III. Методы решения задач

РЕГУЛЯРИЗАЦИЯ ОЦЕНКИ ПЛОХО ОБУСЛОВЛЕННОЙ КОВАРИАЦИОННОЙ МАТРИЦЫ [1]

Голубков В. В.

Институт системного анализа РАН, Москва

Аннотация

В данной работе дается решение задачи нахождения корректной оценки ковариационной матрицы с помощью указанной процедуры регуляризации.

В настоящее время для моделирования и прогнозирования дискретных случайных временных процессов и в частности демографических процессов широко используется вероятностный подход [18], в основе которого лежит методология имитационного моделирования. Одной из задач имитационного моделирования является задача генерирования случайных векторов с заданной ковариационной матрицей и с заданными маргинальными законами распределения. Как правило, ковариационная матрица не известна и в качестве таковой используется ее оценка, получаемая по известным статистическим данным. В случае, когда ковариационная матрица близка к вырожденной, может оказаться, что ее оценка из-за вычислительных ошибок не является положительно полуопределенной матрицей, чего быть не должно. В этом случае можно использовать процедуру регуляризации оценки, согласно которой в качестве корректной оценки ковариационной матрицы берется симметричная неотрицательно определенная матрица, наиболее близкая в некоторой метрике к оценке ковариационной матрицы, полученной в результате обработки статистической информации. В данной работе дается решение задачи нахождения корректной оценки ковариационной матрицы с помощью указанной процедуры регуляризации.

[1] Данная работа выполнена при поддержке РФФИ, грант № 09-001-00273а

Абстрагируясь от физического смысла указанной задачи, решение последней сводится к решению следующей математической задачи. Пусть $K=\{K_{ij}\}(i, j = 1, n)$ заданная вещественная квадратная симметричная n-го порядка. Требуется найти квадратную вещественную симметричную неотрицательно определенную матрицу $P=\{Pji\}(i, j = 1, n)$, наиболее близкую к матрице K в заданной метрике, т.е. в результате решения оптимизационной задачи

$$P = \operatorname{Arg}\min_{P\geq 0} \|K\text{-}P\|^2,\ K = K^T,\ P = P^T \tag{1}$$

где запись $P \geq 0$ означает, что матрица P – неотрицательно определенная, T операция транспонирования, а $\|...\|$- заданная норма. В качестве нормы возьмем матричную норму, которую введем через скалярное произведение матриц (A, B), определяемое для матриц A, B с одинаковыми размерами по формуле

$$(\mathrm{A},\mathrm{B}) = \sum_{i=1}^{n}\sum_{j=1}^{m} A_{ij}B_{ij};\ \mathrm{A} = \{A_{ij}\},\ \mathrm{B} = \{B_{ij}\},\ i = \overline{1,m},\ j = \overline{1,n} \tag{2}$$

Так определенное скалярное произведение матриц обладает всеми свойствами скалярного произведения и, кроме того, имеют место следующие соотношения

$$(A,\ B) = (A^T,\ B^T),\ (ABC,\ D) = (B,\ A^T DC^T),\ (D,\ ABC) = (A^T DC^T,\ B) \tag{3}$$

где размеры матриц такие, что все операции определены. Используя скалярное произведение матриц (2), определим квадрат матричной нормы матрицы по A формуле

$$\|A\|^2 = (A, A)$$

Тогда задача (1) через эту норму запишется в виде

$$\begin{aligned} &\mathrm{P} = Arg \min_{\mathrm{P}\geq 0, \mathrm{P}=\mathrm{P}^T} \Phi(\mathrm{P}) \\ &\Phi(\mathrm{P}) = (\mathrm{K}-\mathrm{P}, \mathrm{K}-\mathrm{P}), \mathrm{K}= \mathrm{K}^T, \mathrm{P} = \mathrm{P}^T \end{aligned} \tag{4}$$

Поскольку любая симметричная неотрицательно определенная матрица, как известно, может быть представлена в виде произведения некоторой матрицы на ее транспонированную матрицу (это следует из спектрального представления матрицы данного типа [9, стр. 280]), а искомая матрица P является по условию симметричной неотрицательно определенной, то ее можно представить в виде

$$P = RR^T, \qquad (5)$$

где R некоторая матрица. Подстановкой выражения для P из (5) в (4), решение задачи (4) сводится к решению следующей задачи

$$R = Arg\min_{R} \Phi(R),\ K = K^T,\ P = RR^T$$
$$\Phi(R) \triangleq \Phi(P) = (K - P, K - P) = (K - RR^T, K - RR^T) \qquad (6)$$

