Ciencia de la
Computación

Gaussian Mixture Model

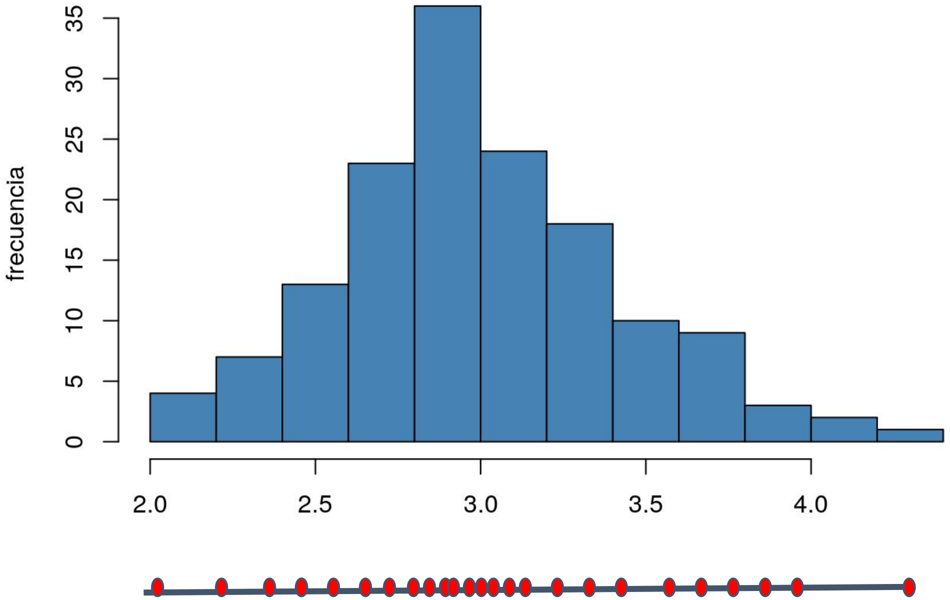
Cristian López Del Alamo
clopezd@utec.edu.pe
IPRODAM3D - Research group

