1. (a) In general, we have $H_{0j} : (\boldsymbol{P}_{j+1} - \boldsymbol{P}_j)\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{0}$ in ANOVA F test, so $\boldsymbol{c}_j$ is any non-zero row of $(\boldsymbol{P}_{j+1} - \boldsymbol{P}_j)\boldsymbol{X}$.

   So, we can obtain using `R` code below as we did on slides 45 and 46 of slide set 6.

$$\boldsymbol{c}_1 = \begin{pmatrix}2, & 1, & 0, & -1, & -2\end{pmatrix}' \quad \text{for linear trend}$$
$$\boldsymbol{c}_2 = \begin{pmatrix}2, & -1, & -2, & -1, & 2\end{pmatrix}' \quad \text{for quadratic trend}$$
$$\boldsymbol{c}_3 = \begin{pmatrix}1, & -2, & 0, & 2, & -1\end{pmatrix}' \quad \text{for cubic trend}$$
$$\boldsymbol{c}_4 = \begin{pmatrix}1, & -4, & 6, & -4, & 1\end{pmatrix}' \quad \text{for quartic trend}$$

```
> d=read.delim("https://dnett.github.io/S510/PlantDensity.txt")
> names(d)=c("x","y")
> n=nrow(d)
> x=(d$x-mean(d$x))/10
> x1=matrix(1,nrow=n,ncol=1)
> x2=cbind(x1,x)
> x3=cbind(x2,x^2)
> x4=cbind(x3,x^3)
> x5=matrix(model.matrix(~0+factor(x)),nrow=n)
> proj <- function(x) {
+   x %*% MASS::ginv(t(x) %*% x) %*% t(x)
+ }
> p1=proj(x1)
> p2=proj(x2)
> p3=proj(x3)
> p4=proj(x4)
> p5=proj(x5)
> ((p2-p1)%*%x5)[1,] *5  ## linear
[1]  2  1  0 -1 -2
> ((p3-p2)%*%x5)[1,] *7  ## quadratic
[1]  2 -1 -2 -1  2
> ((p4-p3)%*%x5)[1,] *10  ## cubic
[1]  1 -2  0  2 -1
> ((p5-p4)%*%x5)[1,] *70  ## quartic
[1]  1 -4  6 -4  1
```

   (b) All $\boldsymbol{c}_i'\boldsymbol{\beta}$ are contrast because $\boldsymbol{c}_i'\boldsymbol{1} = 0$ for $i = 1, 2, 3, 4$.

   (c) By slide 3 of set 9, any two estimable linear combinations $\boldsymbol{c}_i'\boldsymbol{\beta}$ and $\boldsymbol{c}_j'\boldsymbol{\beta}$ are orthogonal if and only if $\boldsymbol{c}_i'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{c}_j = 0$ for $i \neq j$. In the plant density example of slide set 6, the model matrix is

$$\boldsymbol{X} = \begin{bmatrix} \underset{3\times1}{\boldsymbol{1}} & & & & \\ & \underset{3\times1}{\boldsymbol{1}} & & & \\ & & \underset{3\times1}{\boldsymbol{1}} & & \\ & & & \underset{3\times1}{\boldsymbol{1}} & \\ & & & & \underset{3\times1}{\boldsymbol{1}} \end{bmatrix}$$

$$\boldsymbol{X}'\boldsymbol{X} = \begin{bmatrix} 3 & & & & \\ & 3 & & & \\ & & 3 & & \\ & & & 3 & \\ & & & & 3 \end{bmatrix} \text{ and } (\boldsymbol{X}'\boldsymbol{X})^{-1} = \begin{bmatrix} \frac{1}{3} & & & & \\ & \frac{1}{3} & & & \\ & & \frac{1}{3} & & \\ & & & \frac{1}{3} & \\ & & & & \frac{1}{3} \end{bmatrix}$$

Thus in this case, $\boldsymbol{c}_i'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{c}_j = \boldsymbol{c}_i'\boldsymbol{c}_j/3$, so that linear combinations $\boldsymbol{c}_i'\boldsymbol{\beta}$ and $\boldsymbol{c}_j'\boldsymbol{\beta}$ are orthogonal if and only if $\boldsymbol{c}_i'\boldsymbol{c}_j = 0$.
In this problem, all $\boldsymbol{c}_i'\boldsymbol{\beta}$'s are orthogonal because $\boldsymbol{c}_i'\boldsymbol{c}_j = 0$ for all pairs $\{(i,j)|i \neq j\}$ where $i,j = 1,2,3,4$.

