

# Analyse factorielle dynamique : test du nombre de facteurs, estimation et application à l'enquête de conjoncture dans l'industrie

Catherine DOZ, Fabrice LENGART\*

**RÉSUMÉ.** – Nous proposons une procédure en deux étapes pour estimer un modèle à facteurs dynamique. Nous montrons que, dans un cadre dynamique stationnaire, l'analyse factorielle statique fournit des estimateurs convergents des paramètres et permet de construire un test asymptotique du nombre de facteurs à retenir. Une fois ce nombre fixé, le modèle peut être estimé par filtre de KALMAN. Nous appliquons ensuite cette procédure à l'enquête INSEE d'activité dans l'industrie, ce qui permet de construire un indicateur résumé de l'enquête.

---

## Dynamic Factor Analysis: Test of the Number of Factors and Estimation with an Application to the French Industrial Business Survey

**ABSTRACT.** – We suggest a two step procedure to estimate a dynamic factor model. We show that, in a stationary dynamic framework, static factor analysis leads to consistent estimators and allows to build an asymptotic test of the relevant number of factors. Once this number is set, the model can be estimated through a KALMAN filter. We then apply this procedure to the French industrial business survey, in order to build a composite index.

---

\* Catherine Doz, Université de Cergy-Pontoise, UFR d'Économie, THEMA (affiliation), Fabrice LENGART, CREST-INSEE. Nous tenons à remercier Jean-Pierre LAFFARGUE et Eric RENAULT pour leurs commentaires sur une version antérieure de cet article. Nous remercions également les deux rapporteurs de la Revue pour leurs remarques et suggestions. Nous restons seuls responsables des erreurs qui pourraient subsister.

# 1 Introduction

---

Les modèles factoriels dynamiques sont actuellement utilisés dans de nombreux domaines de l'économie. On peut notamment mentionner des exemples en économétrie financière (modèles de taux d'intérêt), mais aussi en macroéconomie (voir STOCK et WATSON [1989, 1993], QUAH et SARGENT [1993], FORNI et REICHLIN [1998] pour citer des travaux récents). Dans de tels modèles, les variables temporelles observées sont supposées dépendre linéairement d'un petit nombre de variables sous-jacentes inobservables, appelées facteurs. Deux méthodes sont principalement utilisées pour leur estimation. La première se situe dans le domaine des fréquences et revient à effectuer une décomposition particulière de la densité spectrale du processus vectoriel constitué par l'ensemble des variables étudiées. La seconde se situe dans le domaine des temps et suppose une modélisation de la dynamique des facteurs, puis une estimation par filtre de KALMAN.

Dans ce papier, on se propose de montrer qu'il est aussi possible d'employer en première approximation, dans un cadre temporel stationnaire, la technique d'analyse factorielle statique (seulement adaptée *a priori* à des données individuelles) et d'en dériver une procédure de test du nombre de facteurs à retenir dans le modèle dynamique. L'intérêt d'une telle procédure est que sa mise en œuvre est très rapide. Elle peut donc utilement précéder la phase de modélisation proprement dite, afin que le modèle estimé par filtre de KALMAN comporte un nombre de facteurs adapté aux données.

Le papier se divise en deux parties. La première est théorique : on y rappelle le principe de la méthode statique d'analyse factorielle, on montre qu'elle fournit des estimateurs convergents des paramètres d'intérêt même lorsque les données temporelles utilisées présentent de l'autocorrélation (mais restent stationnaires), et on expose une procédure de test du nombre de facteurs dans ce cadre. Cette procédure est fondée sur des résultats de GOURIÉROUX et MONFORT [1989] que nous généralisons légèrement pour les rendre applicables dans le cadre qui est le nôtre.

Dans la seconde partie, on applique cette procédure aux données de l'enquête mensuelle d'activité dans l'industrie publiée par l'INSEE. Le test sur le nombre de facteurs est d'abord mis en œuvre : l'hypothèse nulle d'un modèle à un seul facteur n'est pas rejetée. On estime alors par filtre de KALMAN un « indicateur résumé » de l'enquête. Ce faisant, on montre que le cadre des modèles à facteurs dynamiques fournit une sorte de grille d'interprétation de l'enquête mensuelle, passant par l'analyse des évolutions de l'indicateur résumé et par l'utilisation, le cas échéant, des informations supplémentaires apportées par les composantes spécifiques de chaque solde d'opinion.

## 2 Utilisation de l'analyse factorielle statique dans un cadre dynamique

---

### 2.1. Présentation

L'analyse factorielle « classique » a pour but d'offrir une description parcimonieuse d'un ensemble d'observations. Elle tente de représenter les variables étudiées dans un cadre linéaire, en fonction d'un ensemble de variables latentes, appelées *facteurs*, ou *facteurs communs*. C'est une technique qui est appropriée lorsqu'un petit nombre de facteurs peut rendre compte d'une part importante de l'information contenue dans l'ensemble des variables initiales.

Cette méthode a été initialement créée pour l'étude de données *individuelles*. Elle repose sur la modélisation suivante : on note  $n$  le nombre de variables étudiées,  $T$  le nombre d'observations dont on dispose pour chaque variable, et  $y_{it}$  la valeur de la  $t$ -ième observation de la variable  $y_i$  ; le modèle décrivant les variables  $y_1, \dots, y_n$  en fonction de  $p$  facteurs communs  $F_1, \dots, F_p$ ,  $p < n$ , s'écrit :

$$y_{it} = m_i + \lambda_{i1}F_{1t} + \dots + \lambda_{ip}F_{pt} + u_{it}, \text{ pour } i = 1 \text{ à } n \text{ et pour tout } t$$

On suppose que les  $(u_{it})_{t \in T}$  sont indépendants entre eux et indépendants des facteurs :

$$\forall i \neq j \quad \forall (t, \tau) \quad E(u_{it}u_{j\tau}) = 0 \text{ et } \forall (i, j) \quad \forall (t, \tau) \quad E(u_{it}F_{j\tau}) = 0$$

Comme le modèle est destiné à l'étude de données individuelles, on suppose en outre que :

$$\forall i \quad \forall t \neq \tau \quad E(u_{it}u_{i\tau}) = 0 \text{ et } E(F_{it}F_{i\tau}) = 0$$

Un tel modèle n'est donc applicable à l'étude de données *temporelles* ( $y_{it}$ ) que dans le cas très restrictif où *tous les processus* ( $F_{jt}$ ) et ( $u_{it}$ ) *sont supposés sans autocorrélation* (c'est-à-dire sont des bruits blancs au sens faible).

En utilisant des notations matricielles :

$$y_t = (y_{1t}, \dots, y_{nt})', F_t = (F_{1t}, \dots, F_{pt})', \\ u_t = (u_{1t}, \dots, u_{nt})', \Lambda = (\lambda_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$$

le modèle s'écrit de façon plus concise :

$$y_t = m + \Lambda F_t + u_t$$

avec :  $E F_t = 0, E u_t = 0, E(u_t u_t') = D = \text{diag}(d_1, \dots, d_n),$

$$\forall (t, \tau) \quad E(F_t u_\tau') = 0,$$

$$\forall (t, \tau), t \neq \tau, E(F_t F_\tau') = 0,$$

$$\forall (t, \tau), t \neq \tau, E(u_t u_\tau') = 0,$$

Un tel modèle n'a d'intérêt que lorsque les variables présentent une forte corrélation car il permet dans ce cas d'analyser la structure de cette corréla-

tion. Plus précisément, les facteurs rendent compte des corrélations entre variables, alors que chaque  $u_{it}$  représente une source de variation affectant la seule variable  $y_{it}$  : chaque  $u_{it}$  est appelé *composante spécifique* de la variable  $y_{it}$ . Les  $\lambda_{ij}$  sont appelés les *pondérations (loadings)* des facteurs : chaque  $\lambda_{ij}$  représente la contribution du facteur  $F_j$  au comportement de la variable  $y_i$ .

Il apparaît clairement que, dans une telle formulation, les facteurs ne sont définis qu'à une transformation linéaire près, sous réserve de modifier la matrice des pondérations. Dans le modèle le plus classique, les facteurs sont supposés non corrélés entre eux (cette hypothèse peut être ensuite levée) et de variances unitaires (ce qui ne restreint pas la généralité du modèle puisqu'ils n'interviennent qu'à un facteur multiplicatif près). Cependant, même dans ce cas de figure, s'il y a plus d'un facteur à estimer, les facteurs ne sont pas définis de façon unique : ils peuvent être modifiés par n'importe quelle rotation. Dans la pratique, lors de l'estimation, une solution unique est obtenue en ajoutant des contraintes identifiantes sur les facteurs. Cette solution peut ensuite subir une rotation (ou une simple transformation linéaire) si une telle opération facilite l'interprétation des résultats obtenus.

Concentrons-nous maintenant sur le modèle de base, c'est-à-dire sur le cas où les facteurs sont non corrélés entre eux et sont de variance unitaire. Ce modèle conduit à une interprétation simple en terme de variance et covariances des variables. De fait, on rajoute dans ce cas au modèle précédent l'hypothèse  $E(F_t F_t') = Id$ , de sorte qu'on a les relations suivantes :

$$V y_t = \Lambda \Lambda' + D,$$

soit : 
$$V y_{it} = \sum_{j=1}^p \lambda_{ij}^2 + d_i, \text{ pour } i = 1 \text{ à } n.$$

Chaque  $\lambda_{ij}^2$  représente la part de la variance de  $y_i$  qui est expliquée par le facteur  $F_j$ , et  $h_i^2 = \sum_{j=1}^p \lambda_{ij}^2$  représente la contribution totale des facteurs à la variance de  $y_i$  (également appelée « communalité »). La variance de  $u_i, d_i$  est la part de la variance de  $y_i$  qui ne peut pas être expliquée par les facteurs communs.

Il existe deux types de méthode pour estimer le modèle : l'analyse factorielle principale (*principal factor analysis*) et la méthode du maximum de vraisemblance sous hypothèse de normalité (voir LAWLEY et MAXWELL [1971] pour une présentation complète). La première ne nécessite aucune connaissance sur le nombre de facteurs à retenir, alors que ce nombre doit être donné *a priori* pour mettre en œuvre la seconde. En revanche, la méthode de maximum de vraisemblance fournit des estimateurs asymptotiquement efficaces des paramètres, ce qui n'est pas le cas de l'analyse factorielle principale.

En pratique, les deux méthodes sont utilisées. Lors d'une première étape, l'analyse factorielle principale fournit un critère permettant de choisir un nombre de facteurs qui semble pertinent. En fait, sa mise en œuvre revient à effectuer une analyse en composantes principales, mais cette ACP porte non pas sur la matrice de corrélation mais sur la matrice dite de corrélation réduite : les termes hors diagonale de cette matrice sont égaux aux corréla-

tions entre variables, mais les éléments diagonaux (qui correspondent aux  $h_i^2$ ) sont fixés, en première approximation, aux valeurs des corrélations canoniques de chaque variable avec l'ensemble des autres. On est alors conduit à calculer les valeurs propres de cette matrice, et le nombre de facteurs à retenir est choisi en fonction de la taille de ces valeurs propres. Dans une deuxième étape, il est possible de mettre en œuvre la méthode de maximum de vraisemblance. En outre, un test du rapport de vraisemblance permet, *dans ce cadre statique*, de contrôler que le nombre de facteurs retenu est correct.

Signalons enfin que, dans les deux méthodes, les paramètres  $\lambda_{ij}$  sont estimés en premier, et que les valeurs prises par les facteurs communs  $F_{jt}$  (les scores) sont approximés dans un second temps comme des combinaisons linéaires des variables initiales – ceci est fait par le biais de techniques de régression visant à minimiser la variance de l'écart entre chaque facteur et son approximation linéaire.

## 2.2. Convergence, dans un cadre dynamique, des estimateurs obtenus dans l'analyse factorielle statique

Dans toute la suite de cette partie, on suppose que *chacun des processus réels*  $(F_{it})$  et  $(u_{it})$  est stationnaire au sens faible et peut présenter de l'auto-corrélation, mais que le modèle est estimé par une procédure standard de maximum de vraisemblance, comme s'il n'y avait pas d'autocorrélation. La stationnarité des  $(F_{it})$  et  $(u_{it})$  entraîne en particulier celle des  $(y_{it})$ . Le vecteur des paramètres du modèle est  $\theta = \begin{pmatrix} \text{vec } \Lambda \\ d \end{pmatrix}$ , où  $d = (d_1, \dots, d_n)'$ .

L'estimateur  $\hat{\theta}_T$  ainsi obtenu est un M-estimateur de  $\theta$ . Nous montrons ici que cet estimateur est convergent<sup>1</sup>.

Pour tout  $t$  fixé, on note  $x_{it} = y_{it} - \bar{y}_i$  et  $x_t = (x_{1t}, \dots, x_{nt})'$ . Soit  $S = \frac{1}{T} \sum x_t x_t'$  la matrice de variance-covariance empirique des observations et soit  $\Sigma = \Lambda \Lambda' + D$  la matrice de variance-covariance théorique. La pseudo-logvraisemblance du modèle, calculée sous hypothèse de normalité, comme si les facteurs et les composantes spécifiques ne présentaient pas d'autocorrélation, peut s'écrire (à un terme constant et au facteur  $T - 1$  près, voir LAWLEY-MAXWELL [1971]) :

$$\begin{aligned} L_T(x, \Lambda \Lambda' + D) &= \frac{1}{T-1} \sum_{t=1}^T \ln l_t(x, \theta) \\ &= -\frac{1}{2} \ln(\det(\Lambda \Lambda' + D)) - \frac{1}{2} \text{tr}((\Lambda \Lambda' + D)^{-1} S) \end{aligned}$$

1. En toute rigueur, la démonstration qui est donnée montre la convergence de l'estimateur associé de la matrice de variance-covariance, c'est-à-dire de  $\hat{\Lambda}_T \hat{\Lambda}_T' + \hat{D}_T$ . Cependant, si le modèle est localement identifiable, la vraie valeur du paramètre  $\theta_0 = \begin{pmatrix} \text{vec } \Lambda_0 \\ d_0 \end{pmatrix}$  est unique à une rotation près de  $\Lambda_0$ , et on parlera encore de convergence de  $\hat{\theta}_T$  vers  $\theta_0$ , même s'il s'agit d'un abus de langage. Une condition suffisante d'identifiabilité locale a été énoncée par ANDERSON RUBIN (1956, théorème 5.9). Nous énonçons cette condition dans le lemme 5 (cf. ci-dessous) et nous la supposons réalisée dans toute la suite.

