Interprétation de la divergence de Jeffrey entre processus de type autorégressif et/ou à moyenne ajustée

Éric GRIVEL[1], Léo LEGRAND[1,2], Audrey GIREMUS[1]

[1]Université de Bordeaux - Bordeaux INP, ENSEIRB-MATMECA - IMS - UMR CNRS 5218, Talence, France

[2]Thales Systèmes Aéroportés, Mérignac, France

eric.grivel@ims-bordeaux.fr, leo.legrand@ims-bordeaux.fr
audrey.giremus@ims-bordeaux.fr

Résumé – La divergence de Jeffrey (DJ), version symétrique de la divergence de Kullback-Leibler, est utile dans de nombreuses applications pour classifier, détecter des changements, *etc.* Pour cette raison, plusieurs études ont été menées sur la DJ entre processus autorégressifs et/ou à moyenne ajustée. Elles indiquent que la différence entre deux DJs fondées sur k et $k-1$ valeurs successives des processus tend vers une valeur constante quand k augmente, à l'exception de DJs qui impliquent un processus MA d'ordre 1 dont la densité spectrale de puissance (DSP) est caractérisée par l'absence totale d'une fréquence. Cet incrément dit asymptotique est en outre suffisant pour comparer ces processus aléatoires. Bien que de nombreux développements mathématiques aient été menés pour aboutir à ces résultats, une interprétation physique de l'incrément asymptotique serait utile. L'objet de cette communication est d'en fournir une : l'incrément asymptotique de la DJ entre deux processus AR et/ou MA se ramène à calculer la puissance du premier processus filtré par le filtre inverse associé au second et réciproquement. Cela permet d'expliquer les cas atypiques qui ont pu être identifiés dans des travaux passés et de les généraliser à tous les processus ARMA d'ordre quelconque dont la DSP est ponctuellement nulle. Différents exemples sont proposés pour illustrer notre propos.

Abstract – Jeffrey's divergence (JD), which is the symmetric version of the Kullback-Leibler divergence, is useful in many applications for classification, change detection, *etc.* For this reason, several studies were carried out and deal with the JD between processes that are autoregressive (AR) and/or moving average (MA). They state that the difference between two JDs based on k and $k-1$ successive samples tends to a constant value when k increases, except JDs which involve 1st-order MA process whose power spectral density (PSD) is characterized by the total absence of one frequency. This so-called asymptotic increment is also sufficient to compare the processes. The purpose of this communication is to propose an interpretation of the asymptotic increment: it consists in calculating the power of the first process filtered by the inverse filter associated with the second and conversely. This explains the atypical cases that were identified in the past and generalizes them to any ARMA process of any order whose PSD is characterized by the absence of one or more frequencies. Different examples are proposed to illustrate this phenomenon.

1 Introduction

La divergence de Kullback-Leibler (KL), celle de Pearson (DP) ou la divergence de Pearson relative visent à comparer des distributions de jeux de données. De nombreux auteurs les ont analysées dans différents contextes applicatifs, notamment [1] et [2], à des fins de classification ou de détection de changement. Dans ce dernier cas, il s'agit de détecter un changement dans les propriétés statistiques du signal au cours du temps.

Des estimations de la divergence de KL ou de la DP entre deux densités de probabilité non nécessairement gaussiennes ont ainsi été étudiées. Dans ce cas, au lieu d'estimer les deux densités de probabilité à partir des données, leur rapport est directement estimé. En ce qui concerne la divergence de KL, cela donne lieu à la procédure nommée KLIEP et sa variante "Gaussian mixture KLIEP" [3] ou à des méthodes fondées sur des M-estimateurs [4].

