



**HAL**  
open science

# Estimation et détection d'un signal contaminé par un bruit autorégressif

Abdessamad Ezzahar

► **To cite this version:**

Abdessamad Ezzahar. Estimation et détection d'un signal contaminé par un bruit autorégressif. Modélisation et simulation. Université Joseph-Fourier - Grenoble I, 1991. Français. NNT: . tel-00339831

**HAL Id: tel-00339831**

**<https://theses.hal.science/tel-00339831>**

Submitted on 19 Nov 2008

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

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

# THESE

présentée par

**EZZAHAR Abdessamad**

*pour obtenir le titre de*  
**Docteur de l'Université Joseph Fourier-Grenoble I**  
*(arrêté ministériel du 5 Juillet 1984)*  
**"Mathématiques Appliquées"**

## **ESTIMATION ET DETECTION D'UN SIGNAL CONTAMINE PAR UN BRUIT AUTOREGRESSIF**

-----

Thèse soutenue le 31 Octobre 1991 devant la commission d'examen

<b>D.T PHAM</b>	<b>Président</b>
<b>F. BRODEAU</b>	
<b>B. GAREL</b>	<b>Examineurs</b>
<b>G. JOURDAIN</b>	
<b>A. LE BRETON</b>	

Thèse préparée au sein du laboratoire LMC, Institut IMAG



*Ce travail a été réalisé au sein de l'Equipe de Statistique et Modélisation Stochastique du laboratoire L.M.C., Institut I.M.A.G.*

*Mes remerciements les plus sincères sont adressés à Messieurs A. LE BRETON, Professeur à l'Université Joseph Fourier Grenoble I, et D.T. PHAM, Directeur de Recherche au C.N.R.S. Je leur suis particulièrement redevable de leur long et efficace soutien lors de l'élaboration de ce travail.*

*J'adresse mes plus vifs remerciements à Messieurs F. BRODEAU, Professeur à l'université des Sciences Sociales de Grenoble, et B. GAREL, Professeur à l'Institut National Polytechnique de Toulouse, pour avoir accepté d'être rapporteurs sur ce travail et pour leur participation au jury.*

*Je remercie vivement Madame G. JOURDAIN, Professeur à l'Institut National Polytechnique de Grenoble, de m'avoir accueilli au Centre d'Etude des Phénomènes Aléatoires et Géophysiques de Grenoble, lors de mon stage de DEA, et d'avoir accepté de prendre part au jury de cette thèse.*

*J'ai eu la chance d'avoir une formidable famille et des amis qui ont su me soutenir dans les moments difficiles. Que tous soient assurés de ma gratitude.*

*Mes remerciement vont enfin à toute l'équipe du L.M.C. et au service de reprographie de l'Institut I.M.A.G.*



## TABLE DES MATIERES

<b>INTRODUCTION</b> .....	<b>p. 01</b>
<b>CHAPITRE I Estimation et détection en traitement du signal</b> .....	<b>p. 05</b>
<b>§1. Le modèle "signal + bruit"</b> .....	<b>p. 06</b>
<b>§2. Estimation paramétrique</b> .....	<b>p. 08</b>
2.1. Estimation des paramètres du bruit .....	<b>p. 08</b>
2.2. Estimation des paramètres de régression .....	<b>p. 09</b>
2.3. Estimation simultanée .....	<b>p. 10</b>
<b>§3. Détection binaire</b> .....	<b>p. 12</b>
3.1. Les règles de décision .....	<b>p. 13</b>
3.2. Le point de vue de Neyman .....	<b>p. 14</b>
3.3. Détection d'un signal de forme connue dans un bruit gaussien .....	<b>p. 15</b>
<b>CHAPITRE II Estimation paramétrique partielle dans un modèle "signal + bruit AR(1)"</b> .....	<b>p. 24</b>
<b>§1. Le modèle "signal + bruit AR(1)"</b> .....	<b>p. 26</b>
<b>§2. Estimation des paramètres d'un bruit AR(1)</b> .....	<b>p. 28</b>
2.1. Estimation de maximum de vraisemblance exact .....	<b>p. 28</b>
2.2. Estimation de moindres carrés .....	<b>p. 30</b>
2.3. Convergence presque sûre des estimateurs .....	<b>p. 31</b>
2.4. Normalité asymptotique des estimateurs .....	<b>p. 38</b>
2.5. Remarques .....	<b>p. 40</b>



§3. Estimation des paramètres du signal .....	p. 42
3.1. Cas où le coefficient autorégressif du bruit est connu .....	p. 42
3.2. Cas où le coefficient autorégressif du bruit est inconnu .....	p. 52

**CHAPITRE III Estimation paramétrique simultanée  
dans un modèle "signal + bruit AR(1)" .....**

§1. Estimation par maximum de vraisemblance .....	p. 62
1.1. Maximisation de la fonction de vraisemblance .....	p. 62
1.2. Existence, unicité et convergence p.s. de l'estimateur de maximum de vraisemblance .....	p. 64
1.3. Vitesse de convergence et normalité asymptotique de l'estimateur de maximum de vraisemblance .....	p. 70
§2. Estimation de moindres carrés .....	p. 81
2.1. Convergence et vitesse de convergence presque sûre de l'estimateur des moindres carrés .....	p. 81
2.2. Normalité asymptotique de l'estimateur des moindres carrés .....	p. 86
§3. Estimation de maximum de vraisemblance approchée .....	p. 95
3.1. Convergence et vitesse de convergence presque sûre de l'estimateur de maximum de vraisemblance approchée ..	p. 95
3.2. Normalité asymptotique de l'estimateur de maximum de vraisemblance approchée .....	p. 97

**CHAPITRE IV Détection dans un bruit AR(1) .....**

§1. Détection d'un signal connu .....	p. 102
1.1. Cas d'un bruit AR(1) de paramètres connus .....	p. 102
1.2. Cas d'un bruit AR(1) de paramètre autorégressif connu ....	p. 103
1.3. Cas d'un bruit AR(1) de variance d'innovation connue .....	p. 103

1.4. Cas d'un bruit AR(1) de paramètres inconnus .....	p. 107
§2. Détection d'un signal contenant des paramètres inconnus .....	p. 110
2.1. Cas d'un bruit AR(1) de paramètres connus .....	p. 110
2.2. Cas d'un bruit AR(1) de paramètre autorégressif connu ....	p. 112
2.3. Cas d'un bruit AR(1) de variance d'innovation connue .....	p. 114
2.4. Cas d'un bruit AR(1) de paramètres inconnus .....	p. 117
<b>ANNEXE</b> Etude Expérimentale.....	p. 119
I. Cadre de l'expérimentation.....	p. 120
I.1 Modèle de base.....	p. 120
I.2 Plan d'expérimentation.....	p. 120
II. Les résultats.....	p. 120
II.1 Estimation.....	p. 120
II.2 Détection.....	p. 121
Tables des résultats de simulation.....	p. 122
<b>BIBLIOGRAPHIE</b> .....	p. 127

## INTRODUCTION

Détecter un signal dans un bruit consiste à reconnaître la présence d'un signal utile dans un enregistrement perturbé. L'existence du bruit parasite rend impossible cette reconnaissance sans risque d'erreur. C'est pourquoi les problèmes liés à la détection d'un signal dans un bruit relèvent des problématiques de la décision statistique : estimation et tests d'hypothèses. Lorsque la loi de probabilité du bruit est bien spécifiée les problèmes de détection restent très classiques ( cf. par exemple HELSTROM (1967), VAN TREES (1968), WONG (1971), ...). Pourtant cette condition est très rarement satisfaite dans la pratique. On a alors recours à des procédés, dits "adaptatifs", qui consistent à utiliser les structures optimales lorsque la loi du bruit est connue, en remplaçant dans ces structures la loi par son estimation (cf. par exemple MERMOZ (1971),.....). Ces méthodes, bien que naturelles et donnant des résultats honorables dans nombre de situations expérimentales restent assez empiriques, leur valeur théorique n'étant pas clairement élucidée.

Le but de notre travail est d'aborder ces questions dans le contexte de la théorie statistique de l'estimation et des tests d'hypothèses. Le cas d'un signal arbitraire et d'un bruit général à temps discret dont les variables successives ont une corrélation inconnue ne permet sans doute pas des développements à visée d'applications. Pour obtenir des résultats relativement exploitables, on peut envisager de modéliser paramétriquement le signal et de représenter la dépendance des variables du bruit par des modèles linéaires (ARMA) de séries chronologiques. On est alors placé dans le cadre des modèles dits modèles linéaires de régression dans les séries chronologiques. De nombreux articles ont été consacrés dans la littérature à l'analyse statistique de ce type de modèles. Du point de vue de l'estimation des paramètres on peut citer notamment les travaux de HILDRETH (1969), KHOLEVO (1969),

DOROGOVCEV(1973), PHAM (1975) et MALLER (1989). En ce qui concerne les problèmes de tests d'hypothèses, outre PHAM (1975) et MALLER (1989) déjà mentionnés, citons par exemple DURBIN et WATSON (1950), SWENSEN (1985) et GAREL (1989). Ici nous nous plaçons délibérément dans le cas le plus simple des processus autorégressifs d'ordre 1 (AR(1)) pour lesquels les résultats sont très explicites. Nous nous efforçons d'étudier aussi complètement que possible les propriétés d'estimateurs quant à leurs vitesses de convergence presque sûre et leurs lois limites ainsi que les puissances asymptotiques de tests. Il convient de noter que nous faisons l'hypothèse restrictive de normalité du bruit. Toutefois bon nombre des résultats asymptotiques que nous obtenons restent valables sans cette hypothèse dès lors que les variables intervenant possèdent des moments d'ordre suffisamment élevé. Il devrait aussi être possible de les étendre à des modèles à erreurs ARMA non forcément gaussiennes. Ce type de questions a déjà attiré l'attention d'auteurs tels que PESARAN (1973), HARVEY et PHILLIPS (1979), MAGEE, ULLAH et SRIVASTAVA (1987).

Dans le premier Chapitre nous fixons d'abord un cadre assez général d'étude des problèmes d'estimation et de détection dans un modèle "signal + bruit". On traite aussi quelques exemples classiques notamment pour la détection d'un signal déterministe connu à un facteur près, dans un bruit de loi gaussienne de structure du second ordre connue au moins à un facteur près.

Dans le Chapitre II nous considérons un modèle "signal + bruit" où le signal est une combinaison linéaire de suites déterministes données et le bruit est un processus AR(1) stationnaire. Il s'agit donc d'un modèle particulier de régression linéaire déterministe avec erreurs corrélées. Nous examinons des problèmes d'estimation

paramétrique partielle (estimation des paramètres du bruit lorsque le signal est complètement connu, estimation des seuls paramètres du signal lorsque le coefficient autorégressif du bruit est connu ou non) du point de vue de la construction des estimateurs et de leurs propriétés, notamment de leur convergence presque sûre.

Au Chapitre III, on considère le même modèle de régression linéaire multiple déterministe dans lequel le processus d'erreurs est AR(1) stationnaire mais cette fois de structure inconnue. On examine le problème de l'estimation simultanée du paramètre de régression et des paramètres de structure du processus d'erreurs. Dans une première partie nous nous intéressons de manière précise à la maximisation de la fonction de vraisemblance et à l'existence d'un unique estimateur de maximum de vraisemblance dont nous examinons les propriétés asymptotiques. Dans une deuxième partie nous étudions une procédure d'estimation fondée sur une première étape de moindres carrés pour l'estimation du paramètre de régression. On propose enfin une troisième procédure qui fournit des estimateurs facilement calculables. Les propriétés asymptotiques de ces diverses méthodes d'estimation sont étudiées en détail.

Dans le Chapitre IV nous nous intéressons à différents problèmes de détection dans le cadre du modèle "signal + bruit AR(1)". On propose des procédures de détection associées aux diverses méthodes d'estimation étudiées dans les chapitres précédents. Les risques liés à ces procédures sont bien analysés soit, dans les cas les plus simples, pour un nombre d'observations fixé, soit du point de vue asymptotique.

Dans une Annexe nous présentons les résultats d'une étude expérimentale par simulation des performances des méthodes d'estimation et de détection envisagées.

## **Chapitre I**

ESTIMATION ET DETECTION  
EN TRAITEMENT DU SIGNAL

Nous fixons ici le cadre d'étude dont relèvent les différentes situations et problématiques envisagées dans notre travail .

Une fois précisé le modèle " signal + bruit " de base, nous décrivons successivement les problèmes d'estimation paramétrique et de détection binaire d'un signal et nous examinons quelques exemples classiques.

## §1. Le modèle "signal +bruit"

Nous supposons qu'un signal déterministe peut être ou ne pas être présent lors des mesures  $y = (y_t ; t = 1, 2, \dots)$  effectuées sur un écouteur bruité. De façon précise, on suppose que

- si l'hypothèse  $\mathcal{H}_0$  : " bruit seul " est vraie, à tout instant  $t$  on observe

$$(1.0) \quad y_t = u_t$$

où  $u = (u_t ; t = 1, 2, \dots)$  est un processus aléatoire centré "bruit de mesure" dont la loi est éventuellement incomplètement connue,

-si l'hypothèse  $\mathcal{H}_1$  : " signal présent " est vraie , à tout instant  $t$  on observe

$$(1.1) \quad y_t = s_t + u_t$$

où  $s = (s_t ; t = 1, 2, \dots)$  est un signal déterministe dépendant éventuellement d'un paramètre inconnu.

Dans la suite nous notons  $(\mathcal{Y}, \mathcal{B})$  (soit  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  soit  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$  selon que le signal est scalaire ou vectoriel ) l'espace mesurable dans lequel  $s_t$  (et donc aussi  $y_t$  et  $u_t$ ) prend ses valeurs. Selon le contexte,  $y_t$  désigne l'élément courant d'une copie de  $\mathcal{Y}$  ou la  $t$ ème application coordonnée sur  $(\mathcal{Y}^{\otimes \mathbb{N}^*}, \mathcal{B}^{\otimes \mathbb{N}^*})$  ou encore  $(\mathcal{Y}^T, \mathcal{B}^{\otimes T})$  pour un  $T$  entier supérieur ou égal à 1. Enfin  $N(m, v)$  est la loi normale de moyenne  $m$  de variance  $v$ ,  $\mathcal{N}$  (resp.  $\mathcal{N}^{-1}$ ) est la fonction de répartition (resp. inverse) d'une loi normale.

Nous supposons que la loi de  $u = (u_t ; t = 1, 2, \dots)$  est connue (dans la suite, gaussienne centrée) à un paramètre  $\lambda$  près dans un ensemble  $\Lambda$  (en pratique soit la loi de  $u$  est complètement connue, et alors  $\Lambda$  est un singleton, soit  $\Lambda$  est un sous-ensemble d'un espace de type  $\mathbb{R}^1$  paramétrant un ensemble de structures du second ordre). Nous

écrivons  $P_\lambda$  (resp.  $P_\lambda^T$ ) pour la loi de  $u$  (resp.  $\underline{u}_T = (u_1, \dots, u_T)$ ) sur  $(\mathcal{Y}^{\otimes N^*}, \mathcal{B}^{\otimes N^*})$  (resp.  $(\mathcal{Y}^T, \mathcal{B}^{\otimes T})$ ) lorsque la vraie valeur du paramètre est  $\lambda$ .

Nous supposons aussi que le signal  $s = (s_t; t = 1, 2, \dots)$  est connu à un paramètre  $\gamma$  près dans un ensemble  $\Gamma$  (en pratique soit le signal  $s$  est complètement connu et alors  $\Gamma$  est un singleton, soit  $\Gamma$  est un sous ensemble d'un espace de type  $\mathbb{R}^k$ , ensemble des valeurs possibles d'un paramètre de régression). Nous notons  $s(\gamma) = (s_t(\gamma); t = 1, 2, \dots)$  le signal lorsque la vraie valeur du paramètre est  $\gamma$ .

Evidemment nous supposons que pour tout  $\gamma \in \Gamma$  on n'a pas  $s_t(\gamma) \equiv 0, t = 1, 2, \dots$ . Alors adjoignant à  $\Gamma$  un élément  $\gamma_0$  et posant  $s_t(\gamma_0) \equiv 0, t = 1, 2, \dots$ , l'état de la nature définissant les conditions dans lesquelles les mesures  $y_t, t = 1, 2, \dots$  sont effectuées est complètement spécifié par un couple  $(\gamma, \lambda)$  dans  $\Gamma_0 \times \Lambda$  où  $\Gamma_0 = \Gamma \cup \{\gamma_0\}$ . Pour  $(\gamma, \lambda) \in \Gamma_0 \times \Lambda$  nous désignons par  $H_{\gamma, \lambda}$  l'hypothèse : "les observations  $y_t$  sont de la forme (1.1) avec  $s_t = s_t(\gamma)$  et  $\lambda$  est la vraie valeur du paramètre définissant la loi du bruit  $u$ ".

Bien sûr, compte tenu de la convention faite ci-dessus, sous  $H_{\gamma_0, \lambda}$  "les observations  $y_t$  sont de la forme (1.0) et  $\lambda$  est la vraie valeur du paramètre définissant la loi du bruit  $u$ ". En définitive on a

$$\mathcal{H}_0 = \bigvee_{\lambda \in \Lambda} H_{\gamma_0, \lambda} \quad \text{et} \quad \mathcal{H}_1 = \bigvee_{(\gamma, \lambda) \in \Gamma \times \Lambda} H_{\gamma, \lambda}.$$

Sous  $H_{\gamma, \lambda}$ , la loi  $P_{\gamma, \lambda}$  (resp.  $P_{\gamma, \lambda}^T$ ) de  $y = (y_t; t = 1, 2, \dots)$  (resp.  $\underline{y}_T = (y_1, \dots, y_T)$ ) est la loi image de  $P_\lambda$  (resp.  $P_\lambda^T$ ) par l'application  $u \rightarrow (s_t(\gamma) + u_t; t = 1, 2, \dots)$  (resp.  $\underline{u}_T \rightarrow (s_t(\gamma) + u_t; t = \overline{1, T})$ ). En particulier sous  $H_{\gamma_0, \lambda}$  on a  $P_{\gamma_0, \lambda} = P_\lambda$ .

En d'autres termes la structure statistique associée à l'observation  $\underline{y}_T = (y_1, \dots, y_T)$  est de la forme

$$(1.3) \quad (\mathcal{Y}^T, \mathcal{B}^{\otimes T}, \{P_{\gamma, \lambda}^T; \gamma \in \Gamma_0, \lambda \in \Lambda\}).$$

## §2. Estimation paramétrique

Nous distinguons les trois situations classiques en traitement du signal.

### 2.1. Estimation des paramètres du bruit

On suppose ici que l'observateur est assuré de ce que le signal observé  $y$  est une "référence bruit" i.e. que l'hypothèse  $\mathcal{H}_0$  est vraie.

La structure statistique associée à l'observation  $\underline{y}_T = (y_1, \dots, y_T) = (u_1, \dots, u_T)$  est

$$(\underline{y}^T, \mathbf{B}^{\otimes T}, \{P_\lambda^T, \lambda \in \Lambda\})$$

et on s'intéresse au problème d'estimation du paramètre  $\lambda$ .

Notons que bien sûr lorsque l'observation est faite sous l'hypothèse  $H_{\gamma, \lambda}$  pour une valeur  $\gamma$  connue on se ramène au problème précédent par centrage des observations.

#### Exemple 2.1. Estimation de la variance d'un bruit blanc

Supposons que  $u$  est une suite de variables aléatoires réelles indépendantes et identiquement distribuées (i.i.d.) selon une loi  $Q_{\sigma^2}$  centrée admettant une variance  $\sigma^2 > 0$ . Nous dirons alors que  $u$  est un bruit blanc (scalaire). Si  $Q_1$  est donnée et  $Q_{\sigma^2}$  est la loi image de  $Q_1$  par l'application  $z \rightarrow \sigma z$  de  $\mathbb{R}$  dans  $\mathbb{R}$ , la loi de l'observation  $\underline{y}_T$  est connue au paramètre  $\sigma^2$  près dans  $\mathbb{R}_+^*$  au sens que la structure statistique associée est

$$(\underline{y}^T, \mathbf{B}^{\otimes T}, \{Q_{\sigma^2}^{\otimes T}, \sigma^2 > 0\}).$$

Alors une estimation naturelle de  $\sigma^2$  fournie par la méthode des moments (mais aussi par la méthode du maximum de vraisemblance si on suppose par exemple  $Q_1$  gaussienne centrée réduite i.e.  $Q = N(0, 1)$ ) est donnée par

$$(2.1) \quad \hat{\sigma}_T^2 = \frac{1}{T} \sum_{i=1}^T y_i^2.$$

On sait que (2.1) définit un estimateur sans biais (i.e. pour tout  $\sigma^2 > 0$ ,  $E_{\sigma^2}(\hat{\sigma}_T^2) = \sigma^2$  où  $E_{\sigma^2}$  désigne l'espérance mathématique relativement à  $P_{\sigma^2}$ ) et, en vertu de la loi forte des grands nombres pour les variables aléatoires i.i.d. presque sûrement

convergent ( i.e. pour tout  $\sigma^2 > 0$   $P_{\sigma^2} \left\{ \lim_{T \rightarrow +\infty} \hat{\sigma}_T^2 = \sigma^2 \right\} = 1$ ). On sait aussi que si  $Q_1 = N(0, 1)$ , relativement à  $P_{\sigma^2}$ , la loi de  $\frac{T}{\sigma^2} \hat{\sigma}_T^2$  est une loi  $\chi_T^2$  du chi deux à T degrés de liberté.

## 2.2. Estimation des paramètres de régression

On suppose ici que l'observateur est assuré de ce que l'hypothèse  $\mathcal{H}_1$  est vraie et que de plus la loi  $P_*$  du bruit est complètement spécifiée (i.e.  $\Lambda$  est un singleton, éventuellement parce que le paramètre  $\lambda$  a préalablement été bien estimé à partir d'une référence bruit (cf. §.2.1)). La structure statistique associée à l'observation est

$$(\mathbf{y}^T, \mathbf{R}^{\otimes T}, \{P_\gamma, \gamma \in \Gamma\})$$

où  $P_\gamma$  est la loi image de  $P_*$  par l'application  $u \rightarrow (s_t(\gamma) + u_t; t = 1, 2, \dots)$ . On s'intéresse au problème d'estimation du paramètre  $\gamma$ .

### Exemple 2.2. Estimation de l'amplitude d'un signal bruité par un bruit blanc de loi connue

Supposons que  $u$  est un bruit blanc de loi connue. Supposons que  $\Gamma = \mathbb{R}/\{0\}$  et pour  $\gamma \in \Gamma$

$$(2.2) \quad s_t(\gamma) = \gamma s_t; t = 1, 2, \dots$$

où  $s = (s_t; t = 1, 2, \dots)$  est un signal donné tel que  $\sum_{t=1}^T s_t^2 > 0$ .

Alors une estimation naturelle de l'amplitude  $\gamma$  du signal observé (fournie par la méthode des moments mais aussi par la méthode du maximum de vraisemblance si on suppose par exemple que le bruit blanc est gaussien) est donnée par

$$(2.3) \quad \hat{\gamma}_T = \frac{\sum_{t=1}^T y_t s_t}{\sum_{t=1}^T s_t^2}.$$

On sait que ( 2.3 ) définit un estimateur sans biais ( i.e. pour tout  $\gamma \in \Gamma$ ,  $E_\gamma (\hat{\gamma}_T) = \gamma$  ) et, en vertu de la loi des grands nombres pour les martingales (cf. DACUNHA-CASTELLE

et DUFLO (1983) Chap.II ), presque sûrement convergent si  $\lim_{T \rightarrow +\infty} \sum_{t=1}^T s_t^2 = +\infty$ . En

effet d'après (1.1), (2.2) et (2.3) on a

$$\begin{aligned} \hat{\gamma}_T &= \frac{\sum_{t=1}^T (s_t(\gamma) + u_t) s_t}{\sum_{t=1}^T s_t^2} \\ &= \gamma + \frac{\sum_{t=1}^T s_t u_t}{\sum_{t=1}^T s_t^2} \end{aligned}$$

où  $(\sum_{t=1}^T s_t u_t; t \geq 1)$  est une martingale de carré intégrable de processus croissant

associé  $\left\{ \sum_{t=1}^T \sigma^2 s_t^2; t \geq 1 \right\}$ .

On sait aussi que si la loi  $u_t$  est  $N(0, \sigma^2)$  alors la loi de  $\hat{\gamma}_T$  est  $N(\gamma, \sigma^2 (\sum_{t=1}^T s_t^2)^{-1})$ .

### 2.3. Estimation simultanée

On suppose ici que l'observation est faite dans des conditions telles que  $\gamma$  et  $\lambda$  sont inconnus dans  $\Gamma$  et  $\Lambda$  respectivement.

On s'intéresse alors au problème d'estimation du paramètre  $(\gamma, \lambda)$  dans la structure statistique

$$(\mathbf{y}^T, \mathbf{B}^{\otimes T}, \{P_{\gamma, \lambda}^T; \gamma \in \Gamma, \lambda \in \Lambda\})$$

**Exemple 2.3. Estimation simultanée de l'amplitude d'un signal et de la variance d'un bruit blanc**

On suppose ici que  $u$  est un bruit blanc de loi connue au paramètre de variance  $\sigma^2 > 0$  près (cf. Exemple 2.1) et que de plus le signal est de la forme (2.2) (cf. Exemple 2.2),  $\gamma$  étant inconnu. Alors une estimation naturelle de  $(\gamma, \sigma^2)$  (fournie par la méthode des moments mais aussi par la méthode du maximum de vraisemblance si on suppose par exemple que le bruit est gaussien) est donnée par  $(\hat{\gamma}_T, \tilde{\sigma}_T^2)$  où  $\hat{\gamma}_T$  est défini comme en (2.3) et

$$(2.4) \quad \tilde{\sigma}_T^2 = \frac{1}{T} \sum_{t=1}^T (y_t - \hat{\gamma}_T s_t)^2.$$

Bien sûr  $\hat{\gamma}_T$  est un estimateur sans biais et convergent presque sûrement si

$$\sum_{t=1}^{+\infty} s_t^2 = +\infty. \text{ Pour ce qui est de } \tilde{\sigma}_T^2 \text{ on a}$$

$$E_{\gamma, \sigma^2}(\tilde{\sigma}_T^2) = \frac{T-1}{T} \sigma^2$$

et donc  $\frac{T}{T-1} \tilde{\sigma}_T^2$  est un estimateur sans biais de  $\sigma^2$ . De plus, d'après (1.1), (2.3) et

(2.4), on peut écrire

$$\tilde{\sigma}_T^2 = \frac{1}{T} \sum_{t=1}^T u_t^2 - \frac{(\sum_{t=1}^T s_t u_t)^2}{T \sum_{t=1}^T s_t^2}$$

où, d'après la loi forte des grands nombres pour les variables aléatoires indépendantes,

$$\lim_{T \rightarrow +\infty} \frac{1}{T} \sum_{t=1}^T u_t^2 = \sigma^2 \text{ p.s..}$$

Ainsi  $\tilde{\sigma}_T^2$  est un estimateur de  $\sigma^2$  presque sûrement convergent si

$$(2.5) \quad \lim_{T \rightarrow +\infty} \frac{\sum_{t=1}^T s_t u_t}{\left( T \sum_{t=1}^T s_t^2 \right)^{1/2}} = 0 \text{ p.s.}$$

Or, comme  $\left\{ \sum_{t=1}^T s_t u_t ; t \geq 1 \right\}$  est une martingale de carré intégrable de processus

croissant associé  $\left\{ \sigma^2 \sum_{t=1}^T s_t^2 ; t \geq 1 \right\}$ , utilisant NEVEU (1972), Proposition VII-2-4,

on peut voir que si  $\sum_{t=1}^{+\infty} s_t^2 = +\infty$  et s'il existe une fonction croissante  $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  telle que

$$\int_0^{+\infty} \frac{du}{(1+f(u))^2} < +\infty ; f\left(\sum_{t=1}^{+\infty} s_t^2\right) = O\left(\left(T \sum_{t=1}^T s_t^2\right)^{1/2}\right)$$

alors (2.5) est satisfaite. C'est par exemple le cas si  $\sum_{t=1}^{+\infty} s_t^2 = +\infty$  et  $\sum_{t=1}^{+\infty} s_t^2 = O(T^\alpha)$

pour un  $\alpha > 0$ , ou si  $\sum_{t=1}^{+\infty} s_t^2 = +\infty$  et  $\text{Log}\left(\sum_{t=1}^T s_t^2\right) = O(T^\beta)$ , pour un  $\beta < 1$ .

### §3. Détection binaire

Le problème de détection, du moins dans la version simplifiée envisagée ici, consiste à décider au vu de l'observation  $\underline{y}_T = (y_1, \dots, y_T)$  si oui ou non le signal est présent (sous-entendu pour  $t = 1, \dots, T$ ).

En d'autres termes le problème de détection est celui de tester l'hypothèse  $\mathcal{H}_0$  contre l'hypothèse  $\mathcal{H}_1$  au vu de  $\underline{y}_T$  dans la structure statistique (1.3).

### 3.1. Les règles de décision

Une règle de décision -un récepteur- répondant au problème de détection est un test stochastique  $\Phi : \mathcal{Y}^T \rightarrow [0, 1]$  qui à toute observation  $\mathbf{y}_T \in \mathcal{Y}^T$  associe la probabilité  $\Phi(\mathbf{y}_T)$  avec laquelle doit être prise, au vu de  $\mathbf{y}_T$ , la décision de rejeter l'hypothèse  $\mathcal{H}_0$ : "bruit seul" (" $\gamma = \gamma_0$ ") i.e. accepter  $\mathcal{H}_1$ : "signal présent" (" $\gamma \neq \gamma_0$ "). Il sera parfois commode de considérer un test  $\Phi$  comme la probabilité de transition sur  $\mathcal{Y}^T \times \mathcal{P}(\{0, 1\})$ , définie par

$$\Phi(\mathbf{y}_T, \cdot) = \Phi(\mathbf{y}_T) \delta_1(\cdot) + (1 - \Phi(\mathbf{y}_T)) \delta_0(\cdot).$$

Si  $\Phi$  est un test déterministe i.e. à valeurs dans  $\{0, 1\}$ , auquel cas  $\Phi = \mathbb{1}[\Phi = 1]$ ,  $\Phi$  est alors identifié à  $\Phi(\mathbf{y}_T) = \delta_{\Phi(\mathbf{y}_T)}(\cdot)$ .

On appelle "fausse alarme" le fait de décider qu'un signal est présent, alors que seul le bruit est présent i.e., dans le langage de la théorie des tests, une erreur de première espèce liée au rejet de l'hypothèse  $\mathcal{H}_0$  alors qu'elle est vraie. De même on appelle "non détection" le fait de décider qu'un signal est absent alors qu'en réalité il y a un signal i.e. une erreur de seconde espèce liée à l'acceptation de l'hypothèse  $\mathcal{H}_0$  alors qu'elle est fausse.

Alors les "risques de fausse alarme" attachés à  $\Phi$  sont évalués par  $E_{\gamma_0, \lambda} \Phi$ ,  $\lambda \in \Lambda$ ; de même les "risque de non détection" sont donnés par  $E_{\gamma, \lambda}(1 - \Phi)$ ,  $\gamma \in \Gamma$ ,  $\lambda \in \Lambda$ . Ici  $E_{\gamma, \lambda}$  désigne l'espérance relative à la loi  $P_{\gamma, \lambda}$ . Dans le langage de la théorie des tests l'application  $(\gamma, \lambda) \rightarrow \beta_\Phi(\gamma, \lambda) = E_{\gamma, \lambda}(\Phi)$  est appelée fonction puissance du test  $\Phi$ . Notons que bien sûr si  $\Phi$  est un test déterministe i.e. à valeurs dans  $\{0, 1\}$ , auquel cas  $\Phi = \mathbb{1}[\Phi = 1]$ , on a

$$E_{\gamma_0, \lambda}(\Phi) = P_{\gamma_0, \lambda}(\{\Phi = 1\}) \quad ; \quad E_{\gamma, \lambda}(\Phi) = P_{\gamma, \lambda}(\{\Phi = 0\}), \quad \gamma \in \Gamma, \lambda \in \Lambda$$

et les risques sont des probabilités de fausse alarme ou de non détection, le domaine  $R_\Phi = \{\Phi = 1\}$  étant appelé région critique du test  $\Phi$ .

### 3.2. Le point de vue de Neyman

Plaçons-nous dans le cas où une fausse alarme a des conséquences plus fâcheuses qu'une non-détection (dans le cas contraire il convient dans la discussion qui suit d'échanger les rôles de  $\mathcal{H}_0$  et  $\mathcal{H}_1$ ). Selon le point de vue de Neyman on est alors amené à envisager de choisir le test  $\Phi$  dans une classe donnée  $\Psi$  de tests, de manière telle que :

- d'une part, pour un seuil  $\alpha$  fixé ( $\alpha \in ]0, 1[$ ), on ait

$$(3.1) \quad \alpha_\Phi = \sup_{\lambda \in \Lambda} E_{\gamma_0, \lambda}(\Phi) \leq \alpha$$

i.e. un risque maximum de fausse alarme pour  $\Phi$  (appelé niveau de signification) inférieur à  $\alpha$ ,

- d'autre part pour tout autre test  $\Phi' \in \Psi$  vérifiant (3.1) on ait

$$(3.2) \quad E_{\gamma, \lambda}(\Phi) \geq E_{\gamma, \lambda}(\Phi'), \gamma \in \Gamma, \lambda \in \Lambda$$

i.e. des risques de non détection plus forts pour  $\Phi'$  que pour  $\Phi$ .

Si  $\Phi$  est un test vérifiant (3.1) et (3.2) pour tout autre test  $\Phi' \in \Psi$  - i.e. si  $\Phi$  est optimal au sens précédent dans la classe  $\Psi = \mathcal{T}$  de tous les tests - on dit que  $\Phi$  est uniformément le plus puissant (U.P.P.) au seuil  $\alpha$ .

Un test  $\Phi$  est dit sans biais si pour tout  $\gamma \in \Gamma, \lambda \in \Lambda$  on a

$$E_{\gamma, \lambda}(\Phi) \geq \alpha_\Phi.$$

Si  $\Phi$  est un test sans biais optimal dans la classe  $\Psi = \mathcal{S}$  de tous les tests sans biais, on dit que  $\Phi$  est uniformément le plus puissant parmi les tests sans biais (U.P.P.S.B.) au seuil  $\alpha$ .

Un test est dit libre si pour tout  $(\lambda, \lambda') \in \Lambda^2$  on a

$$E_{\gamma_0, \lambda}(\Phi) = E_{\gamma_0, \lambda'}(\Phi), \text{ auquel cas } E_{\gamma_0, \lambda}(\Phi) = \alpha_\Phi, \lambda \in \Lambda.$$

Si  $\Phi$  est un test libre optimal dans la classe  $\Psi = \mathcal{L}$  de tous les tests libres, on dit que  $\Phi$  est uniformément le plus puissant parmi les tests libres (U.P.P.L.) au seuil  $\alpha$ .

### 3.3. Détection d'un signal de forme connue dans un bruit gaussien

Dans ce paragraphe le bruit est supposé gaussien i.e. pour tout  $T \geq 1$ ,  $\underline{u}_T = (u_1, \dots, u_T)$  est un vecteur aléatoire gaussien.

#### 3.3.1. Cas d'un signal connu

Nous considérons d'abord la situation de base où le problème de détection concerne un signal  $s = (s_t, t = 1, 2, \dots)$  complètement connu. On peut donc prendre  $\Gamma = \{1\}$ ,  $\gamma_0 = 0$  et pour  $\gamma \in \Gamma_0 = \{0, 1\}$ ,  $s_t(\gamma) = \gamma s_t$ ,  $t = 1, 2, \dots$

##### a. Cas d'un bruit de structure de covariance connue

Commençons par examiner le cas où  $u$  est un bruit blanc de variance  $v$  connue. Le problème de détection envisagé est celui du test de l'hypothèse simple  $\mathcal{H}_0 = H_{0,v}$  : " $\gamma=0$ " contre l'hypothèse simple  $\mathcal{H}_1 = H_{1,v}$  : " $\gamma=1$ " dans la structure statistique

$$(\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T), \{N_T(\gamma \underline{s}_T, v \mathbb{1}_T); \gamma \in \{0, 1\}\}),$$

où  $\underline{s}_T = (s_1, \dots, s_T)^*$  et  $\mathbb{1}_T$  est la matrice identité  $T \times T$ .