Soit  $\theta_0$  la vraie valeur du paramètre  $\theta$ . On suppose que  $\theta$  varie dans une région de la forme  $R^{np} \times [\alpha, +\infty[^n, \alpha > 0$ , contenant  $\theta_0$ . Ceci entraîne notamment que  $\theta_0$  vérifie lui-même :  $d_{i0} \neq 0$ , pour  $i = 1, \dots, n$ . Cette dernière condition est en particulier une condition suffisante pour que  $\Sigma$  soit inversible, donc pour que la pseudo-vraisemblance soit définie.

La démonstration de la convergence du M-estimateur  $\hat{\theta}_T$  maximisant  $L_T(x, \Lambda \Lambda' + D)$  se fait selon une démarche classique, qui est présentée par exemple dans NEWEY et MACFADDEN [1994]. Cette démonstration, que nous renvoyons en annexe, se fait en plusieurs étapes. On établit d'abord que, pour maximiser cette fonction, on peut se restreindre à supposer que  $\theta$  varie non pas dans  $R^{np} \times [\alpha, +\infty[^n$ , mais dans un compact  $\Theta$  inclus dans  $R^{np} \times [\alpha, +\infty[^n$ . On montre ensuite que la fonction objectif possède des propriétés suffisantes pour assurer la convergence. On obtient alors la proposition suivante :

PROPOSITION 1 : Le M-estimateur  $\hat{\theta}_T$  obtenu en maximisant la pseudo-vraisemblance calculée sous hypothèse de normalité et d'absence d'autocorrélation, est convergent.

*Preuve* : voir annexe 1.

### 2.3. Une procédure de test du nombre de facteurs dans un modèle dynamique

Dans cette partie, nous proposons une procédure pour tester l'hypothèse nulle que les données sont compatibles avec un modèle à  $p$  facteurs communs<sup>2</sup>. Cette procédure est fondée sur la loi asymptotique du vecteur du pseudo-score obtenu à partir de la maximisation de la pseudo-vraisemblance utilisée précédemment.

Soit  $\theta = \begin{pmatrix} \text{vec} \Lambda \\ d \end{pmatrix} \in R^{(p+1)n}$  le vecteur des paramètres du modèle contraint. Comme  $\Sigma$  est symétrique, le vecteur des paramètres du modèle non contraint peut s'écrire :  $\gamma = \text{vech} \Sigma \in R^{n(n+1)/2}$ <sup>3</sup>.

L'hypothèse nulle que nous voulons tester est la compatibilité de  $\Sigma$  avec un modèle à  $p$  facteurs, sous l'hypothèse alternative que  $\Sigma$  est non contrainte. L'hypothèse nulle peut se mettre sous la forme  $H_0 : \text{vech} \Sigma = \text{vech}(\Lambda \Lambda' + D)$ , c'est-à-dire sous la forme  $\gamma = h(\theta)$ , où  $h$  est l'application qui à  $\theta = \begin{pmatrix} \text{vec} \Lambda \\ d \end{pmatrix}$  associe  $\text{vech}(\Lambda \Lambda' + D)$ , lorsque  $D = \text{diag}(d)$ .

Ce test relève du cadre général présenté par GOURIÉROUX et MONFORT [1989a et 1989b]. La procédure de test qu'ils proposent est fondée sur le

2. Bien sûr, cette procédure peut être utilisée pour mener des tests de façon séquentielle avec différentes valeurs de  $p$ . Le modèle est d'autant plus contraint que la valeur de  $p$  retenue est petite.

3. On rappelle que pour toute matrice symétrique  $\Sigma$ ,  
 $\text{vech} \Sigma = (\sigma_{ij}, 1 \leq j \leq i \leq n) = (\sigma_{11}, \sigma_{21}, \dots, \sigma_{n1}, \sigma_{22}, \sigma_{32}, \dots, \sigma_{n2}, \dots, \sigma_{nn})'$

calcul de la loi asymptotique, sous l'hypothèse nulle, du vecteur du score associé à la pseudo-vraisemblance. Elle suppose que la matrice  $\frac{\partial h}{\partial \theta'}$  est de plein rang colonne. Cependant, dans le cas que nous étudions, cette dernière hypothèse n'est pas toujours vérifiée, comme nous le verrons plus loin. De sorte que nous serons amenés à étendre les résultats de GOURIÉROUX et MONFORT au cas où la matrice  $\frac{\partial h}{\partial \theta'}$  n'est pas de plein rang colonne (proposition 4).

Le calcul de la loi asymptotique du vecteur du pseudo-score sous l'hypothèse nulle reposant sur le calcul préalable de la loi asymptotique sous l'hypothèse alternative, nous commençons donc par calculer ce vecteur et cette loi asymptotique (proposition 3).

Dans la suite de ce paragraphe, nous utilisons des notations introduites par MAGNUS et NEUDECKER afin de simplifier les formules de calcul différentiel lorsqu'il est appliqué dans un cadre matriciel (voir par exemple MAGNUS et NEUDECKER [1988]). Ces notations sont les suivantes :

–  $D_n$  est la matrice de *duplication* d'ordre  $n$ , c'est-à-dire la matrice de format  $\left(n^2, \frac{n(n+1)}{2}\right)$  qui vérifie :

pour toute matrice  $M$  symétrique d'ordre  $n$ ,  $\text{vec}M = D_n \text{vech}M$ .

–  $D_n^+ = (D_n' D_n)^{-1} D_n'$  est la pseudo-inverse de  $D_n$ . En particulier, pour toute matrice  $M$  symétrique d'ordre  $n$ ,  $\text{vech}M = D_n^+ \text{vec}M$ .

–  $K_{nn}$  est la matrice de *commutation*, c'est-à-dire la matrice de format  $(n^2, n^2)$  qui vérifie :

pour toute matrice  $M$  carrée d'ordre  $n$ ,  $\text{vec}M' = K_{nn} \text{vec}M$ .

Nous utiliserons en outre diverses propriétés de ces matrices<sup>4</sup> (voir par exemple LÜTKEPOHL [1993] pour un rappel de l'ensemble de leurs propriétés).

Considérons maintenant la pseudo-vraisemblance du modèle non contraint :

$$L_T(x, \Sigma) = -\frac{1}{2} \ln \det \Sigma - \frac{1}{2} \text{tr} \left( \Sigma^{-1} \left( \frac{1}{T} \sum_t x_t x_t' \right) \right).$$

Suivant que l'on prend ou non en compte la symétrie de la matrice  $\Sigma$ , cette fonction peut être considérée comme une fonction de  $\text{vec} \Sigma$  (elle est alors définie sur  $R^{n^2}$ ), ou comme une fonction de  $\text{vech} \Sigma$  (elle est alors définie sur  $R^{\frac{n(n+1)}{2}}$ ).

4. –  $K_{nn}$  est une matrice orthogonale et symétrique

–  $K_{nn} D_n = D_n$

– pour  $M$  et  $N$  matrices carrées d'ordre  $n$  :  $K_{nn} M \otimes N = N \otimes M K_{nn}$ .

Nous utiliserons aussi la propriété suivante :  $\text{vec} MNP = P' \otimes M \text{vec} N$ . En particulier, si  $\Sigma$  est une matrice symétrique inversible :  $\text{vec} \Sigma^{-1} = \Sigma^{-1} \otimes \Sigma^{-1} \text{vec} \Sigma$ .

On a :

$$\begin{aligned}
\frac{\partial L_T}{\partial \text{vec} \Sigma}(x, \Sigma) &= -\frac{1}{2} \text{vec} \left( \Sigma^{-1} - \Sigma^{-1} \left( \frac{1}{T} \sum_t x_t x_t' \right) \Sigma^{-1} \right)' \\
&= -\frac{1}{2} K_{nn} \text{vec} \left( \Sigma^{-1} - \Sigma^{-1} \left( \frac{1}{T} \sum_t x_t x_t' \right) \Sigma^{-1} \right) \\
&= -\frac{1}{2} K_{nn} \Sigma^{-1} \otimes \Sigma^{-1} \text{vec} \left( \Sigma - \frac{1}{T} \sum_t x_t x_t' \right)
\end{aligned}$$

Comme  $\frac{\partial L_T}{\partial \text{vech} \Sigma}(x, \Sigma) = D_n' \frac{\partial L_T}{\partial \text{vec} \Sigma}(x, \Sigma)$  et comme  $D_n' K_{nn} = D_n'$ , on en déduit :

$$\frac{\partial L_T}{\partial \text{vech} \Sigma}(x, \Sigma) = -\frac{1}{2} D_n' \Sigma^{-1} \otimes \Sigma^{-1} \text{vec} \left( \Sigma - \frac{1}{T} \sum_t x_t x_t' \right)$$

L'estimateur du pseudo-maximum de vraisemblance du modèle non contraint est donc l'estimateur usuel  $\hat{\Sigma} = \frac{1}{T} \sum_t x_t x_t'$ , c'est-à-dire la matrice de variance-covariance empirique des données, que nous avons aussi notée  $S$ . Cet estimateur est convergent puisque, par hypothèse, le vecteur  $x$  des variables observées est stationnaire. De plus, on a vu dans la section précédente que l'estimateur obtenu en maximisant la pseudo-vraisemblance du modèle contraint est lui aussi convergent.

LEMME 2 : Soit  $(x_t)$  un processus gaussien stationnaire de taille  $n$ .

On suppose que :  $E x_t = 0, E x_t x_t' = \Sigma = \Gamma(0)$  et  $E x_t x_{t-h}' = \Gamma(h)$ .

Alors :

$\forall h \in Z$

$$\begin{aligned}
E((\text{vec} x_t x_t')(\text{vec} x_{t-h} x_{t-h}')') &= E(x_t x_{t-h}' \otimes x_t x_{t-h}') \\
&= (\text{vec} \Sigma)(\text{vec} \Sigma)' + (\Gamma(h) \otimes \Gamma(h))(I_{n^2} + K_{nn})
\end{aligned}$$

$$\text{cov}(\text{vec}(x_t x_t'), \text{vec}(x_{t-h} x_{t-h}')) = \Gamma(h) \otimes \Gamma(h)(I_{n^2} + K_{nn}).$$

*Preuve* : voir annexe 2.

PROPOSITION 3 : On suppose que  $(x_t)$  est un processus gaussien stationnaire, et l'on note :

$$\forall h \in Z \quad \Gamma(h) = E(x_t x_{t-h}'), \Gamma(0) = \Sigma, \Omega_0 = \sum_{h \in Z} \Gamma(h) \otimes \Gamma(h).$$

$$\text{On pose de plus : } J_0 = E \left( - \frac{\partial^2 L_T}{\partial \text{vech}(\Sigma) \partial \text{vech}'(\Sigma)}(x, \Sigma) \right).$$

Alors :

$$\text{i) } J_0 = \frac{1}{2} D_n' \Sigma^{-1} \otimes \Sigma^{-1} D_n$$



$$\begin{aligned}
\text{ii) } & \frac{\partial L_T}{\partial \text{vech}\Sigma}(x, \Sigma) = -J_0 \text{vech}\left(\Sigma - \frac{1}{T} \sum_t x_t x_t'\right) \\
\text{iii) } & \sqrt{T} \text{vech}\left(\Sigma - \frac{1}{T} \sum_t x_t x_t'\right) \xrightarrow{L} N\left(0, 2D_n^+ \Omega_0 D_n^{+'}\right) \\
\text{iv) } & \sqrt{T} \frac{\partial L_T}{\partial \text{vech}(\Sigma)}(x, \Sigma) \xrightarrow{L} N\left(0, 2J_0 D_n^+ \Omega_0 D_n^{+'} J_0\right)
\end{aligned}$$

*Preuve* : voir annexe 3.

PROPOSITION 4 : On s'intéresse au test d'une hypothèse  $H_0 : \gamma = h(\theta)$ , avec :  $\gamma \in R^m$ ,  $\theta \in R^q$ .

On suppose que :

–  $\gamma_0$  est la vraie valeur du paramètre : sous  $H_0$ ,  $\gamma_0$  vérifie :  $\gamma_0 = h(\theta_0)$ .

–  $\frac{\partial h}{\partial \theta'}$  est une matrice  $(m, q)$  de rang  $q_0$ , avec :  $q_0 \leq q \leq m$ .

– l'estimateur  $\hat{\gamma}_T$  de  $\gamma$  obtenu sous l'hypothèse alternative, est solution du problème :

$$\max_{\gamma \in R^m} Q_T(x, \gamma).$$

– l'estimateur  $\hat{\theta}_T$  de  $\theta$ , obtenu sous l'hypothèse nulle, est solution du problème :

$$\max_{\theta \in R^q} Q_T^*(x, \theta) \text{ avec } Q_T^*(x, \theta) = Q_T(x, h(\theta)).$$

–  $\hat{\gamma}_T$  et  $\hat{\theta}_T$  sont des estimateurs convergents de  $\gamma$  et  $\theta$

–  $\sqrt{T} \frac{\partial Q_T}{\partial \gamma}(x, \gamma_0) \xrightarrow{L} N(0, \Omega)$ , avec  $\Omega$  définie positive.

On note :  $\hat{\gamma}_T^0 = h(\hat{\theta}_T)$  et  $J_0 = p \lim \left( - \frac{\partial^2 Q_T}{\partial \gamma \partial \gamma'}(x, h(\theta_0)) \right)$

On note  $P$  la matrice de la projection  $J_0$ -orthogonale sur l'espace engendré par les colonnes de  $\frac{\partial h}{\partial \theta'}(\theta_0)$ . En désignant par  $(\cdot)^+$  la pseudo-inverse d'une matrice,  $P$  peut s'écrire :

$$P = \frac{\partial h}{\partial \theta'}(\theta_0) \left( \frac{\partial h'}{\partial \theta}(\theta_0) J_0 \frac{\partial h}{\partial \theta'}(\theta_0) \right)^+ \frac{\partial h'}{\partial \theta}(\theta_0) J_0.$$

Soit  $\hat{M} = I_m - P$  et  $(\hat{M}' \hat{\Omega} \hat{M})^-$  une inverse généralisée de  $\hat{M}' \hat{\Omega} \hat{M}$ . On note  $\hat{M}$  un estimateur convergent de  $M$  et  $\hat{\Omega}$  un estimateur convergent de  $\Omega$ .

