

DOI: 10.18721/JPM.11307

УДК 519.24

## ЗАМЕЧАНИЯ К ИСПОЛЬЗОВАНИЮ ГЛАВНЫХ КОМПОНЕНТ В МАТЕМАТИЧЕСКОМ МОДЕЛИРОВАНИИ

**Ю.А. Пичугин**

Санкт-Петербургский государственный университет аэрокосмического приборостроения,  
Санкт-Петербург, Российская Федерация

В статье рассматриваются вопросы, связанные с использованием метода главных компонент (ГК) в математическом моделировании: априорная оценка дисперсии погрешности регрессии на ГК для случаев большой и малой выборки; оценки минимального риска; оценка структурного подобия; проблема пропущенных данных и прогноз нестационарных временных рядов. Во всех случаях автор предлагает свои решения рассматриваемых вопросов или исправляет ранее допущенные неточности.

**Ключевые слова:** главная компонента, априорная оценка, дисперсия погрешности регрессии, малый объем выборки

**Ссылка при цитировании:** Пичугин Ю.А. Замечания к использованию главных компонент в математическом моделировании // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2018. Т. 11. № 3. С. 74–89. DOI: 10.18721/JPM.11307

## NOTES ON USING THE PRINCIPAL COMPONENTS IN THE MATHEMATICAL SIMULATION

**Yu.A. Pichugin**

Saint-Petersburg State University of Aerospace Instrumentation,  
St. Petersburg, Russian Federation

The paper discusses the issues related to the use of principal components analysis (PCA) in mathematical simulation. The paper significantly expands the range of the solved problems using PCA. In particular, the solutions of the following three tasks are given: (i) structural similarity and homogeneity estimation for random Gaussian vectors; (ii) recovery of missing data; (iii) the forecast of non-stationary time series based on the caterpillar method, which is a generalization of PCA for non-stationary time series. To solve the problems, to restore missing data and to predict the data, the author offers an unbiased estimation of the variance of the error of the regression on the PCs base for the cases of large and small samples. All the main statements are formulated in the form of theorems proved by the author.

**Key words:** principal component, variance of the regression error, small sample volume

**Citation:** Yu.A. Pichugin, Notes on using the principal components in the mathematical simulation, St. Petersburg Polytechnical State University Journal. Physics and Mathematics. 11 (3) (2018) 74–89. DOI: 10.18721/JPM.11307

## Введение

Анализ, или метод главных компонент (далее сокращения для главных компонент — PCs, а для метода — PCA (Principal Component Analysis)) представляет собой хорошо известный аппарат математической статистики. Этот метод был предложен К. Пирсоном в 1901 году [1], и суть PCA состоит в следующем.

Если должным образом повернуть (посредством ортогонального преобразования) систему координат  $n$ -мерного пространства так, чтобы оси координат совпадали с главными осями эллипсоида рассеяния, то компоненты нормально распределенного  $n$ -мерного центрированного вектора будут некоррелированными и, в силу нормального закона распределения, независимыми.

В алгебраическом смысле это есть не что иное, как приведение ковариационной матрицы к диагональному виду путем ортогонального преобразования, а квадратичной формы в экспоненте функции плотности многомерного нормального распределения — к каноническому виду. Хорошо известное преобразование Карунена — Лоева [2, 3] есть, по сути, именно это преобразование координат. Переход же к независимым переменным PCs позволяет, как правило, существенно сократить размерность исследуемой задачи с минимальной потерей информации.

В связи с этим в литературе PCs нередко выводятся как решение оптимизационной задачи, хотя все их оптимальные свойства достаточно хорошо видны из самого спектра ковариационной матрицы (спектр показывает, какова доля отбрасываемой дисперсии, см. далее).

В анализе временных рядов PCA известен как анализ сингулярного спектра (SSA — Singular Spectrum Analysis), где посредством данного метода решается проблема избыточности классического спектрального анализа [4 — 6]. Особенность SSA состоит в том, что размерность вектора в этом случае равна  $N$ , а размерность матрицы взаимных ковариаций равна  $N \times N$  ( $N$  — длина исследуемого временного ряда). При этом элементы ковариационной ма-

трицы вычисляются особым способом, когда делитель, независимо от величины сдвига, а соответственно и от числа слагаемых, равен  $N$ . Такие оценки, очевидно, относятся к классу смещенных оценок, но именно они не приводят в спектральном анализе к искажению (в сторону завышения, как отмечают Г. Дженкинс и Д. Ваттс [7]) длины волны. Проблема большой размерности ковариационной матрицы в реализации SSA, как показано в работе [8], легко решается применением итераций фон Мизеса, так как все строки ковариационной матрицы временного ряда могут быть получены из первой строки посредством сдвига, дублирования и перестановки элементов. Собственные значения и собственные векторы ковариационной матрицы получаются последовательностью простых итераций без вращения матрицы размерности  $N \times N$ .

Альтернативой SSA является «метод гусеницы» [9], а также метод, предлагаемый нами далее в настоящей работе (см. разделы «Проблема относительно малой выборки» и «Прогноз нестационарных временных рядов»), где схема прогноза строится на основе метода гусеницы.

Следует отметить, что существует множество методов, близких к PCA, например, в методе независимых компонент (ICA) последние могут подчиняться не только распределению Гаусса, но распределениям Стьюдента, Коши, Дирихле. Отметим, что метод независимых компонент известен также как анализ указанных компонент (ICA — Independent Component Analysis).

Обобщением PCA является метод главных кривых и многообразий. В последнее время PCA широко используется для визуализации и графического представления многомерных данных (рассматривается проекция выборки на плоскость первых двух главных осей [10, 11]). При этом требование нормального распределения исходных данных не возникает.

Здесь мы имеем большое разнообразие достаточно близких по сути методов, таких как многомерное шкалирование, нелинейный маппинг, поиск наилучшей проекции, а также методы нейросетевых задач, такие как метод «узкого горла», самоорганизую-

щиеся карты Кохонена и т. п. Следует также отметить, что графическое представление многомерных данных проекцией на плоскость первых двух главных осей РСs позволяет получить достаточно хорошее начальное приближение разделения выборки в решении задачи классификации в работе [12].

Цель настоящей работы – расширение спектра задач, решаемых на основе метода главных компонент.

В связи с указанной целью в статье рассматриваются задачи анализа структурного подобия, восстановления пропущенных данных, а также прогностическая задача нестационарных рядов. При этом уточняются детали метода главных компонент, непосредственно связанные с задачами восстановления пропусков и прогноза. Вопросы же снижения размерности и визуализации многомерных данных относятся в данном исследовании к второстепенным.

### Краткое описание математического аппарата PCA

Предполагается, что вектор  $\mathbf{y}$  имеет размерность  $m$  ( $\dim \mathbf{y} = m$ ) и подчиняется многомерному нормальному распределению, т. е.  $\mathbf{y} \sim N(\boldsymbol{\theta}_y, \mathbf{V}_y)$ .

Пусть  $\mathbf{P}$  – ортогональная матрица, такая, что

$$\mathbf{P}^T \mathbf{V}_y \mathbf{P} = \Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_m)$$

$$\text{и } \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m,$$

где  $T$  – символ (оператор) транспонирования.

Напомним, что столбцы матрицы  $\mathbf{P}$  есть собственные векторы матрицы  $\mathbf{V}_y$ , а совокупность собственных чисел  $\{\lambda_1, \lambda_2, \dots, \lambda_m\}$  называется спектром этой матрицы. В математической статистике столбцы матрицы  $\mathbf{P}$  называются базисом главных компонент, а главными компонентами называются компоненты вектора  $\mathbf{P}^T(\mathbf{y} - \boldsymbol{\theta}_y)$ . В приложениях параметры распределения  $N(\boldsymbol{\theta}_y, \mathbf{V}_y)$ , как правило, неизвестны. При наличии выборки  $\{\mathbf{y}_j, j = 1, 2, \dots, n\}$  мы можем вычислить несмещенные оценки неизвестных параметров:

$$\hat{\boldsymbol{\theta}}_y = \frac{1}{n} \sum_{j=1}^n \mathbf{y}_j, \tag{1}$$

$$\hat{\mathbf{V}}_y = \frac{1}{n-1} \sum_{j=1}^n (\mathbf{y}_j - \hat{\boldsymbol{\theta}}_y)(\mathbf{y}_j - \hat{\boldsymbol{\theta}}_y)^T.$$

В этом случае в качестве  $\mathbf{P}$  берем ортогональную матрицу  $\hat{\mathbf{P}}$ , которая приводит к диагональному виду оценку  $\hat{\mathbf{V}}_y$ , т. е.

$$\hat{\mathbf{P}}^T \hat{\mathbf{V}}_y \hat{\mathbf{P}} = \hat{\Lambda} = \text{diag}(\hat{\lambda}_1, \hat{\lambda}_2, \dots, \hat{\lambda}_m) \tag{2}$$

$$\text{и } \hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \dots \geq \hat{\lambda}_m.$$

Очевидно, что матрицы  $\mathbf{V}_y$  и  $\hat{\mathbf{V}}_y$  в принципе не равны, следовательно, не равны и матрицы  $\mathbf{P}$  и  $\hat{\mathbf{P}}$ .

