ESTIMATEUR DES MOINDRES CARRÉS GÉNÉRALISÉS (H1, J1)
(21 / 05 / 2020, © Monfort, Dicostat2005, 2005-2020)

(i) La **méthode des moindres carrés généralisés** concerne le **modèle de régression** suivant, exprimé dans l'**espace d'observation** $\mathbf{R}^N$:

(1) $y = z + u$, avec $E\, u = 0$ et $V\, u = \Sigma > 0$.

On appelle **estimateur des moindres carrés généralisés**, ou **prévision des moindres carrés généralisés**, de la **variable endogène** y la solution (en z) du problème de **programmation mathématique** suivant :

(2) $\inf_{z \in V} ||y - z||^2_{\Sigma 1}$,

où $V \subset \mathbf{R}^N$ et $\Sigma 1$ désigne la **matrice inverse** Σ^{-1}. La **distance** $(y, z) \mapsto ||y - z||^2_{\Sigma 1}$ est ainsi calculée avec la **métrique** définie par la **forme quadratique** associée à Σ^{-1}.

(ii) Si (1) est une **régression linéaire** (pr à b), on pose généralement $V = V_K = \{z \in R^N : z = X\, b, \forall\, b \in \mathbf{R}^K\}$. Sous l'hypothèse rg X = K, la solution est :

(3) $y_\Sigma^\# = X\, b_\Sigma^\#$, avec $b_\Sigma^\# = (X'\, \Sigma^{-1}\, X)^{-1}\, X'\, \Sigma^{-1}\, y$.

On appelle ainsi **estimateur des moindres carrés généralisés** (mcg), ou **estimateur de C.F. GAUSS - A.A MARKOV - A.C. AITKEN**, de b le vecteur $b_\Sigma^\#$ explicité en (3).

On appelle « **estimateur** » **des mcg** de la **perturbation** u le vecteur $u_\Sigma^\#$ des **résidus**, défini selon :

(4) $u_\Sigma^\# = y - y_\Sigma^\# = y - X\, b_\Sigma^\# = M_\Sigma\, y$,

avec $M_\Sigma = I_N - X\,(X'\, \Sigma^{-1}\, X)^{-1}\, X'\, \Sigma^{-1}$.

(ii) On montre que $b_\Sigma^\#$ vérifie les propriétés suivantes :

(a) $E\, b_\Sigma^\# = b$ (**estimateur sans biais**), $V\, b_\Sigma^\# = (X'\, \Sigma^{-1}\, X)^{-1}$ (**matrice de dispersion**) ;

(b) $E\, u_\Sigma^\# = 0$ (perturbation nulle en moyenne) ;

(c) $E\, y_\Sigma^\# = E\, y = X\, b$; $V\, y_\Sigma^\# = V\,(X\, b_\Sigma^\#) = X\,(X'\, \Sigma^{-1}\, X)^{-1}\, X'$;

(d) si Σ est connue, la **méthode des mcg**, appliquée au modèle (1), équivaut à la **méthode des mco** appliquée au modèle transformé suivant, parfois dit **modèle « sphéricisé »** (car la **matrice de dispersion** des nouvelles perturbations est une **matrice scalaire**) :

(5) $\Sigma^{-1/2}\, y = \Sigma^{-1/2}\, z + \Sigma^{-1/2}\, u$;

(e) E $(\sigma^2)_\Sigma^\# = \sigma^2$, avec $(\sigma^2)_\Sigma^\# = (N - K)^{-1}\ u'\ \Sigma^{-1}\ u = (N - K)^{-1}\ ||u||^2_{\Sigma-1}$;

(f) si $u \sim \mathcal{N}_N\ (0, \Sigma)$ (**loi normale multidimensionnelle** centrée), alors $b_\Sigma^\# \sim \mathcal{N}_K\ (b, (X'\ \Sigma^{-1}\ X)^{-1})$, et $b_\Sigma^\#$ est aussi l'**estimateur du maximum de vraisemblance** (gaussienne) de b ;

(g) $b_\Sigma^\#$ vérifie le **théorème de GAUSS-MARKOV-AITKEN** (cf **méthode des mcg**).

(iii) Des exemples d'estimateurs généralisés sont donnés par :

(a) le modèle de **régression pondérée** : le plus souvent, il s'agit d'un modèle en **coupe instantanée** dont la **perturbation** est hétéroscédastique (cf **hétéroscédasticité**) ;

(b) le modèle à perturbations autocorrélées : le plus souvent, il s'agit d'un modèle sur **séries temporelles** (cf **autocorrélation**, **modèle à perturbations liées**).

(iv) Si le modèle de régression est non linéaire (pr à b) (cf **modèle non linéaire**), on pose généralement $V = V_Q = \{z \in \mathbf{R}^N : z = F\ (b), \forall\ b \in \mathbf{R}^Q\}$, où $F : \mathbf{R}^Q \mapsto \mathbf{R}^N$ est une fonction vectorielle dépendant de X.

La solution du problème (2) ne s'explicite pas comme en (3) ou en (4). C'est davantage par analogie avec le cas linéaire (i) que les estimateurs des mcg non linéaires sont calculés, plutôt qu'en raison de propriétés d'optimalité (au moins à distance N finie).

Des formules tq les précédentes peuvent être dérivées à partir du **modèle linéaire tangent** (cf **modèle linéarisé**), aussi appelé **modèle affine tangent**. Des méthodes de calcul numérique itératives sont alors possibles.