Examen DEA SICOM, Cours Identification
Luc Pronzato, Décembre 2003

Quelques indications : les parties 1, 2, 3 sont indépendantes du reste ; 6 est un cas particulier d'un exemple traité en cours ; 4 est un cas particulier de ce cas particulier, déjà donné comme exercice ; 5 a également déjà été proposé comme exercice ; 7 et 8 se résolvent comme 4, 5 et 6; la partie 9 est indépendante de tout le reste.

On prendra soin de **souligner les vecteurs** (lettres minuscules en **gras** dans l'énoncé) et de noter les matrices par des lettres capitales.

A chaque fois que l'on parle d'estimateur, c'est du **maximum de vraisemblance** qu'il s'agit. On note avec un ˆ les valeurs estimées.

1)

Soit un modèle de Box et Jenkins correspondant à l'équation de récurrence

$$y(k) = F(q)u(k) + G(q)\epsilon(k)\,,$$

dans laquelle les fonctions de transfert $F(q)$ et $G(q)$ sont des fractions rationnelles en l'opérateur retard q^{-1}. Montrer que cette équation de récurrence peut également être représentée par une structure de type ARMAX, pour laquelle $A(q)y(k) = B(q)u(k) + C(q)\epsilon(k)$ avec $A(q)$, $B(q)$ et $C(q)$ des polynômes en q^{-1} (on explicitera ces polynômes en fonction des numérateurs et dénominateurs de $F(q)$ et $G(q)$).

Dans toute la suite on considère un modèle multivariable, avec deux sorties y_1 et y_2. On suppose que les deux sorties peuvent être décrites comme les réponses de deux modèles de type ARMAX, avec une entrée commune $u(k)$ et deux sources de bruit distinctes $\epsilon_1(k)$ et $\epsilon_2(k)$,

$$\begin{aligned} A_1(q)y_1(k) &= B_1(q)u(k) + C_1(q)\epsilon_1(k) \\ A_2(q)y_2(k) &= B_2(q)u(k) + C_2(q)\epsilon_2(k)\,. \end{aligned}$$

On suppose que

$$\begin{aligned} A_1(q) &= 1 - q^{-1} - a_1 q^{-2} \\ B_1(q) &= q^{-1} + b_1 q^{-2} \\ C_1(q) &= 1 + c_1 q^{-1} \\ A_2(q) &= 1 - q^{-1} - a_2 q^{-2} \\ B_2(q) &= q^{-1} + b_2 q^{-2} \\ C_2(q) &= 1 \end{aligned}$$

On notera $\boldsymbol{\theta} = (a_1,b_1,c_1,a_2,b_2)^\top$, $\boldsymbol{\theta}_1 = (a_1,b_1,c_1)^\top$ et $\boldsymbol{\theta}_2 = (a_2,b_2)^\top$.

2)

Écrire les équations de récurrence satisfaites par $y_1(k)$ et $y_2(k)$.

On suppose que $-1/4 < a_1 < 0$ et $-1/4 < a_2 < 0$. Quelle connaissance sur le système peut justifier cette hypothèse ?

3)

Donner les équations de récurrence satisfaites par les erreurs de prédiction $e_1(\boldsymbol{\theta},k)$ et $e_2(\boldsymbol{\theta},k)$, faisant intervenir les deux couples d'entrée-sortie $(u(k),y_1(k))$ et $(u(k),y_2(k))$, $k = 1,\ldots,N$; ces erreurs de prédiction vont être utilisées pour estimer $\boldsymbol{\theta}$ au sens du maximum de vraisemblance. **Dans toute la suite on notera simplement $e_1(k)$ et $e_2(k)$ ces erreurs (sans oublier pour autant qu'elle dépendent des paramètres...), et $\mathbf{e}(k) = (e_1(k),e_2(k))^\top$.**

4)

On suppose que $\epsilon_1(k)$ et $\epsilon_2(k)$ forment deux suites de variables aléatoires indépendantes, aussi mutuellement indépendantes, distribuées respectivement suivant les lois normales $\mathcal{N}(0,\sigma_1^2)$ et $\mathcal{N}(0,\sigma_2^2)$.

