MÉTHODE DES MOINDRES CARRÉS ORDINAIRES (H3, J1)
(19 / 12 / 2019, © Monfort, Dicostat2005, 2005-2019)

La **méthode des moindres carrés ordinaires** (mco) est une **méthode des moindres carrés** très utilisée pour estimer les paramètres d'une **régression** (cf **régression linéaire**, **régression non linéaire**) dont la **perturbation aléatoire** est homoscédastique (cf **homoscédasticité**) et intervient sous forme additive (cf **additivité**). Cette méthode peut s'adapter à divers **contextes statistiques**.

Elle consiste à minimiser la **distance** (ou la **norme**) euclidienne entre une **variable endogène** y et la partie certaine (**espérance conditionnelle**) z de la régression, laquelle est astreinte à parcourir une partie spécifiée de l'**espace d'observation**. Autrement dit, elle minimise la norme euclidienne de la perturbation du modèle (cf aussi **méthode de moindre norme**).

Il s'agit d'un cas particulier du **théorème de la projection orthogonale** dans un **espace de HILBERT**.

(i) On considère un **modèle de régression** à perturbation additive, exprimé dans l'**espace des variables** $\mathbf{R}^K \times \mathbf{R}$ (K exogènes, G = 1 endogène) :

(1) $\eta = \zeta + \varepsilon$, avec $E\,\varepsilon = 0$, $V\,\varepsilon = \sigma^2$,

et « observé » dans l'**espace des observations** selon :

(2) $y = z + u$, avec $E\,u = 0$, $V\,u = \sigma^2 \,.\, I_N$,

où l'espace ambiant des observations y est $\mathbf{R}^N$.

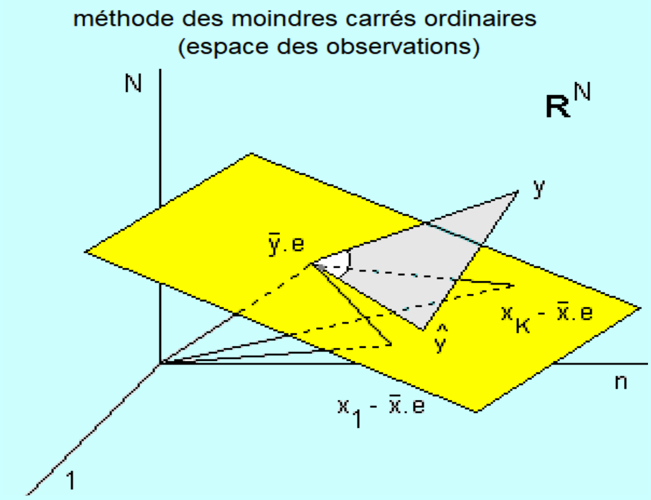
La **méthode des moindres carrés ordinaires** (mco), ou **méthode de C.F. GAUSS - A.A. MARKOV**, consiste à minimiser $q(z) = ||y - z||^2 = (y - z)'\,(y - z)$ (ie à minimiser $||u||^2$) pr à $z \in \mathcal{V}$ (sous-**espace vectoriel** de $\mathbf{R}^N$).

L'interprétation graphique de la méthode peut s'effectuer dans l'**espace des variables** (qui correspond à une « loi scientifique »), mais le plus souvent dans l'**espace des observations**.

La solution ainsi obtenue est appelée **solution des moindres carrés ordinaires** du problème d'optimisation précédent. On la note généralement $z^{\wedge}$.

(ii) Si le modèle est linéaire, $z = X\,b$ est l'équation, paramétrée par $b \in \mathbf{R}^K$, de la variété linéaire $\mathcal{V}$ (de dimension K dans $\mathbf{R}^N$). La méthode revient à minimiser pr à b la **forme quadratique** non homogène (cf représentation graphique ci-dessous) :

(3) $q(b) = ||y - X\,b||^2 = (y - X\,b)'\,(y - X\,b)$.



Cette méthode s'associe à l'hypothèse (stochastique) du second ordre selon laquelle il y a **homoscédasticité** de la perturbation u, ie :

(4) $V\,u = \sigma^2 . I_N$ (matrice scalaire),

Elle équivaut à minimiser pr à b $\in$ $\mathbf{R}^K$ la forme quadratique (3) exprimée dans la **métrique** définie par $\Sigma = V\,u = \sigma^2 . I_N$, soit formellement :

(5) $q_\sigma (b) = ||y - X\,b||_q^2 = (y - X\,b)' (\sigma^{-2} I_N^{-1}) (y - X\,b)$.

Il s'agit donc d'un cas particulier de la **méthode des moindres carrés généralisés**.

On appelle **estimateur des moindres carrés ordinaires linéaire**, ou **estimateur de C.F. GAUSS - A.A MARKOV linéaire**, de b la solution du problème d'optimisation précédent, notée $b^\wedge$ (ou encore β).

(iii) Si le modèle de régression est non linéaire, alors (dans l'espace des variables) $\eta = f(\xi, b) + \varepsilon$, avec $E\,\varepsilon = 0$, $V\,\varepsilon = \sigma^2$. Dans l'espace $\mathbf{R}^N$ des observations, l'écriture est la suivante z = F (b) + u. Par suite, z = F (b) est l'équation, paramétrée par b $\in$ $\mathbf{R}^Q$, de la sous-variété (hyper-surface) $\mathcal{V}$ de $\mathbf{R}^N$ (parfois notée $\mathcal{V}_F$). Cette variété est généralement supposée être localement de dimension Q, ie rg D F (b) = Q dans un voisinage de la **vraie valeur** b* de b.

