# Orthogonal Contrasts and Partition of Squares

by Marek Rychlik

March 13, 2009

# Review of definitions

An abstraction of an experiment is a family of random variables $\{X_{ij}\}_{1 \le i \le t, 1 \le j \le r_i}$, where $t$ is the number of treatment groups, and $r_i$ is the number of observational units in the $i$-th treatment group. The total number of observational units is:

$$N = \sum_{i=1}^{t} r_i$$

Let $X$ be the random vector obtained by stacking $X_{ij}$:

$$X = \begin{pmatrix} X_{11} \\ X_{12} \\ \vdots \\ X_{1r_i} \\ X_{21} \\ X_{22} \\ \vdots \\ X_{2r_2} \\ \vdots \\ X_{tr_t} \end{pmatrix}$$

The group means are:

$$\mu_i = \frac{1}{r_i} \sum_{j=1}^{r_i} X_{ij}$$

This means that group means are random variables as well.

A contrast is a linear function of the group means:

$$C(X) = \sum_{i=1}^{t} c_i \mu_i$$

Contrasts are random variables. Contrasts form a vector space (i.e. they can be added and multiplied by a scalar) and this space can be identified with $\mathbb{R}^t$ using the correspondence:

$$C \leftrightarrow \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_t \end{pmatrix}$$

The estimators of the group means are obtained by plugging in the experimental results coming in the form of sample data $\{y_{ij}\}_{1\le i\le t, 1\le j\le r_i}$:

$$\hat{\mu}_i = \frac{1}{r_i}\sum_{j=1}^{r_i} y_{ij} = \bar{y}_i.$$

Similarly, an estimator of $c(X)$ is defined as:

$$\widehat{C(X)} = \sum_{i=1}^{t} c_i\hat{\mu}_i$$

The covariance of two contrasts $C(X) = \sum c_i\mu_i$ and $D(x) = \sum d_i\mu_i$ is given by:

$$\mathrm{Cov}(C(X), D(X)) = \sigma^2 \sum_{i=1}^{t} \frac{c_i d_i}{r_i}$$

We make the usual assumption of homoskedasticity, and thus $\sigma^2$ is the common variance of all random variables $X_{ij}$.

For convenience and brevity, we define a quadratic form on contrasts:

$$\langle C, D\rangle = \sum_{i=1}^{t} \frac{c_i d_i}{r_i}$$

Thus,

$$\mathrm{Cov}(C(X), D(X)) = \sigma^2 \langle C, D\rangle$$

We observe that $\langle\cdot,\cdot\rangle \colon \mathbb{R}^{\mathfrak{t}} \rightarrow \mathbb{R}$

**Definition 1.** *Contrasts $C(X)$ and $D(X)$ are orthogonal iff $\langle C, D\rangle = 0$.*

# Orthogonality of contrasts to the mean

The overall sample mean is:

$$\mu = \frac{1}{N} \sum_{i=1}^{t} \sum_{j=1}^{r_i} X_{ij}$$

The overall mean can also be computed as a weighted average of the group means:

$$\mu = \frac{1}{N} \sum_{i=1}^{t} r_i \mu_i$$

This means that the overall mean corresponds to this contast:

$$C_0(X) = \sum r_i \mu_i$$

which corresponds to the vector of coefficients:

$$C_0 \leftrightarrow \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_t \end{pmatrix}$$

With this notation, $\mu = \frac{1}{N} C_0(X)$. Also, $\langle C_0, C_0 \rangle = \sum_{i=1}^{t} \frac{r_i r_i}{r_i} = \sum_{i=1}^{t} r_i = N$.

**Lemma 2.** *A contrast $C$ is orthogonal to $C_0$ iff* $\sum_{i=1}^{t} C_i = 0$.

**Proof.** This follows from this identity: $\langle C, C_0 \rangle = \sum_{i=1}^{t} \frac{c_i r_i}{r_i} = \sum_{i=1}^{t} c_i$ □

# A convention

It is customary to require that the sum of the coefficients of a contrast be equal to zero. The contrast $C_0$ does not have this property. In some considerations it is useful to consider contrasts without this property. Most interesting contrasts are orthogonal to $C_0$.

# Contrast sum of squares

**Definition 3.** *The contrast sum of squares of the contrast $C$ is the expression:*

$$SSC = \frac{C(X)^2}{\langle C, C \rangle} = \frac{C(X)^2}{\sum_{i=1}^{t} \frac{c_i^2}{r_i}}$$

# The value of the contrast

**Lemma 4.** *Let $M \in \mathbb{R}^t$ be the vector of the scaled means:*

$$M = \begin{pmatrix} r_1\mu_1 \\ r_2\mu_2 \\ \vdots \\ r_t\mu_t \end{pmatrix}$$

*Then for every contrast $C$:*

$$C(X) = \langle C, M \rangle$$

# Partition of the sum of squares

**Theorem 5.** *Let* $C_1, C_2, \ldots, C_{t-1}$ *be an orthogonal set of contrasts. Moreover, let the sum of the coefficients of every contrast be* 0 *(equivalently,* $\langle C_i, C_0 \rangle = 0$ *for* $i = 1, 2, \ldots, t-1$*). Then:*

$$SST = \sum_{i=1}^{t-1} SSC_i,$$

*where* $SST = \sum_{i=1}^{t} r_i(\mu_i - \mu)^2$.

**Proof.** Let $M = (r_i\mu_i)$ be the scaled vector of the means. By the previous lemma and definition of $SSC$ we have:

$$\sum_{i=0}^{t-1} SSC_i = \sum_{i=1}^{t} \frac{\langle C, M \rangle^2}{\langle C, C \rangle}$$

We also have $\langle C_0, M \rangle = \sum_{i=1}^{t} \frac{r_i r_i \mu_i}{r_i} = \sum_{i=1}^{t} r_i\mu_i = N\mu$, where $\mu$ is the overall mean, and $\langle C_0, C_0 \rangle = N$. Hence,

$$SSC_0 = \frac{(N\mu)^2}{N} = N\mu^2.$$

We observe that $\{C_0, C_1, ..., C_{t-1}\}$ is an orthogonal basis of the inner product space of all contrasts. Hence, The Pythagorean Theorem yields:

$$N\mu^2 + \sum_{i=0}^{t-1} SSC_i = \|M\|^2 = \sum_{i=1}^{t} \frac{r_i^2 \mu_i^2}{r_i} = \sum_{i=1}^{t} r_i \mu_i^2.$$

Therefore,

$$\sum_{i=1}^{t-1} SSC_i = \sum_{i=1}^{t} r_i \mu_i^2 - N\mu^2 = \sum_{i=1}^{t} r_i (\mu_i - \mu)^2 = SST.$$

□

# Discussion

The last theorem demonstrates that $SST$ (sum of squares of treatments) partitions along the directions of orthogonal contrasts (whose sum of coefficients is zero). Therefore, $t - 1$ orthogonal contrasts "explain" the entire variation of the means.

Orthogonal contrasts are independent random variables, if the assumption of normality and homoskedasticity holds.