



Д.А. РАЧКОВСКИЙ

УДК 004.22+004.93'11

**ФОРМИРОВАНИЕ ОТРАЖАЮЩИХ СХОДСТВО
БИНАРНЫХ ВЕКТОРОВ С ИСПОЛЬЗОВАНИЕМ
СЛУЧАЙНЫХ БИНАРНЫХ ПРОЕКЦИЙ**

Аннотация. Предложено преобразование входных вещественных векторов в выходные бинарные векторы путем проецирования бинарной случайной матрицей с элементами $\{0, 1\}$ и порогового преобразования. Исследована скорость сходимости распределения компонентов векторов до бинаризации к гауссовому, а также ошибка оценки угла между входными векторами по бинаризованным выходным векторам. Показано, что при выборе параметров проецирования, для которых достигается близость распределения к гауссовому, экспериментальная и аналитическая ошибки близки.

Ключевые слова: бинарные случайные проекции, сходимость к гауссову распределению, оценка сходства векторов

ВВЕДЕНИЕ

Поиск, классификация, кластеризация по сходству (см. [1–3] и ссылки к ним) широко применяются как в технических системах (поисковые системы интернет и интранет, экспертные системы на основе рассуждений по прецедентам и др.), так и людьми (память, рассуждения по аналогии и т.п.) [1, 4, 5]. Многие виды используемой для этого информации можно представить в векторном формате. Это касается не только данных в виде однородных признаков и полей [6–8], но и сложных структур с иерархическими системами отношений [9–17], которые применяются в базах знаний (в том числе при моделировании рассуждений по аналогии [1, 4, 11, 16, 17]). Векторные операции легко распараллеливаются, что позволяет использовать системы распределенных вычислений для обработки (например, Grid [18, 19]).

Размерность векторов, применяемых для адекватного представления сложных объектов, может быть очень большой; велико и число объектов в базе. Это замедляет обработку, требующую выполнения матрично-векторных операций (вычисление мер сходства–различия векторов, построение линейных моделей, ассоциативная память и т.п.). Одно из направлений повышения эффективности работы с векторными данными — преобразование исходных векторов в другой формат (например, в векторы сокращенной размерности или из вещественных векторов в бинарные), который позволял бы более экономно их хранить и быстрее обрабатывать, получая близкие результаты обработки. При этом необходимо обеспечить малые затраты вычислительных ресурсов и памяти на выполнение непосредственно самого преобразования.

К преобразованиям, позволяющим оценить меры сходства–различия входных векторов по выходным, относится проецирование с помощью случайных матриц, элементы которых — случайно сгенерированные числа из некоторого

© Д.А. Рачковский, 2015

распределения [20–24]. Случайные проекции имеют также множество других применений, например, для устойчивого решения дискретной некорректной обратной задачи [25, 26].

В работах [20–22] исследовано проецирование случайными матрицами с гауссовым распределением, тернарным распределением с элементами из $\{-1, 0, +1\}$ и $\{-1, +1\}$. В [24] предложен и исследован простой и вычислительно эффективный с точки зрения генерации и использования вариант случайной матрицы — бинарная с элементами $\{0, 1\}$. Показана возможность оценки скалярного произведения, евклидова расстояния, нормы входных вещественных векторов по выходным вещественным.

В работах [21–23] исследовано формирование бинарных выходных векторов, по которым можно оценить сходство входных вещественных векторов. В [21, 22] рассмотрены «плотные» бинарные векторы (с половиной единичных компонентов), а в [23] — бинарные векторы с регулируемой долей единичных компонентов (разреженностью). Разреженные бинарные векторы требуются для ряда парадигм обработки информации, например, в ассоциативно-проективных нейронных сетях [11, 12, 27–29], а также для некоторых вариантов [30–34] ассоциативной памяти матричного типа [3, 17, 30–37].

Формирование отражающих сходство бинарных выходных векторов с регулируемой разреженностью посредством применения бинарной случайной матрицы с элементами $\{0, 1\}$ ранее не рассматривалось. Данная статья посвящена исследованию этого вопроса.

1. ПРОЕЦИРОВАНИЕ СЛУЧАЙНОЙ БИНАРНОЙ МАТРИЦЕЙ И ФОРМИРОВАНИЕ БИНАРНЫХ ВЕКТОРОВ

Как и в [24], рассмотрим проецирование векторов случайной бинарной матрицей $\mathbf{R}$ с элементами r_{ij} из множества $\{0, 1\}$. Случайные величины (с.в.), реализацией которых являются элементы $\mathbf{R}$ (единицы и нули), независимы и имеют одинаковое распределение (i.i.d.). Значение 1 каждый r_{ij} принимает с вероятностью q , а значение 0 — с вероятностью $1 - q$. Обозначим $\mathbf{x}, \mathbf{y}$ входные вещественные векторы (размерности D), $\mathbf{u} = \mathbf{R}\mathbf{x}$, $\mathbf{v} = \mathbf{R}\mathbf{y}$ — результат их проецирования (размерности d). Следовательно, размерность $\mathbf{R}$ равна $(d \times D)$.

При проецировании каждый компонент $u_i, i = 1, \dots, d$, вектора $\mathbf{u}$ первоначально формируется как скалярное произведение строки $\mathbf{r}_i$ матрицы $\mathbf{R}$ на $\mathbf{x}$: $u_i = \langle \mathbf{r}_i, \mathbf{x} \rangle = \sum_{j=1}^D r_{ij} x_j$. Поскольку с.в. u_i также являются i.i.d., иногда будем рассматривать $u = \sum_{j=1}^D r_j x_j$, где r_j — элементы строки $\mathbf{R}$, которая умножается на входной вектор. Стандартизация u осуществляется вычитанием математического ожидания (м.о.) (центрирование) и делением на среднеквадратичное отклонение. Введем также с.в. $\xi_j \equiv x_j r_j$. Для $j \neq k$ с.в. ξ_j, ξ_k независимы.

Формирование бинарного выходного вектора осуществляется бинаризирующим пороговым преобразованием $\mathbf{u} \rightarrow \mathbf{z}$:

$$\begin{aligned} z_i &= 1 \text{ при } u_i > t_i, \\ z_i &= 0 \text{ в противном случае,} \end{aligned} \quad i = 1, \dots, d, \quad (1)$$

где t_i — величина порога для i -го компонента выходного вектора. Будем использовать одинаковые величины порогов для всех компонентов, $t_i \equiv t$. Заданная вероятность p единичного компонента z бинарного выходного вектора $\mathbf{z}$ определяется выбором соответствующего порога t_p . Для стандартизованной с.в. u с гауссовым распределением

$$p(z=1) = p(u > t_p) = \frac{1}{\sqrt{2\pi}} \int_{t_p}^{\infty} e^{-\eta^2/2} d\eta = 1 - \Phi(t_p), \quad (2)$$

где Φ — гауссова кумулятивная функция распределения. Величина t_p для обеспечения нужной p определяется как квантиль гауссова распределения, соответствующая $1-p$. Векторы при $p < 0.5$ являются разреженными.