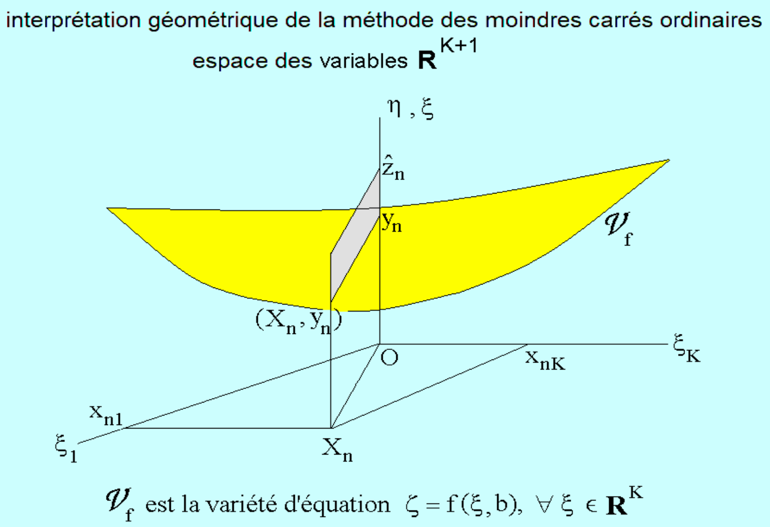
La **méthode des moindres carrés ordinaires non linéaire**, ou **méthode de C.F. GAUSS - A.A. MARKOV non linéaire**, consiste alors à minimiser pr à b la fonction :

(6) $q(b) = ||y - F(b)||^2 = (y - F(b))' (y - F(b))$,

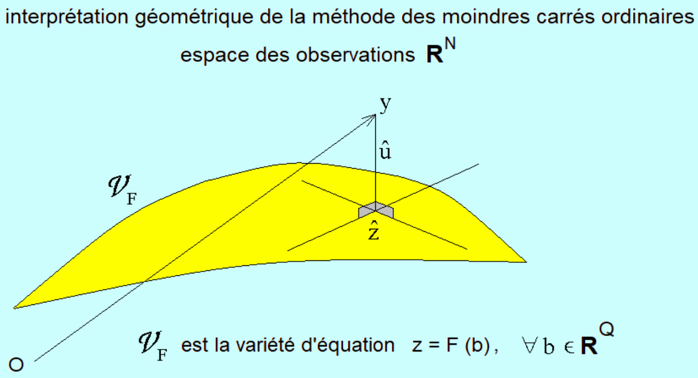
ce qui revient encore à supposer (4) et à minimiser pr à b la fonction (cf graphiques ci-après) :

(7) $q_\sigma (b) = ||y - F(b)||_q^2 = (y - F(b))' (\sigma^{-2} I_N^{-1}) (y - F(b))$.

Géométrie de la méthode des mco dans l'espace des variables



Géométrie de la méthode des mco dans l'espace des observations



Il s'agit encore d'un cas particulier de la **méthode des mcg** non linéaire.

Le minimum obtenu :

(8) $q\,(b^{\wedge}) = ||y - F\,(b^{\wedge})||^2 = ||u^{\wedge}||^2$

pour la fonction q est appelé **somme des carrés des résidus** $u_n^{\wedge}$ (n = 1 ,..., N) (cf **forme quadratique résiduelle**).

(iv) En particulier, si F est linéaire (ie si Q = K et F = X $\in M_{NK}$ (**R**)), on montre que :

(9)

$$E\,q\,(b^{\wedge}) = (N - K)\,\sigma^2,$$

$$V\,q\,(b^{\wedge}) = \sigma^4 \,.\, \{(N - K)\,(\beta - 1) + (\beta - 3)\,.\,(\Sigma_{n=1}^{N}\, g_{nn} - K)\},$$

où $\beta\,\sigma^4$ désigne le moment d'ordre 4 de ε (avec $\beta > 0$) et où $G = X\,(X'\,X)^{-1}\,X' = (g_{\alpha\beta})_{(\alpha,\beta)}$.

(v) Dans le cas d'un **modèle de régression non linéaire implicite**, l'équation s'écrit :

(10) $g(\xi, \eta, b) = \varepsilon$, avec $E\,\varepsilon = 0$,

dans laquelle η est la variable endogène, ξ le vecteur des K variables exogènes, $b \in \mathbf{R}^Q$ le **paramètre principal** (à estimer) et ε une **vars** (inobservable), et lorsque l'équation (10) est « observée » selon :

(11) $g(X_n, y_n, b) = u_n$, $\forall\, n \in N_N^*$,

où y_n est la n-ième observation de η (ou la n-ième coordonnée d'un **vecteur aléatoire** y à valeurs dans $\mathbf{R}^N$), X_n celle de ξ (n-ième ligne d'une matrice $X \in M_{NK}(\mathbf{R})$) et u_n la n-ième **copie** de la va ε (n-ième coordonnée d'un vecteur aléatoire u, inobservable et à valeurs dans $\mathbf{R}^N$), on peut écrire (notations évidentes) :

(12) $G(X, y, b) = u$.

La méthode des moindres carrés ordinaires (non linéaire) consiste alors à estimer b en minimisant $||u||^2 = \Sigma_{n=1}^N (g(X_n, y_n, b))^2$ pr à b. On note encore $b^\wedge$ la solution des mco de ce **problème d'optimisation**, ie l'estimateur du même nom de b.