La version symétrique de la divergence de KL, aussi connue sous le nom de divergence de Jeffrey (DJ), est souvent utilisée. Dans le cas de vecteurs aléatoires gaussiens de taille k, la DJ se ramène à calculer la somme de deux traces de matrices, s'exprimant comme une matrice de covariance de taille $k \times k$ d'un premier processus aléatoire pré-multipliée par l'inverse de la matrice de covariance d'un second. Plus k augmente, plus la complexité calculatoire relative au calcul de la DJ reposant sur l'inversion des matrices de covariance est élevée. Pour pallier ce problème, des approches alternatives au calcul direct pourraient être envisagées. Dans le cas de processus de type autorégressif (AR) et/ou à moyenne ajustée (MA), on pourrait *a priori* opter pour une décomposition en valeurs propres des matrices de covariance lorsque les processus sont d'ordre 1. En effet, des expressions analytiques des valeurs propres et des vecteurs propres pour des processus MA d'ordre 1 existent [5] et des valeurs approximatives des valeurs propres ont été proposées dans le cas des processus AR pour des grandes tailles de matrice de corrélation [6]. Pour traiter des cas où l'ordre p des processus AR est supérieur à 1, une factorisation de Cholesky alternative (appelée en anglais LDL pour faire référence aux structures "lower unit" et "diagonal" des matrices issues de la décomposition) pourrait être envisagée ; elle dépend des paramètres des processus AR dont l'ordre varie de 1 à p. D'autres démarches peuvent être plus pertinentes. En tirant profit du caractère markovien d'un processus AR, la DJ peut se déduire de manière récursive dans le cas de processus AR stationnaires ou variant dans le temps [7]. Cette méthode a ensuite été étendue dans le cas de la classification de plus de deux processus AR [8]. Par ailleurs, l'expression analytique de la DJ pour deux processus MA d'ordre 1, à valeurs réelles ou complexes, sans bruit ou perturbés par un bruit blanc additif gaussien, a été étudiée dans [9]. Dans ce travail, nous nous étions appuyés sur une expression analytique de chaque élément de l'inverse de la matrice de corrélation des processus MA, en exploitant les résultats sur les matrices tridiagonales [10]. Plus récemment, la comparaison entre processus AR et MA d'ordre 1 à partir de la DJ a été abordée dans [11] en se fondant sur un calcul détaillé des inverses des matrices de corrélation [12].

Dans ces différents cas, des liens ont été établis avec la distance de Rao [13] et ont confirmé que le carré de la distance de Rao valait approximativement deux fois la valeur de la DJ sauf lorsque le module du zéro associé à un processus MA d'ordre 1 se rapproche de 1. Nous avons aussi pu constater que l'évolution de la DJ en fonction du nombre de variables k tend vers un régime permanent, c'est-à-dire que la différence entre deux DJs fondées sur k et $k-1$ échantillons tend vers une valeur constante quand k augmente, à l'exception de cas particuliers : si la DJ implique un processus MA d'ordre 1 dont le paramètre MA a un module égal à 1, alors l'évolution de l'incrément de la DJ en fonction du temps ne tend plus vers une constante mais dépend de k. L'incrément asymptotique tend donc vers l'infini.

Nous avons aussi montré que cet incrément asymptotique est suffisant pour comparer ces processus aléatoires ; Alors que la DJ nécessite le choix de k, ce n'est pas le cas de l'incrément asymptotique. Il dépend uniquement des paramètres des processus, ce qui est un avantage.

Dans cette communication, nous proposons une démarche alternative pour obtenir l'incrément asymptotique de la DJ afin d'en faciliter son interprétation et de mieux comprendre les variations de cette grandeur en fonction des paramètres des processus étudiés.

Dans la suite, I_k désigne la matrice identité de taille k et Tr la trace d'une matrice. L'exposant ^{T} définit la transposée. $x_{k_1:k_2} = (x_{k_1}, ..., x_{k_2})$ et $l = 1, 2$.

2 Rappels sur les processus AR/MA

Soit le $l^{\text{ème}}$ processus autorégressif à moyenne ajustée (ARMA) d'ordre p et q respectif. Son $n^{\text{ème}}$ échantillon, x_n^l, est défini comme suit :

$$x_n^l = -\sum_{g=1}^{p} a_g^l x_{n-g}^l + \sum_{h=0}^{q} b_h^l u_{n-h}^l \tag{1}$$

où u_n^l est le processus générateur blanc et gaussien, de moyenne nulle et de variance $\sigma_{u,l}^2$. En posant $a_0^l = 1$, $\{a_g^l\}_{g=0,\dots,p}$ désignent les paramètres AR du $l^{\text{ème}}$ processus. En outre, en prenant $b_0^l = 1$, $\{b_h^l\}_{h=0,\dots,q}$ sont les paramètres MA. Quand $\{b_h^l\}_{h=1,\dots,q} = 0$, le $l^{\text{ème}}$ processus est dit AR ; si $\{a_g^l\}_{g=1,\dots,p} = 0$, il est dit MA.

