Finish Chapter 6 — CI, multiple comparison, contrasts

Start 7 — SS & lack of fit

$$y^T \underbrace{(P_X - P_0)} y = \| (P_X - P_0) y \|^2 = \| \underbrace{P_X y} - \underbrace{P_0 y} \|^2$$

M 6.6
R 3.7

$\frac{z}{\sqrt{\chi^2/\nu}} \sim t(\nu)$

♣ Confidence interval

$\lambda^T\hat{\beta} \sim N(\lambda^T\beta,\ \sigma^2 \lambda^T(X^TX)^-\lambda)$

- Recall for $\lambda^T\beta$ estimable,

$$t = \frac{(\boldsymbol{\lambda}^T\hat{\boldsymbol{\beta}} - \boldsymbol{\lambda}^T\boldsymbol{\beta})}{\hat{\sigma}\sqrt{\boldsymbol{\lambda}(\mathbf{X}^T\mathbf{X})^-\boldsymbol{\lambda}}} \sim t(n - \underbrace{r}_{= r(X)}),$$

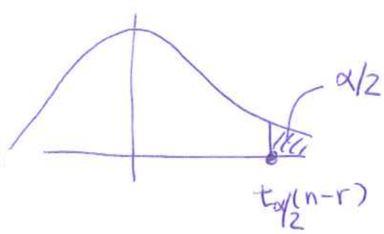
$\frac{SSE}{\sigma^2} \sim \chi^2(n-r)$

where $\hat{\sigma}^2 = MSE = SSE/(n-r)$. Then

$$\mathrm{P}\left(\frac{|\boldsymbol{\lambda}^T\hat{\boldsymbol{\beta}} - \boldsymbol{\lambda}^T\boldsymbol{\beta}|}{\hat{\sigma}\sqrt{\boldsymbol{\lambda}^T(\mathbf{X}^T\mathbf{X})^-\boldsymbol{\lambda}}} \le t_{\alpha/2}(n-r)\right) = 1 - \alpha.$$

We can convert this into an interval, that is,

$$\boxed{\boldsymbol{\lambda}^T\boldsymbol{\beta}} \in \left(\boldsymbol{\lambda}^T\hat{\boldsymbol{\beta}} \pm t_{\alpha/2}(n-r)\hat{\sigma}\sqrt{\boldsymbol{\lambda}^T(\mathbf{X}^T\mathbf{X})^-\boldsymbol{\lambda}}\right)$$



$$\frac{|\lambda^T\hat{\beta} - \lambda^T\beta|}{\hat{\sigma}\sqrt{\lambda^T(X^TX)^-\lambda}} \le t_{\alpha/2}(n-r)$$

$-\lambda^T\hat{\beta}$

$$-t_{\alpha/2}(n-r)\,\hat{\sigma}\sqrt{\lambda^T(X^TX)^-\lambda} \le \lambda^T\hat{\beta} - \lambda^T\beta \le t_{\alpha/2}(n-r)\,\hat{\sigma}\sqrt{\lambda^T(X^TX)^-\lambda}$$

$$\lambda^T\hat{\beta} - t_{\alpha/2}(n-r)\,\hat{\sigma}\sqrt{\lambda^T(X^TX)^-\lambda} \le \lambda^T\beta \le \lambda^T\hat{\beta} + t_{\alpha/2}(n-r)\,\hat{\sigma}\sqrt{\lambda^T(X^TX)^-\lambda}$$

♣ Simultaneous confidence interval

- Recall for $\Lambda^T\beta$ estimable, $\boxed{\Lambda : p \times s}$

$$\underbrace{\Lambda^T\hat{\beta}}_{\equiv\hat{\tau}} \sim N_s(\underbrace{\Lambda^T\beta}_{\equiv\tau},\ \sigma^2\underbrace{\Lambda^T(\mathbf{X}^T\mathbf{X})^{-}\Lambda}_{s \times s} \equiv \sigma^2\mathbf{H}),$$

