

Голубков В. В., Кругляков С. В.

Институт системного анализа РАН, Москва

ЧИСЛЕННЫЕ МЕТОДЫ ГЕНЕРИРОВАНИЯ СЛУЧАЙНЫХ ВЕКТОРОВ С ЗАДАННЫМИ МАРГИНАЛЬНЫМИ РАСПРЕДЕЛЕНИЯМИ И ЗАДАННОЙ КОРРЕЛЯЦИОННОЙ МАТРИЦЕЙ*

В настоящее время для моделирования и прогнозирования случайных временных дискретных процессов и в частности демографических процессов широко используется вероятностный подход, в основе которого лежит методология имитационного моделирования. Одной из задач имитационного моделирования является задача генерирования случайных векторов с заданной корреляционной матрицей и с заданными, вообще говоря, различными маргинальными законами распределения (компоненты случайных векторов могут иметь качественно различные распределения, например, одни компоненты могут быть распределены непрерывно, а другие — дискретно). Решению данной задачи посвящено много теоретических работ [1–10]. В этих работах предложены и обоснованы различные математические методы генерации случайных векторов с заданными свойствами, причем во всех этих работах решение данной задачи сводится к решению системы трансцендентных уравнений, число которых равно числу корреляционных коэффициентов. Целью настоящей работы, инициированной необходимостью вероятностного прогнозирования демографических процессов, является разработка и сравнительный численный анализ эффективных по вычислительным затратам численных методов решения указанной выше задачи. Разработка таких методов применительно к демографическим процессам является достаточно актуальной задачей, поскольку при прогнозировании демографических процессов временной интервал прогнозирования, как правило, составляет порядка 100 лет с шагом прогнозирования один год, что приводит к необходимости генерирования случайных 100-мерных векторов и в конечном счете к решению 495 трансцендентных уравнений, равному в этом случае числу коэффициентов корреляции. В настоящей работе разработка указанных численных методов было основано на теоретических результатах, полученных в наиболее общем случае в работе [11].

* Данная работа выполнена при поддержке РФФИ, грант № 06–01–00282а.

Уравнения для определения коэффициентов корреляции. Методы их решения

Пусть $\mathbf{X} = (x_1 \ x_2 \ \dots \ x_n)^T$ — случайный n -вектор с заданными коэффициентами корреляции r_{ij} ($i, j = \overline{1, n}$) и маргинальными функциями распределения $F_i(x)$ ($i = \overline{1, n}$). Последнее означает, что

$$P(x_i < x) = F_i(x), \quad r_{ij} = \frac{R_{ij} - m_i m_j}{\sigma_i \sigma_j}, \quad r_{ii} = 1,$$

$$R_{ij} = E(x_i x_j), \quad m_i = E(x_i), \quad \sigma_i = \sqrt{E((x_i - m_i)^2)}, \quad (1)$$

$$i, j = \overline{1, n}$$

где $E(\dots)$ и $P(\dots)$ соответственно оператор математического ожидания и вероятность события, указанного в скобках. Поскольку коэффициенты корреляции r_{ij} и маргинальные функции распределения $F_i(x)$ известны, то согласно (1) известны математические ожидания m_i , среднеквадратические отклонения v_i и ковариации R_{ij} . Поэтому сформулированная в данной работе задача сводится к задаче генерирования случайного вектора с заданными ковариациями R_{ij} и маргинальными функциями распределениями $F_i(x)$.

Предлагаемые в настоящей работе численные методы решения поставленной задачи основаны на результатах, полученных в работе [11]. Согласно [11] случайный вектор \mathbf{X} вычисляется следующим образом

$$\mathbf{X} = (x_1 \ x_2 \ \dots \ x_n)^T; \quad x_i = F_i^{-1}(\Phi(z_i)), \quad i = \overline{1, n},$$

$$\mathbf{Z} = (z_1 \ z_2 \ \dots \ z_n)^T \sim N(\mathbf{0}, \mathbf{K}); \quad \mathbf{0} = (\overbrace{0 \ 0 \ \dots \ 0}^n)^T, \quad \mathbf{K} = \{\rho_{ij}\}_{i,j=1}^n, \quad (2)$$

$$\rho_{ii} = 1, \quad i = \overline{1, n}; \quad \Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-t^2/2} dt.$$

