SDS 383D Exercise 2: Bayes and the Gaussian Linear Model

Preston Biro

February 2018

A Simple Gaussian Location Model

A

$$
\begin{aligned}
p(\theta) &= \int p(\theta,\omega) d\omega \\
&\propto \int \omega^{\frac{d+1}{2}-1} e^{-\omega\frac{\kappa(\theta-\mu)^2}{2}} e^{-\omega\frac{\eta}{2}} d\omega \\
&= \int \omega^{\frac{d+1}{2}-1} e^{-\omega\frac{\kappa(\theta-\mu)^2+\eta}{2}} d\omega \\
&= \int \omega^{\alpha-1} e^{-\omega\beta} \\
\text{Where } \alpha = \frac{d+1}{2} &\text{ and } \beta = \frac{\kappa(\theta-\mu)^2+\eta}{2} \\
\implies &= \Gamma(\alpha)\beta^{-\alpha} \\
&\propto (\frac{\kappa(\theta-\mu)^2+\eta}{2})^{-\frac{d+1}{2}} \\
&= (\frac{\kappa(\theta-\mu)^2}{2} + \frac{\eta}{2})^{-\frac{d+1}{2}} \\
&= (1 + \frac{1}{d}\frac{d\kappa(\theta-\mu)^2}{\eta})^{-\frac{d+1}{2}} \\
p(\theta) &= (1 + \frac{1}{\nu}\frac{(\theta-m)^2}{s^2})^{-\frac{\nu+1}{2}}
\end{aligned}
$$

With:

$$
\begin{aligned}
\nu &= d \\
m &= \mu \\
s^2 &= \frac{\eta}{d\kappa}
\end{aligned}
$$

B

$$
\begin{aligned}
p(\theta, \omega|\mathbf{y}) &\propto p(\mathbf{y}|\theta, \omega)p(\theta, \omega)\\
p(\mathbf{y}|\theta, \omega) &= p(\mathbf{y}|\theta, \frac{1}{\sigma^2})\\
p(\mathbf{y}|\theta, \sigma^2) &= \prod_{i=1}^{n} p(y_i|\theta, \sigma^2)\\
&= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(\theta-y_i)^2}{2\sigma^2}}\\
&\propto (\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\theta-y_i)^2}\\
&= (\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}(\sum_{i=1}^{n}(\theta-\bar{y})^2+\sum_{i=1}^{n}(\bar{y}-y_i)^2)}\\
p(\mathbf{y}|\theta, \sigma^2) &\propto (\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}(n(\theta-\bar{y})^2+S_y)}\\
\text{Where } S_y &= \sum_{i=1}^{n}(y_i - \bar{y})^2\\
\implies p(\mathbf{y}|\theta, \omega) &\propto \omega^{n/2} e^{-\frac{\omega}{2}(n(\theta-\bar{y})^2+S_y)}
\end{aligned}
$$

$$
\begin{aligned}
p(\theta, \omega|\mathbf{y}) &\propto p(\mathbf{y}|\theta, \omega)p(\theta, \omega)\\
&\propto \omega^{\frac{n}{2}} e^{-\frac{\omega}{2}(n(\theta-\bar{y})^2+S_y)} \omega^{\frac{d+1}{2}-1} e^{-\frac{\omega}{2}\kappa(\theta-\mu)^2} e^{-\frac{\omega}{2}\eta}\\
&= \omega^{\frac{n+d+1}{2}-1} e^{-\frac{\omega}{2}(n(\theta-\bar{y})^2+\kappa(\theta-\mu)^2)} e^{-\frac{\omega}{2}(S_y+\eta)}
\end{aligned}
$$

Manipulating the exponent on the first exponential term:

$$
\begin{aligned}
-\frac{\omega}{2}(n(\theta-\bar{y})^2+\kappa(\theta-\mu)^2) &= -\frac{\omega}{2}(n\theta^2 - 2n\theta\bar{y} + n\bar{y}^2 + \kappa\theta^2 - 2\kappa\theta\mu + \kappa\mu^2)\\
&= -\frac{\omega}{2}(\theta^2(n+\kappa) - 2\theta(n\bar{y}+\kappa\mu) + n\bar{y}^2 + \kappa\mu^2)\\
&= -\frac{\omega}{2}((n+\kappa)[\theta^2 - 2\theta\frac{(n\bar{y}+\kappa\mu)}{(n+\kappa)} + (\frac{(n\bar{y}+\kappa\mu)}{(n+\kappa)})^2] - (n+\kappa)(\frac{(n\bar{y}+\kappa\mu)}{(n+\kappa)})^2 + n\bar{y}^2 + \kappa\mu^2)\\
&= -\frac{\omega}{2}((n+\kappa)(\theta - \frac{n\bar{y}+\kappa\mu}{n+\kappa})^2 - (n+\kappa)(\frac{n\bar{y}+\kappa\mu}{n+\kappa})^2 + n\bar{y}^2 + \kappa\mu^2)
\end{aligned}
$$

