1. (6.5 from M+S) $y_{ij} | a_i \overset{ind}{\sim} N(\mu + a_i, \sigma^2)$ $i = 1, \ldots, m$

$a_i \overset{iid}{\sim} N(0, \sigma_a^2)$ $j = 1, \ldots, n$

Find a $100(1-\alpha)\%$ CI for $\rho = \dfrac{\sigma_a^2}{\sigma_a^2 + \sigma^2}$

The anova table for this design is

Source	SS	d.f	MS	E(MS)
Groups	SS_{Groups}	$m-1$	$SS_{Groups}/(m-1)$	$n\sigma_a^2 + \sigma^2$
Error	SS_E	$nm - m$	$SS_E/(nm-m)$	σ^2
Total (corrected)	SS_T	$nm - 1$		

where $SS_T = \sum_i \sum_j (y_{ij} - \bar{y}_{..})^2$, $SS_{Groups} = \sum_i \sum_j (\bar{y}_{i.} - \bar{y}_{..})^2$

and $SS_E = SS_T - SS_{Groups}$

It can be shown that $\dfrac{SS_E}{\sigma^2} \sim \chi^2(nm - m)$,

$$\frac{SS_{Groups}}{\sigma^2 + n\sigma_a^2} \sim \chi^2(m-1)$$

and SS_E + SS_{Groups} are independent.

Therefore $MS_{Groups} \sim (\sigma^2 + n\sigma_a^2) \dfrac{\chi^2(m-1)}{m-1}$

$$MS_E \sim \sigma^2 \frac{\chi^2(nm-m)}{nm-m}$$

and MS_{Groups}, MS_E are indep.

$$\Rightarrow \frac{MS_{Groups}}{MS_E} \sim \underbrace{\frac{\sigma^2 + n\sigma_a^2}{\sigma^2}}_{= 1+n\Theta} F(m-1, nm-m) \quad \text{where } \Theta = \frac{\sigma_a^2}{\sigma^2} = \frac{\rho}{1-\rho}$$

$$\Rightarrow \frac{MS_{Groups}/MS_E}{1+n\Theta} \sim F(m-1, nm-m)$$

$$\Rightarrow Pr\left(F_{\alpha/2}(m-1, nm-m) \le \frac{MS_{Groups}/MS_E}{1+n\Theta} \le F_{1-\alpha/2}(m-1, nm-m) \right) = 1-\alpha$$

$$\Rightarrow Pr\left(F_{\alpha/2}(m-1, nm-m)^{-1} \frac{MS_{Groups}}{MS_E} \ge 1+n\Theta \ge F_{1-\alpha/2}(m-1, nm-m)^{-1} \frac{MS_{Groups}}{MS_E} \right) = 1-\alpha$$

$$\Rightarrow Pr\left(U \ge \Theta \ge L \right) = 1-\alpha \quad \text{where } U = \frac{1}{n}\left[F_{\alpha/2}(m-1, nm-m)^{-1} \frac{MS_{Groups}}{MS_E} - 1 \right]$$

$$\text{and } L = \frac{1}{n}\left[F_{1-\alpha/2}(m-1, nm-m)^{-1} \frac{MS_{Groups}}{MS_E} - 1 \right]$$

Since $\rho = \dfrac{\Theta}{1+\Theta}$, This 95% CI for Θ can be transformed to a 95% CI for ρ as follows:

$$\left[\frac{L}{1+L}, \frac{U}{1+U} \right]$$

2. (6.6 in M+S) Write the following models in matrix notation + determine $E(y)$, $var(y)$

a) $y_{ij} \mid a_i \overset{ind}{\sim} N(\mu + a_i + \beta_j, \sigma^2)$ $\quad i = 1, \dots, m$

