L21 ANOVA table for a hypothesis

1. SS table for H_0

   (1) Model, H_0 and reduced model
   Consider linear model $Y \sim N(X\beta,\, \sigma^2\Sigma)$ with $\mathrm{rank}(X) = r$ and consistent $H_0:\ H\beta = b$ with $H\beta_* = b$.

$$\begin{aligned} H\beta = b &\iff H\beta = b = H\beta_* = HH^+H\beta_* = HH^+b \iff H(\beta - H^+b) = 0 \\ &\iff \beta - H^+b \in \mathcal{N}(H) = \mathcal{R}(I - H^+H) = \{(I - H^+H)\gamma:\ \gamma \in R^p\} \\ &\iff \beta - H^+b = (I - H^+H)\gamma \iff \beta = H^+b + (I - H^+H)\gamma. \end{aligned}$$

   The model reduced by H_0, $Y \sim N\left(X[H^+b + (I - H^+H)\gamma,\, \sigma^2\Sigma\right)$, is

$$Y - XH^+b \sim N\left(X(I - H^+H)\gamma,\, \sigma^2\Sigma\right).$$

   (2) SSE_H and SSE
   For the original model,
   $\mathrm{SSE} = (\Sigma^{-1/2}Y)'[I - (\Sigma^{-1/2}X)(\Sigma^{-1/2}X)^+](\Sigma^{-1/2}Y)$ with DF: $n - r$.
   For the reduced model, let $Y_* = \Sigma^{-1/2}(Y - XH^+b)$ and $X_* = \Sigma^{-1/2}X$. Then

$$SSE_H = Y_*'\left\{I - [X_*(I - H^+H)][X_*(I - H^+H)]^+\right\}Y_*$$

   with DF: $n - \mathrm{rank}[(X_*(I - H^+H)] = n - \mathrm{rank}[X(I - H^+H)] \stackrel{def}{=\!=} n - r_1$. Clearly

$$\begin{aligned} Y_*'(I - X_*X_*^+)Y_* &= (\Sigma^{-1/2}Y - X_*H^+b)'(I - X_*X_*^+)(\Sigma^{-1/2}Y - X_*H^+b) \\ &= (\Sigma^{-1/2}Y)'[I - (\Sigma^{-1/2}X)(\Sigma^{-1/2}X)^+](\Sigma^{-1/2}Y) \\ &= \mathrm{SSE}. \end{aligned}$$

   (3) SSD and SS table
   Define $SSD = SSE_H - SSE$. Then

$$\mathrm{SSD} = Y_*'\left\{X_*X_*^+ - [X_*(I - H^+H)][X_*(I - H^+H)]^+\right\}Y_*$$

   with DF: $\mathrm{rank}(X_*) - \mathrm{rank}[X_*(I - H^+H)] = \mathrm{rank}(X) - \mathrm{rank}[X(I - H^+H)] = r - r_1$.
   Thus for the model and H_0 we obtain SS table

Source	SS	DF
Difference	$\mathrm{SSD} = Y_*'\{X_*X_*^+ - [X_*(I - H^+H)][X_*(I - H^+H)]^+\}Y_*$	$r - r_1$
Error	$\mathrm{SSE} = Y_*'(I - X_*X_*^+)Y_*$	$n - r$
Hypothesis	$\mathrm{SSE}_H = Y_*'\{I - [X_*(I - H^+H)][X_*(I - H^+H)]^+\}Y_*$	$n - r_1$

   Here $Y_* = \Sigma^{-1/2}(Y - XH^+b)$, $X_* = \Sigma^{-1/2}X$ and $Y \sim N(X\beta,\, \sigma^2\Sigma)$.

2. ANOVA table for H_0

   (1) Total variations in Y_*
   SSE_H is the variation in Y_* unexplained by H_0. In the problem of testing H_0, this SS is treated as the total variation in Y_*.

   (2) Variation in Y_* unexplained by the model
   SSE is the variation in Y_* unexplained by the model which is part of that in (1).

(3) Variation in Y_* explained by the model.
SSD is the variation in Y_* explained by the model.

Thus the SS table gives the breakdown of the total variation in Y_* into two parts, variation explained, and unexplained by the model.

3. Distributions in the table

(1) $\frac{SSE}{\sigma^2} \sim \chi^2(n-r)$.
Note that $\frac{SSE}{\sigma^2} = Y_*' A Y_*$ where

$$\begin{aligned} Y \sim N(X\beta,\, \sigma^2\Sigma) &\implies Y - XH^+b \sim (X\beta - XH^+b,\, \sigma^2\Sigma) \\ &\implies Y_* \sim N(X_*(\beta - H^+b),\, \sigma^2 I_n). \end{aligned}$$

and $A = \frac{I - X_* X_*^+}{\sigma^2}$. But $A\sigma^2 I A = A = A'$, $[X_*(\beta - H^+b)]' A [X_*(\beta - H^+b)] = 0$ and $\mathrm{tr}(A\sigma^2 I) = n - r$. So $\frac{SSE}{\sigma^2} \sim \chi^2(n-r)$.

