14/14

SDS 386D HW 2

Bowei Yan, Sukyung Park, Michael Zhang, and Maurice Diesendruck

February 15, 2015

1 Problem B.11

Our plan is first constructing an orthonormal basis of the matrix of interest using Gram-Schmidt construction, and then applying Theorem B.35 to achieve the perpendicular projection matrix.

1.1 A

$$A = \begin{bmatrix} 2 & 0 & 4 \\ 1 & 5 & 7 \\ 1 & -5 & -3 \end{bmatrix} \tag{1}$$

The column vectors of A are

$$x_1 = (2,1,1)^T, \quad x_2 = (0,5,-5)^T, x_3 = (4,7,-3)^T$$

By Theorem A.12,

$$y_1 = x_1/\sqrt{x_1'x_1} = (2/\sqrt{6}, 1/\sqrt{6}, 1/\sqrt{6})^T$$
$$w_2 = x_2 - (x_2'y_1) * y_1 = (0,5,-5)^T$$
$$y_2 = w_2/\sqrt{w_2'w_2} = (0, 1/\sqrt{2}, -1/\sqrt{2})^T$$
$$w_3 = x_3 - (x_3'y_1) * y_1 - (x_3'y_2) * y_2 = (0,0,0)^T$$

So the column space of A only has two dimension and

$$O = \begin{bmatrix} 2/\sqrt{6} & 0 \\ 1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{6} & -1/\sqrt{2} \end{bmatrix} \tag{2}$$

By Theorem B.35,

$$P = OO' = \begin{bmatrix} 2/3 & 1/3 & 1/3 \\ 1/3 & 2/3 & -1/3 \\ 1/3 & -1/3 & 2/3 \end{bmatrix} \tag{3}$$

✓

1.2 B

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \tag{4}$$

The column vectors of B are

$$x_1 = (1,0,0)^T, \quad x_2 = (0,0,1)^T, x_3 = (0,1,0)^T$$

We can check that $[x_1, x_2, x_3]$ constructs an orthonormal basis of the column space of B. So the perpendicular projection matrix of B is

$$P = BB' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \checkmark \tag{5}$$

1.3 C

$$C = \begin{bmatrix} 1 & 4 & 1 \\ 2 & 5 & 1 \\ -3 & 0 & 1 \end{bmatrix} \tag{6}$$

The column vectors of C are

$$x_1 = (1,2,-3)^T, \quad x_2 = (4,5,0)^T, x_3 = (1,1,1)^T$$

By Theorem A.12,

$$y_1 = x_1/\sqrt{x_1'x_1} = (1/sqrt14, 2/sqrt14, -3/sqrt14)^T$$
$$w_2 = x_2 - (x_2'y_1) * y_1 = (1,1,1)^T$$
$$y_2 = w_2/\sqrt{w_2'w_2} = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})^T$$
$$w_3 = x_3 - (x_3'y_1) * y_1 - (x_3'y_2) * y_2 = (0,0,0)^T$$

So the column space of C only has two dimension and

$$O = \begin{bmatrix} 1/\sqrt{14} & 1/\sqrt{3} \\ 2/\sqrt{14} & 1/\sqrt{3} \\ -3/\sqrt{14} & 1/\sqrt{3} \end{bmatrix} \tag{7}$$

By Theorem B.35,

$$P = OO' = \frac{1}{42}\begin{bmatrix} 17 & 20 & 5 \\ 20 & \cancel{13}^{26} & -4 \\ 5 & -4 & 41 \end{bmatrix} \checkmark \tag{8}$$

2 Problem B.12

$\checkmark$

$$\sum_i \sum_j m_{ij}^2 = tr(MM) = tr(M) = r(M) \tag{9}$$

Exercise 2.1

i) $$\frac{\lambda'\hat{\beta}-\lambda'\beta}{\sqrt{MSE\,\lambda'(X'X)^-\lambda}} = \frac{\dfrac{\lambda'\hat{\beta}-\lambda'\beta}{\sqrt{\sigma^2\lambda'(X'X)^-\lambda}}}{\sqrt{\dfrac{Y'(I-M)Y}{n-r}\cdot\dfrac{1}{\sigma^2}}} = \frac{N(0,1)}{\sqrt{\chi^2(n-r)/n-r}} \sim t(dfE)$$ ✓

where $dfE = n-r$

ii) $H_0: \lambda'\beta = 0$, reject H_0 if

$$\frac{|\lambda'\hat{\beta}|}{\sqrt{MSE\lambda'(X'X)^-\lambda}} \geq t(1-\frac{\alpha}{2}, n-r)$$ ✓

iii) $(1-\alpha)100\%$ CI : $\lambda'\hat{\beta} \pm t(1-\frac{\alpha}{2}, n-r)\sqrt{MSE\,\lambda'(X'X)^-\lambda}$ ✓

