

FIAP

FIAP GRADUAÇÃO

TDS - FIAP

INTELIGÊNCIA ARTIFICIAL E CHATBOT

Profª . Miguel Bozer da Silva
profmiguel.silva@fiap.com.br



prof. Miguel Bozer da Silva

CLASSIFICADOR BAYESIANO

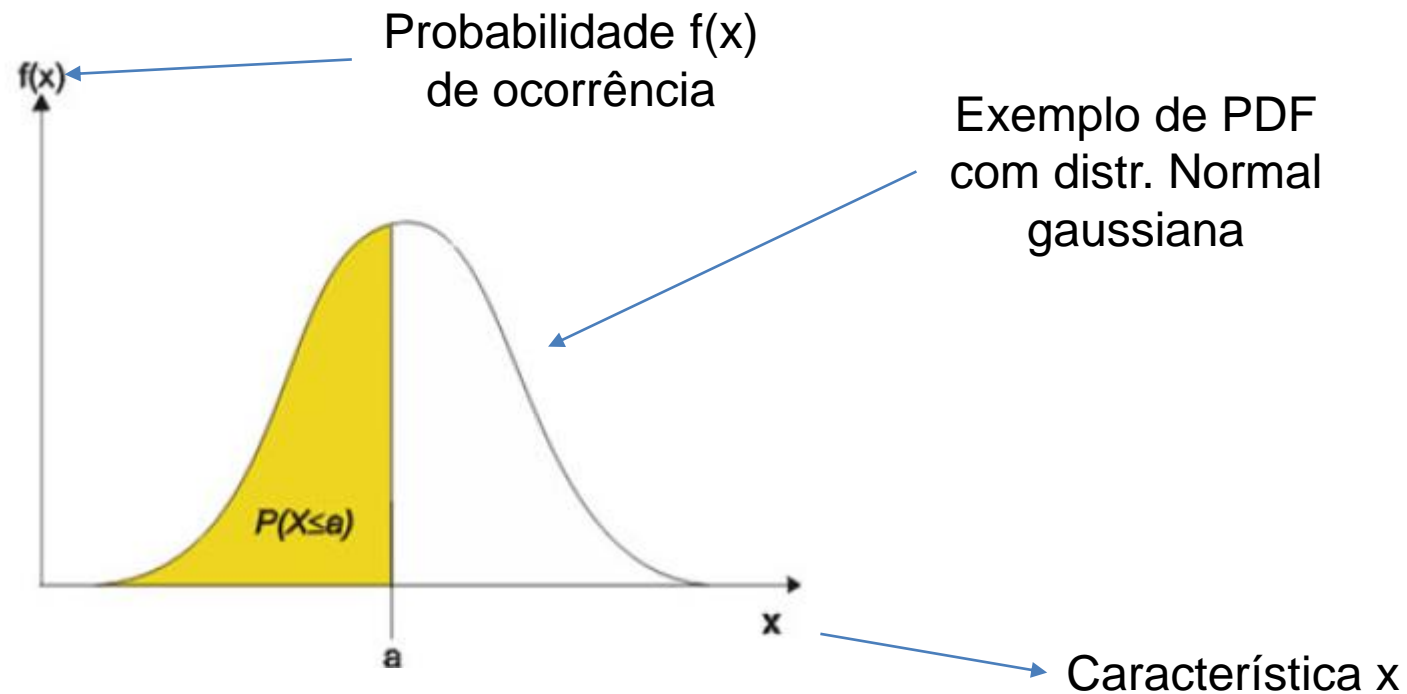
Classificador Bayesiano

- Os classificadores Bayesianos utilizam a regra de Bayes para fazer as classificações. Para isso usamos as Probability density function (PDF's)

