

Optimization in object classification

The multisurface method

Monika Mrozek

11 December 2019

Xcyt

Basic assumptions and functions

The main aim of the software named **Xcyt** was expanding the possibilities of the prognosis and diagnosis of breast cancer. The system was invented at the University of Wisconsin in Madison[2]. It combines many techniques: digital image analysis, inductive machine learning, mathematical programming and statistics.

Assumptions and functions

- analysis of cytological features,
- diagnosis and attempt to estimate the probability of malignancy,
- when the sample is cancerous, prediction of recurrence.

Xcyt

Cellular features

Xcyt makes possible to analyse the exact boundaries of the nuclei. It uses curve-fitting program and examines a lot of parameters:

- area,
- radius,
- perimeter,
- symmetry,
- number and size of concavities,
- fractal dimension,
- compactness,
- smoothness.

Multisurface Method-Tree I

Building of a model

Xcyt is closely connected with **MSM-Tree** (*Multisurface Method-Tree*). Its main idea is separation of the points from matrices **A** and **B**. In these circumstances

- **A** _{$m \times n$} is a matrix containing m vectors $x \in \mathbb{R}^n$ with data about patients who were diagnosed with malignant tumour,
- **B** _{$k \times n$} is a matrix containing k vectors $x \in \mathbb{R}^n$ with data about patients who were diagnosed with benign tumour.

Multisurface Method-Tree II

Building of a model

The separating plane is represented by [1]

$$\gamma = x^T \cdot w \in \mathbb{R}^n$$

if and only with

$$Aw \geq e\gamma + e$$

$$Bw \leq e\gamma - e$$

where

- w - normal to separating plane, $w \in \mathbb{R}^n$
- e - vector of ones of appropriate dimension.

Multisurface Method-Tree

Relationship with linear programming

The problem of separating group of points by piecewise-linear surface is equivalent to solving a linear program

$$\min_{w, \gamma, y, z} \frac{e^T y}{m} + \frac{e^T z}{k}$$

subject to

- $Aw \geq e\gamma + e$
- $Bw \leq e\gamma - e$
- $y, z \geq 0$

Figure: Olvi L. Mangasarian, W. Nick Street, William H. Wolberg: *Breast Cancer Diagnosis via Linear Programming*, Mathematical Programming Technical [1]

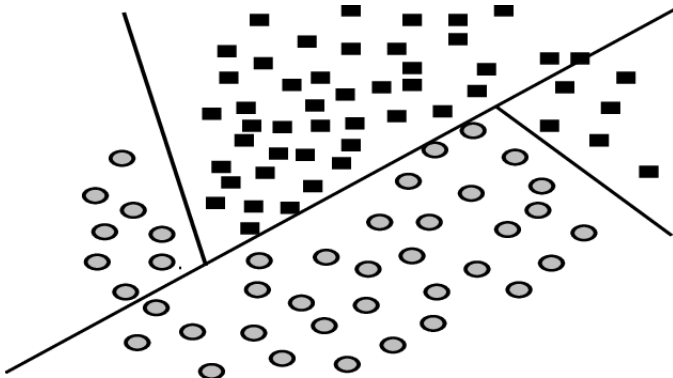


Figure: Separating planes in MSM-T

MSM method has a significant participation in improving breast cancer treatment. The scientists used the diagnosis of breast cytology to demonstrate the applicability of this method to medical diagnosis and decision making.

Figure: *William H. Wolberg, Olvi L. Mangasarian: Multisurface method of pattern separation for medical diagnosis applied to breast cytology [3]*

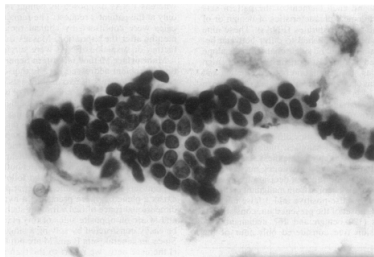


Figure: FNA cytology

Description of procedure I

The experiment was conducted at the University of Wisconsin Hospitals. The scientists analysed 370 samples. Complete classification was obtained after 4 steps.

- $(x_1, x_2, \dots, x_9) \in \mathbb{R}^9$ - the vector of real numbers created for each sample, which contains nine cytological characteristics,
- $(c_{i1}, c_{i2}, \dots, c_{i9}) \in \mathbb{R}^9$ - the normal to the plane, $i = 1, 2, 3, 4$
- a_i, b_i - the distance from the origin to the planes $\in \mathbb{R}^9$

Correct separation was accomplished in 369 of 370 samples (201 benign and 169 malignant).

Description of procedure II

Interpretation of results

For $i = 1, 2, 3$ we have

- $c_{i1}x_1 + c_{i2}x_2 + \dots + c_{i9}x_9 > b_i$ – then tumour is malignant at stage i ,
- $c_{i1}x_1 + c_{i2}x_2 + \dots + c_{i9}x_9 < a_i$ – then tumour is benign at stage i .

Additionally

- $c_{41}x_1 + c_{42}x_2 + \dots + c_{49}x_9 \geq (a_4 + b_4)/2$ – then tumour is malignant at stage 4,
- $c_{41}x_1 + c_{42}x_2 + \dots + c_{49}x_9 < (a_4 + b_4)/2$ – then tumour is benign at stage 4.

Example of decision tree

Figure: Multisurface Method Tree with MATLAB

<http://pages.cs.wisc.edu/~olvi/wmp/msmt.html>

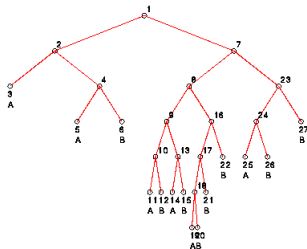


Figure: Graphical illustration of the decision tree for Wisconsin Breast Cancer data

Bibliography



Olvi L. Mangasarian, W. Nick Street, William H. Wolberg: *Breast Cancer Diagnosis via Linear Programming*, Mathematical Programming Technical Report 94-10, Decemeber 19, 1994



W. N. Street: *Xcyt: A System for Remote Cytological Diagnosis and Prognosis of Breast Cancer*, Management Sciences Department, University of Iowa, Iowa City, IA USA



William H. Wolberg, Olvi L. Mangasarian: *Multisurface method of pattern separation for medical diagnosis applied to breast cytology*, Proc. Natl. Acad. Sci. USA, December 1990
<https://pdfs.semanticscholar.org/5053/7a16eb18e2bc4971165258cba7a071f38cb7.pdf>