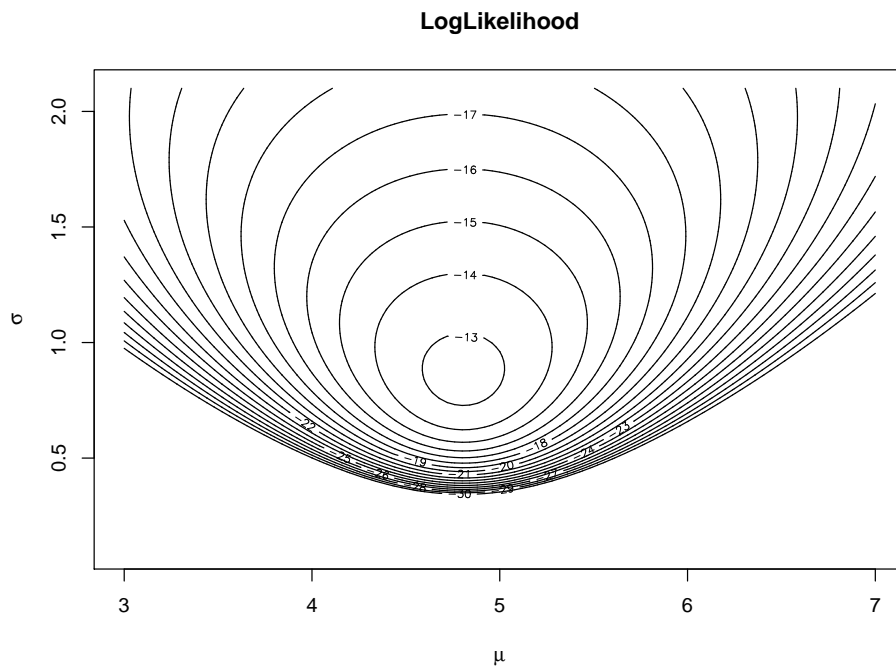
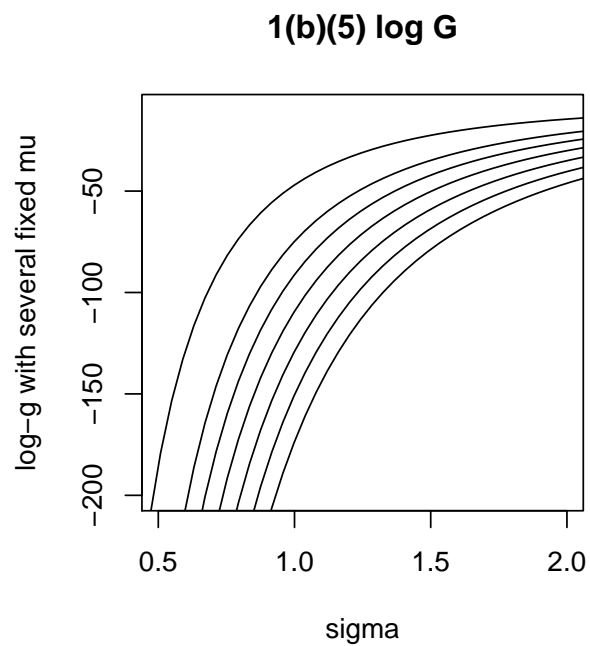
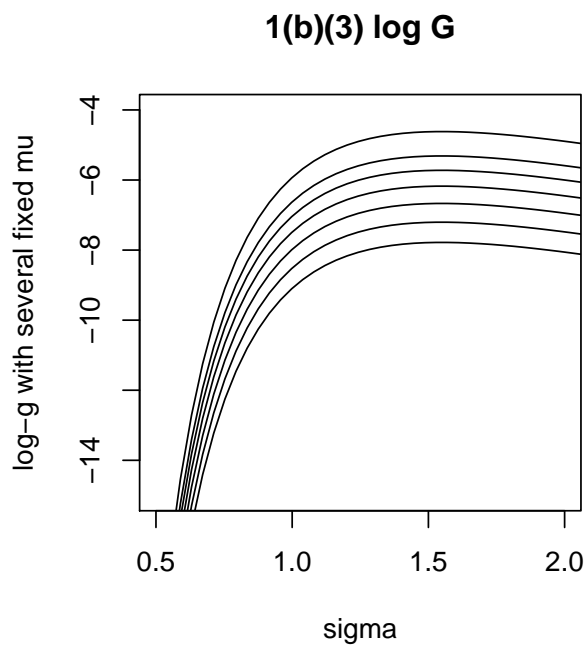


1 (a) The contour plot for the log-likelihood, $\log L(\mu, \sigma | X_1, \dots, X_{10})$.