We will now take these last three terms out of this exponential and group it with the following exponential term:

$$
\begin{aligned}
p(\theta, \omega|\mathbf{y}) &\propto \omega^{\frac{n+d+1}{2}-1} e^{-\frac{\omega}{2}(n(\theta-\bar{y})^2+\kappa(\theta-\mu)^2)} e^{-\frac{\omega}{2}(S_y+\eta)}\\
&= \omega^{\frac{n+d+1}{2}-1} e^{-\frac{\omega}{2}((n+\kappa)(\theta-\frac{n\bar{y}+\kappa\mu}{n+\kappa})^2} e^{-\frac{\omega}{2}(S_y+\eta-(n+\kappa)(\frac{n\bar{y}+\kappa\mu}{n+\kappa})^2+n\bar{y}^2+\kappa\mu^2)}\\
&= \omega^{\frac{n+d+1}{2}-1} e^{-\frac{\omega}{2}((n+\kappa)(\theta-\frac{n\bar{y}+\kappa\mu}{n+\kappa})^2} e^{-\frac{\omega}{2}(S_y+\eta+\frac{n\kappa}{n+\kappa}(\mu-\bar{y})^2)}
\end{aligned}
$$

We can now see this is in the form of:

$$
p(\theta, \omega|\mathbf{y}) \propto \omega^{\frac{d^\star+1}{2}-1} \exp\left\{-\omega \cdot \frac{\kappa^\star(\theta-\mu^\star)^2}{2}\right\} \cdot \exp\left\{-\omega \cdot \frac{\eta^\star}{2}\right\}
$$

With

$$\begin{aligned}
d^\star &= n + d \\
\kappa^\star &= n + \kappa \\
\mu^\star &= \frac{n\bar{y} + \kappa\mu}{n + \kappa} \\
\eta^\star &= \eta + S_y + \frac{n\kappa}{n+\kappa}(\mu - \bar{y})^2
\end{aligned}$$

C

Since

$$p(\theta, \omega|\mathbf{y}) = p(\theta|\omega, \mathbf{y})p(\omega|\mathbf{y})$$

Then $p(\theta|\omega, \mathbf{y})$ must contain all the terms in the posterior containing θ. Thus

$$\begin{aligned}
p(\theta|\omega, \mathbf{y}) &\propto e^{-(\kappa^\star\omega)^{-1}\frac{(\theta-\mu^\star)^2}{2}} \\
&\sim N(\mu^\star, (\kappa^\star\omega)^{-1})
\end{aligned}$$

D

$$\begin{aligned}
p(\omega|y) &= \int p(\theta, \omega|y)d\theta \\
&= \int \omega^{\frac{d^\star+1}{2}-1} \exp\left\{-\omega \cdot \frac{\kappa^\star(\theta-\mu^\star)^2}{2}\right\} \cdot \exp\left\{-\omega \cdot \frac{\eta^\star}{2}\right\} d\theta \\
&= \omega^{\frac{d^\star+1}{2}-1} e^{-\omega\frac{\eta^\star}{2}} \int e^{-\omega\frac{\kappa^\star(\theta-\mu^\star)^2}{2}} d\theta \\
&= \omega^{\frac{d^\star+1}{2}-1} e^{-\omega\frac{\eta^\star}{2}} \int N(\mu^\star, (\omega\kappa^\star)^{-1}) d\theta \\
&\propto \omega^{\frac{d^\star+1}{2}-1} e^{-\omega\frac{\eta^\star}{2}} (\omega\kappa^\star)^{-\frac{1}{2}} \\
p(\omega|y) &\sim Ga(\frac{d^\star}{2}, \frac{\eta^\star}{2})
\end{aligned}$$

E

Since the posterior follows a Normal-Gamma model, the marginal posterior for θ should follow the same distribution as its prior but with the new parameters, thus:

$$\begin{aligned}
p(\theta|y) &\sim t_{v^\star}(m^\star, s^{2\star}) \\
&= (1 + \frac{1}{\nu^\star}\frac{(\theta - m^\star)^2}{s^{2\star}})^{-\frac{\nu^\star+1}{2}}
\end{aligned}$$

Where:

$$\begin{aligned}
\nu^\star &= d^\star \\
&= N + d \\
m^\star &= \mu^\star \\
&= \frac{n\bar{y} + \kappa\mu}{n + \kappa} \\
s^{2\star} &= \frac{\eta^\star}{d^\star \kappa^\star} \\
&= \frac{\eta + S_y + \kappa\mu^2 + n\bar{y}^2 - \frac{(n\bar{y}+\kappa\mu)^2}{n+\kappa}}{(n+d)(n+\kappa)}
\end{aligned}$$

F

False. In the limit of our parameters κ, η, and d approaching zero, our prior distributions are no longer valid probability distributions, as $p(\theta)$ would become 1 for all values of θ, and therefore would not integrate to a finite value, and $p(\omega)$ would become zero.

