

12-548F

C.2

Les modèles pour les séries temporelles



AJSN 1434427
C.2.

STATISTIQUE CANADA
Division des analyses de conjoncture
Série chronologiques générales

LES MODÈLES POUR LES SÉRIES TEMPORELLES

Publication autorisée par
le ministre de l'Industrie et du Commerce



Mai 1974
3-1501-518

Prix: \$1.40

Information Canada
Ottawa

Avant-Propos

Cette monographie sur Modèles pour les Séries Temporelles, a été rédigée par Dr. Estela Bee Dagum, Chef Service de Recherche et Développement de la section des Séries Chronologiques Générales.

L'objet de cette étude est l'identification et l'analyse des hypothèses de base des différents types de modèles construits pour les séries temporelles.

Une attention particulière est apportée aux modèles à une variable pour lesquels existent des méthodes d'estimation bien établies. Par l'examen de leurs hypothèses, on espère rendre les utilisateurs et les chercheurs conscients des limitations de ces modèles.

Une telle étude, repose sur le fait que l'optimalité de toute méthode d'estimation dépend strictement de la vérification de ses hypothèses de base.

En d'autres termes, à chaque méthode d'estimation correspond un modèle pour lequel cette méthode est optimale (le terme "optimale" étant pris dans son acception statistique courante). Mais puisqu'il n'existe aucun modèle unique applicable à toutes les séries temporelles dans tous les cas, il ne faut utiliser aucune méthode d'estimation comme étant la seule fournissant une solution optimale unique sans un examen critique.

L'auteur est la seule responsable des analyses et conclusions dans cette monographie et remercie le Professeur C. Dagum de l'Université d'Ottawa, M. Phillip Smith, de la Division des Analyses de Conjoncture et Mlle. Mary Lennox, Chef de la section des Séries Chronologiques générales pour leurs commentaires très valables.

Le directeur de la
Division des analyses de conjoncture,
P.N. TRIANDAFILLOU

TABLE DES MATIÈRES

	Page
1. Introduction et sommaire	1
2. La série temporelle en tant que réalisation d'échantillonnage d'un processus stochastique	17
3. Modèles de processus stochastiques: processus stationnaires et non stationnaires	23
4. Les modèles d'erreurs	30
5. Les processus stochastiques stationnaires linéaires ...	39
6. Les modèles stochastiques stationnaires non paramétriques: la fonction d'autocovariance et le spectre	46
7. Les fonctions d'autocovariance et d'autocorrélation des processus stationnaires linéaires et leurs transformées de Fourier	59
8. Ergodicité	72
9. Processus stochastiques non stationnaires linéaires homogènes	76
10. Modèles saisonniers	82
11. Conclusions	89
Références	94

1. INTRODUCTION ET SOMMAIRE

Au cours des dix dernières années, de nombreux statisticiens se sont penchés sur la recherche de méthodes optimales d'estimation des séries temporelles. Il est facile de voir pourquoi. Des données statistiques exactes sont essentielles à la prise des décisions. Les responsables des politiques, confrontés aux responsabilités que pose le contrôle de l'activité économique, ne prendront guère de décisions basées sur de mauvaises estimations ou d'estimations sujettes à des révisions appréciables chaque fois que de nouvelles informations paraissent.

On a déjà élaboré plusieurs méthodes d'estimation ainsi que les programmes informatiques correspondants. Dans leur majorité, elles se fondent sur des méthodes classiques d'analyse statistique, telles que les moindres carrés ou les moyennes mobiles, bien que l'on ait envisagé d'autres types de filtres linéaires. De ce fait, les travaux destinés à obtenir des estimations précises et fiables de données historiques et courantes étaient consacrés surtout à des méthodes d'estimation. Mais lorsque l'on utilise ces méthodes d'estimation, on oublie un principe simple, mais essentiel: l'optimalité de toute méthode d'estimation dépend strictement de la vérification de ses hypothèses de base.

En d'autres termes, à chaque méthode d'estimation correspond un modèle pour lequel cette méthode est optimale. Mais puisqu'il n'existe aucun modèle unique applicable à toutes les séries temporelles dans tous les cas, il ne faut utiliser aucune méthode d'estimation comme étant la seule fournissant une solution optimale unique sans un examen critique.

Plus le comportement du processus générateur d'une série temporelle s'écarte des hypothèses de base du modèle, moins sa capacité de décrire le phénomène sera grande, et par conséquent, quelle que soit l'optimalité de la méthode d'estimation correspondante, les résultats finals seront fortement déformés.

Il ne faut pas blâmer uniquement les utilisateurs pour ce genre de négligence; les chercheurs sont aussi blâmables, car ils se sont engagés dans des discussions oiseuses sur la supériorité d'une méthode sur une autre, supériorité souvent illustrée par une série temporelle pour laquelle la méthode en question s'est avérée optimale. Cette critique s'applique aux statisticiens théoriciens, car la plupart des nouvelles méthodes d'estimation sont basées sur des modèles très simples, dont les hypothèses sont parfois tellement générales qu'il est très difficile d'isoler le phénomène réel qui y correspond vraiment. Mais il existe une vaste catégorie de modèles qui peuvent être traités mathématiquement et qui sont mieux adaptés au comportement de nombreux phénomènes variant dans le temps, mais pour lesquels l'on n'a pas encore élaboré de méthodes d'estimation. Nous voulons parler plus particulièrement des modèles qui reposent sur les hypothèses de non-stationnarité et de non-normalité. Nous n'examinerons pas ici ces types de modèles, et nous allons surtout nous concentrer sur les modèles à une variable pour lesquels existent des méthodes d'estimation bien établies. Par l'examen de leurs hypothèses, nous espérons rendre les utilisateurs et les chercheurs conscients de leurs limitations. Notre angle d'approche ne vise que la base empirique théorique des modèles et non pas les méthodes d'estimation ou autres méthodes d'inférence statistique.

Nous espérons ainsi éclairer un peu les hypothèses les plus réalistes et (ou) les plus facilement manipulables mathématiquement relativement au comportement d'un phénomène évoluant dans le temps de manière probabiliste.

Par modèle, nous n'entendons pas un système théorique dépourvu de substance empirique. Au contraire, nous le définissons ici (voir Dagum et Dagum, 1972) comme "un système empirico-théorique, satisfaisant aux propriétés suivantes:

- 1) il doit représenter un univers non contradictoire, mais possible,
- 2) il doit représenter un univers empirique possible, et
- 3) il doit subir et passer des tests d'hypothèses".

Reprenant l'idée de Wold (1938), déjà incluse dans la littérature actuelle, nous considérons ici une série temporelle comme la réalisation d'échantillonnage d'un processus stochastique qui, d'un point de vue non mathématique, est tout processus qui suit des lois de probabilité. Les observations faites à mesure que le processus se développe en tracent l'évolution. À chaque point t dans le temps appartenant à une période T , c'est-à-dire $t \in T$, on considère l'observation faite $x(t)$ comme la réalisation ou la valeur observée d'une variable aléatoire $X(t)$, et une famille de ces variables aléatoires $\{X(t), t \in T\}$ définit un processus stochastique, ou fonction aléatoire.

En posant par hypothèse qu'une série temporelle observée $\{x(t), t \in T\}$ est la réalisation d'échantillonnage d'un processus stochastique $\{X(t), t \in T\}$ on essaie de déduire de la série observée le mécanisme générateur, ou structure probabiliste, du processus.

Pour analyser une série temporelle, il faut donc d'abord supposer un modèle qui doit être entièrement spécifié, à l'exception des valeurs des paramètres que l'on se prépare à estimer sur la base des échantillons observés.

Les modèles pour séries temporelles sont des processus stochastiques, qui peuvent être classés de plusieurs façons.

Nous allons examiner une classification qui nous permettra d'étudier les hypothèses de base de chaque type de processus et de distinguer les processus destinés aux applications pratiques des processus plus théoriques.

Tableau 1

Classification d'un processus stochastique

- | | |
|---|--|
| 1) Selon l'indépendance ou non des propriétés du processus stochastique par rapport à l'origine du temps | { a) Stationnaire
b) Non stationnaire |
| 2) Selon les fonctions de distribution caractérisant le processus | { a) Normal
b) Non normal |
| 3) Selon l'indépendance ou non du comportement du processus par rapport à ses valeurs dans l'intervalle de temps précédent. | { 1) Markovien
2) Non markovien |

A priori, un processus stationnaire est un processus dont la distribution demeure la même dans le temps parce que le mécanisme aléatoire produisant le processus ne change pas avec le temps. En d'autres termes,

toutes les distributions de probabilité dépendent uniquement des positions relatives des instants t_1, t_2, \dots, t_n , mais non pas des valeurs réelles de ces quantités, c'est-à-dire que si l'on déplace l'ensemble des points t_1, t_2, \dots, t_n le long de l'axe t , les distributions restent les mêmes.

Un processus stochastique est entièrement défini par ses distributions de probabilité, mais cette méthode en général ne convient pas, car elle est peu pratique.

Dans la pratique, au lieu de considérer les fonctions de probabilité, on ne spécifie que les deux premiers moments. On dit alors qu'un processus est stationnaire lato sensu, ou stationnaire du second ordre, si la moyenne et la variance sont des constantes et si la fonction de covariance ne dépend que de la différence entre deux points quelconques dans le temps.

Tout processus pour lequel les conditions de stationnarité ne se vérifient pas lato sensu est appelé processus non stationnaire. Il existe une importante sous-catégorie de processus non stationnaires qui a été développée de façon intensive pour des applications pratiques, celle des processus non stationnaires homogènes, également appelés processus aux accroissements stationnaires. Ces processus sont non stationnaires, mais en prenant des différences d'un ordre fini, soit d , les processus deviennent stationnaires dans la d -ième différence. Dans un certain sens, il s'agit d'une généralisation de la théorie des processus stochastiques stationnaires.

La deuxième caractéristique qui peut également servir de base à une classification des fonctions aléatoires est la forme des fonctions de distribution qui définissent le processus. La loi de distribution la plus répandue est la loi normale. Les processus normaux peuvent aussi utiliser des méthodes de calcul qui ne peuvent être appliquées ailleurs. Il est donc avantageux de partager les processus stochastiques en deux groupes: les processus normaux, et les processus non normaux. Il est important de constater que si les fonctions de distribution suivent une loi normale, seuls les moments des deux premiers ordres suffisent à décrire complètement un processus aléatoire normal.

Enfin, lors de la classification des processus stochastiques, on peut prendre en considération la dépendance du comportement d'un processus par rapport à sa valeur dans l'intervalle de temps précédent. On dit qu'un processus stochastique $\{X(t), t \in T\}$ est markovien si pour tout ensemble de n points $t_1 < t_2 < \dots < t_n$, la distribution conditionnelle de $X(t_n)$ pour des valeurs données de $X(t_1), X(t_2), \dots, X(t_{n-1})$, ne dépend que de $X(t_{n-1})$, c'est-à-dire de la valeur connue la plus récente. En d'autres termes, la probabilité de tout comportement futur particulier du processus si sa forme actuelle est connue n'est pas modifiée par notre connaissance empirique de son comportement dans le passé. Inversement, si les caractéristiques probabilistes du processus à un moment donné dépendent de ses valeurs dans les intervalles de temps précédents, on dira que le processus est non markovien.

Ces trois catégories principales sont compatibles, et ce, dans n'importe quelle combinaison. Ainsi, un processus peut être normal et stationnaire de type markovien, ou stationnaire et non normal, ou normal et non stationnaire, et ainsi de suite.

Mais, au niveau des applications pratiques, toutes les combinaisons ne sont pas possibles. Les principales limitations sont l'absence d'une théorie bien développée et l'inexistence de méthodes d'estimation appropriées pour en permettre l'application. Nous allons donc examiner les hypothèses des processus qui peuvent actuellement avoir des applications pratiques. Ces processus sont: 1) les processus non stationnaires (dans la moyenne) normaux, 2) les processus stationnaires linéaires normaux, et 3) les processus non stationnaires linéaires homogènes normaux. On dit qu'un processus stochastique $X(t)$ est linéaire s'il existe un processus purement aléatoire U_t , c'est-à-dire un processus de variables aléatoires indépendantes de distribution identique et une suite de paramètres $\alpha_0, \alpha_1, \alpha_2, \dots$ tels que $X(t) = \sum_{k=0}^{\infty} \alpha_k U_{t-k}$.

Les processus des types 1), 2) et 3) se sont avérés les plus faciles à traiter d'un point de vue mathématique. Ils semblent aussi décrire avec assez de précision les mécanismes générateurs de nombreux problèmes physiques.

Les propriétés qui font que ces types de processus sont très utiles sont les suivantes: a) en vertu de l'hypothèse de normalité, ces processus sont entièrement définis par les deux premiers moments des fonctions de distribution, b) en vertu de l'hypothèse de stationnarité, ou de stationnarité dans les différences, leurs moyennes et leurs variances sont constantes et leurs fonctions d'autocovariance ne dépendent que des décalages temporels. Pour les processus de type 2) et 3) l'hypothèse de linéarité garantit l'ergodicité.

L'importance de la stationnarité d'un processus réside dans le fait que le théorème ergodique et le spectre ont été définis d'abord pour des processus stochastiques stationnaires. Tous les processus stationnaires linéaires sont ergodiques, c'est-à-dire que l'on peut obtenir des estimateurs cohérents de la moyenne et de la fonction de covariance d'un processus donné avec une seule réalisation. En d'autres termes, les moyennes calculées à partir d'un échantillon (un ensemble de données) appelées moyennes temporelles, peuvent être finalement identifiées aux moyennes correspondantes de l'ensemble, c'est-à-dire aux moyennes sur tout l'intervalle des réalisations du processus à un instant donné. La propriété d'ergodicité est très importante lorsque l'on a affaire aux séries temporelles économiques, dont les données proviennent d'une réalisation unique.

Le spectre est la transformée de Fourier de la fonction d'autocovariance, et d'un point de vue mathématique, ils sont équivalents. Les deux opérations fournissent le même type de données probabilistes pour la détermination du mécanisme générateur d'une série temporelle observée, mais la fonction d'autocovariance donne les informations dans le domaine du temps tandis que le spectre le fait dans le domaine de la fréquence.

La représentation spectrale d'un processus stochastique se fait à l'aide des séries de Fourier, ou de l'intégrale de Fourier, selon qu'il s'agisse d'un paramètre temporel discret ou continu. Les fonctions périodiques dans l'analyse de Fourier sont le sinus et le cosinus. Ils possèdent les deux importantes propriétés suivantes: une approximation d'un nombre donné de termes donne l'erreur quadratique moyenne minimale entre la fonction et son approximation; de plus, les fonctions sont orthogonales, de sorte qu'il est possible de déterminer les coefficients indépendents.

La représentation spectrale d'un processus stationnaire se ramène donc à une "décomposition" du processus en couples d'oscillations périodiques non corrélées, et la variance totale du processus est distribuée sur la fréquence. S'il s'agit d'un paramètre temporel discret, il est possible de déterminer la partie de la variance allant à chaque composante avec une fréquence particulière, et dans le cas de processus à paramètre temporel continu, on parlera de l'apport d'une bande de fréquence T.

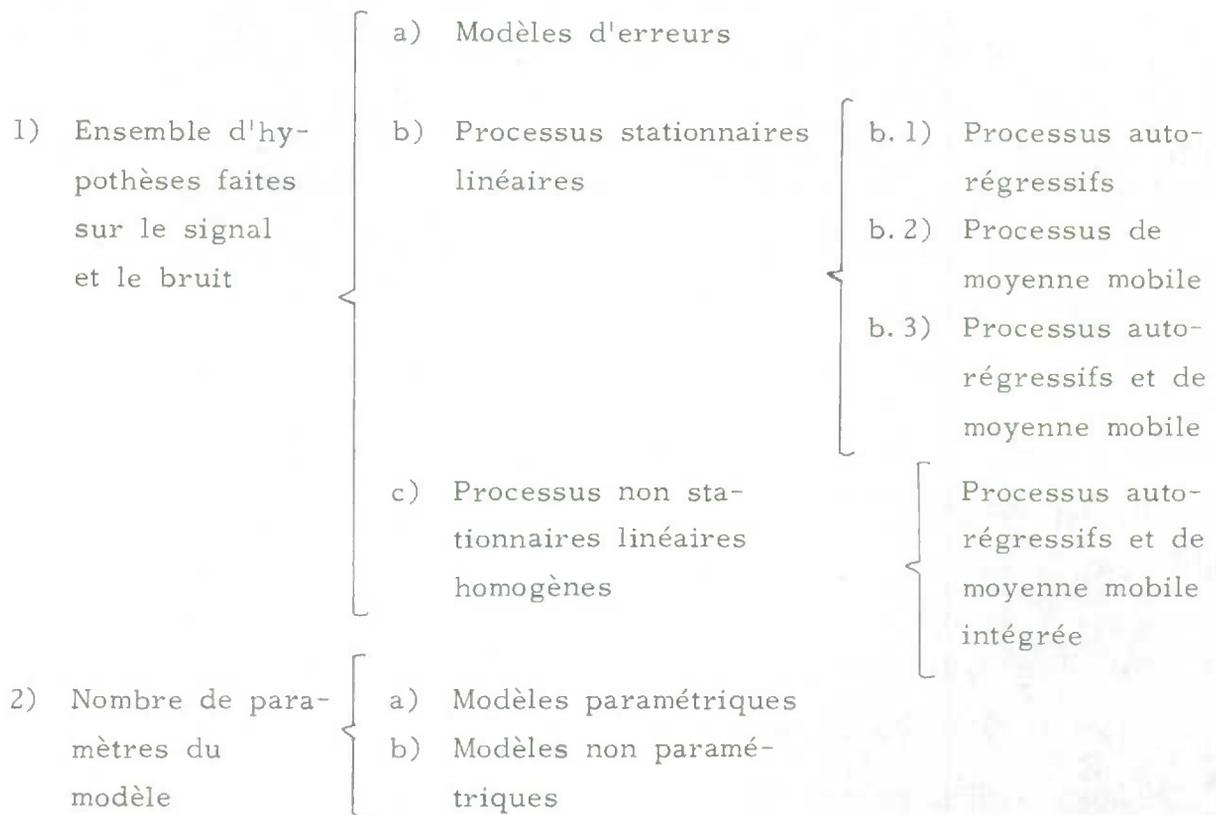
On a surtout utilisé l'analyse spectrale dans le cas de problèmes qui nécessitent des études de réaction de la fréquence, telles que la conception des structures d'avion, ou dans des expériences destinées à optimiser le rendement de procédés industriels. Mais lorsqu'il s'agit des séries temporelles, et en particulier de séries temporelles économiques, la fonction d'autocovariance convient mieux à la construction de modèles. Cependant, nous devons souligner, avec Mandelbrot (1972), que l'analyse de la fonction d'autocovariance n'est efficace avant tout que pour les modèles qui correspondent à des processus stochastiques suivant une loi normale ou quasi-normale avec des fréquences élevées ou de petits retards, c'est-à-dire pour les modèles avec des effets de courte période et presque normaux.

Dans notre étude, nous allons insister surtout sur l'analyse de la fonction d'autocovariance (autocorrélation), mais nous allons aussi donner le spectre correspondant (spectre normalisé) des modèles considérés. Cette dernière opération est effectuée davantage pour donner une image complète que pour son utilité proprement dite.

