Comprehensive Research Journal of Engineering, Technology and Innovation (CRJETI) Vol. 1(1) pp. 001 – 003 November, 2016
Available online https://comprehensiveresearchjournals.org/crjeti

Review

Methods of reduction of information in multidimensional space

Jaroslaw Michalkiewicz

Kwiska 12/5, 54-210 Wrocław, Poland
e-mail:jarekmichal@poczta.onet.pl

Accepted 08 November, 2016

The Author replaces hypersurface with hyperplane in the multidimensional space in order to reduce amount of data in the transmission of information process.

Keywords: multidimansional space, reduction of information

INTRODUCTION

Gathering amount of information cause the necessity of their transformation and classification There also appears a necessity of elimination of a part of this information, which does not carry essential content and reduces unnecessary dimension in mathematical space New branch of knowledge arises named "feature selection" which by means of mathematical methods removes unnecessary information and simplifies model of data transmission. The problem of reduction of data is important in the process of their creation because excess overload of the link lengthens the transmission time. One of selection data technique is PCA, which reduces unnecessary dimension in geometrical space (Belda, 2012). Method of data reduction proposed by Author in the multidimensional space base on replacing hypersurface with hyperplane that leads to data reduction in sending the information. PCA (https: pl.wikipedia.org.wiki analiza glownych skladowych) method is based on an algorithm which:

1. sets mean and standard deviation matrix
2. sets covariance matrix
3. computes eigenvalue covariance matrix
4. reduces number of variables from eigenvalue covariance matrix, and
5. sets eigenvector and data throws on eigenvectors

My method differs from PCA method because it uses a different form of input data.

Mathematical description

In this chapter I replace hypersurface with hyperplane in the multidimensional space. Let M be vector in an n-dimensional space [1] that means

$$M = M(x_1, x_2, ..., x_n) \qquad (1)$$

Let $U_1, U_2, ..., U_n$ be parameters in an n-dimensional space [1]. Let us assume that

$$x_1 = U_1, x_2 = U_2, ..., x_{n-1} = U_{n-1} \qquad (2)$$

$$z = x_n = f(U_1, ..., U_{n-1}) = f(x_1, ..., x_{n-1}) \qquad (3)$$

where $f(x_1, ..., x_{n-1})$ is hypersurface in n-dimensional space. Let us replace hypersurface with hyperplane. In every i elementary hypercube $d_{1,i} < x_1 < e_{1,i}$, $\ldots, d_{n-1,i} < x_{n-1} < e_{n-1,i}$ there is a set of hypersurface and tangent hyperplane. Let us assume in every arbitrary small hypercube is one hyperplane set and a hypersurface tangent to it. We set the vector tangent to a coordinate axis U_1 , U_2, ..., U_{n-1}. Tangent vectors are signed $M_{U_1}, M_{U_2}, ..., M_{U_{n-1}}$ where

$$M_{U_1} = \left[\frac{\partial x_1}{\partial U_1}, \dots . \frac{\partial x_n}{\partial U_1} \right] \quad (4)$$

$$M_{U_{n-1}} = \left[\frac{\partial x_1}{\partial U_1}, \dots . \frac{\partial x_n}{\partial U_{n-1}} \right] \quad (5)$$

Taking under consideration connection between x_i and U_i we receive

$$M_{U_1} = \left[1, 0, \dots, 0, \frac{\partial z}{\partial x_1} \right] \quad (6)$$

$$M_{U_2} = \left[0, 1, \dots, 0, \frac{\partial z}{\partial x_2} \right] \quad (7)$$

$$M_{U_{n-1}} = \left[0, 0, \dots, 1, \frac{\partial z}{\partial x_{n-1}} \right] \quad (8)$$

Vectors $M_{U_1}, \dots, M_{U_{n-1}}$ are linearly independent in n-dimensional space because:

$$\sum_{i=1}^{n-1} \alpha_i M_{U_1} = 0 \Rightarrow \alpha_i = 0, \mathrm{i} = 1, 2, ..., n-1 \quad (9)$$

Because from

$$\sum_{i=1}^{n-1} \alpha_i M_{U_i} = \left[\alpha_1, ..., \alpha_{n-1}, \sum_{i=1}^{n-1} \alpha_i \frac{\partial z}{\partial x_i} \right] = 0 \quad (10)$$