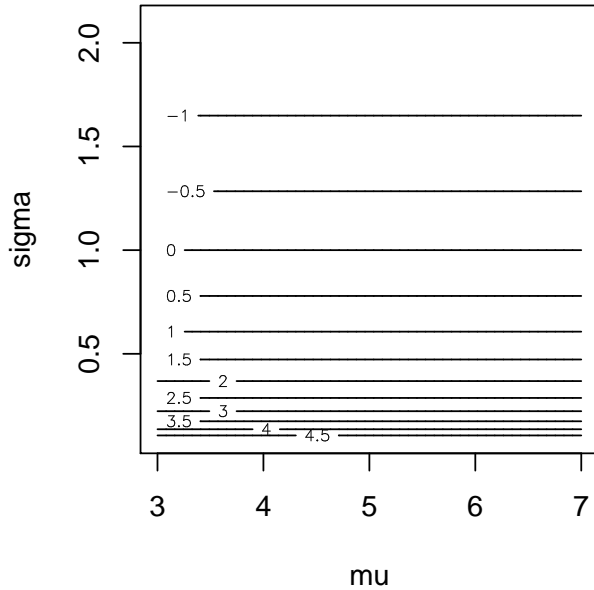
(b)

For independent prior, $\log g(\mu, \sigma^2) = \log g_1(\mu) + \log g_2(\sigma^2)$. When μ is fixed, $\log g(\mu, \sigma) = \log g_2(\sigma^2) + \text{constant}$ if μ and σ are independent. The following figure shows their difference. The lines in the left hand side are parallel but the lines in the right hand side are not.

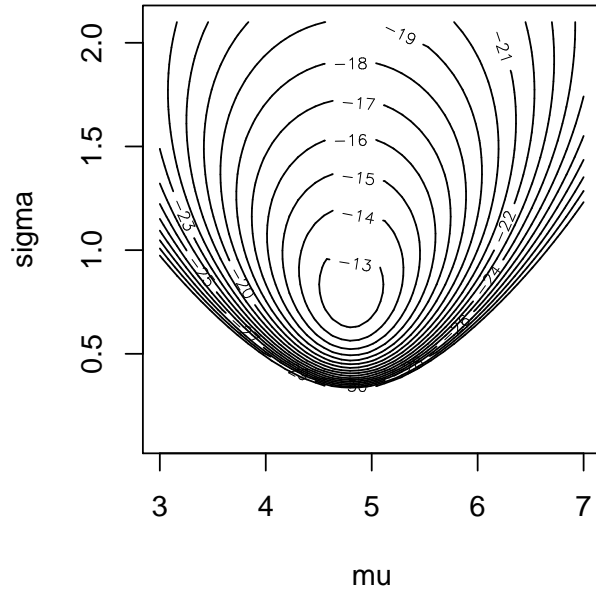


1. The contour plot for the log $g(\mu, \sigma)$ and $\log L(\mu, \sigma) + \log g(\mu, \sigma)$ with Jeffreys improper priors, $g(\mu, \sigma) = \frac{1}{\sigma^2}$.