$\boxed{s \times 1}$ $s \times 1$

$\cancel{MSE}$ $\frac{SSE}{\sigma^2} \sim \chi^2(n-r)$

∗∗ $(1-\alpha)100\%$ confidence region for $\Lambda^T\beta$ is

$$\left\{\mathbf{d} \in \mathbb{R}^s \mid \frac{(\Lambda^T\hat{\beta}-\mathbf{d})^T\mathbf{H}^{-1}(\Lambda^T\hat{\beta}-\mathbf{d})/s}{MSE} \le F_{1-\alpha}(s, n-r)\right\}.$$

∗∗ $(1-\alpha)100\%$ confidence interval for each $\boldsymbol{\lambda}_j^T\beta$, $j = 1, \ldots, s$ is

$$\tau_j \in \left(\hat{\tau}_j \pm t_{\alpha/2}(n-r)\hat{\sigma}\sqrt{\mathbf{H}_{jj}}\right),$$

where $\mathbf{H}_{jj}$ is the (j,j)-element of $\mathbf{H}$.

Economic Dataset Example (contd)

- 95% CI for each β_j

```
> confint(g)
                   2.5 %       97.5 %
(Intercept) 13.753330728 43.378842354
pop15       -0.752517542 -0.169868752
pop75       -3.873977955  0.490982602
dpi         -0.002212248  0.001538444
ddpi         0.014533628  0.804856227
```

β_{pop15} β_{pop75} β_{dpi} β_{ddpi}

$H_0: \beta_{pop15} = 0$ vs $\neq 0$

$0 \notin CI \Rightarrow$ reject H_0

$0 \in CI \Rightarrow$ fail to reject H_0 at 5% significance level

Economic Dataset Example (contd)

- 95% Confidence Region for $(\beta_{pop15}, \beta_{pop75})$

```
> library(ellipse)
> plot(ellipse(g, c(2,3)), type="l", xlim=c(-1,0))
> points(0,0)
> points(coef(g)[2], coef(g)[3], pch=18)
> abline(v=confint(g)[2,], lty=2)
> abline(h=confint(g)[3,], lty=2)
```

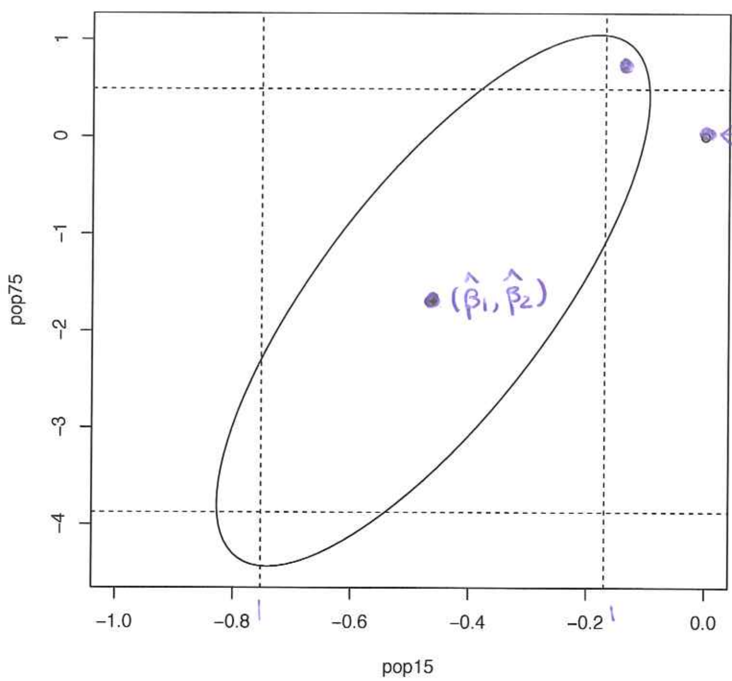
Economic Dataset Example (contd)

- 95% Confidence Region for $(\beta_{pop15}, \beta_{pop75})$

$H_0: \beta_{pop15} = 0$ & $\beta_{pop75} = 0$ vs

$H_1: \beta_{pop15} \neq 0$ or $\beta_{pop75} \neq 0$

$\Rightarrow$ reject H_0



$\leftarrow (0, 0)$

Extreme Case

- Reject both individual hypotheses
- fail to reject the joint

• Multiple Comparison Procedures: comparing fixed-effect means in ANOVA procedures

Factor A

• Ex: Consider the one-way ANOVA;

$$y_{ij} = \mu + \alpha_i + e_{ij}, \quad i = 1, \ldots, a \ , j = 1, \ldots, n_i.$$

$*$ Suppose we reject $H_0 : y_{ij} = \mu + e_{ij}$ (or $H_0 : \alpha_1 = \ldots = \alpha_a$).

