Bayesian Two-Stage method for estimating ODE parameters

김성민

서울대학교
통계학과, 베이즈통계 연구실

2020. 12. 09

CONTENTS

CONTENTS

ODE model

- Suppose that we have a regression model

$$Y = f_\theta(t) + \epsilon,\ \theta \in \Theta \subseteq \mathbb{R}^p$$

- The functional form of f_θ is not known.
- But the function is assumed to satisfy ODE given by

$$\begin{aligned} \frac{df_\theta(t)}{dt} &= F(t, f_\theta(t), \theta) \\ \text{or} \quad F\left(t, f_\theta(t), \frac{df_\theta(t)}{dt}, \ldots, \frac{d^q f_\theta(t)}{dt^q}, \theta\right) &= 0 \end{aligned}$$

- F is a known real-valued function.

Example

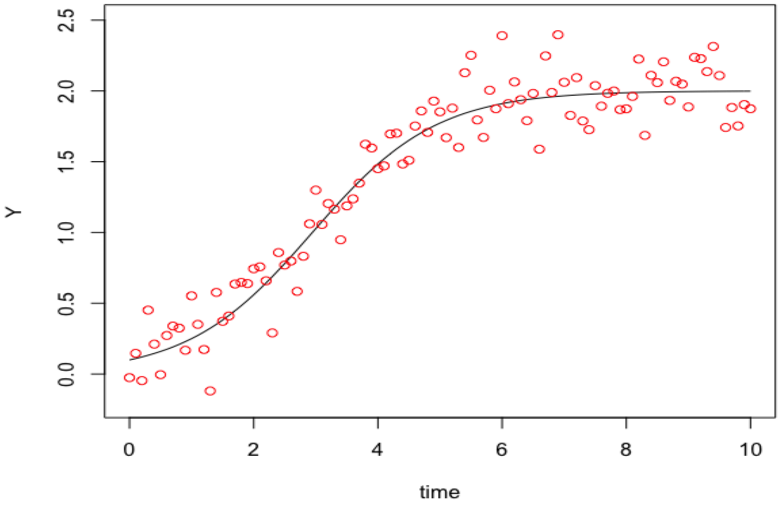
- Consider the following model

$$Y = f_\theta(t) + \epsilon,\ \theta \in \Theta \subseteq \mathbb{R}^2$$

- $f_\theta(t)$ satisfies ODE given by

$$\begin{aligned} \frac{df_\theta(t)}{dt} &= F(t, f_\theta(t), \theta) \\ &= f_\theta(t)(\theta_2 + \theta_1 f_\theta(t)) \end{aligned}$$

Example



CONTENTS

Two-Stage method

- 다음과 같은 모형을 생각한다.

$$Y = f_\theta(t) + \epsilon,$$
$$\frac{df_\theta(t)}{dt} = F(t, f_\theta(t), \theta)$$

- 1단계 : cubic spline 를 이용하여 Y 적합. ($\hat{Y}(t)$)
- 2단계 : $\left|\left|\frac{d\hat{Y}(t)}{dt} - F(t, \hat{Y}(t), \theta)\right|\right|^2$ 를 최소화하는 θ 찾기.

한계점

- 복잡한 모형에 대해서 정확성이 떨어진다.
- 구간 추정을 할 수 없다.

CONTENTS

Bayesian Two-Stage method

- 2단계 방법를 약간 변형
- 구간 추정 가능

1단계

- Penalized B-spline 으로 데이터 적합.
- B : B-spline 기저 행렬 (n by p)
- $\Omega = \ddot{B}^T \ddot{B}$

$$
\begin{aligned}
[\beta|Y] &\propto [Y|\beta][\beta] \\
&\propto \exp\Big[-\frac{1}{2}(Y - B\beta)^T \Sigma^{-1}(Y - B\beta)\Big] \exp\Big[-\frac{1}{2\lambda^2}\beta^T \Omega \beta\Big]
\end{aligned}
$$

$$
\Rightarrow \quad f^{(i)}(t) = B\beta^{(i)} \quad \text{where} \quad \beta^{(i)} \sim [\beta|Y]
$$

- $[\beta|Y] \sim N(\mu_\beta, \Sigma_\beta)$

$$
\text{where} \quad \mu_\beta = (\frac{B^T B}{\sigma^2} + \lambda\Omega)^{-1} B^T Y\,, \quad \Sigma_\beta = (\frac{B^T B}{\sigma^2} + \lambda\Omega)^{-1},
$$

2단계

$$
\begin{aligned}
[\theta|Y] &= \int [\theta, \dot{f}|u]\, d\dot{f} \\
&= \int [\theta|\dot{f}][\dot{f}|Y]\, d\dot{f} \\
&\approx \frac{1}{m}\sum_{i=1}^{m}[\theta|\dot{f}^{(i)}] \\
&\propto \sum_{i=1}^{m}[\dot{f}^{(i)}|\theta][\theta]
\end{aligned}
$$

- 위의 사후분포 계산을 위해서는 $[\dot{f}^{(i)}|\theta]$ 를 알아야 한다.