On souhaite définir un récepteur  $\Phi_\alpha$  tel que le risque de fausse alarme ait une valeur  $\alpha$  donnée dans  $]0, 1[$  et le risque de non détection soit le plus faible possible i.e. un test U.P.P. au seuil  $\alpha$ .

Par application du lemme de Neyman-Pearson (cf. e.g. BARRA (1970), Chap V) on obtient que le test déterministe  $\Phi$  de région critique

$$R = \{y_T : \underline{s}_T^* y_T > c\}$$

est U.P.P. à son niveau de signification  $\alpha_\Phi$ . Comme sous  $\mathcal{H}_0$  la variable aléatoire

$\mathbf{s}_T^* \mathbf{y}_T$  est de loi gaussienne centrée de variance  $v \sum_{t=1}^T s_t^2$ , on a

$$\alpha_\Phi = P_{0, v}(R) = 1 - \mathcal{N}(c / (v \sum_{t=1}^T s_t^2)^{1/2})$$

où  $\mathcal{N}$  désigne la fonction de répartition de la loi  $N(0, 1)$ .

De même, puisque sous  $\mathcal{H}_1$  la variable aléatoire  $\mathbf{s}_T^* \mathbf{y}_T$  est de loi  $N(\sum_{t=1}^T s_t^2, v \sum_{t=1}^T s_t^2)$ ,

on a

$$\beta_{\Phi_\alpha} = P_{1, v}(R) = 1 - \mathcal{N}\left(\frac{(c - \sum_{t=1}^T s_t^2)}{(v \sum_{t=1}^T s_t^2)^{1/2}}\right).$$

Notant  $\mathcal{N}^{-1}$  la fonction réciproque de  $\mathcal{N}$  et choisissant, dans ce qui précède,  $c = c_\alpha$  où

$$c_\alpha = (v \sum_{t=1}^T s_t^2)^{1/2} \mathcal{N}^{-1}(1 - \alpha),$$

on définit un récepteur  $\Phi_\alpha$  optimal (U.P.P.) au niveau  $\alpha$  de probabilité de fausse alarme, par la région critique

$$R_{\Phi_\alpha} = \left\{ \mathbf{y}_T : \frac{\sum_{t=1}^T s_t y_t}{(\sum_{t=1}^T s_t^2)^{1/2}} > v^{1/2} \mathcal{N}^{-1}(1 - \alpha) \right\}.$$

La probabilité de non détection correspondante vaut

$$1 - \beta_{\Phi_\alpha} = \mathcal{N}(\mathcal{N}^{-1}(1 - \alpha) - (\sum_{t=1}^T s_t^2 / \nu)^{1/2}).$$

Notons que le récepteur  $\Phi_\alpha$  ne dépendant que du quotient  $(\sum_{t=1}^T s_t^2 / \nu)$  appelé " rapport signal sur bruit ".

**Exemple 3.1.** Dans le cas particulier où  $s_t = s_1 \neq 0$ ,  $t \geq 2$ , le problème qui vient d'être étudié se réduit au problème classique de "test de l'hypothèse "m = 0" contre l'hypothèse "m = s<sub>1</sub>" pour une variable aléatoire de loi  $N(m, \nu)$  lorsque la

variance  $\nu$  est connue, au vu d'un échantillon". Alors, si  $\bar{y}_T = \frac{1}{T} \sum_1^T y_t$  on est conduit à

$$R_{\Phi_\alpha} = \left\{ \underline{y}_T : (\text{sgn } s_1) \bar{y}_T > (\nu / T)^{1/2} \mathcal{N}^{-1}(1 - \alpha) \right\}$$

et

$$\beta_{\Phi_\alpha} = 1 - \mathcal{N}(\mathcal{N}^{-1}(1 - \alpha) - (T / \nu)^{1/2} |s_1|).$$

Considérons maintenant le cas où  $\underline{y}_T$  est un vecteur gaussien centré de matrice de covariance  $\Lambda_T$  inversible connue. Alors le problème de détection est celui de l'hypothèse simple  $\mathcal{H}_0 = H_{0, \Lambda_T} : " \gamma = 0 "$  contre l'hypothèse simple  $\mathcal{H}_1 = H_{1, \Lambda_T} : " \gamma = 1 "$  dans la structure statistique  $(\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T), \{ N_T(\gamma \underline{s}_T, \Lambda_T); \gamma \in \{0, 1\} \})$ . Si  $\Lambda_T^{1/2}$  est une matrice telle que  $\Lambda_T = \Lambda_T^{1/2} (\Lambda_T^{1/2})^*$ , la statistique  $\Lambda_T^{-1/2} \underline{y}_T$  est exhaustive et induit la structure statistique image  $(\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T), \{ N_T(\gamma \Lambda_T^{-1/2} \underline{s}_T, \mathbf{1}_T); \gamma \in \{0, 1\} \})$ . On est alors ramené au problème étudié précédemment en substituant  $\Lambda_T^{-1/2} \underline{s}_T$  à  $\underline{s}_T$ ,  $\Lambda_T^{-1/2} \underline{y}_T$  à  $\underline{y}_T$  et faisant  $\nu = 1$ . Cela conduit à des tests U.P.P. à leur niveau de signification définis par des régions critiques

$$R = \left\{ \underline{y}_T : \underline{s}_T^* \Lambda_T^{-1} \underline{y}_T > c \right\}.$$

Appliquant les résultats ci-dessus, on définit un récepteur  $\Phi_\alpha$ , optimal au niveau  $\alpha$  de probabilité de fausse alarme, par la région critique

$$R_{\Phi_\alpha} = \left\{ \mathbf{y}_T : \frac{\mathbf{s}_T^* \Lambda_T^{-1} \mathbf{y}_T}{(\mathbf{s}_T^* \Lambda_T^{-1} \mathbf{s}_T)^{\frac{1}{2}}} > \mathcal{N}^{-1}(1-\alpha) \right\}.$$

La probabilité de non détection correspondante vaut

$$1 - \beta_{\Phi_\alpha} = \mathcal{N}(\mathcal{N}^{-1}(1-\alpha) - (\mathbf{s}_T^* \Lambda_T^{-1} \mathbf{s}_T)^{\frac{1}{2}}).$$

Nous appliquerons cela au cas d'un bruit autorégressif AR(1) de paramètres connus au Chapitre IV§.1.1.

### b. Cas d'un bruit de structure de covariance connue à un facteur près

Examinons d'abord le cas où  $u$  est un bruit blanc de variance inconnue. Le problème de détection est celui du test de l'hypothèse multiple  $\mathcal{H}_0 = \bigvee_{v > 0} H_{0,v} : " \gamma =$

$0 "$  contre l'hypothèse multiple  $\mathcal{H}_1 = \bigvee_{v > 0} H_{1,v} : " \gamma = 1 "$  dans la structure statistique

$$(\mathbb{R}^T, \mathfrak{B}(\mathbb{R}^T), \left\{ N_T(\gamma \mathbf{s}_T, v \mathbf{1}_T); \gamma \in \{0, 1\}, v > 0 \right\}).$$

Ecrivant la fonction de vraisemblance il est facile de voir que la statistique  $(v_T, w_T)$  où

$$v_T = \frac{\sum_{t=1}^T y_t s_t}{\sum_{t=1}^T s_t^2}; \quad w_T = \frac{\sum_{t=1}^T y_t^2}{\sum_{t=1}^T s_t^2} - v_T^2,$$

est exhaustive. De plus, posant  $\tilde{y}_t = y_t - \gamma s_t$ , on obtient

$$v_T = \gamma + \tilde{v}_t; w_T = \frac{\sum_{t=1}^T \tilde{y}_t^2}{\sum_{t=1}^T s_t^2} - \tilde{v}_t^2,$$

où

$$\tilde{v}_t = \frac{\sum_{t=1}^T s_t \tilde{y}_t}{\sum_{t=1}^T s_t^2}.$$

Considérons maintenant une matrice orthogonale C de première ligne

$(s_t / (\sum_{u=1}^T s_u^2)^{1/2}; t = 1, T)$  et posons  $\hat{y}_T = C \tilde{y}_T$ . Alors on a

$$\tilde{v}_t = (\sum_{t=1}^T s_t^2)^{-1} \hat{y}_1$$

et

$$\begin{aligned} w_T &= (\sum_{t=1}^T s_t^2)^{-1} \tilde{y}_T^* \tilde{y}_T - \tilde{v}_t^2 \\ &= (\sum_{t=1}^T s_t^2)^{-1} [\hat{y}_T^* \hat{y}_T - \hat{y}_1^2] \end{aligned}$$

soit

$$w_T = (\sum_{t=1}^T s_t^2)^{-1} \sum_{t=2}^T \hat{y}_t^2.$$

Bien sûr, sous  $P_{\gamma, v}$ ,  $\hat{y}_T$  est de loi  $N_T (Q_T, v \mathbb{1}_T)$  et par suite les statistiques  $v_T$  et  $w_T$

sont indépendantes de lois respectives  $N(\gamma, v (\sum_{t=1}^T s_t^2)^{-1})$  et  $\chi(\frac{T-1}{2}, 2v (\sum_{t=1}^T s_t^2)^{-1})$ .

Ainsi, au changement de paramètre  $\sigma^2 = v (\sum_{t=1}^T s_t^2)^{-1}$  près, la structure statistique induite par la statistique  $(v_T, w_T)$  est la même que celle induite par la statistique

$$(\bar{y}_T, S_T^2) = \left( \frac{1}{T} \sum_1^T y_t, \frac{1}{T} \sum_1^T (y_t - \bar{y}_T)^2 \right)$$

à partir de la structure statistique  $(\mathbb{R}^T, \mathfrak{B}(\mathbb{R}^T), \{N_T(\gamma, \nu); \gamma \in \{0, 1\}, \nu > 0\})$ .

Soit  $\Phi$  le test de région critique

$$R = \left\{ \mathcal{Y}_T : \sqrt{T-1} \frac{\nu_T}{\sqrt{w_T}} > c \right\}.$$

Comme sous  $\mathcal{H}_0$  la variable aléatoire  $\sqrt{T-1} \frac{\nu_T}{\sqrt{w_T}}$  suit la loi de Student à T-1

degrés de liberté, si  $\mathcal{F}_{T-1}$  est la fonction de répartition de cette loi, on a

$$\alpha_\Phi = P_{0, \nu}(R) = 1 - \mathcal{F}_{T-1}(c),$$

et  $\Phi$  est un test libre.

Procédant comme dans BARRA (1970), Chapitre V, on montre alors que le test  $\Phi$  est U.P.P.L. pour le problème posé.

Notant  $\mathcal{F}_{T-1}^{-1}$  la fonction réciproque de  $\mathcal{F}_{T-1}$  et choisissant, dans ce qui précède,  $c = c_\alpha$  où

$$c_\alpha = \mathcal{F}_{T-1}^{-1}(1 - \alpha),$$

on définit un récepteur  $\Phi_\alpha$  optimal (U.P.P.L.) au niveau  $\alpha$  de probabilité de fausse alarme, par la région critique

$$R_{\Phi_\alpha} = \left\{ \mathcal{Y}_T : \frac{\sqrt{T-1} \sum_{t=1}^T s_t y_t}{\left[ \left( \sum_{t=1}^T s_t^2 \right) \left( \sum_{t=1}^T y_t^2 \right) - \left( \sum_{t=1}^T s_t y_t \right)^2 \right]^{1/2}} > \mathcal{F}_{T-1}^{-1}(1 - \alpha) \right\}.$$

Les probabilités de non détection correspondantes valent

$$1 - \beta_{\Phi_\alpha}(1, \nu) = \text{Prob} \left[ \frac{\mathcal{F}_{T-1}^{-1}(1 - \alpha) \sqrt{W}}{\sqrt{T-1}} - V \geq \left( \sum_{t=1}^T s_t^2 / \nu \right)^{1/2} \right]$$

où V et W sont des variables aléatoires indépendantes de loi respectives  $N(0, 1)$  et du Chi deux à T-1 degrés de liberté.

**Exemple 3.2.** Dans le cas particulier où  $s_t = s_1 \neq 0$ ,  $t \geq 2$ , le test défini ci-dessus se réduit au test classique de Student de l'hypothèse " $m = 0$ " contre l'hypothèse " $m = s_1$ " pour la moyenne d'une variable aléatoire de loi  $N(m, v)$  lorsque la variance  $v$  est inconnue, au vu d'un échantillon. Alors

$$R_{\Phi_\alpha} = \left\{ \underline{y}_T : (\text{sgn } s_1) \frac{\sqrt{T-1} \bar{y}_T}{S_T} > \mathfrak{t}_{T-1}^{-1}(1-\alpha) \right\}$$

$$1 - \beta_{\Phi_\alpha}(1, v) = \text{Prob} \left[ \frac{\mathfrak{t}_{T-1}^{-1}(1-\alpha) \sqrt{W}}{\sqrt{T-1}} - V \geq (T/v)^{1/2} |s_1| \right].$$

Considérons maintenant le cas où  $\underline{y}_T$  est un vecteur gaussien centré de matrice de covariance  $v\Lambda_T$  où  $\Lambda_T$  est une matrice inversible donnée et  $v > 0$  est un paramètre inconnu. Remplaçant dans ce qui précède  $\underline{s}_T$  et  $\underline{y}_T$  par  $\Lambda_T^{-1/2} \underline{s}_T$  et  $\Lambda_T^{-1/2} \underline{y}_T$  respectivement, on définit un récepteur  $\Phi_\alpha$ , optimal (U.P.P.L.) au niveau  $\alpha$  de probabilité de fausse alarme, par la région critique

$$R_{\Phi_\alpha} = \left\{ \underline{y}_T : \frac{\sqrt{T-1} \underline{s}_T^* \Lambda_T^{-1} \underline{y}_T}{\left[ (\underline{s}_T^* \Lambda_T^{-1} \underline{s}_T)(\underline{y}_T^* \Lambda_T^{-1} \underline{y}_T) - (\underline{s}_T^* \Lambda_T^{-1} \underline{y}_T)^2 \right]^{1/2}} > \mathfrak{t}_{T-1}^{-1}(1-\alpha) \right\}.$$

Les probabilités de non détection correspondantes valent

$$1 - \beta_{\Phi_\alpha}(1, v) = \text{Prob} \left[ \frac{\mathfrak{t}_{T-1}^{-1}(1-\alpha) \sqrt{W}}{\sqrt{T-1}} - V \geq (\underline{s}_T^* \Lambda_T^{-1} \underline{s}_T / v)^{1/2} \right].$$

Nous appliquerons cela au cas d'un bruit autorégressif AR(1) de paramètre autorégressif connu et de variance inconnue au chapitre IV §1.2.

### 3.3.2. Cas d'un signal d'amplitude inconnue

Nous considérons maintenant la situation où le problème de détection concerne un signal de forme connue, disons définie par un signal de base  $s = (s_t, t = 1, 2, \dots)$ , mais d'amplitude inconnue au sens que pour  $\gamma \in \Gamma = \mathbb{R} \setminus \{0\}$  on a

$$s_t(\gamma) = \gamma s_t, \quad t = 1, 2, \dots$$

Ici bien sûr  $\Gamma_0 = \mathbb{R}$  et  $\gamma_0 = 0$ .

Les problèmes à envisager concernent désormais le test de l'hypothèse  $\mathcal{H}_0: " \gamma = 0 "$  contre l'hypothèse  $\mathcal{H}_1: " \gamma \neq 0 "$ .

Nous indiquons rapidement les tests optimaux qui ne sont autres que des versions bilatères des tests décrits au §3.3.1 dont on garde les notations. Pour les preuves on peut utiliser les considérations du §3.3.1 et raisonner de manière analogue à BARRA(1970), ChV.

### a. Cas d'un bruit de structure de covariance connue

Le récepteur  $\Phi_\alpha$  défini par la région critique

$$R_{\Phi_\alpha} = \left\{ \mathbf{y}_T : \frac{\left| \begin{matrix} \mathbf{s}_T^* \Lambda_T^{-1} \mathbf{y}_T \end{matrix} \right|}{\left( \mathbf{s}_T^* \Lambda_T^{-1} \mathbf{s}_T \right)^{1/2}} > \mathcal{N}^{-1} \left( 1 - \frac{\alpha}{2} \right) \right\}$$

est U.P.P.S.B. au niveau  $\alpha$  de probabilité de fausse alarme.

Les probabilités de non détection correspondantes valent

$$1 - \beta_{\Phi}(\gamma) = \mathcal{N} \left( \mathcal{N}^{-1} \left( 1 - \frac{\alpha}{2} \right) - \gamma \left( \mathbf{s}_T^* \Lambda_T^{-1} \mathbf{s}_T \right)^{1/2} \right)$$

$$- \mathcal{N} \left( - \mathcal{N}^{-1} \left( 1 - \frac{\alpha}{2} \right) - \gamma \left( \mathbf{s}_T^* \Lambda_T^{-1} \mathbf{s}_T \right)^{1/2} \right).$$

### b. Cas d'un bruit de structure de covariance connue à un facteur près

Le récepteur  $\Phi_\alpha$  défini par la région critique

$$R_{\Phi_\alpha} = \left\{ \mathbf{y}_T : \frac{\sqrt{T-1} \left| \begin{matrix} \mathbf{s}_T^* \Lambda_T^{-1} \mathbf{y}_T \end{matrix} \right|}{\left[ \left( \mathbf{s}_T^* \Lambda_T^{-1} \mathbf{s}_T \right) \left( \mathbf{y}_T^* \Lambda_T^{-1} \mathbf{y}_T \right) - \left( \mathbf{s}_T^* \Lambda_T^{-1} \mathbf{y}_T \right)^2 \right]^{1/2}} > \mathcal{t}_{T-1}^{-1} \left( 1 - \alpha/2 \right) \right\}$$

est U.P.P.S.B. libre au niveau  $\alpha$  de probabilité de fausse alarme.

Les probabilités de non détection correspondantes valent

$$1 - \beta_{\Phi_\alpha}(\gamma, \nu) = \text{Prob} \left[ \left| V + \left( \frac{\mathbf{s}_T^* \Lambda_T^{-1} \mathbf{s}_T}{\nu} \right)^{1/2} \gamma \right| \leq \frac{\mathbf{t}_{T-1}^{-1} (1 - \alpha/2) \sqrt{W}}{\sqrt{T-1}} \right],$$

où V et W sont telles qu'au §3.3.1.b. ci-dessus.

## **Chapitre II**

ESTIMATION PARAMETRIQUE PARTIELLE  
DANS UN MODELE "SIGNAL + BRUIT AR(1)"



Dans ce chapitre nous considérons un modèle "signal + bruit" où le signal est une combinaison linéaire de suites déterministes données et le bruit est un processus autorégressif du premier ordre stationnaire gaussien AR(1). Il s'agit donc d'un modèle particulier de régression linéaire déterministe avec erreurs corrélées. Nous examinons les problèmes d'estimation paramétrique partielle suivants:

-estimation des paramètres du bruit lorsque le signal est complètement connu,

-estimation des seuls paramètres du signal lorsque les paramètres du bruit sont connus ou non.

Le premier problème se ramène bien sûr à celui de l'estimation des paramètres d'un processus AR(1). Dans une note HASZA (1980) fournit des formules explicites pour les estimateurs du maximum de vraisemblance exact de ce modèle. Nous étudions de manière précise les propriétés asymptotiques de ces estimateurs ainsi que celles des estimateurs de moindres carrés, notamment du point de vue des vitesses de convergence presque sûre.

Le deuxième problème est un problème particulier d'estimation en régression linéaire déterministe avec erreur corrélées. LAI, ROBBINS et WEI (1979) ont étudié la convergence presque sûre de l'estimateur des moindres carrés dans un modèle sous des hypothèses très générales pour le processus d'erreurs (voir aussi LE BRETON et MUSIELA (1987a), (1989)). Ici nous exploitons l'hypothèse AR(1) pour la construction d'estimateurs dont nous étudions complètement le comportement asymptotique.

### §1. Le modèle "signal + bruit AR(1)"

On observe un processus réel  $y = (y_t, t = 1, 2, \dots)$  régi par le modèle

$$(1.1) \quad y_t = \gamma^* z_t + u_t ; t \geq 1,$$

où

.  $z_t = (z_{t1}, \dots, z_{tk})^*$ ;  $t \geq 1$  ) est une suite de valeurs connues d'un régresseur déterministe à valeurs dans  $\mathbb{R}^k$ ,

.  $\gamma = (\gamma_1, \dots, \gamma_k)^*$  est un paramètre, éventuellement inconnu, dans  $\mathbb{R}^k$ ,

.  $(u_t; t \geq 1)$  est un processus "bruit" AR(1) i.e. autorégressif d'ordre 1 stationnaire réel.

On suppose que

$$(1.2) \quad u_{t+1} = \rho u_t + v_{t+1}; t \geq 1,$$

où

.  $\rho$  est un paramètre réel,  $|\rho| < 1$ , éventuellement inconnu,

. les variables aléatoires  $u_1, v_2, \dots, v_t, \dots$  sont indépendantes gaussiennes centrées et

$$E u_1^2 = v / (1 - \rho^2) ; E v_t^2 = v, t \geq 2,$$

le paramètre  $v$  pouvant être inconnu.

Dans la suite de ce travail nous utiliserons souvent les caractéristiques suivantes d'un processus AR(1) tel qu'il vient d'être défini en (1.2).

Notons  $\underline{u}_T = (u_1, \dots, u_T)^*$ ,  $\underline{v}_T = (v_1, \dots, v_T)^*$  et  $\tilde{\underline{v}}_T = (\tilde{v}_1, \dots, \tilde{v}_T)^*$  où

$$(1.3) \quad \tilde{\underline{v}}_T = v^{-1/2} \underline{v}_T = v^{-1/2} M_T(\rho) \underline{u}_T$$

avec

$$(1.4) \quad M_T(\rho) = \begin{bmatrix} (1-\rho^2)^{1/2} & 0 & \dots & 0 \\ -\rho & 1 & \ddots & \vdots \\ 0 \cdot \cdot & \ddots & \ddots & 0 \\ 0 \cdot \cdot & 0 & -\rho & 1 \end{bmatrix}.$$

Alors, d'après les hypothèses,  $\tilde{\mathbf{y}}_T$  est un vecteur aléatoire gaussien centrée de matrice de covariance  $\mathbb{1}_T$ . La matrice de covariance de  $\mathbf{u}_T$  est la matrice de Toeplitz  $\nu A_T(\rho)$  où  $A_T(\rho) = (1-\rho^2)^{-1} \rho^{|t-s|}$ ;  $(s, t) \in \{1, \dots, T\}^2$ ) et, compte tenu de (1.3), on a

$$A_T(\rho) = B_T^{-1}(\rho),$$

où

$$(1.5) \quad B_T(\rho) = M_T^*(\rho) M_T(\rho) = \begin{bmatrix} 1 & -\rho & 0 & \dots & 0 \\ -\rho & 1+\rho^2 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 1+\rho^2 & -\rho \\ 0 & \dots & 0 & -\rho & 1 \end{bmatrix}.$$

Notons que d'après (1.4) et (1.5) on a

$$\det B_T(\rho) = [\det M_T(\rho)]^2 = 1 - \rho^2.$$

En définitive le vecteur aléatoire  $\mathbf{u}_T$  est gaussien centré de matrice de covariance  $\nu B_T^{-1}(\rho)$  et donc de densité

$$(1.6) \quad f_T(\nu, \rho; \mathbf{u}_T) = \frac{(1 - \rho^2)^{1/2}}{(2\pi\nu)^{T/2}} \exp\left\{ -\frac{1}{2\nu} \mathbf{u}_T^* B_T(\rho) \mathbf{u}_T \right\}$$

où  $B_T(\rho)$  est donnée par (1.5). Notons que nous utilisons les mêmes symboles pour les variables aléatoires et les arguments des densités associées.

Plus loin, lorsque  $T$  est fixé, nous supprimons l'indice  $T$  dans  $\mathbf{u}_T$ ,  $M_T(\rho)$ , .... Lorsqu'il s'agit d'études asymptotiques pour  $T$  tendant vers l'infini nous ne faisons pas toujours apparaître cet indice; les convergences ( presque sûre (p.s.), en loi,...) sont sous-entendues relativement à la loi de probabilité sur l'espace de base définie par les vraies valeurs des paramètres toujours notées  $\rho$ ,  $\nu$ ,  $\gamma$ . On note  $r$ ,  $v$  et  $g$  des valeurs arbitraires de ces paramètres.

## §2. Estimation des paramètres d'un bruit AR(1)

Nous supposons ici que le paramètre de régression  $\gamma$  est connu et nous nous intéressons à l'estimation des paramètres  $\rho$  et  $v$  du processus "bruit". Bien sûr on se ramène au problème de l'estimation des paramètres  $\rho$  et  $v$  au vu de l'observation du processus  $u_t = y_t - \gamma^* z_t$ ,  $t \geq 1$  qui est une "référence bruit" AR(1).

### 2.1. Estimation de maximum de vraisemblance exact

En vertu de (1.6) la fonction de Log-vraisemblance associée à l'observation  $\underline{u} = \underline{u}_T$  s'écrit

$$(2.1) \quad \text{Log} f_T(v, r; \underline{u}) = -\frac{T}{2} \text{Log } 2\pi - \frac{T}{2} \text{Log } v + \frac{1}{2} \text{Log}(1 - r^2) - \frac{1}{2v} \underline{u}^* B(r) \underline{u},$$

où  $B(r) = B_T(r)$ . Compte tenu de (1.5) et (2.1), pour estimer le paramètre  $(\rho, v)$  par maximum de vraisemblance, il convient donc de maximiser, par rapport à  $(r, v) \in ]-1, +1[ \times ]0, +\infty[$ , pour  $\underline{u} \in \mathbb{R}^T$  fixé, la fonction

$$(2.2) \quad L(v, r; \underline{u}) = -\frac{T}{2} \text{Log } v + \frac{1}{2} \text{Log}(1 - r^2) - \frac{1}{2v} \left[ (1 - r^2) u_1^2 + \sum_{t=2}^T (u_t - r u_{t-1})^2 \right].$$

Pour tout  $r \in ]-1, +1[$  fixé, la fonction  $v \rightarrow L(v, r; \underline{u})$  est maximum pour

$$v = v_T(r) = \frac{1}{T} \underline{u}^* B(r) \underline{u} \text{ soit}$$

$$(2.3) \quad v_T(r) = \frac{1}{T} \left[ (1 - r^2) u_1^2 + \sum_{t=2}^T (u_t - r u_{t-1})^2 \right],$$

la valeur du maximum étant, à une constante près,

$$(2.4) \quad \hat{L}(r, \underline{u}) = -\frac{T}{2} \text{Log } v_T(r) + \frac{1}{2} \text{Log}(1 - r^2).$$

Il convient donc de maximiser  $\hat{L}$  par rapport à  $r$  dans  $]-1, +1[$ . D'après (2.4) si  $\hat{L}$  a un extremum en  $r \in ]-1, +1[$  on a nécessairement

$$(2.5) \quad r v_T(r) + \frac{T}{2} (1 - r^2) \dot{v}_T(r) = 0,$$

où si  $\varphi$  est une fonction (deux fois) dérivable,  $\dot{\varphi}$  ( $\ddot{\varphi}$ ) désigne sa dérivée (seconde).

En vertu de (2.3) et (2.5) on est amené à s'intéresser aux solutions de l'équation cubique  $g(r) = 0$

où

$$(2.6) \quad g(r) = a_3 r^3 + a_2 r^2 + a_1 r + a_0$$

avec

$$a_0 = \sum_{t=2}^T u_t u_{t-1} \quad ; \quad a_1 = - \sum_{t=1}^{T-1} u_t^2 - \frac{1}{T} \sum_{t=1}^T u_t^2,$$

(2.7)

$$a_2 = - \frac{T-2}{T} \sum_{t=2}^T u_t u_{t-1} \quad ; \quad a_3 = \frac{T-1}{T} \sum_{t=2}^T u_t^2.$$

On peut alors montrer (cf. aussi HASZA (1980)).

**Lemme 1.** Pour  $T \geq 3$ , avec probabilité 1, la fonction  $r \rightarrow \hat{L}(r; \mathbf{u})$  définie par (2.4) sur  $] -1, +1[$  admet un maximum unique en  $\hat{\rho}_T \in ] -1, +1[$  donné par

$$(2.8) \quad \hat{\rho}_T = \frac{2}{3} a_3^{-1} (a_2^2 - 3 a_1 a_3)^{1/2} \text{Cos } \theta - \frac{1}{3} a_2 a_3^{-1}$$

où

$$(2.9) \quad \theta = \frac{1}{3} \text{ArcCos} \left\{ - \frac{1}{2} (2 a_2^3 - 9 a_1 a_2 a_3 + 27 a_0 a_3^2) (a_2^2 - 3 a_1 a_3)^{-3/2} \right\} + \frac{4\pi}{3},$$

$a_0, a_1, a_2$ , et  $a_3$  étant définis par (2.7).

**Preuve.** Remarquons qu'avec probabilité 1 on a, pour  $g$  définie par (2.6)-(2.7):

$$g(+1) = - \frac{1}{T} \sum_{t=2}^T (u_t - u_{t-1})^2 < 0,$$

$$g(-1) = + \frac{1}{T} \sum_{t=2}^T (u_t + u_{t-1})^2 > 0,$$

$$\lim_{r \rightarrow \infty} g(r) = +\infty ; \quad \lim_{r \rightarrow -\infty} g(r) = -\infty.$$

Par suite  $g$  a exactement une racine dans chacun des intervalles  $]-\infty, -1[$ ,  $]-1, +1[$  et  $] +1, +\infty[$  et  $\hat{L}$  admet un extremum unique dans  $]-1, +1[$ . Or on peut voir que le signe de la dérivée de  $\hat{L}$  est le signe de  $g$  ; donc  $\hat{L}$  est croissante au voisinage de  $-1$  et décroissante au voisinage de  $+1$  et finalement  $\hat{L}$  admet un maximum unique sur  $]-1, +1[$  en la racine  $\hat{\rho}$  de  $g(r) = 0$  appartenant à cet intervalle. On peut vérifier que  $\hat{\rho}$  peut s'écrire sous la forme (2.8)-(2.9). 🍏

En définitive on peut énoncer

**Proposition 1.** Pour  $T \geq 3$ , avec probabilité 1, la fonction de Log-vraisemblance définie par (2.2) sur  $]-1, +1[ \times ]0, +\infty[$  admet un maximum unique en  $(\hat{\rho}_T, \hat{v}_T) \in ]-1, +1[ \times ]0, +\infty[$  où  $\hat{\rho}_T$  est donnée par (2.7)-(2.8)-(2.9) et  $\hat{v}_T = \hat{v}_T(\hat{\rho}_T)$  où  $v_T(\cdot)$  est donnée par (2.3).

Remarquons que dans LE BRETON (1989) on trouve un analogue de cette proposition dans le cas d'un processus AR(1) à valeurs complexes.

## 2.2. Estimation de moindres carrés

La méthode des moindres carrés pour l'estimation de  $(\rho, v)$  consiste à estimer

d'abord  $\rho$  par minimisation (sur  $\mathbb{R}$ ) de la quantité  $\frac{1}{T-1} \sum_{t=2}^T (u_t - r u_{t-1})^2$  puis à estimer  $v$

par la valeur du minimum de cette quantité. On peut voir que de manière équivalente cela revient à maximiser la fonction de Log-vraisemblance approchée

$$(2.10) \quad \tilde{L}(v, r; \underline{u}) = - \frac{T-1}{2} \log v - \frac{1}{2v} \sum_{t=2}^T (u_t - r u_{t-1})^2,$$

obtenue à partir de (2.2) en négligeant les termes "liés aux conditions initiales".

L'estimateur des moindres carrés ( $\tilde{\rho}_T, \tilde{v}_T$ ) de  $(\rho, v)$  est alors donné par

$$(2.11) \quad \tilde{\rho}_T = \left\{ \sum_{t=2}^T u_t u_{t-1} \right\} / \left\{ \sum_{t=1}^{T-1} u_t^2 \right\},$$

$$\tilde{v}_T = \tilde{v}_T(\tilde{\rho}_T),$$

où

$$(2.12) \quad \tilde{v}_T(r) = \frac{1}{T-1} \sum_{t=2}^T (u_t - r u_{t-1})^2.$$

Dans ce qui suit nous précisons notamment l'ordre de l'approximation de l'estimateur de maximum de vraisemblance par l'estimateur des moindres carrés.

### 2.3. Convergence presque sûre des estimateurs

Etudions d'abord le comportement de  $v_T(\cdot)$  (cf.(2.3)) et de ses dérivées.

On note  $\dot{\varphi}, \ddot{\varphi}, \dddot{\varphi}$  les dérivées d'ordre successif de  $\varphi$ .

