

Copyright © 2016 by Academic Publishing House *Researcher*

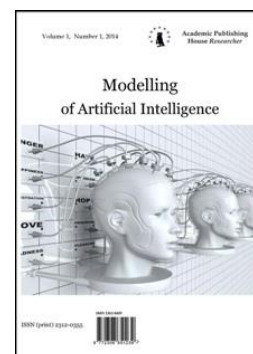
Published in the Russian Federation  
Modeling of Artificial Intelligence  
Has been issued since 2014.

ISSN: 2312-0355

E-ISSN: 2413-7200

Vol. 9, Is. 1, pp. 24-32, 2016

DOI: 10.13187/mai.2016.9.24

[www.ejournal11.com](http://www.ejournal11.com)

UDC 004.93

## Nonlinear Principal Component Analysis Approach to Pattern Recognition

Grigoriy I. Belyavskiy <sup>a, \*</sup>, Evgeniy V. Puchkov <sup>b</sup><sup>a</sup> Southern Federal University, Russian Federation<sup>b</sup> Don State Technical University, Russian Federation

### Abstract

The nonlinear principal component analysis approach to pattern recognition is considered. The attention is given to the nonlinear principal component, which is the sum of subspaces, and the nonlinear principal component, which is realised by the associative neuron network with nonlinear activation functions of output neurons.

**Keywords:** covariation matrix, spectral decomposition, hough transform.

### 1. Введение

Нелинейный метод главных компонент также как и линейный предназначен для извлечения скрытой информации из многомерных наблюдений. Метод базируется на том, что наблюдаемый вектор  $x \in R^N$  с коррелированными компонентами может быть представлен в виде:  $x = F(\alpha) + \varepsilon$ . В этом представлении случайный вектор  $\alpha \in R^m$  с некоррелированными компонентами и с дисперсиями компонент равными единице, вектор  $\varepsilon \in R^n$  - некоррелированный вектор погрешностей, независящий от вектора  $\alpha$ , и одинаковыми дисперсиями компонент. Для линейного случая вектор наблюдений представляется следующим образом:  $x = U\alpha + \varepsilon$ . Матрица  $U$  размера  $(n \times m)$  - матрица с ортогональными столбцами. Для ковариационной матрицы вектора  $x$  -  $C_x$  справедливо равенство:  $C_x = MF(\alpha)F^T(\alpha) + \sigma^2 E$ ,  $E$  - единичная матрица. Для линейного случая соответствующее равенство имеет вид:  $C_x = UU^T + \sigma^2 E$ . Заменяв в последнем равенстве ковариационную матрицу  $C_x$  на выборочную ковариационную матрицу и применив спектральную норму:  $\|A\| = \sup_{\|x\|=1} \|Ax\|$ , получим задачу для определения матрицы  $U$  и дисперсии шума  $\sigma^2$ :  $\inf_{U, \sigma^2} \|C_x^s - UU^T - \sigma^2 E\|$ , при условии: столбцы матрицы  $U$  -

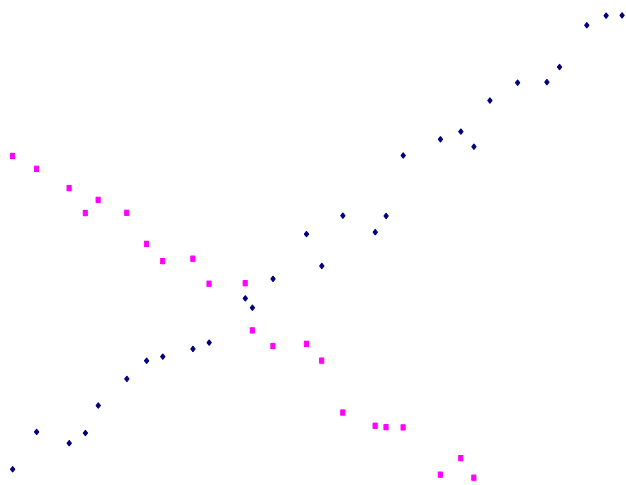
\* Corresponding author

E-mail addresses: [beliavsky@hotmail.com](mailto:beliavsky@hotmail.com) (G.I. Belyavskiy), [puchkoff@gmail.com](mailto:puchkoff@gmail.com) (E.V. Puchkov)