Dans les grandes catégories de processus non stationnaires (dans la moyenne) normaux, des processus stationnaires linéaires normaux et des processus non stationnaires linéaires homogènes normaux, on a construit plusieurs types de modèles pour la description et la prédiction des séries temporelles. Bien que l'hypothèse de normalité ne soit pas toujours posée de façon explicite dans la définition des modèles, elle est néanmoins introduite lorsqu'il s'agit de tester leurs hypothèses. En dépit de sa simplicité et d'un certain chevauchement de ses éléments, la classification ci-dessous nous sera néanmoins utile:

Tableau II

Classification des modèles destinés aux séries temporelles



Dans un modèle d'erreurs, le mécanisme générateur d'une série temporelle est par hypothèse composé d'un élément systématique, ou signal, qui est une fonction déterminable du temps (en général, un polynôme de degré peu élevé) et une composante aléatoire que l'on suppose purement aléatoire, ou bruit blanc, c'est-à-dire distribuée identiquement avec moyenne constante, variance constante et autocorrélation nulle. Cet ensemble d'hypothèses relatives à la composante aléatoire du modèle oblige à inclure tout effet temporel dans la partie déterministe du modèle, ce qui fait que chaque observation est stochastiquement indépendante des précédentes. Les propriétés de la chronique sont ensuite résumées dans l'espérance mathématique du processus, et les variations de la composante aléatoire n'affectent pas ces propriétés. Dans un sens, l'ordre dans lequel sont disposées les observations de la chronique n'a pas d'importance, et l'analyse d'une telle chronique se ramène à une analyse statistique multidimensionnelle. Le processus est non-stationnaire (dans la moyenne) mais si le signal est un polynôme du temps, c'est un processus non stationnaire homogène. Dans ce cas, en prenant un nombre fini de différences, soit d , le processus se réduit à un processus stationnaire linéaire dans la d -ième différence. Ces modèles d'erreurs ont été élaborés d'abord en recherche astronomique. On les a construits en vue de déterminer la position d'une planète à un moment donné. Dans ce cas, les erreurs dans les séries observées sont attribuées aux erreurs d'observation imputables aux conditions atmosphériques ou aux défauts du télescope. Les erreurs ne vont pas affecter les positions ultérieures des planètes, ni les observations de ces dernières.

Ce type de modèle est donc acceptable quand les erreurs peuvent être attribuées à des erreurs de mesure. Mais dans le cas de nombreuses séries temporelles, en particulier les séries économiques, les erreurs sont non seulement imputables à des observations erronées, mais aussi à des causes plus sérieuses, et dès leur apparition, elles s'incorporent au processus et influent sur son évolution future. En d'autres termes, les erreurs sont autocorrélées; l'ordre des observations est essentiel, et la valeur courante de X_t va dépendre du temps écoulé depuis le début du processus.

Pour ces cas, on a ainsi construits plusieurs modèles entrant dans la catégorie des processus stationnaires linéaires normaux et des processus non stationnaires linéaires homogènes normaux. Parmi ceux-ci, le modèle A.R.M.M. autorégressif et de moyenne mobile (p,q) est très utile pour la description et la prédiction lorsque par hypothèse la série suit un processus stationnaire selon une loi normale ou quasi normale avec peu de retards. On obtient un tel processus en égalant un processus autorégressif d'ordre p à un processus de moyenne mobile d'ordre q. On dit qu'un processus X_t est un processus autorégressif d'ordre p si on peut le mettre sous la forme d'une somme pondérée ou d'une combinaison linéaire des p valeurs précédentes du processus, plus une composante purement aléatoire U_t .

Par contre, on dit qu'un processus X_t est un processus de moyenne mobile d'ordre q si on peut le mettre sous la forme d'une somme pondérée ou d'une combinaison linéaire de variables purement aléatoires $U_t, U_{t-1}, \dots, U_{t-q}$.

On considère donc le processus A.R.M.M. (p, q) comme étant l'intrant X_t obtenue à partir d'un extrant purement aléatoire U_t qui a subi une transformation linéaire par l'emploi de coefficients de pondération provenant du quotient de deux polynômes.

Quand le mécanisme générateur ne diffère de l'hypothèse de stationnarité que par les différences observées dans le niveau et (ou) la pente des différents éléments de la série, le processus est donc homogène non stationnaire. Si les erreurs sont autocorrélées, les modèles utilisés dans ces cas sont appelés modèles autorégressifs et de moyenne mobile intégrée (p, d, q) ou A.R.M.M.I., où p est l'ordre du processus autorégressif, d est l'ordre des différences et q , l'ordre du processus de moyenne mobile. Ces modèles contiennent de façon implicite les hypothèses de normalité ou de quasi-normalité et de dépendance à court terme. Les modèles A.R.M.M. (p, q) et A.R.M.M.I. (p, d, q) ont été utilisés avec un certain succès dans la prédiction lorsque les p , d et q ne sont pas supérieurs à deux.

Selon l'hypothèse traditionnelle de l'analyse des séries temporelles économiques, la série observée provient de la superposition de quatre processus représentant différents types d'évolution. Ce sont: a) la tendance, b) le cycle, c) le mouvement saisonnier, et d), les fluctuations accidentelles.

La tendance correspond à une variation dans un sens déterminé qui se maintient pendant une longue période, c'est-à-dire pendant une période qui est longue par rapport au cycle. Le cycle est un mouvement d'allure quasi périodique caractérisé par l'alternance des périodes d'ex-

pansion et de contraction. Le mouvement saisonnier correspond à des variations hebdomadaires, mensuelles ou trimestrielles (fluctuations intra-annuelles répétitives), causées par le climat et l'encadrement institutionnel.

Les fluctuations accidentelles sont des mouvements imprévisibles liés à toutes sortes d'événements, présentant en général une allure aléatoire stable. Pendant la plus grande partie du travail d'analyse, on combine la tendance et le cycle, puisque l'on s'intéresse surtout à l'allure de la série dégagée de la variation saisonnière.

La présence d'un élément tendance-cycle introduit la non-stationnarité dans le processus stochastique. On considère généralement que cet élément est une fonction déterminable du temps (en général, un polynôme de degré peu élevé) dans les modèles d'erreurs, et une fonction sujette à des variations stochastiques dans les modèles A.R.M.M.I.

Le mouvement saisonnier joue un rôle très important dans l'analyse des séries temporelles économiques. Nous allons consacrer la section 10 à la description des modèles qui servent habituellement à décrire la saisonnalité. Nous devons néanmoins faire remarquer qu'il est très rare de trouver le mouvement saisonnier non superposé à une tendance plus ou moins régulière. Par ailleurs, comme toutes les autres composantes d'une série temporelle économique, il ne s'agit pas d'un processus observable. Si l'on suppose que la variation saisonnière est stable et déterminable, elle sera représentée par une fonction strictement périodique du temps, qui se présentera habituellement sous la forme d'un polynôme de sinus et de cosinus du temps, avec des amplitudes et des phases constantes. Mais si l'on suppose que le mouvement saisonnier est stable et stochastique, les amplitudes prendront la forme d'un pro-

cessus purement aléatoire. Lorsque la composante saisonnière évolue dans le temps de façon déterminable, les amplitudes sont une fonction du temps; lorsque l'on suppose qu'elle évolue de façon stochastique, les amplitudes suivent alors un processus stochastique stationnaire non purement aléatoire.

Dans les modèles d'erreurs, on considère que la variation saisonnière est déterminable, tandis que dans les modèles A.R.M.M.I., elle est stochastique.

On suppose que les fluctuations accidentelles suivent un processus purement aléatoire, ou parfois, dans un sens moins restreint, un processus non autocorrélé.

Les modèles d'erreurs, les modèles A.R.M.M. (p, q) et A.R.M.M.I. (p, d, q) ainsi que les modèles saisonniers sont des modèles paramétriques, c'est-à-dire qu'ils font intervenir un nombre fini de paramètres. On peut aussi décrire la structure probabiliste d'une série temporelle par un modèle non paramétrique, ainsi appelé parce qu'il fait appel à un nombre infini de paramètres. Cette classification des modèles paramétriques et non paramétriques pour les séries temporelles ne doit pas prêter à confusion lorsqu'il s'agit des termes "paramétriques" et "non paramétriques" dans l'analyse statistique. Dans ce dernier cas, on dit qu'une méthode est paramétrique si l'on suppose par hypothèse que la variable aléatoire suit une distribution donnée; la méthode est dite non paramétrique dans le cas contraire.

Les modèles non paramétriques que nous allons examiner sont la fonction d'autocovariance et sa transformée de Fourier, c'est-à-dire le spectre. L'étude de l'une ou l'autre de ces deux fonctions est très utile lorsque l'on entreprend un travail de recherche, quand l'on ne connaît que peu de choses sur un phénomène donné. Le but principal de l'analyse des séries temporelles à ce stade est de considérer les données sous divers angles afin d'en dégager les hypothèses susceptibles d'être avancées, et un modèle exploratoire permettrait alors la construction d'un modèle paramétrique approprié.

Finalement, le lecteur ne doit jamais oublier que le choix d'un modèle vraiment approprié à une série temporelle est toujours fonction de la connaissance a priori que nous pouvons avoir de la nature du phénomène considéré et du but de l'analyse. Au niveau des applications empiriques immédiates, ce choix est aussi fonction de l'existence de méthodes numériques de solution pouvant être rapidement mécanisées.

2. LA SÉRIE TEMPORELLE EN TANT QUE RÉALISATION D'ÉCHANTILLONNAGE D'UN PROCESSUS STOCHASTIQUE

D'un point de vue non mathématique, un processus stochastique sera tout processus soumis à des lois probabilistes. Les observations faites à mesure que le processus progresse indiquent la façon dont il évolue.

Dans la majorité des cas, nous nous intéressons aux variations temporelles, et l'état du processus à tout instant donné t est décrit par les valeurs d'un certain nombre de quantités observables qui sont les réalisations particulières des variables aléatoires.

Supposons que le processus soit défini par une seule variable $X(t)$; dans ce cas, pour chaque t , la valeur numérique de $X(t)$ ne sera pas uniquement déterminée comme dans le cas d'un système déterminable, mais dépendra des influences aléatoires qui agissaient sur le processus jusqu'à l'instant t . Alors, pour tout t fixe, $X(t)$ est une variable aléatoire définie dans un espace probabiliste $(\Omega, \mathcal{F}, \mathbb{P})$. Lorsque t varie au cours de l'intervalle de temps considéré, nous obtenons une famille de variables aléatoires $X(t)$ dépendant du paramètre t , et définies dans le même espace probabiliste. Cette famille $\{X(t), t \in T\}$ de variables aléatoires $X(t)$ est une fonction aléatoire, ou processus stochastique. Si T est un ensemble fini, nous avons un nombre fini de variables aléatoires dans le processus, et elles peuvent être représentées par des fonctions de distribution multidimensionnelles. En effet, si $X = \{X_1, X_2, \dots, X_n\}$, nous avons un vecteur aléatoire à n dimensions, qui est défini par sa fonction de distribution multidimensionnelle.

$$F(x_1, x_2, \dots, x_n) = P\{X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n\}.$$

Si T est l'ensemble des nombres entiers, le processus est infini, et on l'appelle alors processus stochastique à paramètre temporel discret (discret se rapporte ici à la nature de l'ensemble T), ou suite stochastique. Si T est l'ensemble des nombres réels, le processus est également infini, mais, non dénombrable, et il est appelé processus stochastique continu ou simplement processus stochastique. Le terme "processus stochastique" est généralement utilisé pour désigner des processus ayant un nombre infini (dénombrable ou non) de variables aléatoires. La spécification complète d'un processus stochastique nécessite plus qu'une simple extension du cas fini.

Pour tout t fixe, disons $t=t_1$, nous avons une variable aléatoire $X(t_1)$ qui est complètement définie par sa fonction de distribution (d,f) désignée

$$(2.1) F_{t_1}(x_1) = P \{X(t_1) \leq x_1\}.$$

Pour tout élément t appartenant à l'ensemble T , nous avons

$$(2.2) F_t(x) = P \{X(t) \leq x\}.$$

Pour tout couple de points t_1, t_2 appartenant à l'ensemble T , nous avons la fonction de distribution suivante:

$$(2.3) F_{t_1, t_2}(x_1, x_2) = P \{X(t_1) \leq x_1, X(t_2) \leq x_2\}.$$

de la variable aléatoire à deux dimensions $X = (X(t_1), X(t_2))$.

De façon générale, pour tout ensemble fini arbitraire de t valeurs, nous avons:

$$(2.4) F_{t_1, t_2, \dots, t_n}(x_1, x_2, \dots, x_n) = P \{X(t_1) \leq x_1, X(t_2) \leq x_2, \dots, X(t_n) \leq x_n\}$$

correspondant à une variable aléatoire à n dimensions $X=(X(t_1), X(t_2), \dots, X(t_n))$.

La famille de toutes ces distributions jointes pour $n=1, 2, \dots$ et toutes les valeurs possibles t_j , constitue la famille de distributions dimensionnelles finies associées au processus stochastique $X = \{X(t), t \in T\}$. Pour que (2.4) puisse définir un processus stochastique, elle doit satisfaire aux deux conditions suivantes:

a) La condition de symétrie, en vertu de laquelle

$$(2.5) F_{t_{j_1}, t_{j_2}, \dots, t_{j_n}}(x_{j_1}, x_{j_2}, \dots, x_{j_n}) = F_{t_1, t_2, \dots, t_n}(x_1, x_2, \dots, x_n)$$

où le j_1, j_2, \dots, j_n est toute permutation des indices 1, 2, ..., n.

On veut dire par là que la condition de symétrie exige que les distributions F à n dimensions données en (2.4) soient symétriques dans tous les couples (x_j, t_j) de sorte que F reste invariable lorsque les t_j et les x_j font l'objet de la même permutation.

b) La condition de compatibilité, en vertu de laquelle

$$(2.6) F_{t_1, t_2, \dots, t_m, t_{m+1}, \dots, t_n}(x_1, x_2, \dots, x_m, \infty, \dots, \infty) = F_{t_1, t_2, \dots, t_m}(x_1, x_2, \dots, x_m)$$

pour tout t_{m+1}, \dots, t_n if $m < n$.

En d'autres termes, la condition de compatibilité exige que

$$\lim_{x_m \rightarrow +\infty} F_{t_1, t_2, \dots, t_m}(x_1, x_2, \dots, x_m) = F_{t_1, t_2, \dots, t_{m-1}}(x_1, x_2, \dots, x_{m-1}).$$

Kolgomorov (1933) a démontré par un célèbre théorème, que (2.4), avec (2.5) et (2.6), définit de façon unique la distribution probabiliste de l'espace échantillon du processus stochastique.

L'inverse est également vrai. On peut considérer toute famille de fonctions de distribution finies de type (2.4) satisfaisant aux conditions (2.5) et (2.6) comme définissant un certain processus stochastique.

Nous avons défini un processus stochastique, ou fonction aléatoire sur T , comme une famille de variables aléatoires $\{X(t), t \in T\}$. Il est important de souligner que, puisque les $X(t)$ sont des variables aléatoires, elles sont des fonctions de nombres réels de l'état ω de l'espace échantillon Ω . C'est pourquoi on écrit parfois un processus stochastique de façon explicite comme un ensemble de fonctions ayant deux arguments, à savoir ω (l'état ou le point) qui est un élément de Ω , et t (le temps), qui est un élément de T . Sous forme symbolique, on a

$$(2.7) X = \{X(\omega, t), \omega \in \Omega, t \in T\}.$$

On ne l'a pas fait ci-dessus parce que dans la théorie des probabilités, la dépendance de ω d'une variable aléatoire X est habituellement supprimée.

Si l'on considère l'expression (2.7), on peut interpréter un processus stochastique de deux façons différentes, selon la variable argument que l'on considère. Pour un point ω , donné, l'expression (2.7) se ramène à une famille de fonctions du temps, indexées sur ω . Ainsi, à chaque état ω d'une épreuve donnée, correspond une fonction réelle bien définie de la variable t . On appelle cette fonction réalisation, ou fonction d'échantillonnage du processus stochastique. Cela présente pour nous un grand intérêt, puisque la réalisation d'un processus stochastique est précisément une série temporelle observée.

Par contre, si t est donné, l'expression (2.7) se ramène à un ensemble de variables aléatoires indexées sur t . Dans ce cas, afin de définir le processus stochastique, nous devons donner la probabilité de concrétisation des diverses réalisations, ce qui conduit à la définition d'une mesure de probabilité P sur l'espace des fonctions des réalisations.

Si t et ω sont tous les deux donnés, l'expression (2.7) se ramène alors à un nombre.

Le processus qui génère l'observation d'une série temporelle est donc considéré comme un processus aléatoire dont l'un des infiniment nombreux états ω , appartenant à Ω , auraient pu être l'état observé à un instant donné t , par exemple t_1 .

Étant donné qu'on fait de même pour tous les éléments t de T , une réalisation (série temporelle observée) du processus est une fonction d'un ensemble doublement infini de fonctions qui auraient pu être générées par le processus stochastique.

Ce qui distingue l'analyse de séries temporelles des autres types d'analyse statistique est que l'on tient explicitement compte de l'ordre dans lequel sont faites les observations. Dans plusieurs domaines d'étude, des observations successives d'une série temporelle sont dépendantes, en particulier dans le cas des séries sociales et économiques. Mais il se peut que les observations soient indépendantes statistiquement, en particulier dans certains problèmes de physique et d'astronomie. L'hypothèse de dépendance ou d'indépendance entre plusieurs observations successives d'une série temporelle détermine le type de modèle à utiliser pour décrire le processus générateur de la chronique.

Lorsque l'on définit un processus stochastique selon la famille des fonctions de distribution finies (2.4) qui satisfont à la condition de symétrie (2.5) et à la condition de compatibilité (2.6), on dit que le modèle qui génère une série temporelle est un modèle non paramétrique parce qu'il fait intervenir un nombre infini de paramètres.

Il existe cependant d'autres moyens de spécifier un processus stochastique. Il est souvent utile de définir une fonction aléatoire grâce à un modèle analytique contenant un nombre fini de paramètres qui sont des variables aléatoires. On dit alors que le modèle est paramétrique. Dans la présente étude, nous analysons les modèles paramétriques et non paramétriques qui se rencontrent le plus souvent dans les séries temporelles. Un des modèles paramétriques qui s'est révélé utile pour la description et la prédiction dans des situations empiriques est ce que l'on appelle le processus autorégressif de moyenne mobile intégrée. Par contre, les fonctions d'autocovariance et d'autocorrélation ainsi que leurs transformées de Fourier, c'est-à-dire la densité spectrale (spectre) et la densité spectrale normalisée (spectre normalisé), sont des moyens non paramétriques pour définir une fonction aléatoire.

3. MODÈLES DE PROCESSUS STOCHASTIQUES: PROCESSUS STATIONNAIRES ET NON STATIONNAIRES

Les plus importantes hypothèses relatives à une série temporelle sont d'une part, que le processus stochastique correspondant est stationnaire, et, d'autre part, qu'on peut décrire un processus stochastique par les deux premiers moments de ses fonctions de distribution. Ces moments comprennent la moyenne, la variance, la fonction de covariance ou sa transformée de Fourier, le spectre. À cet égard, on peut poser qu'un processus stochastique stationnaire peut-être également défini de façon appropriée par un modèle avec quelques paramètres seulement.

Nous commencerons par définir ce que l'on entend par processus stochastique stationnaire.

Intuitivement, on dit qu'un processus est stationnaire s'il est en équilibre statistique, en ce sens que ses propriétés restent invariables dans le temps. En d'autres termes, il s'agit d'un processus sans tendance, et il est possible de récapituler toutes ses caractéristiques en calculant certaines fonctions à partir des données. La fonction d'autocorrélation a été la première fonction de ce type à être étudiée.

D'un point de vue statistique, on définit un processus stochastique

$X = \{X(\omega, t), \omega \in \Omega, t \in T\}$ comme un processus stationnaire ou strictement stationnaire si toutes les fonctions de distribution à dimensions finies (2.4) demeurent les mêmes lorsque l'ensemble de points t_1, t_2, \dots, t_n se déplace le long de l'axe du temps t , c'est-à-dire si

$$(3.1) F_{t_1, t_2, \dots, t_n}^{(x_1, x_2, \dots, x_n)} = F_{t_1+T, t_2+T, \dots, t_n+T}^{(x_1, x_2, \dots, x_n)}$$

pour tout n , t_1, t_2, \dots, t_n et T .

Cela implique en particulier dans le cas d'un processus stochastique stationnaire que toutes les distributions unidimensionnelles de $F_t(x)$ (2.2) doivent être identiques.