Задача снижения размерности, как и другие задачи, рассмотренные в данной работе, непосредственно связана с проверкой следующей гипотезы:

$$H: \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k \geq \lambda_{k+1} = \lambda_{k+2} = \dots = \lambda_m. \tag{3}$$

Принятие этой гипотезы позволяет рассматривать вектор меньшей размерности

$$\hat{\mathbf{P}}_{(k)}^T (\mathbf{y} - \hat{\boldsymbol{\theta}}_y),$$

где матрица  $\hat{\mathbf{P}}_{(k)}$  содержит только первые  $k$  столбцов матрицы  $\hat{\mathbf{P}}$ .

Первым известным тестом для проверки гипотезы  $H$  (см. формулу (3)) был тест Бартлетта [13, 14] (см., например, монографию [15] или справочник [16]). Однако, если первое цитирование теста Бартлетта [15] предполагать безупречным, то во втором обнаруживаются сразу две неточности (знак и множитель).

Действительно, в работе [15]  $\chi^2$ -статистика для проверки гипотезы  $H$  выражается как

$$\gamma_n = n' \left\{ (m-k) \ln \left( \frac{1}{m-k} \sum_{i=k+1}^m \hat{\lambda}_i \right) - \ln \left( \prod_{i=k+1}^m \hat{\lambda}_i \right) \right\},$$

где

$$n' = n - k - \frac{1}{6} \left[ 2(m-k) + 1 + \frac{2}{(m-k)} \right],$$

тогда как в справочном издании [16] эта статистика следует выражению

$$\gamma_{\eta} = (n - 1) \left\{ (m - k) \ln \left( \frac{1}{m - k} \sum_{i=k+1}^m \hat{\lambda}_i \right) + \ln \left( \prod_{i=k+1}^m \hat{\lambda}_i \right) \right\}.$$

Оба издания указывают одинаковое число степеней свободы  $\eta$ .

В работе [15]:

$$\eta = \frac{1}{2} (m - k + 2)(m - k - 1);$$

В работе [16]:

$$\eta = \frac{1}{2} (m - k + 1)(m - k) - 1.$$

В настоящее время широкую известность приобрел тест, называемый правилом «сломанной трости» (англ. Broken stick model, см, например, статью [17]). Согласно этому тесту, следует выбрать максимальное значение  $k$ , при котором выполняется неравенство

$$\frac{\hat{\lambda}_k}{\text{tr} \hat{\mathbf{V}}_y} > \frac{1/k + 1/(k+1) + \dots + 1/m}{m},$$

где  $\text{tr}$  – след матрицы.

Наиболее простым средством определения  $k$  является визуальный анализ графического представления спектра ковариационной матрицы в порядке убывания. В этом случае в качестве  $k$  берется значение, которое предшествует смене относительно быстрого убывания собственных значений на относительно медленное (плавное), что, в сущности, повторяет метод сломанной трости (на интуитивном, неформализованном уровне). Во многих справочных и учебных изданиях предлагается определять  $k$  из соотношения

$$\left( \sum_{i=1}^k \hat{\lambda}_i / \sum_{i=1}^m \hat{\lambda}_i \right) 100\% \approx K\%,$$

где  $K$  – заранее установленный процент общей дисперсии.

Выбор метода определения  $k$ , в конечном итоге, зависит от характера решаемой задачи. Приступим к изложению результатов.

Очевидно, что при снижении размерности возникает ошибка, дисперсия которой

должна быть как-то связана с отбрасываемой частью выборочного спектра. Решению этой задачи посвящены два следующих раздела статьи.

### Априорная оценка дисперсии погрешности регрессии на РСs

Регрессией вектора  $\mathbf{y}$  на РСs (на компоненты вектора  $\mathbf{z}$ ) называется связь, выраженная уравнением вида

$$\mathbf{y} = \boldsymbol{\theta}_y + \mathbf{P}_{(k)} \mathbf{z} + \boldsymbol{\varepsilon}. \quad (4)$$

Согласно общим принципам классического регрессионного анализа, предполагается, что

$$\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}), \quad (5)$$

где  $\mathbf{0}$  – нулевой вектор, а  $\mathbf{I}$  – единичная матрица соответствующей размерности.

Предположение (5) означает, что

$$E(\varepsilon_i) = 0, \quad \text{var}(\varepsilon_i) = \sigma^2 \quad (i = 1, 2, \dots, m)$$

$$\text{и cov}(\varepsilon_i, \varepsilon_j) = 0 \quad (i \neq j),$$

где  $E$ ,  $\text{var}$  и  $\text{cov}$  – операторы математического ожидания, дисперсии и ковариации, соответственно.

Пусть в описанных выше условиях рассматривается некоторая реализация вектора  $\mathbf{y}$ . Это может быть одна из тех реализаций, которые использовались при вычислении оценок параметров  $(\hat{\boldsymbol{\theta}}_y, \hat{\mathbf{V}}_y)$ , т. е.  $\mathbf{y}_j$ , ( $j = 1, 2, \dots, n$ ), или одна из последующих  $\mathbf{y}_{n+l}$ , ( $l \geq 1$ ), что не принципиально. Поэтому нижний индекс пока опускаем. В практических задачах моделирования вместо уравнения (4) мы имеем регрессионную модель вида

$$\mathbf{y} = \hat{\boldsymbol{\theta}}_y + \hat{\mathbf{P}}_{(k)} \mathbf{z} + \boldsymbol{\varepsilon}. \quad (6)$$

Прежде всего, заметим, что, согласно классическому регрессионному анализу, уравнение (6) следовало бы называть регрессией на базис РСs, так как сами РСs (компоненты вектора  $\mathbf{z}$ ) на этапе применения построенной модели (6) являются определяемыми параметрами. Регрессия именно на РСs фигурирует в доказательстве следующей теоремы.

**Теорема 1.** В условиях модели (6), где элементы модели вычислены по формулам

(1), (2), при верной (принятой) гипотезе Н (см. выражение (3)) и предположении (5), априорная несмещенная оценка  $\sigma^2$  выражается как

$$\hat{\sigma}^2 = \frac{n-1}{(m-k)(n-k-1)} \sum_{i=k+1}^m \hat{\lambda}_i. \quad (7)$$

Доказательство. В регрессионной модели (4), которая связывает три нормально распределенных центрированных вектора ( $\mathbf{y} - \hat{\theta}_y$ ,  $\mathbf{z}$  и  $\boldsymbol{\varepsilon}$ ), перейдем к ковариациям. Тогда имеем следующее равенство:

$$\mathbf{V}_y = \mathbf{P}_{(k)} \mathbf{V}_z \mathbf{P}_{(k)}^T + \mathbf{V}_\varepsilon, \quad (8)$$

где

$$\mathbf{V}_\varepsilon = \sigma^2 \mathbf{I}, \quad \mathbf{V}_z = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2),$$

$$\sigma_i^2 = \text{var}(z_i), \quad (i=1, 2, \dots, k).$$

Умножим равенство (8) слева на  $\mathbf{P}^T$ , а справа – на  $\mathbf{P}$ . В результате имеем:

$$\begin{aligned} \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_k, \lambda_{k+1}, \dots, \lambda_m) = \\ = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2, 0, \dots, 0) + \\ + \text{diag}(\sigma^2, \sigma^2, \dots, \sigma^2), \end{aligned} \quad (9)$$

т. е.

$$\lambda_i = \sigma_i^2 + \sigma^2 \quad (i=1, 2, \dots, k),$$

$$\lambda_i = \sigma^2 \quad (i=k+1, k+2, \dots, m).$$

Здесь следует особо отметить, что не только равенство (9) справедливо лишь при верной гипотезе Н (см. (3)), но и предположение (5) в случае большой выборки также возможно лишь при верной гипотезе Н. Это непосредственно видно из равенства (9), где  $\lambda_i = \sigma^2$ . Однако эта логика нарушается в случае малой выборки (см. далее).

Из равенства (9) также следует, что

$$\sigma^2 = \frac{1}{(m-k)} \sum_{i=k+1}^m \lambda_i. \quad (10)$$

Рассмотрим классическую модель регрессии [18]:

$$\begin{aligned} x_j = \beta_0 + \beta_1 z_{j,1} + \beta_2 z_{j,2} + \dots \\ \dots + \beta_{p-1} z_{j,p-1} + \varepsilon_j \quad (j=1, 2, \dots, n). \end{aligned} \quad (11)$$

Для произвольно выбранного значения  $i$  ( $i=1, 2, \dots, m$ ) подставим в совокупность уравнений (11) вместо  $x_j$  исходные значения  $y_{i,j}$ , а в качестве  $z_{j,l}$  ( $l=1, 2, \dots, k; k=p-1$ )

подставим значения выборочных РС, вычисленные по формуле

$$(\hat{z}_{j,1}, \hat{z}_{j,2}, \dots, \hat{z}_{j,k}) = (\mathbf{y}_j - \hat{\theta}_y)^T \hat{\mathbf{P}}_{(k)},$$

где  $\mathbf{y}_j$  –  $j$ -я реализация  $\mathbf{y}$  в исходной выборке.

В последнем случае мы оставляем прежние обозначения. Это РС исходной выборки, но индексы (номера) компоненты и реализации ( $l$  и  $j$ ) меняем местами, приводя в соответствие стандартам регрессионного анализа, т. е. модели (11). В данном контексте эти значения РС

$$\{\hat{z}_{j,1}, \hat{z}_{j,2}, \dots, \hat{z}_{j,k}\}$$

считаем известными и модель (11) здесь действительно есть регрессия на РС. Тогда уравнение (11) будет соответствовать  $i$ -й строке матричного равенства (6).

