- Problem B: Bayesian Image Reconstruction

We study how the ising model can be used as a prior distribution in an image reconstruction setting. We let $y = (y_{ij}, i = 1, \dots, 89, j = 1, \dots, 85)$ be the observed "image.txt". We assume this to be a noisy version of an unobserved image $x = (x_{ij}, i = 1, \dots, 89, j = 1, \dots, 85)$ with $x_{ij} \in \{0, 1\}$. Our goal in this problem is to use the observed y to estimate x. We assume the elements in y to be conditionally independent given x and

$$y_{ij}|x \sim N(\mu_{x_{ij}}, \sigma^2_{x_{ij}}) \tag{1}$$

where μ_0, μ_1 are the mean values for y_{ij} when x_{ij} is zero and one, respectively and σ_0^2 and σ_1^2 are corresponding variances. Apriori we assume x to be independent of $\varphi = (\mu_0, \mu_1, \sigma_0, \sigma_1)$. As prior for x we assume an Ising model with interaction parameter β, i.e.

$$f(x) \propto \exp\Big\{\beta \sum_{(i,j)\sim(k,l)} I(x_{ij} = x_{kl})\Big\} \tag{2}$$

where the sum is over all pairs of neighbour nodes in the 89×85 lattice and the value of β is assumed to be known. To define a prior for φ we follow a procedure used in Austad and Tjelmeland (2017). We first define a reparametrisation to new parameters (m_0, θ, s, τ) by the relations

$$\sigma_0 = s \cdot \tau, \ \sigma_1 = \frac{s}{\tau}, \ \mu_0 = m_0, \text{ and } \mu_1 = m_0 + s\theta$$

The s defines a scale, θ defines the difference between the two mean values in this scale, and τ defines σ_0 and σ_1 using the same scale. We then define a prior for $\varphi = (\mu_0, \mu_1, \sigma_0, \sigma_1)$ implicitly by assigning a prior for (m_0, θ, s, τ)

$$f(\tau) = \begin{cases} \frac{1}{2\tau^2} e^{-(\frac{1}{\tau}-1)} & \text{for } \tau \in (0, 1], \\ \frac{1}{2} e^{-(\tau-1)} & \text{for } \tau > 1. \end{cases}$$

We use the transformation formula to find the corresponding prior for $t = 1/\tau$. We let $t = 1/\tau$, such that $\tau = 1/t = w(t)$. The pdf $g(t)$ for t will then be given by

$$g(t) = f(w(t)) \cdot |w'(t)|$$

which gives

$$g(t) = \begin{cases} \frac{1}{2} t^2 e^{-(t-1)} \cdot \frac{1}{t^2} & \text{for } t \geq 1 \\ \frac{1}{2} e^{-(t-1)} \cdot \frac{1}{t^2} & \text{for } t \in (0, 1) \end{cases} = \begin{cases} \frac{1}{2} e^{-(t-1)} & \text{for } t \geq 1 \\ \frac{1}{2t^2} e^{-(\frac{1}{t}-1)} & \text{for } t \in (0, 1) \end{cases}$$

Notice how the intervals changes and that the expressions for $f(\tau)$ and $g(t)$ are the same. Hence, the priors for τ and t are identical and we can use this to argue that the marginal priors for σ_0 and σ_1 are identical. This is easily shown since $\sigma_0 = s \cdot \tau$ and $\sigma_1 = \frac{s}{\tau}$.

2. We show that the resulting (improper) prior for $\varphi = (\mu_0, \mu_1, \sigma_0, \sigma_1)$ becomes (up to proportionality)

$$f(\varphi) \propto \begin{cases} \frac{(\mu_1-\mu_0)^3}{\sigma_0^3 \sigma_1^2} \exp\{-\Big[\frac{\mu_1-\mu_0}{\sqrt{\sigma_0\sigma_1}} + \sqrt{\frac{\sigma_1}{\sigma_0}}\Big]\} & \text{for } \sigma_0 \leq \sigma_1 \text{ and } \mu_0 < \mu_1 \\ \frac{(\mu_1-\mu_0)^3}{\sigma_0^2 \sigma_1^3} \exp\{-\Big[\frac{\mu_1-\mu_0}{\sqrt{\sigma_0\sigma_1}} + \sqrt{\frac{\sigma_0}{\sigma_1}}\Big]\} & \text{for } \sigma_0 > \sigma_1 \text{ and } \mu_0 < \mu_1 \end{cases}$$

