

Module 11: Lesson 2 Lecture Notes

Callum Arnold

20 July, 2021

Contents

SIR models	1
The modelling process	1
Gamma distribution example	1
Updating separately	2
Block update	2
General (Markov) SIR example	2
Augmented likelihood	3

SIR models

The modelling process

- By Bayes' theorem $\pi(\theta|y) \propto \pi(y|\theta)\pi(\theta)$
 - Posterior \propto likelihood \times prior
- Posterior usually not known explicitly (i.e. only know it up to proportionality)
 - Use MCMC algorithm to get samples
- Typically θ is multidimensional, so we need a way of updating each θ_k
 - If the full conditional density is known explicitly (not just up to proportionality), then we can use it to perform the update for θ_k (known as the **Gibbs step - part of Gibbs sampling**)
 - * $\pi(\theta_k|\theta_1, \dots, \theta_{k-1}, \theta_{k+1}, \dots, \theta_n, y)$
 - If the full conditional density is not known explicitly/can be easily sampled from, use M-H step
- What if the likelihood is unknown (hard/complicated/impossible to compute)
 - Data augmentation
 - * Introduce extra quantities x , such that $\pi(x, y|\theta)$ is tractable
 - Usually to do with infection process that is normally unobserved
 - Give up on MCMC and use something else e.g. Approximate Bayesian Computation

Gamma distribution example

- Suppose we have data on incubation periods $y = y_1, \dots, y_n$ and we wish to fit a Gamma distribution to these data (y_i) is the incubation period for person i
- The Gamma distribution has the probability density function
 - $f(x|\alpha, \beta) = \beta^\alpha x^{\alpha-1} \exp(-\beta x) / \Gamma(\alpha)$ where $x > 0, \alpha > 0, \beta > 0$

- We can write this proportionally as $f(x|\alpha, \beta) \propto x^{\alpha-1} \exp(-\beta x)$
- Assume data are independent draws from this distribution, therefore the likelihood is:

$$\begin{aligned}\pi(y|\alpha, \beta) &= f(y_1|\alpha, \beta) \times \dots \times f(y_n|\alpha, \beta) \\ &= \beta^{n\alpha} \prod_k y_k^{\alpha-1} \exp(\beta \sum y_k) / \{\Gamma(\alpha)\}^n\end{aligned}$$

- We assign independent priors as
 - $\alpha \sim \text{Gamma}(\lambda_\alpha, v_\alpha)$
 - $\beta \sim \text{Gamma}(\lambda_\beta, v_\beta)$
 - * Use Gamma prior as:
 - Flexible so can be very informative/uninformative as we wish
 - Mathematical convenience - for β in particular, gives nice way of evaluation full conditional density
- We can get the posterior density of interest
 - $\pi(\alpha, \beta|y) \propto \pi(y|\alpha, \beta)\pi(\alpha)\pi(\beta)$
 - Define MCMC algorithm to sample from this target density to update α and β

Updating separately

- Find the full conditional densities
 - $\pi(\alpha|\beta, y)$ and $\pi(\beta|\alpha, y)$
 - $\pi(\alpha|\beta, y)$ is only known up to proportionality
 - * Therefore have to update α using a Metropolis-Hastings step
 - * $p = \frac{\pi(y|\alpha^*, \beta)\pi(\alpha^*)q(\alpha|\alpha^*)}{\pi(y|\alpha, \beta)\pi(\alpha)q(\alpha^*|\alpha)}$
 - $\pi(\beta|\alpha, y)$ is the density of a Gamma distribution
 - * Can update using a Gamma distribution

Block update

- Instead of updating separately, put them in a block and update together
- Use M-H and propose (α^*, β^*) from $q(\alpha^*, \beta^*|\alpha, \beta)$ and accept/reject accordingly
- May want to do if α and β are strongly correlated
 - Hard to move one without the other

General (Markov) SIR example

- Suppose we observe n removals at times $r_1 \leq r_2 \leq \dots \leq r_n$
- We want to estimate β and γ
 - Posterior density $\pi(\beta, \gamma|r_1, r_2, \dots, r_n)$
- The likelihood is very hard to compute
 - $\pi(r_1, r_2, \dots, r_n|\beta, \gamma)$
 - Have to think about all possible ways removals could happen (don't know when infected, or the order of the events e.g. inf, inf, inf, rem, rem, rem etc.)