$*$ Which of $\binom{a}{2}$ pairs of means were significantly different?

$\Rightarrow$ **Multiple comparison!**

† Def The **comparisonwise Type I error rate** is defined as the ratio of the number of comparisons incorrectly declared significant to the total number of nonsignificant comparisons tested.

ave. Type I error rate for a comparison

† Def The **experimentwise Type I error rate** is defined as the ratio of the number of experiments with one or more comparison incorrectly declared significant to the total number of experiments with at least two treatment means.

Ex1	Ex2	. . . .
1	1	
0	0	

$P(A \cup B) = P(A) + P(B) - \underbrace{P(A \cap B)}_{\geq 0} \leq P(A) + P(B)$

$E_j = 1 \quad E_j = 1$

$*$ **Bonferroni inequalities:** E_j denotes an error of incorrectly declaring significant to test j.

$$\text{P(at least one error)} = \text{P}(\cup E_j) \leq \sum_j \text{P}(E_j)$$

$$\text{P(all correct)} = 1 - \text{P(at least one error)} \geq 1 - \sum_j \text{P}(E_j)$$

Fisher's protected LSD

① do an α-level F test

② If F is significant, then do the LSD test

$\Rightarrow$ Experimentwise ~~error of~~ error rate is α

$\binom{a}{2}$

* **Fisher's least significant difference (LSD):** Use the t-based CI for each pair. α

$\bar{y}_i \sim N(\alpha_i, \frac{\sigma^2}{n})$

$\bar{y}_{i'} \sim N(\alpha_{i'}, \frac{\sigma^2}{n})$

* The LSD CI of $\alpha_i - \alpha_{i'}$ is

** for the balanced one-way ANOVA,

$$(\bar{y}_i - \bar{y}_{i'}) \pm t_{\alpha/2}(n-r)\frac{\hat{\sigma}\sqrt{2}}{\sqrt{n}}.$$

** for the unbalanced one-way ANOVA,

$$(\bar{y}_i - \bar{y}_{i'}) \pm t_{\alpha/2}(n-r)\hat{\sigma}\sqrt{\frac{1}{n_i}+\frac{1}{n_{i'}}}.$$

* Can we compare all pairs of treatment means based on the usual t tests? Any problem with this?

no significant difference when the null is true

$j = 1, \ldots, s = \binom{a}{2}$

$Pr(\text{all correct}) = Pr(\bigcap_j \text{~~pair j of CI~~} \; 0 \in CI_j \text{ for pair } j)$

$\geq 1 - \sum P(\text{error for pair } j)$

$= 1 - \sum \alpha$

$= 1 - s \times \alpha$

$= 1 - 0.3 = \underline{0.7}$

$a = 4$

$s = \binom{4}{2} = 6$

$\alpha = 0.05$

Experimentwise error rate

$Pr(\text{making at least one mistake}) = 1 - P(\text{all correct})$

$\leq 1 - (1 - s\alpha) = \underline{s\alpha} = 0.30$

$*$ Bonferroini's solution to control the experimentwise Type I error is to replace $t_{\alpha/2}(n-r)$ with $t_{\alpha/(2s)}(n-r)$.

$*$ The Bonferroni CI of $\alpha_i - \alpha_{i'}$ is

$\star\star$ for the balanced one-way ANOVA,

$$(\bar{y}_i - \bar{y}_{i'}) \pm t_{\alpha/(2s)}(n-r)\frac{\hat{\sigma}}{\sqrt{n}}.$$

$\star\star$ for the unbalanced one-way ANOVA,

$$(\bar{y}_i - \bar{y}_{i'}) \pm t_{\alpha/(2s)}(n-r)\hat{\sigma}\sqrt{\frac{1}{n_i} + \frac{1}{n_{i'}}}.$$

Pr(making at least one error) = experimentwise error rate

$= 1 - P(\text{all correct})$

$< 1 - (1 - \cancel{s} \times \frac{\alpha}{\cancel{s}}) = \alpha$

$t = \frac{z}{\sqrt{u/\nu}}$

- Tukey's procedure (Tukey's honest significant difference - HSD)

$*$ Let $z_i \stackrel{iid}{\sim} N(0,1)$, $i = 1, \ldots, a$ and $u \sim \chi^2(\nu)$ independent of z_i's.

