Day 1: Introduction to Bayesian Inference

Exercise 1 (Gibbs Sampler for the Normal Distribution)

- Name the individual model parts:
  i) observation model: $x_i \overset{iid}{\sim} \mathrm{N}(\mu, \sigma^2)$
  ii) priors: $\mu \sim \mathrm{N}(m, s^2)$ and $\sigma^2 \sim \mathrm{IG}(a, b)$
  iii) parameters: μ and σ^2
  iv) hyperparameters: m, s^2, a, b.
- Sketch the generic approach for Gibbs sampling from the joint posterior distribution of μ and σ^2.
  i) Choose starting values $\mu^{[0]}$ and $(\sigma^2)^{[0]}$, set $t = 1$.
  ii) Draw $\mu^{[t]}$ from the full conditional $f(\mu|\boldsymbol{y}, (\sigma^2)^{[t-1]})$
  iii) Draw $(\sigma^2)^{[t]}$ from the full conditional $f(\sigma^2|\boldsymbol{y}, \mu^{[t]})$
  iv) Set $t = t + 1$
  v) Repeat steps ii, iii, and iv until the desired number of samples T is reached, i.e. until $t = T$.
- Derive the full conditional for μ.

  Joint posterior for μ and σ^2 up to a constant factor:

$$f(\mu, \sigma^2|\boldsymbol{y}) \propto f(\boldsymbol{y}|\mu, \sigma^2) f(\mu) f(\sigma^2)$$

  Full conditional for μ up to a constant factor:

$$f(\mu|\boldsymbol{x}, \sigma^2) \propto f(\boldsymbol{y}|\mu, \sigma^2) f(\mu)$$

  Inserting the explicit expressions:

$$f(\mu|\boldsymbol{y}, \sigma^2) \propto \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(y_i - \mu)^2\right) \frac{1}{\sqrt{2\pi s^2}} \exp\left(-\frac{1}{2s^2}(\mu - m)^2\right)$$

  Drop constant factors and simplify product $\prod_{i=1}^{n}$:

$$f(\mu|\boldsymbol{y}, \sigma^2) \propto \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \mu)^2\right) \exp\left(-\frac{1}{2s^2}(\mu - m)^2\right)$$

  Simplify further to

$$f(\mu|\boldsymbol{y}, \sigma^2) \propto \exp\left(-\frac{1}{2}\left[\frac{\sum_{i=1}^{n}(y_i - \mu)^2}{\sigma^2} + \frac{(\mu - m)^2}{s^2}\right]\right)$$

  Expand the squares and split up the fractions:

$$f(\mu|\boldsymbol{y}, \sigma^2) \propto \exp\left(-\frac{1}{2}\left[\frac{\sum_{i=1}^{n} y_i^2}{\sigma^2} + \frac{n\mu^2}{\sigma^2} - \frac{2\mu n\bar{y}}{\sigma^2} + \frac{\mu^2}{s^2} + \frac{m^2}{s^2} - \frac{2\mu m}{s^2}\right]\right)$$

Drop constant factors and write in terms of μ and μ^2:

$$f(\mu|\boldsymbol{y},\sigma^2) \propto \exp\left(-\frac{1}{2}\left[\mu^2\left(\frac{n}{\sigma^2}+\frac{1}{s^2}\right)-2\mu\left(\frac{n\bar{y}}{\sigma^2}+\frac{m}{s^2}\right)\right]\right)$$

Set

$$s^{2*} = \frac{1}{\frac{n}{\sigma^2}+\frac{1}{s^2}}$$

and

$$m^* = \left(\frac{n\bar{y}}{\sigma^2}+\frac{m}{s^2}\right)s^{2*} = \frac{\frac{n\bar{y}}{\sigma^2}+\frac{m}{s^2}}{\frac{n}{\sigma^2}+\frac{1}{s^2}}$$

to simplify to

$$f(\mu|\boldsymbol{y},\sigma^2) \propto \exp\left(-\frac{1}{2s^{2*}}\left[\mu^2-2\mu m^*\right]\right)$$

This is proportional to the probability density function of a normal distribution with mean m^* and variance s^{2*}, so we have found the full conditional for μ as

$$\mu|\cdot \sim \mathrm{N}(m^*, s^{2*})$$

- Derive the full conditional for σ.