Étant donné (1), ce processus ARMA peut être vu comme le filtrage d'un bruit blanc de moyenne nulle et de variance unité dont la fonction de transfert [1] $H_l(z)$ est définie par ses pôles $\{p_g^l\}_{g=1,\dots,p}$ et ses zéros $\{z_h^l\}_{h=1,\dots,q}$. Sa densité spectrale de puissance évaluée à la pulsation normalisée θ vaut :

$$S_{ARMA,l}(\theta) = \sigma_{u,l}^2 \frac{|\sum_{h=0}^{q} b_h^l e^{-jh\theta}|^2}{|\sum_{g=0}^{p} a_g^l e^{-jg\theta}|^2} \tag{2}$$

Le filtre inverse est alors défini par la fonction de transfert $H_l^{-1}(z)$. Dans la suite, pour $|b_1^1| \neq 1$, on opte pour une version stable du filtre inverse. À titre d'exemple, prenons le cas d'un processus MA réel d'ordre 1 et de paramètre MA b_1^1 ; si $|b_1^1| < 1$, le filtre inverse est stable ; si $|b_1^1| > 1$, on tire avantage de l'égalité suivante :

$$b_1^1 \left(1 + \frac{1}{b_1^1} z^{-1}\right) = \frac{b_1^1 + z^{-1}}{1 + b_1^1 z^{-1}} \left(1 + b_1^1 z^{-1}\right) \tag{3}$$

où $\frac{b_1^1+z^{-1}}{1+b_1^1 z^{-1}}$ est un produit de Blaschke, correspondant à une fonction de transfert d'un filtre passe-tout. Dans ce cas, le filtre dit "inverse" est défini par la fonction de transfert $H_1^{-1}(z) = \frac{1}{\sigma_{u,1} b_1^1 \left(1+\frac{1}{b_1^1} z^{-1}\right)}$. Si $|b_1^1| = 1$, le filtre inverse est nécessairement instable.

3 Analyse de la divergence de Jeffrey

3.1 Rappel sur la définition de la divergence de Jeffrey

Pour évaluer la dissimilarité entre les distributions de vecteurs concaténant k valeurs successives de deux processus aléatoires, définies par $p_1(x_{1:k})$ et $p_2(x_{1:k})$, la divergence de KL vérifie [14] :

$$KL_k^{(1,2)} = \int_{x_{1:k}} p_1(x_{1:k}) \ln\left(\frac{p_1(x_{1:k})}{p_2(x_{1:k})}\right) dx_{1:k} \tag{4}$$

Quand les processus sont gaussiens, réels, de moyennes $\mu_{1,k}$ et $\mu_{2,k}$ et de matrices de covariance $Q_{1,k}$ et $Q_{2,k}$, on peut aisément déduire que la divergence de KL satisfait [2] [15] :

$$KL_k^{(1,2)} = \frac{1}{2}\Big[Tr\left(Q_{2,k}^{-1} Q_{1,k}\right) - k - \ln\frac{\det(Q_{1,k})}{\det(Q_{2,k})} + \left(\mu_{2,k} - \mu_{1,k}\right)^T Q_{2,k}^{-1} \left(\mu_{2,k} - \mu_{1,k}\right)\Big] \tag{5}$$

Cependant, comme la divergence de KL n'est pas symétrique, la DJ est souvent utilisée ; elle est définie comme suit :

1. $H_l(z) = \sigma_{u,l} \frac{\prod_{h=1}^{q}(1 - z_h^l z^{-1})}{\prod_{g=1}^{p}(1 - p_g^l z^{-1})}$

2. À noter que dans le cas complexe, ^{T} est remplacé par ^{H} et $\frac{1}{2}$ disparaît dans l'expression (5).