ортогональны. Решение этой задачи выглядит следующим образом: пусть  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$  - спектр выборочной ковариационной матрицы, столбцы матрицы  $U$  -  $u_i, i=1, \dots, m$  - собственные векторы выборочной ковариационной матрицы, нормированные следующим образом:  $(u_i, u_i) = \lambda_i - \sigma^2$ , и  $\sigma^2 = \frac{\lambda_{m+1} + \lambda_m}{2}$ . Значение спектральной нормы для этого

решения  $\|C_x^s - UU^T - \sigma^2 E\| = \sqrt{\frac{\lambda_{m+1} - \lambda_n}{2}}$ . Выражение  $\sqrt{\frac{\lambda_{m+1} - \lambda_n}{2}}$  оценивает качество

представления ковариационной матрицы в линейном случае. Основной смысл метода главных компонент заключается в качественном представлении ковариационной матрицы при как можно наименьшей размерности вектора скрытых факторов  $\alpha$ . В связи с этим рассмотрим пример в пространстве  $R^2$ . Выборка представлена на рис.1.



**Рис. 1.** Выборка с явно выраженными линейными кластерами.

Собственные числа выборочной ковариационной матрицы приблизительно равны. Этот факт означает, что никакая прямая линия не аппроксимирует выборку с приемлемой степенью точности.

## 2. Результаты

*Линейные главные компоненты.* Использование метода главных компонент и факторного анализа в распознавании образов имеет давнюю историю, см., например, работы (Белявский, 1975; Watanabe, 1974). В настоящее время интерес к использованию метода не ослабевает (Chen and Suter, 2006; Nara et al., 2003). Основная идея метода заключается в том, что решающее правило строится на основе вычисления расстояния между образцом, подлежащим распознаванию, и подпространством, представляющим класс. Допустим, что измерением образца является вектор  $x \in R^n$ , а подпространство определяется следующим образом

$$L^m = \left\{ y \in R^n : y = a + \sum_{i=1}^m \alpha_i u_i, a, \alpha \in R^m, u_i \in R^n, (u_i, u_j) = \delta_{i,j} \right\}. \quad \text{Здесь использованы}$$

обозначения:  $(\cdot, \cdot)$  - скалярное произведение и  $\delta_{i,j}$  - символ Кронекера. Расстояние до подпространства определяется следующим образом:  $d(x, L^m) = \|x - \bar{x}\|$ , где

$$\bar{x} = a + \sum_{i=1}^m (x - a, u_i) u_i. \quad \text{Заменим вектор } x \text{ вектором } x - a, \text{ и, не нарушая общности, будем}$$

считать, что  $a = 0$ . Предыдущая формула приобретает более простой вид:  $\bar{x} = \sum_{i=1}^m (x, u_i) u_i$ .

Эту формулу можно представить в матричной форме:

$$\bar{x} = U(U^T x), \quad (1)$$

где матрица  $U$  состоит из  $m$  ортономированных столбцов, то есть произведение  $U^T U$  - единичная матрица. Произведение  $U(U^T x)$  проекция вектора  $x$ . Задача обучения заключается в выборе матрицы  $U$  таким образом, чтобы минимизировать средний по обучающей выборке квадрат расстояния до подпространства

$$\min_U \sum \|x_i - U U^T x_i\|^2. \quad (2)$$

Одно из оптимальных решений задачи (2) матрица  $U$ , столбцы которой первые  $m$  нормированных собственных векторов матрицы  $C = \sum x_i x_i^T$ . Различные алгоритмы обучения связаны с различными способами вычисления  $m$ -первых собственных векторов матрицы  $C$ . В работе Крамера (Kramer, 1991) для вычисления по формуле (1) предложена ассоциативная нейронная сеть, в которой три слоя: входной, выходной и скрытый. Число нейронов входного и выходного слоя совпадает с размерностью пространства признаков -  $n$ , число нейронов скрытого слоя совпадает с размерностью подпространства -  $m$ . Функции активации нейронов скрытого и выходного слоев тождественные функции. Матрица  $U^T$  - матрица весов связей между нейронами входного слоя и нейронами скрытого слоя,  $U$  - матрица весов связей между нейронами скрытого слоя и выходного слоя. Для настройки весов могут быть использованы разнообразные алгоритмы обучения нейронной сети с критерием обучения:  $\sum \|x_i - \bar{x}_i\|^2$ . Для того, чтобы при обучении нейронной сети не отслеживать условие ортономированности столбцов матрицы, формулу (1) следует заменить на более общую формулу:

