

MODÈLE AUTORÉGRESSIF (J8, N)

(24 / 04 / 2020, © Monfort, Dicostat2005, 2005-2020)

(i) Un **modèle autorégressif** est un **modèle** défini par une **équation de récurrence** stochastique. Dans l'**espace des variables**, cette équation est de la forme (cf **processus autorégressif**) :

$$(1) \quad \eta = f(\eta_{-1}, \dots, \eta_{-H}, \xi) + \varepsilon,$$

dans laquelle η est une **variable endogène**, η_{-h} la variable endogène décalée de h « instants » (ie, si L est l'**opérateur retard** $\eta \mapsto L \eta = \eta_{-1}$, alors $\eta_{-h} = L^h \eta$, $\forall h \in \mathbb{N}_H^*$) et où ξ est une liste de K **variables exogènes**.

La fonction f incorpore généralement un **paramètre d'intérêt** θ .

Les variables $\eta_{-1}, \dots, \eta_{-H}$ et ξ sont appelées **variables prédéterminées** du modèle, car elles sont observées dès que η l'est.

Dans le cas particulier où f est linéaire, (1) s'écrit :

$$(2) \quad \eta = \sum_{h=1}^H a_h \cdot \eta_{-h} + \xi' b + \varepsilon = A(L) \eta + \xi' b + \varepsilon,$$

où $A(L)$ désigne le **polynôme formel** $\sum_{h=1}^H a_h L^h$.

(ii) Le principal problème posé par ce type de modèles vient de ce que la **méthode des moindres carrés ordinaires** n'est pas applicable car (cf **biais**) :

$$(3) \quad E \varepsilon \neq 0 \quad (\text{ie } E(\varepsilon / \eta_{-1}, \dots, \eta_{-H}, \xi) \neq 0),$$

ce qui contredit une hypothèse de base de cette méthode.

Des considérations a priori conduisent à ajouter au modèle autorégressif une **hypothèse de stabilité**. Ainsi, dans le cas linéaire (2), on suppose que les racines (réelles ou complexes) du polynôme caractéristique :

$$(4) \quad z^H - a_1 z^{H-1} - \dots - a_H = 0$$

associé à (2) sont de module inférieur à 1.

(iii) En pratique, on dispose d'un T -**échantillon** X , à valeurs dans $M_{TK}(\mathbf{R})$, des K « véritables » variables exogènes constituant ξ et d'un T -échantillon y , à valeurs dans \mathbf{R}^T , de la variable endogène η .

L'équation (2) est alors « observée » dans l'**espace des états** (ou **espace des observations**) selon :

$$(5) \quad y_t = a_1 y_{t-1} + \dots + a_H y_{t-H} + X_t b + u_t, \quad \forall t \in \{H+1, \dots, T\},$$

et les variables y_1, \dots, y_H sont considérées comme des **valeurs initiales**, ou **valeurs d'amorçage**, aléatoires ou non. L'équation (5) relie alors la **trajectoire** $y = (y_t)_{t=1, \dots, T}$ d'un **processus stochastique** réel scalaire à celle $X = (X_t)_{t=1, \dots, T}$ d'un autre processus (**processus vectoriel**).

La propriété (3) se traduit par les inégalités :

$$(6) \quad E(u_t / y_{t-1}, \dots, y_{t-H}, X_t) \neq 0, \quad \forall t \in \{H+1, \dots, T\}.$$

L'estimation de (5) conduit à étudier deux cas :

(a) si $u = (u_t)_{t=1, \dots, T}$ symbolise la trajectoire (inobservable) d'un **processus sans corrélation** propre (ie si $C(u_s, u_t) = 0, \forall (s, t) \text{ tq } t \neq s$), la **méthode des mco** fournit un estimateur biaisé (cf **biais**) :

$$(7) \quad E(\hat{a}, \hat{b}) \neq (a, b), \quad \text{avec } a = (a_1, \dots, a_H),$$

(\hat{a}, \hat{b}) étant l'**estimateur des mco** calculé directement sur le modèle (5).

Cependant, si u est gaussien, cet estimateur est aussi l'**estimateur du mv** et il en possède les **propriétés asymptotiques** optimales : **convergence en probabilité, efficacité asymptotique, normalité asymptotique** ;

(b) si u est la trajectoire d'un processus corrélé (ie s'il existe au moins un couple (s, t) tq $C(u_s, u_t) \neq 0$), la méthode des mco est, non seulement biaisée, mais aussi divergente en probabilité (**inconvergence**) :

$$(8) \quad \text{plim}_T(\hat{a}, \hat{b}) \neq (a, b).$$

De plus, le **test de DURBIN-WATSON** est inefficace.

Si les corrélations $C(u_s, u_t)$ ont une forme connue (ie dépendent de façon analytique connue d'un nombre limité de paramètres θ), la **méthode des moindres carrés généralisés** est applicable (eg en définissant un modèle à perturbations autocorrélées au premier ordre) (cf **modèle à perturbations liées**), de même que la **méthode du mv** (avec les **paramètres principaux** (a, b, θ)).

(iv) On peut toujours passer d'une **forme autorégressive** (2) (ou (5)) à une **forme à retards échelonnés** tq (dans l'espace des variables) :

$$(9) \quad \eta = d + c_0 \xi + c_1 \xi_{-1} + \dots + c_H \xi_{-H} + \varphi,$$

la **constante** d incorporant les valeurs initiales $(y_h)_{h=1, \dots, H}$, considérées comme des variables presque certaines (cf **modèle à retards échelonnés**). Les méthodes propres à ce type de modèle sont alors souvent praticables.

(v) Dans certains cas, on peut aussi transformer un modèle autorégressif en **modèle à erreurs sur les variables**.

(vi) Le modèle précédent a été étendu :

(a) au cas d'un **processus au passé infini**, ie de la forme (dans l'espace des variables) :

$$(10) \quad \eta = f(\eta_{-1}, \dots, \eta_{-h}, \dots, \xi) + \varepsilon,$$

dans laquelle la fonction f incorpore une suite de paramètres $b = (b_h)_{h \in \mathbf{N}^*}$ associée à la suite $\eta = (\eta_{-h})_{h \in \mathbf{N}^*}$, voire aussi un paramètre d'intérêt θ . Cette forme nécessite de réduire le nombre de paramètres à estimer, eg en imposant à la suite b un échelonnement selon une forme analytique simple (cf **modèle à retards échelonnés, parcimonie**) ;

(b) au cas d'un processus en **temps** continu (eg $T = \mathbf{R}_+$), lequel peut s'écrire dans l'espace des variables (eg sous forme non linéaire) :

$$(11) \quad \eta = f\left\{\int_0^\infty b_\tau \eta_{t-\tau} dv(\tau), \xi\right\} + \varepsilon,$$

où v est une **mesure** donnée sur l'**espace du temps** (T, \mathcal{B}_T) .

Dans l'espace des observations (ou **espace des trajectoires**) (11) se réécrit alors :

$$(12) \quad y_t = f\left\{\int_0^\infty b_\tau y_{t-\tau} dv(\tau), X_t\right\} + u_t, \quad \forall t \in T,$$