Здесь \mathbf{Z} — случайный нормально распределенный n -вектор с нулевыми математическими ожиданиями, единичными дисперсиями и корреляционной матрицей \mathbf{K} (ρ_{ij} ($i \neq j$) — коэффициенты корреляции), $F_i^{-1}(z)$ — функция, обратная к функции $F_i(x)$, $\Phi(z)$ — стандартная нормальная функция распределения. Из (2) следует, что маргинальные распределения случайного вектора \mathbf{X} равны заданным функциям $F_i(x)$. Значения коэффициентов корреляции ρ_{ij} ($i \neq j$) находятся из условий равенства ковариаций $E(x_i x_j)$ вектора \mathbf{X} , определенного в (2), заданным значениям R_{ij} . В результате получается следующая система $n(n - 1)/2$ независимых нелинейных уравнений для определения $n(n - 1)/2$ неиз-

вестных коэффициентов корреляции ρ_{ij} ($i = \overline{1, n-1}$; $j = \overline{i+1, n}$)

$$\begin{aligned}
 V_{ij}(\rho_{ij}) &= 0, \\
 V_{ij}(\rho_{ij}) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_i^{-1}(\Phi(z_i)) F_j^{-1}(\Phi(z_j)) \varphi(\rho_{ij}, z_i, z_j) dz_i dz_j - R_{ij}, \\
 \varphi(\rho_{ij}, z_i, z_j) &= \frac{1}{2 \rho_{ij} \sqrt{1 - \rho_{ij}^2}} \exp \left\{ -\frac{1}{2 \sqrt{1 - \rho_{ij}^2}} (z_i^2 + z_j^2 - 2 \rho_{ij} z_i z_j) \right\}, \quad (3) \\
 R_{ij} &= m_i m_j + r_{ij} \sigma_i \sigma_j, \quad m_i = E(x_i), \quad \sigma_i = \sqrt{E(x_i - m_i)^2}, \\
 i &= \overline{1, n-1}; \quad j = \overline{i+1, n}.
 \end{aligned}$$

В работе [11] при достаточно общих ограничениях на свойства функций распределения $F_i(x)$, устанавливается ряд свойств уравнений (3) и, в частности, существование их решений.

Вводя в двойном интеграле новые переменные интегрирования y_i и y_j согласно формулам

$$y_i = \Phi(z_i), \quad y_j = \Phi(z_j)$$

и учитывая свойства функции $\Phi(z)$ уравнения (3) запишем в следующем виде, удобном для построения численных методов решения этих уравнений

$$\begin{aligned}
 V_{ij}(\rho_{ij}) &= 0, \quad V_{ij}(\rho_{ij}) = \int_{1/2}^1 \int_{1/2}^1 g(\rho_{ij}, y_i, y_j) dy_i dy_j - R_{ij}, \\
 g(\rho_{ij}, y_i, y_j) &= (F_i^{-1}(y_i) F_j^{-1}(y_j) + F_i^{-1}(1 - y_i) F_j^{-1}(1 - y_j)) f(\rho_{ij}, z_i, z_j) + \\
 &\quad + (F_i^{-1}(y_i) F_j^{-1}(1 - y_j) + F_i^{-1}(1 - y_i) F_j^{-1}(y_j)) f(-\rho_{ij}, z_i, z_j), \\
 f(\rho_{ij}, z_i, z_j) &= \frac{1}{\sqrt{1 - \rho_{ij}^2}} \exp \left\{ -\frac{\rho_{ij}}{2(1 - \rho_{ij}^2)} (\rho_{ij}(z_i^2 + z_j^2) - 2z_i z_j) \right\}, \\
 z_i &= \Phi^{-1}(y_i), \quad z_j = \Phi^{-1}(y_j); \quad i = \overline{1, n-1}; \quad j = \overline{i+1, n}, \quad (4)
 \end{aligned}$$

где $\Phi^{-1}(y)$ функция, обратная к функции $\Phi(y)$. Для решения системы уравнений (3) или в эквивалентной форме (4) были использованы три подхода.

1. Каждое уравнение $V_{ij}(\rho_{ij}) = 0$ системы (4), зависящее только от одного неизвестного ρ_{ij} , решалось независимо от других уравнений четырьмя численными методами, а именно методом дихотомии, методом секущих, методом парабол и с помощью процедуры поиска

решения (Solver) в приложении Excel 2003. Двойной интеграл вычислялся двумя численными методами наилучшей степени точности (методами Гаусса): двумерными аналогами методов прямоугольников и трапеций [12].