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$\mu = \text{média}$

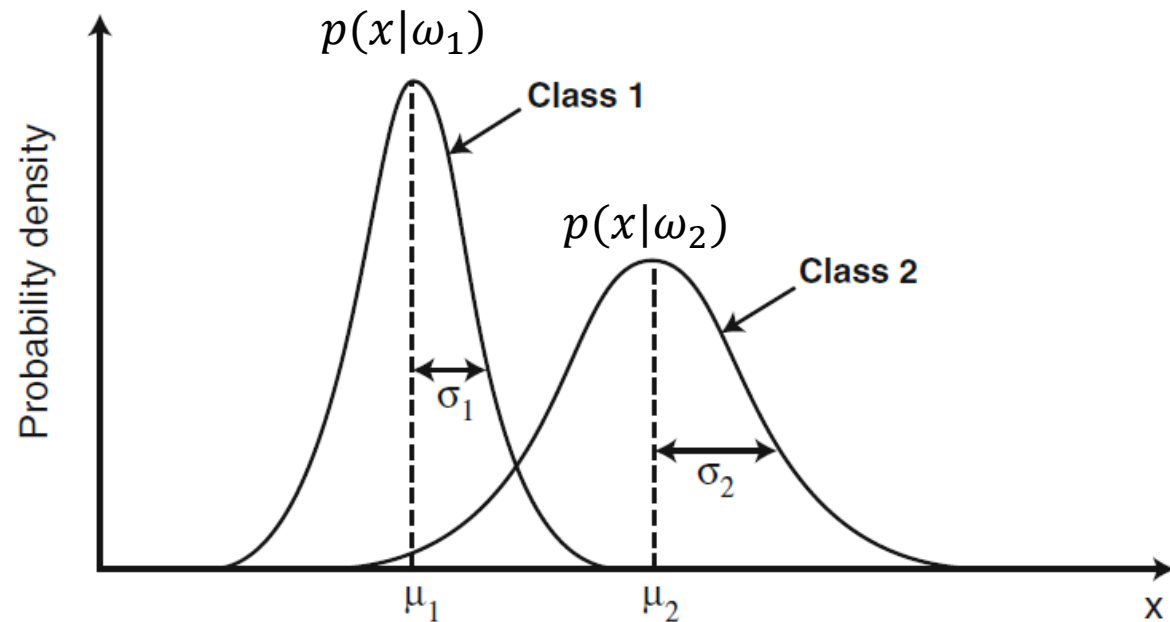
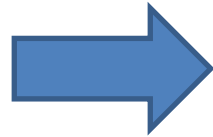
$\sigma^2 = \text{variância dos dados}$



Classificador Bayesiano

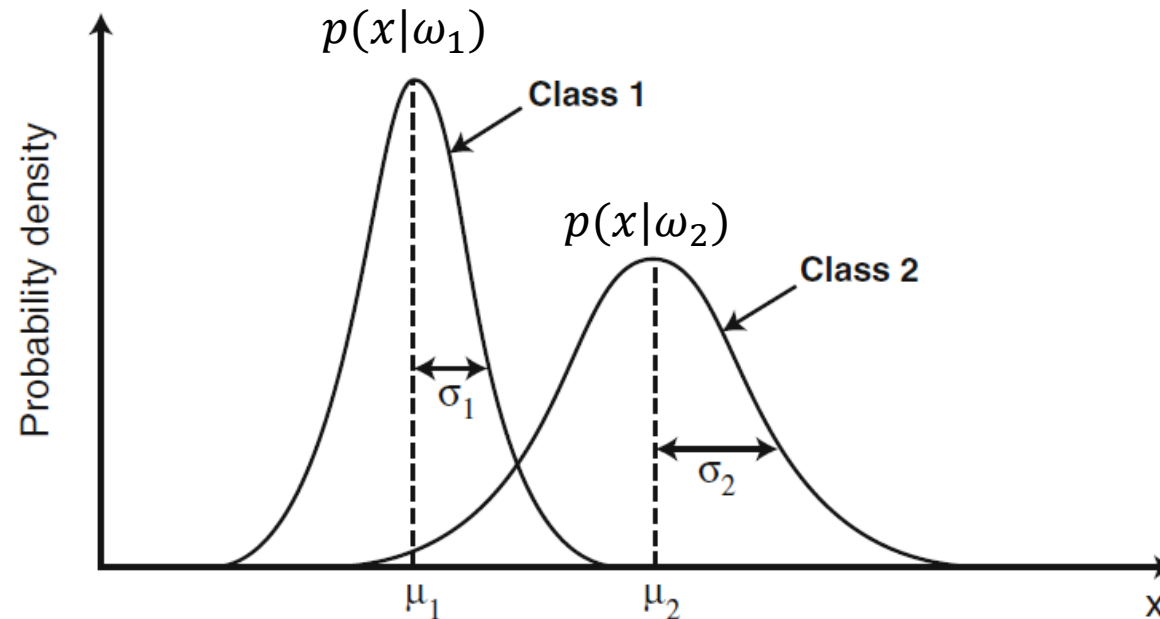
- Para um caso com uma única feature (x) e duas classes (ω_1 e ω_2)
- Com dados de treinamento podemos estimar $p(x|\omega_1)$ e $p(x|\omega_2)$