**Lemme 2.** Soit  $v_T(\cdot)$  défini par (2.3) i.e.  $v_T(r) = \frac{1}{T} \underline{u}_T^* B_T(r) \underline{u}_T$  où  $B_T(r)$  est donné par (1.5). Alors pour tout  $r \in ]-1, +1[$  on a presque sûrement

$$(2.13) \quad \lim_{T \rightarrow +\infty} v_T(r) = v \{ 1 + (1 - \rho^2)^{-1} (r - \rho)^2 \},$$

$$(2.14) \quad \lim_{T \rightarrow +\infty} \dot{v}_T(r) = 2 v (1 - \rho^2)^{-1} (r - \rho),$$

$$(2.15) \quad \lim_{T \rightarrow +\infty} \ddot{v}_T(r) = 2 v (1 - \rho^2)^{-1}.$$

**Preuve.** D'après (1.3)-(1.5) et (2.3) on a

$$v_T(\rho) = \frac{1}{T} \underline{u}_T^* M_T^*(\rho) M_T(\rho) \underline{u}_T$$

$$= \frac{v}{T} \tilde{Y}_T^* \tilde{Y}_T = \frac{v}{T} \sum_1^T \tilde{v}_t^2,$$

où les variables aléatoires  $\tilde{v}_t^2$ ,  $t \geq 1$  sont indépendantes de même loi du Chi deux à 1 degré de liberté. Alors, d'après la loi forte des grands nombres, puisque  $E \tilde{v}_t^2 = 1$ , on a

$$(2.16) \quad \lim_{T \rightarrow +\infty} v_T(\rho) = v \text{ p.s..}$$

D'après (1.2) et (2.3) on a

$$(2.17) \quad \begin{aligned} \dot{v}_T(\rho) &= \frac{1}{T} [-2\rho u_1^2 - 2 \sum_{t=2}^T u_{t-1}(u_t - \rho u_{t-1})] \\ &= -\frac{2}{T} [\rho u_1^2 + M_T], \end{aligned}$$

où

$$(2.18) \quad M_T = \sum_{t=2}^T u_{t-1} v_t; \quad T \geq 2.$$

Notons que  $(M_T, T \geq 2)$  est une martingale de processus croissant  $(\langle M \rangle_T, T \geq 2)$  où

$$(2.19) \quad \langle M \rangle_T = v \sum_2^T u_{t-1}^2; \quad T \geq 2.$$

Comme le processus  $(u_t, t \geq 1)$  est AR(1) gaussien centré stationnaire, il est ergodique et donc

$$(2.20) \quad \lim_{T \rightarrow +\infty} \frac{v}{T} \sum_2^T u_{t-1}^2 = \lim_{T \rightarrow +\infty} \frac{1}{T} \langle M \rangle_T = v^2/(1 - \rho^2) \text{ p.s..}$$

Appliquant la loi forte pour les martingales (cf. e.g. DACUNHA-CASTELLE et DUFLO

(1983), Ch.II) on voit que  $T^{-1}M_T = \frac{\langle M \rangle_T}{T} \frac{M_T}{\langle M \rangle_T}$  tend presque sûrement vers zéro. Par

suite on a

$$(2.21) \quad \lim_{T \rightarrow +\infty} \dot{v}_T(\rho) = 0 \text{ p.s..}$$

D'après (1.2) et (2.3) on a

$$(2.22) \quad \ddot{v}_T(\rho) = \frac{2}{T} \sum_2^T u_{t,1}^2.$$

Alors, appliquant (2.20) on obtient que

$$(2.23) \quad \lim_{T \rightarrow +\infty} \ddot{v}_T(\rho) = 2v/(1-\rho^2) \text{ p.s..}$$

Utilisant le fait que  $v_T(\cdot)$  est quadratique, pour toute valeur  $r \in ]-1, +1[$  on peut écrire

$$\dot{v}_T(r) = \dot{v}_T(\rho) + \ddot{v}_T(\rho) (r - \rho),$$

(2.24)

$$v_T(r) = v_T(\rho) + \dot{v}_T(\rho) (r - \rho) + 1/2 \ddot{v}_T(\rho) (r - \rho)^2.$$

A partir de (2.16), (2.21), (2.23) et (2.24) on obtient immédiatement les assertions du lemme.  $\blacktriangleleft$

Nous pouvons maintenant démontrer :

**Proposition 2.** Les estimateurs  $(\hat{\rho}_T, \hat{v}_T)$  et  $(\tilde{\rho}_T, \tilde{v}_T)$ , définis par (2.7)-(2.8)-(2.9) et (2.11)-(2.12) respectivement, sont presque sûrement convergents. Plus précisément, pour tout  $\beta < 1/2$  on a presque sûrement

$$(2.25) \quad T^\beta (\hat{\rho}_T - \rho) = o(1) \quad ; \quad T^\beta (\hat{v}_T - v) = o(1),$$

$$T^\beta (\tilde{\rho}_T - \rho) = o(1) \quad ; \quad T^\beta (\tilde{v}_T - v) = o(1).$$

En outre on a presque sûrement

$$(2.26) \quad T (\hat{\rho}_T - \tilde{\rho}_T) = O(1) \quad ; \quad T (\hat{v}_T - \tilde{v}_T) = O(1).$$

**Preuve.** Rappelons que  $\hat{\rho}_T$  est la solution, appartenant à  $] -1, +1[$ , de l'équation (2.5) ou, de manière équivalente, de l'équation

$$h_T(r) = \frac{1}{T} r v_T(r) + 1/2 (1 - r^2) \dot{v}_T(r) = 0.$$

Notons qu'en fait  $h_T(r) = -g(r)/T$  où  $g$  est définie par (2.6)-(2.7). L'étude faite dans la preuve du lemme 1 montre que pour  $-1 < r_* < r^* < 1$  on a l'équivalence

$$(2.27) \quad r_* < \hat{\rho}_T < r^* \Leftrightarrow h_T(r_*) < 0 < h_T(r^*).$$

Comme  $h_T(\cdot)$  est cubique, pour toute valeur  $r$  on peut écrire

$$h_T(r) = h_T(\rho) + \dot{h}_T(\rho)(r - \rho) + 1/2 \ddot{h}_T(\rho)(r - \rho)^2 + 1/(3!) \dddot{h}_T(\rho)(r - \rho)^3,$$

où  $\rho$  est la vraie valeur du paramètre et

$$\dot{h}_T(\rho) = \frac{1}{T} v_T(\rho) - \frac{T-1}{T} \rho \dot{v}_T(\rho) + 1/2 (1 - \rho^2) \ddot{v}_T(\rho),$$

$$\ddot{h}_T(\rho) = -\left[ \frac{T-2}{T} \dot{v}_T(\rho) + \frac{2T-1}{T} \rho \ddot{v}_T(\rho) \right] ; \quad \dddot{h}_T(\rho) = -\frac{3(T-1)}{T} \ddot{v}_T(\rho).$$

D'après (2.13)-(2.15) on a presque sûrement

$$\lim_{T \rightarrow +\infty} h_T(\rho) = 0 ; \quad \lim_{T \rightarrow +\infty} \dot{h}_T(\rho) = v,$$

$$\lim_{T \rightarrow +\infty} \ddot{h}_T(\rho) = -4\rho v(1 - \rho^2)^{-1} ; \quad \lim_{T \rightarrow +\infty} \dddot{h}_T(\rho) = -6v(1 - \rho^2)^{-1}.$$

Ainsi on a p.s.

$$\lim_{T \rightarrow +\infty} h_T(r) = v (r - \rho) \frac{1 - r^2}{1 - \rho^2}.$$

Cela montre que pour  $-1 < r_* < r^* < 1$  on a l'équivalence

$$(2.28) \quad r_* < \rho < r^* \Leftrightarrow \lim_{T \rightarrow +\infty} h_T(r_*) < 0 < \lim_{T \rightarrow +\infty} h_T(r^*).$$

Soit alors  $\varepsilon > 0$ . D'après (2.28) il existe un  $T_0$  tel que pour  $T \geq T_0$  on ait

$$h_T(\rho - \varepsilon) < 0 < h_T(\rho + \varepsilon).$$

Par suite, utilisant (2.27), on voit que  $\rho - \varepsilon < \hat{\rho}_T < \rho + \varepsilon$ . Cela entraîne que  $\hat{\rho}_T$  converge presque sûrement vers  $\rho$ .

Comme

$$\hat{v}_T = v_T(\hat{\rho}) = v_T(\rho) + v_T(\rho) (\hat{\rho} - \rho) + 1/2 \ddot{v}_T(\rho) (\hat{\rho} - \rho)^2,$$

on déduit immédiatement de la convergence de  $\hat{\rho}_T$  et de (2.13)-(2.15) que  $\hat{v}_T$  converge presque sûrement vers  $v$ .

D'après (2.11) et (1.2) on peut écrire

$$(2.29) \quad \tilde{\rho}_T - \rho = v \frac{M_T}{\langle M \rangle_T},$$

où  $M$  est la martingale définie par (2.18) de processus croissant  $\langle M \rangle$  donné par (2.19). Alors, compte tenu de (2.20), la loi forte des grands nombres pour les martingales assure que  $\tilde{\rho}_T$  est un estimateur convergent presque sûrement.

Ecrivaint

$$\tilde{v}_T = v_T(\tilde{\rho}_T) = v_T(\rho) + \tilde{v}_T(\rho) (\tilde{\rho}_T - \rho) + 1/2 \ddot{v}_T(\rho) (\tilde{\rho}_T - \rho)^2,$$

où, d'après (2.12) et (1.2)

$$\tilde{v}_T(\rho) = \frac{1}{T-1} \sum_{i=1}^T v_i^2 \quad ; \quad \tilde{v}_T(\rho) = -\frac{2}{T-1} M_T \quad ; \quad \ddot{v}_T(\rho) = \frac{2}{(T-1)v} \langle M \rangle_T,$$

il vient

$$(2.30) \quad \tilde{v}_T - v = \frac{1}{T-1} N_T - v \frac{M_T^2}{(T-1)\langle M \rangle_T},$$

où  $(N_T, T \geq 2)$  est la martingale définie par

$$(2.31) \quad N_T = \sum_2^T (v_t^2 - v).$$

Le processus croissant  $\langle N \rangle$  de la martingale  $N$  est donné par

$$(2.32) \quad \langle N \rangle_T = 2v^2(T-1); T \geq 2.$$

Utilisant (2.20) et (2.32) et appliquant la loi forte des grands nombres aux martingales  $M$  et  $N$  dans (2.30), on voit que  $\tilde{v}_T$  est un estimateur convergent presque sûrement. En fait,

plus précisément on a, pour  $\alpha > 1/2$ ,  $\frac{M_T}{\langle M \rangle_T^\alpha} = o(1)$  et  $\frac{N_T}{\langle N \rangle_T^\alpha} = o(1)$  presque sûrement. On

déduit alors facilement de (2.29) et (2.30) que pour tout  $\beta < 1/2$  on a

$$(2.33) \quad T^\beta(\tilde{\rho}_T - \rho) = o(1) \text{ p.s.}; \quad T^\beta(\tilde{v}_T - v) = o(1) \text{ p.s..}$$

Examinons maintenant l'écart entre les estimateurs de maximum de vraisemblance et ceux de moindres carrés.

Exploitant le fait que  $\tilde{v}_T(\cdot)$  est quadratique, pour toute valeur  $r$  on peut écrire

$$\tilde{v}_T(r) = \tilde{v}_T(\tilde{\rho}_T) + \tilde{v}_T'(\tilde{\rho}_T)(r - \tilde{\rho}_T).$$

Alors, comme  $\tilde{v}_T(\tilde{\rho}_T) = 0$  et  $\tilde{v}_T'(\tilde{\rho}_T) = \frac{2\langle M \rangle_T}{v(T-1)}$ , il vient

$$\tilde{v}_T(r) = \frac{2\langle M \rangle_T}{v(T-1)}(r - \tilde{\rho}_T).$$

En particulier pour  $r = \hat{\rho}_T$  on obtient

$$(2.34) \quad \hat{\rho}_T - \tilde{\rho}_T = v \frac{1}{2} \left[ \frac{(M)_T}{T-1} \right]^{-1} \tilde{v}_T(\hat{\rho}_T).$$

Evalons maintenant  $\tilde{v}_T(\hat{\rho}_T)$ . Comme  $\hat{\rho}_T$  est solution de (2.5) on a

$$(2.35) \quad \dot{v}_T(\hat{\rho}_T) = - \frac{2\hat{\rho}_T v_T(\hat{\rho}_T)}{T(1-\hat{\rho}_T^2)}.$$

D'après (2.3) et (2.12) on a aussi

$$(2.36) \quad \tilde{v}_T(\rho) = \frac{T}{T-1} v_T(\rho) - \frac{1-\rho^2}{T-1} u_1^2,$$

d'où

$$(2.37) \quad \tilde{v}_T(\rho) = \frac{T}{T-1} \dot{v}_T(\rho) + \frac{2\rho}{T-1} u_1^2$$

Utilisant (2.35) et (2.37) on obtient

$$(2.38) \quad \tilde{v}_T(\hat{\rho}_T) = \frac{2}{T-1} \hat{\rho}_T \left[ u_1^2 - \frac{v_T(\hat{\rho}_T)}{1-\hat{\rho}_T^2} \right].$$

Alors, puisque d'une part  $\hat{\rho}_T$  converge presque sûrement vers  $\rho$  (et donc aussi  $v_T(\hat{\rho}_T)$  converge presque sûrement vers  $v$ ) et d'autre part, d'après (2.20),  $\frac{(M)_T}{T-1}$  tend presque

sûrement vers  $\frac{v}{(1-\rho^2)}$ , (2.34) et (2.38) montrent que

$$(2.39) \quad T(\hat{\rho}_T - \tilde{\rho}_T) = O(1) \text{ presque sûrement.}$$

Finalement, utilisant (2.36), on peut écrire

$$\begin{aligned} \tilde{v}_T - v_T &= [\tilde{v}_T(\hat{\rho}_T) - v_T(\hat{\rho}_T)] + [\tilde{v}_T(\hat{\rho}_T) - v_T(\hat{\rho}_T)] \\ &= \tilde{v}_T(\hat{\rho}_T) (\hat{\rho}_T - \tilde{\rho}_T) + 1/2 \tilde{v}_T(\hat{\rho}_T) (\hat{\rho}_T - \tilde{\rho}_T)^2 + \frac{1}{T} [v_T(\hat{\rho}_T) + (1-\hat{\rho}_T^2) u_1^2] \end{aligned}$$

ce qui, compte tenu des propriétés déjà évoquées, montre que

$$(2.40) \quad T(\tilde{\nu}_T - \hat{\nu}_T) = O(1) \text{ presque sûrement.}$$

En définitive (2.33), (2.39) et (2.40) assurent bien que (2.25) et (2.26) sont vérifiées.  $\blacksquare$

Notons que différents auteurs (cf. e.g. YAMAMOTO et KUNITOMO (1984), LE BRETON et PHAM (1989)) se sont intéressés au biais et biais asymptotique des estimateurs.

## 2.4. Normalité asymptotique des estimateurs

Nous pouvons démontrer que les estimateurs sont asymptotiquement gaussiens.

**Proposition 3.** *Les suites  $(T^{1/2}(\hat{\rho}_T, \hat{\nu}_T); T \geq 1)$  et  $(T^{1/2}(\tilde{\rho}_T, \tilde{\nu}_T); T \geq 1)$ , définies par (2.7)-(2.8)-(2.9) et (2.11)-(2.12) respectivement, convergent en loi vers une loi*

*gaussienne centrée de matrice de covariance* 
$$\begin{pmatrix} 1-\rho^2 & 0 \\ 0 & 2\nu^2 \end{pmatrix}.$$

**Preuve.** En vertu de (2.26) il suffit de démontrer le résultat pour l'estimateur  $(\tilde{\rho}_T, \tilde{\nu}_T)$ . Les martingales  $M$  et  $N$  et leur processus croissant  $\langle M \rangle$  et  $\langle N \rangle$  étant définis tels qu'en (2.18)-(2.19) et (2.31)-(2.32), d'après (2.29) et (2.30) nous pouvons écrire

$$(T-1)^{1/2} \begin{pmatrix} \tilde{\rho}_T - \rho \\ \tilde{\nu}_T - \nu \end{pmatrix} = (T-1)^{-1/2} \begin{pmatrix} \frac{\nu(T-1)}{\langle M \rangle_T} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} M_T \\ N_T \end{pmatrix} - \begin{pmatrix} 0 \\ \left( \frac{\langle M \rangle_T}{T-1} \right)^{1/2} \left( \frac{\nu^{1/2} M_T}{\langle M \rangle_T^{3/4}} \right)^2 \end{pmatrix}.$$

D'après (2.20) et le fait que  $\frac{M_T}{\langle M \rangle_T^{3/4}} = o(1)$  p.s., il est clair que le résultat annoncé sera

acquis si on montre que la suite  $(T^{-1/2} \begin{pmatrix} M_T \\ N_T \end{pmatrix}; T \geq 1)$  converge en loi vers une loi

gaussienne centrée de matrice de covariance  $\begin{pmatrix} \frac{v^2}{1-\rho^2} & 0 \\ 0 & 2v^2 \end{pmatrix}$ .

Soit  $Z$  la martingale bidimensionnelle  $\begin{pmatrix} M \\ N \end{pmatrix}$ . Notons que, comme  $E(v_t(v_t^2-v)) = 0$ , la variation quadratique croisée  $\langle M, N \rangle$  est nulle i.e. les martingales  $M$  et  $N$  sont orthogonales et le processus croissant  $\langle Z \rangle$  de  $Z$  est la matrice diagonale  $\begin{pmatrix} \langle M \rangle & 0 \\ 0 & \langle N \rangle \end{pmatrix}$ .

D'après (2.20) et (2.32) il est clair que

$$(2.41) \quad \lim_{T \rightarrow +\infty} (T-1)^{-1} \langle Z \rangle_T = \begin{pmatrix} \frac{v^2}{1-\rho^2} & 0 \\ 0 & 2v^2 \end{pmatrix} = \Gamma(\rho, v) \text{ p.s..}$$

Un calcul facile montre qu'on a aussi

$$E \langle Z \rangle_T = \begin{pmatrix} \frac{(T-1)v^2}{1-\rho^2} & 0 \\ 0 & 2(T-1)v^2 \end{pmatrix},$$

et donc

$$(2.42) \quad \lim_{T \rightarrow +\infty} (T-1)^{-1} E \langle Z \rangle_T = \Gamma(\rho, v).$$

Soit  $l = (a, b)^*$  un vecteur quelconque non nul de  $\mathbb{R}^2$  et considérons la martingale  $y = l^*Z$  de processus croissant  $\langle Y \rangle = l^* \langle Z \rangle l$ .

D'après (2.41)-(2.42) on a presque sûrement

$$(2.43) \quad \lim_{T \rightarrow +\infty} (T-1)^{-1} \langle Y \rangle_T = \lim_{T \rightarrow +\infty} (T-1)^{-1} E \langle Y \rangle_T = l^* \Gamma(\rho, v) l.$$

De plus, compte tenu de ce que  $E u_t^4 = C t^\epsilon < \infty$  et  $E v_t^4 = C t^\epsilon < \infty$  et  $E |v_t^2 - v|^4 = C t^\epsilon < \infty$

pour  $\Delta Y_t = Y_t - Y_{t-1}$ , l'espérance  $E|\Delta Y_t|^4$  est bornée pour  $t \geq 1$ . Alors la condition de Lindeberg

$$\lim_{T \rightarrow +\infty} \frac{1}{\langle Y \rangle_T} \sum_1^T E \left[ \Delta Y_t^2 \mathbb{1} \left[ |\Delta Y_t| \geq \langle Y \rangle_T \varepsilon \right] \right] = 0,$$

est satisfaite pour tout  $\varepsilon > 0$ .

Appliquant le théorème de limite centrale pour les martingales (cf.e.g. DACUNHA-CASTELLE et DUFLO (1983), Ch.II) on obtient alors que la suite  $(\langle Y \rangle_T^{-1/2} Y_T; T \geq 1)$  converge en loi vers une loi gaussienne centrée réduite. En conséquence, compte tenu de (2.43), la suite  $(T^{-1/2} Y_T; T \geq 1)$  converge en loi vers une loi gaussienne centrée de variance  $1 * \Gamma(\rho, \nu)$  où  $\Gamma(\rho, \nu)$  est définie dans (2.41). Ce qui précède étant vrai pour tout  $l$ , la suite  $(T^{-1/2} \begin{pmatrix} M_T \\ N_T \end{pmatrix}; T \geq 1)$  converge en loi vers une loi gaussienne centrée de covariance  $\Gamma(\rho, \nu)$  d'où le résultat.  $\clubsuit$

## 2.5. Remarques

Dans le cas où la variance  $\nu$  de l'innovation est connue, l'estimateur de maximum de vraisemblance  $\hat{\rho}_T^\nu$  du paramètre  $\rho$  est fourni par maximisation de

$$\text{Log}(1 - r^2) - \frac{1}{2\nu} [(1 - r^2) u_1^2 + \sum_{t=2}^T (u_t - r u_{t-1})^2].$$

On est alors amené à s'intéresser aux solutions de l'équation cubique  $k(r) = 0$  où

$$k(r) = \left( \sum_2^{T-1} u_t^2 \right) r^3 - \left( \sum_2^{T-1} u_t u_{t-1} \right) r^2 - \left( \sum_2^{T-1} u_t^2 + \nu \right) r + \left( \sum_2^{T-1} u_t u_{t-1} \right).$$

Ici on a p.s.

$$k(+1) = -\nu < 0 ; \quad k(-1) = 2 \sum_2^{T-1} u_t^2 + \nu > 0$$

$$\lim_{r \rightarrow +\infty} k(r) = +\infty ; \quad \lim_{r \rightarrow -\infty} k(r) = -\infty .$$

Raisonnant comme dans la preuve du lemme 1 on obtient que  $\overset{v}{\rho}_T$  est l'unique solution de  $k(r) = 0$  qui appartient à  $] -1, +1[$ .

Par des preuves semblables à celles des paragraphes précédents on peut montrer que  $\overset{v}{\rho}_T$  est convergent presque sûrement, plus précisément que pour tout  $\beta < 1/2$  on a

$$T^\beta (\overset{v}{\rho}_T - \rho) = o(1) \text{ p.s.,}$$

et en outre

$$T (\overset{v}{\rho}_T - \tilde{\rho}_T) = O(1) \text{ p.s..}$$

Compte tenu de la proposition 3 cela entraîne que la suite  $(T^{1/2}(\overset{v}{\rho}_T - \rho); T \geq 1)$  est asymptotiquement gaussienne centrée de variance  $1 - \rho^2$ .

Dans le cas où le paramètre autorégressif  $\rho$  est connu, l'estimateur de maximum de vraisemblance  $\overset{v}{v}_T$  de  $v$  est bien sûr  $\overset{v}{v}_T = v_T(\rho)$  où  $v_T(\cdot)$  est défini par (2.3). Comme il a été vu dans la preuve du lemme 2 on a alors

$$\overset{v}{v}_T = \frac{v}{T} \sum_1^T \tilde{v}_t^2,$$

où les variables aléatoires  $\tilde{v}_t, t \geq 1$  sont indépendantes gaussiennes centrées réduites.

Ainsi  $\overset{v}{v}_T$  est un estimateur presque sûrement convergent et de plus la loi de  $\overset{v}{v}_T$  est  $\chi^2\left(\frac{T}{2},$

$\frac{2v}{T}\right)$ . En particulier  $\overset{v}{v}_T$  est un estimateur sans biais et convergent en moyenne quadratique.

Evidemment la suite  $(T^{1/2}(\overset{v}{v}_T - v); T \geq 1)$  est asymptotiquement gaussienne centrée de variance  $2v^2$ .

### §3. Estimation des paramètres du signal

#### 3.1. Cas où le coefficient autorégressif du bruit est connu

On suppose ici que dans le modèle (1.1) le paramètre autorégressif  $\rho$  est connu et le paramètre de régression  $\gamma$  est inconnu dans  $\mathbb{R}^k$ , le paramètre de variance  $v$  étant éventuellement inconnu.

##### a. Estimation de maximum de vraisemblance

L'observation  $\underline{y}_T = (y_1, \dots, y_T)^*$  s'écrit

$$(3.1) \quad \underline{y}_T = Z_T \gamma + \underline{u}_T,$$

où  $Z_T$  est la matrice  $T \times k$  dont la  $i$ ème ligne est  $z_i^*$ . D'après (1.6), la fonction de Log-vraisemblance associée à  $\underline{y} = \underline{y}_T$  s'écrit

$$\begin{aligned} \text{Log } f_T(v, \rho; \underline{y} - Z g) \\ = -\frac{T}{2} \text{Log } 2\pi - \frac{T}{2} \text{Log } v + \frac{1}{2} \text{Log}(1 - \rho^2) - \frac{1}{2v} (\underline{y} - Z g)^* B (\underline{y} - Z g), \end{aligned}$$

où  $Z = Z_T$  et  $B = B_T(\rho)$ . Pour estimer  $(\gamma, v)$  par maximum de vraisemblance, il convient donc de minimiser, par rapport à  $(g, v) \in \mathbb{R}^k \times ]0, +\infty[$ , pour  $\underline{y}$  fixé, la fonction

$$(3.2) \quad \begin{aligned} \mathfrak{L}_T(g, v; \underline{y}) = T \text{Log } v \\ + \frac{1}{v} \left[ (1 - \rho^2) (y_1 - g^* z_1)^2 + \sum_2^T [(y_t - \rho y_{t-1}) - g^* (z_t - \rho z_{t-1})]^2 \right]. \end{aligned}$$

Pour tout  $g$  fixé, la fonction  $v \rightarrow \mathfrak{L}(g, v; \underline{y})$  est minimum pour  $v = \bar{v}_T(g)$  où

$$(3.3) \quad \begin{aligned} \bar{v}_T(g) = \frac{1}{T} (\underline{y} - Z g)^* B (\underline{y} - Z g) \\ = \frac{1}{T} \left[ (1 - \rho^2) (y_1 - g^* z_1)^2 + \sum_2^T [(y_t - \rho y_{t-1}) - g^* (z_t - \rho z_{t-1})]^2 \right], \end{aligned}$$

la valeur du minimum étant, à une constante près

$$\widehat{\mathfrak{L}}(g; \underline{y}) = T \text{Log } \bar{v}_T(g).$$

Il convient donc de minimiser  $\bar{v}_T(g)$  par rapport à  $g$ . Cela fournit, sous réserve que la matrice  $Z_T^* B_T(\rho) Z_T$  soit inversible, l'estimateur

$$(3.4) \quad \bar{\gamma}_T = (Z_T^* B_T(\rho) Z_T)^{-1} Z_T^* B_T(\rho) \underline{y},$$

pour  $\gamma$  et l'estimateur

$$(3.5) \quad \bar{v}_T = v_T(\bar{\gamma}_T),$$

pour  $v$ , où  $v_T(\cdot)$  est défini par (3.3).

## b. Propriétés des estimateurs

Commençons par les propriétés à nombre d'observations  $T$  fixé.

En vertu de (3.1), (3.4) et de (1.3) et (1.5) on peut écrire

$$(3.6) \quad \bar{\gamma}_T = \gamma + (Z_T^* B_T(\rho) Z_T)^{-1} Z_T^* M_T^*(\rho) \underline{y}_T,$$

où  $\underline{y}_T = (v_1, \dots, v_T)^* = M_T(\rho) \underline{u}_T$ , les variables aléatoires  $v_t$ ,  $t \geq 1$  étant indépendantes de loi gaussienne centrée de variance  $v$ . Bien sûr (3.6) montre que  $\bar{\gamma}_T$  est un estimateur sans biais de  $\gamma$ ;  $\bar{\gamma}_T$  est en fait de loi gaussienne de matrice de covariance  $v (Z_T^* B_T(\rho) Z_T)^{-1}$ .

Comme  $\bar{v}_T(\cdot)$  est donnée par (3.3), compte tenu de (3.5) et (3.6) on peut écrire

$$(3.7) \quad \begin{aligned} \bar{v}_T &= \frac{1}{T} \underline{u}_T^* B_T(\rho) \underline{u}_T - \frac{2}{T} (\bar{\gamma}_T - \gamma)^* Z_T^* B_T(\rho) \underline{u}_T + \frac{1}{T} (\bar{\gamma}_T - \gamma)^* (Z_T^* B_T(\rho) Z_T)^{-1} (\bar{\gamma}_T - \gamma) \\ &= \frac{1}{T} \underline{y}_T^* [ I - M_T Z_T (Z_T^* B_T(\rho) Z_T)^{-1} Z_T^* M_T^* ] \underline{y}_T. \end{aligned}$$

Remarquons que la matrice  $P_T = M_T Z_T (Z_T^* B_T(\rho) Z_T)^{-1} Z_T^* M_T^*$  est symétrique et idempotente; elle définit en fait l'opérateur de projection orthogonale sur le sous espace  $V_T$  engendré par les colonnes de  $M_T Z_T$ .

De même  $(I - P_T)$  qui apparaît dans (3.7) définit l'opérateur de projection orthogonale sur l'orthogonal  $V_T^\perp$  de  $V_T$ . Comme  $\dim V_T = k$  (puisque l'on suppose que  $Z_T^* B_T(\rho) Z_T$  inversible), utilisant les résultats concernant les vecteurs gaussiens (cf. e.g. MONFORT

(1982), Ch.III) (3.7) montre que  $\bar{v}_T$  est de loi  $\gamma(\frac{T-k}{2}, \frac{2v}{T})$ . En particulier  $E\bar{v}_T = \frac{T-k}{T} v$  et un estimateur sans biais de  $v$  est alors donné par

$$v_T^* = \frac{T-k}{T} \bar{v}_T.$$

De plus comme  $(Z_T^* B_T(\rho) Z_T)^{-1} Z_T^* M_T^* (I - P_T) = 0$ , les variables aléatoires  $(I - P_T)\bar{y}_T$  et  $(Z_T^* B_T(\rho) Z_T)^{-1} Z_T^* M_T^* \bar{y}_T$  sont indépendantes et donc, d'après (3.6) et (3.7), les estimateurs  $\bar{\gamma}_T$  et  $\bar{v}_T$  (resp.  $v_T^*$ ) sont indépendants. Notons aussi qu'on peut montrer que l'estimateur  $(\bar{\gamma}_T, v_T^*)$  est un estimateur sans biais de dispersion minimum (cf. e.g. MONFORT (1982), Ch.XXIV).

Examinons maintenant la convergence des estimateurs.

Notons  $(C(r))$  la condition :

$$(C(r)) : \lim_{T \rightarrow +\infty} (Z_T^* B_T(r) Z_T)^{-1} = 0.$$

Compte tenu de ce que  $E((\bar{\gamma}_T - \gamma)(\bar{\gamma}_T - \gamma)^*) = v (Z_T^* B_T(\rho) Z_T)^{-1}$ , si la condition  $C(\rho)$  est satisfaite, alors  $\bar{\gamma}_T$  est un estimateur convergent en moyenne quadratique. Notons que la condition  $(C(r))$  est équivalente à

$$\lim_{T \rightarrow +\infty} \lambda_{\max}((Z_T^* B_T(r) Z_T)^{-1}) = 0 \text{ ou } \lim_{T \rightarrow +\infty} \lambda_{\min}(Z_T^* B_T(r) Z_T) = +\infty$$

où  $\lambda_{\min}(A)$  (resp.  $\lambda_{\max}(A)$ ) désigne la plus petite (resp. grande) valeur propre de la matrice symétrique  $A$ .

En fait on peut montrer

**Lemme 3:** Si la condition  $(C(\rho))$  est satisfaite alors l'estimateur  $\bar{\gamma}_T$  défini par (3.4) est presque sûrement convergent.

Signalons qu'au §3.2.c. ci-après nous préciserons la vitesse de convergence de  $\bar{\gamma}_T$  vers  $\gamma$

**Preuve.** La matrice  $Z_T^* M_T^*(\rho)$  est la matrice  $T \times k$  de terme général

$$\begin{cases} (1-\rho^2)^{1/2} z_{1j} & \text{si } t=1 \text{ et } j = \overline{1, k}, \\ z_{tj} - \rho z_{t-1, j} & \text{si } t = \overline{2, T} \text{ et } j = \overline{1, k}, \end{cases}$$

où  $\overline{1, k}$  est l'ensemble des entiers  $\{1, 2, \dots, k\}$ .

Alors la  $j^{\text{ème}}$  composante du vecteur  $Z_T^* M_T^*(\rho) \underline{v}_T$  s'écrit

$$(1-\rho^2)^{1/2} z_{1j} v_1 + \sum_2^T (z_{tj} - \rho z_{t-1, j}) v_t .$$

Le processus  $(Z_T^* M_T^*(\rho) \underline{v}_T ; T \geq 1)$  est donc une martingale gaussienne à valeurs dans  $\mathbb{R}^k$ ; son processus croissant est  $(v Z_T^* B_T(\rho) Z_T ; T \geq 1)$ .

Alors, appliquant les résultats LE BRETON et MUSIELA (1987b), on obtient immédiatement l'assertion annoncée compte tenu de la représentation de  $\bar{\gamma}_T - \gamma$  fournie par (3.6). 

Comme  $v_T^*$  est de loi  $\gamma(\frac{T-k}{2}, \frac{2v}{T-k})$  on a  $E(v_T^* - v)^2 = \frac{2v^2}{T-k}$  et donc  $v_T^*$  est un estimateur convergent en moyenne quadratique.

Ecrivant, d'après (3.7)

$$v_T^* = \frac{1}{T-k} [ \underline{v}_T^* \underline{v}_T - \underline{v}_T^* P_T \underline{v}_T ],$$

où  $P_T$  est le projecteur défini ci-dessus, par une démonstration analogue à celle du

théorème 5 dans MONFORT (1982), on peut montrer que  $v_T^*$  est aussi presque sûrement convergent.

Examinons maintenant la loi asymptotique de l'estimateur  $(\bar{\gamma}_T, v_T^*)$ . Les variables aléatoires  $\bar{\gamma}_T$  et  $v_T^*$  étant indépendantes l'étude se réduit aux lois marginales asymptotiques. Bien sûr la suite  $(\sqrt{T} (v_T^* - v); T \geq 1)$  est asymptotiquement gaussienne centrée de variance  $2v^2$ .

Pour tout T, comme  $\bar{\gamma}_T - \gamma$  est de loi gaussienne centrée de matrice de covariance  $v (Z_T^* B_T(\rho) Z_T)^{-1}$ , la loi de  $v^{-1/2} (Z_T^* B_T(\rho) Z_T)^{1/2} (\bar{\gamma}_T - \gamma)$  est gaussienne centrée de matrice de covariance  $\mathbb{1}_k$ .

Envisageons maintenant des hypothèses raisonnables, ne concernant que les régresseurs, sous lesquelles nous préciserons complètement le comportement asymptotique de l'estimateur  $\bar{\gamma}_T$ .

Ecrivons  $B_T(\rho)$ , définie par (1.5), sous la forme

$$(3.8) \quad B_T(\rho) = (1 + \rho^2)\mathbb{1}_T - \rho^2 J_T - \rho K_T.$$

où  $J_T$  est la matrice  $T \times T$  dont tous les termes sont nuls sauf  $J_T(1, 1) = J_T(T, T) = 1$  et  $K_T$  est la matrice  $T \times T$  symétrique telle que  $K_{ij} = \delta_{i, j-1}$ ,  $1 \leq i \leq j \leq T$ . Alors

$$Z_T^* B_T(\rho) Z_T = (1 + \rho^2) Z_T^* Z_T - \rho^2 ((z_{1i} z_{1j} + z_{Ti} z_{Tj}))$$

$$(3.9) \quad - \rho \left\{ \sum_{t=2}^T z_t z_{t-1}^* + \sum_{t=2}^T z_{t-1} z_t^* \right\}.$$

Définissons  $d_{Tj}$  pour  $j = \overline{1, k}$  par

$$(3.10) \quad d_{Tj} = \left( \sum_{t=1}^T z_{tj}^2 \right)^{1/2}$$

i.e.  $d_{Tj}^2$  est le  $j^{\text{eme}}$  terme diagonal de la matrice  $Z_T^* Z_T$ ; notons  $D_T$  la matrice  $k \times k$  diagonale dont les termes diagonaux sont les  $d_{Tj}$ ,  $j = \overline{1, k}$ .

Introduisons l'hypothèse

(A1) : Pour  $j = \overline{1, k}$  on a

$$\lim_{T \rightarrow +\infty} d_{Tj}^{-1} \max_{1 \leq i \leq T} |z_{ij}| = 0,$$

et, pour  $\tau \in \mathbb{N}$ , l'hypothèse

(A<sub>2</sub>( $\tau$ )) : la suite  $(D_T^{-1} \left[ \sum_{t=\tau+1}^T z_t z_{t-\tau}^* \right] D_T^{-1} ; T \geq 1)$  de matrices  $k \times k$  converge vers une limite (finie)  $G(\tau)$ .

Nous pouvons alors énoncer

**Lemme 4.** *Supposons que les conditions (A1) et (A<sub>2</sub>( $\tau$ )),  $\tau = 0, 1$  sont satisfaites. Alors pour toute valeur  $r \in [-1, +1]$  la suite de matrices  $(D_T^{-1} Z_T^* B_T(r) Z_T D_T^{-1} ; T \geq 1)$  converge vers*

$$(3.11) \quad H(r) = (1 + r^2) G(0) - r (G(1) + G^*(1)).$$

**Preuve.** D'après (3.9) on a

$$\begin{aligned} D_T^{-1} Z_T^* B_T(r) Z_T D_T^{-1} &= (1 + r^2) D_T^{-1} Z_T^* Z_T D_T^{-1} \\ &\quad - r^2 ((d_{Ti}^{-1} d_{Tj}^{-1} [z_{1i} z_{1j} + z_{Ti} z_{Tj}])) \\ &\quad - r [D_T^{-1} \left( \sum_{t=2}^T z_t z_{t-1}^* \right) D_T^{-1} + D_T^{-1} \left( \sum_{t=2}^T z_{t-1} z_t^* \right) D_T^{-1}] \end{aligned}$$

La condition (A1) assure que le second terme du second membre tend vers zéro. la condition (A<sub>2</sub>(0)) (resp. (A<sub>2</sub>(1))) assure que le premier (resp. troisième) terme tend vers  $(1 + r^2) G(0)$  (resp.  $-r (G(1) + G^*(1))$ ).  $\blacktriangleleft$

Notons que la matrice  $H(r)$  définie par (3.11) est symétrique et au moins semi-définie positive.

Dans la suite on note, pour  $r \neq 0$ ,  $(D(r))$  la condition

$(D(r)) : (A1) \wedge A_2(0) \wedge A_2(1) \wedge$  "la matrice  $H(r)$  définie par (3.11) est régulière".

et, pour  $r = 0$ ,  $(D(0))$  la condition

$(D(0)) : (A1) \wedge A_2(0) \wedge$  "la matrice  $G(0)$  est régulière".

On note aussi  $(A3)$  la condition

$(A3) : \lim_{T \rightarrow +\infty} d_{Tm} = +\infty$

où  $d_{Tm} = \min_{1 \leq j \leq k} d_{Tj}$ .

Nous pouvons maintenant énoncer le résultat suivant concernant le comportement asymptotique de l'estimateur  $\bar{\gamma}_T$ .