Для оценки угла θ между входными векторами $\mathbf{x}, \mathbf{y}$ в [23] предложено использовать оценки вероятности совпадения единичных компонентов $z_{1,i} = 1$ и $z_{2,i} = 1$ в векторах $\mathbf{z}_1$ и $\mathbf{z}_2$ после бинаризации $\mathbf{u} \rightarrow \mathbf{z}_1$ с порогом t_1 и $\mathbf{v} \rightarrow \mathbf{z}_2$ с порогом t_2 . Для стандартизованных u_i, v_i эту вероятность можно определить как результат вычисления интеграла от двумерного гауссова распределения:

$$\begin{aligned} p(z_{1,i} = 1, z_{2,i} = 1 | \theta, t_1, t_2) &= p(u_i > t_1, v_i > t_2 | \theta) = \\ &= \frac{1}{2\pi(1-\cos^2 \theta)} \int_{t_1}^{\infty} \int_{t_2}^{\infty} \exp\left(-\frac{\eta_1^2 - 2\eta_1\eta_2 \cos \theta + \eta_2^2}{2(1-\cos^2 \theta)}\right) d\eta_1 d\eta_2. \end{aligned} \quad (3)$$

Поэтому θ можно оценить по эмпирической вероятности совпадения единичных компонентов результирующих векторов $\langle \mathbf{z}_1, \mathbf{z}_2 \rangle / d$, где $\langle \mathbf{z}_1, \mathbf{z}_2 \rangle$ — скалярное произведение.

Таким образом, формирование бинарных разреженных векторов с заданной степенью разреженности путем порогового преобразования (1) с заранее вычисленным порогом t_p и анализ их свойств для оценки мер сходства входных векторов согласно подходу [23] основаны на предположении о гауссовом распределении компонентов выходных вещественных векторов до бинаризации пороговым преобразованием. Исследуем это предположение для бинарной случайной матрицы $\mathbf{R}$.

2. СХОДИМОСТЬ К ГАУССОВУ РАСПРЕДЕЛЕНИЮ И СКОРОСТЬ СХОДИМОСТИ

2.1. Сходимость к гауссову распределению. Рассмотрим независимые (и не одинаково распределенные) с.в. $\xi_1, \dots, \xi_D$ с м.о. $E\{\xi_j\}$, дисперсиями $\sigma_j^2 \equiv V\{\xi_j\} = E\{(\xi_j - E\{\xi_j\})^2\}$, абсолютными центральными моментами $E\{|\xi_j - E\{\xi_j\}|^{2+\delta}\}$ порядка $2+\delta$, $\delta > 0$. Обозначим сумму этих с.в. $S_D \equiv u = \sum_{j=1}^D \xi_j$, ее м.о., равное сумме м.о. слагаемых $E\{S_D\} \equiv A_D = \sum_{j=1}^D E\{\xi_j\}$, ее дисперсию, равную сумме дисперсий независимых слагаемых $V\{S_D\} \equiv B_D^2 = \sum_{j=1}^D \sigma_j^2$, сумму абсолютных центральных моментов слагаемых $\sum_{j=1}^D E\{|\xi_j - E\{\xi_j\}|^{2+\delta}\}$.

Центральные предельные теоремы (ЦПТ) теории вероятностей утверждают, что при выполнении определенных условий стандартизованные с.в. $(S_D - A_D) / B_D, B_D \equiv \sqrt{B_D^2}$ (суммы независимых случайных слагаемых) сходятся по распределению к стандартной гауссовой с.в. $\mathcal{N}(0, 1)$ при $D \rightarrow \infty$. Иными словами, в этом случае имеет место равномерная сходимость последовательности кумулятивных функций распределения к стандартной гауссовой Φ .

Условие Ляпунова (обобщенное) [38]. Пусть конечны абсолютные центральные моменты порядка $2+\delta$: $E\{|\xi_j - E\{\xi_j\}|^3\} < \infty$ (соответственно конечны м.о. $E\{\xi_j\}$ и дисперсии $V\{\xi_j\}$). Справедлива ЦПТ, если для некоторого $\delta > 0$ при $D \rightarrow \infty$ выполняется

$$\frac{1}{B_D^{2+\delta}} \sum_{j=1}^D E\{|\xi_j - E\{\xi_j\}|^{2+\delta}\} \rightarrow 0, \quad (4)$$

где выражение в левой части — дробь Ляпунова порядка $2+\delta$. Заметим, что величина этой дроби не изменится при одинаковом масштабировании и сдвиге всех с.в. в сумме.

Условие Ляпунова ограничивает скорость роста абсолютных центральных моментов порядка $2+\delta, \delta>0$.

Условие Линдеберга [38]. Пусть конечны м.о. $E\{\xi_j\}$ и дисперсии $V\{\xi_j\} = E\{(\xi_j - E\{\xi_j\})^2\}$. Справедлива ЦПТ, если для любых $\varepsilon>0$ при $D \rightarrow \infty$ выполняется

$$\frac{1}{B_D^2} \sum_{j=1}^D E\{(\xi_j - E\{\xi_j\})^2\} I\{|\xi_j - E\{\xi_j\}| \geq \varepsilon B_D\} \rightarrow 0, \quad (5)$$

где $I\{\dots\}$ — индикаторная функция, принимающая значение 1 при выполнении условия в скобках и 0 в противном случае.

Смысл условия Линдеберга в том, что ни одно из слагаемых в сумме не доминирует, т.е. его вклад мал по сравнению с суммой.

Условие Ляпунова более сильное, чем условие Линдеберга. Так как $|\xi_j - E\{\xi_j\}| \geq \varepsilon B_D$, то $\left| \frac{\xi_j - E\{\xi_j\}}{\varepsilon B_D} \right|^\delta \geq 1$. Поэтому

$$\begin{aligned} & \frac{1}{B_D^2} \sum_{j=1}^D E\{(\xi_j - E\{\xi_j\})^2\} I\{|\xi_j - E\{\xi_j\}| \geq \varepsilon B_D\} \leq \\ & \leq \frac{1}{\varepsilon^\delta B_D^{2+\delta}} \sum_{j=1}^D E\{|\xi_j - E\{\xi_j\}|^{2+\delta}\} I\{|\xi_j - E\{\xi_j\}| \geq \varepsilon B_D\} \leq \\ & \leq \frac{1}{\varepsilon^\delta B_D^{2+\delta}} \sum_{j=1}^D E\{|\xi_j - E\{\xi_j\}|^{2+\delta}\}. \end{aligned}$$

Если выполняется условие Ляпунова, то при $D \rightarrow \infty$ последнее выражение стремится к нулю при любом $\varepsilon>0$ (постоянная $1/\varepsilon^\delta$ не влияет на предел). Значит, выполняется и условие Линдеберга.

