УДК 621.396

МАКСИМАЛЬНО ПРАВДОПОДОБНАЯ ОЦЕНКА ДИСПЕРСИОННО-КОВАРИАЦИОННОЙ МАТРИЦЫ

Полянский И.С., Патронов Д.Ю.

Академия ФСО России, Орел

В работе сформировано аналитическое выражение, определяющее максимально правдоподобную оценку дисперсионно-ковариационной матрицы наблюдения вектора случайных величин, распределенных по многомерному нормальному закону. Решение основано на получении выражения, определяющего точку экстремума сформированной на основе распределения Уишарта функции правдоподобия. Полученная оценка дисперсионно-ковариационной матрицы позволяет повысить эффективность оценки с точки зрения снижения разброса дисперсии ошибки. Что по существу приводит к увеличению точности оценки дисперсионно-ковариационной матрицы (на величину порядка нескольких тысяч раз) уже на выборках малого объема независимо от разрядности дисперсионно-ковариационной матрицы (размера вектора случайных величин). Причем вычислительные затраты сформированной оценки не хуже существующих. Представлены результаты экспериментальной проверки сформированного аналитического выражения максимально правдоподобной оценки дисперсионно-ковариационной матрицы.

Ключевые слова: максимально правдоподобная оценка, распределение Уишарта, дисперсионно-ковариационная матрица.

MAXIMUM REASONABLE ESTIMATE VARIANCE-COVARIANCE MATRICES

Polyanskiy I.S., Patronov D.Y.

AcademyFSO of Russia, Orel

The work formed the analytic expression that defines the maximum likelihood estimate of variance-covariance matrix of the observation vector of random variables distributed according to multivariate normal distribution. The solution is based on obtaining an expression that identifies the extremum point formed on the basis of the distribution of Wishart likelihood function. The obtained estimate variance-covariance matrix to enhance the evaluation, in terms of reducing the spread of error variance. That substantially increases the accuracy of the estimate variance-covariance matrix (on the order of several thousand) already on the small sample size, regardless of the bit variance-covariance matrix (size of the vector of random variables). Moreover, the computational cost estimates formed better than the existing ones. The results of the experimental verification of the analytical expressions generated maximum likelihood estimation variance-covariance matrix.

Key words: maximum likelihood estimation, Wishart distribution, variance-covariance matrix.

Введение

Подавляющее большинство задач прикладной статистики сводится к необходимости оценки дисперсионно-ковариационной матрицы (ДКМ) наблюдения вектора случайных величин $\vec{X} = \left(X_1,\ldots,X_N\right)$, распределенных по многомерному нормальному закону. Из [1; 7] и многих других источников известно, что существующие аналитические методы определения элементов ДКМ на практике сводятся к оценке [1]:

$$\Sigma_{i,j} = \frac{\sum\limits_{k=1}^{K}\left[\left(X_i\right)_k - \tilde{X}_i\right]\cdot\left[\left(X_j\right)_k - \tilde{X}_j\right]}{K}; \; i,j=\overline{1,N}, \qquad (1)$$

где K – объем выборки элементов вектора случайных величин $\vec{X}$; $\tilde{X}_i$ и $\tilde{X}_j$ – выборочные средние i-го и j-го элементов вектора случайных величин $\vec{X}$, значение которых определяются известным отношением [7]:

$$\tilde{X}_i = \frac{\sum_{k=1}^{K} (X_i)_k}{K}. \tag{2}$$

Однако правило (1) не определяет максимально правдоподобную оценку ДКМ $\hat{\Sigma}$. Другие решения, к примеру [3–5] и др., нахождения ДКМ по известной статистике (конечному набору наблюдений элементов вектора $\vec{X}$) направлены на численное решение градиентными релаксационными методами условных оптимизационных задач по критерию максимума правдоподобия, что в свою очередь не позволяет достигнуть точного определения оптимума целевой функции правдоподобия и в виду итеративного подхода значительно увеличивает вычислительную сложность процесса оценки.