**Proposition 4.** *Supposons que les conditions  $(D(\rho))$  et  $(A3)$  sont satisfaites. Alors l'estimateur  $\bar{\gamma}_T^*$  est convergent presque sûrement et de plus la suite  $\{D_T(\bar{\gamma}_T^* - \gamma) ; T \geq 1\}$  est asymptotiquement gaussienne centrée de matrice de covariance  $v H^{-1}(\rho)$ .*

**Preuve.** En vertu de l'hypothèse  $(D(\rho))$  on a

$$(3.12) \quad \lim_{T \rightarrow +\infty} D_T^{-1} (Z_T^* B_T(\rho) Z_T)^{-1} D_T^{-1} = H^{-1}(\rho),$$

ce qui assure la convergence en loi annoncée. En outre, compte tenu de  $(A3)$ , (3.12) montre que la condition  $(C(\rho))$  du lemme 3 est satisfaite. En conséquence  $\bar{\gamma}_T$  converge presque sûrement.  $\bullet$

Examinons des exemples de régresseurs où les conditions envisagées ci-dessus peuvent être satisfaites.

### Exemples 3.1

Dans le modèle de régression (1.1) du §1, nous avons envisagé les hypothèses (A1), (A<sub>2</sub>(τ)) et (A3) concernant le régresseur (z<sub>t</sub>; t ≥ 1).

A titre d'application, on peut considérer les modèles de régression suivants:

#### (a) : Modèle quasi-ergodique

On dit que la suite (z<sub>t</sub>; t ≥ 0) est quasi-ergodique si les limites  $\frac{1}{T} \sum_{t=\tau}^T z_{ti} z_{t-\tau, j}$ ; (i, j) ∈ {1, .., k} existent et sont strictement positives pour i = j et τ = 0.

Dans ce cas les hypothèses (A<sub>2</sub>(τ)) et (A3) sont satisfaites et de plus (A1) est satisfaite si par exemple les régresseurs sont bornés.

Considérons un cas particulier:

#### Cas du Modèle sinusoïdal

Ici

$$y_t = \sum_{j=1}^k [\gamma_j^1 \sin(\omega_j t) + \gamma_j^2 \cos(\omega_j t)] + u_t; t \geq 1.$$

La suite des régresseurs (z<sub>t</sub> = (z<sub>t</sub><sup>1</sup>, z<sub>t</sub><sup>2</sup>), t ≥ 1) est la suite de vecteurs dans ℝ<sup>2k</sup> z<sub>t</sub><sup>1</sup> = ((sin(ω<sub>1</sub>t), ..., sin(ω<sub>k</sub>t))<sup>\*</sup> ; z<sub>t</sub><sup>2</sup> = ((cos(ω<sub>1</sub>t), ..., cos(ω<sub>k</sub>t))<sup>\*</sup> où les ω<sub>j</sub> sont connus, 0 < ω<sub>j</sub> < π et ω<sub>i</sub> ≠ ω<sub>j</sub>.

Compte tenu de ce qui vient d'être dit pour montrer que (A1), (A<sub>2</sub>(τ)) et (A3) sont satisfaites il suffit de vérifier que les limites des quantités:

$$\frac{1}{T} \sum_{t=\tau+1}^T \sin(\omega_i t) \sin(\omega_j (t - \tau)),$$

$$\frac{1}{T} \sum_{t=\tau+1}^T \cos(\omega_i t) \cos(\omega_j (t - \tau)),$$

$$\frac{1}{T} \sum_{t=\tau+1}^T \sin(\omega_i t) \cos(\omega_j (t - \tau))$$

existent pour i ≠ j et sont strictement positives pour i = j, τ = 0. Nous faisons le calcul pour le premier cas, les autres pouvant être conduits de la même façon.

Pour  $i \neq j$  on a

$$\begin{aligned} \lim_{T \rightarrow +\infty} \frac{1}{T} \sum_{t=\tau+1}^T \sin(\omega_i t) \sin(\omega_j (t - \tau)) &= \cos(\omega_j \tau) \lim_{T \rightarrow +\infty} \frac{1}{T} \sum_{t=\tau+1}^T \sin(\omega_i t) \sin(\omega_j t) \\ &\quad - \sin(\omega_j \tau) \lim_{T \rightarrow +\infty} \frac{1}{T} \sum_{t=\tau+1}^T \sin(\omega_i t) \cos(\omega_j t) \end{aligned}$$

Or, pour  $i \neq j$ ,

$$\lim_{T \rightarrow +\infty} \frac{1}{T} \sum_{t=\tau+1}^T \sin(\omega_i t) \sin(\omega_j t) = \lim_{T \rightarrow +\infty} \frac{1}{2T} \sum_{t=\tau+1}^T [\cos((\omega_i - \omega_j) t) - \cos((\omega_i + \omega_j) t)] = 0$$

et de même

$$\lim_{T \rightarrow +\infty} \frac{1}{T} \sum_{t=\tau+1}^T \sin(\omega_i t) \cos(\omega_j t) = 0.$$

On a donc

$$\lim_{T \rightarrow +\infty} \frac{1}{T} \sum_{t=\tau+1}^T \sin(\omega_i t) \sin(\omega_j (t - \tau)) = 0 \text{ pour } i \neq j, \tau \in \mathbb{N}.$$

Pour  $i = j, \tau = 0$  on a

$$\lim_{T \rightarrow +\infty} \frac{1}{T} \sum_{t=1}^T \sin^2(\omega_i t) = \frac{1}{2} - \lim_{T \rightarrow +\infty} \frac{1}{2T} \sum_{t=1}^T \cos(2\omega_i t) = \frac{1}{2}.$$

Plus généralement si  $\tau \neq 0$  on a

$$\lim_{T \rightarrow +\infty} \frac{1}{T} \sum_{t=\tau+1}^T \sin(\omega_i t) \sin(\omega_i (t - \tau)) = \frac{\cos(\omega_i \tau)}{2}.$$

### **(b) : Modèle polynômial**

Ici

$$y_t = \sum_{j=1}^T \gamma_j t^j + u_t ; t \geq 1.$$

La suite des régresseurs  $(z_t ; t \geq 1)$  est la suite de vecteurs  $z_T^* = (t, t^2, \dots, t^i, \dots, t^k)$  de  $\mathbb{R}^k$ .

Montrons que (A1) est satisfaite.

Pour  $j = \overline{1..k}$  on a

$$d_{Tj}^{-1} \max_{1 \leq t \leq T} |z_{jt}| = \left( \sum_{t=1}^T t^{2j} \right)^{-1/2} \max_{1 \leq t \leq T} |t^j|$$

$$\leq \left( \sum_{t=1}^T t^{2j} \right)^{-1/2} T^j$$

qui tend vers 0, quand T tend vers  $+\infty$ , car le développement

$$\sum_{1 \leq t \leq T} t^r = \frac{1}{r+1} \sum_{0 \leq u \leq r} B_u (r+1) (T+1)^{r+1-u},$$

où les  $B_u$  sont les nombres de Bernoulli, montre que la quantité  $\sum_{t=1}^T t^{2j}$  est un polynôme en  $T^{2j+1}$ . Ainsi (A1) est vérifiée.

Examinons maintenant les conditions  $A_2(\tau)$ .

Pour tout  $(i, j) \in \{1, \dots, k\}^2$  on a

$$\frac{\sum_{t=\tau+1}^T t^i (t-\tau)^j}{\left( \sum_{t=1}^T t^{2i} \right)^{1/2} \left( \sum_{t=1}^T t^{2j} \right)^{1/2}} = \sum_{s=0}^j \frac{j! (-\tau)^{j-s}}{s! (j-s)!} \frac{\sum_{t=\tau+1}^T t^{i+s}}{\left( \sum_{t=1}^T t^{2i} \right)^{1/2} \left( \sum_{t=1}^T t^{2j} \right)^{1/2}}$$

dont tous les termes autres que le terme d'indice j tendent vers zéro, quand T tend vers

$+\infty$ . Utilisant comme ci-dessus les développements des sommes du type  $\sum_{t=1}^T t^r$  on obtient alors que

$$\lim_{T \rightarrow +\infty} \frac{\sum_{t=\tau+1}^T t^{i+j}}{\left( \sum_{t=1}^T t^{2i} \right)^{1/2} \left( \sum_{t=1}^T t^{2j} \right)^{1/2}} = \frac{2 \sqrt{ij}}{i+j}.$$

La condition (A3) est évidemment vérifiée.

Bien sûr dans le modèle " mixte ":

$$y_t = \sum_{j=1}^k \gamma_j t^j + \sum_{j=1}^k [ \gamma_j^1 \sin(\omega_j t) + \gamma_j^2 \cos(\omega_j t) ] + u_t; t \geq 1,$$

les hypothèses sont satisfaites.

### 3.2. Cas où le coefficient autorégressif du bruit est inconnu

On suppose ici que dans le modèle (1.1) le paramètre autorégressif  $\rho$  est inconnu, le paramètre de variance  $v$  étant éventuellement aussi inconnu. On s'intéresse à l'estimation du seul paramètre de régression  $\gamma$ .

#### a. Estimation de moindres carrés

La méthode des moindres carrés consiste à estimer  $\gamma$  par minimisation de la quantité

$\sum_{i=2}^T (y_i - g^* z_i)^2$ . Cela conduit à l'estimateur

$$(3.13) \quad \hat{\gamma}_T = (Z_T^* Z_T)^{-1} Z_T^* y_T,$$

sous réserve que la matrice  $Z_T^* Z_T$  soit inversible.

D'après (3.1) et (3.13) on a

$$(3.14) \quad \hat{\gamma}_T = \gamma + (Z_T^* Z_T)^{-1} Z_T^* u_T.$$

Cela montre que  $\hat{\gamma}_T$  est un estimateur sans biais de  $\gamma$  qui est en fait gaussien de matrice de covariance

$$(3.15) \quad \text{Cov}(\hat{\gamma}_T) = v (Z_T^* Z_T)^{-1} Z_T^* B^{-1}(\rho) Z_T (Z_T^* Z_T)^{-1}.$$

#### b. Comportement asymptotique de l'estimateur des moindres carrés

D'après (3.15) on a

$$\begin{aligned} \lambda_{\max}(\text{Cov}(\hat{\gamma}_T)) &= v \sup_{|\mathbf{x}|=1} \mathbf{x}^* (Z_T^* Z_T)^{-1} Z_T^* B^{-1}(\rho) Z_T (Z_T^* Z_T)^{-1} \mathbf{x} \\ &\leq v \lambda_{\max}(B^{-1}(\rho)) \sup_{|\mathbf{x}|=1} \mathbf{x}^* (Z_T^* Z_T)^{-1} Z_T^* Z_T (Z_T^* Z_T)^{-1} \mathbf{x} \\ &= v \lambda_{\max}(B_T^{-1}(\rho)) \lambda_{\max}((Z_T^* Z_T)^{-1}). \end{aligned}$$

Or  $B_T^{-1}(\rho)$  est la matrice de covariance d'un vecteur de  $T$  variables consécutives d'un processus AR(1) stationnaire de paramètre autorégressif  $\rho$  et de variance d'innovation unité. Donc

$$B_T^{-1}(\rho) = \left( \left( \int_{-\pi}^{+\pi} e^{i\lambda(t-s)} f_\rho(\lambda) d\lambda \right) \right)_{s,t}$$

où  $f_\rho(\lambda)$  est la densité spectrale d'un tel processus i.e.  $f_\rho(\lambda) = (2\pi |1 - \rho e^{i\lambda}|^2)^{-1}$  dont il est facile de voir qu'elle est majorée par  $(2\pi (1 - |\rho|)^{-2})$ . Alors on a pour  $\underline{x} \in \mathbb{R}^T$

$$\begin{aligned} \underline{x}^* B_T^{-1}(\rho) \underline{x} &= \int_{-\pi}^{+\pi} \left| \sum_1^T x_t e^{i\lambda t} \right|^2 f_\rho(\lambda) d\lambda \\ &\leq (1 - |\rho|)^{-2} \frac{1}{2\pi} \int_{-\pi}^{+\pi} \left| \sum_1^T x_t e^{i\lambda t} \right|^2 d\lambda \\ &= (1 - |\rho|)^{-2} |\underline{x}|^2. \end{aligned}$$

Cela montre que  $\lambda_{\max}(B_T^{-1}(\rho)) \leq (1 - |\rho|)^{-2}$ . Par suite, si la condition (C(0)) (i.e. (C(r)) pour  $r = 0$  soit donc  $\lim_{T \rightarrow +\infty} (Z_T^* Z_T)^{-1} = 0$ ) est satisfaite, alors l'estimateur  $\hat{\gamma}_T$  est convergent en moyenne quadratique.

En fait on peut montrer

**Lemme 5.** *Si la condition (C(0)) est satisfaite, alors l'estimateur  $\hat{\gamma}_T$  donné par (3.13) est convergent presque sûrement.*

Signalons qu'au §.3.2.c ci-après nous précisons la vitesse de convergence de  $\hat{\gamma}_T$  vers  $\gamma$ .

**Preuve.** En vertu du théorème 1 de LAI et al (1979) il suffit de démontrer que pour toute

suite déterministe  $(c_t; t \geq 1)$  telle que  $\sum_{t=1}^{\infty} c_t^2 < \infty$ , la série  $\sum_{t=1}^{\infty} c_t u_t$  est presque sûrement convergente. Or cette dernière assertion est vraie pour  $\rho = 0$  i.e.  $u_t = v_t, t \geq 1$

puisqu'alors la suite  $\{\sum_{s=1}^t c_s u_s; t \geq 1\}$  est une martingale de carré intégrable de processus

croissant  $\{v \sum_{s=1}^t c_s^2; t \geq 1\}$ , et cette martingale converge presque sûrement si  $\sum_{s=1}^{\infty} c_s^2 < \infty$ .

Pour  $\rho \neq 0$  dans  $]-1, +1[$  on peut écrire

$$u_t = \rho^{t-1} \left\{ u_1 + \sum_{s=1}^{t-1} \rho^{-s} v_{s+1} \right\}, t \geq 2,$$

d'où

$$(3.16) \quad \sum_1^T c_t u_t = \left\{ \sum_1^T c_t \rho^{t-1} \right\} u_1 + \sum_2^T c_t \rho^{t-1} \left\{ \sum_{s=1}^{t-1} \rho^{-s} v_{s+1} \right\}.$$

D'après la condition  $\sum_{t=1}^{\infty} c_t^2 < \infty$  on a

$$\left\{ \sum_1^{\infty} |c_t \rho^{t-1}| \right\}^2 = \left\{ \sum_{t=1}^{\infty} c_t^2 \right\} \left\{ \sum_1^{\infty} \rho^{2(t-1)} \right\} < \infty,$$

et le premier terme du membre de droite de (3.16) converge presque sûrement. Montrons qu'il en est de même du deuxième terme. Il peut s'écrire

$$(3.17) \quad \sum_2^T \tilde{c}_t \rho^{-t} A_{t-1}^{-1} \left\{ \sum_{s=1}^{t-1} \rho^{-s} v_{s+1} \right\},$$

avec

$$A_t = \sum_1^t \rho^{-2s}; \quad \tilde{c}_s = c_s \frac{\rho(1 - \rho^{2(s-1)})}{1 - \rho^2}.$$

Il est facile de vérifier que la série  $\sum_2^{\infty} \tilde{c}_t^2 \frac{A_t}{A_{t-1}}$  est convergente. Comme nous avons vu

que  $(v_t, t \geq 1)$  vérifie l'assertion souhaitée pour  $(u_t, t \geq 1)$ , par application du lemme 2 de LAI et al (1979), on obtient que la suite définie par (3.17) converge presque sûrement vers une limite finie. Cela achève la démonstration.  $\blacksquare$

Nous sommes alors en mesure de démontrer

**Proposition 5.** *Si les conditions (D(0)) et (A3) sont satisfaites alors l'estimateur  $\hat{\gamma}_T$  défini par (3.13) est convergent presque sûrement. Si en outre les conditions  $(A_2(\tau))$ ,  $\tau \geq 2$  sont satisfaites alors la suite  $(D_T(\hat{\gamma}_T - \gamma); T \geq 1)$  est asymptotiquement gaussienne de matrice de covariance  $v G^{-1}(0) K(\rho) G^{-1}(0)$  où*

$$(3.18) \quad K(\rho) = \lim_{T \rightarrow +\infty} D_T^{-1} Z_T^* B_T^{-1}(\rho) Z_T D_T^{-1}.$$

**Preuve.** En vertu de (A3), la condition (D(0)) garantit que la condition (C(0)) du lemme 5 est vérifiée et donc  $\hat{\gamma}_T$  converge p.s..

La famille de conditions  $(A_2(\tau))$ ,  $\tau \geq 0$  assure la convergence faible de la suite de mesures

$$\left\{ D_T^{-1} \left( \sum_1^T z_t e^{i\lambda t} \right) \left( \sum_1^T z_s e^{-i\lambda s} \right)^* D_T^{-1} \mathbb{1}_{[-\pi, +\pi]}(\lambda) d(\lambda); T \geq 1 \right\} \text{ vers une mesure } \mu(d(\lambda)).$$

Considérons  $f_\rho$  telle qu'au début de ce paragraphe. Alors on a

$$B_T^{-1}(\rho) = \left( \int_{-\pi}^{+\pi} e^{i\lambda(t-s)} f_\rho(\lambda) d\lambda \right)$$

d'où

$$D_T^{-1} Z_T^* B_T^{-1}(\rho) Z_T D_T^{-1} = \int_{-\pi}^{+\pi} D_T^{-1} \left( \sum_1^T z_t e^{i\lambda t} \right) \left( \sum_1^T z_s e^{-i\lambda s} \right)^* D_T^{-1} f_\rho(\lambda) d\lambda.$$

La fonction  $f_\rho$  étant continue bornée on a donc

$$\lim_{T \rightarrow +\infty} D_T^{-1} Z_T^* B_T^{-1}(\rho) Z_T D_T^{-1} = \int_{-\pi}^{+\pi} f_\rho(\lambda) \mu(d\lambda),$$

et la limite  $K(\rho)$  annoncée dans (3.18) existe.

D'après (3.15) on a

$$D_T \text{Cov}(\hat{\gamma}_T) D_T = v D_T (Z_T^* Z_T)^{-1} D_T D_T^{-1} Z_T^* B_T^{-1}(\rho) Z_T D_T^{-1} D_T (Z_T^* Z_T)^{-1} D_T.$$

Alors, en vertu de  $(D(0))$  et de (3.18) il vient

$$\lim_{T \rightarrow +\infty} D_T \text{Cov}(\hat{\gamma}_T) D_T = v G^{-1}(0) K(\rho) G^{-1}(0)$$

d'où la convergence en loi annoncée. 🍏

Notons que dans les exemples 3.1 ci-dessus la famille des conditions  $(A_2(\tau))$ ,  $\tau \geq 0$  est satisfaite.

Dans le chapitre suivant §2 nous envisagerons des procédures d'estimation simultanée de tous les paramètres  $\rho$ ,  $v$ ,  $\gamma$  fondées sur une première étape d'estimation par moindres carrés du paramètre  $\gamma$ .

### c. Etude d'autres estimateurs

Remarquons que l'estimateur  $\bar{\gamma}_T$  de maximum de vraisemblance pour  $\rho$  connu défini par (3.4) et l'estimateur des moindres carrés  $\hat{\gamma}_T$  défini par (3.13) qui viennent d'être étudiés sont des éléments particuliers de la famille  $\{g_T(r); r \in ]-1, +1[ \}$  définie par

$$(3.19) \quad g_T(r) = (Z_T^* B_T(r) Z_T)^{-1} Z_T^* B_T(r) y_T.$$

En fait  $\bar{\gamma}_T = g_T(\rho)$  et  $\hat{\gamma}_T = g_T(0)$ .

Dans la suite on note  $g_T(r) = (g_T^1(r), \dots, g_T^k(r))^*$  et  $\Lambda_T(r) = ((\Lambda_T^{ij}(r)))$  la matrice  $k \times k$

$$(3.20) \quad \Lambda_T(r) = (Z_T^* B_T(r) Z_T)^{-1}.$$

Nous pouvons alors compléter les lemmes 4 et 5 par

**Lemme 6.** Si la condition (C(r)) i.e.  $\lim_{T \rightarrow +\infty} \Lambda_T(r) = 0$  est satisfaite, alors l'estimateur  $g_T(r)$  de  $\gamma$  défini par (3.19) est convergent presque sûrement. Plus précisément, pour tout  $\delta > 0$  on a

$$(3.21) \quad g_T^j(r) - \gamma_j = o\left(\left\{\left|\Lambda_T^{jj}(r)\right| \left|\text{Log } \Lambda_T^{jj}(r)\right|^{1+\delta}\right\}^{1/2}\right) \text{ p.s..}$$

**Preuve.** Remarquons d'abord que  $g_T(r)$  apparaît comme l'estimateur de moindres carrés ordinaires dans le modèle

$$M_T(r) y_T = M_T(r) Z_T \gamma + M_T(r) u_T.$$

Alors, en vertu de théorème 1 de LAI et al (1979) pour assurer l'égalité (3.21), et donc aussi la convergence presque sûrement de  $g_T(r)$  vers  $\gamma$ , il suffit de démontrer que si  $(u_t(r), t \geq 1)$  est défini par

$$(3.22) \quad u_1(r) = (1 - r^2)^{1/2} u_1 \quad ; \quad u_t(r) = u_t - r u_{t-1}, \quad t \geq 2,$$

alors pour toute suite déterministe  $(c_t; t \geq 1)$  telle que  $\sum_{t=1}^{\infty} c_t^2 < \infty$ , la série  $\sum_1^{\infty} c_t u_t(r)$  est convergente presque sûrement. D'après la preuve du lemme 5 cette propriété est vraie pour  $r = 0$  i.e.  $u_t(r) = u_t$ . Au vu de (3.22) il est alors clair qu'elle est aussi vérifiée pour  $r \neq 0$ . 

Sous les hypothèses (D(r)) et (A(3)) la condition (C(r)) est satisfaite et de plus

$$\lim_{T \rightarrow +\infty} d_{Tj}^2 \Lambda_T^{jj}(r) = H_{jj}^{-1}(r).$$

Alors (3.21) assure que pour tout  $\eta > 1/2$

on a

$$g_T^j(r) - \gamma_j = o\left(\left\{d_{Tj}^{-1}(\text{Log } d_{Tj})^\eta\right\}\right) \text{ p.s.}$$

et donc aussi

$$|D_T(g_T(r) - \gamma)| = o((\text{Log } d_{TM})^\eta) \text{ p.s.}$$

où

$$d_{TM} = \max_{1 \leq j \leq k} d_{Tj}.$$

Dans le chapitre suivant nous serons amenés à considérer une situation dans laquelle l'égalité précédente a lieu uniformément par rapport à  $r \in [-1, +1]$ .

**Proposition 6.** *Supposons que la condition (A3) et la condition (D) suivante sont satisfaites :*

$$(D) : (A_1) \wedge (A_2(0)) \wedge (A_2(1)) \wedge " \text{Inf}_{r \in [-1, +1]} \lambda_{\min}(H(r)) > 0 ",$$

la matrice  $H(r)$  étant définie par (3.11). Alors pour tout  $\eta > 1/2$  on a

$$(3.23) \quad \text{Sup}_{r \in [-1, +1]} |D_T(g_T(r) - \gamma)| = o((\text{Log } d_{TM})^\eta) \text{ p.s..}$$

**Preuve.** Compte tenu de (3.1) et de (3.19), on peut écrire

$$\begin{aligned} D_T(g_T(r) - \gamma) &= D_T(Z_T^* B_T(r) Z_T)^{-1} Z_T^* B_T(r) \underline{u}_T \\ &= \Psi_T(r) \Delta_T(r), \end{aligned}$$

où

$$\Psi_T(r) = (D_T^{-1} Z_T^* B_T(r) Z_T D_T^{-1})^{-1},$$

et

$$\Delta_T(r) = D_T^{-1} Z_T^* B_T(r) \underline{u}_T.$$

Alors, comme d'après (D) on a  $\Psi_T(r) = O(1)$  uniformément par rapport à  $r$ , il suffit, pour démontrer l'assertion de la proposition, d'assurer que  $\Delta_T(r) = o((\text{Log } d_{TM})^\eta)$  presque sûrement uniformément par rapport à  $r$ . Pour ce faire il suffit, d'après (3.8), de montrer que pour  $A_T^0 = \underline{1}_T$ ,  $A_T^1 = K_T$  et  $A_T^2 = \underline{1}_T - J_T$  on a  $\Delta_T^i = o((\text{Log } d_{TM})^\eta)$  p.s. pour

$$\Delta_T^i = D_T^{-1} Z_T^* A_T^i \underline{u}_T, \quad i = 0, 2.$$

Dans ce qui suit on note  $\Delta_T^i = (\delta_T^{i1}, \dots, \delta_T^{ik})^*$ . Pour  $i = 0$  on a

$$(\text{Log } d_{TM})^{-\eta} \left| \delta_T^{0j} \right| \leq 2^\eta h_j^{-1}(T) \left| \sum_{t=1}^T z_{tj} u_t \right|,$$

où

$$h_j(t) = d_{tj} (\text{Log } d_{tj})^\eta.$$

On a aussi

$$h_j^{-2}(t) z_{tj}^2 = \frac{d_{tj}^2 - d_{t-1,j}^2}{d_{tj}^2 (\text{Log } d_{tj}^2)^{2\eta}} \leq \int_{d_{t-1,j}^2}^{d_{tj}^2} \frac{ds}{s (\text{Log } s)^{2\eta}},$$

et donc, comme  $2\eta > 1$ ,  $\sum_1^\infty h_j^{-2}(t) z_{tj}^2 < \infty$ . Alors, d'après la preuve du lemme 5, la série

$\sum_1^\infty h_j^{-1}(t) z_{tj}^2$  est convergente presque sûrement. Appliquant le lemme de Kronecker (cf. e.g. DACUNHA-CASTELLE et DUFLO (1983) p.73), dans la mesure où d'après (A3)

on a  $\lim_{T \rightarrow +\infty} h_j(T) = +\infty$ , on obtient que  $(\text{Log } d_{TM})^{-\eta} \delta_T^{0j} = o(1)$  p.s..

Pour  $i = 2$  (resp. 1) on a

$$(\text{Log } d_{TM})^{-\eta} \left| \delta_T^{2j} \right| \leq 2^\eta h_j^{-1}(T) \sum_{t=2}^{T-1} z_{tj} u_t$$

$$\text{(resp. } (\text{Log } d_{TM})^{-\eta} \left| \delta_T^{1j} \right| \leq 2^\eta h_j^{-1}(T) \left( \left| \sum_{t=2}^T z_{tj} u_{t-1} \right| + \left| \sum_{t=2}^T z_{t-1j} u_t \right| \right)$$

et raisonnant comme pour  $\delta_T^{0j}$  on obtient que  $(\text{Log } d_{TM})^{-\eta} \delta_T^{ij} = o(1)$  presque sûrement pour  $i = \overline{1..2}$ .  $\blacktriangle$

Signalons que dans les exemples 3.1 ci-dessus la condition (D) est satisfaite.



### **Chapitre III**

**ESTIMATION PARAMETRIQUE SIMULTANEE  
DANS UN MODELE "SIGNAL + BRUIT AR(1)"**



Dans ce chapitre nous considérons le modèle "signal +bruit AR(1)" tel qu'il est décrit au § 1 du chapitre précédent. Nous nous intéressons à l'estimation simultanée du paramètre de régression et des paramètres de structure du processus d'erreurs.

Plusieurs travaux antérieurs consacrés à ce type d'étude traitent d'estimation par maximum de vraisemblance. L'accent est mis sur les propriétés de convergence en probabilité et de normalité asymptotique des estimateurs sous des hypothèses plus ou moins restrictives concernant les régresseurs. Citons par exemple HILDRETH (1969). La question de l'existence et de l'unicité de l'estimateur de maximum de vraisemblance est peu abordée. Toutefois MALLER (1989) l'examine dans un cas particulier de régresseur bidimensionnel.

D'autres travaux proposent des procédures itératives pour le calcul d'estimateurs mais leurs propriétés asymptotiques sont incomplètement étudiées. Citons COCHRANE et ORCUTT (1949), OBERHOFER et KMENTA (1974), ...

Ici nous mettons l'accent sur l'existence et l'unicité de l'estimateur de maximum de vraisemblance et sur la convergence presque sûre des estimateurs fournis par les différentes procédures d'estimation que nous envisageons.

Dans la première partie du chapitre nous nous intéressons de manière précise à la maximisation de la fonction de vraisemblance et à l'existence d'un unique estimateur de maximum de vraisemblance dont nous examinons les propriétés asymptotiques. Dans la deuxième partie nous étudions une procédure d'estimation fondée sur une première étape de moindres carrés pour l'estimation du paramètre de régression. Dans la dernière partie du chapitre nous construisons une troisième procédure qui fournit des estimateurs facilement calculables dont nous montrons qu'ils possèdent les mêmes propriétés asymptotiques que ceux de maximum de vraisemblance.

## § 1. Estimation par maximum de vraisemblance

### 1.1. Maximisation de la fonction de vraisemblance

L'observation  $\underline{y}_T = (y_1, \dots, y_T)^*$  s'écrit encore

$$\underline{y}_T = Z_T \underline{\gamma} + \underline{u}_T,$$

et, d'après l'expression de la densité de  $\underline{u}_T$  donnée en (1.6) du chapitre II, la fonction de Log-vraisemblance associée à  $\underline{y} = \underline{y}_T$  s'écrit

$$(1.1) \quad L_T(g, r, v; \underline{y}) = -\frac{T}{2} \text{Log} 2\pi - \frac{T}{2} \text{Log} v + \frac{1}{2} \text{Log}(1 - r^2) - (2v)^{-1} (\underline{y} - Z g)^* B(r) (\underline{y} - Z g).$$

Pour estimer les paramètres par maximum de vraisemblance il s'agit donc de maximiser, par rapport à  $(g, r, v) \in \mathbb{R}^{k \times 1} \times ]-1, +1[ \times ]0, +\infty[$  pour  $\underline{y}$  fixé, la quantité  $L(g, r, v; \underline{y})$  donnée par (1.1).

Pour tout  $(r, v)$  fixé dans  $] -1, +1[ \times ]0, +\infty[$ , sous réserve que  $Z^* B(r) Z$  est inversible (ce que nous supposons dans la suite), la fonction  $g \rightarrow L(g, r, v; \underline{y})$  est maximum en  $g = g(r)$  où

$$(1.2) \quad g(r) = (Z^* B(r) Z)^{-1} Z^* B(r) \underline{y},$$

la valeur du maximum étant, à une constante près, égale à

$$(1.3) \quad \tilde{L}(g, r, v; \underline{y}) = -\frac{T}{2} \text{Log} v + \frac{1}{2} \text{Log}(1 - r^2) - (2v)^{-1} (\underline{y} - Z g(r))^* B(r) (\underline{y} - Z g(r)).$$

Il convient désormais de maximiser  $\tilde{L}$  par rapport à  $(r, v)$ . D'après (1.3), pour tout  $r$  fixé dans  $] -1, +1[$ , la fonction  $v \rightarrow \tilde{L}(r, v; \underline{y})$  est maximum pour  $v = v(r)$  où

$$(1.4) \quad v(r) = \frac{1}{T} (\underline{y} - Z g(r))^* B(r) (\underline{y} - Z g(r)),$$

la valeur du maximum étant, à une constante près, égale à

$$(1.5) \quad \hat{L}(r; \underline{y}) = -\frac{T}{2} \text{Log} v(r) + \frac{1}{2} \text{Log}(1 - r^2).$$

Il reste donc à maximiser  $\hat{L}$  par rapport à  $r$  dans  $] -1, +1[$ . D'après (1.4) on a

$$(1.6) \quad \dot{v}(r) = \frac{1}{T} (\underline{y} - Z g(r))^* \dot{B}(r) (\underline{y} - Z g(r)) + \frac{2}{T} (-Z \dot{g}(r))^* B(r) (\underline{y} - Z g(r)),$$

où, d'après (1.2)

$$\dot{g}(r) = (Z^* B(r) Z)^{-1} Z^* \dot{B}(r) \underline{y} - (Z^* B(r) Z)^{-1} (Z^* \dot{B}(r) Z) (Z^* B(r) Z)^{-1} Z^* B(r) \underline{y},$$

soit

$$(1.7) \quad \dot{g}(r) = (Z^* B(r) Z)^{-1} Z^* \dot{B}(r) (\underline{y} - Z g(r)).$$

De (1.6) et (1.7) on déduit facilement que

$$(1.8) \quad \dot{v}(r) = \frac{1}{T} (\underline{y} - Z g(r))^* \dot{B}(r) (\underline{y} - Z g(r)).$$

Alors, d'après (1.5) on a

$$(1.9) \quad \frac{\partial}{\partial r} \hat{L}(r; \underline{y}) = -[2 v(r) (1 - r^2)]^{-1} \{2 r v(r) + T (1 - r^2) \dot{v}(r)\},$$

où  $v(r)$  et  $\dot{v}(r)$  sont donnés par (1.4) et (1.8) respectivement.