Из (6), используя симметричность матрицы K и свойства (3) введенного скалярного произведения матриц, находим первую вариацию $\delta\Phi(R)$ целевой функции $\Phi(R)$

$$\delta\Phi(R) = -2\,(\delta RR^T + R\delta R^T, K - RR^T) = -\,4\,(\delta R, KR - RR^TR)$$

Отсюда, приравнивая первую вариацию $\delta\Phi(R)$ нулю (необходимое условие минимума функции $\Phi(R)$), получаем матричное уравнение, которому должна удовлетворять оптимальное значение матрицы R

$$KR - RR^TR = 0 \qquad (7)$$

Умножая уравнение (7) справа на матрицу R^T и учитывая (5), находим

$$KP - P^2 = 0$$

В силу симметричности матриц K и P отсюда следует

$$KP = PK$$

что означает, что матрицы K и P являются перестановочными (коммутирующими) матрицами. Но тогда для симметричных перестановочных матриц K и P согласно [9, стр. 207, 265] существует такая ортонормированная матрица H, с помощью которой одновременно матрицы K и P приводятся к диагональному виду, т.е.

$$K = H\Lambda^{(k)}H^T;\ P = H\Lambda^{(p)}H^T;\ HH^T = H^TH = E\ ,$$

где Е единичная, а $\Lambda^{(k)}$ и $\Lambda^{(p)}$ – диагональные матрицы с диагональными элементами, равными собственным числам соответствующих матриц. Другими словами, оптимальная матрица Р с необходимостью должна иметь следующую структуру:

$$P = H\Lambda^{(p)} H^T;\ H \in G_H = \{H : K = H\Lambda^{(к)} H^T,\ HH^T = H^T H = E\},$$
$$\Lambda^{(k)} = \mathrm{diag}\,(\lambda_1^{(k)}, \lambda_2^{(k)}, \dots, \lambda_n^{(k)}), \tag{8}$$
$$\Lambda^{(p)} = \mathrm{diag}\,(\lambda_1^{(p)}, \lambda_2^{(p)}, \dots, \lambda_n^{(p)});\ \lambda_i^{(p)} \geq 0,\ i = 1, n.$$

Не нарушая общности, будем считать, что собственные числа $\lambda_i^{(k)}\left(i = \overline{1,n}\right)$ матрицы K пронумерованы таким образом, что имеет место представление матрицы $\Lambda^{(k)}$ в следующем блочном виде:

$$\boldsymbol{\Lambda}^{(k)} = \begin{pmatrix} \boldsymbol{\Lambda}_1^{(k)} & 0 \\ 0 & \boldsymbol{\Lambda}_2^{(k)} \end{pmatrix},\quad \begin{matrix} \boldsymbol{\Lambda}_1^{(k)} = diag\left(\lambda_1^{(k)}, \lambda_2^{(k)}, \dots, \lambda_m^{(k)}\right); \qquad \lambda_i^{(k)} \geq 0,\ i = \overline{1,m} \\ \boldsymbol{\Lambda}_2^{(k)} = diag\left(\lambda_{m+1}^{(k)}, \lambda_{m+2}^{(k)}, \dots, \lambda_n^{(k)}\right);\ \lambda_i^{(k)} < 0,\ i = \overline{m+1,n} \end{matrix}. \tag{9}$$

Используя (8) и свойства скалярного произведения матриц (3), из (6) получаем

$$\Phi(P)=(K\text{-}P,K\text{-}P)=\left(H\Lambda^{(k)}H^T\text{-}H\Lambda^{(p)}H^T, H\Lambda^{(k)}H^T\text{-}H\Lambda^{(p)}H^T\right)=$$
$$=\left(\Lambda^{(k)} - \Lambda^{(p)}, \Lambda^{(k)} - \Lambda^{(p)}\right)=\sum_{i=1}^{n}\left(\lambda_i^{(p)}\text{-}\lambda_i^{(k)}\right)^2=\sum_{i=1}^{m}\left(\lambda_i^{(p)}\text{-}\lambda_i^{(k)}\right)^2+\sum_{i=m+1}^{n}\left(\lambda_i^{(p)}\text{-}\lambda_i^{(k)}\right)^2$$

