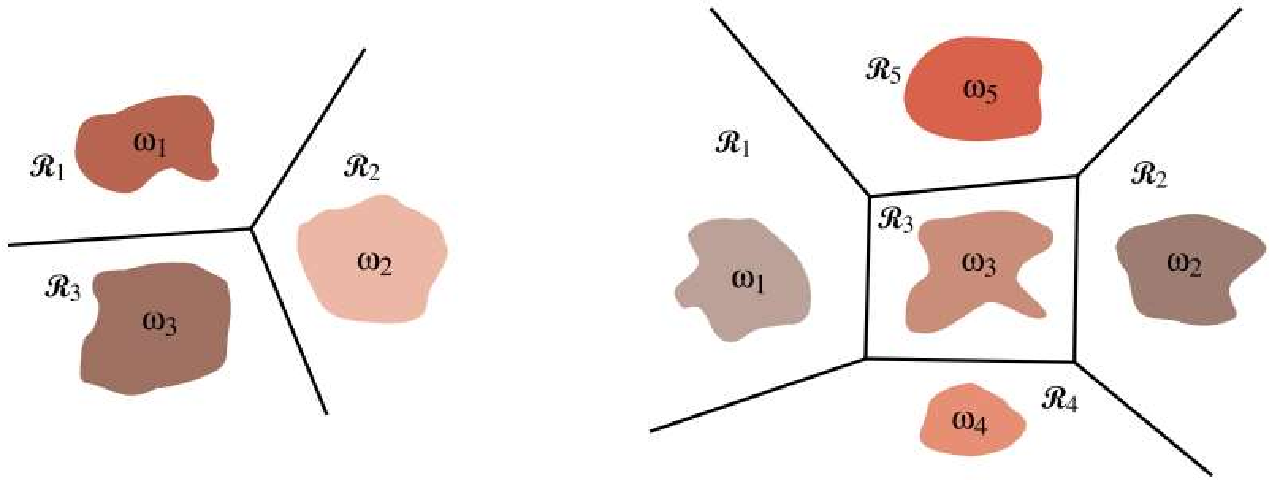
Reconhecimento de Padrões
Discriminantes Lineares

