STAT 460: Assignment 5

David Stenning

For this assignment, please complete the three exercises below. These exercises will require you to write Markov chain Monte Carlo algorithms. You may use the sample code from lecture slides or from BDA3 as a guide, but you should not simply take code from the internet or rely on R packages (or similar for other languages) for implementing the algorithms.

Please also note that **this homework assignment is worth 15 points**, not 10 points as for previous assignments, and will thus be weighted slightly more when computing your final grade.

Exercise 1 (5 points): In the Bayesian Computing (Part 3) lecture slides, we considered a target distribution (i.e., posterior distribution) that was a bivariate normal distribution:

$$p(\theta|y) = \mathrm{N}_2(0, \Sigma_{2\times 2}),$$

where $\Sigma_{2\times 2}$ is the variance-covariance matrix with variances equal to 1 and correlation equal to 0.8.

Construct a Metropolis algorithm to sample from the above target distribution. You may use a starting value of $(0, 0)$ and the jumping distribution from the lecture slides, or you may use another (but still symmetric) jumping distribution. (Of course you are encouraged to try different starting values and jumping distributions to learn about the algorithm, but you only need to turn in one set to be marked.) Obtain 5500 draws from $p(\theta|y)$ using your Metropolis algorithm, and discard the first 500 draws as burn-in. The end result is a sample of 5000 draws from $p(\theta|y)$. Use those draws to produce the figures requested below as necessary.

For this exercise, please turn in the following:

1. A description of your Metropolis algorithm in words, specific to this problem. That is, you should clearly state how proposed values for $\theta = (\theta_1, \theta_2)$ are drawn, what the acceptance ratio is, etc., not just restate the general form of the algorithm from the lecture notes. Or, to put it a third way, a knowledgeable reader should be able to translate your description into code without needing additional information.
2. Trace plots of the draws for θ_1 and θ_2. Also comment on the trace plots. Does the Markov chain appear to have converged? Is the chain mixing well, or is there (visual) evidence of high autocorrelation? (You do not need to make the autocorrelation function plot, just to provide a visual qualitative judgement.)
3. A two-dimensional scatterplot of the draws for θ_1 and θ_2. Comment on whether the draws appear to match the bivariate normal target density. (If they do not, you might consider trying to fix the error, as it is either a bug in your code, or a conceptual misunderstanding of the algorithm!)
4. Your code, which you may provide in an Appendix if you wish. Your code should be commented and readable, particularly if in a language other than R.

Hint: For this exercise, you may want to use the R library `mvtnorm` for obtaining draws from a bivariate normal distribution and evaluating the bivariate normal density.

Exercise 2 (5 points): Repeat Exercise 1, but replace the Metropolis algorithm with an Independence Sampler as follows:

1. Use the sample of θ from Exercise 1 to estimate the variance-covariance matrix of the posterior distribution, $\hat{\Sigma}$. (In R this is done easily using the `var` function.)

2. Approximate the posterior distribution as $\mathrm{N}_2\left(\mathrm{MAP}, \hat{\Sigma}\right)$, where MAP is the MAP estimate (i.e., estimate of the posterior mode). Then use this approximation as the jumping distribution in the Independence Sampler.

You should turn in all four parts that you did for Exercise 1 for this exercise as well, and answer all questions posed therein.

Exercise 3 (5 points): For this exercise we will reconsider the student heights example that we've used previously. Recall that we measure the heights (in inches) of three students and obtain $y_1 = 70$, $y_2 = 75$, and $y_3 = 72$. Previously we assumed that the heights follow a normal distribution with an unknown mean but known variance. For this exercise, we suppose (more realistically) that both the mean and variance are unknown model parameters for which we specify the following prior distributions:

$$\mu \sim \mathrm{N}(\mu_0 = 65, \tau_0^2 = 9)$$
$$\sigma^2 \sim \text{Inv-}\chi^2(\nu_0 = 175, \sigma_0^2 = 16).$$

Construct a Gibbs sampler and use it to obtain a sample from the resulting posterior distribution. Obtain 5000 draws from $p(\mu, \sigma^2|y)$ *after burn-in*. This means that you should run your chains for longer than 5000 iterations such that there are 5000 iterations after the Markov chain has appeared to converge; for this you will have to examine the trace plots. Use the sample of 5000 draws to produce the figures requested below as necessary.

For this exercise, please turn in the following:

1. A description of your Gibbs sampler in words, specific to this problem. For example, you should clearly state the the conditional distributions and how they are used to update μ and σ^2 at each iteration. As before, a knowledgeable reader should be able to translate your description into code without needing additional information.

2. Trace plots of the draws for μ and σ^2. Also comment on the trace plots. Is the chain mixing well, or is there (visual) evidence of high autocorrelation? (You do not need to make the autocorrelation function plot, just to provide a visual qualitative judgement. You also do not need to comment on whether the chain has converged because the 5000 draws are supposed to be *after* convergence.)

3. A two-dimensional scatterplot of the draws for μ and σ^2.

4. Your code, which you may provide in an Appendix if you wish. Your code should be commented and readable, particularly if in a language other than R.

Hints:

1. To obtain a draw θ from an Inv-$\chi^2(\nu, \sigma^2)$ distribution, first draw X from the χ^2_ν distribution and then let $\theta = \nu\sigma^2/X$ (see Appendix A of BDA3).

2. Recall the earlier results for a univariate normal distribution in which only one parameter, either μ or σ^2, was unknown.