- Introduce infection times as extra variables to give tractable augmented likelihood

Augmented likelihood

- Let b be the label of the last removal time i.e. $r_b \geq r_k$ for all $k = 1, \dots, n$
- Given removal data, b is observed and fixed
- Define a as the label of the first infection time $i_a < i_k$ for all $k \neq a$
- Given removal data, a is unknown
- Define:
 - $\vec{r} = r_1, \dots, r_n$
 - $\vec{i} = i_1, \dots, i_{a-1}, i_{a+1}, \dots, i_n$
 - * Note that it does not include i_a , which is the time of infection for the first infected individual a
- Let $f(x|\gamma) = \gamma \exp(-\gamma x) \quad x > 0$
 - Probability density function of the infectious period distribution ($\text{Exp}(\gamma)$)
- The augmented likelihood:

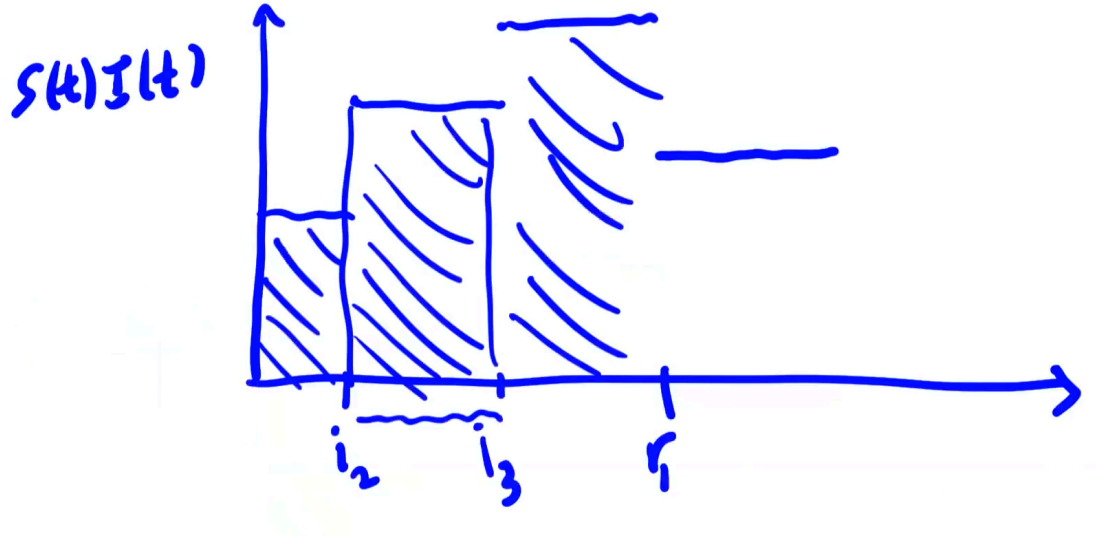
$$\begin{aligned} \pi(\vec{i}, \vec{r} | \beta, \gamma, i_a, a) &= \prod_{j \neq a} \beta N^{-1} I(i_j-) \times \exp\left(-\beta N^{-1} \int S(t) I(t) dt\right) \times \prod_{1 \leq j \leq n} f(r_j - i_j | \gamma) \\ &= \underbrace{\prod_{j \neq a} \beta N^{-1} I(i_j-) \times \exp\left(-\beta N^{-1} \int S(t) I(t) dt\right)}_{\text{how likely that individuals get infected at inf times observed}} \times \underbrace{\gamma^n \exp\left(-\gamma \sum (r_j - i_j)\right)}_{\text{how likely to observe removals}} \end{aligned}$$

- where $I(t-)$ means $I(t)$ **just before** time t
- The first part of the augmented likelihood can be derived by looking at the times to next event distribution

$$\begin{aligned} T_i &\sim \text{Exp}(\beta SI/N + \gamma I) \\ p(\text{inf}) &= \frac{\beta SI/N}{\beta SI/N + \gamma I} \\ p(\text{rem}) &= \frac{\gamma I}{\beta SI/N + \gamma I} \\ \text{given } X &\sim \text{Exp}(\lambda) \rightarrow f(x) = \lambda \exp(-\lambda x) \\ f(T) &= \frac{\beta SI/N}{\beta SI/N + \gamma I} \exp(-(\beta SI/N + \gamma I)(i_2 - i_1)) \times \frac{\beta SI/N}{\beta SI/N + \gamma I} \\ f(T) &= \exp(-(\beta SI/N + \gamma I)(i_2 - i_1)) \times \beta SI/N \end{aligned}$$