1(b)(1) log-g

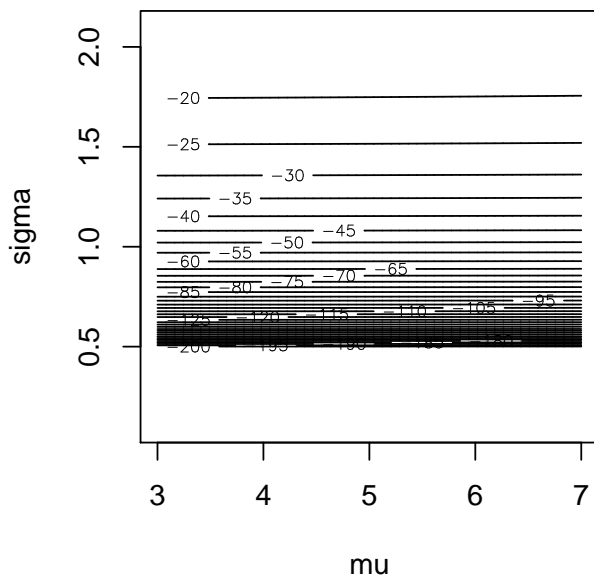


1(b)(1) log-g + log L

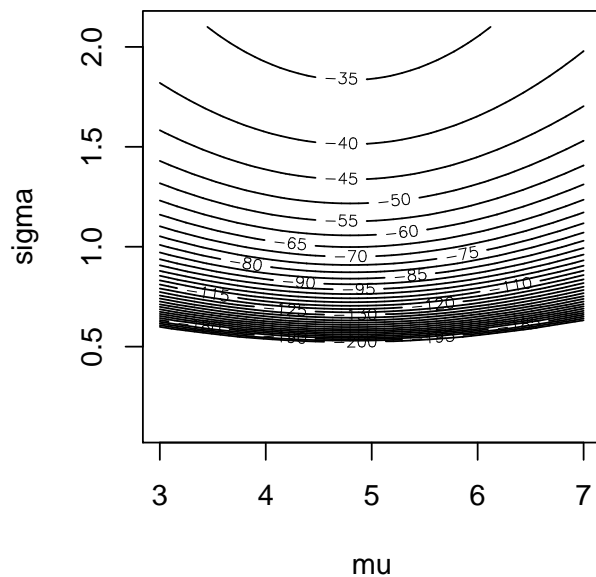


2. The contour plot for the log $g(\mu, \sigma)$ and $\log L(\mu, \sigma) + \log g(\mu, \sigma)$ with a priori $g(\mu, \sigma)$, $\mu \sim N(0, 10^2)$ independent of $\sigma^2 \sim \text{Inv-}\chi^2(1, 10^2)$.