2단계

- 다음을 가정하자.

$$\dot{f}(t) \sim GP(F(t, f(t), \theta), C)$$

- 다음과 같은 Stochastic diffusion model로 볼 수 있다.

$$df(t) = F(t, f(t), \theta))dt + \sigma dW_t$$

- W_t : Brownian motion process
    - $W_0 = 0$
    - W is a process with stationary independent increments
    - $W_{t+s} - W_s \sim N(0, t)$
    - With probability 1, the function $t \mapsto W_t$ is continuous in t.

2단계

- 다음과 같은 간단한 SDE 를 생각해보자.

$$dX(t) = a(X(t), t)dt + dW_t$$

- Q 측도 : $dX_t = dW_t$
- P 측도 : $dX_t = a(X_t, t)dt + dW_t$
- $L(X) = \frac{dP(X)}{dQ(X)}$

가능도

- $t_k = k\Delta t$, $X_k = X(t_k)$, $\Delta W_k = W_{t_{k+1}} - W_{t_k}$
- $\vec{X} = (X_1, \ldots, X_n)$
- 앞의 SDE 를 다음과 같이 근사할 수 있다.

$$X_{k+1} = X_k + a(X_k, t_k)\Delta t + \Delta W_k$$

가능도

- $V(\vec{X})$: Joint density of $\vec{X}$ for Q

$$V(\vec{X}) = \Pi_k \frac{1}{\sqrt{2\pi\Delta t}} exp(-\frac{1}{2\Delta t}(X_{k+1} - X_k)^2)$$

- $G(\vec{X})$: Joint density of $\vec{X}$ for P

$$G(\vec{X}) = \Pi_k \frac{1}{\sqrt{2\pi\Delta t}} exp(-\frac{1}{2\Delta t}(X_{k+1} - X_k - a(X_k, t_k)\Delta t)^2)$$

$$\begin{aligned} \Rightarrow L(X) &= \lim_{\Delta t \to 0} \frac{G(\vec{X})}{V(\vec{X})} \\ &= \exp\left(\int_0^T a(X(t), t)dX(t) - \frac{1}{2}\int_0^T a^2(X(t), t)dt\right) \end{aligned}$$

가능도

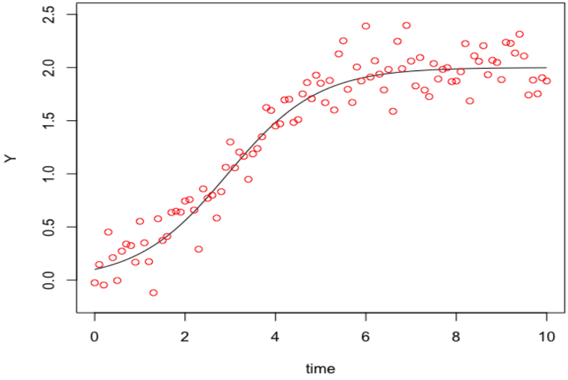
$$
\begin{aligned}
&df(t) = F(t, f(t), \theta)dt + \sigma dW_t \\
\Rightarrow\ &L(\theta|f) = \exp\left(\frac{1}{\sigma^2}\int F(t, f(t), \theta)df(t) - \frac{1}{2\sigma^2}\int F^2(t, f(t), \theta)dt\right)
\end{aligned}
$$

