Approximation of epidemic models by diffusion processes and their statistical inferencedes

Catherine Larédo [1,2]

[1]UR 341, MaIAGE, INRA, Jouy-en-Josas
[2] UMR 7599, LPMA, Université Paris Diderot

21/11/2015 CIMPA 2015

TALK 3-Part 2, CIMPA 2015
Inference for discretely observed epidemic processes

Based on Guy, Laredo, Vergu (2014) Stoch. Proc.Appl.

Back to Epidemics in a close population of size N

Some characteristics of epidemic data

- $\epsilon = \frac{1}{\sqrt{N}}$ (small parameter present in the diffusion term).
- nb of observations n s.t. $n << N$ and $\Delta \geq 1$ (1 to 7 days).
- Framework (1) $\epsilon \to 0$, Δ fixed (n finite) more appropriate.
- Choice of a statistical framework:
- Depends more on the relative magnitudes of T, Δ, N than on their accurate values.
- Interest in studying estimation in both frameworks.
- **Data might change** $\Rightarrow$ Asymptotic framework $\Delta = \Delta_n \to 0$ also appropriate:
- Available data can become more frequent.
- Study over a long time period for recurrent outbreaks.

Recap for the three epidemic models

- *SIR* diffusion approximation on $\mathbb{R}^2$:

$$b_\theta(s,i) = \begin{pmatrix} -\lambda si \\ \lambda si - \gamma i \end{pmatrix} \text{ and } \sigma_\theta(s,i) = \begin{pmatrix} \sqrt{\lambda si} & 0 \\ -\sqrt{\lambda si} & \sqrt{\gamma i} \end{pmatrix}$$

- *SEIR* with temporal transmission rate : diffusion on $\mathbb{R}^3$

$$b(\theta; s,e,i) = \begin{pmatrix} -\lambda(t)si \\ \lambda(t)si - \rho e \\ \rho e - \gamma i \end{pmatrix} \text{ and } \sigma_\theta(t;x) = \begin{pmatrix} \sqrt{\lambda(t)si} & 0 & 0 \\ -\sqrt{\lambda(t)si} & \sqrt{\rho e} & 0 \\ 0 & -\sqrt{\rho e} & \sqrt{\gamma i} \end{pmatrix}$$

- *SIRS* model with seasonal forcing: $\lambda(t) = \lambda_0(1 + \lambda_1 sin(2\pi \frac{t}{T_{per}}))$

$$b_\theta(t;s,i) = \begin{pmatrix} b_{\theta,1}(t;s,i) \\ b_{\theta,2}(t;s,i) \end{pmatrix} = \begin{pmatrix} -\lambda(t)si + \delta(1-s-i) + \mu(1-s) \\ \lambda(t)si - (\gamma+\mu)i \end{pmatrix}.$$

$$\Sigma = \begin{pmatrix} \lambda(t)si + \delta(1-s-i) + \mu(1+s) & -\lambda(t)si \\ -\lambda(t)si & \lambda(t)si + (\gamma+\mu)i \end{pmatrix}.$$

Discrete observations with fixed sampling interval Δ on [0,T]

$dX(t) = b(\alpha, X(t))dt + \epsilon\sigma(\beta, X(t))dB(t), X(0) = x_0$: diffusion sur $\mathbb{R}^p$.

Observations: $X^n := (X(t_k);\ t_k = k\Delta,\ 0 \leq k \leq n)$ with $T = n\Delta$.

No time -dependence in the drift and diffusion terms; $X(0) = x_0$ known.

- $X(t)$ Markov process with transition probabilities : $p_t(x, y) = \mathbb{P}(X(t+s) = y/X(s) = x)$
- $\Rightarrow (X(t_k), k \geq 0)$ Markov chain with state space $\mathbb{R}^p$ and transition kernel $Q_\Delta(x, dy) = p_\Delta(x, y)dy$
- These transition kernels depend on $b(\alpha, x)$ and $\epsilon\sigma(.)$
- $\rightarrow p_\Delta(x, y) = p_{\Delta,\epsilon}(\alpha, \beta; x, y)$
- No analytic expression $\Rightarrow$ untractable likelihood
- impossible to use in practice.

