Л. Н. Слуцкин

Байесовский анализ, когда оцениваемый параметр является случайным нормальным процессом[1]

Рассмотрена задача байесовского оценивания последовательности неизвестных средних значений $\theta_1, \theta_2, ..., \theta_k, ...$ по имеющимся наблюдениям $X_1, X_2, \ldots, X_k, ...$ в ситуации, когда наблюдения $X_1, X_2, \ldots, X_k$ подчиняются многомерному нормальному распределению с вектором средних $(\theta_1, \theta_2, ..., \theta_k)$ и известной ковариационной матрицей. Предполагается, что параметры $\theta_1, \theta_2, ..., \theta_k, ...$ образуют гауссовский процесс. Доказывается сходимость (при $k \to \infty$) ковариационных матриц частного апостериорного распределения последовательности параметров; подробно анализируется пример, в котором размерность наблюдений $X_1, X_2, \ldots, X_k, ...$ полагается равной единице, а последовательность $\theta_1, \theta_2, ..., \theta_k, ...$ образует гауссовский процесс авторегрессии первого порядка.

Ключевые слова: асимптотическая ковариационная матрица, формула Байеса, гауссовский нормальный процесс, частное (маргинальное) апостериорное распределение.

1. Введение[2]

Предположим, что случайный вектор $X_i,\ i = 1, 2, ...$ генерируется некоторым многомерным распределением с соответствующей функцией правдоподобия $L(\theta_i, X_i)$, $X_i \in R^{n_1}$, $\theta_i \in R^{n_2}$ для всех i, где θ_i — случайный вектор, имеющий априорное распределение $p_i(\theta_i)$. Также предположим, что для разных i не только сами θ_i, но и их распределения $p_i(\theta_i)$, могут быть различными. Допустим, что $X_1, X_2, \ldots, X_k, ...$ условно (по $\theta_1, \theta_2, ..., \theta_k, ...$) независимы, и $\theta_1, \theta_2, ..., \theta_k, ...$ есть случайный процесс с совместными априорными распределениями $p_{\theta_1}, ..., p_{\theta_1, ..., \theta_k}, ...$.

Подобные задачи возникают на практике, когда $i = 1, 2, ...$ представляют моменты времени. В таком случае X_i будет случайным многомерным наблюдением из распределения, зависящего от неизвестного многомерного параметра θ_i, который, в свою очередь, при байесовском подходе является случайной величиной. Естественно предположить, что различным моментам времени соответствуют различные значения θ_i, которые, вообще говоря, могут быть взаимозависимыми и представлять собой реализацию некоторого случайного процесса. Как показывает пример гауссовского процесса авторегрессии первого порядка для θ_i, приведенный в разделе 3, асимптотическая апостериорная оценка дисперсии θ_i является более точной, чем при обычном (байесовском) подходе.

1 Автор признателен профессору С. А. Айвазяну за внимательное прочтение рукописи и ценные замечания.

2 Предполагается, что читатель знаком с основными положениями байесовской теории (например, в пределах (Айвазян, 2008) или (Де Гроот, 1974)).

Из взаимной условной независимости $X_1, X_2, \ldots, X_k$ следует, что совместное апостериорное распределение многомерного параметра $(\theta_1, \theta_2, ..., \theta_k)$ при заданных наблюдениях $(X_1, X_2, \ldots, X_k)$ определяется, в соответствии с формулой Байеса, соотношением

$$p_{\theta_1,...,\theta_k|X_1,...,X_k} \propto L(\theta_1; X_1) \cdot ... \cdot L(\theta_k; X_k) \cdot p_{\theta_1,...,\theta_k}. \tag{1.1}$$

В частности, для $i = 1$ имеем:

$$p_{\theta_1|X_1} \propto L(\theta_1; X_1) \cdot p_{\theta_1}. \tag{1.2}$$

Естественно, возникает вопрос, можно ли считать $\theta_1 | X_1, ..., \theta_k | X_k, ...$ случайным процессом с совместными распределениями, заданными формулой (1.1). Ответ, как это следует из примера (1.3), отрицательный. Дело в том, что распределения $p_{\theta_1|X_1}, ..., p_{\theta_1,...,\theta_k|X_1,...,X_k}, ...$ не являются согласованными, т. е. не удовлетворяют условию согласованности Колмогорова для частных (маргинальных) распределений (Боровков, 2009). Однако, в случае, когда $\theta_1, \theta_2, ..., \theta_k, ...$ следуют гауссовскому процессу, соответствующие ковариационные матрицы сходятся. Другими словами, когда число наблюдений $k \to \infty$, мы получаем апостериорную информацию о ковариациях $\operatorname{cov}(\theta_i, \theta_j)$, соответствующих наблюденным значениям $X_1, X_2, \ldots, X_k, ...$ при любых i и j.