Из условия

$$\bar{z}_l = \frac{1}{n} \sum_{j=1}^n z_{j,l} = 0 \quad (l=1, 2, \dots, k)$$

следует, что

$$\hat{\beta}_0 = \bar{x},$$

где  $\bar{x} = \frac{1}{n} \sum_{j=1}^n x_j$ , т. е.  $\hat{\beta}_0 = \frac{1}{n} \sum_{j=1}^n y_{i,j}$  (см. выше).

РСА обеспечивает минимальность остаточной суммы квадратов, как и метод наименьших квадратов (OLS – Ordinary Least Squares), следовательно, вектор-строка OLS-оценок

$$(\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_{p-1})$$

совпадает с  $i$ -й строкой матрицы  $\hat{\mathbf{P}}_{(k)}$ , т. е.

$$(\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_{p-1}) = [\hat{\mathbf{P}}_{(k)}]_i,$$

где  $[\hat{\mathbf{P}}_{(k)}]_i$  –  $i$ -я строка матрицы  $\hat{\mathbf{P}}_{(k)}$ .

Это утверждение нетрудно проверить непосредственным вычислением. Определим следующие матрицы исходных данных:

$$\mathbf{Y} = (\mathbf{y}_1 - \hat{\theta}, \mathbf{y}_2 - \hat{\theta}, \dots, \mathbf{y}_n - \hat{\theta}),$$

$$\mathbf{Z} = \hat{\mathbf{P}}_{(k)}^T \mathbf{Y}.$$

Тогда аналог модели (6) для этих матриц можно записать в виде

$$\mathbf{Y} = \hat{\mathbf{P}}_{(k)} \mathbf{Z} + \mathbf{E},$$

где  $\mathbf{E}$  –  $(m \times n)$ -матрица всех остатков регрессии.

Соответственно, аналог (11) запишется в виде

$$\mathbf{Y}^T = \mathbf{Z}^T \widehat{\mathbf{P}}_{(k)}^T + \mathbf{E}^T.$$

Далее нам нужно проверить равенство

$$\widehat{\mathbf{P}}_{(k)}^T = (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{Z}\mathbf{Y}^T,$$

которое и означает, что матрица  $\widehat{\mathbf{P}}_{(k)}$ , а, соответственно, и все ее строки – суть OLS-оценки.

Подставляя выражение для матрицы  $\mathbf{Z}$  (см. выше), умножая справа на  $\widehat{\mathbf{P}}_{(k)} \widehat{\mathbf{P}}_{(k)}^T$  и расставляя дополнительные скобки, имеем очевидное тождество:

$$\begin{aligned} & (\widehat{\mathbf{P}}_{(k)}^T \widehat{\mathbf{P}}_{(k)}) \widehat{\mathbf{P}}_{(k)}^T = \\ & = (\widehat{\mathbf{P}}_{(k)}^T \mathbf{Y}\mathbf{Y}^T \widehat{\mathbf{P}}_{(k)})^{-1} (\widehat{\mathbf{P}}_{(k)}^T \mathbf{Y}\mathbf{Y}^T \widehat{\mathbf{P}}_{(k)}) \widehat{\mathbf{P}}_{(k)}^T, \end{aligned}$$

так как  $\widehat{\mathbf{P}}_{(k)}^T \widehat{\mathbf{P}}_{(k)} = \mathbf{I}$ .

Корректность умножения на  $\widehat{\mathbf{P}}_{(k)} \widehat{\mathbf{P}}_{(k)}^T$  основана на том, что ранг матриц при этом не снижается:

$$\text{rank}(\widehat{\mathbf{P}}_{(k)} \widehat{\mathbf{P}}_{(k)}^T) = \text{rank} \widehat{\mathbf{P}}_{(k)} = k.$$

Из принятых предположений и равенства (10) следует, что остаточная сумма квадратов  $S^2$  регрессионной модели (11) выражается как

$$S^2 = \frac{n-1}{(m-k)} \sum_{i=k+1}^m \widehat{\lambda}_i.$$

В соответствии с теорией линейной регрессии [18], оценка

$$\widehat{\sigma}^2 = \frac{S^2}{n-p} = \frac{n-1}{(m-k)(n-k-1)} \sum_{i=k+1}^m \widehat{\lambda}_i$$

является несмещенной (напомним, что  $p = k + 1$ ).

Теорема 1 доказана.

Отметим, что оценка (7) и краткий набросок доказательства теоремы 1, содержащий основную идею (см. формулу (9)), были предложены ранее автором настоящей статьи в работе [19]. В ряде задач, касающихся оценки информативности (см. [20]), требуется именно смещенная оценка  $\sigma^2$ . В этом случае необходимо брать оценку

$$\widetilde{\sigma}^2 = \frac{n-1}{(m-k)n} \sum_{i=k+1}^m \widehat{\lambda}_i,$$

которая непосредственно следует из формулы (10). К вопросу об использовании

априорной оценки мы вернемся далее (см. раздел «Восстановление пропущенных данных»).

### Проблема относительно малой выборки

Во многих задачах нередко возникает ситуация, когда объем выборки  $n$  меньше размерности вектора  $m$  ( $n < m$ ). В этом случае

$$\widehat{\lambda}_{n+1} = \widehat{\lambda}_{n+2} = \dots = \widehat{\lambda}_m = 0,$$

что никак не соответствует равенству (9), а, следовательно, и (10). Это обстоятельство, однако, не исключает возможности проверки гипотезы

$$\begin{aligned} H_1: \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k \geq \lambda_{k+1} = \\ = \lambda_{k+2} = \dots = \lambda_n \end{aligned} \quad (12)$$

для дальнейшего рассмотрения модели (6).

Если считать вектор  $\boldsymbol{\theta}_y$  и матрицу  $\widehat{\mathbf{P}}_{(k)}$  известными, то для какой-либо из последующих реализаций вектора  $\mathbf{y}$ , например  $(n+l)$ -й ( $l \geq 1$ ), несмещенная апостериорная оценка  $\sigma^2$  в регрессионной модели (6) следует выражению

$$\widehat{\sigma}^2 = \frac{1}{m-k} \sum_{i=1}^m (y_{i,n+l} - \bar{y}_i - [\widehat{\mathbf{P}}_{(k)}]_i \bar{\mathbf{z}}_{n+l})^2, \quad (13)$$

где  $\bar{y}_i$  –  $i$ -я компонента априорной оценки вектора  $\widehat{\boldsymbol{\theta}}_y$  ( $i = 1, 2, \dots, m$ ), а число оцениваемых параметров равно  $k$ .

Если же считать известной лишь матрицу  $\widehat{\mathbf{P}}_{(k)}$ , то эта оценка будет следующей:

$$\begin{aligned} \widetilde{\sigma}^2 = \frac{1}{m-k-1} \sum_{i=1}^m (y_{i,n+l} - \bar{y}_{n+l} - \\ - [\widehat{\mathbf{P}}_{(k)}]_i \bar{\mathbf{z}}_{n+l})^2, \end{aligned} \quad (14)$$

где  $\bar{y}_{n+l} = \frac{1}{m} \sum_{i=1}^m y_{i,n+l}$ ;  $y_{i,n+l}$  –  $i$ -я компонента вектора  $\mathbf{y}_{n+l}$ .

В обоих случаях, отвечающих формулам (13) и (14), в этих формулах, как и выше,  $[\widehat{\mathbf{P}}_{(k)}]_i$  –  $i$ -я строка матрицы  $\widehat{\mathbf{P}}_{(k)}$ , а

$$\bar{\mathbf{z}}_{n+l} = \widehat{\mathbf{P}}_{(k)}^T (\mathbf{y}_{n+l} - \widehat{\boldsymbol{\theta}}_y).$$

Однако в выражении (13) вектор  $\widehat{\boldsymbol{\theta}}_y$  вычислен по формуле (1) с использованием всех реализаций исходной выборки, а в (14) мы заменяем все компоненты  $\widehat{\boldsymbol{\theta}}_y$

средним значением по компонентам новой реализации  $\mathbf{y}_{n+1}$  ( $\theta_1 = \theta_2 = \dots = \theta_m = \bar{y}_{n+1}$ , см выше). Однако вектор  $\hat{\theta}_y$  используется при вычислении  $\hat{\mathbf{V}}_y$  и  $\hat{\mathbf{P}}$ , что необходимо учесть в априорной оценке  $\sigma^2$ . В связи с этим рассмотрим иной метод вычисления элементов модели (6).

Определим иначе матрицу  $\mathbf{Y}$ . Пусть

$$\mathbf{Y} = (\mathbf{y}_1 - \hat{\theta}_1, \mathbf{y}_2 - \hat{\theta}_2, \dots, \mathbf{y}_n - \hat{\theta}_n),$$

где компоненты каждого вектора  $\hat{\theta}_j$  равны между собой и равны среднему по компонентам реализации  $\mathbf{y}_j$  т. е.

$$\theta_{1j} = \theta_{2j} = \dots = \theta_{mj} = \bar{y}_j = \frac{1}{m} \sum_{i=1}^m y_{i,j} \\ (j=1, 2, \dots, n).$$

Вычислим оценку  $\mathbf{V}_y$  по формуле

$$\hat{\mathbf{V}}_y = (n-1)^{-1} \mathbf{Y} \mathbf{Y}^T. \quad (15)$$

Далее произведем все перечисленные выше действия: вычислим  $\hat{\mathbf{P}}$  и  $\{\hat{\lambda}_1, \hat{\lambda}_2, \dots, \hat{\lambda}_n\}$ ; проверим гипотезу  $H_1$  (см. (12),  $n < m$ ) и определим матрицу  $\hat{\mathbf{P}}_{(k)}$ . При этом могут несколько измениться не только  $\hat{\mathbf{V}}_y$  и  $\hat{\mathbf{P}}$ , но и значение  $k$ .

