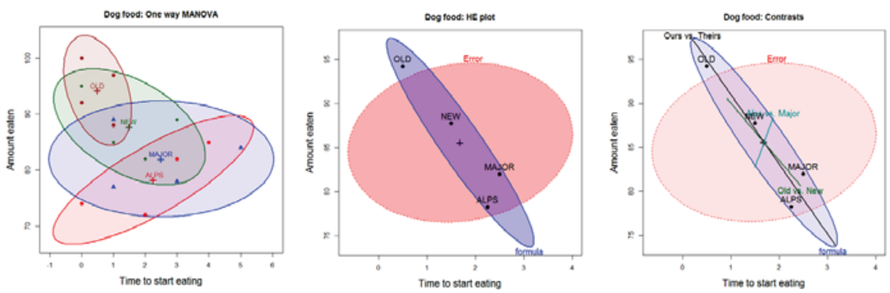



# Multivariate analysis of variance

Michael Friendly

Psychology 6140

---

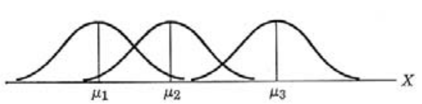
# One-way MANOVA

- *p* responses, 1 “factor” (IV), *g* groups
  - $H_0$: $\underline{\boldsymbol{\mu}}_1 = \underline{\boldsymbol{\mu}}_2 = \ldots \underline{\boldsymbol{\mu}}_g$
  - $H_1$: at least one group centroid is different
- Assumptions:
  - Independent groups, independent observations
  - Responses are independent, multivariate normal w/in each group
  - Pop. covariance matrices are **equal** across groups
    - $H_0$: $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \ldots = \boldsymbol{\Sigma}_g$
    - ($\boldsymbol{\Sigma}$ estimated by E / dfe)
    - (tested by e.g., Box’s test, proc discrim pool=test or **heplots::boxM**)
  - $\rightarrow \mathbf{y}_{ij\ (p \times 1)} \sim N(\boldsymbol{\mu}_j, \boldsymbol{\Sigma})$



---

# One-way ANOVA vs. MANOVA

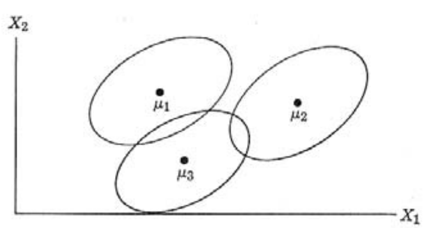


Assume equal within-group variances

How do means differ?

**Figure 8.1.** The simple anova situation, when the differences among the populations are “real.”

source: Cooley & Lohnes ((1971)



Assume equal within-group variance-covariance matrices

How do centroids differ?

How many dimensions?

**Figure 8.2.** The simple manova situation, when the differences among the populations are “real.”



---

# Fundamental ideas

- General linear model

$$\mathbf{Y}_{n\times p} = \mathbf{X}_{n\times q}\ \mathbf{B}_{q\times p} + \mathcal{E}_{n\times p}$$

- Tests: General linear hypothesis

$$H_0: \ \mathbf{L}\ \mathbf{B}\ \mathbf{M} = \mathbf{0} \quad \rightarrow \text{SSP matrices for } \mathbf{H} \ \& \ \mathbf{E}$$

- How big is **H** relative to **E**?
  - Eigenvalues, $\lambda_i$ of $HE^{-1}$ or $\theta_i$ of $H(H+E)^{-1}$
  - → Wilks’ Λ, Pillai & Hotelling trace, Roy’s test
  - # of large dimensions (aspects of responses)
- HE plots: visualize multivariate tests
  - Shows size of dimensions (aspects of responses)
  - Relation to response variables

# GLM: the design matrix (X)

- In the full GLM, the design matrix (**X**) may consist of:
  - Quantitative regressors: age, income, education
  - Transformed regressors: √age, log(income)
  - Polynomial terms: $\text{age}^2$, $\text{age}^3$, …
  - Categorical predictors ("factors", class variables): treatment (control, drug A, drug B), sex
  - Interactions: treatment * sex, age * sex
- ANOVA/MANOVA: Where does **X** come from?