En d'autres termes, elles sont indépendantes des valeurs de t . Dans le cas des fonctions de distribution bidimensionnelles du type (2.3), l'expression (3.1) implique qu'elles ne peuvent dépendre que de la différence temporelle $t_2 - t_1$, mais non pas des valeurs de t_1 et t_2 et, de façon générale, suivant l'expression (3.1), les fonctions de distribution finie à n dimensions ne dépendent que des différences $t - t_j$ ($j = 2, 3, \dots, n$).

Dans la pratique, au lieu de définir un processus comme strictement stationnaire, il est très utile de le considérer comme stationnaire lato sensu, ou stationnaire du second ordre. Dans ce cas, seules sont définies les propriétés des deux premiers moments. Ainsi, on dit qu'un processus stochastique est stationnaire lato sensu si:

i) la moyenne ou moment du premier ordre $\mu(t)$ est une constante.

La représentation symbolique est:

$$(3.2) \mu(t) = E[X(t)] = \int_{-\infty}^{\infty} x dF_t(x) = m$$

ii) la fonction d'autocovariance $\sigma_{XX}(t_1, t_2)$ définie comme:

$$(3.3) \sigma_{XX}(t_1, t_2) = E[X(t_1) - \mu(t_1)][X(t_2) - \mu(t_2)]$$

est finie et dépend d'une seule variable qui est la différence entre deux points dans le temps, t_1 et t_2 , c'est-à-dire, si

$$(3.4) \sigma_{XX}(t_1, t_2) = \sigma_{XX}(t_1 + T, t_2 + T)$$

ce qui, en posant $T = -t_1$, donne

$$(3.5) \sigma_{XX}(t_1, t_2) = \sigma_{XX}(0, t_2 - t_1) = \sigma_{XX}(t_2 - t_1) = \sigma_{XX}(\tau)$$

dans laquelle $\tau = t_2 - t_1$ est le retard entre les deux variables aléatoires. Pour $t_2 - t_1 = 0$, l'expression (3.5) donne la variance $\sigma_{XX}(0)$ du processus, et en divisant l'expression (3.5) pour la variance, nous obtenons les fonctions d'autocorrélation $\rho(\tau)$,

$$(3.6) \rho(\tau) = \rho_{XX}(t_1, t_2) = \frac{\sigma_{XX}(\tau)}{\sigma_{XX}(0)}$$

La moyenne est une caractéristique importante d'un processus stochastique, mais elle ne donne que les propriétés les plus grossières du processus; elle ne sert qu'à mesurer la localisation. La fonction d'autocovariance constitue une meilleure description. Pour $t_1 = t_2$, $\sigma_{XX}(\tau)$ se ramène à la variance qui est une mesure de dispersion dans le sens de la moyenne quadratique, et pour tout $t_2 \neq t_1$, $\sigma_{XX}(\tau)$ est une mesure de l'association linéaire des variables aléatoires dans le temps.

Pour comparer les fonctions d'autocovariance de deux processus stochastiques différents, il est pratique d'utiliser la fonction d'autocorrélation qui élimine l'influence de l'unité de mesure des variables aléatoires utilisées.

Observons qu'un processus stationnaire du second ordre coïncide avec un processus strictement stationnaire si l'on suppose par hypothèse que le processus suit une loi de Gauss, c'est-à-dire une loi normale, avec des moments du second ordre finis.

L'exemple le plus simple d'un processus strictement stationnaire est le processus purement aléatoire ou bruit blanc, pour lequel l'on suppose que les variables aléatoires se caractérisent par l'indépendance sérielle et par une distribution identique. C'est l'hypothèse sous-jacente à la composante aléatoire utilisée dans les modèles d'erreurs qui sont des processus normaux non stationnaires dans la moyenne.

Tout processus qui ne satisfait pas aux conditions i) et ii) est un processus stochastique non stationnaire. Dans la classe des processus non stationnaires, les processus appelés non stationnaires homogènes, ou processus à accroissements stationnaires, ont été étudiés pour la première fois par Kolgomorov (1941) et Yaglom (1955).

Les processus de ce type sont non stationnaires, mais en prenant les différences de façon adéquate, on obtient un processus stationnaire pour un ordre fini de différences. Au niveau des applications empiriques, très souvent la distinction entre une partie de la série observée et une autre partie de la même série ne se fait qu'au niveau local et (ou) à celui de la pente de la courbe; l'ordre de différence ne dépasse donc pas 2.

Les processus non stationnaires homogènes généralisent la théorie des fonctions aléatoires stationnaires. Il est évident que tout processus stationnaire est aussi un processus à accroissements stationnaires.

La dérivée (différence) d'une fonction (séquence) aléatoire d'accroissements stationnaires est un processus stationnaire; inversement, l'intégrale indéfinie (somme infinie) d'un processus stationnaire est un processus d'accroissements stationnaires.

Nous donnons maintenant la définition suivante d'un processus à accroissements stationnaires (Yaglom, 1955): "le processus aléatoire $X(t)$ est appelé processus à accroissements stationnaires si l'espérance mathématique de l'accroissement de $X(t)$ sur tout intervalle de temps est proportionnelle à la longueur de l'intervalle, de sorte que $E[X(s) - X(t)] = a(s-t)$; a étant une constante et la fonction de structure $D(t; u, v)$ du processus $X(t)$ dépend uniquement des différences $u-t$ et $v-t$, c'est-à-dire $D(t; u, v) = D(u-t, v-t) = D(\tau_1, \tau_2)$ ".

La fonction de structure convient mieux que la fonction d'autocovariance à la description de ce type de processus, et elle a été utilisée pour la première fois par Kolgomorov (1941). Ainsi, un processus à accroissements stationnaires est caractérisé par une constante a , que l'on peut poser dans la pratique comme égale à zéro, et par la fonction de structure, qui est une fonction de deux variables:

$$(3.7) \quad D(\tau_1, \tau_2) = E[X(t+\tau_1) - X(t)] [X(t+\tau_2) - X(t)].$$

Dans le cas réel, au lieu de l'expression (3.7), nous avons une fonction d'une variable:

$$(3.8) \quad D(\tau) = E[X(t+\tau) - X(t)]^2.$$

Dans un processus stationnaire réel, \underline{a} est égale à zéro, et $D(\tau)$ peut être exprimée en termes de la fonction d'autocovariance $\sigma(\tau)$ de la façon suivante:

$$(3.9) \quad D(\tau) = 2\sigma(0) - \sigma(\tau) - \sigma(-\tau) = 2\sigma(0) - 2\sigma(\tau).$$

Les modèles d'erreurs que nous examinerons à la section 4 sont des processus non stationnaires homogènes quand la composante déterminable (le signal) est posée comme un polynôme du temps. En prenant les différences successives, le processus se ramène à un processus stationnaire linéaire. Un autre type important de processus stochastique est le processus non stationnaire linéaire homogène. Dans ce cas, la propriété de linéarité vient du fait que le processus est considéré comme l'extrant d'un filtre linéaire dont l'intrant est purement aléatoire (bruit blanc).

Dans ce type de processus, les valeurs précédentes de la variable aléatoire $X(t)$ contribueront à la détermination de sa valeur au moment t . La non-stationnarité peut se présenter dans la moyenne et (ou) dans la variance. Un exemple simple est fourni par le processus promenade au hasard, dans lequel, à mesure que le temps s'écoule, les variables aléatoires tendent à osciller autour de la moyenne (une droite), avec une amplitude croissante. On a utilisé ce type de processus dans les séries temporelles économiques, et notamment pour ajuster les cours de la bourse. Supposons par exemple que W_t est un processus purement aléatoire, et que X_t est un autre processus lié à W_t de la façon suivante:

$$\begin{aligned} X_1 &= W_1 \\ X_2 &= X_1 + W_2 \\ &\vdots \\ X_t &= X_{t-1} + W_t \end{aligned}$$

On peut alors mettre X_t sous la forme d'une combinaison linéaire du processus purement aléatoire W_t dont tous les coefficients de pondération sont égaux à 1. Si la valeur attendue de W_t est μ et la variance σ_W^2 , il s'ensuit que:

$$(3.7) E(X_t) = t \mu$$

et (3.8) $\text{var } X_t = t \sigma_W^2$.

L'autocovariance du processus X_t est:

$$(3.9) \sigma_{XX}(t_1, t_2) = \min(t_1, t_2) \sigma_W^2$$

4. LES MODÈLES D'ERREURS

Les premières études sur les séries temporelles ont été celles faites en recherche astronomique. On a construit des modèles pour déterminer la position d'une planète à un instant donné. La nature des problèmes auxquels se sont heurtés les astronomes les a poussés à construire des modèles très simples, connus aujourd'hui sous le nom de modèles d'erreurs, ou plus simplement de modèles de régression (Pannekoek, 1961).

Les modèles d'erreurs supposent que les observations successives d'une série temporelle sont indépendantes. L'ensemble ordonné $\{x_t, t = \dots -1, 0, 1 \dots\}$ (1) peut être mis sous la forme d'un modèle général composé d'une fonction entièrement déterminable $f(t)$ du temps, qui est appelée signal ou composante systématique du modèle, et d'un élément aléatoire U_t appelé bruit blanc (ou composante purement stochastique) du processus qui est, par hypothèse, indépendant de $f(t)$. Le modèle peut s'écrire sous la forme suivante:

$$(4.1) X_t = f(t) + U_t \quad t = \dots -1, 0, 1, \dots$$

L'hypothèse d'indépendance des observations successives est introduite dans l'ensemble d'hypothèses qui caractérisent la composante aléatoire U_t , que l'on suppose être distribuée identiquement, avec une moyenne nulle, une variance constante à chaque instant et une autocorrélation nulle. En vertu de ces hypothèses, tout effet temporel doit être fait dans le signal $f(t)$.

(1) Nous utiliserons des minuscules lorsque nous dénoterons des quantités observées, et des lettres majuscules pour les variables stochastiques correspondantes. Nous écrirons t en sous-indice lorsque le temps est de nature discrète, et comme argument lorsque le temps est continu.

Dans certains cas, $f(t)$ est une fonction connue du temps ou d'autres quantités observables, et de paramètres. Si l'on suppose que $f(t)$ est linéaire dans les paramètres, elle se ramène à une "fonction de régression", et l'on estime ses paramètres par la méthode des moindres carrés.

Lorsque, par rapport à ses paramètres, le signal n'est pas linéaire, comme dans le cas des fonctions de croissance, l'estimation des paramètres et le test des hypothèses nécessitent des transformations préalables ou l'utilisation de méthodes plus compliquées.

Dans d'autres cas, la forme analytique de la composante systématique n'est pas connue, mais elle peut être assez bien approximée par des combinaisons linéaires de fonctions temporelles connues, comme par exemple des puissances du temps t ou des fonctions trigonométriques de t .

De façon générale, il faut distinguer deux types de fonctions temporelles dans le cas de $f(t)$. L'une est un polynôme d'un degré assez bas qui satisfait à l'hypothèse selon laquelle la composante systématique se déplace lentement, doucement et progressivement dans le temps. L'autre est une combinaison linéaire des sinus et des cosinus du temps avec des coefficients constants (une série finie de Fourier) qui tient compte des fluctuations cycliques, qu'elles soient strictement périodiques ou non.

Dans le cas des séries temporelles économiques, le signal $f(t)$ a très souvent un comportement qui nécessite l'addition ou la multiplication de deux types de fonction. On dit alors que le polynôme représente une tendance générale à long terme, et la série finie de Fourier représente alors les fluctuations cycliques, et les comportements saisonnières.

Dans tous les cas exposés ci-dessus, l'estimation des paramètres du signal se fait grâce aux méthodes de l'analyse de régression (la théorie classique des moindres carrés) ou par des méthodes analogues. En vertu de l'hypothèse de normalité de la distribution des variables aléatoires, il est possible de faire tous les tests d'hypothèse.

Mais il arrive parfois qu'une simple fonction temporelle ne puisse approximer le signal sur toute la longueur de l'intervalle de temps, en dépit d'une forme lisse et de l'absence de fluctuations marquées sur n'importe quel intervalle temporel, aussi petit soit-il. On a alors habituellement recours à des méthodes non paramétriques pour le lissage.

Le lissage d'une séquence x_t signifie le passage à une nouvelle séquence x_t^* provenant de la séquence initiale en formant la moyenne mobile.

$$(4.2) \quad x_t^* = \sum_{k=-n}^n a_k x_{t+k} \quad t=n+1, \dots, T-n.$$

Pour des raisons de simplicité, nous avons écrit une somme finie, mais s'il y a satisfaction à certaines conditions de convergence, nous pouvons étendre la somme à l'infini. ∞ .

De plus, les a_k sont habituellement normalisés, de sorte que $\sum_{k=-n}^n a_k = 1$.

Dans la pratique, on se sert largement du lissage pour filtrer les séquences afin de réduire les effets des erreurs de mesure et autres écarts faiblement corrélés.

En admettant que nous ayons un nombre fini d'observations dans le modèle (4.1), nous obtenons après lissage:

$$(4.3) \quad x_t^* = \sum_{-n}^n a_k f(t+k) + \sum_{-n}^n a_k u_{t+k} \quad t=n+1, \dots, T-n$$

où $E(u_t) = 0$, $E(u_t^2) = \sigma_u^2$ et $E(u_t u_s) = 0$, $t \neq s$. Alors, u_t^* a une variance $\sigma_u^2 \sum_{-n}^n a_k^2$.

La moyenne quadratique de u_t^* diminue considérablement si les coefficients de pondération de lissage a_k sont judicieusement choisis, tandis que le signal $f(t)$ est à peu près le même. La moyenne pondérée des valeurs observées sert donc à estimer la tendance générale, et la courbe assez irrégulière engendrée par les points observés est remplacée par une courbe lisse de la moyenne mobile. Toutefois, les termes successifs de la série lisse sont corrélés. Nous avons:

$$E(u_t^* u_{t+h}^*) = \begin{cases} \sigma_u^2 \sum_{-n}^n a_k a_{k-h} & ; & h = 0, 1, \dots, 2n \\ 0 & ; & h = 2n+1, \dots \end{cases}$$

À noter que l'hypothèse de lissage est une propriété locale, alors que l'hypothèse d'une tendance générale de forme polynômiale intéresse tout l'intervalle de temps $t = 1, 2, \dots, T$. L'hypothèse de lissage ne porte que sur les points voisins d'un instant donné t pour estimer la tendance générale en ce point, alors que l'hypothèse d'une tendance de forme polynômiale implique que toutes les observations servent à estimer la fonction qui représente la tendance tout au long de l'intervalle considéré.

La base générale de la plupart des méthodes de lissage est l'ajustement d'un polynôme(1) à $2n + 1$ observations successives et l'utilisation de ce polynôme pour estimer la valeur de la tendance au centre de l'intervalle de $2n + 1$ observations. Etant donné que les estimations des paramètres du polynôme sont linéaires par rapport aux valeurs observées x_{t+k} , l'estimation de la tendance générale prend la forme de l'expression (4.2).

Supposons que la tendance générale $f(t+k)$ aux instants $t+k = t - n, \dots, t+n$, peut être approximée par un polynôme en k de degré p , c'est-à-dire

$$(4.4) f_t(k) = \alpha_0 + \alpha_1 k + \alpha_2 k^2 + \dots + \alpha_p k^p; \quad k = -n, \dots, n$$

alors pour $k = 0$, $f(t)$ est approximativement $f_t(0) = \alpha_0$.

Nous pouvons estimer les $\hat{\alpha}_j$ d'après les valeurs observées x_{t-n}, \dots, x_{t+n} grâce à la méthode des moindres carrés. Les équations normales des estimations de $\hat{\alpha}_0, \hat{\alpha}_1, \dots, \hat{\alpha}_p$ sont:

$$(4.5) \hat{\alpha}_0 \sum_{-n}^n k^j + \hat{\alpha}_1 \sum_{-n}^n k^{j+1} + \dots + \hat{\alpha}_p \sum_{-n}^n k^{j+p} = \sum_{-n}^n k^j x_{t+k}$$

$$J = 0, 1, \dots, p$$

À noter que les sommes $\sum k^j$ sont des fonctions de n seulement.

(1) Une méthode semblable s'appuie sur l'emploi d'une formule obtenue par l'ajustement d'une fonction harmonique telle que

$$f(t) = \alpha_0 + \sum_{j=1}^p [\beta_j \cos \omega_j k + \gamma_j \sin \omega_j k]$$

où p et les angles ω_j doivent être choisis a priori, et l'ajustement détermine les coefficients α_0, β_j et γ_j .

Par symétrie, la somme de toute puissance impaire de k est nulle. Étant donné que notre estimation de f(t) sera $\hat{\alpha}_0$ nous ne nous intéressons qu'aux équations dans lesquelles j est pair, à savoir

$$(4.6) \quad \hat{\alpha}_0 - \sum_{-n}^n k^j + \hat{\alpha}_2 - \sum_{-n}^n k^{j+2} + \dots + \hat{\alpha}_p - \sum_{-n}^n k^{j+p} = \sum_{-n}^n k^j x_{t+p} \quad j = 0, 2, 4, \dots, p$$

Les équations à résoudre en $\hat{\alpha}_0$ sont les mêmes pour les p impairs que pour la valeur paire inférieure suivante de p. Il nous suffit de considérer seulement $p = 0, 2, 4, \dots$. On peut partager le système (4.6) en deux, une équation pour $j = 0$, et un système d'équations pour $j = 2, 4, \dots, p$. L'expression (4.6) devient alors:

$$(4.7) \quad (2n+1)\hat{\alpha}_0 + 2 \sum_{-1}^n k^2 \hat{\alpha}_2 + \dots + 2 \sum_{-1}^n k^p \hat{\alpha}_p = \sum_{-n}^n x_{t+k}; \quad j=0$$

$$2 \sum_{-1}^n k^j \hat{\alpha}_0 + 2 \sum_{-1}^n k^{j+2} \hat{\alpha}_2 + \dots + 2 \sum_{-1}^n k^{j+p} \hat{\alpha}_p = \sum_{-1}^n k^j (x_{t-k} + x_{t+k})$$

$$j=2, 4, \dots, p$$

En résolvant (4.7) par rapport à $\hat{\alpha}_0$, nous obtenons:

$$(4.8) \quad \hat{\alpha}_0 = \sum_{-n}^n a_k x_{t+k}$$

où $a_{-k} = a_k$ et les a_k sont des fonctions de n et p seulement et des polynômes en k. L'estimation de la tendance générale x_t^* donnée par la formule (4.2) est $x_t^* = \hat{\alpha}_0$, en vertu de l'expression (4.8).

L'ajustement d'un polynôme par la méthode des moyennes mobiles consiste donc à déterminer les coefficients de pondération \underline{a} et à calculer pour chaque ensemble consécutif de $2n + 1$ termes de la série une valeur $\hat{\alpha}_0$ donnée par (4.8).

On peut prouver que pour un p donné, la variance de la séquence plus lisse diminue à mesure que le nombre d'observations ou de points augmente et que pour un nombre donné de points (n), la variance augmente à mesure que p augmente. Le choix de p et de n donnant un ajustement optimum est un problème statistique à décisions multiples (Anderson, T.W., 1971).

L'avantage du lissage dans l'estimation du signal réside dans sa souplesse. Mais, comme la méthode n'est pas basée sur un modèle probabiliste explicite, on ne peut l'exprimer entièrement en termes de statistique mathématique, et l'induction statistique est étroitement limitée. Ainsi $f(t)$ n'est pas déterminé par un petit nombre de paramètres pour lesquels on peut donner une région de confiance. On ne peut non plus tester les hypothèses relatives à la tendance générale, ou relier la tendance estimée à une théorie ou à un modèle pour la génération des séries observées.