Поясним вышесказанное следующим примером. Предположим, что скалярная случайная величина определена как $X_i \sim N(\theta_i, \sigma_X^2)$, $i = 1, 2, \ldots$. Параметр θ является процессом авторегрессии первого порядка $AR(1)$:

$$\theta_{i+1} = \rho\theta_i + e_{i+1}, \;\; i = 1, 2, ..., \tag{1.3}$$

где $|\rho| < 1$, $\rho \neq 0$, $e_i \sim N(0, \sigma_e^2)$ — гауссовский белый шум. Из стационарности θ следует, что $E(\theta) = 0$, $\sigma_\theta^2 = \sigma_e^2 / (1 - \rho^2)$. Рассмотрим первые три члена временного ряда (1.3): θ_1, θ_2, θ_3. Обратными к ковариационным матрицам распределений p_{θ_1}, p_{θ_1,θ_k}, $p_{\theta_1,\theta_2,\theta_3}$ будут, соответственно, матрицы (Айвазян, Иванова, 2007):

$$\Sigma_{\theta_1}^{-1} = [\sigma_\theta^{-2}], \quad \Sigma_{\theta_1,\theta_2}^{-1} = \sigma_e^{-2}\begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix}, \quad \Sigma_{\theta_1,\theta_2,\theta_3}^{-1} = \sigma_e^{-2}\begin{bmatrix} 1 & -\rho & 0 \\ -\rho & 1+\rho^2 & -\rho \\ 0 & -\rho & 1 \end{bmatrix}. \tag{1.4}$$

Так как

$$L(\theta_i, X_i) = \frac{1}{\sqrt{2\pi}\sigma_X} e^{-\frac{(X_i - \theta_i)^2}{2\sigma_X^2}}, \tag{1.5}$$

то из формулы (1.1) следует:

$$\Sigma_1^{-1} = [(\sigma_\theta^{-2} + \sigma_X^{-2})], \quad \Sigma_2^{-1} = \sigma_e^{-2}\begin{bmatrix} 1+c & -\rho \\ -\rho & 1+c \end{bmatrix}, \quad \Sigma_3^{-1} = \sigma_e^{-2}\begin{bmatrix} 1+c & -\rho & 0 \\ -\rho & 1+\rho^2+c & -\rho \\ 0 & -\rho & 1+c \end{bmatrix}, \tag{1.6}$$

где $\Sigma_1, \Sigma_2, \Sigma_3$ — ковариационные матрицы апостериорных распределений $p_{\theta_1|X_1}$, $p_{\theta_1,\theta_2|X_1,X_2}$, $p_{\theta_1,\theta_2,\theta_3|X_1,X_2,X_3}$ и $c = \sigma_e^2 \sigma_X^{-2}$. Отсюда получаем[3]

$$\Sigma_1 = \left[\frac{\sigma_\theta^2 \sigma_X^2}{\sigma_\theta^2 + \sigma_X^2} \right], \ \Sigma_2 = \frac{\sigma_e^2}{(1+c)^2 - \rho^2} \begin{bmatrix} 1+c & \rho \\ \rho & 1+c \end{bmatrix},$$
$$\Sigma_3 = \frac{\sigma_e^2}{\Delta} \begin{bmatrix} (1+c)^2 + \rho^2 c & & \\ & (1+c)^2 & \\ & & (1+c)^2 + \rho^2 c \end{bmatrix}, \tag{1.7}$$

где Δ — определитель матрицы $\sigma_e^2 \Sigma_3^{-1}$.

В частности, из (1.7) следует, что дисперсия $\sigma^2_{\theta_1|X_1} = \dfrac{\sigma_\theta^2 \sigma_X^2}{\sigma_\theta^2 + \sigma_X^2}$ апостериорного распределения $p_{\theta_1|X_1}$ больше, чем соответствующая дисперсия

$$\sigma^2_{\theta_1|X_1,X_2} = \frac{\sigma_e^2 (1+c)}{(1+c)^2 - \rho^2} \tag{1.8}$$

частного распределения $\theta_1 \mid X_1, X_2$ по отношению к $p_{\theta_1,\theta_2|X_1,X_2}$. Тем самым показано, что распределения $p_{\theta_1|X_1}$ и $p_{\theta_1,\theta_2|X_1,X_2}$ не согласованы. Интересно также отметить, что апостериорный коэффициент корреляции

$$\rho_{\theta_1,\theta_2|X_1,X_2} = \frac{\rho}{1+c} \tag{1.9}$$

меньше, чем априорный коэффициент, равный ρ. Особый интерес представляет матрица Σ_3. Так как ее диагональные элементы не равны между собой, можно заключить, что апостериорное распределение $p_{\theta_1,\theta_2,\theta_3|X_1,X_2,X_3}$ не является стационарным.

Мы продолжим апостериорный анализ процесса (1.3) в третьем разделе, где будут получены в явном виде элементы асимптотических ковариационных матриц апостериорных распределений оцениваемых параметров.

2. Гауссовские процессы

Предположим, что $X_i \sim N(\theta_i, \Sigma_X)$, $i = 1, 2, \ldots$, где Σ_X — известная (не зависящая от i) ковариационная матрица порядка n. Пусть $\theta_1, \theta_2, \ldots, \theta_k, \ldots$ — гауссовский процесс, т. е. $\theta_i \sim N(\theta_i^0, \Sigma_{\theta_i})$, θ_i^0 — известный вектор из R^n, Σ_{θ_i} — известная ковариационная матрица порядка n, $i = 1, 2, \ldots$. θ_i^0 и Σ_{θ_i} могут быть различными для разных i. Тогда совместное априорное распределение $\theta_1, \ldots, \theta_k$: $p_{\theta_1,\ldots,\theta_k} = N(\theta^{0,k}, \Sigma_{\theta,k})$, где вектор $\theta^{0,k} = (\theta_1^0, \ldots, \theta_k^0)$, $\theta^{0,k} \in R^{kn}$, и $\Sigma_{\theta,k}$ — известная ковариационная матрица порядка kn, $k = 1, 2, \ldots$. Совместное апостериорное распределение $\theta_1, \ldots, \theta_k$ при заданных $X_1, \ldots, X_k$ определяется формулой

