Department of Statistics

STATS 760: A Survey of Modern Applied Statistics

Assignment 1 2016

Question 1

Suppose we want to fit a linear model with three explanatory variables, two of which are factors (A and B say), and the third (X) is continuous. The model formula is

$$Y \sim A*B*X$$

Discuss how R would build up the model matrix in this case, illustrating your answer with some suitable artificial data. How are the coefficients interpreted? (assume the default contrasts are being used).

How would your answer change if instead of the default contrasts (i.e. using the contrast function `contr.treatment`) you used the "sum contrasts" (i.e. executed the code

```
options(contrasts=c("contr.sum", "contr.poly"))?
```

Question 2

Suppose we have 5 observations $y = (y_1,y_2,y_3,y_4,y_5)$. The first two have mean μ_1, and the last three have mean μ_2. Express the model in the form

$$E(y) = X\beta$$

(where X is a matrix with 5 rows and 2 columns and $\beta=(\beta_1, \beta_2)$ is a vector of two elements) in two different ways. Interpret the beta coefficients in terms of μ_1 and μ_2 for each way.

Email answers to me by Friday March 18.