



Note Technique: Approximation de pdf's à l'aide de mixture de lois connues

Jean-Pierre Cances

► To cite this version:

Jean-Pierre Cances. Note Technique: Approximation de pdf's à l'aide de mixtures de lois connues.
[Rapport de recherche] Xlim UMR CNRS 7252. 2015. hal-01400502

HAL Id: hal-01400502

<https://hal.science/hal-01400502>

Submitted on 22 Nov 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Note Technique: Approximation de pdf's à l'aide de mixtures de lois connues

J.P. Cances, Xlim UMR 7252

Le but de cette note technique est de montrer comment une distribution inconnue dont on a déterminé l'histogramme peut être approximée par un mélange de lois de probabilité connues. On utilise pour ce faire un algorithme de maximum de vraisemblance que l'on approxime à l'aide de l'algorithme E.M. (Expectation Maximization). On donne trois exemples pour des mélanges de lois Gaussiennes, de lois Gama et de lois de Nakagami-m.

- Cas d'un mélange de lois Gaussiennes:

La densité de probabilité approximée par un mélange de lois Gaussiennes prend alors la forme :

$$f_C(x) = \sum_{n=1}^J \pi_n \cdot \frac{1}{\sqrt{2\pi} \cdot \sigma_n} \cdot e^{-(x-m_n)^2 / 2 \cdot \sigma_n^2} = \sum_{n=1}^J \pi_n \cdot N(m_n, \sigma_n^2) \quad (1)$$

Evidemment, la recherche pratique de tous les paramètres J , π_n , m_n et σ_n est une tâche compliquée et on va recourir à un algorithme bien connu (l'algorithme E.M.) pour simplifier le problème. Le problème de l'approximation de la pdf peut se résumer sous la forme suivante : pour un nombre fixe de mixtures J , et en se basant sur les observations issues de l'histogramme des tirages : $\mathcal{E} \equiv \{\xi_i, i=1, \dots, n\}$, comment trouver les paramètres $\theta \equiv \{\pi_j, m_j, \sigma_j, j=1, \dots, J\}$ optimaux (au sens de la maximisation de la vraisemblance de $f_C(x)$). En appelant $\phi(x; \mu, \sigma^2)$ la pdf d'une loi Gaussienne $N(\mu, \sigma^2)$, le problème de cette recherche s'écrit sous la forme :

$$\begin{aligned} \hat{\theta} &= \arg \max_{\theta: \sum_{j=1}^J \pi_j = 1} \log p_{\theta}(\mathcal{E}) \\ &= \arg \max_{\theta: \sum_{j=1}^J \pi_j = 1} \sum_{i=1}^n \log \sum_{j=1}^J \pi_j \cdot \phi(\xi_i; m_j, \sigma_j^2) \end{aligned} \quad (2)$$

L'algorithme E.M., partant d'un état initial $\theta^{(0)}$, résout l'équation (2) de façon itérative en deux étapes : une étape de moyennage et une étape de maximisation. Les observations $\mathcal{E}$ correspondent aux données incomplètes. Les deux étapes s'écrivent :

- **Etape de moyennage**: Calculer

$$Q(\theta | \theta^{(i)}) = E_{\theta^{(i)}} \{ \log p_{\theta}(X) | \mathcal{E} \} \quad (3)$$

- **Etape de maximisation** : Résoudre

$$\theta^{(i+1)} = \arg \max_{\theta} Q(\theta | \theta^{(i)}) \quad (4)$$

On définit alors l'ensemble des données cachées $\mathbf{Z} = \{z_i, i=1, \dots, n\}$ où z_i est un vecteur indicateur de dimension J tel que :

$$z_{i,j} = \begin{cases} 1, & \text{si } \xi_i \cong N(m_j, \sigma_j^2) \\ 0 & \text{sinon} \end{cases} \quad (5)$$

L'ensemble complet des données est alors défini par: $\mathbf{X} \cong (\mathbf{\Xi}, \mathbf{Z})$, et l'on a :

$$p_{\theta}(\mathbf{\Xi}, \mathbf{Z}) = \prod_{i=1}^n \prod_{j=1}^J [\pi_j \cdot \phi(\xi_i; m_j, \sigma_j^2)]^{z_{i,j}} \quad (6)$$