$$\bar{x} = A(B^T x), \quad (3)$$

в которой матрица  $B^T$  размерности  $(m \times n)$ - матрица входных весов, матрица  $A = B(B^T B)^{-1}$  - матрица выходных весов.

**Статистический метод главных компонент.** Еще в ранних работах, см., например (Белявский, 1975), по методу главных компонент при определении главного подпространства рассматривалась статистическая модель. В этих работах предполагалось, что плотность многомерного закона распределения образцов класса выражается следующим образом:

$$p(x) = \int_{R^m} p(x|\alpha) p(\alpha) d\alpha. \quad (4)$$

Задача распознавания заключается в вычислении значения плотности  $p(x)$  для распознаваемого образца. Прделав простые выкладки для нормальных законов распределения

$$p(x|\alpha) = (2\pi\sigma_1)^{-n/2} \exp\left[-\frac{1}{2\sigma_1^2}(x - U\alpha, x - U\alpha)\right],$$

$p(\alpha) = (2\pi\sigma_2^2)^{-n/2} \exp\left[-\frac{1}{2\sigma_2^2}(\alpha, \alpha)\right]$ , получим, что  $p(x) = \varphi(\bar{d}^2(x, L_m))$ , где

$$\bar{d}^2(x, L_m) = \frac{1}{\sigma_1^2} \left\| x - \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} UU^T x \right\|^2 + \frac{\sigma_2^2}{(\sigma_1^2 + \sigma_2^2)^2} \|U^T x\|^2, \quad \varphi(\cdot) - \text{убывающая функция.}$$

Величину  $\bar{d}(x, L_m)$  можно рассматривать как расстояние до подпространства. Другие варианты расстояний до подпространства приведены в работе (Белявский, 1975). Отметим, что  $\lim_{\sigma_2^2 \rightarrow \infty} \bar{d}(x, L_m) = \frac{1}{\sigma_1} d(x, L_m)$ . Применение метода максимального правдоподобия при

обучении приводит к вычислению первых  $m$  собственных векторов матрицы  $C = \sum x_i x_i^T$ .

**Нелинейный метод главных компонент.** Нелинейным главным компонентам посвящен ряд современных исследований [6-9], что говорит о том, что нелинейный метод главных компонент находится в тренде исследований по интеллектуальной обработке данных. В наиболее ранней работе по нелинейному методу главных компонент (Белявский, 1975) было предложено использовать объединение линейных подпространств:  $L = \bigcup L_i^{m_i}$  в качестве главного многообразия. Расстояние между образцом и многообразием вычисляется по формулам:

$$d(x, L) = \min d(x, L_i^{m_i}) \text{ или } d(x, L) = \sum \beta_i d(x, L_i^{m_i}), 0 < \beta_i, \sum \beta_i = 1. \quad (5)$$

Статистический метод в этом случае базируется на распределении с плотностью:

$$p(x) = \int_{R^m} p(x|\alpha) p(\alpha) d\alpha, \quad (6)$$

$$p(x|\alpha) = \sum \beta_i (2\pi\sigma_i^2)^{-n/2} \exp\left[-\frac{1}{2\sigma_i^2}(x - U^i \alpha, x - U^i \alpha)\right],$$

$$p(\alpha) = (2\pi\sigma^2)^{-n/2} \exp\left[-\frac{1}{2\sigma^2}(\alpha, \alpha)\right].$$

Непосредственное вычисление приводит к плотности:

$$p(x) = \text{const} \sum \beta_i \exp\left(-\frac{1}{2\sigma_i^2} \left\| x - \frac{\sigma^2}{\sigma_i^2 + \sigma^2} U^i (U^i)^T x \right\|^2 - \frac{\sigma^2}{2(\sigma_i^2 + \sigma^2)^2} \|(U^i)^T x\|^2\right) \quad (7)$$

В (Белявский, 1975) предложен алгоритм обучения, использующий алгоритм самообучения М.И. Шлезингера (Шлезингер, 1968), который заключается в повторении обучения и распознавания. Для алгоритма Шлезингера необходимо априорное знание числа подпространств и их размерности. Кроме этого алгоритм не гарантирует сходимости к глобальному минимуму среднего расстояния до многообразия. Приятное исключение возникает при объединении двух линейных подпространств. В связи с этим справедлива теорема, которую мы приводим без доказательства.