**Теорема 2.** В условиях модели (6), где элементы модели вычислены по схеме (15), при верной (принятой) гипотезе  $H_1$  (см. (12)) и предположении (5), априорная оценка  $\sigma^2$ , равная

$$\hat{\sigma}^2 = \frac{n-1}{(m-k-1)n} \sum_{i=k+1}^n \hat{\lambda}_i, \quad (16)$$

является несмещенной.

**Доказательство.** В силу линейности оператора математического ожидания, среднее значение несмещенной оценки (14), вычисленное по исходной выборке ( $j=1, 2, \dots, n$ ) и есть несмещенная априорная оценка, равная

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{j=1}^n \frac{1}{m-k-1} (\mathbf{y}_j - \hat{\theta}_j)^T \times \\ \times (\mathbf{I} - \hat{\mathbf{P}}_{(k)} \hat{\mathbf{P}}_{(k)}^T) (\mathbf{y}_j - \hat{\theta}_j) = \\ = \frac{1}{(m-k-1)n} \text{tr}(\mathbf{Y}^T (\mathbf{I} - \hat{\mathbf{P}}_{(k)} \hat{\mathbf{P}}_{(k)}^T) \mathbf{Y}) = \\ = \frac{1}{(m-k-1)n} (\text{tr} \mathbf{Y}^T \mathbf{Y} - \text{tr} \mathbf{Z}^T \mathbf{Z}) =$$

$$= \frac{n-1}{(m-k-1)n} \sum_{i=k+1}^n \hat{\lambda}_i$$

(здесь, как и прежде,  $\mathbf{Z} = \hat{\mathbf{P}}_{(k)}^T \mathbf{Y}$ ).

Теорема 2 доказана.

Заметим, что оценка (16) не требует совпадения математических ожиданий оценок средних ( $\hat{\theta}_j$ ), а требует лишь совпадения математических ожиданий оценок (14). Это означает, что выражение (16) подходит для случая временных рядов, содержащих тренды. В свою очередь, это может быть иным основанием (не только относительная малость объема выборки) к применению формул (15) и (16). Сравнивая (7) и (16), замечаем также, что в случае малой выборки ( $n < m$ ) некорректное использование (7) будет давать завышенное значение оценки  $\sigma^2$ . Смещенным вариантом оценки (16) будет оценка

$$\tilde{\sigma}^2 = \frac{n-1}{mn} \sum_{i=k+1}^n \hat{\lambda}_i.$$

Если принять во внимание тот факт, что индекс  $j$  в математическом моделировании нередко соответствует некоторому временному отсчету, то при достаточной стационарности и несущественном различии средних значений компонент вектора  $\mathbf{y}$ , вычисленных первым способом (см. оценки (1)), можно применять оценку (16), не прибегая к описанной выше коррекции модели (6). Возникающие при этом ошибки и неточности будут несущественны. И, наоборот, в нестационарном случае следует использовать выше изложенный метод оценивания  $\hat{\mathbf{V}}_y$ . При этом подразумевается пересчет  $\hat{\theta}_y$  (см. выше) по реализации  $\mathbf{y}$ , непосредственно фигурирующей в модели (6).

Рассмотренный в этом разделе метод оценивания элементов модели (6) наиболее эффективен, когда метод гусеницы применяется к прогнозу нестационарных временных рядов с явно выраженным трендом (см. раздел «Прогноз нестационарных временных рядов»).

**Важное замечание.** Прежде чем перейти к рассмотрению следующих вопросов, необходимо отметить один очень важный для общего понимания момент. В случае

невыврожденного распределения и большого объема выборки мы можем смотреть на главные компоненты как на регрессоры, а на элементы матрицы  $\hat{\mathbf{P}}_{(k)}$ , представляющей собой базис главных компонент, как на оцениваемые параметры, что мы видим в доказательстве теоремы 1. Случай же малой выборки (преобладание размерности над объемом), наоборот, приводит нас к необходимости рассматривать элементы матрицы  $\hat{\mathbf{P}}_{(k)}$  как регрессоры, а главные компоненты — как оцениваемые параметры (теорема 2). При этом следует осуществлять центрирование вычитанием не среднего значения по реализациям для каждой компоненты, а вычитанием среднего по компонентам для каждой реализации (см. выше). Суть проблемы состоит в том, что при малой выборке равенство (9) не имеет выборочного аналога, так как выборочный спектр не полон и логика теоремы 1 рушится. Поэтому приходится менять точку зрения (обзора ситуации).

И, наконец, отметим, что случай принципиально вырожденного распределения, когда выборочный спектр неполон и большой объем выборки, приводит нас к вычислительной схеме данного раздела и оценке (16).

### Оценки минимального риска

При выборе модели (6) для использования в каких-либо прикладных задачах обычно применяется формула:

$$\hat{\mathbf{y}} = \hat{\boldsymbol{\theta}}_y + \hat{\mathbf{P}}_{(k)} \hat{\mathbf{z}},$$

где  $\hat{\mathbf{z}} = \hat{\mathbf{P}}_{(k)}^T (\mathbf{y} - \hat{\boldsymbol{\theta}}_y)$ .

В справочном издании [21] предлагается применять формулу

$$\tilde{\mathbf{y}} = \hat{\boldsymbol{\theta}}_y + \hat{\mathbf{P}}_{(k)} \mathbf{G} \hat{\mathbf{z}}, \quad (17)$$

где  $\mathbf{G} = \text{diag}(g_1, g_2, \dots, g_k)$ .

При этом значения  $g_i$  определяются из условия минимума квадратичного риска

$$R^2 = E(z_i - g_i \hat{z}_i)^2 \quad (i=1, 2, \dots, k).$$

В этой ситуации компоненты вектора  $\mathbf{G} \hat{\mathbf{z}}$  называются оценками минимального риска. С учетом того, что

$$z_i - g_i \hat{z}_i = z_i(1 - g_i) - g_i(\hat{z}_i - z_i),$$

имеем равенство

$$R^2 = z_i^2(g_i - 1)^2 + g_i^2 \sigma_{\hat{z}_i}^2, \quad (18)$$

где  $E(\hat{z}_i - z_i)^2 = \sigma_{\hat{z}_i}^2$ .

Приравнявая к нулю производную  $R^2$  по  $g_i$ , получаем равенство

$$g_i(z_i^2 + \sigma_{\hat{z}_i}^2) = z_i^2, \quad (19)$$

или  $g_i = \frac{1}{1 + \sigma_{\hat{z}_i}^2 / z_i^2}$ .

Подставляя вместо  $z_i$  оценку минимального риска  $g_i \hat{z}_i$ , получаем простое квадратное уравнение вида

$$g_i^2 - g_i + \delta_i^2 = 0, \quad \delta_i^2 = \sigma_{\hat{z}_i}^2 / \hat{z}_i^2$$

со следующей окончательной формулой для  $g_i$ :

$$g_i = \frac{1}{2} + \sqrt{\frac{1}{4} - \delta_i^2}.$$

В справочном издании [21] авторами предлагается следующий алгоритм вычисления величины  $g_i$ :

$$\begin{cases} g_i = \frac{1}{2} + \sqrt{\frac{1}{4} - \delta_i^2}, & \text{если } \delta_i^2 \leq \frac{1}{4}; \\ g_i = 0, & \text{если } \delta_i^2 > \frac{1}{4}. \end{cases} \quad (20)$$

Автором настоящей статьи в работе [22] предложен альтернативный алгоритм:

$$\begin{cases} g_i = \frac{1}{2} + \sqrt{\frac{1}{4} - \delta_i^2}, & \text{если } \delta_i^2 \leq \frac{1}{4}; \\ g_i = \frac{1}{2}, & \text{если } \frac{1}{4} < \delta_i^2 \leq 1; \\ g_i = \frac{1}{1 + \delta_i^2}, & \text{если } \delta_i^2 > 1. \end{cases} \quad (21)$$

**Теорема 3.** В предположении, что различие между величинами  $\sigma_{\hat{z}_i}^2 / \hat{z}_i^2$  и  $\sigma_{\hat{z}_i}^2 / z_i^2$  пренебрежимо мало, предлагаемый алгоритм (21) обеспечивает меньшее значение квадратичного риска  $R^2$ , по сравнению с алгоритмом (20), для случая  $\delta_i^2 > \frac{1}{4}$ .

Доказательство. При условии  $\delta_i^2 > \frac{1}{4}$  алгоритм (20) дает величину  $R^2 = z_i^2$ . Если  $\frac{1}{4} < \delta_i^2 \leq 1$ , подставляем в (18) значение  $g_i = \frac{1}{2}$ . Учитывая, что  $\sigma_{\hat{z}_i}^2 < z_i^2$ , получаем,



что

$$R^2 = \frac{1}{4} (z_i^2 + \sigma_{z_i}^2) < \frac{z_i^2}{2} < z_i^2.$$

Если же  $\delta_i^2 > 1$ , то, согласно алгоритму (21), значение  $g_i$  таково, что выполняется равенство (19). Подставляя (19) в (18), получаем, что

$$\begin{aligned} R^2 &= g_i^2 (z_i^2 + \sigma_{z_i}^2) - 2g_i z_i^2 + z_i^2 = \\ &= g_i z_i^2 - 2g_i z_i^2 + z_i^2 = z_i^2 (1 - g_i) < z_i^2. \end{aligned}$$

Теорема 3 доказана.