На практике удобнее проверять условие Ляпунова для $\delta=1$. Соответственно найдем дробь Ляпунова 3-го порядка для рассматриваемых далее случайных величин. Так как $E\{\xi_j\} = E\{r_j x_j\} = E\{r_j\} x_j$, то

$$E\{|\xi_j - E\{\xi_j\}|^3\} = E\{|r_j x_j - E\{r_j\} x_j|^3\} = |x_j|^3 E\{|r_j - E\{r_j\}|^3\}.$$

Итак, числитель дроби $\sum_{j=1}^D E\{|\xi_j - E\{\xi_j\}|^3\} = E\{|r_j - E\{r_j\}|^3\} \sum_{j=1}^D |x_j|^3$, знаменатель $B_D^3 = \left(\sum_{j=1}^D \sigma_j^2 \right)^{3/2} = \left(\sum_{j=1}^D E\{|\xi_j - E\{\xi_j\}|^2\} \right)^{3/2} \equiv V^{3/2} \{u\} \equiv V^{3/2} \{S_D\}$. Так как $\sigma_j^2 = E\{|\xi_j - E\{\xi_j\}|^2\} = |x_j|^2 E\{|r_j - E\{r_j\}|^2\}$, то $B_D^3 = \|x\|_2^3 (E\{|r_j - E\{r_j\}|^2\})^{3/2}$.

Таким образом, дробь Ляпунова 3-го порядка для суммы с.в., полученных при умножении входного вектора x на строку случайной матрицы, можно представить как

$$\begin{aligned} L_D &= \frac{E\{|r - E(r)|^3\}}{(E\{|r - E(r)|^2\})^{3/2}} \frac{\sum_{j=1}^D |x_j|^3}{\|x\|_2^3} = L_r L_x, \\ L_r &\equiv \frac{E\{|r - E(r)|^3\}}{(E\{|r - E(r)|^2\})^{3/2}}, \quad L_x \equiv \frac{\sum_{j=1}^D |x_j|^3}{\|x\|_2^3}, \end{aligned} \quad (6)$$

где L_x зависит от входного вектора $\mathbf{x}$ и от D , не зависит от вида случайной матрицы $\mathbf{R}$; L_r зависит от распределения с.в., использованных для генерации элементов $\mathbf{R}$.

Для бинарной $\mathbf{R}$: $E\{|r - E\{r\}|^3\} = E\{|r - q|^3\}$, так как $E\{r\} = 1q + 0(1 - q) = q$. Поэтому $E\{|r - q|^3\} = (1 - q)^3 q + |0 - q|^3 (1 - q) = (q - q^2)((1 - q)^2 + q^2) = (q - q^2) \times (1 - 2(q - q^2))$, $E\{|r - q|^2\} = (1 - q)^2 q + |0 - q|^2 (1 - q) = (q - q^2)$, а также

$$L_r = \frac{1 - 2(q - q^2)}{\sqrt{(q - q^2)}}. \quad (7)$$

Согласно усиленному закону больших чисел эмпирическое (выборочное) среднее D реализаций с.в. ξ_j из одного распределения (i.i.d.) $\sum_{j=1}^D \xi_j / D$ сходится почти наверное (almost surely) к м.о. (если оно существует и конечное, т.е.

$$E\{\xi_j\} < \infty) \text{ при } D \rightarrow \infty. \text{ Представим } L_x \text{ из (6) как } L_x = \frac{1}{\sqrt{D}} \frac{\sum_{j=1}^D |x_j|^3 / D}{\left(\sum_{j=1}^D |x_j|^2 / D\right)^{3/2}}.$$

Будем считать, что компоненты x_j входного вектора $\mathbf{x}$ являются реализациями с.в. с конечным третьим абсолютным центральным моментом $E\{|x_j|^3\} < \infty$. Тогда моменты меньших порядков также конечны, в частности $E\{x_j^2\} \equiv E\{|x_j|^2\} < \infty$.

Применяя усиленный закон больших чисел для $\xi_j \equiv |x_j|^3$ и для $\xi_j \equiv |x_j|^2$, получаем, что при $D \rightarrow \infty$ $\sum_{j=1}^D |x_j|^3 / D \rightarrow E\{|x_j|^3\} < \infty$, $\sum_{j=1}^D |x_j|^2 / D \rightarrow$

$\rightarrow E\{|x_j|^2\} < \infty$. Поэтому (обозначив постоянную $a \equiv \frac{E\{|x_j|^3\}}{(E\{|x_j|^2\})^{3/2}}$) получаем

$$L_x = \frac{1}{\sqrt{D}} \frac{\sum_{j=1}^D |x_j|^3 / D}{\left(\sum_{j=1}^D |x_j|^2 / D\right)^{3/2}} \rightarrow \frac{a}{\sqrt{D}}. \quad (8)$$

Так как $0 < q < 1$, то $0.5 \leq 1 - 2(q - q^2) < 1$ и из (7) $L_r \leq \sqrt{1 / (q - q^2)}$. Поэтому для бинарной случайной матрицы поведение L_D при $D \rightarrow \infty$ определяется выражением $\sqrt{\frac{1 / (q - q^2)}{D}}$, которое стремится к нулю, если $1 / (q - q^2) = o(D)$.

2.2. Скорость сходимости к гауссову распределению. Оценить скорость сходимости кумулятивной функции распределения суммы независимых случайных величин к Φ позволяет неравенство Бери–Эссеена (НБЭ), предложенное в [39, 40]. Скорость сходимости измеряется равномерным расстоянием (расстоянием Колмогорова) Δ_D между $F_D(t)$ — кумулятивной функцией распределения стандартизованной суммы D с.в., и $\Phi(t)$ — стандартной гауссовой кумулятивной функцией распределения: $\Delta_D \equiv \sup_t |F_D(t) - \Phi(t)|$.

При выполнении условия Ляпунова (4) 3-го порядка (т.е. $\delta = 1$) согласно НБЭ для всех D :

$$\Delta_D \equiv \sup_t |F_D(t) - \Phi(t)| \leq C_0 L_D. \quad (9)$$

Абсолютная константа C_0 не зависит от D (но зависит от δ , в данной статье $C_0 \equiv C_0(\delta=1)$). Нижняя оценка (граница) $C_0: C_E$ ($C_0 \geq C_E$) определена в [41]:

$$C_E = \frac{\sqrt{10}+3}{6\sqrt{2\pi}} \approx \frac{1}{\sqrt{2\pi}} + 0.01079 \approx 0.40973.$$

Верхняя оценка (граница) наименьшего возможного значения C_0 определена в [40] как $C_0 \leq 7.59$ и является предметом интенсивных исследований. В последнее время большинство улучшений получено в работах авторов [42–44]. Наименьшие известные значения $C_0 \leq 0.5583$ (0.4690) для неодинаково (одинаково) распределенных слагаемых приведены в [43]. Для симметричных функций распределений оценки незначительно отличаются: $C_0 \leq 0.5582$ (0.4676).