$\quad a_i \overset{iid}{\sim} N(0, \sigma_a^2)$ $\quad j = 1, \dots, n$

$\Rightarrow y_{ij} = \mu + a_i + \beta_j + \varepsilon_{ij}$ where $\varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2)$, $a_i \overset{iid}{\sim} N(0, \sigma_a^2)$

$$
\begin{pmatrix} y_{11} \\ \vdots \\ y_{1n} \\ y_{21} \\ \vdots \\ y_{2n} \\ \vdots \\ y_{m1} \\ \vdots \\ y_{mn} \end{pmatrix}
=
\underbrace{\begin{pmatrix}
1 & 1 & 0 & \dots & 0 \\
\vdots & 0 & 1 & & 0 \\
\vdots & \vdots & \vdots & & \vdots \\
1 & 0 & 0 & & 1 \\
1 & 1 & 0 & & 0 \\
\vdots & 0 & 1 & & 0 \\
\vdots & \vdots & \vdots & & \vdots \\
1 & 0 & 0 & & 1 \\
\vdots & \vdots & \vdots & & \vdots \\
1 & 1 & 0 & & 0 \\
\vdots & 0 & 1 & & 0 \\
\vdots & \vdots & \vdots & & \vdots \\
1 & 0 & 0 & & 1
\end{pmatrix}}_{= (j_N,\ j_m \otimes I_n),\ N = nm}
\begin{pmatrix} \mu \\ \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}
+
\underbrace{\begin{pmatrix}
1 & 0 & \dots & 0 \\
\vdots & \vdots & & \vdots \\
1 & 0 & & 0 \\
0 & 1 & & 0 \\
\vdots & \vdots & & \vdots \\
0 & 1 & & 0 \\
\vdots & \vdots & & \vdots \\
0 & 0 & & 1 \\
\vdots & \vdots & & \vdots \\
0 & 0 & & 1
\end{pmatrix}}_{= I_m \otimes j_n}
\begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix}
+
\begin{pmatrix} \varepsilon_{11} \\ \vdots \\ \varepsilon_{1n} \\ \varepsilon_{21} \\ \vdots \\ \varepsilon_{2n} \\ \vdots \\ \varepsilon_{m1} \\ \vdots \\ \varepsilon_{mn} \end{pmatrix}
$$

or $\underline{y} = \underline{X}\underline{\beta} + \underline{Z}\underline{b} + \underline{\varepsilon}$

$E(y_{ij}) = \mu + \beta_j$, $var(y_{ij}) = \sigma_a^2 + \sigma^2$

$$
cov(y_{ij}, y_{i'j'}) = \begin{cases} \sigma_a^2, & i = i',\ j \neq j' \\ 0, & i \neq i' \end{cases}
$$

$\Rightarrow var(\underline{y}) = blockdiag(\sigma_a^2 J + \sigma^2 I, \dots, \sigma_a^2 J + \sigma^2 I)$
$= I_m \otimes (\sigma_a^2 J_{n,n} + \sigma^2 I)$

b.) $y_{ij} \mid a_i, b_j \overset{ind}{\sim} N(\mu_* + a_i + b_j, \sigma^2)$

$a_i \overset{iid}{\sim} N(0, \sigma_a^2)$, $b_j \overset{iid}{\sim} N(0, \sigma_b^2)$

indep.

Can be written as $y_{ij} = \mu + a_i + b_j + \varepsilon_{ij}$
where $\varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2)$ independent of a_i's, b_j's as given above.

Let $\underline{y}$, $\underline{\varepsilon}$ be as in part (a). Then model in matrix form is

$\underline{y} = \underline{X}\beta + \underline{Z}\underline{b} + \underline{\varepsilon}$, where $\underline{X} = \underline{j}_N$, $N = nm$

$\underline{\beta} = \mu$, $\underline{Z} = (\underline{I}_m \otimes \underline{j}_n, \underline{j}_m \otimes \underline{I}_n)$, $\underline{b} = \begin{pmatrix} a_1 \\ \vdots \\ a_m \\ b_1 \\ \vdots \\ b_n \end{pmatrix}$

$E(\underline{y}) = \underline{X}\beta = \mu \underline{j}_N$, $var(\underline{y}) = var(\underline{Z}_1 \underline{b}_1 + \underline{Z}_2 \underline{b}_2 + \underline{\varepsilon})$

where $\underline{Z}_1 = (I_m \otimes \underline{j}_n)$, $\underline{Z}_2 = j_m \otimes I_n$, $\underline{b}_1 = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix}$, $\underline{b}_2 = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$

$\Rightarrow var(\underline{y}) = \sigma_a^2 \underline{Z}_1 \underline{Z}_1^T + \sigma_b^2 \underline{Z}_2 \underline{Z}_2^T + \sigma^2 I_N$

