ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ВЫСШАЯ ШКОЛА ЭКОНОМИКИ

А.С. Шведов

t-РАСПРЕДЕЛЕНИЕ СЛУЧАЙНОЙ МАТРИЦЫ И ЕГО ПРИМЕНЕНИЕ В РЕГРЕССИОННОЙ МОДЕЛИ

Препринт WP2/2010/01
Серия WP2
Количественный анализ в экономике

Москва
2010

УДК 519.246
ББК 22.172
Ш34

Редактор серии WP2
«Количественный анализ в экономике»
В.А. Бессонов

Ш 34 **Шведов А. С.**
t-Распределение случайной матрицы и его применение в регрессионной модели: Препринт WP2/2010/01. — М.: Государственный университет — Высшая школа экономики, 2010. — 28 с.

В работе вводятся матричные t-распределения, для которых параметр, представляющий число степеней свободы, является вектором. Показана возможность использования таких распределений в эконометрических моделях.

УДК 519.246
ББК 22.172

Классификация JEL: C14, C32.

Ключевые слова: t-распределение случайной матрицы; многомерная регрессионная модель

Shvedov A. S.
Matrix variate t-distribution with application to regression models: Working paper WP2/2010/01. — Moscow: State University — Higher School of Economics, 2010. — 28 p. (in Russian).

This paper is concerned with matrix variate t-distribution. The degrees of freedom parameter of the t-distribution is a vector. The distribution is applied to multivariate regression models.

JEL Classification: C14, C32.

Key phrases: matrix variate t-distribution; multivariate regression models

Препринты Государственного университета – Высшей школы экономики размещаются на сайте http://new.hse.ru/C3/C18/preprintsID/default.aspx

1. Введение

В связи с усложнением используемых эконометрических моделей в последние десятилетия неуклонно возрастает интерес к многомерным и матричным распределениям вероятностей.

Хорошо знакомые из курса математической статистики t-распределения допускают обобщение на многомерный случай, это дает так называемые многомерные t-распределения. Но и многомерные t-распределения, в свою очередь, допускают обобщение для задач, в которых рассматриваются не распределения случайных векторов, а распределения случайных матриц. Это дает матричные t-распределения. Матричные t-распределения подробно изучаются, например, в [8, гл. 4] и играют важную роль в байесовском анализе регрессионных моделей (см., например, [2], [6]). В этих исследованиях тот параметр, который для одномерных t-распределений называется числом степеней свободы, и для многомерных t-распределений, и для матричных t-распределений остается положительным числом.

В работе [4] при изучении модели состояние-наблюдение получено обобщение многомерного t-распределения, для которого указанный параметр является вектором. (Размерность этого вектора совпадает с размерностью случайного вектора, распределение которого рассматривается.) Мы будем называть его t-распределением с вектором степеней свободы. В настоящей работе данное распределение обобщается на матричный случай.

Матричные t-распределения связаны с матричными гамма-распределениями (которые используются и в этой работе) и с матричными бета-распределениями.

Одномерные гамма-распределения образуют двухпараметрическое семейство, где оба параметра – положительные числа. При переходе от одномерных гамма-распределений к гамма-распределениям случайных матриц один из этих параметров становится положительно определенной матрицей. Другой параметр может оставаться положительным числом; если это число полуцелое, то гамма-распределение называется распределением Уишарта. Можно рассматривать и более широкое семейство матричных гамма-распределений, когда указанный параметр является вектором с положительными компонентами (а не положительным числом). Это семейство гамма-распределений изучается в работах [5], [10] (см. также [8, стр. 122 – 123]). Многие принципиальные вопросы, касающиеся этого семейства матричных гамма-распределений остаются неизученными.

Одномерные бета-распределения образуют двухпараметрическое семейство, где оба параметра – положительные числа. Случайная матрица, имеющая бета-распределение может быть построена при помощи двух независимых случайных матриц, имеющих гамма-распределения (см., например, [4]). Хотя известны и попытки использовать для построения бета-распределенной случайной матрицы такие гамма-распределенные случайные матрицы, у которых параметры, являющиеся положительно определенными матрицами, различны (см., например, [7]), в настоящее время можно считать общепринятым подход, когда у обеих используемых при таком построении гамма-распределенных случайных матриц тот параметр, который является положительно определенной матрицей, берется одним и тем же. И бета-распределение оказывается не зависящим от данного параметра.

Но если другой параметр у каждой из гамма-распределенных случайных матриц является лишь положительным числом, то матричные бета-распределения, как и одномерные бета-распределения образуют двухпараметрическое семейство, где оба параметра – положительные числа. Использование таких матричных бета-распределений едва ли может обеспечить достаточную гибкость при решении сложных задач. Более перспективным представляется подход, когда этот параметр у каждой из используемых гамма-распределенных случайных матриц является вектором с положительными компонентами (см. [8, стр 24 – 25], [10], [4]). Такие матричные бета-распределения будем называть бета-распределениями с двумя векторными параметрами.

При моделировании многомерных временных рядов, у которых периоды большой волатильности чередуются с периодами малой волатильности, наряду с моделями $GARCH$ могут применяться модели состояние-наблюдение, основанные на бета-распределении случайных матриц. В [4] показана возможность использования в таких моделях состояние-наблюдение матричных бета-распределений с двумя векторными параметрами.

Также в [4] получено выражение для функции плотности прогнозного распределения для таких моделей состояние-наблюдение. Это и есть многомерное t-распределение с вектором степеней свободы.

