

Linear Classifiers II

Tomáš Svoboda

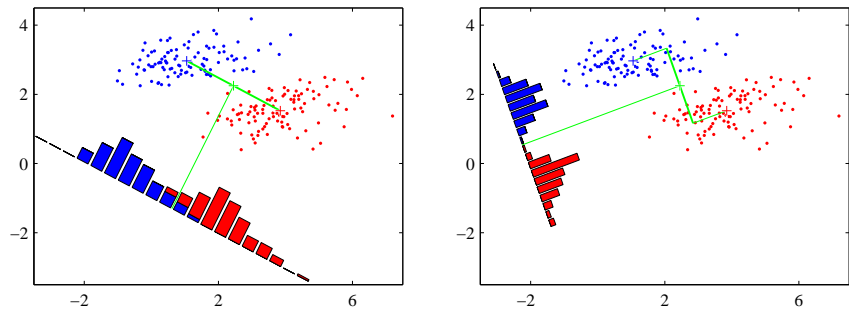
Vision for Robots and Autonomous Systems, Center for Machine Perception
Department of Cybernetics
Faculty of Electrical Engineering, Czech Technical University in Prague

May 22, 2019

Linear Classifiers - supplement lecture

- ▶ Supplement to the lecture about learning Linear Classifiers (perceptron, ...)
- ▶ Better etalons by applying Fischer linear discriminator analysis.
- ▶ LSQ formulation of the learning task.

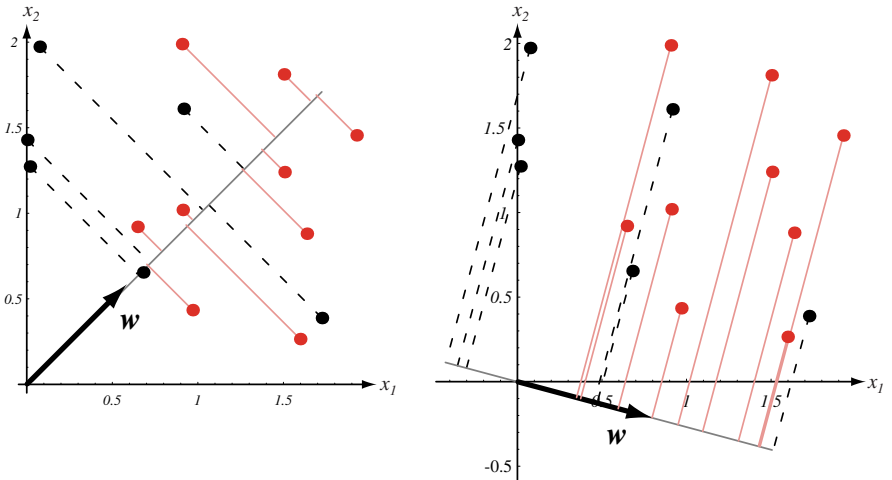
Fischer linear discriminant



- ▶ Dimensionality reduction
- ▶ Maximize distance between means, ...
- ▶ ... and minimize within class variance. (minimize overlap)

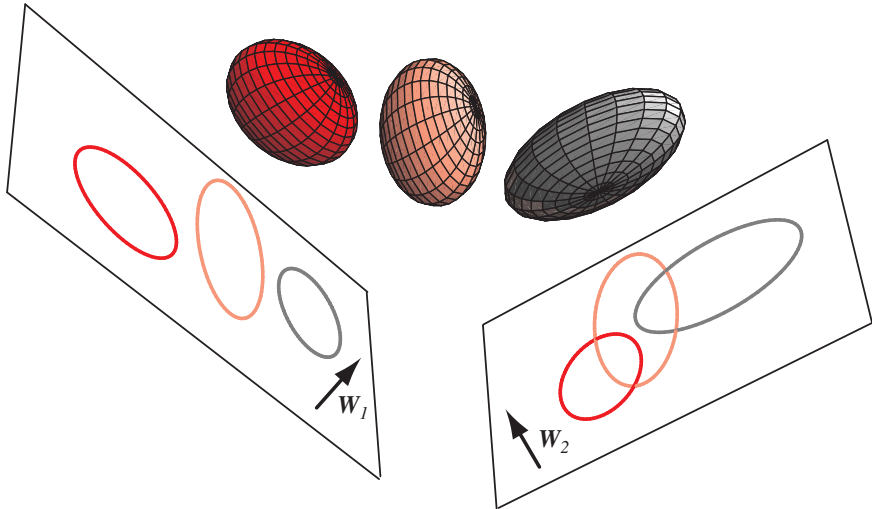
Figures from [1]