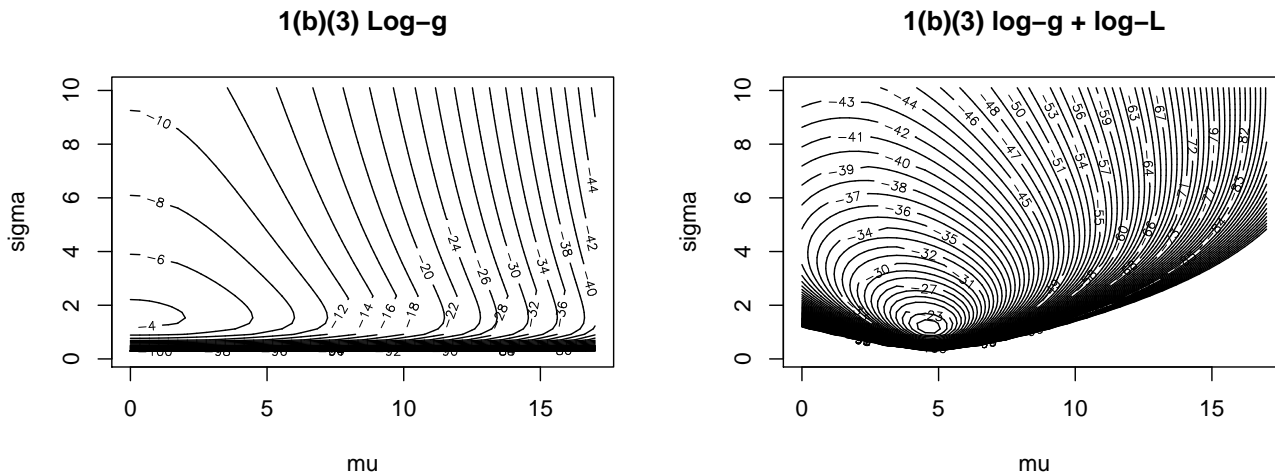
1(b)(2) Log-g



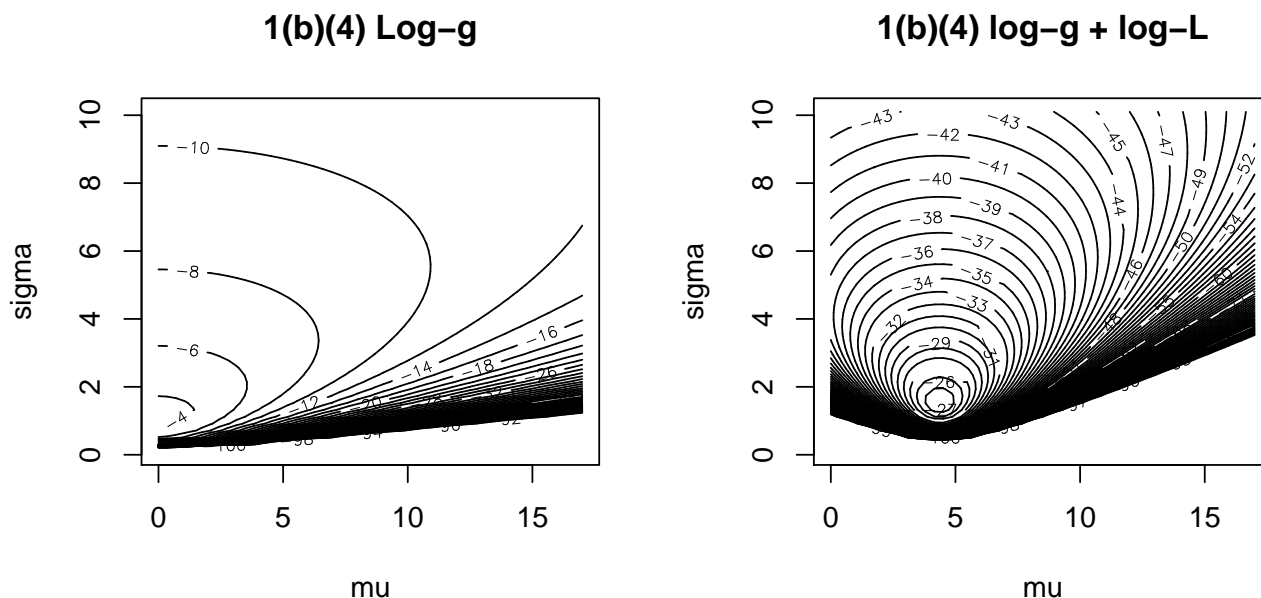
1(b)(2) log-g + log-L



3. The contour plot for the $\log g(\mu, \sigma)$ and $\log L(\mu, \sigma) + \log g(\mu, \sigma)$ with a priori $g(\mu, \sigma)$, $\mu \sim N(0, 2^2)$ independent of $\sigma^2 \sim \text{Inv-}\chi^2(3, 2^2)$.

