Analysis of Covariance

Analysis of covariance models are generalizations of ANOVA models in that the design matrix contains both qualitative and quantitative explanatory variables. The quantitative variable are referred to as regression variables, covariates, or concomitant variables. There are essentially two goals in analysis of covariance (ANCOVA).

i) Compare treatments

ii) Inference on the regression coefficients corresponding to the covariates.

The primary goal is still to compare treatments. The concomitant variables are intended to serve as blocking factors to sharpen the analysis so that differences between treatments can be better assessed. Thus, concomitant variables are introduced to reduce variability. Thus, ANCOVA can be viewed as a variance reduction design, for which the concomitant variable is quantitative and often continuous.

The general ANCOVA model can be written in matrix notation as

$$
\begin{aligned}
Y &= (X, Z)\begin{pmatrix} \beta \\ \gamma \end{pmatrix} + \epsilon \\
&= X\beta + Z\gamma + \epsilon \qquad (5)
\end{aligned}
$$

where X is an $n \times p$ experimental design matrix for an ANOVA, and Z is an $n \times s$ matrix of concomitant variables. To test hypotheses, we assume that

$$\epsilon \sim N_n(0, \sigma^2 I)$$

so that

$$Y \sim N_n(X\beta + Z\gamma, \sigma^2 I) .$$

As we mentioned above, the concomitant variables Z are introduced only to sharpen the analysis, and thus the inference on the ANOVA is done AFTER the regression fit. The regression coefficients for the concomitant variables are tested after the ANOVA tests.

Estimation of γ

We do not need distributional assumptions to estimate γ. To estimate γ, we only need

$$E(\epsilon) = 0 \quad (i.e., E(Y) = X\beta + Z\gamma)$$

and

$$\text{Cov}(\epsilon) = \sigma^2 I \ .$$

To estimate γ, we break up model (5) into two orthogonal parts. Since

$$E(Y) = X\beta + Z\gamma \ ,$$

write

$$\begin{aligned} X\beta + Z\gamma &= X\beta + MZ\gamma + (I-M)Z\gamma \\ &= (X, MZ)\begin{pmatrix} \beta \\ \gamma \end{pmatrix} + (I-M)Z\gamma \ , \end{aligned}$$

where $M = X(X'X)^{-}X'$. Since $C(X) = C(X, MZ)$, the first part of the model is overparameterized. We reparameterize from (β, γ) to δ, and write

$$\begin{aligned} E(Y) &= X\delta + (I-M)Z\gamma \\ &= (X, (I-M)Z)\begin{pmatrix} \delta \\ \gamma \end{pmatrix} \ . \end{aligned}$$

Now the two parts $(X, (I-M)Z)$ are orthogonal since

$$X'(I-M)Z = (X' - X'M)Z = (X' - X')Z = 0 \ .$$

Since, $(X, (I-M)Z)$ are orthogonal, we can do the estimation of δ and γ separately. We justify this in the following theorem.

Theorem

Consider the linear model

$$Y = X_1\beta_1 + X_2\beta_2 + \epsilon \qquad (6)$$

where $E(\epsilon) = 0$, $\text{Cov}(\epsilon) = \sigma^2 I$, X_1 is $n \times p_1$ of rank r_1 and X_2 is $n \times p_2$ of rank r_2. Suppose

$$C(X_1) \perp C(X_2)$$

so that $X_1'X_2 = 0_{p_1 \times p_2}$ and $X_2'X_1 = 0_{p_2 \times p_1}$. Then the least squares estimates of β_1 and β_2 satisfy

$$X_1\hat{\beta}_1 = M_1 Y$$

and

$$X_2\hat{\boldsymbol\beta}_2 = M_2 Y$$

where

$$M_1 = X_1(X_1'X_1)^- X_1'$$

and

$$M_2 = X_2(X_2'X_2)^- X_2' \ .$$

Proof

Write $X = (X_1, X_2)$ and $\boldsymbol\beta = \begin{pmatrix} \boldsymbol\beta_1 \\ \boldsymbol\beta_2 \end{pmatrix}$, so that the linear model in (6) can be written as

$$Y = X\boldsymbol\beta + \boldsymbol\epsilon \ . \tag{7}$$

We know that the least squares solution to $\boldsymbol\beta$ for the model in (7) is given by

$$X\hat{\boldsymbol\beta} = MY.$$

In terms of $X = (X_1, X_2)$,

$$\begin{aligned} M &= X(X'X)^- X' \\ &= (X_1, X_2)\begin{pmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{pmatrix}^- \begin{pmatrix} X_1' \\ X_2' \end{pmatrix} \\ &= X_1(X_1'X_1)^- X_1' + X_2(X_2'X_2)^- X_2' \\ &= M_1 + M_2 \ . \end{aligned}$$