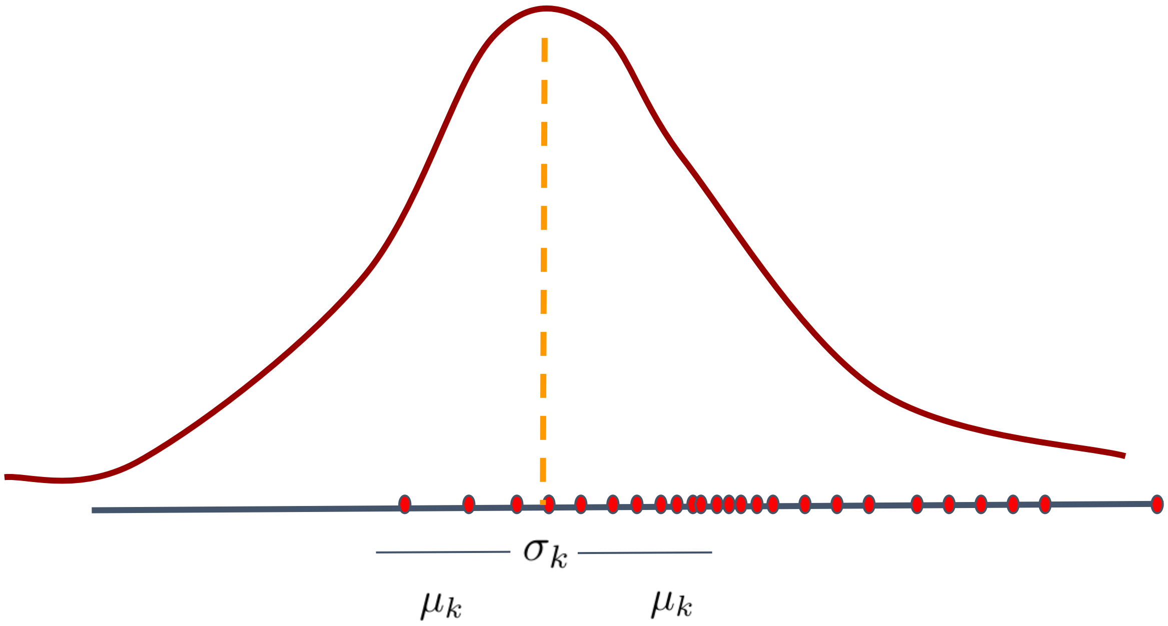
2022

Programa

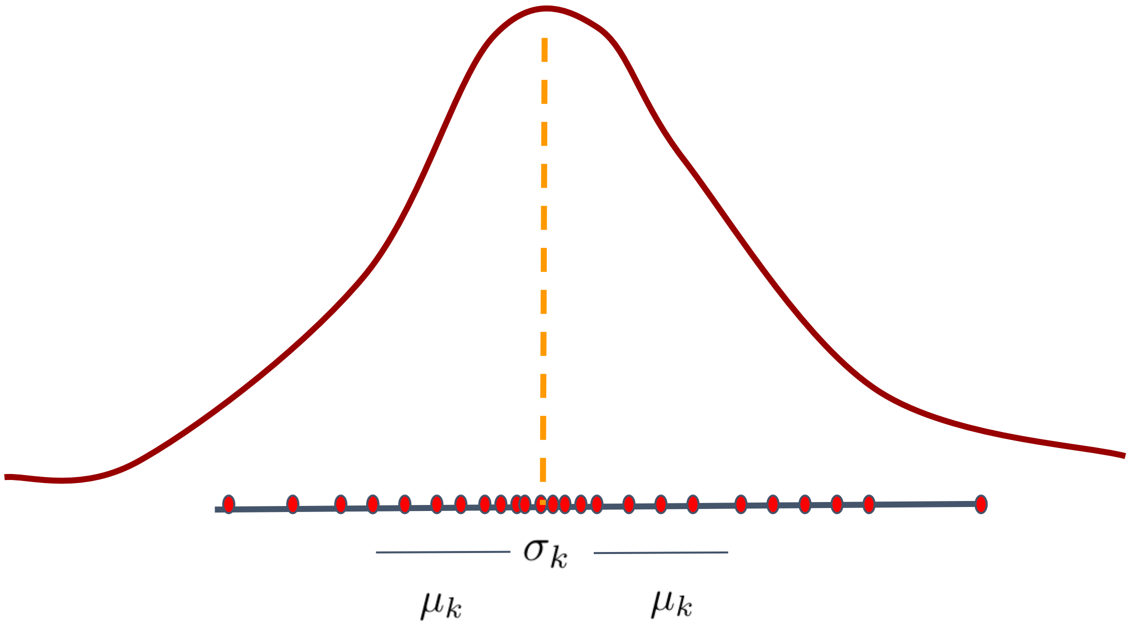
1. Introducción
2. Gaussian Mixtures
3. Gaussian Mixture Models
4. Expectation Maximization
5. Aplicaciones

1 Introducción


frecuencia
0
5
10
15
20
25
30
35
2.0
2.5
3.0
3.5
4.0



$$p(x_i) = \mathcal{N}(x_i|\mu_k, \sigma_k) = \frac{e^{-(x_i-\mu_k)^2/(2\sigma_k^2)}}{\sigma_k\sqrt{2\pi}}$$



$$p(x_i) = \mathcal{N}(x_i|\mu_k, \sigma_k) = \frac{e^{-(x_i-\mu_k)^2/(2\sigma_k^2)}}{\sigma_k\sqrt{2\pi}}$$

Maximun Likelihood

$$ML = argmax \prod_{k=1}^{n} p(x_i|\mu,\sigma)$$

$$p(x) = \prod_{i=1}^{n} p(x_i|\mu,\sigma)$$

$$ln(p(x)) = ln(\prod_{i=1}^{n} p(x_i|\mu,\sigma))$$

$$ln(p(x)) = \sum_{i=1}^{n} ln(p(x_i|\mu,\sigma))$$

$$ln(p(x)) = \sum_{i=1}^{n} ln(\frac{e^{-(x_i-\mu_k)^2/(2\sigma_k^2)}}{\sigma_k\sqrt{2\pi}})$$

$$ln(p(x)) = \sum_{i=1}^{n}(-\frac{1}{2}ln(2\pi) - \frac{1}{2}ln(\sigma) - \frac{1}{2\sigma}(x_i-\mu)^2)$$

- Aplicando la derivada con respecto al parámetro μ

$$ln(p(x)) = \sum_{i=1}^{n}(-\frac{1}{2}ln(2\pi) - \frac{1}{2}ln(\sigma) - \frac{1}{2\sigma}(x_i - \mu)^2)$$