Prof. George von Borries
Departamento de Estatística
Universidade de Brasília

1 - 2024

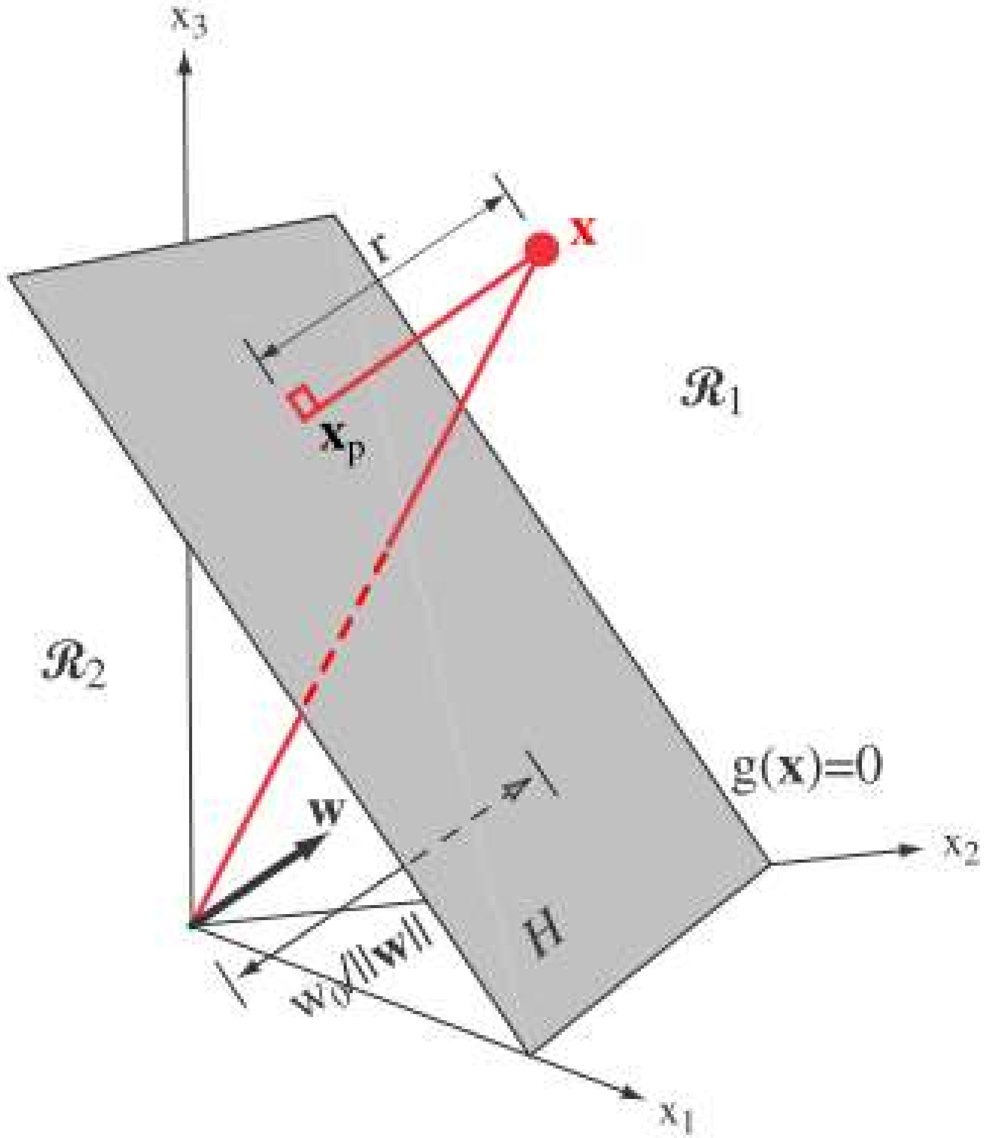
Introdução

- **Discriminantes Lineares** são funções lineares das características (features) em análise.
- Estas funções podem ser *piecewise* lineares, resultando em vários limites de decisão ou ainda aplicadas a várias regiões de classificação.

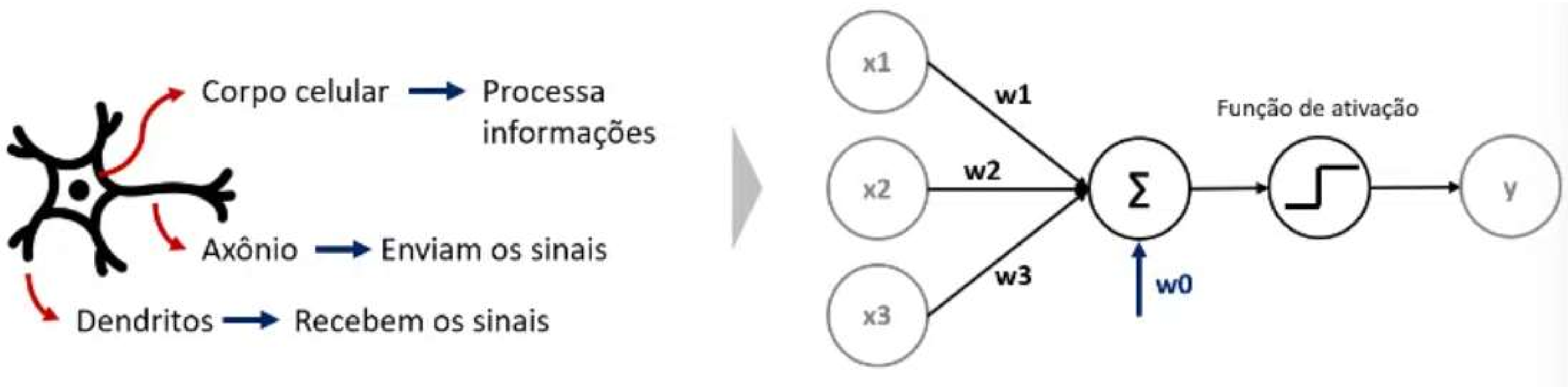


Duda et al. (2001)

- Suposição: limites de decisão são lineares.
  Não existe suposição sobre a distribuição envolvida.



