STAT6011/7611/6111/3317 COMPUTATIONAL STATISTICS (2016 Fall) Assignment 3

Kei Ikegami (u3535947)

October 16, 2016

1

1.1

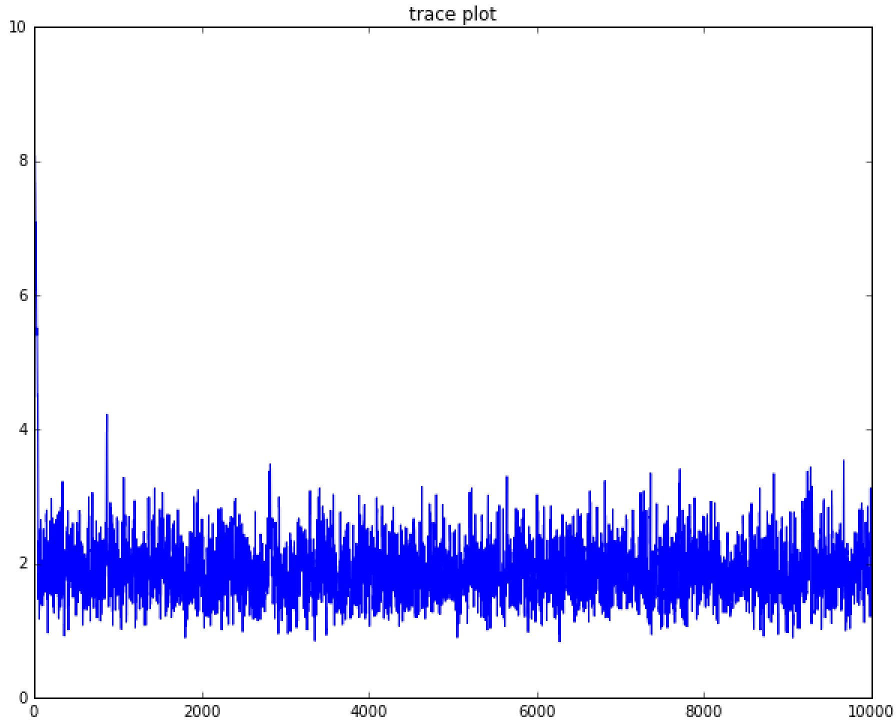
$$
\begin{aligned}
p(\mathbf{y}|\sigma^2) &= \int \cdots \int \Pi_{i=1}^n \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(y_i-\theta_i)^2}{2}\right) \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\theta_i^2}{2\sigma^2}\right) \mathrm{d}\theta_1 \ldots \mathrm{d}\theta_n \\
&= \int \cdots \int \Pi_{i=1}^n \frac{1}{\sqrt{2\pi}\sqrt{\sigma^2+1}} \exp\left(-\frac{y_i}{2} + \frac{\sigma^2 y_i^2}{2(\sigma^2+1)}\right) \frac{\sqrt{\sigma^2+1}}{\sqrt{2\pi}\sigma} \exp\left(\frac{\sigma^2+1}{\sigma^2}\left(\theta_i - \frac{\sigma^2}{\sigma^2+1} y_i\right)\right) \mathrm{d}\theta_1 \ldots \mathrm{d}\theta_n \\
&= \Pi_{i=1}^n \frac{1}{\sqrt{2\pi}\sqrt{\sigma^2+1}} \exp\left(-\frac{y_i}{2} + \frac{\sigma^2 y_i^2}{2(\sigma^2+1)}\right) \\
&= \left(\frac{1}{\sqrt{2\pi}\sqrt{\sigma^2+1}}\right)^n \exp\left(-\frac{\sum_{i=1}^n y_i}{2} + \frac{\sigma^2 \sum_{i=1}^n y_i^2}{2(\sigma^2+1)}\right)
\end{aligned}
$$

1.2

$$
p(\sigma^2|\mathbf{y}) \propto \left(\frac{1}{\sqrt{2\pi}\sqrt{\sigma^2+1}}\right)^n \exp\left(-\frac{\sum_{i=1}^n y_i}{2} + \frac{\sigma^2 \sum_{i=1}^n y_i^2}{2(\sigma^2+1)}\right) \frac{1}{\sigma^2}
$$