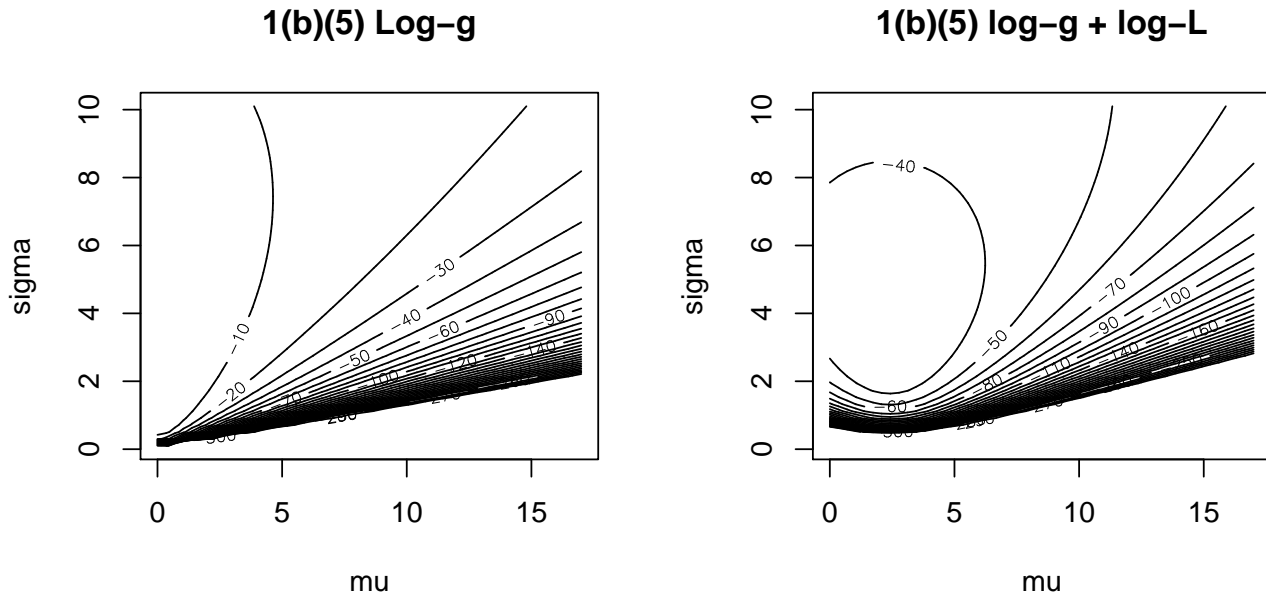


4. The contour plot for the $\log g(\mu, \sigma)$ and $\log L(\mu, \sigma) + \log g(\mu, \sigma)$ with a priori $g(\mu, \sigma)$, $\mu \sim N(0, \sigma^2)$, $\sigma^2 \sim \text{Inv-}\chi^2(1, 2^2)$.



In the 1st to 3rd plot, the shape of marginal slices fixed and only location changed. In the 4th and 5th plot(next page), since the shape of marginal slices changed, the contour plots show fan shape in the right end.

5. The contour plot for the log $g(\mu, \sigma)$ and log $L(\mu, \sigma) + \log g(\mu, \sigma)$ with a priori $g(\mu, \sigma)$, $\mu \sim N(0, \sigma^2/10)$, $\sigma^2 \sim \text{Inv-}\chi^2(1, 2^2)$.



Several ways to find the pdf of a scaled $\text{Inv-}\chi^2$ distribution.

$X \sim \text{Inv-}\chi^2(\nu, s^2)$ with pdf $f(x) = \frac{(s^2\nu/2)^{\nu/2}}{\Gamma(\nu/2)} x^{-\frac{\nu}{2}-1} \exp\left(\frac{-\nu s^2}{2x}\right), \nu > 0, s^2 > 0, x > 0$

$\Leftrightarrow X \sim \text{Inv-}\Gamma\left(\frac{\nu}{2}, \frac{\nu s^2}{2}\right)$

$\Leftrightarrow X = \nu s^2 \frac{1}{Y}, Y \sim \chi^2(\nu)$, get $f(x)$ by the transformation from $f(y)$

```
Rcode:
gsize=50
x<-c(4.90791,4.83425,5.19801,6.85308,4.07085,4.66076,4.16201,5.49752,4.15438,3.72703)
mu<-3+seq(-3,14,length=gsize)
sigma<-.1+seq(0,10,length=gsize)
theta<-expand.grid(mu,sigma)
logL<- function (theta){sum(dnorm(x,theta[1],theta[2], log=TRUE)) }
z<-apply(theta,1,logL)
z<-matrix(z,gsize,gsize)

library(geoR)
logL=function(theta){ sum(dnorm(x,theta[1],theta[2],log=TRUE)) }
logG1=function(theta){ -2*log( theta[2] ) }
logG2=function(theta){ dnorm(theta[1], mean=0, sd=10, log=TRUE) + dinvchisq(theta[2]^2, 1, 100, log=TRUE) }
logG3=function(theta){ dnorm(theta[1], mean=0, sd=2, log=TRUE) + dinvchisq(theta[2]^2, 3, 4, log=TRUE) }
logG4=function(theta){ dnorm(theta[1], mean=0, sd=theta[2], log=TRUE) + dinvchisq(theta[2]^2, 1, 4, log=TRUE) }
logG5=function(theta){ dnorm(theta[1], mean=0, sd=theta[2]/sqrt(10), log=TRUE) + dinvchisq(theta[2]^2, 1, 4, log=TRUE) }

w<-matrix(apply(theta,1, logG1),gsize,gsize)
contour(mu,sigma,w,main=main1 ,xlab="mu", ylab="sigma")
contour(mu,sigma,w+z,main=main2, xlab="mu", ylab="sigma")
# These will give a less detailed contour plot and will provide information to find better values of "level".
```