или

$$\boldsymbol{\Phi}(P) = (K - P, K - P) = \sum^{m}(\lambda_i^{(p)} - \lambda_i^{(k)})^2 + \sum^{n}(\lambda_i^{(p)} - \lambda_i^{(k)})^2.$$

Отсюда с учетом (8) и (9) следует, что минимальное значение целевой функции Ф (Р) достигается на такой матрице Р, у которой собственные числа $\lambda_i^{(P)}$ (i = 1,n) имеют следующие значения:

$$\lambda_i^{(p)} = \lambda_i^{(k)},\ i = \overline{1,m};\ \lambda_i^{(p)} = 0,\ i = \overline{m+1,n}. \tag{10}$$

причем минимальное значение целевой функции в этом случае равно

$$\Phi(\mathrm{P}) = \sum_{i=m+1}^{n}\left(\lambda_i^{(k)}\right)^2$$

Из (8), (9) и (10) получаем следующее выражение для $\Lambda^{(p)}$:

$$\Lambda^{(p)}=\begin{pmatrix}\Lambda_1^{(k)} & 0\\ 0 & 0\end{pmatrix},\ \Lambda_1^{(k)} = \mathrm{diag}\left(\lambda_1^{(k)},\lambda_2^{(k)},...,\lambda_m^{(k)}\right) \quad .$$

Таким образом, установлено, что минимум целевой функции Ф(Р) достигается на матрице Р, вычисляемой по формуле:

$$\mathrm{P} = \mathrm{H}\Lambda^{(p)}\mathrm{H}^{\mathrm{T}},\ \ \mathrm{H} \in \mathrm{G_H} = \left\{\mathrm{H}:\ \mathrm{K} = \mathrm{H}\Lambda^{(k)}\mathrm{H}^{\mathrm{T}},\ \mathrm{HH}^{\mathrm{T}} = \mathrm{H}^{\mathrm{T}}\mathrm{H} = \mathrm{E}\right\}$$

$$\Lambda^{(k)} = \begin{pmatrix}\Lambda_1^{(k)} & 0\\ 0 & \Lambda_2^{(k)}\end{pmatrix},\ \ \Lambda^{(p)} = \begin{pmatrix}\Lambda_1^{(k)} & 0\\ 0 & 0\end{pmatrix} \tag{11}$$

$$\Lambda_1^{(k)} = \mathrm{diag}\left(\lambda_1^{(k)},\lambda_2^{(k)},...,\lambda_m^{(k)}\right);\quad \lambda_i^{(k)} \geq 0,\ i = \overline{1,m}$$

$$\Lambda_2^{(k)} = \mathrm{diag}\left(\lambda_{m+1}^{(k)},\lambda_{m+2}^{(k)},...,\lambda_n^{(k)}\right);\ \lambda_i^{(k)} < 0,\ i = \overline{m+1,n}$$

Поскольку, вообще говоря, матрица $\mathrm{H}\in\mathrm{G_H}$ не является единственной, то возникает вопрос о зависимости оптимальной матрицы Р от Н, т.е. вопрос о единственности матрицы Р. Покажем, что оптимальная матрица Р, вычисляемая по формулам (11), является единственной, т.е. для всех матриц $\mathrm{H}\in\mathrm{G_H}$ значение оптимальной матрицы Р будет одно и тоже. Действительно, возьмем две произвольные матрицы $\mathrm{H_1}\in\mathrm{G_H}$ и $\mathrm{H_2}\in\mathrm{G_H}$ и вычислим по формулам (11) соответствующие матрицы $\mathrm{P_1}$ и $\mathrm{P_2}$

$$\mathrm{P_1} = \mathrm{H_1}\Lambda^{(p)}\mathrm{H_1}^{\mathrm{T}},\ \mathrm{P_2}=\mathrm{H_2}\Lambda^{(p)}\mathrm{H_2}^{\mathrm{T}};\ \mathrm{H_1},\mathrm{H_2} \in \mathrm{G_H}. \tag{12}$$

Так как $\mathrm{H_1}, \mathrm{H_2} \in \mathrm{G_H}$, то согласно (11) имеем

$$\mathrm{H_1}\Lambda^{(k)}\mathrm{H_1}^{\mathrm{T}} = \mathrm{K},\ \mathrm{H_2}\Lambda^{(k)}\mathrm{H_2}^{\mathrm{T}}= \mathrm{K} \Rightarrow \mathrm{H_1}\Lambda^{(k)}\mathrm{H_1}^{\mathrm{T}} = \mathrm{H_2}\Lambda^{(K)}\mathrm{H_2}^{\mathrm{T}},$$

