Example: Balanced Two-Factor Diet-Drug Experiment

It follows that

main diet effect → $\boldsymbol{c}_1^\top \boldsymbol{\beta} = [1, 1, 1, -1, -1, -1]\boldsymbol{\beta}$

diet 1 vs. diet 2

$\frac{1}{3}\ \frac{1}{3}\ \frac{1}{3}\ -\frac{1}{3}\ -\frac{1}{3}\ -\frac{1}{3}$

diet 1 $\mu_{1\cdot}$ vs. $\mu_{2\cdot}$

test for drug main effect

$$\begin{cases} \boldsymbol{c}_2^\top \boldsymbol{\beta} = [1, -1, 0, 1, -1, 0]\boldsymbol{\beta} \\ \boldsymbol{c}_3^\top \boldsymbol{\beta} = [1, 1, -2, 1, 1, -2]\boldsymbol{\beta} \end{cases}$$

drug 1 vs. drug 2 within each diet

$\mu_{11} + \mu_{12} - 2\mu_{13} = 0$

$\frac{\mu_{11}+\mu_{12}}{2} = \mu_{13}$

test for the interaction term

$$\begin{cases} \boldsymbol{c}_4^\top \boldsymbol{\beta} = [1, -1, 0, -1, 1, 0]\boldsymbol{\beta} \\ \boldsymbol{c}_5^\top \boldsymbol{\beta} = [1, 1, -2, -1, -1, 2]\boldsymbol{\beta} \end{cases}$$

// end lecture 18

3-05-25

comprise a set of pairwise orthogonal contrasts.

Connection to the ANOVA Table

Source	Sum of Squares	DF
Diets	$\boldsymbol{y}^\top(\boldsymbol{P}_2 - \boldsymbol{P}_1)\boldsymbol{y}$	$2 - 1 = 1$
Drugs	$\boldsymbol{y}^\top(\boldsymbol{P}_3 - \boldsymbol{P}_2)\boldsymbol{y}$	$4 - 2 = 2$
Diets $\times$ Drugs	$\boldsymbol{y}^\top(\boldsymbol{P}_4 - \boldsymbol{P}_3)\boldsymbol{y}$	$6 - 4 = 2$
Error	$\boldsymbol{y}^\top(\boldsymbol{I} - \boldsymbol{P}_4)\boldsymbol{y}$	$12 - 6 = 6$
C. Total	$\boldsymbol{y}^\top(\boldsymbol{I} - \boldsymbol{P}_1)\boldsymbol{y}$	$12 - 1 = 11$

Connection to the ANOVA Table

drug main effect

diet main effect

SS			DF
$\boldsymbol{y}^\top(\boldsymbol{P}_2-\boldsymbol{P}_1)\boldsymbol{y}$	$=$	$(\boldsymbol{c}_1^\top\widehat{\boldsymbol{\beta}})^2/\boldsymbol{c}_1^\top(\boldsymbol{X}^\top\boldsymbol{X})^-\boldsymbol{c}_1$	1
$\boldsymbol{y}^\top(\boldsymbol{P}_3-\boldsymbol{P}_2)\boldsymbol{y}$	$=$	$(\boldsymbol{c}_2^\top\widehat{\boldsymbol{\beta}})^2/\boldsymbol{c}_2^\top(\boldsymbol{X}^\top\boldsymbol{X})^-\boldsymbol{c}_2+(\boldsymbol{c}_3^\top\widehat{\boldsymbol{\beta}})^2/\boldsymbol{c}_3^\top(\boldsymbol{X}^\top\boldsymbol{X})^-\boldsymbol{c}_3$	2
$\boldsymbol{y}^\top(\boldsymbol{P}_4-\boldsymbol{P}_3)\boldsymbol{y}$	$=$	$(\boldsymbol{c}_4^\top\widehat{\boldsymbol{\beta}})^2/\boldsymbol{c}_4^\top(\boldsymbol{X}^\top\boldsymbol{X})^-\boldsymbol{c}_4+(\boldsymbol{c}_5^\top\widehat{\boldsymbol{\beta}})^2/\boldsymbol{c}_5^\top(\boldsymbol{X}^\top\boldsymbol{X})^-\boldsymbol{c}_5$	2
$\boldsymbol{y}^\top(\boldsymbol{I}-\boldsymbol{P}_4)\boldsymbol{y}$			6
$\boldsymbol{y}^\top(\boldsymbol{I}-\boldsymbol{P}_1)\boldsymbol{y}$			11