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



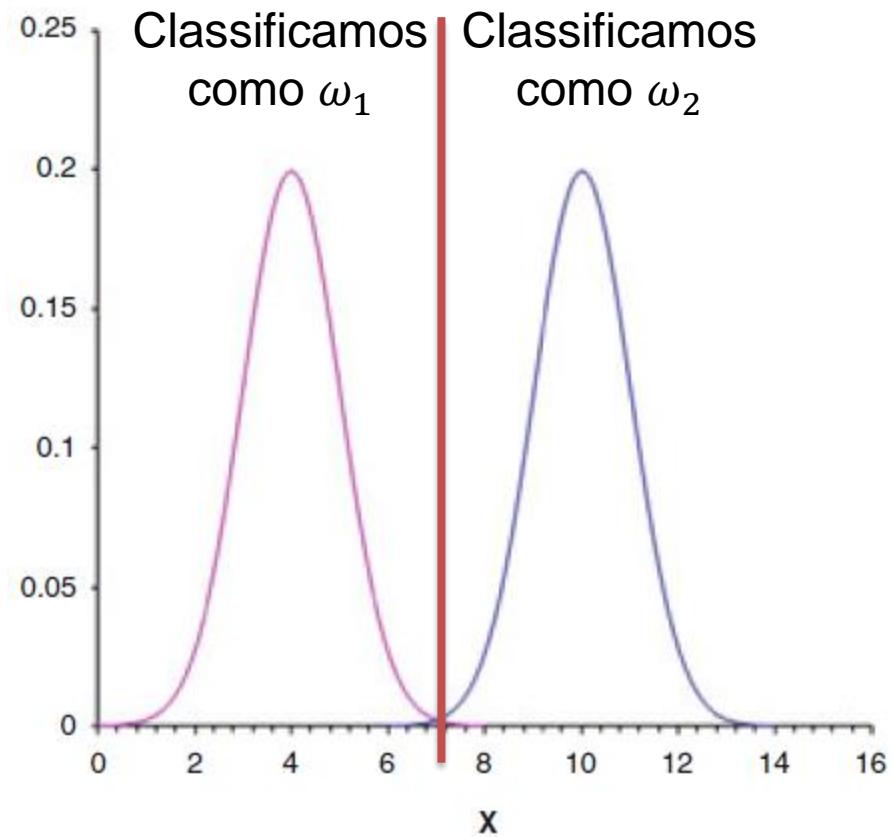
Classificador Bayesiano

- Com as PDF's como podemos classificar novos dados?



