STA244 **1/29/2009**

Homework 4

Due 2/24/2009

For the normal linear model

$$\mathbf{Y} \sim \mathsf{N}(\mathbf{X}\boldsymbol{\beta}, \mathbf{I}_n/\phi) \tag{1}$$

(with $\mathbf{X}$ $n \times p$ of rank p), consider the conjugate Normal-Gamma prior with

$$\boldsymbol{\beta} \mid \phi \sim \mathsf{N}(\boldsymbol{\beta}_0, (\phi\Phi)^{-1}) \tag{2}$$

$$\phi \sim \mathsf{Gamma}(\nu_0/2, SS_0/2) \tag{3}$$

where Φ is a symmetric positive definite matrix of rank p. The conditional posterior for $\boldsymbol{\beta}$ is

$$\boldsymbol{\beta} \mid \phi, \mathbf{Y} \sim \mathsf{N}(\tilde{\boldsymbol{\beta}}, (\phi(\mathbf{X}^T\mathbf{X} + \Phi))^{-1})$$

where $\tilde{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{X} + \Phi)^{-1}(\mathbf{X}^T\mathbf{X}\hat{\boldsymbol{\beta}} + \Phi\boldsymbol{\beta}_0)$ and $\hat{\boldsymbol{\beta}}$ is the MLE. The posterior for ϕ is Gamma with shape parameter $(n + \nu_0)/2$.

1. After completing the square, we showed that the rate parameter for the Gamma distribution for ϕ was

   $$\left(\|\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}\|^2 + SS_0 + \boldsymbol{\beta}_0^T\Phi\boldsymbol{\beta}_0 + \hat{\boldsymbol{\beta}}^T(\mathbf{X}^T\mathbf{X})\hat{\boldsymbol{\beta}} - \tilde{\boldsymbol{\beta}}^T(\mathbf{X}^T\mathbf{X} + \Phi)\tilde{\boldsymbol{\beta}}\right)/2$$

   and then using the projection approach with the prior treated as "prior data" as in Christensen, we showed that the rate parameter was

   $$\left(\|\mathbf{Y} - \mathbf{X}\tilde{\boldsymbol{\beta}}\|^2 + SS_0 + (\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)^T\Phi(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)\right)/2.$$

   Show that the two rate parameters are equal.

2. One version of what is called Zellner's g-prior is based on taking $\boldsymbol{\beta}_0 = 0$, and $\Phi = \mathbf{X}^T\mathbf{X}/g$ for a fixed g and using a reference prior distribution for ϕ, $p(\phi) \propto 1/\phi$. Can this be expressed as a limiting case of the conjugate prior distribution in (2) and (3)? If so. give the limiting values of ν_0 and SS_0. Find the posterior distributions of $\boldsymbol{\beta} \mid \phi$ and ϕ. Show that the posterior distribution of $\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}$ given ϕ may be expressed in terms of the projection matrix $P_{\mathbf{X}} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}$.

3. For $\hat{\boldsymbol{\mu}} = \mathbf{X}\hat{\boldsymbol{\beta}}$ and $\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}$, find the expected loss $\mathsf{E}[\|\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}\|^2$ where the expectation is taken with respect to the distribution of the data $\mathbf{Y}$ given in (1). (Hint: re-expresess such that you may use the results about the expectation of a quadratic form.

4. Define $\tilde{\boldsymbol{\mu}} = \mathbf{X}\tilde{\boldsymbol{\beta}}$ where $\tilde{\boldsymbol{\beta}}$ is the posterior mean under the Zellner g-prior above. Find the sampling distribution of $\tilde{\boldsymbol{\mu}}$. (As a function of $\mathbf{Y}$, what is the distribution of $\tilde{\boldsymbol{\mu}}$ under the model in (1)).

5. Is the posterior mean $\tilde{\boldsymbol{\mu}}$ unbiased for estimating $\boldsymbol{\mu}$? If not, what is the bias?

6. Find $\mathsf{E}[\|\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}\|^2$ assuming model (1) and express as a function of p, g and $\|\mu\|^2$.

7. The Gauss-Markov Theorem showed that out of the class of unbiased linear estimators, the MLE has the smallest loss. If we use the posterior mean above, can it have a smaller loss than the MLE? Explain.