Statistical inference for discretely observed stochastic differential equations with mixed effects

Catherine Larédo (1)
Joint work with Maud Delattre (2) & Valentine Genon-Catalot (3)

(1) Laboratoire MaIAGE, I.N.R.A. and LPMA, Université Denis Diderot, CNRS-UMR 7599,
(2) AgroParisTech, France ,
(3) UMR CNRS 8145, Laboratoire MAP5, Université Paris Descartes, Sorbonne Paris Cité

SAPS XI, July 17-21 2017, Peterhof, Russia

Based on two papers : Delattre, Genon-Catalot & Larédo, 2017 (a),(b)

Mixed effects models

- Longitudinal data widely collected in clinical trials, epidemiology, pharmacokinetic pharmacodynamics experiments on N individuals.
- Interest may focus on population effects among individuals and on individual specific behaviour.
- Random effects $\Rightarrow$ accomodate among inter-individual variability.
- while the same structural model rules the individual dynamics.
- Structural model here: Stochastic Differential Equations $\Rightarrow$ SDE with mixed effects (SDEME).

Main issue for general mixed effects models

- Estimation of parameters in the distribution of random effects.
- Difficult in practice due to the intractable likelihood function.
- Development of computationally intensive numerical methods.
- Large computation times due to iterative settings.

Diffusion process with mixed effects

$$\begin{cases} dX(t) &= b(X(t), \Phi)dt + \sigma(X(t), \Psi)dW(t), \\ X(0) &= x, \\ x \in \mathbb{R}\,,\ t \in [0, T]. \end{cases} \tag{1}$$

Φ, Ψ random variables, independent of $W(.)$,

$W(.)$: Wiener process.

AIM: Estimation of unknown parameters in the distribution of Φ, Ψ.

MIXED EFFECTS: Presence of both fixed effects and random effects.

Remark: Classical problem if $\Phi = \varphi, \Psi = \psi$ are non random, $T \to \infty$.

$\star$ Continuous observations: Kutoyants (1984), Lipster & Shiryaev (2001)

$\star$ Discrete observations: estimation of (φ, ψ): e.g. Kessler(1997) .

Repeated observations of discretized processes

Discrete observations on $[0, T]$ of N *i.i.d.* processes $(X_i(t))$

$$\begin{cases} dX_i(t) & = & b(X_i(t), \Phi_i)dt + \sigma(X_i(t), \Psi_i)dW_i(t), \\ X_i(0) & = & x, \\ x \in \mathbb{R}\ , \ \ t \in [0, T]\ , \ \ i = 1, \ldots, N, \end{cases}$$

- $(\Phi_i, \Psi_i), i = 1, \ldots, N$: *i.i.d.* random variables.
- $W_i, i = 1, \ldots, N$: : independent Wiener processes.
- $((\Phi_i, \Psi_i), i = 1, \ldots, N)$ and $W_i, i = 1, \ldots, N$ independent.

Observations ($N \to \infty$ and $n \to \infty$)

Processes observed at times $t_j = jT/n, j = 1, \ldots, n$; T fixed.
Observations:($X_{i,n} := (X_i(t_j), j = 1, \ldots, n), i = 1, \ldots, N$).
Notation: $\Delta = \Delta_n = T/n$ $(n \to \infty \Rightarrow \Delta_n \to 0)$.

Main issues for SDEME

Likelihood for SDEME

- As for general mixed models: intractable likelihood.
- Additional pb: Discrete observations of each SDE on $[0, T]$.

Two possible sources of randomness in the structural SDE

- Random effects in the drift and/or in the diffusion coef.
- Natural to incorporate in the SDE model a joint distribution.
- Discrete observations of SDE : Different rates of convergence for estimating φ and ψ (Kessler, 1997).
- Important issue: understanding how estimation performs in SDEME.
- Necessary to clarify what to expect from numerical methods.

AIM

Investigate the statistical properties of estimators according to N, nb of individuals and n, nb of observations per individual.

Likelihood for SDE with mixed effects

Parametric distribution for the $(\Phi_i, \Psi_i) := \nu_\vartheta(d\varphi, d\psi)$

Likelihood of the sample $(X_{i,n}, i = 1, \ldots, N)$:

$$L_{N,n}(\vartheta) = \prod_{i=1}^{N} L_n(X_{i,n}, \vartheta), \quad \text{where}$$

$$L_n(X_{i,n}, \vartheta) = \int L_n(X_{i,n}, \varphi, \psi)\nu_\vartheta(d\varphi, d\psi) \quad : \text{Likelihood of } X_{i,n}, \tag{2}$$

$$L_n(X_{i,n}, \varphi, \psi) \quad : \text{Conditional likelihood given } \Phi_i = \varphi . \Psi_i = \psi \tag{3}$$

i.e. likelihood of $dX_i^{\varphi,\psi}(t) = b(X_i^{\varphi,\psi}(t), \varphi)dt + \sigma(X_i^{\varphi,\psi}(t), \psi)\, dW_i(t)$.

