MAXIMUM DE VRAISEMBLANCE LINÉAIRE (A10, A11, H3)

La **méthode du maximum de vraisemblance linéaire** est l'appellation parfois donnée à une **modification** de la **méthode du maximum de vraisemblance** destinée à en faciliter les calculs numériques.

(i) Soit $(X, \theta) \mapsto f(X, \theta)$ la **fonction de vraisemblance** d'un **modèle statistique**, supposée différentiable pr à son deuxième argument (cf **différentiabilité**, **dérivée partielle**), avec $\theta \in \mathbf{R}^Q$.

Sous des **conditions de régularité** classiques, l'**estimateur du mv** du **paramètre** θ est solution de l'**équation de vraisemblance** en θ (cf **équation estimante**) :

(1) $D_2 f(X, \theta) = 0.$

Si $\theta_0 \in \Theta$ est une valeur donnée du **paramètre d'intérêt**, la **modification** (ici **linéarisation**) consiste à remplacer l'équation vectorielle (1) par sa linéarisée (en θ) au voisinage de θ_0 , ie à remplacer $D_2 f(X, \theta) = g(X, \theta)$ par son développement de TAYLOR (coordonnée par coordonnée) au voisinage de θ_0 (cf aussi **formule de TAYLOR stochastique**). Ce développement s'écrit (à l'ordre 1) :

(2) $g(X, \theta) = g(X, \theta_0) + (D_2 g(X, \theta_0)).(\theta - \theta_0) + \varepsilon_X(\theta),$

d'où :

(3) $D_2 f(X, \theta) = D_2 f(X, \theta_0) + (D_2^{(2)} f(X, \theta_0)).(\theta - \theta_0) + \varepsilon_X(\theta).$

La méthode revient alors à résoudre le **système linéaire** suivant pr à $\theta \in \mathbf{R}^Q$:

(4) $(D_2^{(2)} f(X, \theta_0)) . \theta = - D_2 f(X, \theta_0) + (D_2^{(2)} f(X, \theta_0)) . \theta_0 .$

(ii) La méthode suppose notamment que :

(a) θ_0 constitue déjà une **approximation** correcte de θ^* (la **vraie valeur** du paramètre θ) ou de son **estimateur** « exact » (ie celui du mv, donné par (1)) ;

(b) $\varepsilon_X(\theta)$ # 0. Ceci est réalisé eg lorsqu'une **convergence stochastique** tq $\text{plim}_N\, \varepsilon_X(\theta) = 0$ (**convergence en probabilité**), dans laquelle N désigne le nombre d'observations de X, est assez rapide.