Modèles mixtes : vocabulaire et panorama des méthodes d'estimation

Christian Lavergne

IS2 ; Inria Rhône-Alpes

Journée didactique IS2 sur les modèles mixtes
8 mars 2002

Le modèle linéaire LM

$$Y = X\beta + \epsilon$$

- Y vecteur des observations de $\mathbb{R}^n$

- β paramètre inconnu de $\mathbb{R}^p$

- ϵ vecteur des erreurs

- X matrice d'expérience connue.

* Classique et connue de tous :

  - la régression,
  - l'analyse de la variance

* Hypothèses fortes :

  - la forme linéaire,
  - la lois des observations,
  - l'homoscédasticité : la structure de covariance ne dépend pas de variables explicatives.

Repose sur l'hypothèse d'erreur gaussienne et de structure de covariance $\sigma_\epsilon^2\Gamma$ avec Γ connue.

Le cas le plus répandu étant $\Gamma = \mathrm{Id}_n$.

La vraisemblance

$$f(\beta,\theta;y) = \frac{1}{(2\pi\sigma^2)^{N/2}}|\Gamma|^{-1/2}\exp\{-\frac{1}{2\sigma^2}(y-X\beta)'\Gamma^{-1}(y-X\beta)\}$$

et

$$l(\beta,\sigma^2) = -2\log(f(\beta,\theta;y))$$

Estimateur maximun de vraisemblance

$$\widehat{\beta}_{ML} = \arg\min_{\beta}((y-X\beta)'\Gamma^{-1}(y-X\beta))$$

$$\widehat{\sigma}^2_{ML} = \arg\min_{\sigma^2}(\frac{1}{2\sigma^2}(y-X\beta)'\Gamma^{-1}(y-X\beta)+\frac{n}{2}\log(\sigma^2))$$

Fonction score

$$U_\beta = \frac{\partial l(\beta,\sigma^2)}{\partial \beta}\Big|_{\text{indépendant de } \sigma^2} = -2\,X'\Gamma^{-1}(y - X\beta)\,\Big|_{\perp \sigma^2} \tag{1}$$

donc

$$\widehat{\beta}_{ML} = (X'\Gamma^{-1}X)^{-1}X'\Gamma^{-1}Y$$

$$U_{\sigma^2}(\widehat{\beta}_{ML}) = \frac{\partial l(\beta,\sigma^2)}{\partial \sigma^2}\Big|_{\beta=\widehat{\beta}_{ML}} = -\frac{1}{2\sigma^4}(y - X\beta)'\Gamma^{-1}(y - X\beta) + \frac{n}{2\sigma^2}\Big|_{\beta=\widehat{\beta}_{ML}} \tag{2}$$

donc

$$\widehat{\sigma}^2_{ML} = \frac{1}{n}||Y - X\widehat{\beta}_{ML}||^2_{\Gamma^{-1}}$$

Procédure LM de Splus avec $\Gamma = \mathrm{Id}_n$ et $\widehat{\sigma}^2 = \frac{n}{n-p}\widehat{\sigma}^2_{ML}$

Interprétation géométrique et une propriété : W. Kruskal 1968

$$\widehat{\beta} = (X'\Gamma^{-1}X)^{-1}X'\Gamma^{-1}Y = (X'X)^{-1}X'Y$$

$$\Updownarrow$$

l'espace $\mathcal{X}$ est invariant par l'opérateur Γ ($\Gamma X = X$)

$$\widehat{\sigma}^2 = \frac{1}{n}||Y - X\widehat{\beta}_{ML}||^2_{\Gamma^{-1}} = \frac{1}{n}||Y - X\widehat{\beta}_{\mathrm{Id}}||^2$$

$$\Updownarrow$$

les espaces $\mathcal{X}$ et $\mathcal{X}^{\perp}$ sont invariants par l'opérateur Γ

Plusieurs extensions ?