2. В этом случае использовалась система уравнений в виде (3). Каждое уравнение решалось теми же методами, что и в случае 1. При этом двойные интегралы вычислялись имитационным способом (методом Монте-Карло). Использование метода Монте—Карло обусловлено, с одной стороны, простотой его реализации, а с другой стороны, тесной связью его с общей идеологией вероятностного моделирования и прогнозирования.
3. Поскольку в случаях 1 и 2 в результате решения соответственно уравнений (4) и (3) находятся значения ρ_{ij} , то они, будучи по смыслу коэффициентами корреляции, должны удовлетворять условию, согласно которому корреляционная матрица

$$K = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} & \dots & \rho_{1n} \\ \rho_{21} & 1 & \rho_{23} & \dots & \rho_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \rho_{n1} & \rho_{n2} & \rho_{n3} & \dots & 1 \end{pmatrix}, \quad \rho_{ij} = \rho_{ji}, \quad i, j = \overline{1, n}$$

должна быть неотрицательно определенной. Но так как значения ρ_{ij} получаются численными методами, то они вычисляются приближенно, с некоторой точностью и, вообще говоря, могут не удовлетворять указанному условию, т. е. построенная на их основе корреляционная матрица может оказаться не неотрицательно определенной, что исключает возможность генерирования случайных векторов с заданными свойствами. Чтобы гарантировать неотрицательную неопределенность соответствующей корреляционной матрицы в данной работе предлагается (и это одно из основных отличий от других работ) из уравнений (3) находить не сами значения ρ_{ij} , а элементы a_{ij} нижней треугольной матрицы (матрицы Холецкого), определяемой из условия

$$K = AA^T, \\ A = \begin{pmatrix} a_{11} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix} \Rightarrow \sum_{j=1}^i a_{ij}^2 = 1, \quad i = \overline{1, n}. \quad (5)$$

Знание матрицы **A**, как известно, достаточно для генерации случайного вектора **Z**, и, стало быть, согласно (2) и случайного вектора **X**. При данном подходе уже приходится решать систему уравнений (3), а не каждое уравнение в отдельности. Однако решение этой системы уравнений

в силу структуры матрицы \mathbf{A} можно свести к последовательному решению $n(n-1)$ (по числу коэффициентов корреляции ρ_{ij}) скалярных уравнений, в результате решения которых последовательно вычисляются значения

$$a_{11} \rightarrow a_{21} \rightarrow a_{31} \rightarrow \dots \rightarrow a_{n1} \rightarrow a_{22} \rightarrow a_{32} \rightarrow \dots \rightarrow a_{n2} \rightarrow \dots \rightarrow \\ \rightarrow a_{n-1n-1} \rightarrow a_{nn-1} \rightarrow a_{nn}$$

В работе эта редукция системы уравнений (3) не использовалась. Система (3) решалась двумя способами:

- (1) методом Ньютона с выбором шага вдоль направления спуска, основанного на параболической аппроксимации целевой функции вдоль направления спуска,
- (2) с помощью процедуры поиска решения (Solver) в приложении Excel 2003.

При нахождении решения системы (3) методом Ньютона существенно была использована структура матрицы Холецкого \mathbf{A} , что позволило существенно сократить вычислительные затраты. Значения двойных интегралов в обоих методах вычислялись имитационным способом.

Описанные методы были программно реализованы на языке Visual Basic for Applications в среде Excel 2003, где и были сделаны все расчеты. Численный анализ эффективности методов проводился на одном и том же тестовом примере. В этом примере считалось, что все компоненты случайного вектора $\mathbf{X} = (x_1 \ x_2 \ \dots \ x_n)^T$ имеют одну и то же (маргинальную) функцию распределения $F(x)$, равную экспоненциальной функции распределения с единичным математическим ожиданием и единичной дисперсией т. е.

$$F_i(x) = F(x) = 1 - e^{-x}, \quad 0 \leq x < \infty; \quad m_i = 1, \quad \sigma_i = 1; \quad i = \overline{1, n}$$

В качестве коэффициентов корреляции r_{ij} случайного вектора \mathbf{X} были взяты следующие значения

$$r_{ij} = r^{|i-j|}, \quad r = 0,9 \Rightarrow R_{ij} = 1 + r_{ij}; \quad i, j = \overline{1, n}$$

В следующем разделе дается описание полученных результатов численного анализа эффективности численных методов, предложенных для решения поставленной задачи.

Результаты анализа эффективности предложенных численных методов

Все расчеты велись на одном и том же компьютере и в качестве эффективности численных методов бралось время, которое потребовалось им для нахождения коэффициентов корреляции ρ_{ij} . Для первой группы методов решались уравнения (4), в которых двойные интегралы вычислялись, как уже говорилось в предыдущем разделе, двумерными аналогами

Таблица 1

Число компонент вектора X : $n = 10$ Число искомых параметров ρ_{ij} : $\frac{n(n-1)}{2} = 45$ Точность вычисления значений r_{ij} : 0,001		Время счета (сек)			
Итерационный метод		Метод дихотомии	Метод секущих	Метод парабол	Метод Solver
Метод интегрирования	Метод прямоугольников ($N = 100$)	36	14	14	77
	Метод трапеций ($N = 50$)	20	8	7	32

