**Examen DEA SICOM, Cours Identification**
**Luc Pronzato, Décembre 2003**

CORRIGÉ

**1)**

Posons $G = N_G/D_G$, $F = N_F/D_F$, on a $y(k) = N_F(q)/D_F(q)u(k) + N_G(q)/D_G(q)\epsilon(k)$, soit $A(q)y(k) = B(q)u(k) + C(q)\epsilon(k)$, avec $A(q) = D_F(q)D_G(q)$, $B(q) = N_F(q)D_G(q)$ et $C(q) = D_F(q)N_G(q)$.

**2)**

$$
\begin{aligned}
y_1(k) &= y_1(k-1) + a_1 y_1(k-2) + u(k-1) + b_1 u(k-2) + \epsilon_1(k) + c_1\epsilon(k-1) \\
y_2(k) &= y_2(k-1) + a_2 y_2(k-2) + u(k-1) + b_2 u(k-2) + \epsilon_2(k)
\end{aligned}
$$

La stabilité du système dépend des valeurs de $a_1$ et $a_2$ : le système est stable si les valeurs de $z$ pour lesquelles $A_1(z) = 1 - z^{-1} - a_1 z^{-2}$ et $A_2(z) = 1 - z^{-1} - a_2 z^{-2}$ s'annulent sont de module inférieur à 1. On a $A_i(z) = z^{-2}(z^2 - z - a_i)$, $i = 1,2$, qui s'annule en $z = (1 \pm \sqrt{1+4a_i})/2$. Pour $-1/4 < a_i < 0$ le système est donc stable. De plus, les zéros des $A_i(z)$ sont réels et positifs, la réponse indicielle du système est donc non-oscillante.

**3)**

$$
\begin{aligned}
e_1(k) &= -c_1\epsilon(k-1) + y_1(k) - y_1(k-1) - a_1 y_1(k-2) - u(k-1) - b_1 u(k-2) \\
e_2(k) &= y_2(k) - y_2(k-1) - a_2 y_2(k-2) - u(k-1) - b_2 u(k-2)
\end{aligned}
$$

**4)**

$\log \pi(\mathbf{y}|\boldsymbol{\theta},\sigma_1^2,\sigma_2^2) = \log \pi_\epsilon(\mathbf{e}|\boldsymbol{\theta},\sigma_1^2,\sigma_2^2)$ avec $\mathbf{e} = (\mathbf{e}(1),\ldots,\mathbf{e}(N))$ et $\pi_\epsilon$ la densité de probabilité des erreurs. Du fait de l'indépendance,

$$
\begin{aligned}
\log \pi(\mathbf{y}|\boldsymbol{\theta},\sigma_1^2,\sigma_2^2) &= \sum_{i=1,2}\sum_{k=1}^{N} \log\left\{\frac{1}{\sigma_i\sqrt{2\pi}}\exp\left[-\frac{e_i^2(k)}{2\sigma_i^2}\right]\right\} \\
&= \sum_{i=1,2}\left\{-\frac{N}{2}\log(2\pi) - \frac{N}{2}\log(\sigma_i^2) - \frac{1}{2\sigma_i^2}\sum_{k=1}^{N} e_i^2(k)\right\}
\end{aligned}
$$

avec $e_i(k)$ dépendant de $\boldsymbol{\theta}_i$, $i = 1,2$.

Dérivons par rapport à $\sigma_i^2$, cette dérivée doit être nulle à l'optimum, c'est-à-dire pour la valeur estimée des paramètres. On obtient

$$
\hat{\sigma}_i^2 = \frac{1}{N}\sum_{k=1}^{N}\hat{e}_i^2(k)\,,\ i = 1,2\,.
$$

On reporte dans $\log \pi(\mathbf{y}|\boldsymbol{\theta},\sigma_1^2,\sigma_2^2)$, ce qui donne comme critère à maximiser

$$
\sum_{i=1,2}\left\{-\frac{N}{2}\log(2\pi) - \frac{N}{2}\log\left[\frac{1}{N}\sum_{k=1}^{N}e_i^2(k)\right] - \frac{N}{2}\right\}\,,
$$

c'est-à-dire, $\hat{\boldsymbol{\theta}}_i$ minimise $\sum_{k=1}^{N} e_i^2(k)$, $i = 1,2$.