Régression non-linéaire (NLM) $X\beta$ est remplacé par $g(\beta)$, où g est connue

Régression non-paramétrique : le modèle additif (AM)

$$y_i = a + s^1(x_i^1) + s^2(x_i^2) + ... + \epsilon_i$$

Les fonctions s^j seront estimés par des techniques non paramétriques (exemple spline)

Le modèle linéaire généralisé (GLM)

Le modèle linéaire mixte (L2M)

Retour sur le GLM

Les observations à expliquer peuvent être :

Discrètes :

- De type Présence/Absence : 0/1

  Modélisation du risque d'une survenue d'une défaillance, d'une pathologie.
  Les données à expliquer sont du type porteur / non porteur ; contrôlées par des facteurs de risque.

- Des effectifs : des entiers

  Prédiction du taux d'exhautivité d'un registre d'handicap, de la taille d'une population porteuse d'un virus, de la taille d'une population animalière ...
  Nombres de défaillances

Continues , mais où les hypothèses de loi gaussienne sont rejetées

- observations disymétriques
- observations où la variance est une fonction de la moyenne.
  Exemple : Variance = Moyenne^2

Fiabilité du logiciel : modélisation du temps inter défaillance.

Évaluation d'algorithmes évolutionnaires : modélisation de la valeur atteinte avec un nombre fixé d'itération.

Le modèle linéaire généralisé

Trois hypothèses permettent de caractériser un **GLM** :

la distribution de la variable à expliquer

$$f(y_i, \theta_i) = \exp\left(\frac{y_i\theta_i - b(\theta_i)}{a(\phi)} + c(y_i, \phi)\right),$$

θ_i est un paramètre *canonique*
ϕ un paramètre de *dispersion* > 0.
Les fonctions a, b et c sont spécifiques à chaque distribution.

Exemple : la loi Gamma

$$\forall y \in R_+ \,, \quad G(\alpha, \lambda)(y) = \frac{\lambda^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\lambda y} \,, \quad \lambda > 0, \alpha > 0 \,.$$

$$\theta = \frac{\lambda}{\alpha} \,, \quad b(\theta) = \log(\theta) \,, \quad \phi = \frac{1}{\alpha} \,, \quad a(\phi) = -\phi \,.$$

le prédicteur linéaire l'expression de la linéarité mettant en jeu les variables explicatives

$$\eta = X\beta,$$

la fonction de lien la fonction qui relie le prédicteur à l'espérance mathématique de la variable à expliquer ($\mu = b'(\theta)$)

(*fonction de lien* g).

$$X\beta = g(b'(\theta)) \ .$$

Si $g = b'^{-1}$, g est la fonction de lien canonique

Procédure GLM de Splus

La log-vraisemblance $\mathcal{L}$

$$L(\beta; y) = \sum_{i=1}^{N} \left[\frac{y_i \theta_i - b(\theta_i)}{a(\phi)} + c(y_i, \phi) \right] = \sum_{i=1}^{N} L_i(\beta_i; y_i) \ .$$

où $\theta_i = (g \circ b')^{-1} (X\beta)_i$

Les dérivées de vraisemblance pour $i \in \{1, ..., N\}$, $j \in \{1, ..., d\}$:

$$\frac{\partial L_i}{\partial \beta_j} = \frac{\partial \eta_i}{\partial \beta_j} \frac{\partial \mu_i}{\partial \eta_i} \frac{\partial \theta_i}{\partial \mu_i} \frac{\partial L_i}{\partial \theta_i} = X_{ij} \frac{1}{g'(\mu_i)} \frac{1}{b''(\theta_i)} \frac{y_i - \mu_i}{a(\phi)} \ ,$$

$$\frac{\partial L}{\partial \beta_j} = \sum_{i=1}^{N} X_{ij} \frac{1}{g'(\mu_i)^2 Var(Y_i)} g'(\mu_i) \left(y_i - \mu_i \right) \ .$$

Estimateur maximun de vraisemblance

et son équation normale :

$$X'W_\beta^{-1}\frac{\partial\eta}{\partial\mu}(y-\mu)=0$$

avec

$$W_\beta = (Var(Y_i)g'(\mu_i)^2\delta_{ij})_{i,j=1,\ldots,N}$$

L'estimation du paramètre inconnu β par maximum de vraisemblance se fait à l'aide d'un algorithme itératif.