We must use the transformation formula with 4 variables. We thus have

$$f(\varphi) = f(m_0, s, \tau, \theta) = f(m_0) f(s) f(\tau) f(\theta) \cdot |J|$$

because of independence, where J is the 4×4 Jacobi determinant shown in (3). Since m_0 and s are assumed to be improper uniform distributed, we have that

$$f(\varphi) \propto f(\tau)f(\theta)\cdot|J|$$

The Jacobi determinant is given by

$$J = \begin{vmatrix} \frac{\partial m_0}{\partial \mu_0} & \frac{\partial m_0}{\partial \mu_1} & \frac{\partial m_0}{\partial \sigma_0} & \frac{\partial m_0}{\partial \sigma_1} \\ \frac{\partial s}{\partial \mu_0} & \frac{\partial s}{\partial \mu_1} & \frac{\partial s}{\partial \sigma_0} & \frac{\partial s}{\partial \sigma_1} \\ \frac{\partial \tau}{\partial \mu_0} & \frac{\partial \tau}{\partial \mu_1} & \frac{\partial \tau}{\partial \sigma_0} & \frac{\partial \tau}{\partial \sigma_1} \\ \frac{\partial \theta}{\partial \mu_0} & \frac{\partial \theta}{\partial \mu_1} & \frac{\partial \theta}{\partial \sigma_0} & \frac{\partial \theta}{\partial \sigma_1} \end{vmatrix} \tag{3}$$

which becomes

$$J = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2}\sqrt{\frac{\sigma_1}{\sigma_0}} & \frac{1}{2}\sqrt{\frac{\sigma_0}{\sigma_1}} \\ 0 & 0 & \frac{1}{2\sqrt{\sigma_0\sigma_1}} & -\frac{1}{2}\sqrt{\frac{\sigma_0}{\sigma_1^3}} \\ -\frac{1}{\sqrt{\sigma_0\sigma_1}} & \frac{1}{\sqrt{\sigma_0\sigma_1}} & -\frac{(\mu_1-\mu_0)}{2\sqrt{\sigma_0^3\sigma_1}} & -\frac{(\mu_1-\mu_0)}{2\sqrt{\sigma_0\sigma_1^3}} \end{vmatrix}$$

This can now be reduced to

$$J = 1\cdot \begin{vmatrix} 0 & \frac{1}{2}\sqrt{\frac{\sigma_1}{\sigma_0}} & \frac{1}{2}\sqrt{\frac{\sigma_0}{\sigma_1}} \\ 0 & \frac{1}{2\sqrt{\sigma_0\sigma_1}} & -\frac{1}{2}\sqrt{\frac{\sigma_0}{\sigma_1^3}} \\ \frac{1}{\sqrt{\sigma_0\sigma_1}} & -\frac{(\mu_1-\mu_0)}{2\sqrt{\sigma_0^3\sigma_1}} & -\frac{(\mu_1-\mu_0)}{2\sqrt{\sigma_0\sigma_1^3}} \end{vmatrix} = 1\cdot\frac{1}{\sqrt{\sigma_0\sigma_1}}\cdot \begin{vmatrix} \frac{1}{2}\sqrt{\frac{\sigma_1}{\sigma_0}} & \frac{1}{2}\sqrt{\frac{\sigma_0}{\sigma_1}} \\ \frac{1}{2\sqrt{\sigma_0\sigma_1}} & -\frac{1}{2}\sqrt{\frac{\sigma_0}{\sigma_1^3}} \end{vmatrix} = -\frac{1}{2\sqrt{\sigma_0\sigma_1^3}}$$

