MESSL

jmatthieu

18 janvier 2016

Ce document accompagne l'article MESSL, il ne présente cependant pas tout. Les notations sont les mêmes sauf lorsque celà est précisé. Les équations du papier MESSL sont notées (mN) où N correspond au numéro de l'équation dans le papier MESSL.

Le modèle (noté Λ au lieu de Θ) que l'on veut estimer a pour paramètres (voir *III. Parameter estimation from mixtures* et l'équation (m10) pour le détail) :

$$\Lambda = \{\xi(\omega), \sigma_{i,\tau}(\omega), \mu_i(\omega), \nu_i(\omega), \psi_{i,\tau}\}$$

La probabilité de l'observation $\phi(\omega, t); \alpha(\omega, t)$ connaissant les paramètres du modèle Λ est donnée par :

$$p(\phi(\omega, t), \alpha(\omega, t)|\Lambda)$$

La vraisemblance est donnée par (par définition) :

$$L(\Lambda) = \prod_{\omega,t} p(\phi(\omega, t), \alpha(\omega, t)|\Lambda)$$

Et la log-vraisemblance par (par définition) :

$$\begin{aligned}
\mathcal{L}(\Lambda) &= \log L(\Lambda) \\
&= \log \prod_{\omega,t} p(\phi(\omega, t), \alpha(\omega, t)|\Lambda) \\
&= \sum_{\omega,t} \log p(\phi(\omega, t), \alpha(\omega, t)|\Lambda) \qquad (1)
\end{aligned}$$

où l'équation (1) correspond à l'équation (m12).

La variable cachée $z_{i,\tau}(\omega, t)$ correspond à la classe de chaque observation, c'est à dire au couple $(i\,;\,\tau)$ dont est issue l'observation. $z_{i,\tau}(\omega, t) = 1$ si l'observation est issue du couple $(i\,;\,\tau)$, zéro autrement. En marginalisant sur la variable cachée $z_{i,\tau}(\omega,t)$ puis en appliquant la loi de Bayes on obtient à partir de l'équation (1)

$$\begin{aligned}
&p(\phi(\omega,t),\alpha(\omega,t)|\Lambda)\\
=&\sum_{i,\tau} p(\phi(\omega,t),\alpha(\omega,t),z_{i,\tau}(\omega,t)|\Lambda)\\
=&\sum_{i,\tau} p(\phi(\omega,t),\alpha(\omega,t)|z_{i,\tau}(\omega,t),\Lambda)\cdot p(z_{i,\tau}(\omega,t)|\Lambda) \qquad (2)
\end{aligned}$$

En injectant (2) dans (1) on obtient :

$$\mathcal{L}(\Lambda)=\sum_{\omega,t}\log\sum_{i,\tau} p(\phi(\omega,t),\alpha(\omega,t)|z_{i,\tau}(\omega,t),\Lambda)\cdot p(z_{i,\tau}(\omega,t)|\Lambda) \qquad (3)$$

On suppose ϕ et α indépendants, et on modélise leur probabilité connaissant $z_{i,\tau}$ et Λ par une gaussienne, on a donc :

$$\begin{aligned}
&p(\phi(\omega,t),\alpha(\omega,t)|z_{i,\tau}(\omega,t),\Lambda)\\
=&p(\phi(\omega,t)|z_{i,\tau}(\omega,t),\Lambda)\cdot p(\alpha(\omega,t)|z_{i,\tau}(\omega,t),\Lambda)\\
=&\mathcal{N}(\hat{\phi}(\omega,t;\tau)|\xi_{i,\tau}(\omega),\sigma^2_{i,\tau}(\omega))\cdot\mathcal{N}(\alpha(\omega,t;\tau)|\mu_{i,\tau}(\omega),\eta^2_{i,\tau}(\omega)) \qquad (4)
\end{aligned}$$

On pose :

$$\psi_{i,\tau}\equiv p(z_{i,\tau}(\omega,t)|\Lambda) \qquad (5)$$

Et on retombe sur (m13) en injectant (4) et (5) dans (3)

1 fonction auxiliaire

Pour alléger la notation dans cette section nous écrirons $\phi(\omega,t)$, $\alpha(\omega,t)$ et $z_{i,\tau}$ sans leurs dépendances en (ω,t).

On cherche la différence entre la logvraisemblance actuelle et celle calculée au pas précédent s.

$$\begin{aligned}
&\mathcal{L}(\Lambda)-\mathcal{L}(\Lambda_s)\\
=&\sum_{\omega,t}\log\sum_{i,\tau} p(\phi,\alpha|z_{i,\tau},\Lambda)\cdot p(z_{i,\tau}|\Lambda)-\sum_{\omega,t}\log p(\phi,\alpha|\Lambda_s)\\
=&\sum_{\omega,t}\log\sum_{i,\tau} p(\phi,\alpha|z_{i,\tau},\Lambda)\cdot p(z_{i,\tau}|\Lambda)\times\frac{p(z|\omega,\alpha,\Lambda_s)}{p(z|\omega,\alpha,\Lambda_s)}-\sum_{\omega,t}\log p(\phi,\alpha|\Lambda_s)\\
=&\sum_{\omega,t}\log\sum_{i,\tau} p(z|\omega,\alpha,\Lambda_s)\times\frac{p(\phi,\alpha|z_{i,\tau},\Lambda)\cdot p(z_{i,\tau}|\Lambda)}{p(z|\omega,\alpha,\Lambda_s)}-\sum_{\omega,t}\log p(\phi,\alpha|\Lambda_s)\\
\geq&\sum_{\omega,t}\sum_{i,\tau} p(z|\omega,\alpha,\Lambda_s)\times\log\frac{p(\phi,\alpha|z_{i,\tau},\Lambda)\cdot p(z_{i,\tau}|\Lambda)}{p(z|\omega,\alpha,\Lambda_s)}-\sum_{\omega,t}\log p(\phi,\alpha|\Lambda_s) \qquad (6)
\end{aligned}$$

(6) correspond à la formule de Jensen. On rappelle que

$$\sum_{i,\tau} p(z|\omega, \alpha, \Lambda_s) = 1 \tag{7}$$

on peut donc factoriser et d'où (6) est égale à :

$$\begin{aligned}
&\sum_{\omega,t} \sum_{i,\tau} p(z|\omega, \alpha, \Lambda_s) \times \log \frac{p(\phi, \alpha|z_{i,\tau}, \Lambda) \cdot p(z_{i,\tau}|\Lambda)}{p(z|\omega, \alpha, \Lambda_s)} - \sum_{\omega,t} \sum_{i,\tau} p(z|\omega, \alpha, \Lambda_s) \log p(\phi, \alpha|\Lambda_s) \\
=& \sum_{\omega,t} \sum_{i,\tau} p(z|\omega, \alpha, \Lambda_s) \times [\log \frac{p(\phi, \alpha|z_{i,\tau}, \Lambda) \cdot p(z_{i,\tau}|\Lambda)}{p(z|\omega, \alpha, \Lambda_s)} - \log p(\phi, \alpha|\Lambda_s)] \\
=& \qquad (8)
\end{aligned}$$