$$\begin{aligned}\beta^{[t+1]} &= \beta^{[t]} - \left(E\left[\left\{\frac{\partial^2 L}{\partial\beta\partial\beta'}\right\}\right]^{[t]}\right)^{-1}\frac{\partial L}{\partial\beta}^{[t]}\\ &= \beta^{[t]} + \ldots\\ &\quad \left(X'W_{\beta^{[t]}}^{-1}X\right)^{-1}X'W_{\beta^{[t]}}^{-1}\frac{d\eta}{d\mu}^{[t]}\left(Y-\mu^{[t]}\right)\end{aligned}$$

Le modèle linéaire mixte

Pourquoi ? Quelles situations cherche-t-on à modéliser ?

⋆ Tenir compte de mesures répétées pour un même individu alors que l'on ne cherche pas à comparer ces mêmes individus.

⋆ On modèlise un "concept", une "méthode" et on observe des individus soumis à ce concept.

⋆ Modéliser un effet aléatoire caché.

Exemples :

• En science de l'éducation, une classe "scolaire" applique une méthode d'éducation.

• En génétique animale, on sélectionne des races en observant des représentants de la race.

• En fiabilité, la qualité de la correction ou de la réparation est un effet aléatoire non observable.

• Mesures de croissance (sur des arbres) ; donc répétées dans le temps. On cherche à comprendre le phénomène de croissance.
C'est l'analyse de données longitudinales.

Le modèle linéaire gaussien à structure de covariance paramétrée

$$Y = X\beta + \epsilon$$

- $\epsilon \sim \mathcal{N}_{\mathbb{R}^N}(0, \Gamma_\theta)$; avec θ paramètre inconnu de $\mathbb{R}^K$

La vraisemblance

$$f(\beta, \theta; y) = \frac{1}{(2\pi)^{N/2}} |\Gamma_\theta|^{-1/2} \exp\{-\frac{1}{2}(y - X\beta)' \Gamma_\theta^{-1} (y - X\beta)\}$$

et

$$l(\beta, \theta) = -2 \log(f(\beta, \theta; y))$$

Estimateur maximun de vraisemblance

$$\widehat{\beta}_{ML} = \arg\min_{\beta}((y - X\beta)'\Gamma_{\theta}^{-1}(y - X\beta))$$

$$\widehat{\theta}_{ML} = \arg\min_{\theta}((y - X\beta)'\Gamma_{\theta}^{-1}(y - X\beta) + \log(|\Gamma_{\theta}|))$$

Fonction score

$$\begin{aligned} U_{\beta}(\widehat{\theta}_{ML}) &= \left.\frac{\partial l(\beta,\theta)}{\partial \beta}\right|_{\theta=\widehat{\theta}_{ML}} \\ &= -2\,X'\Gamma_{\theta}^{-1}(y - X\beta)\,\Big|_{\theta=\widehat{\theta}_{ML}} \end{aligned} \tag{3}$$

donc $\widehat{\beta}_{ML} = (X'\Gamma_{\widehat{\theta}_{ML}}^{-1}X)^{-1}X'\Gamma_{\widehat{\theta}_{ML}}^{-1}y.$

$$
\begin{aligned}
U_\theta^{ML}(\widehat{\beta}_{ML}) &= \left.\frac{\partial l(\beta,\theta)}{\partial\theta}\right|_{\beta=\widehat{\beta}_{ML}} \\
&= \left. -(y-X\beta)'\Gamma_\theta^{-1}\frac{\partial\Gamma_\theta}{\partial\theta}\Gamma_\theta^{-1}(y-X\beta) + \underline{\mathrm{tr}}(\Gamma_\theta^{-1}\frac{\partial\Gamma_\theta}{\partial\theta})\right|_{\beta=\widehat{\beta}_{ML}} \\
&= -y'P_\theta\frac{\partial\Gamma_\theta}{\partial\theta}P_\theta\, y + \underline{\mathrm{tr}}(\Gamma_\theta^{-1}\frac{\partial\Gamma_\theta}{\partial\theta}) \qquad (4)
\end{aligned}
$$

où

- $P_\theta = \Gamma_\theta^{-1}(\mathrm{Id} - X(X'\Gamma_\theta^{-1}X)^{-1}X'\Gamma_\theta^{-1})$,