Alors la statistique du test du score vérifie :

$$T \left( \frac{\partial Q_T}{\partial \gamma}(x, \hat{\gamma}_T^0) \right)' (\hat{M}' \hat{\Omega} \hat{M})^- \left( \frac{\partial Q_T}{\partial \gamma}(x, \hat{\gamma}_T^0) \right) \xrightarrow{L} \chi^2(m - q_0)$$

*Preuve* : voir annexe 4.

LEMME 5 : On se place sous la condition suffisante d'identifiabilité locale de ANDERSON-RUBIN<sup>5</sup>.

Soit  $\theta = \begin{pmatrix} \text{vec} \Lambda \\ d \end{pmatrix}$  et  $h(\theta) = \text{vech}(\Lambda \Lambda' + D)$  où  $D = \text{diag}(d)$ .

On pose  $\Delta_n = \sum_{i=1}^n e_i \otimes e_i e_i'$  de sorte que  $\Delta_n d = \text{vec} D$ .

Alors :

$$\text{i) } \frac{\partial h}{\partial \theta'}(\theta) = D_n^+ [2\Lambda \otimes I_n \quad \Delta_n]$$

$$\text{ii) } \text{rg} \frac{\partial h}{\partial \theta'}(\theta) = n(p+1) - \frac{p(p-1)}{2}$$

*Preuve* : voir annexe 5.

Remarquons en particulier que le seul cas où  $\frac{\partial h}{\partial \theta'}$  est de plein rang est le cas où  $p = 1$ .

PROPOSITION 6 : On se place sous les hypothèses de la proposition 3 et du lemme 5.

On note  $H_0 : \text{vech} \Sigma_0 = h(\theta_0)$  l'hypothèse nulle du test (il existe une modèle à  $p$  facteurs),  $P = \frac{\partial h}{\partial \theta'}(\theta_0) \left( \frac{\partial h'}{\partial \theta}(\theta_0) J_0 \frac{\partial h}{\partial \theta'}(\theta_0) \right)^+ \frac{\partial h'}{\partial \theta}(\theta_0) J_0$ ,

$M = I_{n(n+1)/2} - P$ , et  $\hat{M}$  un estimateur convergent de  $M$ .

Pour tout  $h \in Z$ , on note  $\hat{\Gamma}(h)$  la matrice des autocorrélations empiriques d'ordre  $h$  des observations et on définit :  $\hat{\Omega}_0 = \sum_h \hat{\Gamma}(h) \otimes \hat{\Gamma}(h)$ .

Enfin, on note  $S = \frac{1}{T} \sum_t x_t x_t'$  la matrice de variance-covariance empirique des observations et  $\hat{\Sigma}_T^0 = \hat{\Lambda}_T \hat{\Lambda}_T' + \hat{D}_T$  l'estimateur contraint obtenu en maximisant la pseudo-vraisemblance sous  $H_0$ .

Alors :

i) la statistique du test du pseudo-score s'écrit :

$$\xi_S = \frac{T}{2} \text{vech}(\hat{\Sigma}_T^0 - S)' \left( \hat{M} D_n^+ \hat{\Omega}_0 D_n^+ \hat{M}' \right)^- \text{vech}(\hat{\Sigma}_T^0 - S)$$

Elle vérifie :  $\xi_S \xrightarrow[H_0]{L} \chi^2(r)$

avec  $r = \frac{1}{2}((n-p)^2 - (n+p))$ .

ii) Elle est identique à la statistique du test de HAUSMAN-WALD de  $H_0$ .

5. Cette condition est la suivante : on définit  $\Phi = D - \Lambda(\Lambda' D^{-1} \Lambda)^{-1} \Lambda' = (\varphi_{ij})$ , et on suppose que la matrice de terme général  $\varphi_{ij}^2$  est inversible.

*Preuve* : voir annexe 6.

On peut remarquer que le nombre de degrés de liberté de la loi de chi-deux obtenue comme loi limite est bien égal au nombre de contraintes testées. En effet, le modèle contraint contient  $np + n$  paramètres, mais la matrice  $\Lambda$  n'est définie qu'à une rotation près. De fait, (cf. LAWLEY-MAXWELL [1971]) la pseudo-vraisemblance est maximisée sous une contrainte identifiante, à savoir la contrainte que la matrice  $\Lambda' D^{-1} \Lambda$  soit diagonale. Ceci impose  $\frac{p(p-1)}{2}$  conditions d'orthogonalité sur les colonnes de  $\Lambda$ . Le nombre de paramètres libres du modèle contraint est donc  $(p+1)n - \frac{p(p-1)}{2}$ , si bien que le nombre de contraintes testées est :

$$\frac{n(n+1)}{2} - \left( (p+1)n - \frac{p(p-1)}{2} \right) = \frac{1}{2}((n-p)^2 - (n+p)) = r.$$

Pour conclure cette partie, il importe de noter que la procédure de maximum de vraisemblance ainsi utilisée sans tenir compte de l'autocorrélation des variables fournit des estimateurs asymptotiquement normaux mais non asymptotiquement efficaces. On pourrait améliorer cet estimateur en changeant la fonction objectif<sup>6</sup>, mais ceci ne permettrait plus d'utiliser la procédure standard de l'analyse factorielle. La procédure de test que nous venons de présenter a l'avantage d'être très simple à utiliser. Mais elle ne peut être conçue que comme une étape préliminaire visant à déterminer le nombre de facteurs pertinent. Dans une deuxième étape, nous serons donc amenés à tenir compte explicitement de la dynamique des variables traitées, de façon à obtenir des estimateurs asymptotiquement efficaces. Nous procéderons alors à une estimation par filtre de KALMAN, fondée sur une écriture état-mesure du modèle. Comme nous l'avons signalé en introduction, cette méthode est en effet employée pour l'estimation des modèles à facteurs dynamiques, lorsque le nombre de facteurs est fixé (voir par exemple STOCK et WATSON [1989, 1993]).

### 3 Application aux données de l'enquête de conjoncture dans l'industrie

---

Les enquêtes de conjoncture constituent une source précieuse pour suivre les fluctuations économiques de court terme, dans la mesure où elles cumulent trois types d'avantages : elles fournissent un message recueilli directement auprès des acteurs économiques ; elles sont disponibles rapidement (environ un mois après l'envoi des questionnaires), à intervalles rapprochés et réguliers ; enfin, elles sont très peu révisées (chaque publication

---

6. CRAGG et DONALD [1995] ont par exemple traité de cette façon le cas de l'hétéroscédasticité des composantes spécifiques.

donne lieu à une révision des soldes d'opinion du seul mois précédent, révision qui, de surcroît, est en général de très faible amplitude).

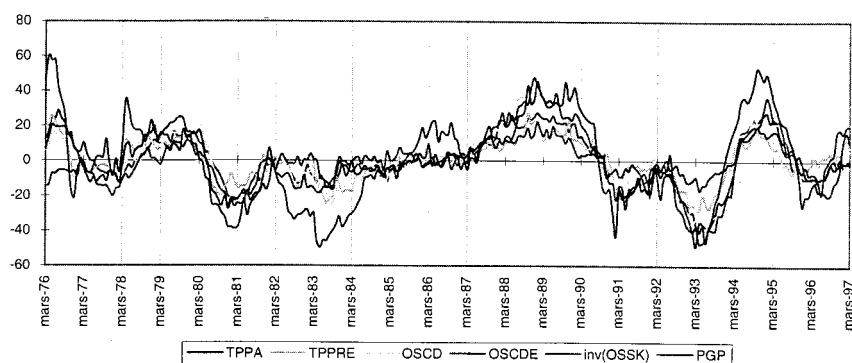
### 3.1. Les données

On considère ici les réponses à six questions de l'enquête de conjoncture de l'INSEE dans l'industrie, sur la période mars 1976-mars 1997. Ces questions sont : opinion sur la tendance passée de la production personnelle (TPPA), opinion sur la tendance prévue de la production personnelle (TPPRE), opinion sur la demande et les carnets de commande globaux (OSCD), opinion sur la demande et les carnets de commande en provenance de l'étranger (OSCDE), opinion sur le niveau des stocks (OSSK) et perspectives générales d'activité (PGP).

Dans cette enquête, pour chaque question, les industriels peuvent répondre suivant trois modalités : le niveau est plus élevé, moins élevé, ou comparable à celui de la période précédente. Les réponses sont agrégées en pondérant à la fois par la taille de l'entreprise répondante et par celle de la branche à laquelle l'entreprise appartient. Le résultat publié est le solde d'opinion, c'est-à-dire la différence entre le pourcentage des réponses « plus élevé » et celui des réponses « moins élevé ».

Les soldes d'opinion sont ici corrigés des variations saisonnières<sup>7</sup> (à l'aide du logiciel X11 ARIMA). Les cinq premiers couvrent l'industrie manufacturière, et le sixième (PGP) couvre l'ensemble de l'industrie (y compris énergie) hormis les industries agro-alimentaires. Un simple coup d'œil au graphique des six soldes (graphique 1) suffit pour se convaincre que les profils de ces séries sont extrêmement similaires. D'où l'idée de chercher à résumer en un seul indice synthétique l'information commune qu'elles contiennent.

GRAPHIQUE 1  
*Soldes d'opinions de l'enquête mensuelle.*



7. Cependant, deux soldes d'opinion (OSCD et OSSK) présentent une saisonnalité trop instable pour être corrigée. Ces soldes, publiés comme CVS par l'INSEE, sont donc en fait égaux à leurs valeurs brutes.

Chaque solde d'opinion présente une dynamique « cyclique » très prononcée, en ce sens qu'il peut être considéré comme stationnaire, mais qu'il présente une structure d'autocorrélation assurant une persistance à ses variations. En effet, les autocorrélogrammes empiriques montrent que les autocorrélations sont significativement positives jusqu'à l'ordre 8.

Néanmoins, ces séries n'en sont pas pour autant intégrées. En effet, le test de DICKEY-FULLER augmenté refuse l'hypothèse nulle de la présence d'une racine unité et, à l'inverse, le test KPSS (KWIATKOWSKI *et alii* [1992]) conduit à accepter à chaque fois l'hypothèse nulle que les soldes sont des variables stationnaires (pour le détail des résultats des tests, voir annexe 7).

### 3.2. Application du test et estimation d'un modèle à un facteur par filtre de KALMAN

Le tableau ci-dessous présente les résultats du test du pseudo-score et les valeurs des seuils correspondants (dans le cas d'un modèle à 6 variables et un facteur, la loi asymptotique sous hypothèse nulle est un  $\chi^2(9)$ ). Le nombre de retards que l'on doit retenir pour le calcul des autocovariances est nécessairement supérieur à 8, puisque les autocorrélations univariées des séries sont significativement différentes de zéro jusqu'à cet ordre. En pratique, la valeur du seuil obtenu est de toute façon quasiment stable au-delà de 8 retards.

nombre de retards	test	Proba
0	83,80	0,00
8	14,65	0,10
12	14,41	0,11

Ainsi, alors qu'un test mené sans tenir compte de l'autocorrélation des séries (nombre de retard fixé à zéro), conduit à rejeter l'hypothèse d'un unique facteur<sup>8</sup>, cette hypothèse est au contraire acceptée pour un seuil de 5 % lorsque l'on modifie le test pour tenir compte du cadre dynamique du modèle. En définitive, il apparaît licite de chercher à construire un indicateur « résumant » l'information commune apportée par les six principaux soldes de l'enquête d'activité.

On peut noter que ce facteur commun suffit à rendre compte de 82 % de la variance totale des séries étudiées. Plus précisément, les pondérations et les « communalités » sont :

Question	Pondération	Communalité
TPPA	0,96	92,6 %
TPPRE	0,89	79,8 %
OSCD	0,97	93,7 %
OSCDE	0,90	81,9 %
OSSK	-0,86	74,5 %
PGP	0,85	73,0 %

8. C'est aussi le résultat que fournit le test du rapport de vraisemblance calculé par le programme standard de SAS dans le cas sans autocorrélation.

Une fois que le choix du nombre de facteurs communs à retenir est arrêté (en l'espèce, un seul), il demeure nécessaire de tenir compte de la dynamique des variables pour estimer le modèle. Dans le cas présent, nous avons essayé par tâtonnement plusieurs spécifications possibles de type ARMA pour le facteur commun et les composantes spécifiques. En nous attachant d'une part à la significativité des paramètres obtenus *in fine* et d'autre part à la « qualité » des différentes séries d'innovations (ces séries, une fois estimées, pouvaient-elles être considérées comme des bruits blancs faibles ?), nous avons été amenés à retenir une dynamique de type ARMA(2,1) pour le facteur commun, et des dynamiques de type AR(1) pour les composantes spécifiques.

Le modèle est alors le suivant :

$$\begin{aligned} y_{it} &= \lambda_i F_t + u_{it}, \text{ pour } i = 1 \text{ à } n \text{ (avec ici } n = 6) \text{ et pour tout } t, \\ F_t &= \varphi_1 F_{t-1} + \varphi_2 F_{t-2} + \varepsilon_t - \theta \varepsilon_{t-1} \text{ pour tout } t, \\ u_{it} &= \rho_i u_{i,t-1} + \varepsilon_{it}, \text{ pour } i = 1 \text{ à } n \text{ et pour tout } t, \end{aligned}$$

où  $\varepsilon_t$  et  $\varepsilon_{it}$  sont les innovations de  $F_t$  et de  $u_{it}$  à chaque date  $t$ , et où les processus  $(\varepsilon_t)$  et  $(\varepsilon_{it})$  sont supposés indépendants.

Ce type de modèle admet une représentation état-mesure. Rappelons que la forme générale des modèles état-mesure est la suivante :

$$\begin{aligned} y_t &= Z_t \alpha_t + d_t + e_t \quad (M) \\ \alpha_t &= A_t \alpha_{t-1} + c_t + R_t \eta_t \quad (T) \end{aligned}$$

dans laquelle les processus  $(e_t)$  et  $(\eta_t)$  sont non autocorrélés, non corrélés entre eux, et vérifient :

$$E e_t = 0, V e_t = H_t, E \eta_t = 0, V \eta_t = Q_t.$$

Dans un tel cadre, le vecteur  $\alpha_t$ , qui est appelé le *vecteur d'état*, est inobservable. L'équation (M) est appelée *équation de mesure* : elle donne la relation, à chaque date  $t$ , entre le vecteur  $y_t$  des variables observées et le vecteur d'état. L'équation (T), ou *équation de transition*, décrit la dynamique du vecteur d'état.