La fonction de vraisemblance logarithmique de l'ensemble complet des données s'écrit :

$$\log p_{\theta}(\mathbf{\Xi}, \mathbf{Z}) = \sum_{i=1}^n \sum_{j=1}^J z_{i,j} \cdot \log \pi_j + \sum_{i=1}^n \sum_{j=1}^J z_{i,j} \cdot \left[-\frac{1}{2} \cdot \log(\sigma_j^2) - \frac{(\xi_i - m_j)^2}{2 \cdot \sigma_j^2} \right] + C \quad (7)$$

Où C est une constante quelconque. L'étape de moyennage prend alors la forme:

$$Q(\theta | \theta') = E'_{\theta} \{ \log p_{\theta}(\mathbf{\Xi}, \mathbf{Z}) | \mathbf{\Xi} \}$$

$$Q(\theta | \theta') = \sum_{i=1}^n \sum_{j=1}^J \hat{z}_{i,j} \cdot [\log \pi_j - \log(\sigma_j^2) - \frac{(\xi_i - m_j)^2}{2 \cdot \sigma_j^2}] + C \quad (8)$$

Où :

$$\begin{aligned} \hat{z}_{i,j} &= E'_{\theta} \{ z_{i,j} | \mathbf{\Xi}, \theta' \} = P'_{\theta} \{ z_{i,j} = 1 | \xi_i \} \\ &= \frac{p'_{\theta}(\xi_i | z_{i,j} = 1) \cdot P'_{\theta}(z_{i,j} = 1)}{\sum_{l=1}^J p'_{\theta}(\xi_i | z_{i,l} = 1) \cdot P'_{\theta}(z_{i,l} = 1)} \\ &= \frac{\phi(\xi_i; m'_j, \sigma'^2_j) \cdot \pi'_j}{\sum_{l=1}^J \phi(\xi_i; m'_l, \sigma'^2_l) \cdot \pi'_l} \end{aligned} \quad (9)$$

De plus, l'étape de maximisation se calcule sous la forme suivante :

- Pour obtenir $\{\pi_j\}$, on calcule la dérivée partielle:

$$\frac{\partial Q(\theta, \theta')}{\partial \pi_j} = 0 \Rightarrow \pi_j = \frac{1}{n} \cdot \sum_{i=1}^n \hat{z}_{i,j}, \quad j = 1, \dots, J \quad (10)$$

De plus, pour les autres paramètres, on obtient en utilisant les dérivées partielles :

Pour $\{m_j\}$, on a:

$$\begin{aligned}
\frac{\partial Q(\theta, \theta')}{\partial m_j} = 0 &\Rightarrow \sum_{i=1}^n \hat{z}_{i,j} \cdot \left[\frac{2 \cdot (\xi_i - m_j)}{2 \cdot \sigma_j^2} \right] = 0, \quad j = 1, \dots, J \\
&\Rightarrow \frac{m_j}{\sigma_j^2} \sum_{i=1}^n \hat{z}_{i,j} = \sum_{i=1}^n \frac{\xi_i \cdot \hat{z}_{i,j}}{\sigma_j^2}, \quad j = 1, \dots, J \\
&\Rightarrow m_j = \frac{\sum_{i=1}^n \xi_i \cdot \hat{z}_{i,j}}{\sum_{i=1}^n \hat{z}_{i,j}}, \quad j = 1, \dots, J \quad (11)
\end{aligned}$$

