L23 Decomposition of SSM

1. Model $y \sim N(X\beta, \sigma^2 I_n)$, $X \in R^{n\times p}$, rank$(X) = p$ and $1_n \notin \mathcal{R}(X)$.

   (1) ANOVA table for testing on $H_0 : \beta = 0$.

Source	SS	DF	MS	F	$P_r > F$
Model	SSM $= y'XX^+y$	p	MSM	MSM/MSE	$P(F(p, n-p) > F_{ob})$
Error	SSE $= y'(I - XX^+)y$	$n-p$	MSE		
U.Total	U.SSTO $= y'I_n y$	n			

   (2) ANOVA table for testing on $H_0 : H\beta = 0$, $H \in R^{q\times p}$, rank$(H) = q$.
   Denote previous SSE = SSE$_r$ − SSE as SSH associated with Hypothesis.
   SSH $= y'\{XX^+ - [X(I - H^+H)][X(I - H^+H)]^+\}y$ with DF $p - (p - q) = q$. Then

Source	SS	DF	MS	F	$P_r > F$
Hypothesis	SSH	q	MSH	MSH/MSE	$P(F(q, n-p) > F_{ob})$
Error	SSE	$n-p$	MSE		

   (3) SSH is part of SSM
   $\{XX^+ - [X(I-H^+H)][X(I-H^+H)]^+\} = XX^+\{XX^+ - [X(I-H^+H)][X(I-H^+H)]^+\}$ implies that SSH is part of SSM. Thus with SSH$^\perp = y'[X(I - H^+H)][X(I - H^+H)]^+y$, SSM = SSH + SSH$^\perp$. We therefore have combined ANOVA table

Source	SS	DF	MS	F	$P_r > F$
Model	SSM	p	MSM	MSM/MSE	$P(F(p, n-p) > F_{ob})$
H	SSH	q	MSH	MSH/MSE	$P(F(q, n-p) > F_{ob})$
H$^\perp$	SSH$^\perp$	$p-q$	MSH$^\perp$		
Error	SSE	$n-p$	MSE		
U.Total	U.SSTO	n			

   Ex1: SSH $= \|\{XX^+ - [X(I - H^+H)][X(I - H^+H)]\}y\|^2$,
   SSH$^\perp = \|[X(I - H^+H)][X(I - H^+H)]^+y\|^2$ and
   $\{XX^+ - [X(I - H^+H)][X(I - H^+H)]\}y \perp [X(I - H^+H)][X(I - H^+H)]^+y$.

   Ex2: Decomposition of SSA$= y'AA^+y$ where rank$(A) = r$.
   For given $r_1 + \cdots + r_k = r$, in the compact form of EVD $AA^+ = PP'$ where $P \in R^{p\times r}$ and $P'P = I_r$, break P as $P = (P_1, ..., P_r)$ where $P_i \in R^{p\times r_i}$. Then $AA^+ = PP' = \sum_{i=1}^k P_iP_i'$. Let $A_i = P_iP_i' = A_iA_i^+$ and $SSA_i = y'A_iA_i^+y$. Then SSA $= \sum_{i=1}^k$ SSA$_i$ where SSA_i, $i = 1, ..., k$, are SSs.
   From $PP' = P_1P_1' + \cdots + P_kP_k'$, $PP'P_iP_i' = P_iP_i'$. So $A_i = AA^+A_i$. Hence SSA_i is part of SSA.

2. Model $y \sim N(X\beta, \sigma^2 I_n)$, $X \in R^{n\times p}$, rank$(X) = p$ and $1_n \in \mathcal{R}(X)$.

   (1) ANOVA table for global F-test.