- $u'A_\theta\frac{\partial\Gamma_\theta}{\partial\theta}A_\theta\, u =$ est le vecteur de composantes $[u'A_\theta\frac{\partial\Gamma_\theta}{\partial\theta_j}A_\theta\, u]_{j=1,\dots,K}$

- resp. $\underline{\mathrm{tr}}(A_\theta\frac{\partial\Gamma_\theta}{\partial\theta}) = [\mathrm{tr}(A_\theta\frac{\partial\Gamma_\theta}{\partial\theta_j})]_{j=1,\dots,K}$

rappel : $\frac{\partial\Gamma_\theta^{-1}}{\partial\theta_j} = -\Gamma_\theta^{-1}\frac{\partial\Gamma_\theta}{\partial\theta_j}\Gamma_\theta^{-1}$ et $\frac{\partial\log(|\Gamma_\theta|)}{\partial\theta_j} = \mathrm{tr}(\Gamma_\theta^{-1}\frac{\partial\Gamma_\theta}{\partial\theta_j})$

Matrice d'information

Posons $\gamma = (\beta, \theta)$ et

$$\mathcal{I}_{\beta,\theta} = E_{\beta,\theta}\left[-\frac{\partial^2 \log f(\beta,\theta;y)}{\partial\gamma\partial\gamma'}\right]$$

alors

$$\mathcal{I}_{\beta,\theta} = E_{\beta,\theta}\left[\frac{1}{2}\frac{l(\beta,\theta)}{\partial\gamma\partial\gamma'}\right] = E_{\beta,\theta}\left[\frac{1}{2}\begin{pmatrix} \dfrac{\partial U_\beta}{\partial\beta'} & \dfrac{\partial U_\beta}{\partial\theta'} \\ \\ \dfrac{\partial U_\theta^{ML}}{\partial\beta'} & \dfrac{\partial U_\theta^{ML}}{\partial\theta'} \end{pmatrix}\right].$$

On vérifie aisément que $\dfrac{\partial U_\beta}{\partial\beta'} = 2\,X'\Gamma_\theta^{-1}X$ et que $E_{\beta,\theta}\left[\dfrac{\partial U_\beta}{\partial\theta'}\right] = 0$.

D'autre part :

$$\left[\frac{\partial U_\theta^{ML}}{\partial \theta'}\right]_{ij} = \mathrm{tr}(\Gamma_\theta^{-1}(\frac{\partial^2 \Gamma_\theta}{\partial\theta_i\partial\theta_j} - \frac{\partial \Gamma_\theta}{\partial\theta_i}\Gamma_\theta^{-1}\frac{\partial \Gamma_\theta}{\partial\theta_j}))$$
$$-(y - X\beta)'\Gamma_\theta^{-1}(\frac{\partial^2 \Gamma_\theta}{\partial\theta_i\partial\theta_j} - 2\,\frac{\partial \Gamma_\theta}{\partial\theta_i}\Gamma_\theta^{-1}\frac{\partial \Gamma_\theta}{\partial\theta_j})\Gamma_\theta^{-1}(y - X\beta)$$

Donc

$$\mathcal{I}_{\beta,\theta} = \begin{pmatrix} \mathcal{I}_\beta = X'\Gamma_\theta^{-1}X & 0 \\ 0 & \mathcal{I}_\theta^{ML} = \left[\frac{1}{2}\mathrm{tr}(\Gamma_\theta^{-1}\frac{\partial \Gamma_\theta}{\partial\theta_i}\Gamma_\theta^{-1}\frac{\partial \Gamma_\theta}{\partial\theta_j})\right]_{i,j=1,\ldots,K} \end{pmatrix}$$