Матричное t-распределение с вектором степеней свободы строится в разделе 2 настоящей работы. В разделе 3 показано, что при соответствующем выборе априорного распределения параметров многомерной регрессионной модели апостериорное распределение этих параметров является матричным t-распределением с вектором степеней свободы. Раздел 4 носит вспомогательный характер, в нем содержатся некоторые известные результаты, относящиеся к матричному нормальному распределению.

2. Матричное t-распределение

Пусть r и s – натуральные числа. Через I_k будем обозначать единичную $k \times k$ матрицу.

Пусть $r \times s$ случайная матрица Z имеет матричное нормальное распределение со средним 0 и с ковариационной матрицей $I_r \otimes B$, где B – положительно определенная $s \times s$ матрица, $\otimes$ – символ кронекерова произведения матриц. В соответствии с формулой (15) (см. раздел 4) это означает, что функция плотности случайной матрицы Z имеет вид

$$g(z) = (2\pi)^{-\frac{1}{2}rs} |B|^{-\frac{1}{2}r} \operatorname{etr}\left(-\frac{1}{2} z B^{-1} z'\right), \qquad (1)$$

где z – произвольная $r \times s$ матрица. Штрих означает транспонирование, $\operatorname{etr}(\cdot) = \exp(\operatorname{tr}(\cdot))$.

Напомним, что для любой $r \times r$ матрицы

$$C = \{c_{ij}\}, \quad 1 \le i, j \le r,$$

через $C^{[k]}$ и $C_{[k]}$ обозначаются подматрицы

$$C^{[k]} = \{c_{ij}\}, \quad 1 \le i, j \le k,$$

$$C_{[k]} = \{c_{ij}\}, \quad r - k + 1 \le i, j \le r,$$

где $k = 1, \ldots, r$.

Пусть положительно определенная $r \times r$ случайная матрица W имеет гамма-распределение с параметрами a и A. Здесь $a = (a_1, \ldots, a_r)$ – r-мерный вектор такой, что

$$a_j > \frac{j-1}{2}, \quad j = 1, \ldots, r.$$

Также введем обозначения $a_0 = 0$, $a_{r+1} = \dfrac{r+1}{2}$. A – положительно определенная $r \times r$ матрица.

Функция плотности случайной матрицы W имеет вид

$$f(w) = \gamma_{a,A} \operatorname{etr}(-Aw) \prod_{j=1}^{r} \left| w_{[j]} \right|^{a_j - a_{j+1}}, \tag{2}$$

где

$$\gamma_{a,A} = \left(\Gamma_r^*(a) \prod_{j=0}^{r-1} \left| A^{[r-j]} \right|^{a_j - a_{j+1}} \right)^{-1}, \tag{3}$$

$$\Gamma_r^*(a) = \pi^{\frac{r(r-1)}{4}} \prod_{j=1}^{r} \Gamma\left(a_j - \frac{j-1}{2} \right),$$

$\Gamma(\cdot)$ – гамма-функция. Функция $f(w)$ задана в области пространства $\mathrm{R}^{\frac{r(r+1)}{2}}$, которую мы будем обозначать $w > 0$, и которая состоит из таких точек

$$w = (w_{11}, \ldots, w_{1r}, w_{22}, \ldots, w_{2r}, \ldots, w_{rr}),$$

что матрица w положительно определена. Подробнее о функции плотности гамма-распределения $f(w)$ см. [5], [10], [4].

Напомним, как изменяется функция плотности случайного вектора при взаимно-однозначном гладком отображении Φ некоторой области N-мерного евклидова пространства на область другого N-мерного евклидова пространства, $t = \Phi(s)$, где через s и t обозначены точки, соответственно, первого и второго евклидова пространства.

Рассмотрим N-мерные случайные вектора S и T, принимающие значения в соответствующих областях, и связанные соотношением $T = \Phi(S)$. Если функция плотности случайного вектора T известна и равна $f_T(t)$, то функция плотности случайного вектора S определяется при помощи формулы

$$f_S(s) = f_T\left(\Phi\left(s\right)\right)\ J\left(t;s\right),$$

где $J\left(t;s\right)$ – абсолютная величина определителя матрицы

$$\frac{\partial(t)}{\partial(s)} = \left(\begin{array}{ccc} \dfrac{\partial t_1}{\partial s_1} & \dots & \dfrac{\partial t_1}{\partial s_N} \\ \dots & \dots & \dots \\ \dfrac{\partial t_N}{\partial s_1} & \dots & \dfrac{\partial t_N}{\partial s_N} \end{array}\right),$$

то есть якобиана отображения Φ.

Известно, что если $r \times s$ матрицы u и z связаны соотношением $z = CuB$, где C и B – невырожденные матрицы порядка r и s, соответственно, то

$$J\left(z;u\right) = \left|B\right|^r \left|C\right|^s \tag{4}$$

(см., например, [9, теорема 2.1.5]).

Пусть $r \times s$ случайная матрица Z имеет функцию плотности $g(z)$, задаваемую формулой (1), положительно определенная $r \times r$ случайная матрица W имеет функцию плотности $f(w)$, задаваемую формулой (2), $r \times r$ случайная матрица L удовлетворяет соотношению $W = L'L$. (Может быть применен любой из трех способов построения такого разложения матрицы W, то есть матрица L может быть либо нижней треугольной с положительными диагональными элементами, либо верхней треугольной с положительными диагональными элементами, либо $L = W^{\frac{1}{2}}$.) Определим $r \times s$ случайную матрицу

$$U = L^{-1}Z + M,$$

где M – произвольная $r \times s$ матрица.