методов прямоугольников и трапеций. При этом отрезок интегрирования по каждой оси равномерно разбивался на N частей. Значение N было взято соответственно для метода прямоугольников, равным 100, а для метода трапеций — 50. Согласно вычислительным схемам методов интегрирования, последнее означает, что общее число значений подынтегральной функции, которое использовалось для вычисления интегралов обоими методами интегрирования одно и то же и равно 10000. Использование одного и того же числа значений подынтегральной функции в обоих методах интегрирования обусловлено необходимостью обеспечения одинаковой точности интегрирования. В расчетах число компонент n вектора X было положено, равным 10. Итерационный процесс поиска значений ρ_{ij} заканчивался, когда разность между расчетными и заданными значениями коэффициентов корреляции r_{ij} по абсолютной величине становилась меньше 0,001. Полученные результаты представлены в табл. 1.

Из табл. 1 следуют следующие выводы.

Методы, использующие численное интегрирование методом трапеций, работают приблизительно в 2 раза быстрее аналогичных итерационных методов с использованием численного интегрирования методом прямоугольников.

Независимо от метода интегрирования самым лучшим (первым) по вычислительным затратам является итерационный метод парабол, вторым — метод секущих (не значительно отличается от метода парабол), третьим — метод дихотомии (в 2,5 раза хуже метода парабол) и четвертым — метод Solver (в 5,5 и 4,5 раза хуже метода парабол при интегрировании соответственно методом прямоугольников и методом трапеций).

Из сказанного вытекает, что по вычислительным затратам по всем параметрам

- (1) самым лучшим является метод секущих,
- (2) метод парабол не значительно отличается от метода секущих,
- (3) метод Solver является самым худшим методом.

Последнее следовало ожидать, поскольку метод Solver является универсальным методом, предназначенным для решения задач нелинейного программирования в общем случае, в то время как другие методы были адаптированы для решения данной конкретной задачи.

Вторая группа методов, как уже говорилось, отличается от первой только тем, что решались уравнения (3), в которых двумерные интегралы вычислялись методом Монте—Карло. Поскольку согласно (3) каждый интеграл представляет собой математическое ожидание случайной величины $F_i^{-1}(\Phi(z_i))F_j^{-1}(\Phi(z_j))$, зависящей от компонент случайного вектора $\mathbf{Z} = (z_i \ z_j)^\top$, распределенного нормально с нулевым математическим ожиданием и ковариационной матрицей

$$\mathbf{K}_z = \begin{pmatrix} 1 & \rho_{ij} \\ \rho_{ij} & 1 \end{pmatrix},$$

то значение интеграла можно вычислить методом Монте—Карло, используя множество реализаций случайных векторов

$$\mathbf{Z}^{(k)} = (z_i^{(k)} \ z_j^{(k)})^\top = \begin{pmatrix} 1 & 0 \\ \rho_{ij} & \sqrt{1 - \rho_{ij}^2} \end{pmatrix} \tilde{\mathbf{Z}}^{(k)}, \quad \tilde{\mathbf{Z}}^{(k)} = (\tilde{z}_i^{(k)} \ \tilde{z}_j^{(k)})^\top,$$

$$\tilde{\mathbf{Z}}^{(k)} \sim \mathbf{N}(0, \mathbf{I}_2), \quad E(\tilde{\mathbf{Z}}^{(k)}\tilde{\mathbf{Z}}^{(l)\top}) = \delta_{kl}\mathbf{I}_2, \quad \mathbf{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; \quad k, l = \overline{1, N_p}$$

где векторы $\tilde{\mathbf{Z}}^{(k)}$ независимо нормально распределены с нулевым математическим ожиданием и единичной ковариационной матрицей. При достаточно общих условиях с учетом независимости распределения векторов $\tilde{\mathbf{Z}}^{(k)}$ согласно соответствующей предельной теореме (закону больших чисел) имеет место соотношение

$$\lim_{N_p \rightarrow \infty} \frac{1}{N_p} \sum_{k=1}^{N_p} F_i^{-1}(\Phi(z_i^{(k)}))F_j^{-1}(\Phi(z_j^{(k)})) =$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_i^{-1}(\Phi(z_i))F_j^{-1}(\Phi(z_j))\varphi(\rho_{ij}, z_i, z_j)dz_i dz_j$$

где сходимость понимается в вероятностном смысле (она может быть в зависимости от свойств функций распределения $F_i(x)$ сходимостью по вероятности, либо в среднеквадратичном, либо с вероятностью единица).