CONTENTS

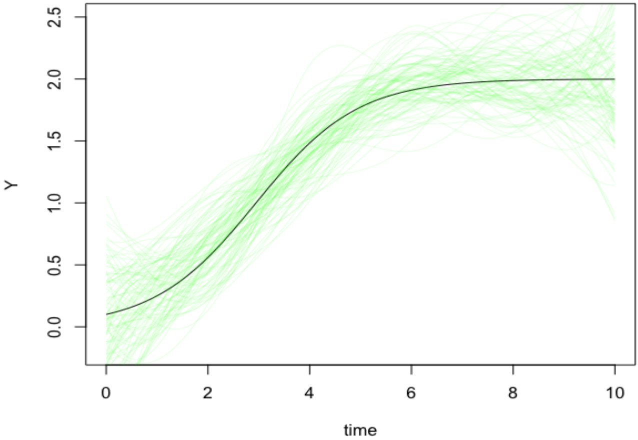
Lotka-Volterra equations

- $\theta_1 = -0.5,\ \theta_2 = 1,\ t \in [0, 10],\ f(0) = 0.1$

$$\frac{df_\theta(t)}{dt} = f_\theta(t)(\theta_2 + \theta_1 f_\theta(t))$$



Lotka-Volterra equations



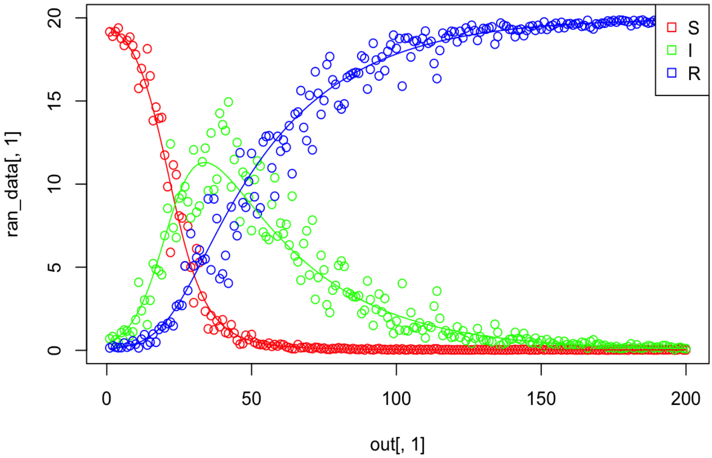
- $\hat{\theta}_1 = -0.487, \ \hat{\theta}_2 = 0.971$

SIR model

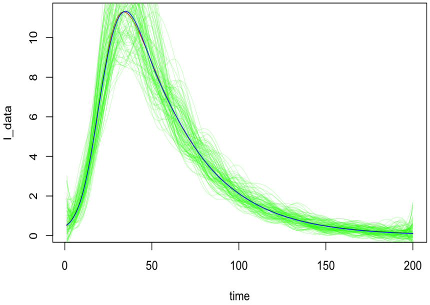
- $\beta = 0.2, \ \gamma = 0.03$
- $S(1) = 19.4, \ I(1) = 0.5, \ R(1) = 0.1$

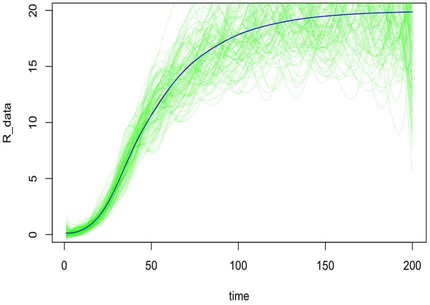
$$\begin{aligned}\frac{dS(t)}{dt} &= -\beta\frac{I(t)}{N}S(t)\\ \frac{dI(t)}{dt} &= \beta\frac{I(t)}{N}S(t) - \gamma I(t)\\ \frac{dR(t)}{dt} &= \gamma I(t)\end{aligned}$$

SIR model



SIR model





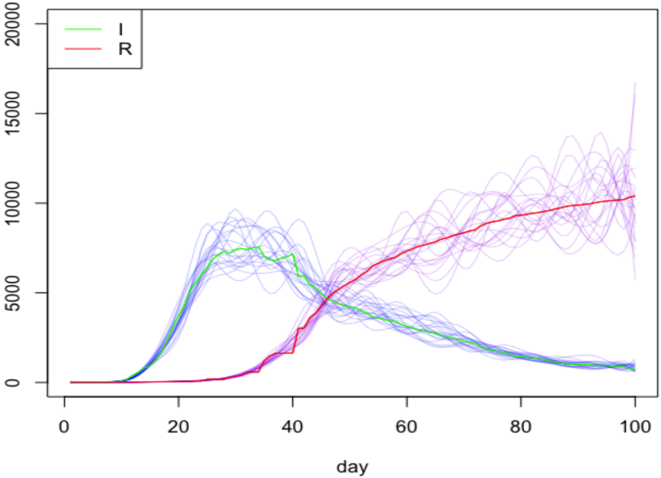
- $\hat{\beta} = 0.1887,\ \hat{\gamma} = 0.0326$

