LINEAR REGRESSION MODELS W4315

Midterm Examination QUESTIONS

March 9, 2010

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

1. (30 points) You are a statistical consultant sent in to help solve a difficult regression analysis problem. You have been asked to fit a simple linear regression model to the following data

$$\{X_1 = 1, Y_1 = 1\},\ \{X_2 = 1, Y_2 = 2\}$$

(a) Using matrix notation, set up the specified regression model. Show the design and response matrices explicitly.

(b) Using matrix rank arguments, prove that the regression problem as specified lacks a unique solution. Explain the intuition behind your answer in no more than two sentences.

(c) In one sentence each, explain two options for solving this problem that don't involve changing the number of parameters in the model .

2. (30 points) Consider the classic regression setup

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

where $\epsilon \sim \mathbf{N}(\mathbf{0}, \sigma^2\mathbf{I})$. 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

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

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")

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

Derive the ridge regression estimate $\hat{\beta}^{ridge}$ in matrix form (show your work), this is equivalent to finding the β which minimizes (2). In no more than 2 sentences, explain why the resulting estimator will work even when $\mathbf{X}'\mathbf{X}$ is singular.

3. (30 points) You are a statistical analysis consultant who has been asked to develop a regression model to analyze data collected in the following way: a team of two food scientists was tasked with developing a new super-fluffy pancake recipe (fluffiness is a made up concept but for our purposes will be defined as measurable scalar quantity denoted f) based on experimenting with including varying levels of a secret new ingredient (the concentration of which is denoted by i). Nearly everything went well, each scientist prepared different recipes with varying levels of the secret new ingredient (i_s^r where $s = 1$ indicates the first scientist and $s = 2$ indicates the second scientist and $r \in [1, \ldots, n_s]$ is an integer that indicates which of the n_s pancake recipes tried by scientist s the input quantity corresponds to. Each scientist measured the fluffiness using their own machine, producing fluffiness values f_s^r. Unfortunately the variance of the fluffiness measurement produced by the scientists two machines was different (this is the reason why you were called). Each scientist produced measurement whose errors were independently distributed but whose variance was dependent on the identify of the scientist (σ_s^2 is the fluffiness measurement variance for scientist s). In other words, the fluffiness measurement errors were iid only for each scientist individually.

(a) Using matrix notation ($\vec{f} = [f_1^1, \ldots, f_1^{n_1}, f_2^1, \ldots, f_2^{n_2}]^T$, etc.) 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 : use a vector like* $\vec{\sigma}^2 = [\sigma_1^2, \ldots, \sigma_1^2, \sigma_2^2, \ldots, \sigma_2^2]^T$ *where* σ_s^2 *is repeated* n_s *times.* 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) Provide the maximum likelihood estimators for both σ_1^2 and σ_2^2.

4. (30 points) Let X_1, X_2, and X_3 be independent and $N(0,1)$-distributed. Set $Y = 8X_1^2 + 5X_2^2 + 5X_3^2 + 4X_1X_2 - 4X_1X_3 + 8X_2X_3$. Show that $\frac{Y}{9}$ is χ^2-distributed, and determine the number of degrees of freedom.

5. (10 points) Describe how you would like the rest of the class to be structured and what you would like to have taught.