Source	SS	DF	MS	F	$P_r > F$
Model	SSM $= y'(XX^+ - 11^+)y$	$p-1$	MSM	MSM/MSE	$P(F(p-1, n-p) > F_{ob})$
Error	SSE $= y'(I - XX^+)y$	$n-p$	MSE		
C.Total	C.SSTO $= y'(I_n - 11^+)y$	$n-1$			

   (2) ANOVA table for testing on $H_0 : H\beta = 0$, $H \in R^{q\times p}$, rank$(H) = q$.
   Denote previous SSE = SSE$_r$ − SSE as SSH associated with Hypothesis.
   SSH $= y'\{XX^+ - [X(I - H^+H)][X(I - H^+H)]^+\}y$ with DF $p - (p - q) = q$. Then

Source	SS	DF	MS	F	$P_r > F$
Hypothesis	SSH	q	MSH	MSH/MSE	$P(F(q,\, n-p) > F_{ob})$
Error	SSE	$n-p$	MSE		

(3) SSH may or may not be part of SSM

Recall: $SSB = y'BB^+y$ is part of $SSA = y'AA^+y \iff B = AT$ for some T $\iff BB^+ = AA^+BB^+$

Thus

$$
\begin{aligned}
& SSH \text{ is part of } SSM \\
\iff & XX^+ - [X(I-H^+H)][X(I-H^+H)]^+ \\
& \quad = (XX^+ - 11^+)\{XX^+ - [X(I-H^+H)][X(I-H^+H)]^+\} \\
\iff & 0 = -11^+ + 11^+[X(I-H^+H)][X(I-H^+H)]^+ \\
\iff & 11^+ = 11^+[X(I-H^+H)][X(I-H^+H)]^+ \\
\iff & 1_n \in \mathcal{R}[X(I-H^+H)].
\end{aligned}
$$

3. Contrast tests

(1) Test on q contrasts

In ANOVA of p treatments with response means μ_i, $i = 1, ..., p$.

Test on $H_0 : H\mu = 0$ where $H \in R^{q\times p}$, rank$(H) = q$ and $H1_p = 0$ is a test on q contrasts.

(2) Testing on a hypothesized equivalent groups in treatments is a testing on a group contrasts.

For example when $p = 4$, the hypothesis on groups $(\mu_1,\, \mu_3)$ and $(\mu_2,\, \mu_4)$ is $H_0 : H\mu = 0$ where $H = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix}$. Clearly $H1_4 = 0$.

(3) Implementation via SAS

Suppose response y and treatment id with values A, B, C and D are stored in file ex.dat.

```
data a;
   infile "D:\ex.dat";
   input y id $ @@;
proc glm;
   class id;
   model y=id/nouni;
   contrast "group" id 1 0 -1 0, id 0 1 0 -1;
   run;
```

The output displays

	MS	DF	F	$Pr > F_{ob}$
Contrast	MSH	2	MSH/MSE	$P(F(2, n-4) > F_{ob})$

L24: Analysis of Covariance model

1. Analysis of covariance model (ANCOVA)

   (1) ANOVA model
   For observed $y \in R^n$, $y = X\beta + \epsilon$, $E(\epsilon) = 0 \in R^n$, is a linear model since $E(y) = X\beta$ is a linear function of β and hence $E(y)$ is in a linear space $S = \mathcal{R}(X)$.
   For ANOVA $y = \mu_i + \epsilon$ with $E(\epsilon) = 0 \in R$, $i = 1, 2, 3$, suppose y_1 and y_2 are from $y = \mu_1 + \epsilon$; y_3 is from $y = \mu_2 + \epsilon$ and y_4 is from $y = \mu_3 + \epsilon$.
   Then $y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix} + \epsilon$, $E(\epsilon) = 0 \in R^4$. Rewrite $y = \mu_. + \alpha_i + \epsilon$ with $E(\epsilon) = 0 \in R$ and $\alpha_1 + \alpha_2 + \alpha_3 = 0$. Then $y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{pmatrix} \begin{pmatrix} \mu_. \\ \alpha_1 \\ \alpha_2 \end{pmatrix} + \epsilon$, $E(\epsilon) = 0 \in R^4$. So one-way ANOVA is a linear model.