[3] Так как данный пример служит исключительно иллюстративным целям, при описании матрицы Σ_3 мы ограничились лишь диагональными элементами, оставляя другие ячейки не заполненными.

$$p_{\theta_1,\dots,\theta_k|X_1,\dots,X_k} = N(Y^k, \Sigma_k), \tag{2.1}$$

где

$$Y^k = \Sigma_k (\Sigma_{X,k}^{-1} X^k + \Sigma_{\theta,k}^{-1} \theta^{0,k}), \tag{2.2}$$

$\Sigma_{X,k} = I_k \otimes \Sigma_X$ — матрица порядка kn, вдоль диагонали которой расположены k копий матрицы Σ_X, а остальные элементы — нули, вектор $X^k = (X_1, \dots, X_k)$, $X^k \in R^{kn}$, и

$$\Sigma_k^{-1} = \Sigma_{X,k}^{-1} + \Sigma_{\theta,k}^{-1}. \tag{2.3}$$

Заметим, что апостериорная ковариационная матрица Σ_k не зависит от наблюдений $X_1, \dots, X_k$.

Формула (2.1) является многомерным аналогом известной в байесовском анализе формулы для случая, когда X и θ — скаляры, а $k=1$ (Де Гроот, 1974):

$$Y = (\sigma_X^{-2} + \sigma_\theta^{-2})^{-1} (\sigma_X^{-2} X + \sigma_\theta^{-2} \theta^0), \quad \sigma^{-2} = \sigma_X^{-2} + \sigma_\theta^{-2}. \tag{2.4}$$

Доказательство формул (2.1)–(2.3) будет дано в Приложении 1[4].

Рассмотрим распределения $p_{\theta_1,\dots,\theta_k}$, $p_{\theta_1,\dots,\theta_{k_1}}$, $k_1 > k$. Матрица Z_{θ,k_1} получается из матрицы $\Sigma_{\theta,k}$ добавлением $(k_1 - k)n$ столбцов и строк снизу и справа от матрицы $\Sigma_{\theta,k}$. $\Sigma_{\theta,k}$ соответствует ведущему главному минору Z_{θ,k_1} порядка kn. Пусть $Z_{k_1}^{(k)}$ будет подматрицей Z_{k_1}, соответствующей ее ведущему главному минору порядка kn. $Z_{k_1}^{(k)}$ является ковариационной матрицей частного апостериорного распределения $\theta_1,\dots,\theta_k | X_1,\dots,X_{k_1}$ по отношению к $p_{\theta_1,\dots,\theta_{k_1}}$. Следующие две леммы следуют из свойств положительно (неотрицательно) определенных матриц (Магнус, Нейдеккер, 2002, глава 1).

Лемма 2.1. *Справедливо неравенство:*

$$\Sigma_k \geq Z_{k_1}^{(k)}. \tag{2.5}$$

*Примечание***.** Везде в дальнейшем неравенство $A \geq B$ означает, что матрица $A - B$ неотрицательно определена.

Смысл леммы 2.1 заключается в том, что, если k фиксировано, то при добавлении новых наблюдений мы получаем более (не менее) точную информацию о совместном апостериорном распределении $\theta_1,\dots,\theta_k$.

Следствие. $\Sigma_k \geq Z_{k+1}^{(k)} \geq Z_{k+2}^{(k)} \geq \dots$.

Лемма 2.2. *Пусть* $A_1 \geq A_2 \geq \dots \geq A_k \geq \dots$ *— последовательность положительно определенных матриц одного и того же порядка* m*. Тогда существует предел* $\Sigma = \lim_{k\to\infty} A_k$*, где матрица* Σ *неотрицательно определена.*

Сформулируем главный результат, который является следствием лемм 2.1 и 2.2.

Теорема. *Для любого фиксированного* k *матрицы* $Z_{k_1}^{(k)}$ *сходятся (при* $k_1 \to \infty$*) к неотрицательно определенной матрице* Ω_k *порядка* k*.*

[4] Заметим, что формулы, подобные (2.1)–(2.3), приводятся (без доказательства) в монографии (Berger, 1985).

3. Пример

Приступим к более детальному анализу примера стационарного априорного процесса, представленного в конце первого раздела. Напомним, что наблюдается скалярная случайная величина $X_i \sim N(\theta_i, \sigma_X^2)$, $i = 1, 2, \ldots$. Параметр θ является процессом авторегрессии первого порядка $AR(1)$:

$$\theta_{i+1} = \rho\theta_i + e_{i+1}, \quad i = 1, 2, \ldots, \tag{3.1}$$

где $|\rho| < 1$, $\rho \neq 0$, $e_i \sim N(0, \sigma_e^2)$ — гауссовский белый шум, $E(\theta) = 0$, $\sigma_\theta^2 = \sigma_e^2 / (1 - \rho^2)$.