(2) $\frac{SSD}{\sigma^2} \overset{H_0}{\sim} \chi^2(r - r_1)$ and $\frac{SSE_H}{\sigma^2} \overset{H_0}{\sim} \chi^2(n - r_1)$.
Note that $\frac{SSD}{\sigma^2} = Y_*' B Y_*$ and $\frac{SSE_H}{\sigma^2} = Y_*' C Y_*$ where $Y_* \overset{H_0}{\sim} N(X_*(I - H^+H)\gamma,\, \sigma^2 I_n)$, $B = \frac{X_* X_*^+ - [X_*(I-H^+H)][X_*(I-H^+H)]}{\sigma^2}$ and $C = \frac{I - [X_*(I-H^+H)][X_*(I-H^+H)]^+}{\sigma^2}$.
But $B\sigma^2 I_n B = B = B'$, $C\sigma^2 I_n C = C = C'$,$[X_*(I - H^+H)\gamma]' B [X_*(I - H^+H)\gamma] = 0$, $[X_*(I - H^+H)\gamma]' C [X_*(I - H^+H)\gamma] = 0$, $\mathrm{tr}(B\sigma^2 I) = r - r_1$ and $\mathrm{tr}(C\sigma^2 I_n) = n - r_1$.
Therefore $\frac{SSD}{\sigma^2} \overset{H_0}{\sim} \chi^2(r - r_1)$ and $\frac{SSE_H}{\sigma^2} \overset{H_0}{\sim} \chi^2(n - r_1)$.

(3) Unbiased estimators for σ^2
$MSE = \frac{SSE}{n-r}$ is UE for σ^2. $MSD = \frac{SSD}{r-r_1}$ and $MSE_H = \frac{SSE_H}{n-r_1}$ are UEs for σ^2 under H_0.

Proof. $E(MSE) = \frac{\sigma^2}{n-r} \cdot E\left(\frac{SSE}{\sigma^2}\right) = \frac{\sigma^2}{n-r} E[\chi^2(n-r)] = \frac{\sigma^2}{n-r} \cdot (n-r) = \sigma^2$.

$$E(MSD) = \frac{\sigma^2}{r-r_1} \cdot E\left(\frac{SSD}{\sigma^2}\right) \overset{H_0}{=\!=} \frac{\sigma^2}{r-r_1} E[\chi^2(r-r_1)] = \frac{\sigma^2}{r-r_1} \cdot (r-r_1) = \sigma^2.$$

$$E(MSE_H) = \frac{\sigma^2}{n-r_1} \cdot E\left(\frac{SSE_H}{\sigma^2}\right) \overset{H_0}{=\!=} \frac{\sigma^2}{n-r_1} E[\chi^2(n-r_1)] = \frac{\sigma^2}{n-r_1} \cdot (n-r_1) = \sigma^2.$$

(4) $\frac{MSD}{MSE} \overset{H_0}{\sim} F(r - r_1,\, n - r)$.
Recall that $\frac{SSE}{\sigma^2} = Y_*' A Y_* \sim \chi^2(n-r)$ and $\frac{SSD}{\sigma^2} = Y_*' B Y_* \overset{H_0}{\sim} \chi^2(r - r_1)$.
With $\mathrm{Cov}(Y_*) = \sigma^2 I$, $A\sigma^2 I B = 0$. So $\frac{SSE}{\sigma^2}$ and $\frac{SSE}{\sigma^2}$ are independent. Thus $F = \frac{\frac{SSD}{\sigma^2}/(r-r_1)}{\frac{SSE}{\sigma^2}/(n-r)} = \frac{MSD}{MSE} \overset{H_0}{\sim} F(r - r_1,\, n - r)$.

(5) A complete ANOVA table
For the model with H_0 we have

Source	SS	DF	MS	F	$Pr > F$
Difference	SSD	$r - r_1$	MSD	MSD/MSE	$P(F(r - r_1, n - r) > F_{ob})$
Error	SSE	$n - r$	MSE		
Hypothesis	SSE_H	$n - r_1$			

L22 A general α-level F-test

1. α-level LRT

   (1) The problem
   For Model $Y \sim N(X\beta, \sigma^2\Sigma)$ with consistent $H_0 : H\beta = b$ where $\text{rank}(X) = r$ and $\text{rank}[X(I - H^+H)] = r_1$, an ANOVA table has been obtained.

Source	SS	DF	MS	F	$Pr > F$
Difference	SSD	$r - r_1$	MSD	MSD/MSE	$P(F(r - r_1, n - r) > F_{ob})$
Error	SSE	$n - r$	MSE		
Hypothesis	SSE_H	$n - r_1$			

   We now need to develop an α-level LRT on H_0.

