}\otimes(\mathbf{Z}'\mathbf{Z})^{-1}) \quad (86)$$
$$\mathbf{\Omega}|\mathbf{X}\sim\mathcal{IW}_n(\overline{\nu}=T-k,\overline{\mathbf{\Psi}}=\widehat{\mathbf{S}}) \quad (87)$$
$$\mathrm{vec}(\mathbf{\Phi})|\mathbf{\Omega},\mathbf{X}\sim\mathcal{N}_{kn}(\overline{\mathbf{\Phi}}=\mathrm{vec}(\widehat{\mathbf{\Phi}}),\overline{\mathbf{V}}_{\boldsymbol{\varphi}}=\mathbf{\Omega}\otimes(\mathbf{Z}'\mathbf{Z})^{-1}) \quad (88)$$
$$\mathrm{vec}(\mathbf{\Phi})|\mathbf{X}\sim t_{kn}(\overline{\nu}=T-k,\mathbf{M}=\widehat{\mathbf{\Phi}},\mathbf{V}=\mathbf{\Omega}\otimes(\mathbf{Z}'\mathbf{Z})^{-1}) \quad (89)$$

□

Утверждение 6. *О выводе апостериорных плотностей параметров VAR-моделей для случая Вишарт-Нормальных априорных распределений. В случае, если в качестве априорных распределений параметров используется Вишарт-нормальные распределения вида:*

$$\mathbf{\Omega}\sim\mathcal{IW}_n(\underline{\nu},\underline{\mathbf{S}}) \quad (90)$$
$$\mathbf{\Omega}^{-1}\sim\mathcal{W}_n(\underline{\nu},\underline{\mathbf{S}}^{-1}) \quad (91)$$
$$\mathbf{\Phi}|\mathbf{\Omega}=\mathbf{\Phi}|\mathbf{\Omega}^{-1}\sim\mathcal{MN}_{kn}(\underline{\mathbf{\Phi}},\mathbf{\Omega},\underline{\mathbf{V}}) \quad (92)$$
$$\mathrm{vec}\left(\mathbf{\Phi}\right)|\mathbf{\Omega}=\mathrm{vec}\left(\mathbf{\Phi}\right)|\mathbf{\Omega}^{-1}\sim\mathcal{N}_{kn}\left(\mathrm{vec}(\underline{\mathbf{\Phi}}),(\mathbf{\Omega}\otimes\underline{\mathbf{V}})\right), \quad (93)$$

то апостериорное распределение параметров будет также принадлежать семейству Вишарт-нормальных распределений со следующим набором гиперпараметров:

$$\mathbf{\Omega},\mathbf{\Phi}|\mathbf{X}\sim\mathcal{W}_n\mathcal{MN}_{kn}(\overline{\nu},\overline{\mathbf{S}},\overline{\mathbf{\Phi}},\overline{\mathbf{V}}) \quad (94)$$
$$\overline{\mathbf{V}}=(\mathbf{Z}'\mathbf{Z}+\underline{\mathbf{V}}^{-1})^{-1} \quad (95)$$
$$\overline{\mathbf{\Phi}}=\overline{\mathbf{V}}(\mathbf{Z}'\mathbf{Y}+\underline{\mathbf{V}}^{-1}\underline{\mathbf{\Phi}}) \quad (96)$$
$$\overline{\nu}=\underline{\nu}+T \quad (97)$$
$$\overline{\mathbf{S}}=\underline{\mathbf{S}}+\widehat{\mathbf{S}}-\overline{\mathbf{\Phi}}'\overline{\mathbf{V}}^{-1}\overline{\mathbf{\Phi}}+\widehat{\mathbf{\Phi}}'\mathbf{Z}'\mathbf{Z}\widehat{\mathbf{\Phi}}+\underline{\mathbf{\Phi}}'\underline{\mathbf{V}}^{-1}\underline{\mathbf{\Phi}}=\underline{\mathbf{S}}+(\mathbf{Y}-\mathbf{Z}\underline{\mathbf{\Phi}})'(\mathbf{I}_T-\mathbf{Z}\underline{\mathbf{V}}\mathbf{Z}')^{-1}(\mathbf{Y}-\mathbf{Z}\underline{\mathbf{\Phi}}) \quad (98)$$

Поэтому

$$\mathbf{\Omega},\mathbf{\Phi}|\mathbf{X}\sim\mathcal{IW}_n\mathcal{MN}_{kn}(\overline{\nu}=T-k,\overline{\mathbf{\Psi}}=\widehat{\mathbf{S}},\overline{\mathbf{\Phi}}=\widehat{\mathbf{\Phi}},\mathbf{\Sigma}=\mathbf{\Omega},\overline{\mathbf{V}}_{\mathbf{\Phi}}=(\mathbf{Z}'\mathbf{Z})^{-1}) \quad (99)$$
$$\mathbf{\Omega}|\mathbf{X}\sim\mathcal{IW}_n(\overline{\nu}=T-k,\overline{\mathbf{\Psi}}=\widehat{\mathbf{S}}) \quad (100)$$
$$\mathbf{\Omega}^{-1}|\mathbf{X}\sim\mathcal{W}_n(\overline{\nu}=T-k,\overline{\mathbf{\Psi}}=\widehat{\mathbf{S}}^{-1}) \quad (101)$$
$$\mathbf{\Phi}|\mathbf{\Omega},\mathbf{X}\sim\mathcal{MN}_{kn}(\overline{\mathbf{\Phi}}=\widehat{\mathbf{\Phi}},\mathbf{\Sigma}=\mathbf{\Omega},\overline{\mathbf{V}}_{\mathbf{\Phi}}=(\mathbf{Z}'\mathbf{Z})^{-1}) \quad (102)$$
$$\mathbf{\Phi}|\mathbf{X}\sim\mathcal{M}t_{kn}(\overline{\nu}=T-k,\mathbf{M}=\widehat{\mathbf{\Phi}},\mathbf{P}=\mathbf{\Omega},\mathbf{Q}=(\mathbf{Z}'\mathbf{Z})^{-1}) \quad (103)$$

Доказательство. Для начала запишем совместную функцию априорного распределения параметров:

$$p(\mathbf{\Phi},\mathbf{\Omega})\propto\left[|\det(\mathbf{\Omega})|^{-\frac{1}{2}(\underline{\nu}+n+1)}\exp\left\{-\frac{1}{2}\operatorname{tr}(\underline{\mathbf{S}}\mathbf{\Omega}^{-1})\right\}\right]\left[|\det(\mathbf{\Omega})|^{-\frac{k}{2}}\exp\left\{-\frac{1}{2}\operatorname{tr}\left(\mathbf{\Omega}^{-1}(\mathbf{\Phi}-\underline{\mathbf{\Phi}})'\underline{\mathbf{V}}^{-1}(\mathbf{\Phi}-\underline{\mathbf{\Phi}})\right)\right\}\right] \quad (104)$$

В силу представления функции правдоподобия в форме, апостериорную функцию распределения можно записать в виде:

$$p(\mathbf{\Phi},\mathbf{\Omega}|\mathbf{X}) \propto \left[|\det(\mathbf{\Omega})|^{-\frac{1}{2}(T+\underline{\nu}+n+1)}\exp\left(-\frac{1}{2}\operatorname{tr}\left(\mathbf{\Omega}^{-1}(\widehat{\mathbf{S}}+\underline{\mathbf{S}})\right)\right)\right]\times$$
$$\times\left[|\det(\mathbf{\Omega})|^{-\frac{k}{2}}\exp\left\{-\frac{1}{2}\operatorname{tr}\left(\mathbf{\Omega}^{-1}(\mathbf{\Phi}-\widehat{\mathbf{\Phi}})'\mathbf{Z}'\mathbf{Z}(\mathbf{\Phi}-\widehat{\mathbf{\Phi}})+\mathbf{\Omega}^{-1}(\mathbf{\Phi}-\underline{\mathbf{\Phi}})'\underline{\mathbf{V}}^{-1}(\mathbf{\Phi}-\underline{\mathbf{\Phi}})\right)\right\}\right] \quad (105)$$

Тогда выражение под экспонентой может быть расписано в следующем виде:

$$(\mathbf{\Phi}-\widehat{\mathbf{\Phi}})'\mathbf{Z}'\mathbf{Z}(\mathbf{\Phi}-\widehat{\mathbf{\Phi}})+(\mathbf{\Phi}-\underline{\mathbf{\Phi}})'\underline{\mathbf{V}}^{-1}(\mathbf{\Phi}-\underline{\mathbf{\Phi}})=$$
$$=\mathbf{\Phi}'\mathbf{Z}'\mathbf{Z}\mathbf{\Phi}-\widehat{\mathbf{\Phi}}'\mathbf{Z}'\mathbf{Z}\mathbf{\Phi}-\mathbf{\Phi}'\mathbf{Z}'\mathbf{Z}\widehat{\mathbf{\Phi}}+\widehat{\mathbf{\Phi}}'\mathbf{Z}'\mathbf{Z}\widehat{\mathbf{\Phi}}+\mathbf{\Phi}'\overline{\mathbf{V}}^{-1}\mathbf{\Phi}-\underline{\mathbf{\Phi}}'\overline{\mathbf{V}}^{-1}\mathbf{\Phi}-\mathbf{\Phi}'\overline{\mathbf{V}}^{-1}\underline{\mathbf{\Phi}}+\underline{\mathbf{\Phi}}'\overline{\mathbf{V}}^{-1}\underline{\mathbf{\Phi}}=$$
$$=\mathbf{\Phi}'\underbrace{(\mathbf{Z}'\mathbf{Z}+\underline{\mathbf{V}}^{-1})}_{\overline{\mathbf{V}}^{-1}}\mathbf{\Phi}-\mathbf{\Phi}'(\underbrace{\mathbf{Z}\widehat{\mathbf{\Phi}}}_{\mathbf{Z}'\mathbf{Y}}+\underline{\mathbf{V}}^{-1}\underline{\mathbf{\Phi}})-(\underbrace{\mathbf{Z}\widehat{\mathbf{\Phi}}}_{\mathbf{Z}'\mathbf{Y}}+\underline{\mathbf{V}}^{-1}\underline{\mathbf{\Phi}})'\mathbf{\Phi}+\widehat{\mathbf{\Phi}}'\mathbf{Z}'\mathbf{Z}\widehat{\mathbf{\Phi}}+\underline{\mathbf{\Phi}}'\underline{\mathbf{V}}^{-1}\underline{\mathbf{\Phi}}=$$
$$=\mathbf{\Phi}'\overline{\mathbf{V}}^{-1}\mathbf{\Phi}-\mathbf{\Phi}'\overline{\mathbf{V}}^{-1}\underbrace{\overline{\mathbf{V}}(\mathbf{Z}'\mathbf{Y}+\underline{\mathbf{V}}^{-1}\underline{\mathbf{\Phi}})}_{\overline{\mathbf{\Phi}}}-\underbrace{(\mathbf{Z}'\mathbf{Y}+\underline{\mathbf{V}}^{-1}\underline{\mathbf{\Phi}})'\overline{\mathbf{V}}}_{\overline{\mathbf{\Phi}}'}\overline{\mathbf{V}}^{-1}\mathbf{\Phi}+\widehat{\mathbf{\Phi}}'\mathbf{Z}'\mathbf{Z}\widehat{\mathbf{\Phi}}+\underline{\mathbf{\Phi}}'\underline{\mathbf{V}}^{-1}\underline{\mathbf{\Phi}}=$$
$$=(\mathbf{\Phi}-\overline{\mathbf{\Phi}})'\overline{\mathbf{V}}^{-1}(\mathbf{\Phi}-\overline{\mathbf{\Phi}})-\overline{\mathbf{\Phi}}'\overline{\mathbf{V}}^{-1}\overline{\mathbf{\Phi}}+\widehat{\mathbf{\Phi}}'\mathbf{Z}'\mathbf{Z}\widehat{\mathbf{\Phi}}+\underline{\mathbf{\Phi}}'\underline{\mathbf{V}}^{-1}\underline{\mathbf{\Phi}} \quad (106)$$

Итого:

$$p(\mathbf{\Phi},\mathbf{\Omega}|\mathbf{X}) \propto \left[|\det(\mathbf{\Omega})|^{-\frac{1}{2}(\overline{\nu}+n+1)}\exp\left(-\frac{1}{2}\operatorname{tr}\left(\mathbf{\Omega}^{-1}\overline{\mathbf{S}}\right)\right)\right]\left[|\det(\mathbf{\Omega})|^{-\frac{k}{2}}\exp\left\{-\frac{1}{2}\operatorname{tr}\left(\mathbf{\Omega}^{-1}(\mathbf{\Phi}-\overline{\mathbf{\Phi}})'\overline{\mathbf{V}}^{-1}(\mathbf{\Phi}-\overline{\mathbf{\Phi}})\right)\right\}\right. \quad (107)$$

$$\overline{\mathbf{V}}=(\mathbf{Z}'\mathbf{Z}+\underline{\mathbf{V}}^{-1})^{-1} \quad (108)$$

$$\overline{\mathbf{\Phi}}=\overline{\mathbf{V}}(\mathbf{Z}'\mathbf{Y}+\underline{\mathbf{V}}^{-1}\underline{\mathbf{\Phi}}) \quad (109)$$

$$\overline{\nu}=\underline{\nu}+T \quad (110)$$

$$\overline{\mathbf{S}}=\underline{\mathbf{S}}+\widehat{\mathbf{S}}-\overline{\mathbf{\Phi}}'\overline{\mathbf{V}}^{-1}\overline{\mathbf{\Phi}}+\widehat{\mathbf{\Phi}}'\mathbf{Z}'\mathbf{Z}\widehat{\mathbf{\Phi}}+\underline{\mathbf{\Phi}}'\underline{\mathbf{V}}^{-1}\underline{\mathbf{\Phi}} \quad (111)$$

Таким образом, апостериорное распределение параметров принадлежит семейству Вишарт-нормальных распределений. Условные и маргинальные функции плотности могут быть выражены подобно утверждению .

Альтернативное выражение для $\overline{\mathbf{S}}$. Все действия, проведённые в данной части доказательства, аналогичны тем, что использовал Sudipto Banerjee в материалах к курсу "Introduction to Bayesian Analysis"для School of Public Health Университета Минесота[2] для случая одномерной регрессии. Рассмотрим $\overline{\mathbf{S}}$ (??). Распишем для начала следующее слагаемое:

$$\widehat{\mathbf{S}}+\widehat{\mathbf{\Phi}}'\mathbf{Z}'\mathbf{Z}\widehat{\mathbf{\Phi}}=\mathbf{Y}'\mathbf{Y}-\underbrace{\widehat{\mathbf{\Phi}}'\mathbf{Z}'\mathbf{Y}}_{\mathbf{Y}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}}-\underbrace{\mathbf{Y}'\mathbf{Z}\widehat{\mathbf{\Phi}}}_{\mathbf{Y}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}}+\underbrace{2\widehat{\mathbf{\Phi}}'\mathbf{Z}'\mathbf{Z}\widehat{\mathbf{\Phi}}}_{\mathbf{Y}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}}=\mathbf{Y}'\mathbf{Y} \quad (112)$$

Теперь распишем матрицы апостериорных параметров (следует помнить, что матрица ковариации симметричная):

$$\overline{\mathbf{S}}-\underline{\mathbf{S}}=\mathbf{Y}'\mathbf{Y}+\underline{\mathbf{\Phi}}'\underline{\mathbf{V}}^{-1}\underline{\mathbf{\Phi}}-\overline{\mathbf{\Phi}}'\overline{\mathbf{V}}^{-1}\overline{\mathbf{\Phi}}=\mathbf{Y}'\mathbf{Y}+\underline{\mathbf{\Phi}}'\underline{\mathbf{V}}^{-1}\underline{\mathbf{\Phi}}-(\mathbf{Z}'\mathbf{Y}+\underline{\mathbf{V}}^{-1}\underline{\mathbf{\Phi}})'\overline{\mathbf{V}}(\mathbf{Z}'\mathbf{Y}+\underline{\mathbf{V}}^{-1}\underline{\mathbf{\Phi}})=$$
$$=\mathbf{Y}'(\mathbf{I}_T-\mathbf{Z}\overline{\mathbf{V}}\mathbf{Z}')\mathbf{Y}-\mathbf{Y}'\mathbf{Z}\overline{\mathbf{V}}\underline{\mathbf{V}}^{-1}\underline{\mathbf{\Phi}}-\underline{\mathbf{\Phi}}'\underline{\mathbf{V}}^{-1}\overline{\mathbf{V}}\mathbf{Z}'\mathbf{Y}+\underline{\mathbf{\Phi}}'(\underline{\mathbf{V}}^{-1}-\underline{\mathbf{V}}^{-1}\overline{\mathbf{V}}\underline{\mathbf{V}}^{-1})\underline{\mathbf{\Phi}} \quad (113)$$

[2] http://www.biostat.umn.edu/ ph7440/

В силу того, что обратная матрица, равная сумме двух матриц может быть расписана как:

$$(\mathbf{A}+\mathbf{BDC})^{-1}=\mathbf{A}^{-1}-\mathbf{A}^{-1}\mathbf{B}(\mathbf{D}^{-1}+\mathbf{CA}^{-1}\mathbf{B})^{-1}\mathbf{CA}^{-1}, \tag{114}$$