interaction

ANOVA Table with Additional Partitioning

Source	SS	DF
Diet	$(\boldsymbol{c}_1^\top \widehat{\boldsymbol{\beta}})^2 / \boldsymbol{c}_1^\top (\boldsymbol{X}^\top \boldsymbol{X})^- \boldsymbol{c}_1$	1
Drug 1 − Drug 2	$(\boldsymbol{c}_2^\top \widehat{\boldsymbol{\beta}})^2 / \boldsymbol{c}_2^\top (\boldsymbol{X}^\top \boldsymbol{X})^- \boldsymbol{c}_2$	1
(Drug 1 + Drug 2)/2 − Drug 3	$(\boldsymbol{c}_3^\top \widehat{\boldsymbol{\beta}})^2 / \boldsymbol{c}_3^\top (\boldsymbol{X}^\top \boldsymbol{X})^- \boldsymbol{c}_3$	1
Diet × (Drug 1 − Drug 2)	$(\boldsymbol{c}_4^\top \widehat{\boldsymbol{\beta}})^2 / \boldsymbol{c}_4^\top (\boldsymbol{X}^\top \boldsymbol{X})^- \boldsymbol{c}_4$	1
Diet × [(Drug 1 + Drug 2)/2 − Drug 3]	$(\boldsymbol{c}_5^\top \widehat{\boldsymbol{\beta}})^2 / \boldsymbol{c}_5^\top (\boldsymbol{X}^\top \boldsymbol{X})^- \boldsymbol{c}_5$	1
Error	$\boldsymbol{y}^\top (\boldsymbol{I} - \boldsymbol{P}_4) \boldsymbol{y}$	6
C.Total	$\boldsymbol{y}^\top (\boldsymbol{I} - \boldsymbol{P}_1) \boldsymbol{y}$	11

Additional Partitioning of ANOVA Sums of Squares

The previous example shows how the Drug and Diet $\times$ Drug sums of squares can each be partitioned into two single-degree of freedom sums of squares corresponding to estimable orthogonal contrasts.

More generally, any ANOVA sum of squares with q degrees of freedom can be partitioned into q single-degree-of-freedom sums of squares corresponding to q estimable orthogonal linear combinations $\boldsymbol{c}_1^\top \boldsymbol{\beta}, \ldots, \boldsymbol{c}_q^\top \boldsymbol{\beta}$.

Proof of ANOVA Partitioning

To see that such a partitioning is always possible, first recall from the last set of slides that an ANOVA sum of squares

$$\boldsymbol{y}^\top(\boldsymbol{P}_{j+1} - \boldsymbol{P}_j)\boldsymbol{y} = \widehat{\boldsymbol{\beta}}^\top \boldsymbol{C}^\top[\boldsymbol{C}(\boldsymbol{X}^\top\boldsymbol{X})^- \boldsymbol{C}^\top]^{-1}\boldsymbol{C}\widehat{\boldsymbol{\beta}},$$

where $\boldsymbol{C}$ is any matrix $q \times p$ matrix of rank $q = r_{j+1} - r_j$ whose row space is the same as the row space of $(\boldsymbol{P}_{j+1} - \boldsymbol{P}_j)\boldsymbol{X}$.

Proof of ANOVA Partitioning (continued)

Now let $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_q$ be an orthogonal basis for $\mathcal{C}(\boldsymbol{P}_{j+1} - \boldsymbol{P}_j)$.