# GLM: Factors and contrasts

- In the GLM, the design matrix **X** can be constructed entirely from contrasts for each factor separately;
- # (contrasts) = df

```
proc glm;
    class Treat Sex;
    model y1 - y3 = Treat Sex Treat*Sex
```

Kronecker product

$$\mathbf{X}_{treat} = \begin{pmatrix}1 & 1 & 0\\ 1 & 0 & 1\\ 1 & -1 & -1\end{pmatrix} \qquad \mathbf{X}_{sex} = \begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix} \qquad \underbrace{\mathbf{X}_{treat\times sex} = \mathbf{X}_{treat} \otimes \mathbf{X}_{sex}}_{2 \times 1 \text{ df}}$$

main effect    2 df    1 df



# What are contrasts?

- For a factor with $r$ levels, a contrast is a weighted sum, L, of the means, with weights, **c**, that sum to zero

$$L = \mathbf{c}'\,\boldsymbol{\mu} = \Sigma\, c_i\, \mu_i \quad \text{such that} \quad \Sigma\, c_i = 0$$

$r$=4

$$\begin{aligned}
L_1 &= (\mu_1 + \mu_2) - (\mu_3 + \mu_4) &&\rightarrow \mathbf{c}_1 = (1 \;\; 1 \;\; -1 \;\; -1)'\\
L_2 &= \mu_1 - \mu_2 &&\rightarrow \mathbf{c}_2 = (1 \;\; -1 \;\; 0 \;\; 0)'\\
L_3 &= \mu_3 - \mu_4 &&\rightarrow \mathbf{c}_3 = (0 \;\; 0 \;\; 1 \;\; -1)'
\end{aligned}$$

- Any $r$-1 linearly independent contrasts → same overall test



# Why contrasts work

- The inner product **c' μ** assesses the degree to which the means in **μ** have the same pattern as the weights in **c**.
  - It is 0 if they are "uncorrelated"
  - It is maximal if they are linearly related

```
# create linear, quadratic and cubic contrasts
c1 <- c(-3, -1, 1, 3)
c2 <- c(1, -1, -1, 1)
c3 <- c(-1, 3, -3, 1)
C <- rbind(c1, c2, c3)
rownames(C) <- c('c1 (lin)', 'c2 (quad)', 'c3 (cubic)')

# data-- means for 4 groups with different patterns
y1 <- c(10, 20, 30, 40)  # linear means
y2 <- c(20, 40, 40, 20)  # quadratic means
y3 <- (y1 + y2)/2        # both
```

```
> C
           [,1] [,2] [,3] [,4]
c1 (lin)     -3   -1    1    3
c2 (quad)     1   -1   -1    1
c3 (cubic)   -1    3   -3    1
```

```
> C %*% y1     # Contrasts with y1
c1 (lin)    100
c2 (quad)     0
c3 (cubic)    0
> C %*% y2      #Contrasts with y2
c1 (lin)      0
c2 (quad)   -40
c3 (cubic)    0
> C %*% y3      #Contrasts with y3
c1 (lin)     50
c2 (quad)   -20
c3 (cubic)    0
```

# Properties of contrasts

- Associated with every contrast is a 1 df sum of squares, $SS_L$ or rank=1 $SSP_H$ = **H** matrix from the GLH
- Two contrasts are orthogonal if $\mathbf{c}'_i \, \mathbf{c}_j = 0$ (and sample sizes are equal)
- For *r-1* orthogonal contrasts, the $SS_{L_i}$ or $SSP_{H_i}$ add to the SS for the overall hypothesis

$$SSP_H = SSP_{H1} + SSP_{H2} + \; \ldots + SSP_{H(r-1)}$$

- Well chosen contrasts facilitate interpretation of group diff<sup>ces</sup> (vs. all pairwise tests)
- *A priori* contrasts can be tested without adjusting α level



The **X** matrix for a factor can always be represented by a set of *r-1* contrasts, combined with the unit vector

$$\mathbf{X}_{(r \times r)} = (\mathbf{1}, \mathbf{C})$$

$$\mathbf{X} = \begin{array}{c} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & -1 & 0 \\ 1 & -1 & 0 & 1 \\ 1 & -1 & 0 & -1 \end{pmatrix} \\ \quad\;\; c1 \;\; c2 \;\; c3 \end{array}$$

Some special contrasts:

**Deviation contrasts**

$$\mathbf{C} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -1 & -1 \end{pmatrix}$$

each treatment vs control or baseline [not orthogonal]

**Helmert contrasts**

$$\mathbf{C} = \begin{pmatrix} 3 & 0 & 0 \\ -1 & 2 & 0 \\ -1 & -1 & 1 \\ -1 & -1 & -1 \end{pmatrix}$$

ordered treatments: each vs all the rest [always orthogonal]

