

CHRISTIAN GOURIÉROUX

Quelques développements récents en séries temporelles

Journal de la société statistique de Paris, tome 131, n° 1 (1990), p. 7-15

http://www.numdam.org/item?id=JSFS_1990__131_1_7_0

© Société de statistique de Paris, 1990, tous droits réservés.

L'accès aux archives de la revue « Journal de la société statistique de Paris » (<http://publications-sfds.math.cnrs.fr/index.php/J-SFDs>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/legal.php>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

QUELQUES DÉVELOPPEMENTS RÉCENTS EN SÉRIES TEMPORELLES

Christian GOURIÉROUX¹

*Professeur à l'ENSAE, Directeur du Centre de Recherche en Statistique
(CRÉST) de l'ENSAE, CEPREMAP*

I. INTRODUCTION

Il est évidemment délicat de faire une présentation synthétique des développements récents en séries temporelles, le domaine est en effet assez vaste, la littérature abondante mais surtout il faut choisir une date charnière à partir de laquelle nous discuterons les développements et les classerons selon quelques grandes orientations.

Nous restreignant pour cerner le sujet aux seules applications économiques et financières, on peut considérer le milieu des années soixante-dix comme une période clef. Avant cette date les séries temporelles sont en effet essentiellement utilisées dans un contexte linéaire et dans des buts descriptifs, ceux-ci étant d'ailleurs assez variés. Ils peuvent concerner les problèmes de prévision [Box-Jenkins (1970)], de filtrage [Kalman (1963)], les méthodologies pour corriger des variations saisonnières [logiciel CENSUS X11, Shiskin, Young, Musgrave (1965), Dagum (1979)], les calculs de multiplicateurs et les recherches de maquettes dans la modélisation macroéconométrique dynamique [Deleau-Malgrange (1978)], l'utilisation de schémas d'anticipation simples comme le schéma adaptatif [Arrow (1958), Muth (1961)]. Les approches suivies pour résoudre ces problèmes apparaissent souvent ad hoc et d'application assez automatique. On recherche des méthodes simples d'emploi permettant des ajustements satisfaisants, ceci sans se préoccuper du caractère explicatif des approches utilisées. La publication du livre de Box-Jenkins (1970), où les auteurs utilisent de façon cohérente dans une optique de prévision les modèles autorégressifs-moyennes mobiles (connus sous l'abréviation ARMA : Autoregressive-Moving Average), étudiés précédemment par Doob (1952), va jouer un rôle important. On va en effet s'apercevoir que les principales approches de désaisonnalisation, de filtrage, de modélisation d'anticipation... ont des interprétations directes dans le cadre de ces modèles et aboutir ainsi à une certaine unification des méthodologies [voir Azencott-Dacunha Castelle (1984), Gouriéroux-Montfort (1985) pour des présentations en français des modèles ARMA].

1. C. Gouriéroux a été lauréat du prix du statisticien d'expression française décerné par les anciens présidents des Sociétés de Statistique de Paris et de France au titre de 1988.

Ces modèles ARMA, au moins dans la façon dont ils sont alors appliqués, apparaissent cependant restrictifs dans deux de leurs aspects. Ils reposent d'abord sur une idée de linéarité, la valeur présente d'une variable étant exprimée comme fonction linéaire de ses valeurs retardées. Ceci les rend inadaptés pour décrire des phénomènes *non linéaires* comme certains types de cycles, comme la présence de variabilité instantanée dépendant de valeurs passées de la série ou comme l'existence de phases descendantes plus lentes que les phases ascendantes [phénomène de friction à la baisse par exemple]. Par ailleurs et bien que divers résultats théoriques aussi bien probabilistes que statistiques soient déjà connus pour le cas *multivarié* [voir *e.g* Hannan (1970)], ces modèles sont essentiellement appliqués aux séries prises une à une. L'écriture de modèles ARMA multidimensionnels conduit en effet à une rapide inflation du nombre de paramètres à estimer, inflation qui ne peut être contournée qu'en introduisant des contraintes sur ces paramètres. La question difficile à résoudre est alors de savoir quelles contraintes imposer et comment les interpréter.