Теорема 1. Функция плотности случайной матрицы U имеет вид

$$f_U\left(u\right) = \left(2\pi\right)^{-\frac{1}{2}\,rs}\,\Gamma_r^*\left(b\right)\left(\Gamma_r^*\left(a\right)\right)^{-1}\left|A\right|^{-\frac{1}{2}\,s}\left|B\right|^{-\frac{1}{2}\,r}\cdot$$

$$\cdot\prod_{j=0}^{r-1}\left|I_{r-j}+\frac{1}{2}\left(A^{[r-j]}\right)^{-1}\left(\left(u-M\right)B^{-1}\left(u-M\right)'\right)^{[r-j]}\right|^{b_j-b_{j+1}},$$

где u – произвольная $r \times s$ матрица,

$$b_j = a_j + \frac{s}{2}, \quad j = 1, \ldots, r,$$

$b = (b_1, \ldots, b_r)$, $b_0 = 0$, $b_{r+1} = \dfrac{r+1}{2}$.

Доказательство. Совместная функция плотности случайных матриц W и Z имеет вид

$$f_{W,Z}(w,z) = f(w)\;g(z).$$

Пусть $r \times r$ матрица l удовлетворяет соотношению $w = l'\,l$ и $r \times s$ матрица u определяется соотношением

$$z = l(u-M).$$

Совместная функция плотности случайных матриц W и U имеет вид

$$f_{W,U}(w,u) = f_{W,Z}(w,z(u))J\left(w,z;w,u\right).$$

Определитель матрицы

$$\frac{\partial(w,z)}{\partial(w,u)} = \begin{pmatrix} \dfrac{\partial(w)}{\partial(w)} & \dfrac{\partial(z)}{\partial(w)} \\ \dfrac{\partial(w)}{\partial(u)} & \dfrac{\partial(z)}{\partial(u)} \end{pmatrix}$$

равен определителю матрицы $\dfrac{\partial(z)}{\partial(u)}$. Следовательно,

$$J\left(w,z;w,u\right)=J\left(z;u\right).$$

В соответствии с (4) $J\left(z;u\right)=|l|^{s}=|w|^{\frac{1}{2}\,s}$. Поэтому

$$f_{W,U}(w,u)=(2\pi)^{-\frac{1}{2}\,rs}\,|B|^{-\frac{1}{2}\,r}\cdot$$
$$\cdot\mathrm{etr}\left(-\frac{1}{2}\,l\,\left(u-M\right)B^{-1}\left(u-M\right)'\,l'\right)\cdot$$
$$\cdot\gamma_{a,A}\,\mathrm{etr}\left(-Aw\right)\prod_{j=1}^{r}\left|w_{[j]}\right|^{a_j-a_{j+1}}\cdot|w|^{\frac{1}{2}\,s}=$$
$$=(2\pi)^{-\frac{1}{2}\,rs}\,|B|^{-\frac{1}{2}\,r}\,\gamma_{a,A}\,\prod_{j=1}^{r}\left|w_{[j]}\right|^{a_j-a_{j+1}}\cdot|w|^{\frac{1}{2}\,s}\cdot$$
$$\cdot\mathrm{etr}\left(-\left(A+\frac{1}{2}\left(u-M\right)B^{-1}\left(u-M\right)'\right)w\right)=$$
$$=(2\pi)^{-\frac{1}{2}\,rs}\,|B|^{-\frac{1}{2}\,r}\,\frac{\gamma_{a,A}}{\gamma_{b,A+\frac{1}{2}(u-M)B^{-1}(u-M)'}}\cdot$$
$$\cdot\gamma_{b,A+\frac{1}{2}(u-M)B^{-1}(u-M)'}\prod_{j=1}^{r}\left|w_{[j]}\right|^{b_j-b_{j+1}}\cdot$$
$$\cdot\mathrm{etr}\left(-\left(A+\frac{1}{2}\left(u-M\right)B^{-1}\left(u-M\right)'\right)w\right).$$

Интегрируя совместную функцию плотности $f_{W,U}(w,u)$ по области $w>0$ и используя вид функции плотности (2), получаем выражение для маргинальной функции плотности случайной матрицы U:

$$f_U(u)=(2\pi)^{-\frac{1}{2}\,rs}\,|B|^{-\frac{1}{2}\,r}\,\frac{\gamma_{a,A}}{\gamma_{b,A+\frac{1}{2}(u-M)B^{-1}(u-M)'}}.$$

Применение (3) дает утверждение теоремы.

Теорема 1 доказана.

Лемма 1. Пусть A – невырожденная $r \times r$ матрица, B – положительно определенная $s \times s$ матрица, x – $r \times s$ матрица. Тогда

$$|A + xBx'| = |A|\,|B|\left|B^{-1} + x'A^{-1}x\right|.$$

Доказательство. Рассмотрим $(r+s) \times (r+s)$ матрицу

$$C = \begin{pmatrix} I_r & xB^{\frac{1}{2}} \\ -B^{\frac{1}{2}}x'A^{-1} & I_s \end{pmatrix}$$

Пользуясь правилом умножения блочных матриц, находим

$$C \begin{pmatrix} A & 0 \\ B^{\frac{1}{2}}x' & I_s \end{pmatrix} = \begin{pmatrix} A + xBx' & xB^{\frac{1}{2}} \\ 0 & I_s \end{pmatrix}. \qquad (5)$$

С другой стороны,

$$\begin{pmatrix} I_r & 0 \\ B^{\frac{1}{2}}x'A^{-1} & I_s \end{pmatrix} C = \begin{pmatrix} I_r & xB^{\frac{1}{2}} \\ 0 & B^{\frac{1}{2}}x'A^{-1}xB^{\frac{1}{2}} + I_s \end{pmatrix}. \qquad (6)$$

Из (5) получаем

$$|C|\,|A| = |A + xBx'|\,.$$

Из (6) получаем

$$|C| = \left|B^{\frac{1}{2}}x'A^{-1}xB^{\frac{1}{2}} + I_s\right|.$$

Лемма 1 доказана.

