Estimation jointe de modèles non linéaires à effets mixtes par vraisemblance pénalisée : application en pharmacocinétique

Edouard Ollier [1] & Vivian Viallon [2]

[1] *Laboratoire de Pharmacologie Toxicologie, CHU Saint-Etienne, F-42055, Saint-Etienne*
E-mail : ed.ollier@gmail.com
[2] *Université de Lyon, F-69622, Lyon, France ; Université Lyon 1, UMRESTTE, F-69373 Lyon ; IFSTTAR, UMRESTTE, F-69675 Bron E-mail : vivian.viallon@univ-lyon1.fr*

Résumé. En pharmacocinétique les données possèdent souvent une structure de groupe : par exemple, les groupes peuvent correspondre à différentes modalités de traitement. Un modèle non linéaire à effets mixtes peut être construit dans chacun de ces groupes, le travail de modélisation consistant alors à identifier les paramètres dont l'estimation varie significativement à travers les groupes.

Dans ce travail, nous nous intéressons à l'estimation jointe de modèles non-linéaires à effets mixtes par une méthode de vraisemblance pénalisée de type fused lasso. Nous étudierons deux types de problèmes : (i) l'estimation jointe des effets fixes et des variances des effets aléatoires; et (ii) l'estimation jointe des effets fixes dans le cas où des covariables doivent être sélectionnées. L'estimation des paramètres sera réalisée par un algorithme de type SAEM. Il sera évalué sur données simulées et appliqué sur les données d'un essai clinique étudiant l'interaction médicamenteuse entre un anticoagulant et un antibiotique.

Mots-clés. Modèles non linéaires à effets mixtes, fused, lasso, SAEM, estimation jointe, sélection de variables.

Abstract. Data in pharmacokinetic often have a group structure, e.g with groups corresponding to different treatment modalities. Each group can be described by a non linear mixed effects model and one goal of the modelisation step is generally to identify parameters whose estimation significantly vary across the groups.

In this work,we propose to jointly estimate multiple non linear mixed effect models by maximizing a log-likelihood penalized with a fused lasso penalty. Two problems will be studied: (i) the joint estimation of fixed effects and random effects variance and (ii) the joint estimation of fixed effects when covariates further need to be selected. Parameter estimation will be performed by a modified SAEM algorithm that solves the penalized maximum likelihood problem. It will be evaluated on simulated data and applied to real data from a clinical trial studying the drug-drug interaction between an anticoagulant and an antibiotic.

Keywords. Non linear mixed effect model, fused, lasso, SAEM, joint modeling, variable selection.

1 Introduction

Les Modèles Non Linéaires à Effets Mixtes (MNLEM) sont utilisés pour modéliser les données longitudinales dans de nombreux domaines comme la pharmacocinétique (PK). Dans certains cas, les données présentent une structure de groupe. Un des objectifs de l'analyse statistique consiste alors à déterminer quels paramètres varient significativement entre les différents groupes. On rencontre ce problème notamment en recherche clinique, où les groupes de patients correspondent à différentes modalités de traitement. Par exemple, on peut vouloir étudier l'existence d'une interaction médicamenteuse entre un médicament d'intérêt et une autre molécule. L'existence d'une telle interaction sera mise en évidence par la variation significative d'un paramètre du modèle entre le groupe ne recevant que le médicament d'intérêt et celui recevant les 2 molécules. De plus, le type de paramètre identifié nous donnera un renseignement sur le mécanisme de l'interaction et la différence d'estimation entre les groupes nous informera sur l'importance de celle-ci.

Dans la littérature, la détection de différences significatives entre les groupes repose généralement sur des tests statistiques. Samson et al (2000) ont analysé les performances du test de Wald et du rapport de vraisemblance à partir des estimations de l'algorithme SAEM (Kuhn et al., 2004). Une alternative consiste à estimer de manière conjointe les modèles dans les différents groupes (Oelker et al. 2012), en encourageant la similarité entre ceux-ci. Plus précisément, on peut chercher à résoudre un problème de régression pénalisée où la pénalité $\text{Pen}(\gamma^1, ..., \gamma^G)$ encourage la similarité entre les paramètres $\gamma^1, ..., \gamma^G$ des différents groupes (G : le nombre de groupes). C'est notamment le cas du fused lasso qui pénalise les différences entre les coefficients et les encourage donc à être égaux.