Par suite,  $\hat{L}$  a un extremum en  $r \in ] -1, +1[$  si et seulement si on a  $k(r) = 0$  où  $k(\cdot)$  est définie par

$$(1.10) \quad k(r) = 2 r v(r) + T (1- r^2) \dot{v}(r).$$

Nous pouvons démontrer

**Lemme 1.** *L'équation  $k(r) = 0$  où  $k(\cdot)$  est définie par (1.10) admet presque sûrement une solution dans l'intervalle  $] -1, +1[$ . Si cette solution est unique et vaut  $\hat{\rho}$ , la fonction de log-vraisemblance  $L$  définie par (1.1) atteint un maximum unique sur le domaine des paramètres  $\mathbb{R}^k \times ] -1, +1[ \times ] 0, +\infty[$  en  $\hat{\gamma}, \hat{\rho}, \hat{v}$  où  $\hat{\gamma} = g(\hat{\rho})$  et  $\hat{v} = v(\hat{\rho})$ ,  $g(\cdot)$  et  $v(\cdot)$  étant définies par (1.2) et (1.4) respectivement.*

**Preuve.** Remarquons d'abord que  $k(-1) = -2v(-1)$  et  $k(+1) = 2v(+1)$  et donc,  $v(r)$  étant par construction positif,

$$k(-1) < 0 \quad \text{et} \quad k(+1) > 0 \text{ p.s..}$$

Alors l'équation  $k(r) = 0$  admet p.s. au moins une solution appartenant à l'intervalle  $] -1, +1[$  et  $\hat{L}(\cdot; \mathbf{y})$  atteint p.s. au moins un extremum dans cet intervalle. Compte tenu de (1.9) le signe de la dérivée de  $\hat{L}(r; \mathbf{y})$  est l'opposé du signe de  $k(\cdot)$  et donc  $\hat{L}(r; \mathbf{y})$  est croissante au voisinage de  $-1$  et décroissante au voisinage de  $+1$ . Ainsi si  $k(r) = 0$  a une seule solution dans  $] -1, +1[$  celle-ci correspond à un maximum (alors unique) de  $\hat{L}(r; \mathbf{y})$ .  $\clubsuit$

## 1.2. Existence, unicité et convergence p.s. de l'estimateur de maximum de vraisemblance

L'équation (1.10) peut être écrite sous la forme  $K_T(r) = 0$  où

$$(1.11) \quad K_T(r) = \frac{r}{T} v_T(r) + \frac{1}{2} (1- r^2) \dot{v}_T(r),$$

avec, d'après (1.4) et (1.8),

$$v_T(r) = \frac{1}{T} \hat{u}_T^*(r) B_T(r) \hat{u}_T(r),$$

(1.12)

$$\dot{v}_T(r) = \frac{1}{T} \hat{u}_T^*(r) \dot{B}_T(r) \hat{u}_T(r),$$

et

$$(1.13) \quad \hat{u}_T(r) = y_T - Z_T g_T(r),$$

$g_T(r)$  étant défini par (1.2). Rappelons la représentation

$$B_T(r) = A_T^0 - r A_T^1 + r^2 A_T^2 \quad ; \quad \dot{B}_T(r) = -A_T^1 + 2r A_T^2,$$

où

$$A_T^0 = \mathbb{1}_T; \quad A_T^1 = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 1 & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & 1 \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}; \quad A_T^2 = \begin{pmatrix} 0 & 0 & \dots & \dots & 0 \\ 0 & 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 & 0 \\ 0 & \dots & \dots & 0 & 0 \end{pmatrix}.$$

Alors (1.11)-(1.12) conduisent à

$$(1.14) \quad K_T(r) = -\frac{(T-1)}{T} q_{2,T}(r) r^3 + \frac{T-2}{2T} q_{1,T}(r) r^2 + [q_{2,T}(r) + \frac{1}{T} q_{0,T}(r)] r - \frac{1}{2} q_{1,T}(r),$$

où

$$(1.15) \quad q_{i,T}(r) = \frac{1}{T} \hat{u}_T^*(r) A_T^i \hat{u}_T(r), \quad i = \overline{0, 2}.$$

Définissons

$$(1.16) \quad q_{i,T} = \frac{1}{T} \hat{u}_T^* A_T^i \hat{u}_T, \quad i = \overline{0, 2}.$$

Nous pouvons alors démontrer

**Lemme 2.** *Supposons que les hypothèses (D) et (A3) sont satisfaites.*

*Alors pour tout  $\lambda > 1$ , si  $q_{i,T}(r)$  et  $q_{i,T}$ ,  $i = \overline{0, 2}$  sont définis respectivement par (1.15) et (1.16) on a*

$$(1.17) \quad \sup_{r \in [-1, +1]} |q_{i,T}(r) - q_{i,T}| = o(T^{-1} (\text{Log} d_{TM})^\lambda) \text{ p.s.}$$

et

$$(1.18) \quad \sup_{r \in [-1, +1]} |q_{i,T}(r)| = o(T^{-1}(\text{Log} d_{TM})^\lambda) \text{ p.s..}$$

**Preuve.** D'après (1.13) on a

$$\hat{u}_T(r) = u_T - Z_T (g_T(r) - \gamma)$$

d'où, par (1.15) et (1.16)

$$\begin{aligned} q_{i,T}(r) - q_{i,T} &= -\frac{2}{T} u_T^* A_T^i Z_T (g_T(r) - \gamma) + \frac{1}{T} (g_T(r) - \gamma)^* Z_T^* A_T^i Z_T (g_T(r) - \gamma) \\ &= -\frac{2}{T} [D_T(g_T(r) - \gamma)]^* \Delta_T^i + \frac{1}{T} [D_T(g_T(r) - \gamma)]^* D_T^{-1} Z_T^* A_T^i Z_T D_T^{-1} [D_T(g_T(r) - \gamma)] \end{aligned}$$

où les  $\Delta_T^i$ ,  $i = \overline{0, 2}$  sont tels que dans la preuve de la proposition II.6. Utilisant le résultat de cette proposition, les éléments de sa preuve concernant l'étude des  $\Delta_T^i$  et le fait que d'après (D) le terme  $D_T^{-1} Z_T^* A_T^i Z_T D_T^{-1}$  converge, on obtient alors immédiatement l'assertion (1.17). Un raisonnement analogue permet de montrer que (1.18) est vraie.  $\clubsuit$

Nous sommes maintenant en mesure de démontrer

**Lemme 3.** *Supposons que les hypothèses (D) et (A3) sont satisfaites et que de plus il existe un  $\varepsilon > 0$  tel que la condition*

$$(E) \quad : \quad T^{-1}(\text{Log} d_{TM})^{1+\varepsilon} = O(1),$$

*est satisfaite. Alors, avec probabilité 1, il existe  $T_0 > 0$  tel que pour tout  $T \geq T_0$  l'équation  $K_T(r) = 0$ , où  $K_T(\cdot)$  est définie par (1.11), admet une solution unique  $\hat{\rho}_T$  dans l'intervalle  $] -1, +1[$  en outre la famille  $\{\hat{\rho}_T ; T \geq 1\}$  converge presque sûrement vers  $\rho$  et la famille  $\{\hat{v}_T ; T \geq 1\}$ , où  $\hat{v}_T = v_T(\hat{\rho}_T)$ ,  $v_T$  étant définie par (1.4), converge p.s. vers  $v$ .*

**Preuve.** En vertu de (1.16), appliquant le lemme II.2 en faisant  $r = 0$  dans II.(2.13)-(2.14)-(2.15) on obtient que p.s.

$$(1.19) \quad \lim_{T \rightarrow +\infty} q_{0,T} = \frac{\nu}{(1-\rho^2)} ; \quad \lim_{T \rightarrow +\infty} q_{1,T} = \frac{2\nu\rho}{(1-\rho^2)} ; \quad \lim_{T \rightarrow +\infty} q_{2,T} = \frac{\nu}{(1-\rho^2)} .$$

Alors, compte tenu de (E), l'assertion (1.17) du lemme 2 fournit, à partir de (1.14) et (1.19), que

$$(1.20) \quad \sup_{r \in [-1, +1]} |K_T(r) - K_\infty(r)| = o(1) \text{ p.s.}$$

où

$$K_\infty(r) = \frac{\nu}{(1-\rho^2)} (1-r^2)(r-\rho).$$

De même, à partir de (1.18) on peut voir que

$$(1.21) \quad \sup_{r \in [-1, +1]} |\dot{K}_T(r) - \dot{K}_\infty(r)| = o(1) \text{ p.s.,}$$

où

$$\dot{K}_\infty(r) = \frac{\nu}{(1-\rho^2)} [-3r^2 + 2\rho r + 1].$$

L'étude de  $K_\infty(\cdot)$  et de sa dérivée  $\dot{K}_\infty(\cdot)$  montre que pour  $\zeta > 0$  et  $\eta > 0$  suffisamment petits on a  $K_\infty(\cdot) \in \mathcal{F}_{\zeta,\eta}$  où  $\mathcal{F}_{\zeta,\eta}$  est l'ensemble des fonctions continûment dérivables sur  $[-1, +1]$  telles que

$$f(r) < 0, -1 + \zeta \leq r \leq \rho - \eta ; f(r) > 0, \rho + \eta \leq r \leq 1 - \zeta .$$

et

$$\dot{f}(r) < 0, -1 \leq r \leq -1 + \zeta ; \dot{f}(r) > 0, \rho - \eta \leq r \leq \rho + \eta ; \dot{f}(r) < 0, 1 - \zeta \leq r \leq 1.$$

Alors, d'après (1.20) et (1.21),

$$\text{Prob} \left( \bigcup_{T \geq 1} \bigcap_{T \geq T_0} [K_T(\cdot) \in \mathcal{F}_{\zeta,\eta}] \right) = 1$$

Pour obtenir les deux premières assertions du Lemme il suffit de montrer que sur l'événement A concerné ci-dessus il existe un  $T_0$  tel que p.s. pour tout  $T \geq T_0$ , on a

$$(1.22) \quad K_T(r) < 0, -1 \leq r \leq -1 + \zeta \quad ; \quad K_T(r) > 0, 1 - \zeta \leq r \leq 1.$$

Dans la mesure où  $K_T(-1) < 0$  p.s. et  $K_T(+1) > 0$  p.s., (1.22) est évidemment vraie puisque pour  $r \in [-1, -1 + \zeta]$  on a

$$K_T(r) = K_T(-1) + (r + 1) \dot{K}_T(\tilde{r}) \text{ avec } \tilde{r} \in [-1, r]$$

et pour  $r \in [1 - \zeta, 1]$  on a

$$K_T(r) = K_T(1) + (r - 1) \dot{K}_T(\tilde{r}) \text{ avec } \tilde{r} \in [r, 1].$$

Examinons maintenant la convergence de  $\hat{v}_T = v_T(\hat{\rho}_T)$ .

Compte tenu de (1.12) et de la décomposition de  $B_T(r)$ , on a

$$v_T(r) = q_{0,T}(r) - r q_{1,T}(r) + r^2 q_{2,T}(r),$$

les  $q_{i,T}(\cdot)$  étant donnés par (1.15). Alors, d'après (1.17) et (1.18) et (E) on a

$$(1.23) \quad \sup_{r \in [-1, +1]} |v_T(r) - \hat{v}_T(r)| = o(1) \text{ p.s.}$$

où

$$\hat{v}_T(r) = q_{0,T} - r q_{1,T} + r^2 q_{2,T},$$

les  $q_{i,T}$  étant donnés par (1.16). Mais, en vertu de (1.19), pour

$$v_\infty(r) = \frac{v}{(1-\rho^2)} (1 - 2r\rho + r^2)$$

on a

$$(1.24) \quad \sup_{r \in [-1, +1]} |\hat{v}_T(r) - v_\infty(r)| = o(1) \text{ p.s..}$$

De plus, comme  $\hat{\rho}_T$  converge p.s. vers  $\rho$  on a

$$(1.25) \quad \lim_{T \rightarrow +\infty} v_{\infty}(\hat{\rho}_T) = v_{\infty}(\rho) = v \text{ p.s.}$$

En définitive (1.23)-(1.24) et (1.25) assurent bien que  $\hat{v}_T = v_T(\hat{\rho}_T)$  converge p.s. vers  $v$ .

Nous pouvons maintenant énoncer

**Proposition 1.** *Supposons que les hypothèses (D) et (A3) sont satisfaites et que de plus il existe un  $\varepsilon > 0$  tel que la condition*

$$(E^*) : (d_{Tm}^2 \wedge T)^{-1} (\text{Log} d_{TM})^{1+\varepsilon} = O(1).$$

Alors, avec une probabilité qui tend vers 1, pour  $T$  suffisamment grand la fonction de Log-vraisemblance  $L$  définie par (1.1) atteint un maximum unique sur le domaine

$\mathbb{R}^k \times ]-1, +1[ \times ]0, +\infty[$  en  $(\hat{\gamma}_T, \hat{\rho}_T, \hat{v}_T)$  où  $\hat{\gamma}_T = g_T(\hat{\rho}_T)$ ,  $\hat{v}_T = v_T(\hat{\rho}_T)$ ,  $g_T(\cdot)$  et  $v_T(\cdot)$  étant définies par (1.2) et (1.4) respectivement et  $\hat{\rho}_T$  étant l'unique solution dans  $]-1, +1[$  de l'équation  $K_T(r) = 0$  où  $K_T(\cdot)$  est définie par (1.11). En outre l'estimateur  $(\hat{\gamma}_T, \hat{\rho}_T, \hat{v}_T)$  est p.s. convergent.

**Preuve.** Dans la mesure où la condition (E\*) assure en particulier que (E) est satisfaite, la première partie de l'énoncé résulte immédiatement des assertions des lemmes 1 et 3. De

même le Lemme 3 assure que  $(\hat{\rho}_T, \hat{v}_T)$  est p.s. convergent. Il reste à montrer que  $\hat{\gamma}_T$  est p.s. convergent.

Comme  $\hat{\gamma}_T = g_T(\hat{\rho}_T)$  et, qu'en vertu de II.(3.23), on a

$$\sup_{r \in [-1, +1]} |g_T(r) - \gamma| = o\left(\frac{(\text{Log}(d_{TM})^{1+\varepsilon})}{d_{TM}^2}\right) \text{ p.s.,}$$

il est clair que  $\hat{\gamma}_T$  converge p.s. vers  $\gamma$  dans la mesure où  $(E^*)$  assure que

$$\frac{(\text{Log}(d_{TM})^{1+\varepsilon})}{d_{TM}^2} = O(1).$$

### 1.3. Vitesse de convergence et normalité asymptotique de l'estimateur de maximum de vraisemblance

Nous nous intéressons d'abord aux dérivées partielles premières et secondes de la fonction

de Log-vraisemblance  $L_T$  définie par (1.1). Soit  $L_T^{(1)} = (\text{grad}_g^* L_T, \frac{\partial}{\partial \rho} L_T, \frac{\partial}{\partial v} L_T)^*$ .

On a

$$\text{grad}_g L_T = v^{-1} Z^* B(r) (y - Z g)$$

$$\frac{\partial}{\partial r} L_T = \frac{-r}{(1-r^2)} - (2v)^{-1} (y - Z g(r))^* \dot{B}(r) (y - Z g(r))$$

$$\frac{\partial}{\partial v} L_T = -\frac{T}{2v} + \frac{v^{-2}}{2} (y - Z g(r))^* B(r) (y - Z g(r))$$

Soit  $L_T^{(2)} = \begin{pmatrix} I_T & K_T \\ K_T^* & J_T \end{pmatrix}$ , où

$$I_T = \left( \left( \frac{\partial^2}{\partial g_i \partial g_j} L_T \right) \right); \quad J_T = \begin{pmatrix} \frac{\partial^2}{\partial r^2} L_T & \frac{\partial^2}{\partial r \partial v} L_T \\ \frac{\partial^2}{\partial v \partial r} L_T & \frac{\partial^2}{\partial v^2} L_T \end{pmatrix}; \quad K_T = (\text{grad}_g \left( \frac{\partial}{\partial r} L_T \right), \text{grad}_g \left( \frac{\partial}{\partial v} L_T \right))$$

avec

$$I_T = -v^{-1} Z^* B(r) Z$$

$$\frac{\partial^2}{\partial r^2} L_T = -\frac{(1+r^2)}{(1-r^2)^2} - (2v)^{-1} (\underline{y} - Z g)^* \ddot{B}(r) (\underline{y} - Z g)$$

$$\frac{\partial^2}{\partial v^2} L_T = \frac{T}{2v^2} - v^{-3} (\underline{y} - Z g)^* B(r) (\underline{y} - Z g)$$

$$\frac{\partial^2}{\partial r \partial v} L_T = \frac{\partial^2}{\partial v \partial r} L_T = \frac{v^{-2}}{2} (\underline{y} - Z g(r))^* \dot{B}(r) (\underline{y} - Z g(r))$$

$$\text{grad}_g \left( \frac{\partial}{\partial r} L_T \right) = v^{-1} Z^* \dot{B}(r) (\underline{y} - Z g)$$

$$\text{grad}_g \left( \frac{\partial}{\partial v} L_T \right) = -v^{-2} Z^* B(r) (\underline{y} - Z g)$$

Le comportement asymptotique de  $L_T^{(1)}(\underline{\gamma}, \rho, v)$  et  $L_T^{(2)}(\underline{\gamma}, \rho, v)$  est décrit dans le

**Lemme 4.** Sous les hypothèses (A1), ((A2( $\tau$ )),  $\tau = 0, 1$ ), (A3) et (E), définissant la matrice  $(k+2) \times (k+2)$   $\tilde{D}_T$  par

$$\tilde{D}_T = \begin{pmatrix} D_T & 0 \\ 0 & \sqrt{T} \mathbb{I}_2 \end{pmatrix},$$

pour  $\theta = (\underline{\gamma}^*, \rho, v)^*$ , on a p.s.

$$(1.26) \quad \lim_{T \rightarrow +\infty} \tilde{D}_T^{-1} L_T^{(2)}(\theta) \tilde{D}_T^{-1} = -\Gamma(\theta) = - \begin{pmatrix} v^{-1} H(\rho) & 0 & 0 \\ 0 & (1-\rho^2)^{-1} & 0 \\ 0 & 0 & (2v^2)^{-1} \end{pmatrix}$$

où  $H(\rho)$  est la matrice définie par II.(3.11).

En outre la famille  $(\tilde{D}_T^{-1} L_T^{(1)}(\theta); T \geq 1)$  est asymptotiquement gaussienne centrée de matrice de covariance  $\Gamma(\theta)$  donnée par (1.26).

**Preuve** .On a

$$\tilde{D}_T^{-1} I_T(\theta) \tilde{D}_T^{-1} = -v^{-1} D_T^{-1} Z_T^* B_T(\rho) Z_T D_T^{-1},$$

d'où, d'après le lemme II.4,  $D_T^{-1} I_T(\theta) D_T^{-1} \rightarrow -v^{-1} H(\rho)$  p.s..

On a

$$T^{-1} \frac{\partial^2}{\partial r^2} L_T(\theta) = -T^{-1} \frac{1 + \rho^2}{(1 - \rho^2)^2} - (2Tv)^{-1} \underline{\mu}_T^* \ddot{B}_T(\rho) \underline{\mu}_T,$$

d'où, d'après II.(2.15),  $T^{-1} \frac{\partial^2}{\partial r^2} L_T(\theta) \rightarrow -(1 - \rho^2)^{-1}$  p.s..

On a

$$T^{-1} \frac{\partial^2}{\partial v^2} L_T(\theta) = (2v^2)^{-1} - (Tv^3)^{-1} \underline{\mu}_T^* B_T(\rho) \underline{\mu}_T,$$

d'où, d'après II.(2.13),  $T^{-1} \frac{\partial^2}{\partial v^2} L_T(\theta) \rightarrow -(2v^2)^{-1}$  p.s..

On a

$$T^{-1} \frac{\partial^2}{\partial r \partial v} L_T(\theta) = (2Tv^2)^{-1} \underline{\mu}_T^* \dot{B}_T(\rho) \underline{\mu}_T,$$

d'où, d'après II.(2.14),  $T^{-1} \frac{\partial^2}{\partial r \partial v} L_T(\theta) \rightarrow 0$  p.s..

On a

$$T^{-1/2} D_T^{-1} \text{grad}_g \left( \frac{\partial}{\partial r} L_T(\theta) \right) = \left[ \frac{(\text{Log} d_{TM})^{1+\epsilon}}{T} \right]^{1/2} v^{-1} (\text{Log} d_{TM})^{-(1+\epsilon)/2} D_T^{-1} Z_T^* \dot{B}_T(\rho) \underline{\mu}_T$$

d'où, en vertu de (A3), (E) et de la preuve de la proposition II.6,

$$T^{-1/2} D_T^{-1} \text{grad}_g \left( \frac{\partial}{\partial r} L_T(\theta) \right) \rightarrow 0 \text{ p.s..}$$

De même  $T^{-1/2} D_T^{-1} \text{grad}_g \left( \frac{\partial}{\partial v} L_T(\theta) \right) \rightarrow 0$  p.s..

Cela achève la démonstration de (1.26).

Montrons maintenant la dernière assertion de l'énoncé.

On peut écrire

$$L_T^{(1)}(\theta) = \begin{pmatrix} X_T^1 \\ X_T^2 \\ X_T^3 \end{pmatrix} = \begin{pmatrix} v^{-1} Z_T^* B_T(\rho) \underline{y}_T \\ -(2v)^{-1} \underline{y}_T^* B_T(\rho) \underline{y}_T - \rho(1-\rho^2)^{-1} \\ (2v^2)^{-1} \underline{y}_T^* B_T(\rho) \underline{y}_T - (2v)^{-1} T \end{pmatrix}$$

En fait

$$X_T^1 = v^{-1} Z_T^* M_T^*(\rho) \underline{y}_T$$

et donc (cf. la preuve du lemme II.3)  $(X_T^1, T \geq 1)$  est une martingale (gaussienne) de processus croissant

$$\langle X^1 \rangle_T = v^{-1} Z_T^* B_T(\rho) Z_T, T \geq 1.$$

De même on peut écrire

$$X_T^2 = v^{-1} \left[ \rho(1-\rho^2)^{-1} (v_1^2 - v) + \sum_2^T u_{t-1} v_t \right]$$

et donc  $(X_T^2, T \geq 1)$  est une martingale de processus croissant

$$\langle X^2 \rangle_T = v^{-2} \left[ 2v\rho^2(1-\rho)^{-2} (v_1^2 - v) + v \sum_2^T u_{t-1}^2 \right].$$

Enfin on peut écrire

$$X_T^3 = (2v^2)^{-1} \underline{y}_T^* v_T - (2v)^{-1} T = (2v^2)^{-1} \sum_{t=1}^T (v_t^2 - v)$$

et donc  $(X_T^3, T \geq 1)$  est une martingale de processus croissant

$$\langle X^3 \rangle_T = (2v^2)^{-1} T.$$

Ainsi  $(L_T^{(1)}(\theta), T \geq 1)$  est une martingale dans  $\mathbb{R}^{k+2}$ .

Compte tenu de ce que  $E(v_t(v_t^2 - v)) = 0$  et  $E u_t = 0$  on peut voir que les variations quadratiques croisées sont

$$\begin{aligned} \langle X^1, X^2 \rangle_T &= v^{-1} \sum_2^T (z_{t, \cdot} - \rho z_{t-1, \cdot}) u_{t-1}, \\ \langle X^1, X^3 \rangle_T &= 0 ; \quad \langle X^2, X^3 \rangle_T = v^{-1} \rho (1 - \rho^2)^{-1}. \end{aligned}$$

Alors, raisonnant comme dans l'étude de  $\tilde{D}_T^{-1} L_T^{(2)}(\theta) \tilde{D}_T^{-1}$  faite ci-dessus, on obtient que le processus croissant  $(L^{(1)}(\theta))_T$  de  $L^{(1)}(\theta)$  est tel que

$$(1.27) \quad \lim_{T \rightarrow +\infty} \tilde{D}_T^{-1} \langle L^{(1)}(\theta) \rangle_T \tilde{D}_T^{-1} = \Gamma(\theta) \text{ p.s.}$$

où  $\Gamma(\theta)$  est définie dans (1.26).

D'après ce qui précède nous pouvons écrire  $L^{(1)}(\theta)$  comme martingale transformée

de la martingale bidimensionnelle  $S = \left( \sum_1^T \begin{pmatrix} v_t \\ v_t^2 - v \end{pmatrix}; T \geq 0 \right)$  relativement à la filtration  $(\mathcal{F}_t = \sigma(v_s, s \leq t), t \geq 1)$ , sous la forme

$$L_T^{(1)}(\theta) = \sum_1^T \xi_{t-1} \Delta S_t, \quad T \geq 1,$$

où les matrices  $(k+2) \times 2$   $\xi_t$  peuvent être identifiées dans les représentations ci-dessus des  $X_T^i$ . Alors, étant donné  $l$  un vecteur dans  $\mathbb{R}^{k+2}$ , on peut écrire

$$l^* \tilde{D}_T^{-1} L_T^{(1)}(\theta) = \sum_1^T (l^* \tilde{D}_T^{-1} \xi_{t-1}) \Delta S_t,$$

où, puisque  $\tilde{D}_T$  est déterministe, la famille  $(\mathfrak{X}_{T,t} = I^* \tilde{D}_T^{-1} \mathfrak{X}_{t-1} \Delta S_t, 1 \leq t \leq T, T \geq 1)$  est une suite triangulaire relativement aux filtrations  $((\mathcal{F}_t^T = \mathcal{F}_t; 1 \leq t \leq T), T \geq 1)$ .

Montrons d'abord que la famille  $(\mathfrak{X}_{T,t})$  est asymptotiquement négligeable au sens que pour tout  $\alpha > 0$  on a

$$(1.28) \quad \lim_{T \rightarrow +\infty} \sum_1^T P(|\mathfrak{X}_{T,t}| \geq \alpha / \mathcal{F}_{t-1}^T) = 0 \text{ en probabilité.}$$

Pour ce faire, notant  $l = (l_1, \dots, l_{k+2})^*$ ,  $\tilde{d}_{Tj}$  le  $j^{\text{eme}}$  terme diagonal de la matrice diagonale  $\tilde{D}_T$ ,  $\mathfrak{X}_t = ((\mathfrak{X}_t^{jm}))$  et  $S_t = (S_t^1, S_t^2)^*$ , il suffit de montrer que pour tout  $j = \overline{1, k+2}$  et  $m = \overline{1, 2}$  on a, pour un  $\beta > 0$ ,

$$\lim_{T \rightarrow +\infty} \sum_1^T E(|\mathfrak{X}_{T,t}^{jm}|^\beta / \mathcal{F}_{t-1}) = 0 \text{ en probabilité,}$$

pour

$$\mathfrak{X}_{T,t}^{jm} = l_j \tilde{d}_{Tj}^{-1} \mathfrak{X}_{t-1}^{jm} \Delta S_t^m.$$

Or

$$E(|\mathfrak{X}_{T,t}^{jm}|^\beta / \mathcal{F}_{t-1}) = |l_j|^\beta \tilde{d}_{Tj}^{-\beta} |\mathfrak{X}_{t-1}^{jm}|^\beta E|\Delta S_t^m|^\beta,$$

et donc, comme  $E|\Delta S_t^m|^\beta = E|\Delta S_1^m|^\beta < \infty$  pour tout  $\beta > 0$  il suffit de vérifier qu'il existe  $\beta > 0$  tel que

$$(1.29) \quad \lim_{T \rightarrow +\infty} \tilde{d}_{Tj}^{-\beta} \sum_1^T |\mathfrak{X}_{T,t}^{jm}|^\beta = 0 \text{ en probabilité.}$$

Mais, on peut voir que pour  $m = \overline{1, 2}$

$$|\mathfrak{X}_{T,t}^{jm}| \leq \text{Cte} (|z_{tj}| + |z_{t-1,j}|); j = \overline{1, k}, t \geq 2,$$

$$|\mathfrak{X}_{T,t-1}^{k+1,m}| \leq \text{Cte} |u_{t-1}|; t \geq 2,$$

$$|\mathfrak{X}_{t-1}^{k+2,m}| \leq \text{Cte}; t \geq 1.$$

Alors, pour  $j = \overline{1, k}$ , comme  $\tilde{d}_{Tj} = d_{Tj} = \left(\sum_1^T z_{t,j}^2\right)^{1/2}$ , choisissant  $\beta = 3$  on a

$$\tilde{d}_{Tj}^{-3} \sum_1^T \left| \xi_{t-1}^{j,m} \right|^3 \leq \text{Cte } d_{Tj}^{-1} \left( \text{Max}_{s \leq t} |z_{s,j}| \right)$$

d'où (1.29) en vertu de (A1).

De même, comme  $\tilde{d}_{k+1,j} = T^{-1/2}$ , choisissant  $\beta = 3$  on a

$$\tilde{d}_{T,k+1}^{-3} \sum_1^T \left| \xi_{t-1}^{k+1,m} \right|^3 \leq \text{Cte } T^{-3/2} \sum_1^T |u_{t-1}|^3$$

$$\tilde{d}_{T,k+2}^{-3} \sum_1^T \left| \xi_{t-1}^{k+2,m} \right|^3 \leq \text{Cte } T^{-1/2}$$

d'où (1.29) pour  $j = k + 1, k + 2$  car  $E|u_{t-1}|^3 = E|u_1|^3 < \infty$ .

Puisque  $E(\xi_{T,t} / \mathcal{F}_{t-1}) = l^* \tilde{D}_T^{-1} \xi_{t-1} E(\Delta S_t / \mathcal{F}_{t-1}) = 0$  p.s. on a bien sûr

**(1.30)**  $\lim_{T \rightarrow +\infty} E(\xi_{T,t} / \mathcal{F}_{t-1}) = 0$  en probabilité.

D'autre part on a

$$\begin{aligned} \sum_1^T E(\xi_{T,t}^2 / \mathcal{F}_{t-1}) &= l^* \tilde{D}_T^{-1} \left\{ \sum_1^T \xi_{t-1}^* E(\Delta S_t \Delta S_t^*) \xi_{t-1} \right\} \tilde{D}_T^{-1} l \\ &= l^* \tilde{D}_T^{-1} \left\langle L^{(1)}(\theta) \right\rangle_T \tilde{D}_T^{-1} l \end{aligned}$$

et donc, d'après (1.27),

**(1.31)**  $\lim_{T \rightarrow +\infty} E(\xi_{T,t}^2 / \mathcal{F}_{t-1}) = l^* \Gamma(\theta) l$  en probabilité.

En définitive, compte tenu de (1.28), (1.30) et (1.31), appliquant le théorème de la limite centrale pour les suites triangulaires (cf. DACUNHA-CASTELLE et DUFLO (1983), théorème 2.8.42), on obtient que la suite  $(l^* \tilde{D}_T^{-1} L_T^{(1)}(\theta); T \geq 1)$  est asymptotiquement gaussienne centrée de variance  $l^* \Gamma(\theta) l$ . Cela étant vrai pour  $l \in \mathbb{R}^{k+2}$ , l'assertion du lemme est démontée.  $\blacksquare$

Nous pouvons maintenant démontrer

**Proposition 2.** *Supposons que les hypothèses (D), (A3) et (E\*) sont satisfaites. Alors si  $\hat{\theta}_T = (\hat{\gamma}_T, \hat{\rho}_T, \hat{v}_T)$  est l'estimateur de maximum de vraisemblance de  $\theta$ , la suite  $(\tilde{D}_T (\hat{\theta}_T - \theta); T \geq 1)$  est asymptotiquement gaussienne centrée de matrice de covariance  $\Gamma^{-1}(\theta)$  où  $\Gamma(\theta)$  est donnée par (1.26). De plus la suite  $(2 [L_T(\hat{\theta}_T) - L_T(\theta)]; T \geq 1)$  est asymptotiquement de loi du Khi deux à  $(k+2)$  degrés de liberté.*

**Preuve.** Ecrivons

$$(1.32) \quad \tilde{D}_T^{-1} L_T^{(1)}(\theta) = \tilde{D}_T^{-1} L_T^{(1)}(\hat{\theta}_T) + \tilde{D}_T^{-1} ((L_{T,ij}^{(2)}(\tilde{\theta}_T(i))) \tilde{D}_T^{-1} \tilde{D}_T (\theta - \hat{\theta}_T))$$

où  $\tilde{\theta}_T(i) = \hat{\theta}_T + \alpha_i(\hat{\theta}_T) (\hat{\theta}_T - \theta)$  avec  $\alpha_i(\hat{\theta}_T) \in [0, 1]$ . D'après la Proposition 1 on a

$\lim_{T \rightarrow +\infty} P([L_T^{(1)}(\hat{\theta}_T) = 0]) = 1$ , et  $\lim_{T \rightarrow +\infty} \tilde{\theta}_T(i) = \theta$  p.s.. Une démonstration analogue à celle

de la première partie du Lemme 4 permet de s'assurer qu'alors on a aussi

$$\lim_{T \rightarrow +\infty} \tilde{D}_T^{-1} L_T^{(2)}(\tilde{\theta}_T) \tilde{D}_T^{-1} = -\Gamma(\theta) \text{ p.s.,}$$

où par convention  $L_T^{(2)}(\tilde{\theta}_T) = ((L_{T,ij}^{(2)}(\tilde{\theta}_T(i)))$ .

Ainsi, compte tenu de la dernière assertion du Lemme 4 et du fait que, d'après (D), la matrice  $\Gamma(\theta)$  est régulière, au vu de (1.32) on obtient la première assertion de l'énoncé.

Ecrivons maintenant, en conservant les notations ci-dessus,

$$L_T(\theta) - L_T(\hat{\theta}_T) = L_T^{(1)}(\hat{\theta}_T) (\theta - \hat{\theta}_T) + (\theta - \hat{\theta}_T) * L_T^{(2)}(\tilde{\theta}_T) (\theta - \hat{\theta}_T)$$

D'après la discussion ci-dessus on a alors p.s.

$$\lim_{T \rightarrow +\infty} \{2 [L_T(\hat{\theta}_T) - L_T(\theta)] - (\theta - \hat{\theta}_T) * \tilde{D}_T \Gamma(\theta) \tilde{D}_T (\theta - \hat{\theta}_T)\} = 0.$$

Ainsi comme  $\tilde{D}_T(\theta - \hat{\theta}_T)$  converge en loi vers la loi  $N_{k+2}(0, \Gamma^{-1}(\theta))$  la dernière assertion de l'énoncé est acquise.  $\blacksquare$

Nous précisons enfin la vitesse de convergence p.s. de l'estimateur de maximum de vraisemblance

**Proposition 3.** *Supposons que les hypothèses (D), (A3) et (E\*) sont satisfaites. Alors si  $\hat{\theta}_T = (\hat{\gamma}_T, \hat{\rho}_T, \hat{v}_T)$  est l'estimateur de maximum de vraisemblance de  $\theta$  on a*

$$(\text{Logd}_{TM})^{-(1+\varepsilon)/2} d_{Tj}(\hat{\gamma}_T^j - \gamma_j) = 0 \text{ p.s. ; } j = \overline{1, k}$$

et, pour  $\beta < \varepsilon / 2(1 + \varepsilon)$ ,

$$T^\beta(\hat{\rho}_T - \rho) = o(1) \text{ p.s. ; } T^\beta(\hat{v}_T - v) = o(1) \text{ p.s..}$$

**Preuve.** De manière analogue à (1.32) écrivons, en convenant encore que  $L_T^{(2)}(\tilde{\theta}_T)$  désigne la matrice  $((L_{T,ij}^{(2)}(\tilde{\theta}_T(i))))$ ,

$$(1.33) \quad \tilde{R}_T^{-1} \tilde{D}_T^{-1} L_T^{(1)}(\theta) = \tilde{R}_T^{-1} \tilde{D}_T^{-1} L_T^{(1)}(\hat{\theta}_T) + \tilde{R}_T^{-1} \tilde{D}_T^{-1} L_T^{(2)}(\tilde{\theta}_T) \tilde{D}_T^{-1} \tilde{R}_T \tilde{R}_T^{-1} \tilde{D}_T(\theta - \hat{\theta}_T)$$

où  $\tilde{R}_T$  est la matrice  $(k+2) \times (k+2)$  diagonale définie par

$$\tilde{R}_T = \begin{pmatrix} R_T & 0 \\ 0 & T^\alpha \mathbb{1}_2 \end{pmatrix} ; \quad R_T = (\text{Logd}_{TM})^{(1+\varepsilon)/2} \mathbb{1}_k ; \quad \alpha > 0.$$

On a, avec les notations de la preuve du Lemme 4,

$$R_T^{-1} D_T^{-1} \text{grad}_g L_T(\theta) = v^{-1} (\text{Logd}_{TM})^{-(1+\varepsilon)/2} D_T^{-1} Z_T^* B_T(\rho) \underline{1}_T$$

$$T^{-(\alpha+1/2)} \frac{\partial L_T(\theta)}{\partial r} = - T^{-(\alpha+1/2)} X_T^2$$

$$T^{-(\alpha+1/2)} \frac{\partial L_T(\theta)}{\partial v} = - T^{-(\alpha+1/2)} X_T^3.$$

D'après la preuve de la Proposition II.6 et le fait que  $T^{-1} \langle X^i \rangle_T$  tend vers une limite strictement positive,  $i = \overline{2, 3}$ , on a donc

$$(1.34) \quad \lim_{T \rightarrow +\infty} \tilde{R}_T^{-1} \tilde{D}_T^{-1} L_T^{(1)}(\theta) = 0 \text{ p.s..}$$

Pour les blocs diagonaux de  $\tilde{R}_T^{-1} \tilde{D}_T^{-1} L_T^{(2)}(\theta) \tilde{D}_T^{-1} \tilde{R}_T$  on a, d'après (1.26),

$$R_T^{-1} D_T^{-1} I_T(\theta) D_T^{-1} R_T = D_T^{-1} I_T(\theta) D_T^{-1} \rightarrow -v^{-1} H(\rho) \text{ p.s.}$$

$$T^{-\alpha} \frac{1}{T} J_T(\theta) T^\alpha = \frac{1}{T} J_T(\theta) \rightarrow \begin{pmatrix} (1 - \rho^2)^{-1} & 0 \\ 0 & (2v^2)^{-1} \end{pmatrix} \text{ p.s..}$$

Examinons maintenant les blocs non diagonaux de  $\tilde{R}_T^{-1} \tilde{D}_T^{-1} L_T^{(2)}(\theta) \tilde{D}_T^{-1} \tilde{R}_T$ . Ils contiennent des vecteurs du type

$$T^{\alpha - 1/2} (\text{Logd}_{TM})^{-(1+\epsilon)/2} D_T^{-1} \text{grad}_g \left( \frac{\partial L_T(\theta)}{\partial r} \right)$$

et

$$T^{-(\alpha + 1/2)} (\text{Logd}_{TM})^{(1+\epsilon)/2} D_T^{-1} \text{grad}_g \left( \frac{\partial L_T(\theta)}{\partial \alpha} \right).$$

Les premiers tendent vers zéro p.s. lorsque  $\alpha \leq 1/2$ ; quant aux seconds, puisque

$$\frac{(\text{Logd}_{TM})^{(1+\epsilon)/2}}{T^{(\alpha + 1/2)}} = \left[ \frac{(\text{Logd}_{TM})^{1+\epsilon}}{T} \right]^{\alpha + 1/2} \times (\text{Logd}_{TM})^{-\alpha(1+\epsilon)}, \text{ ils tendent vers zéros p.s.}$$

lorsque  $\alpha(1 + \epsilon) > 1/2$ .