Two main difficulties

(1) Discrete observations: intractable $L_n(X_{i,n}, \varphi, \psi) \Rightarrow$ Approximation.

(2) Integration in (2) : no closed form for most distributions ν_ϑ

$\Rightarrow$ Choice of specific models $b(.,.), \sigma(.,.)$ and distributions $\nu_\vartheta(.,.)$.

Specific distribution for the random effects

⋆ $\mathcal{L}_n(X_{i,n}, \vartheta)$: approximation of $L_n(X_{i,n}, \vartheta)$ for the i-th path.
⋆ $\psi = \gamma^{-1/2} \Rightarrow \{\vartheta = (\gamma, \varphi)$ and $\nu_\vartheta(d\gamma, d\varphi)\}$.

$$\mathcal{L}_n(X_{i,n}, \vartheta) = \int_{(0,+\infty)\times\mathbb{R}^d} \mathcal{L}_n(X_{i,n}, \gamma, \varphi)\, \nu_\vartheta(d\gamma, d\varphi). \tag{4}$$

FOCUS: Distributions of the random effects such that the integration of (4) is explicit.

Three cases of mixed effects

1. $\Psi_i = \psi = \gamma^{-1/2}$ unknown and $\Phi_i \sim \mathcal{N}_d(\boldsymbol{\mu}, \gamma^{-1}\boldsymbol{\Omega}) \Rightarrow \theta = (\gamma, \boldsymbol{\mu}, \boldsymbol{\Omega})$
2. $\Phi_i = \phi$ unknow ; $\Psi_i = \Gamma_i^{-1/2}$ with $\Gamma_i \sim G(a, \lambda) \Rightarrow \tau = (\lambda, a, \phi)$.
3. Joint distribution for (Φ_i, Ψ_i): $\Psi_i = \Gamma_i^{-1/2}$, $\Gamma_i \sim G(a, \lambda)$. and given $\Gamma_i = \gamma$, $\Phi_i \sim \mathcal{N}_d(\boldsymbol{\mu}, \gamma^{-1}\boldsymbol{\Omega}) \Rightarrow \vartheta = (\lambda, a, \boldsymbol{\mu}, \boldsymbol{\Omega})$

Detail here: Cases (1) and (3).

Case of a joint distribution for the random effects

- Joint distribution for (Φ_i, Ψ_i): $\Psi_i = \Gamma_i^{-1/2}$, $\Gamma_i \sim G(a, \lambda)$. and given $\Gamma_i = \gamma$, $\Phi_i \sim \mathcal{N}_d(\boldsymbol{\mu}, \gamma^{-1}\boldsymbol{\Omega})$.
- Model where the marginal distribution of Φ_i is not Gaussian
  $\Phi_i - \boldsymbol{\mu}$ has a Student distribution

Construction of two distinct approximate likelihoods (a), (b)

(a) Derived from the study of Case (1)
(b) Separation of the inference for (a, λ) and $(\boldsymbol{\mu}, \boldsymbol{\Omega})$

(a) Good approximation of the likelihood.
(b) Easier to implement.

Approximate conditional likelihood (1)

Specific model : Linear mixed effects

- $b(x, \varphi) = \varphi' b(x)$ with $\varphi' = (\varphi_1, \ldots, \varphi_d)$; $\sigma(x, \psi) = \psi\sigma(x)$;
- $b(.) = (b_1(.), \ldots b_d(.))'$ and $\sigma(.)$ known ; x known.

Approximate conditional likelihood

Derived from **Euler scheme** $(Y_{i,n})$ of $(X_i^{\varphi,\psi})$ with $\Delta = T/n$,

$$d\, X_i^{\varphi,\psi}(t) = \varphi b(X_i^{\varphi,\psi}(t))dt + \psi\sigma(X_i^{\varphi,\psi}(t))\, dW_i(t), \quad X_i^{\varphi,\psi}(0) = x$$

Exact $L_n(X_{i,n}, \varphi, \psi)$ replaced by $\mathcal{L}_n$, likelihood of $(Y_{i,n})$:

$$Y_{i,j} - Y_{i,j-1} = \Delta\varphi' b(Y_{i,j-1}) + \sqrt{\Delta}\,\psi\,\sigma(Y_{i,j-1})\,\epsilon_{i,j},$$

with $Y_{i,0} = x$ and $\epsilon_{i,j} = \frac{W_i(t_j) - W_i(t_{j-1})}{\sqrt{\Delta}}$ i.i.d. $\mathcal{N}(0,1)$.

