

# 620-371: Linear Models

Semester 1 Exam, 2009

Tuesday, 9th June

1. [12 marks] Consider the matrices

$$A = \frac{1}{5} \begin{bmatrix} 3 & 4 \\ -4 & 3 \end{bmatrix}, B = \frac{1}{10} \begin{bmatrix} 1 & -3 \\ -3 & 9 \end{bmatrix}.$$

- (a) [2 marks] Show that  $A$  is orthogonal.

**Solution:**

$$\begin{aligned} A^T A &= \frac{1}{25} \begin{bmatrix} 3 & -4 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} 3 & 4 \\ -4 & 3 \end{bmatrix} \\ &= \frac{1}{25} \begin{bmatrix} 25 & 0 \\ 0 & 25 \end{bmatrix} = I. \end{aligned}$$

- (b) [1 mark] Find the trace of  $B$ .

**Solution:**  $tr(B) = \frac{1}{10}(1 + 9) = 1$ .

- (c) [2 marks] Show that  $B$  is idempotent.

**Solution:**

$$\begin{aligned} B^2 &= \frac{1}{100} \begin{bmatrix} 1 & -3 \\ -3 & 9 \end{bmatrix} \begin{bmatrix} 1 & -3 \\ -3 & 9 \end{bmatrix} \\ &= \frac{1}{100} \begin{bmatrix} 10 & -30 \\ -30 & 90 \end{bmatrix} \\ &= B. \end{aligned}$$

- (d) [2 marks] Using your answer from part 1b, find the rank of  $B$ .

**Solution:** Since  $B$  is also symmetric,  $r(B) = tr(B) = 1$ .

- (e) [3 marks] Suppose that  $X$  and  $Y$  are two symmetric matrices with a common diagonalizing matrix, i.e. there exists an orthogonal matrix  $P$  so that both  $P^T X P$  and  $P^T Y P$  are diagonal. Show from first principles that  $XY = YX$ . You may use the property that  $D_1 D_2 = D_2 D_1$  if  $D_1$  and  $D_2$  are diagonal matrices.

**Solution:**

$$\begin{aligned} XY &= P P^T X P P^T Y P P^T \\ &= P (P^T X P) (P^T Y P) P^T \\ &= P (P^T Y P) (P^T X P) P^T \\ &= YX. \end{aligned}$$

(f) [2 marks] Find  $\frac{\partial}{\partial \mathbf{y}}(\mathbf{y}^T A \mathbf{y})$ , where  $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ .

**Solution:**

$$\begin{aligned} \frac{\partial}{\partial \mathbf{y}}(\mathbf{y}^T A \mathbf{y}) &= A \mathbf{y} + A^T \mathbf{y} \\ &= \frac{1}{5} \begin{bmatrix} 6 & 0 \\ 0 & 6 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \\ &= \frac{6}{5} y_1 + \frac{6}{5} y_2. \end{aligned}$$

2. [14 marks] Let  $\mathbf{y}$  be a normal random vector with mean  $\boldsymbol{\mu} = \begin{bmatrix} 0 \\ -4 \\ 1 \end{bmatrix}$  and variance  $I$ . Let

$$A = \frac{1}{3} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}.$$

(a) [3 marks] Find  $E[\mathbf{y}^T A \mathbf{y}]$ .

**Solution:**

$$\begin{aligned} E[\mathbf{y}^T A \mathbf{y}] &= \text{tr}(A) + \boldsymbol{\mu}^T A \boldsymbol{\mu} \\ &= 2 + \frac{1}{3} \begin{bmatrix} 0 & -4 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ -4 \\ 1 \end{bmatrix} \\ &= 2 + \frac{1}{3} \begin{bmatrix} 0 & -4 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ -9 \\ 6 \end{bmatrix} \\ &= 2 + \frac{42}{3} = 16. \end{aligned}$$

(b) [5 marks] Find the distribution of  $\mathbf{y}^T A \mathbf{y}$ .

**Solution:** Since

$$\begin{aligned} A^2 &= \frac{1}{9} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \\ &= \frac{1}{9} \begin{bmatrix} 6 & -3 & -3 \\ -3 & 6 & -3 \\ -3 & -3 & 6 \end{bmatrix} \\ &= A, \end{aligned}$$

and  $r(A) = \text{tr}(A) = 2$ ,  $\mathbf{y}^T A \mathbf{y}$  has a noncentral  $\chi^2$  distribution with 2 degrees of freedom and noncentrality parameter  $\frac{1}{2} \boldsymbol{\mu}^T A \boldsymbol{\mu} = 7$ .

- (c) [4 marks] Let  $\mathbf{z}$  be a normal random vector that is independent of  $\mathbf{y}$ , with mean  $\begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}$  and variance  $I$ . Find the distribution of  $\mathbf{y}^T \mathbf{y} + \mathbf{z}^T \mathbf{z}$ .