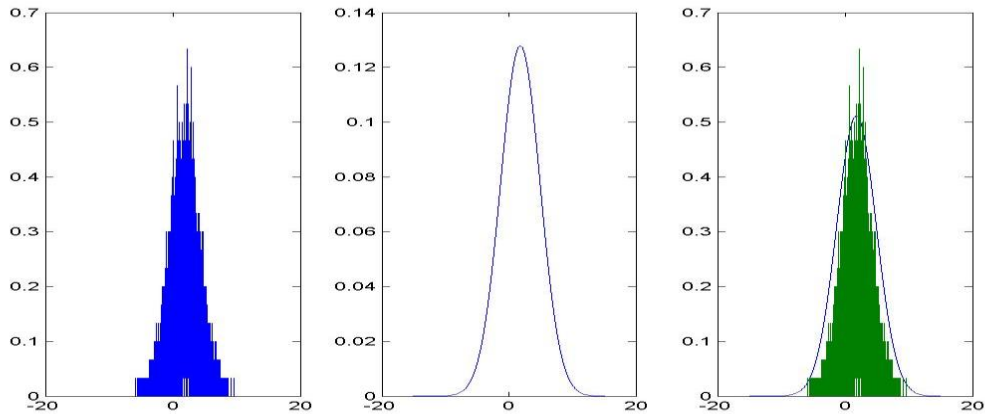
Une fois que les $\{m_j\}$ ont été obtenus, on peut calculer σ_j sous la forme :

$$\begin{aligned}
\frac{\partial Q(\theta, \theta')}{\partial \sigma_j} = 0 &\Rightarrow \sum_{i=1}^n \hat{z}_{i,j} \cdot \left[-\frac{1}{\sigma_j} + \frac{(\xi_i - m_j)^2}{\sigma_j^3} \right] = 0, \quad j = 1, \dots, J \\
&\Rightarrow \sigma_j^2 \cdot \sum_{i=1}^n \hat{z}_{i,j} = \sum_{i=1}^n \hat{z}_{i,j} \cdot (\xi_i - m_j)^2, \quad j = 1, \dots, J \\
&\sigma_j^2 = \frac{\sum_{i=1}^n \hat{z}_{i,j} \cdot (\xi_i - m_j)^2}{\sum_{i=1}^n \hat{z}_{i,j}}, \quad j = 1, \dots, J \quad (12)
\end{aligned}$$

L'algorithme va alors itérer entre ces deux étapes. Pour le choix du nombre de mixtures J on peut adopter le principe dit de la longueur minimale de description qui prend alors la forme :

$$\hat{J}_{MDL} = \operatorname{argmin}_J \left\{ -\log p_{\theta,J}(\mathcal{E}) + \frac{J}{2} \cdot \log(n) \right\} \quad (13)$$

L'algorithme est donné ci-dessous en langage Matlab ainsi qu'un exemple d'exécution pour un mélange de lois Gaussiennes initial.



```

function [MQ, MJ] = ExpectationMjp_Gauss
Nbsamples=100000;
%xx=zeros(1,Nbsamples);
% xx=load('wishart.c')
for ii=1:Nbsamples
    mm=rand(1,1);
    switch mm
        case (mm <=0.1)
            xx(1,ii)=sqrt(4.3)*randn(1,1)+2.3;
        case ((mm > 0.1) & (mm <= 0.3))
            xx(1,ii)=sqrt(1.55)*randn(1,1)+0.85;
        case ((mm > 0.3) & (mm <= 0.8))
            xx(1,ii)=sqrt(3.85)*randn(1,1)+6.85;
        case ((mm > 0.8) & (mm <= 0.9))
            xx(1,ii)=sqrt(2.8)*randn(1,1)+4.4;
        otherwise
            xx(1,ii)=sqrt(3.1)*randn(1,1)+1.8;
    end
end
xa=[-15:30/Nbsamples:15-30/Nbsamples];
%xa;
% for ii=1:Nbsamples
%     ii
%     xa(ii)=(ii-Nbsamples/2)*(30/Nbsamples);
% end;
%HIST(xx)
zu=(1/30)*HIST(xx,xa);
subplot(1,2,1);HIST(xx);
subplot(1,2,2);plot(xa,zu);
size_intp = 6;
max_iter = 10;
za=zeros(length(xx),size_intp);
moy= [1.4 2.5 3.6 4.2 5.1 2.4];
var=[1.2 1.4 2.3 3.1 4.2 5.6];
pond=[0.1 0.2 0.25 0.25 0.1 0.1];
length(xx);
for mm=1:max_iter;
    mm
    mid=zeros(1,length(xx));
    for ii=1:length(xx)
        for jj=1:size_intp
            mid(1,ii)=mid(1,ii)+pond(1,jj)*(1/sqrt(2*pi*var(1,jj)^2))*exp(-(xx(1,ii)-
moy(1,jj))^2/(2*var(1,jj)^2));
        end;
    end
    mid
    %moy
    for ii=1:length(xx)
        for jj=1:size_intp