$$\mu_{ML} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

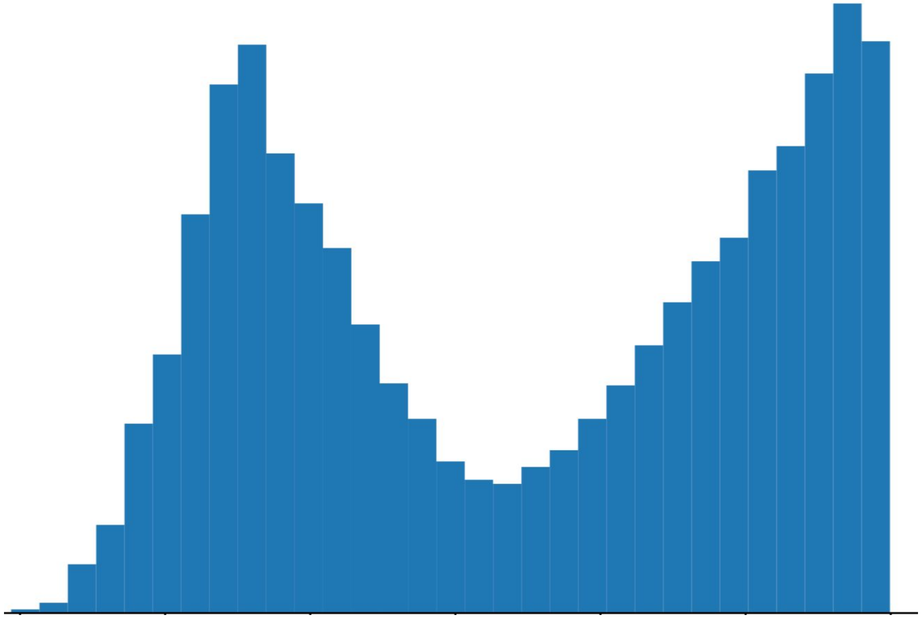
- Aplicando la derivada con respecto al parámetro σ

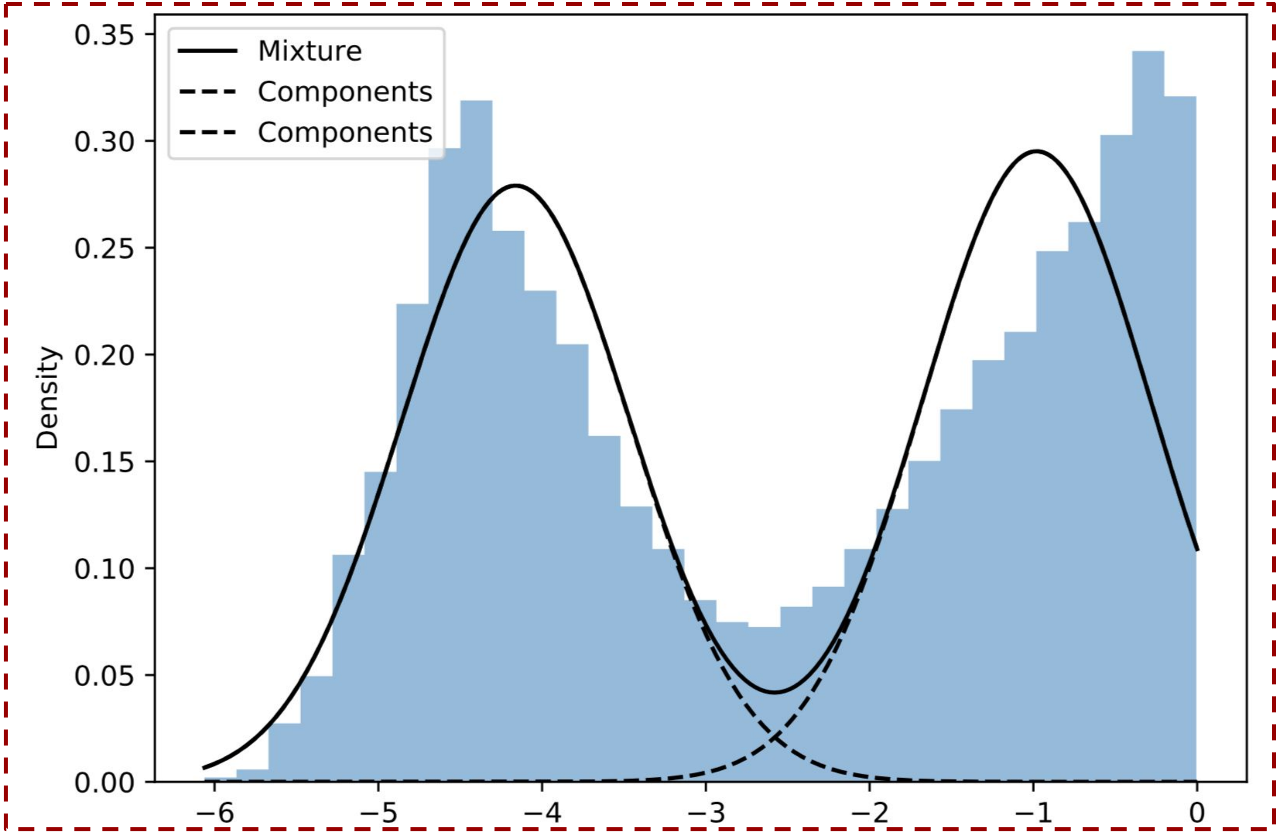
$$ln(p(x)) = \sum_{i=1}^{n}(-\frac{1}{2}ln(2\pi) - \frac{1}{2}ln(\sigma) - \frac{1}{2\sigma}(x_i - \mu)^2)$$