**Solution:**  $\mathbf{y}^T \mathbf{y}$  and  $\mathbf{z}^T \mathbf{z}$  have noncentral  $\chi^2$  distributions with 3 degrees of freedom and noncentrality parameters

$$\lambda_1 = \frac{1}{2}(0 + 16 + 1) = \frac{17}{2}$$

and

$$\lambda_2 = \frac{1}{2}(4 + 0 + 4) = 4$$

respectively. Therefore  $\mathbf{y}^T \mathbf{y} + \mathbf{z}^T \mathbf{z}$  has a noncentral  $\chi^2$  distribution with 6 degrees of freedom and noncentrality parameter  $\frac{25}{2}$ .

- (d) [2 marks] Under what condition on  $B$  is  $\mathbf{y}^T A \mathbf{y}$  independent from  $B \mathbf{y}$ ?

**Solution:** We require  $BVA = \mathbf{0}$ .

3. [24 marks] We wish to fit a linear model to explain the amount of diesel fuel a truck uses, given the number of hours that it runs. A study is performed and the following data collected:

Amount of fuel (litres)	Time motor runs (hours)
5	0.6
10	2
15	2.1
20	2
22	2.4

We fit the linear model  $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ , where  $\mathbf{y}$  is the vector of response values,  $X$  is the design matrix,  $\boldsymbol{\varepsilon}$  is the vector of errors, and

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$$

is the parameter vector, with  $\beta_0$  an intercept term and  $\beta_1$  the parameter associated with the running time.

- (a) [3 marks] What are the common assumptions used to fit a linear model?

**Solution:** We assume that the errors  $\boldsymbol{\varepsilon}$  are jointly normally distributed with mean  $\mathbf{0}$  and variance  $\sigma^2 I$ .

- (b) [4 marks] Calculate the least squares estimators of the parameters.

**Solution:**

$$\mathbf{y} = \begin{bmatrix} 5 \\ 10 \\ 15 \\ 20 \\ 22 \end{bmatrix}, X = \begin{bmatrix} 1 & 0.6 \\ 1 & 2 \\ 1 & 2.1 \\ 1 & 2 \\ 1 & 2.4 \end{bmatrix}.$$

$$X^T X = \begin{bmatrix} 5 & 9.1 \\ 9.1 & 18.53 \end{bmatrix}, (X^T X)^{-1} = \begin{bmatrix} 1.88 & -0.92 \\ -0.92 & 0.51 \end{bmatrix}.$$

$$\mathbf{b} = (X^T X)^{-1} X^T \mathbf{y} = \begin{bmatrix} -0.64 \\ 8.26 \end{bmatrix}.$$

- (c) [2 marks] Calculate the fitted value and residual for the first sample.

**Solution:** The fitted value is  $[1 \ 0.6] \mathbf{b} = 4.32$ , and the residual is  $5 - 4.32 = 0.68$ .

- (d) [3 marks] Show that in a general linear model,  $SS_{Res} = SS_{Total} - \mathbf{y}^T X \mathbf{b}$ , where  $SS_{Res}$  is the residual sum of squares,  $SS_{Total}$  is the total sum of squares, and  $\mathbf{b}$  is the least squares estimator of the parameter vector.

**Solution:**

$$\begin{aligned} SS_{Res} &= \mathbf{y}^T [I - X(X^T X)^{-1} X^T] \mathbf{y} \\ &= \mathbf{y}^T \mathbf{y} - \mathbf{y}^T X \mathbf{b} \\ &= SS_{Total} - \mathbf{y}^T X \mathbf{b}. \end{aligned}$$

- (e) [2 marks] What is the smallest possible variance of a linear estimator of  $\beta_0$ ? You may express this in terms of  $\sigma^2$ , the variance of the errors.

**Solution:** The smallest possible variance is the variance of  $b_0$ , which is  $1.88\sigma^2$ .

- (f) [2 marks] According to the model, what is the average length of time a motor would take to consume 12 litres of fuel?

**Solution:** Let the length of time be  $c$ . Then  $-0.64 + 8.26c = 12$ , so  $c = 1.53$  hours.

- (g) [3 marks] Find a 95% confidence interval for  $\beta_1$ . You may take  $SS_{Res} = 62.86$  and the critical value for a  $t$  distribution with 3 degrees of freedom as  $t_{0.025} = 3.182$ .

**Solution:** We have  $s^2 = \frac{SS_{Res}}{5-2} = 20.95$  and  $c_{11} = 0.51$ , so the confidence interval is

$$8.26 \pm 3.182\sqrt{20.95 \times 0.51} = 8.26 \pm 10.38 = (-2.12, 18.65).$$

- (h) [3 marks] Explain the difference between a confidence interval and a prediction interval for a given set of predictor variables.

