

Approche bayésienne des modèles structurels

Séverine Demeyer
severine.demeyer@lne.fr

LNE-CNAM

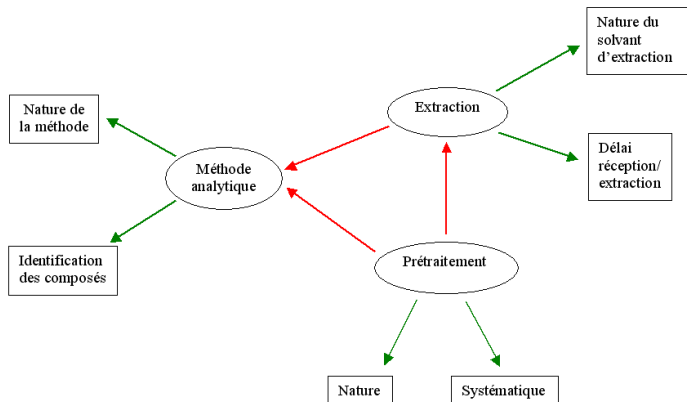
26 novembre 2009

PLAN

- 1 Présentation des modèles structurels
- 2 Modélisation
- 3 Estimation
- 4 Conclusion et perspectives

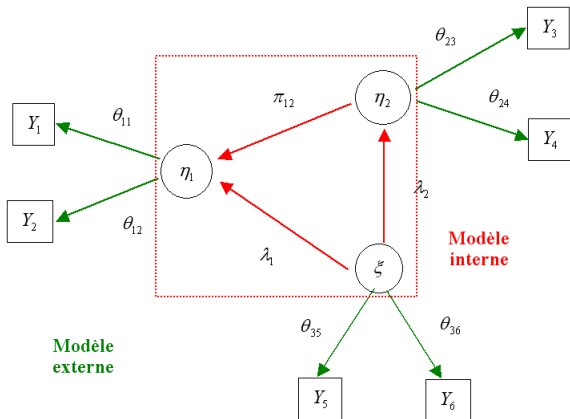
Cas de laboratoires d'analyses

Des questions relatives aux processus de mesures sont groupées dans des blocs qui peuvent être reliés par des liens de causalité.



Introduction des notations

Chaque bloc est résumé par une variable latente (VL) unidimensionnelle.



Ecriture du modèle

Equations du modèle de mesure

$$Y_1 = \theta_{11}\eta_1 + \varepsilon_1$$

$$Y_2 = \theta_{12}\eta_1 + \varepsilon_2$$

$$Y_3 = \theta_{23}\eta_2 + \varepsilon_3$$

$$Y_4 = \theta_{24}\eta_2 + \varepsilon_4$$

$$Y_5 = \theta_{35}\xi + \varepsilon_5$$

$$Y_6 = \theta_{36}\xi + \varepsilon_6$$

Equations du modèle structurel

$$\eta_1 = \pi_{12}\eta_2 + \lambda_1\xi + \delta_1$$

$$\eta_2 = \lambda_2\xi + \delta_2$$

Structure de covariance induite

Soit $Z = \{\eta_1, \eta_2, \xi\}$.

Les équations du modèle intérieur induisent une structure de covariance au niveau des VL.

Soit $\Sigma_Z = \text{cov}(Z)$. On a : $\Sigma_Z = \Sigma_Z(\pi_{12}, \lambda_1, \lambda_2, \delta_1, \delta_2, \phi)$.

Résultats

$$\text{Var}(\eta_1) = (\pi_{12}\lambda_2 + \lambda_1)^2 \phi + \text{Var}(\delta_1) + \pi_{12}^2 \text{Var}(\delta_2)$$

$$\text{Var}(\eta_2) = \lambda_2^2 \phi + \text{Var}(\delta_2)$$

$$\text{Var}(\xi) = \phi$$

$$\text{Cov}(\eta_1, \eta_2) = (\lambda_1\lambda_2 + \pi_{12}\lambda_2^2) \phi + \pi_{12} \text{Var}(\delta_2)$$

$$\text{Cov}(\eta_1, \xi) = (\pi_{12}\lambda_2 + \lambda_1) \phi$$

$$\text{Cov}(\eta_2, \xi) = \lambda_2 \phi$$

PLS Path Modelling

Méthode appartenant à la famille des méthodes PLS, spécifique aux modèles à équations structurelles. Modèle prédéfini

Hypothèses

- Pas d'hypothèses distributionnelles.
- Non corrélation des erreurs et des VL.
- Non corrélation des erreurs entre elles.

Estimation

Algorithme itératif qui alterne la construction de VL sur le modèle externe et sur le modèle interne. Après convergence, les coefficients du modèle intérieur sont estimés par régression.

LISREL

Analyse de la structure de covariance des variables manifestes associées aux VL exogènes.