Les modèles *linéaires* à effets mixtes avec une pénalité l_1 sur les effets fixes ont déjà été étudiés. Bondell et al (2010) ont étudié la pénalisation l_1 des effets fixes et de la variance des effets aléatoires. Leur méthode d'estimation repose sur un algorithme EM dans lequel l'étape de maximisation résout des problèmes de régression pénalisés. A notre connaissance, aucun travail n'a étudié l'applications de *pénalité structurée* (fused lasso, group lasso) à la vraisemblance de modèles *non linéaires* à effets mixtes.

Dans le cadre de cet exposé nous étudierons deux types de problème. Dans un premier temps nous considérerons l'estimation jointe des effets fixes et des variances des effets aléatoires. Pour cela une pénalité de type fused lasso (sans pénalité l_1) sera appliquée au vecteur des effets fixes ainsi qu'à la matrice de concentration des effets aléatoires. Dans un second temps, nous étudierons l'estimation jointe des effets fixes dans le cas où des covariables sont disponibles et doivent être sélectionnées. Pour cela une pénalité de type fused lasso (sans pénalité l_1) sera appliquée au vecteur des effets fixes et une pénalité de type fused lasso (avec pénalité l_1) au vecteur des covariables.

2 Estimation jointe de MNLEM et vraisemblance pénalisée

Soit $y_{i,j}^g$ l'observation au temps $t_{i,j}^g$ ($j \in \{1, ..., n_i\}$) du i-ème patient ($i \in \{1, ..., N_g\}$) du g-ème groupe ($g \in \{1, ..., G\}$). On considère des modèles de la forme :

$$\begin{aligned}
&y_{i,j}^g = f(t_{i,j}^g, \phi_i^g) + h(t_{i,j}^g, \phi_i^g)\epsilon_{i,j}^g \\
&h = af + b \\
&\epsilon_{i,j}^g \sim \mathcal{N}(0,1) \text{ (iid)}
\end{aligned}$$

où f est une fonction non linéaire donnée. La fonction h correspond au modèle d'erreur, avec a et b deux réels à estimer. Soit ϕ_i^g le vecteur de dimension p des paramètres du i-ème sujet dans le groupe g. On suppose par ailleurs que le vecteur des paramètres individuels, ϕ_i^g, se décompose de la façon suivante :

$$\begin{aligned}
\phi_i^g &= \mu^g + A_i \beta^g + \omega_i^g \\
\omega_i^g &\sim \mathcal{N}(0, \Omega^g)
\end{aligned}$$

avec μ^g le vecteur de dimension p des effets fixes du groupe g, A_i^g la matrice de design individuel de dimension $p \times k$ et β^g le vecteur, de dimension k, des coefficients de régression associés aux covariables dans le groupe g. Le terme ω_i^g correspond au vecteur de dimension p des effets aléatoires du i-ième patient ($i \in \{1, ..., N_g\}$) du g-ième groupe. Enfin, Ω^g est la matrice de variance-covariance des effets aléatoires du groupe g.

Sous l'hypothèse que les groupes sont indépendants, l'estimation des paramètres par maximum de vraisemblance revient à résoudre le problème suivant :

$$\hat{\theta} = \underset{\theta}{\text{ArgMin}} \sum_{g=1}^{G} -2 \times \log(p(y^g; \theta^g))$$

où $p(y^g; \theta^g)$ est la vraisemblance des données du groupe g (y^g) par rapport à $\theta = (\theta^1, ..., \theta^G, a, b) = ((\mu^1, \beta^1, \Omega^1), ..., (\mu^G, \beta^G, \Omega^G), a, b)$:

$$p(y^g; \theta^g) = \int p(y^g, \phi^g; \theta^g) d\phi^g$$

avec $p(y^g, \phi^g; \theta^g)$ la vraisemblance des données (y^g, ϕ^g) du groupe g.

