Posterior distribution for correlation coeficient. Bivariate case

October 24, 2014

Consider n observations from $Y \sim N_d(0, \Sigma)$ distribution, the likelihood function can be written as follows:

$$p(y|\mu,\Sigma) \propto |\Sigma|^{-n/2} e^{-\frac{1}{2}\sum_{i=1}^{n}(y_i-\mu)^{'}\Sigma^{-1}(y_i-\mu)} = |\Sigma|^{-n/2} e^{-\frac{1}{2}tr(\Sigma^{-1}S_\mu)} \tag{1}$$

where y_i represents the ith observation from the vector Y. Also $S_\mu = \sum_{i=1}^{n}(y_i-\mu)(y_i-\mu)^{'}$, which can be decompose as $S_\mu = \sum_{i=1}^{n}(y_i-\bar{y})(y_i-\bar{y})^{'} + n(bary-\mu)(\bar{y}-\mu)^{'} = A + M$.

1 Derivaitons

To complete model (1) we set uniform prior for μ, $p(\mu) \propto 1$ and we study alternatives for construct a prior on Σ. This implies,

$$\begin{array}{ll} p(\Sigma|y) & \propto \int p(\Sigma,\mu|y)d\mu \\ & \propto \int |\Sigma|^{-n/2} e^{-\frac{1}{2}tr(\Sigma^{-1}(A+M))} p(\Sigma)d\mu \\ & \propto |\Sigma|^{-n/2} p(\Sigma) e^{-\frac{1}{2}tr(\Sigma^{-1}A)} \int e^{-\frac{1}{2}tr(\Sigma^{-1}M)}d\mu \\ & \propto |\Sigma|^{(-n/2+1/2)} p(\Sigma) e^{-\frac{1}{2}tr(\Sigma^{-1}A)} \end{array} \tag{2}$$

For each of these alternatives we derive the posterior distribution of the correlation coefficient ρ.

1.1 Jeffrey prior

A non informative prior for the covariance matrix is $p(\Sigma) \propto |Sigma|^{-(\frac{d+1}{2})}$, then $p(\Sigma|y) \propto |\Sigma|^{-(\frac{d+n}{2})} e^{-\frac{1}{2}tr(\Sigma^{-1}A)}$ implying that $\Sigma|y \sim IW(n-1, A)$. In the bivariate case ($d = 2$) this implies that

$$p(\sigma_1^2, \sigma_2^2, \sigma_{12}) \propto (\sigma_1^2\sigma_2^2 - \sigma_{12}^2)^{-\frac{n+2}{2}} exp\left[-\frac{1}{2(\sigma_1^2\sigma_2^2 - \sigma_{12}^2)}(a_{11}\sigma_2^2 - 2a_{12}\sigma_{12} + a_{22}\sigma_1^2\right]$$

where $a_{kj} = sum_{i=1}^{n}(y_{ik} - \bar{y}_k)(y_{ij} - \bar{y}_j)$. Then Box-Tiao finds the posterior distribution for ρ using a transformation proposed by Fisher to obtain the sampling distribution of the pearson correlation coeficient, $r = a_{12}/\sqrt{a_{11}a22}$

$$x = (\tfrac{\sigma_1\sigma_2}{a_{11}a_{22}})^{1/2} \quad w = (\tfrac{\sigma_1 a_{22}}{a_{11}\sigma_2})^{1/2} \quad \rho = \tfrac{\sigma_{12}}{\sigma_1\sigma_2} \quad J = \tfrac{2x^2(a_{11}a_{22})^{3/2}}{w}$$

$$
\begin{array}{ll}
p(x,w,\rho) & \propto (x^2(1-\rho^2))^{-\frac{n+2}{2}} \frac{x^2}{w} exp\left[-\frac{1}{2x(1-\rho^2)}(w^{-1}-2r\rho w)\right] \\
p(w,\rho) & \propto \frac{(1-\rho^2)^{-\frac{n+2}{2}}}{w} \int x^{-n} exp\left[-\frac{1}{2x(1-\rho^2)}(w^{-1}-2r\rho w)\right] dx \\
p(w,\rho) & \propto \frac{(1-\rho^2)^{-\frac{n+2}{2}}}{w}(1-\rho^2)^{n-1}(w^{-1}-2r\rho w)^{-(n-1)}
\end{array}
$$

finally

$$
p(\rho|y) \propto (1-\rho^2)^{n/2-2} \int_0^{\infty} w^{-1}(w^{-1}-2r\rho+w)^{-(n-1)} dw \tag{3}
$$

1.2 Conjugate prior

A similar way can be used for the case when $\Sigma \sim IW(\nu, \Lambda)$, in this case we can include $p(\Sigma) \propto |\Sigma|^{-(\frac{\nu+d+1}{2})} e^{-\frac{1}{2}tr(\Lambda\Sigma^{-1})}$ on equation (2) to get $p(\Sigma|y) \propto |\Sigma|^{-(\frac{\nu+d+n}{2})} e^{-\frac{1}{2}tr(\Sigma^{-1}(A+\Lambda)}$ implying that $\Sigma|y \sim IW(n+\nu-1, A+\Lambda)$ as expected.