Listing 1: 1-b

```
import numpy as np
import matplotlib.pyplot as plt
% matplotlib inline
import pandas as pd

y = q1.values[:, 1]
def pos(sigma):
    return np.exp(sigma*np.dot(y, y)/(2*(sigma + 1)) - sum(y)/2) / (sigma*(np.sqrt(2*np.pi * (sigma + 1)))**100)

np.random.seed(1234)
samples = [10]
for i in range(10000):
    u = np.random.uniform()
    propose = samples[-1] + np.random.normal(scale = 0.5)
    prob = min(1, pos(propose)/pos(samples[-1]))

    if prob < u:
        samples.append(samples[-1])
    else:
        samples.append(propose)

plt.figure(figsize = (10, 8))
plt.title("trace plot")
plt.plot(samples)
plt.show()
```



1.3

Let $\Theta = (\theta_1, \theta_2, \dots, \theta_n)$, and let $\Theta_{-j} = (\theta_1, \dots, \theta_{j-1}, \theta_{j+1}, \dots, \theta_n)$.

$$p(\sigma^2|\mathbf{y}, \Theta) \propto \frac{1}{\sigma^2(\sqrt{2\pi\sigma^2})^n} \exp\left(-\frac{\sum_i^n \theta_i^2}{2\sigma^2}\right)$$

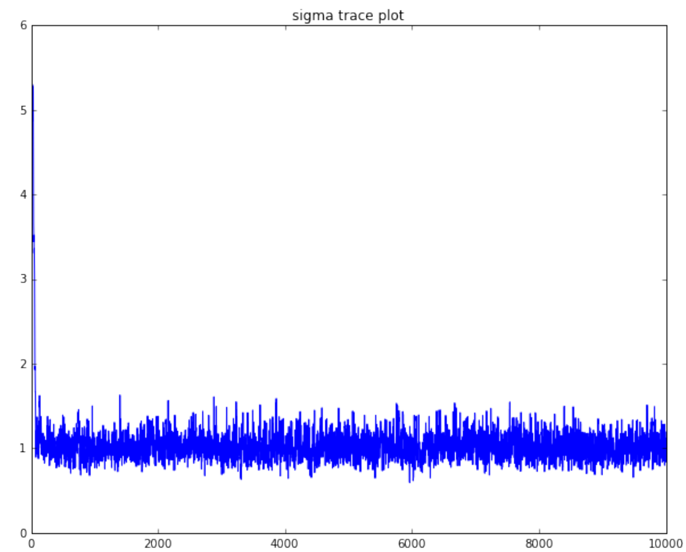
$$p(\theta_j|\mathbf{y}, \Theta_{-j}, \sigma^2) \propto N\left(\frac{\sigma^2 y_j}{\sigma^2+1}, \frac{\sigma^2}{\sigma^2+1}\right) \forall j \in \{1, 2, 3, \dots, n\}$$

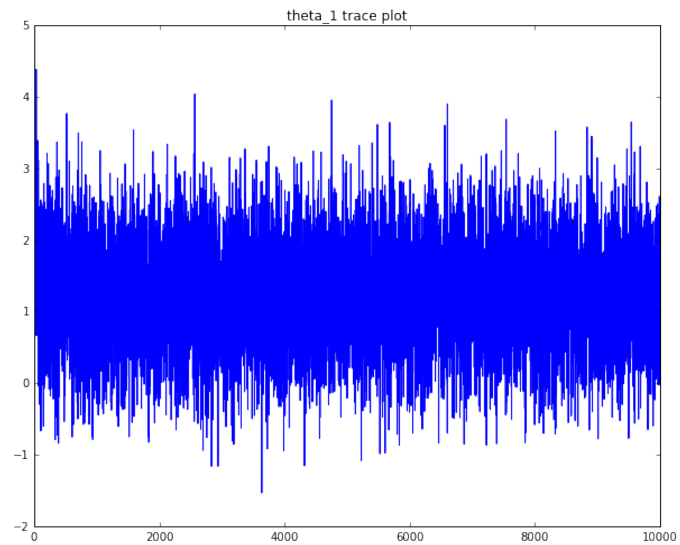
Listing 2: 1-c

```
def pos2(sigma, i):
    return np.exp(−np.dot(theta_sample[:, i], theta_sample[:, i])/(2∗sigma)) / (sigma∗(np.sqrt(2∗np.pi ∗ sigma))∗∗100)

# Gibbs sampler
np.random.seed(1234)

sigma_sample = [5]
theta_sample = np.ones((len(y), 10000))

for i in range(10000):

    if i != 0:
        u = np.random.uniform()
        propose = sigma_sample[−1] + np.random.normal(scale = 0.2)
        prob = min(1, pos2(propose, i)/pos2(sigma_sample[−1], i))

        if prob < u:
            sigma_sample.append(sigma_sample[−1])
        else:
            sigma_sample.append(propose)

        for j in range(100):
            theta_sample[j, i] = np.random.normal(loc = sigma_sample[−1]∗y[j]/(sigma_sample[−1] + 1), scale = np.sqrt(sigma_sample[−1]/(sigma_sample[−1] + 1)))

    else:
        for j in range(100):
```

```
            theta_sample[j, i] = np.random.normal(loc = sigma_sample[-1]*y[j]/(sigma_sample[-1] + 1), scale = np.
                sqrt(sigma_sample[-1]/(sigma_sample[-1] + 1)))


# plot sigma
plt.figure(figsize = (10, 8))
plt.title("sigma trace plot")
plt.plot(sigma_sample)
plt.xlim((0, 10000))
plt.show()

# plot first 3 theta
for i in range(3):
    plt.figure(figsize = (10, 8))
    plt.title("theta_1 trace plot")
    plt.plot(theta_sample[i, :])
    plt.xlim((0, 10000))
    plt.show()
```

