STAT 489-01: Bayesian Methods of Data Analysis

Problem Set 3

Assigned 2017 February 7
Due 2017 February 14

Show your work on all problems! Be sure to give credit to any collaborators, or outside sources used in solving the problems. Note that if using an outside source to do a calculation, you should use it as a reference for the method, and actually carry out the calculation yourself; it's not sufficient to quote the results of a calculation contained in an outside source.

1 Marginalization and the Hessian Matrix

Suppose that the joint posterior on a pair of parameters θ_1, θ_2 is exactly Gaussian, and that the MAP values for the observed data $\mathbf{y}$ happen to be zero:

$$p(\boldsymbol{\theta}|\mathbf{y}, I) \propto \exp\left(-\frac{1}{2}\boldsymbol{\theta}^{\mathrm{T}}\mathbf{H}\boldsymbol{\theta}\right) = \exp\left(-\frac{1}{2}[H_{11}(\theta_1)^2 + 2H_{12}\theta_1\theta_2 + H_{22}(\theta_2)^2]\right) \quad -\infty < \theta_1, \theta_2 < \infty \tag{1.1}$$

(a) Complete the square to write $H_{11}(\theta_1)^2 + 2H_{12}\theta_1\theta_2 + H_{22}(\theta_2)^2$ in the form $[A\theta_2 - B]^2 + C$ (with the explicit forms of A, B and C specified).

(b) Marginalize over θ_2 to obtain the posterior $p(\theta_1|\mathbf{y}, I)$. Show that it is a Gaussian with mean zero and variance $\Sigma_{11} = H_{22}/(H_{11}H_{22} - [H_{12}]^2)$.

(c) Show that the matrix

$$\boldsymbol{\Sigma} = \frac{1}{H_{11}H_{22} - (H_{12})^2}\begin{pmatrix} H_{22} & -H_{12} \\ -H_{12} & H_{11} \end{pmatrix} \tag{1.2}$$

is the matrix inverse of $\mathbf{H}$, i.e., verify that the matrix product $\boldsymbol{\Sigma}\mathbf{H}$ is the identity matrix $\mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

2 McElreath Chapter 3, Exercises 3H1-3H5

3 Multi-Parameter Posterior

Consider a model where the data $\mathbf{y}$ are a sample of size n drawn from a Gamma distribution with unknown shape parameter θ_1 and rate parameter θ_2:

$$p(\mathbf{y}|\boldsymbol{\theta}, I) = \prod_{i=1}^{n} \frac{\theta_2^{\theta_1}}{\Gamma(\theta_1)} y_i^{\theta_1 - 1} e^{-\theta_2 y} \qquad 0 < y_i < \infty \tag{3.1}$$

Assume the improper prior $p(\boldsymbol{\theta}|I) \propto \frac{1}{\theta_2}$, $0 < \theta_1, \theta_2 < \infty$, and let the observed data be $\{y_i\} = \{0.38601, 0.58601, 0.81969, 0.09019, 0.30903\}$.

(a) Evaluate the log-posterior $\ln p(\boldsymbol{\theta}|\mathbf{y}, I)$, up to an additive constant, on a 100×100 grid of points in θ_1 and θ_2, and produce a contour plot. Choose a grid which extends far enough to include all the points where $\ln p(\boldsymbol{\theta}|\mathbf{y}, I) \gtrsim \max(\ln p(\boldsymbol{\theta}|\mathbf{y}, I)) - 3$. This may take a little experimentation, but a good starting point is to work out the method-of-moments estimates for θ_1 and θ_2, and go out about five times as far in each direction.

(b) Evaluate the posterior distribution $p(\boldsymbol{\theta}|\mathbf{y}, I)$, and produce a contour plot.

(c) Marginalize your gridded expressions to obtain posterior distributions $p(\theta_1|\mathbf{y}, I)$ and $p(\theta_2|\mathbf{y}, I)$, and plot each of them.

(d) Draw a sample of size $N = 12345$ from your grid-estimated posterior, with appropriate jitter, and produce a scatter plot of the two-dimensional sample.

(e) Use the `dens()` function from the `rethinking` package to produce estimates of $p(\theta_1|\mathbf{y}, I)$ and $p(\theta_2|\mathbf{y}, I)$ from your sample.

(f) Separately estimate each of the following from your grid-evaluated posterior and your sample: $E(\theta_1|\mathbf{y}, I)$, $E(\theta_2|\mathbf{y}, I)$, $V(\theta_1|\mathbf{y}, I)$, $V(\theta_2|\mathbf{y}, I)$, $\text{Cov}(\theta_1, \theta_2|\mathbf{y}, I)$.

(g) Comment on the behavior of your posterior at the boundaries of your grid, and the implications for defining a practical grid. What method could you use to improve upon this procedure?