Sorry arima, I'm going Bayesian

Pierre Gauthier

École des Mines de Nancy

May 2019

Tuteur : Denis Villemonais

Sommaire

Application numérique : Une alternative à l'algorithme E-M pour le mélange de gaussiennes

Échantillon $\boldsymbol{x} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N)$ suivant une loi $f(\boldsymbol{x}_i, \theta)$ avec des variables latentes $\boldsymbol{z} = (\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_n)$

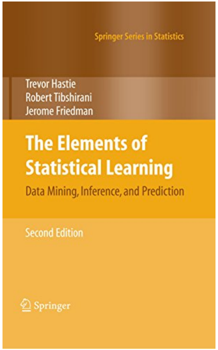
Maximisation de $L(\boldsymbol{\theta}; \mathbf{X}) = p(\mathrm{x}|\boldsymbol{\theta}) = \int p(\mathrm{x}, \mathrm{z}|\boldsymbol{\theta}) d\mathbf{Z}$

$$E\left[L(\mathrm{x};\boldsymbol{\theta})|\boldsymbol{\theta}^{(c)}\right] = E\left[L((\mathrm{x},\mathrm{z});\boldsymbol{\theta})|\boldsymbol{\theta}^{(c)}\right] - E\left[\sum_{i=1}^{n} \log f\left(z_i|\boldsymbol{x}_i, \boldsymbol{\theta}\right)|\boldsymbol{\theta}^{(c)}\right]$$

$$L(\mathrm{x};\boldsymbol{\theta}) = Q\left(\boldsymbol{\theta};\boldsymbol{\theta}^{(c)}\right) - H\left(\boldsymbol{\theta};\boldsymbol{\theta}^{(c)}\right)$$

$\boldsymbol{\theta}^{(c+1)} = \arg\max_{\boldsymbol{\theta}}\left(Q\left(\boldsymbol{\theta},\boldsymbol{\theta}^{(c)}\right)\right)$ fait tendre la suite $L\left(\mathrm{x};\boldsymbol{\theta}^{(c+1)}\right)$ vers un max local

$$p\left(\mathcal{C}_1|x\right) = \frac{p\left(\mathcal{C}_1\right) p\left(x|\mathcal{C}_1\right)}{p\left(\mathcal{C}_1\right) p\left(x|\mathcal{C}_1\right) + p\left(\mathcal{C}_2\right) p\left(x|\mathcal{C}_2\right)}$$



Application : mélange de gaussiennes

- Les données **x** viennent de gaussiennes $\{C_1, \ldots, C_M\}$
- $\boldsymbol{z}_i \in \{1, \ldots, m\}$ si un individu i vient de $C_1, \ldots, C_M$

$$\boldsymbol{x}_i | (\boldsymbol{z}_i = j) \sim \mathcal{N}\left(\mu_j, \sigma_j^2\right) \quad \mathbb{P}(\boldsymbol{z}_i = j) = \pi_j$$

- **E-M** : maximisation de

$$\ell_\theta(x) = \sum_{n=1}^{N} \log\left(\sum_{j=1}^{M} \pi_j\, \mathcal{N}_{\mu_j, \sigma_j^2}(x_n)\right)$$

$$\theta = (\pi, \mu, \sigma^2)$$

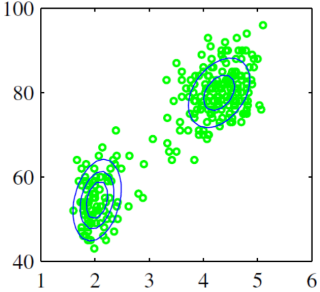


Figure – Mélange de deux gausiennes [F. Sur, Introduction à l'apprentissage automatique]

Application : Échantillonneur de Gibbs pour le mélange de 2 gaussiennes

Pour 2 gaussiennes : $\begin{cases} \boldsymbol{x}_i | (\boldsymbol{z}_i = 1) \sim \mathcal{N}\left(\mu_1, \sigma_1^2\right) & \mathbb{P}\left(\boldsymbol{z}_i = 1\right) = \pi_1 \\ \boldsymbol{x}_i | (\boldsymbol{z}_i = 2) \sim \mathcal{N}\left(\mu_2, \sigma_2^2\right) & \mathbb{P}\left(\boldsymbol{z}_i = 2\right) = \pi_2 = 1 - \pi_1 \end{cases}$