  Joint posterior for μ and σ^2 up to a constant factor:

$$f(\mu,\sigma^2|\boldsymbol{y}) \propto f(\boldsymbol{y}|\mu,\sigma^2)f(\mu)f(\sigma^2)$$

  Full conditional for σ^2 up to a constant factor:

$$f(\sigma^2|\boldsymbol{y},\mu) \propto f(\boldsymbol{y}|\mu,\sigma^2)f(\sigma^2)$$

  Inserting the explicit expressions:

$$f(\sigma^2|\boldsymbol{y},\mu) \propto \prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2\sigma^2}(y_i-\mu)^2\right)\frac{b}{\Gamma(a)}\left(\frac{1}{\sigma^2}\right)^{a+1}\exp\left(-\frac{b}{\sigma^2}\right)$$

  Drop constant factors and simplify product $\prod_{i=1}^{n}$:

$$f(\sigma^2|\boldsymbol{y},\mu) \propto \frac{1}{\sigma^n}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu)^2\right)\left(\frac{1}{\sigma^2}\right)^{a+1}\exp\left(-\frac{b}{\sigma^2}\right).$$

  Simplify fractions and exponentials:

$$f(\sigma^2|\boldsymbol{y},\mu) \propto \sigma^{-2(a+1)-n}\exp\left(-\frac{1}{\sigma^2}\left[\frac{1}{2}\sum_{i=1}^{n}(y_i-\mu)^2+b\right]\right).$$

  Rewrite $\sigma^{-(\cdot)}$ expression:

$$f(\sigma^2|\boldsymbol{y},\mu) \propto \left(\frac{1}{\sigma^2}\right)^{a+\frac{1}{2}n+1}\exp\left(-\frac{1}{\sigma^2}\left[\frac{1}{2}\sum_{i=1}^{n}(y_i-\mu)^2+b\right]\right).$$

  Set $a^* = a+\frac{1}{2}n$ and $b^* = \frac{1}{2}\sum_{i=1}^{n}(y_i-\mu)^2+b$ to further simplify to

$$f(\sigma^2|\boldsymbol{y},\mu) \propto \left(\frac{1}{\sigma^2}\right)^{a^*+1}\exp\left(-\frac{b^*}{\sigma^2}\right)$$

  This expression is proportional to the probability density function of the inverse gamma distribution with parameters a^* and b^*. Thus, we can conclude that the full conditional for σ^2 is

$$\sigma^2|\cdot \sim \mathrm{IG}(a^*, b^*).$$

Exercise 2 (Metropolis-Hastings with Random Walk Proposals)

- Derive the natural logarithm of the acceptance probability for a proposal $\boldsymbol{\beta}^*$ in a Metropolis-Hastings step.

  The acceptance probability is given by

$$\alpha(\boldsymbol{\beta}^*|\boldsymbol{\beta}^{[t-1]}) = \min\left\{1, \underbrace{\frac{p(\boldsymbol{\beta}^*|\boldsymbol{y})}{p(\boldsymbol{\beta}^{[t-1]}|\boldsymbol{y})}\frac{h(\boldsymbol{\beta}^{[t-1]}|\boldsymbol{\beta}^*)}{h(\boldsymbol{\beta}^*|\boldsymbol{\beta}^{[t-1]})}}_{=\alpha^*}\right\}$$

  Taking the log of the big fraction:

$$\log(\alpha^*) = \log\left(\frac{p(\boldsymbol{\beta}^*|\boldsymbol{y})}{p(\boldsymbol{\beta}^{[t-1]}|\boldsymbol{y})}\frac{h(\boldsymbol{\beta}^{[t-1]}|\boldsymbol{\beta}^*)}{h(\boldsymbol{\beta}^*|\boldsymbol{\beta}^{[t-1]})}\right) = \log\left(\frac{p(\boldsymbol{\beta}^*|\boldsymbol{y})}{p(\boldsymbol{\beta}^{[t-1]}|\boldsymbol{y})}\right) + \log\left(\frac{h(\boldsymbol{\beta}^{[t-1]}|\boldsymbol{\beta}^*)}{h(\boldsymbol{\beta}^*|\boldsymbol{\beta}^{[t-1]})}\right)$$

  Rearranging a little:

$$\log(\alpha^*) = \log p(\boldsymbol{\beta}^*|\boldsymbol{y}) - \log p(\boldsymbol{\beta}^{[t-1]}|\boldsymbol{y}) + \log h(\boldsymbol{\beta}^{[t-1]}|\boldsymbol{\beta}^*) - \log h(\boldsymbol{\beta}^*|\boldsymbol{\beta}^{[t-1]})$$

  The log-acceptance probability is thus

$$\log\alpha(\boldsymbol{\beta}^*|\boldsymbol{\beta}^{[t-1]}) = \min\{0, \log(\alpha^*)\}$$

  with $\log(\alpha^*)$ as defined above.