Цель статьи заключается в получении аналитического выражения максимально правдоподобной оценки ДКМ для произвольного размера вектора случайных величин $\vec{X}$ $(N \geq 2)$, распределенных по многомерному нормальному закону.

Теорема о максимально правдоподобной оценке ДКМ

Теорема. *Пусть* $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_k, \dots, \vec{X}_K$ *есть конечная последовательность векторов размерности N случайных величин, определенных в области действительных чисел* **R**, *и распределенных по многомерному нормальному закону. Тогда максимально правдоподобная оценка ДКМ* $\hat{\Sigma}$ *полученной выборки объема K определяется отношением:*

$$\hat{\Sigma} = \frac{2}{K \cdot (K-1)} \cdot \sum_{k=1}^{K} \mathrm{A}(k), \tag{3}$$

где $\mathrm{A}(k)$ *– наблюдаемая ДКМ на выборке вектора случайных величин размера k, определяемая отношением:*

$$\mathrm{A}(k) = (K+1-k) \cdot \left[\vec{X}_k - \tilde{\vec{X}}\right] \cdot \left[\vec{X}_k - \tilde{\vec{X}}\right]^T. \tag{4}$$

Доказательство. Поскольку элементы вектора $\vec{X}$ определены в области действительных чисел **R** и распределены по многомерному нормальному закону, то функция плотности распределения наблюдаемой ДКМ A, формируемой конечной последовательности наблюдаемых векторов $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_k, \dots, \vec{X}_K$, согласно [8] определяется выражением (функция распределения плотности вероятности Уишарта):

$$W(k,\Sigma)=\frac{|\mathrm{A}(k)|^{(k-N)/2}\cdot e^{-\frac{1}{2}\cdot tr\left(\mathrm{A}(k)\cdot\Sigma^{-1}\right)}}{2^{(k-1)N/2}\cdot\pi^{N(N-1)/4}\cdot|\Sigma|^{(k-1)/2}\cdot\prod_{i=1}^{N}\Gamma\left(\frac{k-i}{2}\right)}, \tag{5}$$

где $tr(\cdot)$ – оператор, определяющий величину следа матрицы; $\mathrm{A}(k)$ – наблюдаемая на выборке размера k ДКМ и определяемая известным отношением [1]:

$$\mathrm{A}(k)=\sum_{i=1}^{k}\left[\vec{X}_i-\tilde{\vec{X}}\right]\cdot\left[\vec{X}_i-\tilde{\vec{X}}\right]^T. \tag{6}$$

В выражении (6) $\tilde{\vec{X}}$ – вектор выборочных средних элементов вектора случайных величин $\vec{X}$.

Согласно представленному равенству (5) составим функцию максимально правдоподобной оценки ДКМ $\hat{\Sigma}$ по известным конечным наблюдениям векторов случайных величин $\vec{X}_1,\vec{X}_2,...,\vec{X}_k,...,\vec{X}_K$:

$$f(\Sigma)=\sum_{k=1}^{K}\ln\left[W(k,\Sigma)\right]\to\max_{\Sigma}. \tag{7}$$

Для нахождения точек экстремума определим аналитический дифференциал скалярной функции (7) от матриц по матричному аргументу, в соответствии с правилами, представленными в [2]:

$$\frac{\partial f(\Sigma)}{\partial\Sigma}=\frac{1}{2}\cdot\sum_{k=1}^{K}\left[\Sigma^{-1}\cdot\mathrm{A}(k)\cdot\Sigma^{-1}-(k-1)\cdot\Sigma^{-1}\right]. \tag{8}$$

Приравняв полученную систему уравнений к нулю, определим корни, удовлетворяющие условию:

$$\Sigma^{-1}\cdot\sum_{k=1}^{K}\left[\mathrm{A}(k)\right]\cdot\Sigma^{-1}-\sum_{k=1}^{K}\left[(k-1)\right]\cdot\Sigma^{-1}=0. \tag{9}$$