то верно следующее выражение (если представить, что $\mathbf{A}=\underline{\mathbf{V}}$, $\mathbf{D}=(\mathbf{Z}'\mathbf{Z})^{-1}$, $\mathbf{B}=\mathbf{C}=\mathbf{I}_k$):

$$\underline{\mathbf{V}}^{-1}-\underline{\mathbf{V}}^{-1}\overline{\mathbf{V}}\underline{\mathbf{V}}^{-1}=\underline{\mathbf{V}}^{-1}-\underline{\mathbf{V}}^{-1}(\mathbf{Z}'\mathbf{Z}+\underline{\mathbf{V}}^{-1})^{-1}\underline{\mathbf{V}}^{-1}=\left(\underline{\mathbf{V}}+(\mathbf{Z}'\mathbf{Z})^{-1}\right)^{-1} \tag{115}$$

Это же выражение можно расписать по той же формуле, приняв за $\mathbf{A}=(\mathbf{Z}'\mathbf{Z})^{-1}$, $\mathbf{D}=\underline{\mathbf{V}}$, $\mathbf{B}=\mathbf{C}=\mathbf{I}_k$:

$$\left((\mathbf{Z}'\mathbf{Z})^{-1}+\underline{\mathbf{V}}\right)^{-1}=\mathbf{Z}'\mathbf{Z}-\mathbf{Z}'\mathbf{Z}\underbrace{(\mathbf{Z}'\mathbf{Z}+\underline{\mathbf{V}}^{-1})^{-1}}_{\overline{\mathbf{V}}}\mathbf{Z}'\mathbf{Z}=\mathbf{Z}'(\mathbf{I}_T-\mathbf{Z}\overline{\mathbf{V}}\mathbf{Z}')\mathbf{Z} \tag{116}$$

Можно также заметить, что $\overline{\mathbf{V}}(\underline{\mathbf{V}}^{-1}+\mathbf{Z}'\mathbf{Z})=\mathbf{I}_k$, а значит, $\overline{\mathbf{V}}\underline{\mathbf{V}}^{-1}=\mathbf{I}_k-\overline{\mathbf{V}}\mathbf{Z}'\mathbf{Z}$, то есть:

$$\mathbf{Z}\overline{\mathbf{V}}\underline{\mathbf{V}}^{-1}=\mathbf{Z}-\mathbf{Z}\overline{\mathbf{V}}\mathbf{Z}'\mathbf{Z}=(\mathbf{I}_T-\mathbf{Z}\overline{\mathbf{V}}\mathbf{Z}')\mathbf{Z} \tag{117}$$

Но если рассмотреть транспонированное выражение $\underline{\mathbf{V}}^{-1}\overline{\mathbf{V}}\mathbf{Z}'$ (в силу симметричности $(\mathbf{I}_T-\mathbf{Z}\overline{\mathbf{V}}\mathbf{Z}')$):

$$\overline{\mathbf{V}}\underline{\mathbf{V}}^{-1}\mathbf{Z}'=\mathbf{Z}'(\mathbf{I}_T-\mathbf{Z}\overline{\mathbf{V}}\mathbf{Z}')'=\mathbf{Z}'(\mathbf{I}_T-\mathbf{Z}\overline{\mathbf{V}}\mathbf{Z}') \tag{118}$$

Подставляя (117) в (113), получаем следующее:

$$\mathbf{Y}'(\mathbf{I}_T-\mathbf{Z}\overline{\mathbf{V}}\mathbf{Z}')\mathbf{Y}-\mathbf{Y}'(\mathbf{I}_T-\mathbf{Z}\overline{\mathbf{V}}\mathbf{Z}')\mathbf{Z}\underline{\mathbf{\Phi}}+\underline{\mathbf{\Phi}}'\mathbf{Z}'(\mathbf{I}_T-\mathbf{Z}\overline{\mathbf{V}}\mathbf{Z}')\mathbf{Y}+\underline{\mathbf{\Phi}}'(\mathbf{I}_T-\mathbf{Z}\overline{\mathbf{V}}\mathbf{Z}')\underline{\mathbf{\Phi}}=(\mathbf{Y}-\mathbf{Z}\underline{\mathbf{\Phi}})'(\mathbf{I}_T-\mathbf{Z}\overline{\mathbf{V}}\mathbf{Z}')(\mathbf{Y}-\mathbf{Z}\underline{\mathbf{\Phi}}) \tag{119}$$

При этом:

$$(\mathbf{I}_T-\mathbf{Z}\overline{\mathbf{V}}\mathbf{Z}')^{-1}=\mathbf{I}_T-\mathbf{Z}(\underline{\mathbf{V}}^{-1}+\mathbf{Z}'\mathbf{Z})\mathbf{Z}'=\mathbf{I}_T-\mathbf{Z}\underline{\mathbf{V}}\mathbf{Z}' \tag{120}$$

Итого получаем следующую форму апостериорного гиперпараметра:

$$\overline{\mathbf{S}}=\underline{\mathbf{S}}+(\mathbf{Y}-\mathbf{Z}\underline{\mathbf{\Phi}})'(\mathbf{I}_T-\mathbf{Z}\underline{\mathbf{V}}\mathbf{Z}')^{-1}(\mathbf{Y}-\mathbf{Z}\underline{\mathbf{\Phi}}) \tag{121}$$

□

Утверждение 7. *Распределение Вишарта для диагональных матриц. Пусть случайная диагональная матрица имеет распределение Вишарта* $\mathbf{X}=diag(\mathbf{x})\sim\mathcal{W}_n(\nu,\mathbf{\Psi})$ $(\mathbf{x}=[x_{11},\ldots,x_{nn}])$, *где матрица* $\mathbf{\Psi}=diag(\boldsymbol{\psi})$, *а* $\boldsymbol{\psi}=[\psi_{11},\ldots,\psi_{nn}]'$. *Тогда функцию плотности* $\mathbf{X}$ *можно расписать как произведение плотностей Гамма-распределений:*

$$p(\mathbf{X})=p(x_1)p(x_2|x_1)\ldots p(x_n|x_1\ldots x_{n-1})=p(\mathbf{X})=p(x_1)p(x_2)p(x_3)\ldots p(x_n)=\prod_{i=1}^{n}\gamma\left(\frac{\nu-n+1}{2},\psi_{ii}\right) \tag{122}$$

Доказательство. В силу диагональности $\mathbf{\Psi}^{-1}=diag(\tilde{\boldsymbol{\psi}})$, где $\tilde{\boldsymbol{\psi}}=[\psi_{11}^{-1},\ldots,\psi_{nn}^{-1}]$, при этом $|\det(\mathbf{\Psi})|=\prod_{i=1}^{n}\psi_{ii}$. Таким образом:

$$p(\mathbf{X})\propto\left[|\det(\mathbf{X})|^{\frac{\nu-(n+1)}{2}}\exp\left(-\operatorname{tr}(\mathbf{\Psi}^{-1}\mathbf{S})\right)\right]\propto\prod_{i=1}^{n}\left(x_{ii}^{\frac{\nu-n+1}{2}-1}\exp\left(-\frac{x_{ii}}{\psi_{ii}}\right)\right), \tag{123}$$

что есть ядро функции плотности гамма-распределения с параметрами $x_{ii}\sim\Gamma\left(\frac{\nu-n+1}{2},\psi_{ii}\right)$. □

Утверждение 8. *Маргинальная функция правдоподобия для VAR-модели в редуцированной форме. Предположим, что генерирущий данные процесс описывается некоторой VAR-моделью $\mathcal{M}_s$, априорные распределения параметров которой принадлежат семейству Вишарт-нормальных распределений со следующими гиперпараметрами:*

$$\boldsymbol{\Omega}_s|\mathcal{M}_s \sim \mathcal{IW}(\underline{\nu}_s, \underline{\mathbf{S}}_s) \tag{124}$$

$$\boldsymbol{\Omega}_s^{-1}|\mathcal{M}_s \sim \mathcal{W}(\underline{\nu}_s, \underline{\mathbf{S}}_s^{-1}) \tag{125}$$

$$\boldsymbol{\Phi}_s|\boldsymbol{\Omega};\mathcal{M}_s = \boldsymbol{\Phi}_s|\boldsymbol{\Omega}_s^{-1} \sim \mathcal{MN}_{nk}(\underline{\boldsymbol{\Phi}}_s, \boldsymbol{\Omega}_s, \underline{\mathbf{V}}_s) \tag{126}$$

$$\operatorname{vec}(\boldsymbol{\Phi}_s)\,|\boldsymbol{\Omega}_s;\mathcal{M}_s = \operatorname{vec}(\boldsymbol{\Phi}_s)\,|\boldsymbol{\Omega}_s^{-1};\mathcal{M}_s \sim \mathcal{N}_{nk}\left(\operatorname{vec}(\underline{\boldsymbol{\Phi}}_s), (\boldsymbol{\Omega}_s \otimes \underline{\mathbf{V}}_s)\right), \tag{127}$$

Тогда маргинальная функция правдоподобия при условии модели есть суть функция плотности матричного t-распределения со следующими гиперпараметрами:

$$\mathbf{Y}|\mathbf{Z}_s;\mathcal{M}_s \sim \mathcal{M}t_{Tn}(\underline{\nu}_s, \mathbf{Z}\underline{\boldsymbol{\Phi}}_s, (\mathbf{I}_n - \mathbf{Z}\underline{\mathbf{V}}_s\mathbf{Z}')^{-1}, \underline{\mathbf{S}}_s) \tag{128}$$

$$\mathbf{Y}|\boldsymbol{\Omega}_s^{-1};\mathbf{Z}_s;\mathcal{M}_s \sim \mathcal{M}t_{Tn}(\underline{\nu}_s, \mathbf{Z}\underline{\boldsymbol{\Phi}}_s, (\mathbf{I}_n - \mathbf{Z}\underline{\mathbf{V}}_s\mathbf{Z}')^{-1}, \overline{\mathbf{S}}_s - \underline{\mathbf{S}}_s) \tag{129}$$

$$\overline{\mathbf{V}}_s = (\mathbf{Z}_s'\mathbf{Z}_s + \underline{\mathbf{V}}_s^{-1})^{-1} \tag{130}$$

$$\overline{\boldsymbol{\Phi}}_s = \overline{\mathbf{V}}_s(\mathbf{Z}_s'\mathbf{Y} + \underline{\mathbf{V}}_s^{-1}\underline{\boldsymbol{\Phi}}_s) \tag{131}$$

$$\overline{\nu}_s = \underline{\nu}_s + T \tag{132}$$

$$\overline{\mathbf{S}}_s = \underline{\mathbf{S}}_s + \widehat{\mathbf{S}}_s - \overline{\boldsymbol{\Phi}}_s'\overline{\mathbf{V}}_s^{-1}\overline{\boldsymbol{\Phi}}_s + \widehat{\boldsymbol{\Phi}}_s'\mathbf{Z}_s'\mathbf{Z}_s\widehat{\boldsymbol{\Phi}}_s + \underline{\boldsymbol{\Phi}}_s'\underline{\mathbf{V}}_s^{-1}\underline{\boldsymbol{\Phi}}_s \tag{133}$$

Доказательство. Для начала найдём маргинальную функцию правдоподобия при условии матрицы концентрации. Для этого необходимо найти следующий интеграл:

$$\begin{aligned}
&p(\mathbf{Y}|\boldsymbol{\Omega}_s^{-1};\mathbf{Z},\mathcal{M}_s) = \\
&= \int_{\boldsymbol{\Phi}_s \in S(\boldsymbol{\Phi}_s)} p(\boldsymbol{\Phi}_s|\boldsymbol{\Omega}_s^{-1};\mathcal{M}_s)p(\mathbf{Y}|\boldsymbol{\Phi}_s,\boldsymbol{\Omega}_s^{-1};\mathbf{Z},\mathcal{M}_s)\,\mathrm{d}\boldsymbol{\Phi}_s = C_s|\det(\boldsymbol{\Omega}_s)|^{-\frac{T}{2}}\exp\left(-\frac{1}{2}\operatorname{tr}\left(\boldsymbol{\Omega}_s^{-1}(\overline{\mathbf{S}_s} - \underline{\mathbf{S}_s})\right)\right)\times \\
&\times \int_{\boldsymbol{\Phi}_s \in S(\boldsymbol{\Phi}_s)} (2\pi)^{\frac{1}{2}nk}|\det(\overline{\mathbf{V}}_s)|^{-\frac{n}{2}}|\det(\boldsymbol{\Omega}_s)|^{-\frac{k}{2}}\exp\left\{-\frac{1}{2}\operatorname{tr}\left(\boldsymbol{\Omega}_s^{-1}(\boldsymbol{\Phi} - \overline{\boldsymbol{\Phi}}_s)'\overline{\mathbf{V}}_s^{-1}(\boldsymbol{\Phi} - \overline{\boldsymbol{\Phi}}_s)\right)\right\}\mathrm{d}\boldsymbol{\Phi}_s = \\
&= C_s|\det(\boldsymbol{\Omega}_s)|^{-\frac{T}{2}}\exp\left(-\frac{1}{2}\operatorname{tr}\left(\boldsymbol{\Omega}_s^{-1}(\overline{\mathbf{S}}_s - \underline{\mathbf{S}}_s)\right)\right),
\end{aligned} \tag{134}$$

$$C_s = (2\pi)^{-\frac{1}{2}nT}\left(\frac{|\det(\overline{\mathbf{V}}_s)|}{|\det(\underline{\mathbf{V}}_s)|}\right)^{\frac{n}{2}} \tag{135}$$

Для того, чтобы найти безусловную маргинальную функцию правдоподобия, необходимо найти следующий интеграл:

$$\begin{aligned}
&p(\mathbf{Y}|\mathbf{Z},\mathcal{M}_s) = \int_{\boldsymbol{\Omega}_s^{-1}>0} p(\boldsymbol{\Omega}_s^{-1}|\mathcal{M}_s)p(\mathbf{Y}|\boldsymbol{\Omega}_s^{-1};\mathbf{Z},\mathcal{M}_s)\,\mathrm{d}\boldsymbol{\Omega}_s^{-1} = \\
&= D_s\int_{\mathbf{C}_s>0}\left[|\det(\mathbf{C}_s)|^{\frac{1}{2}(\underline{\nu}-n-1)}\exp\left(-\frac{1}{2}\operatorname{tr}(\mathbf{C}_s\underline{\mathbf{S}}_s)\right)\right]\left[|\det(\mathbf{C}_s)|^{\frac{T}{2}}\exp\left(-\frac{1}{2}\operatorname{tr}\left(\mathbf{C}_s(\overline{\mathbf{S}}_s - \underline{\mathbf{S}}_s)\right)\right)\right]\mathrm{d}\mathbf{C}_s = \\
&= D_s\int_{\mathbf{C}_s>0}|\det(\mathbf{C}_s)|^{\frac{1}{2}(T+\underline{\nu}-n-1)}\exp\left(-\frac{1}{2}\operatorname{tr}\left(\mathbf{C}_s\overline{\mathbf{S}}_s\right)\right)\mathrm{d}\mathbf{C}_s
\end{aligned} \tag{136}$$

$$D_s = |\det(\underline{\mathbf{S}}_s)|^{\frac{1}{2}\underline{\nu}_s}(2)^{-\frac{1}{2}n\underline{\nu}_s}(2\pi)^{-\frac{1}{2}nT}\left(\frac{|\overline{\mathbf{V}}_s|}{|\underline{\mathbf{V}}_s|}\right)^{\frac{n}{2}}\frac{1}{\Gamma_n(\frac{1}{2}\underline{\nu})} \tag{137}$$

Если присмотреться, этот интеграл может быть представлен в виде интеграла (253). Следовательно, решать его можно аналогично:

$$p(\mathbf{Y}|\mathbf{Z},\mathcal{M}_s) = D_s|\det(\overline{\mathbf{S}}_s)|^{-\frac{1}{2}(T+\underline{\nu}_s)}(2)^{\frac{n}{2}(T+\underline{\nu}_s)}\Gamma_n\left(\frac{T+\underline{\nu}_s}{2}\right) \tag{138}$$

Для того, чтобы причесать формулу, используем, что $|\det(\mathbf{A}+\mathbf{BDC})| = |\det(\mathbf{A})||\det(\mathbf{B})||\det(\mathbf{D}^{-1}+\mathbf{CA}^{-1}\mathbf{B})|$:

$$|\det\left(\mathbf{I}_n+\mathbf{Z}\underline{\mathbf{V}}\mathbf{Z}'\right)| = |\det\left(\underline{\mathbf{V}}\right)||\det\left(\underline{\mathbf{V}}^{-1}+\mathbf{Z}'\mathbf{Z}\right)| = |\det\left(\underline{\mathbf{V}}\right)||\det\left(\overline{\mathbf{V}}\right)|^{-1} \tag{139}$$