Ces deux généralisations potentielles : aspect multivarié et non linéarité, se sont développées très rapidement en liaison avec deux types d'applications : la construction et l'analyse des modèles macroéconométriques dynamiques, l'étude des séries financières et monétaires : taux d'intérêt, taux de change, cours d'actions ou d'obligations... L'existence de ces applications spécifiques a naturellement conduit à rechercher des modélisations adaptées et à structurer beaucoup plus les modèles de séries temporelles. En plus de l'obtention de modèles et de résultats mathématiques nouveaux, aspects techniques qui ne seront pas discutés, les principales conséquences sont une *nouvelle réflexion sur des concepts de base* tels les notions de tendance, d'exogénéité, de chocs, de relation de long terme, de volatilité... et l'introduction de *méthodologies cohérentes*, plus descriptives, de *construction des modèles macroéconométriques*.

II. RETOUR SUR LES CONCEPTS DE BASE

a) *Exogénéité et multiplicateurs*

Diverses terminologies sont passées dans le langage usuel des économistes, mais un même terme est souvent employé dans des sens très différents. Un exemple frappant concerne l'exogénéité, pour laquelle on trouve au moins quatre types de définitions. Celles-ci sont les suivantes :

- i) Une variable est dite exogène, si elle peut être fixée à un niveau arbitraire. C'est la notion désormais connue sous le nom de *variable de contrôle*, et qui ne peut être testée à partir des seules évolutions de cette variable et de séries qui lui sont liées.
- ii) Une autre idée d'exogénéité consiste à dire que la série (X_t) est exogène vis à vis de la série (Y_t) dite endogène, si les valeurs de X sont fixées avant

celle de Y . Il s'agit clairement d'une notion différente de la précédente, puisqu'elle n'a de sens que dans une analyse jointe de plusieurs séries.

iii) Une série (X_t) peut aussi être exogène vis-à-vis d'une série (Y_t) , si elle a une évolution autonome, c'est-à-dire non dépendante de l'évolution spécifique de Y . C'est la notion désormais connue sous le nom de *non causalité de Y vers X* .

iv) Finalement une quatrième notion est celle habituellement utilisée dans le cadre du modèle linéaire $Y = Xb + u$. La variable X est exogène, si l'erreur $u = Y - Xb$ est non corrélée avec X . C'est la notion qui assure l'absence de biais de la méthode des moindres carrés ordinaires. Cette définition a un statut très différent des précédentes, puisqu'elle est non seulement relative à plusieurs séries X, Y , mais aussi à un paramètre b ou si l'on préfère à une relation $Y - Xb = 0$. Cette idée a conduit à introduire ce qui est désormais appelé *exogénéité forte* d'une variable vis-à-vis d'un paramètre.

Distinguer ces diverses notions d'exogénéité apparaît comme une nécessité pour la construction et l'analyse des modèles macroéconométriques. Ainsi pour analyser les effets multiplicateurs sur une série Y des chocs effectués sur une série exogène X , nous devons séparer le cas d'une série X contrôlable (définition i)), de celui d'une série non contrôlable d'évolution autonome (définition iii)). Dans le premier cas correspondant au mode de calcul classique des multiplicateurs [Deleau-Malgrange (1978)] on peut faire un choc déterministe sur la valeur présente de la série X en conservant les mêmes valeurs aux composantes passées et futures de cette série (choc transitoire), puis examiner comment se répercute ce choc sur les valeurs présente et futures de l'autre série.

Lorsque la série X n'est pas contrôlable, on peut raisonnablement penser qu'un choc sur la valeur présente de X va avoir un impact sur les valeurs futures de cette même série. Retenant une évolution de la série exogène de type ARMA, on montre que la valeur X_t peut s'écrire comme fonction linéaire des valeurs présente et passées d'un bruit blanc e (suite de variables centrées, non corrélées, de même variance) :

$$X_t = e_t + a_1 e_{t-1} + a_2 e_{t-2} + \dots + a_j e_{t-j} + \dots$$

La variable e_t s'interprète comme la « partie de X_t » spécifique de la date t , appelée *innovation*. La valeur X_t se fixe alors en fonction des divers chocs $e_t, e_{t-1} \dots e_{t-1} \dots$ imposés par l'« environnement » aux dates présente et passées. Le choc relatif à la date t repérable par e_t aura alors un impact sur $X_t, X_{t+1}, X_{t+2} \dots$ et des impacts sur les valeurs présente et passées de Y , ceux-ci dépendant non seulement de la dépendance dynamique de Y et X , mais aussi de la dynamique autonome de X . Des calculs de multiplicateurs fondés sur les innovations des séries ont été initialement introduits par Sims (1980), puis généralisés sous le nom de *mesures de causalité* dans [Geweke (1982), Gouriéroux-Monfort-Renault (1987), Gouriéroux-Monfort (1990)]. Ces mesures peuvent être employées comme d'habitude dans un but descriptif,

mais peuvent aussi servir de base à des procédures de test. Ce sont de telles procédures qui permettent la mise en évidence de formes récursives ou blocs récursives de modèles macroéconométriques [Granger (1980), Geweke (1982)...].