**5)**

$V_i$ correspond au terme de la matrice d'information de Fisher associé au paramètre $\sigma_i^2$, $1/V_i$ est donc la variance asymptotique de l'estimateur de $\sigma_i^2$. La dérivée $\partial \log \pi(\mathbf{y}|\boldsymbol{\theta},\sigma_1^2,\sigma_2^2)/\partial\sigma_i^2$ a été calculée en 4, elle vaut

$$
\frac{\partial \log \pi(\mathbf{y}|\boldsymbol{\theta},\sigma_1^2,\sigma_2^2)}{\partial\sigma_i^2} = -\frac{N}{2\sigma_i^2} + \frac{1}{2\sigma_i^4}\sum_{k=1}^{N}e_i^2(k)\,.
$$

On a donc

$$
\mathrm{E}\left\{\left[\frac{\partial \log \pi(\mathbf{y}|\boldsymbol{\theta},\sigma_1^2,\sigma_2^2)}{\partial\sigma_i^2}\right]^2\middle|\,\boldsymbol{\theta},\sigma_1^2,\sigma_2^2\right\} = \frac{N^2}{4\sigma_i^4} - \frac{2N}{4\sigma_i^6}\mathrm{E}\left\{\sum_{k=1}^{N}e_i^2(k)\middle|\,\boldsymbol{\theta},\sigma_i^2\right\} + \frac{1}{4\sigma_i^8}\mathrm{E}\left\{\left[\sum_{k=1}^{N}e_i^2(k)\right]^2\middle|\,\boldsymbol{\theta},\sigma_i^2\right\}\,,
$$

avec

$$\mathtt{E}\left\{ e_i^2(k) \middle| \boldsymbol{\theta},\sigma_i^2 \right\} = \sigma_i^2 \,,\, \mathtt{E}\left\{ e_i^4(k) \middle| \boldsymbol{\theta},\sigma_i^2 \right\} = 3\sigma_i^4 \,,\ \text{ et pour } j \neq k \,,\, \mathtt{E}\left\{ e_i^2(k)e_i^2(j) \middle| \boldsymbol{\theta},\sigma_i^2 \right\} = \sigma_i^4 \,.$$

Par conséquent,

$$\begin{aligned} \mathtt{E}\left\{ \sum_{k=1}^{N} e_i^2(k) \middle| \boldsymbol{\theta},\sigma_i^2 \right\} &= N\sigma_i^2 \\ \mathtt{E}\left\{ \left[\sum_{k=1}^{N} e_i^2(k)\right]^2 \middle| \boldsymbol{\theta},\sigma_i^2 \right\} &= 3N\sigma_i^4 + N(N-1)\sigma_i^4 \end{aligned}$$

et finalement

$$V_i = \frac{N}{2\sigma_i^4} \,.$$

**6)**

On a, d'après l'indépendance des $\boldsymbol{\epsilon}(k)$,

$$\begin{aligned} \log \pi(\mathbf{y}|\boldsymbol{\theta},\boldsymbol{\Sigma}) &= \sum_{k=1}^{N} \log\left[\frac{1}{2\pi\sqrt{\det \boldsymbol{\Sigma}}} \exp\left(-\frac{1}{2}\,\mathbf{e}(k)^\top \boldsymbol{\Sigma}^{-1}\mathbf{e}(k)\right)\right] \\ &= -N\log(2\pi) - \frac{N}{2}\log\det \boldsymbol{\Sigma} - \frac{1}{2}\sum_{k=1}^{N}\mathbf{e}(k)^\top \boldsymbol{\Sigma}^{-1}\mathbf{e}(k) \,. \end{aligned}$$