```

```

        za(ii,jj)=pond(1,jj)*(1/sqrt(2*pi*var(1,jj)^2))*exp(-(xx(1,ii)-
moy(1,jj))^2/(2*var(1,jj)^2));
    end;
    for jj=1:size_intp
        za(ii,jj)=za(ii,jj)/mid(1,ii);
    end;
end;
za
for jj=1:size_intp
    pond(1,jj)=(1/length(za))*sum(za(:,jj),1);
end
pond
moy=zeros(1,size_intp);
for jj=1:size_intp
    for ii=1:length(xx)
        moy(1,jj)=moy(1,jj)+za(ii,jj)*xx(1,ii);
    end;
    moy(1,jj)=moy(1,jj)/sum(za(:,jj),1);
end;
moy
var=zeros(1,size_intp);
for jj=1:size_intp
    for ii=1:length(xx)
        var(1,jj)=var(1,jj)+za(ii,jj)*(xx(1,ii)-moy(1,jj))^2;
    end;
    var(1,jj)=var(1,jj)/sum(za(:,jj),1);
end;
end
moy
var
pond

MQ=pond;
MJ=moy;
MK=var;
dit=zeros(1,Nbsamples);
for ii=1:Nbsamples
    for jj=1:size_intp
        dit(1,ii)=dit(1,ii)+MQ(1,jj)*(1/sqrt(2*pi*MK(1,jj)^2))*exp(-(xa(ii)-
MJ(1,jj))^2/(2*MK(1,jj)^2));
    end;
end;
dit;
HIST(xx);
zu=(1/30)*HIST(xx,xa)
subplot(1,3,1);plot(xa,zu);
subplot(1,3,2);plot(xa,dit(1,:));
subplot(1,3,3);plot(xa,4*dit(1,:),xa,zu);

```

- Cas d'un mélange de lois Gamma:

On veut réutiliser l'algorithme précédent dans le cas de lois Gamma de la forme :

$$f(x) = K.x^\alpha .e^{-\beta.x}, x > 0, \alpha > 0, \beta > 0 \quad (14)$$

La constante de normalisation K se calcule avec : $\int_0^{+\infty} f(x).dx = 1$, ce qui donne :

$$K. \int_0^{+\infty} x^\alpha .e^{-\beta x}.dx = 1 ; \text{ soit : } K. \int_0^{+\infty} \left(\frac{u}{\beta}\right)^\alpha .e^{-u} .\frac{du}{\beta} = 1 ; \text{ ce qui donne : } K = \frac{\beta^{\alpha+1}}{\Gamma(\alpha+1)} \text{ avec, par}$$

$$\text{définition : } \Gamma(\alpha+1) = \int_0^{+\infty} u^\alpha .e^{-u} .du .$$

Le mélange de lois Gamma prend donc ici la forme :

$$f_C(x) = \sum_{j=1}^J \pi_j . \frac{\beta_j}{\Gamma(\alpha_j + 1)} . x^{\alpha_j} . e^{-\beta_j . x} = \sum_{j=1}^J \pi_n . \gamma(\alpha_j, \beta_j)$$

En reprenant la même démarche que précédemment pour le mélange de Gaussiennes, on arrive ici à (en appelant $\gamma(x, \alpha, \beta)$ la pdf d'une loi Gamma).

$$\begin{aligned} \hat{\theta} &= \arg \max_{\theta: \sum_{j=1}^J \pi_j = 1} \log p_{\theta}(\Xi) \\ &= \arg \max_{\theta: \sum_{j=1}^J \pi_j = 1} \sum_{i=1}^n \log \sum_{j=1}^J \pi_j . \gamma(\xi_i; m_j, \sigma_j^2) \end{aligned} \quad (15)$$