Nous n'insisterons pas davantage sur ce point qui s'attache au problème de l'estimation du signal et non pas au modèle probabiliste qui engendre la série temporelle. Nous allons maintenant récapituler les principales propriétés des modèles d'erreurs présentés dans cette section. Les modèles d'erreurs supposent que la série temporelle observée se compose d'une fonction bien connue du temps, et d'un terme d'erreur, qui est

le bruit blanc. Les observations n'ont donc pas de dépendance sérielle, et l'effet du temps n'est pas compris dans U_t ; on suppose au contraire qu'il n'affecte que $f(t)$. La forme analytique que l'on donne habituellement à $f(t)$ est une combinaison linéaire de fonctions du temps, comme par exemple des puissances du temps t , ou des fonctions trigonométriques de t (sinus et cosinus). L'estimation de la tendance générale se fait soit par la méthode des moindres carrés, soit par une méthode de lissage comme celle des moyennes mobiles.

Ces types de modèles sont acceptables lorsque les erreurs dans les séries temporelles observées peuvent être attribuées uniquement à des erreurs de mesure ou d'observation. Dans ce cas, tout écart est considéré comme purement temporaire, sans aucun effet sur le mouvement futur du processus. Il semble que ce soit le cas dans de nombreux problèmes astronomiques. Ainsi, les erreurs relatives à la position observée des planètes à un instant donné t peut s'expliquer par des erreurs d'observation dues aux conditions atmosphériques ou aux défauts du télescope. De telles erreurs ne vont pas affecter les positions ultérieures des planètes, ni les observations que nous allons en faire. Les propriétés des séries temporelles sont alors récapitulées dans la moyenne du signal, et les variations de la composante aléatoire n'influent pas sur ces propriétés.

Lorsqu'il y a autocorrélation des erreurs, les modèles d'erreurs ne sont plus représentatifs du processus générateur de séries temporelles. Les observations sont maintenant stochastiquement dépendantes dans le temps, et la valeur courante de t dépendra du temps qui s'est

écoulé depuis le début du processus. L'interprétation que l'on donne des erreurs autocorrélées est que les erreurs ne s'expliquent pas simplement par des observations indépendantes fausses, mais aussi par d'autres facteurs systématiques et que, dès leur apparition, elles s'incorporent au processus et influent sur son évolution ultérieure. Il semble que ce soit le cas pour la plupart des séries temporelles, en particulier les séries temporelles sociales et économiques. On a attribué à Yule (1921) (1927) la paternité de cet aspect. Les processus qui prennent en considération l'hypothèse des erreurs autocorrélées sont les processus stationnaires linéaires et les processus non stationnaires linéaires homogènes.

Nous allons examiner dans les sections suivantes ces types de modèles.

5. LES PROCESSUS STOCHASTIQUES STATIONNAIRES LINÉAIRES

On dit qu'un processus stochastique X_t est linéaire s'il existe un processus purement aléatoire U_t et une séquence de paramètres

$\alpha_0, \alpha_1, \alpha_2, \dots$ telle que l'on ait:

$$(5.1) X_{t-m} = U_t + \alpha_1 U_{t-1} + \alpha_2 U_{t-2} + \dots = \sum_{k=0}^{\infty} \alpha_k U_{t-k}; \alpha_0 = 1.$$

Si l'intervalle de temps est continu, la relation (4.2.1) s'écrit:

$$(5.2) X(t)-m = \int_0^{\infty} \alpha(v) U(t-v) dv.$$

Alors $E(U_t) = E[U(t)] = 0$ and $E(X_t) = E[X(t)] = m$.

Si la série, ou l'intégrale, est convergente, l'expression 5.1 (ou 5.2) définit un processus stochastique stationnaire linéaire, dans lequel m est l'espérance mathématique du processus.

Si la série, ou l'intégrale, n'est pas convergente, l'expression (5.1) ou (5.2) définit un processus non stationnaire linéaire, et m n'a pas de signification précise, sauf comme point de référence du niveau du processus. On appelle parfois l'expression (5.1) moyenne mobile infinie, bien que la somme infinie des coefficients ne soit pas posée par hypothèse comme étant égale à 1.

S'il y a un nombre fini de termes, l'expression (5.1) définit un processus stochastique paramétrique linéaire à paramètre temporel discret.

Nous ne nous occuperons dans la présente section que des processus à paramètre temporel discret. (L'analyse des modèles à paramètre temporel continu suit une démarche semblable).

En utilisant l'opération de retard L , où $L^0 = 1$, $LU_t = U_{t-1}$ and $L^n U_t = U_{t-n}$, the l'expression (5.1) pour $m=0$ peut s'écrire sous la forme concise suivante:

$$(5.3) X_t = (1 + \alpha_1 L + \alpha_2 L^2 + \dots) U_t = \sum_{k=0}^{\infty} \alpha_k L^k U_t = \alpha(L) U_t$$

où $\sum_{k=0}^{\infty} \alpha_k L^k = \alpha(L)$, est le filtre linéaire ou l'opération qui transforme la variable d'entrée U_t en la variable de sortie X_t . On l'appelle aussi la fonction de transfert du système linéaire, ou fonction génératrice des coefficients de pondération. Dans ce dernier cas, on peut considérer L comme une variable auxiliaire, dont la k -ième puissance est le coefficient de α_k . Pour que l'expression (5.3) soit stationnaire lato sensu, il faut et il suffit que la fonction de transfert $\alpha(L)$ soit convergente pour $|L| \leq 1$ ce qui implique que $\sum_{k=0}^{\infty} |\alpha_k| < M$, M étant fini.

5.1 Processus autorégressifs (A.R.)

C'est U. Yule (1921) qui parla pour la première fois des processus autorégressifs. Ces derniers sont une sous-classe des processus linéaires. On dit qu'un processus X_t est un processus autorégressif d'ordre p et dénoté AR(p) si on peut le mettre sous la forme d'une combinaison linéaire de p valeurs précédentes du processus, plus un élément purement aléatoire U_t . On peut l'écrire sous la forme suivante:

$$(5.1.1) X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \dots + \beta_p X_{t-p} = U_t.$$

En appliquant l'opérateur de retard L , l'expression (5.1.1) devient:

$$(5.1.2) \quad U_t = (1 + \beta_1 L + \beta_2 L^2 + \dots + \beta_p L^p) X_t = \sum_{r=0}^p \beta_r L^r X_t; \quad \beta_0 = 1.$$

Si le processus X_t est stationnaire, l'expression (5.1.2) peut s'écrire sous la forme d'une moyenne mobile infinie:

$$(5.1.3) \quad X_t = \frac{1}{\sum_{r=0}^p \beta_r L^r} U_t = \beta^{-1}(L) U_t = \sum_{r=0}^{\infty} h_r L^r U_t.$$

En effet, on peut remplacer chaque X_{t-1} , $l=1, \dots, p$ par un processus autorégressif, ce qui nous donne finalement une série infinie de U .

Il existe plusieurs conditions équivalentes pour la stationnarité de (5.1.3). Lorsqu'elle est écrite sous la forme d'une moyenne mobile infinie, nous avons déjà mentionné dans les sections précédentes que la série infinie des coefficients de pondération doit être convergente pour $|L| < 1$, ce qui implique que $\sum_{r=0}^{\infty} |h_r|$ fini.

Ces conditions de convergence sont équivalentes à $\sum_{r=0}^p \beta_r L^r \neq 0$ car si $\sum_{r=0}^p \beta_r L^r = 0$, la fonction de transfert devient infinie.

Voyons maintenant les conditions pour que $\sum_{r=0}^p \beta_r L^r \neq 0$

$$(5.1.4) \quad \sum_{r=0}^p \beta_r L^r = (1 + \beta_1 L + \beta_2 L^2 + \dots + \beta_p L^p) = \beta(L).$$

Notons que $\beta(L)$ est un polynôme en L d'ordre p . En décomposant L^p en facteurs et en posant $L = \frac{1}{G}$ we nous avons:

$$(5.1.5) \quad \left(\frac{1}{G}\right) = \frac{1}{G^p} (G^p + \beta_1 G^{p-1} + \beta_2 G^{p-2} + \dots + \beta_p) =$$

$$\frac{1}{G^p} (G-G_1) (G-G_2) \dots (G-G_p)$$

dans laquelle G_1, G_2, \dots, G_p sont les p racines de l'équation caractéristique de $\beta(\frac{1}{G})$.

Il est évident que l'on peut réécrire l'expression (5.1.5) sous la forme suivante:

$$(5.1.6) \beta(L) = (1-G_1L)(1-G_2L) \dots (1-G_pL).$$

Par conséquent, s'il existe un $G_i = \frac{1}{L}$, $\beta(L) = 0$, le processus (5.1.3) est non stationnaire. Pour qu'il soit stationnaire, les racines de $\beta(L)$, à savoir $L_i = G_i^{-1}$, $i = 1, 2, \dots, p$ doivent être en valeur absolue supérieures à 1 ou $|G_i| < 1$. La condition $|G_i| < 1$ coïncide avec les conditions de stabilité de la composante déterminable de l'expression (5.1.1). À noter que si $U_t = 0$, (5.1.1) se ramène à une équation aux différences finies homogène d'ordre p , dont la solution générale est:

$$(5.1.7) X_t = A_1 G_1^t + A_2 G_2^t + \dots + A_p G_p^t$$

où G_i , $i = 1, 2, \dots, p$ sont les racines de son équation caractéristique, et pour que l'expression (5.1.7) soit stationnaire, 1 pour $|G_i| < 1$, for $i = 1, 2, \dots, p$.

Dans le cas d'un paramètre temporel continu, le processus autorégressif (5.1.1) est défini par une équation différentielle d'ordre p , et pour qu'il soit stationnaire, les racines de son équation caractéristique doivent toutes avoir une partie réelle négative.

5.2 Les processus de moyenne mobile finie (M.M.)

Le modèle autorégressif examiné ci-dessus a exprimé X_t sous la forme d'une somme pondérée finie de p valeurs précédentes du processus X_t , plus une composante purement aléatoire U_t . De même, si le processus est stationnaire, on peut le mettre sous la forme d'une somme pondérée infinie des U_t , dans laquelle $E(U_t) = 0$, $E(U_t)^2 = \sigma_U^2$ et $E(U_t U_{t+\tau}) = 0$ for all $\tau \neq 0$.

Or, on dit qu'un processus X_t est un processus de moyenne mobile finie d'ordre q $MM(q)$ s'il est une combinaison linéaire de variables purement aléatoires $U_t, U_{t-1}, \dots, U_{t-q}$. En d'autres termes,

$$(5.2.1) \quad X_t = U_t + \alpha_1 U_{t-1} + \alpha_2 U_{t-2} + \dots + \alpha_q U_{t-q} = \sum_{r=0}^q \alpha_r U_{t-r}; \quad \alpha_0=1.$$

En utilisant l'opérateur de retard L , l'expression (5.2.1) s'écrit

$$(5.2.2) \quad X_t = \sum_{r=0}^q \alpha_r L^r U_t = \alpha(L)U_t$$

où $\alpha(L) = 1 + \alpha_1 L + \alpha_2 L^2 + \dots + \alpha_q L^q$, est la fonction génératrice des coefficients de pondération.

L'expression (5.2.1) est toujours stationnaire, c'est-à-dire que $\alpha(L)$ est convergent pour $|L| \leq 1$ mais pour que l'expression (5.2.1) soit inversible, c'est-à-dire pour qu'elle soit exprimée sous la forme d'un processus autorégressif infini,

$$(5.2.3) \quad U_t = \left(\sum_{r=0}^q \alpha_r L^r \right)^{-1} X_t = \alpha^{-1}(L)X_t = \sum_{r=0}^{\infty} \delta_r L^r X_t$$

les racines de $\alpha(L) = 0$ doivent se trouver à l'extérieur du cercle unité.

5.3 Processus autorégressifs et de moyenne mobile (A.R.M.M.)

Pour des applications empiriques, une combinaison d'un processus autorégressif disons d'ordre p avec un de moyenne mobile, par exemple d'ordre q , présente l'avantage de ne faire intervenir que très peu de paramètres.

On peut donc écrire un processus A.R.M.M. (p, q) sous la forme suivante:

$$(5.3.1) X_t = U_t + \alpha_1 U_{t-1} + \dots + \alpha_q U_{t-q} - \beta_1 X_{t-1} - \dots - \beta_p X_{t-p}$$

ou encore:

$$(5.3.2) \left(\sum_{r=0}^p \beta_r L^r \right) X_t = \left(\sum_{r=0}^q \alpha_r L^r \right) U_t; \quad \alpha_0 = \beta_0 = 1$$

et, par conséquent:

$$(5.3.3) X_t = \frac{\sum_{r=0}^q \alpha_r L^r}{\sum_{r=0}^p \beta_r L^r} U_t.$$

Le processus A.R.M.M. (p, q) est considéré comme l'extrait X provenant d'un U_t purement aléatoire, ou bruit blanc, où les coefficients de pondération de la fonction de transfert s'obtiennent par la division de deux polynômes. Le nombre de paramètres du modèle (5.3.1) est $p+q+2$, y compris la moyenne de X_t et la variance de U_t . Dans la plupart des cas résolus grâce à ce modèle, p et q ne sont pas supérieurs à 2 (Box et Jenkins, 1970).

Pour que le processus A.R.M.M. (p, q) soit stationnaire, les mêmes conditions que celles examinées dans les sections précédentes doivent être remplies, c'est-à-dire que les racines de l'équation caractéristique $\sum_{r=0}^p \beta_r L^r = \beta(L) = 0$ doivent être en valeur absolue supérieures à 1. L'expression (5.3.3) est inversible si

$\sum_{r=0}^q \alpha_r L^r = \alpha(L) = 0$ a toutes ses racines à l'extérieur du cercle unité. Alors:

$$(5.3.4) \quad U_t = \frac{\sum_{r=0}^p \beta_r L^r}{\sum_{r=0}^q \alpha_r L^r} X_t .$$

6. LES MODÈLES STOCHASTIQUES STATIONNAIRES NON PARAMÉTRIQUES: LA FONCTION D'AUTOCOVARIANCE ET LE SPECTRE

Les modèles que l'on vient d'examiner sont tous les modèles paramétriques, c'est-à-dire qu'ils ont un nombre fini de paramètres. On peut aussi décrire le processus générateur d'une série temporelle stationnaire par des modèles non paramétriques, c'est-à-dire par des modèles qui n'ont pas un nombre fini de paramètres. Parmi les méthodes non paramétriques, l'analyse de l'autocovariance et les fonctions d'autocorrélation et leurs transformées de Fourier, le spectre et le spectre normalisé sont les plus appropriées.

D'un point de vue mathématique, ces fonctions sont des couples de Fourier, et elles sont donc équivalentes. Elles donnent toutes les deux les mêmes types de renseignements en probabilité, en ce sens qu'elles caractérisent tous les moments du second ordre d'un processus stochastique stationnaire. L'utilisation soit de fonction d'autocovariance, soit du spectre, dépend des propriétés particulières des données que l'on désire faire ressortir. Le spectre porte surtout sur le domaine des fréquences, tandis que l'analyse de la fonction d'autocovariance donne les même informations, mais sur le domaine temporel.

Dans la représentation spectrale, un processus stationnaire est considéré comme une combinaison linéaire de composantes oscillant de façon aléatoire, et la variance totale se distribue sur la fréquence. Si le processus est défini en termes d'un paramètre temporel discret, il est possible de déterminer la proportion de la variance attribuable à chaque composante avec une fréquence particulière λ , mais dans le cas des processus de chroniques continus, nous parlerons de l'apport d'une bande de fréquence autour d'une fréquence λ donnée.

Une application importante de la fonction d'autocorrélation et du spectre normalisé est l'identification des filtres linéaires qui minimisent l'erreur quadratique moyenne lorsque la composante systématique d'un processus est perturbée par la présence d'une composante purement aléatoire.

Les deux fonctions se révèlent également très utiles comme indicateurs initiaux lors de la construction d'un modèle probabiliste pour le mécanisme générateur de la série temporelle. Ainsi, par exemple, une fonction d'autocorrélation qui est positive pour des valeurs successives de τ (retards) et qui tend vers zéro à mesure que τ s'accroît, va traduire à la fois un comportement lisse de la chronique et le fait que le processus est davantage de type autorégressif fini que de type purement aléatoire. L'ordre du processus autorégressif peut également être obtenu à partir de la fonction d'autocorrélation partielle. Le domaine de fréquences donne les mêmes informations grâce à une fonction de densité spectrale normalisée, (spectre normalisé) où prédominent les fréquences peu élevées. Par contre, lorsque des valeurs adjacentes de $\rho(\tau)$ sont corrélées négativement, le processus générateur de la série temporelle va présenter une grande fluctuation sur de courtes périodes, et la fonction de densité spectrale normalisée correspondante se caractérisera par une prédominance de fréquences élevées. Mais en dépit de son importance dans la construction de modèles (en particulier en génie et en physique), l'analyse spectrale s'est avérée plus appropriée dans les études de réponse de la fréquence, et dans le domaine de l'élaboration d'expériences pour l'optimisation des procédés industriels.

Dans l'analyse des séries temporelles économiques, les premières méthodes non paramétriques se fondaient sur la fonction d'autocovariance. À la fin des années 50 et au cours des années 60, de telles méthodes ont été presque complètement abandonnées en faveur de l'analyse spectrale. Dernièrement, la fonction d'autocovariance a commencé à connaître un regain de faveur en raison surtout de l'existence de nouveaux programmes informatiques (Box et Jenkins, 1970).

La fonction d'autocovariance d'un processus stationnaire X_t est par définition:

$$(6.1) \quad \sigma_{XX}(\tau) = E(X_t X_{t+\tau}) \quad \tau = \dots, -1, 0, 1, \dots$$

où τ est le décalage temporel, que l'on suppose ici être un nombre entier. Si le paramètre temporel du processus est continu, τ peut prendre toute valeur comprise entre $-\infty$ et $+\infty$.

$$(6.2) \quad \rho_{XX}(\tau) = \frac{\sigma_{XX}(\tau)}{\sigma_{XX}(0)} = \frac{\sigma_{XX}(\tau)}{\sigma_X^2}; \quad \tau = \dots, -1, 0, 1, \dots$$

Notons que

$$(6.3) \quad \sigma_{XX}(0) \rho_{XX}(\tau) = \sigma_{XX}(\tau)$$

ce qui signifie par conséquent que si nous connaissons la fonction d'autocorrélation et la variance du processus X_t , nous avons tous les renseignements fournis par l'autocovariance.

Le graphe de l'expression (6.2) est aussi appelé corrélogramme. Les propriétés fondamentales de la fonction d'autocorrélation d'un processus réel sont (nous supprimons le sous-indice X pour alléger la notation):

(1) $\rho(0)=1$

(2) $\rho(-\tau) = \rho(\tau)$. La fonction est symétrique par rapport à l'origine en raison de l'hypothèse de stationnarité; on ne doit donc la calculer que pour les retards positifs.

(3) $|\rho(\tau)| \leq 1$. C'est la conséquence du fait que la variance d'une variable aléatoire ou d'une combinaison linéaire de variables aléatoires est positive(1).

(4) La matrice d'autocorrélation est positive et semi-définie. En d'autres termes, le déterminant de la matrice d'autocorrélation et tous ses mineurs principaux sont positifs ou nuls. La propriété 4 est une généralisation de la propriété 3 et elle indique que la fonction d'autocorrélation est toujours positive et semi-définie. L'inverse est également vrai, puisque chaque fonction positive semi-définie d'un argu-

(1) Posons en effet $Y_t = \lambda_1 X_t + \lambda_2 X_{t-\tau}$ la variance de Y_t est alors

$$(1) \text{ var. } Y_t = \lambda_1^2 \text{ var. } X_t + \lambda_2^2 \text{ var. } X_{t-\tau} + 2\lambda_1\lambda_2 \text{ cov. } (X_t X_{t-\tau}).$$

Le membre droit est non négatif pour tout λ_1, λ_2 réel et le second membre est une forme quadratique en λ_1, λ_2 . Pour qu'il soit positif, il faut que ses racines soient imaginaires, ce qui implique

$$(2) \text{ var. } X_t \text{ var. } X_{t-\tau} \geq [\text{cov}(X_t X_{t-\tau})]^2$$

ou, de manière équivalente

$$(3) \rho^2(X_t, X_{t-\tau}) = \frac{[\text{cov}(X_t X_{t-\tau})]^2}{\text{var } X_t \text{ var } X_{t-\tau}} \leq 1.$$

Dans le cas d'un processus stationnaire, (3) se ramène à:

$$|\rho(\tau)| = \left| \frac{\sigma(\tau)}{\sigma(0)} \right| \leq 1.$$

ment réel (ou entier) est la fonction d'autocorrélation d'un processus stochastique continu (discret). (Cela a été démontré par Khinchin et Kolmogorov, voir Yaglom, 1962).