2. Given $\boldsymbol{H}$ is a symmetric matrix, by spectral decompositon theorem $\boldsymbol{H} = \sum_{i=1}^{n} \lambda_i\, \boldsymbol{p}_i\, \boldsymbol{p}_i'$, where $\boldsymbol{p}_i$'s are orthonomal eigenvectors of $\boldsymbol{H}$.
"$\Longrightarrow$" part:
By definition, $\boldsymbol{H}$ is nonnegative definite $\Longrightarrow$ $\boldsymbol{p}_i'\boldsymbol{H}\boldsymbol{p}_i \geqslant 0$ for any $\boldsymbol{p}_i$ that $i = 1,\cdots,n$.

$$\begin{aligned}
\boldsymbol{p}_i'\boldsymbol{H}\boldsymbol{p}_i &= \boldsymbol{p}_i'\left(\sum_{j=1}^{n}\lambda_j\,\boldsymbol{p}_j\,\boldsymbol{p}_j'\right)\boldsymbol{p}_i \\
&= \sum_{j=1}^{n}\lambda_j\,\boldsymbol{p}_i'\boldsymbol{p}_j\,\boldsymbol{p}_j'\boldsymbol{p}_i && \boldsymbol{p}_i'\boldsymbol{p}_j = 0 \text{ for all } i \neq j \\
&= \lambda_i\,\boldsymbol{p}_i'\boldsymbol{p}_i\,\boldsymbol{p}_i'\boldsymbol{p}_i && \boldsymbol{p}_i'\boldsymbol{p}_i = 1 \\
&= \lambda_i
\end{aligned}$$

Therefore $\lambda_i \geqslant 0$ for $i = 1,\cdots,n$.
"$\Longleftarrow$" part: given $\lambda_i \geqslant 0$ for $i = 1,\cdots,n$, need to prove $\boldsymbol{y}'\boldsymbol{H}\boldsymbol{y} \geqslant 0$ for any $n \times 1$ vector $\boldsymbol{y}$.

$$\begin{aligned}
\forall \boldsymbol{y},\ \boldsymbol{y}'\boldsymbol{H}\boldsymbol{y} &= \boldsymbol{y}'\left(\sum_{j=1}^{n}\lambda_j\,\boldsymbol{p}_j\,\boldsymbol{p}_j'\right)\boldsymbol{y} && \text{by spectral decompositon} \\
&= \sum_{j=1}^{n}\lambda_j\,\boldsymbol{y}'\boldsymbol{p}_j\,\boldsymbol{p}_j'\boldsymbol{y} && \text{let } \boldsymbol{y}'\boldsymbol{p}_j = x_j \text{ which is a value and } x_j = x_j' \\
&= \sum_{j=1}^{n}\lambda_j\,x_j\,x_j' \\
&= \sum_{j=1}^{n}\lambda_j\,x_j^2 \geqslant 0 && \text{since } \lambda_i \geqslant 0 \text{ and } x_j^2 \geqslant 0 \text{ for } i = 1,\cdots,n
\end{aligned}$$

So, $\boldsymbol{H}$ is nonnegative definite $\Longleftrightarrow$ all its eigenvalues are nonnegative.

3. $y_i = \mu + |x_i|\epsilon_i$ for $i = 1,\cdots,n$ and $\epsilon_i \overset{iid}{\sim} N(0,\sigma^2)$.
we can write the model as

$$\boldsymbol{y} = \underset{n\times 1}{\boldsymbol{1}} \cdot \mu + \boldsymbol{\varepsilon}, \quad \text{where } \boldsymbol{\varepsilon} = \begin{pmatrix} |x_1|\epsilon_1 \\ |x_2|\epsilon_2 \\ \vdots \\ |x_n|\epsilon_n \end{pmatrix} \sim N(\boldsymbol{0}, \sigma^2\boldsymbol{V})$$

$\boldsymbol{V} = diag(x_1^2, x_2^2, \cdots, x_n^2)$ and is positive definite because all x_i's are nonzero. So this is an Aitken model with normal errors.