$\underline{Z}_1 \underline{Z}_1^T = (I_m \otimes \underline{j}_n)(I_m \otimes \underline{j}_n)^T = (I_m \otimes j_n)(I_m^T \otimes \underline{j}_n^T) = (I_n \otimes \underline{j}_n \underline{j}_n^T)$
$= I_m \otimes J_{n,n}$

similarly, $\underline{Z}_2 Z_2^T = (J_{m,m} \otimes I_n)$

c) $y_{ijk} \mid a_i, g_{ij} \overset{ind}{\sim} N(\mu_{\cancel{g}} + a_i + \beta_j + g_{ij}, \sigma^2)$

$a_i \overset{iid}{\sim} N(0, \sigma_a^2)$, $g_{ij} \overset{iid}{\sim} N(0, \sigma_g^2)$, a_i, g_{ij} indep.
$i = 1, \ldots, m$, $j = 1, \ldots, n$, $k = 1, \ldots, r$

Let $\underline{y} = (y_{111}, y_{112}, \ldots, y_{11r}, y_{121}, \ldots, y_{12r}, \ldots, y_{mn1}, \ldots, y_{mnr})^T$
similarly,
$\underline{\varepsilon} = (\varepsilon_{111}, \ldots \quad \ldots, \varepsilon_{mnr})^T$
Then model can be written as $\underline{y} = \underline{X}\beta + \underline{Z}\underline{b} + \underline{\varepsilon}$
where $\underline{X} = (\underline{j}_N, \underline{j}_m \otimes (I_n \otimes \underline{j}_r))$

$$\beta = \cancel{(\beta_1, \ldots, \beta_n)^T} (\mu, \beta_1, \ldots, \beta_n)^T$$

$\underline{Z} = (\underline{Z}_1, \underline{Z}_2)$, $\underline{b} = (\underline{b}_1^T, \underline{b}_2^T)^T$, $\underline{Z}_1 = I_m \otimes \underline{j}_{nr}$

$\underline{Z}_2 = I_{mn} \otimes \underline{j}_r$, $\underline{b}_1 = (a_1, \ldots, a_m)^T$, $\underline{b}_2 = (g_{11}, \ldots, g_{mn})^T$

$Var(\underline{y}) = \sigma_a^2 Z_1 Z_1^T + \sigma_g^2 Z_2 Z_2^T + \sigma^2 I_N$

$\underline{Z}_1 Z_1^T = (I_m \otimes \underline{j}_{nr})(I_m \otimes \underline{j}_{nr})^T = (I_m \otimes \underline{j}_{nr})(I_m \otimes \underline{j}_{nr}^T)$

$= (I_m \otimes \underline{j}_{nr} \underline{j}_{nr}^T) = (I_m \otimes J_{nr,nr})$

similarly,
$\underline{Z}_2 \underline{Z}_2^T = (I_{mn} \otimes J_{r,r})$

$E(\underline{y}) = X\beta \quad (E(y_{ij}) = \mu + \beta_j)$

2. (6.8 in McCulloch et al.)

a) Let $w = K^T y$. With the substitutions given in (6.66), ~~(6.58) becomes~~ the loglikelihood becomes

$$l_1 = -\frac{(n-s)}{2}\log(2\pi) - \frac{1}{2}\log|K^T V K| - \frac{1}{2} y^T K (K^T V K)^{-1} K^T y$$

$$= -\frac{(n-s)}{2}\log(2\pi) - \frac{1}{2}\log|K^T V K| - \frac{1}{2} w^T (K^T V K)^{-1} w$$

b.) $l_2 = \text{constant} - \frac{1}{2}\log|V| - \frac{1}{2}\log|X^T V^{-1} X| - \frac{1}{2} y^T P y$

where

$$P = V^{-1} - V^{-1} X (X^T V^{-1} X)^{-} X^T V^{-1}$$

Show that the quadratic forms in y are the same in l_1 + l_2.

It is sufficient to show

$$K (K^T V K)^{-1} K^T = P = V^{-1} - V^{-1} X (X^T V^{-1} X)^{-} X^T V^{-1}$$

First note that $K (K^T K)^{-1} K^T = P_{C(X)^\perp}$

because $K^T K = I$ + $K K^T = P_{C(X)^\perp}$

~~Sub~~

Our model is $y \sim N(X\beta, V)$. It follows that

$$\tilde{y} \equiv V^{-1/2} y \sim N(V^{-1/2} X\beta, V^{-1/2} V V^{-T/2})$$

$$= N(\tilde{X}\beta, I)$$

where $\tilde{X} = V^{-1/2} X$

If $K^T y$ is a vector of error contrasts for y then for $\tilde{K} = V^{1/2} K$,