Then, $\forall\ k = 1, \ldots, q$,

$$\boldsymbol{a}_k \in \mathcal{C}(\boldsymbol{P}_{j+1} - \boldsymbol{P}_j) \implies \exists\ \boldsymbol{v}_k \in \boldsymbol{a}_k = (\boldsymbol{P}_{j+1} - \boldsymbol{P}_j)\boldsymbol{v}_k.$$

$\forall\ k = 1, \ldots, q$, let $\boldsymbol{c}_k^\top = \boldsymbol{a}_k^\top \boldsymbol{X}$. Then $\boldsymbol{c}_k^\top \boldsymbol{\beta}$ is estimable.

Also, $\forall\ k \neq \ell$,

$$\begin{aligned}
\boldsymbol{c}_k^\top (\boldsymbol{X}^\top \boldsymbol{X})^- \boldsymbol{c}_\ell &= \boldsymbol{a}_k^\top \boldsymbol{X} (\boldsymbol{X}^\top \boldsymbol{X})^- \boldsymbol{X}^\top \boldsymbol{a}_\ell \\
&= \boldsymbol{a}_k^\top \boldsymbol{P}_X \boldsymbol{a}_\ell = \boldsymbol{a}_k^\top \boldsymbol{P}_X (\boldsymbol{P}_{j+1} - \boldsymbol{P}_j)\boldsymbol{v}_\ell \\
&= \boldsymbol{a}_k^\top (\boldsymbol{P}_{j+1} - \boldsymbol{P}_j)\boldsymbol{v}_\ell = \boldsymbol{a}_k^\top \boldsymbol{a}_\ell = 0
\end{aligned}$$

so that $\boldsymbol{c}_1^\top \boldsymbol{\beta}, \ldots, \boldsymbol{c}_q^\top \boldsymbol{\beta}$ are orthogonal linear combinations.

Proof of ANOVA Partitioning (continued)

Let $\boldsymbol{C}^\top = [\boldsymbol{c}_1, \ldots, \boldsymbol{c}_q] = [\boldsymbol{X}^\top \boldsymbol{a}_1, \ldots, \boldsymbol{X}^\top \boldsymbol{a}_q] = \boldsymbol{X}^\top [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_q]$

Is the row space of $\boldsymbol{C}$ the same as the row space of $(\boldsymbol{P}_{j+1} - \boldsymbol{P}_j)\boldsymbol{X}$? Equivalently, is $\mathcal{C}(\boldsymbol{C}^\top) = \mathcal{C}(\boldsymbol{X}^\top(\boldsymbol{P}_{j+1} - \boldsymbol{P}_j))$?

$$
\begin{aligned}
\boldsymbol{C}^\top &= \boldsymbol{X}^\top [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_q] = \boldsymbol{X}^\top [(\boldsymbol{P}_{j+1} - \boldsymbol{P}_j)\boldsymbol{v}_1, \ldots, (\boldsymbol{P}_{j+1} - \boldsymbol{P}_j)\boldsymbol{v}_q] \\
&= \boldsymbol{X}^\top (\boldsymbol{P}_{j+1} - \boldsymbol{P}_j)[\boldsymbol{v}_1, \ldots, \boldsymbol{v}_q] \\
\Longrightarrow \quad & \mathcal{C}(\boldsymbol{C}^\top) \subseteq \mathcal{C}(\boldsymbol{X}^\top (\boldsymbol{P}_{j+1} - \boldsymbol{P}_j)).
\end{aligned}
$$

Also, $\boldsymbol{X}^\top(\boldsymbol{P}_{j+1} - \boldsymbol{P}_j) = \boldsymbol{X}^\top [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_q]\boldsymbol{M} = \boldsymbol{C}^\top \boldsymbol{M}$, for some $q \times n$ matrix $\boldsymbol{M}$ because $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_q$ comprises a basis for $\mathcal{C}(\boldsymbol{P}_{j+1} - \boldsymbol{P}_j)$. Thus, $\mathcal{C}(\boldsymbol{X}^\top(\boldsymbol{P}_{j+1} - \boldsymbol{P}_j)) \subseteq \mathcal{C}(\boldsymbol{C}^\top)$.