G

True. Even though the prior distributions would no longer hold in the limit, the posterior distribution would still be a valid probability distribution, as the parameters for the posterior $\kappa^\star, \eta^\star$, and $d^\star$ would be affected but still be nonzero values, as they are also influenced by the likelihood function.

H

True. As our prior parameters approach zero, our interval turns into the following:

$$\begin{aligned}
\theta &= m \pm t^\star s \\
&= \frac{n\bar{y} + \kappa\mu}{n + \kappa} \pm t^\star \sqrt{\frac{\eta + S_y + \kappa\mu^2 + n\bar{y}^2 - \frac{(n\bar{y}+\kappa\mu)^2}{n+\kappa}}{(n+d)(n+\kappa)}} \\
&= \frac{n\bar{y}}{n} \pm t^\star \sqrt{\frac{S_y + n\bar{y}^2 - \frac{(n\bar{y})^2}{n}}{n^2}} \\
&= \bar{y} \pm t^\star \sqrt{\frac{S_y}{n^2}} \\
&= \bar{y} \pm t^\star s
\end{aligned}$$

Which is the exact formula for a confidence interval in the frequentist setting.

The Conjugate Gaussian Linear Model

A

We will start by deriving the multivariate posterior:

$$
\begin{aligned}
p(\beta,\omega|\mathbf{y}) &\propto p(\mathbf{y}|\beta,\omega)p(\beta,\omega) \\
&\propto p(\mathbf{y}|\beta,\omega)p(\beta|\omega)p(\omega) \\
p(\omega) &\propto \omega^{\frac{d}{2}-1}e^{-\frac{\omega}{2}\eta} \\
p(\beta|\omega) &\propto |\omega K|^{1/2}e^{-\frac{1}{2}(\beta-m)^T\omega K(\beta-m)} \\
p(\mathbf{y}|\beta,\omega) &\propto |\omega\Lambda|^{1/2}e^{-\frac{1}{2}(\mathbf{y}-X\beta)^T\omega\Lambda(\mathbf{y}-X\beta)} \\
\implies p(\beta,\omega|\mathbf{y}) &\propto |\omega\Lambda|^{1/2}e^{-\frac{1}{2}(\mathbf{y}-X\beta)^T\omega\Lambda(\mathbf{y}-X\beta)}|\omega K|^{1/2}e^{-\frac{1}{2}(\beta-m)^T\omega K(\beta-m)}\omega^{\frac{d}{2}-1}e^{-\frac{\omega}{2}\eta} \\
&\propto \omega^{\frac{n}{2}}e^{-\frac{1}{2}(\mathbf{y}-X\beta)^T\omega\Lambda(\mathbf{y}-X\beta)}\omega^{\frac{p}{2}}e^{-\frac{1}{2}(\beta-m)^T\omega K(\beta-m)}\omega^{\frac{d}{2}-1}e^{-\frac{\omega}{2}\eta} \\
&= \omega^{\frac{n+p+d}{2}-1}e^{-\frac{1}{2}(\mathbf{y}-X\beta)^T\omega\Lambda(\mathbf{y}-X\beta)}e^{-\frac{1}{2}(\beta-m)^T\omega K(\beta-m)}e^{-\frac{\omega}{2}\eta} \\
&= \omega^{\frac{n+p+d}{2}-1}e^{-\frac{\omega}{2}((X\beta-\mathbf{y})^T\Lambda(X\beta-\mathbf{y})+(\beta-m)^T K(\beta-m))}e^{-\frac{\omega}{2}\eta}
\end{aligned}
$$