$\tilde{K}^T \tilde{y}$ is a vector of error contrasts for $\tilde{y}$

because

$$E(\tilde{K}^T \tilde{y}) = \tilde{K}^T E(\tilde{y}) = \tilde{K}^T \tilde{X} \beta$$

$$= K^T V^{T/2} V^{-1/2} X\beta$$

$$= K^T X \beta = 0 \ \forall \beta$$

So $P_{C(\tilde{X})^\perp} = \tilde{K}(\tilde{K}^T \tilde{K})^{-1} \tilde{K}^T$

or $I - V^{-1/2} X (X^T V^{-1/2} V^{-1/2} X)^{-} X^T V^{-1/2} = V^{1/2} K (K^T V^{1/2} V^{1/2} K)^{-1} K^T V^{1/2}$

or, if we left multiply and right multiply both sides by $V^{-1/2}$ we get

$$V^{-1} - V^{-1} X (X^T V^{-1} X)^{-} X^T V^{-1} = K(K^T V K)^{-1} K^T \quad \checkmark$$

6.8 (c) If V is a function of t. write $\partial V/\partial t$ as V_t. Show that $\partial l_1/\partial t = \partial l_2/\partial t$

Proof: Since
$$\begin{cases} \dfrac{\partial A^{-1}}{\partial s} = -A^{-1}\dfrac{\partial A}{\partial s}A^{-1} \\ \dfrac{\partial \log|A|}{\partial s} = tr\left(A^{-1}\dfrac{\partial A}{\partial s}\right) \end{cases}$$

$$\frac{\partial l_1}{\partial t} = -\frac{1}{2}\frac{\partial \log|K^TVK|}{\partial t} - \frac{1}{2}\frac{\partial Q_1}{\partial t}, \quad \text{where } Q_1 = y^TK(K^TVK)^{-1}K^Ty$$
$$= -\frac{1}{2}tr\left[(K^TVK)^{-1}K^T\frac{\partial V}{\partial t}K\right] - \frac{1}{2}\frac{\partial Q_1}{\partial t}$$
$$= -\frac{1}{2}tr\left[K(K^TVK)^{-1}K^T\frac{\partial V}{\partial t}\right] - \frac{1}{2}\frac{\partial Q_1}{\partial t}$$
$$= -\frac{1}{2}tr\left[K(K^TVK)^{-1}K^T V_t\right] - \frac{1}{2}\frac{\partial Q_1}{\partial t} = -\frac{1}{2}tr(AV_t) - \frac{1}{2}\frac{\partial Q_1}{\partial t}$$

$$\frac{\partial l_2}{\partial t} = -\frac{1}{2}\frac{\partial \log|V|}{\partial t} - \frac{1}{2}\frac{\partial \log|X^TV^{-1}X|}{\partial t} - \frac{1}{2}\frac{\partial Q_2}{\partial t}, \quad \text{where } Q_2 = y^TPy$$
$$= -\frac{1}{2}tr\left(V^{-1}\frac{\partial V}{\partial t}\right) - \frac{1}{2}tr\left[(X^TV^{-1}X)^{-1}\left(-X^TV^{-1}\frac{\partial V}{\partial t}V^{-1}X\right)\right] - \frac{1}{2}\frac{\partial Q_2}{\partial t}$$
$$= -\frac{1}{2}tr(V^{-1}V_t) + \frac{1}{2}tr\left[V^{-1}X(X^TV^{-1}X)^{-1}X^TV^{-1}\frac{\partial V}{\partial t}\right] - \frac{1}{2}\frac{\partial Q_2}{\partial t}$$
$$= -\frac{1}{2}tr\left[V^{-1}V_t - V^{-1}X(X^TV^{-1}X)^{-1}X^TV^{-1}V_t\right] - \frac{1}{2}\frac{\partial Q_2}{\partial t}$$
$$= -\frac{1}{2}tr\left[\left(V^{-1} - V^{-1}X(X^TV^{-1}X)^{-1}X^TV^{-1}\right)V_t\right] - \frac{1}{2}\frac{\partial Q_2}{\partial t} = \frac{1}{2}tr(PV_t) - \frac{1}{2}\frac{\partial Q_2}{\partial t}$$

We have seen from part (b) that
$$A = K(K^TVK)^{-1}K^T = P = V^{-1} - V^{-1}X(X^TV^{-1}X)^{-1}X^TV^{-1}$$
and $Q_1 = Q_2$.

thus it follows that $\dfrac{\partial l_1}{\partial t} = \dfrac{\partial l_2}{\partial t}$.