Таблица 2

Число компонент вектора X : $n = 10$ Число искомых параметров ρ_{ij} : $\frac{n(n-1)}{2} = 45$ Точность вычисления значений r_{ij} : 0,001 Число случайных реализаций N_p : 10 000				
Итерационный метод	Метод дихотомии	Метод секущих	Метод парабол	Метод Solver
Время счета (сек)	205	70	70	203

Из этого соотношения следует, что при достаточно больших значениях N_p справедливо приближенное равенство

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_i^{-1}(\Phi(z_i)) F_j^{-1}(\Phi(z_j)) \varphi(\rho_{ij}, z_i, z_j) dz_i dz_j \approx \frac{1}{N_p} \sum_{k=1}^{N_p} F_i^{-1}(\Phi(z_i^{(k)})) F_j^{-1}(\Phi(z_j^{(k)})), \quad N_p \gg 1.$$

Основываясь на этом равенстве, вместо уравнений (3) решались уравнения

$$V_{ij}(\rho_{ij}) = 0, \quad V_{ij}(\rho_{ij}) = \frac{1}{N_p} \sum_{k=1}^{N_p} F_i^{-1}(\Phi(z_i^{(k)})) F_j^{-1}(\Phi(z_j^{(k)})) - R_{ij},$$

$$\begin{pmatrix} z_i^{(k)} \\ z_j^{(k)} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \rho_{ij} & \sqrt{1 - \rho_{ij}^2} \end{pmatrix} \begin{pmatrix} \tilde{z}_i^{(k)} \\ \tilde{z}_j^{(k)} \end{pmatrix}, \quad \begin{pmatrix} \tilde{z}_i^{(k)} \\ \tilde{z}_j^{(k)} \end{pmatrix} \sim N(0, I_2),$$

$$i = \overline{2, n}, \quad j = \overline{1, i-1}, \quad k = \overline{1, N_p}.$$

В данном случае расчеты проводились на том же тестовом примере и с теми же данными, что и в предыдущем случае. Чтобы обеспечить приблизительно ту же степень точности вычислений, что и в методах с численным интегрированием, число случайных реализаций N_p было взято, равным 10 000 по числу вычислений значений подынтегральных функций, используемых для вычисления интегралов. Результаты расчетов приведены в табл. 2.

Из табл. 2 видно, что в рассматриваемом случае лучшими методами по вычислительным затратам опять являются методы парабол и секущих, причём они в 3 раза считают быстрее, чем методы дихотомии и Solver.

Сравнение времен счета методов в данном случае с временем счета наилучшего метода в предыдущем случае (табл. 1, метод парабол, вариант интегрирования методом трапеций) показывает, что в данном случае методы дихотомии в Solver по сравнению с наилучшим предыдущим методом затрачивают в 30 раз больше времени, методы секущих и парабол — в 10 раз. Такое значительное различие во временах счета методов с использованием интегрирования и методов, основанных на имитационном подходе, объясняется тем, что при вычислении необходимых для решения уравнений (4) интегралов существенно были использованы структуры подынтегральных функций, позволившие выделить достаточно трудоемкие независимые от искоемых параметров ρ_{ij} части, которые были посчитаны заранее, запомнены и использованы для всех уравнений (4). При реализации же методики Монте—Карло применительно к решению уравнений в форме (3) такой декомпозиции вычислений сделать не удалось. Таким образом итерационные методы секущих и парабол сравнимы и являются лучшими по вычислительным затратам как для случая использования численного интегрирования, так и для случая использования методологии Монте—Карло, причем в первом случае работают как минимум в 10 раз быстрее, чем во втором случае.

Следует отметить, что метод дихотомии работает значительно медленнее методов парабол и секущих. Это объясняется двумя причинами. Во-первых, теоретическая скорость сходимости метода дихотомии ниже, чем у методов парабол и секущих. Во-вторых, и это главное, начальное приближение в методе дихотомии значительно хуже в двух других методах. Поскольку метод дихотомии для своей работы в качестве начального приближения требует двух значений $\rho_{ij}^{(1)}$ и $\rho_{ij}^{(2)}$ коэффициента корреляции ρ_{ij} , таких чтобы левая часть уравнения $V_{ij}(\rho_{ij}) = 0$, из которого определяется ρ_{ij} , имела в точках $\rho_{ij}^{(1)}$ и $\rho_{ij}^{(2)}$ значения разных знаков. Поскольку $-1 \leq \rho_{ij} \leq 1$, то чтобы гарантировать указанное свойство, были взяты $\rho_{ij}^{(1)} = 0,99$, $\rho_{ij}^{(2)} = -0,99$, что практически совпадает с крайними значениями ρ_{ij} . В то же время начальные значения ρ_{ij} для остальных методов, включая и метод Solver, были взяты значение r_{ij} и близкие к нему значения (r_{ij} — заданный коэффициент корреляции вектора \mathbf{X}). Опыт работы показал, что значения ρ_{ij} , полученные в результате решения соответствующих уравнений, действительно достаточно близки к r_{ij} .