Trace plot. σ^2 and θ_1.





2

2.1

$$
\begin{aligned}
p(bfy|\sigma^2) &= \int \cdots \int \Pi_{i=1}^n \left(\Pi_{j=1}^J \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(y_{ij}-\theta_i)^2}{2} \right) \right) \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\theta_i^2}{2\sigma^2} \right) \mathrm{d}\theta_1 \ldots \mathrm{d}\theta_n \\
&= \int \cdots \int \Pi_{i=1}^n \exp\left(-\frac{\sum_{j=1}^J y_{ij}^2}{2} + \frac{\sigma^2 (\sum_{j=1}^J y_{ij})^2}{2(\sigma^2+1)} \right) \frac{1}{\sqrt{2\pi}^J} \frac{1}{\sqrt{J\sigma^2+1}} \\
&\frac{\sqrt{J\sigma^2+1}}{\sqrt{2\pi}\sigma} \exp\left(-\frac{J\sigma^2+1}{2\sigma^2} \left(\theta_i - \frac{\sigma^2}{J\sigma^2+1} \sum_{j=1}^J y_{ij} \right)^2 \right) \mathrm{d}\theta_1 \ldots \mathrm{d}\theta_n \\
&= \Pi_{i=1}^n \frac{1}{\sqrt{2\pi}^J} \frac{1}{\sqrt{J\sigma^2+1}} \exp\left(-\frac{\sum_{j=1}^J y_{ij}^2}{2} + \frac{\sigma^2 (\sum_{j=1}^J y_{ij})^2}{2(\sigma^2+1)} \right) \\
&= \frac{1}{\sqrt{2\pi}^{nJ}} \frac{1}{\sqrt{J\sigma^2+1}^n} \exp\left(-\frac{\sum_{i=1}^n \sum_{j=1}^J y_{ij}^2}{2} + \frac{\sigma^2 \sum_{i=1}^n \left(\sum_{j=1}^J y_{ij} \right)^2}{2(J\sigma^2+1)} \right)
\end{aligned}
$$