Итого, подставив (121) и (139) в (138), получаем:

$$\begin{aligned}
p(\mathbf{Y}|\mathbf{Z},\mathcal{M}_s) = (\pi)^{-\frac{nT}{2}}\frac{\Gamma_n\left(\frac{T+\underline{\nu}_s}{2}\right)}{\Gamma_n\left(\frac{1}{2}\underline{\nu}_s\right)}|\det(\mathbf{I}_n-\mathbf{Z}\underline{\mathbf{V}}_s\mathbf{Z}')|^{-\frac{n}{2}}|\det(\underline{\mathbf{S}}_s)|^{\frac{1}{2}\underline{\nu}_s}\times \\
\times|\det(\underline{\mathbf{S}}_s+(\mathbf{Y}-\mathbf{Z}\underline{\mathbf{\Phi}}_s)'(\mathbf{I}_n-\mathbf{Z}\underline{\mathbf{V}}_s\mathbf{Z}')^{-1}(\mathbf{Y}-\mathbf{Z}\underline{\mathbf{\Phi}}_s))|^{-\frac{1}{2}(T+\underline{\nu}_s)} = (\pi)^{-\frac{nT}{2}}\frac{\Gamma_n\left(\frac{T+\underline{\nu}_s}{2}\right)}{\Gamma_n\left(\frac{1}{2}\underline{\nu}_s\right)}|\det(\mathbf{I}_n-\mathbf{Z}\underline{\mathbf{V}}_s\mathbf{Z}')|^{-\frac{n}{2}}\times \\
\times|\det(\underline{\mathbf{S}}_s)|^{-\frac{T}{2}}|\det\left(\mathbf{I}_n+\underline{\mathbf{S}}_s^{-1}(\mathbf{Y}-\mathbf{Z}\underline{\mathbf{\Phi}}_s)'(\mathbf{I}_n-\mathbf{Z}\underline{\mathbf{V}}_s\mathbf{Z}')^{-1}(\mathbf{Y}-\mathbf{Z}\underline{\mathbf{\Phi}}_s)\right)|^{-\frac{1}{2}(T+\underline{\nu}_s)},
\end{aligned} \tag{140}$$

что есть ничто иное, как матричное t-распределение:

$$\mathbf{Y}|\mathbf{Z},\mathcal{M}_s \sim \mathcal{M}t_{Tn}(\underline{\nu}_s,\mathbf{Z}\underline{\mathbf{\Phi}}_s,(\mathbf{I}_n-\mathbf{Z}\underline{\mathbf{V}}_s\mathbf{Z}')^{-1},\underline{\mathbf{S}}_s) \tag{141}$$

□

4 Декомпозиция Бартлетта и её расширения

Утверждение 9. *О нормировке параметра математического ожидания распределения Вишарта. Если $\mathbf{\Omega}\sim\mathcal{W}_n(\nu,\mathbf{S})$ и матрица $\mathbf{M}$ размерности $k\times n$ ранга k, то $\mathbf{M\Omega M}'\sim\mathcal{W}_n(\nu,\mathbf{MSM}')$.*

Утверждение 10. *О разложении Бартлетта. Пусть $\mathbf{\Omega}\sim\mathcal{W}_n(\nu,\mathbf{I}_n)$, где $\nu\geqslant n$. При этом разложение Холецкого $\mathbf{\Omega}=\mathbf{LL}'$, где $\mathbf{L}$ — нижняя треугольная матрица размерности $n\times n$. Тогда внедиагональные элементы l_{ij} матрицы $\mathbf{L}$, где $1\leqslant i\leqslant j\leqslant n$, независимо одинаково распределены как $l_{ij}\sim\mathcal{N}(0,1)$. При этом квадраты диагональных элементов l_{ij}^2 матрицы $\mathbf{L}$ распределены как $\mathcal{G}((\nu-i+1)/2,2)$, или, что то же самое, $\chi^2_{\nu-i+1}$.*

Доказательство. Рассмотрим распределение матрицы $\mathbf{\Omega}$:

$$p(\mathbf{\Omega}) = 2^{-\frac{n\nu}{2}}\Gamma_n^{-1}\left(\frac{\nu}{2}\right)|\det(\mathbf{\Omega})|^{\frac{\nu-n-1}{2}}\exp\left\{-\frac{1}{2}\operatorname{tr}(\mathbf{\Omega})\right\} \tag{142}$$

При этом

$$\operatorname{tr}(\mathbf{\Omega}) = \operatorname{tr}(\mathbf{LL}') = \sum_{j=1}^{n}\sum_{i=1}^{j}l_{ij}^2 \tag{143}$$

$$\det(\mathbf{\Omega}) = \det(\mathbf{LL}') = (\det(\mathbf{L}))^2 = \prod_{i=1}^{n}l_{ii}^2 \tag{144}$$

$$\Gamma_n\left(\frac{\nu}{2}\right) = \pi^{\frac{n(n-1)}{4}}\prod_{i=1}^{n}\Gamma\left(\frac{\nu-i+1}{2}\right) \tag{145}$$

$$\tag{146}$$

А в силу выводов раздела приложения A.4.4, Якобиан замены исходной матрицы на её разложение Холецкого:

$$\left|\frac{\partial\,\mathrm{vech}(\mathbf{\Omega})}{\partial\,\mathrm{vech}(\mathbf{L})}\right| = 2^n\prod_{i=1}^{n} l_{ii}^{n-i+1}$$

Кроме того, сделаем замену $l_{ii}\to l_{ii}^2$. То есть необходимо также домножить на $2^{-n}\prod_{i=1}^{n} l_{ii}^{-1}$. Тогда функция распределения матрицы $\mathbf{L}$ выглядит как (следует помнить, что в матрице $\mathbf{L}$ ровно $n(n-1)/2$ недиагональных элементов):

$$\begin{aligned} p(\mathbf{L}) = p(\mathbf{\Omega})\left|\frac{\partial\,\mathrm{vech}(\mathbf{\Omega})}{\partial\,\mathrm{vech}(\mathbf{L})}\right| = 2^{\frac{n\nu}{2}}\pi^{-\frac{n(n-1)}{4}}\prod_{i=1}^{n}\Gamma^{-1}\left(\frac{\nu-i+1}{2}\right)\prod_{i=1}^{n} l_{ii}^{\nu-n-1}\exp\left\{-\frac{1}{2}\sum_{j=1}^{n}\sum_{i=1}^{j} l_{ij}^2\right\}\prod_{i=1}^{n} l_{ii}^{n-i} = \\ = \prod_{i=1(i<j)}^{n}\left[(2\pi)^{-\frac{1}{2}}\exp\left(-\frac{1}{2}l_{ij}^2\right)\right]\prod_{i=1}^{n}\left[2^{-(\nu-i+1)/2}\Gamma^{-1}\left(\frac{n-i+1}{2}\right)(l_{ii}^2)^{(n-i+1)/2}\exp\left(-\frac{1}{2}l_{ii}^2\right)\right] \end{aligned} \tag{147}$$

Таким образом, $l_{ij}\sim\mathcal{N}(0,1)$, а $l_{ii}\sim\mathcal{G}((\nu-i+1)/2,2)$. Следует также помнить, что $mathcalG((\nu-i+1)/2,2)=\chi^2_{\nu-i+1}$. □

Утверждение 11. *Об алгоритме AS 53 (Wishart Variate Generator). Пусть $\mathbf{\Omega}\sim\mathcal{W}_n(\nu,\mathbf{S})$. Тогда матрица $\mathbf{\Omega}$ может быть сгенерирована следующим образом.*

1. *Сгенерировать нижнюю треугольную матрицу $\mathbf{L}$, диагональные элементы которой $l_{ii}\sim\chi^2_{\nu-i+1}$, а внедиагональные элементы $l_{ij}\sim\mathcal{N}(0,1)$;*

2. *Рассчитать $\mathbf{\Omega}=\mathbf{MLL'M'}$, где $\mathbf{M}$ — разложение Холецкого матрицы $\mathbf{S}=\mathbf{MM'}$.*

Утверждение 12. *О разложении Бартлетта для смещённого математического ожидания распределённой по Вишарту матрицы. Пусть $\mathbf{\Omega}\sim\mathcal{W}_n(\nu,\mathbf{S})$. Тогда:*

1. *Элементы разложения Холецкого матрицы $\mathbf{M'\Omega M}=\mathbf{RR'}$ (где $\mathbf{S}^{-1}=\mathbf{MM'}$) распределены следующим образом. Недиагональные элементы $r_{ij}, i<j$ распределены независимо нормально $\sim\mathcal{N}(0,1)$. А диагональные элементы $r_{ii}\sim\mathcal{G}((\nu-i+1)/2,2)$,*

2. *Элементы разложения Холецкого матрицы $\mathbf{\Omega}=\mathbf{LL'}$ распределены следующим образом.*

Доказательство. Для доказательства первой части утверждения достаточно сделать замену $\mathbf{\Omega}\to\mathbf{S}^{-1}\mathbf{\Omega}$. Тогда в силу вывода в приложении A.4.3 Якобиан замены будет равен $|\det(\mathbf{S})|^{-(n+1)/2}$:

$$p(\mathbf{S}^{-1}\mathbf{\Omega}) = p(\mathbf{\Omega})\det(\mathbf{S})|^{-(n+1)/2}2^{-\frac{n\nu}{2}} = \Gamma_n^{-1}\left(\frac{\nu}{2}\right)|\det(\mathbf{S}^{-1}\mathbf{\Omega})|^{\frac{\nu-n-1}{2}}\exp\left\{-\frac{1}{2}\,\mathrm{tr}(\mathbf{S}^{-1}\mathbf{\Omega})\right\} \tag{148}$$

Следовательно, $\mathbf{S}^{-1}\mathbf{\Omega}\sim\mathcal{W}_n(\nu,\mathbf{I}_n)$, и доказательство первой части утверждения прямо следует из разложения Бартлетта (утверждение 10).

Для доказательства второй части утверждения нужно расписать функцию распределения следующим образом ($\mathbf{S}^{-1}=\mathbf{MM'},\mathbf{\Omega}=\mathbf{LL'}$):

$$p(\mathbf{\Omega}) = 2^{-\frac{n\nu}{2}}\Gamma_n^{-1}\left(\frac{\nu}{2}\right)|\det(\mathbf{MM'})|^{-\frac{\nu}{2}}|\det(\mathbf{LL'})|^{\frac{\nu-n-1}{2}}\exp\left\{-\frac{1}{2}\,\mathrm{tr}(\mathbf{M'LL'M})\right\} \tag{149}$$

При этом

$$\mathrm{tr}(\mathbf{M}'\mathbf{L}\mathbf{L}'\mathbf{M}) = \sum_{i=1}^{n} \mathbf{m}_i'\mathbf{L}\mathbf{L}'\mathbf{m}_i = \sum_{i=1}^{n}\sum_{s=i}^{n} m_{is} \sum_{r=i}^{n} m_{ir} \sum_{j=1}^{\min(r,s)} l_{jr} l_{js} \tag{150}$$

$$|\det(\mathbf{M}\mathbf{M}')|^{-\frac{\nu}{2}} |\det(\mathbf{L}\mathbf{L}')|^{\frac{\nu-n-1}{2}} = |\det(\mathbf{M})|^{-\nu} |\det(\mathbf{L})|^{\nu-n-1} = \prod_{i=1}^{n} m_{ii}^{-\nu} l_{ii}^{\nu-n-1} \tag{151}$$

$$\Gamma_n\left(\frac{\nu}{2}\right) = \pi^{\frac{n(n-1)}{4}} \prod_{i=1}^{n} \Gamma\left(\frac{\nu-i+1}{2}\right) \tag{152}$$

$$\tag{153}$$

ЗАВЕРШИТЬ!!!ЗАВЕРШИТЬ!!!ЗАВЕРШИТЬ!!!ЗАВЕРШИТЬ!!! □

5 Байесовский анализ SVAR-моделей

Структурная авторегрессионная модель может быть записана в следующей форме:

$$\mathbf{YA} = \mathbf{ZB} + \mathcal{E} \tag{154}$$

В предположении, что $\mathcal{E} \sim \mathcal{MN}_{Tn}(\mathbf{0}, \mathbf{\Omega})$, функцию правдоподобия может быть записана в следующей форме (здесь $\mathbf{\Omega}^{-1} = \mathbf{AA}'$):

$$p(\mathbf{Y}|\mathbf{A},\mathbf{B}) = (2\pi)^{-\frac{T}{2}} |\det(\mathbf{A})|^T \exp\left\{-\frac{1}{2}\,\mathrm{tr}\left((\mathbf{YA}-\mathbf{ZB})(\mathbf{YA}-\mathbf{ZB})'\right)\right\} \tag{155}$$

Утверждение 13. *О видах представления функции правдоподобия для структурных векторных авторегрессий при условии нормального распределения ошибок. Функция правдоподобия для структурных векторных авторегрессий в случае нормального распределения ошибок при условии параметров* $\mathbf{A}, \mathbf{B}$ *может быть представлена одним из следующих способов:*

$$p(\mathbf{Y}|\mathbf{A},\mathbf{B}) = (2\pi)^{-\frac{T}{2}} |\det(\mathbf{A})|^T \exp\left\{-\frac{1}{2}\,\mathrm{tr}\left((\mathbf{YA}-\mathbf{ZB})(\mathbf{YA}-\mathbf{ZB})'\right)\right\} \tag{156}$$

$$p(\mathbf{Y}|\mathbf{A},\mathbf{B}) = (2\pi)^{-\frac{T}{2}} |\det(\mathbf{A})|^T \exp\left\{-\frac{1}{2}\,\mathrm{tr}\left((\mathbf{YA}-\mathbf{Z}\widehat{\mathbf{B}})'(\mathbf{YA}-\mathbf{Z}\widehat{\mathbf{B}})\right)\right\} \exp\left\{-\frac{1}{2}\,\mathrm{tr}\left((\mathbf{B}-\widehat{\mathbf{B}})'\mathbf{Z}'\mathbf{Z}(\mathbf{B}-\widehat{\mathbf{B}})\right)\right\} \tag{157}$$

$$p(\mathbf{Y}|\mathbf{A},\mathbf{B}) = (2\pi)^{-\frac{T}{2}} |\det(\mathbf{A})|^T \exp\left\{-\frac{1}{2}\,\mathrm{tr}\left((\mathrm{vec}(\mathbf{YA}) - (\mathbf{I}_n \otimes \mathbf{Z})\widehat{\mathbf{b}})'(\mathrm{vec}(\mathbf{YA}) - (\mathbf{I}_n \otimes \mathbf{Z})\widehat{\mathbf{b}})\right)\right\} \times \\ \times \exp\left\{-\frac{1}{2}\,\mathrm{tr}\left((\mathbf{b}-\widehat{\mathbf{b}})'(\mathbf{I}_n \otimes \mathbf{Z}'\mathbf{Z})(\mathbf{b}-\widehat{\mathbf{b}})\right)\right\} \tag{158}$$

где $\widehat{mbB} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{YA}$, $\mathbf{b} = \mathrm{vec}(\mathbf{B})$, $\widehat{\mathbf{b}} = \mathrm{vec}(\widehat{\mathbf{B}})$.

Доказательство. Для вывода формы (157) следует расписать функцию правдоподобия в терминах $\widehat{\mathbf{B}} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}\mathbf{A}$:

$$\begin{aligned}
&\operatorname{tr}\{(\mathbf{YA}-\mathbf{ZB})'(\mathbf{YA}-\mathbf{ZB})\} = \operatorname{tr}\left\{(\mathbf{YA}-\mathbf{ZB}+\mathbf{Z}\widehat{\mathbf{B}}-\mathbf{Z}\widehat{\mathbf{B}})'(\mathbf{YA}-\mathbf{ZB}+\mathbf{Z}\widehat{\mathbf{B}}-\mathbf{Z}\widehat{\mathbf{B}}\right\} = \\
&= \operatorname{tr}\left\{(\mathbf{YA}-\mathbf{Z}\widehat{\mathbf{B}})'(\mathbf{YA}-\mathbf{Z}\widehat{\mathbf{B}})\right\} + \operatorname{tr}\left\{(\widehat{\mathbf{B}}-\mathbf{B})'\mathbf{Z}'(\mathbf{YA}-\mathbf{Z}\widehat{\mathbf{B}})\right\} + \operatorname{tr}\left\{(\mathbf{YA}-\mathbf{Z}\widehat{\mathbf{B}})'\mathbf{Z}(\widehat{\mathbf{B}}-\mathbf{B})\right\} + \\
&+ \operatorname{tr}\left\{(\widehat{\mathbf{B}}-\mathbf{B})'\mathbf{Z}'\mathbf{Z}(\widehat{\mathbf{B}}-\mathbf{B})\right\} = \operatorname{tr}\left\{(\mathbf{YA}-\mathbf{Z}\widehat{\mathbf{B}})'(\mathbf{YA}-\mathbf{Z}\widehat{\mathbf{B}})\right\} + \operatorname{tr}\left\{(\widehat{\mathbf{B}}-\mathbf{B})'\mathbf{Z}'\mathbf{Z}(\widehat{\mathbf{B}}-\mathbf{B})\right\} + \\
&+ \operatorname{tr}\left\{(\widehat{\mathbf{B}}-\mathbf{B})'\underbrace{(\mathbf{Z}'\mathbf{YA}-\underbrace{\mathbf{Z}'\mathbf{Z}\widehat{\mathbf{B}}}_{\mathbf{X}'\mathbf{YA}}+\underbrace{\mathbf{Z}'(\mathbf{Y}-\mathbf{Z}\widehat{\mathbf{B}})}_{\mathbf{Z}'\mathbf{YA}-\mathbf{Z}'\mathbf{YA}})\mathbf{Z}'}_{\mathbf{0}}(\widehat{\mathbf{B}}-\mathbf{B})\right\}, \qquad (159)
\end{aligned}$$

что и требовалось доказать.