On veut estimer $\boldsymbol{\theta}$, σ_1^2 et σ_2^2 au sens du maximum de vraisemblance. Construire la log-vraisemblance $\log \pi(\mathbf{y}|\boldsymbol{\theta},\sigma_1^2,\sigma_2^2)$ d'une suite d'observations $\mathbf{y} = (\mathbf{y}(1),\ldots,\mathbf{y}(N))$ avec $\mathbf{y}(k) = (y_1(k),y_2(k))^\top$. Montrer qu'elle se décompose en deux critères indépendants, l'un fonction de $\boldsymbol{\theta}_1$ et σ_1^2, l'autre fonction de $\boldsymbol{\theta}_2$ et σ_2^2.

On notera $\hat{e}_1(k)$ les erreurs de prédiction obtenues pour $\boldsymbol{\theta}_1 = \hat{\boldsymbol{\theta}}_1$, l'estimateur de $\boldsymbol{\theta}_1$ au sens du maximum de vraisemblance. Donner l'expression de l'estimateur $\hat{\sigma}_1^2$ de σ_1^2 à partir des $\hat{e}_1(k)$, $k = 1,\ldots,N$.

Donner l'expression du critère à minimiser pour trouver $\hat{\boldsymbol{\theta}}_1$ (en fonction de la suite des erreurs $e_1(k)$).

Faire de même pour $\hat{\boldsymbol{\theta}}_2$ et $\hat{\sigma}_2^2$.

5)

On caractérise la précision de l'estimation de σ_i^2, $i = 1{,}2$ par $1/V_i$, avec V_i l'espérance conditionnelle

$$V_i = \mathtt{E}\left\{ \left[\frac{\partial \log \pi(\mathbf{y}|\boldsymbol{\theta},\sigma_1^2,\sigma_2^2)}{\partial \sigma_i^2}\right]^2 \middle| \boldsymbol{\theta},\sigma_1^2,\sigma_2^2 \right\} .$$

Justifier ce choix. Calculer la valeur de V_i en fonction de N/σ_i^4, $i = 1{,}2$.

On rappelle que pour x,y,z,t des variables gaussiennes,

$$\mathtt{E}\{xyzt\} = \mathtt{E}\{xy\}\mathtt{E}\{zt\} + \mathtt{E}\{xz\}\mathtt{E}\{yt\} + \mathtt{E}\{xt\}\mathtt{E}\{yz\} .$$

6)

On suppose à présent que les vecteurs $\boldsymbol{\epsilon}(k) = (\epsilon_1(k),\epsilon_2(k))^\top$ forment une suite i.i.d. suivant une loi normale de moyenne $\mathbf{0} = (0{,}0)^\top$ de matrice de covariance $\boldsymbol{\Sigma}$. On rappelle que la densité de cette loi est donnée par

$$\pi_\epsilon(\boldsymbol{\epsilon}) = \frac{1}{2\pi\sqrt{\det \boldsymbol{\Sigma}}} \exp\left(-\frac{1}{2}\,\boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\epsilon}\right) .$$

Calculer la log-vraisemblance $\log \pi(\mathbf{y}|\boldsymbol{\theta},\boldsymbol{\Sigma})$. On notera $\hat{\mathbf{e}}(k)$ les erreurs de prédiction obtenues pour $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$, l'estimateur de $\boldsymbol{\theta}$ au sens du maximum de vraisemblance. Donner l'expression de l'estimateur $\hat{\boldsymbol{\Sigma}}$ de $\boldsymbol{\Sigma}$ à partir des $\hat{\mathbf{e}}(k)$, $k = 1,\ldots,N$. On rappelle que $d(\log\det \boldsymbol{\Sigma})/d\boldsymbol{\Sigma} = \boldsymbol{\Sigma}^{-1}$ et $d(\mathbf{x}^\top\boldsymbol{\Sigma}^{-1}\mathbf{x})/d\boldsymbol{\Sigma} = -\boldsymbol{\Sigma}^{-1}\mathbf{x}\mathbf{x}^\top\boldsymbol{\Sigma}^{-1}$, $\forall \mathbf{x}$.