b) *Co-intégration*

Les modèles macroéconométriques existants ont en grande partie été construits à partir de la constatation que certaines séries étaient fortement corrélées, c'est-à-dire que certaines relations statiques entre séries se révélaient approximativement satisfaites. Certaines des relations exhibées ont reçu des interprétations comme équilibres de long terme, d'autres n'ont pu être expliquées de cette façon et ont parfois été dénommées relations historiques.

La théorie de la co-intégration introduite par Granger (1986) et Engle-Granger (1987) a fourni un cadre cohérent pour chercher de telles relations entre séries comportant des tendances explosives. L'idée simple est de rechercher les combinaisons linéaires de ces séries pour lesquelles le phénomène d'explosion disparaît. Ces combinaisons, lorsqu'elles existent, sont appelées *relations de co-intégration* et en pratique on vérifie que la démarche proposée permet de retrouver les « relations dites de long terme » classiques des modèles macroéconométriques.

En fait, ces relations de co-intégration ont, plus qu'une interprétation de type long terme, une interprétation en terme de chocs. Ce sont les directions pour lesquelles les chocs sur les séries ont le moins d'impact, donc d'une certaine façon les combinaisons linéaires les plus stables. Cette dernière interprétation a permis d'étendre cette idée de combinaisons stables au cas de séries non explosives [voir Box-Tiao (1977), Gouriéroux-Peaucelle (1989)].

c) *Tendances*

A l'heure actuelle la notion de tendance (comme celle de composante saisonnière d'ailleurs) n'est toujours pas définie de façon précise. Cependant on sait que certaines séries ont un comportement explosif en fonction du temps et que généralement quitte à transformer la série brute par logarithme, on peut se ramener à des vitesses d'explosion de type monôme t^a , $a > 0$. Plus qu'une définition précise de la notion de tendance, la question importante est de savoir comment modéliser de tels comportements explosifs. Deux modélisations ont depuis longtemps été utilisées et ont souvent dans le passé été considérées comme à peu près équivalentes. La première, dite *tendance déterministe*, consiste à écrire la série X_t comme somme d'une fonction polynômiale déterministe du temps et d'un terme aléatoire stationnaire; la seconde, dite *tendance aléatoire* consiste à modéliser l'évolution par un modèle autorégressif admettant une racine unité pour le polynôme de retard. L'exemple le plus classique de cette seconde approche étant la marche aléatoire. Les travaux récents

ont permis de séparer les résultats classiques sur les séries avec tendance en deux catégories, ceux qui sont robustes à la modélisation déterministe ou aléatoire de la tendance et ceux qui ne le sont pas. Dans la première catégorie on trouve typiquement les relations de co-intégration, le fait d'estimer facilement celles-ci en utilisant les régressions statiques, ..., dans la seconde un très grand nombre de tests usuels de racine unité, dont la réponse se révèle très sensible à la modélisation retenue [Philips-Durlauf (1986), Sims-Stock-Watson (1986), Gouriéroux-Maurel-Monfort (1987)].

d) Caractéristiques instantanées et caractéristiques historiques

Les modèles de valorisation d'actifs utilisés en finance, qu'ils soient fondés sur des idées d'arbitrage ou de détermination de portefeuilles optimaux, font jouer un rôle important à diverses caractéristiques conditionnelles à l'information disponible : c'est le cas de la notion de volatilité, c'est-à-dire de variabilité instantanée des rendements, de la notion de « bêtas », c'est-à-dire de liens entre rendement des actifs et rendement d'un portefeuille de marché... Bien que ces diverses notions soient conditionnelles, elles sont en pratique estimées sous période par sous période comme si les observations étaient indépendantes de même loi. Ceci conduit à confondre des notions conditionnelles (instantanées) avec des notions marginales (historiques). La nécessité de distinguer ces deux types de notions a conduit à introduire des modèles de séries temporelles non linéaires, où la variance conditionnelle aux observations passées dépend effectivement de ces observations. Ces modèles appelés Autorégressifs Conditionnellement Hétéroscélastiques (ARCH) ont été initialement décrits par Engle (1982), Bollerslev (1987).