Найдем выражения для функции плотности $f_U(u)$ в некоторых частных случаях.

Для случая $s = 1$, $B = 1$, $M = 0$ в соответствии с теоремой 1 имеем

$$f_U\,(u) = (2\pi)^{-\frac{1}{2}\,r}\,\Gamma_r^*\,(b)\,(\Gamma_r^*\,(a))^{-1}\,|A|^{-\frac{1}{2}}\,\cdot$$

$$\cdot \prod_{j=0}^{r-1} \left(\left| A^{[r-j]} \right|^{-1} \left| A^{[r-j]} + \frac{1}{2} u^{[r-j]} \left(u^{[r-j]} \right)' \right| \right)^{b_j - b_{j+1}},$$

где $u^{[k]}$ – k-мерный вектор, состоящий из первых k компонент r-мерного вектора u, $k = 1, \ldots, r$. Применяя лемму 1, получаем

$$f_U(u) = (2\pi)^{-\frac{1}{2}r} \Gamma_r^*(b) \left(\Gamma_r^*(a) \right)^{-1} |A|^{-\frac{1}{2}} \cdot$$

$$\cdot \prod_{j=0}^{r-1} \left(1 + \frac{1}{2} \left(u^{[r-j]} \right)' \left(A^{[r-j]} \right)^{-1} u^{[r-j]} \right)^{b_j - b_{j+1}}$$

– функцию плотности многомерного t-распределения с вектором степеней свободы, найденную другим способом в [4].

Для случая $a_1 = \ldots = a_r > \dfrac{r-1}{2}$, вновь считая для краткости $M = 0$, в соответствии с теоремой 1 имеем

$$f_U(u) =$$

$$= (2\pi)^{-\frac{1}{2}rs} \Gamma_r^*(b) \left(\Gamma_r^*(a) \right)^{-1} |A|^{b_1 - \frac{1}{2}s} |B|^{-\frac{1}{2}r} \left| A + \frac{1}{2} u B^{-1} u' \right|^{-b_1}.$$

Применяя лемму 1, получаем

$$f_U(u) =$$

$$= (2\pi)^{-\frac{1}{2}rs} \Gamma_r^*(b) \left(\Gamma_r^*(a) \right)^{-1} |A|^{-\frac{1}{2}s} |B|^{b_1 - \frac{1}{2}r} \left| B + \frac{1}{2} u' A^{-1} u \right|^{-b_1}$$

– обычную функцию плотности матричного t-распределения.

Здесь надо отметить одно принципиальное отличие обычного матричного t-распределения (то есть матричного t-распределения с числом степеней свободы) от матричного

t-распределения с вектором степеней свободы. Для обычного матричного t-распределения из положительности b_1 и положительной определенности матриц A и B следует, что последний сомножитель в приведенном выше выражении для функции плотности $f_U(u)$ стремится к нулю на бесконечности. Для матричного t-распределения с вектором степеней свободы (функция плотности приводится в формулировке теоремы 1) нет требования, чтобы все разности $b_j - b_{j+1}$ были отрицательными. Поэтому какие-то из сомножителей могут не стремиться к нулю на бесконечности.

Чтобы получить выражение для функции плотности $f_U(u)$ еще в одном частном случае, рассмотрим случайную матрицу

$$V = P^{-1}\,(U - M)\,Q^{-1},$$

где

$$A = P\,P', \quad B = Q'Q,$$

P – $r \times r$ матрица, Q – $s \times s$ матрица, причем матрица P нижняя треугольная.

Функция плотности случайной матрицы V имеет вид

$$f_V(v) = f_U(u(v))\,J\,(u; v)\,.$$

Согласно (4)

$$J\,(u; v) = |P|^s\,|Q|^r = |A|^{\frac{1}{2}\,s}\,|B|^{\frac{1}{2}\,r}\,.$$

Заметим, что

$$(u - M)\,B^{-1}\,(u - M)' = Pvv'P'.$$

Поэтому

$$f_V\,(v) = (2\pi)^{-\frac{1}{2}\,rs}\,\Gamma_r^*\,(b)\,(\Gamma_r^*\,(a))^{-1}\,\cdot$$

$$\cdot \prod_{j=0}^{r-1} \left(\left| A^{[r-j]} \right|^{-1} \left| A^{[r-j]} + \frac{1}{2} \left(Pvv'P' \right)^{[r-j]} \right| \right)^{b_j - b_{j+1}} .$$

Непосредственной проверкой устанавливается, что для любой $r \times r$ матрицы C, для любой нижней треугольной $r \times r$ матрицы L, для любой верхней треугольной $r \times r$ матрицы U и для любого $k = 1, \ldots, r$

$$(LC)^{[k]} = L^{[k]} C^{[k]}, \quad (CU)^{[k]} = C^{[k]} U^{[k]}.$$

Поэтому при $j = 0, \ldots, r-1$

$$A^{[r-j]} + \frac{1}{2} \left(Pvv'P' \right)^{[r-j]} = P^{[r-j]} \left(I_{r-j} + \frac{1}{2} \left(vv' \right)^{[r-j]} \right) \left(P^{[r-j]} \right)' .$$

Отсюда

$$f_V(v) =$$

$$= (2\pi)^{-\frac{1}{2} rs} \, \Gamma_r^*(b) \left(\Gamma_r^*(a) \right)^{-1} \prod_{j=0}^{r-1} \left| I_{r-j} + \frac{1}{2} \left(vv' \right)^{[r-j]} \right|^{b_j - b_{j+1}} .$$

Последнее выражение, разумеется, совпадает с выражением для функции плотности $f_U(u)$ при $A = I_r$, $B = I_s$, $M = 0$ и может рассматриваться как стандартизированная форма матричного t-распределения с вектором степеней свободы.