Для случая, когда все слагаемые могут принимать только два значения и распределение слагаемых не обязательно симметрично, в [45] получено $C_0 = 0.4215$, что близко к C_E и впоследствии улучшено до 0.4138. Однако распределение слагаемых в [45] одинаково в отличие от рассматриваемых распределений с.в. Поэтому будем использовать $C_0 \leq 0.5583$.

Для проецирования бинарной случайной матрицей $L_D \equiv L_r L_x$ определяется (6), (7). Для других случайных матриц, так как C_0 и L_x (для фиксированного входного вектора) постоянны, отличие в величине правой части НБЭ (9) определяется L_r . Для тернарной случайной матрицы с элементами $\{-1, 0, +1\}$

$$L_r \equiv \frac{E\{|r - E(r)|^3\}}{(E\{|r - E(r)|^2\})^{3/2}} = \frac{1}{\sqrt{q}}, \quad (10)$$

так как $E\{r\} = 0$ и $E\{|r|^3\} = 1^3 q/2 + |-1|^3 q/2 + 0^3(1-q) = q$, $E\{|r|^2\} = 1^2 q/2 + |-1|^2 q/2 + 0^2(1-q) = q$.

Интересно также проверить, какой результат дает НБЭ (9) для гауссовой случайной матрицы, т.е. при L_r для сумм стандартных гауссовых с.в. ($E\{r\} = 0$, $E\{|r|^2\} = 1$), которые, как известно, имеют гауссово распределение для любого D . Центральные (относительно среднего, т.е. не зависящие от среднего) абсолютные моменты гауссовой с.в. определяются как в [46, с. 120]

$$E\{|r|^k\} = \begin{cases} (k-1)!! \sigma^k & \text{для четных } k, \\ \sqrt{2/\pi} (k-1)!! \sigma^k & \text{для нечетных } k, \end{cases}$$

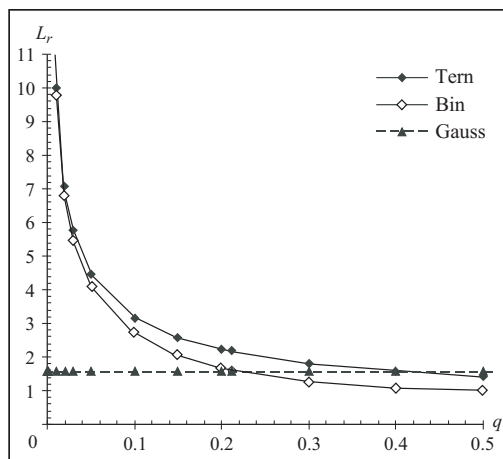


Рис. 1. Зависимость величины сомножителя L_r в дроби Ляпунова 3-го порядка от вероятности q ненулевого значения бинарной (Bin) и тернарной (Tern) с.в.; Gauss — значение L_r для гауссовой величины

где двойной факториал $k!!$ — произведение всех натуральных чисел в $[1, k]$, имеющих такую же четность, что и k . Для $k = 3$ получаем

$$L_r = 2\sqrt{2/\pi}. \quad (11)$$

На рис. 1 приведены зависимости (7), (10) L_r от q для с.в. из $\{0, 1\}$ (Bin) и из $\{-1, 0, +1\}$ (Tern). Показано также значение L_r (11) для гауссова стандартного распределения. Видно, что L_r для бинарной и тернарной случайных матриц близки для всех q (для Bin несколько меньше). Для малых q величина L_r сильно растет, и можно ожидать значительного увеличения отличий распределений от гауссова при тех же D .

Заметим, что аналитическое и экспериментальное исследования скорости сходимости распределения (u_i, v_i) к двумерному гауссову распределению (3) гораздо более трудоемки, чем для случая одномерного распределения, и в данной работе не рассмотрены. Аналитические результаты получены в работах по многомерной ЦПТ [47–49]. Наилучшая скорость сходимости для многомерного распределения размерности a : $O(a^{1/4})$ найдена в [47], однако включает неопределенную абсолютную положительную константу. Экспериментальную оценку левой части НБЭ (9) для этого случая также определить сложно, так как это требует достаточно точной оценки эмпирической двумерной функции распределения.

3. ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ

3.1. Скорость сходимости к гауссову распределению. Проведено сравнительное экспериментальное исследование поведения правой части и оценки левой части НБЭ в зависимости от D для различных типов случайных матриц. Компоненты входных векторов разной размерности D генерировались из равномерного распределения в $[-D, +D]$.

Величина L_x (6) зависит от размерности D , а при фиксированном D — от реализации входного вектора x . На рис. 2 приведены экспериментальные зависимости L_x от D , полученные по 10000 реализаций входных векторов x , для среднего значения L_x (mean), минимального (min), максимального (max), mean-std и mean+std (std — среднеквадратичное отклонение). Видно, что диапазон разброса значений L_x невелик и быстро уменьшается с ростом D . Поэтому в последующих экспериментах значение правой части НБЭ ($C_0 L_D$) и оценка значения левой части НБЭ вычислялись для одного x для каждого D . На рис. 2 также показана величина $\sqrt{D} L_x$ для среднего значения L_x (mean). В соответствии с (8) она (приблизительно) постоянна при различных D .

На рис. 3 приведены теоретические зависимости скорости сходимости $C_0 L_D$ (правой части НБЭ) от D (обозначения кривых начинаются с Т-), вычисленные для проецирования случайными матрицами с элементами из $\{-1, 0, 1\}$ (Tern), из $\{0, 1\}$ (Bin) с вероятностями ненулевого элемента $q = \{0.5, 0.1, 0.01\}$, а также для гауссовой случайной матрицы.

Для вычисления левой части НБЭ Δ_D (9) необходимо знать кумулятивную функцию распределения F_D , аналитическое определение которой для различных D , x и распределений элементов случайных матриц трудоемко. Поэтому будем определять и исследовать оценку Δ_D :

$$\Delta_D^* \equiv \Delta_{D,N} \equiv \sup_u |F_{D,N}(u) - \Phi(u)|, \quad (12)$$

где $F_{D,N}$ — эмпирическая (выборочная) кумулятивная функция распределения, которую будем получать методом Монте-Карло.