Полученный в данном разделе результат можно отнести к уточнению деталей. При исследовании различных многомерных процессов и явлений, в особенности природных, не всегда использование РС подчинено задаче снижения размерности, а может быть направлено на исследование внутренней структуры явления, что показано в следующем разделе данной работы.

### Структурное подобие и однородность

Предположим, что кроме вектора  $y$  исследуется и вектор  $x$  такой же размерности:

$$\dim x = \dim y = m.$$

Предположим, что имеется выборка реализаций этого вектора. Объемы выборок реализаций векторов  $x$  и  $y$  могут быть различны. Используя формулы (1) для вектора  $x$ , вычислим оценки параметров распределения  $\hat{\theta}_x$  и  $\hat{V}_x$ . Пусть ортогональная матрица  $\hat{Q}$  такова, что выполняется равенство

$$\hat{Q}^T \hat{V}_x \hat{Q} = \text{diag}(\hat{\mu}_1, \hat{\mu}_2, \dots, \hat{\mu}_m), \quad (22)$$

$$\hat{\mu}_1 \geq \hat{\mu}_2 \geq \dots \geq \hat{\mu}_m. \quad (23)$$

В этой ситуации мы имеем два множества статистических характеристик РС обоих векторов:

$$\{\hat{\lambda}_i, \hat{p}_i\}_{i=1}^m \text{ и } \{\hat{\mu}_i, \hat{q}_i\}_{i=1}^m,$$

причем  $\hat{q}_i, \hat{p}_i$  есть  $i$ -е столбцы матриц  $\hat{Q}$  и  $\hat{P}$ , соответственно, т. е.

$$\hat{Q} = (\hat{q}_1, \hat{q}_2, \dots, \hat{q}_m),$$

$$\hat{P} = (\hat{p}_1, \hat{p}_2, \dots, \hat{p}_m).$$

Оценкой коэффициента структурного

подобия векторов  $x$  и  $y$  назовем величину

$$\hat{s}_{xy} = \frac{\sum_{i=1}^m \sqrt{\hat{\mu}_i \hat{\lambda}_i} |\hat{q}_i^T \hat{p}_i|}{\sqrt{\text{tr} \hat{V}_x \text{tr} \hat{V}_y}}. \quad (24)$$

Этот коэффициент показывает, насколько структуры колебаний исследуемых векторов согласуются в относительных долях дисперсии. В некоторых случаях  $\hat{s}_{xy}$  целесообразно вычислять по формуле

$$\hat{s}_{xy} = \max_{\varphi(i)} \frac{\sum_{i=1}^m \sqrt{\hat{\mu}_{\varphi(i)} \hat{\lambda}_i} |\hat{q}_{\varphi(i)}^T \hat{p}_i|}{\sqrt{\text{tr} \hat{V}_x \text{tr} \hat{V}_y}}, \quad (25)$$

где  $\varphi(i)$  есть перестановка индексов, т. е. варьируется порядок статистических характеристик РС одного из векторов (здесь это  $x$ ).

Необходимость применения формулы (25) может возникнуть в случае обнаружения близких собственных значений хотя бы в спектре одного из векторов. Если верна (тестирована) гипотеза

$$H_x: \mu_1 \geq \mu_2 \geq \dots \geq \mu_l \geq \mu_{l+1} = \mu_{l+2} = \dots = \mu_m,$$

то можно рассматривать отфильтрованный коэффициент структурного подобия в виде

$$\hat{s}_{xy}^f = \frac{\sum_{i=1}^p \sqrt{\hat{\mu}_i \hat{\lambda}_i} |\hat{q}_i^T \hat{p}_i|}{\sqrt{\sum_{i=1}^p \hat{\mu}_i \sum_{i=1}^p \hat{\lambda}_i}},$$

где  $p = \min(k, l)$ ,

или относительный коэффициент структурного подобия вида

$$\hat{s}_{xy}^r = \frac{\sum_{i=1}^p \sqrt{\hat{\mu}_i \hat{\lambda}_i} |\hat{q}_i^T \hat{p}_i|}{\sqrt{\text{tr} \hat{V}_x \text{tr} \hat{V}_y}}.$$

Оценки коэффициентов  $\hat{s}_{xy}^f$  и  $\hat{s}_{xy}^r$  в случаях, когда соседние собственные значения мало различимы, целесообразно находить по формулам, аналогичным формуле (25).

В различных исследованиях коэффициент структурного подобия может использоваться для сравнения метеорологических, климатических и океанографических полей, а также для анализа полей экологического и медицинского мониторинга об-

ластей и районов. В микроэлектронном производстве, когда в каждой ячейке кристаллической пластины производится несколько типов микроэлектронных приборов (см., например, работу [23]), коэффициент структурного подобия удобен для оценки, насколько погрешности изготовления различных приборов связаны между собой и зависит ли эта погрешность от положения ячейки на кристаллической пластине. Такие вопросы могут возникать и в процессе настройки оборудования. В анализе временных рядов базой для применения коэффициента структурного подобия является упомянутый во введении анализ сингулярного спектра (SSA). Первые попытки построения коэффициента структурного подобия были предприняты автором данной работы именно при анализе временных рядов [24].

Если в качестве выборки вектора  $x$  используется выборка значений того же вектора  $y$ , а второй набор характеристик PCs  $\{\hat{\mu}_i, \hat{q}_i\}_{i=1}^m$  вычисляется по оценке корреляционной матрицы  $\hat{R}_y$ , т. е.

$$\hat{Q}^T \hat{R}_y \hat{Q} = \text{diag}(\hat{\mu}_1, \hat{\mu}_2, \dots, \hat{\mu}_m),$$

то тогда коэффициент структурного подобия, во всех его вариантах, становится коэффициентом однородности вектора  $y$ . В анализе временных рядов его применение возможно только в сочетании с методом гусеницы (см. выше), где нормирование может оказать влияние на формы собственных векторов, так как формы собственных векторов автоковариационной и автокорреляционной матриц, в принципе, различны.

Возникает естественный вопрос о возможности проверки гипотезы о равенстве коэффициента структурного подобия нулю, т. е.  $H_s: s_{xy} = 0$ . Прежде всего, отметим, что на практике мы всегда имеем лишь оценку  $\hat{s}_{xy}$ , а истинное значение  $s_{xy}$  мы имели бы лишь в случае использования в наших вычислениях матриц  $V_y, V_x, P$  и  $Q$ , что возможно лишь гипотетически.

Рассмотрим два составных вектора:

$$Y_s = (\sqrt{\hat{\lambda}_1} \hat{p}_1^T, \sqrt{\hat{\lambda}_2} \hat{p}_2^T, \dots, \sqrt{\hat{\lambda}_m} \hat{p}_m^T)^T = (Y_1, Y_2, \dots, Y_M)^T;$$

$$X_s = (\sqrt{\hat{\mu}_1} \hat{q}_1^T, \sqrt{\hat{\mu}_2} \hat{q}_2^T, \dots, \sqrt{\hat{\mu}_m} \hat{q}_m^T)^T = (X_1, X_2, \dots, X_M)^T,$$

где  $M = m \times m$ .

Нумерацию подвекторов (субвекторов)

$$\sqrt{\hat{\mu}_i} \hat{q}_i \quad (i=1, 2, \dots, m)$$

составного вектора  $X_s$ , при необходимости, можно установить в соответствии с формулой (25). Это может нарушить лишь условие (23), что не имеет для нас значения. Знаки некоторых столбцов матрицы

$$\hat{Q} = (\hat{q}_1, \hat{q}_2, \dots, \hat{q}_m)$$

нужно поменять на противоположные так, чтобы все произведения  $\hat{q}_i^T \hat{p}_i$  или  $\hat{q}_{\Phi(i)}^T \hat{p}_i$  были положительными. Эта смена знаков не нарушит равенства (22). Тогда, если все эти условия выполнены, из ортогональности матриц  $P$  и  $Q$  следует, что

$$\hat{s}_{xy} = \frac{X_s^T Y_s}{\sqrt{X_s^T X_s Y_s^T Y_s}} = \frac{\sum_{i=1}^M X_i Y_i}{\sqrt{\sum_{i=1}^M X_i^2 \sum_{i=1}^M Y_i^2}}. \quad (26)$$

Рассмотрим два регрессионных уравнения, связывающие компоненты составных векторов  $Y_s$  и  $X_s$ :

$$Y_i = \beta X_i + \varepsilon_i; \quad (27)$$

$$Y_i = b_0 + b_1 X_i + e_i \quad (i=1, 2, \dots, M),$$

где предполагается, что остатки (погрешности) каждой из регрессий подчиняются нормальному закону распределения, взаимно независимы и имеют одинаковую дисперсию.