Заметим, что $\sum_{k=1}^{K}\left[(k-1)\right]=\frac{K\cdot(K-1)}{2}$, тогда получим равенство (9) в виде:

$$\Sigma^{-1}\cdot\sum_{k=1}^{K}\left[\mathrm{A}(k)\right]\cdot\Sigma^{-1}-\frac{K\cdot(K-1)}{2}\cdot\Sigma^{-1}=0. \tag{10}$$

Помножив обе части равенства (10) на Σ, получим:

$$\sum_{k=1}^{K}\left[\mathrm{A}(k)\right]\cdot\Sigma^{-1}-\frac{K\cdot(K-1)}{2}\cdot\mathbf{E}=0, \tag{11}$$

где $\mathbf{E}$ – единичная матрица размерности $N \times N$.

Выразив Σ в выражении (11), получим:

$$\Sigma = \left(\left(\sum_{k=1}^{K} \left[\mathrm{A}(k) \right] \right)^{-1} \cdot \frac{K \cdot (K-1)}{2} \right)^{-1}. \tag{12}$$

Раскрыв в правой части равенства (12) скобки, получим:

$$\Sigma = \frac{2}{K \cdot (K-1)} \cdot \sum_{k=1}^{K} \left[\mathrm{A}(k) \right]. \tag{13}$$

Заметим, что для уменьшения вычислительной сложности определения ДКМ выражением (13) наблюдаемую матрицу $\mathrm{A}(k)$ на выборки объемом *k*, рассчитываемую в соответствии с выражением (6), можно справедливо представить в виде:

$$\mathrm{A}(k) = (K+1-k) \cdot \left[\vec{X}_k - \tilde{\tilde{\vec{X}}} \right] \cdot \left[\vec{X}_k - \tilde{\tilde{\vec{X}}} \right]^T. \tag{14}$$

Очевидно, что ввиду линейности системы уравнений (9) решение (13) является для целевой функции (7) единственной точкой экстремума, а из унимодальности функции (5) следует, что найденной решение (13) соответствует точке, удовлетворяющей правилу (7), и максимально правдоподобная оценка ДКМ справедливо запишется в виде (3). ***Теорема доказана.***

Аналогичным образом определяется максимально правдоподобная оценка ДКМ $\hat{\dot{\Sigma}}$, формируемой $\vec{X}_1, \vec{X}_2, ..., \vec{X}_k, ..., \vec{X}_K$ конечной последовательностью векторов размерности *N* случайных величин, определенных в области комплексных чисел **C** и распределенных по многомерному нормальному закону. В данном случае ДКМ будет рассчитываться выражением:

$$\hat{\dot{\Sigma}} = \frac{2 \cdot \sum_{k=1}^{K} \left[(K+1-k) \cdot \left[\dot{\vec{X}}_k - \tilde{\tilde{\dot{\vec{X}}}} \right] \cdot \left[\dot{\vec{X}}_k - \tilde{\tilde{\dot{\vec{X}}}} \right]^* \right]}{K \cdot (K-1)}, \tag{15}$$

где $[\cdot]^*$ – оператор комплексного сопряжения.

Оценка полученных результатов

Для определения качества оценки по полученному выражению (3) оценки ДКМ определим зависимость нижней границы дисперсии оценки $\overline{\chi}$ от объема выборки *K* и размерности *N* вектора наблюдений $\vec{X}$ из многомерного неравенства Крамера-Рао [7]:

$$\overline{\chi} = tr\left[\mathbf{I}(\Sigma)^{-1}\right], \tag{16}$$

где $tr\left[\mathbf{I}(\Sigma)^{-1}\right]$ – след обратной информационной матрицы Фишера, определяемой соотношением [7]:

$$\mathbf{I}(\Sigma) = E\left[\frac{\partial^2 f(\Sigma)}{\partial \Sigma^2}\right], \tag{17}$$

где $E[\cdot]$ – оператор математического ожидания.