Для одного и того же входного вектора x размерности D получим выборку N стандартизованных значений u путем проецирования x на N случайных векторов — строк матрицы R с исследуемым распределением элементов, а также получим эмпирическую кумулятивную функцию распределения $F_{D,N}$. Для этого упорядочим N значений u по возрастанию. Далее для каждого значения u вычислим значение $F_{D,N}(u)$ как $F_{D,N}(u) =$

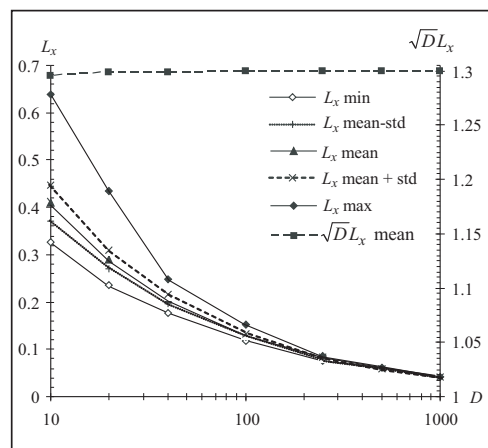


Рис. 2. Зависимость значений L_x (6) (min, max, mean и др.) от D ; зависимость $\sqrt{D} L_x$ (mean) от D

$= \sum_i I\{u_i \leq u\} / N$ ($I\{\dots\}$ — индикатор, как в (5)). В качестве u используем по очереди все значения u_i . Величина $\sum_i I\{u_i \leq u\}$ определяется индексом i значения u

в упорядоченном массиве. При наличии нескольких одинаковых значений u можно использовать наибольшее значение индекса i . Согласно усиленному закону больших чисел эмпирическая кумулятивная функция распределения $F_{D,N}$ сходится почти наверное к теоретической кумулятивной функции распределения F_D при $N \rightarrow \infty$, а по теореме Гнеденко–Кантелли сходимость является равномерной. Заметим, что для каждого D можно выбирать большие N ($N \gg D$).

По полученной эмпирической кумулятивной функции распределения $F_{D,N}(u)$ можно вычислить оценку расстояния Колмогорова $\Delta_{D,N}$ (12) по всем u_i , для которых имеются значения данной функции. Для другой реализации N значений u получим (в общем случае) другие значения в $F_{D,N}(u)$ и для $\Delta_{D,N}$. Поэтому для оценки левой части НБЭ будем использовать Q реализаций эмпирической кумулятивной функции распределения $F_{D,N,l}$, $l=1, \dots, Q$, с соответствующими $\Delta_{D,N,l} \equiv \sup_u |F_{D,N,l}(u) - \Phi(u)|$. В качестве оценки левой части НБЭ будем использовать

$$\Delta_{D,N,Q} = \sup_l \Delta_{D,N,l}. \quad (13)$$

Значения N, Q выберем следующим образом. Пусть F_N — эмпирическая кумулятивная функция распределения, F — кумулятивная функция распределения; ∂ — малая величина; α — число, близкое к 1, вероятность того, что $\sup |F_N(\eta) - F(\eta)| \leq \partial$: $P(\sup |F_N(\eta) - F(\eta)| \leq \partial) = \alpha$. Соответственно вероятность $1-\alpha$ — число, близкое к 0, определяется как $P(\sup |F_N(\eta) - F(\eta)| > \partial) = 1-\alpha$. Вероятность α называют уровнем значимости или доверительной вероятностью. Ей соответствует квантиль распределения Колмогорова

K_α . Для α , близких к 1, $K_\alpha \approx \sqrt{-\frac{1}{2} \ln \frac{1-\alpha}{2}}$. Для определения объема выборки N

используем соотношение $\sqrt{N}\partial = K_\alpha$. Таким образом, $N = (K_\alpha / \partial)^2$. Пусть $\partial = 0.01$, $\alpha = 0.995$, тогда $K_\alpha \approx 1.73$, $N \approx 30000$. Соответствие этих параметров эксперименту можно проверить для известного распределения и сгенерированной по нему выборки, например, для гауссовых эмпирических кумулятивных функций распределения Φ_N и кумулятивных функций распределения Φ . Для оценки α будем использовать $Q = 10000$. Понятно, что если для некоторого N при увеличении D значение (13) становится соизмеримым с ∂ , то дальнейшее увеличение D не имеет смысла (вследствие наличия ошибки моделирования эмпирических кумулятивных функций распределения для данного N).

В экспериментах при $N = 30000$, $Q = 10000$ получены следующие результаты (см. рис. 3, обозначения кривых начинаются с Е-): для проецирования случайной гауссовой матрицей (Gauss), как и предполагалось, оценка (13) левой части НБЭ не зависит от D ввиду гауссовости суммы для любого D ; для бинарной и тернарной случайных проекций ошибки (13) очень близки для всех значений параметров q, D . Для $q = 0.5$ и $D \geq 20$ ошибка не превышает 0.014 и находится на уровне ошибки для гауссовой матрицы (т.е. минимальной для выбранного N). Для $q = 0.1$ и $D \geq 100$ ошибка не превышает 0.017, а при $D = 1000$ находится на уровне ошибки для гауссовой матрицы (т.е. приблизительно 0.01). Для $q = 0.01$ и $D = 1000$ значение ошибки не превышает 0.015. Заметим, что при этих параметрах строки бинарной и тернарной матриц содержат в среднем только 10 ненулевых элементов, т.е. в сумме $u = \sum_{j=1}^D r_j x_j$ имеется только 10 компонентов входного вектора. Соответственно для $q = 0.01$ и $D = 100$ в сумме в среднем имеется только один компонент входного вектора, т.е. распределение входного вектора определяет распределение u .

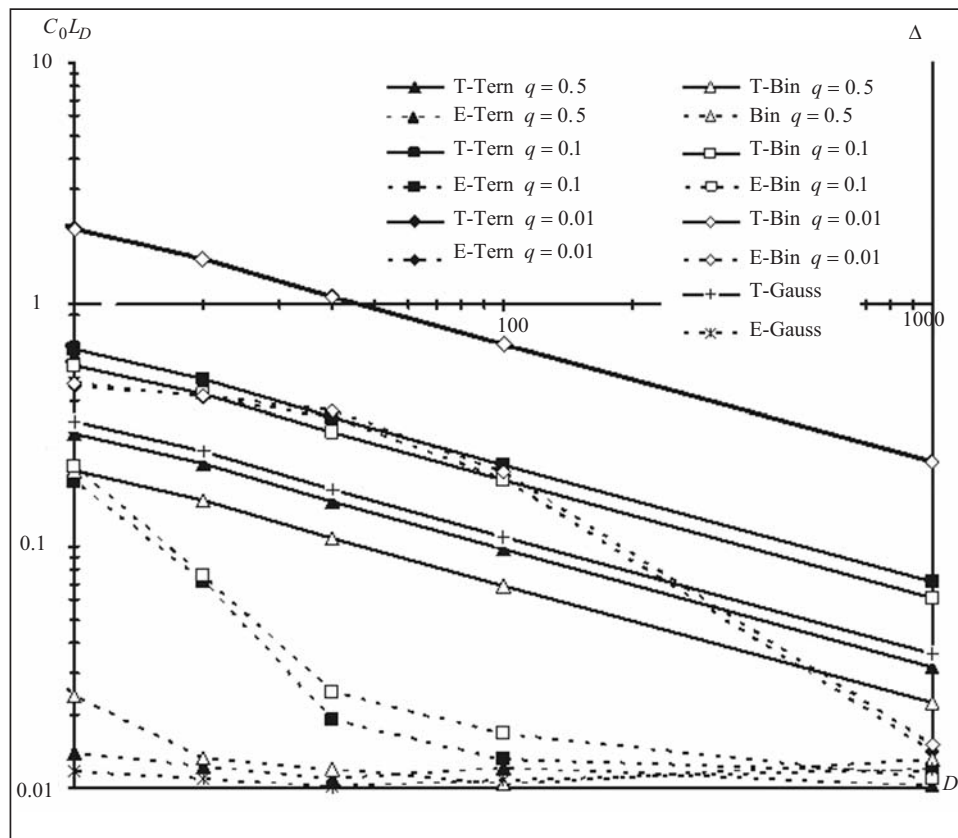


Рис. 3. Зависимости значений $C_0 L_D$ (9) (Т-) и оценок левой части Δ (13) (Е-) от числа слагаемых D

Согласно рис. 3 при исследованных параметрах экспериментальная оценка ошибки гауссовой аппроксимации эмпирического распределения не только не превосходит правой части НБЭ (9), но и является намного меньшей. Заметим также, что для гауссовой матрицы правая часть НБЭ уменьшается с ростом D , в то время как реальное распределение суммы гауссово для любого D . Это не удивительно, так как правая часть НБЭ рассчитана на худший случай.