Proof of ANOVA Partitioning (continued)

Thus, we have $\mathcal{C}(\boldsymbol{X}^\top(\boldsymbol{P}_{j+1} - \boldsymbol{P}_j)) = \mathcal{C}(\boldsymbol{C}^\top)$.

It follows that

$$\begin{aligned}
\boldsymbol{y}^\top(\boldsymbol{P}_{j+1} - \boldsymbol{P}_j)\boldsymbol{y} &= \widehat{\boldsymbol{\beta}}^\top \boldsymbol{C}^\top[\boldsymbol{C}(\boldsymbol{X}^\top\boldsymbol{X})^-\boldsymbol{C}^\top]^{-1}\boldsymbol{C}\widehat{\boldsymbol{\beta}} \\
&= \sum_{k=1}^{q} \widehat{\boldsymbol{\beta}}^\top \boldsymbol{c}_k[\boldsymbol{c}_k^\top(\boldsymbol{X}^\top\boldsymbol{X})^-\boldsymbol{c}_k]^{-1}\boldsymbol{c}_k^\top\widehat{\boldsymbol{\beta}} \\
&= \sum_{k=1}^{q} (\boldsymbol{c}_k^\top\widehat{\boldsymbol{\beta}})^2/\boldsymbol{c}_k^\top(\boldsymbol{X}^\top\boldsymbol{X})^-\boldsymbol{c}_k. \qquad \square
\end{aligned}$$

ANOVA Partitioning is Not Always Necessary

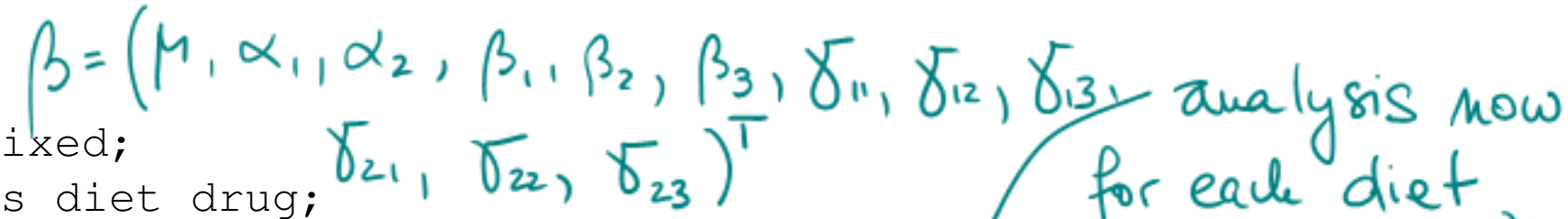
Just because we can partition ANOVA sums of squares does not mean we need to partition ANOVA sums of squares.

The goals of an analysis typically involve constructing estimates or conducting tests of scientific interest.

The tests of scientific interest do not necessarily involve orthogonal linear combinations.

For example, suppose the goal of the researchers who conducted the diet-drug study is to determine which of the three drugs is best for enhancing weight gain of pigs on each diet.

SAS Code




```
proc mixed;
  class diet drug;
  model weightgain=diet drug diet*drug;
  lsmeans diet*drug / slice=diet;
  estimate 'drug 1 - drug 2 for diet 1'
           drug 1 -1  0 diet*drug 1 -1  0  0  0  0 / cl;
  estimate 'drug 1 - drug 3 for diet 1'
           drug 1  0 -1 diet*drug 1  0 -1  0  0  0;
  estimate 'drug 2 - drug 3 for diet 1'
           drug 0  1 -1 diet*drug 0  1 -1  0  0  0;
  estimate 'drug 1 - drug 2 for diet 2'
           drug 1 -1  0 diet*drug 0  0  0  1 -1  0;
  estimate 'drug 1 - drug 3 for diet 2'
           drug 1  0 -1 diet*drug 0  0  0  1  0 -1;
  estimate 'drug 2 - drug 3 for diet 2'
           drug 0  1 -1 diet*drug 0  0  0  0  1 -1;
run;
```

effect size

SAS Output

The Mixed Procedure

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
diet	1	6	61.96	0.0002
drug	2	6	6.04	0.0365
diet*drug	2	6	5.01	0.0526