By manipulating the exponent:

$$
\begin{aligned}
(X\beta-\mathbf{y})^T\Lambda(X\beta-\mathbf{y})+(\beta-m)^TK(\beta-m) &= (\beta^TX^T-\mathbf{y}^T)\Lambda(X\beta-\mathbf{y})+(\beta^T-m^T)K(\beta-m) \\
&= \beta^TX^T\Lambda X\beta-2\beta^TX^T\Lambda\mathbf{y}+\mathbf{y}^T\Lambda\mathbf{y}+\beta^TK\beta-2\beta^TKm+m^TKm \\
&= \beta^T(X^T\Lambda X+K)\beta-2\beta^T(X^T\Lambda\mathbf{y}+Km)+\mathbf{y}^T\Lambda\mathbf{y}+m^TKm
\end{aligned}
$$

We will combine the last two terms with the next exponential term and complete the square on the first two terms as follows:

$$
\begin{aligned}
\beta^T(X^T\Lambda X+K)\beta-2\beta^T(X^T\Lambda\mathbf{y}+Km) &= (\beta-(X^T\Lambda\mathbf{y}+Km)(X^T\Lambda X+K)^{-1})^T(X^T\Lambda X+K)\cdot \\
&\quad\cdot(\beta-(X^T\Lambda\mathbf{y}+Km)(X^T\Lambda X+K)^{-1}) \\
&\quad-(X^T\Lambda\mathbf{y}+Km)^T(X^T\Lambda X+K)^{-1}(X^T\Lambda\mathbf{y}+Km) \\
&= (\beta-\mu^\star)^T\kappa^\star(\beta-\mu^\star)
\end{aligned}
$$

Where we take the last term and combine it with the final exponential term. This leaves our posterior in the following form:

$$
p(\beta,\omega|\mathbf{y}) = \omega^{\frac{d^\star}{2}-1}e^{-\frac{\omega}{2}(\beta-\mu^\star)^T\kappa^\star(\beta-\mu^\star)}e^{-\frac{\omega}{2}\eta^\star}
$$

Where our new parameters are:

$$
\begin{aligned}
d^\star &= n+p+d \\
\mu^\star &= (X^T\Lambda X+K)^{-1}(X^T\Lambda\mathbf{y}+Km) \\
\kappa^\star &= (X^T\Lambda X+K) \\
\eta^\star &= \eta+\mathbf{y}^T\Lambda\mathbf{y}+m^TKm-(X^T\Lambda\mathbf{y}+Km)^T(X^T\Lambda X+K)^{-1}(X^T\Lambda\mathbf{y}+Km)
\end{aligned}
$$

Now, to find $p(\beta|\omega,\mathbf{y})$, since we know $p(\beta,\omega|\mathbf{y}) = p(\beta|\omega,\mathbf{y})p(\omega|\mathbf{y})$, $p(\beta|\omega,\mathbf{y})$, we see that $p(\beta|\omega,\mathbf{y}) = \omega^{\frac{p}{2}}e^{-\frac{\omega}{2}(\beta-\mu^\star)^T\kappa^\star(\beta-\mu^\star)} \sim MVN(\mu^\star,(\omega\kappa^\star)^{-1})$.

B

Similarly, since we know $p(\beta,\omega|\mathbf{y}) = p(\beta|\omega,\mathbf{y})p(\omega|\mathbf{y})$, $p(\beta|\omega,\mathbf{y})$, and we know the distribution of $p(\beta|\omega,\mathbf{y})$, this means $p(\omega|\mathbf{y}) \propto \omega^{\frac{n+d}{2}-1}e^{-\frac{\omega}{2}\eta^\star} \sim Ga(\frac{n+d}{2},\frac{\eta^\star}{2})$.

C

$$
\begin{aligned}
p(\beta|\mathbf{y}) &= \int p(\beta,\omega|\mathbf{y})d\omega \\
&= \int \omega^{\frac{d^\star}{2}-1}e^{-\frac{\omega}{2}(\beta-\mu^\star)^T\kappa^\star(\beta-\mu^\star)}e^{-\frac{\omega}{2}\eta^\star}d\omega \\
&= \frac{\Gamma(\frac{n+d+p}{2})}{(\frac{(\beta-\mu^\star)^T\kappa^\star(\beta-\mu^\star)}{2}+\frac{\eta^\star}{2})^{\frac{n+d+p}{2}}}\int \frac{((\beta-\mu^\star)^T\kappa^\star(\beta-\mu^\star)+\eta^\star)^{\frac{n+d+p}{2}}}{\Gamma(\frac{n+d+p}{2})}\omega^{\frac{d^\star}{2}-1}e^{-\frac{\omega}{2}(\beta-\mu^\star)^T\kappa^\star(\beta-\mu^\star)}e^{-\frac{\omega}{2}\eta^\star}d\omega \\
&= \frac{\Gamma(\frac{n+d+p}{2})}{(\frac{(\beta-\mu^\star)^T\kappa^\star(\beta-\mu^\star)}{2}+\frac{\eta^\star}{2})^{\frac{n+d+p}{2}}} \\
&\propto (\frac{(\beta-\mu^\star)^T\kappa^\star(\beta-\mu^\star)}{2}+\frac{\eta^\star}{2})^{-\frac{n+d+p}{2}} \\
&= (1+(\beta-\mu^\star)^T\kappa^\star(\beta-\mu^\star)(\eta^\star)^{-1})^{-\frac{n+d+p}{2}} \\
&= (1+\frac{n+d}{n+d}(\beta-\mu^\star)^T\kappa^\star(\beta-\mu^\star)(\eta^\star)^{-1})^{-\frac{n+d+p}{2}} \\
&\sim t_{n+d}(\mu^\star,\frac{\eta^\star}{\kappa^\star(n+d)})
\end{aligned}
$$