Кроме того, как видно на рис. 3 и в соответствии с зависимостью для L_x (6) на рис. 2, правая часть НБЭ несколько (незначительно) меньше для бинарной матрицы по сравнению с тернарной. По экспериментальной оценке левой части НБЭ (13) бинарная матрица дает результаты не лучше тернарной. Это может являться следствием того, что центрирование реальных сумм u_i (для различных i) осуществляется вычитанием одной и той же величины м.о. $q \sum_{j=1}^D x_j$, а реальная часть единичных элементов в строках $\mathbf{R}$ может заметно отличаться от q , особенно для малых величин qD .

3.2. Ошибка оценки угла. Чтобы проверить, как размерность входных векторов D влияет на ошибку оценки угла между ними, проведен эксперимент. Результаты (табл. 1) приведены для одной пары входных векторов для каждого D . Выбирались пары с углом приблизительно 60° . Выходные бинарные векторы генерировались пороговым преобразованием результатов проецирования (1). Порог устанавливался по (2) в предположении гауссова стандартного распределения, чтобы обеспечить вероятность единичного компонента в выходных векторах $p = \{0.5, 0.1\}$ (т.е. использовались квантили $t_p = \{0.0, 1.282\}$ соответственно). Угол между входными векторами оценивался по оценкам вероятности совпадения единичных компонентов (3) (см. также [23]).

Таблица 1. Значения среднего квадрата ошибки определения угла (град²) между входными векторами (Е-) и расчетной дисперсии угла (Т-)

Параметры преобразования входных векторов	Значения ошибки определения угла (в град ²) для D						
	10	20	40	100	250	500	1000
T-Tern $q = 0.5, p = 0.5$	28.7	28.5	28.1	27.8	29.0	28.8	28.8
E-Tern $q = 0.5, p = 0.5$	28.1	29.3	28.7	27.2	31.6	28.9	29.2
T-Bin $q = 0.5, p = 0.5$	28.8	28.5	28.0	27.8	29.0	28.8	28.8
E-Bin $q = 0.5, p = 0.5$	29.0	28.4	28.4	27.5	28.8	30.4	29.4
T-Tern $q = 0.5, p = 0.1$	36.6	35.8	35.2	35.0	36.6	36.4	36.4
E-Tern $q = 0.5, p = 0.1$	39.7	38.6	36.8	32.9	37.4	35.5	37.6
T-Bin $q = 0.5, p = 0.1$	34.7	35.8	35.1	34.9	36.6	36.3	36.3
E-Bin $q = 0.5, p = 0.1$	113.9	37.4	35.6	37.5	37.7	38.6	35.6
T-Tern $q = 0.1, p = 0.5$	31.7	31.3	29.3	27.9	29.0	28.8	28.8
E-Tern $q = 0.1, p = 0.5$	7430.6	898.1	120.5	27.9	29.8	27.8	27.0
T-Bin $q = 0.1, p = 0.5$	31.8	31.4	28.8	28.2	29.2	28.9	28.8
E-Bin $q = 0.1, p = 0.5$	7103.3	970.2	61.3	36.5	31.8	27.1	27.9
T-Tern $q = 0.1, p = 0.1$	41.0	36.5	36.2	35.1	36.6	36.3	36.4
E-Tern $q = 0.1, p = 0.1$	385.5	42.7	59.6	36.2	35.1	37.9	36.1
T-Bin $q = 0.1, p = 0.1$	33.8	33.9	35.1	34.5	36.3	36.4	36.4
E-Bin $q = 0.1, p = 0.1$	487.9	233.0	38.3	46.2	41.3	34.0	35.4
T-Tern $q = 0.01, p = 0.5$	7.5	13.9	20.9	31.7	31.3	29.3	28.9
E-Tern $q = 0.01, p = 0.5$	47633.6	39165.4	29886.3	8993.8	688.5	46.6	27.4
T-Bin $q = 0.01, p = 0.5$	8.1	14.6	21.3	31.8	31.1	29.1	28.9
E-Bin $q = 0.01, p = 0.5$	46822.0	38201.8	29259.1	8408.0	539.3	35.2	26.7
T-Tern $q = 0.01, p = 0.1$	33.7	37.3	44.3	40.6	38.2	36.9	36.5
E-Tern $q = 0.01, p = 0.1$	526.2	77.8	1162.9	639.2	82.6	42.6	34.6
T-Bin $q = 0.01, p = 0.1$	33.9	33.9	40.2	36.9	36.4	36.8	36.7
E-Bin $q = 0.01, p = 0.1$	284.2	276.7	469.5	135.2	41.2	40.6	38.5
T-Gauss $p = 0.5$	28.8	28.6	28.1	27.9	29.0	28.823	28.8
E-Gauss $p = 0.5$	29.1	26.6	28.2	27.8	28.1	30.785	28.6
T-Gauss $p = 0.1$	36.3	35.9	35.3	35.0	35.8	36.3	36.4
E-Gauss $p = 0.1$	39.0	36.1	36.0	33.7	36.6	37.8	40.6

Размерность входных вещественных векторов составляла $D = \{10, 20, 40, 100, 250, 500, 1000\}$, а выходных бинарных векторов составляла $d = 1000$. В табл. 1 приведены экспериментальные значения среднего квадрата ошибки оценки угла (Е-) для 1000 пар выходных векторов, полученных для одной входной пары векторов по 1000 реализаций случайной матрицы, а также дисперсии значений углов для выбранных параметров (Т-), вычисленные по выражениям из [23].

Результаты, сильно отличающиеся для Т- и Е-, в табл. 1 даны жирным шрифтом. Сравнение с результатами рис. 3 показывает, что в целом соответствие Т- и Е- существует при малых экспериментальных отличиях эмпирических распределений от гауссовых. Для $D = 1000$ экспериментальная (Е-) и теоретическая (Т-) ошибки соответствуют для всех исследованных q и p . Для $D \geq 100$ соответствие наблюдается при $q = \{0.1, 0.5\}$ и $p = \{0.5, 0.1\}$, а для $q = 0.5$ — при $D \geq 20$ и $p = \{0.5, 0.1\}$. Заметим, что в практических приложениях ожидаемые значения D — тысячи и миллионы (по крайней мере сотни), а меньшие значения D исследовались здесь для иллюстрации особенностей сходимости эмпирических распределений к гауссову.