**Теорема 4.** Если принята (не отвергнута) гипотеза  $H_\varepsilon: E(\varepsilon) = 0$  (или эквивалентная гипотеза  $H_0: b_0 = 0$ ), то, при верной гипотезе  $H_s: s_{xy} = 0$ , величина

$$\frac{\hat{s}_{xy} \sqrt{M-1}}{\sqrt{1 - \hat{s}_{xy}^2}} \sim t_{M-1}, \quad (28)$$

т. е. подчиняется распределению Стьюдента с числом степеней свободы  $M-1$ .

Способы проверки гипотез  $H_0: b_0 = 0$  и  $H_\varepsilon: E(\varepsilon) = 0$  хорошо известны. Поэтому рассмотрим простое доказательство, которое мало чем отличается от хорошо известного для обычного коэффициента корреляции.

Доказательство. Из общей теории метода наименьших квадратов, предположения  $E(\varepsilon) = 0$  (см. гипотезу  $H_\varepsilon$ ) и формулы (26) следует, что OLS-оценка параметра регрессии (27) имеет вид

$$\hat{\beta} = \frac{\sum_{i=1}^M X_i Y_i}{\sum_{i=1}^M X_i^2} = \hat{s}_{xy} \frac{\sqrt{\sum_{i=1}^M Y_i^2}}{\sqrt{\sum_{i=1}^M X_i^2}}, \quad (29)$$

$$\text{var}(\hat{\beta}) = \frac{\sigma^2}{\sum_{i=1}^M X_i^2},$$

где  $\sigma^2 = \text{var}(\varepsilon_i)$ , а

$$\sum_{i=1}^M X_i \varepsilon_i = 0. \quad (30)$$

Из равенства (30) следует, что несмещенная оценка  $\sigma^2$  выражается как

$$\begin{aligned} \hat{\sigma}^2 &= \frac{1}{M-1} \sum_{i=1}^M (Y_i - \hat{\beta} X_i)^2 = \\ &= \frac{1}{M-1} \left( \sum_{i=1}^M Y_i^2 - \hat{\beta}^2 \sum_{i=1}^M X_i^2 \right). \end{aligned} \quad (31)$$

Очевидно, что равенство  $s_{xy} = 0$  эквивалентно  $\beta = 0$  (см. (29)). Следовательно, принимая во внимание выражение (31), имеем:

$$\begin{aligned} t_{M-1} &\sim \frac{\hat{\beta} \sqrt{\sum_{i=1}^M X_i^2}}{\hat{\sigma}} = \frac{\hat{s}_{xy} \sqrt{\sum_{i=1}^M Y_i^2} \sqrt{M-1}}{\sqrt{\sum_{i=1}^M Y_i^2 - \hat{\beta}^2 \sum_{i=1}^M X_i^2}} = \\ &= \frac{\hat{s}_{xy} \sqrt{M-1}}{\sqrt{1 - \hat{s}_{xy}^2}}. \end{aligned}$$

Теорема 4 доказана.

Как было отмечено выше, смена всех знаков матрицы  $\hat{\mathbf{Q}}$  не нарушает формулы (22), но меняет знак  $\hat{s}_{xy}$  и знак (28) на противоположный. Это вполне соответствует симметрии распределения Стьюдента. Нетрудно заметить, что гипотеза  $H_s$  может быть проверена и для отфильтрованного коэффициента структурного подобия  $s_{xy}^f$ . Различие сводится лишь к тому, что в этом случае  $M = m \times p$ , а из  $s_{xy}^f = 0$  следует и

$s_{xy}^r = 0$ . Невозможность проверить гипотезу  $H_s$  (гипотеза  $H_\varepsilon$  отвергнута) не является принципиальным препятствием для вычисления  $\hat{s}_{xy}$ . В большинстве приложений мы скорее заинтересованы в том, чтобы отвергнуть гипотезу  $H_s$ , т. е. установить структурное подобие. Исключение составляет случай, когда мы собираемся установить неоднородность некоторого поля или временного ряда вышеуказанным способом.

Перейдем к рассмотрению задач практического характера, где непосредственно используются результаты разделов «Априорная оценка дисперсии погрешности регрессии на РС», «Проблема относительно малой выборки» и «Оценки минимального риска».

#### Восстановление пропущенных данных

Рассмотрим ситуацию, когда распределение невырожденное и объем выборки достаточно превышает размерность. Для уменьшения количества индексов и прочих знаков перепишем уравнение (6) в виде

$$\mathbf{y} = \bar{\mathbf{y}} + \mathbf{Fz} + \varepsilon, \quad (32)$$

где  $\bar{\mathbf{y}} = \hat{\boldsymbol{\theta}}_y$ ,  $\mathbf{F} = \hat{\mathbf{P}}_{(k)}$ .

Пусть  $\mathbf{y}_{n+l}$  ( $l \geq 1$ ) – какая-либо очередная реализация вектора  $\mathbf{y}$ , которая имеет  $u$  измерений и  $v$  пропусков ( $m = u + v$ ).

Тогда, при соответствующей нумерации, имеем следующие разбиения на блоки:

$$\mathbf{y}_{n+l} = (\mathbf{y}_1^T, \mathbf{y}_2^T)^T, \quad \bar{\mathbf{y}} = (\bar{\mathbf{y}}_1^T, \bar{\mathbf{y}}_2^T)^T,$$

$$\mathbf{F} = (\mathbf{F}_1^T, \mathbf{F}_2^T)^T,$$

где блоки  $\mathbf{y}_1$ ,  $\bar{\mathbf{y}}_1$  и  $\mathbf{F}_1$  соответствуют измеренным компонентам  $\mathbf{y}_{n+l}$ , а блоки  $\mathbf{y}_2$ ,  $\bar{\mathbf{y}}_2$  и  $\mathbf{F}_2$  – пропущенным данным.

Вычислим оценку  $\mathbf{y}_2$  по формуле

$$\hat{\mathbf{y}}_2 = \bar{\mathbf{y}}_2 + \mathbf{F}_2 \mathbf{G} (\mathbf{F}_1^T \mathbf{F}_1)^{-1} \mathbf{F}_1^T (\mathbf{y}_1 - \bar{\mathbf{y}}_1), \quad (33)$$

где матрица  $\mathbf{G}$  (см. формулу (17)) имеет размерность  $k \times k$ , а ее компоненты вычисляются по алгоритму (21).

При этом в алгоритме (21)

$$\sigma_{\hat{z}_i}^2 = \hat{\sigma}^2 [(\mathbf{F}_1^T \mathbf{F}_1)^{-1}]_{i,i},$$

где  $[\dots]_{i,j}$  – оператор взятия элемента матрицы с указанными номерами строки ( $i$ ) и столбца ( $j$ );  $\hat{z}_i$  –  $i$ -я компонента вектора  $\hat{\mathbf{z}}$

OLS-оценок:

$$\hat{z} = (\mathbf{F}_1^T \mathbf{F}_1)^{-1} \mathbf{F}_1^T (\mathbf{y}_1 - \bar{\mathbf{y}}_1) \quad (i=1, 2, \dots, k).$$

Из-за отсутствующих строк столбцы матрицы  $\mathbf{F}_1$  не ортогональны, матрица  $\mathbf{F}_1^T \mathbf{F}_1$  не диагональна, а компоненты вектора  $\hat{z}$  зависимы между собой:

$$\mathbf{V}_{\hat{z}} = \sigma^2 (\mathbf{F}_1^T \mathbf{F}_1)^{-1}.$$

Апостериорная оценка  $\sigma^2$  в этом случае имеет вид (см. выражение (13)):

$$\hat{\sigma}^2 = \frac{1}{u-k} (\mathbf{y}_1 - \bar{\mathbf{y}}_1)^T \times (\mathbf{I} - \mathbf{F}_1 (\mathbf{F}_1^T \mathbf{F}_1)^{-1} \mathbf{F}_1^T) (\mathbf{y}_1 - \bar{\mathbf{y}}_1).$$

Однако именно в этой ситуации естественно воспользоваться более надежной априорной оценкой (7) как вычисленной по выборке существенно большего объема. Эта оценка используется во всех следующих формулах данного раздела.

Здесь следует отметить, что  $\hat{\mathbf{y}}_2$  есть, в сущности, центр условного распределения  $\mathbf{y}_2$ . Действительно, безусловное или априорное распределение  $\mathbf{y}_2$  в некотором приближении (при достаточно большом  $n$ ) есть  $N(\bar{\mathbf{y}}_2, \hat{\mathbf{V}}_2)$ , где  $\hat{\mathbf{V}}_2$  получается из  $\hat{\mathbf{V}}_y$  удалением строк и столбцов, которые не имеют отношения к  $\mathbf{y}_2$ . Условное или апостериорное распределение  $\mathbf{y}_2$ , также лишь в некотором приближении, есть  $N(\hat{\mathbf{y}}_2, \hat{\sigma}^2 \mathbf{I})$ , что непосредственно следует из соотношения (5) и уравнений (32) и (33).

Для пропущенных в количестве  $v$  компонент вектора  $\mathbf{y}$  можно указать относительно точное распределение Стьюдента:

$$\frac{y_i - \hat{y}_i}{\hat{\sigma} \sqrt{1 + \frac{1}{n} + \mathbf{f}_i \mathbf{G} (\mathbf{F}_1^T \mathbf{F}_1)^{-1} \mathbf{G} \mathbf{f}_i^T}} \sim t_{\eta}, \quad (34)$$

где число степеней свободы  $\eta = n - k - 1$ , если используется оценка (7) и  $\eta = m - k - 1$ , если по какой-либо причине используется оценка (16);  $\mathbf{f}_i$  —  $i$ -я строка  $\mathbf{F}$  ( $i = u + 1, u + 2, \dots, m$ ; т. е. строка  $\mathbf{f}_i$  относится к блоку  $\mathbf{F}_2$ ).