- (5) Si le processus est continu, $\rho(\tau)$ est défini pour toute valeur de τ comprise entre $+\infty$ et $-\infty$, et c'est une condition nécessaire et suffisante que la fonction soit continue au point $\tau = 0$ puisque ceci implique que le processus est continu en tous points (Yaglom, 1962). Si l'on suppose que le processus est purement aléatoire, cette propriété de continuité pose des problèmes.

Dans le cas d'un processus purement aléatoire discret U_t , la fonction d'autocorrélation est $\rho(0) = 1$ and $\rho(\tau) = 0$, pour tout $\tau \neq 0$. Si U_t est un processus temporel continu, nous aurions un point de singularité pour $\tau = 0$. Pour l'éviter, on redéfinit la fonction d'autocovariance de la façon suivante:

$$(6.4) \sigma_{UU}(\tau) = \sigma(0) \delta(\tau) = \sigma_U^2 \delta(\tau)$$

où $\delta(\tau)$ est une fonction delta de Dirac(1) qui prend la valeur zéro pour $\tau \neq 0$ et la valeur ∞ pour $\tau = 0$. Alors, la covariance entre des points voisins est nulle, mais la variance du processus devient alors infinie (Jenskins et Watts, 1969).

(1) On définit une fonction delta comme une séquence de fonction $\delta_n(t)$ telle que $\int_{-\infty}^{\infty} \delta_n(t) dt = 1$, pour tout n et dans la limite quand $n \rightarrow \infty$

$$\delta(t) \left\{ \begin{array}{l} = 0 \quad t \neq 0 \\ = \infty \quad t = 0 \end{array} \right.$$

Puisque toute fonction analytique, qu'elle soit périodique ou non, peut être approximée à n'importe quel degré par l'emploi de n'importe quelle classe de fonctions périodiques, la représentation spectrale d'un processus stochastique peut être effectuée en utilisant les séries de Fourier ou l'intégrale de Fourier, selon que le paramètre temporel est discret ou continu. Dans l'analyse de Fourier, les fonctions périodiques sont des sinus et des cosinus. Elles possèdent les propriétés importantes suivantes: une approximation d'un nombre donné de termes donne l'erreur quadratique moyenne minimale entre la fonction et son approximation; elles sont également orthogonales, de sorte qu'il est possible de déterminer les coefficients indépendamment les uns des autres.

L'utilisation des séries de Fourier pour décrire des phénomènes évoluant dans le temps a été proposée dans plusieurs études par Lagrange (1772-1778), Buys-Ballot (1847) et Stokes (1879), mais il convient surtout de mentionner à cet égard la méthode du périodogramme utilisé par Schuster (1898) dans le cadre de la recherche de périodicités cachées dans les données sur les taches du soleil. Dans le domaine économique, ce périodogramme a été utilisé par Moore (1914) et Beveridge (1922). L'utilisation du périodogramme pour la description des séries temporelles a échoué en raison des hypothèses de fixité des amplitudes des fréquences et des phases dans les composantes de Fourier. L'analyse spectrale moderne utilise les séries de Fourier (ou les intégrales de Fourier) en supposant que les amplitudes et les phases sont des variables aléatoires.

On a montré (Yaglom, 1962) qu'il est possible d'approximer tout processus stochastique stationnaire $X(\omega, t)$ par une combinaison linéaire ou par des oscillations harmoniques de la forme

$$(6.5) X(\omega, t) = X_k(\omega) f(t) = X_k(\omega) \operatorname{Re} e^{i(\lambda_k t + \theta)}$$

où $X_k(\omega)$ est une variable aléatoire indépendante du temps et $f(t)$ une fonction numérique de t . Le facteur numérique $\operatorname{Re} e^{i\theta}$ peut être inclus dans la variable aléatoire $X_k(\omega)$, et le produit $X_k(\omega) \operatorname{Re} e^{i\theta}$ sera simplement dénoté ici par X_k . L'expression (6.5) devient alors:

$$(6.6) X_k(t) = X_k e^{i\lambda_k t}$$

où X_k est une variable aléatoire complexe, dont la moyenne est nulle, et pour laquelle λ est une constante.

Chaque élément de la forme (6.6) va ainsi décrire une oscillation périodique de fréquence angulaire $\lambda_k(1)$, avec une amplitude aléatoire R et une phase aléatoire.

Si l'on définit le processus pour un paramètre temporel discret, nous pouvons le mettre sous la forme suivante:

$$(6.7) X_t = \sum_{k=1}^{\infty} X_k e^{i\lambda_k t}$$

et dans le cas d'un paramètre temporel continu:

$$(6.8) X(t) = \int_{-\infty}^{\infty} e^{i\lambda t} d\mathfrak{Z}(\lambda)$$

où $\mathfrak{Z}(\lambda)$ est un processus stochastique indexé sur λ .

(1) La fréquence angulaire $\lambda = 2\pi f = 2\pi \frac{1}{T}$, est le nombre de cycles autour du cercle unité par unité de temps; f est simplement la fréquence, et elle est la réciproque de la période T , c'est-à-dire de la durée nécessaire à une oscillation complète.

l'expression (6.7) est la représentation spectrale d'un processus stationnaire avec un spectre discret, et l'ensemble de nombres $\{\lambda_1, \lambda_2, \dots\}$ est appelé le spectre du processus.

L'expression (6.8) est la représentation spectrale d'un processus stationnaire, où $Z(\lambda)$ est un spectre continu dont la moyenne est nulle, et qui possède des accroissements non corrélés. La possibilité d'une telle représentation de processus stationnaires arbitraires a été démontrée pour la première fois par Kolmogorov (Yaglom, 1962).

La représentation spectrale d'un processus stationnaire est par conséquent une "décomposition" du processus en couples distincts d'oscillations périodiques non corrélées. Il est possible de séparer les composantes spectrales correspondant à des parties distinctes du spectre en utilisant des opérateurs ou filtres linéaires judicieusement choisis(2). Dans la pratique, les filtres utilisés sont de trois sortes: les filtres basses fréquences, qui laissent passer toutes les oscillations dont la fréquence est inférieure à une certaine fréquence critique λ_0 , les filtres hautes fréquences qui laissent passer toutes les oscillations supérieures à λ_0 , et les filtres bandes passantes qui laissent passer les oscillations de fréquences λ sur un intervalle donné (bande passante) $[\lambda_0, \lambda_1]$.

Nous allons montrer maintenant que les informations contenues dans la fonction d'autocovariance sont équivalentes à celles fournies par la transformée de Fourier-Stieljes, c'est-à-dire la fonction de distribution spectrale.

(2) En génie, un filtre est un dispositif qui laisse passer des oscillations harmoniques dans une certaine bande de fréquence (bande passante) tout en supprimant les oscillations ayant des fréquences différentes.

Puisque tout processus stationnaire arbitraire peut avoir une représentation spectrale, sa fonction d'autocovariance correspondante peut être mise également sous forme spectrale.

Ainsi, dans le cas d'un processus tel que (6.7), que l'on suppose être stationnaire et par conséquent, $E(X_k \bar{X}_1) = 0$, $k \neq 1$ (par \bar{X} nous dénotons la conjuguée de X), la fonction d'autocovariance est:

$$(6.9) \quad \sigma(\tau) = \sum_{k=1}^{\infty} E|X_k|^2 e^{i\lambda_k \tau} = \sum_{k=1}^{\infty} b_k e^{i\lambda_k \tau}; \quad b_k > 0.$$

La fonction d'autocovariance (6.9) existe seulement si la série est convergente, c'est-à-dire si

$$(6.10) \quad \sum_{k=1}^{\infty} E|X_k|^2 = \sum_{k=1}^{\infty} b_k < \infty.$$

Slutsky (1938) a également montré que l'inverse est vrai, c'est-à-dire que tout processus stochastique stationnaire ayant une fonction d'autocovariance de la forme (6.9) peut se mettre sous la forme (6.7), avec $E(X_k \bar{X}_1) = 0$ pour $k \neq 1$. En posant $\tau = 0$, l'expression (6.9) devient:

$$(6.11) \quad \sigma(0) = \sum_{k=1}^{\infty} E|X_k|^2 = \sum_{k=1}^{\infty} b_k$$

ce qui montre qu'en superposant des oscillations périodiques non corrélées, la variance totale du processus est égale à la somme des variances des différentes composantes périodiques.

Notons que les variances b_k des composantes périodiques séparées sont les moyennes des carrés de l'amplitude X_k des composantes harmoniques $X_k e^{i\lambda_k \tau}$ du processus X_t . Khinchin (1934) a généralisé la formule (6.9) et il a démontré que la fonction d'autocovariance de tout processus stochastique stationnaire peut se mettre sous la forme d'une intégrale:

$$(6.12) \quad \sigma(\tau) = \int_{-\infty}^{\infty} e^{i\lambda\tau} dG(\lambda)$$

où $G(\lambda)$ est la fonction de distribution spectrale, ou la transformée de Fourier-Stieljes de la fonction d'autocovariance $\sigma(\tau)$. La fonction de distribution spectrale est une fonction non décroissante monotone, symétrique par rapport à l'origine et bornée, $G(-\infty)=0$ and $G(\infty)=\sigma(0)$.

Lorsque $G(\lambda)$ est normalisée, c'est-à-dire divisée par la variance,

$$F(\lambda) = \frac{G(\lambda)}{\sigma(0)} \quad \text{est appelée fonction de distribution spectrale}$$

normalisée, qui est la transformée de Fourier-Stieljes de la fonction d'autocorrélation $\rho(\tau)$. c'est-à-dire

$$(6.13) \quad \rho(\tau) = \int_{-\infty}^{\infty} e^{i\lambda\tau} dF(\lambda).$$

La fonction de distribution normalisée $F(\lambda)$ est également non décroissante, symétrique par rapport à l'origine et bornée, $F(-\infty)=0$ et $F(\infty)=1$. On peut la décomposer sous la forme suivante:

$$(6.14) \quad F(\lambda) = F_1(\lambda) + F_2(\lambda) + F_3(\lambda)$$

où $F_1(\lambda)$, $F_2(\lambda)$ et $F_3(\lambda)$ sont toutes trois non décroissantes, $F_1(\lambda)$ est une fonction escalier pure, $F_2(\lambda)$ est absolument continue, c'est-à-dire que $F_2(\lambda) = \int_{-\infty}^{\lambda} F_2^1(u) du$ et $F_3(\lambda)$ est une fonction singulière continue, avec $F_3^1(\lambda) = 0$ presque partout.

On peut donc considérer $F(\lambda)$ comme une fonction de distribution et $\rho(\tau)$ comme sa fonction caractéristique. Puisque $G(\lambda)$ et $F(\lambda)$ sont des fonctions impaires, pour tout processus faisant intervenir des nombres réels, les expressions (6.12) et (6.13) sont des intégrales de nombres réels et peuvent se mettre sous la forme suivante:

$$(6.15) \quad \sigma(\tau) = \int_{-\infty}^{\infty} \cos \lambda \tau dG(\lambda) = \int_0^{\infty} \cos \lambda \tau dG_1(\lambda)$$

où $G_1(\lambda) = 2G(\lambda)$, et

$$(6.16) \quad \rho(\tau) = \int_{-\infty}^{\infty} \cos \lambda \tau dF(\lambda) = \int_0^{\infty} \cos \lambda \tau dF_1(\lambda)$$

où $F_1(\lambda) = 2F(\lambda)$.

Si $F(\lambda)$ et $G(\lambda)$ ont des dérivées (les cas qui nous intéressent le plus souvent), on a

$$(6.17.a) \quad dF(\lambda) = f(\lambda)d\lambda$$

$$(6.17.b) \quad dG(\lambda) = g(\lambda)d\lambda$$

et $f(\lambda)$ est appelée alors la fonction de densité spectrale normalisée, ou spectre normalisé, et $g(\lambda)$ est la fonction de densité spectrale, ou spectre. Les transformées de Fourier inverses de $g(\lambda)$ et $f(\lambda)$ sont respectivement $\sigma(\tau)$ et $\rho(\tau)$.

Si τ est discret, la fonction de densité spectrale normalisée $f(\lambda)$ est la transformée de Fourier d'une suite d'autocorrélations, et nous obtenons:

$$(6.20) \quad f(\lambda) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} \rho(\tau) e^{-i\lambda\tau}; \quad -\pi \leq \lambda \leq \pi$$

et dans le cas d'un paramètre temporel continu:

$$(6.21) \quad f(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \rho(\tau) e^{-i\lambda\tau} d\tau; \quad -\infty \leq \lambda \leq \infty.$$

De même, le spectre $g(\lambda)$ d'un processus discret sera:

$$(6.22) \quad g(\lambda) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} \rho(\tau) e^{-i\lambda\tau}; \quad -\pi \leq \lambda \leq \pi$$

et pour un processus continu:

$$(6.23) \quad g(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \rho(\tau) e^{-i\lambda\tau} d\tau \quad -\infty \leq \lambda \leq \infty.$$

Puisque $F(\lambda) = \int_{-\infty}^{\lambda} f(u) du$, l'intégration de (6.20) et de (6.21) nous donne la fonction de distribution spectrale normalisée $F(\lambda)$.

Dans le cas d'un processus réel X_t , l'expression (6.20) se ramène à:

$$(6.24) \quad f(\lambda) = \frac{\rho(0)}{2\pi} + \frac{1}{\pi} \sum_{\tau=1}^{\infty} \rho(\tau) \cos \lambda \tau = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} \rho(\tau) \cos \lambda \tau; \quad -\pi \leq \lambda \leq \pi$$

et l'expression (6.21) prend la forme suivante:

$$(6.25) \quad f(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \rho(\tau) \cos \lambda \tau d\tau \quad -\infty \leq \lambda \leq \infty.$$

Les fonctions de distribution spectrale normalisée correspondantes sont alors:

$$(6.26) \quad F(\lambda) = \frac{\rho(0)}{2\pi} + \frac{1}{\pi} \sum_{\tau=1}^{\infty} \frac{\rho(\tau) \sin \lambda \tau}{\tau} \quad -\pi \leq \lambda \leq \pi$$

et

$$(6.27) \quad F(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\lambda} \int_{-\infty}^{\infty} \rho(\tau) \cos \lambda \tau d\tau d\lambda; \quad -\infty \leq \lambda \leq \infty.$$

On utilise une méthode semblable pour calculer $G(\lambda)$.

Dans la section suivante, nous étudierons la fonction d'autocorrélation et la densité spectrale normalisée correspondant aux processus linéaires (5), (5.1), (5.2) et (5.3), et nous illustrerons le tout par quelques exemples théoriques.

7. LES FONCTIONS D'AUTOCOVARIANCE ET D'AUTOCORRÉLATION DES PROCESSUS STATIONNAIRES ET LEURS TRANSFORMÉES DE FOURIER

Nous avons vu à la section (5) qu'un processus stationnaire linéaire X_t peut être interprété comme l'extrait obtenu à partir d'un intrant U_t (processus purement aléatoire) passée par une fonction de transfert $\alpha(L)$, qui est une somme infinie convergente de coefficients de pondération α_k pour $|L| \leq 1$. En d'autres termes:

$$(7.1) \quad X_t = \sum_{k=0}^{\infty} \alpha_k U_{t-k} = \sum_{k=0}^{\infty} \alpha_k L^k U_t = \alpha(L) U_t; \quad \sum_{k=0}^{\infty} |\alpha_k| < M; \quad M \text{ fini.}$$

Dans cette expression, L est l'opérateur de retard de la fonction de transfert. Lorsque l'on considère la fonction de transfert comme une fonction génératrice des coefficients de pondération α_k on tient L pour une variable auxiliaire dans laquelle α_k est le coefficient de la k -ième puissance de L . En appliquant ensuite la formule à la fonction d'autocovariance, et en se rappelant que nous avons un processus stationnaire, nous obtenons:

$$(7.2) \quad \sigma_{XX}(\tau) = E(X_t X_{t+\tau}) = \sigma_U^2 \sum_{k=0}^{\infty} \alpha_k \alpha_{k+\tau}.$$

La variance de X_t est alors:

$$(7.3) \quad \sigma_{XX}(0) = \sigma_U^2 \sum_{k=0}^{\infty} \alpha_k^2.$$

$$(7.4) \quad \rho_{XX}(\tau) = \frac{\sigma_{XX}(\tau)}{\sigma_{XX}(0)} = \frac{\sum_{k=0}^{t-\tau} \alpha_k \alpha_{k+\tau}}{\sum_{k=0}^{t-0} \alpha_k^2}.$$

On peut calculer plus facilement la fonction d'autocovariance en se servant de la fonction génératrice d'autocovariance, que l'on peut également utiliser pour obtenir le spectre du processus.

La fonction génératrice d'autocovariance est:

$$(7.5) \quad \sigma_{XX}(L) = \sum_{\tau=-\infty}^{\infty} \sigma_{XX}(\tau) L^{\tau}, \quad \tau = 0, \pm 1, \pm 2, \dots$$

Etant donné que pour un processus stationnaire $\sigma_{XX}(\tau)$ est une fonction paire, $\sigma_{XX}(k) = \sigma_{XX}(-k)$, est le coefficient de L^k et de L^{-k} .

Dans le cas du processus linéaire (7.1), la fonction génératrice d'autocovariance est:

$$(7.6) \quad \sigma_{XX}(L) = \sigma_U^2 \alpha(L) \alpha(L^{-1}).$$

Si l'on pose $L = e^{-i\lambda}$, l'expression (7.5) prend la forme suivante:

$$(7.7) \quad \sigma_{XX}(L) = \sum_{\tau=-\infty}^{\infty} \sigma_{XX}(\tau) e^{-i\lambda\tau}; \quad \tau = 0, \pm 1, \pm 2, \dots$$

Si nous comparons l'expression (7.7) au spectre $g_{XX}(\lambda)$ de l'expression (6.22), soit $g_{XX}(\lambda) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} \sigma_{XX}(\tau) e^{-i\lambda\tau}$, nous voyons que $2\pi g_{XX}(\lambda) = \sigma_{XX}(L)$. Si nous posons λ non négatif seulement, $\pi g_{XX}(\lambda) = \sigma_{XX}(L)$.