D

Using the model described above, we fit a bayesian linear model to the gdpgrowth data set provided. The following graph shows the fit created using a prior mean of .025 and prior variance of .001. Code is included at the end of the document.

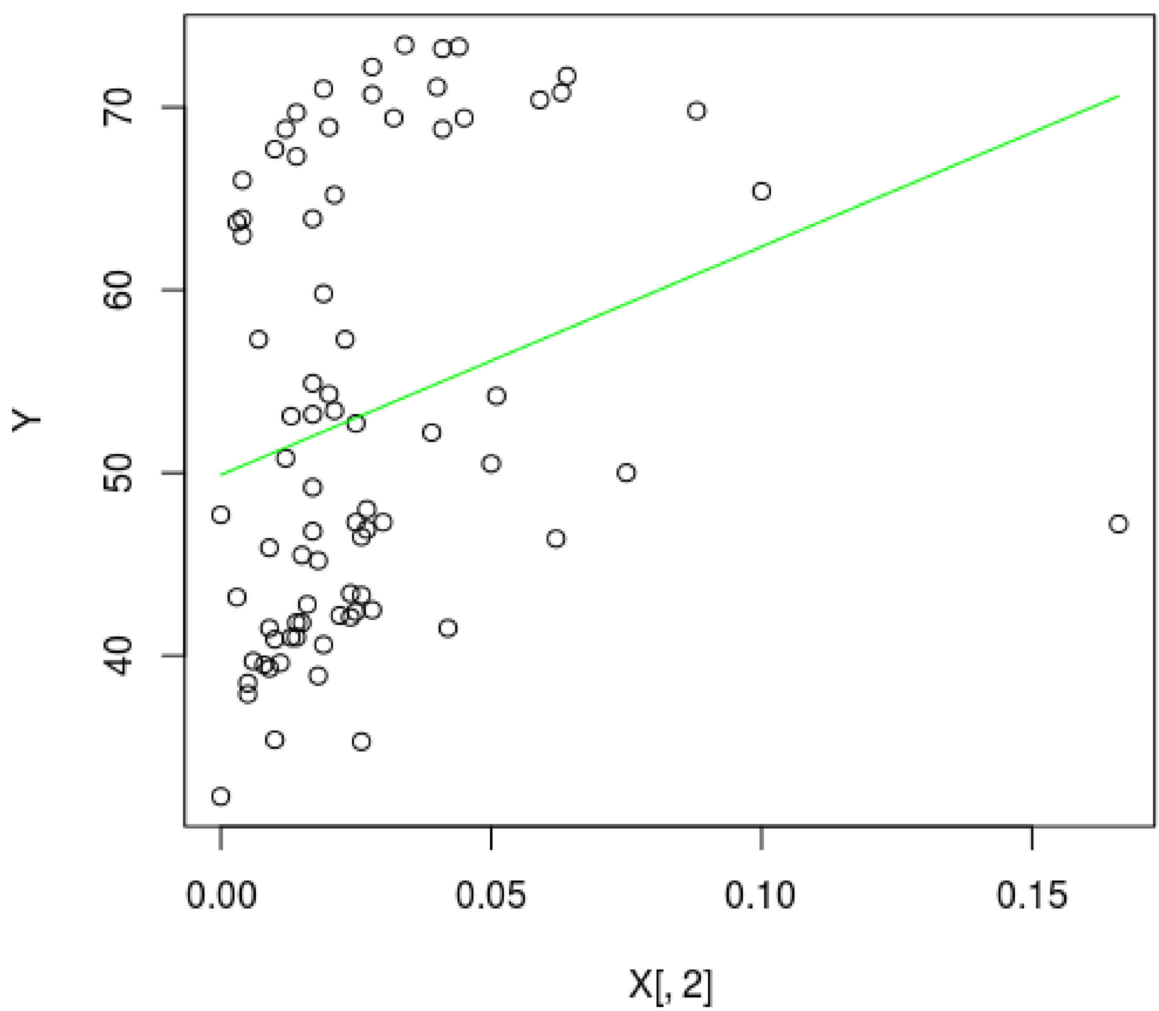

Figure 1: Bayesian Linear Model for Defense Spending as a Fraction of GDP

As you can see, the fit of a straight line does not seem to be a reasonable solution for this dataset. A nonlinear model would be better for this data, however the model still can be used.