Thus $MY = (M_1 + M_2)Y = M_1 Y + M_2 Y$. Now

$$X\hat{\boldsymbol\beta} = (X_1, X_2)\begin{pmatrix} \hat{\boldsymbol\beta}_1 \\ \hat{\boldsymbol\beta}_2 \end{pmatrix} = X_1\hat{\boldsymbol\beta}_1 + X_2\hat{\boldsymbol\beta}_2 \ .$$

Thus

$$X_1\hat{\boldsymbol\beta}_1 + X_2\hat{\boldsymbol\beta}_2 = M_1 Y + M_2 Y \ . \tag{8}$$

Now multiplying both sides by X_2', we get

$$X_2'X_1\hat{\boldsymbol\beta}_1 + X_2'X_2\hat{\boldsymbol\beta}_2 = X_2'M_1 Y + X_2'M_2 Y$$

which implies

$$X_2'X_2\hat{\boldsymbol\beta}_2 = X_2'Y \ .$$

since $X_2'X_1 = 0$. These are the normal equations for $\hat{\boldsymbol\beta}_2$, which we know are equivalent to

$$X_2\hat{\boldsymbol\beta}_2 = M_2Y \ .$$

Multiplying both sides of (8) by X_1' leads to

$$X_1\hat{\boldsymbol\beta}_1 = M_1Y \ .$$

This completes the proof.

The normal equations for $(\boldsymbol\delta, \boldsymbol\gamma)$ are given by

$$(X, (I - M)Z)'(X, (I - M)Z)\begin{pmatrix} \boldsymbol\delta \\ \boldsymbol\gamma \end{pmatrix} = (X, (I - M)Z)'Y \ ,$$

Carrying through the multiplication and noting that $X'(I - M)Z = 0$, the normal equations reduce to

$$\begin{pmatrix} X'X & 0 \\ 0 & Z'(I - M)Z \end{pmatrix}\begin{pmatrix} \boldsymbol\delta \\ \boldsymbol\gamma \end{pmatrix} = \begin{pmatrix} X'Y \\ Z'(I - M)Y \end{pmatrix}$$

Thus, a least squares estimate of $\boldsymbol\gamma$ is

$$\hat{\boldsymbol\gamma} = (Z'(I - M)Z)^{-}Z'(I - M)Y.$$

Since Z consists of concomitant variables, $Z'(I - M)Z$ will be typically nonsingular. In this case, the unique least squares estimate of $\boldsymbol\gamma$ is

$$\hat{\boldsymbol\gamma} = (Z'(I - M)Z)^{-1}Z'(I - M)Y.$$

Since X is the design matrix for an ANOVA, $\boldsymbol\delta$ (or $\boldsymbol\beta$) is usually not estimable. Thus we will concern ourselves with the estimation of $X\boldsymbol\delta$ (or $X\boldsymbol\beta$). The least squares equations for $\boldsymbol\delta$ are

$$X\hat{\boldsymbol\delta} = MY \ .$$

Thus, if we want the least squares estimate of $X\boldsymbol\beta$, we note that

$$X\boldsymbol\delta + (I - M)Z\boldsymbol\gamma = X\boldsymbol\beta + MZ\boldsymbol\gamma + (I - M)Z\boldsymbol\gamma$$

so that

$$X\hat{\boldsymbol\delta} + (I - M)Z\hat{\boldsymbol\gamma} = X\hat{\boldsymbol\beta} + MZ\hat{\boldsymbol\gamma} + (I - M)Z\hat{\boldsymbol\gamma} \ .$$

subtracting $(I - M)Z\hat{\boldsymbol\gamma}$ from both sides gives

$$X\hat{\boldsymbol\delta} = X\hat{\boldsymbol\beta} + MZ\hat{\boldsymbol\gamma} \ .$$