M

$*$ Define the studentized range statistic, w.

$$w = \frac{\max_i z_i - \min_i z_i}{\sqrt{u/\nu}} = \frac{\max_{i,i'}(z_i - z_{i'})}{\sqrt{u/\nu}}.$$

range when all z_i's have the same mean

$*$Tukey derived the distribution of w.

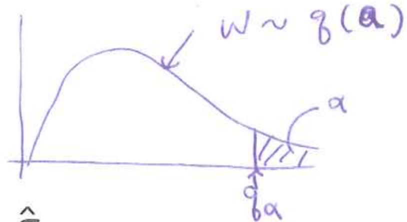
$\Rightarrow$ We can find $P(\ w > q^{\star}_{\alpha}(a, \nu)\ ) = \alpha$.

$t_{\alpha/2}$

$t_{\alpha/(2S)}$

• Application of the distribution of the studentized range statistic to the *balanced one-way* ANOVA.

$*$ Let $z_i = \alpha_i - \hat{\alpha}_i \sim \mathrm{N}(0, \sigma^2)$ and $u = MSE \sim \chi^2(N-a)$ independent of z_i.



$*$ The Tukey CI of $\alpha_i - \alpha_{i'}$ is

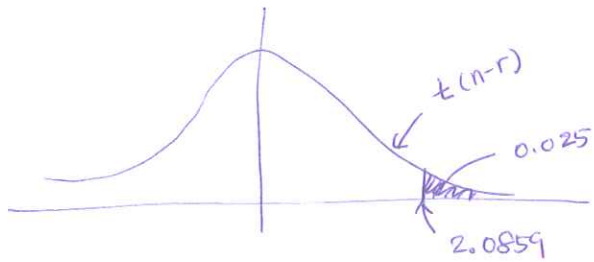
$$(\bar{y}_i - \bar{y}_{i'}) \pm q_\alpha(a, n-r)\frac{\hat{\sigma}}{\sqrt{n}}.$$

← # of replates

• Adjustment for *unbalanced* one-way ANOVA.

$$(\bar{y}_i - \bar{y}_{i'}) \pm q_\alpha(a, n-r)\frac{\hat{\sigma}}{\sqrt{2}}\sqrt{\frac{1}{n_i} + \frac{1}{n_{i'}}}.$$

• **Caution!** when the sample sizes are very unequal, Tukey's method may become too conservative.



\# of levels

$\binom{4}{2} = 6$ pairs

$\bar{y}_{1\cdot} = 61 = \hat{\mu} + \hat{\alpha}_1$

$\bar{y}_{2\cdot} = 66 = \hat{\mu} + \hat{\alpha}_2$

$\Rightarrow$

Coagulation Example– revisit

CI for $(\alpha_1 - \alpha_2)$

Fisher LSD $(61-66) \pm 2.085 \times \sqrt{\frac{1}{4}+\frac{1}{6}} \times \sqrt{5.6}$ $\hat{\sigma}$

Bonferonni's $(61-66) \pm 2.927 \times \sqrt{\frac{1}{4}+\frac{1}{6}} \times \sqrt{5.6}$

```
> qt(0.975, 20)
[1] 2.085963
> qt(1-0.05/12, 20)
[1] 2.927119

> TukeyHSD(aov(coag ~ diet, coagulation))
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = coag ~ diet, data = coagulation)

$diet
    diff         lwr       upr     p adj
B-A    5   0.7245544  9.275446 0.0183283
C-A    7   2.7245544 11.275446 0.0009577
D-A    0  -4.0560438  4.056044 1.0000000
C-B    2  -1.8240748  5.824075 0.4766005
D-B   -5  -8.5770944 -1.422906 0.0044114
D-C   -7 -10.5770944 -3.422906 0.0001268

>
```

- More methods for multiple comparison

∗ Scheffé : The number of comparison pairs does not need to be specified (check p 144 of M). Construct a confidence interval for any linear combination.