Обратной к ковариационной матрице процесса (3.1) для $\theta_1, \ldots, \theta_k$ будет следующая матрица (Айвазян, Иванова, 2007):

$$\Sigma_{\theta,k}^{-1} = \sigma_e^{-2} \begin{bmatrix} 1 & -\rho & 0 & 0 & \cdots & \cdots & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 & \cdots & \cdots & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & -\rho & \cdots & \cdots & 0 & 0 \\ 0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \cdots & -\rho & 1+\rho^2 & -\rho & 0 \\ 0 & 0 & \cdots & \cdots & 0 & -\rho & 1+\rho^2 & -\rho \\ 0 & 0 & \cdots & \cdots & 0 & 0 & -\rho & 1 \end{bmatrix}$$

Из (2.3) следует, что

$$\Sigma_k^{-1} = \sigma_X^{-2} I + \Sigma_{\theta,k}^{-1} = \sigma_e^{-2} cI + \Sigma_{\theta,k}^{-1}, \tag{3.2}$$

где I — единичная матрица порядка k,

$$c = \sigma_e^2 \sigma_X^{-2}. \tag{3.3}$$

Из (3.2) получаем:

$$\Sigma_k^{-1} = \sigma_e^{-2} \begin{bmatrix} 1+c & -\rho & 0 & 0 & \cdots & \cdots & 0 & 0 \\ -\rho & 1+\rho^2+c & -\rho & 0 & \cdots & \cdots & 0 & 0 \\ 0 & -\rho & 1+\rho^2+c & -\rho & \cdots & \cdots & 0 & 0 \\ 0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \cdots & -\rho & 1+\rho^2+c & -\rho & 0 \\ 0 & 0 & \cdots & \cdots & 0 & -\rho & 1+\rho^2+c & -\rho \\ 0 & 0 & \cdots & \cdots & 0 & 0 & -\rho & 1+c \end{bmatrix}. \tag{3.4}$$

Пусть $z_{i,j}^{(k)}$ — элемент матрицы Σ_k. Из результатов раздела 2 следует, что существует

$$z_{i,j} = \lim_{k\to\infty} z_{i,j}^{(k)}. \tag{3.5}$$

Лемма 3.1. *Рассмотрим асимптотическую апостериорную ковариационную матрицу процесса* (3.1) $\Sigma = \begin{bmatrix} z_{1,1} & z_{1,2} & \dots \\ z_{2,1} & z_{2,2} & \dots \\ \dots & \dots & \dots \end{bmatrix}$. *Тогда:*

а)
$$z_{1,1}^{-1} = \sigma_e^{-2}\left(1 + c - \frac{\rho^2}{x}\right), \tag{3.6}$$

где $x = \dfrac{1+\rho^2+c+\sqrt{(1+\rho^2+c)^2-4\rho^2}}{2}$, $1+c < x < 1+\rho^2+c$;

б) *при* $i > 1$
$$z_{i,i}^{-1} = z_{1,1}^{-1}\frac{x^{i-1}}{P_{i-1}}, \tag{3.7}$$

где P_{i-1} *— ведущий главный минор матрицы* Σ_k^{-1} $(k > i-1)$ *порядка* $(i-1)$;

в) $z_{i,i}^{-1}$, $i = 1, 2, \dots$ *является строго монотонно возрастающей последовательностью положительных чисел, т. е.*

$$z_{1,1}^{-1} < z_{2,2}^{-1} < \dots < z_{i,i}^{-1} < \dots; \tag{3.8}$$

г) *при* $j > i$

$$z_{i,j} = \frac{z_{1,1}\rho^{j-1}}{x^{j-1}},\ i = 1,\ z_{i,j} = \frac{z_{1,1}P_{i-1}\rho^{j-i}}{x^{j-1}},\ i > 1. \tag{3.9}$$

Примечание. Из (3.8) следует, что Σ не является ковариационной матрицей для стационарного процесса.

Нетрудно показать, что $\sigma_e^{-2}\left(1+c-\dfrac{\rho^2}{1+c}\right) \le z_{1,1}^{-1} \le \sigma_e^{-2}\left(1+c-\dfrac{\rho^2}{1+\rho^2+c}\right)$, откуда следует $\sigma_X^{-2} + \sigma_\theta^{-2} < z_{1,1}^{-1} < \sigma_X^{-2} + \sigma_e^{-2}$.

Если сравнить последнее неравенство с (2.4), можно убедиться, что асимптотическое значение является более точным.

Наконец, изучим асимптотический апостериорный коэффициент корреляции между θ_i и θ_j: $\rho'_{\theta_i,\theta_j} = z_{i,j}\big/\sqrt{z_{i,i}z_{j,j}}$.

Следующая лемма легко следует из формул (3.7) и (3.9).

Лемма 3.2. *Пусть* $j > i$. *Тогда*

$$\rho'_{\theta_i,\theta_j} = P_{j-1}^{-1/2}\rho_1^{j-1},\ i = 1;\ \rho'_{\theta_i,\theta_j} = P_{i-1}^{1/2}P_{j-1}^{-1/2}\rho_1^{j-i},\ i > 1, \tag{3.10}$$

где $\rho_1 = \rho\big/\sqrt{x}$.

Следствие.

$$\rho'_{\theta_i,\theta_j} < \rho_1^{|j-i|} < \rho^{|j-i|}. \tag{3.11}$$

Примечание. Так как миноры $P_1, P_2, ..., P_{i-1}, P_i, ...$ образуют строго возрастающую последовательность ($1 < P_1 < P_2 < ... < P_{i-1} < P_i < ...$), то из формул (3.10) и (3.11) следует, что асимптотический апостериорный коэффициент корреляции меньше и сходится к нулю быстрее (при $|i-j| \to \infty$), чем соответствующий априорный коэффициент.