Le problème décrit ci-dessus revient à estimer les paramètres séparément pour chacun des groupes. Il se peut que certains groupes partagent des caractéristiques communes, et donc que certains paramètres varient peu ou pas entre ces groupes. En introduisant une pénalité de type fused lasso dans le problème de maximum de vraisemblance on encourage les paramètres dont l'estimation varie faiblement entre deux groupes à être égaux.

Nous utiliserons les pénalités suivantes pour les effets fixes :

$$P_{FIX}^{\lambda_F,0}(\mu^1,...,\mu^G) = \lambda_F \sum_{i=1}^{p} \sum_{(g_1,g_2)\in\mathcal{E}} |\mu_i^{g_1} - \mu_i^{g_2}|$$

$$P_{FIX}^{\lambda_F,\lambda_L}(\beta^1,...,\beta^G) = \lambda_F \sum_{i=1}^{k} \sum_{(g_1,g_2)\in\mathcal{E}} |\beta_i^{g_1} - \beta_i^{g_2}| + \lambda_L \sum_{g=1}^{G} \|\beta^g\|_1$$

où $\mathcal{E}$ correspond à l'ensemble des arêtes d'un graphe donné. Dans ce graphe, deux groupes sont connectés s'ils sont supposés avoir des estimations comparables. Cette pénalité encourage donc les paramètres à avoir les mêmes estimations entre 2 groupes connectés. De plus la pénalité $\|\beta^g\|_1$ permet de sélectionner les covariables influentes au sein de chaque groupe, en encourageant la nullité des composantes.

En ce qui concerne les variances des effets aléatoires, l'idéal serait de pénaliser la matrice de variance covariance des effets aléatoires mais le problème d'optimisation à résoudre lors de l'étape de maximisation de SAEM est alors non convexe. Sous l'hypothèse d'indépendance des paramètres (usuelle en pratique), une solution équivalente est de pénaliser l'inverse de la matrice de variance-covariance (la matrice de concentration). Le problème à résoudre lors de l'étape de maximisation est alors convexe. La pénalité considérée ici est donc la suivante :

$$P_{VAR}^{\lambda}(\Omega^{1^{-1}},...\Omega^{G^{-1}}) = \lambda \sum_{i=1}^{p} \sum_{(g_1,g_2)\in\mathcal{E}} |\Omega_{ii}^{g_1^{-1}} - \Omega_{ii}^{g_2^{-1}}|$$

Dans ce travail, nous étudierons 2 problèmes :

1. La sélection des effets fixes et de la variance des effets aléatoires :

$$\hat{\theta} = \underset{\theta}{\text{ArgMin}} \sum_{g=1}^{G} -2 \times \log(p(y^g;\theta^g)) + P_{FIX}^{\lambda_F,0}(\mu^1,...,\mu^G) + P_{VAR}^{\lambda}(\Omega^{1^{-1}},...\Omega^{G^{-1}})$$

2. La sélection des effets fixes dans le cas où des covariables doivent être sélectionnées :

$$\hat{\theta} = \underset{\theta}{\text{ArgMin}} \sum_{g=1}^{G} -2 \times \log(p(y^g;\theta^g)) + P_{FIX}^{\lambda_F,0}(\mu^1,...,\mu^G) + P_{FIX}^{\lambda_F,\lambda_L}(\beta^1,...,\beta^G)$$

3 Algorithme SAEM et vraisemblance pénalisée

Dans le cas des modèles linéaires mixtes, Bondell et al. (2010) ont développé un algorithme EM dans lequel seule l'étape de maximisation a été modifiée. Cette dernière correspondait à un problème de régression pénalisée. L'algorithme SAEM (Kuhn et al. 2004) développé dans ce travail reprend le même principe. Soit Q_k l'espérance conditionelle de la vraisemblance complète : $Q_k(\theta) = E(\log[p(y,\phi;\theta)]|y,\theta_{k-1})$. À l'itération k, l'étape de maximisation correspond à :