Третья группа методов, как и вторая группа, основана на идеологии имитационного моделирования с той лишь разницей, генерировались не двумерные, а n -мерные случайные векторы $\mathbf{Z}^{(k)}$ ($k = \overline{1, n}$) и определялись не коэффициенты корреляции ρ_{ij} , а не нулевые элементы a_{ij} ($i = \overline{2, n}$, $j = \overline{1, i}$) матрицы \mathbf{A} , связанные с ρ_{ij} соотношениями (5), из которых в частности следует, что не нарушая общности, можно положить $a_{11} = 1$. Случайные векторы $\mathbf{Z}^{(k)}$ генерировались по следующей схеме

$$\begin{aligned} \mathbf{Z}^{(k)}(\mathbf{A}) &= (z_1^{(k)}(\mathbf{A}) z_2^{(k)}(\mathbf{A}) \dots z_n^{(k)}(\mathbf{A}))^T = \mathbf{A}\tilde{\mathbf{Z}}^{(k)}, \\ \tilde{\mathbf{Z}}^{(k)} &= (\tilde{z}_1^{(k)} \tilde{z}_2^{(k)} \dots \tilde{z}_n^{(k)})^T, \\ \tilde{\mathbf{Z}}^{(k)} &\sim \mathbf{N}(0, \mathbf{I}_n), \quad E(\tilde{\mathbf{Z}}^{(i)}\tilde{\mathbf{Z}}^{(k)T}) = \delta_{ik}\mathbf{I}_n, \quad \mathbf{I}_n = \{\delta_{\alpha\beta}\}_{\alpha\beta=1}^n, \\ & \quad i, k = \overline{1, N_p}, \end{aligned}$$

где векторы $\tilde{\mathbf{Z}}^{(k)}(\mathbf{A})$ независимо нормально распределены с нулевым математическим ожиданием и единичной ковариационной матрицей, \mathbf{A} — матрица в (5), а $\delta_{\alpha\beta}$ — символ Кронекера. Затем не нулевые элементы матрицы \mathbf{A} находились в результате решения следующей системы нелинейных уравнений

$$\begin{aligned} V_{ij}(\mathbf{A}) &= 0, \\ V_{ij}(\mathbf{A}) &= \frac{1}{N_p} \sum_{k=1}^{N_p} F_i^{-1}(\Phi(z_i^{(k)}(\mathbf{A}))) F_i^{-1}(\Phi(z_j^{(k)}(\mathbf{A}))) - R_{ij}; \\ & \quad i = \overline{2, n}, \quad j = \overline{1, i-1}, \\ & \quad a_{11} = 1, \quad \sum_{j=1}^i a_{ij}^2 = 1; \quad i = \overline{2, n}, \end{aligned} \tag{6}$$

состоящей из двух подсистем, первая из которых была получена из тех же соображений, что и в предыдущем случае, а вторая является следствием соотношений (5). Данная система уравнений представляет собой систему $n(n+1)/2 - 1$ нелинейных уравнений с $n(n+1)/2 - 1$ неизвестными $a_{21}, a_{31}, \dots, a_{n1}, a_{22}, a_{32}, \dots, a_{n2}, \dots, a_{n-1n-1}, a_{nn-1}, a_{nn}$. Как уже говорилось, в данной работе для решения этой системы уравнений были предложены и протестированы два метода: метод Ньютона с выбором шага вдоль направления спуска, основанного на параболической аппроксимации целевой функции вдоль этого направления, и стандартный метод, реализованный в процедуре поиска решения (Solver) в приложении Excel 2003. Метод Ньютона был программно реализован с учетом специфики задачи на языке Visual Basic for Application. Тестирование проводилось на том же примере и с теми же данными, что и в предыдущем случае. В качестве начального приближения $\mathbf{A}^{(0)}$ для матрицы \mathbf{A} была взята нижняя треугольная матрица Холецкого для корреляционной матрицы случайного вектора \mathbf{X} , т. е.

$$\mathbf{A}^{(0)} = \mathbf{A}_x = \begin{pmatrix} a_{x11} & 0 & 0 & \dots & 0 \\ a_{x21} & a_{x22} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{xn1} & a_{xn2} & a_{xn3} & \dots & a_{xnn} \end{pmatrix},$$