**Теорема.** Пусть  $p(x)$  вычисляется по формуле (4), в которой

$$p(x|\alpha) = \beta (2\pi\sigma_1^2)^{-n/2} \exp\left[-\frac{1}{2\sigma_1^2}(x - U^1 \alpha, x - U^1 \alpha)\right] +$$

$$(1 - \beta) (2\pi\sigma_2^2)^{-n/2} \exp\left[-\frac{1}{2\sigma_2^2}(x - U^2 \alpha, x - U^2 \alpha)\right], \quad 0 < \beta < 1 \text{ и}$$

$p(\alpha) = (2\pi\sigma^2)^{-n/2} \exp\left[-\frac{1}{2\sigma^2}(\alpha, \alpha)\right]$ . Алгоритм М.И. Шлезингера сходится к

$\beta^*, \sigma^*, \sigma_1^*, \sigma_2^*, (U^1)^*, (U^2)^* \in \text{Aarg} \min_{\beta, \sigma, \sigma_1, \sigma_2, U^1, U^2} d_{av}(\beta, \sigma, \sigma_1, \sigma_2, U^1, U^2)$ , где

$$d_{av}(\beta, \sigma, \sigma_1, \sigma_2, U^1, U^2) = \sum \ln \left[ \beta \exp \left[ -\frac{\left\| x_i - \frac{\sigma^2}{\sigma_1^2 + \sigma^2} U^1 (U^1)^T x_i \right\|^2}{2\sigma_1^2} - \frac{\sigma^2 \left\| (U^1)^T x_i \right\|^2}{2(\sigma_1^2 + \sigma^2)^2} \right] + \right. \\ \left. (1 - \beta) \exp \left[ -\frac{\left\| x_i - \frac{\sigma^2}{\sigma_2^2 + \sigma^2} U^2 (U^2)^T x_i \right\|^2}{2\sigma_2^2} - \frac{\sigma^2 \left\| (U^2)^T x_i \right\|^2}{2(\sigma_2^2 + \sigma^2)^2} \right] \right] - \text{среднее расстояние до}$$

многообразия.

Особое место в нелинейном методе главных компонент занимает задача спрямления кривой на плоскости. В научной литературе этой задаче уделено достаточно внимания ([Hastie and Stuetzle, 1989](#); [Kegl et al., 2000](#)). Методы заключаются в повторении двух этапов: кластеризации и вычисления главных прямых. Методы отличаются друг от друга различными способами кластеризации данных, чтобы затем для каждого кластера построить главную прямую и, на основе семейства главных прямых произвести новую кластеризацию. Таким образом, главное многообразие состоит из сегментов главных прямых по одной для каждого кластера. В принципе, каждый из методов близок к уже упомянутой процедуре, использующей алгоритм самообучения Шлезингера.

С задачей кластеризации успешно справляется преобразование Хафа ([Duda and Hart, 1972](#)), которое каждой точке  $(x, y)$  ставит в соответствие синусоиду согласно формуле:

$$\rho = x \cos \varphi + y \sin \varphi, \quad -\pi \leq \varphi \leq \pi, \quad \rho \geq 0. \quad (8)$$

Прежде чем привести алгоритм вычисления главного многообразия рассмотрим связь между главной компонентой и преобразованием Хафа. Рассмотрим выборку точек на плоскости:  $\{(x_i, y_i)\}$ , вычислим параметры  $(\rho, \varphi)$  таким образом, чтобы средний квадрат расстояния от выборки до прямой был минимальным. Чтобы не загромождать выкладки предположим, что выборка нормирована таким образом, чтобы выборочная дисперсия равнялась единице. Непосредственно проверяется, что оптимальные значения выражаются следующим образом:

$$\rho^* = \bar{x} \cos \varphi^* + \bar{y} \sin \varphi^*, \quad \varphi^* = \arctg \frac{u_2}{u_1}. \quad (9)$$