По формуле (34) и выбранному уровню доверия  $1 - \alpha$  ( $\alpha$  — уровень значимости) для этих значений получаем границы  $(1 - \alpha)\%$ -х стьюдентовских доверительных интервалов:

$$\hat{y}_i \pm t_{\eta}^{\alpha/2} \hat{\sigma} \sqrt{1 + \frac{1}{n} + \mathbf{f}_i \mathbf{G} (\mathbf{F}_1^T \mathbf{F}_1)^{-1} \mathbf{G} \mathbf{f}_i^T}, \quad (35)$$

где  $t_{\eta}^{\alpha/2}$  — квантиль распределения Стьюдента с учетом двустороннего характера критерия (берется  $\alpha / 2$  в силу симметрии распределения Стьюдента), ( $i = u + 1, u + 2, \dots, m$ ).

Целесообразность применения модели (32) и формулы (33) тем выше, чем существеннее неравенство

$$t_{\eta} \hat{\sigma} \sqrt{1 + \frac{1}{n} + \mathbf{f}_i \mathbf{G} (\mathbf{F}_1^T \mathbf{F}_1)^{-1} \mathbf{G} \mathbf{f}_i^T} < t_{n-1} \sqrt{\left(1 + \frac{1}{n}\right) [\hat{\mathbf{V}}_y]_{i,i}}, \quad (36)$$

т. е. чем меньше ширина доверительного интервала, построенного по формуле (35), по сравнению с соответствующей шириной, где доверительный интервал построен по выборочным оценкам параметров распределения ( $\hat{\mathbf{V}}_y$ ), т. е. меньше неопределенность относительно восстанавливаемых данных.

### Прогноз нестационарных временных рядов

Модель (32) и формула (33) могут использоваться для прогноза нестационарных временных рядов [25, 26]. Здесь модель (32) строится методом гусеницы, когда  $\mathbf{y}$  есть скользящий отрезок (гусеница) временного ряда  $\{y_i\}_{i=1}^N$ , а  $\dim \mathbf{y} = m$  есть длина гусеницы.

При вычислении оценки ковариационной матрицы выборка строится пошаговым сдвигом, т. е.

$$\begin{aligned} \mathbf{y}_1 &= (y_1, y_2, \dots, y_m)^T, \\ \mathbf{y}_2 &= (y_2, y_3, \dots, y_{m+1})^T, \\ &\dots \\ \mathbf{y}_n &= (y_{N-m+1}, y_{N-m+2}, \dots, y_N)^T, \end{aligned}$$

где  $n = N - m + 1$ ,  $N$  — общая длина временного ряда.

Оценку ковариационной матрицы следует вычислять по алгоритму вычисления оценки (15). При определении длины гусеницы окончательное значение (здесь предварительные расчеты будут весьма полезны)  $m$  следует брать не меньше, чем период

волны, несущей наибольшую долю дисперсии, т. е. соответствующей  $\hat{\lambda}_1$ . Стремление удовлетворить указанному требованию может приводить к ситуации малого объема выборки ( $n < m$ ). В разделе «Проблема относительно малой выборки» было отмечено, что применение оценок (15) и (16) может быть обусловлено двумя причинами: малый объем выборки или нестационарность.

Авторы метода гусеницы [9] придерживаются классической схемы вычислений оценки ковариационной матрицы (см. разделы «Краткое описание математического аппарата РСА» и «Проблема относительно малой выборки») и ошибочно определяют число степеней свободы по минимальному размеру выборочной матрицы исходных данных (выбор  $\eta = n - k - 1$  или  $\eta = m - k - 1$  (см. выше) определяется через  $\min(m, n)$ ). В методе гусеницы строки и столбцы матрицы исходных данных имеют одну и ту же природу, в связи с чем двойственность, которая изложена в конце раздела «Проблема относительно малой выборки», становится очевиднее. Однако в работе [9] метод гусеницы использовался преимущественно для фильтрации временных рядов, и данный вопрос о числе степеней свободы не стоял столь принципиально, как в случае прогноза по рассматриваемой здесь схеме.

При прогнозировании (в формуле (33))  $y_1$  есть вектор последних  $u$  значений временного ряда:

$$y_1 = (y_{N-u+1}, y_{N-u+2}, \dots, y_N)^T,$$

( $u < m$ , так как  $m = u + v$ , см. «Восстановление пропущенных данных»), а  $y_2$  – вектор прогнозируемых значений:

$$y_2 = (y_{N+1}, y_{N+2}, \dots, y_{N+v})^T.$$

Из вышеизложенного (см. раздел «Проблема относительно малой выборки») следует, что в качестве априорной оценки  $\sigma^2$  необходимо использовать оценку (16). Подлинный смысл прогнозирования заключается не столько в вычислении значений  $\hat{y}_2$ , сколько в построении достаточно узкой доверительной полосы (см. выше) для компонент  $y_2$ . При этом в формулах (35) и (36)  $\eta = m - k - 1$ , а вместо  $1/n$  будет

стоять значение  $1/m$ , если равные между собой компоненты (см. раздел «Проблема относительно малой выборки») вектора  $\bar{y}$  вычислять как среднее по последним  $m$  значениям временного ряда, или  $1/u$ , если вычислять среднее по последним  $u$  значениям (по компонентам), что вполне допустимо. В любом случае окончательный выбор параметров схемы прогноза определяется неравенством (36).

### Заключение

В результате проведенного анализа существующих методов решения задач на основе главных компонент и предлагаемой модификации методов, можно сформулировать следующие основные итоги.

1. Получены оценки дисперсии погрешности регрессии на главные компоненты для случаев большой и малой (относительно размерности задачи) выборки и доказана несмещенность этих оценок. Полученные оценки являются важной деталью в схемах восстановления пропущенных данных и прогноза нестационарных рядов, предложенных автором настоящей работы. При этом несмещенность оценок служит непременным условием в построении доверительных интервалов для восстановленных или прогнозируемых значений (см. далее).

2. Теоретически обоснованно уточнены ранее известные оценки минимального риска, которые также использованы в перечисленных выше практических задачах.

3. Введен в рассмотрение коэффициент структурного подобия и теоретически обоснована статистика для проверки гипотезы о равенстве этого коэффициента нулю.

4. Предложены схемы восстановления пропущенных данных и прогноза нестационарных рядов. Указаны критерии применимости и доверительные интервалы для восстанавливаемых или прогнозируемых значений.

В заключение отметим, что в построении статистических моделей нам приходится выбирать, какие элементы модели наделять свойством статистической устойчивости и включить в модель, а какие нет. От адекватности этого выбора во многом зависит успех применения построенных моделей на прак-

тике, возможно даже в большей степени, чем от точности применяемых формул. В настоящей работе предполагается, что базис главных компонент как раз и является наиболее статистически устойчивой частью модели.

Автор настоящей работы надеется, что полученные здесь оценки и представленные решения задач найдут практическое применение в исследованиях широкого спектра предметных областей.