Again for bivariate case, posterior distribution of the correlation coefficient can be obtain letting $A^{'} = A+\Lambda$ and applying the same transformation as before to obtain

$$
p(\rho|y) \propto (1-\rho^2)^{\frac{n+\nu}{2}-2} \int_0^{\infty} w^{-1}(w^{-1}-2r^{'}\rho+w)^{-(n+\nu-1)} dw \tag{4}
$$

where the only differences with (3) are the effect of ν and we use $r^{'} = \frac{a_{12}+\lambda_{12}}{\sqrt{(a_{11}+\lambda_{11})(a_{22}+\lambda_{22})}}$ instead of the sample correlation.

1.3 Separation strategy

It is known the implied marginal prior distribution from $IW(\nu = d+1, \Lambda = I)$ are $\sigma_i^2 \sim IG(1, 1/2)$ and $\rho \sim Unif(-1,1)$. A posible separation strategy that mach the marginal prior of the conjugate case is then $p(\Sigma) \propto (\sigma_1^2\sigma_2^2)^{-2} exp\left[-\frac{1}{2\sigma_1^2} - \frac{1}{2\sigma_2^2}\right]$ which gives a posterior for Σ as follows

$$
\begin{array}{ll}
p(\Sigma|y) & \propto |\Sigma|^{-\frac{n+1}{2}}(\sigma_1^2\sigma_2^2)^{-2} exp\left[-\frac{1}{2}\left(tr(\Sigma^{-1}A)+\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}\right)\right] \\
& \propto (\sigma_1^2\sigma_2^2(1-\rho^2))^{-\frac{n+1}{2}}(\sigma_1^2\sigma_2^2)^{-2} exp\left[-\frac{1}{2}\left(\frac{1}{(1-\rho^2)}\left(\frac{a_{11}}{\sigma_2^2} - \frac{2\rho a_{12}}{\sigma_1\sigma_2} + \frac{a_{22}}{\sigma_1^2}\right) + \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)\right] \\
& \propto (\sigma_1^2\sigma_2^2(1-\rho^2))^{-\frac{n+1}{2}}(\sigma_1^2\sigma_2^2)^{-2} exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{a_{11}+1-\rho^2}{\sigma_2^2} - \frac{2\rho a_{12}}{\sigma_1\sigma_2} + \frac{a_{22+1-\rho^2}}{\sigma_1^2}\right)\right]
\end{array}
$$

applying a similar transformation we can obtain the posterior distribution for ρ, this is

$$
x = (\frac{\sigma_1\sigma_2}{a_{11}a_{22}})^{1/2} \quad w = (\frac{\sigma_1 a_{22}}{a_{11}\sigma_2})^{1/2} \quad \rho = \rho \quad J = \frac{2xa_{11}a_{22}}{w}
$$

$$
\begin{array}{ll}
p(x,w,\rho) & \propto (x^2(1-\rho^2))^{-\frac{n+1}{2}} \frac{x^{-3}}{w} exp\left[-\frac{1}{2x(1-\rho^2)}\left(\frac{a_{11}+1-\rho^2}{wa_{11}} - 2\rho r + \frac{a_{22+1-\rho^2}}{a_{22}}w\right)\right] \\
p(w,\rho) & \propto (1-\rho^2)^{\frac{n+3}{2}} \frac{1}{w}\left(\frac{a_{11}+1-\rho^2}{wa_{11}} - 2r\rho + \frac{a_{22+1-\rho^2}}{a_{22}}w\right)^{-(n+2)}
\end{array}
$$