В формуле (9) вектор  $u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$  - собственный вектор ковариационной матрицы

$C = \begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix}$ , соответствующий меньшему собственному значению  $\lambda_{\min} = 1 - r$ ,  $\bar{x}$  и  $\bar{y}$  выборочные средние соответствующих координат. Максимальное собственное значение матрицы  $C$   $\lambda_{\max} = 1 + r$  и отношение  $\frac{1-r}{1+r}$  может быть использовано в качестве теста на

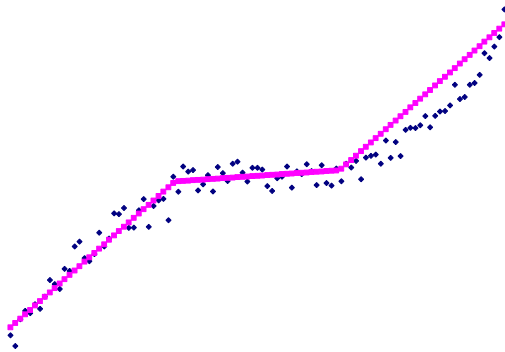
линейность. Применив преобразование Хафа к выборке  $V$ , получим матрицу  $H = (h_{i,j})$ , где

$h_{i,j} = \#\{(x, y) : x \cos \varphi_j + y \sin \varphi_j \in [\rho_{i-1}, \rho_i]\}$ . Здесь  $\rho_i = i\Delta\rho, \varphi_j = -\pi + j\Delta\varphi$ . На основе преобразования Хафа вычисляется главная компонента, составленная из отрезков прямых, в результате применения следующего алгоритма.

Алгоритм Хафа.

1. Полагаем  $k = 1$
2. Определяем  $k$  первых локальных экстремумов матрицы  $H$ .
3. При помощи прямых, соответствующих экстремумам разбиваем выборку на  $k$  кластеров.
4. Для каждого кластера вычисляем прямую, используя формулы (9) и проверяем тест на линейность.
5. Если для каждого кластера тест на линейность выполняется, то переходим к 7
6. Полагаем  $k = k + 1$  и переходим к 2.
7. Остановка.

Пример работы алгоритма приведен на [рис. 1](#).



**Рис. 2.** Главная компонента из отрезков прямых линий. Алгоритм Хафа.

С помощью преобразования Хафа можно вычислять главные компоненты, составленные из более сложных кривых, например, из дуг окружностей. Для окружностей преобразование Хафа следует применять для пары точек на плоскости, а именно, пара точек  $P_1$  и  $P_2$  отображается в прямую линию параметрического пространства:

$(P_1 - P_2, D) = \frac{\|P_1\|^2 - \|P_2\|^2}{2}$ . В остальном алгоритм работает так же, как и в случае прямых линий.

Вернемся к ассоциативной нейронной сети и рассмотрим сеть с нелинейными функциями активации внутренних нейронов. Преобразование входного вектора  $x$  в выходной вектор  $\bar{x}$  будет выглядеть следующим образом:

$$\bar{x} = UF(U^T x), \tag{10}$$

где  $F$  - векторнозначная функция активации нейронов промежуточного слоя:  $F(U^T x) = (F_i((u_i, x)))$ . При обучении преследуются две цели: определение матрицы  $U$  и функции активации  $F$ . Критерий обучения средний квадрат отклонения выходных значений от входных значений.

Обучение при фиксированной матрице  $U$ . Является справедливым следующее очевидное утверждение.

*Утверждение.* Для преобразования (10) оптимальной функцией активации нейронов внутреннего слоя является тождественная функция:  $F_i(z) = z$ .

Из утверждения следует, что преобразование (10) приобретает вид:  $\bar{x} = UU^T x$ . Для этого преобразования задача выбора оптимальной матрицы  $U$  рассмотрена ранее.