∗ Newman-Keuls, Duncan's multiple range ...

chapter 5 of Ronald's book
(table 5.1 p. 120)

$\sum c_i \alpha_i = \sum c_i(\mu + \alpha_i)$

$\alpha_1 - \frac{\alpha_2+\alpha_3+\alpha_4}{3}$

$= (\mu + \alpha_1) - \frac{(\mu+\alpha_2) + \cdots + (\mu+\alpha_4)}{3}$

$\Lambda = \begin{bmatrix} 0 & 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & \cdots & 0 \end{bmatrix}$

$\alpha_1 - \alpha_2$

$= \underbrace{(\mu+\alpha_1)}_{E(y_1)} - \underbrace{(\mu+\alpha_2)}_{E(y_2)}$

- Contrast – One-Way ANOVA

$$y_{ij} = \mu + \alpha_i + e_{ij}, i = 1, \ldots, a, \quad j = 1, \ldots, n_i.$$

* A contrast is a function $\cancel{\sum_i c_i \alpha_i = 0}$. $\underbrace{\sum c_i \alpha_i = 0}_{i}$, $\sum c_i = 0$

** $\alpha_1 - \alpha_2$. $\lambda = [0 \ 1 \ -1 \ 0 \cdots 0]^T$ ($c_1 \ c_2 \ c_3 \cdots c_a$) $H_0: \alpha_1 - \alpha_2 = 0$

** $\alpha_1 - (\alpha_2 + \alpha_3 + \alpha_4)/3$. $H_0: \alpha_1 - \frac{(\alpha_2+\alpha_3+\alpha_4)}{3} = 0$

$\lambda = [0 \ 1 \ -\frac{1}{3} \ -\frac{1}{3} \ -\frac{1}{3} \ 0 \cdots 0]$

* We can easily check this is estimable.

* How to construct a test statistic? First,

$$\sum_{i=1}^{a} c_i \hat{\alpha}_i = \sum_{i=1}^{a} c_i(\hat{\mu} + \hat{\alpha}_i) = \sum_{i=1}^{a} c_i \bar{y}_{i\cdot} \sim \mathrm{N}(\sum_{i=1}^{a} c_i \alpha_i, \sigma^2 \sum_{i=1}^{a} \frac{c_i^2}{n_i}).$$

$\sum_{i=1}^{a} c_i \hat{\mu} = \hat{\mu} \sum_{i=1}^{a} c_i = 0$ $\qquad = 0$

$\Rightarrow$ Reject H_0 if

$$\frac{|\sum_{i=1}^{a} c_i \bar{y}_{i\cdot}|}{\sqrt{MSE \sum_{i=1}^{a} \frac{c_i^2}{n_i}}} > t_{\alpha/2}(N - r).$$

- Two contrasts, $\mathbf{c}_1$ and $\mathbf{c}_2$ are orthogonal if $\sum_{i=1}^{a} c_{1i}c_{2i}/n_i = 0$.

  * It becomes simpler for balanced designs: $\sum_{i=1}^{a} c_{1i}c_{2i} = 0$

- Contrast when we have more than one factors:

  * For main effects: Ignore the fact that other treatment exits and do exactly the same as in the one-way ANOVA.

  * For interactions: Become complicate. Read Ronald 7.2.1 for interaction contrasts in the two-way ANOVA.

$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk}$

AMS 256
Monahan Chapter 7: Further Topics in Testing
Ronald Chapter 3: Testing Hypothesis

Spring 2016

M& 7.3

R : 3.6

- **Sequentially** break a sum of squares (SS) into independent components.

- Called Type I sequential sum of squares.

- **Caution!** We follow the notation in M (little different from our previous notation).

- Recall the one-way ANOVA example.

Consider the model; $y_{ij} = \mu + \alpha_i + e_{ij}$, $i = 1, \dots, a$ and $b = 1, \dots, n_i$.