from where we get

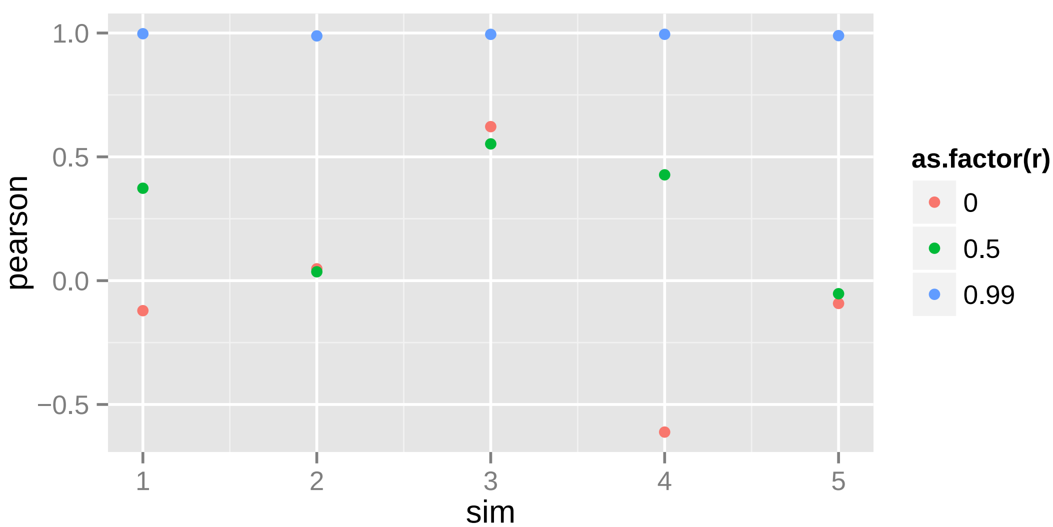
$$p(\rho|y) \propto (1-\rho^2)^{\frac{n+3}{2}} \int \frac{1}{w} \left(\frac{a_{11}+1-\rho^2}{wa_{11}} - 2r\rho + \frac{a_{22+1-\rho^2}}{a_{22}} w \right)^{-(n+2)} \tag{5}$$

2 simulations

Here we simulate 5 bivariate normally distributed data set, from the model

$$\binom{Y_1}{Y_2} \sim N\left(\binom{0}{0}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \right)$$

with specific values $\sigma_1 = \sigma_2 = 0.01$, and $\rho = 0$, sample size is $n = 10$.