- This represents the likelihood that the next event after i_1 is an infection and occurs at time i_2
 - This can be multiplied for all successive infection event times, which gives the product term
 - The term $(i_2 - i_1)$ is specifying the time until the infection
 - $(\beta SI/N + \gamma I) \exp(-(\beta SI/N + \gamma I)(i_2 - i_1))$ is how likely it is to draw the time $i_2 - i_1$ from the exponential infection time distribution
 - $\frac{\beta SI/N}{\beta SI/N + \gamma I}$ is the likelihood that the event is an infection, if an infection was drawn

- * If the next event was a removal, the final term would instead be $\frac{\gamma I}{\beta SI/N + \gamma I}$ and we would instead have $(r_1 - i_1)$
- The integral symbol comes because we are actually calculating the area when multiplying $\beta SI/N$ in the exponential by the time difference $(i_2 - i_1)$



- There's an analogous one for the removal times
- **NB:** S and I refer to the numbers in the respective compartments at time t
 - We end up dropping the S and I terms due to the differences between the simulation and calculating the likelihood
 - * The simulation describes the number of susceptibles and the number of infecteds
 - * Calculating the likelihood, we label the individuals already e.g. when infection occurs, we say it's a particular individual who gets infected
 - Results in slight difference
- The target posterior density $\pi(\beta, \gamma, \vec{i}, i_a, a | \vec{r}) \propto \pi(\vec{i}, \vec{r} | \beta, \gamma, i_a, a) \pi(\beta, \gamma, i_a, a)$
 - Define independent priors

$$\alpha \sim \text{Gamma}(\lambda_\alpha, \nu_\alpha)$$

$$\beta \sim \text{Gamma}(\lambda_\beta, \nu_\beta)$$

$$a \sim \text{U}[1, n]$$

$$i_a \sim \text{U}[-\infty, r_1]$$

- **NB:** we know the first infection must happen before the first removal, but we don't know exactly when, hence $\text{U}[-\infty, r_1]$
- If we can find the first full conditional distribution, we can update using that if it's a standard distribution, and use MH if not
 - We know the posterior for β is a Gamma distribution, which we can sample from
 - * Sampling directly from the full conditional distribution is often called the **Gibbs step**
 - * We can show this in the following steps

First let's compute the full joint density function up to proportionality. To do this, we will use Bayes' theorem, and the specification of the likelihood $\pi(y|\alpha, \beta)$, which stems from the probability density function of the Gamma distribution (see Section 1). From here, we can multiply the likelihood by the prior (a Gamma distribution specified above), to get the posterior (up to proportionality).

$$\begin{aligned}
\pi(\alpha, \beta|y) &\propto \pi(y|\beta, \dots) \cancel{\pi(\alpha)} \pi(\beta) \\
\pi(y|\alpha, \beta) &\propto \prod_{j \neq a} \frac{\beta}{N} I(i_j -) e^{-\frac{\beta}{N} \int S(t) I(t) dt} \gamma^n e^{-\gamma \sum (r_j - i_j)} \\
\pi(\alpha, \beta|y) &\propto \prod_{j \neq a} \cancel{\frac{\beta}{N} I(i_j -)} e^{-\frac{\beta}{N} \int S(t) I(t) dt} \cancel{\gamma^n e^{-\gamma \sum (r_j - i_j)}} \times \frac{\nu_\beta^{\lambda_\beta}}{\Gamma(\lambda_\beta)} \beta^{\lambda_\beta - 1} e^{-\nu_\beta \beta} \\
&\propto \beta^n e^{-\frac{\beta}{N} \int S(t) I(t) dt} \times \beta^{\lambda_\beta - 1} e^{-\nu_\beta \beta} \\
&\propto \beta^{n + \lambda_\beta - 1} e^{-\beta(\frac{1}{N} \int S(t) I(t) dt + \nu_\beta)} \\
&\quad \vdots \\
f(\beta|\dots) &\sim \Gamma\left(n + \lambda_\beta - 1, \frac{1}{N} \int S(t) I(t) dt + \nu_\beta\right)
\end{aligned}$$