Для определения элементов информационной матрицы Фишера $\mathbf{I}(\Sigma)$ представим её в виде блочной матрицы:

$$\mathbf{I}(\Sigma) = \begin{bmatrix} \mathbf{I}_{1,1}(\Sigma) & \ldots & \mathbf{I}_{1,N}(\Sigma) \\ \ldots & \ldots & \ldots \\ \mathbf{I}_{N,1}(\Sigma) & \ldots & \mathbf{I}_{N,N}(\Sigma) \end{bmatrix}. \tag{18}$$

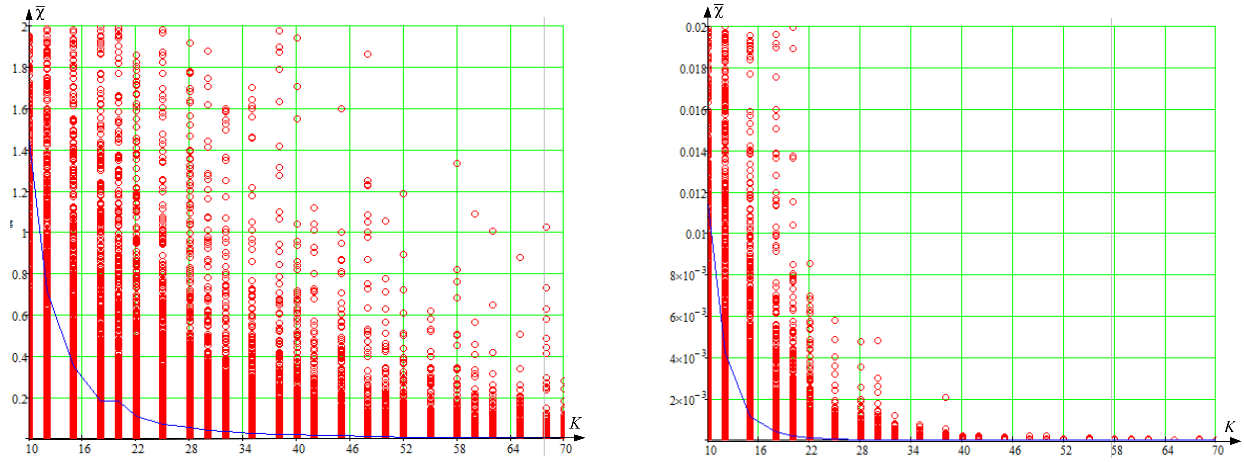
Затем, воспользовавшись правилами дифференцирования матричной функции по матричному аргументу, сформированными в [6], определим *i*-е, *j*-е элементы (подматрицы) блочной матрицы (17):

$$\mathbf{I}_{i,j}(\Sigma) = \frac{1}{2} \cdot \sum_{k=1}^{K} \begin{bmatrix} -\left(\Sigma^{-1} \cdot dF(i,j) \cdot \Sigma^{-1} \cdot \mathrm{A}(k) \cdot \Sigma^{-1}\right)^T - \\ -\left(\Sigma^{-1} \cdot \mathrm{A}(k) \cdot \Sigma^{-1} \cdot dF(i,j) \cdot \Sigma^{-1}\right)^T \end{bmatrix}, \tag{19}$$

где $dF(i,j)$ – матричная функция размерности $N \times N$, *n*-е, *m*-е $\left(n, m = \overline{1, N}\right)$ элементы которой определяются согласно выражению:

$$dF_{n,m}(i,j) = \begin{cases} 1, \text{ если } (n = i) \wedge (m = j); \\ 0, \text{ если } (n \neq i) \vee (m \neq j). \end{cases} \tag{20}$$