$$\sigma_{ML} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu_{ML})^2$$

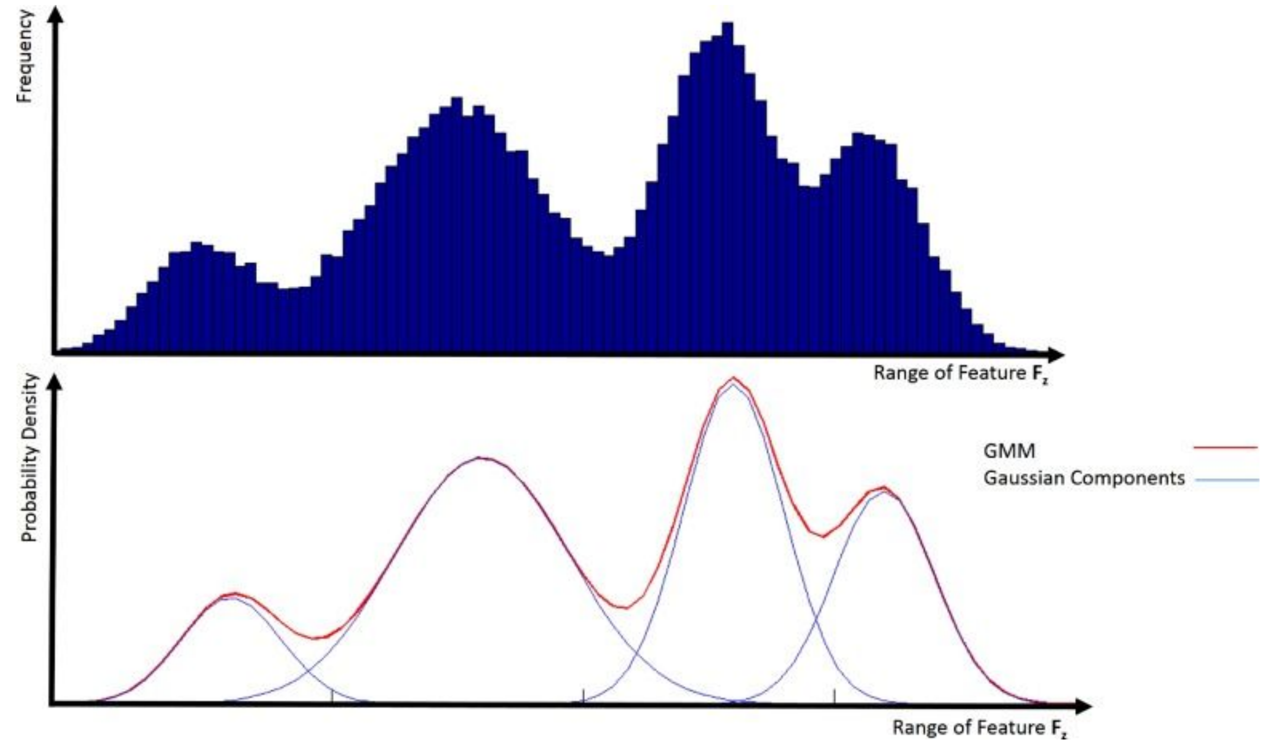
2 Gaussian Mixture

¿Qué ocurre la dispersión de los datos muestran un patrón como este?





Gaussian Mixtures



Gaussian Mixtures

Gaussian Mixtures: Es una superposición Lineal de Gaussianas

$$p(x_j) = \sum_{i=1}^{K} \pi_i \mathcal{N}(x_j|\mu_i, \Sigma_i)$$

Restricciones

$$\sum_{i=1}^{K} \pi_i = 1 \qquad 0 \leq \pi_i \leq 1$$

Al igual que en el ejemplo anterior, se trata de maximizar la probabilidad conjunta de los datos

$$p(x) = \prod_{j=1}^{n} p(x_j)$$

Dado que:

$$\boxed{p(x) = \prod_{j=1}^{n} p(x_j)}$$

y

$$\boxed{p(x_j) = \sum_{i=1}^{K} \pi_i \mathcal{N}(x_j|\mu_i, \Sigma_i)}$$

$$\Rightarrow \quad p(x) = \prod_{j=1}^{n} (\sum_{i=1}^{K} \pi_i \mathcal{N}(x_j|\mu_i, \Sigma_i))$$

$$\Rightarrow \quad ln(p(x)) = ln(\prod_{j=1}^{n} (\sum_{i=1}^{K} \pi_i \mathcal{N}(x_j|\mu_i, \Sigma_i)))$$

$$\Rightarrow \quad \boxed{ln(p(x)) = \sum_{j=1}^{n} (ln(\sum_{i=1}^{K} \pi_i \mathcal{N}(x_j|\mu_i, \Sigma_i)))}$$