Estimation REML

La vraisemblance restreinte : c'est la vraisemblance marginale après intégration sur β

$$f_{RE}(\theta; y) = \frac{1}{(2\pi)^{(N-p)/2}} |\Gamma_\theta|^{-1/2} |X'\Gamma_\theta^{-1} X|^{-1/2} \exp\{-\frac{1}{2} y' P_\theta\, y\}$$

et

$$l_{RE}(\theta) = -2 \log(f_{RE}(\theta; y))$$

Estimateur maximun de vraisemblance restreinte

$$\widehat{\theta}_{REML} = \arg\min_\theta (y' P_\theta\, y + \log(|\Gamma_\theta|) + \log(|X'\Gamma_\theta^{-1} X|)$$

Fonction score

$$\begin{aligned} U_\theta^{REML} &= \frac{\partial l_{RE}(\theta)}{\partial \theta} \\ &= -y' P_\theta \frac{\partial \Gamma_\theta}{\partial \theta} P_\theta \, y + \underline{\mathrm{tr}}(\Gamma_\theta^{-1} \frac{\partial \Gamma_\theta}{\partial \theta}) \\ & \quad - \underline{\mathrm{tr}}((X'\Gamma_\theta^{-1} X)^{-1} X' \Gamma_\theta^{-1} \frac{\partial \Gamma_\theta}{\partial \theta} \Gamma_\theta^{-1} X) \\ &= -y' P_\theta \frac{\partial \Gamma_\theta}{\partial \theta} P_\theta \, y + \underline{\mathrm{tr}}(P_\theta \frac{\partial \Gamma_\theta}{\partial \theta}) \end{aligned} \qquad (5)$$

et $\widehat{\beta}_{REML}$ est solution de l'équation $U_\beta(\widehat{\theta}_{REML}) = 0$.

Matrice d'information associée au vecteur y Posons

$$\mathcal{I}_\theta^{REML} = E_\theta \left[-\frac{\partial^2 \log f_{RE}(\theta; y)}{\partial\theta\partial\theta'} \right]$$

alors

$$\mathcal{I}_\theta^{REML} = E_\theta \left[\frac{1}{2} \frac{l_{RE}(\theta)}{\partial\theta\partial\theta'} \right] = E_\theta \left[\frac{1}{2} \frac{\partial U_\theta^{REML}}{\partial\theta'} \right].$$

Un calcul similaire au cas précédent

$$\mathcal{I}_\theta^{REML} = \left[\frac{1}{2} \mathrm{tr}(P_\theta \frac{\partial\Gamma_\theta}{\partial\theta_i} P_\theta \frac{\partial\Gamma_\theta}{\partial\theta_j}) \right]_{i,j=1,\ldots,K}$$

Remarque

• En résumé :

$\widehat{\theta}_{ML}$ est donc solution de la minimisation du critère :

$$\text{MixteML} = y' P_\theta\, y + \log(|\Gamma_\theta|)\ ;$$

$\widehat{\theta}_{REML}$ est solution de la minimisation du critère :

$$\text{MixteREML} = \text{MixteML} + \log(|X'\Gamma_\theta^{-1} X|)$$

et $\widehat{\beta}_{.} = (X'\Gamma_{\widehat{\theta}_{.}}^{-1} X)^{-1} X'\Gamma_{\widehat{\theta}_{.}}^{-1} y.$

• Remarquant que $E_\theta(P_\theta\, y) = 0$ et $P_\theta \Gamma_\theta P_\theta = P_\theta$ alors

$$E_\theta(y' P_\theta \frac{\partial \Gamma_\theta}{\partial \theta} P_\theta\, y) = \underline{\text{tr}}(P_\theta \frac{\partial \Gamma_\theta}{\partial \theta}).$$

La fonction score U_θ^{REML} est donc une statistique centrée contrairement à la fonction score U_θ^{ML}.

Le modèle linéaire mixte : L2M

$$Y = X\beta + \boldsymbol{Z}U + \epsilon \ = \ X\beta + \sum_{k=1}^{K} Z_k U_k + \epsilon$$

• $\epsilon \sim \mathcal{N}_{I\!R^N}(0, R = \theta_0 V_0)$, V_0 est une matrice connue.