Таблица 3

Число компонент вектора X : $n = 10$ Число искомых параметров a_{ij} : $\frac{n(n+1)}{2} - 1 = 54$ Точность вычисления значений r_{ij} : 0,001 Число случайных реализаций N_p : 10 000		
Итерационный метод	Метод Ньютона	Метод Solver Конечно-разностный аналог метода Ньютона
Время счета (сек)	26	11 398

$$\mathbf{A}_x \mathbf{A}_x^T = \mathbf{K}_x = \begin{pmatrix} 1 & r_{21} & r_{31} & \dots & r_{n1} \\ r_{21} & 1 & r_{32} & \dots & r_{n2} \\ \dots & \dots & \dots & \dots & \dots \\ r_{n1} & r_{n2} & r_{n3} & \dots & 1 \end{pmatrix}.$$

В табл. 3 приведены значения времени счета, которые потребовались для решения системы уравнений методом Ньютона, а также конечно-разностными аналогами метода Ньютона и сопряженных градиентов, реализованных в процедуре поиска решения (метод Solver) в Excel 2003.

Из приведенных в табл. 3 данных следует, что метод Ньютона считает в 438 раз быстрее конечно-разностного аналога метода Ньютона. Такое значительное отличие во временах счета для метода Ньютона и его конечно-разностного аналога на наш взгляд можно объяснить следующими несколькими причинами.

1. В методе Ньютона с выбором шага вдоль направления спуска нужные первые производные от левых частей уравнений по искомым параметрам считались по аналитическим формулам. Следствием этого, как известно, является более высокая теоретическая и, следовательно, практическая скорость сходимости метода Ньютона по сравнению с его конечно-разностными аналогами.

2. При выводе формул для производных и их программной реализации существенно была использована специфика решаемой задачи, а именно тот факт, что разные уравнения зависят от разного числа искомых параметров, что обусловлено треугольной структурой матрицы Холецкого A . Это позволило существенно повысить быстродействие работы метода Ньютона.

3. В процессе программной реализации метода Ньютона были выявлены значительные по вычислительным затратам общие части, используемые как при вычислении левых частей уравнений, так при вычислении производных, что тоже позволило значительно ускорить работу метода Ньютона.

4. Был реализован экономичный алгоритм поиска минимума целевой функции (сумма квадратов невязок уравнений) вдоль направления спуска, основанный на аппроксимации целевой функции вдоль этого направления квадратной параболой. Этот алгоритм был ориентирован на случай, когда шаг вдоль направления спуска близок к 1. Алгоритм был устроен так, что, если при начальном шаге вдоль направления спуска, равном 1, значение целевой функции было меньше первоначального значения, то делалась обычная итерация по методу Ньютона с шагом 1 без дополнительных пересчетов левых частей уравнений. Практический опыт работы показал, что подавляющее число итераций оказались именно такими, что означает, что данный метод практически работает как обычный метод Ньютона с шагом 1.

5. Процедура поиска решения Solver, реализованная в среде Excel 2003, не умеет решать систему уравнений, а может решать либо одно трансцендентное уравнение, либо задачу нелинейного программирования, т. е. задачу оптимизации целевой функции при ограничениях на параметры, заданных в виде равенств и неравенств. Поэтому решение системы уравнений (6) с помощью процедуры Solver было сведено к решению задачи минимизации целевой функции в виде суммы квадратов невязок уравнений (6) по ненулевым элементам матрицы A . Поскольку метод Ньютона, реализованный в процедуре Solver, является его конечно-разностным аналогом, в котором на каждой итерации разностным методом вычисляются производные, то на выполнение каждой итерации этому методу требуется, как минимум, $n(n + 1)/2$ вычислений целевой функции, а, стало быть, и значений левых частей уравнений (6). В тоже время для метода Ньютона с аналитическим вычислением производных требуется одно вычисление левых частей уравнений (6) и производных, на расчет которых в силу сказанного выше, уходит существенно меньше времени, чем на вычисление левых частей уравнений (6). К тому же, как показали расчеты, число итераций, которое затрачивалось для решения системы уравнений с заданной точностью конечно-разностным аналогом Ньютона, было на порядок больше чем для метода Ньютона с аналитическим вычислением производных (в тестовом примере Метод Ньютона сделал 3 итерации, а его конечно-разностный аналог — 30). При этом наблюдалось существенное ухудшение сходимости конечно-разностного метода Ньютона по мере приближения к решению, особенно для больших размерностей случайных векторов. Последнее наиболее вероятно является следствием конечно-разностной аппроксимации производных.

Сравнение метода Ньютона с методами второй группы (табл. 2), показывает, что в среднем метод Ньютона считает в 2,5 раза быстрее методов секущих и хорд и 8 раз быстрее методов дихотомии и Solver. Такое различие во временах счета объясняется тем, что в методах второй группы решается серия нелинейных скалярных уравнений и поэтому каждый раз

при решении очередного уравнения приходится запускать итерационный процесс снова, что требует дополнительных вычислительных затрат, тогда как в методе Ньютона итерационный процесс стартует один раз и все искомые параметры вычисляются одновременно. К тому же все методы второй группы являются конечно-разностными методами. Что касается методов первой группы, то лучшие методы этой группы (методы секущих и парабол с численным интегрированием методом трапеций) считают в среднем в 3,5 раза быстрее метода Ньютона, что обусловлено декомпозицией вычислительного процесса в этих методах.