De même :

$$\begin{aligned} \log p_{\theta}(\Xi, \mathbf{Z}) &= \sum_{i=1}^n \sum_{j=1}^J \hat{z}_{i,j} . \log \pi_j + \sum_{i=1}^n \sum_{j=1}^J \hat{z}_{i,j} . [(\alpha_j + 1) . \log(\beta_j) - \log(\Gamma(\alpha_j + 1)) + \dots \\ &\quad + \alpha_j . \log(\xi_i) - \beta_j . \xi_i + C \end{aligned} \quad (16)$$

Pour le calcul des paramètres α_j, β_j , on calcule les dérivées partielles :

$$\begin{aligned} \frac{\partial \log p_{\theta}(\Xi, \mathbf{Z})}{\partial \alpha_j} &= \sum_{i=1}^n \hat{z}_{i,j} . [\log(\beta_j) - \frac{\partial \Gamma(\alpha_j + 1) / \partial \alpha_j}{\Gamma(\alpha_j + 1)} + \log(\xi_i)] \\ \frac{\partial \log p_{\theta}(\Xi, \mathbf{Z})}{\partial \alpha_j} &= 0 \Rightarrow \log(\beta_j) . \sum_{i=1}^n \hat{z}_{i,j} = \sum_{i=1}^n (1 - \log(\xi_i)) \hat{z}_{i,j} \end{aligned}$$

D'où, finalement :

$$\log(\beta_j) = \frac{\sum_{i=1}^n (1 - \log(\xi_i)) \hat{z}_{i,j}}{\sum_{i=1}^n \hat{z}_{i,j}} \quad (17)$$

En effet, on a : $\Gamma(\alpha_j + 1) = \int_0^{+\infty} u^{\alpha_j} . e^{-u} . du = \alpha_j . \Gamma(\alpha_j)$ par une simple intégration par parties.

et $\frac{\partial \Gamma(\alpha_j + 1)}{\partial \alpha_j} = \alpha_j . \int_0^{+\infty} u^{\alpha_j - 1} . e^{-u} . du = \alpha_j . \Gamma(\alpha_j)$

$$\frac{\partial \log p_{\theta}(\Xi, \mathbf{Z})}{\partial \beta_j} = \sum_{i=1}^n \hat{z}_{i,j} . [(\alpha_j + 1) . \frac{1}{\beta_j} - \xi_i]$$

$$\frac{\partial \log p_{\theta}(\Xi, \mathbf{Z})}{\partial \beta_j} = 0 \Rightarrow \alpha_j = \frac{\beta_j . \sum_{i=1}^n \xi_i . \hat{z}_{i,j} - \sum_{i=1}^n \hat{z}_{i,j}}{\sum_{i=1}^n \hat{z}_{i,j}} \quad (18)$$