Projections to lower dimensions $y = \mathbf{w}^\top \mathbf{x}$



Figures from [2]

Projections to lower dimensions $y = \mathbf{w}^\top \mathbf{x}$



Figures from [2]

Finding the best projection

$$y = \mathbf{w}^\top \mathbf{x}$$

thresholding $y \geq -w_0$ C_1 , otherwise C_2

$$\mathbf{m}_i = \frac{1}{n_i} \sum_{\mathbf{x} \in C_i} \mathbf{x}$$

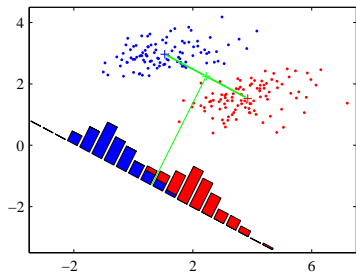
$$\mathbf{m}_2 - \mathbf{m}_1 = \mathbf{w}^\top (\mathbf{m}_2 - \mathbf{m}_1)$$

Within class scatter of projected samples

$$s_i^2 = \sum_{y \in C_i} (y - m_i)^2$$

Fischer criterion:

$$J(\mathbf{w}) = \frac{(m_2 - m_1)^2}{s_1^2 + s_2^2}$$



$$S_i = \sum_{\mathbf{x} \in C_i} (\mathbf{x} - \mathbf{m}_i)(\mathbf{x} - \mathbf{m}_i)^\top$$

$$S_W = S_1 + S_2$$

$$S_B = (\mathbf{m}_2 - \mathbf{m}_1)(\mathbf{m}_2 - \mathbf{m}_1)^\top$$

$$J(\mathbf{w}) = \frac{\mathbf{w}^\top S_B \mathbf{w}}{\mathbf{w}^\top S_W \mathbf{w}}$$

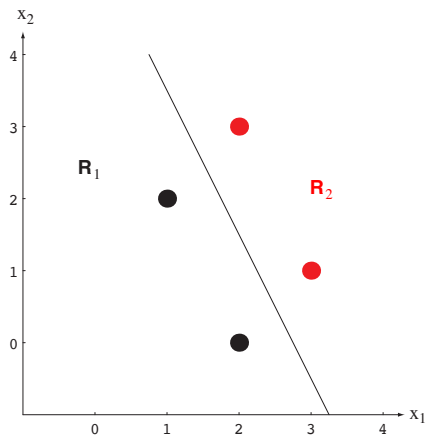
S_B stands for the *between* class scatter matrix

After some derivation on blackboard:

$$\mathbf{w} = S_W^{-1}(\mathbf{m}_2 - \mathbf{m}_1)$$

LSQ approach to linear classification

$$X\mathbf{w} = \mathbf{b}$$



Linear least squares not guaranteed to correctly classify everything on the training set. It's objective function not perfect for classification.
Outliers can shift the decision boundary.

References I

Further reading: Chapter 4 of [1], or chapter 3 and 5 of [2].

[1] Christopher M. Bishop.

Pattern Recognition and Machine Learning.

Springer Science+Business Media, New York, NY, 2006.

PDF freely downloadable.

[2] Richard O. Duda, Peter E. Hart, and David G. Stork.

Pattern Classification.

John Wiley & Sons, 2nd edition, 2001.