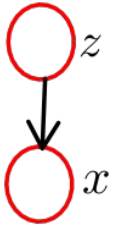
3 Gaussian Mixture Model

Gaussian Mixture Models

- Input : $X = \{x_1, x_2, ..., x_n\}, x_i \in R^n$

- Objetivo : partir el conjunto de datos X, en K cluster maximizando el likelihood del modelo probabilístico

$$p(x) = \sum_z p(x,z) = \sum_z p(x|z)p(z)$$



$$p(x,z) = p(z)p(x|z)$$

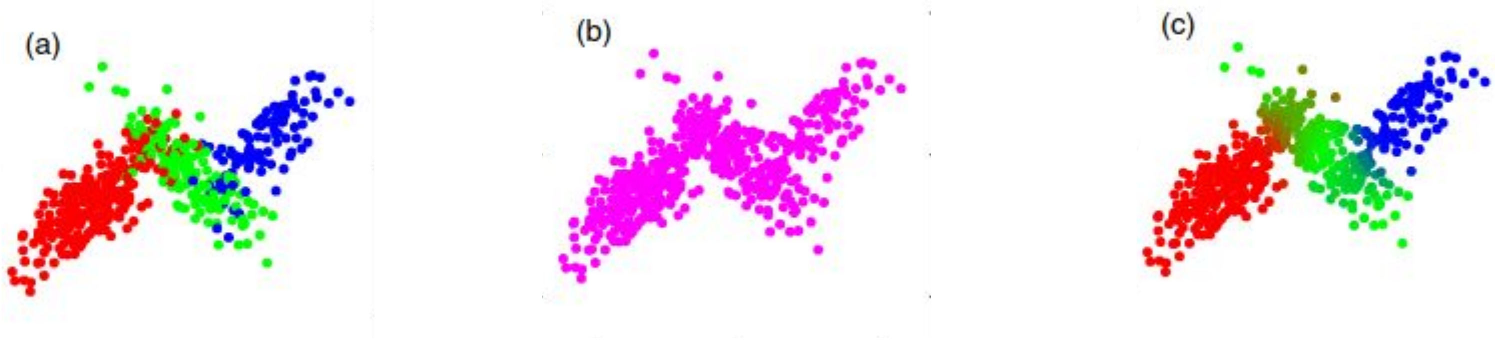
Gaussian Mixture Models

$$p(z_i = 1) = \pi_i, \pi_i \in [0,1], \sum_{i=1}^{K} \pi_i = 1$$

$$p(x|z_i = 1) = \mathcal{N}(x|\mu_i, \Sigma_i)$$

$$p(x, z_i = 1) = p(x|z_i = 1)p(z_i = 1) = \pi_i \mathcal{N}(x|\mu_i, \Sigma_i)$$

$$p(x) = \sum_{z} p(x, z) = \sum_{i} \pi_i \mathcal{N}(x|\mu_i, \Sigma_i)$$



$$x, z \sim p(x, z)$$

$$x \sim p(x)$$

$$z \sim p(z|x)$$

Probabilidad a posteriori

$$p(z_i = 1|x) = \frac{p(z_i = 1)p(x|z_i = 1)}{p(x)}$$

$$p(z_i = 1|x) = \frac{p(z_i = 1)p(x|z_i = 1)}{\sum_j p(z_j = 1)p(x|z_j = 1)}$$

$$p(z_i = 1|x) = \frac{\pi_i \mathcal{N}(x|\mu_i, \Sigma_k)}{\sum_j \pi_j \mathcal{N}(x|\mu_j, \Sigma_j)} = \gamma(z_i)$$

Log Likelihood

$$p(x) = \sum_z p(x,z) = \sum_i \pi_i \mathcal{N}(x|\mu_i, \Sigma_i)$$

$$\boxed{ln(p(x)) = \sum_{j=1}^{n} (ln(\sum_{i=1}^{K} \pi_i \mathcal{N}(x_j|\mu_i, \Sigma_i)))}$$

4 Expectation Maximization

Expectation - Maximization Algorithms (EM)

- Necesitamos maximizar el likelihood con respecto a π_i, μ_i y Σ_i para i = 1,2,...K

$$\boxed{\sum_{j=1}^{n}(ln(\sum_{i=1}^{K}\pi_i\mathcal{N}(x_j|\mu_i,\Sigma_i)))}$$

- No podemos encontrar una solución cerrada.