**Solution:** A confidence interval puts bounds on the mean value of the response variable given the predictor variables; a prediction interval puts bounds on the value of the response variable for a specific sample which has the given predictor variables.

- (i) [2 marks] Simultaneously taking 95% confidence intervals for all the parameters gives us a joint confidence region for the parameter vector. What is the confidence level of this region?

**Solution:** The probability that each parameter lies within its confidence interval is 0.95, so the probability that all parameters lie in the confidence region is  $0.95^p$ , where  $p$  is the number of parameters.

4. [14 marks] Consider the study in question 3.

- (a) [3 marks] Test for model adequacy,  $H_0 : \boldsymbol{\beta} = \mathbf{0}$ , at the 5% level. You may take the critical value of an  $F$  distribution with 2 and 3 degrees of freedom as  $f_{0.05} = 9.55$ .

**Solution:**  $\mathbf{y}^T \mathbf{y} = 1234$ , so  $SS_{Reg} = 1234 - 62.86 = 1171.14$ . The  $F$  statistic is

$$\frac{1171.14/2}{62.86/3} = 27.95.$$

Therefore we reject the hypothesis of model inadequacy.

- (b) [3 marks] Find  $R(\beta_0|\beta_1)$ , the regression sum of squares of  $\beta_0$  in the presence of  $\beta_1$ .

**Solution:**

$$\begin{aligned} R(\beta_1) &= \mathbf{y}^T X_1 (X_1^T X_1)^{-1} X_1^T \mathbf{y} = 1170.92 \\ R(\beta_0|\beta_1) &= 1234 - 1170.92 = 63.07. \end{aligned}$$

- (c) [3 marks] What is a difference between a partial test of  $H_0 : \beta_0 = 0$  and a sequential test of  $H_0 : \beta_0 = 0$ , assuming that the sequential test tests the parameters in the order  $\beta_0, \beta_1$ ?

**Solution:** The partial test tests  $\beta_0 = 0$  in the presence of the parameter  $\beta_1$  (in the reduced model). The sequential test tests  $\beta_0 = 0$  in the presence of no other parameters.

- (d) [2 marks] Express the hypothesis  $5\beta_0 = \beta_1 = 1$  as a general linear hypothesis in matrix form.

**Solution:**  $H_0 : \begin{bmatrix} 5 & -1 \\ 0 & 1 \end{bmatrix} \boldsymbol{\beta} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ .

- (e) [3 marks] Give an advantage of mutually orthogonal design.

**Solution:** In a mutually orthogonal design, the regression sum of squares attributed to each parameter does not change in the presence or absence of other parameters. This makes partial tests both easier to calculate and more meaningful in the context of variable selection, as we can test if a variable should be selected regardless of what variables are already in the model.

5. [24 marks] A study of two major coal seams in a region is conducted, to compare the sulfur content of coal drawn from each seam. The following data is collected (in terms of percentage of sulfur):

Seam	
1	2
1.51	1.69
1.92	0.64
1.08	0.9

The linear model that we use is

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}, i = 1, 2, j = 1, 2, 3,$$

where  $\tau_i$  is the effect on the sulfur content associated with seam  $i$ , and  $\varepsilon_{ij}$  is the error for the  $j$ th sample from seam  $i$ .

- (a) [2 marks] Write down the linear model in a matrix form.

**Solution:**  $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ , where

$$\mathbf{y} = \begin{bmatrix} 1.51 \\ 1.92 \\ 1.08 \\ 1.69 \\ 0.64 \\ 0.9 \end{bmatrix}, X = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \end{bmatrix}, \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{13} \\ \varepsilon_{21} \\ \varepsilon_{22} \\ \varepsilon_{23} \end{bmatrix}.$$

- (b) [4 marks] Find two different conditional inverses for  $X^T X$ , where  $X$  is the design matrix.

**Solution:**

$$X^T X = \begin{bmatrix} 6 & 3 & 3 \\ 3 & 3 & 0 \\ 3 & 0 & 3 \end{bmatrix}.$$

Two different conditional inverses for  $X^T X$  are

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix} \text{ and } \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ -\frac{1}{3} & \frac{1}{3} & 0 \end{bmatrix}.$$

- (c) [3 marks] Under what conditions does a matrix have 0, 1, or an infinite number of conditional inverses?

**Solution:** A matrix never has 0 conditional inverses. It has 1 conditional inverse if it is square and nonsingular. Otherwise it has an infinite number of conditional inverses.

- (d) [2 marks] What is the best linear unbiased estimator for  $\mu$ ?

**Solution:** There is no best linear unbiased estimator for  $\mu$ , as it is not estimable.

- (e) [2 marks] Is  $2\mu + \tau_1 + \tau_2$  estimable? Why or why not?