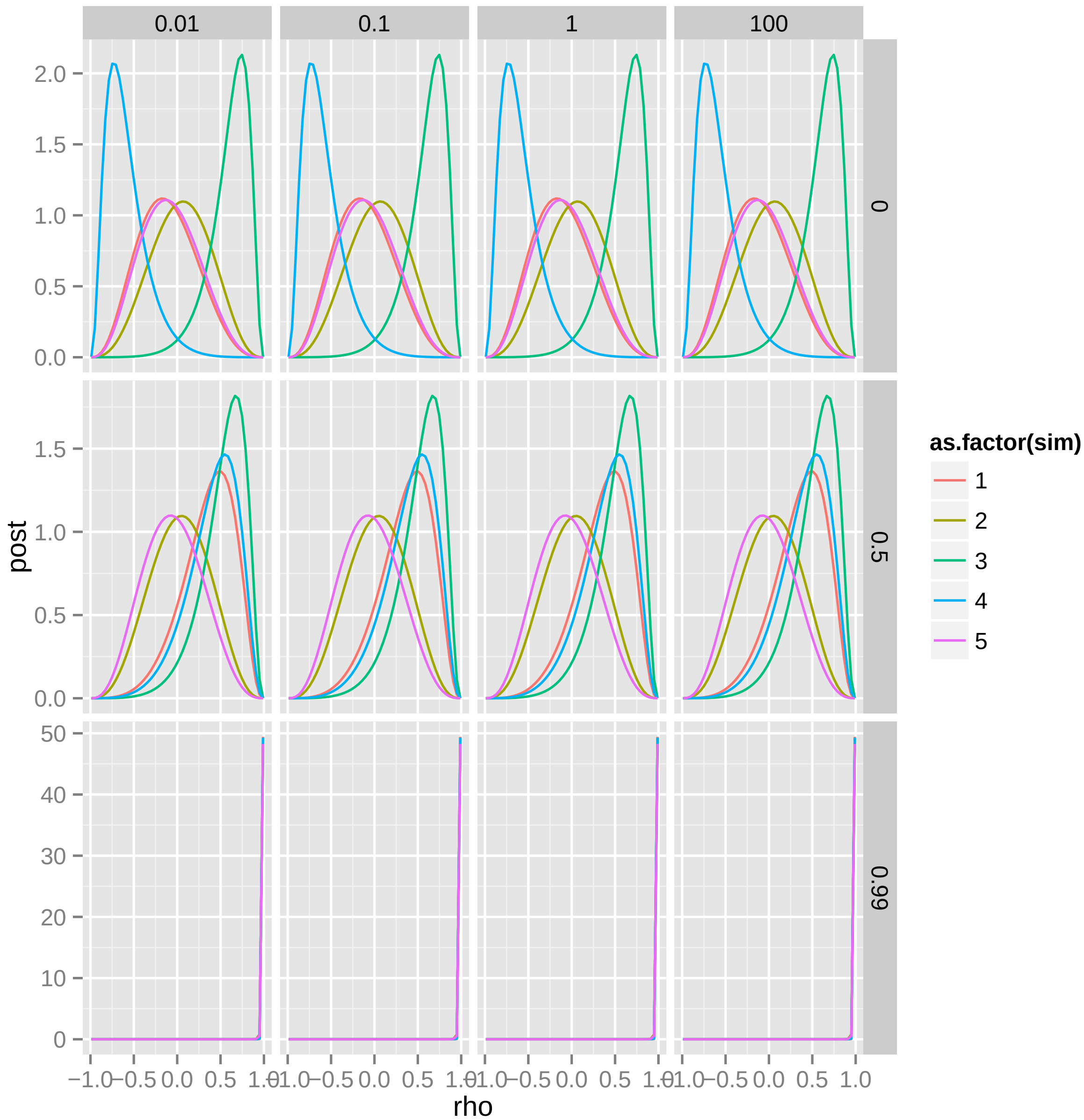


Figure 1: ρ posterior density for Jeffrey prior

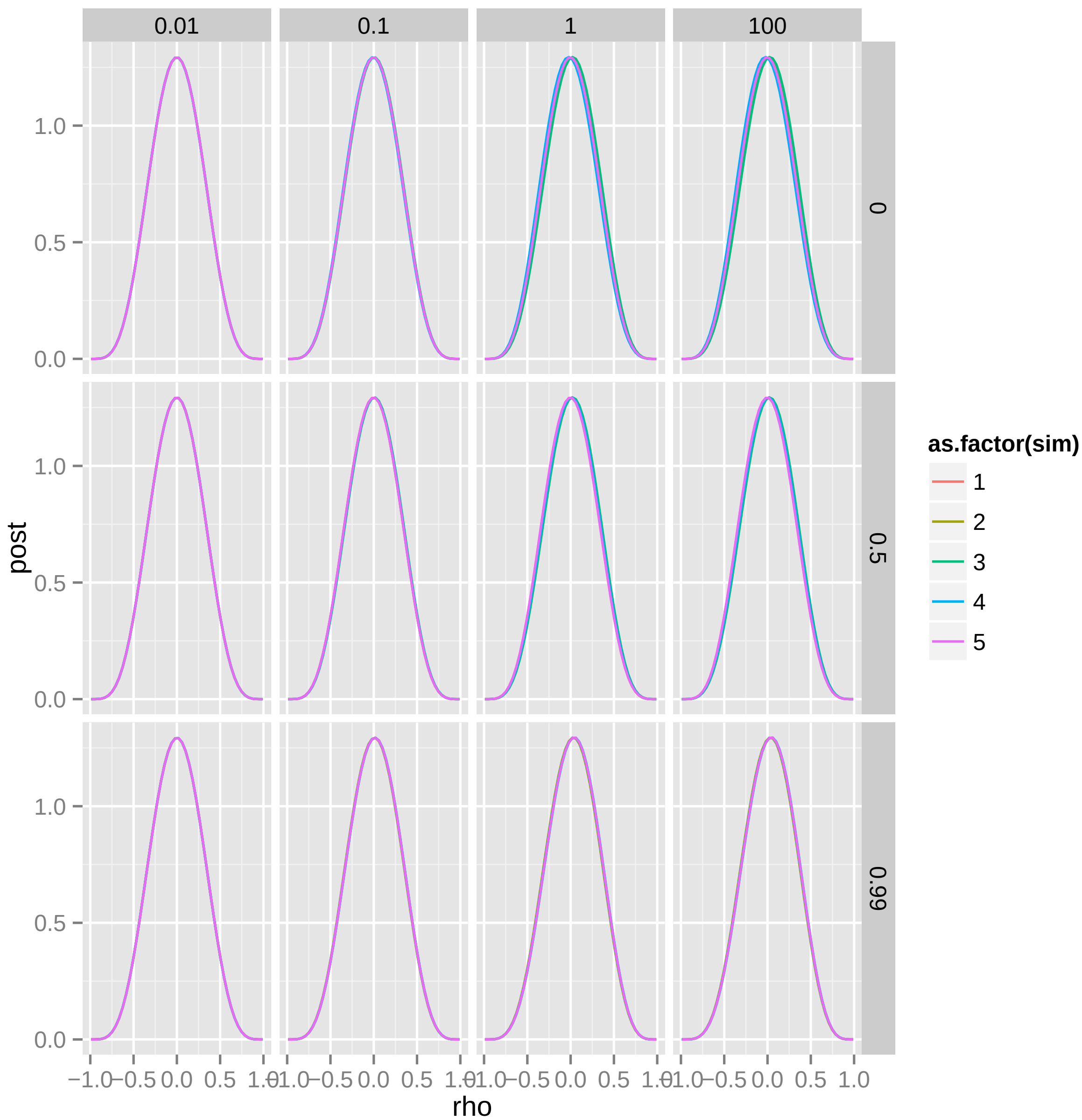