Par conséquent, le produit de la multiplication de la fonction génératrice d'autocovariance par $\frac{1}{\pi}$ nous donne le spectre du processus. En appliquant ensuite l'expression (7.6), le spectre du processus linéaire en (5) peut aussi s'écrire sous la forme:

$$(7.8) \quad g_{XX}(\lambda) = \frac{\sigma_U^2}{\pi} |\alpha(e^{-i\lambda})|^2; \quad -\pi \leq \lambda \leq \pi$$

On appelle $|\alpha(e^{-i\lambda})|^2$ le gain du filtre, qui est égal au carré de la fonction de transfert du filtre. L'expression (7.8) montre que le spectre de l'extrait X_t du processus linéaire (7.1) peut s'obtenir à partir du spectre uniforme constant $\frac{\sigma_U^2}{\pi}$ d'un processus de bruit

blanc U_t multiplié par un facteur (le gain du filtre) qui ne dépend que des caractéristiques du filtre. Les fréquences pour lesquelles $|\alpha(e^{-i\lambda})|^2$ est élevé sont amplifiées, celles pour lesquelles le gain du filtre est faible sont amorties. La variance de X_t est:

$$(7.9) \quad \sigma_U^2 = \int_0^\pi g_{XX}(\lambda) d\lambda = \frac{\sigma_U^2}{\pi} \int_0^\pi |\alpha(e^{-i\lambda})|^2 d\lambda.$$

En divisant (7.8) par (7.9), nous obtenons la fonction de densité spectrale normalisée $f_{XX}(\lambda)$ qui est la transformée de Fourier de la fonction d'autocorrélation $\rho_{XX}(\tau)$. Par conséquent:

$$(7.10) \quad f(\lambda) = \frac{|\alpha(e^{-i\lambda})|^2}{\int_0^\pi |\alpha(e^{-i\lambda})|^2 d\lambda}; \quad 0 \leq \lambda \leq \pi$$

Il est facile de montrer que les fonctions d'autocovariance et d'autocorrélation du processus autorégressif AR(p) dans l'expression (5.1.1) répondent à la même forme d'équation aux différences finies qui correspond à la partie déterminable. En effet, si on multiplie (5.1.1) par $X_{t-\tau}$ et si l'on applique (5.1.3), nous obtenons:

$$(7.11) \quad \sum_{r=0}^p \beta_r X_{t-r} X_{t-\tau} = \sum_{r=0}^{\infty} h_r U_t U_{t-\tau-r}.$$

Etant donné que $E(X_{t-r} X_{t-\tau}) = \sigma_{XX}(\tau-r)$; $E(U_t^2) = \sigma_U^2$; $E(U_t U_s) = 0$ pour tout t, s ; l'espérance mathématique des deux membres de l'expression (7.11)

satisfait, pour $\tau=0$ et pour $\tau>0$, respectivement:

$$(7.12) \quad \sum_{r=0}^p \beta_r \sigma_{XX}(-r) = \sigma_U^2$$

et

$$(7.13) \quad \sum_{r=0}^p \beta_r \sigma_{XX}(\tau-r) = 0 \quad \tau = 1, 2, \dots$$

On appelle souvent ces deux expressions les équations de Yule-Walker. Ainsi, la suite $\sigma_{XX}(\tau-1), \sigma_{XX}(\tau-2), \dots, \sigma_{XX}(\tau-p), \tau = 1, 2, \dots$, résout l'équation aux différences finies homogène (7.13). En divisant (7.13) par σ_X^2 , nous obtenons la fonction d'autocorrélation qui résout aussi une équation aux différences finies homogène analogue à celle du processus X_t lui-même. Nous pouvons écrire l'expression (7.13) en nous servant de l'opérateur de retard L sous la forme suivante:

$$(7.14) \quad \sum_{r=0}^p \beta_r L^r \sigma_{XX}(\tau) = \beta(L) \sigma_{XX}(\tau) = 0; \quad \tau = 1, 2, \dots$$

et, de la même façon, pour la fonction d'autocorrélation:

$$(7.15) \quad \sum_{r=0}^p \beta_r L^r \rho_{XX}(\tau) = \beta(L) \rho_{XX}(\tau) = 0; \quad \begin{array}{l} \tau = 1, 2, \dots \\ \beta_0 = 1 \end{array}$$

dans laquelle L s'applique à τ .

Les mêmes conditions de stationnarité nécessaires aux processus autorégressifs finis de la forme (5.1.1) s'appliquent ici. Si toutes les racines $|G_i| < 1$ sont distinctes, nous avons deux possibilités:

- 1) Une racine G_i est un nombre réel; dans ce cas, $A_i G_i^s$ (voir l'expression 5.1.7) décroît géométriquement en tendant vers zéro à mesure que s augmente. Si la racine est positive, nous aurons une fonction exponentielle décroissante, et si elle est négative, nous aurons une fonction exponentielle dont le signe va alterner et décroissante en valeur absolue.

2) Un couple de racines conjuguées complexes G_1, G_j ; dans ce cas, elles génèrent un terme qui est une fonction trigonométrique oscillante, décroissante en valeur absolue, et dont la période d'oscillation dépend de l'argument des racines complexes.

La variance $\sigma_{XX}(0)$ d'un processus autorégressif X_t peut être obtenue à partir de l'expression (7.12); elle peut être également exprimée en termes de la fonction d'autocorrélation en divisant (7.12) par $\sigma_{XX}(0)$ et en rendant $\sigma_{XX}(-r) = \sigma_{XX}(r)$. Nous avons alors:

$$(7.16) \quad \sigma_X^2 = \frac{\sigma_U^2}{1 + \beta_1 \rho(1) + \beta_2 \rho(2) + \dots + \beta_p \rho(p)}$$

Le spectre $g_{XX}(\lambda)$ pour le processus AR(p) peut se calculer en utilisant l'expression (7.8), où le gain de filtre est $|\beta(e^{-i\lambda})|^{-2}$.

Alors:

$$(7.17) \quad g_{XX}(\lambda) = \frac{\sigma_U^2}{\pi |\beta(e^{-i\lambda})|^2} = \frac{\sigma_U^2}{\pi |1 + \beta_1 e^{-i\lambda} + \dots + \beta_p e^{-ip\lambda}|^2}; \quad 0 \leq \lambda \leq \pi.$$

On obtient la fonction de densité spectrale normalisée $f(\lambda)$ en divisant $g_{XX}(\lambda)$ par σ_X^2 .

Bien que la fonction d'autocorrélation d'un processus AR(p) soit infinie, de par sa nature même elle peut être décrite en termes de p fonctions non nulles des autocorrélations. Cette information est donnée par la fonction d'autocorrélation partielle, qui aide à déterminer l'ordre d'un processus autorégressif pour une série temporelle observée. Dans le cas d'un processus autorégressif d'ordre p, la fonction d'autocorrélation partielle présente un point limite après le retard p.

En dénotant par β_{kj} le j-ième coefficient d'un processus AR(k), de telle façon que β_{kk} est le dernier coefficient, à partir de (7.15) β_{kk} satisfait l'ensemble d'équations:

$$(7.18) \quad \rho(j) = \beta_{k1}\rho(j-1) + \beta_{k2}\rho(j-2) + \dots + \beta_{kk}\rho(j-k); \quad j=1, 2, \dots, k$$

ce qui nous mène aux équations de Yule-Walker, qui peuvent se mettre sous la forme suivante:

$$(7.19) \quad \begin{bmatrix} 1 & \rho & \dots & \rho(k-1) \\ \rho & 1 & \dots & \rho(k-2) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(k-1) & \rho(k-2) & \dots & 1 \end{bmatrix} \begin{bmatrix} \beta_{k1} \\ \beta_{k2} \\ \vdots \\ \beta_{kk} \end{bmatrix} = \begin{bmatrix} \rho(1) \\ \rho(2) \\ \vdots \\ \rho(k) \end{bmatrix}$$

En résolvant (7.19) pour $k=1, 2, 3, \dots$, nous trouvons les auto-corrélations partielles $\beta_{11}, \beta_{22}, \beta_{33}, \dots$. De façon générale, β_{kk} est la fonction d'autocorrélation partielle du retard k.

Dans le processus de moyenne mobile finie d'ordre q (MM-q) en (5.2.1), la fonction d'autocovariance est:

$$(7.20) \quad \sigma_{XX}(\tau) = E(X_t X_{t+\tau}) = \sigma^2 \sum_{k=0}^{q-\tau} \alpha_k \alpha_{k+\tau}; \quad \tau \leq q$$

et

$$\sigma_{XX}(\tau) = 0; \quad \tau > q.$$

La variance du processus MM(q) sera alors:

$$(7.21) \quad \sigma_X^2 = \sigma_U^2 \sum_{k=0}^q \alpha_k^2$$

et la fonction d'autocorrélation:

$$(7.22) \quad \rho_{XX}(\tau) = \frac{\sum_{k=0}^{q-\tau} \alpha_k \alpha_{k+\tau}}{\sum_{k=0}^q \alpha_k^2}; \quad \tau \leq q$$

et

$$\rho_{XX}(\tau) = 0; \quad \tau > q.$$

Par conséquent, le corrélogramme d'un processus MM(q) est nul à partir de $\tau=q+1$. Suivant l'expression (4.7.8), le spectre d'un tel MA(q) processus est:

$$(7.23) \quad g_{XX}(\lambda) = \frac{\sigma_U^2}{\pi} |\alpha(e^{-i\lambda})|^2 = \frac{\sigma_U^2}{\pi} \left| \sum_{k=0}^q \alpha_k e^{-i\lambda k} \right|^2; \quad \alpha_0=1; \quad 0 \leq \lambda \leq \pi$$

et la fonction de densité spectrale normalisée $f_{XX}(\lambda)$ est $g_{XX}(\lambda)/\sigma_X^2$.

Dans le cas du processus A.R.M.M. (p q) X_t défini par l'expression (4.5.1), on peut calculer de la même façon, $\sigma_{XX}(\tau)$, $\rho_{XX}(\tau)$, $g_{XX}(\lambda)$ et $f_{XX}(\lambda)$. Ainsi:

$$(7.24) \quad \sigma_{XX}(\tau) = \sigma_{XU}(\tau) + \alpha_1 \sigma_{XU}(\tau-1) + \dots + \alpha_q \sigma_{XU}(\tau-q) - \beta_1 \sigma_{XX}(\tau-1) - \dots - \beta_p \sigma_{XX}(\tau-p)$$

où $\sigma_{XU}(\tau)$ est la fonction de covariance croisée entre X_t et U_t , définie par $\sigma_{XU}(\tau) = E(X_{t-\tau} U_t)$.

Etant donné que $X_{t-\tau}$ ne dépend que des composantes aléatoires qui se sont présentées jusqu'à l'instant $t-\tau$ et qui ne sont pas corrélées avec U_t , il s'ensuit que $\sigma_{XU}(\tau) = 0$ pour $\tau > 0$ et $\sigma_{XU}(\tau) \neq 0$, pour $\tau \leq 0$.

L'expression (7.24) se ramène à :

$$(7.25) \quad \sigma_{XX}(\tau) = -\beta_1 \sigma_{XX}(\tau-1) - \dots - \beta_p \sigma_{XX}(\tau-p); \quad \tau \geq q+1.$$

La fonction d'autocorrélation est donc :

$$(7.26) \quad \rho_{XX}(\tau) = -\beta_1 \rho_{XX}(\tau-1) - \dots - \beta_p \rho_{XX}(\tau-p), \quad \tau \geq q+1$$

ou encore :

$$\beta(L) \rho_{XX}(\tau) = 0; \quad \tau \geq q+1$$

dans laquelle L s'applique à τ .

Par conséquent, dans le cas d'un processus A.R.M.M. (p, q) , il y aura q autocorrélations dont les valeurs dépendront du choix des q paramètres de moyennes mobiles α , ainsi que des p paramètres autorégressifs β . Or, pour $\tau > q+1$, les p autocorrélations déjà obtenues donnent les valeurs initiales de l'équation aux différences finies homogène $\beta(L) \rho_{XX}(\tau) = 0$ qui détermine alors entièrement les autocorrélations aux retards plus élevés.

Si $q < p$, toute la fonction d'autocorrélation sera composée d'un mélange de fonctions exponentielles et(ou) sinusoidales amorties. Si $q \geq p$, les $q-p+1$ autocorrélations utilisées comme valeurs initiales ne suivront pas cette tendance générale.

Pour $\tau=0$, l'expression (7.24) donne la variance du processus:

$$(7.27) \quad \sigma_X^2 = \sigma_U^2 + \alpha_1 \sigma_{XU}(-1) + \dots + \alpha_q \sigma_{XU}(-q) - \beta_1 \sigma_{XX}(1) - \dots - \beta_p \sigma_{XX}(p)$$

qu'il faut résoudre ainsi que les p équations (7.24) pour $\tau=1, 2, \dots, p$ pour obtenir $\sigma_X^2, \sigma_{XX}(1), \dots, \sigma_{XX}(p)$.

Le spectre du processus est:

$$(7.28) \quad g_{XX}(\lambda) = \frac{\sigma_U^2}{\pi} \frac{\left| \sum_{r=0}^q \alpha_r e^{-i\lambda r} \right|^2}{\left| \sum_{k=0}^p \beta_k e^{-i\lambda k} \right|^2}; \quad 0 \leq \lambda \leq \pi$$

et la fonction de densité spectrale normalisée $f_{XX}(\lambda)$ est le quotient de $g_{XX}(\lambda)$ and σ_X^2 .

Nous allons maintenant illustrer par quelques exemples théoriques adaptés de A.A. Sveshnikov (1966) le lien qui existe entre la fonction d'autocorrélation et la fonction de densité spectrale normalisée.

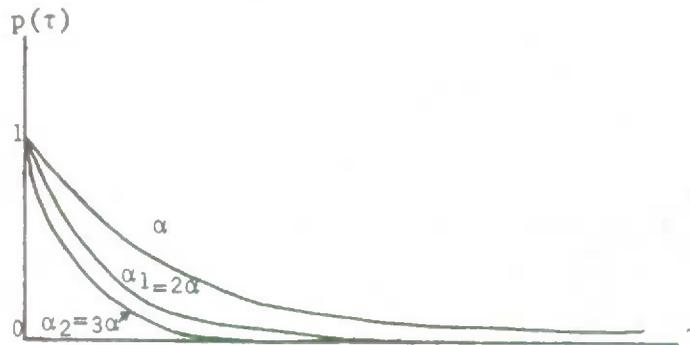
Exemple 1

Soit une fonction d'autocorrélation de la forme

$$1) \quad \rho(\tau) = e^{-\alpha |\tau|}$$

où $0 < \alpha < 1$ sert à mesurer la rapidité avec laquelle $\rho(\tau)$ décroît à mesure que le retard temporel τ augmente. Nous voyons à la figure 1 que plus α , est grand, plus la fonction d'autocorrélation est amortie, ce qui implique un processus stochastique moins lisse.

Figure 1



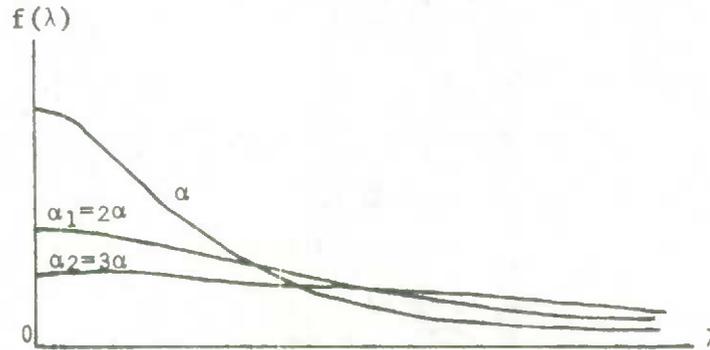
La fonction de densité spectrale normalisée correspondante $f(\lambda)$ est:

$$2) f(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \rho(\tau) e^{-i\lambda\tau} d\tau = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\lambda\tau - \alpha|\tau|} d\tau = \frac{1}{\pi} \frac{\alpha}{\lambda^2 + \alpha^2} .$$

Elle donne les mêmes informations que la fonction en 1), mais dans le domaine des fréquences. La figure 2 présente la fonction de densité spectrale normalisée. Nous pouvons voir que, pour α petit, le spectre normalisé a une prédominance de basses fréquences, ce qui signifie α que le processus est lisse, tandis qu'à mesure que λ croît, la courbe se tasse vers l'axe λ tout en devenant plus plate. Ce type de comportement de la fonction $f(\lambda)$ nous permet d'illustrer le comportement d'un processus purement aléatoire ou bruit blanc dont la fonction de densité spectrale normalisée est une constante égale à $\frac{1}{2\pi}$, pour $-\pi \leq \lambda \leq \pi$. Nous observons que l'ordonnée de $f(\lambda)$ à l'origine est $\frac{1}{\pi\alpha}$, et à mesure que α croît, l'intersection diminue. En fait,

on suppose que α peut prendre des valeurs très élevées, et la fonction $\rho(\tau)$ se transforme en une fonction en dents de scie, qui ne diffère de zéro que dans un très petit intervalle autour de $\tau=0$.

Figure 2



La variance totale du processus est l'aire sous $f(\lambda)$. Pour α , petit, une bande de basse fréquence représente la plus grande partie de la variance, tandis que pour α grand, la variance est distribuée presque uniformément dans la bande de fréquence capable d'influer sur le processus considéré. Un bruit absolument blanc ne peut exister, car pour que la densité spectrale demeure constante sur tout l'intervalle de variation de λ , il faudrait que la fonction d'autocovariance pour $\tau=0$, ait une variance infinie, ce qui est impossible dans n'importe quel processus à nombres réels. En effet:

$$\sigma_{XX}(0) = \int_{-\infty}^{\infty} g(\lambda) d\lambda = c \int_{-\infty}^{\infty} d\lambda = \infty .$$

Exemple 2

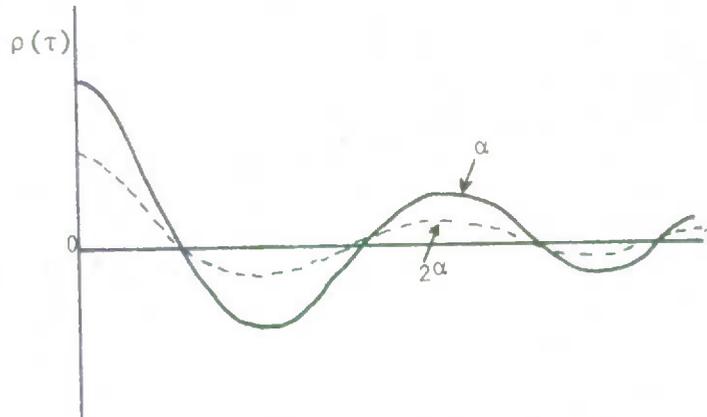
Soit maintenant une fonction d'autocorrélation $\rho(\tau)$ de la forme:

$$3) \rho(\tau) = e^{-\alpha |\tau|} \cos \omega \tau$$

qui diffère de (1) par la présence du facteur $\cos \omega\tau$ et qui donne à la $\rho(\tau)$ fonction la forme d'une oscillation harmonique amortie, comme on peut le voir à la figure 3.

Si nous observions le processus générateur, on remarquerait une certaine périodicité.

Figure 3



La fonction de densité spectrale normalisée correspondante s'obtient en remplaçant $\cos \omega\tau$ par $\frac{1}{2}(e^{i\omega\tau} + e^{-i\omega\tau})$ et λ par $(\lambda - \omega)$ et $(\lambda + \omega)$ dans les intégrales de $f(\lambda)$. D'où:

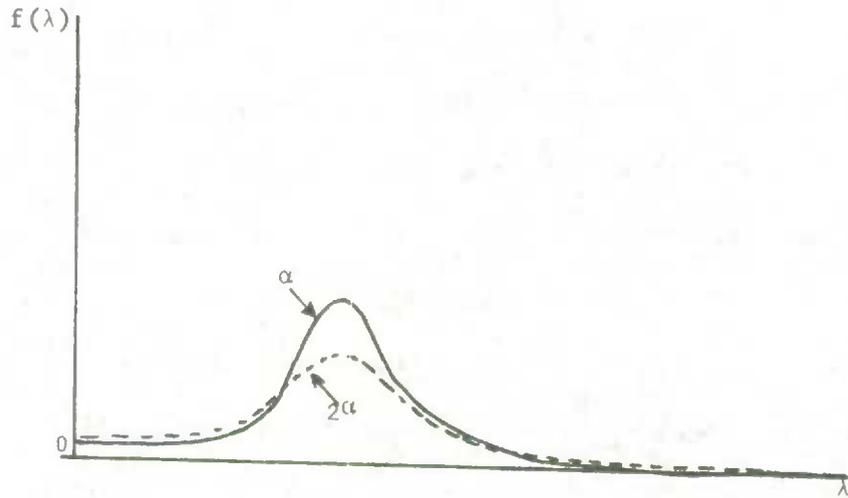
$$4) f(\lambda) = \frac{1}{2\pi} \left| \frac{\alpha}{(\lambda - \omega)^2 + \alpha^2} + \frac{\alpha}{(\lambda + \omega)^2 + \alpha^2} \right| = \frac{\alpha}{\pi} \left| \frac{\lambda^2 + \alpha^2 + \omega^2}{(\lambda^2 - \omega^2 - \alpha^2)^2 + 4\alpha^2 \lambda^2} \right|.$$

La représentation graphique (Fig. 4) de la fonction $f(\lambda)$ se caractérise par des sommets dans le voisinage de la fréquence angulaire ω . Pour

$\omega = \frac{\pi}{3}$, la fréquence saisonnière fondamentale sur l'intervalle temporel correspondrait à une période de 6 mois, pour $\omega = \frac{\pi}{6}$, à une période de 12 mois. Le modèle correspondant pour le processus générateur aurait une composante saisonnière oscillatoire dont l'amplitude et la phase changeraient lentement par rapport à la période saisonnière fondamentale. Plus α est grand, c'est-à-dire plus le taux de variation

est élevé, plus la composante saisonnière est obscurcie, et moins aigu est par conséquent le sommet de la fonction de densité spectrale normalisée.