Limite de decisão linear H, em que $g(\mathbf{x}) = \mathbf{w}^\top \mathbf{x} + w_0 = 0$ separa o espaço de características em dois espaços $\mathcal{R}_1$ $(g(\mathbf{x}) > 0)$ e $\mathcal{R}_2$ $(g(\mathbf{x}) < 0)$.
Duda et al. (2001)



(https://www.hashtagtreinamentos.com/o-perceptron-ciencia-de-dados)

Frank Rosenblatt
1928–1969

Rosenblatt's perceptron played an important role in the history of machine learning. Initially, Rosenblatt simulated the perceptron on an IBM 704 computer at Cornell in 1957, but by the early 1960s he had built special-purpose hardware that provided a direct, parallel implementation of perceptron learning. Many of his ideas were encapsulated in "Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms" published in 1962. Rosenblatt's work was criticized by Marvin Minksy, whose objections were published in the book "Perceptrons", co-authored with Seymour Papert. This book was widely misinterpreted at the time as showing that neural networks were fatally flawed and could only learn solutions for linearly separable problems. In fact, it only proved such limitations in the case of single-layer networks such as the perceptron and merely conjectured (incorrectly) that they applied to more general network models. Unfortunately, however, this book contributed to the substantial decline in research funding for neural computing, a situation that was not reversed until the mid-1980s. Today, there are many hundreds, if not thousands, of applications of neural networks in widespread use, with examples in areas such as handwriting recognition and information retrieval being used routinely by millions of people.

(Bishop, 2006)