$$DJ_k^{(1,2)} = \frac{1}{2}\left(KL_k^{(1,2)} + KL_k^{(2,1)}\right) \tag{6}$$

Dans le cas de processus aléatoires à moyenne nulle et étant donné (6), l'expression de la DJ se réduit alors à :

$$DJ_k^{(1,2)} = -k + \frac{1}{2}\left[Tr\left(Q_{2,k}^{-1}Q_{1,k}\right) + Tr\left(Q_{1,k}^{-1}Q_{2,k}\right)\right] \tag{7}$$

3.2 Comportement asymptotique de la DJ et interprétation

L'expression (7) se ramène au calcul de deux traces. Dans la suite, nous étudions le comportement de la DJ quand k augmente. En particulier, notre objectif est d'évaluer l'incrément asymptotique de la DJ, défini comme suit :

$$\Delta DJ^{(1,2)} = -1 + \frac{1}{2}\left[\Delta T_{2,1} + \Delta T_{1,2}\right] \tag{8}$$

où les incréments asymptotiques $\Delta T_{i,j}$ des traces vérifient la relation suivante -le couple (i,j) étant égal à $(1,2)$ ou $(2,1)$- :

$$\Delta T_{i,j} = \lim_{k\to+\infty}\left[Tr\left(Q_{i,k}^{-1}Q_{j,k}\right) - Tr\left(Q_{i,k-1}^{-1}Q_{j,k-1}\right)\right] \tag{9}$$

Afin d'évaluer $\Delta T_{i,j}$, nous proposons dans cette communication une interprétation facilitant les développements mathématiques : supposons que le processus ARMA soit associé à une matrice de covariance $Q_{2,k}$ qui admet une décomposition en valeurs propres de la forme $P_2 D_2 P_2^T$, où P_2 désigne la matrice unitaire des vecteurs propres de $Q_{2,k}$ et D_2 la matrice diagonale des valeurs propres. Il vient alors :

$$Tr\left(Q_{2,k}^{-1}Q_{1,k}\right) = Tr\left(D_2^{-1/2}P_2^T Q_{1,k} P_2 D_2^{-1/2}\right) \tag{10}$$

L'interprétation suivante peut éviter de nombreux calculs fastidieux. Soit $X_{l,k}$ le vecteur aléatoire associé à $Q_{l,k}$. Prémultiplier $X_{2,k}$ par $D_2^{-1/2}P_2^T$ revient à blanchir le processus $X_{2,k}$ et à faire que le signal résultant soit de puissance unité. Quand k tend vers l'infini, cela revient à effectuer un filtrage "inverse" de fonction de transfert $H_2^{-1}(z)$ sur tous les échantillons stockés dans $X_{2,k}$.

De manière similaire, prémultiplier $X_{1,k}$ par $D_2^{-1/2}P_2^T$ revient quand k tend vers l'infini, à effectuer un filtrage "inverse" de fonction de transfert $H_2^{-1}(z)$ sur tous les échantillons stockés dans $X_{1,k}$.

Par conséquent, l'incrément asymptotique $\Delta T_{2,1}$ de la trace $Tr\left(Q_{2,k}^{-1}Q_{1,k}\right)$ correspond à la puissance véhiculée par le processus dont les échantillons sont stockés dans $X_{1,k}$ -avec k tendant vers l'infini- et filtré par $H_2^{-1}(z)$.

Dans la prochaine partie, appliquons ce principe à différents cas afin de confirmer les résultats obtenus dans nos travaux précédents [7], [9], [11] où la DJ était exprimée pour toute valeur de k, la différence entre deux DJs consécutives était calculée et enfin la limite de cet incrément quand k tend vers l'infini était analysée.