Discrete observations on a finite time interval

Use an estimation function (or contrast) derived from the Euler scheme:
Euler scheme associated with $(X(t))$
$X(t_k) \simeq X(t_{k-1}) + \Delta b(X(t_{k-1})) + \epsilon\sqrt{\Delta}\sigma(X(t_{k-1}))\,\eta_k$
with $(\eta_k) \sim \mathcal{N}(0,1)$ i.i.d random variables.
$\Rightarrow$ Markov chain model with explicit transition kernels

$$p_{\epsilon,\Delta}(\alpha,\beta;x,y)dy \sim \mathcal{N}(\Delta b(\alpha;x)), \epsilon^2\Delta\Sigma(\beta;x)).$$

Contrast process:

$$U_{\epsilon,\Delta}\left(\alpha,\beta;(X(t_k))\right) = \sum_{k=1}^{n} log(det(V_k(\beta))) + \frac{1}{\epsilon^2\Delta}\,{}^tB_k(\alpha)V_k^{-1}(\beta)B_k(\alpha).$$

$B_k(\alpha) = X(t_k) - X(t_{k-1}) - \Delta b(\alpha, X(t_{k-1}))$,
$V_k(\beta) = \Sigma(\beta, X(t_{k-1})) = \sigma(\beta, X(t_{k-1}))\,{}^t\sigma(\beta, X(t_{k-1}))$
Pb Δ fixed: The Euler scheme is not a good approximation of $(X(t))$.
Pb 2:$\Delta = \Delta_n \to 0$: ϵ and Δ_n are linked in this approach.

Choosing estimating functions

Other approach

• Use another approximation for the sample paths of $X(t)$

• Base the estimation method on this approximation.

• Theorem (Ventsell &Freidlin,19??, Small pertubations of dynamical systems)

$$X_\epsilon(t) = x(\alpha; t) + \epsilon g(\alpha, \beta; t) + \epsilon R_\epsilon(t).$$

$\star$ $x(\alpha; .)$ satisfies $\frac{\partial x}{\partial t}(t) = b(\alpha; x(t))dt; x(0) = x_0$
$\star$ $g(\alpha, \beta; t) = \int_0^t \Phi(\alpha; t, s)\sigma(\beta, x(\alpha; s))dB(s)$:
$\star$ $\Phi(\alpha; t, u)$: $\frac{\partial \Phi}{\partial t}(t, u) = \nabla b(\alpha; x(\alpha; t))\Phi(t, u)$; $\Phi(u, u) = I_p$.
$\star$ $sup_{t \leq T} \parallel \epsilon R_\epsilon(t) \parallel \rightarrow 0$ in probability as $\epsilon \rightarrow 0$,
(Extension of Genon-Catalot (1990))

Estimation functions

We have that $X(t_k) \sim Y(t_k)$ with $Y(t) = x(\alpha; t) + \epsilon g(\alpha, \beta; t)$.

A noteworthy property of g

- $g(\alpha, \beta; t_k) = \Phi(\alpha; t_k, t_{k-1}) g(\alpha, \beta; t_{k-1}) + \sqrt{\Delta}\, Z_k$
- $Z_k = \int_{t_{k-1}}^{t_k} \Phi(\alpha; t_k, s) \sigma(\beta; x(\alpha; s)) dB(s)$
- $(Z_k, k = 1, \ldots, n)$ independent $\mathbb{R}^p$ Gaussian r.v. with covariance
- $S_k(\alpha, \beta) = \int_{t_{k-1}}^{t_k} \Phi(\alpha; t_k, s) \Sigma(\beta; x(\alpha; s))\, {}^t\Phi(\alpha; t_k, s) ds$

Define the function $\mathbb{R}^p \longrightarrow \mathbb{R}^p$

$x \longrightarrow B_k(\alpha; x) = x(\alpha; t_k) + \Phi(\alpha; t_k, t_{k-1})(x - x(\alpha, t_{k-1}))$.

$Z_k = \frac{1}{\epsilon\sqrt{\Delta}}(Y(t_k) - B_k(\alpha, Y(t_{k-1}))$.

$(Y(t_k), k \geq 0)$ Markov chain with explicit transition kernel

$$p_{\epsilon,\Delta}(\alpha, \beta; x, y) dy \sim \mathcal{N}_p(B_k(\alpha; x), \epsilon^2 \Delta S_k(\alpha, \beta)\ ).$$

Parametric inference for fixed sampling interval

Asymptotic framework

$\epsilon \to 0$, and T, Δ fixed : $\Rightarrow$ finite nb n of observations;

Only parameters in the drift can be consistently estimated

Ex:Brownian motion with drift: $dX(t) = \alpha dt + \epsilon\beta dB(t)$; $X(0) = 0$.