откуда следует, что $\Lambda^{(k)}\mathrm{H_1}^{\mathrm{T}}\mathrm{H_2} = \mathrm{H_1}^{\mathrm{T}}\mathrm{H_2}\Lambda^{(k)}$

или

$$\Lambda^{(k)}\,\widehat{\mathrm{H}}\text{-}\widehat{\mathrm{H}}\Lambda^{(k)}=0,\ \ \widehat{\mathrm{H}}=\mathrm{H_1}^{\mathrm{T}}\mathrm{H_2} \tag{13}$$

Запишем матрицу $\widehat{\mathrm{H}}$ в следующем блочном виде

$$\widehat{\mathrm{H}}=\begin{pmatrix}\widehat{\mathrm{H}}_{11} & \widehat{\mathrm{H}}_{12}\\ \widehat{\mathrm{H}}_{21} & \widehat{\mathrm{H}}_{22}\end{pmatrix}.$$

где блоки $\widehat{\mathrm{H}}_{\mathrm{ij}}\,(\mathrm{i}, \mathrm{j} = 1, 2)$ имеют размеры одноименных блоков матрицы $\Lambda^{(\mathrm{k})}$ в (11). Отсюда и из (11) находим

$$\Lambda^{(\mathrm{k})}\widehat{\mathrm{H}}\text{-}\widehat{\mathrm{H}}\Lambda^{(\mathrm{k})}=\begin{pmatrix}\Lambda_1^{(\mathrm{k})} & 0\\ 0 & \Lambda_2^{(\mathrm{k})}\end{pmatrix}\begin{pmatrix}\widehat{\mathrm{H}}_{11} & \widehat{\mathrm{H}}_{12}\\ \widehat{\mathrm{H}}_{21} & \widehat{\mathrm{H}}_{22}\end{pmatrix}-\begin{pmatrix}\widehat{\mathrm{H}}_{11} & \widehat{\mathrm{H}}_{12}\\ \widehat{\mathrm{H}}_{21} & \widehat{\mathrm{H}}_{22}\end{pmatrix}\begin{pmatrix}\Lambda_1^{(\mathrm{k})} & 0\\ 0 & \Lambda_2^{(\mathrm{k})}\end{pmatrix}=$$

$$=\begin{pmatrix}\Lambda_1^{(\mathrm{k})}\widehat{\mathrm{H}}_{11}\text{-}\widehat{\mathrm{H}}_{11}\Lambda_1^{(\mathrm{k})} & \Lambda_1^{(\mathrm{k})}\widehat{\mathrm{H}}_{12}\text{-}\widehat{\mathrm{H}}_{12}\Lambda_2^{(\mathrm{k})}\\ \Lambda_2^{(\mathrm{k})}\widehat{\mathrm{H}}_{21}\text{-}\widehat{\mathrm{H}}_{21}\Lambda_1^{(\mathrm{k})} & \Lambda_2^{(\mathrm{k})}\widehat{\mathrm{H}}_{22}\text{-}\widehat{\mathrm{H}}_{22}\Lambda_2^{(\mathrm{k})}\end{pmatrix}. \qquad (14)$$

что совместно с (13) в частности дает

$$\Lambda_1^{(\kappa)}\widehat{\mathrm{H}}_{11}-\widehat{\mathrm{H}}_{11}\Lambda_1^{(\kappa)}=0,\ \Lambda_1^{(\kappa)}\widehat{\mathrm{H}}_{12}-\widehat{\mathrm{H}}_{12}\Lambda_2^{(\kappa)}=0,\ \Lambda_2^{(\kappa)}\widehat{\mathrm{H}}_{21}-\widehat{\mathrm{H}}_{21}\Lambda_1^{(\kappa)}=0. \qquad (15)$$

Из двух последних соотношений и свойств элементов диагональных матриц $\Lambda_1^{(\mathrm{k})}$ и $\Lambda_2^{(\mathrm{k})}$ в (11) следует, что

$$\widehat{\mathrm{H}}_{12}=0,\ \widehat{\mathrm{H}}_{21}=0. \qquad (16)$$

Действительно, второе соотношение в (15) в скалярной форме можно записать в виде

$$(\Lambda_1^{(k)}\widehat{H}_{12})_{ij}-(\widehat{H}_{12}\Lambda_2^{(k)})_{ij}=\sum_{r=1}^{m}(\Lambda_1^{(k)})_{ir}(\widehat{H}_{12})_{rj}-\sum_{r=1}^{n-m}(\widehat{H}_{12})_{ir}(\Lambda_2^{(k)})_{rj}=0, \qquad (17)$$
$$i=\overline{1,m},\ j=\overline{1,n-m}.$$