**Polynomial contrasts**

$$\mathbf{C} = \begin{array}{c} \begin{pmatrix} -3 & 1 & -1 \\ -1 & -1 & 3 \\ 1 & -1 & -3 \\ 3 & 1 & 1 \end{pmatrix} \\ \text{lin} \;\; \text{quad} \;\; \text{cubic} \end{array}$$

quantitative treatment levels [orthogonal]



# Polynomial contrasts

Orthogonal polynomial contrasts are constructed by orthogonalizing the matrix

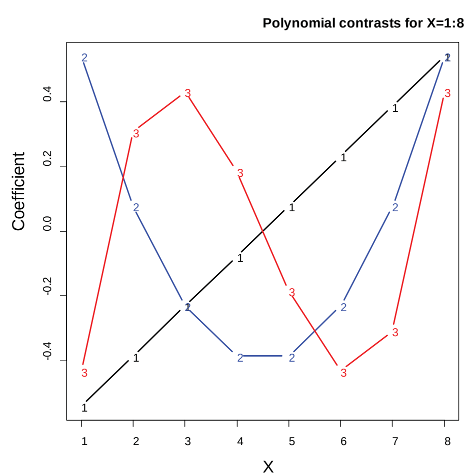
$$\mathbf{C} = (1, X, X^2, X^3, \ldots)$$

e.g. using Gram-Schmidt

These provide tests of trends, similar to poly. regression.

```
> outer(1:8, 0:3, `^`)
     [,1] [,2] [,3] [,4]
[1,]    1    1    1    1
[2,]    1    2    4    8
[3,]    1    3    9   27
[4,]    1    4   16   64
[5,]    1    5   25  125
[6,]    1    6   36  216
[7,]    1    7   49  343
[8,]    1    8   64  512
```

lin quad cubic






# Using contrasts in R

- R has 4 basic functions for generating contrasts for a factor
  - Dummy coding, aka “reference level”, “treatment” contrasts
  - Deviation coding, aka “sum-to-zero” constraints
  - Polynomial contrasts for an ordered/quantitative factor
  - Helmert contrasts for ordered factor comparisons
- Defaults are set separately for unordered and ordered factors
- Define your own by assigning a matrix to `contrasts(myfactor)`
- These affect the tests of coefficients, but not overall tests

```
> contr.treatment(4)
  2 3 4
1 0 0 0
2 1 0 0
3 0 1 0
4 0 0 1
```

```
> contr.sum(4)
  [,1] [,2] [,3]
1    1    0    0
2    0    1    0
3    0    0    1
4   -1   -1   -1
```

```
> contr.poly(4)
             .L   .Q         .C
[1,] -0.6708  0.5 -0.2236
[2,] -0.2236 -0.5  0.6708
[3,]  0.2236 -0.5 -0.6708
[4,]  0.6708  0.5  0.2236
```

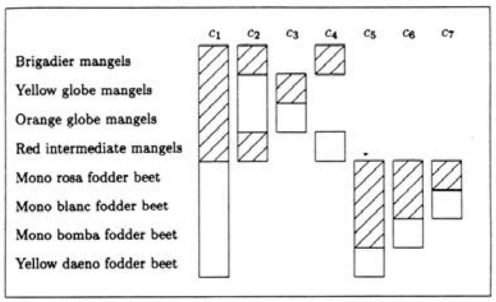
```
> options("contrasts")
$contrasts
        unordered           ordered
"contr.treatment"      "contr.poly"
```

```
> contr.helmert(4)
  [,1] [,2] [,3]
1   -1   -1   -1
2    1   -1   -1
3    0    2   -1
4    0    0    3
```

See:http://www.ats.ucla.edu/stat/r/library/contrast_coding.htm

# Nested dichotomies

- Orthogonal contrasts can always be generated as nested dichotomies
- They correspond to independent research questions
- Sums of squares decompose the overall effect