- Podemos encontrar un **mínimo local** mediante un **algoritmo iterativo**:
  - Actualización alternativa de la probabilidad a posteriori esperada (**Expectation**)
  - Maximización (**Maximization**) de los parámetros π, μ y Σ

Expectation - Maximization Algorithms (EM)

- Para un $\gamma(z_{ji})$ fijo, podemos estimar los parámetros π_i, μ_i y Σ_i

$$\boxed{\sum_{j=1}^{n}\left(ln\left(\sum_{i=1}^{K}\pi_i\mathcal{N}(x_j|\mu_i,\Sigma_i)\right)\right)}$$

Expectation - Maximization Algorithms (EM): Actualización de Parámetros

$$\mu_i = \frac{1}{N_i}\sum_{j=1}^{n}\gamma(z_{ji})x_i$$

$$N_i = \sum_{j=1}^{n}\gamma(z_{ji})$$

$$\Sigma_i = \frac{1}{N_i}\sum_{j=1}^{n}\gamma(z_{ji})(x_j - u_i)(x_j - u_i)^T$$

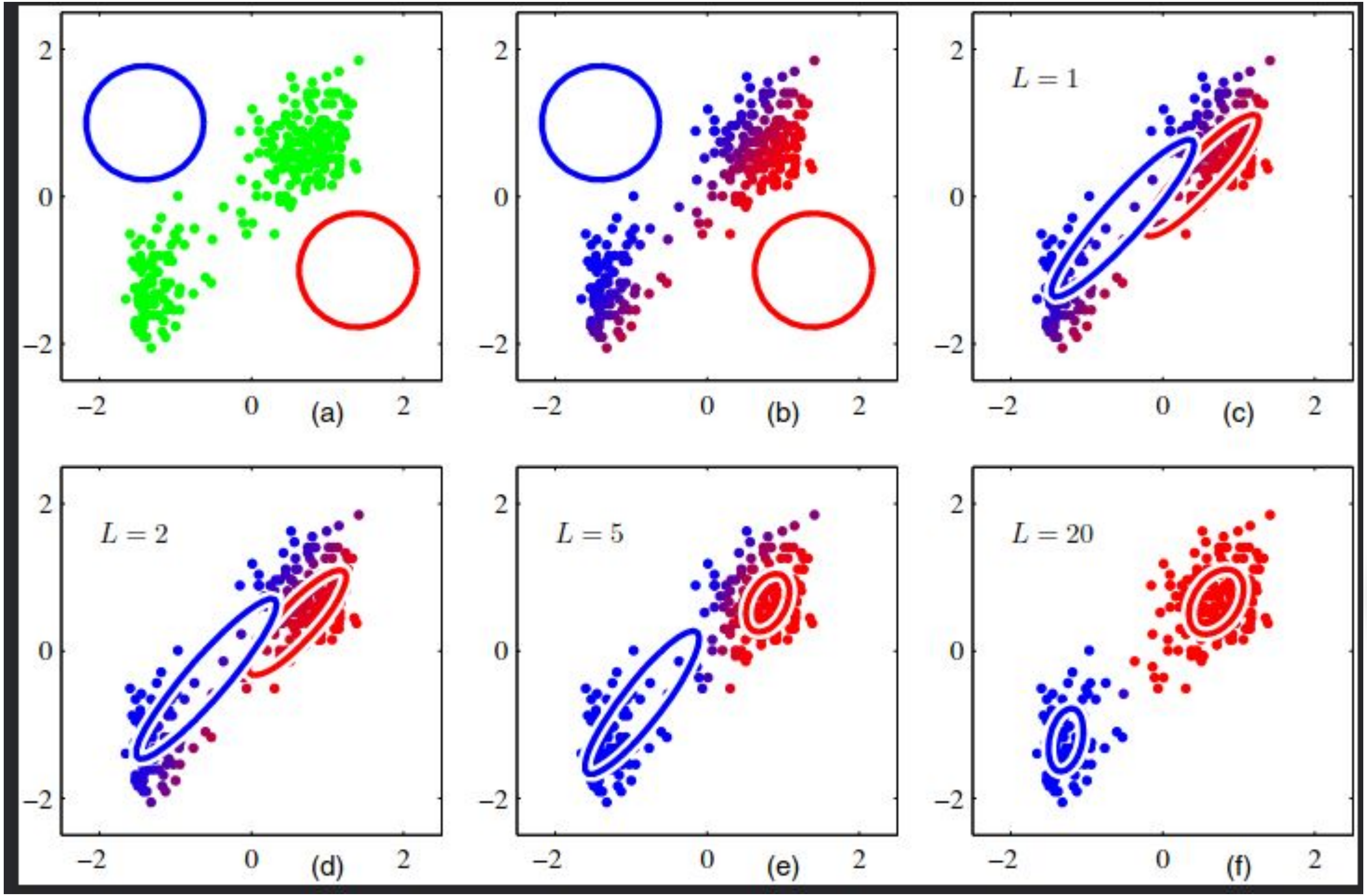
$$\pi_i = \frac{N_i}{N}$$

	c_1	c_2	c_3
x_1	0.5	3.3	2.1
x_2	1.5	2.3	1.3
x_3	0.5	3.3	2.1
x_4	4.5	6.3	7.3
x_5	3.5	3.5	2.1
x_n	0.5	3.3	2.1

	z_1	z_2
γ_1	0.3	0.7
γ_2	0.6	0.4
γ_3	0.2	0.8
γ_4	0.1	0.9
γ_n	0.9	0.1

μ_1	μ_{11}	μ_{12}	μ_{13}
μ_2	μ_{21}	μ_{22}	μ_{23}

π_1	π_2
0.3	0.7



2
0
−2
−2
0
2
(a)
(b)
(c)
$L = 1$
$L = 2$
$L = 5$
$L = 20$
(d)
(e)
(f)