2 (a)

The WinBUGS code specifies a set of bivariate Normal data $Y_i = (Y_{1i}, Y_{2i})^T \sim MVN(\boldsymbol{\mu}, \boldsymbol{\Sigma} = R^{-1} = D)$ for $i = 1, \dots, 10$. The prior of $\boldsymbol{\mu}$ is $\boldsymbol{\mu} \sim MVN\left(\text{Alpha} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \boldsymbol{\Sigma}_0 = \begin{pmatrix} 100 & 0 \\ 0 & 100 \end{pmatrix}\right)$. (i.e. $\boldsymbol{\tau} = \begin{pmatrix} 0.01 & 0 \\ 0 & 0.01 \end{pmatrix}$). The prior of $\boldsymbol{\Sigma}$ is $\boldsymbol{\Sigma} = D = R^{-1} \sim \text{Inv-Wishart}_{\nu=4}\left(\text{Lambda}^{-1} = \Delta = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}^{-1}\right)$. Therefore, $\boldsymbol{\Sigma}$ has mean $(\nu - k - 1)^{-1} \text{Lambda} = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}$.

The following are statistic result from **WinBUGS** for 5 parameters, with 4 chains, initial values are generated by default, and beg from 90000 to end 100000 with thin 10 to get 4000 samples.

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
mu[1]	4.892	0.5352	0.009465	3.865	4.877	6.009	90000	4000
mu[2]	4.254	0.5776	0.009552	3.04	4.273	5.384	90000	4000
rho	0.5822	0.2813	0.004229	-0.171	0.6583	0.9181	90000	4000
sig1	1.43	0.4093	0.006233	0.8715	1.352	2.425	90000	4000
sig2	1.626	0.4085	0.006277	1.048	1.56	2.63	90000	4000

(b)

The following are statistic result from **WinBUGS** for 5 parameters, set $\nu = 20$ and $\text{Lambda} = \begin{pmatrix} 68 & 0 \\ 0 & 68 \end{pmatrix}$ such that hold the same prior mean for $\boldsymbol{\Sigma}$, with 4 chains, initial values are generated by default, and beg from 90000 to end 100000 with thin 10 to get 4000 samples.

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
mu[1]	4.664	0.6907	0.0108	3.327	4.666	6.077	90000	4000
mu[2]	4.456	0.6725	0.01183	3.111	4.457	5.759	90000	4000
rho	0.06784	0.1975	0.002921	-0.3219	0.07103	0.4437	90000	4000
sig1	1.823	0.2731	0.004557	1.38	1.794	2.43	90000	4000
sig2	1.871	0.2726	0.00455	1.423	1.837	2.485	90000	4000

Since the prior suggest higher σ , and the degrees of freedom $\nu = 20$ is increased to hold the same prior mean for $\boldsymbol{\Sigma}$ when changing the Lambda , that induced higher mean of σ and smaller ρ and also induced wider credible intervals for mu , narrower credible intervals for σ and smaller ρ .

(c)

The following are statistic result from **WinBUGS** for 5 parameters, set $\boldsymbol{\Sigma}_0 = \tau_0 = I_{2 \times 2}$, with 4 chains, initial values are generated by default, and beg from 90000 to end 100000 with thin 10 to get 4000 samples.