- On veut obtenir $\Theta = (\boldsymbol{z}, \mu) = (\boldsymbol{z_1}, \ldots, \boldsymbol{z_1}, \mu)$ $(\pi, \sigma$ fixés à $\hat{\pi}, \hat{\sigma})$
- $z|\mu$: utiliser $\mathbb{P}\left(\boldsymbol{z}_i = j | \mu, \boldsymbol{x}\right) = \dfrac{\hat{\pi}_j p_{\mathcal{N}}(\boldsymbol{x}_i, \mu_j, \hat{\sigma}_j)}{\hat{\pi}_1 p_{\mathcal{N}}(\boldsymbol{x}_i, \mu_1, \hat{\sigma}_1) + \hat{\pi}_2 p_{\mathcal{N}}(\boldsymbol{x}_i, \mu_2, \hat{\sigma}_2)}$ $\quad j \in \{1, 2\}$
- $\mu|z$:

(à priori)

$$\boldsymbol{x_i}|\mu, \boldsymbol{z_i} = j \sim \mathcal{N}\left(\mu_j, 1/\hat{\tau}_j\right)$$
$$\mu_j \sim \mathcal{N}\left(\mu_0, 1/\tau_0\right)$$

$\left(\tau_j = \mathbf{1}/\sigma_j^{\mathbf{2}},\ \bar{x} = \frac{\mathbf{1}}{N} \sum_{i=\mathbf{0}}^{N} x_i\right)$

(à postériori)

$$p(\mu_j|\boldsymbol{z}_i, \mathrm{x}) \propto p(\boldsymbol{x}|\mu, \boldsymbol{z_i} = j)p(\mu_j)$$
$$\sim \mathcal{N}(\mu_0', 1/\tau_0')$$

(= maj des coefs)

$$\tau_{j_0}' = \tau_0 + N\hat{\tau}_j$$
$$\mu_{j_0}' = \frac{N\ \hat{\tau}_j\ \bar{x} + \tau_0 \mu_0}{N\hat{\tau}_j + \tau_0}$$

Échantilloneur de Gibbs pour une mixture gaussienne

1. Prendre les valeurs initiales $\Theta_0 = (\boldsymbol{z}^{(0)}, \mu_1^{(0)}, \mu_2^{(0)}),\ \mu_0, \tau_0$
2. Pour t de à ...
   - Pour i de 1 à N tirer

     $\boldsymbol{z}_i^{(t+1)} \in \{1,2\}$ avec $\mathbb{P}(\boldsymbol{z}_i^{(t+1)} = 1) = \dfrac{\hat{\pi}_1 p_{\mathcal{N}}(\boldsymbol{x}_i, \mu_1^{(t)}, \hat{\sigma}_1)}{\hat{\pi}_1 p_{\mathcal{N}}(\boldsymbol{x}_i, \mu_1^{(t)}, \hat{\sigma}_1) + \hat{\pi}_2 p_{\mathcal{N}}(\boldsymbol{x}_i, \mu_2^{(t)}, \hat{\sigma}_2)}$
   - Pour $j \in \{1,2\}$

     Estimer $\tau'_{jo} = \tau_0 + N\hat{\tau}_j \quad \mu'_{jo} = \dfrac{N\ \hat{\tau}_j\ \bar{x} + \tau_0\mu_0}{N\hat{\tau}_j + \tau_0}$

