MCMC 2: Lecture 2
Coding and output

Phil O'Neill Theo Kypraios
School of Mathematical Sciences
University of Nottingham

Contents

- 1. General (Markov) epidemic model
- 2. Non-Markov epidemic model
- 3. Debugging tips
- 4. What to do with MCMC output

Contents

1. General epidemic model

Population of N individuals

At time t there are:

S_t susceptibles

I_t infectives

R_t recovered / immune individuals

Thus $S_t + I_t + R_t = N$ for all t.

Initially $(S_0, I_0, R_0) = (N-1,1,0)$.

1. General epidemic model

- Each infectious individual remains so for a length of time $T_I \sim \text{Exp}(\gamma)$.
- During this time, infectious contacts occur with each susceptible according to a Poisson process of rate β / N.
- Thus overall infection rate is $\beta S_t I_t / N$.
- Two model parameters, β and γ.

1. General epidemic model

- Let infection times be

$$i_1 \leq i_2 \leq i_3 \leq \ldots \leq i_n,$$

where i_1 is time that initial infective begins their infectious period.

- Note that n = total number infected.
- Define $\mathbf{i} = (i_2, i_3, \ldots, i_n)$

1. General epidemic model

- Let removal times be

$$r_1 \leq r_2 \leq r_3 \leq \ldots \leq r_n.$$

- Note that kth infection time need not correspond to the kth removal.

- Define $\mathbf{r} = (r_1, r_2, \ldots, r_n)$

1. General epidemic model

- Recall the standard inference problem: we observe removal times and wish to perform Bayesian inference for β and γ.

- Solution [as discussed in MCMC I] is to use MCMC, treat missing infection times i_1 and $\mathbf{i} = (i_2, i_3, \ldots, i_n)$ as "latent" variables.

1. General epidemic model

- The target posterior density is

$$\pi(\beta, \gamma, i_1, i \mid r) \propto \pi(i, r \mid \beta, \gamma, i_1)\, \pi(\beta, \gamma, i_1)$$

$$\text{posterior} \propto \text{likelihood} \times \text{prior}$$

1. General epidemic model

- Recall the likelihood:

π (i, r | β , γ, i_1) =

$$\left(\prod_{j=2}^{n}(\beta/N)S_{ij}I_{ij}\right)\left(\prod_{j=1}^{n}\gamma I_{rj}\right)\exp\left(-\int_{i_1}^{r_n}\{(\beta/N)S_tI_t+\gamma I_t\}dt\right)$$

product terms integral term

Note that S_{ij} here means S <u>just before</u> time i_j.

1. General epidemic model

- Recall: if β and $\gamma \sim$ Gamma *a priori* then both have Gamma full conditional distributions.
- e.g.

$$\beta \mid \gamma, i_1, i, r \sim \Gamma\left(m_\beta + n - 1, \lambda_\beta + N^{-1} \int SI \right)$$

where $\beta \sim \Gamma(m_\beta , \lambda_\beta)$ *a priori* .

- Thus β and γ can be updated using a “Gibbs step” - i.e. according to their full conditional distributions - during an MCMC algorithm.

1. General epidemic model

- Recall that the unknown infection times are updated using a Metropolis-Hastings step.
- The acceptance probability requires us to calculate

$$\text{posterior} \propto \text{likelihood} \times \text{prior}$$

1. General epidemic model

- Thus to write an MCMC algorithm, it is necessary to be able to evaluate both the product and (integral) terms in the likelihood.

1. General epidemic model

Product terms

- First note that

$$\left(\prod_{j=2}^{n}\beta N^{-1}\mathrm{S}_{\mathrm{ij}}\mathrm{I}_{\mathrm{ij}}\right)\left(\prod_{j=1}^{n}\gamma \mathrm{I}_{\mathrm{rj}}\right)\propto \beta^{n-1}\gamma^{n}\left(\prod_{j=2}^{n}\mathrm{S}_{\mathrm{ij}}\mathrm{I}_{\mathrm{ij}}\right)\left(\prod_{j=1}^{n}\mathrm{I}_{\mathrm{rj}}\right)$$

$$=\beta^{n-1}\gamma^{n}(\mathrm{N}-1)(\mathrm{N}-2)\ldots(\mathrm{N}-\mathrm{n}+1)\left(\prod_{j=2}^{n}\mathrm{I}_{\mathrm{ij}}\right)\left(\prod_{j=1}^{n}\mathrm{I}_{\mathrm{rj}}\right)$$

1. General epidemic model

Product terms

- Thus only the products of numbers of infectives may potentially change when updating the infection times.
- The product is most easily evaluated “directly”, i.e. by keeping track of changes to I_t at each infection or removal event.