2.2

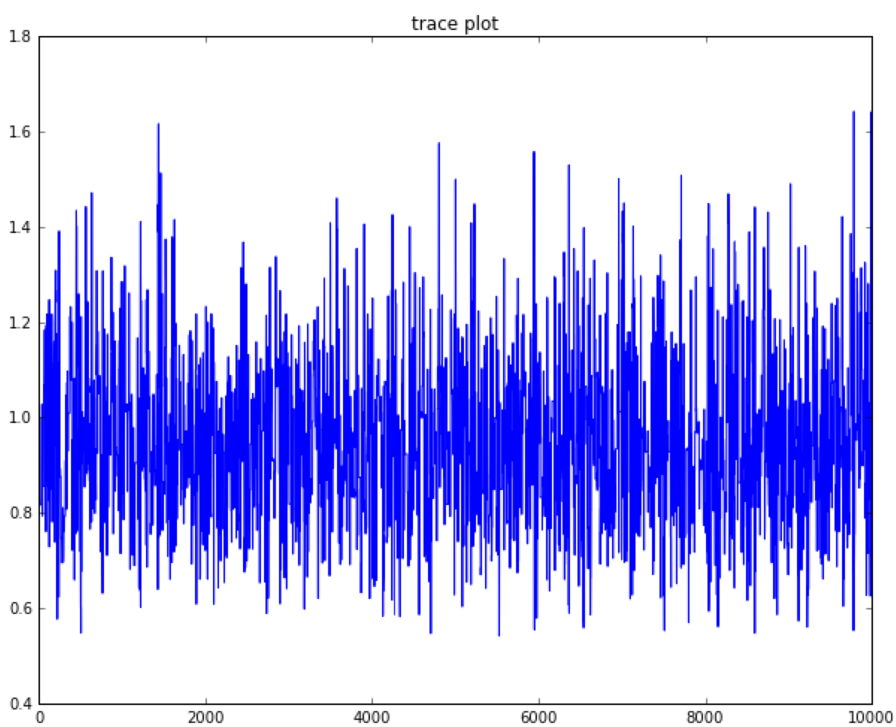
2.3

By problem 2-a, we can derive the below posterior distribution.

$$p(\sigma^2|\mathbf{y}) \propto \frac{1}{\sigma^2}\frac{1}{\sqrt{J\sigma^2+1}^n}\exp\left(\frac{\sigma^2\sum_{i=1}^n\left(\sum_{j=1}^J y_{ij}\right)^2}{2(J\sigma^2+1)}\right)$$

Listing 3: 2-c

```
q2 = pd.read_csv('q2.csv')
y = q2.values

summention = np.dot(np.reshape(y[:, 2], (100, 5)).sum(axis = 1), np.reshape(y[:, 2], (100, 5)).sum(axis = 1))
def pos3(sigma):
    return np.exp(summention*sigma/(2*(5*sigma + 1))) / (sigma*(sigma*5 + 1)**50)

np.random.seed(1234)
samples = [1]
for i in range(10000):
    u = np.random.uniform()
    propose = samples[-1] + np.random.normal(scale = 1)
    prob = min(1, pos3(propose)/pos3(samples[-1]))

    if prob < u:
        samples.append(samples[-1])
    else:
        samples.append(propose)

plt.figure(figsize = (10, 8))
plt.title("trace plot")
plt.plot(samples)
plt.show()
```



2.4

Use the previous notations. And by the same way we get the below posterior distributions.

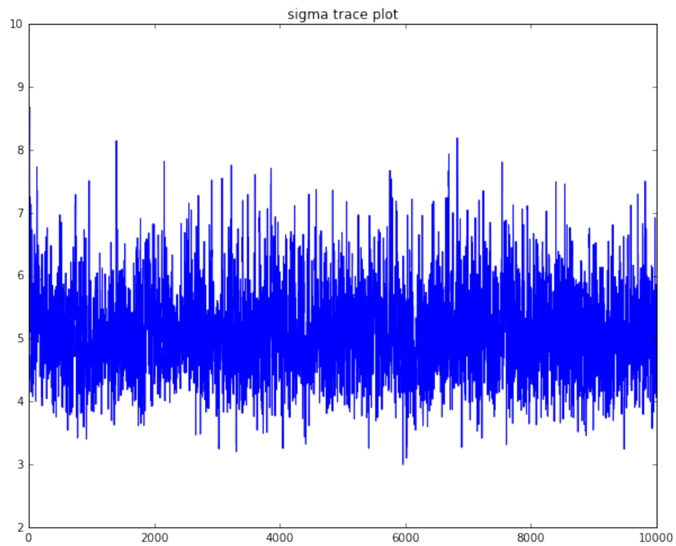
$$p(\sigma^2|\Theta, \mathbf{y}) \propto \frac{1}{\sigma^2\sqrt{2\pi\sigma^2}^n} \exp\left(-\frac{\sum_{i=1}^n \theta_i^2}{2\sigma^2}\right)$$