Therefore,

$$X\hat{\boldsymbol\beta} \quad = \quad X\hat{\boldsymbol\delta} - MZ\hat{\boldsymbol\gamma}$$

$$
\begin{aligned}
&= MY - MZ(Z'(I-M)Z)^{-}Z'(I-M)Y \\
&= M(Y - Z\hat{\boldsymbol{\gamma}})
\end{aligned}
$$

Notice that if $Z'(I-M)Z$ is singular, then $X\boldsymbol{\beta}$ is not estimable. Note that

$$
\begin{aligned}
& E(X\hat{\boldsymbol{\beta}}) \\
= \;& E(MY - MZ\hat{\boldsymbol{\gamma}}) \\
= \;& ME(Y) - MZE(\hat{\boldsymbol{\gamma}}) \\
= \;& M(X\boldsymbol{\beta} + Z\boldsymbol{\gamma}) - MZE(\hat{\boldsymbol{\gamma}}) \\
= \;& X\boldsymbol{\beta} + MZ\boldsymbol{\gamma} - MZ(Z'(I-M)Z)^{-}Z'(I-M)(X\boldsymbol{\beta} + Z\boldsymbol{\gamma}) \\
= \;& X\boldsymbol{\beta} + MZ\boldsymbol{\gamma} - MZ(Z'(I-M)Z)^{-}Z'(I-M)Z\boldsymbol{\gamma}
\end{aligned}
$$

If $Z'(I-M)Z$ is nonsingular, then of course $\boldsymbol{\gamma}$ is estimable. Also, in this case, $X\boldsymbol{\beta}$ is estimable. In particular, if $Z'(I-M)Z$ is nonsingular, then any function of $\boldsymbol{\beta}$ that's estimable in the ANOVA model is also estimable in the ANCOVA model.

When $Z'(I-M)Z$ is singular, we need to characterize the estimable functions of $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$. We are led to the following theorem.

Theorem

$\boldsymbol{\xi}'\boldsymbol{\gamma}$ is estimable if and only if $\boldsymbol{\xi}' = \boldsymbol{\rho}'(I-M)Z$ for some vector $\boldsymbol{\rho} \in R^n$.

This theorem says that the estimable functions of $\boldsymbol{\gamma}$ are those that are linear functions of $(I-M)Z$.

Proof

If $\boldsymbol{\xi}'\boldsymbol{\gamma}$ is estimable, then there exists a $\boldsymbol{\rho}$ such that

$$
\boldsymbol{\xi}'\boldsymbol{\gamma} = \boldsymbol{\rho}'(X, Z)\begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\gamma} \end{pmatrix}
$$

so that

$$
\boldsymbol{\rho}'(X, Z) = (0, \boldsymbol{\xi}')
$$

and $\boldsymbol{\rho}'X = 0$. Therefore, $\boldsymbol{\xi}' = \boldsymbol{\rho}'Z = \boldsymbol{\rho}'(I-M)Z$. Conversely, if $\boldsymbol{\xi}' = \boldsymbol{\rho}'(I-M)Z$, then

$$
\boldsymbol{\rho}'(I-M)(X, Z)\begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\gamma} \end{pmatrix} = \boldsymbol{\xi}'\boldsymbol{\gamma}\ .
$$

Corollary

The unique BLUE of $\boldsymbol{\rho}'(I - M)Z\boldsymbol{\gamma}$ is

$$\boldsymbol{\rho}' M_{(I-M)Z} Y$$

where

$$M_{(I-M)Z} = (I - M)Z(Z'(I - M)Z)^{-}Z'(I - M)$$

is the orthogonal projection operator onto $C((I - M)Z)$.