1. General epidemic model

Integral terms

- $\int I_t \, dt = \Sigma_{1 \le k \le n} (r_k - i_k)$
- $\int S_t I_t \, dt = \Sigma_{1 \le k \le n} \Sigma_{1 \le j \le N} [(r_k \wedge i_j) - (i_k \wedge i_j)]$

Here, "$a \wedge b$" denotes "minimum of a,b".

Also $i_j = \infty$ for $j > n$, i.e. for those individuals never infected.

1. General epidemic model

Integral terms

Explanation:

$$\int I_t \, dt = \int \Sigma_{1 \le k \le n} \, 1_{\{k \text{ is infective at time } t\}} \, dt$$

$$= \Sigma_{1 \le k \le n} \int 1_{\{k \text{ is infective at time } t\}} \, dt$$

where $1_A = 1$ if event A occurs

$= 0$ otherwise

1. General epidemic model

Integral terms

Recall that individual k is removed at r_k .

Suppose their infection time is $i_{L(k)}$.

Then $\int 1_{\{k \text{ is infective at time } t\}}\ dt$

= total time k is infective

= $(r_k - i_{L(k)})$

1. General epidemic model

Integral terms

Then $\int I_t \, dt = \Sigma_{1 \le k \le n} (r_k - i_{L(k)})$

$= \Sigma_{1 \le k \le n} r_k - \Sigma_{1 \le k \le n} i_{L(k)}$

$= \Sigma_{1 \le k \le n} r_k - \Sigma_{1 \le k \le n} i_k$

$= \Sigma_{1 \le k \le n} (r_k - i_k)$

1. General epidemic model

Integral terms

- $\int S_t I_t \, dt = \Sigma_{1 \le k \le n} \Sigma_{1 \le j \le N} [(r_k \wedge i_j) - (i_k \wedge i_j)]$

Similar arguments used to derive this…

1. General epidemic model

Integral terms

$[(r_k \wedge i_j) - (i_k \wedge i_j)]$ = time that k is infective and j is susceptible

$[(r_k \wedge i_j) - (i_k \wedge i_j)] = r_k - i_k$ if $i_j > r_k$ (e.g. $i_j = \infty$)

$[(r_k \wedge i_j) - (i_k \wedge i_j)] = i_j - i_k$ if $i_k < i_j < r_k$

$[(r_k \wedge i_j) - (i_k \wedge i_j)] = 0$ if $i_j < i_k$

Contents

- 1. General (Markov) epidemic model
- 2. Non-Markov epidemic model
- 3. Debugging tips
- 4. What to do with MCMC output

2. Non-Markov epidemic model

- We now consider a generalisation of the basic (Markov) SIR model in which the infectious periods are no longer exponentially distributed.
- This apparently minor change has a material impact on the likelihood calculations.

2. Non-Markov epidemic model

- Each infectious individual remains so for a length of time T_I drawn from some specified distribution with parameter vector θ
- During this time, infectious contacts occur with each susceptible according to a Poisson process of rate β / N.
- Thus overall infection rate is $\beta S_t I_t / N$.
- Two model parameters, β and θ.

2. Non-Markov epidemic model

Likelihood

- Assume population contains N individuals of whom n ever become infective.
- Label the n infectives 1, 2, ..., n and the other individuals n+1, n+2, ..., N.
- Define r_k and i_k as the removal and infection times of individual k. Note these = ∞ if k never becomes infected.

2. Non-Markov epidemic model

Likelihood

- Let b be the label of the last removal time, i.e. $r_b \geq r_k$ for all k = 1, ..., n.
- Given removal data, b is observed and fixed for any given labelling.
- Define a as the label of the first infection time, i.e. $i_a \leq i_k$ for all k = 1, ..., n.
- Given removal data, a is unknown.

2. Non-Markov epidemic model

Likelihood

- Define $\mathbf{r} = (r_1, r_2, \ldots, r_n)$
- Define $\mathbf{i} = (i_1, i_2, \ldots, i_{a-1}, i_{a+1}, \ldots, i_n)$

- Let $f(x \mid \theta)$ denote the probability density function (or mass function if appropriate) of the infectious period distribution with parameter vector θ.

2. Non-Markov epidemic model

Likelihood

- π (i, r | β , θ, a, i_a) =

$$\left(\prod_{j=2}^{n} (\beta / N) \mathrm{I}_{ij} \right) \exp\left(- \int_{i_a}^{r_b} \{ (\beta / N) \mathrm{S}_t \mathrm{I}_t \} \mathrm{d}t \right) \prod_{j=1}^{n} f(r_j - i_j \mid \theta)$$

2. Non-Markov epidemic model

Bayesian inference

- π (β , θ, a, i_a , i | r)

  $\propto \pi$ (i, r | β , θ, a, i_a) π (β , θ, a, i_a)

- Thus we must specify a prior distribution for β , θ, a, and i_a .