Для вывода формы (158) необходимо воспользоваться теоремами 1 и 2 и выразить $\mathbf{b} = \operatorname{vec}(\mathbf{B})$ и $\widehat{\mathbf{b}} = \operatorname{vec}(\widehat{\mathbf{B}})$:

$$\begin{aligned}
&\operatorname{tr}\left\{(\mathbf{YA}-\mathbf{Z}\widehat{\mathbf{B}})'(\mathbf{YA}-\mathbf{Z}\widehat{\mathbf{B}})\right\} + \operatorname{tr}\left\{(\widehat{\mathbf{B}}-\mathbf{B})'\mathbf{Z}'\mathbf{Z}(\widehat{\mathbf{B}}-\mathbf{B})\right\} = \\
&= \operatorname{vec}(\mathbf{YA}-\mathbf{Z}\widehat{\mathbf{B}})'\operatorname{vec}(\mathbf{YA}-\mathbf{Z}\widehat{\mathbf{B}}) + \operatorname{vec}\left(\mathbf{Z}(\mathbf{B}-\widehat{\mathbf{B}})\right)'\operatorname{vec}\left(\mathbf{Z}(\mathbf{B}-\widehat{\mathbf{B}})\right) = \\
&= (\operatorname{vec}(\mathbf{YA}) - (\mathbf{I}_n\otimes\mathbf{Z})\widehat{\mathbf{b}})'(\operatorname{vec}(\mathbf{YA}) - (\mathbf{I}_n\otimes\mathbf{Z})\widehat{\mathbf{b}}) + (\mathbf{b}-\widehat{\mathbf{b}})'(\mathbf{I}_n\otimes\mathbf{Z})(\mathbf{I}_n\otimes\mathbf{Z})(\mathbf{b}-\widehat{\mathbf{b}}) \qquad (160)
\end{aligned}$$

□

Утверждение 14. *О выводе апостериорных плотностей параметров SVAR-моделей вида Sims-Zha в случае неинформативных априорных распределений.*

Утверждение 15. *О выводе апостериорных плотностей параметров SVAR-моделей вида Waggoner-Zha.*

Утверждение 16. *О выводе апостериорных плотностей параметров SVAR-моделей вида Baumeister-Gamilton.*

Утверждение 17. *О выводе формы функции правдоподобия структурной векторной авторегрессии для случая нормальных ошибок и верхней треугольной матрицы* $\mathbf{A}$. *Я премного благодарен Ирине Дмитриевне Козловцевой за помощь в выводе формулы. Для случая, когда* $\mathbf{A}$ *— верхняя треугольная, функцию правдоподобия можно представить в других формах. Квадраты диагональных элементов этой матрицы обозначим* $d_{ii}, i = 1\dots n$. *Обозначим также* $\mathbf{S}(\mathbf{A},\mathbf{B}) = (\mathbf{Y}-\mathbf{B}\mathbf{A}^{-1})'(\mathbf{Y}-\mathbf{B}\mathbf{A}^{-1})$. *При этом пусть* $\widetilde{\mathbf{a}}_j = \operatorname{vechd}(\mathbf{A}, j)$. *При этом* $\widetilde{\mathbf{s}}_i = \operatorname{vechd}(\mathbf{S}, j)$, $\mathbf{S}_j$ *— верхняя левая подматрица матрицы* $\mathbf{S}$ *размерности* $j\times j$, *а* s_{ij} *— её* (i,j)*-элемент. Определим* $v_i = s_{ii} - \widetilde{\mathbf{s}}_i'\mathbf{S}_i^{-1}\widetilde{\mathbf{s}}_i$ *(по соглашению* $v_{11} = s_{11}$*). Тогда функцию прадоподобия можно записать в следующей форме:*

$$p(\mathbf{Y}|\mathbf{A},\mathbf{B}) = (2\pi)^{-\frac{T}{2}}\prod_{i=1}^{n}\left(|d_{ii}|^{\frac{T}{2}}\right)\exp\left\{-\frac{1}{2}\left(\sum_{i=1}^{n} d_{ii}v_i + \sum_{j=2}^{n}(\widetilde{\mathbf{a}}_j + d_{jj}^{\frac{1}{2}}\mathbf{S}_{j-1}^{-1}\widetilde{\mathbf{s}}_j)'\mathbf{S}_{j-1}(\widetilde{\mathbf{a}}_j + d_{jj}^{\frac{1}{2}}\mathbf{S}_{j-1}^{-1}\widetilde{\mathbf{s}}_j)\right)\right\} \qquad (161)$$

Доказательство. Данное утверждение напрямую будет следовать из утверждения 13, если доказать, что:

$$\operatorname{tr}(\mathbf{A}'\mathbf{S}\mathbf{A}) = \sum_{i=1}^{n} d_{ii}v_i + \sum_{j=2}^{n}(\widetilde{\mathbf{a}}_j + d_{jj}^{\frac{1}{2}}\mathbf{S}_{j-1}^{-1}\widetilde{\mathbf{s}}_j)'\mathbf{S}_{j-1}(\widetilde{\mathbf{a}}_j + d_{jj}^{\frac{1}{2}}\mathbf{S}_{j-1}^{-1}\widetilde{\mathbf{s}}_j) \qquad (162)$$

Приступим. Обозначим $\mathbf{a}_i$ как i-столбец матрицы $\mathbf{A}$ (здесь предполагается, что $\mathbf{g}_1 = \emptyset$). Обозначим также через $\mathbf{S}_{r\times k}$ верхнюю подматрицу матрицы $\mathbf{S}$ размерности $r \times k$. Тогда:

$$\mathrm{tr}(\mathbf{A}'\mathbf{S}\mathbf{A}) = \sum_{i=1}^{n} \mathbf{a}_i'\mathbf{S}\mathbf{a}_i = \sum_{i=1}^{n} \begin{bmatrix} \mathbf{g}_i' & d_{ii}^{\frac{1}{2}} & 0 & \dots & 0 \end{bmatrix} \mathbf{S} \begin{bmatrix} \widetilde{\mathbf{a}}_i \\ d_{ii}^{\frac{1}{2}} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \sum_{i=1}^{n} \begin{bmatrix} \begin{bmatrix} \widetilde{\mathbf{a}}_i' & d_{ii}^{\frac{1}{2}} \end{bmatrix} \mathbf{S}_{i\times n} \end{bmatrix} \begin{bmatrix} \widetilde{\mathbf{a}}_i \\ d_{ii}^{\frac{1}{2}} \\ 0 \\ \vdots \\ 0 \end{bmatrix} =$$

$$= \sum_{i=1}^{n} \begin{bmatrix} \widetilde{\mathbf{a}}_i' & d_{ii}^{\frac{1}{2}} \end{bmatrix} \mathbf{S}_i \begin{bmatrix} \widetilde{\mathbf{a}}_i \\ d_{ii}^{\frac{1}{2}} \end{bmatrix} = \sum_{i=1}^{n} \begin{bmatrix} \widetilde{\mathbf{a}}_i' & d_{ii}^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} \mathbf{S}_{i-1} & \widetilde{\mathbf{s}}_i' \\ \widetilde{\mathbf{s}}_i & s_{ii} \end{bmatrix} \begin{bmatrix} \widetilde{\mathbf{a}}_i \\ d_{ii}^{\frac{1}{2}} \end{bmatrix} = \sum_{i=1}^{n} \widetilde{\mathbf{a}}_i'(\mathbf{S}_{i-1}\mathbf{g}_i + d_{ii}^{\frac{1}{2}}\widetilde{\mathbf{s}}_i) + d_{ii}^{\frac{1}{2}}(\widetilde{\mathbf{s}}_i{}'\widetilde{\mathbf{a}}_i + d_{ii}^{\frac{1}{2}} s_{ii}) =$$

$$= d_{11}s_{11} + \sum_{j=2}^{n} d_{jj}s_{jj} + \widetilde{\mathbf{a}}_j'\mathbf{S}_{j-1}\widetilde{\mathbf{a}}_j + 2d_{jj}^{\frac{1}{2}}\widetilde{\mathbf{a}}_j'\widetilde{\mathbf{s}}_j = d_{11}s_{11} + \sum_{j=2}^{n} d_{jj}s_{jj} + \widetilde{\mathbf{a}}_j'\mathbf{S}_{j-1}\widetilde{\mathbf{a}}_j + 2d_{jj}^{\frac{1}{2}}\widetilde{\mathbf{a}}_j'\mathbf{S}_{j-1}\mathbf{S}_{j-1}^{-1}\widetilde{\mathbf{s}}_j =$$

$$= d_{11}s_{11} + \sum_{j=2}^{n} d_{jj}s_{jj} + (\widetilde{\mathbf{a}}_j + d_{jj}^{\frac{1}{2}}\mathbf{S}_{j-1}^{-1}\widetilde{\mathbf{s}}_j)'\mathbf{S}_{j-1}(\widetilde{\mathbf{a}}_j + d_{jj}^{\frac{1}{2}}\mathbf{S}_{j-1}^{-1}\widetilde{\mathbf{s}}_j) - d_{jj}\widetilde{\mathbf{s}}_j'\mathbf{S}_{j-1}^{-1}\widetilde{\mathbf{s}}_j =$$

$$= \sum_{i=1}^{n} d_{ii}v_i + \sum_{j=2}^{n} (\widetilde{\mathbf{a}}_j + d_{jj}^{\frac{1}{2}}\mathbf{S}_{j-1}^{-1}\widetilde{\mathbf{s}}_j)'\mathbf{S}_{j-1}(\widetilde{\mathbf{a}}_j + d_{jj}^{\frac{1}{2}}\mathbf{S}_{j-1}^{-1}\widetilde{\mathbf{s}}_j) \quad (163)$$

□

Утверждение 18. ***О выводе апостериорных плотностей параметров SVAR-моделей вида George et. al. Случай поиска ограничений только на верхнюю треугольную матрицу A.*** *Рассмотрим функцию правдоподобия в форме, выведенной в утверждении 17. При этом поиск ограничений ведётся по m элементам из $n(n-1)/2$ недиагональных элементов матрицы $\mathbf{A}$. При этом $\mathbf{\Phi} = \mathbf{B}\mathbf{A}^{-1}$, $\boldsymbol{\varphi} = \mathrm{vec}(\mathbf{\Phi})$, Для следующего набора априорных распределений:*

$$\mathbb{P}(\omega_{ij} = 1) = p_{ij}, \forall i, j: \quad 1 \leqslant i < j < n \tag{164}$$
$$\boldsymbol{\varphi}|\mathbf{A}, \boldsymbol{\omega} \sim \mathcal{N}_{nk}(\underline{\boldsymbol{\varphi}}, \underline{\mathbf{V}}) \tag{165}$$
$$d_{ii}|\boldsymbol{\varphi}, \boldsymbol{\omega} \sim \mathcal{G}(\underline{\alpha}_i^{\omega}, \underline{\beta}_i^{\omega}) \tag{166}$$
$$\widetilde{\mathbf{a}}_i|\boldsymbol{\varphi}, d_{ii}, \boldsymbol{\omega} \sim \mathcal{N}_{i-1}(\mathbf{0}, \underline{\mathbf{H}}_i^{\omega}\underline{\mathbf{R}}_i^{\omega}\underline{\mathbf{H}}_i^{\omega}) \tag{167}$$
$$\underline{\mathbf{H}}_i^{\omega} = \mathrm{diag}([h_{12}^{\omega}, \ldots, h_{(n-1)n}^{\omega}]) \tag{168}$$
$$h_{ij}^{\omega} = \begin{cases} c \to 0, \textit{ если } \omega_s(\widetilde{(a)}_{ij}) = 1, \\ d \to \infty, \textit{ если } \omega_s(\widetilde{(a)}_{ij}) = 0 \end{cases} \tag{169}$$

апостериорные плотности вероятности параметров выглядят следующим образом:

$$d_{ii}|\boldsymbol{\varphi}, \boldsymbol{\omega}; \mathbf{X} \sim \mathcal{G}(\overline{\alpha}_i^{\omega}, \overline{\beta}_i^{\omega}) \tag{170}$$
$$\widetilde{\mathbf{a}}_i|\boldsymbol{\varphi}, \boldsymbol{\omega}; \mathbf{X} \sim \mathcal{N}_{i-1}(\overline{\mathbf{a}}_i, \overline{\mathbf{M}}_i^{\omega}) \tag{171}$$
$$\boldsymbol{\varphi}|\mathbf{A}, \boldsymbol{\omega}; \mathbf{X} \sim \mathcal{N}_{nk}(\overline{\boldsymbol{\varphi}}, \overline{\mathbf{V}}) \tag{172}$$

где

$$\overline{\mathbf{V}} = \left(\mathbf{A}\mathbf{A}' \otimes \mathbf{Z}'\mathbf{Z} + \underline{\mathbf{V}}^{-1}\right)^{-1} \tag{173}$$
$$\overline{\boldsymbol{\varphi}} = \overline{\mathbf{V}}\left((\mathbf{A}\mathbf{A}' \otimes \mathbf{Z}'\mathbf{Z})\,\widehat{P\boldsymbol{\varphi}} + \underline{\mathbf{V}}^{-1}\underline{\boldsymbol{\varphi}}\right) \tag{174}$$
$$\overline{\mathbf{M}}_i^{\omega} = \left((\underline{\mathbf{H}}_j^{\omega}\underline{\mathbf{R}}_j^{\omega}\underline{\mathbf{H}}_j^{\omega})^{-1} + \mathbf{S}_{j-1}\right)^{-1} \tag{175}$$
$$\widehat{\boldsymbol{\varphi}} = \mathrm{vec}((\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}) = \left[\mathbf{I}_n \otimes ((\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}')\right]\mathrm{vec}(\mathbf{Y}) \tag{176}$$
$$\overline{\mathbf{a}} = -d_{jj}^{\frac{1}{2}}\overline{\mathbf{M}}_j^{\omega}\widetilde{\mathbf{s}}_j \tag{177}$$
$$\overline{\alpha}_i^{\omega} = \underline{\alpha}_i^{\omega} + \frac{T}{2} \tag{178}$$
$$\overline{\beta}_i^{\omega} = \begin{cases} \underline{\beta}_i^{\omega} + \frac{1}{2}s_{ii}, \textit{ если } i = 1 \\ \underline{\beta}_i^{\omega} + \frac{1}{2}s_{ii} + \frac{1}{2}\widetilde{\mathbf{s}}_i'\overline{\mathbf{M}}_i^{\omega}\widetilde{\mathbf{s}}_i, \textit{ если } i \neq 1 \end{cases} \tag{179}$$

Вектор-индикатор ограничений имеет распределение Бернулли:

$$\omega_{ij}|\boldsymbol{\varphi}, \mathbf{D}, \boldsymbol{\omega}_k; \mathbf{X} \sim \mathrm{Bernoulli}\left(\frac{1}{1+U}\right) \tag{180}$$

В общем случае:

$$U = \frac{p(\widetilde{\mathbf{a}}_j|\boldsymbol{\varphi}, \mathbf{D}, \boldsymbol{\omega}_{-s}, \omega_s(\widetilde{a}_{ij}) = 0)}{p(\widetilde{\mathbf{a}}_j|\boldsymbol{\varphi}, \mathbf{D}, \boldsymbol{\omega}_{-s}, \omega_s(\widetilde{a}_{ij}) = 1)}\frac{p(\omega_s(\widetilde{a}_{ij}) = 0)}{p(\omega_s(\widetilde{a}_{ij}) = 1)} = \frac{u_{ij}^0}{u_{ij}^1} \tag{181}$$
$$u_{ij}^1 = |\det\left(\underline{\mathbf{H}}_i^1\right)|^{-2(j-1)}|\det\left(\underline{\mathbf{R}}_i^1\right)|^{-(j-1)}\exp\left\{-\frac{1}{2}\widetilde{\mathbf{a}}_j'\left(\underline{\mathbf{H}}_i^1\underline{\mathbf{R}}_i^1\underline{\mathbf{H}}_i^1\right)^{-1}\widetilde{\mathbf{a}}_j\right\}p_{ij} \tag{182}$$
$$u_{ij}^0 = |\det\left(\underline{\mathbf{H}}_i^0\right)|^{-2(j-1)}|\det\left(\underline{\mathbf{R}}_i^0\right)|^{-(j-1)}\exp\left\{-\frac{1}{2}\widetilde{\mathbf{a}}_j'\left(\underline{\mathbf{H}}_i^0\underline{\mathbf{R}}_i^0\underline{\mathbf{H}}_i^0\right)^{-1}\widetilde{\mathbf{a}}_j\right\}(1-p_{ij}) \tag{183}$$

В случае, если $\mathbf{R}_j = \mathbf{I}_{j-1}$*:*

$$u_{ij}^1 = \frac{1}{\underline{h}_{ij}^1} \exp\left\{-\frac{1}{2}\left(\frac{\widetilde{a}_{ij}}{\underline{h}_{ij}^1}\right)^2\right\} p_{ij} \tag{184}$$

$$u_{ij}^0 = \frac{1}{\underline{h}_{ij}^0} \exp\left\{-\frac{1}{2}\left(\frac{\widetilde{a}_{ij}}{2\underline{h}_{ij}^0}\right)^2\right\} (1-p_{ij}) \tag{185}$$