Classificador Bayesiano

- Podemos criar uma fronteira de decisão

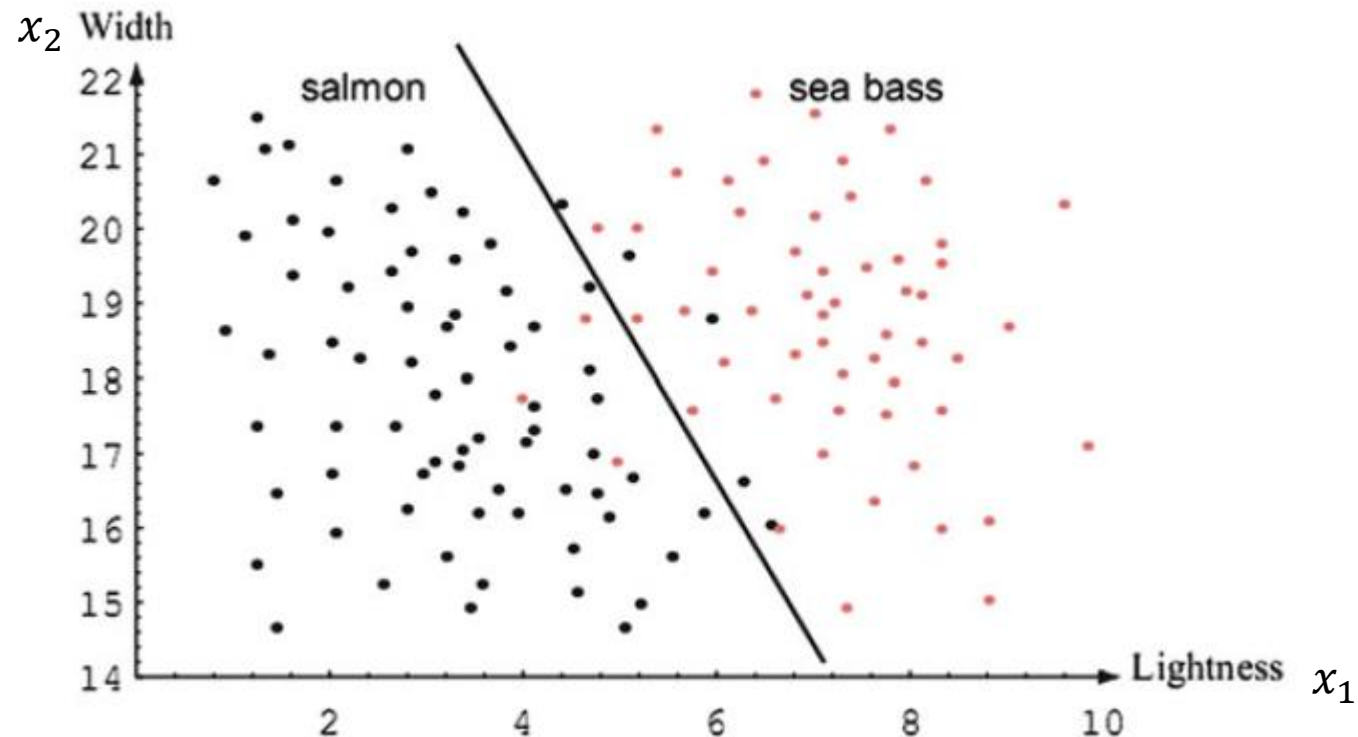


Classificador Bayesiano

- Como proceder quando temos mais features?

Exemplo: Classificação entre Salmão e Robalo:

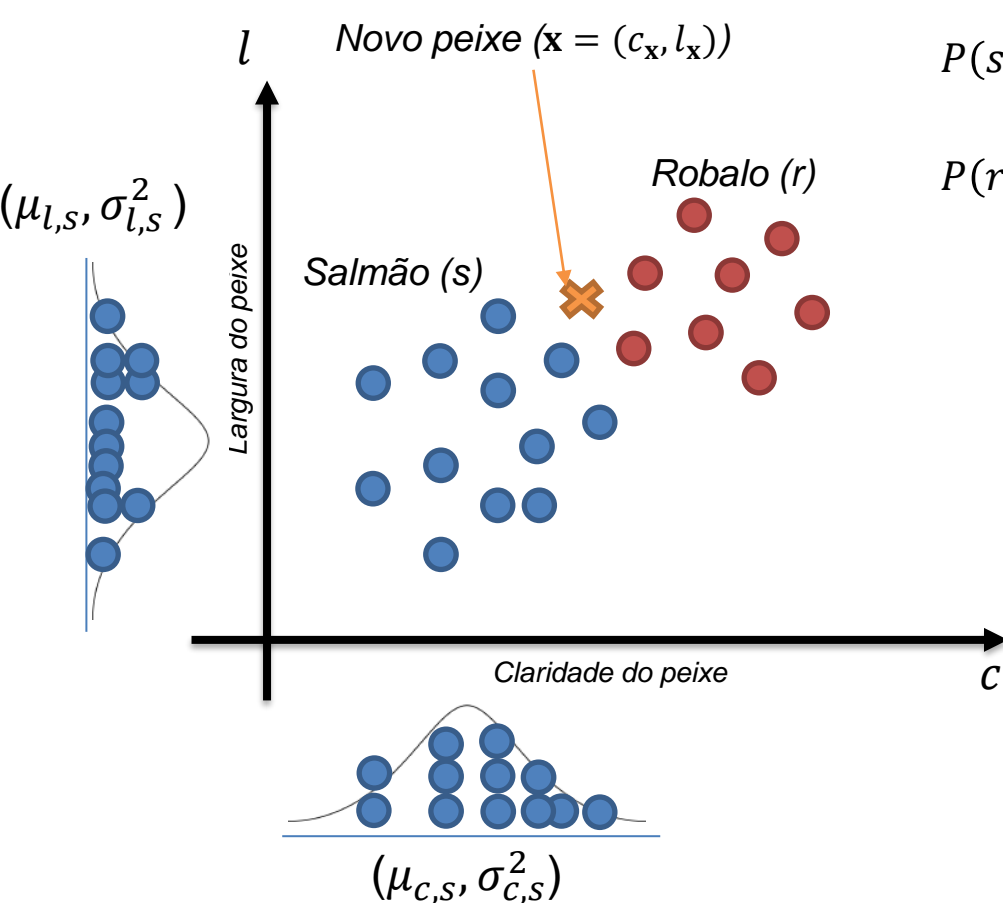
x_1 claridade da pele do peixe; x_2 largura do peixe



