

On a un modèle génératif M complexe qui se décompose en trois morceaux $M = (M_1, M_2, M_3)$ comme suit :

$$M_1 : Y|X^{I_1}, X^{I_2}, Z \sim X^{I_1} A_{I_1} + X^{I_2} A_{I_2} + \varepsilon \quad (1)$$

$$M_2 : X^{I_2}|X^{I_1}, Z \sim X^{I_1} B_{I_1}^{I_2} + \varepsilon_{X^{I_2}} \quad (2)$$

$$M_3 : X^{I_1}|Z \sim GM(\alpha) \quad (3)$$

On pose $\theta = (\alpha, \Sigma, \Sigma_{X^{I_2}}, B_{I_1}^{I_2}, A_{I_1}, A_{I_2})$ Pour un Z donné on a donc

$$P_{M,Z}(Y, X^{I_1}, X^{I_2}|\theta) = P_{M,Z}(X^{I_1}|\theta)P_{M,Z}(X^{I_2}|X^{I_1};\theta)P_{M,Z}(Y|X^{I_2}, X^{I_1};\theta) \quad (4)$$

$$= P_{M_3,Z}(X^{I_1}|\alpha)P_{M_1,M_2,Z}(X^{I_2}|X^{I_1};\Sigma, \Sigma_{X^{I_2}}, B_{I_1}^{I_2}, A_{I_1}, A_{I_2})P_{M_1,M_2,Z}(Y|X^{I_2}, X^{I_1};\Sigma, \Sigma_{X^{I_2}}, B_{I_1}^{I_2}, A_{I_1}, A_{I_2}) \quad (5)$$

$P_{M_3,Z}(X^{I_1}|\alpha)$ vit tout seul et est estimé par Mixmod. On sait également que

$$P_{M_2,Z}(X^{I_2}|X^{I_1};\Sigma, \Sigma_{X^{I_2}}, B_{I_1}^{I_2}, A_{I_1}, A_{I_2}) = P_{M_2,Z}(X^{I_2}|X^{I_1};\Sigma_{X^{I_2}}, B_{I_1}^{I_2}) \text{ et} \quad (6)$$

$$P_{M_1,Z}(Y|X^{I_2}, X^{I_1};\Sigma, \Sigma_{X^{I_2}}, B_{I_1}^{I_2}, A_{I_1}, A_{I_2}) = P_{M_1,Z}(Y|X^{I_2}, X^{I_1};\Sigma, A_{I_1}, A_{I_2}) \quad (7)$$

De manière générale les gens s'arrêtent là, et obtiennent le modèle complet... qui ne tient pas compte de la structure. Mais nous, nous ne sommes pas juste sous M_2 pour les sous-régression ni juste sous M_1 pour Y . Donc la dépendance persiste (c'est ce qui permet que la contrainte aie une utilité, sinon elle serait inutile).

Du coup, il faut définir proprement $P_{M_1,M_2,Z}(X^{I_2}|X^{I_1};\Sigma, \Sigma_{X^{I_2}}, B_{I_1}^{I_2}, A_{I_1}, A_{I_2})$ et $P_{M_1,M_2,Z}(Y|X^{I_2}, X^{I_1};\Sigma, \Sigma_{X^{I_2}}, B_{I_1}^{I_2}, A_{I_1}, A_{I_2})$

On a $Y|Z \sim X^{I_1} A_{I_1} + X^{I_2} A_{I_2} + \varepsilon$ mais le modèle M signifie aussi que :

$$Y|Z \sim X^{I_1} A_{I_1} + X^{I_1} B_{I_1}^{I_2} A_{I_2} + \varepsilon_{X^{I_2}} A_{I_2} + \varepsilon \quad (8)$$

on sait que quand on connaît X_{I_1} , X_{I_2} , et $B_{I_1}^{I_2}$, on connaît aussi (par M_2) $\varepsilon_{X^{I_2}}$ de manière exacte (pas son paramètre $\Sigma_{X^{I_2}}$ mais bien la valeur de la variable elle-même).

On se demande si $P_{M_1,Z}(Y|X^{I_2}, X^{I_1};\Sigma, A_{I_1}, A_{I_2})$ est différent de $P_{M_1,M_2,Z}(Y|X^{I_2}, X^{I_1};\Sigma, \Sigma_{X^{I_2}}, B_{I_1}^{I_2}, A_{I_1}, A_{I_2})$. donc si (sous l'hypothèse du modèle M_2) $\Sigma_{X^{I_2}}$ et $B_{I_1}^{I_2}$ sont indépendants de $(Y|X^{I_2}, X^{I_1};\Sigma, A_{I_1}, A_{I_2})$.

Pour moi, il y a dépendance mais je n'arrive pas à l'explicitier. C'est a priori pour cela qu'on a besoin de faire du séquentiel en s'appuyant sur le A explicatif. Je vais continuer avec les formules d'indépendance, ne sachant pas faire autrement.

$\Sigma_{X^{I_2}}$ matrice diagonale constituée des $\sigma_{X,j}$ (pour la variable j à gauche).

$$P_{M_3,Z}(X^{I_1}|\alpha) \quad \text{maximisé de manière indépendante (du reste et de la contrainte) par mixmod} \quad (9)$$

$$P_{M_2,Z}(X^{I_2}|X^{I_1};\Sigma_{X^{I_2}}, B_{I_1}^{I_2}) = \prod_j \frac{1}{\sigma_{X,j}\sqrt{2\pi}} e^{-\frac{(X^j - X^{I_1} B_{I_1}^j)^2}{2\sigma_{X,j}^2}} \quad (10)$$

$$P_{M_1,Z}(Y|X^{I_2}, X^{I_1};\Sigma, A_{I_1}, A_{I_2}) = \quad (11)$$