4 Applications

4.1 DJ entre deux bruits blancs

Même si ce cas est trivial en partant de la définition (7), on obtient en utilisant la démarche proposée ci-dessus que l'incrément asymptotique dans le cas de bruits blancs (BB) gaussiens de moyenne nulle et de variance $\sigma^2_{(u,1)}$ et $\sigma^2_{(u,2)}$ vaut :

$$\Delta DJ^{(BB1,BB2)} = -1 + \frac{1}{2}\left(\frac{\sigma^2_{(u,1)}}{\sigma^2_{(u,2)}} + \frac{\sigma^2_{(u,2)}}{\sigma^2_{(u,1)}}\right) \tag{11}$$

En effet, le terme $\frac{\sigma^2_{(u,1)}}{\sigma^2_{(u,2)}}$ s'explique ainsi : le premier bruit blanc de variance $\sigma^2_{(u,1)}$ est filtré par un filtre passe-tout de fonction de transfert $\frac{1}{\sigma_{(u,2)}}$. La puissance du signal filtré vaut alors $\frac{\sigma^2_{(u,1)}}{\sigma^2_{(u,2)}}$. On peut opérer de même pour $\frac{\sigma^2_{(u,2)}}{\sigma^2_{(u,1)}}$.

4.2 DJ entre un processus AR d'ordre 1 et un bruit blanc

En tirant avantage de l'expression de l'inverse de la matrice de corrélation d'un processus AR d'ordre 1 [12], on peut exprimer l'incrément asymptotique de la DJ comme suit :

$$\Delta DJ^{(AR1,BB2)} = -1 + \frac{1}{2}\left[\frac{\sigma^2_{(u,1)}}{\sigma^2_{(u,2)}}\frac{1}{(1-(a_1^1)^2)} + \frac{\sigma^2_{(u,2)}}{\sigma^2_{(u,1)}}(1+(a_1^1)^2)\right] \tag{12}$$

En suivant la démarche proposée, $\frac{\sigma^2_{(u,1)}}{\sigma^2_{(u,2)}}\frac{1}{(1-(a_1^1)^2)}$ correspond bien à la puissance d'un processus AR d'ordre 1 filtré par un filtre passe-tout de fonction de transfert $\frac{1}{\sigma_{(u,2)}}$ et $\frac{\sigma^2_{(u,2)}}{\sigma^2_{(u,1)}}(1+(a_1^1)^2)$ correspond à la puissance d'un bruit blanc centré de variance $\sigma^2_{(u,2)}$ filtré par un filtre de fonction de transfert $\frac{1}{\sigma_{(u,1)}}(1+a_1^1 z^{-1})$.

4.3 DJ entre deux processus AR d'ordre 1

Dans [7], nous avions proposé une approche récursive pour déduire la DJ comme suit :

$$\Delta DJ^{(AR1,AR2)} = A + B \tag{13}$$

avec

$$\begin{cases} A = -1 + \frac{1}{2}\left(R_{ar} + \frac{1}{R_{ar}}\right) \\ B = \frac{(a_1^2 - a_1^1)^2}{2}\left[\frac{1}{1-(a_1^1)^2}\frac{1}{R_{ar}} + \frac{1}{1-(a_1^2)^2}R_{ar}\right] \\ R_{ar} = \frac{\sigma^2_{(u,2)}}{\sigma^2_{(u,1)}} \end{cases} \tag{14}$$

Or, en réorganisant les termes, on obtient :

$$\Delta DJ^{(AR1,AR2)} = -1 + \frac{1}{2}\Big(R_{ar}\frac{1-2a_1^1a_1^2+(a_1^1)^2}{1-(a_1^2)^2} \tag{15}$$

$$+ \frac{1}{R_{ar}}\frac{1-2a_1^1a_1^2+(a_1^2)^2}{1-(a_1^1)^2}\Big)$$

En suivant notre démarche, nous pouvions obtenir le même résultat rapidement puisque $R_{ar}\frac{1-2a_1^1a_1^2+(a_1^1)^2}{1-(a_1^2)^2}$ correspond à la puissance du second processus AR filtré par le 1er filtre inverse. $\frac{1}{R_{ar}}\frac{1-2a_1^1a_1^2+(a_1^2)^2}{1-(a_1^1)^2}$ vaut quant à lui la puissance du premier processus AR filtré par le 2nd filtre inverse.