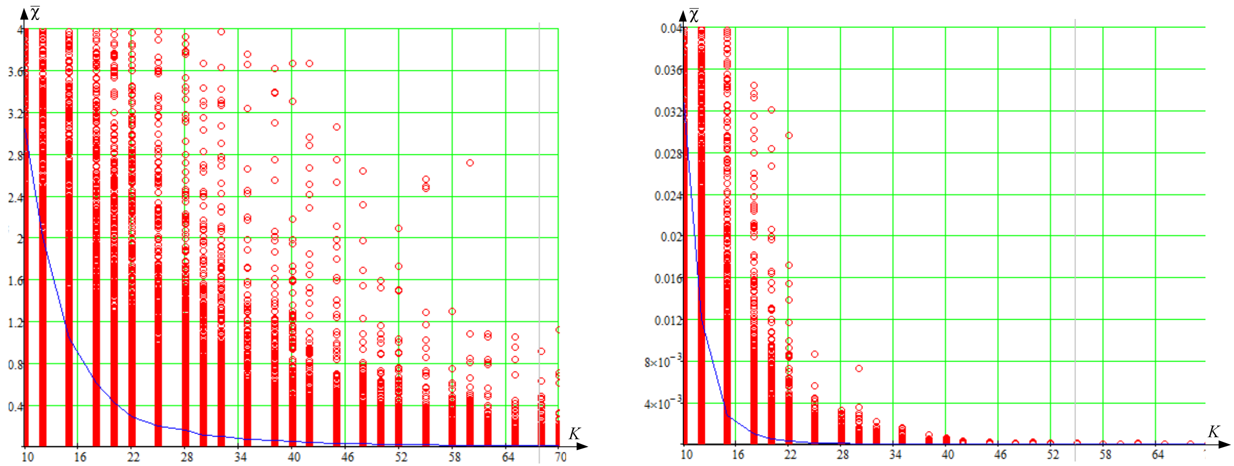
На рисунках 1–4 представлена зависимость дисперсии оценки $\overline{\chi}$ ДКМ, произведенной по известному выражению (1) и полученному – (3), от объема выборки K для вектора наблюдений $\vec{X}$, распределенного по многомерному нормальному закону, размерности 2, 4, 6, 8 соответственно.



а *б*

Рисунок 1 – Зависимость дисперсии оценки $\overline{\chi}$ от объема выборки K для размера вектора наблюдений, равного 2 (N = 2)

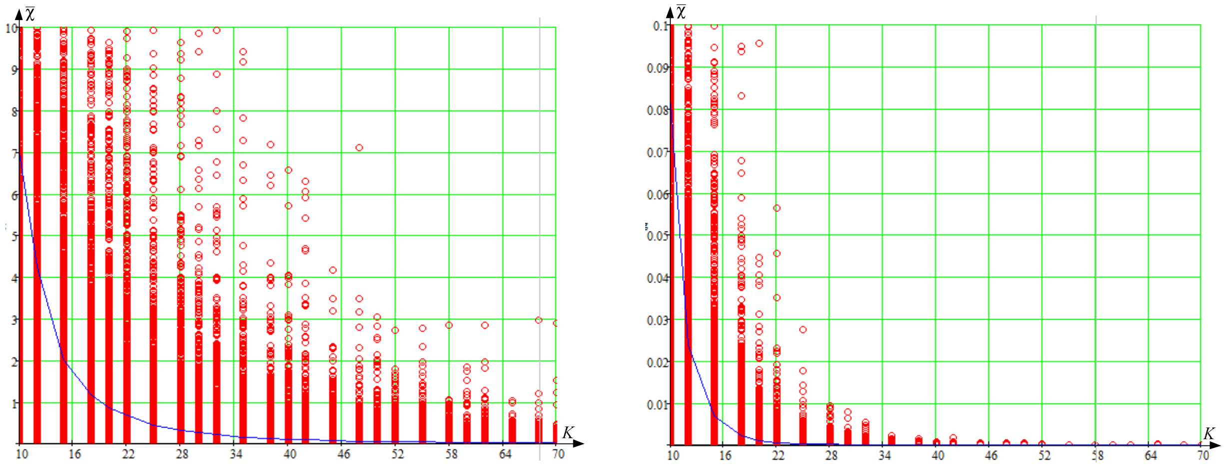
(*а* – оценка ДКМ выражением (1); *б* – оценка ДКМ выражением (3)).