Finally, we note that if $Z'(I - M)Z$ is nonsingular, so that $(X\boldsymbol{\beta}, \boldsymbol{\gamma})$ are both estimable, then

$$\mathrm{Cov}\begin{pmatrix} X\hat{\boldsymbol{\beta}} \\ \hat{\boldsymbol{\gamma}} \end{pmatrix} = \boldsymbol{\sigma}^2 \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

where

$A_{11} = M + MZ(Z'(I - M)Z)^{-1}Z'M$

$A_{12} = -MZ(Z'(I - M)Z)^{-1}$

$A_{21} = A_{12}'$

and

$A_{22} = (Z'(I - M)Z)^{-1}$. Estimation of $\boldsymbol{\sigma}^2$

Let P denote the orthogonal projection operator onto $C(X, Z)$. Then, the error sum of squares (SSE) for the ANCOVA model is given by

$$SSE = \|(I - P)Y\|^2 = Y'(I - P)Y \ .$$

We would like to express P in terms of X and Z. We have established that

$$C(X, Z) = C(X, (I - M)Z) \ .$$

Since

$$C(X) \perp C((I - M)Z) \ ,$$

the orthogonal projection operator for $C(X, (I - M)Z)$ is the sum of the orthogonal projection operators onto $C(X)$ and $C((I - M)Z)$, respectively. Thus

$$P = M + M_{(I-M)Z}$$

where

$$M = X(X'X)^{-}X'$$

and

$$M_{(I-M)Z} = (I - M)Z(Z'(I - M)Z)^{-}Z'(I - M) \ .$$

Thus,

$$SSE = \|\left(I - \left[M + (I-M)Z(Z'(I-M)Z)^{-}Z'(I-M)\right]\right)Y\|^2 .$$

Define the notation

$$E_{AB} = A'(I-M)B$$

where A and B are arbitrary matrices. Thus, we can write

$$\begin{aligned} \|(I-P)Y\|^2 &= Y'(I-P)Y \\ &= E_{yy} - E_{yz}E_{zz}^{-}E_{zy}. \end{aligned}$$

To prove this identity, note that

$$\begin{aligned} & Y'(I-P)Y \\ = \; & Y'(I-M-(I-M)Z(Z'(I-M)Z)^{-}Z'(I-M)Y \\ = \; & Y'(I-M)Y - Y'(I-M)Z(Z'(I-M)Z)^{-}Z'(I-M)Y \\ = \; & E_{yy} - E_{yz}E_{zz}^{-}E_{zy} . \end{aligned}$$

Tests of hypotheses for ANCOVA are obtained by considering the reductions in the SSE's for the models being tested. We note here that the primary interest in tests of hypothesis are tests concerning the treatment effects. Hypotheses tests concerning $\boldsymbol{\gamma}$ are usually not of interest in ANCOVA.

If X_0 is such that $C(X_0) \subset C(X)$, we may wish to test the reduced model

$$Y = X_0\boldsymbol{\gamma}_0 + Z\boldsymbol{\gamma} + \boldsymbol{\epsilon}$$

against

$$Y = X\boldsymbol{\beta} + Z\boldsymbol{\gamma} + \boldsymbol{\epsilon} .$$

We can write the hypotheses as

$$\begin{aligned} H_0 &: E(Y) \in C(X_0, Z) \\ H_a &: E(Y) \in C(X, Z) \cap (C(X_0, Z))^c \end{aligned}$$

The F test for nested models applies here and is given by

$$\begin{aligned} F &= \frac{\|(P-P_0)Y\|^2/r(P-P_0)}{\|(I-P)Y\|^2/r(I-P)} \\ &\sim F(r(P-P_0), r(I-P), \boldsymbol{\gamma}^*) \end{aligned}$$

where

$$\boldsymbol{\gamma}^* = \frac{\|(P-P_0)(X\boldsymbol{\beta} + Z\boldsymbol{\gamma})\|^2}{2\boldsymbol{\sigma}^2}$$

and P_0 is the orthogonal projection operator onto $C(X_0, Z)$.

Thus,

$$P_0 = M_0 + (I - M_0)Z(Z'(I - M_0)Z)^- Z'(I - M_0)$$

where

$$M_0 = X_0(X_0'X_0)^- X_0'$$

is the orthogonal projection operator onto $C(X_0)$. The F test above can also be written as

$$\begin{aligned} F &= \frac{Y'(P - P_0)Y/r(P - P_0)}{Y'(I - P)Y/r(I - P)} \\ &= \frac{[Y'(I - P_0)Y - Y'(I - P)Y]/(r(P) - r(P_0))}{Y'(I - P)Y/(n - r(P))} \end{aligned}$$

where n is the total number of observations. We see that numerator of the F test can be written as a difference in the error sums of squares between the nested models.