Classificador Bayesiano

- Como proceder quando temos mais features?

Exemplo: Para checarmos isso, vamos analisar um caso mais simples:



$$P(s) = \frac{12}{12 + 8} = 0,6$$

$$P(r) = 0,4$$

$$p(c_x|s) = \frac{1}{\sqrt{2\pi\sigma_{c,s}^2}} \exp\left[-\frac{1}{2}\left(\frac{(c_x - \mu_{c,s})}{\sigma_{c,s}}\right)^2\right]$$

$$p(l_x|s) = \frac{1}{\sqrt{2\pi\sigma_{l,s}^2}} \exp\left[-\frac{1}{2}\left(\frac{(l_x - \mu_{l,s})}{\sigma_{l,s}}\right)^2\right]$$

$$p(c_x|r) = \frac{1}{\sqrt{2\pi\sigma_{c,r}^2}} \exp\left[-\frac{1}{2}\left(\frac{(c_x - \mu_{c,r})}{\sigma_{c,r}}\right)^2\right]$$

$$p(l_x|r) = \frac{1}{\sqrt{2\pi\sigma_{l,r}^2}} \exp\left[-\frac{1}{2}\left(\frac{(l_x - \mu_{l,r})}{\sigma_{l,r}}\right)^2\right]$$

$$p(\mathbf{x}|s) = p(c_x|s)p(l_x|s)$$

$$p(\mathbf{x}|r) = p(c_x|r)p(l_x|r)$$

$$p(s|\mathbf{x}) = \frac{p(\mathbf{x}|s)P(s)}{p(\mathbf{x}|s)P(s) + p(\mathbf{x}|r)P(r)}$$

Explicando o denominador:
Teorema da probabilidade Total



prof. Miguel Bozer da Silva

FUNÇÕES DISCRIMINANTES

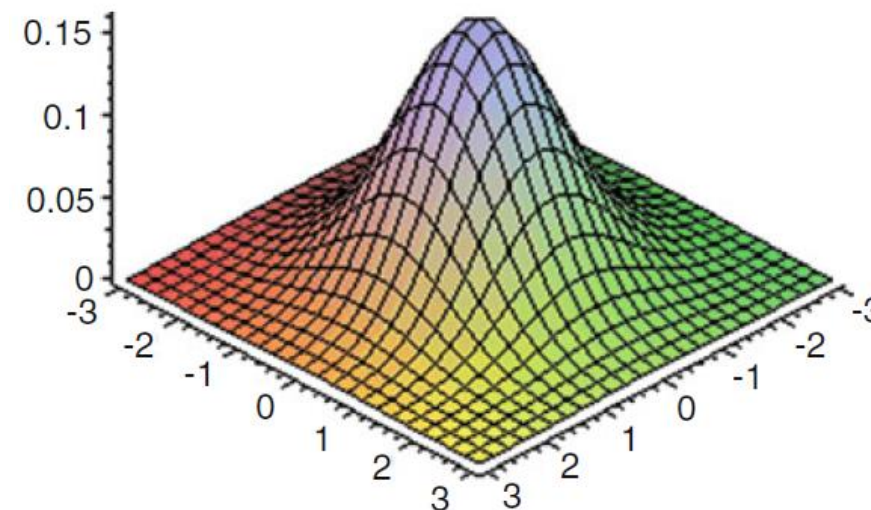
Classificador Bayesiano

- Como proceder quando temos mais features?
- Podemos usar a equação para multivariáveis Gaussiana:

$$f(\mathbf{X}) = \frac{1}{((2\pi)^{n/2} |\Sigma|^{1/2})} \exp\left(-\frac{1}{2}(\mathbf{X} - \mu)^T \Sigma^{-1}(\mathbf{X} - \mu)\right)$$