Приложение 1
Доказательство формул (2.1)–(2.3)

Из формулы (1.1) следует, что

$$p_{\theta_1,...,\theta_k|X_1,...,X_k} \propto \exp\left\{-\frac{1}{2}\left[\sum(\theta_i - X_i)'\Sigma_X^{-1}(\theta_i - X_i) + (\theta^k - \theta^{0,k})'\Sigma_{\theta,k}^{-1}(\theta^k - \theta^{0,k})\right]\right\}$$

$$\propto \exp\left\{-\frac{1}{2}\left[(\theta^k - X^k)'\Sigma_{X,k}^{-1}(\theta^k - X^k) + (\theta^k - \theta^{0,k})'\Sigma_{\theta,k}^{-1}(\theta^k - \theta^{0,k})\right]\right\}$$

$$\propto \exp\left\{-\frac{1}{2}\left[(\theta^k)'\left(\Sigma_{X,k}^{-1} + \Sigma_{\theta,k}^{-1}\right)\theta^k - 2(\theta^k)'\left(\Sigma_{X,k}^{-1}X^k + \Sigma_{\theta,k}^{-1}\theta^{0,k}\right)\right]\right\}$$

$$\propto \exp\left\{-\frac{1}{2}\left[(\theta^k)'\Sigma_k^{-1}\theta^k - 2(\theta^k)'\Sigma_k^{-1}\Sigma_k\left(\Sigma_{X,k}^{-1}X^k + \Sigma_{\theta,k}^{-1}\theta^{0,k}\right)\right]\right\}$$

$$\propto \exp\left\{-\frac{1}{2}\left[(\theta^k - Y^k)'\Sigma_k^{-1}(\theta^k - Y^k)\right]\right\},$$

где $\Sigma_k^{-1} = \Sigma_{X,k}^{-1} + \Sigma_{\theta,k}^{-1}$, $Y^k = \Sigma_k\left(\Sigma_{X,k}^{-1}X^k + \Sigma_{\theta,k}^{-1}\theta^{0,k}\right)$.

Приложение 2
Доказательство результатов раздела 2

Доказательство леммы 2.1

Матрица Z_{θ,k_1} может быть записана в виде

$$Z_{\theta,k_1} = \begin{bmatrix} Z_{\theta,k} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix}. \tag{П2.1}$$

Запишем

$$Z_{\theta,k_1}^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \tag{П2.2}$$

и

$$Z_{k_1}^{-1} = \begin{bmatrix} A'_{11} & A_{12} \\ A_{21} & A'_{22} \end{bmatrix}. \tag{П2.3}$$

Из (2.3) следует, что

$$A'_{11} = A_{11} + \Sigma^{-1}_{X,k}\,,\ A'_{22} = A_{22} + \Sigma^{-1}_{X,k_1-k}\,, \tag{П2.4}$$

где $\Sigma_{X,k} = I_k \otimes \Sigma_X$.

Рассмотрим матрицу:

$$Z_{k_1} = \begin{bmatrix} Z^{(k)}_{k_1} & Z'_{12} \\ Z'_{21} & Z'_{22} \end{bmatrix}. \tag{П2.5}$$

Покажем, что

$$Z^{(k)}_{k_1} \le \Sigma_k\,. \tag{П2.6}$$

Достаточно доказать, что

$$\Sigma^{-1}_k \le \left(Z^{(k)}_{k_1}\right)^{-1}. \tag{П2.7}$$

Используем следующее представление для $\left(Z^{(k)}_{k_1}\right)^{-1}$:

$$\left(Z^{(k)}_{k_1}\right)^{-1} = -A_{12}\left(A'_{22}\right)^{-1} A_{21} + A'_{11}\,. \tag{П2.8}$$

Подставив (П2.4) в (П2.8), получаем

$$\left(Z^{(k)}_{k_1}\right)^{-1} = -A_{12}\left(A_{22} + \Sigma^{-1}_{X,k_1-k}\right)^{-1} A_{21} + A_{11} + \Sigma^{-1}_{X,k}\,. \tag{П2.9}$$

Так как

$$\Sigma^{-1}_{\theta,k} = -A_{12}A^{-1}_{22}A_{21} + A_{11}\,, \tag{П2.10}$$

то из (2.3) получаем

$$\Sigma^{-1}_k = \Sigma^{-1}_{X,k} + \Sigma^{-1}_{\theta,k} = \Sigma^{-1}_{X,k} - A_{12}A^{-1}_{22}A_{21} + A_{11}\,. \tag{П2.11}$$

Рассмотрим разность $\left(Z^{(k)}_{k_1}\right)^{-1} - \Sigma^{-1}_k$. Из (П2.9) и (П2.11) следует, что

$$\left(Z^{(k)}_{k_1}\right)^{-1} - \Sigma^{-1}_k = -A_{12}\left(A_{22} + \Sigma^{-1}_{X,k_1-k}\right)^{-1} A_{21} + A_{12}A^{-1}_{22}A_{21} = A_{12}[A^{-1}_{22} - \left(A_{22} + \Sigma^{-1}_{X,k_1-k}\right)^{-1}]A_{21}\,. \tag{П2.12}$$

С другой стороны,

$$A_{22} + \Sigma^{-1}_{X,k_1-k} > A_{22} \quad \Rightarrow \quad A^{-1}_{22} > \left(A_{22} + \Sigma^{-1}_{X,k_1-k}\right)^{-1}. \tag{П2.13}$$

Так как $(A_{12})' = A_{21}$, то из (П2.12) и (П2.13) следует, что $\left(Z^{(k)}_{k_1}\right)^{-1} - \Sigma^{-1}_k$ является неотрицательно определенной матрицей.