**Solution:** Yes. Elements of  $X\boldsymbol{\beta}$  are estimable, and  $\mu + \tau_1$  and  $\mu + \tau_2$  are elements of  $X\boldsymbol{\beta}$ . Sums of estimable quantities are estimable, so  $2\mu + \tau_1 + \tau_2$  is estimable.

- (f) [3 marks] Estimate  $\tau_1 - \tau_2$ .

**Solution:** A solution of the normal equations is

$$\begin{aligned}
 \mathbf{b} &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1.51 \\ 1.92 \\ 1.08 \\ 1.69 \\ 0.64 \\ 0.9 \end{bmatrix} \\
 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 7.74 \\ 4.51 \\ 3.23 \end{bmatrix} \\
 &= \begin{bmatrix} 0 \\ 1.5 \\ 1.08 \end{bmatrix}.
 \end{aligned}$$

An estimator of  $\tau_1 - \tau_2$  is  $1.5 - 1.08 = 0.43$ .

(g) [4 marks] Calculate  $s^2$ , the estimator for the variance of the errors.

**Solution:**

$$\begin{aligned}
 \mathbf{e} &= \mathbf{y} - X\mathbf{b} \\
 &= \begin{bmatrix} 1.51 \\ 1.92 \\ 1.08 \\ 1.69 \\ 0.64 \\ 0.9 \end{bmatrix} - \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1.5 \\ 1.08 \end{bmatrix} \\
 &= \begin{bmatrix} 0.0067 \\ 0.42 \\ -0.42 \\ 0.61 \\ -0.44 \\ -0.18 \end{bmatrix} \\
 SS_{Res} &= \mathbf{e}^T \mathbf{e} = 0.95 \\
 s^2 &= \frac{0.95}{6 - 2} = 0.238.
 \end{aligned}$$

(h) [4 marks] Find a 95% confidence interval for  $\tau_1 - \tau_2$ . You may take the critical value for a  $t$  distribution with 4 degrees of freedom as  $t_{0.025} = 2.78$ .

**Solution:** We have  $\mathbf{t}^T = [0 \ 1 \ -1]$ . So a confidence interval is

$$\begin{aligned} \mathbf{t}^T \mathbf{b} \pm t_{\alpha/2} s \sqrt{\mathbf{t}^T (X^T X)^c \mathbf{t}} &= 0.43 \pm 2.78 \times \sqrt{0.238 \times [0 \ 1 \ -1] \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}} \\ &= 0.43 \pm 2.78 \times \sqrt{0.238 \times \frac{2}{3}} \\ &= 0.43 \pm 1.1 = (-0.68, 1.53). \end{aligned}$$

6. [12 marks] We are interested in modelling the life span of paint in years, based on two different types of undercoating and three different types of paint formula. A study is conducted with one sample from each pair of factor levels and the following data is obtained:

		Formula (factor 1)		
		1	2	3
Undercoat (factor 2)	1	1.09	1.35	1.6
	2	1.16	1.38	2.18

- (a) [2 marks] Express this data as a two-factor linear model without interaction, in matrix form.

**Solution:**  $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ , where

$$\mathbf{y} = \begin{bmatrix} 1.09 \\ 1.16 \\ 1.35 \\ 1.38 \\ 1.6 \\ 2.18 \end{bmatrix}, X = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \\ \beta_1 \\ \beta_2 \end{bmatrix}, \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{21} \\ \varepsilon_{22} \\ \varepsilon_{31} \\ \varepsilon_{32} \end{bmatrix}.$$

- (b) [2 marks] We wish to test the hypothesis that the formula used has no effect on the lifespan of the paint. How many degrees of freedom do we use in the  $F$  distribution that we are testing against?

**Solution:** We have 2 independent hypotheses,  $r(X) = 4$ , and 6 samples, so we test against an  $F$  distribution with 2 and 2 degrees of freedom.

- (c) [2 marks] Express the hypothesis that neither formula nor undercoat has an effect on paint life in matrix form.

**Solution:**  $H_0 : C\boldsymbol{\beta} = \mathbf{0}$ , where

$$C = \begin{bmatrix} 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}.$$

- (d) [2 marks] Is this hypothesis testable? Justify your answer.

**Solution:** Yes, the hypothesis is testable as it consists entirely of testing whether treatment contrasts are 0.

- (e) [2 marks] Write down the design matrix if we were to express this as a two-factor linear model with interaction.

**Solution:**

$$X = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

- (f) [2 marks] Suppose the true means are:

		Formula (factor 1)		
		1	2	3
Undercoat (factor 2)	1	1	1.3	1.5
	2	1.1	1.4	2

Is there interaction between the factors? Why or why not?

**Solution:** Yes, there is interaction among the factors. The difference between undercoat levels for the first paint formula is different to the difference between undercoat levels for the third paint formula.