• $U_k \sim \mathcal{N}_{I\!R^{q_k}}(0, \theta_k G_k)$ pour tout $k = 1, \ldots, K$; les G_k sont des matrices connues.

$U_1, U_2, \ldots, U_k$ sont les effets aléatoires non observés ; ils sont indépendants entre eux et indépendants de ϵ.

• $U \sim \mathcal{N}_{I\!R^q}(0, G_\theta)$ où $q = \sum_{k=1}^{K} q_k$

$$\text{Var}(Y) = \Gamma_\theta = R + \boldsymbol{Z} G_\theta \boldsymbol{Z}' = \sum_{k=0}^{K} \theta_k V_k \text{ où } V_k = Z_k G_k Z_k'$$

les θ_k ($\in I\!R^+$) sont appelés les composantes de la variance.

Simplification des fonctions scores U_θ^{ML} et U_θ^{REML}

Dans le cas des modèles mixtes on a : $\frac{\partial \Gamma_\theta}{\partial \theta_k} = V_k$:

$$\begin{array}{rcrcl} U_\theta^{ML} & : & y' P_\theta V_k P_\theta \, y & = & \text{tr}(\Gamma_\theta^{-1} V_k) \text{ pour } k = 0, \ldots, K \\ U_\theta^{REML} & : & y' P_\theta V_k P_\theta \, y & = & \text{tr}(P_\theta V_k) \text{ pour } k = 0, \ldots, K \end{array}$$

Comme pour $k = 0, \ldots, K$

$$\begin{array}{rrcl} & \text{tr}(\Gamma_\theta^{-1} V_k) & = & \sum_{j=1}^{K} \text{tr}(\Gamma_\theta^{-1} V_k \Gamma_\theta^{-1} V_j) \theta_j \\ \text{et} & \text{tr}(P_\theta V_k) & = & \sum_{j=1}^{K} \text{tr}(P_\theta V_k P_\theta V_j) \theta_j \end{array}$$

On obtient ainsi 2 systèmes linéaires identiques :

$2\mathcal{I}_\theta \, \theta = F_\theta$ où F_θ est le vecteur $\{y' P_\theta V_k P_\theta \, y\}$

L'algorithme EM

Principe : étant donné un vecteur aléatoire observé Y et un vecteur aléatoire non observé (ici les effets aléatoires) U, on cherche des statistiques exhaustives, fonction des données complètes $x = (y', u')'$ qui permettent de réaliser l'estimation des paramètres inconnus γ : notées $t(x)$.

On obtient un algorithme qui à chaque itération $[t]$ pour une valeur courante du paramètre $\gamma^{[t]}$ se décompose en 2 étapes :

• étape E : calculer l'espérance conditionnelle de $t(x)$ sachant les données observées y et la valeur $\gamma^{[t]}$.

• étape M : maximiser la vraisemblance des données complètes en remplaçant les statistiques exhaustives par l'espérance conditionnelle et obtenir ainsi $\gamma^{[t+1]}$.

Le modèle linéaire mixte : L2M

$$Y = X\beta + \sum_{k=1}^{K} Z_k U_k + \epsilon \,;\, \epsilon \sim \mathcal{N}_{\mathbb{R}^N}(0, \theta_0 V_0),\ U_k \sim \mathcal{N}_{\mathbb{R}^{q_k}}(0, \theta_k G_k).$$

Donc si les U_k étaient observées, les estimations des paramètres seraient obtenues par :

$$\widehat{\theta_k} = \frac{U_k G_k^{-1} U_k'}{q_k} \text{ et } \widehat{\beta} = (X' V_0^{-1} X)^{-1} X' V_0^{-1} (y - \sum_{k=1}^{K} Z_k U_k)$$

On vérifie alors aisément que :

$$E(U_k G_k^{-1} U_k' | y) = \theta_k^2 (y - X\beta)' \Gamma_\theta^{-1} V_j \Gamma_\theta^{-1} (y - X\beta) + q_k \theta - \theta^2 \mathrm{tr}(\Gamma_\theta^{-1} V_j)$$

$$E(y - \sum_{k=1}^{K} Z_k U_k | y) = X\beta + \theta_0 V_0 \Gamma_\theta^{-1} (y - X\beta)$$