Approximate conditional likelihood given $\Phi_i = \varphi, \Psi_i = \psi$

Set $\Psi_i = \psi = \gamma^{-1/2}$:

$$\mathcal{L}_n(X_{i,n}, \gamma, \varphi) = \gamma^{n/2} \exp\left[-\frac{\gamma}{2}(S_{i,n} + \varphi' V_{i,n}\varphi - 2\varphi' U_{i,n})\right], \quad \text{where}$$

$$S_{i,n} = S_i = \frac{1}{\Delta}\sum_{j=1}^{n} \frac{(X_i(t_j) - X_i(t_{j-1}))^2}{\sigma^2(X_i(t_{j-1}))},$$

$$V_{i,n} = V_i = \left(\sum_{j=1}^{n} \Delta \frac{b_k(X_i(t_{j-1}))b_\ell(X_i(t_{j-1}))}{\sigma^2(X_i(t_{j-1}))}\right)_{1\leq k,\ell\leq d},$$

$$U_{i,n} = U_i = \left(\sum_{j=1}^{n} \frac{b_k(X_i(t_{j-1}))(X_i(t_j) - X_i(t_{j-1}))}{\sigma^2(X_i(t_{j-1}))}\right)_{1\leq k\leq d},$$

Preliminary discretization results

Three Sufficient statistics

Discretization results: As $n \to \infty$

$$\frac{S_{i,n}}{n} \quad \to \quad \Psi_i^2 = \Gamma_i^{-1} \quad \text{in prob. (quadratic variations),}$$

$$V_{i,n} \quad \to \quad V_i(T) = \left(\int_0^T \frac{b_k(X_i(t)) b_l(X_i(t))}{\sigma^2(X_i(t))} \, dt \right)_{1 \le k,l \le d} \qquad a.s. \text{ (Rieman),}$$

$$U_{i,n} \quad \to \quad U_i(T) = \left(\int_0^T \frac{b_k(X_i(t))}{\sigma^2(X_i(t))} \, dX_i(t) \right)_{1 \le k \le d} \text{in prob. (stoch. integral)}$$

Assumptions for the statistical study

(H1): Assume that $b(.), \sigma(.)$ C^2 bounded; $\sigma(.) \geq \sigma_0 > 0$.
(H2) $V_i(T)$ positive definite a.s.

Results of Case (1): $\Psi_i = \psi$ fixed unknown

$$\begin{aligned} dX_i(t) &= \Phi_i' b(X_i(t))dt + \Psi_i \sigma(X_i(t))dW_i(t), \\ X_i(0) &= x \in \mathbb{R}\,,\ t \in [0,T]\,, T = n\Delta\,;\ i = 1,\ldots,N. \end{aligned}$$

$\Psi_i = \psi = \gamma^{-1/2}$ unknown and $\Phi_i \sim \mathcal{N}_d(\boldsymbol{\mu}, \gamma^{-1}\boldsymbol{\Omega}) \Rightarrow \theta = (\gamma, \boldsymbol{\mu}, \boldsymbol{\Omega})$

Proposition: Explicit approximate likelihood for $(X_{i,n})$

$$\begin{aligned} \mathcal{L}_n(X_{i,n},\theta) &= \gamma^{n/2}(det(\boldsymbol{I}_d + V_{i,n}\,\boldsymbol{\Omega}))^{-1/2} \exp -\frac{\gamma}{2}(S_{i,n} + T_{i,n}(\boldsymbol{\mu},\boldsymbol{\Omega})), \\ T_{i,n}(\boldsymbol{\mu},\boldsymbol{\Omega}) &= (\boldsymbol{\mu} - V_{i,n}^{-1}U_{i,n})'\, R_{i,n}^{-1}\,(\boldsymbol{\mu} - V_{i,n}^{-1}U_{i,n}) - U_{i,n}'\, V_{i,n}^{-1}\, U_{i,n} \\ R_{i,n} &= V_{i,n}^{-1} + \boldsymbol{\Omega}. \end{aligned}$$

⋆ Formula $\mathcal{L}_n(X_{i,n},\theta)$ holds if $\boldsymbol{\Omega}$ is singular.

⋆ Possibility to have both fixed and random effects in the drift coefficient.

Asymptotic properties of estimators

- $L_{N,n}(\theta) = \prod_{i=1}^{N} \mathcal{L}_n(X_{i,n}, \theta) \Rightarrow \ell_{N,n}(\theta) = \log L_{N,n}(\theta)$ (loglikelihood)
- Define estimators $\tilde{\theta}_{N,n}$ maximizing $\ell_{N,n}(\theta)$.
- Different rates of convergence for γ and $\boldsymbol{\mu}, \boldsymbol{\Omega}$.

$$D_{N,n} = \begin{pmatrix} \frac{1}{\sqrt{Nn}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{1}{\sqrt{N}} \boldsymbol{I}_d & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \frac{1}{\sqrt{N}} \boldsymbol{I}_{d\times d} \end{pmatrix}, \quad \mathcal{I}(\theta) = \left(\begin{array}{c|c} \frac{1}{2\gamma^2} & \mathbf{0} \\ \hline \mathbf{0} & I(\theta) \end{array} \right), \qquad (5)$$

$I(\theta)$ explicit covariance matrix.