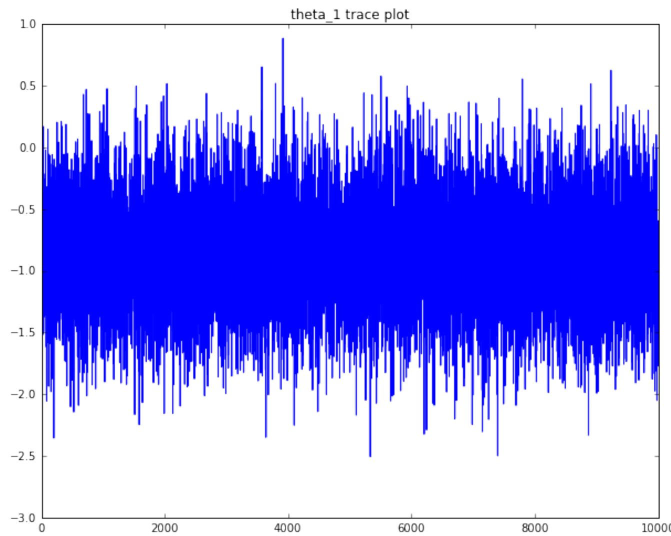
$$p(\theta_i|\Theta_{-i}, \mathbf{y}, \sigma^2) \propto N\left(\frac{\sigma^2 \sum_{j=1}^J y_{ij}}{1+J\sigma^2}, \frac{\sigma^2}{1+J\sigma^2}\right) \forall i \in \{1, 2, 3, \ldots, n\}$$

Listing 4: 2-d

```python
def pos4(sigma, i):
    return np.exp(-np.dot(theta_sample[:, i], theta_sample[:, i])/(2*sigma))/(sigma*(2*np.pi*sigma)**50)

np.random.seed(1234)

sigma_sample = [10]
theta_sample = np.ones((len(y), 10000))

for i in range(10000):

    if i != 0:
        u = np.random.uniform()
        propose = sigma_sample[-1] + np.random.normal(scale = 1)
        prob = min(1, pos4(propose, i)/pos4(sigma_sample[-1], i))

        if prob < u:
            sigma_sample.append(sigma_sample[-1])
        else:
            sigma_sample.append(propose)

        for j in range(100):
            theta_sample[j, i] = np.random.normal(loc = sigma_sample[-1]*sum(y[j*5:j*5+5, 2])/(5*sigma_sample[-1]
                + 1), scale = np.sqrt(sigma_sample[-1]/(5*sigma_sample[-1] + 1)))

    else:
        for j in range(100):
            theta_sample[j, i] = np.random.normal(loc = sigma_sample[-1]*sum(y[j*5:j*5+5, 2])/(5*sigma_sample[-1]
                + 1), scale = np.sqrt(sigma_sample[-1]/(5*sigma_sample[-1] + 1)))

# plot sigma
plt.figure(figsize = (10, 8))
plt.title("sigma trace plot")
plt.plot(sigma_sample)
plt.xlim((0, 10000))
plt.show()


# plot first 3 thetas
for i in range(3):
    plt.figure(figsize = (10, 8))
    plt.title("theta_1 trace plot")
    plt.plot(theta_sample[i, :])
    plt.xlim((0, 10000))
    plt.show()
```

Trace plot. σ^2 and θ_1.





3

3.1

Let $\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$ and $\hat{\epsilon} = \mathbf{y} - \hat{\mathbf{y}}$.

$$\begin{aligned}
p(\beta, \sigma^2|\mathbf{X}, \mathbf{y}) &\propto \frac{1}{\sigma^2} p(\mathbf{y}|\mathbf{X}, \beta, \sigma^2) \\
&\propto (\sigma^2)^{-(\frac{n}{2}+1)} \exp\left(\frac{1}{2\sigma^2}(\mathbf{y} - \mathbf{X}\beta)^T(\mathbf{y} - \mathbf{X}\beta)\right) \\
&= (\sigma^2)^{-(\frac{n}{2}+1)} \exp\left(\frac{1}{2\sigma^2}(\hat{\epsilon}^T\hat{\epsilon} + (\beta - \hat{\beta})^T(\mathbf{X}^T\mathbf{X})(\beta - \hat{\beta}))\right)
\end{aligned}$$