It appears that $\alpha i = 0$ for $i = 1, ..., n-1$ that ends the proof. 16

Vector $(\alpha_1, ..., \alpha_n)$ perpendicular to M_{U_1} , ..., $M_{U_{n-1}}$ sets the hyperplane

$$(x_1 - x_{1,0})\alpha_1 + ... + (x_n - x_{n,0})\alpha_n = 0 \quad (11)$$

tangent to hypersurface z. Vector $(\alpha_1 , ..., \alpha_n)$ (perpendicular to vectors tangent to hypersurface) must fulfil equations

$$\alpha_1 1 + \dots + \alpha_n \frac{\partial z}{\partial x1} = 0 \quad (12)$$

$$\alpha_1 0 + \dots + \alpha_n \frac{\partial z}{\partial x_{n-1}} = 0 \quad (13)$$

In order to solve equations let us assume that $\alpha_n = 1$. Then the vector

$$(\alpha_1, ..., \alpha_n) = \left(- \frac{\partial z}{\partial x_1}, \dots, - \frac{\partial z}{\partial x_{n-1}}, 1\right) \quad (14)$$

Experimental Research

Experimental research was carried out for a 4-domensional space. Examined hypersurface is expressed by the formula.

$$z = f(x_1, ..., x_3) = x_1^2 + x_2^2 + x_3^2 \quad (15)$$

Derivative z after the variables x_i is

$$\frac{\partial z}{\partial x_i} = 2x_i \quad (16)$$

Thus the vector

$$(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = (-2x_1, -2x_2, -2x_3, 1) \quad (17)$$

Sets hyperplane

$$(x_1 - x_{1,0})\alpha_1 + \dots (x_4 - x_{4,0})\alpha 4 = 0 \quad (18)$$

Let us divide segment [0, 1] to 2 parts. We receive 8 hypercubes. Then $x_{i,0} = 0,25 or 0,75$ and $z = x_4$. We receive 8 parts approximating a paraboloid. For $0 \le x_i \le 0,5$ $z_0 = x_{4,0} = 3 * 0,25^2 = \frac{3}{10}$ we have the hyperplane

$$(x_1 - 0,25)(-2 * 0,25) + (x_2 - 0,25)(-2 * 0,25) + (x_3 - 0,25)(-2 * 0,25) + (z - z_0) = 0 \qquad (19)$$

Convergence and computational complexity

Approximation error of the continuous function will aspire to arbitrarily small value by decreasing the dimension hypercube fields. It appears from that variability of continuous function around the hyperplane can be decreased to the arbitrarily small value by decreasing the dimension field. It saves the amount of sending the data by using proposed algorithm. Let us assume that field [0, 1] is divided by 10 parts for 3-dimensional function. We have 10^3 cubes so we need 7x1000 = 7000 numbers. Without the algorithm, when we assume the field divided into 100 parts, we have 100^3 = 1000000 numbers.

CONCLUSION

This paper has the purpose to replace the function of many variables with a hyperplane in an n-dimensional space. Performance of this operation enables data reduction and leads to simplification of a multivariable model of function. Methods of sending the data used now include escessive amount of information that makes the data transmission difficult. Using proposed algorithm simplifies the whole operation and reduces the cost of signal transports. Methods proposed by the Author are simpler than PCA methods. Instead of deleting variables measure data are replaced with hyperplane which contains the most important information data. Technique of data input and output is different. Input data are points in multidimensional space in PCA methods. Input data are an n variable function in my method. In the PCA method we have data in the form of points in reduction of dimension. In my method we have a function based on decreased number of points.

REFERENCE

Finikow S.,P.:Geometria różniczkowa. PWN, Warszawa 1956.
Belda I.:Umysł, maszyny I matematyka. Seria: świat jest matematyczny. RBA.
Fichtenholz G.,M.:Rachunek różniczkowy i całkowy. Tom 1-3. PWN Warszawa 1978.
Gleichgewicht B.:Algebra. PWN Warszawa 1976.
https: pl.wikipedia.org wiki analiza glownych skladowych