Theorem

Assume (H1)-(H2), $I(\theta)$ invertible. If $N, n \to \infty$, $N/n \to 0$, there exists a solution $\tilde{\theta}_{N,n}$ with prob. tending to 1 which is consistent and s.t.

$$D_{N,n}^{-1}(\tilde{\theta}_{N,n} - \theta) \to_{\mathcal{D}} \mathcal{N}_q(0, \mathcal{I}^{-1}(\theta)) \text{ under } \mathbb{P}_\theta,$$

$(q = 1 + d + d(d+1)/2.)$

Comments

1. Theorem holds if Ω singular:
   Possible to include mixed effects in the drift coefficient.
2. Constraint $N/n \to 0$
3. Fixed and random effects in the drift: Same rates of convergence.
4. No loss of information from the discrete observations:
   Continuous observations $(X_i(t), i = 1, \dots N)$ (γ known, $d = 1$):
   Delattre et al.(2013): M.L.E $\hat{\theta}_{N,n}$ strongly consistent and same asymptotic variance for $\hat{\theta}_{N,n}$ and $\tilde{\theta}_{N,n}$.
5. Loss of efficiency w.r.t. direct observations of $\Phi_i \sim \mathcal{N}_2(\mu, \gamma^{-1}\omega^2)$

Joint distribution for random effects (Φ_i, Ψ_i)

Assumption: (Φ_i, Ψ_i) are dependent.

$$\Psi_i = \frac{1}{\Gamma_i^{1/2}}, \quad \Gamma_i \sim G(a, \lambda), \text{ and given } \Gamma_i = \gamma, \quad \Phi_i \sim \mathcal{N}_d(\boldsymbol{\mu}, \gamma^{-1}\boldsymbol{\Omega}).$$

Construction of two distinct approximate likelihoods (1), (2)

(1) Derived from the study of Case (1).
(2) Separation of the inference for (a, λ) and $(\boldsymbol{\mu}, \boldsymbol{\Omega})$.

(1) Good approximation of the likelihood.
(2) Easier to implement.

Remark: Φ_i has a non Gaussian marginal distribution ($\Phi_i - \boldsymbol{\mu} \sim$ Student).

Approximate likelihood for one path

$\vartheta = (\lambda, a, \boldsymbol{\mu}, \boldsymbol{\Omega})$

$$\Lambda_n(X_{i,n}, \vartheta) = \int_0^{+\infty} \mathcal{L}_n(X_{i,n}, \gamma, \boldsymbol{\mu}, \boldsymbol{\Omega}) \gamma^{a-1} \exp(-\lambda\gamma) \frac{\lambda^a}{\Gamma(a)} d\gamma, \qquad (6)$$

where $\mathcal{L}_n(X_{i,n}, \gamma, \boldsymbol{\mu}, \boldsymbol{\Omega})$ likelihood obtained for fixed γ:

$$\begin{aligned} \mathcal{L}_n(X_{i,n}, \gamma, \boldsymbol{\mu}, \boldsymbol{\Omega}) &= \gamma^{n/2}(det(\boldsymbol{I}_d + V_{i,n}\,\boldsymbol{\Omega}))^{-1/2} \exp -\frac{\gamma}{2}(S_{i,n} + T_{i,n}(\boldsymbol{\mu}, \boldsymbol{\Omega})) \\ T_{i,n}(\boldsymbol{\mu}, \boldsymbol{\Omega}) &= (\boldsymbol{\mu} - V_{i,n}^{-1} U_{i,n})'\, R_{i,n}^{-1}\, (\boldsymbol{\mu} - V_{i,n}^{-1} U_{i,n}) - U_{i,n}'\, V_{i,n}^{-1}\, U_{i,n}. \end{aligned}$$

Difficulty: (6) requires $2\lambda + S_{i,n} + T_{i,n}(\boldsymbol{\mu}, \boldsymbol{\Omega}) > 0$.

Need to define the random sets:

$$E_{i,n}(\vartheta) = \{S_{i,n} + T_{i,n}(\boldsymbol{\mu}, \boldsymbol{\Omega}) > 0\} \;\;;\;\; E_{N,n}(\vartheta) = \cap_{i=1}^{N} E_{i,n}(\vartheta).$$

Method 1 for estimating $\vartheta = (\lambda, a, \boldsymbol{\mu}, \boldsymbol{\Omega})$

Explicit approximate likelihood

$$\text{On } E_{N,n}(\vartheta), \quad \Lambda_{N,n}(\vartheta) = \prod_{i=1}^{N} \Lambda_n(X_{i,n}, \vartheta), \qquad \text{where} \tag{7}$$

$$\Lambda_n(X_{i,n}, \vartheta) = \frac{\lambda^a \Gamma(a + (n/2))}{\Gamma(a)(\lambda + \frac{S_{i,n}}{2} + \frac{T_{i,n}(\boldsymbol{\mu}, \boldsymbol{\Omega})}{2})^{a+(n/2)}} \frac{1}{(det(I_d + V_{i,n}\boldsymbol{\Omega}))^{1/2}} \tag{8}$$

Problem: Find a sufficient. condition ensuring $\forall \theta_0, \theta, \mathbb{P}_{\vartheta_0}(E_{N,n}(\vartheta)) \to 1$.