Здесь и в дальнейшее запись $(\mathrm{A})_{\mathrm{ij}}$ означает элемент матрицы А, стоящий на пересечении *i*-й строки и *j*-го столбца. Поскольку согласно (11)

$$\begin{aligned}&(\Lambda_1^{(\mathrm{k})})_{ij}=\lambda_i^{(\mathrm{k})}\delta_{ij},\ \lambda_i^{(\mathrm{k})}\geq 0;\ \ i=\overline{1,m},\ j=\overline{1,m}\\ &(\Lambda_2^{(\mathrm{k})})_{ij}=\lambda_{m+i}^{(\mathrm{k})}\delta_{ij},\ \lambda_{m+i}^{(\mathrm{k})}<0;\ i=\overline{1,n-m},\ j=\overline{1,n-m}\end{aligned}\qquad \delta_{ij}=\begin{cases}1,\ i=j\\ 0,\ i\neq j.\end{cases} \qquad (18)$$

то отсюда и из (17) получаем

$$\sum_{r=1}^{m}\lambda_i^{(\mathrm{k})}\delta_{ir}(\widehat{\mathrm{H}}_{12})_{rj}-\sum_{r=1}^{n-m}(\widehat{\mathrm{H}}_{12})_{ir}\lambda_{m+r}^{(\mathrm{k})}\delta_{rj}=(\lambda_i^{(\mathrm{k})}-\lambda_{m+j}^{(\mathrm{k})})(\widehat{\mathrm{H}}_{12})_{ij}=0,$$
$$i=\overline{1,m},\ j=\overline{1,n-m}.$$

Учитывая, что в силу (18) значения

$$(\lambda_i^{(k)} - \lambda_{m+j}^{(k)}) > 0 \left(i = \overline{1,m};\, j = \overline{1,n-m} \right)$$

отсюда непосредственно следует

$$(\widehat{\mathbf{H}}_{12})_{ij} = 0;\ i = \overline{1,m},\ j = \overline{1,n-m}\ .$$

что доказывает справедливость первого соотношения в (16). Транспонируя третье соотношение в (15) и умножая полученное соотношение на -1, находим

$$\Lambda_1^{(k)} \widehat{\mathrm{H}}_{21}^{\mathrm{T}} - \widehat{\mathrm{H}}_{21}^{\mathrm{T}} \Lambda_2^{(k)} = 0\ .$$

Далее, дословно повторяя выкладки, использованные при доказательстве первого соотношения в (16), с заменой в них $\widehat{\mathrm{H}}_{12}$ на $\widehat{\mathrm{H}}_{21}^{\mathrm{T}}$, получаем, $\widehat{\mathrm{H}}_{21}^{\mathrm{T}}=0$ а, значит и $\widehat{\mathrm{H}}_{21}^{\mathrm{T}} = 0$.

Таким образом, справедливость соотношений (16) доказана.

Из (12) имеем

$$\mathrm{P}_1\text{-}\mathrm{P}_2 = \mathrm{H}_1 \Lambda^{(p)} \mathrm{H}_1{}^{\mathrm{T}} - \mathrm{H}_2 \Lambda^{(p)} \mathrm{H}_2{}^{\mathrm{T}} = \mathrm{H}_1 (\Lambda^{(p)} \mathrm{H}_1{}^{\mathrm{T}} \mathrm{H}_2 - \mathrm{H}_1{}^{\mathrm{T}} \mathrm{H}_2 \Lambda^{(p)}) \mathrm{H}_2{}^{\mathrm{T}}$$

или с учетом (13)

$$\mathrm{P}_1 - \mathrm{P}_2 = \mathrm{H}_1 (\boldsymbol{\Lambda}^{(p)} \widehat{\mathrm{H}} - \widehat{\mathrm{H}} \boldsymbol{\Lambda}^{(p)}) \mathrm{H}_2{}^{T} \qquad (19)$$