L'algorithme est donné ci-dessous en langage Matlab ainsi qu'un exemple d'exécution pour une distribution de la trace de l'inverse d'une matrice de Wishart. .

```
function [MQ, MJ] = ExpectationM_loigama
Nbsamples=100000;
%xx=zeros(1,Nbsamples);
xx=load('wishart.c')
%for ii=1:Nbsamples
%mm=rand(1,1);
%switch mm
%case (mm <=0.1)
%xx(1,ii)=abs(sqrt(4.3)*randn(1,1)+2.3);
%case ((mm > 0.1) & (mm <= 0.3))
%xx(1,ii)=abs(sqrt(1.55)*randn(1,1)+0.85);
%case ((mm > 0.3) & (mm <= 0.8))
%xx(1,ii)=abs(sqrt(3.85)*randn(1,1)+6.85);
%case ((mm > 0.8) & (mm <= 0.9))
%xx(1,ii)=abs(sqrt(2.8)*randn(1,1)+4.4);
%otherwise
%xx(1,ii)=abs(sqrt(3.1)*randn(1,1)+1.8);
%end
%end
step=0.001;
lim_inf=0.01;
lim_sup=6;
xa=[lim_inf:step:lim_sup];
Nbpoints=1+(lim_sup-lim_inf)/step;
length(xa);
%xa;
% for ii=1:Nbsamples
% ii
% xa(ii)=(ii-Nbsamples/2)*(30/Nbsamples);
% end;
%HIST(xx)
zu=(100/(2*Nbsamples))*HISTC(xx,xa);
subplot(1,2,1);HIST(xx);
subplot(1,2,2);plot(xa,zu);
```

```

size_intp = 6;
max_iter = 10;
za=zeros(length(xx),size_intp);
alpha = [9.4 3.5 3.6 4.2 5.1 12.4];
beta =[5.3 8.4 2.3 3.1 4.2 5.6];
pond=[0.1 0.2 0.25 0.25 0.1 0.1];
length(xx);
for mm=1:max_iter;
    mm
    mid=zeros(1,length(xx));
    for ii=1:length(xx)
        for jj=1:size_intp

mid(1,ii)=mid(1,ii)+pond(1,jj)*(beta(1,jj)^(1+alpha(1,jj)))*(xx(1,ii)^(alpha(1,jj)))*(exp(-
beta(1,jj)*xx(1,ii)))/(gamma(1+alpha(1,jj)));
            end;
        end;
        mid;
        %moy
        for ii=1:length(xx)
            for jj=1:size_intp
                za(ii,jj)=pond(1,jj)*(beta(1,jj)^(1+alpha(1,jj)))*(xx(1,ii)^(alpha(1,jj)))*(exp(-
beta(1,jj)*xx(1,ii)))/(gamma(1+alpha(1,jj)));
            end;
            for jj=1:size_intp
                za(ii,jj)=za(ii,jj)/mid(1,ii);
            end;
        end;
        za;
        for jj=1:size_intp
            pond(1,jj)=(1/length(za))*sum(za(:,jj),1);
        end
        pond;
        zeta=zeros(1,size_intp);
        for jj=1:size_intp
            for ii=1:length(xx)
                zeta(1,jj)=zeta(1,jj)+za(ii,jj)*(1-log(xx(1,ii)));
            end;
            zeta(1,jj)=zeta(1,jj)/sum(za(:,jj),1);
        end;
        for jj=1:size_intp
            beta(1,jj)=exp(zeta(1,jj));
        end
        alpha=zeros(1,size_intp);
        for jj=1:size_intp
            for ii=1:length(xx)
                alpha(1,jj)=alpha(1,jj)+beta(1,jj)*za(ii,jj)*xx(1,ii)-za(ii,jj);
            end;
            alpha(1,jj)=alpha(1,jj)/sum(za(:,jj),1);
        end;
    end;
end;

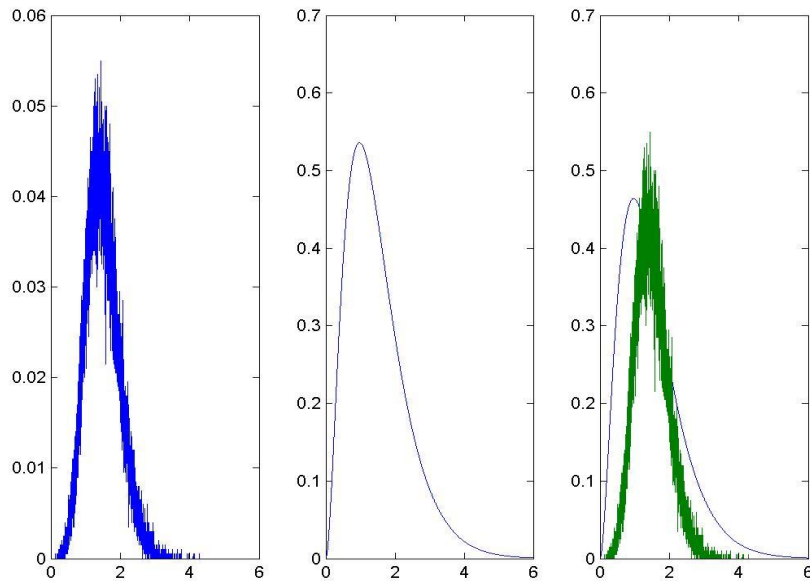
```

```

end
alpha
beta
pond

MQ=pond;
MJ=alpha;
MK=beta;
dit=zeros(1,Nbpoints);
for ii=1:Nbpoints
    for jj=1:size_intp
        dit(1,ii)=dit(1,ii)+MQ(1,jj)*(MK(1,jj)^(1+MJ(1,jj)))*(xa(1,ii)^(MJ(1,jj)))*(exp(-
MK(1,jj)*xa(1,ii)))/(gamma(1+MJ(1,jj)));
    end;
end;
dit;
HIST(xx);
%zu=HIST(xx,xa);
subplot(1,3,1);plot(xa,zu);
subplot(1,3,2);plot(xa,dit(1,:));
subplot(1,3,3);plot(xa,(0.45/0.52)*dit(1,:),xa,10*zu);