Доказательство. Для начала распишем совместную функцию плотности параметров $\{d_{ii}\}_{i=1}^n$ и $\{\widetilde{\mathbf{a}}_j\}_{j=2}^n$ при условии данных $\mathbf{X}$ и ограничений $\boldsymbol{\omega}$:

$$\begin{aligned}
&p(\{d_{ii}\}_{i=1}^n, \{\widetilde{\mathbf{a}}_j\}_{j=2}^n | \boldsymbol{\varphi}, \boldsymbol{\omega}; \mathbf{X}) \propto \\
&\propto \left[\prod_{i=1}^n \left((d_{ii})^{\underline{\alpha}_i^\omega - 1} \exp\left\{-\underline{\beta}_i^\omega d_{ii}\right\}\right)\right] \left[\prod_{j=2}^n \exp\left\{-\frac{1}{2}\widetilde{\mathbf{a}}_j'(\underline{\mathbf{H}}_j^\omega \underline{\mathbf{R}}_j^\omega \underline{\mathbf{H}}_j^\omega)^{-1}\widetilde{\mathbf{a}}_j\right\}\right] \prod_{i=1}^n \left(|d_{ii}|^{\frac{T}{2}}\right) \exp\left\{-\frac{1}{2}\sum_{i=1}^n d_{ii} v_i\right\} \times \\
&\times \exp\left\{-\frac{1}{2}\sum_{j=2}^n (\widetilde{\mathbf{a}}_j + d_{jj}^{\frac{1}{2}}\mathbf{S}_{j-1}^{-1}\widetilde{\mathbf{s}}_j)'\mathbf{S}_{j-1}(\widetilde{\mathbf{a}}_j + d_{jj}^{\frac{1}{2}}\mathbf{S}_{j-1}^{-1}\widetilde{\mathbf{s}}_j)\right\} = \left[(d_{11})^{\overline{\alpha}_1^\omega - 1} \exp\left\{-(\underline{\beta}_1^\omega + \frac{1}{2}s_{11})d_{11}\right\}\right] \times \\
&\times \left[\prod_{j=2}^n \left((d_{ii})^{\underline{\alpha}_j^\omega + \frac{T}{2} - 1} \exp\left\{-(\underline{\beta}_j^\omega + v_i)d_{ii}\right\}\right)\right] \times \\
&\times \left[\prod_{j=2}^n \exp\left\{-\frac{1}{2}(\widetilde{\mathbf{a}}_j' \underbrace{((\underline{\mathbf{H}}_j^\omega \underline{\mathbf{R}}_j^\omega \underline{\mathbf{H}}_j^\omega)^{-1} + \mathbf{S}_{j-1})}_{\left(\overline{\mathbf{M}}_j^\omega\right)^{-1}} \widetilde{\mathbf{a}}_j + \underbrace{2d_{jj}^{\frac{1}{2}}\widetilde{\mathbf{a}}_j'\widetilde{\mathbf{s}}_j}_{2d_{jj}^{\frac{1}{2}}\widetilde{\mathbf{a}}_j'\left(\overline{\mathbf{M}}_j^\omega\right)^{-1}\overline{\mathbf{M}}_j^\omega\widetilde{\mathbf{s}}_j} + d_{jj}\widetilde{\mathbf{s}}_j'\mathbf{S}_{j-1}\widetilde{\mathbf{s}}_j)\right\}\right] = \\
&= \left[(d_{11})^{\overline{\alpha}_1^\omega - 1} \exp\left\{-(\underline{\beta}_1^\omega + \frac{1}{2}s_{11})d_{11}\right\}\right] \left[\prod_{j=2}^n \left((d_{ii})^{\overline{\alpha}_j^\omega - 1} \exp\left\{-\underbrace{\left(\underline{\beta}_j^\omega + \frac{v_i}{2} + \frac{1}{2}\widetilde{\mathbf{s}}_j'\mathbf{S}_{j-1}\widetilde{\mathbf{s}}_j - \frac{1}{2}\widetilde{\mathbf{s}}_j'\left(\overline{\mathbf{M}}_j^\omega\right)^{-1}\widetilde{\mathbf{s}}_j\right)}_{\overline{\beta}_i^\omega = \underline{\beta}_i^\omega + \frac{1}{2}\left(s_{ii} + \widetilde{\mathbf{s}}_j'\left(\overline{\mathbf{M}}_j^\omega\right)^{-1}\widetilde{\mathbf{s}}_j\right)} d_{ii}\right\}\right)\right] \\
&\times \left[\prod_{j=2}^n \exp\left\{-\frac{1}{2}(\widetilde{\mathbf{a}}_j + \underbrace{d_{jj}^{\frac{1}{2}}\overline{\mathbf{M}}_j^\omega\widetilde{\mathbf{s}}_j}_{\overline{\mathbf{a}}_j})'\left(\overline{\mathbf{M}}_j^\omega\right)^{-1}(\widetilde{\mathbf{a}}_j + d_{jj}^{\frac{1}{2}}\overline{\mathbf{M}}_j^\omega\widetilde{\mathbf{s}}_j)\right\}\right] = \\
&= \left[\prod_{i=1}^n d_{ii}^{\overline{\alpha}_i^\omega} \exp\left\{-\overline{\beta}_i^\omega\right\}\right] \left[\prod_{j=2}^n \exp\left\{-\frac{1}{2}(\widetilde{\mathbf{a}}_j - \overline{\mathbf{a}}_j)'\left(\overline{\mathbf{M}}_j^\omega\right)^{-1}(\widetilde{\mathbf{a}}_j - \overline{\mathbf{a}}_j)\right\}\right]
\end{aligned} \tag{186}$$

Для того, чтобы выразить апостериорную функцию плотности $\boldsymbol{\varphi}$ при условии параметров $\mathbf{A}$ и $\boldsymbol{\omega}$, необходимо расписать следующую функцию:

$$p(\boldsymbol{\varphi}|\mathbf{A}, \boldsymbol{\omega}; \mathbf{X}) \propto \exp\left\{-\frac{1}{2}(\boldsymbol{\varphi} - \underline{\boldsymbol{\varphi}})'\underline{\mathbf{V}}^{-1}(\boldsymbol{\varphi} - \underline{\boldsymbol{\varphi}})\right\} \exp\left\{-\frac{1}{2}(\mathbf{Y} - \mathbf{Z\Phi})\mathbf{AA}'(\mathbf{Y} - \mathbf{Z\Phi})'\right\} \tag{187}$$

что аналогично утверждению 3. Следовательно, можно получить апостериорную плотность, аналогичную описанному в настоящем утверждении.

Для того, чтобы выразить апостериорную при условии остальных параметров функцию плотности параметра $\boldsymbol{\omega}$, контролирующего ограничения на матрицу $\mathbf{A}$, рассмотрим апостериорную условную функцию плотности каждого отдельного его элемента:

$$
\begin{aligned}
p(\omega_s(\widetilde{a}_{ij})=1|\boldsymbol{\varphi},\mathbf{D},\widetilde{\mathbf{A}},\boldsymbol{\omega}_{-s};\mathbf{X}) &= \frac{p(\omega_s(\widetilde{a}_{ij})=1)p(\boldsymbol{\varphi},\mathbf{D},\widetilde{\mathbf{A}},\boldsymbol{\omega}_{-s};\mathbf{X}|\omega_s(\widetilde{a}_{ij})=1)}{p(\boldsymbol{\varphi},\mathbf{D},\widetilde{\mathbf{A}},\boldsymbol{\omega}_{-s};\mathbf{X})} = \\
&= \frac{p(\omega_s(\widetilde{a}_{ij})=1)p(\boldsymbol{\varphi},\mathbf{D},\widetilde{\mathbf{A}},\boldsymbol{\omega}_{-s};\mathbf{X}|\omega_s(\widetilde{a}_{ij})=1)}{p(\omega_s(\widetilde{a}_{ij})=1)p(\boldsymbol{\varphi},\mathbf{D},\widetilde{\mathbf{A}},\boldsymbol{\omega}_{-s};\mathbf{X}|\omega_s(\widetilde{a}_{ij})=1)+p(\omega_s(\widetilde{a}_{ij})=0)p(\boldsymbol{\varphi},\mathbf{D},\widetilde{\mathbf{A}},\boldsymbol{\omega}_{-s};\mathbf{X}|\omega_s(\widetilde{a}_{ij})=0)}
\end{aligned} \quad (188)
$$

Следует обратить внимание на два факта. Во-первых, функция плотности парметров при условии ограничения на конкретный элемент может быть расписана следующим образом:

$$
p(\boldsymbol{\varphi},\mathbf{D},\widetilde{\mathbf{A}},\boldsymbol{\omega}_{-s};\mathbf{X}|\omega_s(\widetilde{a}_{ij})) = p(\boldsymbol{\omega}_{-s}|\omega_s(\widetilde{a}_{ij}))p(\boldsymbol{\varphi}|\boldsymbol{\omega})p(\mathbf{D}|\boldsymbol{\varphi},\boldsymbol{\omega})p(\widetilde{\mathbf{A}}|\mathbf{D},\boldsymbol{\varphi},\boldsymbol{\omega})p(\mathbf{X}|\widetilde{\mathbf{A}},\mathbf{D},\boldsymbol{\varphi},\boldsymbol{\omega}) \quad (189)
$$

При этом если переменные $\mathbf{Z}$ являются слабо экзогенными, то $p(\mathbf{X}|\boldsymbol{\varphi},\widetilde{\mathbf{A}},\mathbf{D},\boldsymbol{\gamma}) = p(\mathbf{Y}|\boldsymbol{\varphi},\widetilde{\mathbf{A}},\mathbf{D},\boldsymbol{\gamma};\mathbf{Z})p(\mathbf{Z})$.

Во-вторых, для случаев $\omega_s(\widetilde{a}_{ij})=1$ и $\omega_s(\widetilde{a}_{ij})=0$ различия заключаются лишь в априорной функции распределения конкретного элемента матрицы $\widetilde{\mathbf{A}}$. Это свзано с тем, что набор ограничений ω не ограничивает в явном виде значения функции правдоподобия, но ограничивает исключительно параметры априорных распределений диагональных элементов, в связи с чем $p(\mathbf{Y}|\widetilde{\mathbf{A}},\mathbf{D},\varphi,\omega;\mathbf{Z}) = p(\mathbf{Y}|\widetilde{\mathbf{A}},\mathbf{D},\varphi;\mathbf{Z})$. В силу независимости строк матрицы $\widetilde{\mathbf{A}}$ уравнение (188) можно переписать как:

$$
\begin{aligned}
&p(\omega_s(\widetilde{a}_{ij})=1|\boldsymbol{\varphi},\mathbf{D},\widetilde{\mathbf{A}},\boldsymbol{\omega}_{-s};\mathbf{X}) = \\
&\quad= \frac{p(\omega_s(\widetilde{a}_{ij})=1)p(\widetilde{\mathbf{a}}_j|\mathbf{D},\boldsymbol{\varphi},\boldsymbol{\omega}_{-s}|\omega_s(\widetilde{a}_{ij})=1)}{p(\omega_s(\widetilde{a}_{ij})=1)p(\widetilde{\mathbf{a}}_j|\mathbf{D},\boldsymbol{\varphi},\boldsymbol{\omega}_{-s}|\omega_s(\widetilde{a}_{ij})=1)+p(\omega_s(\widetilde{a}_{ij})=0)p(\widetilde{\mathbf{a}}_j|\mathbf{D},\boldsymbol{\varphi},\boldsymbol{\omega}_{-s}|\omega_s(\widetilde{a}_{ij})=0)}
\end{aligned} \quad (190)
$$

Более того, можно обратить внимание, что каждое из этих распределений можно нормировать на константу, что даёт доказательство этой части утверждения.

В случае, если $\mathbf{R}_j = \mathbf{I}_{j-1}$, эту же формулу можно переписать в терминах функции плотности конкретного (i,j)-элемента:

$$
\begin{aligned}
&p(\omega_s(\widetilde{a}_{ij})=1|\boldsymbol{\varphi},\mathbf{D},\widetilde{\mathbf{A}},\boldsymbol{\omega}_{-s};\mathbf{X}) = \\
&\quad= \frac{p(\omega_s(\widetilde{a}_{ij})=1)p(\widetilde{a}_{ij}|\mathbf{D},\boldsymbol{\varphi},\boldsymbol{\omega}_{-s}|\omega_s(\widetilde{a}_{ij})=1)}{p(\omega_s(\widetilde{a}_{ij})=1)p(\widetilde{a}_{ij}|\mathbf{D},\boldsymbol{\varphi},\boldsymbol{\omega}_{-s}|\omega_s(\widetilde{a}_{ij})=1)+p(\omega_s(\widetilde{a}_{ij})=0)p(\widetilde{a}_{ij}|\mathbf{D},\boldsymbol{\varphi},\boldsymbol{\omega}_{-s}|\omega_s(\widetilde{a}_{ij})=0)}
\end{aligned} \quad (191)
$$

Сокращение в функциях плотности $p(\widetilde{a}_{ij}|\mathbf{D},\boldsymbol{\varphi},\boldsymbol{\omega}_{-s}|\omega_s(\widetilde{a}_{ij}))$ одинаковых констант даёт доказательство утверждения. □

Утверждение 19. *О выводе апостериорных плотностей параметров SVAR-моделей вида George et. al. Случай поиска ограничений только на матрицу $\boldsymbol{\Phi}$ при условии верхней треугольной матрицы $\mathbf{A}$.* *Рассмотрим функцию правдоподобия в форме, выведенной в утверждении 17. При этом поиск ограничений ведётся по f элементам из nk элементов матрицы $\boldsymbol{\Phi} = \mathbf{B}\mathbf{A}^{-1}$. Вектор параметров при лаговых переменных $\boldsymbol{\varphi} = \mathrm{vec}(\boldsymbol{\Phi})$ может быть разделён на два подвектора параметров: вектора параметров, по которому производится поиск ограничений $\boldsymbol{\varphi}_R$ размерности $m \times 1$ и вектора параметров $\boldsymbol{\varphi}_{UR}$ размерности $nk - m \times 1$, имеющих одинаковое априорное распределение для любого набора ограничений. При этом априори авторы статьи предполагают, что $\boldsymbol{\varphi}_R \perp \boldsymbol{\varphi}_{UR}$ (это связано с тем, что а приори предполагается, что $\gamma_{ij} = 0$ соответствует ограничению, а $\gamma_{ij} = 1$ соответствует неинформативным априорным предположениям).*

Помимо этого предположения можо рассматривать аналог g-prior Зельнера для векторных авторегрессий для случая $\boldsymbol{\gamma}_{i:s} = 0$: $[\mathbf{V}^{\gamma} = \mathbf{K}^{-1} \otimes (\mathbf{Z}'\mathbf{Z})]_{(s-i)\times(s-i)}$. Однако в этом случае не обязательно $\boldsymbol{\varphi}_R \perp \boldsymbol{\varphi}_{UR}$, и следует рассматривать совместную функцию плотности $\boldsymbol{\varphi} = [\boldsymbol{\varphi}'_R, \boldsymbol{\varphi}'_{UR}]'$. В случае, если $\boldsymbol{\gamma}_{i:s} = 1$, логично предположить независимость ограниченных параметров. Таким образом, для *следующего набора априорных распределений:*

$$\mathbb{P}(\gamma_{ij}=1)=q_{ij}, \forall i,j: \;\; 1 \leqslant i < j < n \tag{192}$$

$$\boldsymbol{\varphi}_R|\mathbf{A},\boldsymbol{\gamma} \sim \mathcal{N}_{nk}(\underline{\boldsymbol{\varphi}}_R^{\gamma}, \underline{\mathbf{G}}_R^{\gamma}\underline{\mathbf{P}}_R^{\gamma}\underline{\mathbf{G}}_R^{\gamma}) \tag{193}$$

$$\boldsymbol{\varphi}_{UR}|\mathbf{A},\boldsymbol{\gamma} \sim \mathcal{N}_{nk}(\underline{\boldsymbol{\varphi}}_{UR}, \underline{\mathbf{V}}_{UR}^{\gamma}) \tag{194}$$

$$\boldsymbol{\varphi}|\mathbf{A},\boldsymbol{\gamma} \sim \mathcal{N}_{nk}(\underline{\boldsymbol{\varphi}}^{\gamma}, \underline{\mathbf{V}}^{\gamma}) \tag{195}$$

$$d_{ii}|\boldsymbol{\varphi},\boldsymbol{\gamma} \sim \mathcal{G}(\underline{\alpha}_i, \underline{\beta}_i) \tag{196}$$

$$\widetilde{\mathbf{a}}_i|\boldsymbol{\varphi}, d_{ii}, \boldsymbol{\gamma} \sim \mathcal{N}_{i-1}(\mathbf{0}, \underline{\mathbf{H}}_i\underline{\mathbf{R}}_i\underline{\mathbf{H}}_i) \tag{197}$$

$$\underline{\mathbf{G}}_R^{\gamma} = \mathrm{diag}([g_{12}^{\gamma}, \ldots, g_m^{\gamma}]) \tag{198}$$

$$g_{ij}^{\gamma} = \begin{cases} c \to 0, \textit{если } \gamma_s(\widetilde{\boldsymbol{\varphi}}_{ij}) = 1, \\ d \to \infty, \textit{если } \gamma_s(\widetilde{\boldsymbol{\varphi}}_{ij}) = 0 \end{cases} \tag{199}$$

$$\mathbf{V}_{\mathbf{R}}^{\gamma} = \begin{array}{c c} & \begin{array}{ccc} \boldsymbol{\gamma}=1 & \boldsymbol{\gamma}=0 & UR \end{array} \\ \begin{array}{c} \boldsymbol{\gamma}=1 \\ \boldsymbol{\gamma}=0 \\ UR \end{array} & \begin{bmatrix} \underline{\mathbf{G}}_R^{\gamma}\underline{\mathbf{P}}_R^{\gamma}\underline{\mathbf{G}}_R^{\gamma} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \underline{\mathbf{V}}_0^{\gamma} & \underline{\mathbf{V}}_{0.UR}^{\gamma} \\ \mathbf{0} & \underline{\mathbf{V}}_{UR.0}^{\gamma} & \underline{\mathbf{V}}_{UR}^{\gamma} \end{bmatrix} \end{array} \tag{200}$$

$$\underline{\boldsymbol{\varphi}}_R^0 = \mathbf{0} \tag{201}$$

$$\underline{\boldsymbol{\varphi}}_R^1 = \underline{\boldsymbol{\varphi}}_R^1 \tag{202}$$