- $(X(t_k) - X(t_{k-1}), k = 1, \ldots n$ i.i.d $\mathcal{N}(\alpha\Delta, \epsilon^2\beta^2\Delta)$.
- Explicit likelihood $\to$ Explicit M.L.E.
- $\hat{\alpha}_\epsilon = \frac{X(T)}{T})$, $\hat{\beta}_\epsilon^2 = \frac{1}{\Delta\epsilon^2}\sum_{i=1}^n (U_k - \Delta\hat{\alpha}_\epsilon)^2$.
- Under $\mathbb{P}^\epsilon_{\alpha_0,\beta_0}$, $\hat{\alpha}_\epsilon = \alpha_0 + \epsilon\beta_0\frac{B(T)}{T}$,
  $\hat{\beta}_\epsilon^2 = \beta_0^2(\frac{1}{n}\sum_{i=1}^n U_k^2 - \frac{1}{n}\frac{B^2(T)}{T})$ where (U_k) i.i.d. $\mathcal{N}(0,1)$.
- $\hat{\alpha}_\epsilon \to \alpha_0$ and $\epsilon^{-1}(\hat{\alpha}_\epsilon - \alpha_0) \sim \mathcal{N}(0, \beta_0^2/T)$.
- $\hat{\beta}_\epsilon^2$ fixed random variable independent of ϵ,
- $E_{\alpha_0,\beta_0}(\hat{\beta}_\epsilon^2) = \beta_0^2(1 - 1/n) \neq \beta_0^2$: biased estimator.

Inference for fixed sampling interval Δ

$dX(t) = b(\alpha, X(t))dt + \epsilon\sigma(\beta, X(t))dB(t), X(0) = x_0$: diffusion on $\mathbb{R}^p$.

- Parameter β in $\sigma(\beta, x)$ cannot be consistently estimated.
- Only α in the drift term $b(\alpha, x)$ can be estimated.
- Diffusion approximations of epidemic models: $\beta = \alpha$.

Two-stage approach in framework (1)

Estimation of parameter α assuming β unknown.
Use of $\beta = \alpha$ to improve the estimator.

First step (General case):
Estimation of α: Approximate Conditional Least Squares

$$\tilde{U}_\epsilon(\alpha, \beta, X^{(n)}) = \frac{1}{\epsilon^2\Delta}\sum_{k=1}^{n} {}^t(X(t_k) - B_k(\alpha; X(t_{k-1})))\,(X(t_k) - B_k(\alpha; X(t_{k-1}))).$$

where $B_k(\alpha; x) = x(\alpha; t_k) + \Phi(\alpha; t_k, t_{k-1})(x - x(\alpha, t_{k-1}))$.

Diffusion approximation of epidemic models: $\beta = \alpha, \epsilon = \frac{1}{\sqrt{N}}$

$S_k(\alpha, \beta) = \int_{t_{k-1}}^{t_k} \Phi(\alpha; t_k, s)\Sigma(\beta; x(\alpha; s))\, {}^t\Phi(\alpha; t_k, s) ds$

Contrast Process

$$U_{\epsilon,\Delta}(\alpha; X^{(n)}) = \frac{1}{2}\sum_{k=1}^{n} log(det(S_k(\alpha, \alpha)) +$$

$$\frac{1}{2\epsilon^2\Delta}\sum_{k=1}^{n} {}^t(X(t_k) - B_k(\alpha; X(t_{k-1})))S_k^{-1}(\alpha, \alpha)\,(X(t_k) - B_k(\alpha; X(t_{k-1}))).$$

$$\hat{\alpha}_\epsilon \text{ such that } U_{\epsilon,\Delta}(\hat{\alpha}_\epsilon, X^{(n)}) = inf\{U_{\epsilon,\Delta}(\alpha, X^{(n)}), \alpha \in \Theta\}.$$

Three conditions to check as $\epsilon \to 0$

(1) $\epsilon^2 U_{\epsilon,\Delta}(\alpha, \beta, X^{(n)}) \to K_\Delta(\alpha_0, \alpha)$ $P^\epsilon_{\alpha_0}$ a.s. uniformly on Θ