| Treatment | | $c_1$ | $c_2$ | $c_3$ | $c_4$ | $c_5$ | $c_6$ | $c_7$ |
|---|---|---|---|---|---|---|---|---|
| Brigadier mangels | $\mu_1$ | 1 | 1 | 0 | 1 | 0 | 0 | 0 |
| York globe mangels | $\mu_2$ | 1 | −1 | 1 | 0 | 0 | 0 | 0 |
| Orange globe mangels | $\mu_3$ | 1 | −1 | −1 | 0 | 0 | 0 | 0 |
| Red intermediate mangels | $\mu_4$ | 1 | 1 | 0 | −1 | 0 | 0 | 0 |
| Mono rosa fodder beet | $\mu_5$ | −1 | 0 | 0 | 0 | 1 | 1 | 1 |
| Mono blanc fodder beet | $\mu_6$ | −1 | 0 | 0 | 0 | 1 | 1 | −1 |
| Mono bomba fodder beet | $\mu_7$ | −1 | 0 | 0 | 0 | 1 | −2 | 0 |
| Yellow daeno fodder beet | $\mu_7$ | −1 | 0 | 0 | 0 | −3 | 0 | 0 |



More examples

**Three treatments**

| | | $c_1$ | $c_2$ |
|---|---|---|---|
| 1. Control sheep | (No drench) | −2 | 0 |
| 2. Sheep drenched once | (Drench) | 1 | −1 |
| 3. Sheep drenched twice | (Drench) | 1 | 1 |

Drench vs No drench
Drench once vs twice

**Four treatments**

| | | $c_1$ | $c_2$ | $c_3$ |
|---|---|---|---|---|
| 1. Fan heater, brand A | (Convection) | 1 | 1 | 0 |
| 2. Fan heater, brand B | (Convection) | 1 | −1 | 0 |
| 3. Bar heater, brand P | (Radiation) | −1 | 0 | 1 |
| 4. Bar heater, brand Q | (Radiation) | −1 | 0 | −1 |

Conv vs Rad
A vs B
P vs Q

| | | | | | |
|---|---|---|---|---|---|
| 1. Lupins | (Legume) | | −3 | 0 | 0 |
| 2. Mustard | (Non legume) | (Noncereal) | 1 | −2 | 0 |
| 3. Barley | (Non legume) | (Cereal) | 1 | 1 | −1 |
| 4. Oats | (Non legume) | (Cereal) | 1 | 1 | 1 |

Like Helmert

**Five Treatments**

| | | | $c_1$ | $c_2$ | $c_3$ | $c_4$ |
|---|---|---|---|---|---|---|
| 1. Dacron | (Synthetic Fibre) | | 3 | −1 | 0 | 0 |
| 2. Terylene | (Synthetic Fibre) | | 3 | 1 | 0 | 0 |
| 3. Cotton | (Natural Fibre) | (Plant fibre) | −2 | 0 | −2 | 0 |
| 4. Angora | (Natural Fibre) | (Animal fibre) | −2 | 0 | 1 | 1 |
| 5. Wool | (Natural Fibre) | (Animal fibre) | −2 | 0 | 1 | −1 |

2 subsets

| | | | | | |
|---|---|---|---|---|---|
| 1. Control | (No herbicide) | −4 | 0 | 0 | 0 |
| 2. Systemic herbicide A | (Herbicide) | 1 | 1 | 1 | 0 |
| 3. Systemic herbicide B | (Herbicide) | 1 | 1 | −1 | 0 |
| 4. Contact herbicide X | (Herbicide) | 1 | −1 | 0 | 1 |
| 5. Contact herbicide Y | (Herbicide) | 1 | −1 | 0 | −1 |

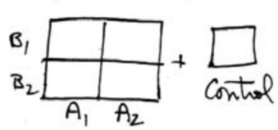
2x2 + control



# Factorial designs + control group(s)

- Contrasts provide a way to analyze complex designs: A x B + control group(s)

eg. 2x2 + 1




| | $C_1$ | $C_2$ | $C_3$ | $C_4$ |
|---|---|---|---|---|
| Control | −4 | 0 | 0 | 0 |
| $A_1B_1$ | 1 | 1 | 1 | 1 |
| $A_1B_2$ | 1 | 1 | −1 | −1 |
| $A_2B_1$ | 1 | −1 | 1 | −1 |
| $A_2B_2$ | 1 | −1 | −1 | 1 |
| | | A | B | AB |

Treat as a one-way, 5 group design.