а *б*

Рисунок 2 – Зависимость дисперсии оценки $\overline{\chi}$ от объема выборки K для размера вектора наблюдений, равного 4 (N = 4)

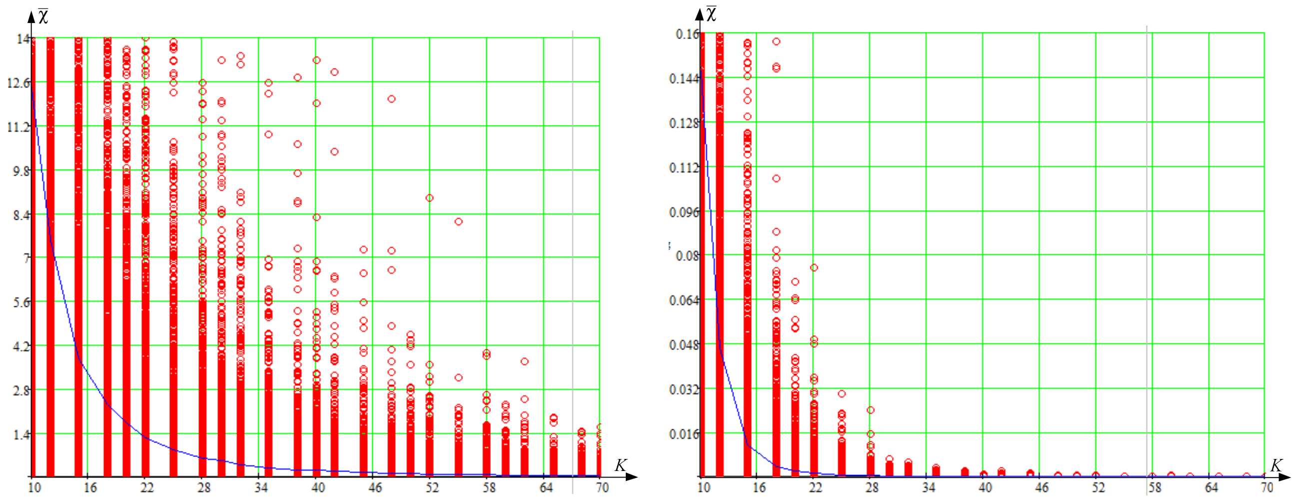
(*а* – оценка ДКМ выражением (1); *б* – оценка ДКМ выражением (3)).



а *б*

Рисунок 3 – Зависимость дисперсии оценки $\overline{\chi}$ от объема выборки K для размера вектора наблюдений, равного 6 (N = 6)

(*а* – оценка ДКМ выражением (1); *б* – оценка ДКМ выражением (3)).



а *б*

Рисунок 4 – Зависимость дисперсии оценки $\overline{\chi}$ от объема выборки K для размера вектора наблюдений, равного 8 (N = 8)

(*а* – оценка ДКМ выражением (1); *б* – оценка ДКМ выражением (3)).

На рисунке 5 представлены графики зависимости отношения дисперсии оценки ДКМ выражением (3) к дисперсии оценки ДКМ выражением (1) от объема выборки K, для N равном 2, 4, 6.