We split the sum of squares $\mathbf{y}^T\mathbf{y}$ into two parts to test that the effects α_i are all equal.

$$
\begin{aligned}
\mathbf{y}^T\mathbf{y} &= \overbrace{\mathbf{y}^T\mathbf{P}_x\mathbf{y}} + \overbrace{\mathbf{y}^T(I - \mathbf{P}_x)\mathbf{y}}^{SSE} \\
&= \underbrace{\mathbf{y}^T\mathbf{P}_1\mathbf{y}} + \underbrace{\mathbf{y}^T(\mathbf{P}_x - \mathbf{P}_1)\mathbf{y}} + \underbrace{\mathbf{y}^T(I - \mathbf{P}_x)\mathbf{y}}.
\end{aligned}
$$

X_1^*

X_0^*

X_0 $\qquad X_1$

$$
X = \begin{bmatrix} 1_{n_1} & 1_{n_1} & \cdots & 0_{n_1} \\ 1_{n_2} & 0_{n_2} & \cdots & 0_{n_2} \\ \vdots & \vdots & & \vdots \\ 1_{n_a} & 0_{n_a} & & 1_{n_a} \end{bmatrix}
$$

1_N

$N = \sum n_i$

- By Cochran's theorem,

$$\frac{1}{\sigma^2}\mathbf{y}^T\mathbf{P}_1\mathbf{y} \sim \chi^2(r(\mathbf{P}_1), \frac{(\mathbf{X}\beta)^T\mathbf{P}_1\mathbf{X}\beta}{2\sigma^2})$$

$$\frac{1}{\sigma^2}\mathbf{y}^T(\mathbf{P}_x - \mathbf{P}_1)\mathbf{y} \sim \chi^2(r(\mathbf{P}_x - \mathbf{P}_1), \frac{(\mathbf{X}\beta)^T(\mathbf{P}_x - \mathbf{P}_1)\mathbf{X}\beta}{2\sigma^2})$$

$$\frac{1}{\sigma^2}\mathbf{y}^T(I - \mathbf{P}_x)\mathbf{y} \sim \chi^2(r(I - \mathbf{P}_x), \underbrace{\frac{(\mathbf{X}\beta)^T(I - \mathbf{P}_x)\mathbf{X}\beta}{2\sigma^2}}_{=0})$$

Furthermore, $\frac{1}{\sigma^2}\mathbf{y}^T\mathbf{P}_1\mathbf{y}$, $\frac{1}{\sigma^2}\mathbf{y}^T(\mathbf{P}_x - \mathbf{P}_1)\mathbf{y}$ and $\frac{1}{\sigma^2}\mathbf{y}^T(I - \mathbf{P}_x)\mathbf{y}$ are independent.

To test $\alpha_1 = \cdots = \alpha_a = 0$

$$F = \frac{Y^T(P_x - P_1)Y/(a-1)}{Y^T(I-P_x)Y/(n-r)} \sim F(a-1, n-r)$$

- Following the notation in M, rewrite what we have.
- Let
  ** $\mathbf{X} = [\mathbf{X}_0 \mid \mathbf{X}_1]$.

  ** $\mathbf{X}_0^\star = \mathbf{X}_0$ and $\mathbf{X}_1^\star = [\mathbf{X}_0 \mid \mathbf{X}_1] = \mathbf{X}$.

  ** $\mathbf{P}_{x_0^\star}$ and $\mathbf{P}_{x_1^\star}$: orthogonal projection matrices onto $C(\mathbf{X}_0^\star)$ and $C(\mathbf{X}_1^\star)$, respectively.