3.2

Let $\theta = \frac{2}{\hat{\epsilon}^T\hat{\epsilon} + (\beta - \hat{\beta})^T(\mathbf{X}^T\mathbf{X})(\beta - \hat{\beta}))}$

$$\begin{aligned}
p(\beta|\mathbf{X}, \mathbf{y}) &= \int \left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}+1} \exp\left(\frac{1}{\sigma^2\theta}\right) \mathrm{d}\sigma^2 = \int |\sigma^{-4}| \left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}+1} \exp\left(\frac{1}{\sigma^2\theta}\right) \mathrm{d}\sigma^{-2} \\
&= \int \left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}-1} \exp\left(\frac{1}{\sigma^2\theta}\right) \mathrm{d}\sigma^{-2} = \Gamma\left(\frac{n}{2}\right)\theta^{\frac{n}{2}} \int \frac{1}{\Gamma(\frac{n}{2})}\left(\frac{1}{\theta}\right)^{\frac{n}{2}} \left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}-1} \exp\left(\frac{1}{\sigma^2\theta}\right) \mathrm{d}\sigma^{-2} \\
&= \Gamma\left(\frac{n}{2}\right)\left(\frac{1}{\theta}\right)^{-\frac{n}{2}} \propto \left(\frac{1}{\theta}\right)^{-\frac{n}{2}} \\
&= \left(\frac{\hat{\epsilon}^T\hat{\epsilon}}{2} \frac{\hat{\epsilon}^T\hat{\epsilon} + (\beta - \hat{\beta})^T(\mathbf{X}^T\mathbf{X})(\beta - \hat{\beta}))}{\hat{\epsilon}^T\hat{\epsilon}}\right)^{-\frac{n}{2}} \propto \left(1 + \frac{(\beta - \hat{\beta})^T(\mathbf{X}^T\mathbf{X})(\beta - \hat{\beta}))}{\hat{\epsilon}^T\hat{\epsilon}}\right)^{-\frac{n}{2}}
\end{aligned}$$

This means the posterior distribution of β is $t_{n-k}(\hat{\beta}, S_\epsilon^2(\mathbf{X}^T\mathbf{X})^{-1})$, where $S_\epsilon^2 = \frac{\hat{\epsilon}^T\hat{\epsilon}}{n-k}$.

Next we calculate the posterior distribution of σ^2.

$$
\begin{aligned}
p(\sigma^2|\mathbf{X},\mathbf{y}) &= \int (\sigma^2)^{-(\frac{n}{2}+1)} \exp\left(-\frac{\hat{\epsilon}^T\hat{\epsilon} + (\beta-\hat{\beta})^T(\mathbf{X}^T\mathbf{X})(\beta-\hat{\beta})}{2\sigma^2}\right) \mathrm{d}\beta \\
&= (2\pi)^{\frac{n}{2}} \left| \left(\frac{\mathbf{X}^T\mathbf{X}}{\sigma^2}\right)^{-1} \right| (\sigma^2)^{-(\frac{n}{2}+1)} \exp\left(-\frac{\hat{\epsilon}^T\hat{\epsilon}}{2\sigma^2}\right) \\
&\int \frac{1}{(2\pi)^{\frac{n}{2}} \left| \left(\frac{\mathbf{X}^T\mathbf{X}}{\sigma^2}\right)^{-1} \right|} \exp\left(-\frac{1}{2}\left((\beta-\hat{\beta})^T((\frac{\mathbf{X}^T\mathbf{X}}{\sigma^2})^{-1})^{-1}(\beta-\hat{\beta})\right)\right) \mathrm{d}\beta \\
&= \frac{(2\pi)^{\frac{n}{2}}}{|\mathbf{X}^T\mathbf{X}|} (\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{\hat{\epsilon}^T\hat{\epsilon}}{2\sigma^2}\right) \\
&\propto (\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{\hat{\epsilon}^T\hat{\epsilon}}{2\sigma^2}\right)
\end{aligned}
$$

The last term tells us that the posterior distribution of σ^2 is inverse gamma distribution whose parameters are $(\frac{n}{2}, \frac{\hat{\epsilon}^T\hat{\epsilon}}{2})$. Thus, the posterior mean of β is $\hat{\beta}$ and one of σ^2 is $\frac{\frac{\hat{\epsilon}^T\hat{\epsilon}}{2}}{\frac{n}{2}-1} = \frac{\hat{\epsilon}^T\hat{\epsilon}}{n-2}$.