апостериорные плотности вероятности параметров выглядят следующим образом:

$$\boldsymbol{\varphi}|\mathbf{A},\boldsymbol{\gamma};\mathbf{X} \sim \mathcal{N}_{nk}(\overline{\boldsymbol{\varphi}}^{\gamma}, \overline{\mathbf{V}}^{\gamma}) \tag{203}$$

$$d_{ii}|\boldsymbol{\gamma};\mathbf{X} \sim \mathcal{G}(\overline{\alpha}_i, \overline{\beta}_i) \tag{204}$$

$$\widetilde{\mathbf{a}}_i|\boldsymbol{\gamma};\mathbf{X} \sim \mathcal{N}_{i-1}(\overline{\mathbf{a}}_i, \overline{\mathbf{M}}_i) \tag{205}$$

$$\tag{206}$$

где

$$\overline{\mathbf{V}}^{\gamma} = \left(\mathbf{A}\mathbf{A}' \otimes \mathbf{Z}'\mathbf{Z} + (\underline{\mathbf{V}}^{\gamma})^{-1}\right)^{-1} \tag{207}$$

$$\overline{\boldsymbol{\varphi}}^{\gamma} = \overline{\mathbf{V}}^{\gamma}\left((\mathbf{A}\mathbf{A}' \otimes \mathbf{Z}'\mathbf{Z})\,\widehat{\boldsymbol{\varphi}} + (\underline{\mathbf{V}}^{\gamma})^{-1}\underline{\boldsymbol{\varphi}^{\gamma}}\right) \tag{208}$$

$$\overline{\mathbf{M}}_i = \left((\underline{\mathbf{H}}_j\underline{\mathbf{R}}_j\underline{\mathbf{H}}_j)^{-1} + \mathbf{S}_{j-1}\right)^{-1} \tag{209}$$

$$\overline{\mathbf{a}} = -d_{jj}^{\frac{1}{2}}\overline{\mathbf{M}}_j\widetilde{\mathbf{s}}_j \tag{210}$$

$$\overline{\alpha}_i = \underline{\alpha}_i + \frac{T}{2} \tag{211}$$

$$\overline{\beta}_i = \begin{cases} \underline{\beta}_i + \frac{1}{2}s_{ii}, \; \textit{если } i = 1 \\ \underline{\beta}_i + \frac{1}{2}s_{ii} + \frac{1}{2}\widetilde{\mathbf{s}}_i'\overline{\mathbf{M}}_i\widetilde{\mathbf{s}}_i, \; \textit{если } i \neq 1 \end{cases} \tag{212}$$

Вероятность ограничения может быть выражена как:

$$\gamma_{ij}|\boldsymbol{\varphi},\mathbf{D},\boldsymbol{\gamma}_k;\mathbf{X} \sim \mathrm{Bernoulli}\left(\frac{1}{1+V}\right) \tag{213}$$

В общем случае:

$$V=\frac{p(\boldsymbol{\varphi}|\mathbf{D},\widetilde{\mathbf{A}},\boldsymbol{\gamma}_{-s},\gamma_s(\varphi_{ij})=0)}{p(\boldsymbol{\varphi}|\mathbf{D},\widetilde{\mathbf{A}},\boldsymbol{\gamma}_{-s},\gamma_s(\varphi_{ij})=1)}\frac{p(\gamma_s(\varphi_{ij})=0)}{p(\gamma_s(\varphi_{ij})=1)}=\frac{v_{ij}^0}{v_{ij}^1} \tag{214}$$

$$v_{ij}^1=|\det(\underline{\mathbf{V}}^1)|^{-nk}\exp\left\{-\frac{1}{2}(\boldsymbol{\varphi}-\underline{\boldsymbol{\varphi}}^1)'\left(\underline{\mathbf{V}}^1\right)^{-1}(\boldsymbol{\varphi}-\underline{\boldsymbol{\varphi}}^1)\right\}p_{ij} \tag{215}$$

$$v_{ij}^0=|\det(\underline{\mathbf{V}}^0)|^{-nk}\exp\left\{-\frac{1}{2}(\boldsymbol{\varphi}-\underline{\boldsymbol{\varphi}}^0)'\left(\underline{\mathbf{V}}^0\right)^{-1}(\boldsymbol{\varphi}-\underline{\boldsymbol{\varphi}}^0)\right\}(1-p_{ij}) \tag{216}$$

В случае, если $\mathbf{P}_j=\mathbf{I}_{j-1}$ *и* $\mathbf{V}_{0.UR}=\mathbf{0}$*:*

$$v_{ij}^1=\frac{1}{\underline{g}_{ij}^1}\exp\left\{-\frac{1}{2}\left(\frac{\varphi_{ij}}{\underline{g}_{ij}^1}\right)^2\right\}p_{ij} \tag{217}$$

$$v_{ij}^0=\frac{1}{\underline{g}_{ij}^0}\exp\left\{-\frac{1}{2}\left(\frac{\varphi_{ij}}{\underline{g}_{ij}^0}\right)^2\right\}(1-p_{ij}) \tag{218}$$

Доказательство. Все апостериорные распределения выводятся аналогично распределениям из утверждения 18. Поэтому остановимся на выводе апостериорных вероятностей индикаторов ограничений. Распишем апостериорную вероятность индикаторной функции:

$$\begin{aligned}p(\gamma_s(\varphi_{ij})=1|\boldsymbol{\varphi},\mathbf{D},\widetilde{\mathbf{A}},\boldsymbol{\omega}_{-s};\mathbf{X})&=\frac{p(\gamma_s(\varphi_{ij})=1)p(\boldsymbol{\varphi},\mathbf{D},\widetilde{\mathbf{A}},\boldsymbol{\gamma}_{-s};\mathbf{X}|\gamma_s(\varphi_{ij})=1)}{p(\boldsymbol{\varphi},\mathbf{D},\widetilde{\mathbf{A}},\boldsymbol{\gamma}_{-s};\mathbf{X})}=\\&=\frac{p(\gamma_s(\varphi_{ij})=1)p(\boldsymbol{\varphi},\mathbf{D},\widetilde{\mathbf{A}},\boldsymbol{\gamma}_{-s};\mathbf{X}|\gamma_s(\varphi_{ij})=1)}{p(\gamma_s(\varphi_{ij})=1)p(\boldsymbol{\varphi},\mathbf{D},\widetilde{\mathbf{A}},\boldsymbol{\gamma}_{-s};\mathbf{X}|\gamma_s(\varphi_{ij})=1)+p(\gamma_s(\varphi_{ij})=0)p(\boldsymbol{\varphi},\mathbf{D},\widetilde{\mathbf{A}},\boldsymbol{\gamma}_{-s};\mathbf{X}|\gamma_s(\varphi_{ij})=0)}\end{aligned} \tag{219}$$

Распишем совместную плотность с помощью цепной формулы:

$$p(\boldsymbol{\varphi},\mathbf{D},\widetilde{\mathbf{A}},\boldsymbol{\gamma}_{-s};\mathbf{X}|\gamma_s(\varphi_{ij})=j))=p(\boldsymbol{\gamma}_{-s}|\gamma_s(\varphi_{ij})=j)p(\mathbf{D}|\boldsymbol{\gamma})p(\widetilde{\mathbf{A}}|\mathbf{D},\boldsymbol{\gamma})p(\boldsymbol{\varphi}|\widetilde{\mathbf{A}},\mathbf{D},\boldsymbol{\gamma})p(\mathbf{X}|\boldsymbol{\varphi},\widetilde{\mathbf{A}},\mathbf{D},\boldsymbol{\gamma}) \tag{220}$$

При этом если переменные $\mathbf{Z}$ являются слабо экзогенными, то $p(\mathbf{X}|\boldsymbol{\varphi},\widetilde{\mathbf{A}},\mathbf{D},\boldsymbol{\gamma})=p(\mathbf{Y}|\boldsymbol{\varphi},\widetilde{\mathbf{A}},\mathbf{D},\boldsymbol{\gamma};\mathbf{Z})p(\mathbf{Z})$. Подобно 18, функции плотности для $\gamma_s=1$ и $\gamma_s=0$ отличаются лишь априорными плотностями параметров $\boldsymbol{\varphi}$:

$$\begin{aligned}&p(\boldsymbol{\varphi},\mathbf{D},\widetilde{\mathbf{A}},\boldsymbol{\gamma}_{-s};\mathbf{X}|\gamma_s(\varphi_{ij})=j))=\\&\quad=\frac{p(\gamma_s(\varphi_{ij})=1)p(\boldsymbol{\varphi}|\mathbf{D},\widetilde{\mathbf{A}}\boldsymbol{\gamma}_{-s},\gamma_s(\varphi_{ij})=1)}{p(\gamma_s(\varphi_{ij})=1)p(\boldsymbol{\varphi}|\mathbf{D},\widetilde{\mathbf{A}},\boldsymbol{\gamma}_{-s},\gamma_s(\varphi_{ij})=1)+p(\gamma_s(\varphi_{ij})=0)p(\boldsymbol{\varphi}|\mathbf{D},\widetilde{\mathbf{A}},\boldsymbol{\gamma}_{-s},\gamma_s(\varphi_{ij})=0)}\end{aligned} \tag{221}$$

В случае, если предполагается независимость ограничиваемых параметров (случай $\underline{\mathbf{V}}_{0.UR}^{\gamma}=\mathbf{0}$, $\mathbf{P}_R^0=\mathbf{P}_R^1=\mathbf{I}_f$), вероятность индикаторов ограничений расписывается следующим образом:

$$\begin{aligned}&p(\boldsymbol{\varphi},\mathbf{D},\widetilde{\mathbf{A}},\boldsymbol{\gamma}_{-s};\mathbf{X}|\gamma_s(\varphi_{ij})=j))=\\&\quad=\frac{p(\gamma_s(\varphi_{ij})=1)p(\varphi_{ij}|\mathbf{D},\widetilde{\mathbf{A}}\boldsymbol{\gamma}_{-s},\gamma_s(\varphi_{ij})=1)}{p(\gamma_s(\varphi_{ij})=1)p(\varphi_{ij}|\mathbf{D},\widetilde{\mathbf{A}},\boldsymbol{\gamma}_{-s},\gamma_s(\varphi_{ij})=1)+p(\gamma_s(\varphi_{ij})=0)p(\varphi_{ij}|\mathbf{D},\widetilde{\mathbf{A}},\boldsymbol{\gamma}_{-s},\gamma_s(\varphi_{ij})=0)}\end{aligned} \tag{222}$$

□

A Выводы из линейной алгебры

Большая часть материала этого раздела — это компиляция и осмысление материала статьи (Magnus and Neudecker, 1986).

A.1 Произведение Кронекера и его свойства

$$(\mathbf{A} \otimes \mathbf{B}) = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \otimes \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} \end{bmatrix}$$

$$\mathbf{A} \otimes (\mathbf{B} \otimes \mathbf{C}) = (\mathbf{A} \otimes \mathbf{B}) \otimes \mathbf{C} \tag{223}$$

$$\mathbf{A} \otimes (\mathbf{B} + \mathbf{C}) = (\mathbf{A} \otimes \mathbf{B}) + (\mathbf{A} \otimes \mathbf{C}) = (\mathbf{A} + \mathbf{B}) \otimes \mathbf{C} = (\mathbf{A} \otimes \mathbf{C}) + (\mathbf{B} \otimes \mathbf{C}) \tag{224}$$

$$(\mathbf{A} \otimes \mathbf{B})' = \mathbf{A}' \otimes \mathbf{B}' \tag{225}$$

$$(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1} \tag{226}$$

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD} \tag{227}$$

$$|\det(\underbrace{\mathbf{A}}_{p\times p} \otimes \underbrace{\mathbf{B}}_{m\times m})| = |\det(\mathbf{A})|^m |\det(\mathbf{B})|^p \tag{228}$$

A.2 Оператор векторизации и его связь с произведением Кронекера

Оператор векторизации определим следующим образом: это оператор, преобразующий заданную матрицу в вектор-столбец, состоящий из векторов-столбцов исходной матрицы: $\mathrm{vec}(\mathbf{A}) = [\mathbf{a}_{\bullet,1},\ \mathbf{a}_{\bullet,2},\ \ldots,\ \mathbf{a}_{\bullet,n}]$. Так, например, для матрицы размера (3×3) данная операция выглядит следующим образом:

$$\mathrm{vec}(\mathbf{A}) = \mathrm{vec}\left(\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\right) = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ a_{12} \\ a_{22} \\ a_{32} \\ a_{13} \\ a_{23} \\ a_{33} \end{bmatrix}$$

Теорема 1. *Магическая теорема:* $\mathrm{vec}(\mathbf{AXB}) = (\mathbf{B}' \otimes \mathbf{A})\,\mathrm{vec}(\mathbf{X})$

Доказательство. Пусть $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n]$ и $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m]$. Тогда k-столбец $\mathbf{AXB}$ выглядит следующим образом:

$$(\mathbf{AXB})_{\bullet,k} = \mathbf{AXb}_k = \mathbf{A}\sum_{i=1}^{m} \mathbf{x}_i b_{ik} = \sum_{i=1}^{m} b_{ik}\mathbf{Ax}_i = \begin{bmatrix} b_{1k}\mathbf{A} & b_{2k}\mathbf{A} & \ldots b_{mk}\mathbf{A} \end{bmatrix} \underbrace{\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \ldots \\ \mathbf{x}_m \end{bmatrix}}_{\mathrm{vec}(\mathbf{X})} = (\mathbf{b}_k' \otimes \mathbf{A})\,\mathrm{vec}(\mathbf{X})$$

Собираем всё воедино:

$$\begin{bmatrix} \mathbf{AXB}_{\bullet,1} \\ \mathbf{AXB}_{\bullet,2} \\ \dots \\ \mathbf{AXB}_{\bullet,n} \end{bmatrix} = \begin{bmatrix} \mathbf{b}_1' \otimes \mathbf{A} \\ \mathbf{b}_2' \otimes \mathbf{A} \\ \dots \\ \mathbf{b}_n' \otimes \mathbf{A} \end{bmatrix} \operatorname{vec}(\mathbf{X}) = (\mathbf{B}' \otimes \mathbf{A}) \operatorname{vec}(\mathbf{X}), \text{ q.e.d.}$$

□

Теорема 2. *2я Магическая теорема:* $\operatorname{tr}(\mathbf{B}'\mathbf{B}) = \operatorname{vec}(\mathbf{B})' \operatorname{vec}(\mathbf{B})$.