3. Байесовская многомерная регрессионная модель

Пусть y – это $n \times m$ матрица, элементы которой представляют n наблюдений, каждое наблюдение – m-мерный вектор. Таким образом, наблюдение представляется строкой матрицы y. Также через y будем обозначать случайную матрицу той же размерности при построении вероятностной модели для данного набора наблюдений. Наконец, через

y будет обозначаться аргумент функции плотности случайной матрицы y. Предположим, что вероятностная модель для рассматриваемого набора наблюдений имеет вид

$$y = x\,\theta + \varepsilon,$$

где

x – $n \times q$ матрица, состоящая из известных элементов,

θ – $q \times m$ случайная матрица,

ε – $n \times m$ случайная матрица,

$q \leq n$.

Будем предполагать, что строки случайной матрицы ε – это независимые m-мерные случайные вектора, и что каждый из этих векторов имеет m-мерное нормальное распределение со средним 0 и с ковариационной матрицей T^{-1}. Здесь T – положительно определенная $m \times m$ матрица, называемая матрицей точности.

Нетрудно увидеть, что то же условие можно выразить словами, что ε имеет матричное нормальное распределение со средним 0 и с ковариационной матрицей $I_n \otimes T^{-1}$. В соответствии с формулой (15) из раздела 4 это означает, что функция плотности случайной матрицы ε имеет вид

$$(2\pi)^{-\frac{1}{2}\,mn}\,|T|^{\frac{1}{2}\,n}\,\mathrm{etr}\left(-\frac{1}{2}\,\varepsilon\,T\varepsilon'\right).$$

Здесь обозначение ε используется не только для случайной матрицы, но и для аргумента ее функции плотности.

Если матрицы θ и T известны, то функция плотности случайной матрицы y имеет вид

$$(2\pi)^{-\frac{1}{2}\,mn}\,|T|^{\frac{1}{2}\,n}\,\mathrm{etr}\left(-\frac{1}{2}\,(y - x\,\theta)\,T(y - x\,\theta)'\right) =$$

$$= (2\pi)^{-\frac{1}{2}\,mn}\,|T|^{\frac{1}{2}\,n}\,\mathrm{etr}\left(-\frac{1}{2}\,(y - x\,\theta)'(y - x\,\theta)\,T\right).$$

Допустим, что $q \times q$ матрица $x'x$ невырожденная, и рассмотрим $q \times m$ матрицу

$$\hat{\theta} = (x'x)^{-1}\,x'y.$$

Разумеется, матрица $\hat{\theta}$ случайная, когда матрица y случайная.

Нетрудно убедиться, что

$$(y - x\,\theta)'(y - x\,\theta) = P + Q,$$

где

$$\begin{aligned} P &= (y - x\,\hat{\theta})'(y - x\,\hat{\theta}), \\ Q &= (\theta - \hat{\theta})'x'x(\theta - \hat{\theta}). \end{aligned} \tag{7}$$

Соответственно, для условной функции плотности случайной матрицы y получаем выражение

$$f(y|\theta, T) = (2\pi)^{-\frac{1}{2}\,mn}\,|T|^{\frac{1}{2}\,n}\,\mathrm{etr}\left(-\frac{1}{2}\,(P + Q)\,T\right). \tag{8}$$

Подчеркнем, что матрица P от матрицы θ не зависит.

Если рассматривать матрицы θ и T как случайные, то для совместной функции плотности случайных матриц y, θ и T могут быть записаны два различных выражения через условные и маргинальные функции плотности:

$$f(y, \theta, T) = f(y|\theta, T) f(\theta, T) = f(\theta, T|y) f(y).$$

Отсюда

$$f(\theta, T|y) \propto f(\theta, T) f(y|\theta, T). \tag{9}$$

Последняя запись является выражением апостериорной функции плотности случайных матриц θ, T через априорную функцию плотности и через условную функцию плотности.

Будем считать априорную функцию плотности расплывчатой относительно θ и информативной относительно T (см. [2]),

$$f(\theta, T) = f(T), \tag{10}$$

а относительно функции $f(T)$ сделаем предположение, что она имеет вид (2), то есть является функцией плотности гамма-распределения с векторным параметром:

$$f(T) = \gamma_{a,R}\, \mathrm{etr}\,(-RT) \prod_{j=1}^{m} \left|T_{[j]}\right|^{a_j - a_{j+1}}, \tag{11}$$

где R – положительно определенная $m \times m$ матрица, компоненты m-мерного вектора $a = (a_1, \ldots, a_m)$ удовлетворяют условию

$$a_j > \frac{j-1}{2}, \quad j = 1, \ldots, m,$$

$a_0 = 0$, $a_{m+1} = \dfrac{m+1}{2}$.