(6.10 in M+S) Derive ML + REML solutions for the following models

a) $E(y_{ij}) = \mu$, $\underline{V} = \sigma^2 \underline{\underline{I}}_n$ $\quad i=1,\dots,m,\ j=1,\dots,n$

Here, $\underset{\beta}{\underline{X}\underline{\beta}} = \mu \underset{\underline{X}}{\underline{j}_N}$, $\theta = \sigma^2$

The REML Estimating equation for θ is given on the top of p. 84 of the notes and the ML EE is given on p. 73 (0)

For these equations, we need $\frac{\partial V}{\partial \sigma^2} = I_n$

$$\beta_{\sigma^2} = (X^T(\sigma^2 I)^{-1}X)^- X^T(\sigma^2 I)^{-1}y = (X^TX)^{-1}X^Ty = \underline{\beta}_{OLS} = \bar{y}_{..}$$

$$Q = (\sigma^2 I)^{-1}\left[I - X(X^T(\sigma^2 I)^{-1}X)^- X^T(\sigma^2 I)^{-1}\right]$$

full rank so $=$ to regular inverse

$$= \sigma^{-2}\left[I - X(X^TX)^- X^T\right] = \sigma^{-2}\left[I - \underline{j}_N(\underline{j}_N^T\underline{j}_N)^{-1}\underline{j}_N^T\right]$$

$$= \sigma^{-2}\left[I - \frac{1}{N}J_{N,N}\right]$$

Now the REML E.E. for σ^2 is

$$\frac{1}{2}(\underline{y} - \bar{y}_{..}\underline{j}_N)^T(\sigma^2 I)^{-1} I (\sigma^2 I)^{-1}(\underline{y} - \bar{y}_{..}\underline{j}_N)$$

$$= \frac{1}{2}\operatorname{tr}\left(\sigma^{-2}\left[I - \frac{1}{N}J_{N,N}\right] I\right)$$

Multiplying through by $2\sigma^4$, we get

$$(\underline{y} - \bar{y}_{\cdot\cdot}\underline{j}_N)^T(\underline{y} - \bar{y}_{\cdot\cdot}\underline{j}_N) = \sigma^2 tr\left(P_{C(X)^\perp}\right)$$

or $$\sum_i\sum_j (y_{ij} - \bar{y}_{\cdot\cdot})^2 = \sigma^2(N-1)$$

$$\Rightarrow \hat{\sigma}^2_{REML} = \frac{1}{N-1}\sum_i\sum_j (y_{ij} - \bar{y}_{\cdot\cdot})^2$$

The ML equation for σ^2 is the same, w/ Q replaced by V^{-1}
$\Rightarrow$ the ML equation becomes

$$\sum_i\sum_j (y_{ij} - \bar{y}_{\cdot\cdot})^2 = 2\sigma^4 \frac{1}{2} tr\left(V^{-1}\frac{\partial V}{\partial \sigma^2}\right)$$
$$= \sigma^2 tr(I) = N\sigma^2$$

$$\Rightarrow \hat{\sigma}^2_{ML} = \frac{1}{N}\sum_i\sum_j (y_{ij} - \bar{y}_{\cdot\cdot})^2$$

b.) $E(y_{ij}) = \mu + \alpha_i$, $V = \sigma^2 I_N$

Now $$\underline{X}\underline{\beta} = (\underline{j}_N, I_M \otimes \underline{j}_n)\begin{pmatrix}\mu \\ \alpha_1 \\ \vdots \\ \alpha_M\end{pmatrix}, \quad \theta = \sigma^2$$

$\beta_{\sigma^2} = \hat{\beta}_{OLS}$ again and $Q = \sigma^{-2}[I - X(X^TX)^- X^T]$
$= \sigma^{-2} P_{C(X)^\perp}$ again