```



- Cas d'un mélange de lois Nakagami- m :

On veut réutiliser l'algorithme précédent dans le cas de lois Nakagami- m de la forme :

$$f(x) = \frac{2 \cdot \mu^\mu}{\Gamma(\mu) \cdot \omega^\mu} x^{2\mu-1} \cdot e^{-\frac{\mu}{\omega} x^2}, \mu > 0, \omega > 0 \quad (19)$$

Le mélange de lois Nakagami- m prend alors la forme :

$$f_C(x) = \sum_{j=1}^J \pi_j \frac{2 \cdot \mu_j^{\mu_j}}{\Gamma(\mu_j) \cdot \omega_j^{\mu_j}} x^{2 \cdot \mu_j - 1} \cdot e^{-\frac{\mu_j}{\omega_j} \cdot x^2} \quad (20)$$

On trouve maintenant :

$$\begin{aligned} \log p_{\theta}(\Xi, \mathbf{Z}) = & \sum_{i=1}^n \sum_{j=1}^J \hat{z}_{i,j} \cdot \log \pi_j + \sum_{i=1}^n \sum_{j=1}^J z_{i,j} \cdot [\log(2 \cdot \mu_j^{\mu_j}) - \log(\Gamma(\mu_j)) + \dots \\ & - \mu_j \cdot \log(\omega_j) + (2 \cdot \mu_j - 1) \cdot \log(\xi_i) - \frac{\mu_j}{\omega_j} \cdot \xi_i^2] \end{aligned} \quad (21)$$

Pour le calcul des dérivées partielles, on trouve :

$$\begin{aligned} \frac{\partial \log p_{\theta}(\Xi, \mathbf{Z})}{\partial \omega_j} &= \sum_{i=1}^n \left(-\frac{\mu_j}{\omega_j} + \frac{\mu_j}{\omega_j^2} \cdot \xi_i^2 \right) \cdot \hat{z}_{i,j} \\ \omega_j &= \frac{\sum_{i=1}^n \xi_i^2 \cdot \hat{z}_{i,j}}{\sum_{i=1}^n \hat{z}_{i,j}} \end{aligned} \quad (22)$$

Et :

$$\begin{aligned} \frac{\partial \log p_{\theta}(\Xi, \mathbf{Z})}{\partial \mu_j} &= \sum_{i=1}^n \left(\log(\mu_j) + 1 - 1 - \log(\omega_j) + 2 \cdot \log(\xi_i) - \frac{\xi_i^2}{\omega_j} \right) \cdot \hat{z}_{i,j} \\ \frac{\partial \log p_{\theta}(\Xi, \mathbf{Z})}{\partial \mu_j} = 0 &\Rightarrow \log(\mu_j) = \frac{\sum_{i=1}^n \hat{z}_{i,j} \cdot (\log(\omega_j) - 2 \cdot \log(\xi_i) + \xi_i^2 / \omega_j)}{\sum_{i=1}^n \hat{z}_{i,j}} \end{aligned} \quad (23)$$