A Heavy Tailed Error Model

A

Similar to our past normal gamma posteriors, the conditional distribution $y_i|X, \beta, \omega$ with λ_i marginalized out can be found by integrating the likelihood times the data distribution with respect to λ_i:

$$\begin{aligned}
p(y_i|\beta,\omega) &\propto \int p(y_i|\beta,\omega,\lambda_i)p(\lambda_i)d\lambda_i \\
&= \int N(X\beta,(\omega\Lambda)^{-1}Ga(\frac{h}{2},\frac{h}{2}))d\lambda_i \\
&\propto \int (\omega\lambda_i)^{\frac{1}{2}} e^{-\frac{\omega\lambda_i}{2}(y_i - X_i^T\beta)^2} \lambda_i^{\frac{h}{2}-1} e^{-\lambda_i \frac{h}{2}} d\lambda_i \\
&\propto \int \lambda_i^{\frac{h+1}{2}-1} e^{-\frac{\omega\lambda_i}{2}(y_i - X_i^T\beta)^2 + \lambda_i\frac{h}{2}} d\lambda_i \\
&\propto \int Ga(\frac{h+1}{2}, \frac{\omega}{2}(y_i - X_i^T\beta)^2 + \frac{h}{2})d\lambda_i \\
&= (\frac{\omega}{2}(y_i - X_i^T\beta)^2 + \frac{h}{2})^{-\frac{h+1}{2}} \\
p(y_i|\beta,\omega) &= (1 + \frac{\omega}{h}(y_i - X_i^T\beta)^2)^{-\frac{h+1}{2}}
\end{aligned}$$

Which is clearly a scaled and shifted t-distribution, as we have seen in the past, where in this case the new degrees of freedom is h, the new mean parameter is $X_i^T\beta$, and the new scale parameter is $\frac{1}{\omega}$.

B

Since

$$p(\lambda_i|\mathbf{y},\beta,\omega) \propto p(y_i|\beta,\omega,\lambda_i)p(\lambda_i)$$

it must retain all parts of the likelihood and prior that pertain to λ_i. But we showed what distribution the right side of the equation follows, which is

$$p(y_i|\beta,\omega,\lambda_i)p(\lambda_i) \sim Ga(\frac{h+1}{2}, \frac{\omega}{2}(y_i - X_i^T\beta)^2 + \frac{h}{2})$$

C

A Gibbs sampler was created to fit the same model from the gdpgrowth dataset as before. The following plot shows the fit made using the average beta values provided from the gibbs sampler. Gibbs is in blue and the previous fit is in green.