Therefore, the REML E.E. for σ^2 is

$$\frac{1}{2\sigma^4}(\underline{y} - \underline{X}\hat{\beta}_{OLS})^T(\underline{y} - \underline{X}\hat{\beta}_{OLS}) = \frac{1}{2} tr\left(\sigma^{-2} P_{C(X)^\perp} I\right)$$

or $(\underline{y}-\underline{X}\hat{\beta}_{OLS})^T(\underline{y}-\underline{X}\hat{\beta}_{OLS}) = \sigma^2 tr(P_{C(\underline{X})^\perp}) = \sigma^2(N - \text{rank}(\underline{X}))$
$= \sigma^2(N-m)$

$\Rightarrow \hat{\sigma}^2_{REML} = \frac{1}{N-m}\sum_i\sum_j(y_{ij}-\bar{y}_{i\cdot})^2$

since $\underline{X}\hat{\beta}_{OLS} = p(\underline{y}|C(\underline{X})) = \begin{pmatrix}\bar{y}_{1\cdot}\underline{j}_n\\ \vdots \\ \bar{y}_{m\cdot}\underline{j}_n\end{pmatrix}$

Similarly, the ML E.E. for σ^2 ~~gets~~ becomes

$$\sum_i\sum_j(y_{ij}-\bar{y}_{i\cdot})^2 = 2\sigma^4\frac{1}{2}tr\left(V^{-1}\frac{\partial V}{\partial\sigma^2}\right)$$
$$= \sigma^2 tr(I_N) = \sigma^2 N$$

$\Rightarrow \hat{\sigma}^2_{ML} = \frac{1}{N}\sum_i\sum_j(y_{ij}-\bar{y}_{i\cdot})^2$

d.) $\mathcal{E}(y_{i}) = \mu + \beta x_{i}$, $\underline{V} = \sigma^2 I_N$

Here, $\underline{X}\underline{\beta} = (\underline{j}_N, \underline{x})\begin{pmatrix}\mu \\ \beta\end{pmatrix}$ where $\underline{x} = (x_1, \dots, x_N)^T$, $\theta = \sigma^2$

As in parts (a) + (b), $\beta_{\sigma^2} = \hat{\beta}_{OLS}$ and $Q = \sigma^{-2} P_{C(X)^\perp}$ and $\frac{\partial V}{\partial \sigma^2} = I$

Therefore, the REML E.E. for σ^2 is

$$\frac{1}{2\sigma^4}(\underline{y} - \underline{X}\hat{\beta}_{OLS})^T(\underline{y} - \underline{X}\hat{\beta}_{OLS}) = \frac{1}{2} tr(\sigma^{-2} P_{C(X)^\perp} I)$$

or $(\underline{y} - \underline{X}\hat{\beta}_{OLS})^T(\underline{y} - \underline{X}\hat{\beta}_{OLS}) = \sigma^2(N - rank(\underline{X}) = \sigma^2(N-2)$

$$\Rightarrow \hat{\sigma}^2_{REML} = \frac{\sum_i (y_i - \hat{y}_i)^2}{N-2} = SS_E / (N-2)$$

where $\hat{y}_i$ is ~~the~~ $\hat{\mu} + \hat{\beta}x_i$ where $\hat{\mu}, \hat{\beta}$ are the usual OLS estimates of intercept + slope in a simple linear regression

Similarly, the ML E.E. for σ^2 becomes

$$SS_E = 2\sigma^4 \frac{1}{2} tr\left(V^{-1} \frac{\partial V}{\partial \sigma^2}\right) = \sigma^2 tr(I_N) = \sigma^2 N$$

$$\Rightarrow \hat{\sigma}^2_{ML} = SS_E / N$$

Additional Problem #1:

#(6.1 in Davis' book) $\underline{y}_i = X\beta + \underline{Z}\gamma_i + \underline{\varepsilon}_i$ where

$\underline{Z} = \underline{j}_t$, γ_i's $\overset{ind}{\sim} N(0, \sigma^2)$, $\underline{\varepsilon}_i \overset{ind}{\sim} N(\underline{0}, \sigma_e^2 \underline{I}_t)$

indep

$var(\underline{y}_i) = \underline{Z}\, var(\gamma_i)\, \underline{Z}^T + var(\underline{\varepsilon}_i) = \underline{j}_t \sigma^2 \underline{j}_t^T + \sigma_e^2 \underline{I}_t$

$$= \sigma^2 \underline{J}_{t,t} + \sigma_e^2 \underline{I}_t = \begin{pmatrix} \sigma^2+\sigma_e^2 & \sigma^2 & \sigma^2 & \cdots & \sigma^2 \\ & \sigma^2+\sigma_e^2 & \sigma^2 & \cdots & \sigma^2 \\ & \text{symmetric} & & \ddots & \vdots \\ & & & & \sigma^2+\sigma_e^2 \end{pmatrix}$$

