LINEAR REGRESSION MODELS W4315

Midterm Examination QUESTIONS

November 9, 2010

Instructor: Frank Wood (10:35-11:50)

1. (30 points) The data below give weight X (kg) and height Y (cm) of 5 teenagers.

X (kg)	Y (cm)
50	160
60	160
70	170
80	170
90	180

1. (10 pts) Draw a scatter plot of height Y versus weight X.
2. (10 pts) Fit a simple linear regression $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$ by finding the least squares estimators b_0 and b_1.
3. (10 pts) In what sense are your estimators optimal? Please specify the minimum assumptions required for the optimality to hold.

2. (30 points) Consider the same data set and the simple linear regression model specification as given in the preceding problem. Suppose, in addition, ϵ_i are independent, normally distributed with mean 0 and variance σ^2.

1. (5 pts) Find an unbiased estimator of σ^2.
2. (5 pts) Find a 95% confidence interval for β_1.
3. (5 pts) Test one-sided hypothesis $H_0 : \beta_1 \leq 0$ versus $H_a : \beta_1 > 0$ (level of significance $\alpha = 0.05$)
4. (5 pts) Find a 95% Bonferroni Joint Confidence Interval of β_0 and β_1.

Source of Variation	SS	df	MS
Regression			
Error			
Total			

5. (10 pts) Produce an ANOVA table and perform an F-test for $H_0 : \beta_1 = 0$ versus $H_a : \beta_1 \neq 0$. (level of significance: $\alpha = 0.05$).

3. (30 points) Consider the classical regression setup

$$\mathbf{y} = \mathbf{X}\beta + \epsilon$$

where $\epsilon \sim \mathbf{N}(\mathbf{0}, \sigma^{\mathbf{2}}\mathbf{I})$. We know that the least square estimator for the parameter β minimizes the residual sum of squares, which in matrix terms can be written $(\mathbf{y} - \mathbf{X}\beta)'(\mathbf{y} - \mathbf{X}\beta)$. The value of β which minimizes this expression has the following analytic form

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{X}'\mathbf{y}. \tag{1}$$

Suppose, though, that $\mathbf{X}'\mathbf{X}$ is not invertible. In this case, this estimator can't be used. To get around this problem we define a penalized residual sum of squares (this is called "ridge regression")

$$(\mathbf{y} - \mathbf{X}\beta)'(\mathbf{y} - \mathbf{X}\beta) + \lambda\beta'\beta. \tag{2}$$

(a) (20 pts) Derive the ridge regression estimate $\hat{\beta}^{ridge}$ in matrix form (show your work), this is equivalent to finding the β which minimizes (2).

(b) (10 pts) In no more than 2 sentences, explain why the resulting estimator will work even when $\mathbf{X}'\mathbf{X}$ is singular.

4. (40 points) You, a statistical analysis consultant, are asked to analyze a regression problem. Food scientists invent a new secret ingredient to increase the fluffiness of pancakes (fluffiness is a made up concept but for our purposes will be defined as measurable scalar quantity). They know that the concentration of the new ingredient affects the fluffiness, and they want to figure out how large the effect is. We assume that a linear regression model is appropriate and the error comes from measurement of fluffiness. Two factories experimentally produce the pancakes with varying concentrations of the new ingredient. It could have been a simple linear regression problem, but the difficulty here is that the variances of fluffiness measurement of the pancakes produced by the two factories are DIFFERENT

(that is why you are called). In other words, the fluffiness measurement errors were iid only for each scientist individually. Denote fluffiness as $y = [y_1^1, y_1^2, \ldots, y_1^{n_1}, y_2^1, y_2^2, \ldots, y_2^{n_2}]$ and concentration level as $x = [x_1^1, x_1^2, \ldots, x_1^{n_1}, x_2^1, x_2^2, \ldots, x_2^{n_2}]$, where subscripts denote factories.

(a) (20 pts) Using matrix notation to set up a normal regression problem and derive the maximum likelihood estimates for the regression coefficients $\vec{\beta} = [\beta_0, \beta_1]^T$ under the given assumptions. *Hint : here because the i.i.d. assumption is no longer valid, instead of only one parameter σ^2, we need to use two parameters σ_1^2 and σ_2^2 in our model.* You may find it easier, once the problem is expressed in matrix form, to simplify the likelihoods by expressing them in scalar form before seeking the ML estimates for some of the variables.

(b) (20 pts) Provide the maximum likelihood estimators for both σ_1^2 and σ_2^2.