Рассмотрим ассоциативную сеть с нелинейной функцией активации выходных нейронов. Реализуемое преобразование будет иметь вид:

$$\bar{x} = F(AB^T x) \quad (11)$$

Пусть  $F_i \in L_2([0, M])$  и семейство функций  $(H_i)_{i \geq 1}$  образуют базис в  $L_2$ . Рассмотрим аппроксимацию  $F_i: \bar{F}_i = \sum_{j=1}^k \alpha_{i,j} H_j$ . Целью обучения является вычисление и матрицы  $B$ .

Обучение заключается в повторении двух процедур до тех пор, пока не будет получено удовлетворительное приближение в среднеквадратическом смысле.

Первая процедура предназначена для вычисления коэффициентов  $(\alpha_{i,j})$  при фиксированной матрице  $B$ . Она заключается в решении задач регрессионного типа:

$$\min_{\alpha_i} \sum_j \left( x_{i,j} - \sum_l \alpha_{i,l} H_l \left( (a_i^T, B^T x_j) \right) \right)^2. \quad (12)$$

В (12) использованы следующие обозначения:  $x_{i,j}$  -  $i$ -я координата вектора  $x_j$ ,  $a_i^T$  -  $i$  строка матрицы  $A$ . Решение задач получается в результате решения семейства систем линейных алгебраических уравнений с матрицами коэффициентов

$$G_i = \left( \sum_j H_l \left( (a_i^T, B^T x_j) \right) H_r \left( (a_i^T, B^T x_j) \right) \right) \text{ и правыми частями } g_i = \left( \sum_j H_l \left( (a_i^T, B^T x_j) \right) x_{i,j} \right).$$

Вторая процедура обучения нейронной сети заключается в выборе весовых коэффициентов нейронов входного и выходного слоев при фиксированной функции активации нейронов выходного слоя, которая является стандартной задачей обучения нейронной сети. На каждой итерации происходит уменьшение критерия обучения, критерий ограничен снизу, поэтому алгоритм сходится. Остановка происходит при достижении требуемой точности либо при наступлении стабилизации.

### 3. Заключение

В связи с нелинейным методом главных компонент основное внимание в работе уделено двум подходам. Первый подход использует кластеризацию выборки, причем для каждого кластера вычисляется аппроксимирующее линейное подпространство. При втором подходе используется ассоциативная нейронная сеть с нелинейными функциями активации нейронов выходного слоя. Причем предполагается, что функции активации заранее неизвестны и настраиваются, также как и матрицы весов нейронов скрытого и выходного слоев в процессе обучения нейронной сети. Единственное предположение, которое используется при этом, это то, что функции активации хорошо аппроксимируются линейными комбинациями конечного набора выбранных заранее базисных функций. Каждый из этих методов обладает общим недостатком. А именно, отсутствует гарантия сходимости к оптимальному главному многообразию.

### 4. Благодарности

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 14-01-00579 а.

### Литература

**Белявский, 1975** - Белявский Г.И. Метод линейных подпространств в распознавании образов, сб. Распознавание образов, ИК АНУССР, Киев, 1975.

**Шлезингер, 1968** - Шлезингер М.И. Взаимосвязь обучения и самообучения в распознавании образов. Кибернетика, 2, 1968.

[Chen and Suter, 2006](#) - Chen, P. and Suter, D An analysis of linear subspace approaches for computer vision and pattern recognition. // International Journal of Computer Vision. 68 (1), 83–106, 2006.

[Duda and Hart, 1972](#) - Duda, R. O. and P. E. Hart, Use of the Hough Transformation to Detect Lines and Curves in Pictures. *Comm. ACM, Vol. 15*, pp. 11–15, 1972.

[Hastie and Stuetzle, 1989](#) - Hastie, T. and Stuetzle, W. Principal curves. Journal of the American Statistical Association 84 (406), 502–516, 1989.

[Kegl et al., 2000](#) - Kegl, B., Krzyzak, A., Linder, T., and Zeger, K. Learning and design of principal curves IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (3), 281–297, 2000.

[Kim et al., 2002](#) - Kim, K.I., Jung, K., and Kim, H. J. Face recognition using kernel principal component analysis. IEEE Signal Processing Letters, 9 (2), 40–42, 2002.