La dérivée par rapport à $\boldsymbol{\Sigma}$ doit être nulle à l'optimum (valeur estimée des paramètres), c'est-à-dire

$$-\frac{N}{2}\hat{\boldsymbol{\Sigma}}^{-1} + \frac{1}{2}\hat{\boldsymbol{\Sigma}}^{-1}\left[\sum_{k=1}^{N}\hat{\mathbf{e}}(k)\hat{\mathbf{e}}^\top(k)\right]\hat{\boldsymbol{\Sigma}}^{-1} = \mathbf{0} \,,$$

soit

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{N}\sum_{k=1}^{N}\hat{\mathbf{e}}(k)\hat{\mathbf{e}}^\top(k) \,.$$

On reporte dans l'expression de $\log \pi(\mathbf{y}|\boldsymbol{\theta},\boldsymbol{\Sigma})$ qui devient

$$-N\log(2\pi) - \frac{N}{2}\log\det\left[\frac{1}{N}\sum_{k=1}^{N}\hat{\mathbf{e}}(k)\hat{\mathbf{e}}^\top(k)\right] - \frac{1}{2}\text{trace}(N\mathbf{I}_2)$$

avec $\mathbf{I}_2$ la matrice identité de dimension 2, et $\hat{\boldsymbol{\theta}}$ doit minimiser $\log\det[\sum_{k=1}^{N}\mathbf{e}(k)\mathbf{e}^\top(k)]$.

**7)**

On commence par calculer $\det \boldsymbol{\Sigma} = \sigma^4(1-\alpha^2)$ et

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

On reprend l'expression trouvée en 6,

$$\begin{aligned} \log \pi(\mathbf{y}|\boldsymbol{\theta},\sigma^2,\alpha) &= -N\log(2\pi) - \frac{N}{2}\log\sigma^4 - \frac{N}{2}\log(1-\alpha^2) - \frac{1}{2\sigma^2(1-\alpha^2)}\sum_{k=1}^{N}[e_1^2(k)+e_2^2(k)-2\alpha e_1(k)e_2(k)] \\ &= -N\log(2\pi) - N\log\sigma^2 - \frac{N}{2}\log(1-\alpha^2) - \frac{1}{2\sigma^2(1-\alpha^2)}(S_N - 2\alpha P_N) \,. \end{aligned}$$

La dérivée par rapport à $\sigma^2$ doit être nulle à l'optimum, soit

$$-\frac{N}{\hat{\sigma}^2} + \frac{1}{2\hat{\sigma}^4(1-\hat{\alpha}^2)}(\hat{S}_N - 2\hat{\alpha}\hat{P}_N) = 0$$

et donc

$$\hat{\sigma}^2 = \frac{1}{2N(1-\hat{\alpha}^2)}(\hat{S}_N - 2\hat{\alpha}\hat{P}_N) \,.$$

On reporte dans $\log \pi(\mathbf{y}|\boldsymbol{\theta},\sigma^2,\alpha)$, qui devient

$$\begin{aligned}
\log \pi(\mathbf{y}|\boldsymbol{\theta},\alpha) &= -N\log(2\pi) + N\log[2N(1-\alpha^2)] - N\log(S_N - 2\alpha P_N) - \frac{N}{2}\log(1-\alpha^2) - N \\
&= -N\log(2\pi) + N\log(2N) - N - N\log(S_N - 2\alpha P_N) + \frac{N}{2}\log(1-\alpha^2)\,.
\end{aligned}$$

Sa dérivée par rapport à $\alpha$ doit être nulle à l'optimum, soit

$$\frac{\hat\alpha}{1-\hat\alpha^2} = \frac{2\hat P_N}{\hat S_N - 2\hat\alpha \hat P_N}$$

d'où $\hat\alpha = 2\hat P_N/\hat S_N$.

