Bayesian Data Analysis EC543

Instructor: M.A. Rahman

Submitted by: Mohit Shukla 11435

Date: March 27, 2015

Problem:

Using attached data set (i.e. sales price of 546 houses sold in Winsor, Canada), perform regression with conjugate prior on multiple variable to illustrate Bayesian Inference to find out which factor affects the house price.

Solution:

Analysis had 11 explanatory variable and one dependent variable.
For posterior calculation both informative and non-informative priors were used. Results showed that our prior is relatively non-informative since posterior results based on informative prior are quite similar to non-informative prior.

Analysis is performed to illustrate the use of Gibbs-sampling for Bayesian inference in normal linear regression model with independent Normal-Gamma prior.
Results showed similar mean and standard deviation as in the previous case reflecting the fact that we have used similar

informative prior.

Results also showed that model comparison results are more different than the posterior mean.

Code:

```
%% Mohit Shukla


%Informative Prior

data=xlsread('C:/Users/caped_crusader/Bayesian Data
Analyses/Assignments/HousePrice.xlsx');

%Hyperparameters
V_pr=[2.4 0 0 0 0;0 6*10^-7 0 0 0;0 0 0.15  0 0; 0 0 0 0.6 0; 0 0 0 0
0.6];

b_pr=[0;10;5000;10000;10000];       %Prior beta

h=4*10^-8;
N=546;
n_pr=5;     %Prior nu
k=5;        %No. of parameters

var_b_pr=n_pr*V_pr/((n_pr-2)*h);        %Variance of prior beta
sd_b_pr=sqrtm(var_b_pr);                %Std Dev of prior beta

n_post=N-n_pr;

y=data(1:546,1);   %Dependent Variable

x=ones(546,1);      %Covariate X1

x(1:546,2:5)=data(1:546,2:5); %Covariates

V_post=(V_pr^-1 + transpose(x)*x)^-1;   %Posterior V

xsq=(transpose(x)*x);

b_ols=(((transpose(x))*x)^-1)*(transpose(x)*y);   %OLS beta

b_post=V_post*((V_pr^-1)*b_pr + transpose(x)*x*b_ols);  %Posterior
beta

v=N-k;                                                  %Degree of freedom

sse=(transpose(y-x*b_ols))*(y-x*b_ols);     %Std deviation of error
from OLS estimate
```

```
h_post=n_post*(n_pr*(h^-1) + sse+ (transpose(b_ols -
b_pr))*((V_pr+xsq^-1)^-1)*(b_ols - b_pr))^-1; %h posterior/inv of post
s sq

var_b_post=n_post*V_post/((n_post-2)*h_post);      %Variance of
posterior beta

sd_b_post=sqrtm(var_b_post);                      %Std dev of posterior
beta


%%Non Informative Prior

b_posterior=b_ols;         %Posterior beta

V_posterior=(transpose(x)*x)^-1;    %Posterior V




%% Mohit Shukla


%Informative Prior

data=xlsread('C:/Users/caped_crusader/Bayesian Data
Analyses/Assignments/HousePrice.xlsx');

%Hyperparameters
V_pr=[2.4 0 0 0 0;0 6*10^-7 0 0 0;0 0 0.15  0 0; 0 0 0 0.6 0; 0
0 0 0 0.6];

b_pr=[0;10;5000;10000;10000];       %Prior beta
h=4*10^-8;
N=546;
n_pr=5;     %Prior nu
k=5;        %No. of parameters
var_b_pr=n_pr*V_pr/((n_pr-2)*h);        %Variance of prior beta

sd_b_pr=sqrtm(var_b_pr);                %Std Dev of prior beta

n_post=N-n_pr;
y=data(1:546,1);   %Dependent Variable
x=ones(546,1);      %Covariate X1

x(1:546,2:5)=data(1:546,2:5); %Covariates

b_ols=(((transpose(x))*x)^-1)*(transpose(x)*y);   %OLS beta
```

```
xsq=(transpose(x)*x);

s_sq_inv=h^-1;
G=10000;
sum_b_exp=0;

for i=1:G
   V_posterior=(V_pr^-1 + h*xsq)^-1;
   b_posterior=V_posterior*(V_pr^-1*b_pr + h*transpose(x)*y);
   b=transpose(mvnrnd(b_posterior,V_posterior));   %Random draw
for beta
   n_posterior=N+n_pr;
   s_sq_inv=((transpose(y-x*b))*(y-x*b)+n_pr*(h^-
1))/n_posterior;
   h=gamrnd(n_posterior/2,s_sq_inv);        %Random draw for h
   if i>=0.1*G                             %Accepting values
after 1000 burns
   sum_b_exp=sum_b_exp+b;
   end
end

b_exp=sum_b_exp/(0.9*G)    %Beta obtained from simulation
```