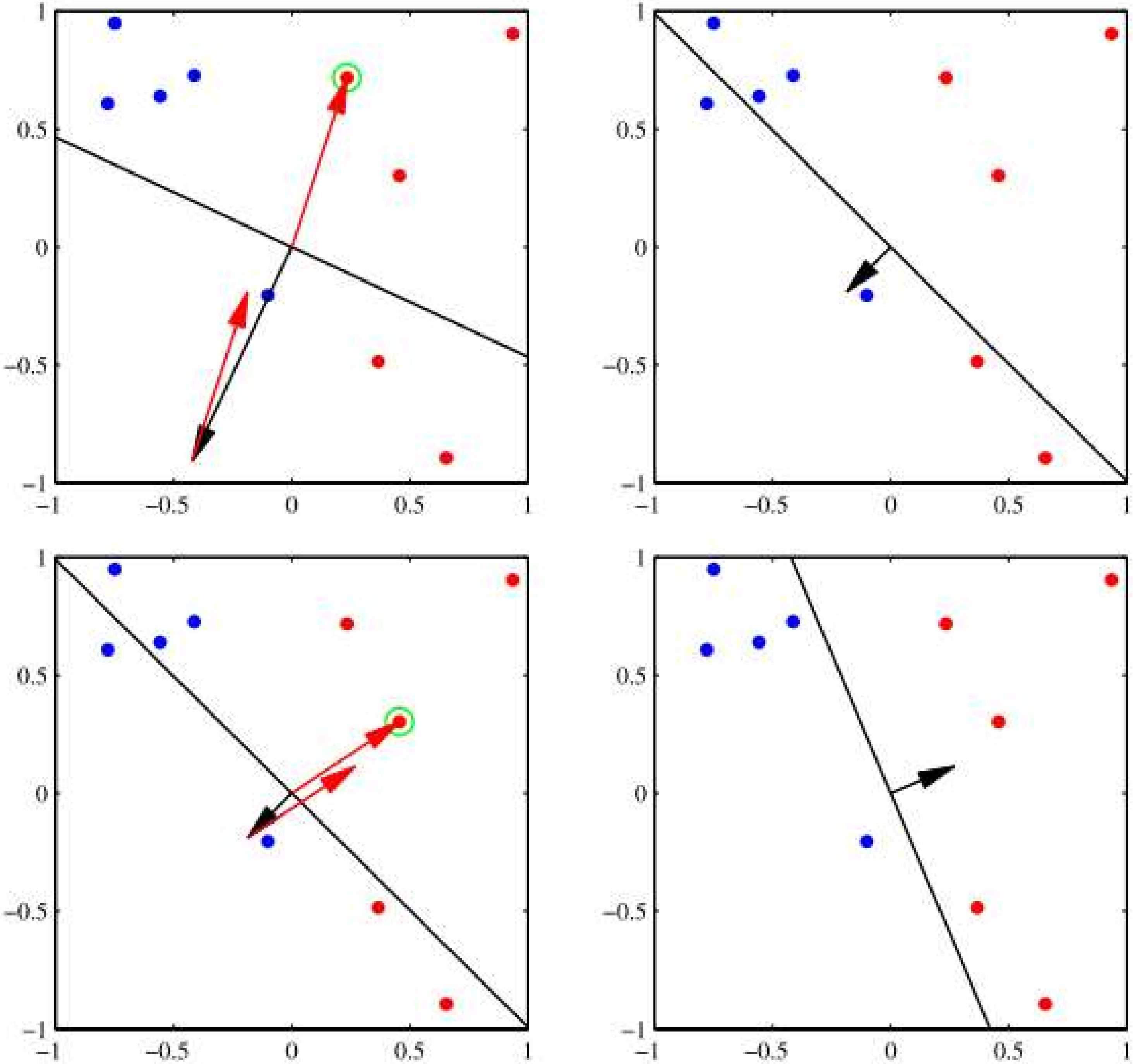
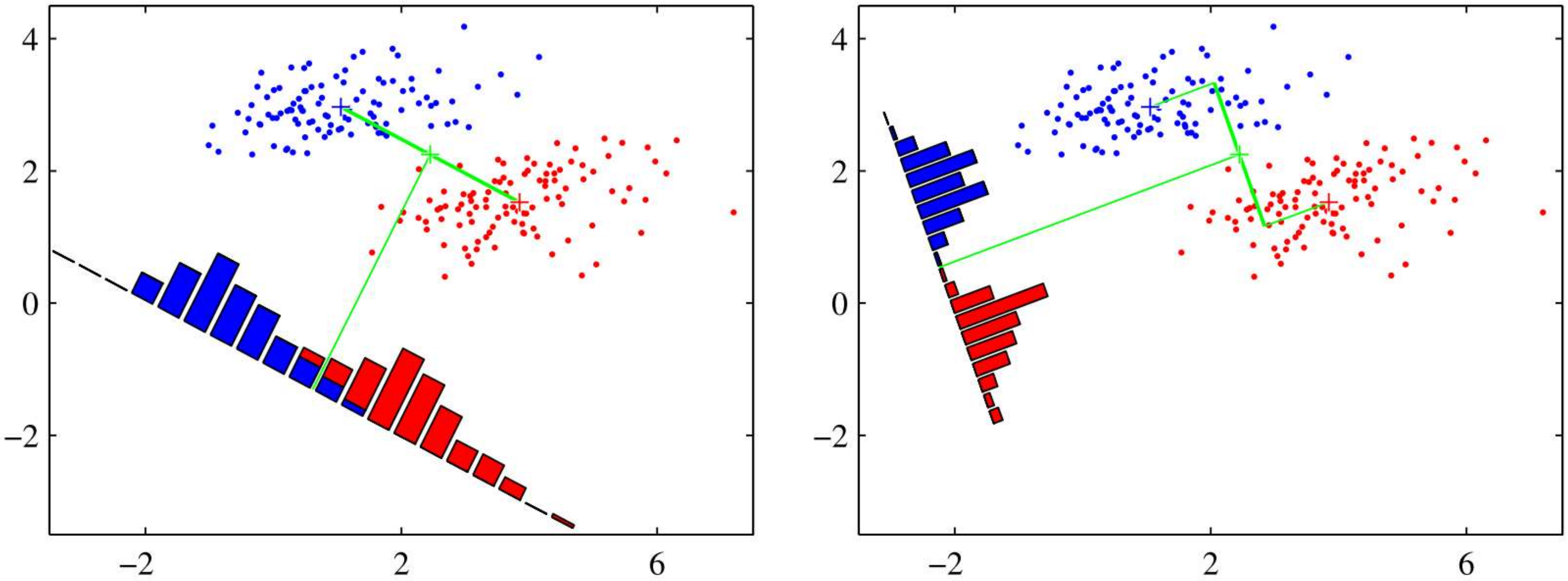
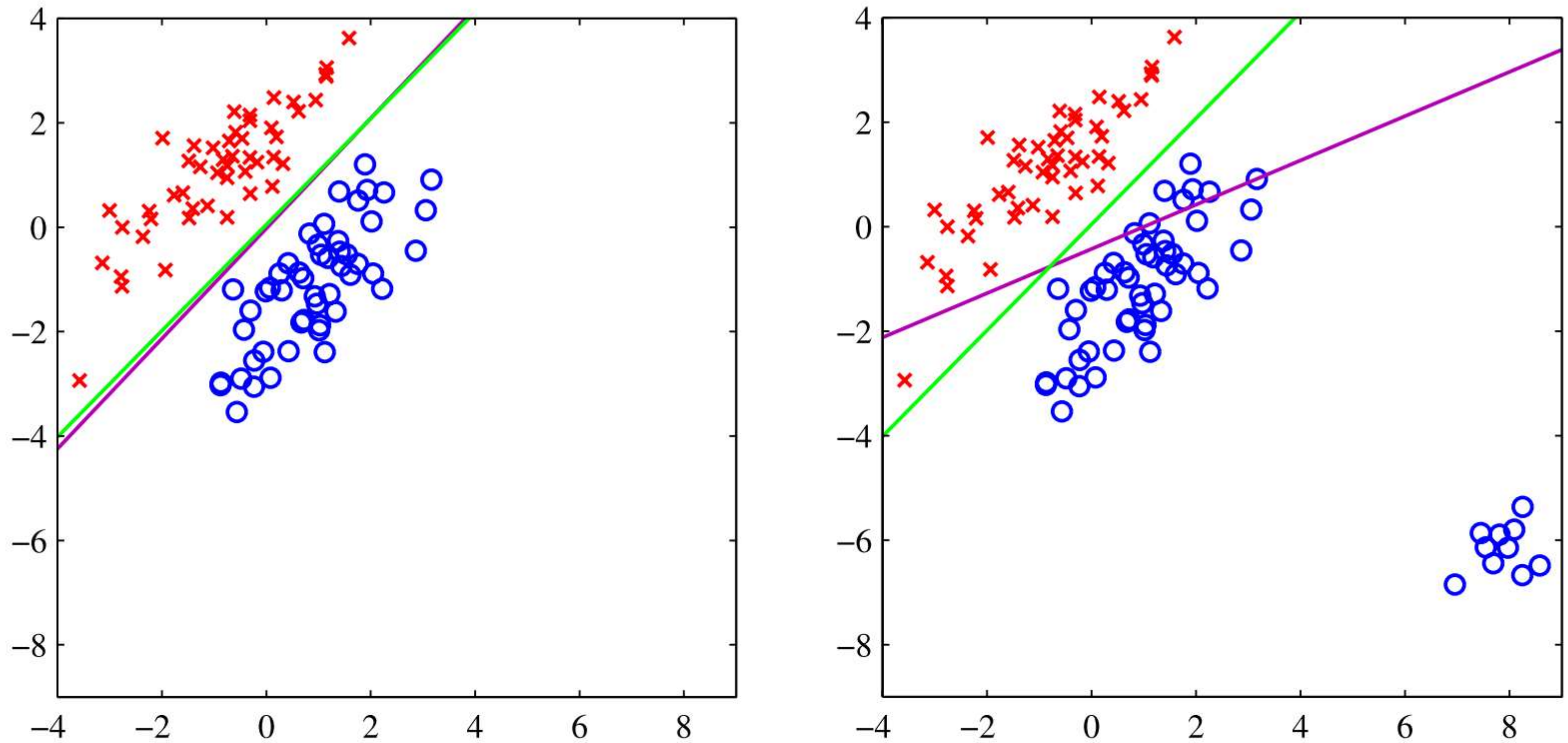