Cette nouvelle modélisation permet de mieux comprendre l'existence de queues de distribution épaisses pour les lois des rendements (effet leptokurtique), d'évaluer l'importance des biais d'interprétation lorsqu'une caractéristique instantanée était abusivement remplacée par une caractéristique historique. Elle a aussi montré que certaines hypothèses de type efficience de marché qui paraissaient ne pouvoir être rejetées dans un contexte statique l'était beaucoup plus souvent lorsque l'hétéroscélasticité conditionnelle était prise en compte.

III. MODÉLISATION DU GÉNÉRAL SPÉCIFIQUE

a) Deux approches de construction de modèles

On peut de façon simplifiée considérer qu'il y a deux approches extrêmes de construction de modèles, en particulier de modèles macroéconométriques dynamiques, appelées ascendante ou explicative, descendante ou descriptive ou « du général au spécifique ».

Dans la première approche, on part d'un modèle restreint, c'est-à-dire comportant un petit nombre de variables, d'équations et très structuré. Ce modèle de départ est typiquement déduit de la théorie macroéconomique. Cet aspect restreint va souvent entraîner de mauvais ajustements aux données disponibles de sorte qu'il faudra étendre le modèle initial, en ajoutant des variables, des équations, en relâchant des contraintes sur les paramètres, en introduisant de la dynamique... . A chaque étape de l'approche ascendante, on effectuera des tests pour savoir dans quelle(s) direction(s) étendre prioritairement le modèle. Ces tests sont en pratique faits à partir des résidus en utilisant des procédures du type multiplicateur de Lagrange. On obtient ainsi une suite de modèles emboîtés, de moins en moins contraints, la démarche s'arrêtant lorsque la qualité d'ajustement est jugée suffisante, sans perte importante de l'intérêt des interprétations.

A l'inverse, la démarche descendante prend pour base un modèle très large et très peu contraint, qui dans un cadre dynamique linéaire est généralement une formulation vectorielle autorégressive à plusieurs variables (modèle VAR). On cherche alors à chaque étape à contraindre le modèle afin de le rendre plus structuré. Les tests sont habituellement menés en utilisant des procédures de Wald. On obtient une suite de modèles emboîtés de plus en plus contraints. La démarche est arrêtée lorsque la structure mise en évidence permet des interprétations intéressantes, sans perte importante au niveau de la qualité de l'ajustement.

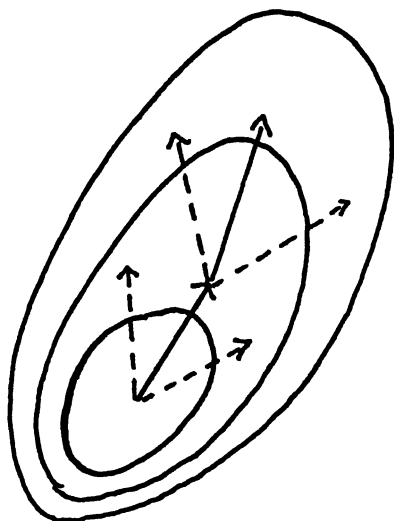
En fait dans la pratique ces démarches ne sont pas envisagées indépendamment l'une de l'autre. Ainsi après une construction ascendante, on recherche souvent une version simplifiée du modèle (maquette) afin de la comparer au modèle de base. Cette recherche de maquette est souvent menée par une approche descendante. De façon similaire il est clair que la recherche des contraintes dans l'approche descendante n'est pas menée de façon quelconque, mais souvent en s'aidant des contraintes sous-jacentes aux théories macroéconomiques pertinentes.

Ces deux démarches sont visualisées graphiquement ci-dessous, représentant chaque modèle par un « patateïde », les phases de tests par les ensembles de flèches, les tests conduisant à l'acceptation par des flèches en trait plein, au refus par des flèches en traits pointillés.