$K(\alpha_0, \alpha)$ continuous deterministic + unique global minimum at α_0

(2) $\epsilon\nabla_\alpha U_{\epsilon,\Delta}(\alpha_0) \to \mathcal{N}(0, J_\Delta(\alpha_0))$ in distribution under $P^\epsilon_{\alpha_0}$

(3) There exists a non-singular matrix $I_\Delta(\alpha_0)$ such that

$lim_{\delta\to 0} lim_{\epsilon\to 0} sup\{\| \epsilon^2\nabla^2_\alpha U_\epsilon(\alpha) - I_\Delta(\alpha_0) \|, \| \alpha - \alpha_0 \| \leq \delta\} = 0$ $P^\epsilon_{\alpha_0}$ a.s.

Checking conditions (1)-(3)

Condition (1): ensures consistency of $\hat{\alpha}_\epsilon$
Assumption: $\alpha, \alpha' \in \Theta, \alpha \neq \alpha' \Rightarrow x(\alpha, t_k) \neq x(\alpha', t_k)$ for some k.

Condition (2): $\frac{1}{\epsilon\sqrt{\Delta}}(X(t_k) - B_k(\alpha_0, X(t_{k-1})))$: approximately Gaussian
$\Rightarrow \epsilon \nabla_\alpha U_{\epsilon,\Delta}(\alpha_0) \to \mathcal{N}(0, J_\Delta(\alpha_0))$ with
$J_\Delta(\alpha_0) = \Delta \sum_{i=1}^{n} {}^tD_k(\alpha_0) S_k(\alpha_0)^{-1} D_k(\alpha_0)$ where
$D_k(\alpha) = \frac{1}{\Delta}(\nabla_\alpha x(\alpha, t_k) - \Phi(\alpha; t_k, t_{k-1}) \nabla x(\alpha, t_{k-1}))$

Condition (3) $\epsilon^2 \nabla^2_\alpha U_\epsilon(\alpha_0) \to I_\Delta(\alpha_0)$, with $I_\Delta(\alpha_0) = J_\Delta(\alpha_0)$.
Assumption: $J_\Delta(\alpha_0)$ non singular.

Result: $\epsilon^{-1}(\hat{\alpha}_\epsilon - \alpha_0) \to \mathcal{N}(0, I_\Delta^{-1}(\alpha_0))$

Remark: Approximate Conditional Least Squares
$J_\Delta(\alpha_0, \beta_0) = \Delta \sum_{i=1}^{n} {}^tD_k(\alpha_0) S_k(\alpha_0), \beta_0) D_k(\alpha_0)$;
$I_\Delta(\alpha_0) = \Delta \sum_{i=1}^{n} {}^tD_k(\alpha_0) D_k(\alpha_0)$
$\epsilon^{-1}(\tilde{\alpha}_\epsilon - \alpha_0) \to \mathcal{N}(0, I_\Delta^{-1}(\alpha_0) J_\Delta(\alpha_0) I_\Delta^{-1}(\alpha_0))$

Inference for small sampling interval $\Delta = \Delta_n \rightarrow$

$dX(t) = b(\alpha; t, X(t))dt + \epsilon\sigma(\beta; t, X(t))dB(t), X(0) = x_0$: diffusion on $\mathbb{R}^p$.

Asymptotics: $\epsilon \rightarrow 0$, $\Delta = \Delta_n \rightarrow 0$ with $T = n\Delta_n$ fixed;

$\Delta = T/n \Rightarrow$ Notation $(\epsilon, \Delta) = (\epsilon, n)$.