Les matrices  $Z_t, A_t, d_t, c_t, R_t, H_t$  et  $Q_t$  sont non stochastiques, mais elles peuvent dépendre du temps de façon déterministe. Elles peuvent aussi dépendre des paramètres du modèle.

Dans le cas qui nous concerne, on peut réécrire le modèle sous une forme état-mesure en posant le vecteur d'état  $\alpha_t = (F_t, F_{t-1}, \varepsilon_t, u_{1t}, \dots, u_{nt})'$  :

$$y_t = \begin{pmatrix} \lambda_1 & 0 & 0 & & \\ \vdots & \vdots & \vdots & I_n & \\ \lambda_n & 0 & 0 & & \end{pmatrix} \begin{pmatrix} F_t \\ F_{t-1} \\ \varepsilon_t \\ u_{1t} \\ \vdots \\ u_{nt} \end{pmatrix}$$

$$\begin{pmatrix} F_t \\ F_{t-1} \\ \varepsilon_t \\ u_{1t} \\ \vdots \\ u_{nt} \end{pmatrix} = \left( \begin{array}{ccc|ccc} \varphi_1 & \varphi_2 & -\theta & & & \\ 1 & 0 & 0 & & & 0 \\ 0 & 0 & 0 & & & \\ \hline & & & \rho_1 & & \\ & & & & \ddots & \\ & & 0 & & & \rho_n \end{array} \right) \begin{pmatrix} F_{t-1} \\ F_{t-2} \\ \varepsilon_{t-1} \\ u_{1t-1} \\ \vdots \\ u_{nt-1} \end{pmatrix} + \left( \begin{array}{c|c} 1 & \\ 0 & 0 \\ \hline 1 & \\ 0 & \\ \vdots & I_n \\ 0 & \end{array} \right) \begin{pmatrix} \varepsilon_t \\ \varepsilon_{1t} \\ \vdots \\ \varepsilon_{nt} \end{pmatrix}$$

Un tel modèle peut être estimé de diverses façons. Dans ce papier, nous utilisons une procédure du maximum de vraisemblance par filtre de KALMAN. Rappelons brièvement les aspects principaux de cette méthode - une présentation complète figure, par exemple, dans HARVEY [1989] ou GOURIÉROUX et MONFORT [1989b]. Pour tout ensemble fixé de valeurs des paramètres, le filtre de KALMAN en tant que tel consiste à déterminer la meilleure prévision du vecteur d'état à chaque date, compte tenu de l'information disponible à cette date. Pour tout  $t$ , les prévisions optimales :  $a_{t|t-1} = E(\alpha_t | I_{t-1})$  et  $a_t = E(\alpha_t | I_t)$ , sont ainsi calculées par une procédure itérative.

Sous hypothèse de normalité des perturbations, il est alors possible de calculer explicitement la vraisemblance du modèle en fonction de ces prévisions optimales du vecteur d'état et des matrices de variance-covariance des erreurs de prévision associées. La vraisemblance peut ainsi être obtenue en tout point de l'espace des paramètres, et maximisée à l'aide d'une procédure quelconque d'optimisation, de façon à obtenir les estimateurs du maximum de vraisemblance des paramètres. Sous des hypothèses standard, ces estimateurs sont asymptotiquement normaux.

Lorsqu'on estime le modèle à facteur dynamique présenté précédemment, les résultats sont les suivants.

Les estimations obtenues pour les paramètres du facteur commun conduisent à l'équation :

$$F_t = \underset{(0,05)}{1,83} F_{t-1} - \underset{(0,05)}{0,85} F_{t-2} + \varepsilon_t - \underset{(0,11)}{0,48} \varepsilon_{t-1}$$

Il faut ici noter que la valeur de  $\sigma_\varepsilon$  a été fixée arbitrairement à 0,5. En effet, le modèle est défini à un paramètre d'échelle près : il faut donc fixer la valeur de l'un des paramètres pour le rendre identifiable.

On peut remarquer que l'estimation obtenue pour le polynôme autorégressif du facteur commun a des racines proches de l'unité. Cependant, il semble difficile de tester directement la présence d'une racine unité dans un tel contexte. En effet, sous l'hypothèse nulle, la loi est non standard et les tests de DICKEY-FULLER ne peuvent être directement appliqués dans le cadre non-linéaire des modèles à composantes inobservables.

Cependant, les tests de racines unité menés sur les séries initiales ont conduit à rejeter l'hypothèse de non-stationnarité, et le facteur commun est ici nécessairement de même nature puisque les composantes spécifiques sont clairement stationnaires. Ce type de résultat n'est d'ailleurs pas spécifique à notre étude ; au contraire, il semble que l'on obtienne souvent des résultats analogues lorsqu'on étudie des séries présentant un caractère cyclique, ceci étant dû au caractère fortement persistant de telles séries<sup>9</sup>.

Les estimations des autres paramètres figurent dans le tableau ci-dessous

Question	$\lambda_i$		$\rho_i$		$\sigma_{\varepsilon_i}$	
	estimation	écart-type	estimation	écart-type	estimation	écart-type
TPPA	2,76	0,24	0,33	0,04	2,55	0,15
TPPRE	1,96	0,21	0,81	0,03	3,23	0,15
OSCD	3,17	0,30	0,88	0,03	2,55	0,19
OSCDE	3,25	0,36	0,87	0,03	3,94	0,20
OSSK	-1,61	0,18	0,82	0,04	2,83	0,14
PGP	3,69	0,55	0,89	0,03	5,95	0,25

Si l'on tient compte du fait que, dans le cas présent, les soldes et le facteur commun obtenu ne sont pas normalisés, les pondérations obtenues sont compatibles avec celles que fournissait l'analyse factorielle statique<sup>10</sup>.

Comme nous travaillons dans un cadre stationnaire, les statistiques usuelles peuvent être utilisées. Cependant, le problème de la significativité des paramètres d'écart-type est un peu plus complexe puisque la valeur zéro est sur la frontière de l'espace admissible des paramètres. Dans ce cas, le carré de la statistique de STUDENT suit un mélange de lois :  $\frac{1}{2}\chi_0^2 + \frac{1}{2}\chi_1^2$ <sup>11</sup>. L'étude des résultats conduit, ici encore, à conclure que ces paramètres sont tous très significatifs. En particulier, les composantes spécifiques de chaque solde ne peuvent être assimilées à de simples bruits blancs (le coefficient autorégressif d'ordre 1 est significativement différent de zéro), ce qui laisse penser que la réponse à chaque question de l'enquête contient bien une information supplémentaire par rapport aux autres et par rapport aux enseignements qu'on peut tirer du facteur commun.

Il convient également, pour juger de la qualité de l'estimation, de savoir si les différentes innovations obtenues peuvent être assimilées à des bruits blancs. On a donc procédé à une analyse des autocorrélogrammes direct, inverse et partiel de ces innovations et on a calculé des tests du portmanteau

9. Par exemple, WATSON [1986] estime la composante cyclique du PIB américain sous une forme AR(2) dans le cadre d'un modèle à composantes inobservables et il obtient :  $C_t = 1,50C_{t-1} - 0,58C_{t-2} + u_t$ . Ici aussi,  $\varphi_1 + \varphi_2 \neq 1$ , ce qui signifie que les racines du polynôme retard sont en module proches de l'unité. WATSON relève ce fait, mais ne teste par l'hypothèse de non stationnarité.

10. Plus précisément, si l'on multiplie les pondérations normalisées – obtenues par analyse factorielle statique – par les écart-types des soldes correspondants et si l'on les divise par l'écart-type du facteur commun estimé par filtre de KALMAN (ce afin de rendre les résultats comparables entre les deux méthodes), on obtient :

TPPA : 2,74 ; TPPRE : 1,96 ; OSCD : 3,33 ; OSCDE : 3,20 ; OSSK : -1,66 ; PGP : 4,08.

11. Voir par exemple HARVEY (1989) : un test de niveau 5 % est obtenu en comparant le STUDENT à 1,6.

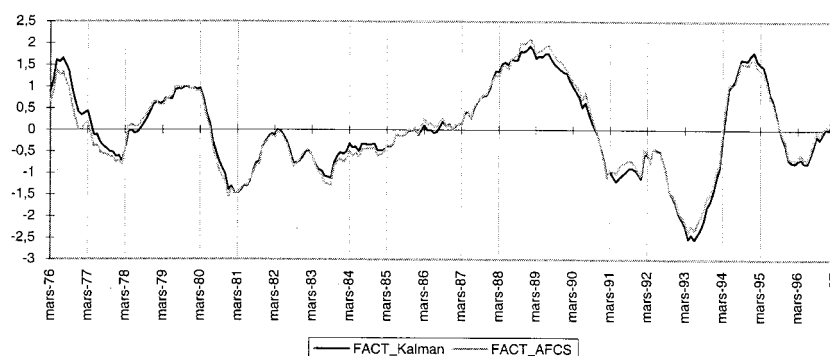


(bien qu'ils ne soient pas directement adaptés au cadre des modèles à composantes inobservables). Ces derniers conduisent à accepter au seuil de 5 % l'hypothèse que les innovations du facteur commun et des composantes spécifiques sont des bruits blancs, pour toutes les séries sauf pour celle qui correspond au solde portant sur la demande et les carnets globaux (OSCD). L'examen de l'autocorrélogramme de l'innovation correspondante permet en effet de constater l'existence d'une autocorrélation significativement différente de zéro à l'ordre 12 (de l'ordre de 0,20). Ce phénomène est sans doute attribuable au fait que le solde correspondant n'est pas corrigé des variations saisonnières lorsqu'il est publié par l'INSEE (pour des raisons de mauvaise stabilité des coefficients saisonniers éventuels). En définitive, on est conduit à juger que l'estimation obtenue est globalement satisfaisante.

On peut enfin noter que le facteur commun tiré du filtre de KALMAN est extrêmement proche de celui obtenu par analyse factorielle statique (cf. graphique 2). Ce résultat est surprenant *a priori* puisque cette dernière technique ne tient aucun compte de la dynamique des séries utilisées. Il semble suggérer que l'analyse factorielle classique fournit des résultats qui demeurent, en pratique, utilisables dans le domaine temporel lorsque les données observées présentent une très forte corrélation instantanée.

GRAPHIQUE 2

*Comparaison des facteurs communs obtenus par les deux méthodes.*

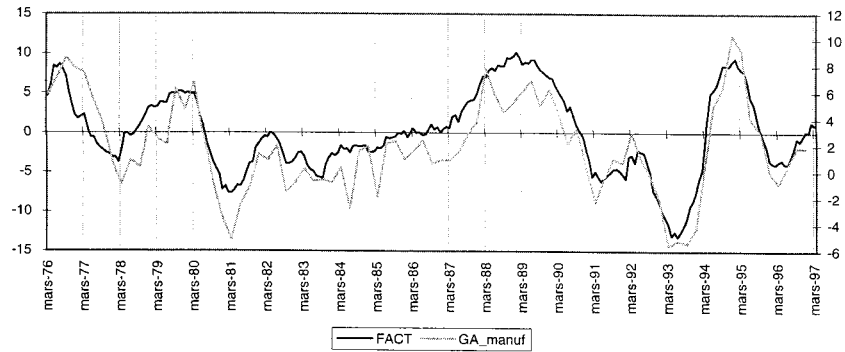


### 3.3. Une grille d'interprétation pour l'enquête mensuelle dans l'industrie

Lorsque l'on rapproche l'évolution du facteur commun obtenu du glissement annuel de la production manufacturière (tirée des comptes trimestriels), on constate que les retournements des deux séries se produisent le plus souvent à des dates très voisines (cf. graphique 3). Il ne faudrait pas, cependant, se méprendre sur le sens de cette comparaison : on se borne à constater visuellement que l'appréciation portée par les industriels sur leur activité (au travers de leurs réponses à l'enquête mensuelle) offre un reflet fidèle des grands mouvements de la conjoncture industrielle. En outre, les variations au mois le mois du facteur commun apparaissent moins heurtées que celles des

GRAPHIQUE 3

**Indicateur composite et glissement annuel de la production manufacturière.**



soldes d'opinion relatifs aux différentes questions de l'enquête, ce qui permet en principe de détecter plus facilement et plus rapidement les inflexions qu'elle connaît.

Sur la période d'étude dont nous disposons, on peut lire en particulier le retournement conjoncturel à la baisse du début de 1980 – consécutif à la forte hausse du prix du pétrole tout au long de l'année précédente (second choc pétrolier) – de même que la forte expansion de la fin des années 80 faisant suite au contre-choc pétrolier de 1986. L'indicateur rend compte du ralentissement enregistré au début des années 90, marqué par une stabilisation de la mi-91 à la mi-92, puis par la sévère récession de 1993. Enfin, la nette ascension du facteur commun en 1994 et sa rapide redescente en 1995 attestent de la brièveté du dernier cycle.

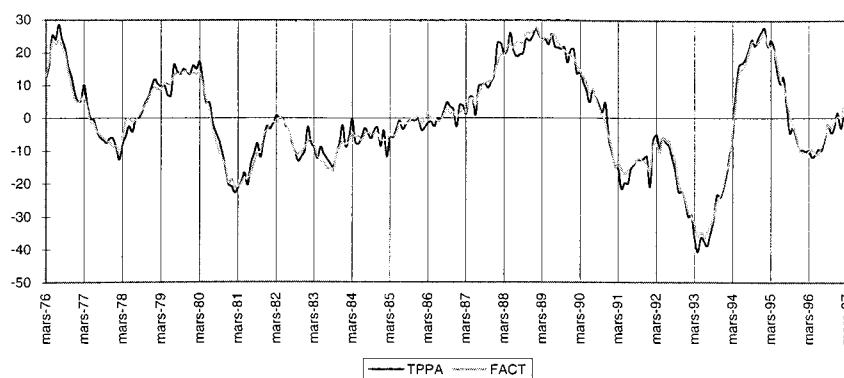
L'indicateur résumé ainsi calculé apparaît au total comme une sorte de climat général, rendant compte de l'information *commune* contenue dans les différents soldes. Cependant, l'analyse factorielle présente également l'intérêt *d'isoler la part d'information spécifiquement apportée par chaque solde d'opinion* (chacune des composantes spécifiques estimée ne pouvant être assimilée à un simple bruit blanc). Ce faisant, elle fournit une sorte de grille de lecture complète de l'enquête. On se contentera ici de montrer que, durant les périodes où la composante spécifique d'un solde d'opinion présente un mouvement prononcé et durable, il est souvent possible d'interpréter cette information au moins d'un point de vue qualitatif.