1. Mise à jour des effets fixes :

$$\left(\mu^1_{k+1},...,\mu^G_{k+1}\right) = \underset{\mu^1,...,\mu^G,\beta^1,...,\beta^G}{\text{ArgMax}} \begin{cases} \sum_{g=1}^G Q_k(\mu^g,\beta^g,\Omega^g_k,a_k,b_k) \\ \quad + \quad P_F^{\lambda_F,0}(\mu^1,...,\mu^G) \\ \quad + \quad P_{FIX}^{\lambda_F,\lambda_L}(\beta^1,...,\beta^G) \end{cases}$$

2. Mise à jour des variances des effets aléatoires :

$$(\Omega^1_{k+1},...\Omega^G_{k+1}) = \underset{\Omega^1,...\Omega^G}{\text{ArgMax}} \begin{cases} \sum_{g=1}^G Q_k(\mu^g_{k+1},\beta^g_{k+1}\Omega^g,a_k,b_k) \\ \quad + \quad P^{\lambda}_{VAR}(\Omega^{1^{-1}},...,\Omega^{G^{-1}}) \end{cases}$$

3. Mise à jour des paramètres du modèle d'erreur : inchangé car non pénalisé .

Le problème de l'étape 2 à été précédement étudié par Danaher et al (2011) dans le cadre de l'estimation jointe de modèles graphiques gaussiens. La mise à jour peut donc être effectuée avec l'algorithme développé dans cet article.

4 Sélection du modèle optimal

L'estimateur proposé dans ce travail dépend des paramètres de sparsité λ_F, λ_L et λ. Ceux-ci sont évalués sur une grille ad hoc. Les valeurs optimales sont sélectionées grâce au critère BIC (Delattre et al. 2012) en adaptant la procédure lasso-OLS hybride (sections 2.9, 2.10 de Buhlmann, Van de Geer 2011).

5 Evaluation

L'algorithme est évalué sur des données simulées et appliqué sur un jeu de données réelles provenant d'un essai clinique étudiant l'interaction médicamenteuse entre la Clarithromycine (antibiotique) et le Dabigatran (anticoagulant oral) (Delavenne et al., 2013).

Bibliographie

[1] Bondell, H. D., Krishna, A., et Ghosh, S. K. (2010). *Joint Variable Selection for Fixed and Random Effects in Linear Mixed Effects Models. Biometrics*, 66(4), 10691077.
[2] Kuhn, E., et Lavielle, M. (2004). *Coupling a stochastic approximation version of EM with an MCMC procedure*. ESAIM: Probability and Statistics, 8(1), 115131.
[3] Delavenne, X., Ollier, E., Basset, T., Bertoletti, L., Accassat, S., Garcin, A., et al. (2013). *A semi-mechanistic absorption model to evaluate drug-drug interaction with dabigatran: application with clarithromycin*. British Journal of Clinical Pharmacology, 76(1), 107113. doi:10.1111/bcp.12055
[4] Samson, A., Lavielle, M., et Mentré, F. (2000). *The SAEM algorithm for group comparison tests in longitudinal data analysis based on non-linear mixed-effects model* . Statist. Med. 00, 16.
[5] Danaher, P., Wang, P., et Witten, D. M. (2013). The joint graphical lasso for inverse covariance estimation across multiple classes. arXiv.org.
[6] Buhlmann, P. L., Van de Geer, S. A. (2011). Statistics dor high-dimensional data. Springer.
[7] Delattre, M., Lavielle, M., et Porsat, M. A. (2012). BIC selection procedures in mixed effects models.
[8] Oelker, M.-R., Gertheiss, J., Tutz, G. (2012). Regularization and Model Selection with Categorial Predictors and Effect Modifiers in Generalized Linear Models.