SAS Output

Some evidence that drug might have an effect on weightgain

Tests of Effect Slices

Effect	diet	Num DF	Den DF	F Value	Pr > F
diet*drug	1	2	6	9.59	0.0135
diet*drug	2	2	6	1.46	0.3042

while for diet 2 we lack evidence that the drugs differ in their effect

SAS Output

Drug 1 vs. Drug 3 appear to be more different.

Least Squares Means

Effect	diet	drug	Estimate	Standard Error
diet*drug	1	1	42.5000	0.7832
diet*drug	1	2	40.0500	0.7832
diet*drug	1	3	37.6500	0.7832
diet*drug	2	1	35.7000	0.7832
diet*drug	2	2	33.9500	0.7832
diet*drug	2	3	35.4500	0.7832

very similar

SAS Output

because these contrasts are not orthogonal, adjustments for multiple comparisons might be in place: Tukey, Bonferoni: $\alpha^* = \frac{0.05}{6}$

Estimates

Label		Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha
drug 1-drug 2	diet 1	2.4500	1.1075	6	2.21	0.0689	0.05
drug 1-drug 3	diet 1	4.8500	1.1075	6	4.38	0.0047	0.05
drug 2-drug 3	diet 1	2.4000	1.1075	6	2.17	0.0734	0.05
drug 1-drug 2	diet 2	1.7500	1.1075	6	1.58	0.1652	0.05
drug 1-drug 3	diet 2	0.2500	1.1075	6	0.23	0.8289	0.05
drug 2-drug 3	diet 2	-1.5000	1.1075	6	-1.35	0.2244	0.05

Label		Lower	Upper
drug 1-drug 2	diet 1	-0.2601	5.1601
drug 1-drug 3	diet 1	2.1399	7.5601
drug 2-drug 3	diet 1	-0.3101	5.1101
drug 1-drug 2	diet 2	-0.9601	4.4601
drug 1-drug 3	diet 2	-2.4601	2.9601
drug 2-drug 3	diet 2	-4.2101	1.2101

include zero

Main Conclusions

For pigs on diet 1, treatment with drug 1 led to significantly greater mean weight gain than treatment with drug 3.

No other differences in mean weight gain between drugs within either diet were statistically significant.

Comments on the Analysis

Note that the main analysis focuses on pairwise comparisons of drugs within each diet.

This involves a set of six contrasts, but the contrasts are not pairwise orthogonal within either diet.

The sums of squares for these contrasts do not add up to any ANOVA sums of squares, but they are the contrasts that best address the researchers' questions.

If we want to control the probability of one or more type I errors, we could use Bonferroni's method. In this case, the adjustment for multiple testing would not change the conclusions.

Comments on the Cell Means vs. Additive Model

We used the cell means model for analysis even though the interactions were not significant at the 0.05 level.

I tend to prefer the cell means model in experiments with a full-factorial treatment design even if interactions are not significant.

The cell means model is less restrictive than an additive model.

The cell means model estimator of error variance σ^2 is not inflated by incorrectly specifying an additive mean structure when the additive mean structure is too restrictive.

Comments on the Cell Means vs. Additive Model

Using the cell means model honors the treatment structure.

Using the cell means model avoids problems with using the data once to select a model and a second time to perform inference.

Some other statisticians may favor a different strategy, especially in experiments with many factors or few degrees of freedom for error.