Figure 2: ρ posterior density for IW prior

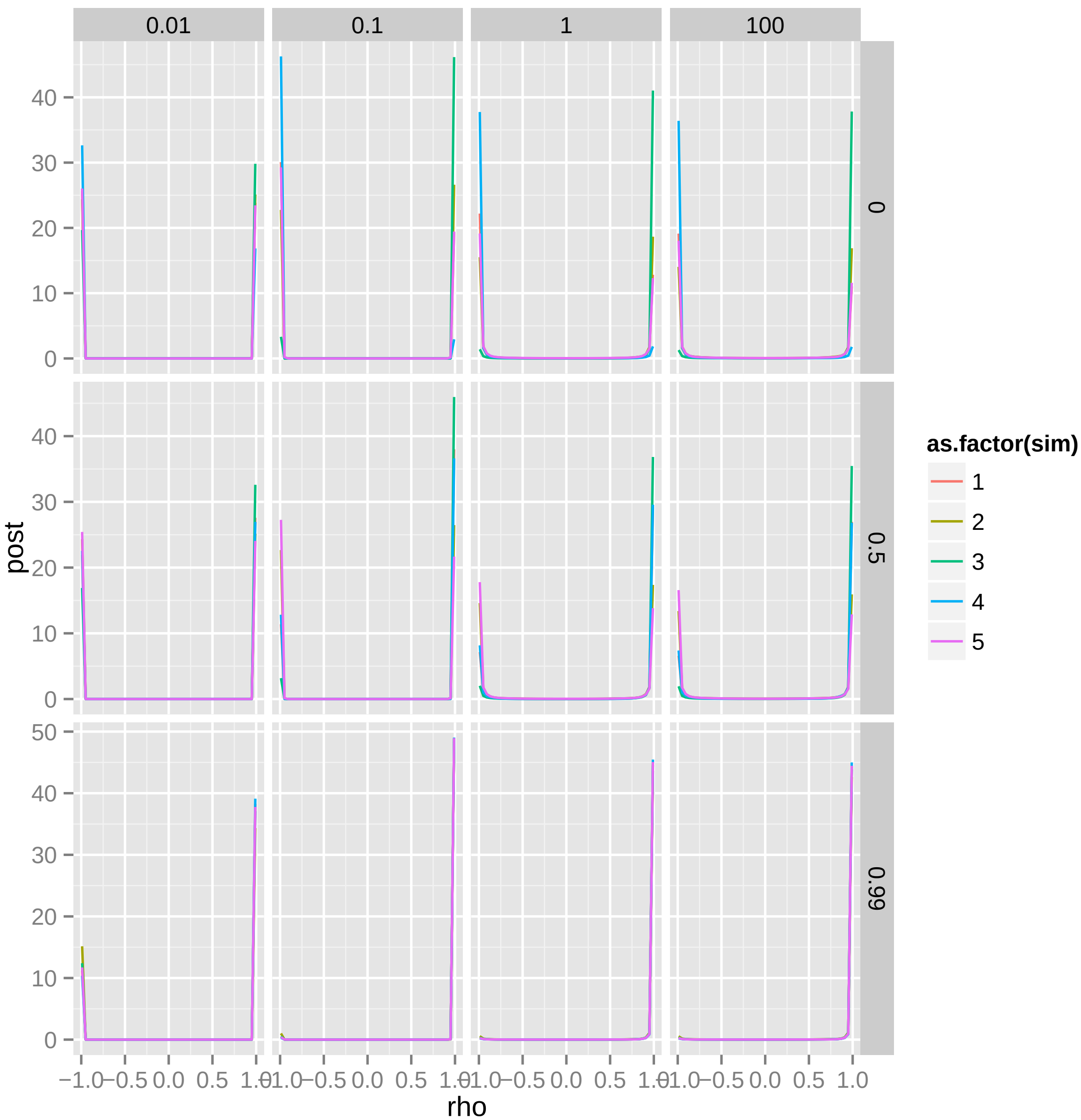


Figure 3: ρ posterior density for $SS[IG, IW]$ prior