Donner l'expression du critère à minimiser pour trouver $\hat{\boldsymbol{\theta}}$ en fonction de la suite des erreurs $\mathbf{e}(k)$, $k = 1,\ldots,N$.

7)

On suppose que $\boldsymbol{\Sigma}$ a la forme suivante

$$\boldsymbol{\Sigma} = \sigma^2 \begin{pmatrix} 1 & \alpha \\ \alpha & 1 \end{pmatrix} .$$

Donner l'expression de la log-vraisemblance $\log \pi(\mathbf{y}|\boldsymbol{\theta},\sigma^2,\alpha)$. On notera

$$S_N = \sum_{k=1}^{N} e_1^2(k) + e_2^2(k)\,,\; P_N = \sum_{k=1}^{N} e_1(k)e_2(k) \text{ et } \hat{S}_N = \sum_{k=1}^{N} \hat{e}_1^2(k) + \hat{e}_2^2(k)\,,\; \hat{P}_N = \sum_{k=1}^{N} \hat{e}_1(k)\hat{e}_2(k)$$

avec $\hat{e}_1(k)$ et $\hat{e}_2(k)$ les erreurs pour $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$.

Exprimer $\hat{\sigma}^2$, estimateur de σ^2, en fonction de $\hat{\alpha}$, estimateur de α, et de $\hat{S}_N,\hat{P}_N$. Reporter dans l'expression de $\log \pi(\mathbf{y}|\boldsymbol{\theta},\sigma^2,\alpha)$ trouvée précédemment, et donner l'expression de $\log \pi(\mathbf{y}|\boldsymbol{\theta},\alpha)$ à maximiser par rapport à $\boldsymbol{\theta}$ et α.

Exprimer l'estimateur $\hat{\alpha}$ en fonction de $\hat{S}_N$ et $\hat{P}_N$. Quelles sont les valeurs minimale et maximale possibles pour $\hat{\alpha}$? En quoi est-ce naturel ?

Reporter $\hat{\alpha}$ dans l'expression de $\hat{\sigma}^2$ trouvée précédemment, et exprimer $\hat{\sigma}^2$ en fonction de $\hat{S}_N/N$.

Reporter $\hat{\alpha}$ dans $\log \pi(\mathbf{y}|\boldsymbol{\theta},\alpha)$ et donner enfin l'expression du critère à minimiser pour trouver $\hat{\boldsymbol{\theta}}$, en fonction des $e_1(k)$ et $e_2(k)$, $k = 1,\ldots,N$.

8)

De même que pour la question 5, on caractérise la précision de l'estimation de σ^2 par $1/V$, avec

$$V = \mathtt{E}\left\{ \left[\frac{\partial \log \pi(\mathbf{y}|\boldsymbol{\theta},\sigma^2,\alpha)}{\partial \sigma^2}\right]^2 \middle| \boldsymbol{\theta},\sigma^2,\alpha \right\} .$$

Calculer V quand α prend la valeur $\alpha = 0$ (normalement, la dérivée $\partial \log \pi(\mathbf{y}|\boldsymbol{\theta},\sigma^2,\alpha)/\partial\sigma^2$ a déjà été calculée à la question 7).

Comparer avec le résultat de la question 5 concernant la précision de l'estimation de σ_1^2 et σ_2^2, et justifier ce résultat.

9)

On effectue une série de $N = 1000$ observations. La structure de modèle présentée en 1) a conduit à la valeur $\log \pi(\mathbf{y}|\hat{\boldsymbol{\theta}},\hat{\sigma}^2,\hat{\alpha}) = 23{,}9$. En supposant que $c_1 = 0$, c'est-à-dire en estimant seulement $\boldsymbol{\theta}' = (a_1,b_1,a_2,b_2)^\top$, σ^2 et α, on trouve $\log \pi(\mathbf{y}|\hat{\boldsymbol{\theta}}',\hat{\sigma}^2,\hat{\alpha}) = 23{,}2$. Quelle structure de modèle faut-il retenir suivant le critère AIC d'Akaike ?