1. Initialize the means $\boldsymbol{\mu}_k$, covariances $\boldsymbol{\Sigma}_k$ and mixing coefficients π_k, and evaluate the initial value of the log likelihood.
2. **E step**. Evaluate the responsibilities using the current parameter values

$$\gamma(z_{nk}) = \frac{\pi_k \mathcal{N}(\mathbf{x}_n|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}{\sum_{j=1}^{K} \pi_j \mathcal{N}(\mathbf{x}_n|\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)}. \tag{9.23}$$

3. **M step**. Re-estimate the parameters using the current responsibilities

$$\boldsymbol{\mu}_k^{\text{new}} = \frac{1}{N_k} \sum_{n=1}^{N} \gamma(z_{nk}) \mathbf{x}_n \tag{9.24}$$

$$\boldsymbol{\Sigma}_k^{\text{new}} = \frac{1}{N_k} \sum_{n=1}^{N} \gamma(z_{nk}) \left(\mathbf{x}_n - \boldsymbol{\mu}_k^{\text{new}}\right) \left(\mathbf{x}_n - \boldsymbol{\mu}_k^{\text{new}}\right)^{\mathrm{T}} \tag{9.25}$$

$$\pi_k^{\text{new}} = \frac{N_k}{N} \tag{9.26}$$

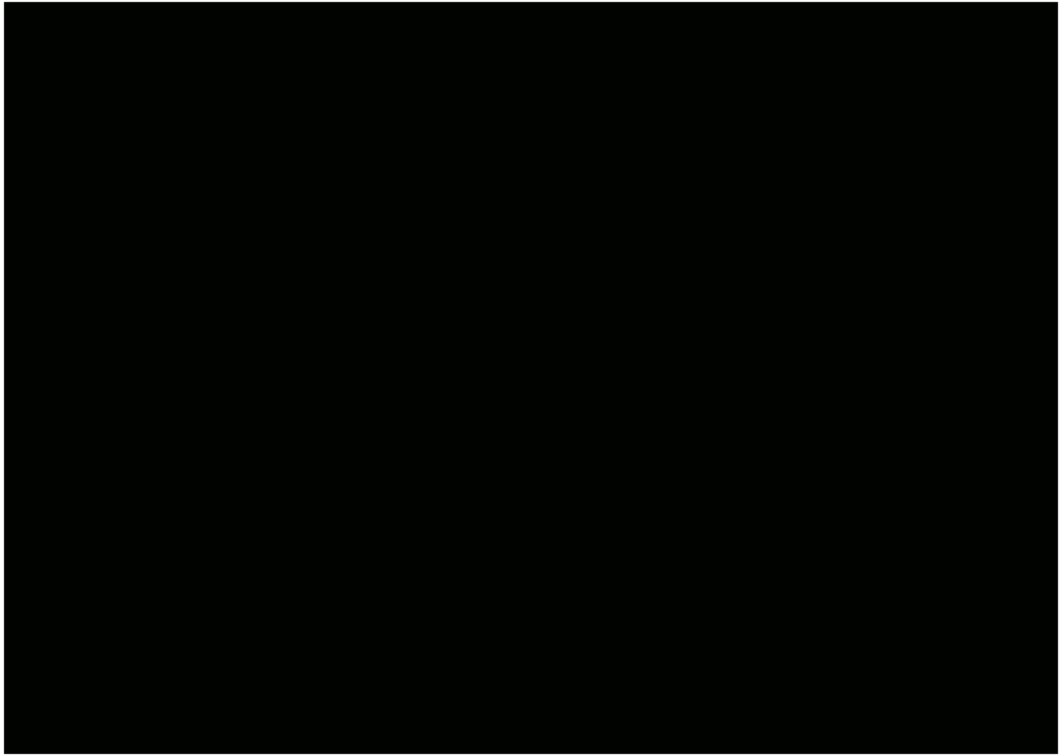
where

$$N_k = \sum_{n=1}^{N} \gamma(z_{nk}). \tag{9.27}$$

4. Evaluate the log likelihood

$$\ln p(\mathbf{X}|\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\pi}) = \sum_{n=1}^{N} \ln \left\{ \sum_{k=1}^{K} \pi_k \mathcal{N}(\mathbf{x}_n|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \right\} \tag{9.28}$$

and check for convergence of either the parameters or the log likelihood. If the convergence criterion is not satisfied return to step 2.