Example

Consider the balanced two-way ANOVA model with no replication, (and hence no interaction), and one covariate. The ANCOVA model can be written as

$$Y_{ij} = \mu + \alpha_i + \eta_j + \gamma z_{ij} + \epsilon_{ij}$$

for $i = 1, \ldots, a$ and $j = 1, \ldots, b$. We can write this model in the matrix form

$$Y = X\boldsymbol{\beta} + Z\gamma + \epsilon$$

where X is the usual $n \times (a + b + 1)$ design matrix for a two-way ANOVA without interaction with $n = abN$, $N = 1$,

$$Z = \begin{pmatrix} z_{11} \\ \vdots \\ z_{1b} \\ z_{21} \\ \vdots \\ z_{ab} \end{pmatrix}_{ab\times 1}, \quad \boldsymbol{\beta} = \begin{pmatrix} \mu \\ \alpha_1 \\ \vdots \\ \alpha_a \\ \eta_1 \\ \vdots \\ \eta_b \end{pmatrix}_{(a+b+1)\times 1},$$

and γ is a scalar.

We can write the sums of squares for this ANCOVA model as follows.

$$E_{yy} = Y'(I - M)Y = \sum_{i=1}^{a}\sum_{j=1}^{b}(Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})^2 ,$$

$$\begin{aligned} E_{yz} &= Y'(I - M)Z \\ &= \sum_{i=1}^{a}\sum_{j=1}^{b}(Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})(Z_{ij} - \bar{Z}_{i.} - \bar{Z}_{.j} + \bar{Z}_{..}) . \end{aligned}$$

We note here that since Z is a vector in this example,
$Y'(I - M)Z = Z'(I - M)Y$, so that $E_{yz} = E_{zy}$.

Finally, we have

$$E_{zz} = Z'(I - M)Z = \sum_{i=1}^{a}\sum_{j=1}^{b}(Z_{ij} - \bar{Z}_{i.} - \bar{Z}_{.j} + \bar{Z}_{..})^2 ,$$

and

$$SSE = Y'(I - P)Y = E_{yy} - E_{yz}E_{zz}^{-}E_{zy} .$$

Suppose we are interested in testing

$$\begin{aligned} H_0 &: \gamma = 0 \\ H_a &: \gamma \neq 0 . \end{aligned}$$

The reduced model is given by

$$Y_{ij} = \mu + \alpha_i + \eta_j + \epsilon_{ij} .$$

Thus, X_0 is the design matrix for the two-way ANOVA without interaction.
Thus, $P_0 = X_0(X_0'X_0)^{-}X_0'$, and

$$Y'(I - P_0)Y = \sum_{i=1}^{a}\sum_{j=1}^{b}(Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})^2 .$$

Note here that $r(P_0) = 1 + a - 1 + b - 1 = a + b - 1$ and, thus, the F test takes the form

$$\begin{aligned} F &= \frac{\|(P - P_0)Y\|^2 / r(P - P_0)}{\|(I - P)Y\|^2 / r(I - P)} \\ &= \frac{Y'(P - P_0)Y / r(P - P_0)}{Y'(I - P)Y / r(I - P)} \end{aligned}$$

$$
\begin{aligned}
&= \frac{[Y'(I-P_0)Y - Y'(I-P)Y]/r(P-P_0)}{Y'(I-P)Y/r(I-P)} \\
&= \frac{(E_{yy} - \left[E_{yy} - E_{yz}E_{zz}^{-1}E_{zy}\right])/r(P-P_0)}{Y'(I-P)Y/r(I-P)} \\
&= \frac{E_{yz}E_{zz}^{-1}E_{zy}/1}{Y'(I-P)Y/(n-a-b)} \\
&\sim F(1, n-a-b, \gamma^*)
\end{aligned}
$$

