

## HW3 Due: April 22th

### 1 Bayesian perspective of Lasso regression

Recall that the Lasso regression can be formulated as the following optimization problem:

$$\hat{\beta}^{lasso} = \arg \min_{\beta \in \mathbb{R}^p} \left\{ \sum_{i=1}^n (y_i - \beta_0 - \mathbf{x}_i^T \beta)^2 + \lambda \sum_{j=1}^p |\beta_j| \right\}$$

where  $\lambda > 0$  and  $\beta = (\beta_1, \dots, \beta_p)^T$ . Now consider its Bayesian counterpart where we have

$$y_i \stackrel{iid}{\sim} \mathcal{N}(\beta_0 + \mathbf{x}_i^T \beta, \sigma^2), \quad i = 1, \dots, n$$
$$\beta_j \stackrel{iid}{\sim} \text{Laplace}(0, \tau^2), \quad j = 1, \dots, p$$

1. Write out the log density of the following terms:

- $\log p(y_i | \beta, \sigma^2, \tau^2)$
- $\log p(\beta_j | \tau^2)$

2. Show that the posterior density  $p(\beta | y, \sigma^2, \tau^2) \propto \prod_{i=1}^n p(y_i | \beta, \sigma^2, \tau^2) \prod_{j=1}^p p(\beta_j | \tau^2)$

3. Show that the maximum a posterior (MAP) estimate  $\beta^{MAP}$

$$\hat{\beta}^{MAP} = \arg \max_{\beta \in \mathbb{R}^p} \{p(\beta | y, \sigma^2, \tau^2)\}$$

is the same as the Lasso estimate for any fixed  $\sigma^2 > 0$  and  $\tau^2 > 0$  when  $\lambda = \frac{2\sigma^2}{\tau^2}$ .

### 2 Standard error estimate of ANOVA

Consider the one way ANOVA model

$$Y_{jk} = \mu + \alpha_j + \epsilon_{jk}$$

with the sum-to-zero constraint  $\sum_{j=1}^m \alpha_j = 0$  and normal distributed error  $\epsilon_{jk} \stackrel{iid}{\sim} \mathcal{N}(0, \sigma^2)$ . And the linear model with dummy variables

$$Y_{jk} = \gamma_0 + \sum_{i=1}^{m-1} \gamma_i D_{ijk} + \epsilon_{jk}$$

where  $D_{ijk}$  is the dummy variable such that  $D_{ijk} = 1$  if and only if  $i = j$ . We refer to the latter model as the dummy-variable model.

1. Show that  $\hat{\gamma}_i + \hat{\gamma}_0 = \hat{\mu} + \hat{\alpha}_i$  for  $i = 1, \dots, m-1$  and  $\hat{\gamma}_0 = \hat{\mu} + \hat{\alpha}_m$
2. Show that  $\widehat{SE}(\hat{\mu}_j)$  can be constructed by using the variance-covariance matrix  $(\widehat{\Sigma})$  estimated from the dummy-variable model.

### 3 JF exercise 8.10

Testing contrasts using group means: Suppose that we wish to test a hypothesis concerning a contrast of group means in a one-way ANOVA:

$$H_0 : c_1\mu_1 + c_2\mu_2 + \cdots + c_m\mu_m = 0$$

where  $c_1 + c_2 + \cdots + c_m = 0$ . Define the sample value of the contrast as

$$C \equiv c_1\bar{Y}_1 + c_2\bar{Y}_2 + \cdots + c_m\bar{Y}_m$$

and let

$$C'^2 \equiv \frac{C^2}{\frac{c_1^2}{n_1} + \frac{c_2^2}{n_2} + \cdots + \frac{c_m^2}{n_m}}$$

$C'^2$  is the sum of squares for the contrast.

Show that under the null hypothesis

1.  $E(C) = 0$
2.  $Var(C) = \sigma^2\left(\frac{c_1^2}{n_1} + \frac{c_2^2}{n_2} + \cdots + \frac{c_m^2}{n_m}\right)$
3.  $t_0 = C'/S_E$  follows a  $t$ -distribution with  $n - m$  degrees of freedom. [Hint: The  $\bar{Y}_j$  are independent, and each is distributed as  $N(\mu_j, \sigma^2/n_j)$ ]