Доказательство леммы 2.2

Сначала покажем, что абсолютные значения элементов всех матриц ограничены константой C. Так как диагональные элементы матрицы A_k не превосходят соответствующих элементов матрицы A_1, то они ограничены наибольшим диагональным элементом A_1. С другой стороны, для любого элемента a_{ij} матрицы A_k выполняется неравенство $a^2_{ij} \le a_{ii}a_{jj}$.

Отсюда следует, что последовательность $A_1, A_2, ..., A_k, ...$ содержит сходящуюся подпоследовательность. Ее пределом будет неотрицательно определенная матрица. Достаточно показать, что, если $A_{m_1}, A_{m_2}, ..., A_{m_k}, ...$ и $A_{n_1}, A_{n_2}, ..., A_{n_k}, ...$ — две сходящиеся подпоследовательности из $A_1, A_2, ..., A_k, ...,$ то их пределы совпадают. Обозначим соответствующие пределы как A' и A'' и рассмотрим подпоследовательность $A_{l_1}, A_{l_2}, ..., A_{l_k}, ...$ последовательности $A_{m_1}, A_{m_2}, ..., A_{m_k}, ...,$ так чтобы $l_k > n_k$ для любого k. Тогда подпоследовательность $A_{l_1}, A_{l_2}, ..., A_{l_k}, ...$ также сходится к A'. Поскольку $A_{l_k} \geq A_{n_k}$ для любого k, то матрица $A' - A''$ неотрицательно определенная. Из тех же соображений можно заключить, что матрица $A'' - A'$ также неотрицательно определенная. Отсюда следует, что $A' = A''$.

Приложение 3
Доказательство результатов раздела 3[5]

а) *Доказательство формулы* (3.6).

Пусть D_k обозначает определитель матрицы $\sigma_e^2 \Sigma_k^{-1}$ в (3.4), и P_m $(m < k)$ — ведущий главный минор $\sigma_e^2 \Sigma_k^{-1}$. Величины D_k и P_m — положительные числа, т. к. Σ_k^{-1} — положительно определенная матрица. Поскольку $\sigma_e^2 \Sigma_k^{-1}$ является матрицей с центральной симметрией, то P_i будет главным минором, соответствующим подматрице, полученной из $\sigma_e^2 \Sigma_k^{-1}$ вычеркиванием ее первых $(k - m)$ столбцов и строк. Для вычисления определителей можно применить формулу Лапласа, сначала для первой строки $\sigma_e^2 \Sigma_k^{-1}$, а затем для первого столбца подматрицы $\sigma_e^2 \Sigma_k^{-1}$, соответствующей минору $M_{1,2}$. Таким образом, получим

$$D_k = (1 + c)P_{k-1} - \rho^2 P_{k-2}\,. \tag{П3.1}$$

Аналогично, получается следующая рекуррентная формула:

$$P_m = (1 + \rho^2 + c)P_{m-1} - \rho^2 P_{m-2}\,,\ \ 2 < m < k\,. \tag{П3.2}$$

Следующие две формулы потребуются в дальнейшем

$$P_1 = |1 + c| = 1 + c,\ \ P_2 = \begin{vmatrix} 1 + c & -\rho \\ -\rho & 1 + \rho^2 + c \end{vmatrix} = (1 + c)^2 + \rho^2 c\,. \tag{П3.3}$$

Лемма П3.1. *При* $m > 1$

$$1 + c < \frac{P_m}{P_{m-1}} < 1 + \rho^2 + c. \tag{П3.4}$$

Доказательство проведем методом индукции. Из (П3.3) следует, что (П3.4) верно для m=2 (напомним, что $c = \sigma_e^2 \sigma_X^{-2} > 0$). По индукции из (П3.4) следует, что

$$\frac{P_{m-2}}{P_{m-1}} < 1\,. \tag{П3.5}$$

[5] Для краткости доказательства результатов третьего раздела приведены в сокращенном виде.

Из (П3.2) получаем

$$\frac{P_m}{P_{m-1}} = 1 + \rho^2 + c - \rho^2 \frac{P_{m-2}}{P_{m-1}}. \qquad \text{(П3.6)}$$

Таким образом, (П3.4) следует из (П3.5) и (П3.6).

Следствие. *Ведущие главные миноры матрицы* $\sigma_e^2 \Sigma_k^{-1}$ *образуют возрастающую последовательность, т. е.* $1 < P_1 < P_2 < \ldots < P_{i-1} < P_i < \ldots$.

Доказательство следует из (П3.3) и (П3.4).

Лемма П3.2.

$$1 + c - \frac{\rho^2}{1+c} < \frac{D_k}{P_{k-1}} < 1 + c - \frac{\rho^2}{1+\rho^2+c}. \qquad \text{(П3.7)}$$

Доказательство. Из (П3.1) следует, что

$$\frac{D_k}{P_{k-1}} = 1 + c - \rho^2 \frac{P_{k-2}}{P_{k-1}}. \qquad \text{(П3.8)}$$

Таким образом, (П3.7) следует из (П3.4) и (П3.8).