Proposition: If $a > 4$, there exists a subset $F_{N,n}$ containing $E_{N,n}(\vartheta)$ satisfying $P_{\vartheta_0}(F_{N,n}) \to 1$ as $N, n \to \infty$.

Formula (8) holds for non invertible $\boldsymbol{\Omega}$ $\Rightarrow$ Mixed effects for Φ .

Simplifying the likelihood: Method 2

For large n, using $S_{i,n}/n \rightarrow \Gamma_i^{-1}$

$$\begin{aligned}
\log \Lambda_n(X_{i,n}, \vartheta) &\simeq \mathbf{V}_n^{(1)}(X_{i,n}, \lambda, a) + \mathbf{V}_n^{(2)}(X_{i,n}, \boldsymbol{\mu}, \boldsymbol{\Omega}), \\
\mathbf{V}_n^{(1)}(X_{i,n}, \lambda, a) &= \log\left(\frac{\lambda^a \Gamma(a + (n/2))}{\Gamma(a)(\lambda + \frac{S_{i,n}}{2})^{a+(n/2)}}\right) \\
\mathbf{V}_n^{(2)}(X_{i,n}, \boldsymbol{\mu}, \boldsymbol{\Omega}) &= -\frac{n}{2S_{i,n}} T_{i,n}(\boldsymbol{\mu}, \boldsymbol{\Omega}) - \frac{1}{2}\log(det(I_d + V_{i,n}\boldsymbol{\Omega}))
\end{aligned}$$

$\mathbf{V}_n^{(1)}(X_{i,n}, \lambda, a)$: appr. likelihood for $b \equiv 0$ (Delattre et al., 2015).

New approximate loglikelihood for the ith path

$$\mathbf{V}_n(X_{i,n}, \vartheta) = \mathbf{V}_n^{(1)}(X_{i,n}, \lambda, a) + \mathbf{V}_n^{(2)}((X_{i,n}, \boldsymbol{\mu}, \boldsymbol{\Omega}) \quad (9)$$

Interest of $\mathbf{V}_n$: estimation of (λ, a) and $(\boldsymbol{\mu}, \boldsymbol{\Omega})$ performed separately.

Random effects in both coef.: estimating equations

$\star$ Two different approximations of the loglikelihood:

$$\mathbf{U}_{N,n}(\vartheta) = \sum_{i=1}^{N} \log \Lambda_n(X_{i,n}, \vartheta), \quad \mathbf{V}_{N,n}(\vartheta) = \sum_{i=1}^{N} \mathbf{V}_n(X_{i,n}, \vartheta),$$

$\star$ Two estimating equations leading to estimators:

$$\nabla_\vartheta \mathbf{U}_{N,n}(\tilde{\vartheta}_{N,n}) = 0, \quad \nabla_\vartheta \mathbf{V}_{N,n}(\bar{\vartheta}_{N,n}) = 0.$$

Random effects in both coef.: properties of estimators

The estimators $\tilde{\vartheta}_{N,n}$ and $\bar{\vartheta}_{N,n}$ of $\theta = (\lambda, a, \boldsymbol{\mu}, \boldsymbol{\Omega})$ satisfy:

Theorem

Assume $a > 6$, $N/n \to 0$ and that $J(\vartheta)$ is invertible. Then, a solution $\tilde{\vartheta}_{N,n}$ exists with probability tending to 1 which is consistent and such that, under $\mathbb{P}_\vartheta$,

$$\sqrt{N}(\tilde{\vartheta}_{N,n} - \vartheta) \to_{\mathcal{D}} \mathcal{N}_q(0, \mathcal{J}^{-1}(\vartheta))$$

where $q = 2 + d + d(d+1)/2)$. For the first two components of $\sqrt{N}(\tilde{\vartheta}_{N,n} - \vartheta)$, the constraint N/n^2 is enough.
The same properties hold for $\bar{\vartheta}_{N,n}$. These two estimators are asymptotically equivalent.

$\star$ $E_\vartheta(\Gamma_i^{-1}) = \lambda/(a-1)$.
$\star$ $a > 6$: moment property for Γ_i^{-1}(*i.e.* quadratic variations of $X_i(t)$).

$\mathcal{J}(\vartheta)$ block diagonal matrix precised in the next slide.

