MODÈLE AVEC AUTOCORRÉLATION TEMPORELLE (D2, J1, J8, N)
(23 / 04 / 2020, © Monfort, Dicostat2005, 2005-2020)

(i) Soit :

(1) $y = F(b) + u$, avec $E\,u = 0$, $V\,u = \Sigma$, $b \in \mathbf{R}^Q$,

un **modèle de régression multiple** non linéaire écrit dans un **espace d'observation** donné. On suppose que le **processus** en **temps** discret $U = (U_t)_{t \in T}$ (avec $T \subset \mathbf{Z}$), qui engendre la **série temporelle** (inobservable) $u = (u_t)_{t=1,\dots,T}$, est un **processus stationnaire en covariance**.

On dit que (1) est un **modèle avec autocorrélation temporelle**, ou **modèle avec perturbations autocorrélées**, ssi U est :

(a) soit un **processus autorégressif** ar (p), ie :

(2) $U_t = \Sigma_{i=1}^{p} \rho_i U_{t-i} + V_t$, $\forall\, t \in T$,

avec $E\,V_t = 0$ et $C(V_s, V_t) = \delta_{st} \,.\, \sigma_V^2$, $\forall\,(s, t) \in T^2$;

(b) soit un **processus de moyenne mobile** mm (q), ie :

(3) $U_t = \Sigma_{j=1}^{q} \rho_j V_{t-j} + V_t$, $\forall\, t \in T$,

avec $E\,V_t = 0$ et $C(V_s, V_t) = \delta_{st}\, \sigma_V^2$, $\forall\,(s, t) \in T^2$.

(ii) Un exemple élémentaire est celui du **processus de MARKOV** (autorégressif d'ordre 1) :

(4) $U_t = \rho_1 \,.\, U_{t-1} + V_t$, où $|\rho_1| \leq 1$,

avec $E\,V_t = 0$ et $C(V_s, V_t) = \delta_{st} \,.\, \sigma_V^2$.

La **matrice des covariances** de U est alors de la forme :

(5) $V\,U = (1 - \rho_1)^{-1} \,.\, \sigma_V^2 \,.\, \Omega$ (notée $\sigma_u^2\, \Omega$),

où $\sigma_u^2 = (1 - \rho_1)^{-1} \,.\, \sigma_V^2$ et où $\Omega = (\omega_{st})_{(s,t)}$ est une **matrice de TOEPLITZ** tq :

$$
(6) \quad \omega_{st} = \begin{cases} 1 & \text{si } t = s \\ \rho_1 & \text{si } t = s+1 \text{ ou si } s = t+1 \\ \rho_1^2 & \text{si } t = s+2 \text{ ou si } s = t+2 \\ \dots & \\ \rho_1^{T-1} & \text{si } t = 1 \text{ et } s = T\text{, ou si } s = 1 \text{ et } t = T. \end{cases}
$$

On distingue deux cas :

(a) si Ω est connue, ie si ρ_1 est donné, la **méthode des moindres carrés généralisés** est applicable, avec :

(7) $(1 - \rho_1^2)\,\Omega^{-1} = H'\,H,$

où $H = (h_{st})_{(s,t)}$ est définie par :

$$
(8) \quad h_{st} = \begin{cases} (1 - \rho_1^2)^{1/2} & \text{si } s = t = 1, \\ -\rho_1 & \text{si } s = t + 1, \\ 0 & \text{si } t > s \text{ ou si } s > t + 1, \\ 1 & \text{si } t = s \text{ et } t > 1. \end{cases}
$$

Autrement dit, on transforme le modèle (1) à l'aide de H selon :

(9) $H\,y = H\,.\,F\,(b) + H\,u,$

et l'on applique à (8) la **méthode des moindres carrés ordinaires** ;

(b) si Ω n'est pas connue (cas général), on utilise les mêmes alternatives que dans le cas où la **méthode des mcg** s'applique. Si la forme (1) est linéaire (avec $Q = K$ et $F = X \in M_{NK}$ (**R**)), on peut procéder en deux étapes :

(b_1) estimation du **coefficient d'autocorrélation** ρ_1 :

(b_{11}) soit par la **méthode des mco**, ce qui conduit à l'**estimateur** :

(10) $\rho_1^{\wedge} = \{\Sigma_{t=2}^{T}\, u_t^{\wedge}\, u_{t-1}^{\wedge}\} / \Sigma_{t=1}^{T}\, (u_t^{\wedge})^2$;

(b_{12}) soit par la **méthode de J. DURBIN**, qui consiste à appliquer la méthode des mco sur le modèle (non linéaire en (b, ρ_1)) suivant, parfois appelé **modèle quasi-différencié** (cf **quasi-différentiation**) :

(11) $y_t = \rho_1\, y_{t-1} + X_t\, b - \rho_1\, X_{t-1}\, b + v_t$.

Ces méthodes fournissent un estimateur $\rho_1^{\wedge}$ convergeant en probabilité vers ρ_1 (cf **convergence en probabilité**) ;

(b_2) estimation de b en remplaçant dans (9) la fonction matricielle $H = H\,(\rho_1)$ par la matrice estimée à l'aide de la première étape, ie par $H^{\wedge} = H\,(\rho_1^{\wedge})$ (**procédure de D. COCHRANE - G.H. ORCUTT**).

Pour estimer (b, ρ_1), il est aussi possible d'utiliser l'une des méthodes suivantes :

(a) **méthode du maximum de vraisemblance** appliquée au modèle (11) précédent ;

(b) « balayage » du segment [-1, +1] par ρ_1 : ie on fait varier ρ_1 par **pas** de longueur 1 / D (où le nombre D >> 0 est donné a priori) dans le segment précédent. A chaque valeur ρ_1 (D) obtenue pour ρ_1 , on calcule $b^\wedge$ par la méthode des mco ; la procédure s'arrête lorsque eg $||y - y^\wedge||^2$ est minimum, ce qui fournit une **approximation** de la valeur de $b^\wedge$ cherchée. Pour tester l'**hypothèse nulle** de non autocorrélation des perturbations suivante :

(12) $H_0 : \rho_1 = 0$,

on utilise souvent le **test de DURBIN-WATSON**.