     Tirer $\mu_j^{(t+1)} \sim \mathcal{N}\left(\mu'_{jo}, 1/\tau'_{jo}\right)$

```
Z_given_mu <- function(X,Z,mu,pi_1){
    # Z variables latentes

    remove_i <- (c(1:length(X)) !=i) # enleve variable i
    estimate_sigma_1 <- sd(X[(remove_i & Z == 1)]) # classe 1
    estimate_sigma_2 <- sd(X[(remove_i & Z == 2)]) # classe 2

    for (i in 1:length(Z)){

        proba1 <- pi_1 * dnorm(X[i],mu[1],estimate_sigma_1) /
            (pi_1 * dnorm(X[i],mu[1],estimate_sigma_1) +
            (1-pi_1) * dnorm(X[i],mu[2],estimate_sigma_2))

        Z[i] = sample(1:2, size=1,prob=c(proba1, 1-proba1),
            replace=TRUE)
    }
    return(Z)
}
```

```
mu_given_Z = function(X, Z, mu_prior){
    # Z variables latentes
    # mu_prior contient parametres loi a priori de mu

    mu = rep(0,2) ; sigma = rep(0,2)

    for(j in 1:2){

        sample_j_size = sum(Z==j)
        sample_j_mean = mean(X[Z==j])
        sigma[j] = sd(X[Z==j]) ; precision_j = 1 / sigma[j]^2

        precision_post = sample_j_size * precision_j +
            mu_prior$precision

            mean_post = (sample_j_mean * sample_j_size *
            precision_j + mu_prior$mean *
            mu_prior$precision ) / precision_post

        mu[j] = rnorm(1,mean_post,sqrt(1/precision_post)) # on
               tire mu selon la loi normale a posteriori
    }
    return(list(mu = mu, sigma = sigma))
    }
```

```
echantillonneur_gibbs <- function(X,N_simu){
    # X sont les donnees
    # initalisation
    .
    .

    for (k in 1:N_simu){

        Z <- Z_given_mu(X,Z,mu,pi_1) ; param_post <- mu_given_Z(X, Z, mu_prior)
        mu = param_post$mu
    .
    .
    .
    }
    return(list(Z, mu_1, mu_2))
}
```

Listing 1 – Implémentation en R

Blocky block

Just a Block

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Blocky block

Example Block

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Blocky block

Alert Block

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

How to use frame-breakings ?

In this template, and only this, I defined a "breakingframe" template frame that should not hold any useful information. The background of this frame is pinkish solid and it is not countable as a separate frame. You can use this as a transitioning page between different topics or for any funny funky stuff to release the tense of the poor audience during your presentation.

```
        \breakingframe{
            Put your contents here, such as images, text ..etc. Be as silly as
                possible .. or not!
        }
```

Look at the next slide, in code, as an example !

Motivation : Q1

1. Stone masonry walls are usually not homogeneous through the thickness
2. Leaf-separation effects on the strength capacity
3. In-plane and out-of-plane behaviours interaction
4. Internal cracking onsets and 3D crack paths (cannot be captured experimentally)

Motivation : Q1

1. Stone masonry walls are usually not homogeneous through the thickness
2. Leaf-separation effects on the strength capacity
3. In-plane and out-of-plane behaviours interaction
4. Internal cracking onsets and 3D crack paths (cannot be captured experimentally)

Motivation : Q1

1. Stone masonry walls are usually not homogeneous through the thickness
2. Leaf-separation effects on the strength capacity
3. In-plane and out-of-plane behaviours interaction
4. Internal cracking onsets and 3D crack paths (cannot be captured experimentally)

Motivation : Q1

1. Stone masonry walls are usually not homogeneous through the thickness
2. Leaf-separation effects on the strength capacity
3. In-plane and out-of-plane behaviours interaction
4. Internal cracking onsets and 3D crack paths (cannot be captured experimentally)

The study main phases

How to arrange stones ?

Objective function

Packing objective

$$\text{Minimize } F(\vec{X}_i)_i = || \vec{S}_i - \vec{S}_{i-1} ||$$

$$Fitness\left(F(\vec{X}_i)\right)_i = F(\vec{X}_i)_i(1 + \xi_1 P_A)^{\xi_2}$$

- $S_i, S_{i-\mathbf{1}}$: locations of i and $i - 1$ stones
- $\xi_{\mathbf{1}}$: penalty multiplier
- $\xi_{\mathbf{2}}$: penalty exponent
- P_A : penalties summation

Merci de votre attention