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
mu[1]	2.257	0.9878	0.01849	0.1471	2.318	3.971	90000	4000
mu[2]	1.414	0.9284	0.01534	-0.4374	1.413	3.205	90000	4000
rho	0.8887	0.1214	0.002136	0.5471	0.9276	0.9855	90000	4000
sig1	3.061	1.257	0.02328	1.276	2.842	6.134	90000	4000
sig2	3.071	1.001	0.01477	1.573	2.917	5.451	90000	4000

Since the prior for μ suggest a more condensed distribution, the posterior means are closer to the prior means zero. That also increases the standard deviation and induce a wider credible interval.

(d)

stats result from **WinBUGS** for 5 parameters, set $Lambda = \begin{pmatrix} 4 & -3 \\ -3 & 4 \end{pmatrix}$, with 4 chains, initial values are generated by default, and beg from 90000 to end 100000 with thin 10 to get 4000 samples.

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
mu[1]	4.748	0.5364	0.009367	3.703	4.732	5.837	90000	4000
mu[2]	4.383	0.5922	0.009807	3.153	4.397	5.548	90000	4000
rho	0.2609	0.3664	0.005364	-0.5033	0.3024	0.8277	90000	4000
sig1	1.379	0.3938	0.00565	0.8584	1.295	2.378	90000	4000
sig2	1.612	0.4059	0.006606	1.038	1.539	2.621	90000	4000

Comparing to (b), this prior suggest a negative correlation between 2 variables. The posterior distribution reflects this change.

(e)

stats result from **WinBUGS** for 5 parameters, set $Lambda = \begin{pmatrix} 40 & 0 \\ 0 & 40 \end{pmatrix}$, with 4 chains, initial values are generated by default, and beg from 90000 to end 100000 with thin 10 to get 4000 samples.

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
mu[1]	4.638	0.9746	0.01672	2.644	4.643	6.634	90000	4000
mu[2]	4.432	0.9329	0.01547	2.527	4.443	6.297	90000	4000
rho	0.09748	0.3043	0.004519	-0.4965	0.1128	0.6533	90000	4000
sig1	2.55	0.6948	0.009826	1.622	2.412	4.342	90000	4000
sig2	2.584	0.6362	0.01011	1.656	2.473	4.15	90000	4000

Compare to (b) the prior of σ suggest higher expectation. The posterior distribution reflects this change.

The following table shows the estimate of means and covariance structure and the corresponding credible intervales and length of these intervales.

Note	(a)Mean	2.50	97.50	Length	(b)Mean	2.50	97.50	Length	(c)Mean	2.50	97.50	Length
mu[1]	4.892	3.865	6.009	2.144	4.664	3.327	6.077	2.75	2.257↓	0.1471	3.971	3.8239
mu[2]	4.254	3.04	5.384	2.344	4.456	3.111	5.759	2.648	1.414↓	-0.4374	3.205	3.6424
rho	0.5822	-0.171	0.9181	1.0891	0.06784 ↓	-0.3219	0.4437	0.7656	0.8887 ↑	0.5471	0.9855	0.4384
sig1	1.43	0.8715	2.425	1.5535	1.823	1.38	2.43	1.05	3.061 ↑	1.276	6.134	4.858
sig2	1.626	1.048	2.63	1.582	1.871	1.423	2.485	1.062	3.071 ↑	1.573	5.451	3.878
Note	(d)Mean	2.50	97.50	Length	(e)Mean	2.50	97.50	Length				
mu[1]	4.748	3.703	5.837	2.134	4.638	2.644	6.634	3.99				
mu[2]	4.383	3.153	5.548	2.395	4.432	2.527	6.297	3.77				
rho	0.2609 ↓	-0.5033	0.8277	1.331	0.09748 ↓	-0.4965	0.6533	1.1498				
sig1	1.379	0.8584	2.378	1.5196	2.55	1.622	4.342	2.72				
sig2	1.612	1.038	2.621	1.583	2.584	1.656	4.15	2.494				