Доказательство. Распишем матрицу $\mathbf{B}$ в терминах столбцов: $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n]$. Тогда:

$$\operatorname{tr}(\mathbf{B}'\mathbf{B}) = \sum_{j=1}^{n} \mathbf{b}_j'\mathbf{b}_j = \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}^2 = \operatorname{vec}(\mathbf{B})' \operatorname{vec}(\mathbf{B})$$

□

A.3 Матрицы дупликации и элиминации. Оператор полувекторизации

Рассмотрим симметричную матрицу $\mathbf{A}$. В силу того, что для $i \neq j$ $\{i,j\}$ элементы симметричной матрицы по построению совпадают с $\{i,j\}$ элементами, может быть важно рассмотреть вектор, состоящий из по построению не повторяющихся элементов матрицы. Для этого используется операция полувекторизации $\operatorname{vech}(\bullet)$, которая, например, для матрицы 3×3 выглядит следующим образом:

$$\operatorname{vech}(\mathbf{A}) = \operatorname{vech}\left(\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}\right) = \begin{bmatrix} a_{11} \\ a_{12} \\ a_{13} \\ a_{22} \\ a_{23} \\ a_{33} \end{bmatrix}$$

Очевидно, что операцию полувекторизации можно свести к операции векторизации, домножив его слева на матрицу $\mathbf{D}_n$, состоящую из 0 и 1:

$$\mathbf{D}_n \operatorname{vech}(\mathbf{A}) = \operatorname{vec}(\mathbf{A}) \tag{229}$$

Матрица $\mathbf{D}_n$ называется **матрицей дупликации**. Для случая матрицы размерности 3×3 операция (229) выглядит следующим образом:

$$\underbrace{\left[\begin{array}{ccc|cc|c} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right]}_{\mathbf{D}_n} \underbrace{\begin{bmatrix} a_{11} \\ a_{12} \\ a_{13} \\ a_{22} \\ a_{23} \\ a_{33} \end{bmatrix}}_{\operatorname{vech}(\mathbf{A})} = \underbrace{\begin{bmatrix} a_{11} \\ a_{12} \\ a_{13} \\ a_{12} \\ a_{22} \\ a_{23} \\ a_{13} \\ a_{23} \\ a_{33} \end{bmatrix}}_{\operatorname{vec}(\mathbf{A})} \tag{230}$$

К сожалению, матрица дупликации всегда несимметричная ($\mathrm{rank}(\mathbf{D}_n) = 0.5n(n+1)$). Поэтому обратная к ней матрица не определена, и вместо обычной обратной матрицы пользуются псевдообратной матрицей Мура-Пенроуза:

$$\mathbf{D}_n^+ = (\mathbf{D}_n'\mathbf{D}_n)^{-1}\mathbf{D}_n' \tag{231}$$

Для случая 3×3 эта матрица выглядит следующим образом:

$$(\mathbf{D}_n^+)' = \left[\begin{array}{ccc|cc|c} 0.5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0 \\ \hline 0 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 0 \\ \hline 0 & 0 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

По определению псевдообратной матрицы верно следующее выражение:

$$\mathbf{D}_n^+\mathbf{D}_n = \mathbf{I}_{n(n-1)/2} \tag{232}$$

Лемма 1. *Пусть* $\mathbf{A}$ *— диагональная (верхняя треугольная или нижняя треугольная) матрица размерности* $n \times n$ *с диагональными элементами* $a_{11}, a_{22}, \ldots, a_{nn}$. *Тогда матрица* $\mathbf{D}_n^+(\mathbf{A} \otimes \mathbf{A})\mathbf{D}_n$ *также диагональная (верхняя треугольная или нижняя треугольная) с диагональными элементами* $a_{ii}a_{jj} (1 \leqslant j \leqslant i \leqslant n)$.

Доказательство. Доказательство представлено в работе (Magnus and Neudecker, 1986), стр. 174. □

Лемма 2. *Пусть* $\mathbf{A}$ *— диагональная (верхняя треугольная или нижняя треугольная) матрица размерности* $n \times n$ *с собственными значениями* $\lambda_{11}, \lambda_{22}, \ldots, \lambda_{nn}$. *Тогда собственные значения матрицы* $\mathbf{D}_n^+(\mathbf{A} \otimes \mathbf{A})\mathbf{D}_n$ *есть* $\lambda_i\lambda_j (1 \leqslant j \leqslant i \leqslant n)$ *и верны следующие соотношения:*

$$\mathrm{tr}(\mathbf{D}_n^+(\mathbf{A} \otimes \mathbf{A})\mathbf{D}_n) = \frac{\mathrm{tr}(\mathbf{A})^2 + \mathrm{tr}(\mathbf{A}^2)}{2}$$
$$|\det(\mathbf{D}_n^+(\mathbf{A} \otimes \mathbf{A})\mathbf{D}_n)| = |\det(\mathbf{A})|^{n+1}$$

Доказательство. Доказательство представлено в работе (Magnus and Neudecker, 1986), стр. 174-175.

По теореме Shur, существует невырожденная матрица $\mathbf{S}$, такая, что $\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \mathbf{M}$, где матрица $\mathbf{M}$ — верхняя треугольная матрица со собственными значениями $\lambda_{ii}, i \geqslant 1$ на её диагонали. Следовательно, по свойствам кронекеровского произведения,

$$\begin{aligned}\mathbf{B} = \mathbf{D}_n^+(\mathbf{M} \otimes \mathbf{M})\mathbf{D}_n = \mathbf{D}_n^+\left((\mathbf{S}^{-1}\mathbf{A}\mathbf{S}) \otimes (\mathbf{S}^{-1}\mathbf{A}\mathbf{S})\right)\mathbf{D}_n = \mathbf{D}_n^+\left(\mathbf{S}^{-1} \otimes \mathbf{S}^{-1}\right)(\mathbf{A} \otimes \mathbf{A})(\mathbf{S} \otimes \mathbf{S})\mathbf{D}_n = \\ = \underbrace{\mathbf{D}_n^+\left(\mathbf{S}^{-1} \otimes \mathbf{S}^{-1}\right)\mathbf{D}_n}_{\widetilde{\mathbf{S}}^{-1}}\underbrace{\mathbf{D}_n^+\,(\mathbf{A} \otimes \mathbf{A})\,\mathbf{D}_n}_{\widetilde{\mathbf{M}}}\underbrace{\mathbf{D}_n^+\,(\mathbf{S} \otimes \mathbf{S})\,\mathbf{D}_n}_{\widetilde{\mathbf{S}}}\end{aligned} \tag{233}$$

Следует обратить внимание на тот факт, что исходная матрица может быть представлена декомпозицией Шура $\mathbf{B} = \widetilde{\mathbf{S}}^{-1}\widetilde{\mathbf{M}}\widetilde{\mathbf{S}}$. А значит, матрицы $\mathbf{B}$ и $\widetilde{\mathbf{M}}$ имеют одинаковый набор собственных значений. По лемме **??** на диагональных элементах матрицы $\mathbf{B}$ стоят $\mu_s = \lambda_i\lambda_j (1 \leqslant j \leqslant n), 1 \leqslant s \leqslant n(n+1)/2$, которые являются собственными значениями матрицы $\mathbf{B}$, а значит, и матрицы $\widetilde{\mathbf{M}} = \mathbf{B} = \mathbf{D}_n^+(\mathbf{A} \otimes \mathbf{A})\mathbf{D}_n$.

Поскольку след матрицы есть сумма её собственных знчений, то:

$$\mathrm{tr}(\mathbf{D}_n^+(\mathbf{A} \otimes \mathbf{A})\mathbf{D}_n) = \sum_{s=1}^{n(n+1)/2} \mu_s = \sum_i^n \sum_{j \leqslant i} \lambda_i\lambda_j = \frac{1}{2}\sum_s^n \lambda_s^2 + \frac{1}{2}\sum_i^n \sum_{j \leqslant i} \lambda_i\lambda_j = \frac{\mathrm{tr}(\mathbf{A})^2 + \mathrm{tr}(\mathbf{A}^2)}{2}$$

Поскольку определитель матрицы есть произведение её собственных значений, то:

$$|\mathbf{D}_n^+(\mathbf{A}\otimes\mathbf{A})\mathbf{D}_n| = \prod_{i=1}^{n}\prod_{j=1}^{i}(\lambda_i\lambda_j) = \prod_{i}^{n}\left(\lambda_i^i\prod_{j=1}^{i}\lambda_j\right) = \prod_{i}^{n}\lambda_i^{i+1+\sum_{j=i+1}^{n}1} = \prod_{i}^{n}\lambda_i^{n+1} = |\mathbf{A}|^{n+1}$$

□

A.4 Якобианы для трансформации некоторых матриц

A.4.1 Якобиан замены элементов обратной матрицы на элементы исходной

Рассмотрим некоторую матрицу $\mathbf{A}$ размерности $n\times n$, для которой обратная определена как $\mathbf{C}=\mathbf{A}^{-1}$. Но это означает, что выполняется $\mathbf{CA}=\mathbf{I}_n$. Следовательно, взяв дифференциал по матрицам, можно получить:

$$\begin{aligned} \mathrm{d}\mathbf{C}\mathbf{A}+\mathbf{C}\,\mathrm{d}\mathbf{A} &= \mathbf{0}\\ \mathrm{d}\mathbf{C} &= -\mathbf{C}\,\mathrm{d}\mathbf{A}\mathbf{C}\end{aligned}$$

Таким образом, производная обратной матрицы по каждому элементу исходной выглядит как:

$$\frac{\partial C_{kl}}{\partial A_{ij}} = C_{ki}C_{jl} \tag{234}$$

Тогда Якобиан рассматриваемой замены выглядит следующим образом:

$$\left|\frac{\partial\,\mathrm{vec}(\mathbf{C})}{\partial\,\mathrm{vec}(\mathbf{A})}\right| = \left| - \begin{array}{c|c} & \begin{matrix} A_{11} & \dots & A_{n1} & A_{12} & \dots & A_{n2} & \dots & A_{1n} & \dots & A_{nn} \end{matrix} \\ \begin{matrix} C_{11}\\ \vdots \\ C_{n1} \\ C_{12} \\ \vdots \\ C_{n2} \\ \vdots \\ C_{1n} \\ \vdots \\ C_{nn}\end{matrix} & \begin{bmatrix} C_{11}C_{11} & & C_{1n}C_{11} & C_{11}C_{21} & & C_{1n}C_{21} & & C_{11}C_{n1} & & C_{1n}C_{n1} \\ \\ C_{n1}C_{11} & & C_{nn}C_{11} & C_{n1}C_{21} & & C_{nn}C_{21} & & C_{n1}C_{n1} & & C_{nn}C_{n1} \\ C_{11}C_{12} & & C_{1n}C_{12} & C_{11}C_{22} & & C_{1n}C_{22} & & C_{11}C_{n2} & & C_{1n}C_{n2} \\ \\ C_{n1}C_{12} & & C_{nn}C_{12} & C_{n1}C_{22} & & C_{nn}C_{22} & & C_{n1}C_{n2} & & C_{nn}C_{n2} \\ \\ C_{11}C_{1n} & & C_{1n}C_{1n} & C_{11}C_{2n} & & C_{1n}C_{2n} & & C_{11}C_{nn} & & C_{1n}C_{nn} \\ \\ C_{n1}C_{1n} & & C_{nn}C_{1n} & C_{n1}C_{2n} & & C_{nn}C_{2n} & & C_{n1}C_{nn} & & C_{nn}C_{nn} \end{bmatrix}\end{array}\right| =$$

$$= \left| - \begin{bmatrix} C_{11}\mathbf{C} & C_{21}\mathbf{C} & \dots & C_{n1}\mathbf{C} \\ C_{12}\mathbf{C} & C_{22}\mathbf{C} & \dots & C_{n2}\mathbf{C} \\ \dots & \dots & \dots & \dots \\ C_{1n}\mathbf{C} & C_{2n}\mathbf{C} & \dots & C_{nn}\mathbf{C}\end{bmatrix}\right| = |\mathbf{C}'\otimes\mathbf{C}| = |\mathbf{C}'|^n|\mathbf{C}|^n = |\mathbf{C}|^{2n} = |\mathbf{A}|^{-2n} \tag{235}$$

Тот же самый результат можно получить, просто векторизовав $\mathrm{d}\mathbf{C}$ и $\mathrm{d}\mathbf{A}$:

$$\mathrm{vec}(\mathrm{d}\mathbf{C}) = -(\mathbf{C}'\otimes\mathbf{C})\,\mathrm{vec}(\mathrm{d}\mathbf{A}) \tag{236}$$

$$\left|\frac{\partial\,\mathrm{vec}(\mathbf{C})}{\partial\,\mathrm{vec}(\mathbf{A})}\right| = |\mathbf{C}'\otimes\mathbf{C}| = |\mathbf{A}|^{-2n} \tag{237}$$

A.4.2 Якобиан замены элементов обратной матрицы на элементы исходной для симметричной матрицы

Для симметричных матриц ситуация осложняется тем, что теперь в матрицах $\mathbf{A}$ и $\mathbf{C}$ ровно $n(n+1)/2$ уникальных неповторяющихся элементов. Это значит, что количество замен сокращается ровно на $n(n-1)/2$ штук. Поэтому вместо того, чтобы рассматривать замену векторизации матрицы $\mathbf{C}$ на векторизацию матрицы $\mathbf{A}$, можно рассмотреть замену векторизации нижней треугольной подматрицы $\mathbf{C}$ на векторизацию нижней треугольной подматрицы $\mathbf{A}$.

Таким образом, этот якобиан можно вывести из (235) , используя вместо векторизации операцию полувекторизации. Но так как полувекторизация может быть получена из векторизации умножением на матрицу дупликации $\mathbf{D}$, то из леммы 2 следует, что:

$$\mathbf{D}\,\mathrm{vech}(\mathrm{d}\mathbf{C}) = -(\mathbf{C}'\otimes\mathbf{C})\mathbf{D}\,\mathrm{vech}(\mathrm{d}\mathbf{A}) \tag{238}$$

$$\left|\frac{\partial\,\mathrm{vech}(\mathbf{C})}{\partial\,\mathrm{vech}(\mathbf{A})}\right| = |\mathbf{D}_n^+(\mathbf{C}'\otimes\mathbf{C})\mathbf{D}_n| = |\mathbf{C}|^{n+1} = |\mathbf{A}|^{-(n+1)} \tag{239}$$

A.4.3 Якобиан замены элементов матрицы на элементы произведения матрицы на другую матрицу

Рассмотрим замену матрицы $\mathbf{A}$ на матрицу $\mathbf{B}=\mathbf{AC}$. Векторизуем матрицы $\mathbf{A}$ и $\mathbf{B}$ по строкам. Тогда очевидно, что:

$$\left|\frac{\partial\,\mathrm{vec}(\mathbf{A})}{\partial\,\mathrm{vec}(\mathbf{B})}\right| = \left|\frac{\partial\,\mathrm{vec}(\mathbf{B})}{\partial\,\mathrm{vec}(\mathbf{A})}\right|^{-1} = \left[\det\begin{pmatrix}\mathbf{C}' & \mathbf{0} & \dots & \mathbf{0}\\ \mathbf{0} & \mathbf{C}' & \dots & \mathbf{0}\\ \dots & \dots & \dots & \dots\\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{C}'\end{pmatrix}\right]^{-1} = |\mathbf{I}_n\otimes\mathbf{C}'|^{-1} = |\mathbf{I}_n|^{-n}|\mathbf{C}'|^{-n} = |\mathbf{C}|^{-n} \tag{240}$$

Для симметричных матриц подобная замена не определена: если матрица $\mathbf{A}$ симметрична, матрица $\mathbf{B}$ несимметрична, поэтому Якобиан для тако замены не определён. Зато для такой матрицы определен Якобиан замены $\mathbf{A}\to\mathbf{B}=\mathbf{CAC}$, где $\mathbf{C}$ — невырожденная матрица размерности $n\times n$:

$$\begin{aligned}\mathrm{d}\mathbf{B} &= \mathbf{C}\,\mathrm{d}\mathbf{A}\mathbf{C}\\ \mathrm{vec}(\mathrm{d}\mathbf{B}) &= (\mathbf{C}'\otimes\mathbf{C})\,\mathrm{vec}(\mathrm{d}\mathbf{A})\\ \mathrm{vech}(\mathrm{d}\mathbf{B}) &= \mathbf{D}_n^+(\mathbf{C}'\otimes\mathbf{C})\mathbf{D}_n\,\mathrm{vech}(\mathrm{d}\mathbf{A})\end{aligned}$$

Таким образом, в силу леммы 2 Якобиан такой замены выглядит следующим образом:

$$\left|\frac{\partial\,\mathrm{vech}(\mathbf{B})}{\partial\,\mathrm{vech}(\mathbf{A})}\right| = |\mathbf{C}'\otimes\mathbf{C}| = |\mathbf{C}|^{n+1} \tag{241}$$

A.4.4 Якобиан замены симметричной положительно определённой матрицы на матрицу разложения Холецкого для неё