Time-Varying SIR model

$$
\begin{aligned}
\frac{dS(t)}{dt} &= -\beta(t)\frac{I(t)}{N}S(t) \\
\frac{dI(t)}{dt} &= \beta(t)\frac{I(t)}{N}S(t) - \gamma(t)I(t) \\
\frac{dR(t)}{dt} &= \gamma(t)I(t)
\end{aligned}
$$

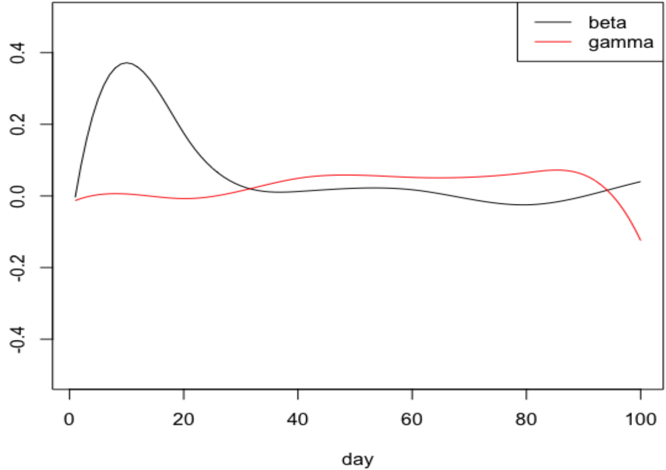
- 2/11 부터 5/20까지의 국내 데이터 이용.
- I, R 의 식을 이용.
- $\beta(t) = B_1(t)\theta_\beta, \quad \gamma(t) = B_1(t)\theta_\gamma$

Time-Varying SIR model

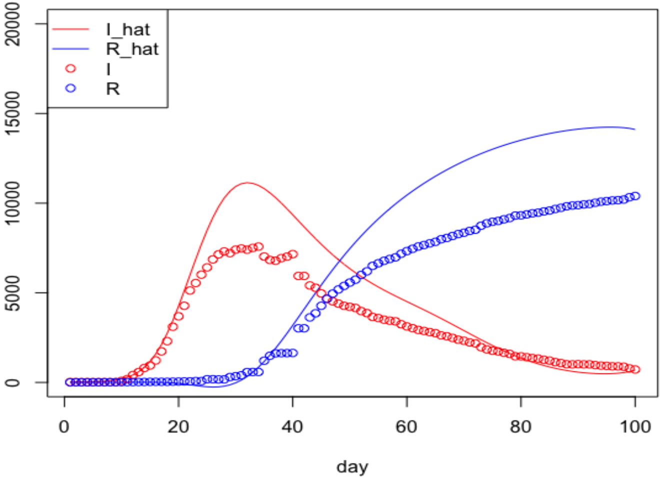


Result

- B1 자유도 : 8

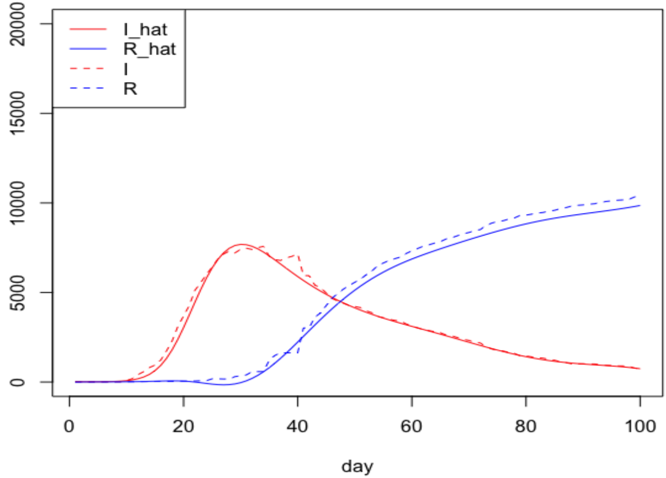


Result



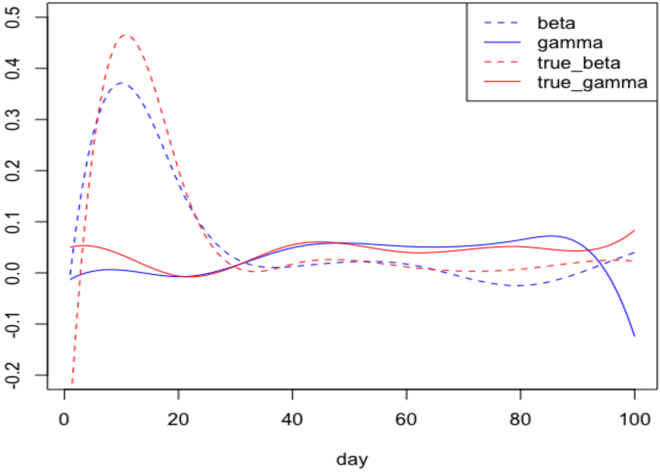
Result

- 실제로는 거의 완벽하게 fitting 이 된다.