Il est très important ici de ne pas se laisser abuser par les expressions employées : le terme d'information spécifique désigne l'information complémentaire apportée par une question donnée ; en particulier, il ne signifie pas qu'il faille se fonder uniquement sur elle pour analyser les résultats fournis par cette question. Bien au contraire, on a vu que l'indicateur résumé, cette information commune à toutes les questions, rend compte d'une part très importante de l'évolution du profil de chacune d'entre elles. De sorte que, pour prendre un exemple, si, en période de climat favorable, une composante spécifique est négative de façon durable, on jugera, non pas que l'opinion des industriels sur le sujet abordé est défavorable, mais qu'elle est plutôt moins favorable qu'elle ne pourrait – ou ne devrait – l'être dans le contexte conjonc-

turel étudié. Cette précaution prise, on peut maintenant tenter, à la lumière des réponses aux questions de l'enquête mensuelle, de décrire plus précisément les fluctuations conjoncturelles que l'industrie française a traversées.

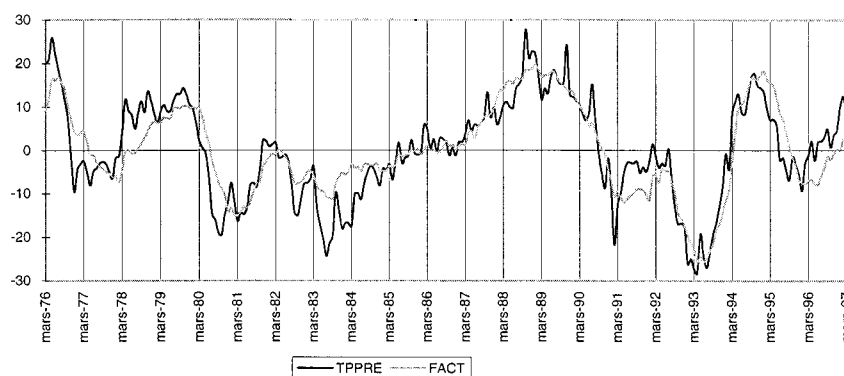
La réponse relative au **niveau de la production passée** (cf. graphique 4) est extrêmement proche du climat général, si bien qu'elle constitue à elle seule une bonne approximation de l'information commune sous-jacente aux six questions de l'enquête, à condition de la lisser quelque peu. Cet avantage a pour contrepartie évidente le fait que l'information spécifique apportée par l'opinion sur l'activité passée, animée de mouvements de faible amplitude, généralement irréguliers et contradictoires, ne semble pas interprétable.

GRAPHIQUE 4  
*Tendance de la production passée.*



En ce qui concerne les **perspectives personnelles de production** (cf. graphique 5), si l'on veut voir dans l'information spécifique fournie par cette question une variable anticipée de la conjoncture, le bilan que l'on peut dresser est pour le moins contrasté. En fait, il semble, en première approximation, que, à la fin des années 70, cette information spécifique tendait à positionner le solde d'opinion au dessus du climat général en période de conjoncture ascendante et en dessous dans le cas contraire, si bien que le

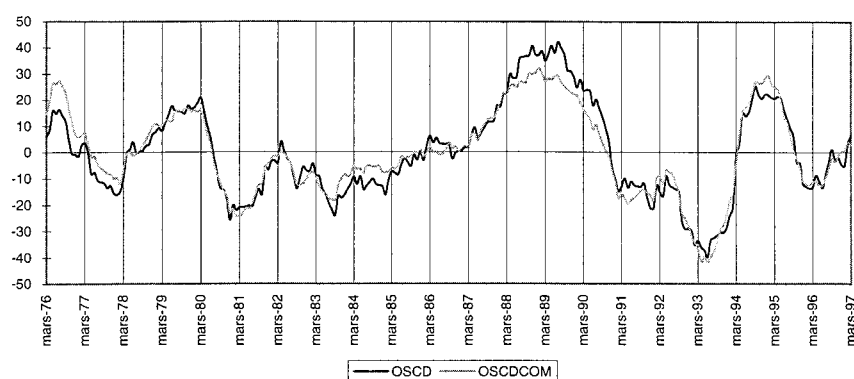
GRAPHIQUE 5  
*Tendance de la production prévue.*



croisement des deux courbes signalait qu'un véritable retournement était en cours ; en ce sens, on pouvait donc estimer que les grandes inflexions de l'activité pouvaient être décelées au travers des anticipations des entrepreneurs. La situation semble s'être obscurcie au début des années 80. Néanmoins, on peut noter que la composante spécifique semble avoir retrouvé un certain pouvoir prédictif lors des dernières années (ralentissement de 1995, accélération qui a suivi).

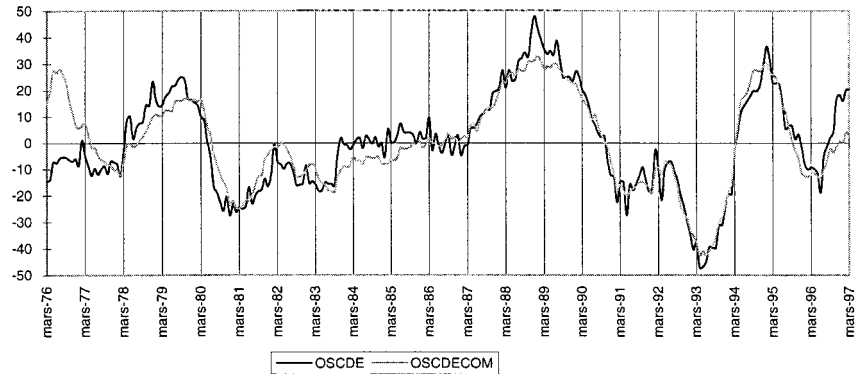
C'est généralement en période de haut ou de bas de cycle que l'information spécifique apportée par la réponse sur le **niveau de la demande et des carnets globaux** (cf. graphique 6) se distingue du climat général. Une interprétation possible est la suivante : lorsque l'information spécifique est positive, cela signifie que l'état de la demande adressée aux industriels interrogés est jugé meilleur que ce que laisserait prévoir le contexte conjoncturel qu'ils décrivent : ils font donc sans doute preuve d'un pessimisme personnel relatif lorsqu'ils rendent compte du climat général. Ce cas de figure semble avoir prévalu lors des années de forte expansion de 1989 et 1990. Si ce point de vue est juste, il faut mettre en regard cette période avec l'année 1994, au cours de laquelle la composante est, au contraire, faiblement négative ; ceci signifierait que les industriels ont perçu, lors de la dernière reprise, une réalité plutôt plus favorable que ce qu'indiquaient les carnets, se montrant cette fois plutôt optimistes.

GRAPHIQUE 6  
*Carnet de commande globaux.*



L'information spécifique liée à la question sur la **demande et les carnets étrangers** (cf. graphique 7) offre un éclairage intéressant sur les décalages conjoncturels éventuels entre la France et l'étranger. On y voit par exemple que la composante spécifique est durablement positive en 1984, année de reprise pour l'ensemble des économies occidentales (États-Unis et pays d'Europe de l'Ouest) : ceci illustre le fait que la croissance française bénéficiait alors d'une contribution positive de l'extérieur, marque tout à la fois de la bonne tenue des exportations et d'une relative faiblesse de la demande intérieure. Autrement dit, le comportement de cette composante spécifique constituait un signe que la France ne participait pas pleinement au mouvement de reprise mondiale, même si elle en bénéficiait. De même, cet indicateur suggère que l'activité industrielle est tirée par l'extérieur depuis la

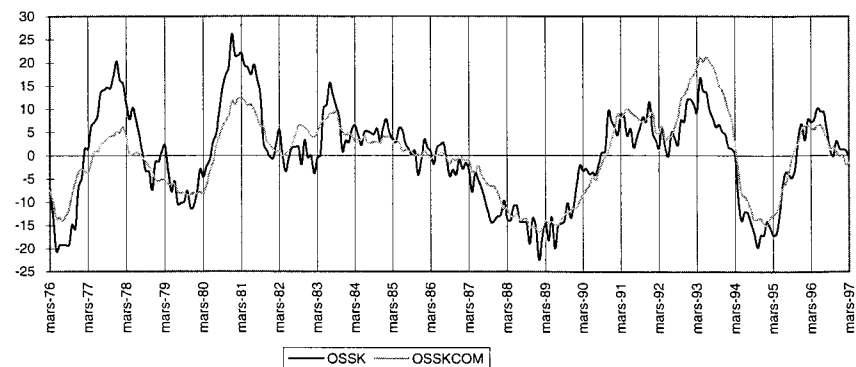
GRAPHIQUE 7  
*Carnets de commande étrangers.*



mi 1996. Toutefois, à l'échelle des vingt dernières années, on peut noter que les évolutions du solde d'opinion sur la demande étrangère semblent se rapprocher de celles suivies par le climat général. Ce fait peut sans doute s'interpréter comme un signe tangible de l'ouverture progressive de l'économie française, et plus spécifiquement du secteur industriel, vers l'extérieur, si bien que les fluctuations conjoncturelles majeures suivies par le pays apparaissent de plus en plus reliées à celles que connaissent ses principaux partenaires.

Un intérêt de la décomposition en informations commune et spécifique du solde d'opinion relatif au **niveau des stocks** (cf. graphique 8) pourrait être de proposer une lecture plus précise de la réponse à cette question qui est, en effet, toujours difficile à analyser. On demande aux industriels s'ils jugent le niveau de leurs stocks supérieur, inférieur ou conforme à la normale. La difficulté porte donc sur l'interprétation à donner à ce terme : la « normale » désigne-t-elle un niveau moyen de stocks, constant au cours du temps, ou bien un niveau de stocks variable, dépendant de la conjoncture du moment ? Dans le schéma présenté ici, les mouvements dus aux fluctuations du climat général traduiraient le comportement de stockage lié au cycle conjoncturel proprement dit : en phase ascendante, pour des raisons tenant à la fois à une

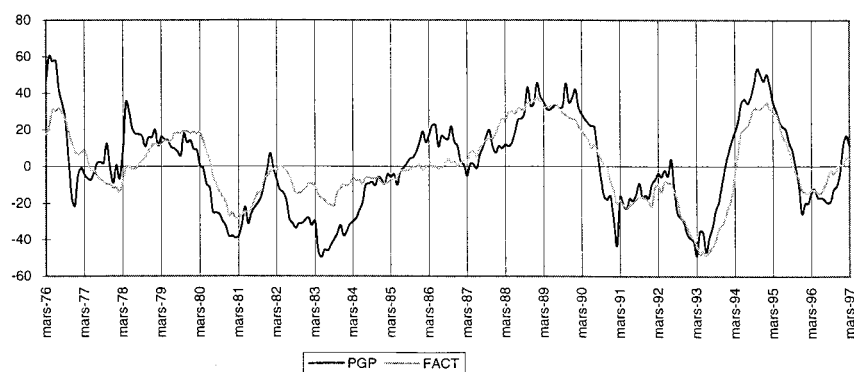
GRAPHIQUE 8  
*Opinion sur les stocks.*



activité passée et à des anticipations de demande favorables, le niveau des stocks a tendance à être jugé de plus en plus léger et les industriels sont donc plutôt enclins à reconstituer leurs stocks, alors que les variations inverses se produisent en phase descendante. Dans ces conditions, l'information spécifiquement apportée par le solde d'opinion relatif à la question sur les stocks fournirait un renseignement sur l'écart de comportement par rapport à ce scénario de référence.

Les fluctuations de l'information spécifiquement apportée par la réponse sur les **perspectives générales de production** (cf. graphique 9) sont amples et persistantes. Il semble qu'en faisant passer le solde d'opinion au dessus ou en dessous du climat général qui se dégage de l'enquête, elle constitue une sorte de miroir du discours ambiant sur l'état de la conjoncture, et de ses excès éventuels. La composante spécifique est ainsi nettement négative de 1982 à la mi-84, période marquée par l'instauration de deux plans de rigueur (juin 1982 et mars 1983) qui sont sans doute à l'origine de diagnostics particulièrement pessimistes ; 1983 fut certes une année de croissance très faible, mais la France ne tomba pas dans la récession, contrairement aux craintes qui s'exprimaient à l'époque dans les médias. De même, alors que l'activité connaissait une reprise vigoureuse qui ne se démentait pas depuis le printemps 1987, la survenue du crack boursier en octobre affecta durablement et spécifiquement le solde portant sur les perspectives générales.

GRAPHIQUE 9  
*Perspectives générales de production.*



## 4 Conclusion

---

La méthode de test du nombre de facteurs dynamiques proposée dans ce papier a l'avantage d'être fondée sur la procédure standard de l'analyse factorielle classique et d'être très simple à mettre en œuvre. Toutefois, il importe de souligner que la procédure standard fournit des estimateurs qui ne sont pas

asymptotiquement efficaces en présence d'autocorrélation, et doivent donc être considérés comme des estimateurs préliminaires.

Dans ce papier, après avoir fixé le nombre de facteurs à retenir à l'aide du test proposé, nous avons choisi de modéliser la dynamique des diverses composantes du modèle, et d'utiliser ensuite le filtre de KALMAN pour obtenir dans une deuxième étape des estimateurs asymptotiquement efficaces. Il se trouve que, sur les données que nous avons étudiées ici, les résultats obtenus sont quasiment identiques à ceux de la première étape. Il serait évidemment intéressant, dans des travaux ultérieurs portant sur d'autres données, de comparer les résultats obtenus en utilisant les deux démarches.