#### СПИСОК ЛИТЕРАТУРЫ

1. **Pearson K.** On lines and planes of closest fit to systems of points in space // The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science. 1901. Ser. 6. Vol. 2. No. 11. Pp. 559–572.
2. **Karhunen K.** Über lineare Methoden in der Wahrscheinlichkeitsrechnung // Ann. Acad. Sci. Fennicae. Ser. A. I. Math.-Phys. 1947. Vol. 1947. No. 37. Pp. 1–79.
3. **Лоев М.** Теория вероятностей. М.: Изд-во иностр. лит-ры, 1962. 719 с.
4. **Broomhead D.S., King G.P.** Extracting qualitative dynamics from experimental data // Physica. D: Nonlinear Phenomena. 1986. Vol. 20. No. 2–3. Pp. 217–236.
5. **Broomhead D.S., King G.P.** On the qualitative analysis of experimental dynamical systems // Nonlinear Phenomena and Chaos (Malvern physics series). 1<sup>st</sup> edition. Sarkar S. (Ed.). Bristol: CRC Press, 1986. Pp. 113–144.
6. **Ghil M., Vautard R.** Interdecadal oscillations and the warming trend in global temperature time series // Nature. 1991. Vol. 350. No. 6316. Pp. 324–327.
7. **Дженкинс Г., Ваттс Д.** Спектральный анализ и его приложения. Вып.2. М.: Мир, 1972. 287 с.
8. **Пичугин Ю.А.** Итерационный анализ сингулярного спектра в оценке естественных циклическостей метеорологических наблюдений // Метеорология и гидрология. 2001. № 10. С. 34–39.
9. Главные компоненты временных рядов: метод «Гусеница». Под ред. Д.Л. Данилова и А.А. Жиглявского. СПб.: Пресском, 1997. 308 с.
10. The transform and data compression handbook. Eds. K. Rao, P. Yip. Boca Raton (USA): CRC Press, 2001.
11. **Muresan D.D., Parks T.W.** Adaptive principal components and image denoising // Proceedings of IEEE International Conference on Image Processing (ICIP). 14–17 Sept. 2003. Vol. 1. Pp. I-101– I-104.
12. **Пичугин Ю.А.** О классификации летних режимов погоды в Санкт-Петербурге // Метеорология и гидрология. 2000. № 5. С. 31–39.
13. **Bartlett M.S.** The effect of standardization on a  $\chi^2$  approximation in factor analysis // Biometrika. 1951. Vol. 38. No. 3–4. Pp. 337–344.
14. **Bartlett M.S.** A note on the multiplying factor for various 2 approximations // J. Roy. Statist. Soc. 1954. Vol. B16. Pp. 296–298.
15. **Лоули Д., Максвелл А.** Факторный анализ как статистический метод. М.: Мир, 1967. 144 с.
16. **Айвазян С.А., Бухштабер В.М., Енюков И.С., Мешалкин Л.Д.** Прикладная статистика. Классификация и снижение размерности. М.: Финансы и статистика, 1989. 607 с.
17. **Jackson D.** Stopping rules in principal components analysis: A comparison of heuristical and statistical approaches // Ecology. 1993. Vol. 74. No. 8. Pp. 2204–2214.
18. **Себер Дж.** Линейный регрессионный анализ. М.: Мир, 1980. 456 с.
19. **Пичугин Ю.А.** К проблеме статистического контроля данных наблюдений за приземной температурой на отдаленных станциях // Метеорология и гидрология. 2000. № 10. С. 18–24.
20. **Пичугин Ю.А.** Экологический мониторинг и методы многомерной математической статистики // Астраханский вестник экологического образования. 2012. № 2. С. 101–105.
21. **Айвазян С.А., Бухштабер В.М., Енюков И.С., Мешалкин Л.Д.** Прикладная статистика. Исследование зависимостей. М.: Финансы и статистика, 1985. 487 с.
22. **Пичугин Ю.А.** Учет сезонных эффектов в задачах прогноза и контроля данных о приземной температуре воздуха // Метеорология и гидрология, 1996. № 4. С. 52–64.
23. **Михальчук А.С., Пичугин Ю.А.** Дисперсионный анализ погрешностей технологических процессов микроэлектроники // Моделирование и ситуационное управление качеством сложных систем. Сб. докл. СПб.: ГУАП, 2017. С. 35–38.
24. **Пичугин Ю.А.** Естественные составляющие годового хода приземной температуры воздуха // Метеорология и гидрология. 1994. № 12. С. 34–43.
25. **Пичугин Ю.А., Малафеев О.А.** Оптимизация и прогноз в динамической модели управления портфелем ценных бумаг // Матер. секц. заседаний симп. «Нобелевские лауреаты по экономике и российские экономические школы».

СПб., 2003 г. СПб.: СПбГУ, 2003. С. 183–185.

26. **Пичугин Ю.А.** Главные компоненты многомерных временных рядов: анализ и прогноз

// Сб. докл. XIII Междунар. научн. конф. по мягким вычислениям. СПб., 2010 г. Т. 1. СПб.: СПбГЭТУ «ЛЭТИ», 2010. С. 160–163.

*Статья поступила в редакцию 26.03.2018, принята к публикации 21.06.2018.*

#### СВЕДЕНИЯ ОБ АВТОРЕ

**ПИЧУГИН Юрий Александрович** – доктор физико-математических наук, профессор Института инноватики и базовой магистерской подготовки Санкт-Петербургского государственного университета аэрокосмического приборостроения, Санкт-Петербург, Российская Федерация.

190000, Российская Федерация, г. Санкт-Петербург, Большая Морская ул., 61.

yury-pichugin@mail.ru

#### REFERENCES

- [1] **K. Pearson**, On lines and planes of closest fit to systems of points in space, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, Ser. 6. 2 (11) (1901) 559–572.
- [2] **K. Karhunen**, Über lineare Methoden in der Wahrscheinlichkeitsrechnung, Ann. Acad. Sci. Fennicae, Ser. A, I. Math.-Phys. 1947 (37) (1947) 1–79.
- [3] **M. Loève**, Probability theory, Vol. II, 4<sup>th</sup> ed., Graduate texts in mathematics, Springer-Verlag. 46 (1978).
- [4] **D.S. Broomhead, G.P. King**, Extracting qualitative dynamics from experimental data, Physica. D: Nonlinear Phenomena. 20 (2–3) (1986) 217–236.
- [5] **D.S. Broomhead, G.P. King**, On the qualitative analysis of experimental dynamical systems, Nonlinear Phenomena and Chaos, S. Sarkar. (Ed.), CRC Press, Bristol (1986) 113–144.
- [6] **M. Ghil, R. Vautard**, Interdecadal oscillations and the warming trend in global temperature time series, Nature. 350 (6316) (1991) 324–327.
- [7] **G.M. Jenrins, D.G. Watts**, Spectral analysis and its applications, Holden-Day, San Francisco – Cambridge – London – Amsterdam, 1969.
- [8] **Yu.A. Pichugin**, Iterative singular-spectrum analysis in estimating natural cyclicities in meteorological observation data, Meteorology and Hydrology. 10 (2001) 34–39.
- [9] Glavnyye komponenty vremennykh ryadov: metod “Gusenitsa” [Principal components of time series: Caterpillar method], D.L. Danilov, A.A. Zhiglyavskiy (Eds.), SPbSU, St. Petersburg, 1997.
- [10] The transform and data compression handbook, K. Rao, P. Yip (Eds.), CRC Press LLC, Boca Raton, USA, 2001.
- [11] **D.D. Muresan, T.W. Parks**, Adaptive principal components and image denoising, Proceedings of IEEE International Conference on Image Processing (ICIP), 14–17 Sept. 1 (2003) I-101–I-104.
- [12] **Yu.A. Pichugin**, Classification of summer weather regions in St. Petersburg, Meteorology and Hydrology. 5 (2000) 31–39.
- [13] **M.S. Bartlett**, The effect of standardization on a  $\chi^2$  approximation in factor analysis, Biometrika. 38 (3–4) (1951) 337–344.
- [14] **M.S. Bartlett**, A note on the multiplying factor for various 2 approximations, J. Roy. Statist. Soc. B16 (1954) 296–298.
- [15] **D.N. Lawley, A.E. Maxwell**, Factor analysis as a statistical method, Butterworths, London, 1963.
- [16] **S.A. Ayzvazyan, V.M. Bukhshtaber, I.S. Enyukov, L.D. Meshalkin**, Prikladnaya statistika. Klassifikatsiya i snizheniye razmernosti [Applied Statistics. Classification and dimension reduction], Finansy i statistika, Moscow, 1989.
- [17] **D. Jackson**, Stopping rules in principal components analysis: A comparison of heuristic and statistical approaches, Ecology. 74 (8) (1993) 2204–2214.
- [18] **G.A.F. Seber**, Linear regression analysis, John Wiley & Sons, New York, London, Sydney, Toronto (1977).
- [19] **Yu.A. Pichugin**, The problem of statistical control of observation data on surface temperature at distant stations, Meteorology and Hydrology. 10 (2000) 18–24.
- [20] **Yu.A. Pichugin**, Ekologicheskii monitoring i metody mnogomernoy matematicheskoy statistiki [Environmental quality monitoring and multivariate mathematical statistics], Astrakhanskiy vestnik ekologicheskogo obrazovaniya. (2) (2012) 101–105.
- [21] **S.A. Ayzvazyan, V.M. Bukhshtaber, I.S. Enyukov, L.D. Meshalkin**, Prikladnaya statistika. Issledovaniye zavisimostey [Applied Statistics. Relation studies]. Finansy i statistika, Moscow, 1985.
- [22] **Yu.A. Pichugin**, Consideration of seasonal effects in problem of SAT forecasting and data control, Meteorology and Hydrology. 4 (1996) 52–64.
- [23] **A.S. Mikhalechuk, Yu.A. Pichugin**, Dispersionnyy analiz pogreshnostey tekhnologicheskikh



protsestov mikroelektroniki [The variance analysis of errors in the microelectronics technological processes], In collection of papers: Modelirovaniye i situatsionnoye upravleniye kachestvom slozhnykh sistem: sbornik dokladov [Simulation and quality control of complicated systems], SUAI, St. Petersburg (2017) 35–38.

[24] **Yu.A. Pichugin**, Empirical components of annual march of surface temperature, *Meteorology and Hydrology*. 12 (1994) 34–43.

[25] **Yu.A. Pichugin, O.A. Malafeyev**, Optimizatsiya i prognoz v dinamicheskoy modeli upravleniya portfelem tsennykh bumag [Optimization and prediction in the dynamic model of investment

portfolio governance], In: Materialy sektionnykh zasedaniy simpoziuma «Nobelevskiye laureaty po ekonomike i rossiyskiye ekonomicheskiye shkoly» [In: Proceedings of symp. “Nobel Prize winners in economics and Russian economic schools of sciences”], SPbSU, St. Petersburg (2003) 183–185.

[28] **Yu.A. Pichugin**, Glavnyye komponenty mnogomernykh vremennykh ryadov: analiz i prognoz [Principal components of multivariate time series: Analysis and prediction], In: Collection of papers of “The 13th International Youth Scientific Conf. on Soft Computing”, Vol. 1, The 1<sup>st</sup> Electrotechnical University «LETI», St. Petersburg (2010) 160–163.

*Received 26.03.2018, accepted 21.06.2018.*

#### THE AUTHOR

**PICHUGIN Yury A.**

*Saint-Petersburg State University of Aerospace Instrumentation*

61 Bolshaya Morskaya St., St. Petersburg, 190000, Russian Federation

yury-pichugin@mail.ru