where $n = ab$ and

$$\gamma^* = \frac{\|(P-P_0)(X\beta + Z\gamma)\|^2}{2\sigma^2} .$$

We note here that $r(P_0) = 1 + (a-1) + (b-1) = a + b - 1,$

and $r(P) = 1 + (a-1) + (b-1) + 1 = a + b,$

so that

$$
\begin{aligned}
r(P-P_0) &= r(P) - r(P_0) \\
&= a + b - (a + b - 1) = 1,
\end{aligned}
$$

and

$$
\begin{aligned}
r(I-P) &= r(I) - r(P) \\
&= ab - (a+b) = ab - a - b \\
&= n - a - b .
\end{aligned}
$$

Sums of Squares for the General ANCOVA Model

Consider the general ANCOVA model

$$Y = X\beta + Z\gamma + \epsilon .$$

We wish to construct the general ANCOVA table. To do so, we need to construct sums of squares for Y, denoted S_{YY}, the sums of squares for Z, denoted S_{ZZ}, and the sum of cross products of Y and Z, denoted S_{YZ}. To illustrate how this can be done, suppose the ANOVA part of the model is a two-way balanced ANOVA with interaction. We have a total of $n = abN$ observations. Also,

$$C(X) = C(M_\mu) + C(M_\alpha) + C(M_\eta) + C(M_{\alpha\eta}) .$$

We can write the ANCOVA table as

source	df	SS_{yy}	SS_{yz}	SS_{zz}
μ	1	$Y'M_{\mu}Y$	$Y'M_{\mu}Z$	$Z'M_{\mu}Z$
α	$a-1$	$Y'M_{\alpha}Y$	$Y'M_{\alpha}Z$	$Z'M_{\alpha}Z$
η	$b-1$	$Y'M_{\eta}Y$	$Y'M_{\eta}Z$	$Z'M_{\eta}Z$
$\alpha\eta$	$(a-1)(b-1)$	$Y'M_{\alpha\eta}Y$	$Y'M_{\alpha\eta}Z$	$Z'M_{\alpha\eta}Z$
Error	$n-ab$	$Y'(I-M)Y$	$Y'(I-M)Z$	$Z'(I-M)Z$
Total	n	$Y'Y$	$Y'Z$	$Z'Z$

Thus, in ANCOVA, we still partition R^n into a sum of orthogonal subspaces.

To test for no interaction for this model, our hypothesis is

$$H_0 : (\alpha\eta)_{11} = \ldots = (\alpha\eta)_{ab} .$$

To construct the F test, we need P_0 and P. We have

$$P_0 = M_0 + (I-M_0)Z(Z'(I-M_0)Z)^{-}Z'(I-M_0)$$

where

$$M_0 = M - M_{\alpha\eta} .$$

Thus, the numerator sums of squares for the F test is

$$\begin{aligned}
& \|(P-P_0)Y\|^2 \\
= \;& Y'(P-P_0)Y \\
= \;& Y'(M-M_0)Y \\
+ \;& Y'\left[(I-M)Z(Z'(I-M)Z)^{-}Z'(I-M)\right]Y \\
- \;& Y'\left[(I-M_0)Z(Z'(I-M_0)Z)^{-}Z'(I-M_0)\right]Y \\
= \;& Y'M_{\alpha\eta}Y + Y'(I-M)Z(Z'(I-M)Z)^{-}Z'(I-M)Y \\
- \;& Y'(I-M+M_{\alpha\eta})Z(Z'(I-M+M_{\alpha\eta})Z)^{-}Z'(I-M+M_{\alpha\eta})Y \\
= \;& Y'M_{\alpha\eta}Y + Y'(I-M)Z(Z'(I-M)Z)^{-}Z'(I-M)Y \\
- \;& \left(Y'(I-M)Z + Y'M_{\alpha\eta}Z\right) \\
\times \;& \left(Z'(I-M)Z + Z'M_{\alpha\eta}Z\right)^{-}\left(Z'(I-M)Y + Z'M_{\alpha\eta}Y\right)
\end{aligned}$$

The decomposition of $Y'(I-P)Y$ into terms involving only Y, M, and Z is similar. Note that all of these terms can be obtained from the ANCOVA table.