Pour ce qui concerne la présente étude, nous avons mis en évidence qu'un modèle avec un seul facteur commun peut décrire de façon satisfaisante les soldes d'opinion de l'enquête mensuelle d'activité dans l'industrie : ceci permet de justifier l'utilisation de ce facteur comme un indicateur résumé, sorte de climat général. De plus, dans ce cadre, l'apport supplémentaire des différentes questions peut être analysé à travers l'évolution des composantes spécifiques. En définitive, on peut espérer qu'un tel outil constitue une aide dans le travail d'interprétation délicat que doivent mener les conjoncturistes.

## ANNEXE 1

---

### Preuve de la proposition 1

Dans ce paragraphe, on note :  $\|M\| = \max_{i,j} |m_{ij}|$  pour toute matrice  $M$ .

La matrice de variance-covariance du processus s'écrit  $\Sigma = \Lambda \Lambda' + D$ . On définit :  $Q_T^*(x, \theta) = L_T(x, \Lambda \Lambda' + D)$ .

La preuve s'articule en quatre étapes.

*Etape 1 :*  $\|\Sigma\| \rightarrow +\infty \implies Q_T^*(x, \theta) \rightarrow -\infty$

On montre tout d'abord que :  $\|\Sigma\| \rightarrow +\infty \implies \det \Sigma \rightarrow +\infty$ .

En effet, comme on a supposé que  $d_i \neq 0$ , pour  $i = 1, \dots, n$ , la matrice  $D$  est définie positive. Comme, en outre,  $\Sigma \geq D$ ,  $\Sigma$  est aussi définie positive. Pour montrer que  $\det \Sigma \rightarrow +\infty$ , il suffit donc de montrer que l'une au moins des valeurs propres de  $\Sigma$  tend vers  $+\infty$ .

Or, on sait que :  $\|M\|_2^2 = \max_{y \neq 0} \frac{y' M' M y}{y' y}$  définit une autre norme sur l'espace des matrices, et que, si l'on désigne par  $l_i, i = 1, \dots, n$  les valeurs propres de la matrice symétrique définie positive  $\Sigma$ , cette norme vérifie :  $\|\Sigma\|_2 = \max_i l_i$ . Par équivalence des normes sur l'espace des matrices, on obtient ainsi le résultat.

Montrons maintenant que  $\text{tr} \Sigma^{-1} S$  reste borné. Tout d'abord, il est clair que :

$$|\text{tr} \Sigma^{-1} S| \leq n^2 \|\Sigma^{-1}\| \|S\|.$$

Si l'on note :  $d = \min_i d_i$ , comme  $\Sigma \geq D > 0$ , on sait que les valeurs propres de  $\Sigma$  sont toutes supérieures à  $d$ . D'après le résultat précédemment cité sur l'équivalence des normes de matrices, il existe donc un réel  $k$  tel que :

$\|\Sigma^{-1}\| \leq \frac{k}{d}$ . Comme on a, de plus, supposé qu'il existe un réel  $\alpha > 0$ , tel

que :  $\forall \theta \in \Theta, \forall i = 1, \dots, n \quad d_i \geq \alpha$ , on obtient :  $|\text{tr} \Sigma^{-1} S| \leq \frac{kn^2}{\alpha} \|S\|$ .

Comme le processus  $(x_t)$  est stationnaire,  $\|S\|$  est p.s. bornée quand  $T$  varie et on obtient le résultat.

L'étape 1 découle du fait que  $\|\Sigma\| \rightarrow +\infty \implies \det \Sigma \rightarrow +\infty$  et que  $\text{tr} \Sigma^{-1} S$  reste borné. Il en résulte que, lorsqu'on cherche à maximiser  $Q_T^*(x, \theta)$  sur une région de la forme  $R^{np} \times [\alpha, +\infty[^n$ , on peut se restreindre au cas où  $\theta$  varie dans un compact inclus dans  $R^{np} \times [\alpha, +\infty[^n$ .

*Dans toute la suite, on suppose donc que  $\theta$  appartient à un compact  $\Theta$  contenant  $\theta_0$  et inclus dans une région de la forme  $R^{np} \times [\alpha, +\infty[^n$ , où  $\alpha$  est un réel strictement positif fixé.*

*Etape 2 :*  $Q_T^*(x, \theta)$  converge en probabilité, uniformément pour  $\theta \in \Theta$ , vers  $Q_0^*(\theta)$



$$\text{Posons : } Q_0^*(\theta) = \frac{1}{2} \ln(\det \Sigma) - \frac{1}{2} \text{tr}(\Sigma^{-1} \Sigma_0)$$

On a :

$$\begin{aligned} \forall \theta \in \Theta \mid Q_T^*(x, \theta) - Q_0^*(\theta) &= \left| \text{tr} \Sigma^{-1} \left( \frac{1}{T} \sum_t x_t x_t' \right) - \text{tr} \Sigma^{-1} \Sigma_0 \right| \\ &= \left| \text{tr} \Sigma^{-1} \left( \frac{1}{T} \sum_t x_t x_t' - \Sigma_0 \right) \right| \\ &\leq n^2 \left\| \Sigma^{-1} \right\| \left\| \frac{1}{T} \sum_t x_t x_t' - \Sigma_0 \right\| \\ &\leq \frac{kn^2}{\alpha} \left\| \frac{1}{T} \sum_t x_t x_t' - \Sigma_0 \right\| \end{aligned}$$

Cette majoration étant indépendante de  $\theta$ , l'étape 2 s'en déduit.

*Etape 3 : la fonction  $Q_0^*$  admet un maximum unique au point  $\theta_0$*

On considère maintenant un processus  $(z_t)$  non autocorrélé, tel que pour tout  $t$ , la loi de  $z_t$  est de la forme  $N(0, \Sigma)$  avec  $\Sigma = \Lambda \Lambda' + D$  définie comme précédemment. Soit  $\Sigma_0 = \Lambda_0 \Lambda_0' + D_0$  la matrice de variance-covariance associée à la vraie valeur des paramètres. On note  $f(z_t, \theta)$  et  $f(z_t, \theta_0)$  les densités associées.

Les propriétés de l'information de KULLBACK conduisent à l'inégalité suivante :

$$E_{\theta_0} \left( \ln \left( \frac{f(z_t, \theta)}{f(z_t, \theta_0)} \right) \right) < 0 \quad \text{si } \theta \neq \theta_0.$$

Pour  $\theta \neq \theta_0$ , cette inégalité entraîne que :

$$E_{\theta_0} \left( \frac{1}{T-1} \sum_t \ln(f(z_t, \theta)) \right) < E_{\theta_0} \left( \frac{1}{T-1} \sum_t \ln(f(z_t, \theta_0)) \right)$$

c'est-à-dire que :

$$-\frac{1}{2} \ln(\det \Sigma) - \frac{1}{2} \text{tr}(\Sigma^{-1} \Sigma_0) < -\frac{1}{2} \ln(\det \Sigma_0) - \frac{1}{2} \text{tr}(\Sigma_0^{-1} \Sigma_0),$$

ou encore que :  $Q_0^*(\theta) < Q_0^*(\theta_0)$ <sup>12</sup>. On en déduit la troisième étape.

*Etape 4 : la fonction  $Q_0^*$  est continue sur  $\Theta$*

Ceci résulte de la continuité de l'application qui à  $\theta$  associe  $\Sigma$ , des applications *trace* et *déterminant*, ainsi que de l'application qui à  $\Sigma$  associe  $\Sigma^{-1}$ .

Grâce à l'étape 1, on a vu qu'on pouvait supposer que  $\theta \in \Theta$ , où  $\Theta$  est un compact inclus dans  $R^{np} \times (R^{+*})^n$ . Dans ces conditions, les trois étapes suivantes permettent d'appliquer les théorèmes généraux sur la convergence des M-estimateurs (voir par exemple NEWEY et MacFADDEN [1994], théorème 2.1) et de conclure que le M-estimateur  $\hat{\theta}_T$  est convergent.

12. On notera que le raisonnement qui est fait ici consiste à appliquer l'inégalité de l'information de KULLBACK à la pseudo-loi, en faisant momentanément comme si c'était la vraie loi. Cette inégalité est en effet une propriété de la loi étudiée, qu'elle soit ou non la vraie loi des observations.

### Preuve du lemme 2

Montrons d'abord la propriété pour  $h = 0$ .

Comme  $\text{vec}(x_t x_t') = x_t \otimes x_t$ , il est clair que :

$$(\text{vec} x_t x_t')(\text{vec} x_t x_t')' = x_t \otimes x_t x_t' \otimes x_t' = x_t x_t' \otimes x_t x_t'.$$

En notant  $(e_i)$  les vecteurs de la base canonique de  $R^n$ , et  $x_t = \sum_{i=1}^n x_{it} e_i$ , on obtient :

$$\begin{aligned} E(x_t x_t' \otimes x_t x_t') &= \sum_{i,j,k,l} E(x_{it} x_{jt} x_{kt} x_{lt}) e_i e_j' \otimes e_k e_l' \\ &= \sum_{i,j,k,l} E(x_{it} x_{jt} x_{kt} x_{lt}) e_i \otimes e_k \cdot e_j' \otimes e_l' \\ &= \sum_{i,j,k,l} (\sigma_{ij} \sigma_{kl} + \sigma_{ik} \sigma_{jl} + \sigma_{il} \sigma_{jk}) e_i \otimes e_k \cdot e_j' \otimes e_l' \end{aligned}$$

(voir par exemple Anderson [1984] page 49, pour le calcul de  $E(x_{it} x_{jt} x_{kt} x_{lt})$ ). Calculons chacun des trois termes de cette somme.

$$\begin{aligned} \sum_{i,j,k,l} \sigma_{ij} \sigma_{kl} e_i \otimes e_k \cdot e_j' \otimes e_l' &= \sum_{i,j,k,l} \sigma_{ij} \sigma_{kl} e_i e_j' \otimes e_k e_l' \\ &= \left( \sum_{i,j} \sigma_{ij} e_i e_j' \right) \otimes \left( \sum_{k,l} \sigma_{kl} e_k e_l' \right) = \Sigma \otimes \Sigma \end{aligned}$$

$$\begin{aligned} \sum_{i,j,k,l} \sigma_{ik} \sigma_{jl} e_i \otimes e_k \cdot e_j' \otimes e_l' &= \left( \sum_{i,k} \sigma_{ik} \text{vec}(e_k e_l') \right) \left( \sum_{j,l} \sigma_{jl} \text{vec}(e_l e_j') \right)' \\ &= \text{vec} \Sigma' \cdot (\text{vec} \Sigma')' = \text{vec} \Sigma \cdot (\text{vec} \Sigma)' \end{aligned}$$

Enfin, comme :

$\forall u \in R^n, \forall v \in R^n : K_{nn} u \otimes v = K_{nn} \text{vec}(vu') = \text{vec}(uv') = v \otimes u$ , on a :

$$\begin{aligned} \sum_{i,j,k,l} \sigma_{il} \sigma_{jk} e_i \otimes e_k \cdot e_j' \otimes e_l' &= K_{nn} \sum_{i,j,k,l} \sigma_{il} \sigma_{jk} e_k \otimes e_i \cdot e_j' \otimes e_l' \\ &= K_{nn} \sum_{i,j,k,l} \sigma_{il} \sigma_{jk} e_k e_j' \otimes e_i e_l' \\ &= K_{nn} \left( \sum_{j,k} \sigma_{jk} e_k e_j' \right) \otimes \left( \sum_{i,l} \sigma_{il} e_i e_l' \right) \\ &= K_{nn} \Sigma' \otimes \Sigma = K_{nn} \Sigma \otimes \Sigma \end{aligned}$$

Comme  $K_{nn}\Sigma \otimes \Sigma = \Sigma \otimes \Sigma K_{nn}$ , on a ainsi obtenu :

$$(1) \quad \begin{aligned} E((\text{vec}x_t x_t')(\text{vec}x_t x_t')') &= E(x_t x_t' \otimes x_t x_t') \\ &= (\text{vec}\Sigma)(\text{vec}\Sigma)' + (\Sigma \otimes \Sigma)(I_{n^2} + K_{nn}) \end{aligned}$$

En outre, comme  $E(\text{vec}x_t x_t') = \text{vec}\Sigma$ , on obtient de façon immédiate :

$$(2) \quad V(\text{vec}x_t x_t') = E(\text{vec}(x_t x_t' - \Sigma)\text{vec}(x_t x_t' - \Sigma)') = (\Sigma \otimes \Sigma)(I_{n^2} + K_{nn})$$

Montrons maintenant la propriété pour  $h > 0$ . On sait que :

$$\begin{aligned} E(x_t x_t' / x_{t-h}) &= V(x_t / x_{t-h}) + E(x_t / x_{t-h})E(x_t / x_{t-h})' \\ &= \Sigma - \Gamma(h)\Sigma^{-1}\Gamma(h)' + \Gamma(h)\Sigma^{-1}x_{t-h}x_{t-h}'\Sigma^{-1}\Gamma(h)' \end{aligned}$$

On en déduit :

$$\begin{aligned} E((\text{vec}x_t x_t')(\text{vec}x_{t-h}x_{t-h}')') &= E(E(\text{vec}x_t x_t' / x_{t-h})(\text{vec}x_{t-h}x_{t-h}')') \\ &= E(\text{vec}(\Sigma - \Gamma(h)\Sigma^{-1}\Gamma(h)') \\ &\quad + \Gamma(h)\Sigma^{-1}x_{t-h}x_{t-h}'\Sigma^{-1}\Gamma(h)')(\text{vec}x_{t-h}x_{t-h}')') \\ &= \text{vec}(\Sigma - \Gamma(h)\Sigma^{-1}\Gamma(h)')(\text{vec}\Sigma)' \\ &\quad + E(\text{vec}(\Gamma(h)\Sigma^{-1}x_{t-h}x_{t-h}'\Sigma^{-1}\Gamma(h)')(\text{vec}x_{t-h}x_{t-h}')') \end{aligned}$$