ЗАКЛЮЧЕНИЕ

Для преобразования входных вещественных векторов (в формате с плавающей запятой) в выходные бинарные векторы предложено использовать проецирование случайной матрицей с бинарными элементами $\{0, +1\}$ и пороговое преобразование с регулируемым порогом. Выходные бинарные векторы позволяют оценивать меры сходства–различия исходных векторов. Величину (косинуса) угла можно оценить, например, по величине скалярного произведения выходных бинарных векторов.

Аналитически и экспериментально исследована скорость сходимости распределения компонентов вещественного вектора — результата проецирования — к гауссову. Для бинарных и тернарных случайных матриц экспериментальные оценки отличия от гауссовости для всего диапазона исследованных параметров значительно меньше предельных аналитических значений, рассчитанных по правой части неравенства Бери–Эссеена. Экспериментальные и аналитические результаты для бинарных и тернарных матриц близки и свидетельствуют о том, что при размерности входного вектора $D = 1000$ для всех исследованных вероятностей ненулевых элементов в случайных матрицах $q = \{0.5, 0.1, 0.01\}$ отличие эмпирических распределений от гауссового становится близко к уровню ошибки вследствие оценки эмпирических распределений по конечной выборке.

Отметим, что хотя аналитически несколько лучшие результаты скорости сходимости распределения к гауссову получены для проецирования бинарными матрицами, в эксперименте несколько лучшие результаты наблюдались для тернарных матриц. Предположительно это может быть связано с использованием для центрирования компонентов выходного вектора, получаемых при проецировании бинарной случайной матрицы, теоретического значения их м.о. которое может отличаться от реального среднего значения.

Для изучения влияния оценок отличия распределения компонентов выходного вещественного вектора от гауссового на ошибку оценки угла между входными векторами по бинаризованным выходным векторам проведены экспериментальные исследования при одинаковых параметрах случайных матриц различного типа. Полученные результаты показывают, что для параметров, для которых незначительно отличие эмпирических распределений от гауссового, мало и отличие экспериментальной и аналитической ошибок в определении угла. При размерности входного вектора $D = 1000$ экспериментальная и аналитическая ошибки близки для всех исследованных параметров, что требует выполнения условия $Dq > 10$ в приложениях.

Очевидно, чем меньше q , тем больше возможность ускорения реализации проецирования векторов. При этом проецирование бинарной матрицей потенциально более эффективно, чем тернарной и гауссовой. Разработка технических решений для эффективной реализации проецирования (и оперирования бинарными векторами) является одним из направлений перспективных работ. Представляет также интерес дальнейшее сравнительное исследование оценок мер сходства–различия исходных векторов по бинарным векторам в целях повышения их точности.

СПИСОК ЛИТЕРАТУРЫ

1. Rachkovskij D.A., Slipchenko S.V. Similarity-based retrieval with structure-sensitive sparse binary distributed representations // Computational Intelligence. — 2012. — **28**, N 1. — P. 106–129.
2. Reznik A.M., Galinskaya A.A., Dekhtyarenko O.K., Nowicki D.W. Preprocessing of matrix QCM sensors data for the classification by means of neural network // Sensors and Actuators B. — 2005. — **106**. — P. 158–163.
3. Frolov A.A., Husek D., Polyakov P.Yu. Recurrent-neural-network-based boolean factor analysis and its application to word clustering // IEEE Transactions on Neural Networks. — 2009. — **20**, N 7. — P. 1073–1086.
4. Нейросетевые распределенные представления для интеллектуальных информационных технологий и моделирования мышления / В.И. Гриценко, Д.А. Рачковский, А.Д. Гольцев, В.В. Лукович, И.С. Мисуно, Е.Г. Ревунова, С.В. Слипенченко, А.М. Соколов, С.А. Талаев // Кибернетика и вычислит. техника. — 2013. — Вып. 173. — С. 7–24.

5. Slipchenko S.V., Rachkovskij D.A. Analogical mapping using similarity of binary distributed representations // Intern. J. Inform. Theories and Appl. — 2009. — **16**, N 3. — P. 269–290.
6. Мисуно И.С., Рачковский Д.А., Слипченко С.В. Векторные и распределенные представления, отражающие меру семантической связи слов // Математические машины и системы. — 2005. — № 3. — С. 50–67.
7. Sokolov A. LIMSI: learning semantic similarity by selecting random word subsets // Proc. of the Sixth Intern. Workshop on Semantic Evaluation (SEMEVAL'12). — Montreal (Canada): Association for Computational Linguistics, 2012. — P. 543–546.
8. Sokolov A., Riezler S. Task-driven greedy learning of feature hashing functions // Proceedings of the NIPS'13 Workshop "Big Learning: Advances in Algorithms and Data Management", Lake Tahoe (USA), 2013. — P. 1–5.
9. Kussul E.M., Rachkovskij D.A. Multilevel assembly neural architecture and processing of sequences // Neurocomputers and Attention: Vol. II. Connectionism and neurocomputers / A.V. Holden, V.I. Kryukov (Eds.). — Manchester; New York: Manchester Univ. Press, 1991. — P. 577–590.
10. Computing with large random patterns / P. Kanerva, G. Sjodin, J. Kristoferson, R. Karlsson, B. Levin, A. Holst, J. Karlgren, M. Sahlgren // Foundations of Real-World Intelligence. — Stanford (Calif.): CSLI Publ., 2001. — P. 251–311.
11. Rachkovskij D.A. Representation and processing of structures with binary sparse distributed codes // IEEE Transactions on Knowledge and Data Engineering. — 2001. — **13**, N 2. — P. 261–276.
12. Rachkovskij D.A., Slipchenko S.V., Kussul E.M., Baidyk T.N. Binding procedure for distributed binary data representations // Cybernetics and Systems Analysis. — 2005. — **41**, N 3. — P. 319–331.
13. Letichevsky A., Letychevsky A. Jr, Peschanenko V. Insertion modeling system // Lecture Notes in Computer Science. — 2011. — **7162**. — P. 262–274.
14. The properties of predicate transformer in VRS system / A. Letichevsky, A. Godlevsky, A. Letichevsky Jr, S. Potienko, V. Peschanenko // Cybernetics and Systems Analysis. — 2010. — **46**, N 4. — P. 521–532.
15. Gallant S.I., Okaywe T.W. Representing objects, relations, and sequences // Neural Computation — 2013. — **25**, N 8. — P. 2038–2078.
16. Rachkovskij D.A. Some approaches to analogical mapping with structure sensitive distributed representations // Journal of Experimental and Theoretical Artificial Intelligence — 2004. — **16**, N 3. — P. 125–145.
17. Emruli B., Gayler R.W., Sandin F. Analogical mapping and inference with binary spatter codes and sparse distributed memory // Intern. Joint Conf. on Neural Networks (IJCNN), 4–9 Aug 2013 (Dallas, TX). — IEEE, 2013. — P. 1–8.
18. The data fusion Grid infrastructure: Project objectives and achievements / N. Kussul, A. Shelestov, S. Skakun, O. Kravchenko, Y. Gripich, L. Hluchy, P. Kopp, E. Lupian // Computing and Informatics. — 2012. — **29**, N 2. — P. 319–334.
19. The wide area grid testbed for flood monitoring using earth observation data / N.N. Kussul, A.Y. Shelestov, S.V. Skakun, Guoqing Li, O.M. Kussul // IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. — 2012. — **5**, N 6. — P. 1746–1751.
20. Achlioptas D. Database-friendly random projections: Johnson–Lindenstrauss with binary coins // J. Comp. and System Sci. — 2003. — **66**, N 4. — P. 671–687.
21. Charikar M. Similarity estimation techniques from rounding algorithms // ACM Symposium on Theory of Computing. — Montreal (Canada): ACM, 2002. — V. 1. — P. 380–388.
22. Li P., Hastie T.J., Church K.W. Very sparse random projections // 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. — Philadelphia (USA): ACM Press, 2006. — P. 287–296.
23. Rachkovskij D.A., Misuno I.S., Slipchenko S.V. Randomized projective methods for construction of binary sparse vector representations // Cybernetics and Systems Analysis. — 2012. — **48**, N 1. — P. 146–156.
24. Rachkovskij D.A. Vector data transformation using random binary matrices // Cybernetics and Systems Analysis. — 2014. — **50**, N 6. — P. 960–968.
25. Revunova E.G., Rachkovskij D.A. Using randomized algorithms for solving discrete ill-posed problems // Information Theories and Applications. — 2009. — **16**, N 2. — P. 176–192.
26. Rachkovskij D.A., Revunova E.G. Randomized method for solving discrete ill-posed problems // Cybernetics and Systems Analysis. — 2012. — **48**, N 4. — P. 621–635.
27. Нейрокомпьютеры и интеллектуальные роботы / Н.М. Амосов, Т.Н. Байдык, А.Д. Гольцев, А.М. Касаткин, Л.М. Касаткина, Э.М. Куусуль, Д.А. Рачковский — Киев: Наук. думка, 1991. — 269 с.