Random effects in both coef.: Covariance matrix

$$\mathcal{J}(\vartheta) = \left(\begin{array}{c|c} I_0(\lambda, a) & \mathbf{0} \\ \hline \mathbf{0} & J(\vartheta) \end{array} \right), \quad I_0(\lambda, a) = \begin{pmatrix} \frac{a}{\lambda^2} & -\frac{1}{\lambda} \\ -\frac{1}{\lambda} & \psi'(a) \end{pmatrix},$$

$J(\vartheta)$ is explicit (depends on $\Gamma_i, U_i(T), V_i(T),$).

$I_0(\lambda, a)$: Fisher information matrix for direct observations $\Gamma(a, \lambda)$.
$\Psi(a)$ Digamma function

Comments

1. Random effects in the drift or/and in the diffusion coefficient: Same rates of convergence $\sqrt{N}$.
2. Estimation of (λ, a): no loss of efficiency w.r.t. direct observations of N *i.i.d.* Γ_i and weaker constraint $N/n^2 \to 0$
3. $a > 6$: moment condition on Γ_i^{-1} $(\mathbb{E}_\vartheta \Gamma_i^{-6} < +\infty)$.
4. Note that the marginal distribution of Φ_i is not Gaussian in this set-up.

Remark: Similar results for Case (2) :
Fixed effects in the drift and random effects in the diffusion coefficients.

Assessment on various examples

Joint distribution for (Φ_i, Γ_i): $\Gamma_i \sim \Gamma(a, \lambda)$, $\Phi_i | \Gamma_i = \gamma \sim \mathcal{N}(\boldsymbol{\mu}, \gamma^{-1}\boldsymbol{\Omega})$.

(1) Mixed effect Brownian motion:

$$dX_i(t) = \Phi_i dt + \Gamma_i^{-1} dW_i(t), X_i(0) = 0 \ .$$

(2) Mixed O.U process: $\Gamma_i = \gamma$ or $\Gamma_i \sim G(a, \lambda)$

$$dX_i(t) = (\rho - \Phi_i X_i(t))dt + \Gamma_i^{-1/2} dW_i(t), X_i(0) = 0 \ ,$$

(3) Mixed O.U. with two ind. random effects: $\Gamma_i = \gamma$ or $\Gamma_i \sim G(a, \lambda)$

$$dX_i(t) = (\Phi_{i,1} - \Phi_{i,2} X_i(t))dt + \Gamma_i^{-1/2} dW_i(t), X_i(0) = 0 \ ,$$

(4) Mixed C.I.R. process $(\rho > 0 \Rightarrow \forall t \geq 0, \ \ X_i(t) \geq 0.)$

$$dX_i(t) = (\rho X_i(t)) - \Phi_i)dt + \Gamma_i^{-1/2} \sqrt{|X_i(t)|} dW_i(t), X_i(0) = x > 0 \ ,$$

(5) Implementation on a real data set of membrane potential paths

Simulation design and results

- Each SDE : 100 data sets with N subjects on $[0, T]$ with $T = 5$.
- Simulation: First draw the random effect.
- Exact simulation for Models 1-3 ; Alfonsi scheme for (4).
- Two sampling intervals $\Delta = 0.01; 0.005$.
- Two values for N : $N = 50, 100$.
- Preliminaries: standard estimation results from direct observations of N i.i.d. sample of (Φ_i, Γ_i)

Results

For Case 3 (joint distribution), similar results with Methods 1 and 2; Method 2: easier to implement $\Rightarrow$ Only Method 2 results presented.

Drift with random effect, fixed or random diff. coef.

$dX_i(t) = \Phi_i dt + \gamma^{-1/2} dW_i(t)$; $\Phi_i \sim \mathcal{N}(\mu, \gamma^{-1}\omega^2)$

	value	$\Delta = 0.01$ ($n = 500$)	$\Delta = 0.005$ ($n = 1000$)
μ	0	0.01 (0.04)	0.01 (0.03)
ω^2	0.1	0.10 (0.06)	0.1 (0.04)
γ	4.00	4.01 (0.04)	4.00 (0.02)

$dX_i(t) = \Phi_i dt + \Gamma_i^{-1/2} dW_i(t)$; $\Gamma_i \sim G(a, \lambda)$

	value	$\Delta = 0.025$ ($n = 200$)	$\Delta = 0.005$ ($n = 1000$)
μ	-0.50	- 0.49 (0.06)	- 0.51 (0.06)
ω^2	0.50	0.45 (0.11)	0.49 (0.12)
m	4.00	3.85 (0.22)	3.96 (0.23)
t	1.32	1.24 (0.06)	1.30 (0.06)

$N = 50$, $T = 5$; $(a, \lambda) = (8, 2)$, $m = a/\lambda$, $t = \psi(a) - \log\lambda$; $E(\Gamma_i^{-1}) = 2/7$.

Drift with mixed effects, fixed or random diff. coef.