Figure 4



8. ERGODICITE

Les applications de la théorie de la probabilité portent d'ordinaire sur des évènements qui se répètent de nombreuses fois. Par conséquent, comme moyenne d'une variable aléatoire X décrivant un évènement nous pouvons prendre la moyenne arithmétique de toutes les valeurs observées X_j de X . De même, pour déterminer la moyenne et la fonction d'autocovariance $\sigma(t, s)$ d'un processus stochastique $X(t)$, nous devons disposer d'un grand nombre de réalisations du processus $X(t)$, c'est-à-dire $X_1(t), \dots, X_N(t)$, et calculer ensuite la moyenne pour chaque valeur t et la fonction d'autocovariance pour chaque couple des valeurs t et s . Mais dans la pratique, les données de nombreuses séries temporelles, et en particulier les séries temporelles économiques, proviennent d'une seule expérience. Dans ce cas, il est nécessaire de préciser les conditions auxquelles une seule réalisation du processus permet de calculer des estimateurs cohérents pour toutes les caractéristiques de la distribution du processus. Cette dernière opération est possible parce que le théorème ergodique (ou loi des grands nombres) s'applique à une classe de processus aléatoires stationnaires.

En vertu de ce théorème, l'espérance mathématique de $X(t)$ et de $X(t)X(s)$ calculée en prenant la moyenne des quantités correspondantes sur tout l'espace échantillon Ω (appelé la moyenne de l'ensemble, ou parfois la moyenne spatiale) peut être remplacée par les moyennes temporelles des mêmes quantités.

Etant donné un processus stochastique stationnaire $X(t)$, la moyenne temporelle du moment de premier ordre de $X(t)$ est définie par:

$$(8.1) \hat{\mu}_T = \frac{1}{T} \int_0^T X(t) dt; \quad 1 \leq t \leq T.$$

Quant T tend vers $T \rightarrow \infty$, la moyenne temporelle $\hat{\mu}_T$ converge en moyenne quadratique (aussi en probabilité vers la moyenne de l'ensemble μ (espérance mathématique) si et seulement si:

$$(8.2) \quad \lim_{T \rightarrow \infty} E(\hat{\mu}_T - \mu)^2 = 0$$

puisque

$$(8.3) \quad \lim_{T \rightarrow \infty} E(\hat{\mu}_T - \mu)^2 = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \sigma(\tau) d\tau; \quad \tau = s-t$$

alors l'expression (8.2) sera vérifiée si et seulement si

$$(8.4) \quad \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \sigma(\tau) d\tau = 0.$$

C'est Slutsky (1938) qui le premier a démontré cette condition. Tout processus stationnaire qui remplit la condition (8.4) est dit ergodique dans la moyenne. Pour l'ergodicité dans les moments de second ordre, nous avons besoin de la moyenne temporelle de la fonction d'autocovariance, soit:

$$(8.5) \quad \hat{\sigma}_T(\tau) = \frac{1}{T} \int_0^T \{ [X(t+\tau) - \mu] [X(t) - \mu] \} dt$$

qui doit converger en moyenne quadratique vers la moyenne d'ensemble de la fonction d'autocovariance $\sigma(\tau)$. En d'autres termes:

$$(8.6) \quad \lim_{T \rightarrow \infty} E[\hat{\sigma}_T(\tau) - \sigma(\tau)]^2 = 0.$$

Dans le cas d'un processus normal, l'expression (8.6) se ramène à:

$$(8.7) \quad \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T |\sigma(\tau)|^2 d\tau = 0.$$

L'existence des moyennes temporelles et de leurs propriétés de convergence a été démontrée dans le célèbre théorème ergodique de Birkhoff et Khinchin (Genedenko, 1966).

Il est facile de montrer que tous les processus stochastiques stationnaires linéaires sont ergodiques. Cependant, tous les processus stationnaires ne sont pas ergodiques. Prenons par exemple le cas d'un processus harmonique simple, dans lequel a et b sont des variables aléatoires suivant une loi normale avec des moyennes égales à zéro et une même variance σ^2 .

$$(8.8) \quad X(t) = a \cos \lambda t + b \sin \lambda t.$$

La fonction d'autocovariance $\sigma(\tau)$ est:

$$(8.9) \quad \sigma(\tau) = \sigma^2 \cos \lambda \tau.$$

En utilisant l'expression (8.4), nous pouvons montrer que ce processus est ergodique dans la moyenne. Nous avons:

$$(8.10) \quad \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \sigma(\tau) d\tau = \lim_{T \rightarrow \infty} \sigma^2 \frac{\sin \lambda T}{T} = 0.$$

Mais le processus n'est pas ergodique dans la fonction d'autocovariance.

En appliquant (8.7), nous obtenons:

$$(8.11) \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T |\sigma(\tau)|^2 d\tau = \lim_{T \rightarrow \infty} \left(\frac{\sigma^4}{2} + \sigma^4 \frac{\sin 2\lambda T}{T} \right) = \frac{\sigma^4}{2} .$$

9. PROCESSUS NON STATIONNAIRES LINÉAIRES HOMOGÈNES

9.1 Modèles autorégressifs et de moyenne mobile intégrée (A.R.M.M.I.)

Le modèle A.R.M.M. (p, q) examiné en (5.3) peut être généralisé en incluant des processus qui sont non stationnaires au niveau local et(ou) au niveau de la pente. En d'autres termes, nous allons considérer le cas où l'opérateur linéaire qui transforme une entrée U_t purement aléatoire en une sortie X_t , est non stationnaire homogène. En supposant que X_t suit un processus stationnaire dans la d-ième différence et que U_t suit un processus de moyenne d'ordre q, nous pouvons donc écrire:

$$(9.1.1) \quad \phi(L) X_t = \alpha(L)U_t$$

où $\phi(L)$ est un opérateur autorégressif non stationnaire et $\alpha(L)$ un opérateur de moyenne mobile inversible.

Étant donné que le processus X_t est stationnaire dans la d-ième différence d, des racines de $\phi(L) = 0$ sont égales à 1, tandis que les autres, disons p, se trouvent à l'intérieur du cercle unité. On peut donc mettre l'opérateur $\phi(L)$ sous la forme suivante:

$$(9.1.2) \quad \phi(L) = \beta(L)(1-L)^d = \beta_p(L)\Delta^d$$

où $1-L = \Delta$ est l'opérateur de différence, les sous-indices indiquant l'ordre des opérateurs, à savoir p+d pour l'opérateur autorégressif généralisé $\phi(L)$ et p pour l'opérateur autorégressif stationnaire $\beta(L)$.

En supposant que l'ordre de l'opérateur de moyenne mobile est q , l'expression (9.1.1) devient:

$$(9.1.3) \quad \phi_{p+d}(L)X_t = \beta_p(L)\Delta^d X_t = \alpha_q(L)U_t$$

et elle s'appelle processus autorégressif de moyenne mobile intégrée - A.R.M.M.I. (p, d, q) . Notons que pour $d=0$, l'expression (9.1.3) définit un processus A.R.M.M. (p, q) et qu'en posant $\Delta^d X_t = W_t$, la même expression est stationnaire en W_t , ou, ce qui revient au même, stationnaire dans la d -ième différence de X_t .

On peut mettre le modèle A.R.M.M.I. (p, d, q) sous la forme explicite d'une équation aux différences finies comme suit:

$$(9.1.4) \quad X_t = U_t + \alpha_1 U_{t-1} + \dots + \alpha_q U_{t-q} - \phi_1 X_{t-1} - \dots - \phi_{p+d} X_{t-p-d};$$

$$\alpha_0 = \phi_0 = 1.$$

On se sert généralement d'une telle équation pour le calcul des prévisions.

Étant donné que dans l'expression (9.1.3), $\beta_p(L)$ et $\alpha_q(L)$ sont des opérateurs dont l'un est stationnaire et l'autre inversible, il est possible de mettre le modèle sous d'autres formes, à savoir: 1) en fonction des valeurs actuelles et précédentes du processus aléatoire U_t , et 2) en fonction des valeurs précédentes de X_t , plus la valeur présente du U_t . Dans le premier cas, nous avons:

$$(9.1.5) \quad X_t = \phi_{p+d}^{-1}(L)\alpha_q(L)U_t = \psi(L)U_t$$

où

$$(9.1.6) \quad \psi(L) = 1 + \psi_1 L + \psi_2 L^2 + \dots$$

On peut calculer les coefficients de pondération ψ en égalant les coefficients de L dans l'expression développée de $\phi(L)\psi(L)U_t = \alpha_q(L)U_t$.

En d'autres termes:

$$(9.1.7) \quad (1 + \phi_1 L + \dots + \phi_{p+d} L^{p+d}) (1 + \psi_1 L + \psi_2 L^2 + \dots) = 1 + \alpha_1 L + \alpha_2 L^2 + \dots + \alpha_q L^q.$$

Pour j supérieur à la plus grande valeur de $p+d-1$ ou q , les coefficients de pondération ψ satisfont l'équation aux différences finies suivante:

$$(9.1.8) \quad \phi_{p+d}(L)\psi_j = \beta_p(L)\Delta^d \psi_j = 0$$

dans laquelle L s'applique à l'indice j .

Dans le deuxième cas, où X_t est fonction des valeurs antérieures de X_t , plus la valeur présente de U_t , nous avons:

$$(9.1.9) \quad \phi_{p+d}(L)\alpha_q^{-1}(L)X_t = \Pi(L)X_t = U_t$$

où

$$(9.1.10) \quad \Pi(L) = 1 + \Pi_1 L + \Pi_2 L^2 + \dots$$

Étant donné que $\alpha_q(L)$ est un opérateur inversible, $\Pi(L)$ est convergent sur le cercle unité, ou à l'intérieur de celui-ci.

Pour calculer les coefficients de pondération Π , nous procédons de la même façon que dans le cas des coefficients de pondération ψ .
 Nous égalons les coefficients de L dans

$$(9.1.11) \quad \phi_{p+d}(L) = \alpha_q(L) \Pi(L)$$

et pour j supérieur à la plus grande valeur de p+d-1 ou q, les coefficients Π satisfont l'équation aux différences finies:

$$(9.1.12) \quad \alpha_q(L) \Pi_j = 0$$

dans laquelle L s'applique à j.

Enfin, il est possible de modifier le processus A.R.M.M.I. (p, d, q) dans (9.1.1) en ajoutant une fonction déterminable du temps f(t), habituellement un polynôme d'un degré égal à l'ordre de l'opérateur de différence, qui représente une tendance générale déterminable.

Il peut également arriver qu'il y ait davantage de bruit dans le processus, en ce sens que nous n'observons pas en fait X_t , mais un autre processus aléatoire $Z_t = X_t + U_t$, dans lequel U_t peut être un processus purement aléatoire ou un processus aléatoire corrélé. Par conséquent, si le processus observé est Z_t

$$(9.1.13) \quad \beta_p(L) \Delta^d Z_t = \beta_p(L) \Delta^d (X_t + U_t^1) = \alpha_q(L) U_t + \beta_p(L) \Delta^d U_t^1.$$

Or, si nous supposons que U_t^1 est un processus A.R.M.M. (p_1, q_1) , c'est-à-dire si:

$$\beta_{p_1}(L)U_t^1 = \alpha_{q_1}(L)E_t$$

où E_t est un processus purement aléatoire indépendant de U_t^1 , l'expression (9.1.13) devient:

$$(9.1.14) \quad \beta_{p_1}(L)\beta_p(L)\Delta^d Z_t = \beta_{p_1}(L)\alpha_q(L)U_t + \beta_p(L)\alpha_{q_1}(L)\Delta^d E_t.$$

Soit $P = p_1 + p$ et posons Q égal à la valeur la plus élevée de $p_1 + q$ ou de $p + q_1 + d$, l'expression (9.1.14) peut alors se mettre sous la forme:

$$(9.1.15) \quad \beta_P(L)\Delta^d Z_t = \alpha_Q(L)V_t$$

dans laquelle V_t est un processus purement aléatoire et Z_t un processus A.R.M.M.I. (P, d, Q) , puisque la somme des deux processus de moyenne mobile indépendants donne un autre processus de moyenne mobile dont l'ordre est celui du processus dont l'ordre est plus élevé.

9.2 Modèles de moyenne mobile intégrée (M.M.I.)

Pour $p = 0$, le processus A.R.M.M.I. (p, d, q) défini par l'expression (9.1.3) se ramène à un processus de moyenne mobile intégrée d'ordre $(0, d, q)$, c'est-à-dire que le processus stochastique X_t est considéré comme étant le résultat de d applications de l'opérateur de somme infinie à un processus de moyenne mobile finie d'ordre q . En représentation symbolique:

$$(9.2.1) \quad \Delta^d X_t = \alpha_q(L)U_t$$

ou encore:

$$(9.2.2) \quad X_t = \Delta^{-d} \alpha_q(L) U_t = S^d \alpha_q(L) U_t$$

dans laquelle $S = (1-L)^{-1} = \Delta^{-1}$ est l'opérateur de sommation infinie et l'indice en puissance d indique que le processus (9.2.1) est sommé d fois. (1)

On peut mettre le modèle (9.2.1) sous la forme de l'équation aux différences finies suivante:

$$(9.2.3) \quad X_t = U_t + \alpha_1 U_{t-1} + \dots + \alpha_q U_{t-q} + dX_{t-1} - \frac{1}{2}d(d-1)X_{t-2} + \dots + (-1)^{d+1}X_{t-d}.$$

On peut également l'expliciter mathématiquement en fonction de la valeur présente de U_t et de ses valeurs précédentes ou en fonction des valeurs précédentes de X_t , plus la valeur présente de U_t .

(1) À noter que

$$(1) \quad S U_t = \sum_{h=-\infty}^t U_h = (1+L+L^2+\dots) U_t = (1-L)^{-1} U_t$$

$$(2) \quad S^2 U_t = S(S U_t) = \sum_{i=-\infty}^t \sum_{h=-\infty}^i U_h = S U_t + S U_{t-1} + \dots$$

et ainsi de suite.

10. MODÈLES SAISONNIERS

L'analyse temporelle des phénomènes économiques a distingué depuis longtemps divers types d'évolution, soit a) la tendance, b) le cycle, c) le mouvement saisonnier et d) les mouvements irréguliers.

La tendance correspond à une variation lente s'effectuant dans un sens déterminé et qui se maintient pendant de longues années. Le cycle est un mouvement d'allure quasi périodique comportant une phase croissante et une phase décroissante. Le mouvement saisonnier correspond à des variations s'effectuant régulièrement pendant l'année. Il tient aux saisons et aux éléments de type institutionnel et technologique. Ces trois mouvements sont d'un comportement systématique dans le temps. Ils sont les signaux du processus stochastique.

Par contre, les mouvements irréguliers sont des mouvements erratiques, à fréquences élevées, présentant une allure générale plus ou moins stable. Ils résultent des influences que toutes sortes d'événements exercent sur la variable en cause.

Dans la plupart des travaux d'analyse, la tendance et le cycle sont traités ensemble car c'est la série sans variations saisonnières sur laquelle se concentre notre intérêt. L'information donnée par une série temporelle corrigée des variations saisonnières joue un grand rôle dans la prise des décisions à moyen terme et à long terme. Il est alors très important d'identifier les modèles saisonniers. Dans cette section, nous examinerons les modèles les plus couramment utilisés pour la description des mouvements saisonniers.

L'hypothèse la plus simple et la plus étudiée est celle selon laquelle la composante saisonnière a un comportement stable, c'est-à-dire qu'elle est une fonction strictment périodique. On peut alors représenter la composante saisonnière s_t dans le cas de données mensuelles par une suite de douze constantes, donc une par mois, dont la somme est égale à zéro. Désignons par a_k ces constantes. Nous avons donc:

$$(10.1) \quad s_t = \begin{cases} a_k; & \text{pour } t=k \text{ ou } t-k \text{ divisible par } 12 \\ 0; & \text{autrement} \end{cases}$$
$$\sum_{k=1}^{12} a_k = 0$$

l'expression (10.1) peut aussi être représentée avec des fréquences au lieu de périodes:

$$(10.2) \quad s_t = \sum_{j=1}^6 (\alpha_j \cos \lambda_j t + \beta_j \sin \lambda_j t); \quad \lambda_j = \frac{2\pi j}{12}$$

où les λ_j sont les fréquences saisonnières d'une série mensuelle et pour laquelle l'on n'observe pas des composantes d'une périodicité inférieure à 2 mois. Cependant, on peut facilement adapter tous les résultats à des séries hebdomadaires ou autres.

A partir des expressions (10.1) et (10.2), nous obtenons la relation suivante:

$$(10.3) \quad a_k = \sum_{j=1}^6 (\alpha_j \cos \lambda_j k + \beta_j \sin \lambda_j k); \quad k=1, \dots, 12.$$

Bien que s_t soit par hypothèse une fonction strictement périodique, elle n'est pas une fonction déterminable tant que l'on considérera α_j et β_j comme des variables purement aléatoires dont la moyenne est zéro et

$$E(\alpha_j \alpha_k) = E(\beta_j \beta_k) = \begin{cases} \sigma_j^2 & ; \\ 0 & ; \end{cases} \begin{matrix} j=k \\ j \neq k \end{matrix}$$

et $E(\alpha_j \beta_k) = 0$ pour tout j et tout k .

L'expression (10.3) sera alors un modèle saisonnier stable qui suit une loi de processus stationnaire dans la covariance, mais qui ne sera pas ergodique aux moments de second ordre si les variables aléatoires sont par hypothèse distribuées suivant une loi normale (voir la section 8).

La représentation spectrale d'un processus s_t aura un sommet de hauteur σ_j^2 à chaque λ_j . Nous trouverons une concentration de masse spectrale précisément aux niveau des fréquences saisonnières. Bien que cela ne se vérifie jamais dans la pratique, on a une bonne approximation si les sommets saisonniers du spectre sont étroits.

On a tenté plusieurs fois de produire des modèles saisonniers qui soient un peu plus réalistes. En raison de changements institutionnels, technologiques et autres changements de nature exogène dans une structure économique, les mouvements saisonniers des séries temporelles économiques tendent à varier dans le temps. Dans ce cas, la meilleure marche à suivre dépendra du type de comportement suivi par l'évolution du phénomène. Diverses possibilités ont été examinées, notamment les suivantes:

- 1) Il y a un changement subit qui partage la série temporelle observée en deux parties assez uniformes. Dans ce cas, il faut manifestement estimer un modèle saisonnier pour chacune de ces deux parties. O. Lange (1963) et M. Abel (1962) en donnent des exemples.
- 2) L'amplitude peut varier, alors que la phase demeure la même. En d'autres termes, les coefficients α_j et β_j de l'expression (10.3) dépendent de t . Le cas le plus simple consiste à supposer qu'il s'agit de polynômes de degré peu élevé. En cas de linéarité, le s_t est:

$$(10.4) \quad s_t = \sum_{j=1}^6 \{(\alpha_{1j} + \alpha_{2j}t) \cos \lambda_j t + (\beta_{1j} + \beta_{2j}t) \sin \lambda_j t\}.$$

Ces types de modèles ont été étudiés par Hannan (1963), Nettheim (1965) et N.M. Rosenblatt (1963).