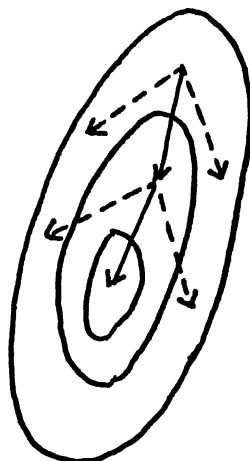
b) Un exemple d'approche du général au spécifique

L'utilisation des résultats récents concernant les déterminations de relations, les calculs de multiplicateurs,... peut permettre d'envisager des procédures cohérentes d'approches descendantes de construction de modèles. Un exemple de telle approche cohérente pourrait être le suivant en se restreignant pour simplifier à un cadre linéaire.

- i) On commencerait par estimer une formulation autorégressive à partir des séries brutes éventuellement non stationnaires.



Approche ascendante
ou structurelle



Approche descendante
ou descriptive

ii) On utiliserait alors la théorie de la co-intégration pour chercher les relations stables entre ces séries et on examinerait les interprétations éventuelles des relations ainsi mises en évidence.

iii) La démarche serait alors poursuivie en utilisant le théorème de représentation d'Engle-Granger (1987). Celui-ci établit que, si les séries initiales sont co-intégrées, le modèle dynamique peut être réécrit sous une forme équivalente où apparaissent clairement les relations stables et les ajustements autour de ses relations, *forme dite à correction d'erreurs* ou E.C.M. [voir Davidson-Hendry-Srba-Yeo (1978)].

iv) On peut alors passer à l'examen des variables d'ajustements, voir s'il existe des relations « stables » entre ces variables désormais stationnaires, déterminer quelles sont pour ces relations les variables pouvant être considérées comme exogènes, celles pouvant être considérées comme endogènes, voir si certains effets dynamiques ne pourraient être regroupés et interprétés par exemple par la présence d'anticipation... .

v) Une phase ultérieure devrait être la comparaison des structures mises en évidence sur les composantes non stationnaires des séries et sur les composantes d'ajustement.

Il est clair qu'une telle démarche ambitieuse n'a pas encore été menée de façon complète sur des exemples concrets. Cependant certaines étapes ont déjà été appliquées : recherche de relations stables entre monnaie et revenu,

entre taux d'inflation et taux de change, analyse de causalités et de récursivité sur les prix de diverses matières premières, sur les variables de la boucle prix-salaires... .

c) Recherche de facteurs

La démarche précédente repose fondamentalement sur la recherche de relations ou de façon équivalente de direction de « stabilité » et de « direction d'ajustement ». Il s'agit essentiellement d'une analyse de corrélation canonique prenant explicitement en compte l'aspect dynamique du problème.

Cette méthodologie de recherche de facteurs n'a pas d'intérêt que pour la modélisation macroéconomique, mais se révèle aussi utile pour l'autre domaine d'application aux séries financières. Ce sont de telles méthodes qui paraissent pouvoir être employées pour déterminer quelques évolutions de base (ou facteurs), dont dépendent principalement les rendements des divers actifs. Ce dernier problème n'a pour l'instant été que partiellement abordé dans la littérature [voir Engle-Ng-Rothschild (1989), Nerlove-Diebold-Van Beeck-Cheung (1988)]. La détermination des facteurs est en effet rendue plus difficile par la nécessité de prendre en compte l'hétéroscédasticité des données et vraisemblablement de distinguer des facteurs relatifs à la moyenne (rendement moyen anticipé) et des facteurs relatifs à la variance (volatilité).

IV. CONCLUSION

La présentation qui vient d'être faite est évidemment par nature courte et partielle. Elle permet cependant, je l'espère, de voir clairement que l'accent de cette recherche est mis sur une meilleure adéquation des modèles de séries temporelles avec les problèmes posés et sur une utilisation jointe d'approches descriptives fondées sur les données et d'approches explicatives fondées sur les théories économiques.

RÉFÉRENCES

- ARROW K.J. (1959) "Towards a Theory of Price Adjustment", dans Abramowitz, M. et al. : *The Allocation of Economic Resources*, Stanford University Press, 49-51.
- AZENCOTT R. et D. DACUNHA-CASTELLE (1984) *Séries d'observations Irrégulières*, Masson.
- BOLLERSLEV T. (1987) "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics*, 31, 309-328.
- BOX G. et G. JENKINS (1970) *Time Series Analysis : Forecasting and Control*, Holden Day.
- BOX G. et G. TIAO (1977) "A Canonical Analysis of Multiple Time Series", *Biometrika*, 64, 355-365.