μ is obviously estimable, so the BLUE is

$$\begin{aligned}
\hat{\mu} &= \left(\mathbf{1}'\boldsymbol{V}^{-1}\mathbf{1}\right)^{-}\mathbf{1}'\boldsymbol{V}^{-1}\boldsymbol{y} \\
&= \left([x_1^{-2}, x_2^{-2}, \cdots, x_n^{-2}]\mathbf{1}\right)^{-1}\left([x_1^{-2}, x_2^{-2}, \cdots, x_n^{-2}]\boldsymbol{y}\right) \\
&= \frac{\sum\limits_{i=1}^{n} x_i^{-2} y_i}{\sum\limits_{i=1}^{n} x_i^{-2}}
\end{aligned}$$

4. (a) Rewrite these hypotheses in terms of the parameter vector.

$$H_0 : \boldsymbol{\alpha}_2 = \mathbf{0} \text{ vs. } H_A : \boldsymbol{\alpha}_2 \neq \mathbf{0}$$

(b) If we want to fit a reduced model corresponding to the null hypothesis, we could use $\boldsymbol{W}_1$ because $\boldsymbol{W}\boldsymbol{\alpha} = \boldsymbol{W}_1\boldsymbol{\alpha}_1 + \boldsymbol{W}_2\boldsymbol{\alpha}_2$ reduces down to $\boldsymbol{W}_1\boldsymbol{\alpha}_1$ under the null hypothesis $H_0 : \boldsymbol{\alpha}_2 = \mathbf{0}$.

(c) Provide a matrix $\boldsymbol{C}$ for testing the main effect of time.
The main effect of time is

$$\frac{1}{2}(\mu_{11} + \mu_{12}) - \frac{1}{2}(\mu_{21} + \mu_{22}) = \left(\frac{1}{2}, \ \frac{1}{2}, \ -\frac{1}{2}, \ -\frac{1}{2}\right)\begin{pmatrix}\mu_{11} \\ \mu_{12} \\ \mu_{21} \\ \mu_{22}\end{pmatrix}$$

So one possible matrix $\boldsymbol{C}$ is $\left(\frac{1}{2}, \ \frac{1}{2}, \ -\frac{1}{2}, \ -\frac{1}{2}\right)$.

(d) Provide a matrix $\boldsymbol{A}$ so that the rank of $\boldsymbol{B}$ is 4. Many answers are possible. One choice that works is

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

(e) Provide a model matrix for a reduced model that corresponds to the null hypothesis of no time main effect.

$$\begin{aligned}
\boldsymbol{W} = \boldsymbol{X}\boldsymbol{B}^{-1} &= \begin{bmatrix} \underset{2\times 1}{\mathbf{1}} & & & \\ & \underset{3\times 1}{\mathbf{1}} & & \\ & & \underset{4\times 1}{\mathbf{1}} & \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}^{-1} \\
&= \begin{bmatrix} \underset{2\times 1}{\mathbf{1}} & & & \\ & \underset{3\times 1}{\mathbf{1}} & & \\ & & \underset{4\times 1}{\mathbf{1}} & \\ 1 & 1 & -1 & -2 \end{bmatrix}
\end{aligned}$$

By part (b), the model matrix of the reduced model is the first $p-q=3$ columns of $\boldsymbol{W}$, i.e.,

$$\boldsymbol{W}_1 = \begin{bmatrix} \underset{2\times 1}{\mathbf{1}} & & \\ & \underset{3\times 1}{\mathbf{1}} & \\ & & \underset{4\times 1}{\mathbf{1}} \\ 1 & 1 & -1 \end{bmatrix}.$$

(f) Find the error sum of squares for the reduced and full models.
$SS_{Reduced} = \boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{W}_1})\boldsymbol{y} = 24$
$SS_{Full} = \boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{W}})\boldsymbol{y} = 12$

(g) Find the degrees of freedom associated with the sums of squares in part (f).
$df_{Reduced} = rank(\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{W}_1}) = 10 - 3 = 7$
$df_{Full} = rank(\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{W}}) = 10 - 4 = 6$