- γ posterior is also Gamma distribution so can use Gibbs steps
- For the infection times, the full conditional distribution is non-standard
 - Update using M-H algorithm with M-H acceptance ratio
 - * We're mainly interested in β and γ , so may just want to update 10% of the infection times between each update of β and γ , otherwise may be too much computational work
 - * ratio of the likelihood x ratio of proposal densities
 - * Accept with M-H acceptance ratio:

$$A_{i_k i_k^*} = \min \left\{ \frac{\pi(\vec{i}^*, \vec{r}|\beta, \gamma, i_{a^*}, a^*)}{\pi(\vec{i}, \vec{r}|\beta, \gamma, i_a, a^*)} \cdot \frac{q(i_k|i_k^*)}{q(i_k^*|i_k)}, 1 \right\}$$

- * **Note:** If $i_k^* < i_a$ then $a^* = k$ (update identity of the initial infective), otherwise a is unchanged (first infected individual)
- How to propose infection times
 - Propose a new time using proposal density $q(i_k^*|i_k)$
 - Must be before removal time
 - Multiple options:
 - * $i_k^* = r_k - \text{Exp}(\gamma)$
 - * $i_k^* = r_k - \text{Exp}(\mu)$ where μ is fixed
 - * $i_k^* \sim \mathcal{N}(i_k, \sigma^2)$
- Evaluating the likelihood
 - $\pi(\vec{i}, \vec{r}|\beta, \gamma, i_a, a) = \prod_{j \neq a} \beta N^{-1} I(i_j -) \times \exp(-\beta N^{-1} \int S(t) I(t) dt) \times \gamma^n \exp(-\gamma \sum (r_j - i_j))$
 - The product term and the integral term are the difficult parts to evaluate
 - Product term
 - * $\prod_{j \neq a} \frac{\beta}{N} I(i_j -) = \left(\frac{\beta}{N}\right)^{n-1} \prod_{j \neq a} I(i_j -)$

- * Can't simplify product term any more
- * Need to write code to evaluate the number of infectives at each infection event
 - Calculate number at each event
- Integral term
 - * $\int S(t)I(t)dt = \sum_{1 \leq k \leq n} \sum_{1 \leq j \leq N} [(r_k \wedge i_j) - (i_k \wedge i_j)]$ where: $a \wedge b = \min(a, b)$
 - j is over all individuals in the population, and k goes across only those who are/were ever infected
 - $i_j = \infty$ for $j > n$ i.e. individuals never infected
 - * Can derive this term by rewriting the number of susceptibles and infectives by using sums of indicator functions (1 if conditions met, 0 otherwise)
 - **Note:** For $I(t)$ indicator function, we only look across **all** individuals who have ever been infected, and count those who are presently infected (i.e. we're at a time after they've been infected and before they've been remove)
 - * Can imagine the integral as the total amount of time during infected individual is ability to infect a susceptible individual, summed over all pairs, i.e. total time available for infection

$$\begin{aligned}
 S(t) &= \sum_{1 \leq j \leq N} 1_{\{j \text{ is susceptible at time } t\}} \\
 &= \sum_{1 \leq j \leq N} 1_{\{i_j < t\}} \\
 I(t) &= \sum_{1 \leq k \leq n} 1_{\{i_k < t < r_k\}} \\
 \int S(t)I(t)dt &= \int \sum_{1 \leq k \leq n} \sum_{1 \leq j \leq N} 1_{\{i_j < t\}} 1_{\{i_k < t < r_k\}} dt \\
 &= \sum_{1 \leq k \leq n} \sum_{1 \leq j \leq N} \int 1_{\{i_k < t < r_k \text{ and } i_j < t\}} dt \\
 &= \sum_{1 \leq k \leq n} \sum_{1 \leq j \leq N} [(r_k \wedge i_j) - (i_k \wedge i_j)] \\
 \text{as } 1_{\{i_j < t\}} 1_{\{i_k < t < r_k\}} &= 1_{\{i_k < t < r_k \text{ and } i_j < t\}} **
 \end{aligned}$$

** **Note:** This line is true because if either one of the terms = 0, the product = 0

In practise, our MCMC algorithm would look like this:

- Initialize $\beta, \gamma, \vec{i}, i_a, a$
- Loop:
 - Update β
 - Update γ
 - Update a fraction of the infection times i_k
 - Record current values of β, γ and a summary statistic of infection times e.g. sum of the infection times or infectious period
 - * Helps make sure there isn't an error in the code, for example, infection times next update
- The output is a sequence $(\beta_1, \gamma_1), (\beta_2, \gamma_2), \dots, (\beta_M, \gamma_M)$ where M is the number of iterations in loop