On peut aussi définir le processus stochastique qui génère α_j et β_j . Hannan, Terrel et Tuckwell (1970) et Terrel et Tuchwell (1970) se sont livrés à des travaux empiriques en s'appuyant sur l'hypothèse selon laquelle α_j et β_j suivent un modèle autorégressif stationnaire d'ordre 1. Grether et Nerlove (1970) ont envisagé un angle d'approche semblable.

Il n'y a cependant pas d'objection, sauf du point de vue des méthodes d'estimation, à supposer que α_j et β_j sont décrits par d'autres types de processus.

(3) Le modèle saisonnier peut aussi varier pour ce qui est de la phase; dans ce cas, non seulement les amplitudes, mais aussi la phase, doivent être considérées comme générées par un processus stochastique. L'expression (10.2) mise maintenant sous la forme d'une fonction cosinus devient:

$$(10.5) \quad s_t = \sum_{j=1}^6 R_{jt} (\cos \lambda_j t + \theta)$$

où $R_{jt} = \sqrt{\alpha_{jt}^2 + \beta_{jt}^2}$ est l'amplitude aléatoire du processus
 et $\theta = \arctan \frac{\alpha_{jt}}{\beta_{jt}}$ est la phase, et aussi une variable aléatoire.

Pour la prévision de séries temporelles dans lesquelles pourrait se trouver un comportement saisonnier évolutif, Box et Jenkins (1970) ont proposé un autre type de modèle appelé modèle saisonnier multiplicatif. Ce dernier appartient à la catégorie des processus non stationnaires linéaires homogènes, et il repose sur l'hypothèse selon laquelle, dans le cas de données périodiques, on s'attend à des relations: 1) entre les observations mensuelles successives au cours d'une année, et 2) entre des observations portant sur le même mois au cours d'années successives. Nous avons ainsi deux périodes, l'une pour laquelle le retard L est d'une période, et l'autre pour laquelle la période du retard L est donnée par la saisonnalité. En supposant que la variation saisonnière est de périodicité \underline{s} , nous aurons alors $L^s X_t = X_{t-s}$. En considérant les observations pour chaque mois au cours d'années successives, on propose un modèle A.R.M.M.I. (P,D,Q) pour les décrire, c'est-à-dire:

$$(10.6) \quad B_p(L^s) \Delta_s^D X_t = A_q(L^s) E_t$$

dans lequel les sous-indices P, D et Q dénotent respectivement l'ordre de l'opérateur autorégressif stationnaire $B(L^S)$, de l'opérateur de différence $\Delta_S = 1-L^S$ et de l'opérateur de moyenne mobile inversible $A(L^S)$. E_t représente un processus aléatoire autocorrélé. En d'autres termes, on s'attend à ce que l'observation pour chaque mois soit aussi reliée aux mois précédents, et, par conséquent, que E_t soit relié à E_{t-1} , E_{t-2} et ainsi de suite. Afin de prendre en compte cette relation, on introduit un second modèle A.R.M.M.I. (p, d, q):

$$(10.7) \quad B_p(L)\Delta^d E_t = \alpha_q(L)U_t$$

où U_t est maintenant un processus purement aléatoire et $\beta_p(L)$ est un opérateur stationnaire autorégressif d'ordre p, $\Delta = 1-L$ est l'opérateur de différence d'ordre d, et $\alpha_q(L)$ est une moyenne mobile inversible d'ordre q. En combinant (10.6) et (10.7), on obtient:

$$(10.8) \quad \beta_p(L)B_p(L^S)\Delta^d \Delta_S^D X_t = \alpha_q(L)A_q(L^S)U_t$$

qui est le modèle saisonnier multiplicatif général d'ordre (p, d, q) $(P, D, Q)_S$.

L'expression (10.8) est utilisée pour la prédiction des séries temporelles lorsqu'il peut s'y trouver une composante saisonnière évolutive. K.R. Brewer (1969) a utilisé les fonctions de prédiction, ou bases de prédiction, correspondant à différentes identifications de (10.8) et il les a partagées en tendance et en mouvement saisonnier pour désaisonnalisation. Cette idée est très intéressante

et bien élaborée d'un point de vue théorique, mais elle ne peut être appliquée dans la pratique au niveau de la production. De sérieuses limitations émanent à la fois de la capacité des modèles de Box et Jenkins de décrire convenablement divers types de séries temporelles économiques et de la méthode d'estimation pour la décomposition des fonctions de prédiction.

11. CONCLUSIONS

Les modèles pour séries temporelles sont des processus stochastiques, c'est-à-dire des processus qui suivent des lois probabilistes. D'un point de vue théorique, il est possible de construire une grande classe de modèles pour décrire le comportement des phénomènes qui changent dans le temps d'une façon probabiliste. Cependant, pour des applications empiriques, on n'a élaboré qu'une sous-classe restreinte de modèles. Ces modèles appartiennent aux catégories suivantes: 1) processus non stationnaires (dans la moyenne) normaux, 2) processus stationnaires linéaires normaux et 3) processus non stationnaires linéaires homogènes normaux.

Parfois on n'introduit pas explicitement dans un modèle l'hypothèse de normalité, auquel cas la stationnarité est définie lato sensu.

La construction de modèles de séries temporelles dépend du comportement du phénomène et du but de l'analyse, mais pour ce qui est des applications empiriques immédiates, elle dépend aussi de l'existence de méthodes optimales d'estimation.

Si l'analyse vise à décrire la structure probabiliste d'une série temporelle, des modèles paramétriques (nombre fini de paramètres, habituellement peu élevé) et non paramétriques (nombre infini, ou très grand, de paramètres) sont également acceptables.

Parmi les modèles paramétriques, le modèle d'erreurs qui est un processus normal non stationnaire (dans la moyenne), et le modèle autorégressif et de moyenne mobile intégrée, qui est un processus normal non stationnaire linéaire homogène, sont le plus fréquemment utilisés.

Les modèles non paramétriques les plus communs sont la fonction d'autocovariance et le spectre. Ils ne se révèlent cependant pas utiles dans les problèmes de prévision, qui ne peuvent être résolus que par des modèles paramétriques.

Dans un modèle d'erreurs, le processus générateur d'une série temporelle est décomposé en une composante systématique (signal) et un processus purement aléatoire (bruit blanc). Les propriétés de la série temporelle sont ensuite résumées dans l'espérance mathématique du processus, et elles ne sont pas affectées par les variations de la composante aléatoire. En d'autres termes, tout effet temporel sur le processus se répercute sur la partie déterminable, ce qui fait que chaque observation est indépendante stochastiquement des précédentes.

Ce modèle est le modèle classique pour l'analyse des séries temporelles économiques, où l'on suppose que le signal a une composante de longue durée, la tendance (composante systématique se déplaçant lentement, doucement et progressivement dans le temps), un mouvement quasi-périodique, le cycle, et une composante saisonnière (composante oscillatoire qui tend à se répéter avec une certaine régularité chaque année).

Cependant, l'hypothèse d'une composante sous la forme d'une erreur non autocorrélée restreint sérieusement la validité de ce type de modèle quand il s'agit de décrire le comportement des séries économiques. Les erreurs que présentent les valeurs observées des séries économiques sont dues non seulement à des observations indépendantes erronées, mais aussi à d'autres écarts plus sérieux et, dès leur apparition, elles s'incorporent au processus et influent sur son évolution future. Yule (1921, 1927) a apparemment été le premier à signaler cet aspect de la question qui semble avoir été presque entièrement laissé de côté dans l'analyse des séries économiques jusqu'à ces derniers temps.

La décomposition du signal en une composante de tendance-cycle et une composante saisonnière a une importance considérable pour celui qui prend des décisions économiques, car devant faire face au problème du contrôle du niveau de l'activité économique, il ne veut pas confondre un mouvement saisonnier et un mouvement de longue durée.

Un modèle qui répond à ces désirs et qui en même temps introduit des hypothèses plus réalistes sur le comportement de la composante aléatoire serait celui qui verrait la série temporelle décomposée en un signal (tendance-cycle et saisonnalité) et en une erreur, qui est autocorrélée. Bien que d'un point de vue théorique cela soit possible, il n'existe aucune méthode d'estimation suffisamment développée pour les séries temporelles qui suivent ce modèle.

Des programmes informatiques pour l'estimation de modèles ayant des résidus autorégressifs basés sur la méthode des moindres carrés généralisés ou sur le maximum de vraisemblance ont été mis au point au cours des dernières années, mais ils ne peuvent être utilisés que lorsque la tendance et le cycle suivent un polynôme de degré relativement peu élevé du temps, et lorsque les facteurs saisonniers sont aussi des fonctions polynômiales du temps sur tout l'intervalle temporel. Malheureusement, ces conditions se vérifient très rarement dans le cas des séries temporelles économiques.

Par contre, le modèle autorégressif et de moyenne mobile intégrée suppose que le mécanisme générateur d'une série temporelle a des erreurs autocorrélées, et qu'il est non stationnaire au niveau local et(ou) au niveau de la pente. Par conséquent, en différenciant convenablement le processus, on obtient un processus stationnaire dans les différences d'un ordre fini. On suppose que ce processus stationnaire est l'extrait obtenu à partir de l'intrant qu'est le bruit blanc auquel on a fait subir une transformation linéaire. On généralise ce modèle en incluant un opérateur saisonnier de manière multiplicative ou additive.

On identifie les paramètres de ce type de modèle en analysant la fonction d'autocorrélation. Toutefois, il a été démontré que l'analyse de la fonction d'autocorrélation est efficace avant tout pour les modèles qui correspondent à des processus stochastique suivant une loi normale ou quasi normale avec de hautes fréquences ou de petits retards, c'est-à-dire des modèles d'effets à court terme quasi gaussiens (Mandelbrot, 1972).

De plus, les estimations finales de ces modèles donnent le total pour le signal et elles ne partagent pas ce total en composantes correspondant au mouvement tendance-cycle et au mouvement saisonnier, de sorte qu'elles sont utiles pour d'autres utilisations que celles du modèle d'erreurs. On les utilise principalement pour la production d'estimations finales des données brutes et la prévision. On peut également les employer pour estimer et prévoir des séries déjà désaisonnalisées, mais il faut alors faire très attention lors de l'identification et de l'estimation, puisque la correction des données initiales peut introduire des distorsions qui contredisent les hypothèses de base de ces types de modèles.

Cela montre que pour des applications empiriques, et notamment des applications économiques, les types de modèles de séries temporelles pour lesquels existent déjà des méthodes optimales d'estimation sont très limités, et qu'ils sont loin d'être les meilleurs. Il faudra encore beaucoup de recherches théoriques et empiriques pour que l'on puisse parvenir à une compréhension profonde des problèmes que pose l'analyse des séries temporelles économiques.

REFERENCES

Ouvrages

- (1) Anderson, T.W. (1971) The Statistical Analysis of Time Series, Wiley, New York
- (2) Bartlett, M.S. (1955) Stochastic Processes, Cambridge University Press
- (3) Bensoussan, A. (1971) Filtrage Optimal des systèmes linéaires, Dunod, Paris
- (4) Box, G. and G. Jenkins (1970) Time Series Analysis Forecasting and Control, Holden Day, California
- (5) Bucy, R.S. and P.D. Joseph (1968) Filtering for Stochastic Processes with Applications to Guidance, Interscience, Wiley, N.Y.
- (6) Cramer, H. and M.R. Leadbetter (1967) Stationary and Related Stochastic Processes, Wiley, N.Y.
- (7) Doob, J.L. (1953) Stochastic Processes, Wiley, N.Y.
- (8) Dynkin, E.B. (1965) Theory of Markov Processes, Pergamon Press
- (9) Fishman, G. (1969) Spectral Methods in Econometrics, Harvard University Press
- (10) Genedenko, B. (1966) The Theory of Probability, Chelsea Pub. Co. 3rd ed.
- (11) Gikhman, I. and A.V. Skorokhod (1965) Introduction to the Theory of Random Processes, Nanka Moscow (en russe)
- (12) Gikhman, I. and A.V. Skorokhod (1968) Stochastic Differential Equations, Kiev (en russe)
- (13) Granger, C.W. and M. Hatanaka, (1964) Spectral Analysis of Economic Time Series, Princeton University Press
- (14) Grenander, V. and M. Rosenblatt (1957) Statistical Analysis of Stationary Time Series, Wiley, N.Y.
- (15) Hannan, E.J. (1960) Time Series Analysis, Methuen and Co., London
- (16) Hannan, E.J. (1970) Multiple Time Series, Wiley, N.Y.
- (17) Harris, B. (ed) (1967) Spectral Analysis of Time Series, Wiley, N.Y.

Ouvrages- fin

- (18) Jazwinski, A. (1970) Stochastic Processes and Filtering Theory, Academic Press
- (19) Jenkins, G. and D. Watts (1969) Spectral Analysis and Its Applications, Holden Day, California
- (20) Kolmogorov, A. (1933) Foundations of the theory of probability (en russe) traduit en anglais par Chelsea Pub. Co. Londres 3^e éd. (1956)
- (21) Kushner, H.J. (1971) Introduction to Stochastic Control, Holt, Rinehart and Winston
- (22) Lagrange, (1772-78) Oevres, Vol. 6
- (23) Lange, O.C. (1963) Introduction to Econometrics, MacMillan
- (24) Lions, J.L. (1969) Quelques Méthodes de Résolution des Problèmes aux Limites Non-Linéaires, Dunod, Paris
- (25) Moore, H.L. (1914) Economic Cycles: Their Law and Causes, MacMillan Co.
- (26) Pannekoek, A. (1961) A History of Astronomy, Interscience, New York.
- (28) Pugachev, U.S. (1966) Theory of Random Functions and its Applications to Control Problems, Pergamon Press
- (29) Skorokhod, A.V. (1965) Studies in the Theory of Random Processes, Addison Wesley
- (30) Stuart, R.D. (1961) An Introduction to Fourier Analysis, Methuen, London
- (31) Sveshnikov, A.A. (1966) Applied Methods of the Theory of Random Functions, Pergamon Press
- (32) Wainstein, L.A. and V.O. Zubakov (1961) Extraction of Signals from Noise, Prentice Hall
- (33) Wiener, N. (1949) Extrapolation, Interpolation and Smoothing of Stationary Time Series, Wiley N.Y.
- (34) Wold, H. (1938) A Study in the Analysis of Stationary Time Series. Almqvist and Wiksell, Stochholm, 2nd ed. (1954)
- (35) Yaglom, A.M. (1962) Theory of Stationary Random Functions, Prentice Hall

Articles

- (1) Abel, M.E. (1962) "Harmonic Analysis of Seasonal Variation with an Application to Hog Production" J.A.S.A., 57, p. 655.
- (2) Beveridge, W.H. (1922) "Wheat Prices and Rainfall in Western Europe" J. Roy. Stat. Soc. 85, p. 412.
- (3) Brewer, K.R.W. (1969) "Box-Jenkins Models For Prediction and Seasonal Adjustment", Commonwealth, Bureau of Census and Statistics, Canberra.
- (4) Buys-Ballot, G. (1847) "Les changements périodique de température" Utrecht.
- (5) Dagum, C. and E.B. Dagum (1972) "The Meaning of Theory and Model in Social Sciences", Working Paper Series No. 72-9, The University of Iowa, 39 pages.
- (6) Grether, D.M. and M. Nerlove (1970) "Some properties of Optimal Seasonal Adjustment", Econometrica, V. 38 - N. 5 p. 682.
- (7) Hannan, E.J. (1960) "The Estimation of Seasonal Variation" the Australian Jour. of Stat. II p. 1.
- (8) Hannan, E.J. (1963) "The Estimation of Seasonal Variation in Economic Time Series" J.A.S.A. LVIII - p. 31.
- (9) Hannan, E.J. (1964) "The Estimation of a changing Seasonal Pattern" J.A.S.A. LIX p.p. 1063-77.
- (10) Hannan, E.J., R.D. Terrell and N.E. Tuckwell (1970) "The Seasonal Adjustment of Economic Time Series" Intern. Ec. Review p. 24.
- (11) Kalman, R.E. (1960) "A New Approach to Linear Filtering and Prediction Problems" Jour. Basic. Engineering Series D82-35 p. 35.
- (12) Kalman, R.E. and R.S. Bucy (1961) "New Results in Linear Filtering and Prediction Theory" Jour. Basic. Engineering, Series D83-V. 35 p. 95.
- (13) Khinchin, A.Y. (1934) "Korrelations Theorie der Stationarem Stochastischen Prozesse" Math. Ann. 109, p. 604.
- (14) Khinchin, A.Y. (1938) "The Theory of Damped Spontaneous Effects" Izv. Akad. Nauk. SSR, Ser. 3 p. 313.
- (15) Kolgomorov, A.N. (1939) "Sur l'interpolation et l'extrapolation des suites stationnaires" Compt. tend. acad. sci. Paris, v. 208 p. 2043.

Articles-fin

- (16) Kolgomorov, A.N. (1941) "The Local Structure of Turbulence in Incompressible Viscous Fluid for Very Large Reynolds numbers, Dokl, Akad, Nauk. SSR, Vol. 30, p. 301 - reproduit dans: Turbulence: Classic Papers on Statistical Theory, publié par S.K. Friedlander et L. Topping, Interscience Pub. N.Y. (1961).
- (17) Mandelbrot, B. (1972) "Statistical Methodology for Non-periodic Cycles from the Covariance to R/S Analysis, Annals of Ec. and Soc. Measurement, NBER, V. 1 No. 3, p. 259.
- (18) Nerlove, M. (1964) "Spectral Analysis of Seasonal Adjustment Procedures" *Econometrica*, 32, No. 3, p. 241.
- (19) Nerlove, M. (1965) "A Comparison of a Modified Hannan and the BLS Seasonal Adjustment Filters" *J.A.S.A.* 60, No. 310, p. 442.
- (20) Nettheim, N.F. (1965) "Fourier Methods for Evolving Seasonal Patterns" *J.A.S.A. Co. N.* 310, p. 492.
- (21) Parzen, R. (1967) "The Role of Spectral Analysis in Time Series Analysis" *Rev. of the I.S.I.*, 35, No. 2.
- (22) Rosenblatt, H.M. (1963) "Spectral Analysis and Parametric Methods for Seasonal Adjustment of Economic Time Series" *Proc. Bus. Ec. Stat. Section of A.S.A.*, p. 94.
- (23) Schuster, A. (1898) "On the Investigation of Hidden Periodicities with Application to a Supposed Twenty-six Day Period of Meteorological Phenomena", *Terrestrial Magnetism*, Vol. 3, p. 13.
- (24) Stokes, G.C. (1879) "Note on Searching for Periodicities" *Proc. Roy. Soc.* V. 29, p. 122.
- (25) Slutsky, E.E. (1938) "Sur les fonctions aléatoires presque périodiques et sur la décomposition de fonctions aléatoires stationnaires en composantes" *Actualités Scient. et Ind.* 738, Paris.
- (26) Terrel, R.D. and N.E. Tuckwell (1971) "Application of an Evolving Seasonal Pattern to an Australian Economic Time Series" *Australian Economic Papers*, Vol. 9, p. 186.
- (27) Yule, V. (1921) "On the Time Correlation Problem" *J. Roy. Stat. Soc.* Vol. 84, p. 497.
- (28) Yule, V. (1927) "On a Method of Investigating Periodicities in Disturbed Series" *Trans. Roy. Soc. (A)* p. 226.
- (29) Zadeh, L. and J. Ragazzini (1950) "An Extension of Wiener's Theory of Prediction" *J. of Applied Physics*, Vol. 21, p. 645.
- (30) Yaglom, A.M. (1955) "The Correlation Theory of Processes whose n th difference constitute a Stationary Process" *Matem. Sb.* 37(79) p. 141.

Statistics Canada Library
Bibliothèque Statistique Canada



1010013734