Exercise 3 (Hamiltonian Monte Carlo)

- Sketch a Hamiltonian Monte Carlo algorithm for drawing samples from the posterior of $\boldsymbol{\beta}$.
  - Set tuning parameters: Mass matrix $\boldsymbol{M}$, step size ϵ, number of steps L.
  - Choose starting values $\boldsymbol{\beta}^{[0]}$, set $t = 1$.
  - Draw momentum vector $\boldsymbol{p}^{[t]} = (p^{[t]}_{\beta_0}, p^{[t]}_{\beta_1})'$ from $\boldsymbol{p} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{M})$
  - Find proposal $\boldsymbol{\beta}^*$ and momentum $\boldsymbol{p}^*$ by conducting L leapfrog steps.
  - Compute acceptance probability and conduct acceptance step for proposal
  - If accepted, set $\boldsymbol{\beta}^{[t]} \leftarrow \boldsymbol{\beta}^*$
  - Repeat steps iii - vi until the target number of samples t is reached, i.e. until $t = T$.
- Find the derivative of the log posterior with respect to $\boldsymbol{\beta}$

  Note that β_0 has a constant prior. From the posterior

$$f(\boldsymbol{\beta}|\boldsymbol{y}, \sigma^2) \propto f(\boldsymbol{y}|\boldsymbol{\beta}, \sigma^2) f(\beta_1)$$

  we get to the log posterior

$$\log f(\boldsymbol{\beta}|\boldsymbol{y}, \sigma^2) \propto \log f(\boldsymbol{y}|\boldsymbol{\beta}, \sigma^2) + \log f(\beta_1).$$

  Inserting the explicit expressions yields

$$\log f(\boldsymbol{\beta}|\boldsymbol{y}, \sigma^2) \propto -\frac{1}{2\sigma^2}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) - \frac{1}{2s^2}\beta_1^2.$$

  Expand the square:

$$\log f(\boldsymbol{\beta}|\boldsymbol{y}, \sigma^2) \propto -\frac{1}{2\sigma^2}(\boldsymbol{y}'\boldsymbol{y} - 2\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{y} + \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta}) - \frac{1}{2s^2}\beta_1^2.$$

  Take the derivative:

$$\frac{\partial}{\partial\boldsymbol{\beta}}\log f(\boldsymbol{\beta}|\boldsymbol{y}, \sigma^2) = \frac{1}{\sigma^2}\boldsymbol{X}'\boldsymbol{y} - \frac{1}{\sigma^2}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} - \frac{1}{s^2}\beta_1.$$

Exercise 4 (IWLS)

- Derive the quantities $\boldsymbol{F}_s^*$ and $\boldsymbol{s}_s$ for $\boldsymbol{\beta}$ using the model from the previous exercises.

  We have already derived $\boldsymbol{s}_s$ in the previous exercise:

  $$\boldsymbol{s}_{\boldsymbol{\beta}} = \frac{\partial}{\partial \boldsymbol{\beta}} \log f(\boldsymbol{\beta}|\boldsymbol{y}, \sigma^2) = \frac{1}{\sigma^2}\boldsymbol{X}'\boldsymbol{y} - \frac{1}{\sigma^2}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} - \frac{1}{s^2}\beta_1$$

  To get $\boldsymbol{F}_{\boldsymbol{\beta}}^*$, we start by differentiating a second time:

  $$\frac{\partial^2}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'} \log f(\boldsymbol{\beta}|\boldsymbol{y}, \sigma^2) = \frac{\partial}{\partial \boldsymbol{\beta}'}\boldsymbol{s}_{\boldsymbol{\beta}} = -\frac{1}{\sigma^2}\boldsymbol{X}'\boldsymbol{X} - \frac{1}{s^2}$$

  Multiplying with -1 and taking the expectation, we get

  $$\boldsymbol{F}_{\boldsymbol{\beta}}^* = \mathbb{E}\left(\frac{1}{\sigma^2}\boldsymbol{X}'\boldsymbol{X} - \frac{1}{s^2}\right) = \frac{1}{\sigma^2}\boldsymbol{X}'\boldsymbol{X} + \frac{1}{s^2}$$