Par définition, $S_N$ est positif et $S_N + 2P_N \geq 0$ et $S_N - 2P_N \geq 0$, donc $-1 \leq \hat\alpha \leq 1$, ce qui est était à prévoir puisque $\alpha$ est le coefficient de corrélation entre $\epsilon_1(k)$ et $\epsilon_2(k)$.

On reporte la valeur de $\hat\alpha$ dans celle de $\hat\sigma^2$, ce qui donne $\hat\sigma^2 = \hat S_N/(2N)$.

On reporte la valeur de $\hat\alpha$ dans $\log\pi(\mathbf{y}|\boldsymbol{\theta},\alpha)$, qui devient

$$\begin{aligned}
& -N\log(2\pi) + N\log(2N) - N - \frac{N}{2}\log(S_N^2 - 4P_N^2) \\
&= -N\log(2\pi) + N\log(2N) - N - \frac{N}{2}\log(S_N - 2P_N) - \frac{N}{2}\log(S_N + 2P_N) \\
&= -N\log(2\pi) + N\log(2N) - N - \frac{N}{2}\log\left[\sum_{k=1}^{N}(e_1(k)-e_2(k))^2\right] - \frac{N}{2}\log\left[\sum_{k=1}^{N}(e_1(k)+e_2(k))^2\right]\,.
\end{aligned}$$

Par conséquent, $\hat{\boldsymbol{\theta}}$ doit minimiser $\log[\sum_{k=1}^N (e_1(k)-e_2(k))^2] + \log[\sum_{k=1}^N (e_1(k)+e_2(k))^2]$.

**8)**

On a pour $\alpha = 0$

$$\frac{\partial \log\pi(\mathbf{y}|\boldsymbol{\theta},\sigma^2,\alpha)}{\partial\sigma^2} = -\frac{N}{\sigma^2} + \frac{1}{2\sigma^4}S_N$$

et donc

$$\mathtt{E}\left\{\left[\frac{\partial\log\pi(\mathbf{y}|\boldsymbol{\theta},\sigma^2,\alpha)}{\partial\sigma^2}\right]^2\middle|\,\boldsymbol{\theta},\sigma^2,0\right\} = \frac{N^2}{\sigma^4} - \frac{N}{\sigma^6}\mathtt{E}\left\{S_N\middle|\,\boldsymbol{\theta},\sigma^2\right\} + \frac{1}{4\sigma^8}\mathtt{E}\left\{S_N^2\middle|\,\boldsymbol{\theta},\sigma^2\right\}\,,$$

avec

$$\mathtt{E}\left\{S_N\middle|\,\boldsymbol{\theta},\sigma^2\right\} = 2N\sigma^2\,,\ \mathtt{E}\left\{S_N^2\middle|\,\boldsymbol{\theta},\sigma^2\right\} = 8N\sigma^4 + 4N(N-1)\sigma^4\,.$$

Par conséquent, $V = N/\sigma^4$, tandis qu'en 4, $V_i = N/(2\sigma_i^2)$. La variance est donc deux fois plus faible à présent, ce qui s'explique par le fait que deux fois plus de données (les $e_1(k)$ *et* les $e_2(k)$) sont utilisées pour estimer $\sigma^2$.

**9)**

Le critère d'Akaike fait intervenir $-\log\pi(\mathbf{y}|\hat{\boldsymbol{\theta}},\hat\sigma^2,\hat\alpha) + p$ avec $p$ le nombre de paramètres dans le modèle. Dans le premier cas, il y a 5+2=7 paramètres (5 pour $\boldsymbol{\theta}$, 2 pour $\sigma^2$ et $\alpha$), dans le second seulement 4+2=6. Le critère (à minimiser) conduit donc à préférer la structure la plus simple, c'est-à-dire la seconde.