Dans ces conditions on a donc

$$\tilde{R}_T^{-1} \tilde{D}_T^{-1} L_T^{(2)}(\theta) \tilde{D}_T^{-1} \tilde{R}_T \rightarrow -\Gamma(\theta) \text{ p.s.}$$

où  $\Gamma(\theta)$  est donnée par (1.26). Une démonstration analogue permet de voir que comme

$\tilde{\theta}_T \rightarrow \theta$  p.s. on a aussi

$$(1.35) \quad \tilde{R}_T^{-1} \tilde{D}_T^{-1} L_T^{(2)}(\tilde{\theta}_T) \tilde{D}_T^{-1} \tilde{R}_T \rightarrow -\Gamma(\theta) \text{ p.s..}$$

Alors, d'après (1.34)-(1.35), le fait que  $P([L_T^{(1)}(\hat{\theta}_T) = 0]) \rightarrow 1$  et l'inversibilité de  $\Gamma(\theta)$ , l'égalité (1.33) assure que

$$\tilde{R}_T^{-1} \tilde{D}_T^{-1} (\hat{\theta}_T - \theta) \rightarrow 0 \text{ p.s..}$$

On obtient ainsi les assertions de la Proposition.  $\clubsuit$

**Remarque.** Il convient de noter qu'en fait les résultats des propositions 2 et 3 restent valables si on substitue l'hypothèse (D(p)) à l'hypothèse (D).

## §2. Estimation de moindres carrés

En pratique l'estimateur de maximum de vraisemblance exact est difficilement calculable et on peut souhaiter disposer de procédures plus explicites d'estimation des paramètres  $\gamma$ ,  $\rho$  et  $v$ . Nous avons déjà annoncé au §II.3.2.b. qu'il semble raisonnable d'envisager la procédure consistant à estimer d'abord  $\gamma$  par moindres carrés i.e. par

$$(2.1) \quad \tilde{\gamma}_T = (Z_T^* Z_T)^{-1} Z_T^* y_T,$$

(dont on sait, par le Lemme II.5, qu'il s'agit, sous (D(0)) et (A3), d'un estimateur presque sûrement convergent) puis à estimer  $\rho$  et  $v$ , comme on le ferait si  $\gamma$  était connu égal à  $\tilde{\gamma}_T$ , par moindres carrés, à savoir (cf.II.§.2.2) par

$$(2.2) \quad \tilde{\rho}_T = \left\{ \sum_{i=2}^T (y_i - \tilde{\gamma}_T^* z_i) (y_{i-1} - \tilde{\gamma}_T^* z_{i-1}) \right\} / \left\{ \sum_{i=1}^{T-1} (y_i - \tilde{\gamma}_T^* z_i)^2 \right\},$$

et

$$(2.3) \quad \tilde{v}_T = \frac{1}{T-1} \sum_{i=2}^T ((y_i - \tilde{\gamma}_T^* z_i) - \tilde{\rho}_T (y_{i-1} - \tilde{\gamma}_T^* z_{i-1}))^2.$$

L'estimateur  $\tilde{\theta}_T = (\tilde{\gamma}_T^*, \tilde{\rho}_T, \tilde{v}_T)^*$  de  $\theta = (\gamma^*, \rho, v)^*$  défini par (2.1)-(2.3) est dans la suite appelé "estimateur des moindres carrés".

### 2.1. Convergence et vitesse de convergence presque sûre de l'estimateur des moindres carrés.

Comme  $\tilde{\gamma}_T$  défini par (2.1) est en fait donné par

$$(2.4) \quad \tilde{\gamma}_T = g_T(0),$$

où  $g_T(\cdot)$  est fourni par (3.19), d'après les considérations du II§3.2.c., nous savons que, si les conditions (D(0)) et (A3) sont vérifiées, alors, pour tout  $\eta > 1/2$ , on a

$$(2.5) \quad \tilde{\gamma}_T^j - \gamma_j = o\left\{\left(d_{Tj}^{-1} (\text{Log} d_{Tj})^\eta\right)\right\} \text{ p.s..}$$

Vu (2.4), avec les notations du §1.2 ci-dessus (cf. (1.13) et (1.15)), nous pouvons écrire

$$\begin{aligned} \sum_{t=1}^{T-1} (y_t - \tilde{\gamma}_T^* z_t)^2 &= \hat{u}_T^*(0) \hat{u}_T^*(0) - (y_T - \tilde{\gamma}_T^* z_T)^2 \\ &= T q_{0,T}(0) - (y_T - \tilde{\gamma}_T^* z_T)^2 \end{aligned}$$

et

$$\begin{aligned} \sum_{t=2}^T (y_t - \tilde{\gamma}_T^* z_t)^2 &= \hat{u}_T^*(0) \hat{u}_T^*(0) - (y_1 - \tilde{\gamma}_T^* z_1)^2 \\ &= T q_{0,T}(0) + (y_1 - \tilde{\gamma}_T^* z_1)^2. \end{aligned}$$

De même on a

$$\begin{aligned} \sum_{t=2}^T (y_t - \tilde{\gamma}_T^* z_t) (y_{t-1} - \tilde{\gamma}_T^* z_{t-1}) &= \frac{1}{2} \hat{u}_T^*(0) A_T^1 \hat{u}_T^*(0) \\ &= \frac{T}{2} q_{1,T}(0). \end{aligned}$$

Par suite, compte tenu de (2.2) et (2.3) on a

$$(2.2)' \quad \tilde{\rho}_T = \frac{q_{1,T}(0)}{2\{q_{0,T}(0) - T^{-1} (y_T - \tilde{\gamma}_T^* z_T)^2\}}$$

et

$$(2.3)' \quad \tilde{v}_T = \frac{1}{T-1} \{T (1 + \tilde{\rho}_T^2) q_{0,T}(0) - (y_1 - \tilde{\gamma}_T^* z_1)^2 - \tilde{\rho}_T^2 (y_T - \tilde{\gamma}_T^* z_T)^2 - T \tilde{\rho}_T q_{1,T}(0)\}.$$

Nous pouvons alors démontrer

**Proposition 4** . Supposons que les conditions (D(0)), (A3) et (E) sont satisfaites. Alors l'estimateur des moindres carrés  $\tilde{\theta}_T$  défini par (2.1), (2.2) et (2.3) est presque sûrement convergent. Plus précisément on a

$$(2.6) \quad (\text{Log}d_{Tj})^{-(1+\varepsilon)/2} d_{Tj} (\tilde{\gamma}_T^j - \gamma_j) = o(1) \text{ p.s. ; } j = \overline{1, k}$$

et pour  $\beta < 1/2$

$$(2.7) \quad (\text{Log}d_{TM})^{-\beta(1+\varepsilon)} T^\beta (\tilde{\rho}_T - \rho) = o(1) \text{ p.s.}$$

$$(\text{Log}d_{TM})^{-\beta(1+\varepsilon)} T^\beta (\tilde{v}_T - v) = o(1) \text{ p.s..}$$

**Preuve** . D'après (2.5) on sait déjà que  $\tilde{\gamma}_T$  est convergent p.s. et plus précisément que (2.6) vraie. Une démonstration analogue à celle du Lemme 2 ci-dessus, fondée sur (2.5) montre que, pour  $q_{i,T}$ ,  $i = \overline{1, 2}$  défini par (1.16) on a

$$(2.8) \quad |q_{i,T(0)} - q_{i,T}| = o(T^{-1}(\text{Log}d_{TM})^{(1+\varepsilon)}) \text{ p.s., } i = \overline{0, 1} .$$

et donc, puisque (E) est satisfaite,

$$(2.9) \quad |q_{i,T(0)} - q_{i,T}| = o(1) \text{ p.s., } i = \overline{0, 1} .$$

On a aussi pour  $t = 1$  ou  $T$

$$(2.10) \quad T^{-1} (y_t - \tilde{\gamma}_T^* z_t)^2 \leq 2 T^{-1} \left\{ u_t^2 + \sum_{j=1}^k z_{tj}^2 (\tilde{\gamma}_T^j - \gamma_j)^2 \right\}$$

et donc

$$(2.11) \quad T^{-1} (y_t - \tilde{\gamma}_T^* z_t)^2 \xrightarrow{T \rightarrow +\infty} 0 \text{ p.s., } t = 1 \text{ ou } T .$$

En effet pour  $t = 1$  cela résulte immédiatement de ce que  $\tilde{\gamma}_T \rightarrow \gamma$  p.s.. Il en est de même pour  $t = T$  dans la mesure où  $T^{-1} u_T^2 \rightarrow 0$  p.s. et, d'après (2.5) et le fait que  $z_{Tj}^2 = o(d_{Tj}^2)$ ,

$$(2.12) \quad T^{-1} (\tilde{\gamma}_T^j - \gamma_j)^2 z_{Tj}^2 = o(T^{-1} (\text{Log}d_{TM})^{(1+\varepsilon)}) .$$

Appliquant alors (2.9), (2.11) et (1.19) dans (2.2)' et (2.3)', on obtient facilement que  $\tilde{\rho}_T$  et  $\tilde{v}_T$  convergent p.s. vers  $\rho$  et  $v$  respectivement.

En fait, d'après (2.2)', (2.8), (2.10) et (2.12), on peut écrire

$$\tilde{\rho}_T = \frac{q_{1,T} + \varepsilon_{1,T}}{2q_{0,T} + \varepsilon_{0,T}}$$

où  $\varepsilon_{0,T} = O(T^{-1} u_T^2) + o(T^{-1}(\text{Logd}_{TM})^{(1+\varepsilon)})$  p.s. ,  $\varepsilon_{1,T} = o(T^{-1}(\text{Logd}_{TM})^{(1+\varepsilon)})$  p.s.

et , d'après (1.16),

$$\begin{aligned} q_{1,T} &= \frac{2}{T} \sum_2^T u_{t-1} u_t = \frac{2}{T} \rho \sum_2^T u_{t-1}^2 + \frac{2}{T} \sum_2^T u_{t-1} v_t \\ &= \rho [2q_{0,T} + \varepsilon_{0,T}] - \frac{2\rho}{T} u_T^2 - \rho \varepsilon_{0,T} + \frac{2}{T} \sum_2^T u_{t-1} v_t , \end{aligned}$$

Alors

$$(2.13) \quad T^\beta (\tilde{\rho}_T - \rho) =$$

$$\left\{ 2q_{0,T} + \varepsilon_{0,T} \right\}^{-1} \left\{ -2\rho T^{-(1-\beta)} u_T^2 + T^\beta (\varepsilon_{1,T} - \rho \varepsilon_{0,T}) + 2T^{-(1-\beta)} \sum_2^T u_{t-1} v_t \right\}$$

où le premier facteur converge p.s. vers  $(2v/(1 - \rho^2))^{-1}$ . Comme  $\sum_1^\infty t^{-2(1-\beta)} < \infty$ ,

d'après la preuve du Lemme II.5 on a  $T^{-(1-\beta)} u_T^2 \rightarrow 0$  p.s.. Comme pour la martingale

$$(M_T = \sum_2^T u_{t-1} v_t, T \geq 2) \text{ on a } \frac{\langle M \rangle_T}{T} \rightarrow v/(1 - \rho^2) \text{ p.s., on a aussi } T^{-(1-\beta)} \sum_2^T u_{t-1} v_t \rightarrow 0$$

p.s.. Ainsi (2.13) montre que  $T^\beta (\tilde{\rho}_T - \rho) = o(1) + o(T^{-(1-\beta)} (\text{Logd}_{TM})^{(1+\varepsilon)})$  p.s.

d'où (2.7) pour  $\tilde{\rho}_T$ .

De même, d'après (2.3)', (2.8), (2.10), (2.12) et (2.7) pour  $\tilde{\rho}_T$ , on peut écrire

$$\tilde{v}_T = (1 + \rho^2) q_{0,T} - \rho q_{1,T} + \varepsilon_{3,T}$$

où

$$\varepsilon_{3,T} = O(T^{-1} u_T^2) + o(T^{-\beta} \text{Log} d_{TM})^{\beta(1+\varepsilon)} \text{ p.s.}$$

et, d'après (1.16),

$$\begin{aligned} (1 + \rho^2) q_{0,T} - \rho q_{1,T} &= \frac{1}{T} \underline{y}_T^* B_T(\rho) \underline{y}_T - \frac{1}{T} \rho^2 (u_1^2 + u_T^2) \\ &= \frac{1}{T} \underline{y}_T^* \underline{y}_T - \frac{1}{T} \rho^2 (u_1^2 + u_T^2) \end{aligned}$$

Ainsi

$$\tilde{v}_T - v = \frac{1}{T} \sum_1^T (v_i^2 - v) + \varepsilon_{4,T}$$

où

$$\varepsilon_{4,T} = O(T^{-1} u_T^2) + o(T^{-\beta} \text{Log} d_{TM})^{\beta(1+\varepsilon)} \text{ p.s..}$$

Cela permet encore de conclure à (2.7) pour  $\tilde{v}_T$  dans la mesure où, comme pour la martingale

$$(N_T = \sum_1^T (v_i^2 - v) ; T \geq 1) \text{ on a } \frac{\langle N \rangle_T}{T} \rightarrow 2v^2 \text{ p.s., on a aussi } T^{-(1-\beta)} \sum_1^T (v_i^2 - v) \rightarrow 0 \text{ p.s..} \quad \blacktriangleleft$$

## 2.2. Normalité asymptotique de l'estimateur des moindres carrés

Notons d'abord que  $\tilde{\theta}_T = (\tilde{\gamma}_T, \tilde{\rho}_T, \tilde{v}_T)^*$  est solution du système d'équations  $l_T(g, r, v) = 0$  où  $l_T = (l_{1,T}^*, l_{2,T}, l_{3,T})^*$  avec

$$l_{1,T}(g, r, v) = Z_T^* (y_T - Z_T g)$$

$$(2.14) \quad l_{2,T}(g, r, v) = - (y_T - Z_T g)^* \dot{B}_T(r) (y_T - Z_T g) - 2 r (y_1 - g^* z_1)^2$$

$$l_{3,T}(g, r, v) = - (T-1) v + (y_T - Z_T g)^* B_T(r) (y_T - Z_T g) - (1 - r^2) (y_1 - g^* z_1)^2.$$

Désignons par  $l_T^{(1)}$  la matrice  $(k+2) \times (k+2)$

$$(2.15) \quad l_T^{(1)} = \begin{pmatrix} \text{grad}_g l_T^* \\ \frac{\partial}{\partial r} l_T^* \\ \frac{\partial}{\partial v} l_T^* \end{pmatrix}.$$

Le comportement asymptotique de  $l_T(\theta)$  et  $l_T^{(1)}(\theta)$  est décrit dans:

**Lemme 5.** *Sous les hypothèses (D(0)), (A3) et (E), définissant  $\tilde{D}_T$  comme dans l'énoncé du Lemme 4, pour  $\theta = (\gamma^*, \rho, v)^*$  on a p.s.*

$$(2.16) \quad \lim_{T \rightarrow +\infty} \tilde{D}_T^{-1} l_T^{(1)}(\theta) \tilde{D}_T^{-1} = \begin{pmatrix} G(0) & 0 & 0 \\ 0 & 2 v (1 - \rho^2)^{-1} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Si en outre les conditions  $(A_2(\tau), \tau \geq 2)$  sont satisfaites alors la famille  $(\tilde{D}_T^{-1} l_T^{(1)}(\theta)) ; T \geq 1)$  est asymptotiquement gaussienne centrée de matrice de covariance

$$(2.17) \quad \Gamma(\theta) = \begin{pmatrix} v^2 K(\rho) & 0 & 0 \\ 0 & 4 v^2 (1 - \rho^2)^{-1} & 0 \\ 0 & 0 & 2v^2 \end{pmatrix}$$

où  $K(\rho)$  est la matrice donnée par II.(3.18).

Preuve. On a,

$$D_T^{-1} \text{grad}_g l_{1,T}^*(\theta) D_T^{-1} = - D_T^{-1} Z_T^* \dot{B}_T D_T^{-1}$$

qui, d'après le Lemme II.4, converge vers  $-G(0)$ .

On a

$$T^{-1} \frac{\partial}{\partial r} l_{2,T}(\theta) = - T^{-1} \underline{u}_T^* \dot{B}_T(\rho) \underline{u}_T - 2T^{-1} (y_1 - \gamma^* z_1)^2$$

qui, d'après II.(2.15), converge p.s. vers  $2v(1 - \rho^2)^{-1}$ .

On a

$$T^{-1} \frac{\partial}{\partial v} l_{3,T}(\theta) = - T^{-1} (T - 1) \rightarrow -1$$

On a

$$T^{-1} \frac{\partial}{\partial v} l_{2,T}(\theta) = 0$$

et

$$T^{-1} \frac{\partial}{\partial r} l_{3,T}(\theta) = T^{-1} \underline{u}_T^* \dot{B}_T(\rho) \underline{u}_T + 2\rho T^{-1} (y_1 - \gamma^* z_1)^2$$

qui, d'après II.(2.14), converge presque sûrement vers zéro.

On a aussi

$$T^{-1/2} D_T^{-1} \text{grad}_g l_{2,T}(\theta) = \left[ \frac{(\text{Log}(d_{TM})^{1+\epsilon})^{1/2}}{T} \right] (\text{Log}(d_{TM})^{-(1+\epsilon)/2} D_T^{-1} Z_T^* \dot{B}_T(\rho) \underline{u}_T + o(1))$$

qui, en vertu de (A3), (E) et de la preuve de la Proposition II.6, converge p.s. vers zéro.

On montre de même que

$$T^{-1/2} D_T^{-1} \text{grad}_g l_{3T}(\theta) \rightarrow 0 \text{ p.s.}$$

$$T^{-1/2} D_T^{-1} \frac{\partial}{\partial \alpha} l_{1,T}(\theta) \rightarrow 0 \text{ p.s.}$$

et

$$T^{-1/2} D_T^{-1} \frac{\partial}{\partial v} l_{1,T}(\theta) \rightarrow 0 \text{ p.s.}$$

Cela achève la démonstration de (2.16). Montrons maintenant la dernière assertion de l'énoncé. On a

$$l_T(\theta) = \begin{pmatrix} Y_T^1 \\ Y_T^2 \\ Y_T^3 \end{pmatrix} = \begin{pmatrix} Z_T^* \underline{u}_T \\ -\underline{u}_T^* \dot{B}_T(\rho) \underline{u}_T - 2\rho (y_1 - \gamma^* z_1)^2 \\ \underline{u}_T^* B_T(\rho) \underline{u}_T - (T-1)v - (1-\rho^2) (y_1 - \gamma^* z_1)^2 \end{pmatrix}$$

Comme  $\underline{u}_T = v^{1/2} M_T^{-1}(\rho) \underline{y}_T$ , et, avec les notations de la preuve du Lemme 4,

$$Y_T^2 = 2v X_T^2 + O(1) \quad ; \quad Y_T^3 = 2v^2 X_T^3 + O(1),$$

pour montrer le résultat il suffit d'assurer que pour

$$\tilde{l}_T(\theta) = \begin{pmatrix} Z_T^* M_T^{-1}(\rho) \underline{y}_T \\ X_T^2 \\ X_T^3 \end{pmatrix}$$

on a convergence en loi de la suite  $\{\tilde{D}_T^{-1} \tilde{l}_T(\theta); T \geq 1\}$  vers la loi gaussienne centrée de

matrice de covariance

$$\begin{pmatrix} vK(\rho) & 0 & 0 \\ 0 & (1 - \rho^2)^{-1} & 0 \\ 0 & 0 & (2v^2)^{-1} \end{pmatrix}.$$

Pour ce faire nous suivons la démarche utilisée dans la preuve du Lemme 4 en utilisant des notations analogues. Pour  $l \in \mathbb{R}^{k+2}$  nous écrivons

$$\tilde{l}_T(\theta) = \sum_1^T \tilde{\xi}_{T,t}, \quad T \geq 1.$$

où les variables  $\tilde{\xi}_{T,t}, 1 \leq t \leq T, T \geq 1$ , peuvent être identifiées en utilisant notamment le fait que

$$(M_T^{-1}(\rho))_{\lambda} \cdot \underline{y}_T = u_t = \rho^{t-1} \left\{ \left( \frac{v}{1-\rho^2} \right)^{1/2} v_1 + \sum_{s=1}^{t-1} \rho^{-s} v_{s+1} \right\}.$$

En fait on a

$$\tilde{\xi}_{T,t} = \sum_{j=1}^{k+2} l_j \tilde{d}_{Tj}^{-1} \sum_{m=1}^2 \eta_{Tt}^{jm} \Delta S_t^m$$

où, pour  $j = \overline{1, k}$

$$\eta_{Tt}^{jm} = \begin{cases} \left[ \left( \frac{v}{1-\rho^2} \right)^{1/2} z_{1j} + \sum_{s=2}^T z_{sj} \rho^{t-s} \right] & \text{si } t = 1, m = 1 \\ \sum_{s=t}^T z_{sj} \rho^{t-s} & \text{si } t \geq 2, m = 1 \\ 0 & \text{si } t \geq 1, m = 2, \end{cases}$$

et, pour  $j \in \{k+1, k+2\}, m \in \{1, 2\}, \eta_{Tt}^{jm} = \xi_{t-1}^{jm}$ , les  $\xi_{t-1}^{jm}$  étant tels que dans la preuve du Lemme 4, à savoir

$$\eta_{T,t}^{k+1,m} = \begin{cases} 0 & \text{si } t = 1, m = 1 \\ v^{-1}u_{t-1} & \text{si } t \geq 2, m = 1 \\ \rho v^{-1}(1-\rho^2)^{-1} & \text{si } m = 2, t = 1 \\ 0 & \text{si } m = 2, t \geq 2 \end{cases}$$

et

$$\eta_{T,t}^{k+2,m} = \begin{cases} 0 & \text{si } m = 1, t \geq 1 \\ (2v^2)^{-1} & \text{si } m = 2, t \geq 1. \end{cases}$$

Il est clair que la famille  $(\tilde{\xi}_{T,t}, 1 \leq t \leq T, T \geq 1)$  est une suite triangulaire. Pour montrer que  $(\tilde{\xi}_{T,t})$  est asymptotiquement négligeable il suffit de montrer que pour  $j = \overline{1, k+2}$  et  $m = \overline{1, 2}$  on a pour un  $\beta > 0$

$$(2.18) \quad \lim_{T \rightarrow +\infty} \sum_1^T E \left( |\eta_{T,t}^{j,m}|^\beta / \mathcal{F}_{t-1} \right) = 0 \text{ en probabilité.}$$

Pour  $j = k+1, k+2$  cela a déjà été vu dans la preuve du Lemme 4.

C'est aussi acquis pour  $j = \overline{1, k}$  et  $m=2$  puisque  $\eta_{T,t}^{j,2} = 0$ . Pour  $j = \overline{1, k}$  et  $m = 1$  on a

$$E \left( |\eta_{T,t}^{j,1}|^\beta / \mathcal{F}_{t-1} \right) = |l_j|^\beta d_{Tj}^{-\beta} |\eta_{T,t}^{j,1}|^\beta E |\Delta S_t^1|^\beta, t \geq 1$$

et donc, comme  $E |\Delta S_t^1|^\beta = E |v_t|^\beta < \infty$  pour tout  $\beta > 0$ , il suffit de vérifier qu'il existe un  $\beta > 0$  tel que

$$(2.19) \quad \lim_{T \rightarrow +\infty} d_{Tj}^{-\beta} \sum_{t=2}^T \left( \sum_{s=t}^T |z_{sj}| |\rho|^{t-s} \right)^\beta = 0.$$

Or, appliquant l'inégalité de Hölder, on a pour  $\beta > 1$

$$\begin{aligned} \sum_{s=t}^T |z_{sj}| |\rho|^{t-s} &= \sum_{s=t}^T \left\{ (s+1-t)^{-(\beta-1)/\beta} |z_{sj}| \right\} \left\{ (s+1-t)^{(\beta-1)/\beta} |\rho|^{t-s} \right\} \\ &\leq \left\{ \sum_{s=t}^T (s+1-t)^{-(\beta-1)} |z_{sj}|^\beta \right\}^{1/\beta} \left\{ \sum_{s=t}^T (s+1-t) |\rho|^{\frac{\beta}{\beta-1}(t-s)} \right\}^{(\beta-1)/\beta} \\ &\leq C^e \left\{ \sum_{s=t}^T (s+1-t)^{-(\beta-1)} |z_{sj}|^\beta \right\}^{1/\beta}. \end{aligned}$$

Alors pour  $\beta \geq 2$

$$\begin{aligned} (2.20) \quad d_{Tj}^{-\beta} \sum_{t=2}^T \left( \sum_{s=t}^T |z_{sj}| |\rho|^{t-1} \right)^\beta &\leq C^{te} d_{Tj}^{-\beta} \left\{ \sum_{t=2}^T \sum_{s=t}^T (s+1-t)^{-(\beta-1)} |z_{sj}|^\beta \right\} \\ &= C^{te} d_{Tj}^{-\beta} \left\{ \sum_{s=2}^T \left\{ \sum_{t=2}^s (s+1-t)^{-(\beta+1)} \right\} |z_{sj}|^\beta \right\} \\ &\leq C^{te} d_{Tj}^{-\beta} \sum_{s=1}^T |z_{sj}|^\beta \\ &\leq C^{te} \left\{ d_{Tj}^{-1} \left( \text{Max}_{s \leq T} |z_{sj}| \right) \right\}^{\beta-2} \end{aligned}$$

d'où (2.19) pour  $\beta > 2$  en vertu de (A1).

Puisque  $E(\tilde{\xi}_{T,t} / \mathcal{F}_{t-1}) = 0$  p.s. on a bien sûr

$$(2.21) \quad \lim_{T \rightarrow +\infty} E(\tilde{\xi}_{T,t} / \mathcal{F}_{t-1}) = 0 \text{ en probabilité.}$$

D'autre part on a

$$\sum_1^T E(\tilde{\xi}_{T,t}^2 / \mathcal{F}_{t-1}) = I^* \tilde{D}_T^{-1} \left\{ \sum_1^T \eta_{T,t} E(\Delta S_t \Delta S_t^*) \eta_{T,t} \right\} \tilde{D}_T^{-1} I$$

où

$$\eta_{T,t} = ((\eta_{T,t}^{jm})) \text{ et } E(\Delta S_t \Delta S_t^*) = \begin{pmatrix} v & 0 \\ 0 & 2v^2 \end{pmatrix}.$$

Le calcul montre que notant  $\sum_T = \sum_1^T \eta_{T,t} E(\Delta S_t \Delta S_t^*) \eta_{T,t}$ , on a

$$\left( \left( \sum_T^j ; j, l = \overline{1, k} \right) \right) = v Z_T^* B_T^{-1}(\rho) Z_T$$

$$\sum_T^{k+1, k+1} = \langle X^2 \rangle_T ; \sum_T^{k+2, k+2} = \langle X^3 \rangle_T.$$

Quant aux blocs non diagonaux de  $\sum_T$  ils contiennent des termes de la forme

$$\sum_T^{j1} = v \sum_{t=1}^T \eta_{T,t}^{j1} \eta_{T,t}^{11} + 2v^2 \sum_{t=1}^T \eta_{T,t}^{j2} \eta_{T,t}^{12} \text{ pour } j \neq 1.$$

Pour  $j = k + 1, l = k + 2$  il s'agit de la variation croisée  $\langle X^2, X^3 \rangle_T$ .

Pour  $j = k + 1, l = \overline{1, k}$  on a

$$\sum_T^{k+1, l} = v^{-1} \sum_{t=2}^T u_{t-1} \left( \sum_{s=t}^T z_{sl} \rho^{t-s} \right)$$

et pour  $j = k + 2, l = \overline{1, k}$  on a  $\sum_T^{k+2, l} = 0$ .

En définitive on a

$$\sum_1^T E(\tilde{\mathfrak{F}}_{T,1}^2 / \mathfrak{F}_{t-1}) = I^* \tilde{D}_T^{-1} \begin{pmatrix} v Z_T^* B_T^{-1}(\rho) Z_T & 0 & 0 \\ 0 & \langle X^2 \rangle_T & 0 \\ 0 & 0 & \langle X^2 \rangle_T \end{pmatrix} \tilde{D}_T^{-1} I$$

$$+ 2 I_{k+1} I_{k+2} T^{-1} \langle X^2, X^3 \rangle_T + 2 I_{k+1} \sum_{u=1}^k I_u T^{-1/2} d_{Tu}^{-1} \sum_T^{k+1, u}.$$

D'après la preuve du Lemme 4 et II.(3.18) le premier terme du second membre tend p.s.

vers  $1^* \begin{pmatrix} v K(\rho) & 0 & 0 \\ 0 & (1-\rho^2)^{-1} & 0 \\ 0 & 0 & (2v^2)^{-1} \end{pmatrix} 1$  et le second tend p.s. vers zéro. Nous allons

montrer que le dernier terme tend en probabilité vers zéro. Pour ce faire il suffit de montrer que pour  $j = \overline{1, k}$  on a

$$(2.22) \quad \lim_{T \rightarrow +\infty} E \left| T^{-1/2} d_{Tj}^{-1} \sum_{t=2}^T u_{t-1} \left( \sum_{s=t}^T z_{sj} \rho^{t-s} \right) \right|^2 = 0.$$

Or posant  $\alpha_{Tt} = \sum_{s=t}^T z_{sj} \rho^{t-s}$ , on a

$$\begin{aligned} E \left| T^{-1/2} d_{Tj}^{-1} \sum_{t=2}^T u_{t-1} \alpha_{Tt} \right|^2 &= T^{-1} d_{Tj}^{-2} \sum_{t,t'=2}^T \alpha_{Tt} \alpha_{Tt'} \text{Cov}(u_{t-1}, u_{t'-1}) \\ &= T^{-1} d_{Tj}^{-2} \frac{v}{1-\rho^2} \sum_{t,t'=2}^T \alpha_{Tt} \alpha_{Tt'} \rho^{|t-t'|} \\ &= T^{-1} d_{Tj}^{-2} \frac{v}{1-\rho^2} \sum_{j=2-T}^{T-2} \rho^{|j|} \sum_{\substack{2 \leq t \leq T \\ 2 \leq t-j \leq T}} \alpha_{T,t} \alpha_{T,t-j} \\ &\leq T^{-1} d_{Tj}^{-2} \frac{v}{1-\rho^2} \left( \sum_{j=2-T}^{T-2} \rho^{|j|} \right) \left( \sum_{t=2}^T \alpha_{T,t}^2 \right), \end{aligned}$$

la dernière inégalité s'obtenant par application de l'inégalité de Schwarz. Mais, d'après

(2.21) pour  $\beta = 2$ ,  $d_{Tj}^{-2} \sum_{t=2}^T \alpha_{T,t}^2$  est borné. Ainsi, la série  $\sum_{j=-\infty}^{+\infty} \rho^{|j|}$  étant convergente,

l'assertion (2.22) est bien vérifiée.

En définitive on a

$$(2.23) \quad \lim_{T \rightarrow +\infty} \sum_1^T E(\tilde{\xi}_{T,1}^2 / \mathcal{F}_{t-1}) = I^* \begin{pmatrix} v K(\rho) & 0 & 0 \\ 0 & (1 - \rho^2)^{-1} & 0 \\ 0 & 0 & (2v^2)^{-1} \end{pmatrix} I.$$

Alors, compte tenu de (2.18), (2.21) et (2.23), on peut appliquer le théorème de la limite centrale pour les suites triangulaires (cf. DACUNHA-CASTELLE et DUFLO (1983), théorème 2.8.42) qui assure que la suite  $\{I^* \tilde{D}_T^{-1} \tilde{I}_T(\theta); T \geq 1\}$  est asymptotiquement gaussienne centrée de variance égale au second membre de (2.23). Cela étant vrai pour tout  $I \in \mathbb{R}^{k+2}$ , l'assertion du Lemme est démontrée. ♣

Nous pouvons maintenant démontrer

**Proposition 5.** *Supposons que les hypothèses (D(0)), (A3), (A2(τ), τ ≥ 2) et (E) sont satisfaites. Alors si  $\tilde{\theta}_T$  est l'estimateur des moindres carrés de θ défini par (2.1), (2.2), et (2.3), la suite  $(\tilde{D}_T(\tilde{\theta}_T - \theta); T \geq 1)$  est asymptotiquement gaussienne centrée de matrice de covariance donnée par*

$$(2.24) \quad \tilde{\Gamma}(\theta) = \begin{pmatrix} v G^{-1}(0) K(\rho) G^{-1}(0) & 0 & 0 \\ 0 & 1 - \rho^2 & 0 \\ 0 & 0 & 2v^2 \end{pmatrix}.$$

**Preuve.** On procède de manière analogue à la preuve de la Proposition 2 en écrivant

$$\tilde{D}_T^{-1} I_T(\theta) = \tilde{D}_T^{-1} \left( \left( I_{T,ij}^{(1)}(\theta_T^*(i)) \right) \right) \tilde{D}_T^{-1} \tilde{D}_T (\tilde{\theta}_T - \theta)$$

avec

$$\theta_T^*(i) = \tilde{\theta}_T + \alpha_i(\tilde{\theta}_T)(\theta - \tilde{\theta}_T), \alpha_i(\tilde{\theta}_T) \in [0, 1].$$

On vérifie, comme dans la preuve du Lemme 5 que, compte tenu de ce que  $\theta_T^* \rightarrow \theta$  p.s., la convergence p.s. de  $\tilde{D}_T^{-1} \left( \left( I_{T,ij}^{(1)}(\theta_T^*(i)) \right) \right) \tilde{D}_T^{-1}$  vers le second membre de (2.16) est valide. Appliquant alors la dernière assertion du Lemme 5 et utilisant le fait que, par (D(0)), le second membre de (2.16) est une matrice régulière, on obtient immédiatement le résultat annoncé. ♣

### §3. Estimation de maximum de vraisemblance approchée

Dans ce paragraphe nous proposons une troisième procédure d'estimation des paramètres  $\gamma$ ,  $\rho$  et  $v$ . Elle consiste d'abord à estimer  $\gamma$ , comme on le ferait si  $\rho$  était connu égal à son estimateur des moindres carrés presque sûrement convergent (cf. §2.(2.2)), par maximum de vraisemblance i.e. par

$$(3.1) \quad \check{\gamma}_T = [Z_T^* B_T(\tilde{\rho}_T) Z_T]^{-1} Z_T^* B_T(\tilde{\rho}_T) y_T,$$

puis à estimer  $\rho$  et  $v$ , comme on le ferait si  $\gamma$  était connu égal à  $\check{\gamma}_T$ , par moindres carrés, à savoir (cf. II. §.2.2) par

$$(3.2) \quad \check{\rho}_T = \left\{ \sum_{t=2}^T (y_t - \check{\gamma}_T^* z_t)(y_{t-1} - \check{\gamma}_T^* z_{t-1}) \right\} / \left\{ \sum_{t=1}^{T-1} (y_t - \check{\gamma}_T^* z_t) \right\}^2,$$

et

$$(3.3) \quad \check{v}_T = \frac{1}{T-1} \sum_{t=2}^T \left\{ (y_t - \check{\gamma}_T^* z_t) - \check{\rho}_T (y_{t-1} - \check{\gamma}_T^* z_{t-1}) \right\}^2.$$

L'estimateur  $\check{\theta}_T = (\check{\gamma}_T^*, \check{\rho}_T, \check{v}_T)^*$  de  $\theta = (\gamma^*, \rho, v)^*$  défini par (3.1)-(3.3) est dans la suite appelé "estimateur de maximum de vraisemblance approchée".