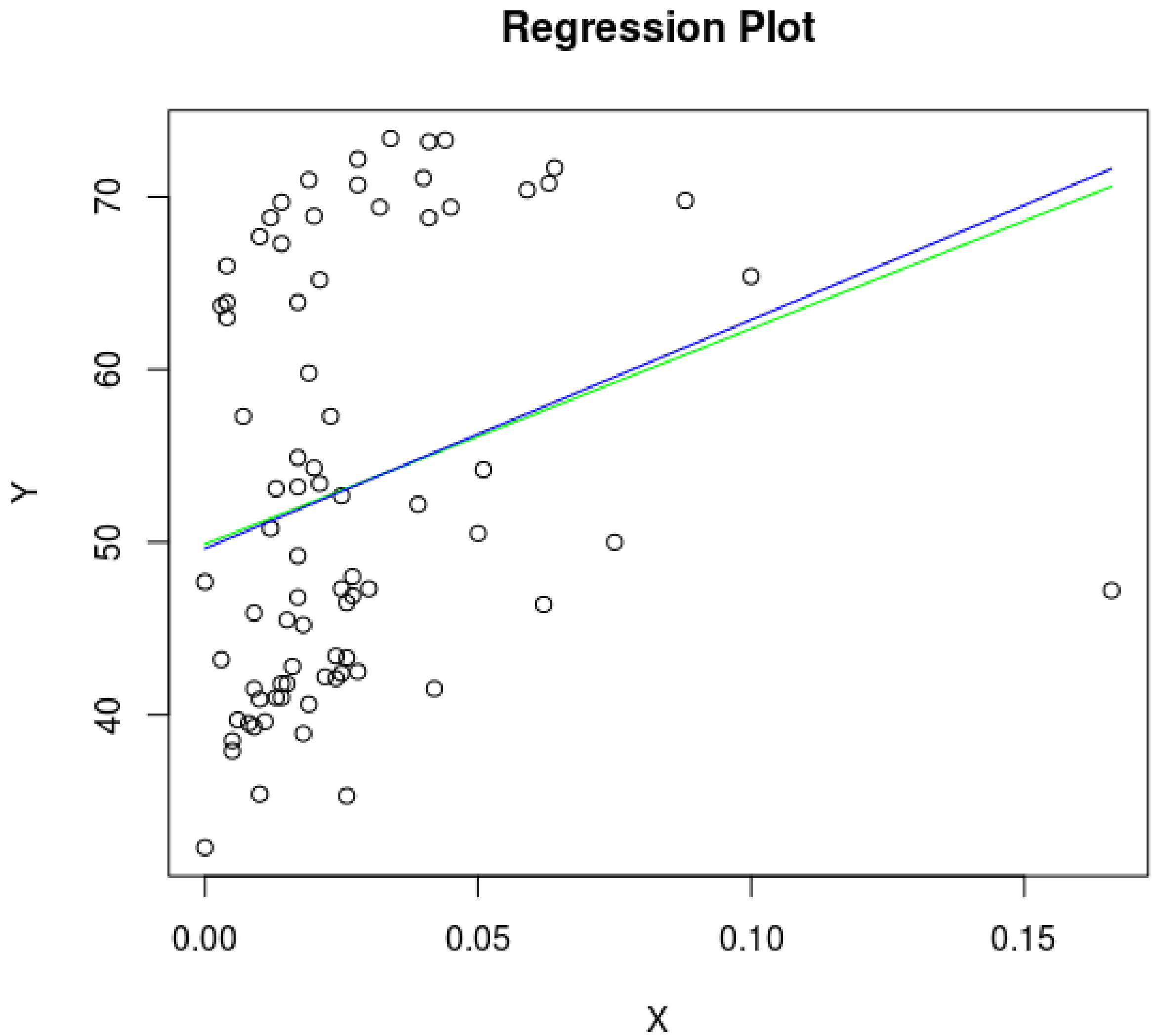

Figure 2: Gibbs Sampler Fit for Defense Spending as a Fraction of GDP

Once again, the a straight line does not fit the data well. However, this particular line seems to be reasonable for the set of data, showing the sampler did its job.