(h) Compute the F -statistic for testing the null hypothesis of no time main effect using the sums of squares and degrees of freedom computed in parts (f) and (g).
From the code and output below,

$$F = \frac{(SSE_{Reduced} - SSE_{Full})/(df_{Reduced} - df_{Full})}{SSE_{Full}/df_{Full}} = 6$$

```
> X=model.matrix(~ 0 + factor(c(rep(1, 2), rep(2, 3), rep(3, 4), 4)))
> B=rbind(cbind(diag(3), 0), .5 * c(1, 1, -1, -1))
> W=X %*% solve(B)
> W1=W[, -4]
> proj=function(x) {
+    x %*% MASS::ginv(t(x) %*% x) %*% t(x)
+ }
> Pw=proj(W)
> Pw1=proj(W1)
> y=c(3, 5, 11, 13, 15, 5, 6, 6, 7, 16)
> I10=diag(rep(1, length(y)))
> SS.red=t(y) %*% (I10 - Pw1) %*% y
> SS.full=t(y) %*%(I10 - Pw) %*% y
> (SS.red - SS.full) / SS.full * 6
```

5. (a) Note that $E(\boldsymbol{a}'\boldsymbol{y}) = E(a_1y_1 + a_2y_2) = a_1E(y_1) + a_2E(y_2) = (a_1 + a_2)\mu$. In order for $\boldsymbol{a}'\boldsymbol{y} = a_1y_1 + a_2y_2$ to be an unbiased estimator of μ, $a_1 + a_2 = 1$ because $(a_1 + a_2)\mu$ must be μ for all μ in $\mathbb{R}$.

(b)

$$Var(\boldsymbol{a}'\boldsymbol{y}) = \boldsymbol{a}'Var(\boldsymbol{y})\boldsymbol{a} = (a_1, a_2)\begin{pmatrix} 1/4 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \frac{1}{4}a_1^2 + a_2^2.$$

(c) Note that $a_1 + a_2 = 1$ in part (a). Using this fact and the result in part (b),

$$\begin{aligned} Var(\boldsymbol{a}'\boldsymbol{y}) &= \frac{1}{4}a_1^2 + a_2^2 = \frac{1}{4}a_1^2 + (1 - a_1)^2 \\ &= \frac{1}{4}a_1^2 + 1 - 2a_1 + a_1^2 = \frac{5}{4}a_1^2 - 2a_1 + 1 \end{aligned}$$

(d) To be the BLUE of μ, $\boldsymbol{a}'\boldsymbol{y}$ must be an unbiased estimator with the minimum variance. Using parts (a) through (c), an unbiased estimator of μ has the variance of the form in terms of a single variable a_1 as follows:

$$Var(\boldsymbol{a}'\boldsymbol{y}) = \frac{5}{4}a_1^2 - 2a_1 + 1 \overset{set}{=} f(a_1)$$

To find the minimum variance, we need to check the following:

$$\begin{aligned} \frac{d}{da_1}f(a_1) &= \frac{10}{4}a_1 - 2 \overset{set}{=} 0, \\ \frac{d^2}{da_1^2}f(a_1) &= \frac{10}{4} > 0. \end{aligned}$$

$f(a_1)$ achieves the minimum at $a_1 = \frac{8}{10}$. Therefore, $a_2 = 1 - a_1 = 1 - \frac{8}{10} = \frac{2}{10}$ and $\frac{8}{10}y_1 + \frac{2}{10}y_2 = \frac{4}{5}y_1 + \frac{1}{5}y_2$ is the BLUE of μ.

(e) Consider the following model:

$$\boldsymbol{y} = \boldsymbol{X}\mu + \boldsymbol{\epsilon},\ E(\boldsymbol{\epsilon}) = \boldsymbol{0} \text{ and } Var(\boldsymbol{\epsilon}) = \sigma^2\boldsymbol{V}$$

where $\boldsymbol{X} = \underset{2\times 1}{\boldsymbol{1}}$, $\sigma^2 = 1$ and $\boldsymbol{V} = \begin{pmatrix} 1/4 & 0 \\ 0 & 1 \end{pmatrix}$ is a positive definite variance matrix. This model becomes the Aitken model on slide 8 of slide set 10. Then, using the result on slide 12 of slide set 10, $\hat{\mu}_{GLS}$ becomes the BLUE of estimable μ where

$$\begin{aligned} \hat{\mu}_{GLS} &= \left(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y} \\ &= \left(\underset{2\times 1}{\boldsymbol{1}'}\begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}\underset{2\times 1}{\boldsymbol{1}}\right)^{-1}\underset{2\times 1}{\boldsymbol{1}'}\begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\ &= (5)^{-1}(4y_1 + y_2) \\ &= \frac{4}{5}y_1 + \frac{1}{5}y_2 \end{aligned}$$

which is the same result in part (d).