Рассмотрим симметричную положительно определённую матрицу $\mathbf{X}$ размерности $n\times n$, разложение Холецкого $\mathbf{T}$ для которой можно записать как $\mathbf{X}=\mathbf{TT}'$. При этом распишем $\mathbf{T}'=[\mathbf{t}_1,\mathbf{t}_2,\dots,\mathbf{t}_n]$, где $\mathbf{t}_i$ — соответствующий вектор-столбец матрицы $\mathbf{T}'$. Следует обратить внимание на то, что, поскольку матрица $\mathbf{T}'$ верхняя треугольная, в векторе $\mathbf{t}_j$ элементы l_{ij} для $i\geqslant j+1$ являются нулевыми. Тогда исходную матрицу можно расписать как:

$$
\mathbf{X}=\begin{bmatrix}
x_{11} & x_{12} & \dots & x_{1(n-1)} & x_{1n}\\
x_{12} & x_{22} & \dots & x_{2(n-1)} & x_{2n}\\
\dots & \dots & \dots & \dots & \dots\\
x_{1(n-1)} & x_{2(n-1)} & \dots & x_{(n-1)(n-1)} & x_{(n-1)n}\\
x_{1n} & x_{2n} & \dots & x_{(n-1)n} & x_{nn}
\end{bmatrix}=\begin{bmatrix}
\mathbf{t}_1'\mathbf{t}_1 & \mathbf{t}_1'\mathbf{t}_2 & \dots & \mathbf{t}_1'\mathbf{t}_{n-1} & \mathbf{t}_1'\mathbf{t}_n\\
\mathbf{t}_2'\mathbf{t}_1 & \mathbf{t}_2'\mathbf{t}_2 & \dots & \mathbf{t}_2'\mathbf{t}_{n-1} & \mathbf{t}_2'\mathbf{t}_n\\
\dots & \dots & \dots & \dots & \dots\\
\mathbf{t}_{n-1}'\mathbf{t}_1 & \mathbf{t}_{n-1}'\mathbf{t}_2 & \dots & \mathbf{t}_{n-1}'\mathbf{t}_{n-1} & \mathbf{t}_{n-1}'\mathbf{t}_n\\
\mathbf{t}_n'\mathbf{t}_1 & \mathbf{t}_n'\mathbf{t}_2 & \dots & \mathbf{t}_n'\mathbf{t}_{n-1} & \mathbf{t}_n'\mathbf{t}_n
\end{bmatrix}=
$$

$$
=\begin{bmatrix}
t_{11}^2 & t_{11}t_{12} & t_{11}t_{13} & \dots & t_{11}t_{1(n-1)} & t_{11}t_{1n}\\
t_{12}t_{11} & \sum_{i=1}^{2} t_{i2}^2 & \sum_{i=1}^{2} t_{i2}t_{i3} & \dots & \sum_{i=1}^{2} t_{i2}t_{i(n-1)} & \sum_{i=1}^{2} t_{i2}t_{in}\\
t_{13}t_{11} & \sum_{i=1}^{2} t_{i3}t_{i2} & \sum_{i=1}^{3} t_{i3}^2 & \dots & \sum_{i=1}^{3} t_{i3}t_{i(n-1)} & \sum_{i=1}^{3} t_{i3}t_{in}\\
\dots & \dots & \dots & \dots & \dots & \dots\\
t_{1(n-1)}t_{11} & \sum_{i=1}^{2} t_{i(n-1)}t_{i2} & \sum_{i=1}^{3} t_{i(n-1)}t_{i3} & \dots & \sum_{i=1}^{n-1} t_{i(n-1)}^2 & \sum_{i=1}^{n-1} t_{i(n-1)}t_{in}\\
t_{1n}t_{11} & \sum_{i=1}^{2} t_{in}t_{i2} & \sum_{i=1}^{3} t_{in}t_{i3} & \dots & \sum_{i=1}^{n-1} t_{in}t_{i(n-1)} & \sum_{i=1}^{n} t_{in}^2
\end{bmatrix} \tag{242}
$$

Следовательно:

$$
x_{ij}=\begin{cases}\sum_{s=1}^{\min(i,j)} t_{si}t_{sj}, \text{ если } i\neq j\\ \sum_{s=1}^{i} t_{si}^2, \text{ если } i=j\end{cases} \tag{243}
$$

Производная диагональных элементов матрицы $\partial x_{ii}/\partial t_{sl}=2t_{ii}, \forall s\leqslant l, \forall l=i$ и $\partial x_{ii}/\partial t_{sl}=0$ иначе. Тогда для диагональных элементов матрицы $\mathbf{X}$ вектор производных по элементам матрицы $\mathbf{T}$ выглядит следующим образом:

$$
\left(\frac{\partial x_{ii}}{\partial\,\mathrm{vech}(\mathbf{T})}\right)'= x_{ii}\begin{array}{c}
\begin{matrix} t_{11} & \dots & t_{1i} & \dots & t_{1n} & t_{22} & \dots & t_{2i} & \dots & t_{ii} & \dots & t_{nn}\end{matrix}\\
\begin{bmatrix} 0 & 0 & 2t_{1i} & 0 & 0 & 0 & 0 & 2t_{2i} & \dots & 2t_{ii} & 0 & 0\end{bmatrix}
\end{array} \tag{244}
$$

Для несимметричных же элементов сложнее. В силу того, что матрица $\mathbf{X}$ симметричная, можно без ограничения общности рассматривать только элементы $x_{ij}: i\leqslant j$. Тогда $\min(i,j)=i$. Следует обратить внимание на тот факт, что производная элемента x_{ij} по любому диагональному элементу x_{ss}, кроме $s=i\leqslant j$, равна нулю: $\partial x_{ij}/\partial t_{ss}=0, \forall s\neq i$, а $\partial x_{ij}/\partial t_{ii}=t_{ij}$. Более того, $\partial x_{ij}/\partial t_{sl}=0, \forall l>j$, $\partial x_{ij}/\partial t_{si}=t_{sj}, \forall s\leqslant i$, $\partial x_{ij}/\partial t_{sj}=t_{si}, \forall s\leqslant i$. Тогда для элементов матрицы $\mathbf{X}$ вне главной диагонали вектор производных по элементам матрицы $\mathbf{T}$ выглядит следующим образом:

$$
\left(\frac{\partial x_{ii}}{\partial\,\mathrm{vech}(\mathbf{T})}\right)'= x_{ij}\begin{array}{c}
\begin{matrix} t_{11} & \dots & t_{1i} & \dots & t_{1j} & \dots & t_{1n} & t_{22} & \dots & t_{ii} & \dots & t_{ij} & \dots & t_{nn}\end{matrix}\\
\begin{bmatrix} 0 & 0 & t_{1j} & 0 & t_{1i} & 0 & 0 & 0 & \dots & t_{ij} & 0 & t_{ii} & 0 & 0\end{bmatrix}
\end{array} \tag{245}
$$

Следовательно, Якобиан рассматриваемой замены может быть записан в следующей форме:

$$\left|\frac{\partial\,\mathrm{vech}(\mathbf{X})}{\partial\,\mathrm{vech}(\mathbf{T})}\right| =$$

$$= \left| \begin{array}{cc} & \begin{array}{ccccccccccccccc} t_{11} & t_{12} & t_{13} & \dots & t_{1n} & t_{22} & \dots & t_{2n} & t_{33} & \dots & t_{3n} & t_{44} & \dots & t_{(n-1)n} & t_{nn} \end{array} \\ \begin{array}{c} x_{11} \\ x_{12} \\ x_{13} \\ \vdots \\ x_{1n} \\ x_{22} \\ \vdots \\ x_{2n} \\ x_{33} \\ \vdots \\ x_{3n} \\ \vdots \\ x_{(n-1)n} \\ x_{nn} \end{array} & \begin{bmatrix} 2t_{11} & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 \\ t_{12} & t_{11} & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 \\ t_{13} & 0 & t_{11} & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \ddots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ t_{1n} & 0 & 0 & \dots & t_{11} & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 \\ 0 & 2t_{12} & 0 & \dots & 0 & 2t_{22} & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \ddots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & t_{1n} & 0 & \dots & t_{12} & t_{2n} & \dots & t_{22} & 0 & \dots & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 2t_{13} & \dots & 0 & 0 & \dots & 0 & 2t_{33} & \dots & 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \ddots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & t_{1n} & \dots & t_{13} & 0 & \dots & t_{23} & t_{3n} & \dots & t_{33} & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \ddots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & t_{1(n-1)} & 0 & \dots & t_{2(n-1)} & 0 & \dots & t_{3(n-1)} & 0 & \dots & t_{n-1.n-1} & 0 \\ 0 & 0 & 0 & \dots & 2t_{1n} & 0 & \dots & 2t_{2n} & 0 & \dots & 2t_{3n} & 0 & \dots & 0 & 2t_{nn} \end{bmatrix} \end{array} \right| \quad (24$$

Очевидно, что это определитель треугольной матрицы. Следовательно, определитель равен произведению её диагональных элементов. Среди диагональных элементов 2 встречается ровно n раз (количество диагональных элементов в матрице $\mathbf{X}$), x_{11} встречается n раз (количество элементов первого вектора-столбца матрицы $\mathbf{X}$), x_{22} встречается $n-1$ раз (количество элементов второго вектора-столбца матрицы $\mathbf{X}$), и т.д. Тогда Якобиан может быть выражен следующим образом:

$$\left|\frac{\partial\,\mathrm{vech}(\mathbf{X})}{\partial\,\mathrm{vech}(\mathbf{T})}\right| = 2^n \prod_{i=1}^{n} t_{ii}^{n-i+1} \tag{247}$$

A.5 Многомерная гамма-функция

Многомерная гамма-функция — это многомерный интеграл по положительно определённой матрице $\mathbf{S} > 0$ размерности $n \times n$ следующей формы (смотри Siegel, 1935):

$$\Gamma_n(a) = \int_{\mathbf{S}>0} |\det(\mathbf{S})|^{a-(n+1)/2} \exp\left\{-\operatorname{tr}(\mathbf{S})\right\} \mathrm{d}\mathbf{S} \tag{248}$$

Данный интеграл, подобно одномерной гамме-функции, можно вывести аналитически:

$$\Gamma_n(a) = \pi^{n(n-1)/2} \prod_{j=1}^{n} \Gamma\left(a + \frac{j-1}{2}\right) \tag{249}$$

Для некоторых выводов нам также может быть полезен следующий интеграл, который может быть выражен через многомерную гамма-функцию:

$$\int_{\mathbf{S}>0} |\det(\mathbf{S})|^{\nu} \exp\left\{-\frac{1}{2}\operatorname{tr}(\mathbf{S}\mathbf{K})\right\} \mathrm{d}\mathbf{S}, \tag{250}$$

где $\mathbf{K}$ — некоторая постоянная матрица размерности $n \times n$. Для того, чтобы привести эту функцию к (249), сделаем замену переменных $\mathbf{B} = 0.5\mathbf{S}\mathbf{K}$. Учитывая результат, полученный в подразделе A.4.3, данный интеграл превращается в следующее выражение:

$$\begin{aligned}
&\int_{\mathbf{S}>0} |\det(\mathbf{2BK}^{-1})|^{\nu+[(n+1)/2]-(n+1)/2} \exp\left\{-\operatorname{tr}(\mathbf{B})\right\} \mathrm{d}\mathbf{B} \left|\frac{\partial\operatorname{vec}(\mathbf{S})}{\partial\operatorname{vec}(\mathbf{B})}\right| = \\
&= |\det(\mathbf{K})|^{-(n+\nu)}(2)^{n+\nu} \int_{\mathbf{S}>0} |\det(\mathbf{B})|^{\nu+[(n+1)/2]-(n+1)/2} \exp\left\{-\frac{1}{2}\operatorname{tr}(\mathbf{B})\right\} \mathrm{d}\mathbf{B} = \\
&= |\det(\mathbf{K})|^{-(n+\nu)}(2)^{n+\nu}\Gamma_n\left(\nu + \frac{n+1}{2}\right)
\end{aligned} \tag{251}$$

В случае, если матрицы $\mathbf{S}$ и $\mathbf{K}$ симметричные, положительно определённые и невырожденные, следует сделать разложение Холецкого для постоянной матрицы $\mathbf{K}$: $\mathbf{K} = \mathbf{L}\mathbf{L}'$. Известно, что $\operatorname{tr}(\mathbf{SK}) = \operatorname{tr}(\mathbf{SLL}') = \operatorname{tr}(\mathbf{L}'\mathbf{SL})$. Следовательно, нужно сделать замену $\mathbf{S} \to \mathbf{L}'\mathbf{SL}$.

Тогда в силу вывода (241) и того факта, что $|\mathbf{L}| = |\mathbf{K}|^{\frac{1}{2}}$, Якобиан может быть найден следующим образом:

$$\left|\frac{\partial\operatorname{vech}(\mathbf{S})}{\partial\operatorname{vech}(\mathbf{B})}\right| = \left|\det(\frac{1}{2}\mathbf{K})\right|^{-(n+1)/2} \tag{252}$$

Тогда формула (251) преобразуется в следующее выражение (следует помнить, что для положительной константы c и матрицы $\mathbf{A}$ размерности $n \times n$ верно, что $|c\mathbf{A}| = c^n|\mathbf{A}|$):

$$\begin{aligned}
&\int_{\mathbf{S}>0} |\det(\mathbf{2BK}^{-1})|^{\nu+[(n+1)/2]-(n+1)/2} \exp\left\{-\operatorname{tr}(\mathbf{B})\right\} \mathrm{d}\mathbf{B} \left|\frac{\partial\operatorname{vech}(\mathbf{S})}{\partial\operatorname{vech}(\mathbf{B})}\right| = \\
&= |\det(\mathbf{K})|^{-((n+1)/2+\nu)}(2)^{(\nu+(n+1)/2)} \int_{\mathbf{S}>0} |\mathbf{B}|^{\nu+[(n+1)/2]-(n+1)/2} \exp\left\{-\frac{1}{2}\operatorname{tr}(\mathbf{B})\right\} \mathrm{d}\mathbf{B} = \\
&= |\det(\mathbf{K})|^{-(\nu+(n+1)/2)}(2)^{n(\nu+(n+1)/2)}\Gamma_n\left(\nu + \frac{n+1}{2}\right)
\end{aligned} \tag{253}$$

B Распределения

B.1 Распределение Вишарта

Пусть $\mathbf{X}$ есть $n \times \nu$ $(n > \nu - 1)$ случайная матрица, которую можно записать через векторы-строки: $\mathbf{X}' = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n]$, причём каждый вектор $\mathbf{x}_i$ $i.i.d \sim \mathcal{N}_\nu(\mathbf{0}, \mathbf{\Psi})$. Тогда будем говорить, что матрица $\mathbf{S} = \mathbf{X}'\mathbf{X} \sim \mathcal{W}_n(\nu, \mathbf{\Psi})$.

Обозначение: $\mathbf{S} \sim \mathcal{W}_n(\nu, \mathbf{\Psi})$

Функция плотности распределения:

$$p(\mathbf{S}) = (2\pi)^{-\frac{\nu n}{2}} \Gamma_n^{-1}\left(\frac{\nu}{2}\right) |\det(\mathbf{\Psi})|^{-\frac{\nu}{2}} \left[|\det(\mathbf{S})|^{\frac{\nu-(n+1)}{2}} \exp\left(-\operatorname{tr}(\mathbf{\Psi}^{-1}\mathbf{S})\right) \right] \tag{254}$$

B.2 Обратное Распределение Вишарта

Если некоторая матрица имеет распределение Вишарта $\mathbf{S} \sim \mathcal{W}_n(\nu, \mathbf{\Psi})$, то обратная к ней имеет обратное распределение Вишарта: $\mathbf{C} = \mathbf{S}^{-1} \sim \mathcal{IW}_n(\nu, \mathbf{\Xi})$, где $\mathbf{\Xi} = \mathbf{\Psi}^{-1}$.

Обозначение: $\mathbf{C} \sim \mathcal{IW}_n(\nu, \mathbf{\Xi})$

Функция плотности распределения:

$$p(\mathbf{S}) = (2\pi)^{-\frac{\nu n}{2}} \Gamma_n^{-1}\left(\frac{\nu}{2}\right) |\det(\mathbf{\Xi})|^{\frac{\nu}{2}} \left[|\det(\mathbf{C})|^{-\frac{\nu-(n+1)}{2}} \exp\left(-\operatorname{tr}(\mathbf{\Xi}\mathbf{C}^{-1})\right) \right] \tag{255}$$

B.3 Матричное нормальное распределение

Если случайная матрица $\mathbf{X}$ размерности $k \times n$ имеет матричное нормальное распределение, $\mathbf{M}$ — матрица размерности $k \times n$, $\mathbf{U}$ — матрица размерности $k \times k$, а $\mathbf{W}$ — матрица размерности $n \times n$ то:

Обозначение: $\mathbf{X} \sim \mathcal{MN}_{kn}(\mathbf{M}, \mathbf{U}, \mathbf{W})$

Функция плотности распределения:

$$p(\mathbf{X}) = (2\pi)^{-\frac{nk}{2}} |\mathbf{U}|^{-\frac{n}{2}} |\det(\mathbf{W})|^{-\frac{k}{2}} \exp\left[-\frac{1}{2} \operatorname{tr}\left(\mathbf{U}^{-1}(\mathbf{X}-\mathbf{M})'\mathbf{W}^{-1}(\mathbf{X}-\mathbf{M}) \right) \right] \tag{256}$$

B.4 Вишарт-нормальное распределение

B.5 Матричное t-распределение

Если случайная матрица $\mathbf{X}$ размерности $k \times n$ имеет матричное нормальное распределение, $\mathbf{M}$ — матрица размерности $k \times n$, $\mathbf{Q}$ — матрица размерности $k \times n$ и с $\nu \geqslant n$ степенями свободы.

Обозначение: $\mathbf{X} \sim \mathcal{M}t_{kn}(\nu, \mathbf{M}, \mathbf{P}, \mathbf{Q})$

Функция плотности распределения:

$$p(\mathbf{X}) = (\pi)^{-\frac{kn}{2}} \Gamma_n\left(\frac{\nu+k}{2}\right) \Gamma_n^{-1}\left(\frac{\nu}{2}\right) |\det(\mathbf{P})|^{-\frac{k}{2}} |\det(\mathbf{Q})|^{-\frac{n}{2}} \left|\det\left(\mathbf{I}_k + \mathbf{Q}^{-1}(\mathbf{X}-\mathbf{M})'\mathbf{P}^{-1}(\mathbf{X}-\mathbf{M})\right)\right|^{-(\nu+k)/2} \tag{257}$$

Список литературы

Magnus, J. R. and Neudecker, H. (1986). Symmetry, 0-1 matrices and jacobians: A review. *Econometric Theory*, 2(02):157–190.