Modèle mixte

C. Ambroise

Laboratoire de Mathématiques et Modélisation d'Évry
UMR CNRS 8071

Plan

Le modèle mixte I

Ronald Fisher a introduit les effets aléatoires dans les modèles linéaires pour etudier les corrélations entre prédicteurs :
Un modèle linéaire mixte est un modèle linéaire tel que

$$\boldsymbol{y} = X\beta + \epsilon,\ \epsilon \sim \mathcal{N}(0, V),$$

avec V différent de la matrice identité.

Le modèle mixte II

Le vecteur bruit ϵ est considéré comme une combinaison linéaire de variable aléatoires structurales (qui structurent la matrice de variance covariance)

$$\varepsilon = \sum_{k=0}^{K} Z_k \boldsymbol{u}_k,$$

avec les Z_k des matrices connues de dimension $n \times q_k$ et les vecteurs $\boldsymbol{u}_k = \{u_{kl}\}_l = 1 \cdots q_k$ correspondant à des variables structurales. Notons $q_+ = \sum_k q_k$ la dimension du vecteur $\boldsymbol{u}$ concaténation des vecteur $\boldsymbol{u}_k$ et Z la matrice à n lignes q_+ colonnes concaténation des matrices Z_k.

Le modèle mixte III

On suppose que

$$\boldsymbol{u} \sim \mathcal{N}(0, \boldsymbol{\Sigma}_u = \sum_{m=1}^{M} \theta_m F_m),$$

où les F_m sont des matrices carré de dimension q_+.
Par définition on a $V = Z\boldsymbol{\Sigma}_u Z' = \sum_m ZF_m Z\theta_m = \sum_m V_m\theta_m$ car

$$var(\boldsymbol{y}) = var(Z\boldsymbol{u}) = Z\boldsymbol{\Sigma}_u Z'.$$

Les effets aléatoires apparaissent en définitive comme un moyen de structurer la matrice de variance-covariance $V = \sum_m V_m\theta_m$ des observations.

Le modèle mixte IV

Un bruit résiduel peut être incorporé ou non dans la structure générale de $\boldsymbol{u}$ en considérant $\boldsymbol{u}_0 = \boldsymbol{e}$ et $Z_0 = I_n$, la structure d'un modèle mixte est souvent présenté sous la forme

$$\boldsymbol{y} = X\beta + Z\boldsymbol{u} + \boldsymbol{e}. \tag{1}$$

où cette fois ci les Z_k sont des sous matrices et les u_k des sous vecteurs dont les indices varient entre 1 et K.

Le modèle mixte V

Dans cette dernière présentation nous avons

$$var(\boldsymbol{y}|\boldsymbol{u}) = R \text{ et } var(\boldsymbol{u}) = D \tag{2}$$

ce qui nous donne

$$\boldsymbol{y} \sim (X\beta, V = ZDZ' + R) \tag{3}$$

Exemple I

Considérons des notes à un examen de mathématique donné à quatre classes de chacun des quinze écoles de New York city. Notons

- β_t est l'effet fixe pour le sexe t,
- s_i est l'effet aléatoire pour l'école (s pour school), $i = 1, \cdots, 15$,
- c_{ij} est l'effet aléoire pour la classe dans l'école $j = 1, \cdots, 4$.

Exemple II

Dans cet exemple le vecteur β possède trois éléments (l'intercept, β_f, le coefficient pour les filles, et β_m le coefficient pour les garçons.) et nous avons donc deux sous vecteurs

- $\boldsymbol{u}_1$ de taille 15 pour les effets aléatoires des écoles,
- $\boldsymbol{u}_2$ de taille 4×15 pour les effets aléatoires des classes.

Exemple III

La matrice $Z = [Z_1 Z_2]$ aura donc $15 + 60$ colonnes et $n = 1200$ lignes. C'est une matrice binaire qui permet de spécifier pour chaque élève son école et sa classe.
L'équation du modèle est

$$E[y_{tijk|s_i,c_{ij}}] = \beta_t + s_i + c_{ij}.$$

$$\boldsymbol{u} = \begin{bmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \end{bmatrix}$$

et $D = var(\boldsymbol{u}) = \begin{bmatrix} var(\boldsymbol{u}_1) & cov(\boldsymbol{u}_1, \boldsymbol{u}_2{}') \\ cov(\boldsymbol{u}_2, \boldsymbol{u}_1{}') & var(\boldsymbol{u}_2) \end{bmatrix} = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix}$

Exemple IV

Dans cet exemple, il semble raisonnable de choisir les matrices D_{21} et D_{12} avec des entrées nulles. Elles représentent les covariances entre les écoles et les classes.
Si on prends des covariances nulles entres écoles et entres classes, ainsi qu'une même variance pour toutes les écoles (σ_s et toutes les classes (σ_c) alors notre modèle de variance s'écrit :

$$V = \sigma_s^2 Z_1 Z_1' + \sigma_c^2 Z_2 Z_2' + R.$$

En prenant $R = \sigma^2 I_n$, il vient

$$var(y_{tijk}) = \sigma_s^2 + \sigma_c^2 + \sigma^2.$$

Plan

Maximum de vraisemblance restreinte (RMLE) I

L'idée de la méthode consiste à maximiser la vraisemblance de combinaisons linéaire des éléments de **y** de manière à ce que ces combinaisons ne contiennent pas d'effets fixes : $\boldsymbol{k}'X = 0$. Le maximum de combinaisons $(n - rank(X))$ sont utilisées. Les vecteurs $\boldsymbol{k}$ sont donc des vecteurs qui forment une base de l'espace orthogonal à l'espace engendré par les colonnes de X.

Maximum de vraisemblance restreinte (RMLE) II

Comment trouver ces vecteur ?
Ces vecteurs sont regroupés dans la matrice
$K = [\boldsymbol{k}_1 \boldsymbol{k}_2 \cdots \boldsymbol{k}_{n-r_X}]$,
En prenant un modèle gaussien, il vient que

$$\boldsymbol{K}' y \sim \mathcal{N}(0, \boldsymbol{K}' VK). \tag{4}$$

Maximum de vraisemblance restreinte (RMLE) III

On résout les équations du maximum de vraisemblance et l'annulation du vecteur gradient amène

$$P = K(K'VK)^{-1}K' \tag{5}$$

qui contient la matrice V à estimer