1
0.5
0
−0.5
−1
−1
−0.5
0
0.5
1

Figure 4.7 Illustration of the convergence of the perceptron learning algorithm, showing data points from two classes (red and blue) in a two-dimensional feature space (ϕ_1, ϕ_2). The top left plot shows the initial parameter vector $\mathbf{w}$ shown as a black arrow together with the corresponding decision boundary (black line), in which the arrow points towards the decision region which classified as belonging to the red class. The data point circled in green is misclassified and so its feature vector is added to the current weight vector, giving the new decision boundary shown in the top right plot. The bottom left plot shows the next misclassified point to be considered, indicated by the green circle, and its feature vector is again added to the weight vector giving the decision boundary shown in the bottom right plot for which all data points are correctly classified.

(Bishop, 2006)



Esquerda: projeção com superposição das classes. Direita: projeção segundo critério de Fisher. (Bishop, 2006)



O critério de mínimos quadrados é fortemente influenciado por *outliers*. Linha rosa indica limite obtido por mínimos quadrados e linha verde por regressão logística. (Bishop, 2006)

- Se $t_i = t_1$ (constante) para todo $\mathbf{x}_i \in \omega_1$ e $t_i = t_2$ (constante) para todo $\mathbf{x}_i \in \omega_2$, a solução de mínimos quadrados corresponde ao discriminante de Fisher.
- Observação: usual fazer $t_1 + t_2 \neq 0$.