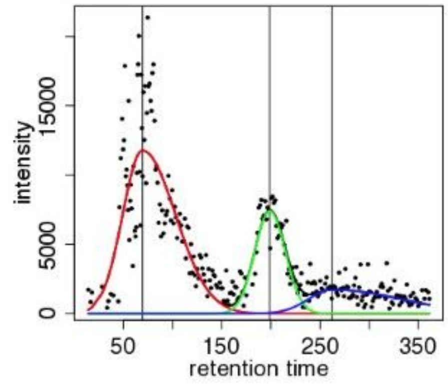


Fuente: [click]

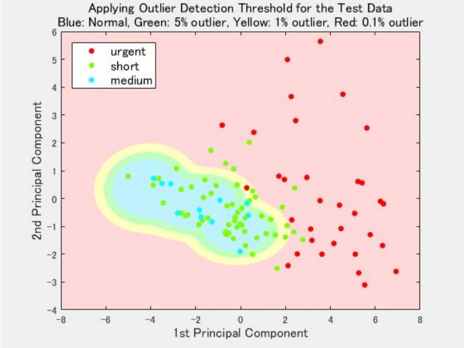
5 Aplicaciones

Aplicaciones

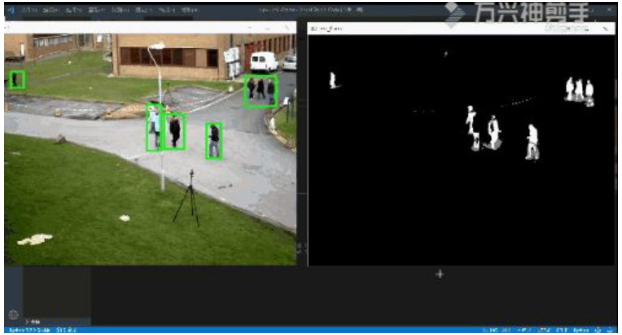
Procesamiento de señales



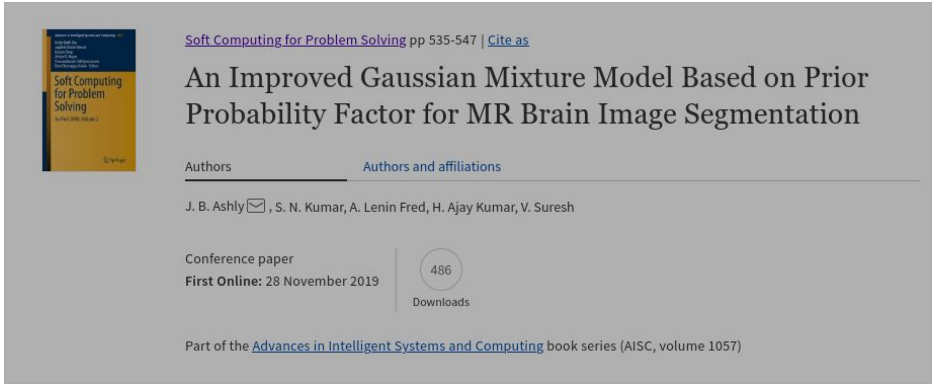
Detección de Anomalías



Seguimiento de objetos en video



Aplicaciones



Soft Computing for Problem Solving pp 535-547 | Cite as

An Improved Gaussian Mixture Model Based on Prior Probability Factor for MR Brain Image Segmentation

Authors Authors and affiliations

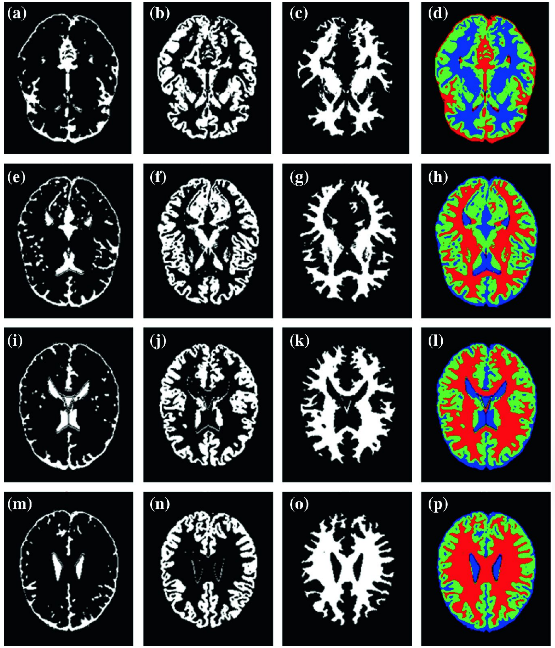
J. B. Ashly, S. N. Kumar, A. Lenin Fred, H. Ajay Kumar, V. Suresh

Conference paper

First Online: 28 November 2019

486 Downloads

Part of the Advances in Intelligent Systems and Computing book series (AISC, volume 1057)



Gracias

Cristian López Del Alamo

UTEC
UNIVERSIDAD DE INGENIERÍA
Y TECNOLOGÍA