4.4 DJ entre deux processus MA d'ordre 1 réels

Dans [9], en nous appuyant sur l'expression de l'inverse de la matrice de corrélation d'un processus MA d'ordre 1 réel, nous avons montré que l'incrément asymptotique de la DJ est constant si $|b_1^l| \neq 1$ et vaut :

$$\Delta DJ^{(MA1,MA2)} = -1 + \frac{1}{2}\left(\frac{r_1^1}{r_1^2} + \frac{r_1^2}{r_1^1}\right) + \frac{1}{2}\left[\left(\frac{r_0^1}{r_0^2} - \frac{r_1^1}{r_1^2}\right)\frac{1+(b_1^2)^2}{|1-(b_1^2)^2|} + \left(\frac{r_0^2}{r_0^1} - \frac{r_1^2}{r_1^1}\right)\frac{1+(b_1^1)^2}{|1-(b_1^1)^2|}\right] \tag{16}$$

avec $r_0^l = (1+(b_1^l)^2)\sigma_{u,l}^2$ et $r_1^l = b_1^l\sigma_{u,l}^2$ avec $l = 1, 2$.
De plus, si $|b_1^l| = 1$, l'incrément est infini.
Au regard de l'interprétation que nous proposons de la DJ dans cette communication, nous sommes en mesure d'expliquer à présent pourquoi ces différents cas apparaissent.
En développant et réorganisant les termes, on montre que :

$$\Delta T_{MA2,MA1} = \frac{r_1^1}{r_1^2} + \left(\frac{r_0^1}{r_0^2} - \frac{r_1^1}{r_1^2}\right)\frac{1+(b_1^2)^2}{|1-(b_1^2)^2|} \tag{17}$$
$$= \begin{cases} \frac{\sigma_{u,1}^2}{\sigma_{u,2}^2}\frac{1+(b_1^1)^2-2b_1^1b_1^2}{1-(b_1^2)^2} \text{ si } |b_1^2| < 1 \\ \frac{\sigma_{u,1}^2}{\sigma_{u,2}^2}\frac{1+(b_1^1)^2-2\frac{b_1^1}{b_1^2}}{(b_1^2)^2-1} \text{ si } |b_1^2| > 1 \end{cases}$$

En étudiant la fonction d'autocorrélation du premier processus MA filtré par le filtre inverse associé au second, on montre que la puissance du premier processus filtré par le filtre inverse du second vaut bien $\Delta T_{MA2,MA1}$. Quand le zéro associé à un processus MA est sur le cercle unité dans le plan complexe, le filtre inverse résultant admet un pôle de module unité. Par conséquent, la puissance du processus filtré par ce filtre inverse est infinie. Cela explique la valeur infinie de l'incrément asymptotique qui avait été identifiée pour $|b_1^l| = 1$.

4.5 DJ entre un processus AR et un processus MA

Dans [11], nous avons analysé la DJ entre deux processus AR et MA d'ordre 1 dont les processus générateurs étaient de variance respective $\sigma_{u,1}^2$ et $\sigma_{u,2}^2$. L'incrément de la DJ vérifiait :

$$\Delta DJ_k^{(AR,MA)} = -1 + \frac{1}{2}\left[\Delta T_k^{(MA,AR)} + \Delta T_k^{(AR,MA)}\right] \tag{18}$$

avec

$$\Delta T_k^{(AR,MA)} = \frac{\sigma_{u,2}^2}{\sigma_{u,1}^2}[2a_1^1b_1^2 + (1+(a_1^1)^2)(1+(b_1^2)^2)] \tag{19}$$

et

$$\Delta T^{(MA,AR)} = \lim_{k\to+\infty}\Delta T_k^{(MA,AR)} = \tag{20}$$
$$\begin{cases} \frac{\sigma_{u,1}^2}{\sigma_{u,2}^2(1-(a_1^1)^2)}\left[\frac{1}{1-(b_1^2)^2} + 2\frac{a_1^1b_1^2}{(1-a_1^1b_1^2)(1-(b_1^2)^2)}\right] & \text{si } |b_1^2| < 1 \\ \frac{\sigma_{u,1}^2}{\sigma_{u,2}^2(1-(a_1^1)^2)}\left[\frac{1}{(b_1^2)^2-1} + 2\frac{a_1^1/b_1^2}{(1-a_1^1/b_1^2)(1-(b_1^2)^2)}\right] & \text{si } |b_1^2| > 1 \end{cases}$$

ou encore si $|b_1^2| = 1$

$$\Delta T_k^{(MA,AR)} = \frac{\sigma_{u,1}^2}{\sigma_{u,2}^2(1-(a_1^1)^2)}\left[\frac{2k+3}{6} + \frac{2(a_1^1b_1^2)^k}{k+1} + \frac{1}{3}\sum_{j=1}^{k}\frac{(a_1^1b_1^2)^j(k+1-j)(k+2-j)(2k+3+j)}{(k+1)(k+2)}\right] \tag{21}$$