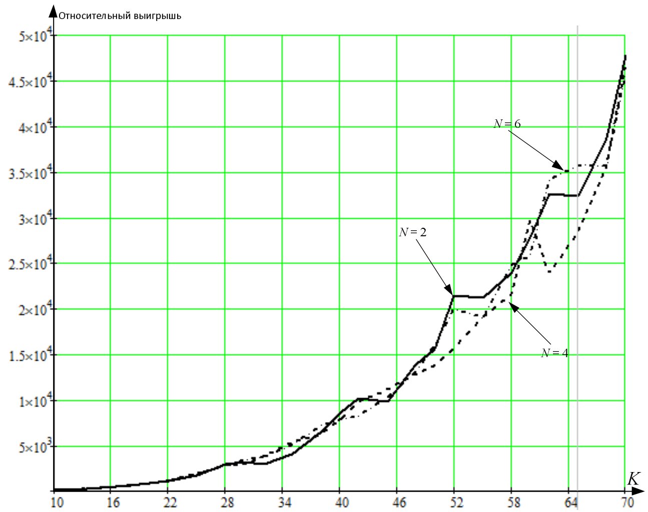


Рисунок 5 – Зависимость отношения дисперсии оценки ДКМ выражением (3) к дисперсии оценки ДКМ выражением (1) от объема выборки K, для N = 2, 4, 6.

Выводы

Из полученных результатов следует, что наряду с очевидным улучшением точности оценки ДКМ выражениями (1) и (3) с ростом объема выборки вектора случайных величин предложенное выражение (3) максимально правдоподобной оценки ДКМ приводит к существенному увеличению точности оценки ДКМ (на величину порядка нескольких тысяч раз) уже на выборках малого объема независимо от разрядности ДКМ (размера вектора случайных величин). Такая ситуация обуславливает очевидную предпочтительность оценки ДКМ вектора случайных величин $\vec{X}$ произвольной размерности $(N \geq 2)$, распределенных по многомерному нормальному закону, полученным аналитическим выражением (3) по сравнению с известным и применяемым на практике отношением (1).

Список литературы

1. Андерсон Т. Введение в многомерный статистический анализ / пер. с англ. Ю.Ф. Кичатова, Е.С. Кочеткова, Н.С. Райбмана ; под ред. Б.В. Гнеденко. – М. : Гос. изд-во физико-математической литературы, 1963. – 500 с.
2. Амосов А.А. Скалярно-матричное дифференцирование и его применение к конструктивным задачам теории связи / А.А. Амосов, В.В. Колпаков // Проблемы передачи информации. – 1972. – С. 3–15.
3. Батенков А.А. Оценка корреляционной матрицы сигналов на входе приемной фазируемой антенной решетки на основе многомерного распределения Уишарта / А.А. Батенков, К.А. Батенков, В.А. Пискун // Телекоммуникации. – 2012. – № 8. – С. 40–43.

4. Берг Дж. П. Оценивание ковариационных матриц с заданной структурой / Д.Г. Люнбергер, Д.Д. Венгер // ТИИЭР. – 1982. – Т. 70. – № 9. – С. 63–76.

5. Леховецкий Д.И. Ленточно-диагональная регуляризация МП оценок корреляционных матриц гауссовских помех в алгоритмах адаптации антенных решеток / Д.И. Леховецкий [и др.] // Сборник материалов докладов Всероссийской конференции «Радиолокация и радиосвязь». – М. : ИРЭ РАН, 2010. – С. 336–358.

6. Магнус Я.Р. Матричное дифференциальное исчисление с приложениями к статистике и экономике / Я.Р. Магнус, Х. Нейдеккер ; пер. с англ. ; под ред. С.А. Айвозяна. – М. : ФИЗМАТЛИТ, 2002. – 496 с.

7. Рао С.Р. Линейные статистические методы и их применения / пер. с англ. А.М. Кагана, В.М. Калинина, К.П. Латышева ; под ред. академика Ю.В. Линника. – М. : Наука, 1968. – 548 с.

8. Wishart John. The generalized product moment distribution in samples from a normal multivariate population / Biometrika 20A. – 1928. – P. 32–52.

Рецензенты:

Архипов Н.С., доктор технических наук, доцент, сотрудник Академии ФСО России, г. Орел.

Сычев К.И., доктор технических наук, доцент, сотрудник Академии ФСО России, г. Орел.