Or, d'une part :

$$\begin{aligned} &\text{vec}(\Sigma - \Gamma(h)\Sigma^{-1}\Gamma(h)')(\text{vec}\Sigma)' \\ &= (\text{vec}\Sigma - \Gamma(h) \otimes \Gamma(h)\text{vec}\Sigma^{-1})(\text{vec}\Sigma)' \\ &= (\text{vec}\Sigma - \Gamma(h) \otimes \Gamma(h)\Sigma^{-1} \otimes \Sigma^{-1}\text{vec}\Sigma)(\text{vec}\Sigma)' \\ &= (I_{n^2} - \Gamma(h)\Sigma^{-1} \otimes \Gamma(h)\Sigma^{-1})(\text{vec}\Sigma)(\text{vec}\Sigma)' \end{aligned}$$

d'autre part, en appliquant (1) :

$$\begin{aligned} &E(\text{vec}(\Gamma(h)\Sigma^{-1}x_{t-h}x_{t-h}'\Sigma^{-1}\Gamma(h)')(\text{vec}x_{t-h}x_{t-h}')') \\ &= \Gamma(h)\Sigma^{-1} \otimes \Gamma(h)\Sigma^{-1} E(\text{vec}(x_{t-h}x_{t-h}')(\text{vec}x_{t-h}x_{t-h}')') \\ &= \Gamma(h)\Sigma^{-1} \otimes \Gamma(h)\Sigma^{-1} ((\text{vec}\Sigma)(\text{vec}\Sigma)' + (\Sigma \otimes \Sigma)(I_{n^2} + K_{nn})) \end{aligned}$$

Finalement, on obtient :

$$(3) \quad \begin{aligned} &E((\text{vec}x_t x_t')(\text{vec}x_{t-h}x_{t-h}')') \\ &= (\text{vec}\Sigma)(\text{vec}\Sigma)' + (\Gamma(h)\Sigma^{-1} \otimes \Gamma(h)\Sigma^{-1})(\Sigma \otimes \Sigma)(I_{n^2} + K_{nn}) \\ &= (\text{vec}\Sigma)(\text{vec}\Sigma)' + (\Gamma(h) \otimes \Gamma(h))(I_{n^2} + K_{nn}) \end{aligned}$$

Comme  $E(\text{vec}x_t x_t') = E(\text{vec}x_{t-h}x_{t-h}') = \text{vec}\Sigma$ ,  
il s'ensuit de façon immédiate :

$$(4) \quad E(\text{vec}(x_t x_t' - \Sigma)\text{vec}(x_{t-h}x_{t-h}' - \Sigma)') = (\Gamma(h) \otimes \Gamma(h))(I_{n^2} + K_{nn})$$

**Preuve de la proposition 3**

i) En notant  $S = \frac{1}{T} \sum_t x_t x_t'$ , on sait que :

$$\frac{\partial L_T}{\partial \text{vec} \Sigma}(x, \Sigma) = -\frac{1}{2} K_{nn} \text{vec} \left( \Sigma^{-1} - \Sigma^{-1} S \Sigma^{-1} \right)$$

On en déduit (voir par exemple LÜTKEPOHL [1993], page 471, pour les formules de dérivation utilisées) :

$$\begin{aligned} \frac{\partial^2 L_T}{(\partial \text{vec} \Sigma)(\partial \text{vec} \Sigma)'}(x, \Sigma) &= -\frac{1}{2} \frac{\partial}{\partial \text{vec} \Sigma} \left[ \text{vec} \Sigma^{-1} - \text{vec}(\Sigma^{-1} S \Sigma^{-1}) \right]' K_{nn} \\ &= -\frac{1}{2} \left[ \frac{\partial \text{vec} \Sigma^{-1}}{(\partial \text{vec} \Sigma)'} - (I_n \otimes \Sigma^{-1} S) \frac{\partial \text{vec} \Sigma^{-1}}{(\partial \text{vec} \Sigma)'} - (\Sigma^{-1} S \otimes I_n) \frac{\partial \text{vec} \Sigma^{-1}}{(\partial \text{vec} \Sigma)'} \right]' K_{nn} \\ &= \frac{1}{2} \left[ \Sigma^{-1} \otimes \Sigma^{-1} - (I_n \otimes \Sigma^{-1} S)(\Sigma^{-1} \otimes \Sigma^{-1}) \right. \\ &\quad \left. - (\Sigma^{-1} S \otimes I_n)(\Sigma^{-1} \otimes \Sigma^{-1}) \right]' K_{nn} \\ &= \frac{1}{2} \left[ \Sigma^{-1} \otimes \Sigma^{-1} - \Sigma^{-1} \otimes \Sigma^{-1} S \Sigma^{-1} - \Sigma^{-1} S \Sigma^{-1} \otimes \Sigma^{-1} \right] K_{nn} \end{aligned}$$

Comme  $\frac{\partial^2 L_T}{(\partial \text{vech} \Sigma)(\partial \text{vech} \Sigma)'}(x, \Sigma) = D_n' \frac{\partial^2 L_T}{(\partial \text{vec} \Sigma)(\partial \text{vec} \Sigma)'}(x, \Sigma) D_n$ ,

il s'ensuit :

$$\begin{aligned} \frac{\partial^2 L_T}{(\partial \text{vech} \Sigma)(\partial \text{vech} \Sigma)'}(x, \Sigma) \\ &= \frac{1}{2} D_n' \left[ \Sigma^{-1} \otimes \Sigma^{-1} - \Sigma^{-1} \otimes \Sigma^{-1} S \Sigma^{-1} - \Sigma^{-1} S \Sigma^{-1} \otimes \Sigma^{-1} \right] D_n \end{aligned}$$

Enfin, puisque  $E(S) = E\left(\frac{1}{T} \sum_t x_t x_t'\right) = \Sigma$ , on obtient :

$$J_0 = E \left[ -\frac{\partial^2 L_T}{(\partial \text{vech} \Sigma)(\partial \text{vech} \Sigma)'}(x, \Sigma) \right] = \frac{1}{2} D_n' \Sigma^{-1} \otimes \Sigma^{-1} D_n$$

ii) Résulte de i) de façon immédiate puisque :

$$\begin{aligned} \frac{\partial L_T}{\partial \text{vech} \Sigma}(x, \Sigma) &= -\frac{1}{2} D_n' \Sigma^{-1} \otimes \Sigma^{-1} \text{vec} \left( \Sigma - \frac{1}{T} \sum_t x_t x_t' \right) \\ &= -\frac{1}{2} D_n' \Sigma^{-1} \otimes \Sigma^{-1} D_n \text{vech} \left( \Sigma - \frac{1}{T} \sum_t x_t x_t' \right) \end{aligned}$$

iii) Le vecteur  $z_t = \text{vec}(\Sigma - x_t x_t')$  étant stationnaire d'espérance nulle, on sait que :  $\sqrt{T} \frac{1}{T} \sum_t z_t \xrightarrow{L} N\left(0, \left(\sum_{h \in Z} E(z_t z_{t-h}')\right)\right)$ .

Par application du lemme 2, on sait que :

$$\forall h \in Z \quad E(z_t z_{t-h}') = \Gamma(h) \otimes \Gamma(h)(I_{n^2} + K_{nn})$$

En notant  $\Omega_0 = \sum_{h \in Z} \Gamma(h) \otimes \Gamma(h)$ , on a ainsi :

$$\sqrt{T} \text{vec}\left(\Sigma - \frac{1}{T} \sum_t x_t x_t'\right) \xrightarrow{L} N\left(0, \Omega_0(I_{n^2} + K_{nn})\right)$$

Comme  $\text{vech}\left(\Sigma - \frac{1}{T} \sum_t x_t x_t'\right) = D_n^+ \text{vec}\left(\Sigma - \frac{1}{T} \sum_t x_t x_t'\right)$  et comme  $K_{nn} D_n^{+'} = D_n^{+'}$ , on obtient :

$$\sqrt{T} \text{vech}\left(\Sigma - \frac{1}{T} \sum_t x_t x_t'\right) \xrightarrow{L} N\left(0, 2D_n^+ \Omega_0 D_n^{+'}\right)$$

iv) résulte de ii) et iii) de façon immédiate.

### Preuve de la proposition 4

Comme les estimateurs  $\hat{\gamma}_T$  et  $\hat{\theta}_T$  sont convergents, on peut effectuer les mêmes développements limités sous l'hypothèse nulle que ceux qui sont faits par GOURIÉROUX et MONFORT [1989]. De fait, on peut écrire :

$$(1) \quad \begin{aligned} \frac{\partial Q_T}{\partial \gamma}(x, \hat{\gamma}_T^0) &\# \frac{\partial Q_T}{\partial \gamma}(x, \gamma_0) + \frac{\partial^2 Q_T}{\partial \gamma \partial \gamma'}(x, \gamma_0)(\hat{\gamma}_T^0 - \gamma_0) \\ &\# \frac{\partial Q_T}{\partial \gamma}(x, \gamma_0) - J_0 \frac{\partial h}{\partial \theta'}(\theta_0)(\hat{\theta}_T - \theta_0) \end{aligned}$$

$$(2) \quad \begin{aligned} \frac{\partial Q_T^*}{\partial \theta}(x, \hat{\theta}_T) = 0 &\# \frac{\partial Q_T^*}{\partial \theta}(x, \theta_0) + \frac{\partial^2 Q_T^*}{\partial \theta \partial \theta'}(x, \theta_0)(\hat{\theta}_T - \theta_0) \\ &\# \frac{\partial h'}{\partial \theta}(\theta_0) \frac{\partial Q_T}{\partial \gamma}(x, \gamma_0) - \frac{\partial h'}{\partial \theta}(\theta_0) J_0 \frac{\partial h}{\partial \theta'}(\theta_0)(\hat{\theta}_T - \theta_0) \end{aligned}$$

Dans le cas où la matrice  $\frac{\partial h}{\partial \theta'}(\theta_0)$  est de plein rang colonne, GOURIÉROUX et MONFORT montrent que ces deux relations entraînent :

$$(3) \quad \frac{\partial Q_T}{\partial \gamma}(x, \hat{\gamma}_T^0) \# M' \frac{\partial Q_T}{\partial \gamma}(x, \gamma_0).$$

Pour montrer que ce résultat reste vrai lorsque  $\frac{\partial h}{\partial \theta'}(\theta_0)$  n'est pas de plein rang, il suffit, compte tenu de la relation (1) et du fait que  $M = I_m - P$ , de montrer que :

$$(4) \quad J_0 \frac{\partial h}{\partial \theta'}(\theta_0)(\hat{\theta}_T - \theta_0) \# P' \frac{\partial Q_T}{\partial \gamma}(x, \gamma_0)$$

Or, en utilisant la définition de  $P$  et la relation (2), on peut écrire :

$$\begin{aligned} P' \frac{\partial Q_T}{\partial \gamma}(x, \gamma_0) &= J_0 \frac{\partial h}{\partial \theta'}(\theta_0) \left( \frac{\partial h'}{\partial \theta}(\theta_0) J_0 \frac{\partial h}{\partial \theta'}(\theta_0) \right)^+ \frac{\partial h'}{\partial \theta}(\theta_0) \frac{\partial Q_T}{\partial \gamma}(x, \gamma_0) \\ &\# J_0 \frac{\partial h}{\partial \theta'}(\theta_0) \left( \frac{\partial h'}{\partial \theta}(\theta_0) J_0 \frac{\partial h}{\partial \theta'}(\theta_0) \right)^+ \frac{\partial h'}{\partial \theta}(\theta_0) J_0 \frac{\partial h}{\partial \theta'}(\theta_0)(\hat{\theta}_T - \theta_0) \\ &= J_0 P \frac{\partial h}{\partial \theta'}(\theta_0)(\hat{\theta}_T - \theta_0) \end{aligned}$$

Comme  $P$  est la matrice de la projection  $J_0$ -orthogonale sur l'espace engendré par les colonnes de  $\frac{\partial h}{\partial \theta'}(\theta_0)$ , on a :  $P \frac{\partial h}{\partial \theta'}(\theta_0) = \frac{\partial h}{\partial \theta'}(\theta_0)$  et on obtient ainsi le résultat.

Comme on a supposé :  $\sqrt{T} \frac{\partial Q_T}{\partial \gamma}(x, \gamma_0) \xrightarrow{L} N(0, \Omega)$ , il en résulte que sous  $H_0$  :

$$\sqrt{T} \frac{\partial Q_T}{\partial \gamma}(x, \hat{\gamma}_T^0) \xrightarrow{L} N(0, M' \Omega M).$$

Puisque  $rg(M' \Omega M) = rg(M) = m - q_0$ , on obtient :

$$T \left( \frac{\partial Q_T}{\partial \gamma}(x, \hat{\gamma}_T^0) \right)' (M' \Omega M)^{-1} \left( \frac{\partial Q_T}{\partial \gamma}(x, \hat{\gamma}_T^0) \right) \xrightarrow{L} \chi^2(m - q_0),$$

et par conséquent :

$$T \left( \frac{\partial Q_T}{\partial \gamma}(x, \hat{\gamma}_T^0) \right)' (\hat{M}' \hat{\Omega} \hat{M})^{-1} \left( \frac{\partial Q_T}{\partial \gamma}(x, \hat{\gamma}_T^0) \right) \xrightarrow{L} \chi^2(m - q_0).$$

**Preuve du lemme 5**

i) Notons :  $\theta_0 = \begin{pmatrix} \text{vec} \Lambda_0 \\ d_0 \end{pmatrix}$  et  $\theta = \begin{pmatrix} \text{vec} \Lambda \\ d \end{pmatrix}$ .

$$\begin{aligned} h(\theta_0 + \theta) - h(\theta_0) &= \text{vech}((\Lambda_0 + \Lambda)(\Lambda_0 + \Lambda)' + (D_0 + D)) \\ &\quad - \text{vech}(\Lambda_0 \Lambda_0' + D_0) \\ &= \text{vech}(\Lambda_0 \Lambda' + \Lambda \Lambda_0' + \Lambda \Lambda' + D) \end{aligned}$$

Comme l'application  $\theta \mapsto \text{vech}(\Lambda_0 \Lambda' + \Lambda \Lambda_0' + D)$  est linéaire et comme  $\|\text{vech}(\Lambda \Lambda')\| = o(\|\Lambda\|)$ , on en déduit :

$$\begin{aligned} \frac{\partial h}{\partial \theta'}(\theta_0) \cdot \theta &= \text{vech}(\Lambda_0 \Lambda' + \Lambda \Lambda_0' + D) \\ &= D_n^+ \text{vec}(\Lambda_0 \Lambda' + \Lambda \Lambda_0' + D) \\ &= D_n^+ \left[ (I_n^2 + K_{nn}) \text{vec}(\Lambda \Lambda_0') + \text{vec} D \right] \end{aligned}$$

Or :  $\text{vec}(\Lambda \Lambda_0') = (\Lambda_0 \otimes I_n) \text{vec}(\Lambda)$  et :

$$\begin{aligned} \text{vec} D &= \text{vec} \left( \sum_{i=1}^n d_i e_i e_i' \right) = \sum_{i=1}^n \text{vec}(e_i e_i' d e_i') \\ &= \sum_{i=1}^n (e_i \otimes e_i e_i') \text{vec}(d) = \Delta_n d \end{aligned}$$

On obtient donc :

$$\frac{\partial h}{\partial \theta'}(\theta_0) = D_n^+ [(I_n + K_{nn})(\Lambda_0 \otimes I_n) \Delta_n] = D_n^+ [2\Lambda_0 \otimes I_n \Delta_n]$$

puisque  $D_n^+ K_{nn} = D_n^+$ .

ii)  $\frac{\partial h}{\partial \theta'}(\theta_0)$  est une matrice  $(n(n+1)/2, n(p+1))$ . Pour étudier son rang, il suffit de déterminer la dimension de son noyau.