   (2) Regression model
   Regression $y = \beta_1 x_1 + \beta_2 x_2 + \epsilon$, $E(\epsilon) = 0 \in R$, with data can be expressed as $y = X\beta + \epsilon$ where $E(\epsilon) = 0 \in R^n$, $\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$ and the two columns of X are observed x_1 and x_2.

   (3) ANCOVA model
   Suppose one initially has ANOVA model $y = X\delta + \epsilon$ and later adds regression part $Z\gamma$ into the model to have $y = X\delta + Z\gamma + \epsilon$. This model is called ANCOVA model since the later added regressors are called covariates.

   Comment: Write ANCOVA $y = X\delta + Z\gamma + \epsilon$ as $y = (X, Z)\beta + \epsilon$ where $\beta = \begin{pmatrix} \delta \\ \gamma \end{pmatrix}$.
   Clearly this a linear model.

2. Estimable parameters and BLUE

   (1) Recall: LSE and estimable parameters
   For linear model $y = X\beta + \epsilon$, $E(\epsilon) = 0$

   $$\begin{aligned} \widehat{\beta} \text{ is LSE for } \beta \quad &\overset{def}{\iff} \quad \|y - X\beta\|^2 \geq \|y - X\widehat{\beta}\|^2 \text{ for all } \beta \iff X\widehat{\beta} = \pi(y|\mathcal{R}(X)) \\ &\iff \quad X\widehat{\beta} = XX^+y \iff \widehat{\beta} \in X^+y + \mathcal{N}(X) \end{aligned}$$

   So $\text{LSE}(\beta) = X^+y + \mathcal{N}(X)$.

   $$\begin{aligned} H\beta \text{ is estimable} \quad &\overset{def}{\iff} \quad \exists\, L \text{ such that } E(Ly) = H\beta \iff \exists\, L \text{ such that } LX\beta = H\beta \\ &\iff \quad LX = H \iff \exists\, L \text{ such that } H\beta = L(X\beta). \end{aligned}$$

   So $E(y) = X\beta$ is estimable since $X = IX$.

   Comment: $H\beta$ is estimable $\iff H\widehat{\beta}$ is unique. Clearly the unique value is HX^+y.

   (2) Estimator for σ^2 and BLUE
   Suppose $Y = X\beta + \epsilon$, $\epsilon \sim N(0, \sigma^2 I_n)$. Then $\text{SSE} = \|y - X\widehat{\beta}\|^2 = y'(I - XX^+)y$. $\frac{SSE}{n - \text{rank}(X)}$ is UE for σ^2, $\frac{SSE}{n}$ is MLE for σ^2. For estimable $H\beta$, the unique $H\widehat{\beta}$ is a BLUE.

(3) ANCOVA model

For $y = (X,\, Z)\beta + \epsilon$ where $\beta = \begin{pmatrix} \delta \\ \gamma \end{pmatrix}$ and $\epsilon \sim N(0,\, \sigma^2 I_n)$, $E(y) \in S = \mathcal{R}[(X,\, Z)]$.

$\widehat{\beta}$ is a LSE for $\beta \iff (X,\, Z)\widehat{\beta} = (X,\, Z)(X,\, Z)^+ y$.

SSE $= y'[I - (X,\, Z)(X,\, Z)^+]y$.

Comment: Fit the linear model framework, X is replaced by $(X,\, Z)$. Thus we need to explore more on $\mathcal{R}[(X,\, Z)]$ and $(X,\, Z)(X,\, Z)^+$.

3. Some specifics for ANCOVA

(1) For $y = (X,\, Z)\beta + \epsilon$

(i) $E(y) \in \mathrm{R}[(X,\, Z)] = \mathcal{R}(X) + \mathcal{R}(Z)$

(ii) $\mathcal{R}[(X,\, Z)] = \mathcal{R}(X) \dot{\oplus} \mathcal{R}[(X,\, (I - XX^+)Z)]$

(iii) $\mathcal{R}[(X,\, Z)] = \mathcal{R}[(I - ZZ^+)X] \dot{\oplus} \mathcal{R}(Z)$.