Можно дать оценку времени работы лучших методов первой группы и метода Ньютона, когда число компонент случайных векторов $n = 100$, исходя из полученных результатов и того факта, что в этих методах время счета приблизительно пропорционально числу искомых параметров, которое в первом случае равно $n(n-1)/2$, а во втором случае — $n(n+1)/2-1$. Такие оценки были сделаны. В результате получилось, что методы секущих и парабол с численным интегрированием методом трапеций будут затрачивать на решение данной задачи при $n = 100$ раз в 110 больше времени, чем при $n = 10$ (это порядка 13 минут счета на данной машине), а метод Ньютона — в 93 раза (порядка 40 минут счета на данной машине). Хотя эти оценки, сделанные для конкретной функции распределения $F(x)$ случайного вектора \mathbf{X} , зависят от сложности этой функции, но эта зависимость достаточно слабая.

Резюмируя выше изложенное, заключаем, что самыми эффективными в смысле вычислительных затрат являются методы секущих и парабол с численным интегрированием двумерным аналогом метода трапеций с использованием квадратурной формулы наилучшей степени точности (метод Гаусса), а также разработанный и программно реализованный авторами применительно к данной задаче метод Ньютона с аналитическим вычислением производных. Хотя метод Ньютона считает медленнее, чем два других метода, но он имеет по сравнению с ними одно преимущество, согласно которому он в отличие от них заведомо гарантирует положительную определенность корреляционной матрицы \mathbf{K} вектора \mathbf{Z} , без чего генерирование случайных векторов невозможно. В первых двух методах из-за приближенности вычислений матрица \mathbf{K} , вообще говоря, может не оказаться положительно определенной и, стало быть, задача генерирования случайных векторов не может быть решена. Указанные методы позволяют на современных вычислительных машинах решать задачу генерирования случайных векторов с данными маргинальными распределениями и корреляционной матрицей за вполне приемлемое время даже в случае генерирования случайных векторов большой размерности.

В заключение авторы выражают благодарность за инициализацию данной работы сотруднику Венского института демографии Австрийской академии наук Щербову С. Я.

Литература

1. *Cario M. C., B. I. Nelson.* Autoregressive to Anything: Time Series Input Processes for Simulation // *Operations Research Letters.* 1996. 19. P. 51–58.
2. *Cario M. C., B. I. Nelson.* Numerical Methods for Fitting and Simulating Autoregressive-to-Anything Processes // *INFORMS Journal on Computing,* forthcoming. 1997.
3. *Cook R. D., Johnson M. F.* A family of distributions for modeling non-elliptically symmetric multivariate data // *Journal of the Royal Statistical Society. B.* 1981. 43. 210–218.
4. *Devroye I.* *Nov-Uniform Random Variate Generation.* New York, Springer-Verlag, 1986.
5. *Hill R. R., Reilly C. H.* Composition for multivariate random vectors // *Proceedings of the 1991 Winter Simulation Conference.* 1994. P. 332–339.
6. *Johnson, M. F.* *Multivariate Statistical Simulation.* New York, John Wiley, 1987.
7. *Li S. T., J. I. Hammond.* Generation of pseudorandom numbers with specified univariate distributions and correlation coefficients // *IEEE Transactions on Systems, Man and Cybernetics.* 1975. 5. P. 557–561.
8. *Mardia K. V.* A translation family of bivariate distributions and Frechet's bounds. *Sankhya. A.* 1970. 32. P. 119–122.
9. *Rubinstein R. Y., Samorodnitsky G., Shaked M.* Antithetic Variates. *Multivariate Dependence and Simulation of Stochastic Systems* // *Management Science.* 1985. 31. P. 66–77.
10. *Schweizer, B.* Thirty years of copulas // *Advances in Probability Distributions with Given Marginals: Beyond the Copulas* (G. Dall'Aglio, S. Kotz and G. Salinetti eds.), 13–50. Boston, Kluwer, 1991.
11. *Cario M. C., Nelson B. L.* Modeling and Generating Random Vectors with Arbitrary Marginal Distributions and Correlation Matrix. Technical Report, Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, Illinois, USA. 1997.
12. *Березин И. С., Жидков П.* Методы вычислений. Т. 1. М.: Наука, ГРФ-МЛ, 1966. С. 247–254.