#### 3.1. Convergence et vitesse de convergence presque sûre de l'estimateur de maximum de vraisemblance approchée

Comme  $\check{\gamma}_T$  défini par (3.1) est en fait donné par

$$(3.4) \quad \check{\gamma}_T = g_T(\tilde{\rho}_T),$$

où  $g_T(\cdot)$  est fourni par (3.19), d'après les considérations du II§3.2.c, nous savons que si les conditions (D) et (A3) sont vérifiées, alors pour tout  $\eta > 1/2$  on a

$$(3.5) \quad \check{\gamma}_T^j - \gamma_j = o(d_{Tj}^{-1} (\text{Log} d_{Tj})^\eta) \text{ p.s.}$$

Vu (3.4), avec les notations du §1.2 ci-dessus (cf. (1.13) et (1.15), nous pouvons écrire

$$\begin{aligned} \sum_{t=1}^{T-1} (y_t - \check{\gamma}_T^* z_t)^2 &= \hat{u}_T^*(\tilde{\rho}_T) \hat{u}_T(\tilde{\rho}_T) - (y_T - \check{\gamma}_T^* z_T)^2 \\ &= T q_{0,T}(\tilde{\rho}_T) - (y_T - \check{\gamma}_T^* z_T)^2 \end{aligned}$$

et

$$\begin{aligned} \sum_{t=2}^T (y_t - \check{\gamma}_T^* z_t)^2 &= \hat{u}_T^*(\tilde{\rho}_T) \hat{u}_T(\tilde{\rho}_T) - (y_1 - \check{\gamma}_T^* z_1)^2 \\ &= T q_{0,T}(\tilde{\rho}_T) - (y_1 - \check{\gamma}_T^* z_1)^2. \end{aligned}$$

De même on a

$$\begin{aligned} \sum_{t=2}^T (y_t - \check{\gamma}_T^* z_t) (y_{t-1} - \check{\gamma}_T^* z_{t-1}) &= \frac{1}{2} \hat{u}_T^*(\tilde{\rho}_T) A_T^1 \hat{u}_T(\tilde{\rho}_T) \\ &= \frac{T}{2} q_{1,T}(\tilde{\rho}_T). \end{aligned}$$

Par suite, compte tenu de (3.2) et (3.3) on a

$$(3.2)' \quad \check{\rho}_T^v = \frac{q_{1,T}(\tilde{\rho}_T)}{2 \{ q_{0,T}(\tilde{\rho}_T) - T^{-1} (y_T - \check{\gamma}_T^* z_T)^2 \}}$$

et

$$(3.3)' \quad \check{v}_T = \frac{1}{T-1} \{ T (1 + \check{\rho}_T^v)^2 q_{0,T}(\tilde{\rho}_T) - (y_1 - \check{\gamma}_T^* z_1)^2 - \check{\rho}_T^v (y_T - \check{\gamma}_T^* z_T)^2 - T \check{\rho}_T^v q_{1,T}(\tilde{\rho}_T) \}.$$

Par une preuve analogue à celle de la proposition 4 du §2.1, nous pouvons démontrer

**Proposition 6.** *Supposons que les conditions (D), (A3) et (E) sont satisfaites. Alors l'estimateur de maximum de vraisemblance approchée  $\check{\theta}_T^v$  défini par (3.1)-(3.2) et (3.3) est presque sûrement convergent. Plus précisément on a*

$$(3.6) \quad (\text{Log} d_{Tj})^{-(1+\varepsilon)2} d_{Tj}^{\nu_j} (\check{\gamma}_T^{\nu_j} - \gamma_j) = o(1) p.s.; \quad j = \overline{1, k}$$

et pour tout  $\beta < 1/2$

$$(\text{Log} d_{TM})^{-\beta(1+\varepsilon)} T^\beta (\check{\rho}_T - \rho) = o(1) p.s.;$$

$$(\text{Log} d_{TM})^{-\beta(1+\varepsilon)} T^\beta (\check{\nu}_T - \nu) = o(1) p.s..$$

### 3.2. Normalité asymptotique de l'estimateur de maximum de vraisemblance approchée

Supposons les conditions (D), (A3) et (E) sont satisfaites.

Notons d'abord qu'à partir de (3.1), (3.2)' et (3.3)' on peut écrire

$$(3.7) \quad D_T (\check{\gamma}_T - \gamma) = [D_T^{-1} Z_T^* B_T(\rho) Z_T D_T^{-1} + R_{0,T}]^{-1} [D_T^{-1} Z_T^* B_T(\rho) \underline{u}_T + R_{1,T}],$$

$$(3.8) \quad T^{1/2} (\check{\rho}_T - \rho) =$$

$$\{2q_{0,T} + R_{3,T}\}^{-1} \{-2\rho T^{-1/2} u_T^2 + T^{1/2}(R_{2,T} - \rho R_{3,T}) + 2T^{-1/2} \sum_{t=2}^T u_{t-1} \nu_t\},$$

$$(3.9) \quad T^{1/2} (\check{\nu}_T - \nu) = \frac{T}{T-1} \{T^{-1/2} \sum_{t=1}^T (\nu_t^2 - \nu) + R_{4,T}\},$$

avec

$$R_{0,T} = (\check{\rho}_T - \rho) D_T^{-1} Z_T^* \{A_T^1 + (\check{\rho}_T + \rho) A_T^2\} Z_T D_T^{-1},$$

$$R_{1,T} = (\check{\rho}_T - \rho) D_T^{-1} Z_T^* \{A_T^1 + (\check{\rho}_T + \rho) A_T^2\} \underline{u}_T,$$

$$R_{2,T} = T^{-1} (D_T (\check{\gamma}_T - \gamma))^* \{-2D_T^{-1} Z_T^* A_T^1 \underline{u}_T + (D_T^{-1} Z_T^* A_T^1 Z_T D_T^{-1}) (D_T (\check{\gamma}_T - \gamma))\},$$

$$R_{3,T} = T^{-1} (D_T (\check{\gamma}_T - \gamma))^* \{-4D_T^{-1} Z_T^* A_T^0 \underline{u}_T + 2(D_T^{-1} Z_T^* A_T^0 Z_T D_T^{-1}) (D_T (\check{\gamma}_T - \gamma))\} - 2T^{-1} (y_T - \check{\gamma}_T^* z_T)^2,$$

$$R_{4,T} = T^{-1/2} \{ -(\overset{v}{\rho}_T - \rho) \underline{u}_T^* A_T^1 \underline{u}_T + (\overset{v}{\rho}_T^2 - \rho^2) \underline{u}_T^* A_T^2 \underline{u}_T - 2(D_T(\overset{v}{\gamma}_T - \gamma))^* D_T^{-1} B_T(\rho) \underline{u}_T + 2R_{1,T} \\ + (D_T(\overset{v}{\gamma}_T - \gamma))^* \{ (D_T^{-1} Z_T^* B_T(\rho) Z_T D_T^{-1}) + R_{0,T} \} (D_T(\overset{v}{\gamma}_T - \gamma)) - (1 - \overset{v}{\rho}_T^2) (y_1 - \overset{v}{\gamma}_T^* z_1)^2 \}.$$

En vertu de la convergence presque sûre de  $\tilde{\rho}_T$  vers  $\rho$ , celle des termes  $D_T^{-1} Z_T^* A_T^i Z_T D_T^{-1}$  et le fait que  $D_T^{-1} Z_T^* A_T^i \underline{u}_T = O(1)$  en moyenne quadratique on a  $R_{0,T} = o(1)$  p.s. et  $R_{1,T} = o(1)$  en probabilité. On a aussi,  $D_T(\overset{v}{\gamma}_T - \gamma) = O(1)$  en probabilité, vue la convergence en loi de la suite  $\{D_T(\overset{v}{\gamma}_T - \gamma); T \geq 1\}$  d'après la Proposition 4ChII§3.b. D'autre part, il est facile de vérifier que  $R_{3,T} = o(1)$  en probabilité et  $-2\rho T^{-1/2} + T^{1/2} (R_{2,T} - \rho R_{3,T}) = o(1)$  en probabilité. Enfin puisque  $\{T^{-1/2} \underline{u}_T^* A_T^i \underline{u}_T; T \geq 1\}$  est bornée en moyenne quadratique on a  $R_{4,T} = o(1)$  en probabilité.

Nous sommes en mesure de montrer

**Proposition 7.** *Supposons que les hypothèses (D), (A3) et (E) sont satisfaites. Alors si  $\overset{v}{\theta}_T$  est l'estimateur de maximum de vraisemblance approchée de  $\theta$  défini par (3.1), (3.2) et (3.3), la suite  $\{\tilde{D}_T(\overset{v}{\theta}_T - \theta); T \geq 1\}$  est asymptotiquement gaussienne centrée de matrice de covariance  $\Gamma(\theta)^{-1}$ , où  $\Gamma(\theta)$  est donnée par (1.26).*

**Preuve.** D'après (3.7)-(3.9) on peut écrire

$$(3.10) \quad \tilde{D}_T(\overset{v}{\theta}_T - \theta) = Q_T \tilde{D}_T^{-1} S_T + R_T,$$

où  $Q_T$  est la  $T \times T$  matrice diagonale par blocs dont les blocs diagonaux sont

$$Q_T^{11} = [D_T^{-1} Z_T^* B_T(\rho) Z_T D_T^{-1} + R_{0,T}]^{-1} \quad ; \quad Q_T^{22} = (2q_{0,T} + T^{-1} R_{3,T})^{-1} \quad ; \quad Q_T^{33} = \frac{T}{T-1},$$

$$S_T = ((Z_T^* B_T(\rho) \underline{u}_T)^*, 2 \sum_{t=2}^T u_{t-1} v_t, \sum_{t=1}^T (v_t^2 - v))^*$$

et  $\{R_T; T \geq 1\}$  est une suite de vecteurs aléatoires dont les composantes sont faciles à identifier à partir de (3.7)-(3.9) et convergent en probabilité vers 0, quand  $T$  tend vers l'infini.

La dernière assertion de l'énoncé du Lemme 4 assure que la famille  $\{\tilde{D}_T^{-1} S_T; T \geq 1\}$  est asymptotiquement gaussienne centrée de matrice de covariance

$$\Gamma(\theta) = \begin{pmatrix} vH(\rho) & 0 & 0 \\ 0 & 4v^2(1-\rho^2)^{-1} & 0 \\ 0 & 0 & 2v^2 \end{pmatrix}.$$

D'autre part on a

$$\lim_{T \rightarrow +\infty} [D_T^{-1} Z_T^* B_T(\rho) Z_T D_T^{-1} + R_{0,T}]^{-1} = vH^{-1}(\rho)$$

et

$$\lim_{T \rightarrow +\infty} (2 q_{0,T} + T^{-1} R_{3,T}) = (2v/(1-\rho^2))^{-1} \text{ p.s..}$$

On en déduit immédiatement, à partir de (3.10), l'assertion de la proposition.



## **Chapitre IV**

DETECTION DANS UN BRUIT AR(1)



Le problème de détection d'un signal contaminé par un bruit est plus au moins classique selon que la loi du bruit est bien spécifiée ou non. Pour le premier cas citons par exemple HELSTROM (1968), VAN TREES (1971), WONG (1971), PICINBONO (1987),... . Dans le deuxième cas des procédures "adaptatives" ont été envisagées (cf. par exemple MERMOZ (1971),....) mais leur valeur théorique n'est pas bien maîtrisée.

Dans ce Chapitre nous considérons le problème de détection d'un signal combinaison linéaire, à coefficients éventuellement inconnus, de suites déterministes données, susceptible d'être présent lors des mesures effectuées sur un écouteur bruité par un processus d'erreurs AR(1). Nous nous plaçons dans le cadre du formalisme décrit dans les chapitres I et II. Nous nous efforçons ici de bien analyser les risques associés à des procédures de détection qui sont naturelles compte tenu des discussions menées au Chapitre I ou des résultats concernant l'estimation obtenus dans les Chapitres II et III. Nous examinons le problème selon que les paramètres du bruit sont connus, partiellement inconnus ou complètement inconnus. Dans la première partie nous traitons le cas d'un signal complètement connu. Dans la deuxième partie nous traitons le cas où les paramètres du signal sont inconnus.

## §1. Détection d'un signal connu

On considère ici le cas où le signal  $s = (s_t, t = 1, 2, \dots)$  susceptible d'être présent est complètement connu, par exemple de la forme  $s_t = \gamma^* z_t, t = 1, 2, \dots$  où  $z = (z_t, t = 1, 2, \dots)$  est un signal vectoriel donné à valeurs dans  $\mathbb{R}^k$  et  $\gamma$  un paramètre connu dans  $\mathbb{R}^k$ . Le problème de détection est donc celui du test de l'hypothèse  $\mathcal{H}_0$  "y = u" contre l'hypothèse  $\mathcal{H}_1$  "y = s + u".

### 1.1. Cas d'un bruit AR(1) de paramètres connus

Rappelons que, d'après l'équation (1.5) du chapitre II, on peut écrire l'inverse de la matrice de covariance du bruit  $\underline{y}_T = (u_1, \dots, u_T)^*$  sous la forme

$$(1.1) \quad v^{-1} B_T(\rho) = (v^{-1/2} M_T^*(\rho))(v^{-1/2} M_T(\rho))$$

où  $M_T(\rho)$  est donnée par

$$(1.2) \quad M_T(\rho) = \begin{pmatrix} (1-\rho^2)^{1/2} & 0 \dots & 0 \\ -\rho & 1 \dots & \vdots \\ 0 \dots 0 & -\rho & 1 \end{pmatrix}.$$

Alors, appliquant les considérations du §3.3.1.a du Chapitre I, on définit un récepteur  $\Phi_\alpha$  optimal (U.P.P.) au niveau  $\alpha$  de probabilité de fausse alarme par la région critique

$$(1.3) \quad R_{\Phi_\alpha} = \left\{ \underline{y}_T : v^{-1/2} \frac{(M_T(\rho) \underline{s}_T)^* (M_T(\rho) \underline{y}_T)}{|M_T(\rho) \underline{s}_T|} > \mathcal{N}^{-1}(1-\alpha) \right\},$$

où  $|x|$  désigne la norme euclidienne d'un vecteur.

La probabilité de non détection correspondante vaut

$$(1.4) \quad 1 - \beta_{\Phi_\alpha} = \mathcal{N}(\mathcal{N}^{-1}(1 - \alpha) - v^{-1/2} |M_T(\rho) \mathfrak{s}_T|).$$

## 1.2. Cas d'un bruit AR(1) de paramètre autorégressif connu

On suppose ici que seul le paramètre  $\rho$  du bruit est connu. On est alors placé dans la situation où la structure de covariance du bruit est connue au facteur près  $v$ . Alors appliquant (1.1) et les considérations du §3.3.1.b du Chapitre I, on définit un récepteur optimal (U.P.P.L.) au niveau  $\alpha$  de probabilité de fausse alarme, par la région critique

$$(1.5) \quad R_{\Phi_\alpha} = \left\{ \mathfrak{y}_T : \frac{\sqrt{T-1} (M_T(\rho) \mathfrak{s}_T)^* (M_T(\rho) \mathfrak{y}_T)}{\left[ |M_T(\rho) \mathfrak{s}_T|^2 |M_T(\rho) \mathfrak{y}_T|^2 - \left\{ (M_T(\rho) \mathfrak{s}_T)^* (M_T(\rho) \mathfrak{y}_T) \right\}^2 \right]^{1/2}} > \mathfrak{c}_{T-1}^{-1}(1 - \alpha) \right\}.$$

Les probabilités de non détection correspondantes valent

$$(1.6) \quad 1 - \beta_{\Phi_\alpha(v)} = \text{Prob} \left[ \frac{\mathfrak{c}_{T-1}^{-1}(1 - \alpha) \sqrt{W}}{\sqrt{T-1}} - V \geq \frac{|M_T(\rho) \mathfrak{s}_T|}{v^{1/2}} \right]$$

où  $V$  et  $W$  sont des variables aléatoires indépendantes de lois respectives  $N(0, 1)$  et  $\chi_{T-1}^2$  du chi deux à  $T-1$  degrés de liberté.

## 1.3. Cas d'un bruit AR(1) de variance d'innovation connue

On suppose ici que seul le paramètre  $v$  du bruit est connu. Compte tenu des considérations du §1.1. où  $\rho$  est aussi connu, il semble naturel d'envisager un récepteur  $\Phi_T$  défini par une région critique de la forme

$$(1.7) \quad R_{\Phi_T} = \left\{ \mathfrak{y}_T : v^{-1/2} V_T(\rho_T) > \mathcal{N}^{-1}(1 - \alpha) \right\}$$

où  $V_T(r)$  est donné par

$$V_T(r) = \frac{v^{-1/2} (M_T(r) \underline{s}_T)^* (M_T(r) \underline{y}_T)}{|M_T(r) \underline{s}_T|}$$

(1.8)

$$= \frac{\underline{s}_T^* B_T(r) \underline{y}_T}{(\underline{s}_T^* B_T(r) \underline{s}_T)^{1/2}},$$

et  $\rho_T$  est un estimateur de  $\rho$  qui converge au moins en probabilité dans les conditions asymptotiques envisagées ci-après (cf. Ch. II et III).

Pour  $i = \overline{0, 1}$ , sous l'hypothèse  $\mathcal{H}_i$  on peut écrire

$$V_T(\rho) = \frac{\underline{s}_T^* B_T(\rho) \underline{y}_T}{(\underline{s}_T^* B_T(\rho) \underline{s}_T)^{1/2}} + i (\underline{s}_T^* B_T(\rho) \underline{s}_T)^{1/2}.$$

La loi de  $v^{-1/2} V_T(\rho)$  est évidemment  $\mathcal{N}(i v^{-1/2} (\underline{s}_T^* B_T(\rho) \underline{s}_T)^{1/2}, 1)$ .

Plus généralement, si pour  $\lambda \geq 0$   $\mathcal{H}_{T, \lambda}$  est l'hypothèse

$$"y_t = \lambda s_t + u_t; t = 1, \dots, T",$$

sous  $\mathcal{H}_{T, \lambda}$  on a

$$(1.9) \quad V_T(\rho) = \frac{\underline{s}_T^* B_T(\rho) \underline{y}_T}{(\underline{s}_T^* B_T(\rho) \underline{s}_T)^{1/2}} + \lambda (\underline{s}_T^* B_T(\rho) \underline{s}_T)^{1/2}$$

et la loi de  $v^{-1/2} V_T(\rho)$  est  $\mathcal{N}(\lambda v^{-1/2} (\underline{s}_T^* B_T(\rho) \underline{s}_T)^{1/2}, 1)$ .

Nous nous proposons d'étudier les niveaux asymptotiques de probabilité de fausse alarme et de non détection attachés au test  $\Phi_T$  défini par (1.7)-(1.8) ainsi que la puissance asymptotique de ce test pour une suite d'hypothèses alternatives  $\{\mathcal{H}_{T, \lambda}; T \geq 1\}$ , contigues à  $\mathcal{H}_0$ .

Introduisons la condition

$$(L) \quad \lim_{T \rightarrow +\infty} d_T = +\infty; \quad \lim_{T \rightarrow +\infty} d_T^{-1} \max_{1 \leq t \leq T} |s_t| = 0; \quad \lim_{T \rightarrow +\infty} d_T^{-2} \sum_{t=2}^T s_t s_{t-1} = g(1) \in ]-1, +1[$$

où  $d_T = \left( \sum_1^T s_T^2 \right)^{1/2}$ . On pose  $h(\rho) = 1 + \rho^2 - 2\rho g(1)$ . Notons qu'en vertu de ce que  $|g(1)| < 1$  on a  $h(\rho) > 0$ .

Démontrons d'abord

**Lemme 1.** *Supposons que la condition (L) est satisfaite. Alors, sous l'hypothèse  $\mathcal{H}_{T, \lambda}$  on peut écrire*

$$(1.10) \quad V_T(\rho) - V_T(\rho_T) = (1 + \lambda d_T) \varepsilon_T$$

où  $\{\varepsilon_T ; T \geq 1\}$  est une suite de variables aléatoires qui tend vers 0 au sens de la convergence en probabilité.

**Preuve.** Utilisant (1.8)-(1.9) et le fait que

$$B_T(r) = \mathbb{1}_T - r A_T^1 + r^2 A_T^2$$

(cf. Ch.III§1.2) et donc

$$B_T(\rho_T) = B_T(\rho) - (\rho_T - \rho) A_T^1 + (\rho_T^2 - \rho^2) A_T^2,$$

on obtient facilement que sous  $\mathcal{H}_{T, \lambda}$  on peut écrire

$$(1.11) \quad \begin{aligned} V_T(\rho) - V_T(\rho_T) = & \\ & [1 - (1 - \delta_{1,T})^{-1/2}] \left( s_T^* B_T(\rho) s_T \right)^{-1/2} s_T^* B_T(\rho) s_T + (1 - \delta_{1,T})^{-1/2} \delta_{2,T} \\ & + \lambda \left( s_T^* B_T(\rho) s_T \right)^{+1/2} [1 - (1 - \delta_{1,T})^{+1/2}] \end{aligned}$$

avec

$$\delta_{1,T} = (\rho_T - \rho) \left( s_T^* B_T(\rho) s_T \right)^{-1} \left[ s_T^* A_T^1 s_T - (\rho_T - \rho) s_T^* A_T^2 s_T \right]$$

et

$$\delta_{2, T} = (\rho_T - \rho) (\hat{s}_T^* B_T(\rho) \hat{s}_T)^{-1/2} [ \hat{s}_T^* A_T^1 \underline{u}_T - (\rho_T + \rho) \hat{s}_T^* A_T^2 \underline{u}_T ] .$$

D'après (L), par une preuve analogue à celle du Lemme II.4, on peut voir que  $d_T^{-2} \hat{s}_T^* B_T(\rho) \hat{s}_T$  converge quand  $T \rightarrow +\infty$  vers  $h(\rho) = (1 + \rho^2) - 2\rho g(1) > 0$ . De même les quantités  $d_T^{-2} \hat{s}_T^* A_T^i \hat{s}_T$  ;  $i = \overline{1, 2}$  convergent quand  $T \rightarrow +\infty$  et les quantités  $d_T^{-1} \hat{s}_T^* A_T^i \underline{u}_T$ ,  $i = \overline{1, 2}$  sont bornées en moyenne quadratique. En définitive on obtient que  $\delta_{1, T} = o(1)$  et  $\delta_{2, T} = o(1)$  en probabilité, d'où le résultat du Lemme d'après (1.11).  $\clubsuit$

Nous pouvons alors étudier le niveau asymptotique du test  $\Phi_T$ . Faisant  $\lambda = 0$  dans (1.10), il est clair que si (L) est satisfaite alors, sous l'hypothèse  $\mathcal{H}_0$ ,  $v^{-1/2} V_T(\rho_T)$  a pour loi asymptotique la loi de  $v^{-1/2} V_T(\rho)$  i.e.  $N(0, 1)$ . Par suite les probabilités de fausse alarme attachées au test  $\Phi_T$  défini par (1.7) tendent vers  $\alpha$  quand  $T \rightarrow +\infty$ .

Montrons maintenant que si (L) est satisfaite, alors le test  $\Phi_T$  est convergent i.e. que quand  $T \rightarrow +\infty$  les probabilités de non détection attachées à  $\Phi_T$  tendent vers zéro. En fait, il suffit de montrer que sous  $\mathcal{H}_1$  on a pour  $v \in \mathbb{R}$

$$\lim_{T \rightarrow +\infty} \text{Prob} ([V_T(\rho_T) \leq v]) = 0.$$

Or, faisant  $\lambda = 1$  dans (1.10), on a

$$d_T^{-1} V_T(\rho_T) = d_T^{-1} [V_T(\rho) - (\hat{s}_T^* B_T(\rho) \hat{s}_T)^{1/2}] + d_T^{-1} [(\hat{s}_T^* B_T(\rho) \hat{s}_T)^{1/2} - (1 + d_T) \varepsilon_T]$$

où le premier terme du second membre tend en probabilité vers 0, le dernier terme tendant en probabilité vers  $h^{1/2}(\rho) > 0$ . Alors  $\{d_T^{-1} V_T(\rho_T); T \geq 1\}$  converge en probabilité vers  $h^{1/2}(\rho) > 0$  et donc

$$\begin{aligned} & \lim_{T \rightarrow +\infty} \text{Prob} ([V_T(\rho_T) \leq v]) \\ &= \lim_{T \rightarrow +\infty} \text{Prob} ([d_T^{-1} V_T(\rho_T) \leq d_T^{-1} v]) = 0. \end{aligned}$$

Considérons enfin la suite d'hypothèses alternatives  $\{ \mathcal{H}_{T, d_T^{-1}} ; T \geq 1 \}$  contigues à  $\mathcal{H}_0$  et étudions la puissance asymptotique de  $\Phi_T$  le long de cette suite. Il s'agit d'étudier la limite quand  $T \rightarrow +\infty$  de la quantité

$$1 - \beta_T = \text{Prob} \left( \left[ v^{-1/2} V_T(\rho_T) \leq \mathcal{N}^{-1}(1 - \alpha) \right] \right)$$

où Prob désigne ici la probabilité évaluée sous  $\mathcal{H}_{T, d_T^{-1}}$  i.e. pour  $V_T(\rho_T)$  donné par (1.9) avec  $\lambda = d_T^{-1}$ . Utilisant le Lemme 1 on obtient immédiatement que  $V_T(\rho) - V_T(\rho_T)$  tend vers zéro en probabilité. De plus la loi asymptotique de  $V_T(\rho)$  donnée par (1.9) avec  $\lambda = d_T^{-1}$  est gaussienne  $\mathcal{N}(h^{1/2}(\rho), v)$ . Par suite on a

$$\lim_{T \rightarrow +\infty} (1 - \beta_T) = \mathcal{N} \left( \mathcal{N}^{-1}(1 - \alpha) - v^{-1/2} h^{1/2}(\rho) \right).$$

#### 1.4 Cas d'un bruit AR(1) de paramètres inconnus

On suppose ici que les paramètres  $\rho$  et  $v$  du bruit sont tous les deux inconnus. Compte tenu des considérations du §1.2 où seul  $\rho$  est connu, il est naturel d'envisager un récepteur  $\Phi_T$  défini par une région critique de la forme

$$R_{\Phi_T} = \left\{ y_T : X_T(\rho_T) > \mathcal{N}^{-1}(1 - \alpha) \right\}$$

où  $X_T(r) = \frac{\sqrt{T-1} V_T(r)}{(W_T(r))^{1/2}}$ ,  $V_T(r)$  et  $W_T(r)$  sont donnés respectivement par (1.8)

et

$$(1.12) \quad w_T(r) = y_T^* B_T(r) y_T - (s_T^* B_T(r) s_T)^{-1} (s_T^* B_T(r) y_T)^2$$

et  $\rho_T$  est encore un estimateur de  $\rho$  qui converge dans les conditions asymptotiques

envisagées.

Avec des notations identiques à celles du §1.3, on peut voir par des considérations analogues à celles du Ch.I §3.3.b., que sous l'hypothèse  $\mathcal{H}_{T,\lambda}$  les variables aléatoires  $V_T(\rho)$  et  $W_T(\rho)$  sont indépendantes et que de plus  $V_T(\rho)$  est de loi gaussienne

$\mathcal{N}(\lambda (\mathbf{s}_T^* B_T(\rho) \mathbf{s}_T)^{1/2}, \nu)$  et  $\nu^{-1} W_T(\rho)$  de loi du  $\chi_{T-1}^2$ .

Alors la variable aléatoire  $\frac{\sqrt{T-1} (V_T(\rho) - \lambda (\mathbf{s}_T^* B_T(\rho) \mathbf{s}_T)^{1/2})}{(W_T(\rho))^{1/2}}$  est de loi de student  $\mathcal{T}_{T-1}$  et a pour loi asymptotique la loi  $N(0,1)$ . Notons que  $[\nu(T-1)]^{-1} W_T(\rho)$  converge en probabilité vers 1 et que  $\mathcal{T}_{T-1}(a)$  tend vers  $\mathcal{N}(a)$ .

Nous montrons d'abord

**Lemme 2.** *Supposons que la condition (L) est satisfaite. Alors sous l'hypothèse  $\mathcal{H}_{T,\lambda}$  on a*

$$(T-1)^{-1} [W_T(\rho) - W_T(\rho_T)] = (1 + \lambda d_T + \lambda_T^2 d_T^2) \eta_T$$

où  $\{\eta_T; T \geq 1\}$  est une suite de variables aléatoires qui converge en probabilité vers zéro.

**Preuve.** En réutilisant les notations et considérations de la preuve du Lemme 1.§1.3, on obtient que sous  $\mathcal{H}_{T,\lambda}$  on peut écrire

$$\begin{aligned} W_T(\rho) - W_T(\rho_T) = & \\ & (1 - \delta_{1,T})^{-1} [(\mathbf{s}_T^* B_T(\rho) \mathbf{s}_T)^{-1/2} (\mathbf{s}_T^* B_T(\rho) \mathbf{u}_T) - \delta_{2,T}]^2 + \delta_{3,T} \\ & - (\mathbf{s}_T^* B_T(\rho) \mathbf{s}_T)^{-1} (\mathbf{s}_T^* B_T(\rho) \mathbf{u}_T)^2 \\ (1.13) \quad & - 2\lambda (\mathbf{s}_T^* B_T(\rho) \mathbf{s}_T)^{1/2} (\mathbf{s}_T^* B_T(\rho) \mathbf{s}_T)^{-1/2} (\mathbf{s}_T^* B_T(\rho) \mathbf{u}_T) - \lambda^2 (\mathbf{s}_T^* B_T(\rho) \mathbf{s}_T) \end{aligned}$$

avec  $\delta_{1,T} = o(1)$  et  $\delta_{2,T} = o(1)$  en probabilité et

$$\delta_{3,T} = (\rho_T - \rho) \mathbb{1}_T^* \left[ A_T^1 + (\rho_T + \rho) A_T^2 \right] \mathbb{1}_T.$$

En vertu de la convergence presque sûre de  $\rho_T$  vers  $\rho$  et du fait que la suite de variables aléatoires  $\left\{ T^{-1/2} \mathbb{1}_T^* A_T^i \mathbb{1}_T; T \geq 1 \right\}$  est bornée en moyenne quadratique on a  $\delta_{3,T} = o(1)$  en probabilité. Le résultat du Lemme est alors acquis d'après (1.13).  $\blacktriangleleft$

Raisonnant comme au §1.3. et exploitant les lemmes 1 et 2 on obtient encore que si la condition (L) est satisfaite alors les risques de fausse alarme et de non détection attachés au test  $\Phi_T$  tendent respectivement vers  $\alpha$  et 0.

De même, dans l'asymptotique de la suite d'hypothèses  $\left\{ \mathcal{H}_{T, d_T^{-1}}; T \geq 1 \right\}$  on obtient que pour

$$1 - \beta_T = \text{Prob} \left( \left[ X_T(\rho_T) \leq \mathcal{N}^{-1}(1 - \alpha) \right] \right),$$

où Prob désigne la probabilité évaluée sous  $\mathcal{H}_{T, d_T^{-1}}$ , on a encore

$$\lim_{T \rightarrow +\infty} (1 - \beta_T) = \mathcal{N} \left( \mathcal{N}^{-1}(1 - \alpha) - v^{-1/2} h^{1/2}(\rho) \right).$$

## §2. Détection d'un signal contenant des paramètres inconnus

Ici on suppose que le signal susceptible d'être présent est de la forme  $s_t = \gamma^* z_t$ ,  $t = 1, 2, \dots$  où  $z$  est un signal vectoriel donné à valeurs dans  $\mathbb{R}^k$  et  $\gamma$  un paramètre inconnu dans  $\mathbb{R}^k$ . Le problème de détection est donc celui du test de l'hypothèse  $\mathcal{H}_0 : \gamma = 0$  contre l'hypothèse  $\mathcal{H}_1 : \gamma \neq 0$ .

### 2.1. Cas d'un bruit AR(1) de paramètres connus

Rappelons (cf. §3.1.b du Chapitre II) que

$$(2.1) \quad \bar{\gamma}_T = (Z_T^* B_T(\rho) Z_T)^{-1} Z_T^* B_T(\rho) y_T$$

est un estimateur sans biais de dispersion minimum de  $\gamma$  qui est de loi gaussienne de matrice de covariance  $v(Z_T^* B_T(\rho) Z_T)^{-1}$ . La statistique  $v^{1/2} (Z_T^* B_T(\rho) Z_T)^{1/2} \bar{\gamma}_T$  est ainsi de loi gaussienne de moyenne  $v^{1/2} (Z_T^* B_T(\rho) Z_T)^{1/2} \gamma$  et de matrice de covariance  $\mathbb{I}_k$ .

Il est alors naturel de considérer un test déterministe  $\Phi$  de région critique  $R$  de la forme

$$R = \{ y_T : \left| (Z_T^* B_T(\rho) Z_T)^{1/2} \bar{\gamma}_T \right|^2 > c \}.$$

Comme sous  $\mathcal{H}_0$ ,  $v^{-1} \left| (Z_T^* B_T(\rho) Z_T)^{1/2} \bar{\gamma}_T \right|^2$  est de loi  $\chi_k^2$  on a

$$\alpha_\Phi = 1 - \mathcal{K}_k(c/v),$$

où  $\mathcal{K}_k$  désigne la fonction de répartition de la loi  $\chi_k^2$ .

De même puisque sous  $\mathcal{H}_1$  la variable aléatoire  $v^{-1/2} (Z_T^* B_T(\rho) Z_T)^{1/2} (\bar{\gamma}_T - \gamma)$  est de

loi  $N_k(0, \mathbb{I}_k)$  on a

$$\beta_{\Phi}(\gamma) = N_k(Q, \mathbb{1}_k) \{ \underline{y} \in \mathbb{R}^k : | \underline{y} + v^{-1/2} (Z_T^* B_T(\rho) Z_T)^{1/2} \gamma |^2 > c/v \}.$$

En d'autre termes (cf. BARRA (1970) Ch. VII) on a

$$\beta_{\Phi}(\gamma) = 1 - \mathcal{K}_k \left( v^{-1} \gamma^* Z_T^* B_T(\rho) Z_T \gamma ; c/v \right)$$

où  $\mathcal{K}_k(v ; \cdot)$  désigne la fonction de répartition de loi du Chi deux décentrée, de paramètre de décentrage  $v$ , à  $k$  degrés de liberté.

Notant  $\mathcal{K}_k^{-1}$  la fonction réciproque de  $\mathcal{K}_k$  et choisissant, dans ce qui précède  $c = c_{\alpha}$  où

$$c_{\alpha} = v \mathcal{K}_k^{-1}(1 - \alpha),$$

on définit un récepteur  $\Phi_{\alpha}$  au niveau  $\alpha$  de probabilité de fausse alarme, par la région critique

$$R_{\Phi_{\alpha}} = \{ \underline{y}_T : \underline{y}_T^* B_T(\rho) Z_T (Z_T^* B_T(\rho) Z_T)^{-1} Z_T^* B_T(\rho) \underline{y}_T > v \mathcal{K}_k^{-1}(1 - \alpha) \}.$$

Les probabilités de non détection correspondantes valent

$$\begin{aligned} 1 - \beta_{\Phi_{\alpha}}(\gamma) &= N_k(Q, \mathbb{1}_k) \{ \underline{y} \in \mathbb{R}^k : | \underline{y} + v^{-1/2} (Z_T^* B_T(\rho) Z_T)^{1/2} \gamma |^2 \leq \mathcal{K}_k^{-1}(1 - \alpha) \}. \\ &= \mathcal{K}_k \left( v^{-1} \gamma^* Z_T^* B_T(\rho) Z_T \gamma ; \mathcal{K}_k^{-1}(1 - \alpha) \right). \end{aligned}$$

**Cas particulier.** Lorsque  $k = 1$  puisque  $\mathcal{K}_1^{-1}(1 - \alpha) = (\mathcal{N}^{-1}(1 - \alpha/2))^2$  le test  $\Phi_{\alpha}$  a pour région critique

$$R_{\Phi_{\alpha}} = \left\{ \underline{y}_T : \frac{| (M_T(\rho) \underline{y}_T)^* (M_T(\rho) \underline{z}_T) |}{| M_T(\rho) \underline{z}_T |^{1/2}} > v^{1/2} \mathcal{N}^{-1}(1 - \alpha/2) \right\}$$

et les probabilités de non détection s'écrivent

$$1 - \beta_{\Phi_{\alpha}}(\gamma) = \mathcal{N}(\mathcal{N}^{-1}(1 - \alpha/2) - v^{-1/2} \gamma | M_T(\rho) \underline{z}_T |) - \mathcal{N}(-\mathcal{N}^{-1}(1 - \alpha/2) - v^{-1/2} \gamma | M_T(\rho) \underline{z}_T |).$$

Ce récepteur coïncide bien sûr avec celui que l'on obtient par application des considérations du §3.3.2.a du chapitre I avec

$$\Lambda_T^{-1} = v^{-1} M_T^*(\rho) M_T(\rho) ; s_t = z_t.$$

Il s'agit en fait d'une version bilatère du test défini au §1.1. ci-dessus.

## 2.2. Cas d'un bruit AR(1) de paramètre autorégressif connu

On suppose ici que seul le paramètre  $\rho$  du bruit est connu. Nous avons vu au §3.1.b que les estimateurs  $\bar{\gamma}_T$  et  $v_T^*$  de  $\gamma$  et  $v$  définis respectivement par (2.1) et

$$v_T^* = \frac{1}{T-k} (y_T - Z_T \bar{\gamma}_T)^* B_T(\rho) (y_T - Z_T \bar{\gamma}_T)$$

sont indépendants, de lois respectives  $N_k(\gamma, v(Z_T^* B_T(\rho) Z_T)^{-1})$  et  $\gamma(\frac{T-k}{2}, \frac{2v}{T-k})$ .