Onde:

- \mathbf{X} é o vetor com as features
- μ é o vetor com as médias
- Σ é a matriz de covariância



Classificador Bayesiano

- Como proceder quando temos mais features?
- Podemos usar a equação para multivariáveis Gaussiana:

$$f(\mathbf{X}) = \frac{1}{((2\pi)^{n/2} |\Sigma|^{1/2})} \exp\left(-\frac{1}{2}(\mathbf{X} - \mu)^T \Sigma^{-1} (\mathbf{X} - \mu)\right)$$

Onde:

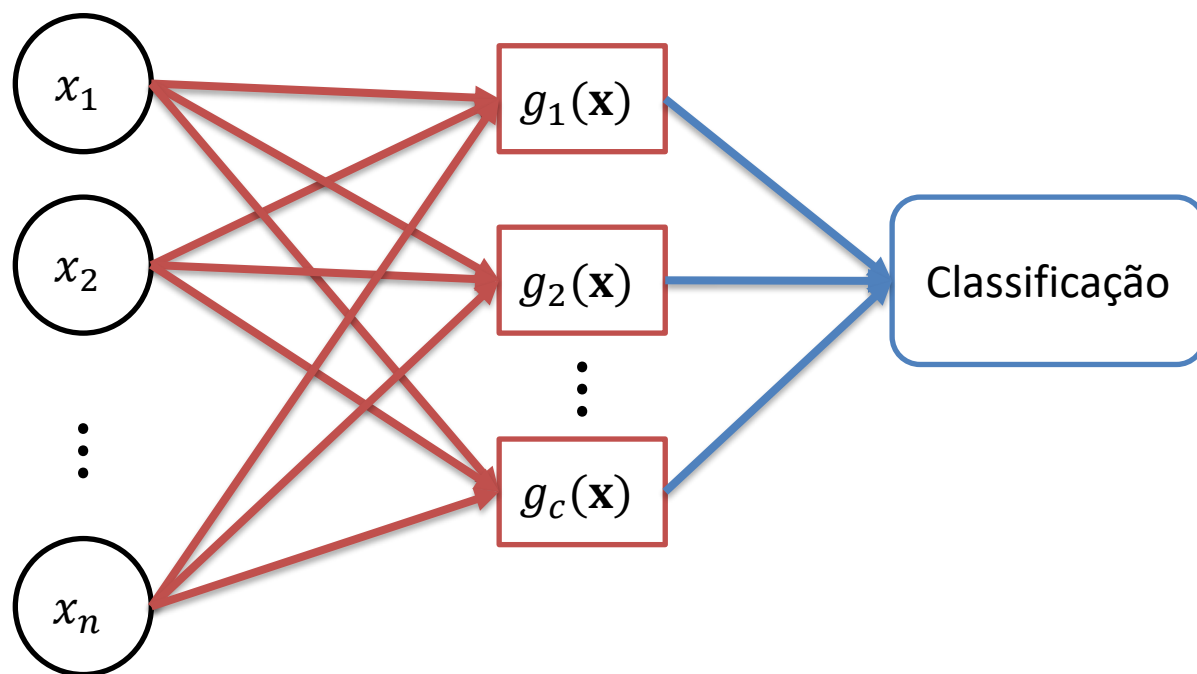
- \mathbf{X} é o vetor com as features
- μ é o vetor com as médias
- Σ é a matriz de covariância

Em um casos com
duas features

$$\xrightarrow{\text{Em um casos com duas features}} \mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \Sigma = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$$

Funções Discriminantes

- Considerem o caso de múltiplas classes $\omega_i, i = 1, 2, \dots, c$
- Dado um vetor com as nossas features $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ podemos criar um sistema de classificação que utiliza as *funções discriminantes* $g_i(\mathbf{x})$