Prove that $cov(\underline{C}\hat{\beta}, \hat{\underline{b}}) = \underline{0}$

$$cov(\underline{C}\hat{\beta}, \hat{\underline{b}}) = cov(\underline{C}(\underline{X}^T\underline{V}^{-1}\underline{X})^- \underline{X}^T\underline{V}^{-1}\underline{y}, \underline{D}\underline{Z}^T\underline{P}\underline{y})$$

$$= C(X^TV^{-1}X)^- X^T V^{-1} \underbrace{cov(\underline{y}, \underline{y})}_{=V} \underline{P}\underline{Z}\underline{D} = C(X^TV^{-1}X)^- X^T P Z D$$

$$= C(X^TV^{-1}X)^- X^T V^{-1} Z D - C(X^TV^{-1}X)^- X^T V^{-1} X (X^TV^{-1}X)^- X^T V^{-1} Z D = (*)$$

Since $C\beta$ is estimable, $C = AX$ for some A

therefore $(*) = C(X^TV^{-1}X)^- X^T V^{-1} Z D - A\underbrace{X(X^TV^{-1}X)^- X^T V^{-1}}_{\text{a (non-orthogonal) projection}} X (X^TV^{-1}X)^- X^T V^{-1} Z D$

a (non-orthogonal) projection matrix onto $C(X)$, so it projects X onto itself

$$\Rightarrow (*) = C(X^TV^{-1}X)^- X^T V^{-1} Z D - \underbrace{AX}_{=C}(X^TV^{-1}X)^- X^T V^{-1} Z D = 0$$

Additional Problem #2:

Prove $\text{var}(\hat{\psi}) = (c^T \Sigma c)\left[a^T (X^T X)^{-1} a\right]$ where $\hat{\psi} = \underline{a}^T \underline{\hat{B}} \underline{c}$

no $\underline{\hat{B}} = (X^T X)^{-1} X^T Y = \left[(X^T X)^{-1} X^T \underline{y}_1, \ldots, (X^T X)^{-1} X^T \underline{y}_t\right] = (\hat{\beta}_1, \ldots, \hat{\beta}_t)$

where $\underline{y}_i$ is i^{th} column of Y.

$$\hat{\psi} = \underline{a}^T (\underline{\hat{B}} \underline{c}) = \underline{a}^T \sum_{j=1}^{t} c_j \hat{\beta}_t$$

$$\Rightarrow \text{var}(\hat{\psi}) = \underline{a}^T \text{var}\left(\sum_{j=1}^{t} c_j \hat{\beta}_t\right) \underline{a}$$

$$= \underline{a}^T \left[\sum_{j=1}^{t} \sum_{k=1}^{t} c_j c_k \text{cov}(\hat{\beta}_j, \hat{\beta}_k)\right] \underline{a}$$

$$= \underline{a}^T \left[\sum_{j}^{t} \sum_{k}^{t} c_j c_k (X^T X)^{-1} X^T \text{cov}(\underline{y}_j, \underline{y}_k) X (X^T X)^{-1}\right] \underline{a}$$

$$= \underline{a}^T (X^T X)^{-1} X^T \left[\sum_j \sum_k c_j c_k \underbrace{\text{cov}(\underline{y}_j, \underline{y}_k)}_{= \sigma_{jk} I}\right] X (X^T X)^{-1} \underline{a}$$

$\leftarrow j,k^{th}$ element of Σ

$$= \underline{a}^T (X^T X)^{-1} X^T \left[\sum_j \sum_k c_j c_k \sigma_{jk}\right] I X (X^T X)^{-1} a$$

$$= \underline{a}^T (X^T X)^{-1} X^T \underbrace{\underline{c}^T \Sigma \underline{c}}_{\text{a scalar}} X (X^T X)^{-1} a = \underline{c}^T \Sigma \underline{c} \; a^T (X^T X)^{-1} \cancel{X^T X} \cancel{(X^T X)^{-1}}$$