On considère alors le test libre  $\Phi$  de région critique

$$R = \left\{ y_T : \frac{|(Z_T^* B_T(\rho) Z_T)^{1/2} \bar{\gamma}_T|}{v_T^*} > c \right\}.$$

Sous  $\mathcal{H}_0$ , comme la variable aléatoire  $v^{-1} |(Z_T^* B_T(\rho) Z_T)^{1/2} \bar{\gamma}_T|^2$  est de loi  $\chi_k^2$  et la variable aléatoire  $v^{-1} (T-k) v_T^*$  est de loi  $\chi_{T-k}^2$ , compte tenu de l'indépendance, la variable aléatoire

$$\frac{1}{k} \frac{|(Z_T^* B_T(\rho) Z_T)^{1/2} \bar{\gamma}_T|^2}{v_T^*},$$

suit la loi  $F_{k, T-k}$  de Fisher-Snedecor à  $k$  et  $T-k$  degrés de liberté.

On a alors

$$\alpha_{\Phi} = 1 - \mathcal{F}_{k, T-k}(c/k),$$

où  $\mathcal{F}_{k, T-k}$  désigne la fonction de répartition de la loi  $F_{k, T-k}$ .

De même, puisque sous  $\mathcal{H}_1$  les variables aléatoires  $v^{1/2} (Z_T^* B_T(\rho) Z_T)^{1/2} (\bar{y}_T - \gamma)$  et  $v^{-1} (T-k) v_T^*$  sont indépendantes de lois respectives  $N_k(\Omega, \mathbb{1}_k)$  et  $\chi_{T-k}^2$ , on peut écrire

$$\beta_{\Phi}(\gamma, v) = \text{Prob} \left[ \frac{|X + v^{1/2} (Z_T^* B_T(\rho) Z_T)^{1/2} \gamma|^2}{Y} > \frac{c}{T-k} \right],$$

où  $X$  et  $Y$  sont deux variables indépendantes de loi  $N_k(\Omega, \mathbb{1}_k)$  et  $\chi_{T-k}^2$  respectivement.

Notons  $\mathcal{F}_{k, T-k}^{-1}$  la fonction réciproque de  $\mathcal{F}_{k, T-k}$  et choisisant, dans ce qui précède  $c = c_{\alpha}$  où

$$c_{\alpha} = k \mathcal{F}_{k, T-k}^{-1}(1 - \alpha),$$

on définit un récepteur  $\Phi_{\alpha}$  libre au niveau  $\alpha$  de probabilité de fausse alarme, par la région critique

$$R_{\Phi_{\alpha}} = \left\{ y_T : \frac{y_T^* B_T(\rho) Z_T (Z_T^* B_T(\rho) Z_T)^{-1} Z_T^* B_T(\rho) y_T}{(y_T - Z_T \bar{y}_T)^* B_T(\rho) (y_T - Z_T \bar{y}_T)} > \frac{k}{T-k} \mathcal{F}_{k, T-k}^{-1}(1 - \alpha) \right\}.$$

Les probabilités de non détection associées valent

$$1 - \beta_{\Phi}(\gamma, v) = \text{Prob} \left[ \frac{|X + v^{1/2} (Z_T^* B_T(\rho) Z_T)^{1/2} \gamma|^2}{Y} \leq \frac{k}{T-k} \mathcal{F}_{k, T-k}^{-1}(1 - \alpha) \right].$$

**Cas particulier.** Lorsque  $k = 1$ , puisque  $\mathcal{F}_{1,T-1}^{-1}(1-\alpha) = (\mathcal{C}_{T-1}^{-1}(1-\alpha/2))^2$ , le test  $\Phi_\alpha$  a pour région critique

$$R_{\Phi_\alpha} = \left\{ \mathbf{y}_T : \frac{|(M_T(\rho) \Sigma_T)^* (M_T(\rho) \mathbf{y}_T)|}{\left[ |M_T(\rho) \Sigma_T|^2 |M_T(\rho) \mathbf{y}_T|^2 - \left\{ (M_T(\rho) \Sigma_T)^* (M_T(\rho) \mathbf{y}_T) \right\}^2 \right]^{1/2}} > \frac{1}{\sqrt{T-1}} \mathcal{C}_{T-1}^{-1}(1-\alpha/2) \right\}$$

et les probabilités de non détection sont données par

$$1 - \beta_{\Phi_\alpha}(\gamma, \nu) = \text{Prob} \left[ \left| V + \nu^{-1/2} \gamma |M_T(\rho) \mathbf{z}_T| \right| \leq \frac{\mathcal{C}_{T-1}^{-1} \sqrt{W}}{\sqrt{T-1}} \right]$$

où  $V$  et  $W$  sont deux variables aléatoires indépendantes de lois respectives  $N(0, 1)$  et  $\chi_{T-1}^2$ .

Ce récepteur coïncide avec celui que l'on obtient par application des éléments du §3.3.2.b du chapitre I avec

$$\Lambda_T^{-1} = M_T^*(\rho) M_T(\rho); s_t = z_t.$$

Il s'agit d'une version bilatère du test défini au §1.2. ci-dessus.

### 2.3. Cas d'un bruit AR(1) de variance d'innovation connue

On suppose ici que seul le paramètre  $\nu$  du bruit est connu. Compte tenu des considérations du §2.1, où  $\rho$  est aussi connu, il est naturel d'envisager un récepteur  $\Phi_T$  défini par une région critique de la forme

$$(2.2) \quad R_{\Phi_T} = \left\{ \mathbf{y}_T : \Delta_T(\rho_T) > \nu \mathcal{K}_k^{-1}(1-\alpha) \right\}$$

où  $\Delta_T(r)$  est donné par

$$(2.3) \quad \Delta_T(r) = \left| \left( Z_T^* B_T(r) Z_T \right)^{-1/2} Z_T^* B_T(r) \mathbf{y}_T \right|^2$$

et  $\rho_T$  est un estimateur de  $\rho$  qui converge au moins en probabilité dans les conditions asymptotiques envisagées ci-après (cf. Ch.II et III).

Notons  $\mathcal{H}_{T, \gamma}$  l'hypothèse " $y_t = \gamma^* z_t + u_t ; t = \overline{1, T}$ ". Alors sous  $\mathcal{H}_{T, \gamma}$  on a

$$(2.4) \quad \Delta_T(\tau) = \underline{u}_T^* B_T(\tau) Z_T \left( Z_T^* B_T(\tau) Z_T \right)^{-1} Z_T^* B_T(\tau) \underline{u}_T \\ + 2 \underline{u}_T^* B_T(\tau) Z_T \gamma + \gamma^* Z_T^* B_T(\tau) Z_T \gamma.$$

Bien sûr (cf. §2.1) sous  $\mathcal{H}_{T, \gamma}$ ,  $v^{-1} \Delta_T(\rho)$  suit la loi du chi deux décentrée, de paramètre de décentrage  $v^{-1} \gamma^* Z_T^* B_T(\rho) Z_T \gamma$ , à  $k$  degrés de liberté.

Dans la suite nous faisons appel à la condition ( $\mathcal{L}$ ) : "(D)  $\wedge$  (A3)", où (D) et (A3) ont été définies au Chapitre II, qui généralise au cas vectoriel la condition (L). Nous pouvons alors énoncer

**Lemme 3** . Supposons que la condition ( $\mathcal{L}$ ) est satisfaite. Alors sous l'hypothèse  $\mathcal{H}_{T, \gamma}$  on a

$$(2.5) \quad \Delta_T(\rho) - \Delta_T(\rho_T) = \varepsilon_T + \eta_T^* D_T \gamma + \gamma^* D_T \Sigma_T D_T \gamma$$

où  $(\varepsilon_T ; T \geq 1)$  (resp.  $(\eta_T ; T \geq 1)$ ,  $(\Sigma_T ; T \geq 1)$ ) est une suite (resp. de vecteurs, matrices) aléatoires qui convergent vers zéro en probabilité.

**Preuve** . Utilisant (2.4) on obtient facilement que sous  $\mathcal{H}_{T, \gamma}$  on peut écrire

$$(2.6) \quad \Delta_T(\rho) - \Delta_T(\rho_T) = \left| \left( Z_T^* B_T(\rho) Z_T \right)^{-1/2} Z_T^* B_T(\rho) \underline{u}_T \right|^2 \\ - \left| \left[ \mathbb{1}_k - \omega_{1, T} \right]^{-1/2} \left[ \left( Z_T^* B_T(\rho) Z_T \right)^{-1/2} Z_T^* B_T(\rho) \underline{u}_T - \omega_{2, T} \right] \right|^2 \\ + 2(\rho_T - \rho) \underline{u}_T^* A_T^1 Z_T^* D_T^1 D_T \gamma - (\rho_T^2 - \rho^2) \gamma^* D_T D_T^1 Z_T^* A_T^2 Z_T D_T^1 D_T \gamma$$

avec

$$\omega_{1, T} = (\rho_T - \rho) D_T D_T^1 Z_T^* (A_T^1 + (\rho_T + \rho) A_T^2) Z_T D_T^1 (D_T^1 Z_T^* B_T(\rho) Z_T D_T^1)^{-1} D_T^1$$

et

$$\omega_{2,T} = (\rho_T - \rho) D_T^{-1/2} (D_T^{-1} Z_T^* B_T(\rho) Z_T D_T^{-1})^{-1/2} D_T^{-1/2} Z_T^* (A_T^1 + (\rho_T + \rho) A_T^2) \underline{u}_T$$

D'après (L),  $D_T^{-1} Z_T^* B_T(\rho) Z_T D_T^{-1}$  converge quand  $T \rightarrow +\infty$  vers  $H(\rho)$  définie dans le Lemme II.4 et  $H(\rho)$  est régulière. De même les quantités  $D_T^{-1} Z_T^* A_T^i Z_T D_T^{-1}; i = \overline{1..2}$  convergent quand  $T \rightarrow +\infty$  et les quantités  $D_T^{-1} Z_T^* A_T^i \underline{u}_T; i = \overline{1..2}$  sont bornées en moyenne quadratique. En définitive on obtient que  $\omega_{1,T} = o(1)$  p.s. et  $\omega_{2,T} = o(1)$  en probabilité, d'où le résultat du Lemme d'après (2.6).  $\clubsuit$

Faisant  $\gamma = 0$  dans (2.5) on voit que si (L) est satisfaite, alors sous  $\mathcal{H}_0$   $v^{-1} \Delta_T(\rho_T)$  a pour loi asymptotique la loi de  $v^{-1} \Delta_T(\rho)$  à savoir la loi  $\chi_k^2$ . Par suite les probabilités de fausse alarme attachées au test  $\Phi_T$  défini par (2.2) tendent vers  $\alpha$  quand  $T \rightarrow +\infty$ .

Etudions maintenant les risques asymptotiques de non détection. Pour ce faire plaçons-nous dans une suite d'hypothèses alternatives  $\{\mathcal{H}_{T, \gamma_T}; T > 1\}$  où

$\{\gamma_T; T > 1\}$  est une suite de vecteurs telle que  $\lim_{T \rightarrow +\infty} D_T^{-1} \gamma_T = l$  où  $l$  est un vecteur de  $\mathbb{R}^k$ . Alors, d'après le Lemme 3,  $\Delta_T(\rho) - \Delta_T(\rho_T)$  tend vers zéro en probabilité. De plus, dans la mesure où

$$\begin{aligned} & v^{-1} \gamma_T^* Z_T^* B_T(\rho) Z_T \gamma_T \\ &= v^{-1} \gamma_T^* D_T (D_T^{-1} Z_T^* B_T(\rho) Z_T D_T^{-1}) D_T \gamma_T \end{aligned}$$

tend alors vers  $v^{-1} l^* H(\rho) l$  où  $H(\rho)$  est donnée par II.(3.11), la loi asymptotique de  $v^{-1} \Delta_T(\rho)$  est une loi du chi deux décentrée, avec paramètre de décentrage  $v^{-1} l^* H(\rho) l$ , à  $k$  degrés de liberté. Par suite les risques asymptotiques de non détection sont donnés par

$$\mathcal{K}_k \left( v^{-1} l^* H(\rho) l ; \mathcal{K}_k^{-1} (1 - \alpha) \right).$$

## 2.4. Cas d'un bruit AR(1) de paramètres inconnus

On suppose ici que les paramètres  $\rho$  et  $v$  du bruit sont tous les deux inconnus. Compte tenu des considérations du §2.2 où seul  $\rho$  est connu, il est naturel d'envisager un récepteur  $\Phi_T$  défini par une région critique de la forme

$$(2.7) \quad R_{\Phi_T} = \left\{ y_T : \frac{\Delta_T(\rho_T)}{\nabla_T(\rho_T)} > \mathcal{K}_k^{-1}(1-\alpha) \right\}$$

où  $\Delta_T(r)$  et  $\nabla_T(r)$  sont donnés respectivement par (2.3) et

$$(2.8) \quad \nabla_T(r) = (T - k)^{-1} (y_T - Z_T \bar{\gamma}_T(r))^* B_T(r) (y_T - Z_T \bar{\gamma}_T(r))$$

et  $\rho_T$  est encore un estimateur de  $\rho$  qui converge dans les conditions asymptotiques envisagées.

Avec des notations identiques à celles du §2.3, on peut voir par des considérations analogues à celles des paragraphes §2.1 et §2.2, que sous l'hypothèse  $\mathcal{H}_{T,\gamma}$  les variables aléatoires  $\Delta_T(\rho)$  et  $\nabla_T(\rho)$  sont indépendantes et que de plus  $v^{-1} \Delta_T(\rho)$  suit la loi du chi deux décentrée, de paramètre de décentrage  $v^{-1} \gamma^* Z_T^* B_T(\rho) Z_T \gamma$ , à  $k$  degrés de

liberté et  $\nabla_T(\rho)$  est de loi  $\chi^2((T - k)/2, 2v/(T - k))$ . Alors la variable aléatoire  $\frac{1}{k} \frac{\Delta_T(\rho)}{\nabla_T(\rho)}$  suit la loi  $F_{k, T-k}$  de Fisher-Snedecor à  $k$  et  $T - k$  degrés de liberté et la variable

aléatoire  $\frac{\Delta_T(\rho)}{\nabla_T(\rho)}$  a pour loi asymptotique la loi du  $\chi_k^2$ .

Notons que  $v^{-1} \Delta_T$  converge en probabilité vers 1.

Par des techniques analogues à celles de la preuve du Lemme 3, on peut montrer à partir de (2.8)

**Lemme 4** . Supposons que la condition  $(\mathcal{L})$  est satisfaite. Alors sous l'hypothèse  $\mathcal{H}_{T, \gamma}$   $\Delta_T(\rho_T) - \Delta_T(\rho)$  vérifie une égalité analogue à (2.5) du Lemme 3.

Raisonnant comme au §2.3 et exploitant les Lemmes 3 et 4 on obtient encore que si la condition  $(\mathcal{L})$  est satisfaite, alors les probabilités de fausse alarme attachées au test  $\Phi_T$  défini par (2.7) tendent vers  $\alpha$  quand  $T \rightarrow +\infty$  .

De même, dans l'asymptotique de la suite d'hypothèses  $\{\mathcal{H}_{T, \gamma_T} ; T > 1\}$  telle qu'au §2.3, on obtient que les risques asymptotiques de non détection attachés au test  $\Phi_T$  défini par (2.7)-(2.8) sont donnés par

$$\mathcal{K}_k \left( v^{-1} l^* H(\rho) l ; \mathcal{K}_k^{-1} (1 - \alpha) \right).$$



ANNEXE

Dans cette partie annexe nous présentons les résultats d'une étude expérimentale par simulation des performances des méthodes d'estimation par moindres carrés (cf.(2.1)-(2.3) chap.III page 80) et par maximum de vraisemblance approché (cf (3.1)-(3.3) chap. III page 94), et de la procédure de détection définie en (2.7) chap.IV page 116 (cas où on suppose la structure du bruit inconnue).

## I. Cadre de l'expérimentation

### I.1 Modèle de base

Cette étude expérimentale est basée sur le modèle particulier suivant :

$$y_t = \gamma_1 \cos(\pi t / 3) + \gamma_2 \sin(\pi t / 3) + u_t ; t > 0,$$

où ( $u_t ; t > 0$ ) est un processus "bruit" AR(1) i.e. autorégressif d'ordre 1 stationnaire réel de paramètre d'autorégression  $\rho$  et de variance d'innovation  $v$ .

### I.2 Plan d'expérimentation

Pour ce modèle, on considère plusieurs situations expérimentales caractérisées par des valeurs différentes des paramètres  $\gamma_1$ ,  $\gamma_2$ ,  $\rho$  et  $v$ . On choisit  $\rho = 0.1$  ou  $0.8$ ,  $v = 0.1$  ou  $0.5$ , et des valeurs de  $\gamma_1$  et  $\gamma_2$  telles que le rapport signal sur bruit varie dans un domaine de valeurs raisonnable. La valeur (asymptotique) de ce rapport est ici :  $(\gamma_1^2 + \gamma_2^2)(1 - \rho^2)/v$ . Pour chaque jeu de valeurs des paramètres on simule 1000 répétitions d'une séquence d'observations de taille  $T=50$  ou  $100$  du processus  $y = (y_t, t = 1, 2, \dots)$  issu du modèle.

## II. Les résultats

### II.1 Estimation

Pour chaque jeu de valeurs des paramètres et pour  $T = 50$  ou  $100$  nous calculons la moyenne empirique et la racine carrée de l'erreur quadratique moyenne empirique des estimateurs obtenus pour  $\gamma_1$ ,  $\gamma_2$ ,  $\rho$  et  $v$  au cours des 1000 répétitions par la méthode des moindres carrés (cf.(2.1)-(2.3) Chap. III) et la méthode de maximum de

vraisemblance approchée (cf.(3.1)-(3.3) Chap. III). Les résultats sont présentés dans les TABLES I à V, colonnes 2 à 5.

On note une grande cohérence des qualités des estimations des paramètres fournies par la méthode des moindres carrés et par celle de maximum de vraisemblance approché. On peut toutefois signaler un léger avantage de la méthode de maximum de vraisemblance approché en ce qui concerne la dispersion des estimateurs des paramètres  $\gamma_1$  et  $\gamma_2$  du signal, l'inverse étant vrai pour le biais.

## II.2 Détection

Pour chaque jeu de valeurs des paramètres et pour  $T = 50$  ou  $100$  nous calculons pour chacune des 1000 répétitions la valeur de la statistique définissant le test (2.7) du chapitre IV à partir des estimations de maximum de vraisemblance approché des paramètres  $\gamma_1$ ,  $\gamma_2$ ,  $\rho$  et  $v$ . On compare cette valeur à la valeur critique correspondant au seuil asymptotique  $\alpha = 0.1$ , i.e. la valeur 4.61 obtenue dans la table de la loi du chi-deux à deux degrés de liberté (cf. §2.4 Chap. IV).

Selon qu'il n'y a pas de signal (i.e.  $\gamma_1 = \gamma_2 = 0$ ) ou qu'un signal est présent (en fait  $\gamma_1 = \gamma_2 \neq 0$  ou  $\gamma_1 = -\gamma_2 \neq 0$ ) on calcule :

-les fréquences empiriques de fausse alarme au cours des 1000 répétitions)(cf. TABLES I et III, Colonnes 6)

-ou les fréquences empiriques de non détection au cours des 1000 répétitions (cf. TABLES II, IV et V, Colonnes 6; les colonnes 7 de ces tableaux fournissent la valeur du rapport signal sur bruit).

On peut faire les remarques suivantes :

En l'absence de signal on obtient une fréquence de fausse alarme très proche du seuil asymptotique théorique  $\alpha = 0.1$ ; cela valide l'utilisation de la loi asymptotique du chi deux à deux degrés de liberté pour le calcul du seuil pour une taille d'échantillon assez faible.

En présence d'un signal nous constatons évidemment que la fréquence de non détection est d'autant plus faible que le rapport signal sur bruit est important.

**TABLE I**

**Estimation de moindres carrés, de maximum de vraisemblance approchée et Détection**

Nombre de répétitions = 1000

Longueur de la série observée = 50

Le seuil = 0.1

Modèle	Moyenne Empirique Moindres carrés	Racine Erreur quadratique moyenne Moindres carrés	Moyenne Empirique Maximum Vraisemblance approché	Racine Erreur quadratique moyenne maximum vraisemblance approché	Fréquence Empirique de fausse alarme
$\gamma_1 = 0$	-.0007	.0665	-.0007	.0663	.096
$\gamma_2 = 0$	-.0002	.0614	-.0003	.0608	
$\rho = 0.1$	.0749	.1512	.0753	.1515	
$\nu = 0.1$	.0937	.0209	.0937	.0209	
$\gamma_1 = 0$	.0002	.1544	.0006	.1549	.133
$\gamma_2 = 0$	-.0011	.1470	-.0020	.1463	
$\rho = 0.1$	.0766	.1466	.0785	.1468	
$\nu = 0.5$	.4696	.1043	.4694	.1044	
$\gamma_1 = 0$	-.0001	.0732	-.0006	.0704	.109
$\gamma_2 = 0$	.0012	.0695	.0029	.0658	
$\rho = 0.8$	.7668	.1036	.7664	.1038	
$\nu = 0.1$	.0991	.0213	.0992	.0214	
$\gamma_1 = 0$	-.0066	.1619	-.0048	.1575	.123
$\gamma_2 = 0$	.0023	.1678	.0011	.1601	
$\rho = 0.8$	.7762	.0981	.7763	.0984	
$\nu = 0.5$	.4851	.1060	.4856	.1064	

**TABLE II**

**Estimation de moindres carrés, de maximum de vraisemblance approchée et Détection**

Nombre de répétitions = 1000  
 Longueur de la série observée = 50  
 Le seuil = 0.1

Modèle	Moyenne Empirique Moindres carrés	Racine Erreur quadratique moyenne Moindres carrés	Moyenne Empirique Maximum Vraisemblance approché	Racine Erreur quadratique moyenne maximum vraisemblance approché	Fréquence Empirique de non détection	Rapport Signal sur Bruit
$\gamma_1 = .15$	.1537	.0661	.1528	.0660	.098	4.65
$\gamma_2 = .15$	.1458	.0683	.1387	.0688		
$\rho = 0.1$	.0763	.1538	.0765	.1539		
$\nu = 0.1$	.0933	.0210	.0933	.0210		
$\gamma_1 = .15$	.1511	.0686	.1505	.0668	.138	1.8
$\gamma_2 = .15$	.1507	.0710	.1465	.0697		
$\rho = 0.8$	.7717	.0975	.7717	.0978		
$\nu = 0.1$	.0977	.0207	.0978	.0208		
$\gamma_1 = .15$	.1527	.1469	.1518	.1467	.645	.99
$\gamma_2 = .15$	.1478	.1425	.1392	.1405		
$\rho = 0.1$	.0771	.1423	.0769	.1425		
$\nu = 0.5$	.4736	.1025	.4734	.1026		
$\gamma_1 = .15$	.1496	.1626	.1496	.1572	.662	.36
$\gamma_2 = .15$	.1480	.1626	.1437	.1553		
$\rho = 0.8$	.7752	.0993	.7753	.0996		
$\nu = 0.5$	.4891	.1093	.4895	.1098		

**TABLE III**

**Estimation de moindres carrés, de maximum de vraisemblance approchée et Détection**

Nombre de répétitions = 1000

Longueur de la série observée = 100

Le seuil = 0.1

Modèle	Moyenne Empirique Moindres carrés	Racine Erreur quadratique moyenne Moindres carrés	Moyenne Empirique Maximum Vraisemblance approché	Racine Erreur quadratique moyenne maximum vraisemblance approché	Fréquence Empirique de fausse alarme
$\gamma_1 = 0$	.0011	.0459	.0011	.0460	.114
$\gamma_2 = 0$	-.0007	.0467	-.0008	.0465	
$\rho = 0.1$	.0859	.0996	.0858	.0997	
$\nu = 0.1$	.0969	.0138	.0969	.0138	
$\gamma_1 = 0$	.0000	.1068	-.0001	.1067	.105
$\gamma_2 = 0$	-.0040	.1052	-.0036	.1036	
$\rho = 0.1$	.0920	.1042	.0919	.1043	
$\nu = 0.5$	.4864	.0709	.4864	.0709	
$\gamma_1 = 0$	.0014	.0504	.0013	.0489	.106
$\gamma_2 = 0$	.0003	.0481	.0002	.0479	
$\rho = 0.8$	.7867	.0621	.7868	.0622	
$\nu = 0.1$	.0988	.0142	.0988	.0142	
$\gamma_1 = 0$	.0031	.1127	.0031	.1094	.106
$\gamma_2 = 0$	.0006	.1076	.0003	.1072	
$\rho = 0.8$	.7867	.0621	.7868	.0622	
$\nu = 0.5$	.4940	.0711	.4941	.0712	

**TABLE IV**

**Estimation de moindres carrés, de maximum de vraisemblance approchée et Détection**

Nombre de répétitions = 1000

Longueur de la série observée = 100

Le seuil = 0.1

Modèle	Moyenne Empirique Moindres carrés	Racine Erreur quadratique moyenne Moindres carrés	Moyenne Empirique Maximum vraisemblance approché	Racine Erreur quadratique moyenne Vmaximum vraisemblance approché	Fréquence Empirique de non détection	Rapport Signal sur Bruit
$\gamma_1 = .15$	.1500	.0477	.1495	.0477	.005	4.65
$\gamma_2 = .15$	.1482	.0471	.1441	.0466		
$\rho = 0.1$	.0920	.1042	.0919	.1043		
$\nu = 0.1$	.0973	.0142	.0973	.0142		
$\gamma_1 = .15$	.1496	.0703	.1493	.0681	.137	1.8
$\gamma_2 = .15$	.1496	.0727	.1447	.0692		
$\rho = 0.8$	.7711	.0981	.7708	.0987		
$\nu = 0.1$	.0976	.0204	.0977	.0205		
$\gamma_1 = .15$	.1470	.1038	.1464	.1038	.463	.99
$\gamma_2 = .15$	.1519	.1015	.1473	.1009		
$\rho = 0.1$	.0929	.0998	.0929	.0999		
$\nu = 0.5$	.4880	.0724	.4879	.0724		
$\gamma_1 = .15$	.1457	.1083	.1437	.1053	.510	.36
$\gamma_2 = .15$	.1454	.1057	.1452	.1080		
$\rho = 0.8$	.7901	.0651	.7902	.0652		
$\nu = 0.5$	.4935	.0695	.4936	.0695		

**TABLE V**

**Estimation de moindres carrés, de maximum de vraisemblance approchée et Détection**

Nombre de répétitions = 1000

Longueur de la série observée = 100

Le seuil = 0.1

Modèle	Moyenne Empirique Moindres carrés	Racine Erreur quadratique moyenne Moindres carrés	Moyenne Empirique Maximum Vraisemblance approché	Racine Erreur quadratique moyenne maximum vraisemblance approché	Fréquence Empirique de non détection	Rapport Signal sur Bruit
$\gamma_1 = -.2$	-.1931	.1141	-.1939	.1102	.267	0.115
$\gamma_2 = .2$	.2022	.1056	.2010	.1049		
$\rho = 0.8$	.7844	.0666	.7845	.0666		
$\nu = 0.5$	.4960	.0711	.4960	.0712		
$\gamma_1 = .1$	.1006	.0474	.1004	.0474	.144	1.98
$\gamma_2 = -.1$	-.0989	.0470	-.0978	.0468		
$\rho = 0.1$	.0898	.1041	.0898	.1041		
$\nu = 0.1$	.0971	.0143	.0971	.0143		
$\gamma_1 = -.1$	-.0974	.0471	-.0976	.0470	.137	1.98
$\gamma_2 = .1$	.1014	.0474	.1006	.0469		
$\rho = 0.1$	.0951	.1040	.0951	.1041		
$\nu = 0.1$	.0967	.0143	.0967	.0143		
$\gamma_1 = .15$	.1493	.0464	.1492	.0465	0.007	4.65
$\gamma_2 = -.15$	-.1505	.0468	-.1493	.0463		
$\rho = 0.1$	.0830	.1069	.0830	.1070		
$\nu = 0.1$	.0974	.0144	.0974	.0144		

## BIBLIOGRAPHIE

- J.R. BARRA (1971) : Notion Fondamentales de Statistique Mathématique. Dunod, Paris.
- D. COCHRANE et G.H. ORCUTT (1949) : "Application of least squares regression to relationships containing autocorrelated error terms". J. Amer. Statist. Assoc., Vol. 44, 32-61.
- D. DACUNHA-CASTELLE et M. DUFLO (1983) : Probabilités et Statistiques, tome 2. Problèmes à temps mobile. Masson, Paris.
- A.J. DOROGOVCEV (1973) : On asymptotical properties of least squares estimates for regression coefficients. Problemy Predachi Informatsii IX, 4, 49-57.
- J. DURBIN et G.S. WATSON (1950) : "Testing for serial correlation in least squares regression I". Biometrika, 37, 409-428.
- F.J. FERGUSON (1967) : Mathematical statistics : a decision theoretic approach. Academic Press, New York.
- B. GAREL (1989) : "The asymptotic distribution of the likelihood ratio for M.A. processes with a regression trend". Statistics and Decisions, N°1-2, 167-184.
- A.C. HARVEY et G.D.A PHILLIPS (1979) : "Maximum likelihood estimation of regression models with autoregressive-moving average disturbances". Biometrika., Vol. 66, 49-58.
- D. HASZA (1980) : "A note on maximum likelihood estimation for the first-order autoregressive process". Comm. Statist. - Theo. Math., A9 (13), 1411-1415.
- C.W. HELSTROM (1968) : Statistical theory of signal detection. Pergamon Press, New York.
- C. HILDRETH (1969) : "Asymptotic distribution of maximum likelihood estimators in a linear model with autoregressive disturbances". Ann. Math. Statist., Vol. 46, 583-594.

M.G. KENDALL (1954) : "Note on biais in the estimation of autocorrelation".  
Biometrika , Vol. 41.

A.S. KHOLEVO (1969) : "On estimates of regression coefficients". Theory. Probab.  
Appl. XIV 1, 79-104.

T.L. LAI, H.ROBBINS et C.Z. WEI (1979) : "Strong consistency of least squares  
estimates in multiple regression II". J. Multiv. Anal., Vol. 9, N°3, 343-361.

A. LE BRETON (1989) : "A note on maximum likelihood estimation for the complex-  
valued first-order autoregressive process". Statistics and Probability letters, 7, 171-173.

A. LE BRETON et M. MUSIELA (1987a) : "Strong consistency of least squares  
estimates in multiple regression models driven by semimartingales". J. Multivariate  
Anal., 23, 77-92.

A. LE BRETON et M. MUSIELA (1987b) : "A strong law of large numbers for vector  
gaussian martingales and a statistical application in linear regression". Statistics and  
Probability Letters, 5, 71-73.

A. LE BRETON et M. MUSIELA (1989) : "Order of convergence of regression  
parameter estimates in models with infinite variance". J. Multivariate Anal., 31, 59-68.

A. LE BRETON et D.T. PHAM (1989) : "On the biais of the least squares estimator for  
the first order autoregressive process". Ann. Inst. Statist. Math. N°3, 555-563.

E.L. LEHMANN (1959) : Testing statistical hypotheses. Wiley, New York.

L. MAGEE ; A. ULLAH et V.K SRIVASTAVA (1987) : "Efficiency of estimators in  
the regression model with first order autoregressive errors". In : Specification analysis  
in the linear model. King et Giles Ed. p. 81-98.

R.A. MALLER (1989) : "Regression with autoregressive errors- Some asymptotic  
results". Statistics, Vol. 20, 23-39.

P. MERMOZ (1971) : Antennes de détection optimales et adaptatives. Collection  
technique et scientifique du CNET.

A. MONFORT (1982) : Cours de Statistique Mathématique. Economica, Paris.

- J. NEVEU (1970) : Bases mathématiques du calcul des probabilités. Masson, Paris.
- J. NEVEU (1972) : Martingales à temps discret. Masson, Paris.
- W. OBERHOFER et J. KMENTA (1974) : "A general procedure for obtaining maximum likelihood estimates in generalised regression models". *Econometrica*, Vol. 42, 579-590.
- M.H. PESARAN (1973) : "Exact maximum likelihood estimation of regression models with a first order moving average errors". *Rev. Econ. Studics.*, Vol. 40, 529-535.
- D.T. PHAM (1975) : Estimation et tests dans les modèles paramétriques de processus stationnaires. Thèse, Université de Grenoble.
- B. PICINBONO (1987) : "Détection et estimation : problèmes et perspectives". Présenté au XI-ème colloque GRETSI, Nice.
- A.R. SWENSEN (1985) : "The asymptotic distribution of the likelihood ratio for autoregressive time series with a regression trend". *J. Multivariate. Anal.*, 16, 54-70.
- H.L. VAN TREES (1971) : Detection, estimation and modulation theory. Wiley, New York.
- G. VEZZOSI (1976) : Détection d'un signal dans un bruit de loi incertaine. Thèse, Université de Paris Sud.
- A.D. WHALEN (1971) : Detection of signals in noise. Academic press, New York.
- J.S. WHITE (1961) : "Asymptotic expansion for the mean and variance of the serial correlation coefficient". *Biometrika*, Vol. 48, 85-95.
- J.S. WHITE (1958) : "The limiting distribution of the serial correlation coefficient in the explosive case". *Ann. Statist.*, Vol. 29, 1188-1197.
- C.K WONG (1971) : Stochastic processes in information and dynamical systems. Mc. Graw-Hill.
- T. YAMAMOTO et N. KUNITOMO (1984) : "Asymptotic bias of least squares estimator of multivariate autoregressive models". *Ann. Inst. Statist. Math.*, Vol. 36, 419-430.



AUTORISATION DE SOUTENANCE

DOCTORAT DE L'UNIVERSITE JOSEPH FOURIER - GRENOBLE 1

Vu les dispositions de l'Arrêté du 5 juillet 1984,

Vu les rapports de M<sup>r</sup>... François... BRODEAU.....

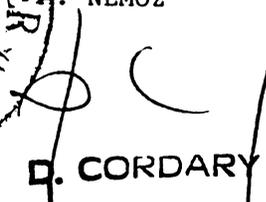
M<sup>r</sup>... Bernard... GAREL.....

M<sup>r</sup>... Al-de ssamad... EZZAHAR..... est autorisé(e)  
à présenter une thèse en vue de l'obtention du doctorat... de.....  
l'Université... Joseph... FOURIER.....

Grenoble, le ..... 8 OCT. 1991

Le Président de l'Université  
Joseph Fourier - Grenoble 1



A. NEMOZ  
  
D. CORDARY  
Vice-Président

## **Résumé**

*L'estimation et la détection d'un signal à partir d'un enregistrement bruité relève des problématiques de la décision statistique : estimation de paramètres et test d'hypothèses.*

*Nous considérons un modèle "signal plus bruit" particulier où le signal est une combinaison linéaire de suites déterministes données et est contaminé par un bruit additif autorégressif d'ordre un stationnaire.*

*Nous étudions d'abord des problèmes d'estimation partielle. On analyse les propriétés asymptotiques d'estimateurs de maximum de vraisemblance ou de moindres carrés pour les paramètres du bruit lorsque le signal est complètement connu ou pour les paramètres du signal lorsque l'un des paramètres du bruit est connu.*

*Puis nous examinons le problème de l'estimation simultanée des paramètres du signal et du bruit. On montre l'existence et l'unicité de l'estimateur de maximum de vraisemblance dont on étudie le comportement asymptotique. De même on considère une méthode d'estimation fondée sur une première étape de moindres carrés pour l'estimation des paramètres du signal, et une procédure de maximum de vraisemblance approché.*

*On construit ensuite des tests pour la détection du signal à partir des méthodes d'estimation envisagées précédemment. Les risques associés à ces tests sont analysés de manière précise.*

*Enfin une étude expérimentale par simulation des performances des diverses méthodes d'estimation et de détection est menée.*

## **Mots clés**

*Régression linéaire - Processus autorégressif - Estimation - Maximum de vraisemblance - Moindres carrés - Convergence presque sûre - Normalité asymptotique - Détection.*