- $B_k(\alpha; x) = x(\alpha; t_k) + \Phi(\alpha; t_k, t_{k-1})(x - x(\alpha, t_{k-1}))$
- $S_k(\alpha, \beta) = \int_{t_{k-1}}^{t_k} \Phi(\alpha; t_k, s)\Sigma(\beta; s, x(\alpha; s))\, {}^t\Phi(\alpha; t_k, s)ds$
- Since $\Delta_n \rightarrow 0$, $S_k(\alpha, \beta) \sim \Sigma(\beta, t_{k-1}, X_{t_{k-1}})$. $S_k^\theta \sim \Sigma(\beta, X_{t_{k-1}}) \Rightarrow$

$$U_{\epsilon,n}(\alpha, \beta; X^{(n)}) = \sum_{k=1}^{n} (log(det(\Sigma(\beta, X_{t_{k-1}}) +$$

$$\frac{1}{\epsilon^2 \Delta}\, {}^t(X(t_k) - B_k(\alpha; X(t_{k-1})))\Sigma^{-1}(\beta; X(t_{k-1})(X(t_k) - B_k(\alpha; X(t_{k-1}))).$$

Estimators: $(\hat{\alpha}_{\epsilon,n}, \hat{\beta}_{\epsilon,n})$ such that

$$U_{\epsilon,n}(\hat{\alpha}_{\epsilon,n}, \hat{\beta}_{\epsilon,n}; X^{(n)}) = inf\{U_{\epsilon,n}(\alpha, \beta; (X^{(n)})), (\alpha, \beta) \in \Theta\}$$

Checking additional conditions

(a): $\epsilon^2 U_{\epsilon,n}(\alpha, \beta; X^{(n)}) \to K(\alpha_0, \alpha, \beta)$ a.s. under P_{α_0,β_0} as $\epsilon \to 0, n \to \infty$,
$K(.)$: Deterministic continuous, unique minimum at α_0 for all β
(b) Uniform convergence of (a) for all β.
(c) Uniform bound in P_{θ_0}-probability for $\epsilon^{-1}(\hat{\alpha}_{\epsilon,n} - \alpha_0)$.
(d): Additional condition

$$I_b(\alpha, \beta) = \left(\int_0^T {}^t\frac{\partial b}{\partial \alpha_i}(\alpha; t, x(\alpha, t)) \Sigma^{-1}(\beta; t, x(\alpha, t)) \frac{\partial b}{\partial \alpha_j}(\alpha; t, x(\alpha, t)) dt \right)_{i,j}$$

$$I_\sigma(\alpha, \beta) = \left(\int_0^T Tr\left(\frac{\partial \Sigma}{\partial \beta_k} \Sigma^{-1}\right) \frac{\partial \Sigma}{\partial \beta_l}\right)(\beta; t, x(\alpha, t)) dt \Big)_{k,l}$$

Theorem

Under $P^{\epsilon}_{\alpha_0,\beta_0}$

$$\begin{pmatrix} \epsilon^{-1}(\hat{\alpha}_{\epsilon,n} - \alpha_0) \\ \sqrt{n}\,(\hat{\beta}_{\epsilon,n} - \beta_0) \end{pmatrix} \underset{n\to\infty,\epsilon\to 0}{\longrightarrow} \mathcal{N}\left(0, \begin{pmatrix} I_b^{-1}(\alpha_0, \beta^0) & 0 \\ 0 & I_\sigma^{-1}(\alpha_0, \beta_0) \end{pmatrix}\right)$$

Additional conditions

The function K is equal to
$K(\alpha_0,\alpha,\beta)=\int_0^T {}^t\Gamma(\alpha_0,\alpha;t)\Sigma^{-1}(\beta,x(\alpha_0,t))\Gamma(\alpha_0,\alpha;t)dt$ with
$\Gamma(\alpha_0,\alpha;t)=$
$b(\alpha_0;t,x(\alpha_0,t))-b(\alpha;t,x(\alpha,t))-\nabla_x b(\alpha;t,x(\alpha,t))(x(\alpha_0,t)-x(\alpha,t))$.

Additional condition on U:
Let $V(\alpha_0,\beta_0,\beta)(t)=\Sigma^{-1}(\beta,t,x(\alpha_0,t))\Sigma(\beta_0,tx(\alpha_0,t))$ $K'(\alpha_0,\beta_0,\beta)=$
$\frac{1}{T}\int_0^T Tr(V((\alpha_0,\beta_0,\beta)(t)))dt-\frac{1}{T}\int_0^T logdetV(\alpha_0,\beta_0,\beta)(t))dt-p$.
$|\frac{1}{n}(U_{\epsilon,n}(\hat\alpha_{\epsilon,n},\beta)-U_{\epsilon,n}(\hat\alpha_{\epsilon,n},\beta_0))-K'(\alpha_0,\beta_0,\beta))|\rightarrow 0$ uniformly w.r.t. β.