   (2) Maximized likelihood function
   For $Y \sim N(X\beta, \sigma^2\Sigma)$ let $\widehat{\beta} \in \text{GLSE}_{\Sigma^{-1}}(\beta)$ and SSE$= \|Y - X\widehat{\beta}\|^2_{\Sigma^{-1}}$. Then

   $$\begin{aligned} L(\beta, \sigma^2\Sigma) &= \frac{1}{(2\pi)^{n/2}|\sigma^2\Sigma|^{1/2}} \exp\left[-\frac{1}{2\sigma^2}(Y - X\beta)'\Sigma^{-1}(Y - X\beta)\right] \le L(\widehat{\beta}, \sigma^2) \\ &= \frac{1}{(2\pi)^{n/2}(\sigma^2)^{n/2}|\Sigma|} \exp\left(-\frac{SSE}{2\sigma^2}\right) \le L\left(\widehat{\beta}, \frac{SSE}{n}\right) \\ &= \left(\frac{n}{2\pi e}\right)^{n/2} \cdot \frac{1}{|\Sigma|^{1/2}} \cdot SSE^{-n/2}. \end{aligned}$$

   Thus $\max[L(\beta, \sigma^2) : \beta, \sigma^2] = \left(\frac{n}{2\pi e}\right)^{n/2} \cdot \frac{1}{|\Sigma|^{1/2}} \cdot SSE^{-n/2}$.

   (3) α-level LRT

   > $H_0 : H\beta = b$ vs $H_a : H\beta \neq b$
   > Text Statistic: $F = \frac{MSD}{MSE}$
   > Reject H_0 if $F > F_\alpha(r - r_1, n - r)$

   is an α-level LRT

   Proof. With $F = \frac{MSD}{MSE}$, the likelihood ratio

   $$\begin{aligned} \text{LR} &= \frac{\max[L(\beta, \sigma^2) : \beta, \sigma^2]}{\max[L(\beta, \sigma^2) : H_0]} = \left(\frac{SSE_H}{SSE}\right)^{n/2} = \left(1 + \frac{SSE_H - SSE}{SSE}\right)^{n/2} \\ &= \left(1 + \frac{MSD}{MSE} \cdot \frac{r - r_1}{n - r}\right)^{n/2} = \left(1 + F \cdot \frac{r - r_1}{n - r}\right)^{n/2} \end{aligned}$$

   is an increasing function of F. So a test that rejects H_0 when $F > c$ is a LRT.
   For α-level test, c is determined by $\alpha = P(F > c|H_0) = P(F(r - r_1, n - r) > c)$.
   So $c = F_\alpha(r - r_1, n - r)$.

   Comment: p-value: $P(F(r - r_1, n - r) > F_{ob})$.
   Thus the key for implementing the test is to create the ANOVA table.

2. Implementation in regressions

   (1) For $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \epsilon$ with $H_0 : H\beta = \begin{pmatrix} 1 \\ 5 \end{pmatrix}$ where $H = \begin{pmatrix} 2 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$.

   ```
   proc reg; model y=x1 x2 x3;
        test 2*intercept+x3=1, x1+x2=5;
        run;
   ```

displays

	MS	F	$Pr > F$	
Numerator	MSD	2	MSD/MSE	$P(F(2, n-4) > F_{ob})$
Denominator	MSE	$n-4$		

Comments: Output for "test x1=0;" verifies the p-value in parameter table.
Output for "test x1=0, x2=0, x3=0;" verifies the p-value in Global ANOVA table.

(2) For $y = \beta_1 x_1 + \beta_2 x2 + \beta_3 x_3 + \epsilon$ with $H_0 : H\beta = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ where $H = \begin{pmatrix} 1 & -2 & 1 \\ 0 & 2 & 3 \end{pmatrix}$.

```
proc reg; model y=x1 x2 x3/noint;
     test x1-2*x2+x3=1, 2*x2+3*x3=2;
     run;
```

displays

	MS	F	$Pr > F$	
Numerator	MSD	2	MSD/MSE	$P(F(2, n-3) > F_{ob})$
Denominator	MSE	$n-3$		

Comments: Output for "test x1=0;" verifies the p-value in parameter table.
Output for "test x1=0, x2=0, x3=0;" verifies the p-value in Global ANOVA table.

3. Decomposition of SS

(1) SS

A function $g(Y)$ of $Y \in R^n$ is a sum of squares (SS)
$\overset{df}{\Longleftrightarrow}$ There exists a 1-1 mapping $Y \longleftrightarrow Y_*$ such that $g(Y) = Y_*' AA^+ Y_*$.

Comment: With $Z = AA^+ Y_* \in R^n$, $g(Y) = \|Z\|^2 = \sum_i Z_i^2$ is a sum of squares.

(2) A part of SS
$Y'BB^+Y$ is part of $Y'AA^+Y \overset{df}{\Longleftrightarrow} B = AT$ for some T.
Comment: (i) $B = AT$ for some T $\Longrightarrow$ (ii) $BB^+ = (AT)(AT)^+$
$\Longrightarrow$ (iii) $AA^+BB^+ = BB^+ \Longrightarrow$ (i)

(3) Decomposition of SS
If $Y'BB^+Y$ is part of $Y'AA^+Y$, then $Y'AA^+Y$ has SS decomposition

$$Y'AA^+Y = Y'BB^+Y + Y'(AA^+ - BB^+)(AA^+ - BB^+)^+ Y.$$

Comment: Under (iii) $AA^+ - BB^+ = (AA^+ - BB^+)(AA^+ - BB^+)^+ = AA^+ - BB^+$.