Code

```
#SDS 383D
#Exercise 2
library(mvtnorm)
setwd("~/Documents/SDS 383D")
```

```
data = read.csv('Data/gdpgrowth.csv')

X = data.matrix(data[4:11])
Y = data$GR6096
m = matrix(c(3.5,3,.03,.1,.025,6,.03,42))

fullBayesianReg <- function(xvals,yreal,priorM){
  #xvals should just be n X p data for x, no intercept
  #yreal should be the actual output for the regression
  #priorM should be a vector of guesses on the means of the columns of the xvals
  #Outputs a graph of the data with the estimated data values using a bayesian linear mode
  #with p components. Will plot the data with each of the p components on the x axis with
  #the actual y values plotted with black points and estimated with the red points yhat. T
  #will also be the line of best fit for the simple linear regression for the independent
  #Will also print out the pairs of the individual components in order
  X = xvals
  Y = yreal
  n = length(Y)

  p = length(X[1,])

  Lambda = diag(n)
  Kappa = .001* diag(p)
  m = priorM

  mu = solve(t(X)%*%Lambda%*%X + Kappa) %*%(t(X) %*% Lambda %*% Y + Kappa%*%m)
  betahat = mu[,1]

  Yhat = X %*% betahat

  betaMat = matrix(NA,2,p)

  for(i in 1:8){
    betaMat[,i] = simpleBayesianReg(X[,i],Y,m[i])
  }
  par(mfrow=c(2,ceiling(p/2)))
  for(i in 1:i){
    plot(X[,i],Y)
    points(X[,i],Yhat,col='red')
    xtemp = matrix(c(rep(1,n),X[,i]),ncol=2)
    lines(X[,i],xtemp %*% betaMat[,i],col='green')
  }
}

simpleBayesianReg <- function(xvals,yreal,priorM){
  #xvals should just be 1-D data for x, no intercept
  #yreal should be the actual output for the regression
  #priorM should be a guess on the mean of x
  #Outputs a graph of the data with the Bayesian line of best fit over it along with
  #the yhat guessed values for the lines
  #Will also print out the beta coefficients
```

```
    X = matrix(c(rep(1,length(xvals)),xvals),ncol=2)
    Y = yreal
    m = c(0,priorM)

    n = length(data$CODE)
    p = 2

    Lambda = diag(n)
    Kappa = matrix(c(.001,0,0,.001),nrow=2)

    mu = solve(t(X)%*%Lambda%*%X + Kappa) %*% (t(X) %*% Lambda %*% Y + Kappa%*%m)

    Yhat = X %*% mu
    par(mfrow=c(1,1))
    plot(X[,2],Y)
    lines(X[,2],Yhat,col='green')
    return(mu)
}

X = data$DEF60
m = .025
simpleBayesianReg(X,Y,m)

Xfull = data.matrix(data[4:11])
Y = data$GR6096
mfull = matrix(c(3.5,3,.03,.1,.025,6,.03,42))

fullBayesianReg(Xfull,Y,mfull)

GibbsSamp <- function(X,Y,h,m0,d,eta,k){
    dataX = matrix(c(rep(1,length(X)),X),ncol=2)
    dataY = Y
    m = c(0,m0)
    k = matrix(c(k,0,0,k),nrow=2)
    #p = length(dataX[1,])
    p = 1
    n = length(dataY)
    dstar = n + d + p

    lambdas = rgamma(n,h/2,h/2)
    Lambda = diag(lambdas)
    mustar = solve(t(dataX)%*%Lambda%*%dataX + k) %*%(t(dataX) %*% Lambda %*% dataY + k%*%m)
    kappastar = t(dataX)%*%Lambda%*%dataX + k
    etastar = eta + t(dataY) %*% Lambda %*% dataY + t(m) %*% k %*% m -
      t(t(dataX) %*% Lambda %*% dataY + k%*%m) %*% mustar
    omega = rgamma(1,d/2,eta/2)

    Burn = 200
    for(i in 1:Burn){
      beta = matrix(rmvnorm(1,mustar,solve(omega * kappastar)))
      omega = rgamma(1,dstar/2,etastar/2)
```

```
    #print((dataY - dataX%*%beta)^2)
    lambdas = rgamma(n,(h+1)/2,(omega*(dataY - dataX%*%beta)^2 + h)/2)
    Lambda = diag(lambdas)
    mustar = solve(t(dataX)%*%Lambda%*%dataX + k) %*%(t(dataX) %*% Lambda %*% dataY + k%*%
    kappastar = t(dataX)%*%Lambda%*%dataX + k
    etastar = eta + t(dataY) %*% Lambda %*% dataY + t(m) %*% k %*% m -
      t(t(dataX) %*% Lambda %*% dataY + k%*%m) %*% mustar
    #print(lambdas)
  }

  B = 1000
  lambdaStore = matrix(NA,B,n)
  betaStore = matrix(NA,B,2)
  for(i in 1:B){
    beta = matrix(rmvnorm(1,mustar,solve(omega * kappastar)))
    omega = rgamma(1,dstar/2,etastar/2)
    lambdas = rgamma(n,(h+1)/2,(omega*(dataY - dataX%*%beta)^2 + h)/2)
    Lambda = diag(lambdas)
    mustar = solve(t(dataX)%*%Lambda%*%dataX + k) %*%(t(dataX) %*% Lambda %*% dataY + k%*%
    kappastar = t(dataX)%*%Lambda%*%dataX + k
    etastar = eta + t(dataY) %*% Lambda %*% dataY + t(m) %*% k %*% m -
      t(t(dataX) %*% Lambda %*% dataY + k%*%m) %*% mustar
    lambdaStore[i,] = lambdas
    betaStore[i,] = t(beta)
  }
  gibbBeta = colMeans(betaStore)

  betaSimp = simpleBayesianReg(X,Y,m0)
  plot(X,Y,main='Regression Plot')
  lines(X,dataX %*% betaSimp,col='green')
  lines(X,dataX %*% gibbBeta,col='blue')
  #legend('bottomright',c('Bayesian Linear Model','Gibbs Sampling Model'),col=c('green','b
  return(colMeans(lambdaStore))
}

GibbsSamp(X,Y,h,m,d,eta,k)

#X = data.matrix(data[4:11])
X = data$DEF60
Y = data$GR6096
h = 40
#m = matrix(c(3.5,3,.03,.1,.025,6,.03,42))
m = .025
d = 5
eta = 20
#k = .01*diag(diag(var(X)))
k = .001
GibbsSamp(X,Y,h,m,d,eta,k)
```