Из (8), (9), (10), (11) получаем выражение для совместной апостериорной функции плотности случайных матриц θ и T:

$$f(\theta, T|y) \propto \mathrm{etr}\left(-\left(R + \frac{1}{2}(P+Q)\right)T\right) \prod_{j=1}^{m} \left|T_{[j]}\right|^{c_j - c_{j+1}}, \tag{12}$$

где $c = (c_1, \ldots, c_m)$,

$$c_j = a_j + \frac{n}{2}, \quad j = 1, \ldots, m,$$

$c_0 = 0$, $c_{m+1} = \dfrac{m+1}{2}$. Будем использовать также обозначение $d = (d_1, \ldots, d_m)$, где

$$d_j = c_j - \frac{q}{2}, \quad j = 1, \ldots, m,$$

$d_0 = 0$, $d_{m+1} = \dfrac{m+1}{2}$. Введем обозначения

$$A = R + \frac{1}{2}\,P, \quad B = \left(x'x\right)^{-1}.$$

Теорема 2. Маргинальное апостериорное распределение случайной матрицы θ является матричным t-распределением с вектором степеней свободы

$$f\left(\theta|y\right) = \left(2\pi\right)^{-\frac{1}{2}\,mq}\,\Gamma_m^*\left(c\right)\left(\Gamma_m^*\left(d\right)\right)^{-1}\left|A\right|^{-\frac{1}{2}\,q}\left|B\right|^{-\frac{1}{2}\,m}\cdot$$

$$\cdot\prod_{j=0}^{m-1}\left|I_{m-j} + \frac{1}{2}\left(A^{[m-j]}\right)^{-1}\left(\left(\theta - \hat{\theta}\right)' B^{-1}\left(\theta - \hat{\theta}\right)\right)^{[m-j]}\right|^{c_j - c_{j+1}}.$$

Доказательство. Интегрируя функцию (12) по области $T > 0$ и пользуясь видом функции плотности гамма-распределения с векторным параметром (2), а затем выражением (3), получаем

$$\int_{T>0} \operatorname{etr}\left(-\left(R + \frac{1}{2}\left(P+Q\right)\right)T\right)\prod_{j=1}^{m}\left|T_{[j]}\right|^{c_j - c_{j+1}} dT =$$

$$= \frac{1}{\gamma_{c,R+\frac{1}{2}\,(P+Q)}} = \Gamma_m^*\left(c\right)\prod_{j=0}^{m-1}\left|\left(A + \frac{1}{2}\,Q\right)^{[m-j]}\right|^{c_j - c_{j+1}}.$$

Из (12) и (7) следует, что

$$f(\theta|y) \propto$$

$$\propto \prod_{j=0}^{m-1} \left| I_{m-j} + \frac{1}{2} \left(A^{[m-j]}\right)^{-1} \left(\left(\theta - \hat{\theta}\right)' B^{-1} \left(\theta - \hat{\theta}\right)\right)^{[m-j]} \right|^{c_j - c_{j+1}} .$$

Теорема 2 доказана.

Теорема 3. Маргинальное апостериорное распределение случайной матрицы T является матричным гамма-распределением с векторным параметром

$$f\left(T|y\right) = \gamma_{d,A} \operatorname{etr}\left(-AT\right) \prod_{j=1}^{m} \left|T_{[j]}\right|^{d_j - d_{j+1}} .$$

Доказательство. Интегрируя правую часть выражения (12) по θ (интегрирование ведется по mq-мерному евклидову пространству) и пользуясь формулой (7), получаем

$$\operatorname{etr}\left(-AT\right) \prod_{j=1}^{m} \left|T_{[j]}\right|^{c_j - c_{j+1}} \cdot$$

$$\cdot \int \operatorname{etr}\left(-\frac{1}{2}\left(\theta - \hat{\theta}\right)' B^{-1} \left(\theta - \hat{\theta}\right) T\right) d\theta \propto$$

$$\propto \operatorname{etr}\left(-AT\right) \prod_{j=1}^{m} \left|T_{[j]}\right|^{c_j - c_{j+1}} \left|T\right|^{-\frac{1}{2}q} ,$$

при последнем переходе использовано выражение (15) из раздела 4 для функции плотности матричного нормального распределения. Поэтому

$$f\left(T|y\right) \propto \operatorname{etr}\left(-AT\right) \prod_{j=1}^{m} \left|T_{[j]}\right|^{d_j - d_{j+1}} .$$

Теорема 3 доказана.

4. Матричное нормальное распределение

Будем считать, что $r \times s$ матрица имеет вид

$$X = \begin{pmatrix} x_{11} & \dots & x_{1s} \\ \dots & \dots & \dots \\ x_{r1} & \dots & x_{rs} \end{pmatrix},$$

то есть первый индекс у каждого элемента матрицы обозначает строку, в которой этот элемент находится, а второй индекс – столбец.

Определим кронекерово произведение $r \times r$ матрицы C и $s \times s$ матрицы B:

$$C \otimes B = \begin{pmatrix} c_{11}B & c_{12}B & \dots & c_{1r}B \\ c_{21}B & c_{22}B & \dots & c_{2r}B \\ \dots & \dots & \dots & \dots \\ c_{r1}B & c_{r2}B & \dots & c_{rr}B \end{pmatrix}.$$

Таким образом, кронекерово произведение – это $rs \times rs$ матрица. (Кронекерово произведение можно определять и для неквадратных матриц C и B, см., например, [3, стр. 19].)

В данном разделе будем считать, что индексы i и j меняются от 1 до r; индексы k и l меняются от 1 до s; индексы m и n меняются от 1 до rs; и между этими индексами существует связь

$$m = (i-1)s + k; \quad n = (j-1)s + l. \tag{13}$$

Элемент (m, n) матрицы $C \otimes B$ – это $c_{i\,j} b_{k\,l}$.