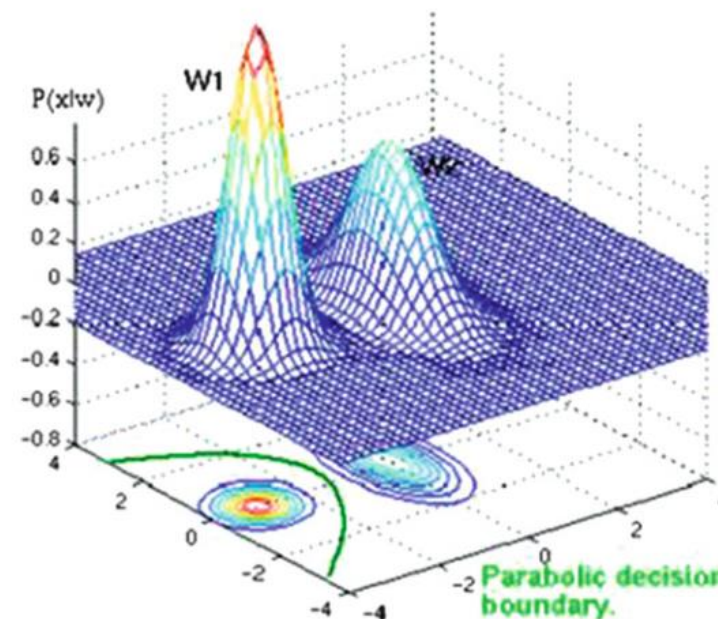
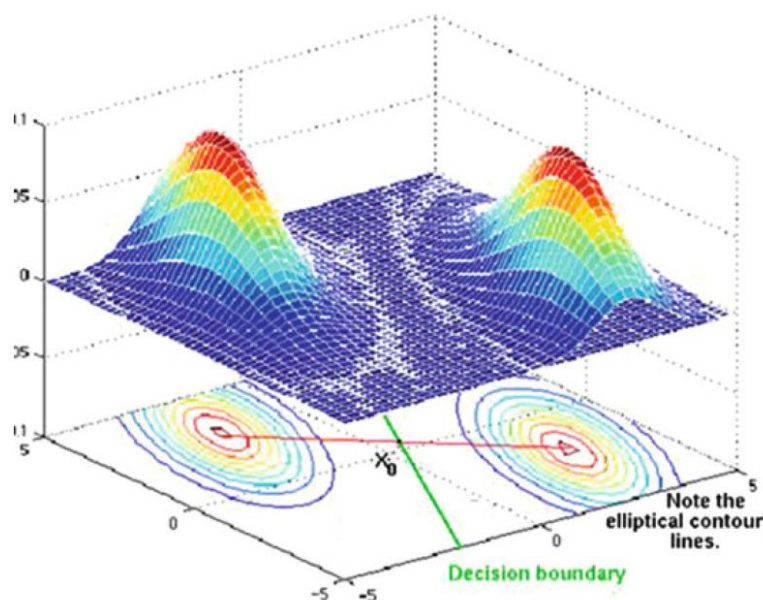


Sendo a classificação dada por:
Se $g_i(\mathbf{x}) > g_j(\mathbf{x})$ com $i \neq j$, classificar
como classe ω_i

Funções Discriminantes

Podemos usar a regra de Bayes como função discriminante gaussiana (**gaussian naive bayes - GNB**);

Também podemos definir outras funções discriminantes para o classificador bayesiano, a partir de regras que irão providenciar um discriminante linear (**Linear Discriminant Analysis – LDA**) ou um discriminante quadrático (**Quadratic Discriminant Analysis - QDA**).



- Para fechar a aula uma pergunta a todos: Qual o melhor classificador entre todos que vimos até o dado momento?
 - *No free lunch theorem:* **NÃO EXISTE UM CLASSIFICADOR IDEAL PARA TODOS OS TIPOS DE PROBLEMA**

- DOUGHERTY, Geoff. **Pattern Recognition and Classification: an introduction**. New York: Springer International Publishing, 2013.

Copyright © 2021 Prof. **Miguel Bozer da Silva**

Todos direitos reservados. Reprodução ou divulgação total ou parcial deste documento é expressamente proibido sem o consentimento formal, por escrito, do Professor (autor).