Soit  $\theta = \begin{pmatrix} \text{vec} \Lambda \\ d \end{pmatrix}$  tel que  $\frac{\partial h}{\partial \theta'}(\theta_0) \cdot \theta = 0$ . On a de façon équivalente :

$$\text{vech}(\Lambda_0 \Lambda' + \Lambda \Lambda_0' + D) = 0 \text{ et donc : } \Lambda_0 \Lambda' + \Lambda \Lambda_0' + D = 0.$$

Posons  $P_{\Lambda_0} = \Lambda_0 (\Lambda_0' D_0^{-1} \Lambda_0)^{-1} \Lambda_0' D_0^{-1}$ .

La relation précédente entraîne :

$$(I_n - P_{\Lambda_0})(\Lambda_0 \Lambda' + \Lambda \Lambda_0' + D)(I_n - P_{\Lambda_0}') = 0$$

et donc :  $(I_n - P_{\Lambda_0})D(I_n - P_{\Lambda_0}') = 0$ . Or,  $I_n - P_{\Lambda_0} = \Phi_0 D_0^{-1}$ , où  $\Phi_0$  est la matrice qui intervient dans la condition suffisante d'identifiabilité locale. On



obtient alors :  $\Phi_0 D_0^{-1} D D_0^{-1} \Phi_0' = 0$  et la nullité des termes diagonaux fournit :

$$\forall i \quad \sum_j \varphi_{ij}^2 \frac{d_j}{d_{0j}^2} = 0$$

Comme la condition de ANDERSON RUBIN est supposée vérifiée,  $(\varphi_{ij}^2)$  est par hypothèse une matrice inversible. Ceci entraîne que  $D = 0$ , donc que  $\Lambda_0 \Lambda' + \Lambda \Lambda_0' = 0$ .

On obtient alors :  $(I_n - P_{\Lambda_0})(\Lambda_0 \Lambda' + \Lambda \Lambda_0') = 0$ , d'où :  $(I_n - P_{\Lambda_0})\Lambda \Lambda_0' = 0$ . Comme  $\Lambda_0$  est une matrice  $(n, p)$  de rang  $p$ , on a aussi :  $(I_n - P_{\Lambda_0})\Lambda = 0$ . Il en résulte qu'il existe une matrice carrée  $M$  de taille  $p$  telle que  $\Lambda = \Lambda_0 M$ .

La relation  $\Lambda_0 \Lambda' + \Lambda \Lambda_0' = 0$  se réécrit alors :

$$\Lambda_0 M' \Lambda_0' + \Lambda_0 M \Lambda_0' = \Lambda_0 (M' + M) \Lambda_0' = 0$$

Comme  $\Lambda_0$  est une matrice  $(n, p)$  de rang  $p$ , on en déduit :  $M' + M = 0$ , c'est-à-dire que  $M$  est une matrice antisymétrique. On vient ainsi de montrer que :

$$\ker \frac{\partial h}{\partial \theta'}(\theta_0) \subset \left\{ \theta = \begin{pmatrix} \text{vec} \Lambda_0 M \\ 0 \end{pmatrix} / M + M' = 0 \right\}$$

Comme l'inclusion réciproque est triviale, on en déduit que la dimension de  $\ker \frac{\partial h}{\partial \theta'}(\theta_0)$  est celle de l'ensemble des matrices  $(p, p)$  antisymétriques, c'est-à-dire :  $\frac{p(p-1)}{2}$ . On en déduit :  $rg \frac{\partial h}{\partial \theta'}(\theta) = n(p+1) - \frac{p(p-1)}{2}$ .

### Preuve de la proposition 6

i) Il suffit d'appliquer la proposition 4 avec :  $\theta = \begin{pmatrix} \text{vec } \Lambda \\ d \end{pmatrix}$ ,  $\gamma = \text{vech } \Sigma$  et :  
 $Q_T(x, \gamma) = L_T(x, \Sigma)$ .

On sait en effet dans ce cas que :

- $\hat{\gamma}_T$  et  $\hat{\theta}_T$  sont des estimateurs convergents de  $\gamma$  et  $\theta$  ;
- $\frac{\partial L_T}{\partial \text{vech } \Sigma}(x, \hat{\Sigma}_T^0) = -J_0 \text{vech} \left( \hat{\Sigma}_T^{0-1} - A \right)$  (cf. propriété 3 ii) )
- $\text{vech}(\hat{\Sigma}_T^0) = h(\hat{\theta}_T) = \text{vech}(\hat{\Lambda}_T \hat{\Lambda}'_T + \hat{D}_T)$
- $\sqrt{T} \frac{\partial L_T}{\partial \text{vech}(\Sigma)}(x, \Sigma) = \sqrt{T} \frac{\partial Q_T}{\partial \gamma}(x, \gamma) \xrightarrow{L} N \left( 0, 2J_0 D_n^+ \Omega_0 D_n^{+'} J_0 \right)$   
 (cf. propriété 3 iii) )

c'est-à-dire, en reprenant les notations de la proposition 4 :

$$\Omega = 2J_0 D_n^+ \Omega_0 D_n^{+'} J_0.$$

- $m = n(n+1)/2$  et  $\Omega$  est de plein rang  $n(n+1)/2$ , donc définie positive.
- $rg P = q_0 = n(p+1) - \frac{p(p-1)}{2}$  d'après le lemme 5, puisque  $P$  est un projecteur sur le sous-espace engendré par les colonnes de  $\frac{\partial h}{\partial \theta'}(\theta_0)$ . On en déduit que :

$$\begin{aligned} rg M &= m - q_0 = \frac{n(n+1)}{2} - \left( (p+1)n - \frac{p(p-1)}{2} \right) \\ &= \frac{1}{2}((n-p)^2 - (n+p)). \end{aligned}$$

Pour appliquer la proposition 4, il reste donc à calculer une inverse généralisée de  $M' \Omega M$ . Or, on a :

$$\begin{aligned} M' \Omega M &= 2M' J_0 D_n^+ \Omega_0 D_n^{+'} J_0 M \\ &= 2J_0 M D_n^+ \Omega_0 D_n^{+'} M' J_0 \end{aligned}$$

car  $M$  est une matrice de projection  $J_0$ -orthogonale. On vérifie alors de façon immédiate que, si  $\left( M D_n^+ \Omega_0 D_n^{+'} M' \right)^-$  est une inverse généralisée de  $M D_n^+ \Omega_0 D_n^{+'} M'$ , la matrice  $\left( M' \Omega M \right)^- = \frac{1}{2} J_0^{-1} \left( M D_n^+ \Omega_0 D_n^{+'} M' \right)^- J_0^{-1}$  est une inverse généralisée de  $M' \Omega M$ .

Alors :

$$\begin{aligned}
T \left( \frac{\partial L_T}{\partial \text{vech} \Sigma} (x, \widehat{\Sigma}_T^0) \right)' (M' \Omega M)^{-} & \left( \frac{\partial L_T}{\partial \text{vech} \Sigma} (x, \widehat{\Sigma}_T^0) \right) \\
& = T \cdot \text{vech} \left( \widehat{\Sigma}_T^{0-1} - S \right)' J_0 \frac{1}{2} J_0^{-1} \left( M D_n^+ \Omega_0 D_n^{+'} M' \right)^{-} \\
& \qquad \qquad \qquad J_0^{-1} J_0 \text{vech} \left( \widehat{\Sigma}_T^{0-1} - S \right) \\
& = \frac{T}{2} \text{vech} \left( \widehat{\Sigma}_T^{0-1} - S \right)' \left( M D_n^+ \Omega_0 D_n^{+'} M' \right)^{-} \text{vech} \left( \widehat{\Sigma}_T^{0-1} - S \right)
\end{aligned}$$

ii) On sait que  $\text{vech}(\Sigma - S) = -J_0^{-1} \frac{\partial L_T}{\partial \text{vech} \Sigma} (x, \Sigma)$ .

On a donc :

$$\sqrt{T} \text{vech} \left( \widehat{\Sigma}_T^0 - S \right) \xrightarrow[H_0]{L} N \left( 0, 2 M D_n^+ \Omega_0 D_n^{+'} M' \right),$$

si bien que :

$$\xi_S = T \left( \text{vech}(\widehat{\Sigma}_T^0 - S) \right)' \left[ V_{as} \left( \sqrt{T} \text{vech} \left( \widehat{\Sigma}_T^0 - S \right) \right) \right]^{-} \text{vech}(\widehat{\Sigma}_T^0 - S)$$

Comme  $S = \frac{1}{T} \sum_t x_t x_t'$  est l'estimateur non contraint, on a bien écrit une statistique de HAUSMAN-WALD.

## ANNEXE 7

---

### Résultat des tests de racine unité et de stationnarité appliqués aux soldes d'opinion

Les six soldes d'opinion utilisés n'étant pas centrés, on effectue un test de DICKEY-FULLER augmenté avec constante (les soldes étant nécessairement compris entre + 100 % et - 100 %, introduire une tendance déterministe n'a *a priori* aucun sens pour faire le test). Pour déterminer le nombre de retards utilisés, on choisit de commencer par en mettre un nombre fixe (en l'espèce, six), puis on teste pas à pas la nullité du dernier retard, ce jusqu'à ce que le coefficient du dernier retard soit significativement non nul. La valeur limite du test à 5 % est de - 2,89. Par ailleurs, on teste l'hypothèse nulle que chaque solde est stationnaire à l'aide du test KPSS (KWIATKOWSKI-PHILLIPS-SCHMIDT-SHIN). Là encore, on teste dans le cas avec constante (et non avec constante et tendance déterministe). La valeur limite à 5 % est de 0,463. Le nombre de retards choisi pour le calcul de la fenêtre de BARTLETT est choisi conformément à la procédure recommandée par NEWEY-WEST [1994]. Les deux types de tests conduisent sans ambiguïté à retenir l'hypothèse que les séries utilisées sont stationnaires.

	TPPA	TPPRE	OSCD	OSCDE	OSSK	PGP
nb de retards du DFA	4	0	3	5	2	1
Student du dernier retard	3,27		2,21	2,90	2,22	2,67
test DFA	- 4,21	- 3,00	- 2,97	- 3,49	- 3,33	- 3,25
test KPSS	0,12	0,20	0,12	0,14	0,21	0,17

## • Références bibliographiques

- ANDERSON, T.W. (1984) – *An Introduction to Multivariate Statistical Analysis*, 2nd ed., Wiley.
- CRAGG, J.G., DONALD, S.G. (1995) – Factor analysis under more general conditions with reference to heteroskedasticity of unknown form, in *Advances in econometrics and quantitative economics: Essays in honor of Professor C.R. Rao, G.S. Maddala, P.C.B. Phillips, and T.N. Srinivasan* ed. Blackwell, Cambridge MA.
- FORNI, M. , REICHLIN, L. (1998) – Let's get real: a dynamic factor analytical approach to disaggregated business cycle, *Review of Economic Studies*.
- GOURIÉROUX, C., MONFORT, A. (1989a) – A general framework for testing a null hypothesis in a "mixed" form, *Econometric Theory*, 5, pp. 63-82.
- GOURIÉROUX, C., MONFORT, A. (1989b) – *Statistique et Modèles Econométriques*, vol. 2, Economica.
- HARVEY, A.C. (1989) – *Forecasting, structural time series models and the Kalman filter*, Cambridge University Press.
- KWIATKOWSKI, D., PHILLIPS, P.C.B., SCHMIDT, P., SHIN, S. (1992) – Testing the null hypothesis of stationarity against the alternative of a unit root, *Journal of Econometrics*, 54, pp. 159-178.
- LAWLEY, D.N., MAXWELL, A.E. (1971) – *Factor Analysis as a Statistical Method*, New York Macmillan Publishing Co., Inc.
- LÜTKEPOHL, H., (1993) – *Introduction to Multiple Time Series Analysis*, 2nd ed, Springer Verlag.
- MAGNUS, J.R., NEUDECKER, H. (1988) – *Matrix Differential Calculus with Applications in Statistics and Econometrics*, Chichester, J. Wiley.
- NEWBY W.K., MACFADDEN, D. (1994) – Large Sample Estimation and Hypothesis testing, in *Handbook of Econometrics*, vol. 4, pp. 2113-2245.
- NEWBY W.K., WEST, K.D. (1994) – Automatic Lag Selection in Covariance Matrix Estimation, *Review of Economic Studies*, 61, pp. 631-653.
- QUAH D., SARGENT, T.J. (1993) – A dynamic index model for large cross-sections, in *Business cycles, indicators and forecasting*, J.H. Stock and M.W. Watson Ed., University of Chicago Press.
- STOCK, J.H., WATSON, M.W. (1989) – New indexes of coincident and leading indicators, in *NBER Macroeconomics Annual*, Blanchard and Fisher Ed. MIT Press, Cambridge.
- STOCK, J.H., WATSON, M.W. (1993) – A procedure for predicting recessions with leading indicators: econometric issues and recent experience, in *Business cycles, indicators and forecasting*, J.H. Stock and M.W. Watson Ed., University of Chicago Press.
- WATSON, M.W., (1986) – Univariate detrending methods with stochastic trends, *Journal of Monetary Econometrics*, 18, pp. 49-75.