Proof. (i) $\mathcal{R}[(X,\, Z)] = \left\{ (X,\, Z) \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} : \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} \right\} = \{Xr_1 + Zr_2 : r_1,\, r_2\}$

$= \{Xr_1 : r_1\} + \{Zr_2 : r_2\} = \mathcal{R}(X) + \mathcal{R}(Z)$.

For (ii) note that $(X,\, Z) = [X,\, (I - XX^+)Z] \begin{pmatrix} I & X^+Z \\ 0 & I \end{pmatrix}$

and $[X,\, (I - XX^+)Z] = (X,\, Z) \begin{pmatrix} I & -X^+Z \\ 0 & I \end{pmatrix}$. So

$\mathcal{R}[(X,\, Z)] = \mathcal{R}[(X,\, (I - XX^+)Z)] = \mathcal{R}(X) + \mathcal{R}[(I - XX^+)Z] = \mathcal{R}(X) \dot{\oplus} \mathcal{R}[(I - XX^+)Z]$.

The last equal sign holds since $X'(I - XX^+)Z = 0$. The proof of (iii) is skipped.

(2) $(X,\, Z)(X,\, Z)^+ = XX^+ + [(I - XX^+)Z][(I - XX^+)Z]^+$ and
$(X,\, Z)(X,\, Z)^+ = [(I - ZZ^+)Z][(I - ZZ^+)Z]^+ + ZZ^+$.

Proof. First $(X,\, Z)(X,\, Z)^+ = [X,\, (I - XX^+)Z][X,\, (I - XX^+)Z]^+$ since they are the projection matrices onto the same space$\mathcal{R}[(X,\, Z)] = \mathcal{R}[(X,\, (I - XX^+)Z)]$.

Secondly $[X,\, (I - XX^+)Z]^+ = \begin{pmatrix} X^+ \\ [(I - XX^+)Z]^+ \end{pmatrix}$ since $X'(I - XX^+)Z = 0$.

Consequently $(X,\, Z)(X,\, Z)^+ = XX^+ + [(I - XX^+)Z][(I - XX^+)Z]^+$. The second equation can be proved similarly.

(3) If $\mathcal{R}(X) \cap \mathcal{R}(Z) = \{0\}$, then

(i) $\mathrm{rank}(X) = \mathrm{rank}[(I - ZZ^+)X]$ and $\mathrm{rank}(Z) = \mathrm{rank}[(I - XX^+)Z]$

(ii) $\mathcal{R}(X') = \mathcal{R}[X'(I - ZZ^+)]$ and $\mathcal{R}(Z') = \mathcal{R}[Z'(I - XX^+)]$

Proof. We first show the second equation in (i). Under $\mathcal{R}(X) \cap \mathcal{R}(Z) = \{0\}$,

$$
\begin{aligned}
\mathrm{rank}[(X,\, Z)] &= \dim[\mathcal{R}[(X,\, Z)]] = \dim[\mathcal{R}(X) + \mathcal{R}(Z)] \\
&= \dim[R(X)] + \dim[R(Z)] - \dim[R(X) \cap R(Z)] \\
&= \mathrm{rank}(X) + \mathrm{rank}(Z).
\end{aligned}
$$

$$
\begin{aligned}
\text{But } \mathrm{rank}[(X,\, Z)] &= \dim[\mathcal{R}[(X,\, (I - XX^+)Z)]] = \dim[\mathcal{R}(X) \dot{\oplus} \mathcal{R}((I - XX^+)Z)] \\
&= \dim[R(X)] + \dim[R((I - XX^+)Z)] \\
&= \mathrm{rank}(X) + \mathrm{rank}[(I - XX^+)Z].
\end{aligned}
$$

Thus $\mathrm{rank}(Z) = \mathrm{rank}[(I - XX^+)Z]$.

For the first equation in (ii) note that $\mathcal{R}[X'(I - ZZ^+)] \subset \mathcal{R}(X')$. But by (i) the two spaces share the same dimension. Thus $\mathcal{R}(X') = \mathcal{R}[X'(I - ZZ^+)]$.