Exercise 2.2

$$\begin{bmatrix} Y_{11} \\ Y_{12} \\ \vdots \\ Y_{1r} \\ Y_{21} \\ Y_{22} \\ \vdots \\ Y_{2s} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ \vdots & \vdots \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \begin{bmatrix} e_{11} \\ \vdots \\ \vdots \\ e_{2s} \end{bmatrix}, \quad e_{ij} \overset{i.i.d}{\sim} N(0,\sigma^2)$$ ✓

$$Y = X\mu + e$$

i) $\hat{\mu}_1 = \lambda_1\hat{\mu} = [1\ 0](X'X)^-X'Y = [1\ 0]\begin{bmatrix} \frac{1}{r} & 0 \\ 0 & \frac{1}{s} \end{bmatrix}\begin{bmatrix} \sum_{i=1}^{r} Y_{1i} \\ \sum_{i=1}^{s} Y_{2i} \end{bmatrix} = \frac{\sum_{i=1}^{r} Y_{1i}}{r} = \bar{Y}_1$ ✓

$\hat{\mu}_2 = \lambda_2\hat{\mu} = [0\ 1](X'X)^-X'Y = \frac{\sum_{i=1}^{s} Y_{2i}}{s} = \bar{Y}_2$ ✓

$\widehat{\mu_1-\mu_2} = \lambda_3\hat{\mu} = [1\ -1](X'X)^-X'Y = \bar{Y}_1 - \bar{Y}_2$ ✓

$MSE = \hat{\sigma}^2 = \frac{Y'(I-M)Y}{r+s-2}$, $M = X(X'X)^-X'$

$= \frac{\sum_{i=1}^{r}(Y_{1i}-\bar{Y}_1)^2 + \sum_{i=1}^{s}(Y_{2i}-\bar{Y}_2)^2}{r+s-2}$ ✓

✓

ii) $\alpha = 0.01$, $H_0: \mu_1-\mu_2 = 0$, reject H_0 if $\frac{|\bar{Y}_1-\bar{Y}_2|}{\sqrt{MSE(\frac{1}{r}+\frac{1}{s})}} \geq t(0.995, \underbrace{dfE}_{r+s-2})$

95% CI for $\mu_1-\mu_2$: $\bar{Y}_1-\bar{Y}_2 \pm t(0.975, r+s-2)\sqrt{MSE(\frac{1}{r}+\frac{1}{s})}$

" μ_1 : $\bar{Y}_1 \pm t(0.975, r+s-2)\sqrt{MSE/r}$

α, $H_0: \mu_1-\mu_2 = \Delta$, reject H_0 if $\frac{|\bar{Y}_1-\bar{Y}_2-\Delta|}{\sqrt{MSE(\frac{1}{r}+\frac{1}{s})}} \geq t(1-\frac{\alpha}{2}, r+s-2)$

iii) Same as the usual analysis for two independent samples with same variance ✓

Ex. 2.3. We define our linear model as $Y = \mu + \epsilon$ where $\epsilon \sim N(0, \sigma^2)$. Thus to form a hypothesis test about our true value of μ we develop a test statistic from our observations given a null hypothesis $H_0 : \mu_0$. Let this test statistic be

$$t_{\hat{\mu}} = \frac{\hat{\mu} - \mu_0}{\hat{\sigma}/\sqrt{n}}$$

which will be t-distributed because we do not know the population variance. In order to run a hypothesis test then we need to calculate the probability that $P(|t| \geq t_{\hat{\mu}} | H_0)$, this is our p-value. If our p-value is less than some cutoff α, which is .01 for this problem, then we reject the null hypothesis. For a confidence interval, we generate an estimate for μ with $\hat{\mu}$ and estimate the standard error for $\hat{\mu}$ with $\sqrt{\frac{\hat{\sigma}^2}{n}}$. Thus our confidence interval is

$$\hat{\mu} \pm t^{df}_{\alpha/2} \sqrt{\frac{\hat{\sigma}^2}{n}}$$

where $\alpha = .05$ in this problem and t is the t-value for $\alpha/2$ with some given degrees of freedom. These results are nearly identical to how we generally perform one-sample tests.

Exercise 2.10.1

(a) From the distributions of $\hat{\beta}$ and y_0, $x_0(\hat{\beta}-\beta)$ and $y_0 - x_0'\beta$ have the following distributions:

$$x_0(\hat{\beta}-\beta) \sim N(0, \sigma^2 x_0'(X'X)^- x_0)$$
$$y_0 - x_0'\beta \sim N(0, \sigma^2)$$

Therefore,

$$\frac{y_0 - x_0'\hat{\beta}}{\sqrt{MSE(1 + x_0'(X'X)^- x_0)}} = \frac{\frac{(y_0 - x_0'\beta) - x_0(\hat{\beta}-\beta)}{\sqrt{\sigma^2(1+x_0'(X'X)^- x_0)}}}{\sqrt{\frac{Y'(I-M)Y}{\sigma^2}/(n-r)}} = \frac{N(0,1)}{\sqrt{\chi^2(n-r)/n-r}} \sim t(n-r)$$