3. Refer to p. 70 of the class notes. Show that $PVP = P$

Proof:

$$
\begin{aligned}
PVP &= \left(V^{-1} - V^{-1}X(X^TV^{-1}X)^-X^TV^{-1}\right) V \left(V^{-1} - V^{-1}X(X^TV^{-1}X)^-X^TV^{-1}\right) \\
&= \left(I - V^{-1}X(X^TV^{-1}X)^-X^T\right)\left(V^{-1} - V^{-1}X(X^TV^{-1}X)^-X^TV^{-1}\right) \\
&= V^{-1} - V^{-1}X(X^TV^{-1}X)^-X^TV^{-1} - V^{-1}X(X^TV^{-1}X)^-X^TV^{-1} \\
&\quad + V^{-1}X(X^TV^{-1}X)^-X^TV^{-1}X(X^TV^{-1}X)^-X^TV^{-1} \\
&= V^{-1} - 2V^{-1}X(X^TV^{-1}X)^-X^TV^{-1} + V^{-1}X(X^TV^{-1}X)^-X^TV^{-1} \\
&= V^{-1} - V^{-1}X(X^TV^{-1}X)^-X^TV^{-1} \\
&= P
\end{aligned}
$$

4. Refer to p. 64 of the class notes. Complete the proof of the theorem on this page by showing that $E[\{y_0 - m_{blp}(Y)\}\{m_{blp}(Y) - t(Y)\}] = 0$

Proof:

$$
\begin{aligned}
m_{blp}(Y) &= \mu_{y_0} + (\gamma^*)^T(Y - \mu_Y) = \mu_{y_0} + V_{Yy_0}^T V_{YY}^{-1}(Y - \mu_Y) \\
t(Y) &= \gamma_0 + Y^T\gamma = \gamma_0 + \gamma^T Y
\end{aligned}
$$

$$
\begin{aligned}
&E\left[\{y_0 - m_{blp}(Y)\}\{m_{blp}(Y) - t(Y)\}\right] \\
&= E\left[\{y_0 - \mu_{y_0} - V_{Yy_0}^T V_{YY}^{-1}(Y - \mu_Y)\}\{\mu_{y_0} + V_{Yy_0}^T V_{YY}^{-1}(Y - \mu_Y) - \gamma_0 - \gamma^T Y\}\right] \\
&= E\left[\{(y_0 - \mu_{y_0}) - V_{Yy_0}^T V_{YY}^{-1}(Y - \mu_Y)\}\{(\mu_{y_0} - \gamma_0 - \gamma^T\mu_Y) + V_{Yy_0}^T V_{YY}^{-1}(Y - \mu_Y) - \gamma^T(Y - \mu_Y)\}\right] \\
&= E\left[(y_0 - \mu_{y_0})(\mu_{y_0} - \gamma_0 - \gamma^T\mu_Y)\right] - E\left[V_{Yy_0}^T V_{YY}^{T}(Y - \mu_Y)(\mu_{y_0} - \gamma_0 - \gamma^T\mu_Y)\right] \\
&\quad + E\left[(y_0 - \mu_{y_0})(V_{Yy_0}^T V_{YY}^{-1} - \gamma^T)(Y - \mu_Y)\right] - E\left[V_{Yy_0}^T V_{YY}^{-1}(Y - \mu_Y)(V_{Yy_0}^T V_{YY}^{-1} - \gamma^T)(Y - \mu_Y)\right] \\
&= 0 - 0 + (V_{Yy_0}^T V_{YY}^{-1} - \gamma^T) E\left[(y_0 - \mu_{y_0})(Y - \mu_Y)\right] - E\left[V_{Yy_0}^T V_{YY}^{-1}(Y - \mu_Y)(Y - \mu_Y)^T(V_{Yy_0}^T V_{YY}^{-1} - \gamma^T)^T\right] \\
&= 0 + (V_{Yy_0}^T V_{YY}^{-1} - \gamma^T) V_{Yy_0} - V_{Yy_0}^T V_{YY}^{-1} E\left[(Y - \mu_Y)(Y - \mu_Y)^T\right](V_{Yy_0}^T V_{YY}^{-1} - \gamma^T)^T \\
&= 0 + V_{Yy_0}^T V_{YY}^{-1} V_{Yy_0} - \gamma^T V_{Yy_0} - V_{Yy_0}^T V_{YY}^{-1} V_{YY}(V_{YY}^{-1} V_{Yy_0} - \gamma) \\
&= V_{Yy_0}^T\gamma - \gamma^T V_{Yy_0} \\
&= 0
\end{aligned}
$$

a scalar