Hypothèses

- Normalité et indépendance des observations
- La matrice de covariance estimée s'exprime en fonction des paramètres à estimer.
- Grands échantillons

Estimation

Maximum de vraisemblance

Pourquoi une approche bayésienne ?

- Possibilité d'apporter de l'information en plus via les distributions a priori.
- Bénéficier de la distribution complète des paramètres (a posteriori).
- Inscire le modèle structurel au sein d'un autre modèle et bénéficier d'une estimation simultanée de tous les paramètres et des variables latentes.

Description de la démarche bayésienne

- Travail sur données et VL centrées, réduites.
- Le modèle est prédéfini.
- Les données vérifient les hypothèses de normalité et d'indépendance.
- Les VL sont traitées comme des données supplémentaires, manquantes.
- Sous les hypothèses adéquates, la distribution conditionnelle a posteriori des paramètres et des VL se factorise.

Estimation

L'algorithme de Gibbs alterne les tirages dans les distributions conditionnelles a posteriori des VL et des paramètres.

Ré-écriture matricielle du modèle

Modèle de mesure : individus \times variables manifestes

$$\begin{array}{ccccccc}
 Y & = & Z & \times & \theta & + & \varepsilon \\
 (n \times p) & & (n \times q) & & (q \times p) & & (n \times p)
 \end{array}$$

Le niveau suivant modélise les relations entre les variables latentes :

Modèle structurel : individus \times VL dépendantes

$$\begin{array}{ccccccc}
 H & = & Z & \Lambda & + & \delta \\
 (n \times q_1) & & (n \times q) & (q \times q_1) & & (n \times q_1)
 \end{array}$$

$$\begin{array}{ccccccccc}
 H & = & H & \Pi & + & \Xi & \Gamma & + & \delta \\
 (n \times q_1) & & (n \times q_1) & (q_1 \times q_1) & & (n \times q_2) & (q_2 \times q_1) & & (n \times q_1)
 \end{array}$$

Notations

- Σ_{ε} est la matrice de variance-covariance des variables manifestes,
- Σ_{δ} est la matrice de variance-covariance des VL endogènes (H),
- Φ est la matrice de variance-covariance des VL exogènes (Ξ)

Écriture matricielle de l'exemple

$$\theta = \begin{pmatrix} \theta_{11} & \theta_{12} & 0 & 0 & 0 & 0 \\ 0 & 0 & \theta_{23} & \theta_{24} & 0 & 0 \\ 0 & 0 & 0 & 0 & \theta_{35} & \theta_{36} \end{pmatrix}$$

$$\Lambda = \begin{pmatrix} 0 & 0 \\ \pi_{12} & 0 \\ \lambda_1 & \lambda_2 \end{pmatrix}$$

Hypothèses

- Chaque VM est rattachée à une unique VL (chaque colonne de θ contient un unique élément non nul et $q - 1$ zéros).
- Les VL sont unidimensionnelles.
- $\Xi_i \sim \mathcal{N}(0, \Phi)$
- $Z_i \parallel \varepsilon_i$
- $\Xi_i \parallel \delta_i$
- $\varepsilon_i \sim \mathcal{N}(0, \Sigma_\varepsilon)$
- $\delta_i \sim \mathcal{N}(0, \Sigma_\delta)$
- Σ_ε et Σ_δ sont diagonales

Distributions a priori conjuguées

Soit θ_k et Λ_k les éléments non nuls de la k ème colonne de θ .

$(\theta_k, \Sigma_{\varepsilon k})$ et $(\Lambda_k, \Sigma_{\delta k})$ sont distribués selon le modèle Normal-Gamma.

$$\begin{aligned}\theta_k | \Sigma_{\varepsilon k} &\sim \mathcal{N}(\theta_{0k}, \Sigma_{\varepsilon k} \Sigma_{\varepsilon 0k}) \\ (\Sigma_{\varepsilon k})^{-1} &\sim \mathcal{G}(\alpha_{0\varepsilon k}, \beta_{0\varepsilon k})\end{aligned}$$

$$\begin{aligned}\Lambda_k | \Sigma_{\delta k} &\sim \mathcal{N}(\Lambda_{0k}, \Sigma_{\delta k} \Sigma_{\delta 0k}) \\ (\Sigma_{\delta k})^{-1} &\sim \mathcal{G}(\alpha_{0\delta k}, \beta_{0\delta k})\end{aligned}$$

(Ξ, Φ) est distribué selon le modèle Normal-Wishart.

$$\begin{aligned}\Xi_i | \Phi &\sim \mathcal{N}(0, \Phi) \\ [\Phi] &\propto |\Phi|^{-\frac{1}{2}(q_2+1)}\end{aligned}$$

