Simulation

Platon
Tsonogeorgis
P3180088

Assignment 8

① The first part is implemented almost entirely via a built in library which calculates successfully the beta hats and the $\hat{\Sigma}$ (7x7) covariance matrix.
The most "manual" operation on the first part is to successfully load the data but this does not concern the theoretical part of the project.

② The second part is Random Walk Metropolis. Since we have analyzed the algorithm in previous projects, I will just enumerate the differences in this case.

• The target function is the log likelihood. The basic formula is this:

$$\prod_{i=1}^{24} \frac{e^{-\mu_i} \cdot \mu_i^{y_i}}{y_i}, \quad \mu_i = \exp\{\beta_1 + \beta_2 x_2 + \ldots + \beta_7 \cdot x_7\}$$

If we get the log of this then we will get:

$$\log\left(\frac{e^{-\mu_1} \cdot \mu_1^{y_1}}{y_1!} \cdot \frac{e^{-\mu_2} \cdot \mu_2^{y_2}}{y_2!} \cdot \ldots \cdot \frac{e^{-\mu_{24}} \cdot \mu_{24}^{y_{24}}}{y_{24}!}\right)$$

$$\log\left(\frac{e^{-\mu_1} \cdot \mu_1^{y_1}}{y_1!}\right) + \log\left(\frac{e^{-\mu_2} \cdot \mu_2^{y_2}}{y_2!}\right) + \ldots + \log\left(\frac{e^{-\mu_{24}} \cdot \mu_{24}^{y_{24}}}{y_{24}!}\right)$$

We can skip the calculations of the factorials by removing them since its just a constant.

$$\log\left(e^{-\mu_1} \cdot \mu_1^{y_1}\right) + \log\left(e^{-\mu_2} \cdot \mu_2^{y_2}\right) + \ldots + \log\left(\frac{e^{-\mu_{24}} \cdot \mu_{24}^{y_{24}}}{\cancel{y_{24}}}\right)$$

$$\left(-\mu_1 + y_1 \cdot \log(\mu_1)\right) + \left(-\mu_2 + y_2 \cdot \log(\mu_2)\right) + \ldots + \left(-\mu_{24} + y_{24} \cdot \log(\mu_{24})\right)$$

So our target function will be this sum, which is directly input in this form because operations like these create overflow errors and the numbers are too big/small.

- The proposal function q is the multivariate normal distribution and we use the covariance matrix we got from part 1 to scale it. The constant c is 1.02. This is because it delivers acceptance rate 23% which is optimal.
- Finally, we calculate the ratio via:

$$ratio = p(Y) - p(X)$$

due to the log likelihood and accept when:

$$a > \log(u)$$

where a is the $\min\{0, ratio\}$

③ The third part is Independent Metropolis. The implementation is almost identical with the second part.

- The target function is exactly the same
- The proposal function q is the multivariate student-t distribution, since the multivariate normal does not cover entirely the target distribution, probably due to its thin tails.
- Also, the ratio now includes q:

$$ratio = p(Y) + q(X) - p(X) - q(Y)$$

The rest of the details are identical.

④ The last part is Gibbs sampling. The idea is that we get a sample from a 1D-distribution of the form:

$$p(b0 \mid b1, \ldots, b7)$$
$$p(b1 \mid b0', b2, \ldots, b7)$$
$$\vdots$$
$$p(b7 \mid b0', \ldots, b6')$$

The tricky part is that we don't know how to take (easily at least) a sample from these distributions. Each of the above distributions is of the form:

Example for $p(b2 \mid b0', b1', \ldots, b7)$

$$\sum_{i=1}^{24} \left(-e^{-x_2 \cdot b_2 + \lambda} + y_i (x_2 \cdot b_2 + \lambda) \right) \quad (1)$$

where everything is constant besides b_2 which is the variable of the example.

This is the moment that adaptive rejection sampling shows up. Using tangents we "restrict" the distribution and we accept/reject when the samples are/are not samples belonging to the target distribution, which is (1). This has to happen dynamically because the distributions keep shifting, hence the "adaptive" part. This was proposed by Gilks, Wild in 1992.