Лемма П3.3.

$$\sigma_e^{-2}\left(1 + c - \frac{\rho^2}{1+c}\right) \le z_{1,1}^{-1} \le \sigma_e^{-2}\left(1 + c - \frac{\rho^2}{1+\rho^2+c}\right). \qquad \text{(П3.9)}$$

Доказательство. Из результатов, полученных во втором разделе, следует, что существует

$$z_{1,1} = \lim z_{1,1}^{(k)} = \lim \sigma_e^2 \frac{P_{k-1}}{D_k}. \qquad \text{(П3.10)}$$

Лемма теперь следует из (П3.7).

Следствие 1. $z_{1,1}$ — положительное число.

Следствие 2.

$$\sigma_X^{-2} + \sigma_\theta^{-2} < z_{1,1}^{-1} < \sigma_X^{-2} + \sigma_e^{-2}. \qquad \text{(П3.11)}$$

Доказательство. Правое неравенство в (П3.11) следует из (П3.9). Чтобы доказать левое неравенство, заметим, что

$$\sigma_\theta^{-2} = \sigma_e^{-2}(1-\rho^2). \qquad \text{(П3.12)}$$

С помощью элементарных преобразований получаем

$$\sigma_X^{-2} + \sigma_e^{-2}(1-\rho^2) < \sigma_e^{-2}\left(1 + c - \frac{\rho^2}{1+c}\right).$$

Лемма П3.4. $\lim\limits_{k\to\infty}\dfrac{P_{k-1}}{P_{k-2}}=x$, *где* x *— больший корень квадратного уравнения*

$$x^2-(1+\rho^2+c)x+\rho^2=0\text{, т. е. } x=\frac{1+\rho^2+c+\sqrt{(1+\rho^2+c)^2-4\rho^2}}{2}. \qquad (\text{П}3.13)$$

Доказательство. Из (П3.8), (П3.10) и следствия 1 к лемме П3.3 следует, что существует $y=\lim\dfrac{P_{k-2}}{P_{k-1}}$, $k\to\infty$. Из леммы П3.1 заключаем, что $y>0$. Пусть $x=y^{-1}$. Перейдя к пределам в обеих частях (П3.6), получим квадратное уравнение:

$$x^2-(1+\rho^2+c)x+\rho^2=0. \qquad (\text{П}3.14)$$

Поскольку дискриминант $\Delta=(1+\rho^2+c)^2-4\rho^2>0$, уравнение имеет два действительных корня. Из (П3.4) следует, что $x\ge 1+c$. Легко проверить, что меньший корень уравнения (П3.14) будет меньше, чем $1+c$.

Следствие.

$$1+c<x<1+\rho^2+c. \qquad (\text{П}3.15)$$

Заметим, что нельзя получить строгие неравенства в (П3.15), просто взяв пределы в обеих частях формулы (П3.4). Однако, переходя к пределам в обеих частях (П3.8) и применив (П3.10), получим

$$z_{1,1}^{-1}=\sigma_e^{-2}\left(1+c-\frac{\rho^2}{x}\right). \qquad (\text{П}3.16)$$

б) *Доказательство формулы* (3.7)

Прежде всего, заметим, что из (3.4) следует

$$z_{i,i}^{(k)}=\sigma_e^2\frac{M_{i,i}}{D_k}=\sigma_e^2\frac{P_{i-1}P_{k-i}}{D_k},\ i>1. \qquad (\text{П}3.17)$$

Таким образом, из (П3.10) и (П3.17) получаем

$$\frac{z_{1,1}}{z_{2,2}}=\lim_{k\to\infty}\frac{P_{k-1}}{D_k}\cdot\frac{D_k}{P_1P_{k-2}}=\frac{x}{P_1} \qquad (\text{П}3.18)$$

и

$$\frac{z_{i-1,i-1}}{z_{i,i}}=\lim_{k\to\infty}\frac{P_{i-2}P_{k-(i-1)}}{D_k}\cdot\frac{D_k}{P_{i-1}P_{k-i}}=\frac{xP_{i-2}}{P_{i-1}},\ i>2. \qquad (\text{П}3.19)$$

Запишем

$$\frac{z_{1,1}}{z_{i,i}}=\frac{z_{1,1}}{z_{2,2}}\cdot\frac{z_{2,2}}{z_{3,3}}\cdot\ldots\cdot\frac{z_{i-1,i-1}}{z_{i,i}}=\frac{x}{P_1}\cdot\frac{xP_1}{P_2}\cdot\ldots\cdot\frac{xP_{i-2}}{P_{i-1}}=\frac{x^{i-1}}{P_{i-1}}.$$

Окончательно имеем

$$z_{i,i}^{-1}=z_{1,1}^{-1}\frac{x^{i-1}}{P_{i-1}}. \qquad (\text{П}3.20)$$

Следствие. *Для любого* $i\ge 1$, $z_{i,i}$ *— положительное число.*

в) *Доказательство формулы* (3.8)

