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.

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

```
> 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 contrasts 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 positive definite $\Longrightarrow$ $\boldsymbol{p}_i'\boldsymbol{H}\boldsymbol{p}_i > 0$ for any $\boldsymbol{p}_i$ that $i = 1, \cdots, n$.

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

Therefore $\lambda_i > 0$ for $i = 1, \cdots, n$.
"$\Longleftarrow$" part: given $\lambda_i > 0$ for $i = 1, \cdots, n$, need to prove $\boldsymbol{y}'\boldsymbol{H}\boldsymbol{y} > 0$ for any non-zero $n \times 1$ vector $\boldsymbol{y}$. Let $\boldsymbol{y} \in \mathbb{R}^n \setminus \{\boldsymbol{0}\}$ be arbitrary. By the Spectral Decomposition Theorem, $\boldsymbol{H} = \boldsymbol{P}\text{diag}(\lambda_1, \ldots, \lambda_n)\boldsymbol{P}'$, where $\boldsymbol{P} = [\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n]$ and $\boldsymbol{P}\boldsymbol{P}' = \boldsymbol{P}'\boldsymbol{P} = \boldsymbol{I}$, which implies $\boldsymbol{P}'$ is nonsingular. Thus, the columns of $\boldsymbol{P}'$ are linearly independent so that

$$\begin{aligned}
\boldsymbol{y} \neq \boldsymbol{0} &\implies \boldsymbol{P}'\boldsymbol{y} \neq \boldsymbol{0} \\
&\implies \boldsymbol{p}_j'\boldsymbol{y} \neq 0 \text{ for at least some } j \in \{1, \ldots, n\}.
\end{aligned}$$

For $j = 1, \ldots, n$, let $x_j = \boldsymbol{p}_j'\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{p}_j$. Note that by the argument above, $x_j \neq 0$ for at least some $j \in \{1, \ldots, n\}$.

$$\begin{aligned}
\boldsymbol{y}'\boldsymbol{H}\boldsymbol{y} &= \boldsymbol{y}' \left(\sum_{j=1}^{n} \lambda_j \, \boldsymbol{p}_j \, \boldsymbol{p}_j' \right) \boldsymbol{y} && \text{by spectral decompositon} \\
&= \sum_{j=1}^{n} \lambda_j \, \boldsymbol{y}'\boldsymbol{p}_j \, \boldsymbol{p}_j'\boldsymbol{y} \\
&= \sum_{j=1}^{n} \lambda_j \, x_j^2 > 0 && \begin{array}{l}\text{because } \lambda_j > 0 \text{ for } j = 1, \ldots, n \text{ and} \\ x_j^2 > 0 \text{ for some } j \in \{1, \ldots, n\}\end{array}
\end{aligned}$$

So, $\boldsymbol{H}$ is positive definite $\iff$ all its eigenvalues are positive.

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

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

4. (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 $\boldsymbol{X} = \underset{2\times1}{\boldsymbol{1}}$, $\sigma^2 = 1$ 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

$$\begin{aligned}
\hat{\mu}_{GLS} &= \left(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y} \\
&= \left(\underset{2\times1}{\boldsymbol{1}'}\begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}\underset{2\times1}{\boldsymbol{1}}\right)^{-1}\underset{2\times1}{\boldsymbol{1}'}\begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\
&= (5)^{-1}(4y_1 + y_2) \\
&= \frac{4}{5}y_1 + \frac{1}{5}y_2
\end{aligned}$$

which is the same result in part (d).

5. The Aitken Model with normal errors described on slide 18 of slide set 10 can be transformed to $\boldsymbol{z} = \boldsymbol{W}\boldsymbol{\beta} + \boldsymbol{\delta}$, $\boldsymbol{\delta} \sim N(\boldsymbol{0}, \sigma^2\boldsymbol{I})$, where $\boldsymbol{z}$=$\boldsymbol{V}^{-1/2}\boldsymbol{y}$, $\boldsymbol{W}$=$\boldsymbol{V}^{-1/2}\boldsymbol{X}$ and $\boldsymbol{\delta}$=$\boldsymbol{V}^{-1/2}\boldsymbol{\epsilon}$. With this transformation, we can apply all the results we have established previously to the Gauss-Markov model with normal errors. Thus, the 95% confidence interval for estimable $\boldsymbol{c}'\boldsymbol{\beta}$ is $\boldsymbol{c}'(\boldsymbol{W}'\boldsymbol{W})^-\boldsymbol{W}'\boldsymbol{z} \pm t_{n-rank(W)\ 0.975}\sqrt{\frac{\boldsymbol{z}'(\boldsymbol{I}-\boldsymbol{P}_W)\boldsymbol{z}}{n-rank(\boldsymbol{W})}\boldsymbol{c}'(\boldsymbol{W}'\boldsymbol{W})^-\boldsymbol{c}}$.

   Replacing $\boldsymbol{W}$ with $\boldsymbol{V}^{-1/2}\boldsymbol{X}$ and $\boldsymbol{z}$ with $\boldsymbol{V}^{-1/2}\boldsymbol{y}$ and simplifying yields

   $\boldsymbol{c}'(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y} \pm t_{n-r,0.975} \times \sqrt{\frac{(\boldsymbol{y}-\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y})'\boldsymbol{V}^{-1}(\boldsymbol{y}-\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y})}{n-r}\boldsymbol{c}'(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^-\boldsymbol{c}}$

6. Let θ be the expected value of yield for plots receiving 4 pounds of fertilizer per acre. From the Exam 1 solution, we know that $\hat{\theta}$ can be expressed as $\bar{y}_{..}$ with $E(\hat{\theta}) = (1/4)(160+180+200+\mu_4)$ and $Var(\hat{\theta}) = \frac{9}{5}$. Moreover, we also know that mean squared error of $\hat{\mu}_3$, the estimator of θ obtained from the fit of a cell-means model, is $\frac{36}{5}$. To have lower $MSE(\bar{y}_{..})$ than $MSE(\hat{\mu}_3)$,

$$\begin{aligned}
MSE(\bar{y}_{..}) < MSE(\hat{\mu}_3) &\iff \left(\frac{160+180+200+\mu_4}{4} - 200\right)^2 + \frac{9}{5} < \frac{36}{5} \\
&\iff \left|\frac{540+\mu_4}{4} - 200\right| < \sqrt{\frac{27}{5}} \\
&\iff 260 - 4\sqrt{\frac{27}{5}} < \mu_4 < 260 + 4\sqrt{\frac{27}{5}}
\end{aligned}$$

   Thus, μ_4 should be in (250.7048, 269.2952). Note that when $\mu_4 = 260$, then the simple linear model (1) holds. The point of this problem is that fitting the incorrect, but simpler, Model (1) provides an estimator of θ with less variance, more bias, and lower mean squared error than the unbiased estimator of θ from the correct model, provided that Model (1) is not too far from correct (i.e, as long as μ_4 is not too far from 260).