Hence, we get

$$f(\varphi) \propto f(\tau)f(\theta)\cdot|J| \propto \begin{cases} \frac{1}{2}\frac{\sigma_1}{\sigma_0}e^{-\left(\sqrt{\frac{\sigma_1}{\sigma_0}}-1\right)}\cdot\left(\frac{\mu_1-\mu_0}{\sqrt{\sigma_0\sigma_1}}\right)^3 e^{-\frac{\mu_1-\mu_0}{\sqrt{\sigma_0\sigma_1}}}\cdot\frac{1}{2\sqrt{\sigma_0\sigma_1^3}} & \text{for } \sigma_0 \le \sigma_1 \text{ and } \mu_0 < \mu_1 \\ \frac{1}{2}e^{-\left(\sqrt{\frac{\sigma_0}{\sigma_1}}-1\right)}\cdot\left(\frac{\mu_1-\mu_0}{\sqrt{\sigma_0\sigma_1}}\right)^3 e^{-\frac{\mu_1-\mu_0}{\sqrt{\sigma_0\sigma_1}}}\cdot\frac{1}{2\sqrt{\sigma_0\sigma_1^3}} & \text{for } \sigma_0 > \sigma_1 \text{ and } \mu_0 < \mu_1 \end{cases}$$

$$f(\varphi) \propto \begin{cases} \frac{(\mu_1-\mu_0)^3}{\sigma_0^3\sigma_1^2}\exp\{-\left[\frac{\mu_1-\mu_0}{\sqrt{\sigma_0\sigma_1}}+\sqrt{\frac{\sigma_1}{\sigma_0}}\right]\} & \text{for } \sigma_0 \le \sigma_1 \text{ and } \mu_0 < \mu_1 \\ \frac{(\mu_1-\mu_0)^3}{\sigma_0^2\sigma_1^3}\exp\{-\left[\frac{\mu_1-\mu_0}{\sqrt{\sigma_0\sigma_1}}+\sqrt{\frac{\sigma_0}{\sigma_1}}\right]\} & \text{for } \sigma_0 > \sigma_1 \text{ and } \mu_0 < \mu_1 \end{cases} \tag{4}$$

which is what we wanted to show.

3. We find (up to proportionality) a formula for the posterior distribution $f(x,\varphi|y)$. We thus have

$$f(x,\varphi|y) \propto f(y|x,\varphi)\cdot f(x,\varphi) \propto f(y|x)\cdot f(x)\cdot f(\varphi)$$

All of these distributions are known to us from equations (1), (2) and (4).

$$f(x,\varphi|y) \propto \frac{1}{\sqrt{2\pi}\sigma_{x_{ij}}}e^{-\frac{1}{2\sigma_{x_{ij}}^2}\left(y_{ij}-\mu_{x_{ij}}\right)^2}\cdot\exp\left\{\beta\sum_{(i,j)\sim(k,l)} I(x_{ij}=x_{kl})\right\}$$

$$\cdot\begin{cases} \frac{(\mu_1-\mu_0)^3}{\sigma_0^3\sigma_1^2}\exp\{-\left[\frac{\mu_1-\mu_0}{\sqrt{\sigma_0\sigma_1}}+\sqrt{\frac{\sigma_1}{\sigma_0}}\right]\} & \text{for } \sigma_0 \le \sigma_1 \text{ and } \mu_0 < \mu_1 \\ \frac{(\mu_1-\mu_0)^3}{\sigma_0^2\sigma_1^3}\exp\{-\left[\frac{\mu_1-\mu_0}{\sqrt{\sigma_0\sigma_1}}+\sqrt{\frac{\sigma_0}{\sigma_1}}\right]\} & \text{for } \sigma_0 > \sigma_1 \text{ and } \mu_0 < \mu_1 \end{cases}$$

4. We define and implement a Metropolis-Hastings algorithm for simulating from $f(x, \varphi|y)$

```
y = read.table("./image.txt", header = FALSE, sep = " ")
nrows = dim(y)[1]
ncolumns = dim(y)[2]
x = matrix(rbinom(nrows * ncolumns, 1, 0.5), ncol = ncolumns, nrow = nrows)
i = ceiling(nrows*runif(1))
j = ceiling(ncolumns*runif(1))

i = 1
j = 1
right = x[i,j] == x[i+1,j]
left = x[i,j] == x[i-1,j]
up = x[i,j] == x[i,j+1]
down = x[i,j] == x[i,j-1]
indicator = right + left + up + down

beta = 1
x_prop = 1 - x[i,j]
```