À nouveau, notre nouvelle interprétation de la DJ permet d'expliquer les différents cas qui apparaissent et les raisons pour lesquelles l'incrément asymptotique n'est pas constant dans le cas d'un paramètre MA de module égal à 1.

5 Conclusions et perspectives

Sans développer des calculs fastidieux, notre interprétation de l'incrément asymptotique de la DJ permet d'expliquer les cas atypiques qui ont pu être identifiés dans des travaux précédents pour lesquels l'incrément asymptotique n'est pas une constante finie. Elle permet aussi de généraliser notre analyse sur la DJ entre processus AR, ARMA, MA *etc.* Ainsi, tous les processus ARMA d'ordre quelconque dont la DSP est caractérisée par l'absence totale d'une ou de plusieurs fréquences donnent lieu à des incréments asymptotiques non constants.

Références

[1] L. Bombrun, N.-E. Lasmar, Y. Berthoumieu et G. Verdoolaege, Multivariate texture retrieval using the SIRV representation and the geodesic distance, ICASSP, pp. 865-868, 2011.

[2] R. Murthy, I. Pavlidis et P. Tsiamyrtzis, Touchless monitoring of breathing function, IEEE EMBS, pp. 1196-119, 2004.

[3] M. Sugiyama, T. Suzuki, S. Nakajima, P. von Bunau et M. Kawanabe, Direct importance estimation for covariate shift adapation, Annals of the Institute, Vol. 60, n°4, pp. 699-746, 2008.

[4] X. Nguyen, M. J. Wainwright et M. I. Jordan, Estimating divergence functionals and the likelihood ratio by convex risk minimization, IEEE Trans. on Information Theory, Vol. 56, n°11, pp. 5847-5861, 2010.

[5] S. Noschese, L. Pasquini et L. Reichel, Tridiagonal Toeplitz matrices : properties and novel applications, Numerical Linear Algebra with Applications, Vol. 20, n°2, pp. 302-326, 2013.

[6] R. J. Stroeker, Approximations of the eigenvalues of the covariance matrix of a first-order autoregressive process, Journal of Econometrics, pp. 269-281, 1983.

[7] C. Magnant, A. Giremus et E. Grivel, On computing Jeffrey's divergence between time-varying autoregressive models, IEEE Signal Processing Letters, Vol. 22, issue 7, pp. 915-919, 2014.

[8] C. Magnant, E. Grivel, A. Giremus, B. Joseph et L. Ratton, Jeffrey's divergence for state space model comparison, Elsevier Signal Processing, Vol. 114, pp.61-74, 2015.

[9] L. Legrand et E. Grivel, Jeffrey's divergence between moving-average models that are real or complex, noise-free or disturbed by additive white noises, Elsevier Signal Processing, Vol. 131, pp. 350-363, 2017.

[10] W.-C. Yueh, Explicit Inverses Of Several Tridiagonal Matrices, Applied Mathematics E-Notes, pp. 74-83, 2006.

[11] L. Legrand et E. Grivel, Jeffrey's divergence between moving-average and autoregressive models, ICASSP, 2017.

[12] B. Cernuschi-Frias, A derivation of the Gohberg-Semencul relation [signal analysis], IEEE Trans. on Signal Processing, Vol. 39, n°1, pp. 190-192, 1991.

[13] C. Rao, Information and the accuracy attainable in the estimation of statistical parameters, Bull. Calcutta Math. Soc., Vol. 37, pp. 81-89, 1945.

[14] S. Kullback et R. A. Leibler, On Information and sufficiency, The annals of mathematical statistics, Vol. 22, n°1, pp. 79-86, 1951.

[15] C. E. Rasmussen et C. K. I. Williams, Gaussian processes for machine learning, MIT Press, 2006.