Лемма П3.5. *При* $m>1$

$$1<\frac{P_m}{P_{m-1}}<x. \tag{П3.21}$$

Доказательство. Левое неравенство следует из леммы П3.1. Доказательство правой части будет осуществлено методом индукции. Для $m=2$ лемма следует из формул (П3.3) и (П3.13). Предположим, что $\frac{P_{m-1}}{P_{m-2}}<x$. Тогда из (П3.6) получаем

$$\frac{P_m}{P_{m-1}}=1+\rho^2+c-\rho^2\frac{P_{m-2}}{P_{m-1}}<1+\rho^2+c-\frac{\rho^2}{x}=x. \tag{П3.22}$$

Докажем теперь, что

$$z_{i,i}^{-1}<z_{i+1,i+1}^{-1}. \tag{П3.23}$$

Если $i=1$, то из (П3.17) заключаем, что

$$\frac{z_{2,2}^{-1}}{z_{1,1}^{-1}}=\frac{x}{P_1}=\frac{x}{1+c}>1.$$

Заметим, что последнее неравенство следует из (П3.15).

Если $i>1$, то из формул (П3.20) и (П3.21) получаем:

$$\frac{z_{i+1,i+1}^{-1}}{z_{i,i}^{-1}}=\frac{x^i}{P_i}\cdot\frac{P_{i-1}}{x^{i-1}}=x\frac{P_{i-1}}{P_i}>1.$$

г) *Доказательство формулы* (3.9)

Поскольку Σ_k является симметрической матрицей, достаточно рассмотреть случай, когда $j>i$.

1) $i=1$. Запишем $(1,j)$-ый минор $\sigma_e^2\Sigma_k^{-1}$ в следующем виде

$$M_{1,j}=\begin{vmatrix} -\rho & 1+\rho^2+c & & & & & & \\ 0 & -\rho & 1+\rho^2+c & & & & & \\ \dots & \dots & \dots & \dots & \dots & & & \\ 0 & 0 & 0 & -\rho & -\rho & & & \\ 0 & 0 & 0 & 0 & 1+\rho^2+c & -\rho & & \\ 0 & 0 & 0 & 0 & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & -\rho & 1+\rho^2+c & -\rho \\ 0 & 0 & 0 & 0 & 0 & 0 & -\rho & 1+c \end{vmatrix}. \tag{П3.24}$$

Отсюда следует, что

$$z^{(k)}{}_{1,j}=(-1)^{1+j}\sigma_e^2\frac{(-\rho)^{j-1}P_{k-j}}{D_k}=\sigma_e^2\frac{\rho^{j-1}P_{k-j}}{D_k}. \tag{П3.25}$$

Из (П3.17) заключаем, что

$$z^{(k)}{}_{1,j}=\frac{z^{(k)}_{j,j}\rho^{j-1}}{P_{j-1}}. \tag{П3.26}$$

Переходя к пределам (при $k\to\infty$) в обеих частях (П3.26) и затем, воспользовавшись формулой (П3.20), получим

$$z_{1,j}=\frac{z_{j,j}\rho^{j-1}}{P_{j-1}}=\frac{z_{1,1}\rho^{j-1}}{x^{j-1}}. \tag{П3.27}$$

2) $i>1$. (i,j)-ый минор матрицы $\sigma_e^2\Sigma_k^{-1}$ записывается как

$$M_{i,j}=\begin{vmatrix}
1+c & -\rho & & & & & & & & & & & \\
-\rho & 1+\rho^2+c & -\rho & & & & & & & & & & \\
\dots & \dots & \dots & \dots & & & & & & & & & \\
0 & 0 & -\rho & 1+\rho^2+c & -\rho & & & & & & & & \\
0 & 0 & 0 & 0 & -\rho & 1+\rho^2+c & & & & & & & \\
0 & 0 & 0 & 0 & 0 & -\rho & 1+\rho^2+c & & & & & & \\
\dots & \dots & \dots & \dots & .. & \dots & \dots & \dots & & & & & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -\rho & -\rho & & & & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1+\rho^2+c & -\rho & & & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\rho & 1+\rho^2+c & -\rho & & \\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\rho & 1+\rho^2+c & -\rho \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\rho & 1+c
\end{vmatrix}. \tag{П3.28}$$

Отсюда заключаем, что

$$z^{(k)}{}_{i,j}=(-1)^{i+j}\sigma_e^2\frac{P_{i-1}(-\rho)^{j-i}P_{k-j}}{D_k}=\sigma_e^2\frac{P_{i-1}\rho^{j-i}P_{k-j}}{D_k}. \tag{П3.29}$$

Повторяя в точности рассуждения, сделанные в 1), получаем

$$z_{i,j}=\frac{z_{1,1}P_{i-1}\rho^{j-i}}{x^{j-1}}. \tag{П3.30}$$

Список литературы

Айвазян С. А. (2008). Байесовский подход в эконометрическом анализе. *Прикладная эконометрика,* 1 (9), 93–130.

Айвазян С. А., Иванова С. С. (2007). *Эконометрика.* М.: Маркет ДС.

Боровков А. А. (2009). *Теория вероятностей*. М.: Книжный дом «ЛИБРОКОМ».

Де Гроот М. (1974). *Оптимальные статистические решения*. М.: Мир.

Магнус Я. Р., Нейдеккер Х. (2002). *Матричное дифференциальное исчисление с приложениями к статистике и эконометрике*. М.: Физматлит.

Berger J. O. (1985). *Statistical decision theory and Bayesian analysis*, 2nd ed., New York: Springer-Verlag.