Лемма 2. Пусть C и D – $r \times r$ матрицы, B и E – $s \times s$ матрицы. Тогда

$$(C \otimes B)\,(D \otimes E) = CD \otimes BE.$$

Доказательство. Пусть $F = C \otimes B$; $G = D \otimes E$; $H = FG$. Тогда

$$h_{mn} = \sum_{\tau=1}^{rs} f_{m\tau} g_{\tau n}.$$

Если

$$\tau = (\alpha - 1)s + \gamma,$$

где $\alpha = 1, \ldots, r$; $\gamma = 1, \ldots, s$; а выражения для m и n через i, j, k, l даются формулами (13), то

$$f_{m\tau} = c_{i\alpha} b_{k\gamma}, \quad g_{\tau n} = d_{\alpha j} e_{\gamma l}.$$

Поэтому

$$h_{mn} = \sum_{\alpha=1}^{r} \sum_{\gamma=1}^{s} c_{i\alpha} b_{k\gamma} d_{\alpha j} e_{\gamma l} = \left(\sum_{\alpha=1}^{r} c_{i\alpha} d_{\alpha j} \right) \left(\sum_{\gamma=1}^{s} b_{k\gamma} e_{\gamma l} \right).$$

Если ввести в рассмотрение матрицы $U = CD$, $V = BE$, то полученное соотношение можно записать в виде

$$h_{mn} = u_{i\,j} v_{k\,l}$$

Лемма 2 доказана.

Лемма 3. Если матрицы C и B невырожденные, то и матрица $C \otimes B$ невырожденная и

$$(C \otimes B)^{-1} = C^{-1} \otimes B^{-1}.$$

Доказательство. По лемме 2

$$(C \otimes B) \left(C^{-1} \otimes B^{-1} \right) = I_r \otimes I_s = I_{rs}.$$

Лемма 3 доказана.

Рассмотрим подстановку q чисел

$$\begin{pmatrix} 1 & 2 & \dots & q \\ \alpha_1 & \alpha_2 & \dots & \alpha_q \end{pmatrix}$$

и соответствующую $q \times q$ перестановочную матрицу P с элементами

$$p_{mn} = \begin{cases} 1 & \text{при } n = \alpha_m \\ 0 & \text{при других } n \end{cases}$$

Нетрудно увидеть, что для любой $q \times q$ матрицы A с элементами a_{mn} элемент (m, n) матрицы PAP' – это $a_{\alpha_m \alpha_n}$.

Рассмотрим подстановку rs чисел, которая каждому числу $(k-1)r+i$ ставит в соответствие число $(i-1)s+k$; здесь $i = 1, \dots, r$; $k = 1, \dots, s$. Нетрудно убедиться, что при соответствующей этой подстановке перестановочной матрице P для любой $r \times r$ матрицы C имеет место соотношение

$$P\left(I_s \otimes C\right)P' = C \otimes I_s. \tag{14}$$

Лемма 4. Для определителей имеет место соотношение

$$|C \otimes B| = |C|^s\,|B|^r\,.$$

Доказательство. По лемме 2

$$C \otimes B = \left(C \otimes I_s\right)\left(I_r \otimes B\right).$$

$I_r \otimes B$ – это блочно-диагональная матрица, определитель которой равен произведению определителей блоков, поэтому

$$|I_r \otimes B| = |B|^r\,.$$

Аналогично $|I_s \otimes C| = |C|^s$. Поскольку любая перестановочная матрица является ортогональной, и ее определитель равен 1, из (14) находим

$$|C \otimes I_s| = |C|^s\,.$$

Лемма 4 доказана.

С каждой $r \times s$ матрицей X (см. начало этого раздела) свяжем $rs \times 1$ вектор

$$\mathrm{vec}(X) = \begin{pmatrix} x_1 \\ \vdots \\ x_s \end{pmatrix},$$

где x_k – k-й столбец матрицы X.

Используя соотношение (13), нетрудно получить, что

$$\mathrm{vec}(X')' \, (C \otimes B) \, \mathrm{vec}(X') = \sum_{i=1}^{r} \sum_{j=1}^{r} \sum_{k=1}^{s} \sum_{l=1}^{s} x_{ik} c_{ij} b_{kl} x_{jl}.$$

Из последней формулы вытекает, что имеет место следующая лемма.

Лемма 5.

$$\mathrm{vec}(X')' \, (C \otimes B) \, \mathrm{vec}(X') = \mathrm{tr} \, (BX'C'X) \, .$$

Нетрудно увидеть, что если $c_1, \ldots, c_r$ – собственные числа матрицы C; $b_1, \ldots, b_s$ – собственные числа матрицы B, то rs чисел $c_i b_k$, где $i = 1, \ldots, r$; $k = 1, \ldots, s$; являются собственными числами матрицы $C \otimes B$. Действительно, приведем матрицы C и B к жордановой нормальной форме

$$C = P_C J_C P_C^{-1}, \quad B = P_B J_B P_B^{-1}$$

(см., например, [1, гл. 6]). Здесь P_C и J_C – $r \times r$ матрицы, P_B и J_B – $s \times s$ матрицы. Поскольку матрицы J_C и J_B верхние треугольные, матрица $J_C \otimes J_B$ также является верхней треугольной. На главной диагонали матрицы J_C стоят числа $c_1, \ldots, c_r$. На главной диагонали матрицы J_B стоят числа $b_1, \ldots, b_s$. Поэтому на главной диагонали матрицы

$J_C \otimes J_B$ стоят числа $c_i b_k$, $i = 1, \ldots, r$; $k = 1, \ldots, s$, которые и являются собственными числами этой матрицы. Используя леммы 2 и 3, получаем

$$C \otimes B = \left(P_C J_C P_C^{-1}\right) \otimes \left(P_B J_B P_B^{-1}\right) =$$

$$= \left(P_C \otimes P_B\right) \left(J_C P_C^{-1} \otimes J_B P_B^{-1}\right) =$$

$$= \left(P_C \otimes P_B\right) \left(J_C \otimes J_B\right) \left(P_C^{-1} \otimes P_B^{-1}\right) =$$

$$= \left(P_C \otimes P_B\right) \left(J_C \otimes J_B\right) \left(P_C \otimes P_B\right)^{-1}.$$

Отсюда следует, что собственные числа матрицы $C \otimes B$ совпадают с собственными числами матрицы $J_C \otimes J_B$.