2. Non-Markov epidemic model

MCMC algorithm

- β is updated as for the Markov model (i.e. Gibbs step, assuming β has a Gamma prior)
- Infection times updated using a M-H step. One option is to propose $(r_k - i_k)$ from distribution of infectious period.
- θ updates depend on particular choice of infectious period distribution.

Contents

- 1. General (Markov) epidemic model
- 2. Non-Markov epidemic model
- 3. Debugging tips
- 4. What to do with MCMC output

3. Debugging tips

1. Test each piece of code separately

Most MCMC algorithms in this field involve various components, e.g.

- Gibbs updates
- Metropolis-Hastings updates
- Likelihood

It is good practice to check each component works before proceeding.

3. Debugging tips

2. Validate output using simulations

As discussed in Lecture 1, one way to test MCMC code (e.g. for SIR model) is

- Simulate SIR model M times (e.g. M=1000)
- Run MCMC on each output to infer parameters
- Average parameter estimates from MCMC should be close to the known true values

3. Debugging tips

2. Validate output using simulations

If the MCMC code is time-consuming to run then an alternative is use simulation output that gives a single large epidemic - idea being that this should give reasonable information about the model parameters.

3. Debugging tips

3. Beware Zeroes

Some languages allow operations such as "0/0" without reporting an error.

3. Debugging tips

4. Try a very small data set

Sometimes it is possible to test MCMC code by using a very small data set where one can work out the required inference by hand. This can then be checked against the MCMC output.

3. Debugging tips

5. Use log likelihood

Many likelihoods require calculation of products which can in turn lead to numerical instabilities and run-time errors.

One way to tackle this issue is to instead work with the log likelihood, since

$$\log(A_1 \times A_2 \times \ldots \times A_m) = \log(A_1) + \ldots + \log(A_m)$$

3. Debugging tips

5. Use log likelihood (cont)

The likelihood may involve the calculation of Beta or Gamma functions.

R has built-in functions to compute such functions, i.e. `beta`, `gamma`; but if we are working on the log scale, instead of doing something like `log(gamma(k))` we could use another built in function `lgamma(k)` to ensure numerical stability, especially if `k` is large.

Contents

- 1. General (Markov) epidemic model
- 2. Non-Markov epidemic model
- 3. Debugging tips
- 4. What to do with MCMC output

4. What to do with MCMC output

In this section, for illustration it is assumed that we have MCMC output from the Markov SIR model removal-data-observed scenario:

$$(\beta_1, \gamma_1), (\beta_2, \gamma_2), \ldots, (\beta_M, \gamma_M),$$

where M is large (e.g. $M=10^6$).

Each pair (β_k, γ_k) is (approx) a sample from the joint posterior density $\pi(\beta, \gamma \mid r)$.

4. What to do with MCMC output

Marginal summaries

Quantities such as the marginal mean, median, variance etc of β and of γ can be readily obtained using the package R.

It is also useful to plot the marginal posterior density of each parameter.

4. What to do with MCMC output

Joint summaries

It can be useful to assess the extent to which β and γ can be estimated separately.

The posterior correlation and a scatterplot of the samples against axes β and γ provide such information.

4. What to do with MCMC output

Functions of model parameters

The quantity R_0 is of enormous interest in mathematical epidemic theory. It is (roughly) defined as the average number of secondary cases caused by a typical infective in an infinitely large population of susceptibles.

If $R_0 \leq 1$, epidemics are unlikely to take off.

4. What to do with MCMC output

Functions of model parameters

For the (general) SIR model,

$$R_0 = \beta E(T_I),$$

where $E(T_I)$ is the mean infectious period.

This follows from the fact that each infective causes new infections at (Poisson) rate β during a period of time T_I.

4. What to do with MCMC output

Functions of model parameters

For the Markov model we have

$$R_0 = \beta\, E(T_I) = \beta / \gamma\,,$$

since $T_I \sim \text{Exp}(\gamma)$.

Thus given the MCMC output we can create a new file containing

$$(\beta_1 / \gamma_1), (\beta_2 / \gamma_2), \ldots, (\beta_M / \gamma_M),$$

i.e. samples from the posterior density of R_0.

4. What to do with MCMC output

Functions of model parameters

R_0 can be summarised in the usual ways (mean, variance etc): also interesting to find the posterior probability that $R_0 \leq 1$.

4. What to do with MCMC output

Functions of model parameters

Can also be interesting to translate inference for rates into inference for probabilities.

e.g. 1 - exp(- β / N) is the probability that one infective individual infects a given susceptible in one time unit.