28. Rachkovskij D.A., Kussul E.M., Baidyk T.N. Building a world model with structure-sensitive sparse binary distributed representations // *Biologically Inspired Cognitive Architectures*. — 2013. — **3**. — P. 64–86.
29. Омельченко Р.С. Программа проверки орфографии (spellchecker) на основе распределенных представлений // *Проблемы программирования*. — 2013. — № 4. — С. 35–42.
30. Frolov A., Kartashov A., Goltsev A., Folk R. Quality and efficiency of retrieval for Willshaw-like autoassociative networks. I. Correction // *Network: Computation in Neural Systems*. — 1995. — **6**, N 4. — P. 513–534.
31. Frolov A., Kartashov A., Goltsev A., Folk R. Quality and efficiency of retrieval for Willshaw-like autoassociative networks. II. Recognition // *Network: Computation in Neural Systems*. — 1995. — **6**, N 4. — P. 535–549.
32. Frolov A.A., Husek D., Muraviev I.P. Informational capacity and recall quality in sparsely encoded Hopfield-like neural network: Analytical approaches and computer simulation // *Neural Networks*. — 1997. — **10**, N 5. — P. 845–855.
33. Frolov A.A., Rachkovskij D.A., Husek D. On information characteristics of Willshaw-like auto-associative memory // *Neural Network World*. — 2002. — **12**, N 2. — P. 141–158.
34. Frolov A. A., Husek D., Rachkovskij D. A. Time of searching for similar binary vectors in associative memory // *Cybernetics and Systems Analysis*. — 2006. — **42**, N 5. — P. 615–623.
35. Nowicki D.W., Dekhtyarenko O.K. Averaging on Riemannian manifolds and unsupervised learning using neural associative memory // *Proceedings of 13th European Symposium on Artificial Neural Networks (ESANN 2005) (April 27–29)*. — Bruges (Belgium) 2005. — P. 181–189.
36. Nowicki D., Verga P., Siegelmann H. Modeling reconsolidation in kernel associative memory // *PloS one*. — 2013. — **8**, N 8. — e68189.doi:10.1371/journal.pone.0068189.
37. Різник О.М., Дзюба Д.О. Динамічна асоціативна пам'ять на основі відкритої рекурентної нейромережі // *Математичні машини і системи*. — 2010. — № 2. — С. 50–60.
38. Central limit theorem. — http://en.wikipedia.org/wiki/Central_sub_limit_sub_theorem, accessed 2013.12.25.
39. Berry A.C. The accuracy of the Gaussian approximation to the sum of independent variates // *Transactions of the American Mathematical Society*. — 1941. — **49**. — P. 122–136.
40. Esseen C.G. On the Liapunov limit of error in the theory of probability // *Arkiv für Matematik, Astronomi och Fysik*. — 1942. — **28A**, N 9. — P. 1–19.
41. Esseen C.G. A moment inequality with an application to the central limit theorem // *Skandinavisk Aktuarietidskrift*. — 1956. — **39**. — 160–170.
42. Korolev V., Shevtsova I. An improvement of the Berry–Esseen inequality with applications to Poisson and mixed Poisson random sums // *Scandinavian Actuarial Journ.* — 2012. — N 2. — P. 81–105.
43. Shevtsova I.G. On the absolute constants in the Berry–Esseen-type inequalities // *Doklady Mathematics*. — 2014. — **89**, N 3. — P. 378–381.
44. Тюрин И.С. Уточнение остаточного члена в теореме Ляпунова // *Теория вероятностей и ее применение*. — 2011. — **56**, № 4. — С. 808–811.
45. Nagaev S.V., Chebotarev V.I. On the bound of proximity of the binomial distribution to the normal one // *Doklady Mathematics*. — 2011. — **83**, N 1. — P. 19–21.
46. Walck C. Hand-book on statistical distributions for experimentalists // *Internal Report SUF-PFY/96-01 (last modification 10 Sept. 2007)*. — Fysikum University of Stockholm: Particle Physics Group, 2007. — 190 p.
47. Bentkus V. A Lyapunov type bound in $\mathbb{R}^d$ // *Theory of Probability & Its Applications*. — 2005. — **49**. — P. 311–323.
48. Bhattacharya R.N., Holmes S. An exposition of Gotze's estimation of the rate of convergence in the multivariate central limit theorem // *Eprint arXiv:1003.4254*. — 2010.
49. Chen L.H.Y., Fang X. Multivariate normal approximation by Stein's method: the concentration inequality approach // *Eprint arXiv:1111.4073*. — 2011.

Поступила 24.12.2013