D'où l'algorithme EM pour ML:

$$\left\{\begin{array}{lcl} q_k\theta_k^{[t+1]} & = & \theta_k^{2[t]}(y - X\beta^{[t]})'\Gamma_{\theta^{[t]}}^{-1}V_j\Gamma_{\theta^{[t]}}^{-1}(y - X\beta^{[t]}) \\ & & +q_k\theta^{[t]} - \theta^{2[t]}\mathrm{tr}(\Gamma_{\theta^{[t]}}^{-1}V_j) \\ X\beta^{[t+1]} & = & (X'V_0^{-1}X)^{-1}X'V_0^{-1}[X\beta^{[t]} + \theta_0^{[t]}V_0\Gamma_{\theta^{[t]}}^{-1}(y - X\beta^{[t]})] \end{array}\right.$$

La 1^{re} partie de l'algorithme peut aussi s'écrire :

$$\left\{\begin{array}{lcl} q_k\theta_k^{[t+1]} & = & \theta_k^{2[t]}y'P_{\theta^{[t]}}V_jP_{\theta^{[t]}}y + q_k\theta^{[t]} - \theta^{2[t]}\mathrm{tr}(\Gamma_{\theta^{[t]}}^{-1}V_j) \end{array}\right.$$

et en remplaçant $\Gamma_{\theta^{[t]}}^{-1}$ par $P_{\theta^{[t]}}$ on obtient l'algorithme EM pour REML.

autres terminologies :

• La méthode d'Henderson : équation et algorithme propre aux modèles mixtes

• La méthode MINQUE : approche géométrique.

• L'ANOVA à effet(s) aléatoire(s) : cas des plans équilibrés.

Les modèles d'ANOVA à un ou deux facteurs, avec effet aléatoire

- Les effets fixes possibles (paramètres inconnus) sont :

  μ : paramètre réel

  $\underline{a}$ paramètre de $I\!R^I$ de composantes a_i.

- Les effets aléatoires possibles sont :

  α : vecteur aléatoire de $I\!R^I$ centré, de variance $\sigma_1^2 \mathcal{I}d_I$.

  β : vecteur aléatoire de $I\!R^J$ centré, de variance $\sigma_2^2 \mathcal{I}d_J$.

  γ : vecteur aléatoire de $I\!R^C$ centré, de variance $\sigma_3^2 \mathcal{I}d_C$ de composantes γ_{ij}.

- ε vecteur aléatoire des erreurs de variance $\sigma^2 \mathcal{I}d_n$.

Modèle 1 : modèle d'analyse de la variance à un facteur aléatoire ;

$$y_{ik} = \mu + \alpha_i + e_{ik}$$

Modèle 2 : modèle d'analyse de la variance à deux facteurs aléatoires emboîtés ;

$$y_{ijk} = \mu + \alpha_i + \gamma_{ij} + e_{ijk}.$$

Modèle 3 : modèle d'analyse de la variance à deux facteurs aléatoires croisés, sans interaction ;

$$y_{ijk} = \mu + \alpha_i + \beta_j + e_{ijk},$$

Modèle 4 : modèle d'analyse de la variance à deux facteurs aléatoires croisés, avec interaction ;

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk},$$

Modèle 5 : modèle d'analyse de la variance à un facteur à effet fixe : ce modèle n'est évidemment pas un modèle à effet aléatoire ; on le mentionne ici pour garder l'analogie avec le modèle 1 ;

$$y_{ik} = a_i + e_{ik}$$

Modèle 6 : modèle mixte d'analyse de la variance à deux facteurs emboîtés ;

$$y_{ijk} = a_i + \gamma_{ij} + e_{ijk}.$$

Modèle 7 : modèle mixte d'analyse de la variance à deux facteurs croisés, sans interaction ;

$$y_{ijk} = a_i + \beta_j + e_{ijk},$$

Modèle 8 : modèle mixte d'analyse de la variance à deux facteurs croisés, avec interaction ;

$$y_{ijk} = a_i + \beta_j + \gamma_{ij} + e_{ijk},$$