[Kramer, 1991](#) - Kramer, M.A. Nonlinear principal component analysis using autoassociative neural networks. *AICHE Journal*, 37 (3), 233–243, 1991.

[Nara et al., 2003](#) - Nara, Y., Jianming Y., and Suematsu, Y. Face recognition using improved principal component analysis. In: Proceedings of 2003 International Symposium on Micromechatronics and Human Science (IEEE Cat. No.03TH8717), 77–82, 2003.

[Shawe-Taylor and Cristianini, 2004](#) - Shawe-Taylor, J. and Cristianini, N. Kernel Methods for Pattern Analysis. Cambridge University Press, West Nyack, NY, 2004.

[Watanabe, 1974](#) - Watanabe S. A subspace representation of classes in pattern recognition, Seventh Prague conference on information theory, statistical decision function and processes, Prague, 1974.

#### References:

[Belyavskii, 1975](#) - Belyavskii G.I. Metod lineinykh podprostranstv v raspoznavanii obrazov [The method of linear subspaces in pattern recognition], sb. Raspoznavanie obrazov, IK ANUSSR, Kiev, 1975.

[Shlezinger, 1968](#) - Shlezinger M.I. Vzaimosvyaz' obucheniya i samoobucheniya v raspoznavanii obrazov [The relationship of learning and self-learning in pattern recognition]. Kibernetika, 2, 1968.

[Chen and Suter, 2006](#) - Chen, P. and Suter, D An analysis of linear subspace approaches for computer vision and pattern recognition. // International Journal of Computer Vision. 68 (1), 83–106, 2006.

[Duda and Hart, 1972](#) - Duda, R. O. and P. E. Hart, Use of the Hough Transformation to Detect Lines and Curves in Pictures. *Comm. ACM, Vol. 15*, pp. 11–15, 1972.

[Hastie and Stuetzle, 1989](#) - Hastie, T. and Stuetzle, W. Principal curves. Journal of the American Statistical Association 84 (406), 502–516, 1989.

[Kegl et al., 2000](#) - Kegl, B., Krzyzak, A., Linder, T., and Zeger, K. Learning and design of principal curves IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (3), 281–297, 2000.

[Kim et al., 2002](#) - Kim, K.I., Jung, K., and Kim, H. J. Face recognition using kernel principal component analysis. IEEE Signal Processing Letters, 9 (2), 40–42, 2002.

[Kramer, 1991](#) - Kramer, M.A. Nonlinear principal component analysis using autoassociative neural networks. *AICHE Journal*, 37 (3), 233–243, 1991.

[Nara et al., 2003](#) - Nara, Y., Jianming Y., and Suematsu, Y. Face recognition using improved principal component analysis. In: Proceedings of 2003 International Symposium on Micromechatronics and Human Science (IEEE Cat. No.03TH8717), 77–82, 2003.

[Shawe-Taylor and Cristianini, 2004](#) - Shawe-Taylor, J. and Cristianini, N. Kernel Methods for Pattern Analysis. Cambridge University Press, West Nyack, NY, 2004.

[Watanabe, 1974](#) - Watanabe S. A subspace representation of classes in pattern recognition, Seventh Prague conference on information theory, statistical decision function and processes, Prague, 1974.



УДК 004.93

## Метод нелинейных главных компонент в распознавании образов

Григорий Исаакович Белявский <sup>a,\*</sup>, Евгений Владимирович Пучков <sup>b</sup>

<sup>a</sup> Южный федеральный университет, Российская Федерация

<sup>b</sup> Донской государственный технический университет, Российская Федерация

**Аннотация.** Рассматривается нелинейный метод главных компонент применительно к задачам распознавания образов. Внимание уделяется нелинейной главной компоненте, являющейся объединением главных линейных подпространств, и нелинейной главной компоненте, реализуемой ассоциативной нейронной сетью с нелинейными функциями активации нейронов выходного слоя.

**Ключевые слова:** ковариационная матрица, спектральное разложение, преобразование Хафа.

---

\* Корреспондирующий автор

Адреса электронной почты: [beliavsky@hotmail.com](mailto:beliavsky@hotmail.com) (Г.И. Белявский),  
[puchkoff@gmail.com](mailto:puchkoff@gmail.com) (Е.В. Пучков)