$dX_i(t) = (\rho X_i(t) + \Phi_i)dt + \gamma^{-1/2} dW_i(t)$; $\Phi_i \sim \mathcal{N}(\mu, \gamma^{-1}\omega^2)$;

	value	$\Delta = 0.01$ ($n = 500$)	$\Delta = 0.005$ ($n = 1000$)
ρ	-0.1	-0.11 (0.02)	-0.10 (0.01)
μ	1	1.01 (0.08)	1.02 (0.06)
ω^2	0.4	0.40 (0.14)	0.40 (0.12)
γ	4.00	4.01 (0.04)	4.01 (0.02)

$dX_i(t) = (\rho - \Phi_i X_i(t))dt + \Gamma_i^{-1/2} dW_i(t)$; $\Gamma_i \sim G(a, \lambda)$

	value	$\Delta = 0.025$ ($n = 200$)	$\Delta = 0.005$ ($n = 1000$)
ρ	1.00	0.99 (0.07)	1.00 (0.08)
μ	0.50	0.49 (0.06)	0.50 (0.06)
ω^2	0.10	0.08 (0.04)	0.09 (0.04)
m	4.00	3.95 (0.21)	3.99 (0.21)
t	1.32	1.26 (0.05)	1.30 (0.05)

$N = 50, T = 5$; $m = a/\lambda, t = \psi(a) - \log\lambda$ $(E(\Gamma_i^{-1}) = 2/7)$.

Drift with 2 random effects, fixed or random diff. coef.

$$dX_i(t) = (\Phi_{i,1}X_i(t) + \Phi_{i,2})dt + \gamma^{-1/2}dW_i(t)$$

	value	$\Delta = 0.01$ ($n = 500$)	$\Delta = 0.005$ ($n = 1000$)
μ_1	-0.1	-0.1 (0.04)	-0.11 (0.02)
μ_2	1.00	0.98 (0.08)	1.01 (0.06)
ω_1^2	0.04	0.04 (0.02)	0.04 (0.01)
ω_2^2	0.4	0.37 (0.15)	0.37 (0.13)
γ	4.00	4.02 (0.03)	4.00 (0.03)

$$dX_i(t) = (\Phi_{i,1}X_i(t) + \Phi_{i,2})dt + \Gamma_i^{-1/2}dW_i(t);\ \Gamma_i \sim G(a\lambda)$$

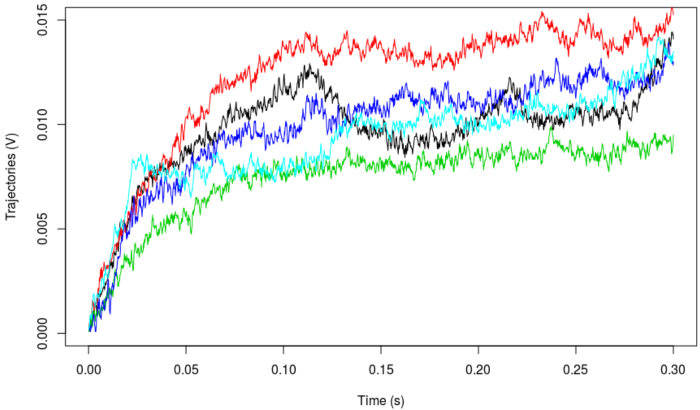
	value	$\Delta = 0.025$ ($n = 200$)	$\Delta = 0.005$ ($n = 1000$)
μ_1	-0.50	-0.48 (0.07)	-0.49 (0.07)
μ_2	1.00	0.97 (0.11)	0.99 (0.11)
ω_1^2	0.10	0.09 (0.06)	0.09 (0.06)
ω_2^2	0.50	0.49 (0.20)	0.52 (0.23)
m	4.00	3.96 (0.22)	4.00 (0.21)
t	1.32	1.27 (0.06)	1.31 (0.05)

Comments on these simulation results

- Pb with the sets $F_{i,n}$: too stringent truncations $\Rightarrow$ Method 2.
- Two est. methods: similar good performances (whatever n and N).
- Bias and s.d of estimates decrease as N increases.
- For N fixed, increasing n decreases the bias but no impact on the s.d . of estimates
- Agrees with theoretical results
- For $n = 200$, condition $N/n \to 0$ not fullfilled , but good results.
- Ass. (H1)(H2) not satisfied : does not deteriorate perfomances.
- Singular $\mathbf{\Omega}$ (mixed effects in the drift) : parameters still well estimated

Neuronal data (from Picchini et al.,2008)

Data set: $N = 240$ membrane potential trajectories (Volts)
Observations: $n = 2000$ points on $[0, T]$ with $T = 0.3$ s
SDE MODELS: O.U.(Picchini et al, 2008,2010; Dion,2016); C.I.R (Hoepfner,2007).



Modeling neuronal data with SDE with mixed effects

$$dX_i(t) = (\Phi_{i,1}X_i(t) + \Phi_{i,2})dt + \Psi_i\sigma(X_i(t))dW_i(t), \quad X_i(0) = x_i$$

Includes $\sigma(x) \equiv 1$ (O.U.) and $(x) = \sqrt{|x|}$ (CIR).