(b) According to the result of (a), we can construct a 95% prediction interval as follows:

$$\begin{aligned}
0.95 &= P\left(-t(0.975, n-r) < \frac{y_0 - x_0'\hat{\beta}}{\sqrt{MSE(1+x_0'(X'X)^- x_0)}} < t(0.975, n-r)\right) \\
&= P\left(-t(0.975, n-r)\sqrt{MSE(1+x_0'(X'X)^- x_0)} \le y_0 - x_0'\hat{\beta} \le t(0.975, n-r)\sqrt{MSE(1+x_0'(X'X)^- x_0)}\right) \\
&= P\left(x_0'\hat{\beta} - t(0.975, n-r)\sqrt{MSE(1+x_0'(X'X)^- x_0)} \le y_0 \le x_0'\hat{\beta} + t(0.975, n-r)\sqrt{MSE(1+x_0'(X'X)^- x_0)}\right)
\end{aligned}$$

Therefore, 95% prediction interval for y_0 is:

$$x_0'\hat{\beta} \pm t(0.975, n-r)\sqrt{MSE(1+x_0'(X'X)^- x_0)}$$

(c) $x_0'\hat{\beta} - \gamma(\eta)$ and $(x_0'\hat{\beta} - \gamma(\eta))/\sigma$ have the following distributions:

$$x_0'\hat{\beta} - \gamma(\eta) = x_0'(\hat{\beta}-\beta) - z(\eta)\sigma \sim N(-z(\eta)\sigma, \sigma^2 x_0'(X'X)^- x_0)$$
$$\frac{x_0'\hat{\beta} - \gamma(\eta)}{\sigma} = \frac{x_0'(\hat{\beta}-\beta)}{\sigma} - z(\eta) \sim N(-z(\eta), x_0'(X'X)^- x_0)$$

and thus,

$$\frac{x_0'\hat{\beta} - \gamma(\eta)}{\sqrt{MSE x_0'(X'X)^- x_0}} = \frac{\left(\frac{x_0'(\hat{\beta}-\beta)}{\sigma} - z(\eta)\right) / \sqrt{x_0'(X'X)^- x_0}}{\sqrt{\frac{Y'(I-M)Y}{\sigma^2}/(n-r)}} \sim t(n-r, -z(\eta)),$$

Therfore, we can find a $(1-\alpha)100\%$ lower confidence bound as follows:

$$x_0'\hat{\beta} - t\left(1 - \frac{\alpha}{2}, n-r, -z(\eta)\right)\sqrt{MSE x_0'(X'X)^- x_0}$$

Exercise 2.10.4

(a) Estimate β_1, β_2, and σ^2.

Given $X'X$, $X'Y$, and $Y'Y$ defined as follows:

$$X'X = \begin{pmatrix} 15 & 374.5 \\ 374.5 & 9482.75 \end{pmatrix}$$

$$X'Y = \begin{pmatrix} 6.03 \\ 158.25 \end{pmatrix}$$

$$Y'Y = 3.03$$

Normal equations allow estimation of $\hat{\beta}$ and $\hat{\sigma^2}$, using the following equations:

$$\hat{\beta} = \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = (X'X)^{-1}X'Y = \begin{pmatrix} -1.04653064 \\ 0.05801858 \end{pmatrix}$$

$$\hat{\sigma^2} = MSE = \frac{Y'(1-M)Y}{n-r} = 0.01224145$$

Where $n = 15$ and $r = R(X) = 2$.

(b) Give 98% confidence intervals for β_2 and $\beta_2 - \beta_1$.

If errors are uncorrelated and homoscedastic, β has distribution $\hat{\beta} \sim N(\beta, \sigma^2(X_1'X_1)^{-1})$, and the partitioned form (for β_1, β_2) borrows from the following result in Exercise 2.1, with λ_2 to select for $\hat{\beta_2}$, and λ_{dif} to select for $\hat{\beta_2} - \hat{\beta_2}$.

$$\frac{\lambda'\hat{\beta} - \lambda'\beta}{\sqrt{MSE\lambda'(X'X)^{-}\lambda}} \sim t(dfE)$$

$$\lambda_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\lambda_{dif} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

For the given n and r, the 98% confidence interval for a t-distribution uses the two-tailed t-score of 2.65, and the interval is generally defined as the following, for each λ:

$$\lambda'\beta = \lambda'\hat{\beta} \pm \sqrt{MSE\lambda'(X'X)^{-}\lambda}(2.65)$$

The confidence intervals for β_2 and $\beta_2 - \beta_1$ are therefore as follows:

$$CI_{\beta_2} = [0.03256949, 0.08346768]$$
$$CI_{dif} = [0.4393989, 1.7696995]$$

(c) Perform an $\alpha = 0.05$ test for $H_0 : \beta_1 = 0.5$.

Using a similar form as in part (b), with $\lambda'\hat{\beta}$ unknown, with $\lambda'\beta = 0.5$, and with λ_1 to isolate β_1, the t-score is -6.404873. This is far below the 95% t-score that marks the 95% cutoff for a two-tailed t-distribution ($t = \pm 2.16$), so the null hypothesis is rejected.