Result

- 초기값(day=0) 에 의해서 차이가 발생하는 것으로 보인다.



Remind

$$
\begin{aligned}
[\theta | Y] &= \int [\theta, I, R | Y] \, dI \, dR \\
&= \int [\theta | I, R][I, R | Y] \, dI \, dR \\
&\approx \frac{1}{m} \sum_{i=1}^{m} [\theta | I^{(i)}, R^{(i)}] \\
&\approx \frac{1}{m} \sum_{i=1}^{m} [I^{(i)} | \theta][R^{(i)} | \theta][\theta]
\end{aligned}
$$

Remind

$$[I^{(i)}|\theta] \propto \frac{\exp(-\frac{1}{2\sigma^2}\sum_t (I^{(i)}(t+1) - I^{(i)}(t) - v_I(t))^2)}{\exp(-\frac{1}{2\sigma^2}\sum_t (I^{(i)}(t+1) - I^{(i)}(t))^2)}$$

$$[R^{(i)}|\theta] \propto \frac{\exp(-\frac{1}{2\sigma^2}\sum_t (R^{(i)}(t+1) - R^{(i)}(t) - v_R(t))^2)}{\exp(-\frac{1}{2\sigma^2}\sum_t (R^{(i)}(t+1) - R^{(i)}(t))^2)}$$

where $\quad v_I(t) = \beta(t)\dfrac{I(t)}{N}S(t) - \gamma(t)I(t), \quad v_R(t) = \gamma(t)I(t),$

$$\beta(t) = B_1(t)\theta_\beta, \quad \gamma(t) = B_1(t)\theta_\gamma$$

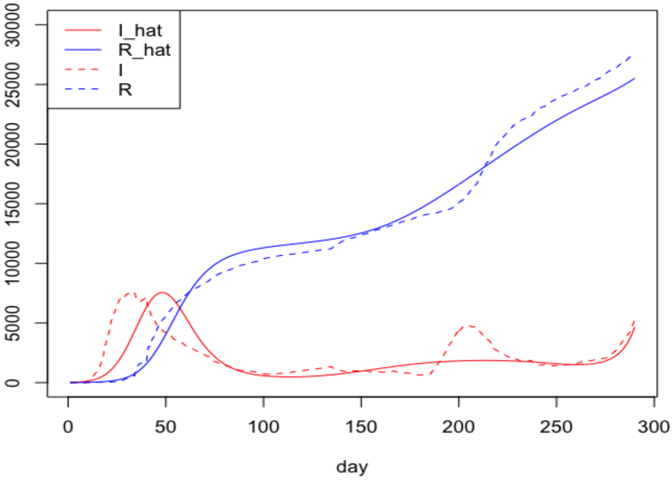
Discussion

- t가 작을 때 모수 추정값이 약간 차이가 발생하면, 이 차이가 t=T 까지 누적된다.
- $\beta(\hat{1})$이 실제보다 크게 추정되어서 integration 한 $\hat{I}(2)$가 실제 $I(2)$보다 커지면 그 뒤의 모수가 정확히 추정되어도 integration 결과가 왜곡될 수 있다.

- 적은 수의 기저를 사용해서 데이터를 근사하는 것에 어려움이 있다.
- 1단계에서의 샘플 추출과 $v(t)$에서 $I(t), S(t)$ 의 선택을 수정해야 한다.

Approximation

- 기저 10개를 이용하여 실제 데이터를 근사하였다.



Reference I

Bhaumik, P. and Ghosal, S. (2017).
Efficient Bayesian estimation and uncertainty quantification in ordinary differential equation models.
Bernoulli, 23 3537–3570.

Varah, J. M. (1982).
spline least squares method for numerical parameter estimation in differential equations
SIAM J. Scient.Computn, 3, 28–46.

Huzak, M. (1998)
Parameter estimation of diffusion models
Math. Commun. 3, 129– 134.