Try to answer two questions

- Which of the OU or the CIR is the most appropriate SDE?
- Which parameter should be fixed or random?

Four models tested with $\Psi_i = \Gamma_i^{-1/2}$ with $\Gamma_i \sim G(a, \lambda)$

1. O.U.: $\Phi_{i,1} = \mu_1$ fixed; $\Phi_{i,2}$ random s.t. $\Phi_{i,2}|\Gamma_i \sim \mathcal{N}(\mu_2, \omega_2^2/\Gamma_i)$.
2. O.U.: $\Phi_{i,1}$ random s.t. $\Phi_{i,1}|\Gamma_i \sim \mathcal{N}(\mu_1, \omega_1^2/\Gamma_i)$; $\Phi_{i,2} = \mu_2$ fixed.
3. O.U.: $\Phi_{i,1}, \Phi_{i,2}$ random independent;
   $\Phi_{i,1}|\Gamma_i \sim \mathcal{N}(\mu_1, \omega_1^2/\Gamma_i)$, $\Phi_{i,2}|\Gamma_i \sim \mathcal{N}(\mu_2, \omega_2^2/\Gamma_i)$.
4. CIR: $\Phi_{i,1} = \mu_1$ fixed; $\Phi_{i,2}$ random s.t. $\Phi_{i,2}|\Gamma_i \sim \mathcal{N}(\mu_2, \omega_2^2/\Gamma_i)$.

Results

Comparison between various models comparing BIC values

$$BIC = -2\log\hat{\mathcal{L}} + d\log N$$

⋆ $\hat{\mathcal{L}}$: likelihood of the observations evaluated at the parameter estimate

⋆ d= nb of model parameters.

⋆ Likelihood not explicit $\Rightarrow$ use of the best approximation (Method 1)

Results

- Best model O.U. (1): $dX_i(t) = (\mu_1 X_i(t) + \Phi_{i,2})dt + \Gamma^{-1/2}dW_i(t)$
- Parameter estimates: $(\mu_1 = -0.037, \mu_2 = 0.38, \omega_2^2 = 0.015)$; $(a = 16.2, \lambda = 2.93)$ (corresponds to $E(\Gamma^{-1}) = 0.19$).

⋆ Consistent with previous results.

⋆ Comparison with models with fixed diffusion coefficient?

⋆ No longer possible to use BIC $\Rightarrow$ two solutions:

⋆ Use of the Q-BIC of Eguchi& Masuda (2016) for M-estimators.

⋆ Develop rigorous testing approach (parameters located at a boundary).

Concluding remarks

- Mixed effects for stochastic differential equations on $\mathbb{R}$.
- Linear multidimensional mixed effect in the drift and multiplicative random effect Γ in the diff. coef. with joint distribution $\nu_\theta(.,.)$.
- Discrete observations (n) of N paths observed on $[0, T]$ fixed.
- Good results over all of estimators under conditions linking N and n.
- It works well in pratice even if it is not satisfied.

Extensions

- Time -dependent drift and diffusion coefficients: $b(t,x), \sigma(t,x)$: O.K.
- Assumption $X_i(0) = x$ done for simplicity $\Rightarrow X_i(0) = x_i$ is O.K.
- Multimensional mixed effects diffusion models on $\mathbb{R}^k$:
  $dX_i(t) = B(X_i(t))\Phi_i dt + \Psi_i \Sigma(X_i(t)) dW_i(t);\ X_i(0) = x.$
- Adding fixed unknown parameters in b, σ: $b(\alpha, x), \sigma(\beta, x)$?
- Testing methods to decide which parameters are fixed or random?

THANKS FOR YOUR ATTENTION !

General References

- Nie & Yang (2005), Nie (2006, 2007). Theoretical likelihood study. Rely on many abstract assumptions impossible to check in practice.
- Donnet, S. & Samson, A. (2008). Review for mixed effects SDEs.
- Picchini, De Gaetano & Ditlevsen (2008, 2010); Picchini & Ditlevsen (2011) (approximations of the likelihood, no theoretical results)
- Delattre, Genon-Catalot & Samson (2012, 2015, 2016); Genon-Catalot & Larédo (2016); Delattre, Genon-Catalot & Larédo (Preprints 2016, 2017) (parametric, likelihood methods).
- Comte, Genon-Catalot & Samson (2013), Genon-Catalot & Dion (2015) (non parametric)
- Große Ruse, Samson & Ditlevsen (Preprint 2017) (Multidimensional diff., Mixed effects in the drift+ covariates.)

References on applications based on data sets of PK/PD dynamics

- Overgaard,Jonsson, Tornøe & Madsen (2005)
- Berglund, Sunnake, Adiels, Jirstrand & Wennberg (2011)
- Leander, J., Almquist, J., Ahlstrom, C., Gabrielsson, J. & Jirstrand, M. (2015) (concrete models and real data + many references therein)