  ** $\mathbf{P}_{x_1^\star} - \mathbf{P}_{x_0^\star}$: orthogonal projection matrices onto $C(\mathbf{P}_{x_1^\star} - \mathbf{P}_{x_0^\star}) = C(\mathbf{X}_0^\star)^{\perp}_{X_1^\star}$.

$$\Rightarrow \quad \mathbf{y}^T \mathbf{P}_x \mathbf{y} = \mathbf{y}^T \mathbf{P}_{x_1^\star} \mathbf{y} = \mathbf{y}^T \mathbf{P}_{x_0^\star} \mathbf{y} + \mathbf{y}^T (\mathbf{P}_{x_1^\star} - \mathbf{P}_{x_0^\star}) \mathbf{y}$$

• $y = X\beta + e$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix}$$

$\underbrace{\quad}_{X_0}\ \underbrace{\quad\quad}_{X_1}$

$\underbrace{\quad}_{X_0^*}$

$\underbrace{\quad\quad\quad}_{X_1^*}$

$$P_{X_1^*} = X_1^* (X_1^{*T} X_1^*)^- X_1^{*T}$$

vs $y = X_0\beta_0 + e$

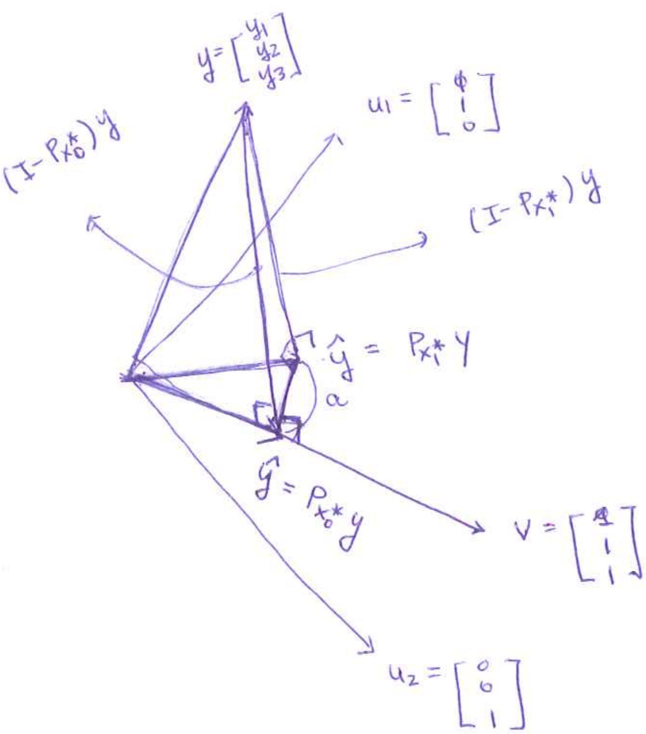
$$P_{X_0^*} = X_0^* (X_0^{*T} X_0^*)^- X_0^{*T}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \mu + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix}$$

basis vectors for $\mathcal{C}(X_1^*)$: $u_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $u_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

basis vector for $\mathcal{C}(X_0^*)$: $v = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = u_1 + u_2$

$\mathcal{C}(X_0^*) \subset \mathcal{C}(X_1^*)$



$$\overbrace{\|(I - P_{X_0^*})y\|^2}^{\text{SSE under the reduced model}} = a^2 + \overbrace{\|(I - P_{X_1^*})y\|^2}^{\text{SSE under the full model}}$$

$$\Rightarrow a^2 = y^T(I - P_{X_0^*})y - y^T(I - P_{X_1^*})y$$

$$= y^T(P_{X_1^*} - P_{X_0^*})y$$

$$\Rightarrow a^2 = y^T(P_{X_1^*} - P_{X_0^*})y$$

: reduction in SSE

by having (α_1, α_2)

or by having X_1

$$\| P_{X_1^*} y \|^2 = \| (P_{X_1^*} - P_{X_0^*}) y \|^2 + \| P_{X_0^*} y \|^2$$

$$\Leftrightarrow \underbrace{y^T P_{X_1^*} y}_{\text{SS for regressing on } (\mu, \alpha)} = \underbrace{y^T (P_{X_1^*} - P_{X_0^*}) y}_{\text{addition in SSR by having } (\alpha_1, \alpha_2) \text{ or by having } X_1} + \underbrace{y^T P_{X_0^*} y}_{\text{SS for regressing on } \mu}$$

$$\underbrace{y^T y}_{\text{total}} = \underbrace{y^T P_{X_1^*} y}_{\text{explained}} + \underbrace{y^T (I - P_{X_1^*}) y}_{\text{unexplained}}$$