Distributions conditionnelles a posteriori des VL

Soit $\Theta = \{\theta, \Sigma_\varepsilon, \Lambda, \Sigma_\delta, \Phi\}$ l'ensemble des paramètres du modèle.

$$[Z|Y, \Theta] \propto \prod_{i=1}^n [Z_i|Y_i, \Theta] \propto \prod_{i=1}^n [Y_i|Z_i, \Theta] [Z_i|\Theta]$$

- $[Y_i|Z_i, \Theta] = [Y_i|Z_i, \theta, \Sigma_\varepsilon]$
- $[Z_i|\Theta] = [Z_i|\Lambda, \Sigma_\delta, \Phi] \sim \mathcal{N}(0, \Sigma_Z)$ où Σ_Z est la matrice de covariance des VL induite par le modèle structurel.

Résultat

$$Z_i|Y_i, \theta, \Sigma_\varepsilon, \Lambda, \Sigma_\delta, \Phi \sim \mathcal{N}(D\theta\Sigma_\varepsilon^{-1}Y_i, D)$$

où $D = \theta\Sigma_\varepsilon^{-1}\theta^t + \Sigma_Z^{-1}$.

Factorisation de la distribution conditionnelle a posteriori des paramètres

Soit $\Theta = \{\theta, \Sigma_\varepsilon, \Lambda, \Sigma_\delta, \Phi\}$ l'ensemble des paramètres du modèle.

$$\begin{aligned} [\Theta|Y, Z] &\propto [Y, Z|\Theta] [\Theta] \\ &\propto [Y|Z, \Theta] [Z|\Theta] [\Theta] \end{aligned}$$

- $[Y|Z, \Theta] = [Y|Z, \theta, \Sigma_\varepsilon]$
- $[Z|\Theta] = [Z|\Lambda, \Sigma_\delta, \Phi] = [H|\Xi, \Lambda, \Sigma_\delta] [\Xi|\Phi]$
- $[\Theta] = [\theta, \Sigma_\varepsilon] [\Lambda, \Sigma_\delta] [\Phi]$

$$[\Theta|Y, Z] \propto \underbrace{[Y|Z, \theta, \Sigma_\varepsilon] [\theta, \Sigma_\varepsilon]}_{[\theta, \Sigma_\varepsilon|Y, Z]} \underbrace{[H|\Xi, \Lambda, \Sigma_\delta] [\Lambda, \Sigma_\delta]}_{[\Lambda, \Sigma_\delta|Y, Z]} \underbrace{[\Xi|\Phi] [\Phi]}_{[\Phi|Z]}$$

Distributions conditionnelles a posteriori : résultats

Distribution conditionnelle a posteriori de $(\theta_k, \Sigma_{\varepsilon k})$

$$\begin{aligned}\theta_k | Y, Z, \Sigma_{\varepsilon} &\sim \mathcal{N}(D_k A_k, \Sigma_{\varepsilon k} D_k) \\ D_k &= (Z^t Z + \Sigma_{0\varepsilon}^{-1})^{-1} \\ A_k &= Z^t Y_k + \Sigma_{0\varepsilon}^{-1} \theta_{0k}\end{aligned}$$

Distribution conditionnelle a posteriori de $(\Lambda_k, \Sigma_{\delta k})$

$$\begin{aligned}\Lambda_k | Z, \Sigma_{\delta k} &\sim \mathcal{N}(D_k A_k, \Sigma_{\delta k} D_k) \\ D_k &= (Z^t Z + \Sigma_{0\delta}^{-1})^{-1} \\ A_k &= Z^t H_k + \Sigma_{0\delta}^{-1} \Lambda_{0k}\end{aligned}$$

Distribution a posteriori de Φ

$$\Phi | \Xi \sim \mathcal{IW}(\Xi^t \Xi, n)$$

Algorithme de Gibbs

Répéter les étapes suivantes jusqu'à convergence :

Tirage dans la distribution conditionnelle a posteriori des VL

$$[Z_i | Y_i, \theta, \Sigma_\varepsilon, \Lambda, \Sigma_\delta, \Phi]$$

Tirage dans la distribution conditionnelle a posteriori des paramètres

- $[\theta, \Sigma_\varepsilon | Y, Z]$
- $[\Lambda, \Sigma_\delta | Y, Z]$
- $[\Phi | Z]$

Conclusion

- Approche complémentaire des approches PLS-PM et LISREL, avec une portée plus générale.
- Mise en oeuvre aisée de l'algorithme de Gibbs (distributions conjuguées).
- Problèmes d'identifiabilité du modèle.

Perspectives

- Adaptation de l'algorithme de Gibbs pour données catégorielles (introduction d'étapes MH).
- Validation du modèle structurel par Posterior Predictive p-values.
- Question : Winbugs introduit-il des hypothèses d'indépendance conditionnelle lors de l'estimation ?