Если матрицы C и B имеют положительные собственные числа, то, следовательно, и собственные числа матрицы $C \otimes B$ положительны. Если матрицы C и B симметричные, то, очевидно, и матрица $C \otimes B$ симметричная. Поэтому имеет место следующая лемма.

Лемма 6. Если матрицы C и B положительно определенные, то и матрица $C \otimes B$ положительно определенная.

Пусть μ – произвольная $r \times s$ матрица. Матрицы C и B положительно определенные.

Определение. $r \times s$ случайная матрица Z имеет матричное нормальное распределение со средним μ и с ковариационной матрицей $C \otimes B$, если $rs \times 1$ случайный вектор $\mathrm{vec}\,(Z')$ имеет rs-мерное нормальное распределение со средним $\mathrm{vec}\,(\mu')$ и с ковариационной матрицей $C \otimes B$.

Пусть $r \times s$ случайная матрица Z имеет матричное нормальное распределение со средним μ и с ковариационной матрицей $C \otimes B$. Применяя леммы 3, 4, 5 и учитывая, что след произведения двух матриц не меняется при изменении порядка сомножителей, получаем выражение для функции плотности случайной матрицы Z:

$$
\begin{aligned}
g\,(z) = (2\pi)^{-\frac{1}{2}\,rs}\,|C\otimes B|^{-\frac{1}{2}}\,\cdot \\
\cdot\exp\left(-\frac{1}{2}\,\mathrm{vec}((z-\mu)')'\,(C\otimes B)^{-1}\,\mathrm{vec}((z-\mu)')\right) = \\
= (2\pi)^{-\frac{1}{2}\,rs}\,|C|^{-\frac{1}{2}\,s}\,|B|^{-\frac{1}{2}\,r}\exp\left(-\frac{1}{2}\,\mathrm{vec}((z-\mu)')'\,\cdot\right. \\
\left.\cdot\left(C^{-1}\otimes B^{-1}\right)\mathrm{vec}((z-\mu)')\right) = \qquad (15)\\
= (2\pi)^{-\frac{1}{2}\,rs}\,|C|^{-\frac{1}{2}\,s}\,|B|^{-\frac{1}{2}\,r}\,\mathrm{etr}\left(-\frac{1}{2}\,B^{-1}(z-\mu)'\,C^{-1}(z-\mu)\right) = \\
= (2\pi)^{-\frac{1}{2}\,rs}\,|C|^{-\frac{1}{2}\,s}\,|B|^{-\frac{1}{2}\,r}\,\mathrm{etr}\left(-\frac{1}{2}\,C^{-1}(z-\mu)\,B^{-1}(z-\mu)'\right).
\end{aligned}
$$

Библиографический список

[1] Гантмахер Ф.Р. Теория матриц. 2-е изд. М.: Наука, 1966.

[2] Зельнер А. Байесовские методы в эконометрии. М.: Статистика, 1977.

[3] Маркус М., Минк Х. Обзор по теории матриц и матричных неравенств. М.: Наука, 1972.

[4] Шведов А.С. Бета-распределение случайной матрицы и его применение в модели состояние-наблюдение. Препринт WP2/2009/01. М.: ГУ ВШЭ, 2009.

[5] Bellman R. A generalization of some integral identities due to Ingham and Siegel // Duke Math. J., 23 (1956), 571 – 577.

[6] Box G.E.P., Tiao G.C. Bayesian inference in statistical analysis. Reading (MA): Addison-Wesley, 1973.

[7] de Waal D.J. Distributions connected with a multivariate beta statistics // Annals of Mathematical Statistics, 41 (1970), 1091 – 1095. Correction 42 (1971), 2165 – 2166.

[8] Gupta A.K., Nagar D.K. Matrix variate vistributions. N.Y.: Chapman & Hall, 1999.

[9] Muirhead R.J. Aspects of multivariate statistical theory. Wiley-Interscience, 2005.

[10] Olkin I. A class of integral identities with matrix argument // Duke Math. J., 26 (1959), 207 – 213.

Препринт WP2/2010/01
Серия WP2
Количественный анализ в экономике

А.С. Шведов

t-Распределение случайной матрицы и его применение в регрессионной модели

Выпускающий редактор *А.В. Заиченко*
Технический редактор *Ю.Н. Петрина*

Отпечатано в типографии Государственного университета –
Высшей школы экономики с представленного оригинал-макета.
Формат 60×84 1/16. Бумага офсетная. Тираж 150 экз. Уч.-изд. л. 1,6
Усл. печ. л. 1,62. Заказ № . Изд. № 1161

Государственный университет – Высшая школа экономики.
125319, Москва, Кочновский проезд, 3
Типография Государственного университета – Высшей школы экономики.

Тел.: (495) 772-95-71; 772-95-73

Для заметок