Используя структуру матрицы $\Lambda^{(p)}$ в (11) и повторяя выкладки при выводе формулы (14), в которых надо положить $\Lambda_2^{(k)} = 0$, получаем

$$\Lambda^{(p)} \widehat{\mathrm{H}} - \widehat{\mathrm{H}} \Lambda^{(p)} = \begin{pmatrix} \Lambda_1^{(k)} \widehat{\mathrm{H}}_{11} - \widehat{\mathrm{H}}_{11} \Lambda_1^{(k)} & \Lambda_1^{(k)} \widehat{\mathrm{H}}_{12} \\ -\widehat{\mathrm{H}}_{21} \Lambda_1^{(k)} & 0 \end{pmatrix}.$$

Отсюда, принимая во внимание первое соотношение в (15) и соотношения (16), следует, что

$$\Lambda^{(p)} \widehat{\mathrm{H}} - \widehat{\mathrm{H}} \Lambda^{(p)} = 0\,.$$

Последнее вместе с (19) дает

$$\mathrm{P}_1 - \mathrm{P}_2 = 0 \ \Rightarrow\ \mathrm{P}_1 = \mathrm{P}_2$$

что и доказывает единственность оптимальной матрицы P. Из этого следует вывод: для вычисления матрицы P по формулам (11) можно использовать любую матрицу $H \in G_H$.

В случае, когда матрица K является неотрицательно определенной ($K \geq 0$), то тогда m = n и, стало быть, согласно (11) в этом случае матрица P совпадает с матрицей K, т.е. P = K. Для вычисления матрицы P по формулам (11) необходимо знать спектральное представление матрицы K, а именно надо знать матрицы $\Lambda^{(k)}$ и H, которые формируются соответственно по собственные числам и собственным векторам матрицы K. Для нахождения матриц $\Lambda^{(k)}$ и H можно использовать численный итерационный метод вращений [10, стр. 598]. Этот метод является достаточно эффективным, как по вычислительным затратам, так и по скорости сходимости.

В заключение следует отметить, что в задачах по вероятностному прогнозированию имитационным методом дискретных случайных временных процессов используется не сама ковариационная матрица P, а матрица R, через которую по формуле (5) выражается матрица P. Коль скоро матрицы $\Lambda^{(k)}$ и H вычислены, то в качестве матрицы R можно взять арифметический квадратный корень $P^{1/2}$ из матрицы P

$$R = P^{1/2} = H(\Lambda^{(p)})^{1/2} H^T,\ (\Lambda^{(p)})^{1/2} = \mathrm{diag}\left(\sqrt{\lambda_1^{(p)}}, \sqrt{\lambda_2^{(p)}}, \ldots, \sqrt{\lambda_n^{(p)}}\right)$$

или матрицу вида

$$R = H(\Lambda^{(p)})^{1/2},\ (\Lambda^{(p)})^{1/2} = \mathrm{diag}\left(\sqrt{\lambda_1^{(p)}}, \sqrt{\lambda_2^{(p)}}, \ldots, \sqrt{\lambda_n^{(p)}}\right).$$

Второй вариант более предпочтителен, поскольку при его реализации требуется меньше вычислений, что важно для ковариационных матриц большой размерности.

Литература

1. *Pfaumer P.* Confidence intervals for population projections based on Monte Carlo methods. International Journal of Forecasting, volume 4, issue 1, 1988, 135-142.
2. *Alho J.V.* Stochastic methods in population forecasting, International Journal of Forecasting 6: 521-530, 1990.
3. *Keilman N., Pham D.Q.,* Predictive Intervals for Age-Specific Fertility, European Journal of Population 16: 41-66, 2000.

4. *Keilman N., Pham D.Q., Hetland A.,* Why population forecasts should be probabilistic – illustrated by the case of Norway, Demographic Research, v.6, 2002.
5. *Lutz W., Sanderson W, Scherbov S.* Probabilistic population projections based on expert opinion, pp 397-428 in Lutz W. (ed). The future population of the world: What can we assume today? London: Earthscan (rev.ed.), 1996.
6. *Lutz W., Sanderson W, Scherbov S.* The end of world population growth, Nature 412: 543545, 2001.
7. *Oeppen J., Vaupel J.W.* Broken limits to life expectancy. Science 296: 1029-1031, 2002.
8. *Голубков В.В., Кругляков С.В.,* Прогнозирование случайных процессов с использованием параметрических моделей автокорреляционных функций. Динамика неоднородных систем, Москва, Изд. URSS, 2006, № 10.
9. *Гантмахер Ф.Р. ,* Теория матриц, «Наука», М.: 1966.
10. *Березин И.С. , Н.П. Жидков,* Методы вычислений т. 1, «Наука», М.: 1966.