Use contrasts to analyze the A, B and A*B effects



# Interactions from main effects

- For any factorial design, contrasts & **X** matrix columns for interactions are generated from those for the main effects
  - $df_{A*B} = df_A * df_B = (a-1)(b-1)$
  - Contrasts for A*B are the (a-1)(b-1) products of *each* contrast for A with *each* contrast for B
  - They represent differences of differences

2 x 3 design:

| | **(a) Main effect contrasts** | | | **(b) Interaction contrasts** | |
|---|---|---|---|---|---|
| | Factor A | Factor B | | $A \times B$ | |
| | $c_1$ | $c_2$ | $c_3$ | $c_4 = c_1 \times c_2$ | $c_5 = c_1 \times c_3$ |
| $\mu_1$ | −1 | −2 | 0 | 2 | 0 |
| $\mu_2$ | −1 | 1 | −1 | −1 | 1 |
| $\mu_3$ | −1 | 1 | 1 | −1 | −1 |
| $\mu_4$ | 1 | −2 | 0 | −2 | 0 |
| $\mu_5$ | 1 | 1 | −1 | 1 | −1 |
| $\mu_6$ | 1 | 1 | 1 | 1 | 1 |

# Interactions from main effects

- Mathematically, this is generated by the Kronecker product (⊗) of the one-way contrasts

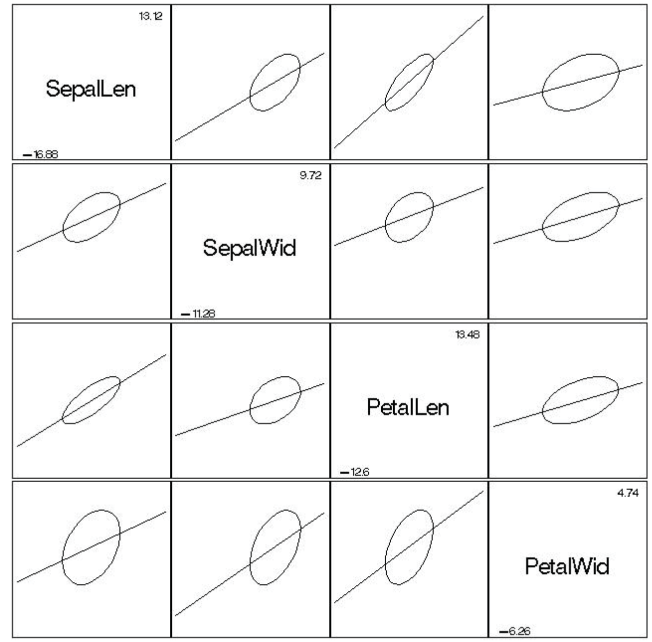
$$C_A \otimes C_B = \begin{pmatrix}-1\\1\end{pmatrix} \otimes \begin{pmatrix}-2&0\\1&-1\\1&1\end{pmatrix} = \begin{pmatrix}-1\times C_B\\1\times C_B\end{pmatrix} = \begin{pmatrix}2&0\\-1&1\\-1&-1\\-2&0\\1&-1\\1&1\end{pmatrix}$$

- The full **X** matrix for any factorial design is the Kronecker product of all one-way **X** matrices

$$\mathbf{X}_{ABCD} = (\mathbf{1}, \mathbf{C}_A) \otimes (\mathbf{1}, \mathbf{C}_B) \otimes (\mathbf{1}, \mathbf{C}_C) \otimes (\mathbf{1}, \mathbf{C}_D)$$



# Partial plots: visualize within-group scatter alone

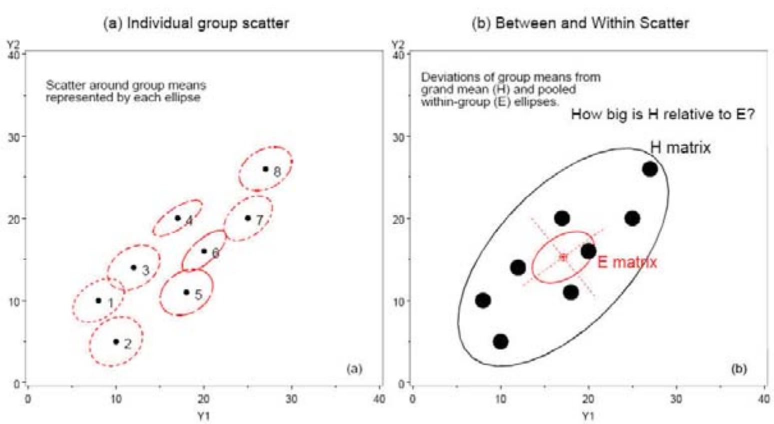



Data ellipses of the *residuals* from the linear model

These show only within-group covariation --- pooling them gives the pooled **E** matrix



# HE plots: visualizing H & E (co)variation

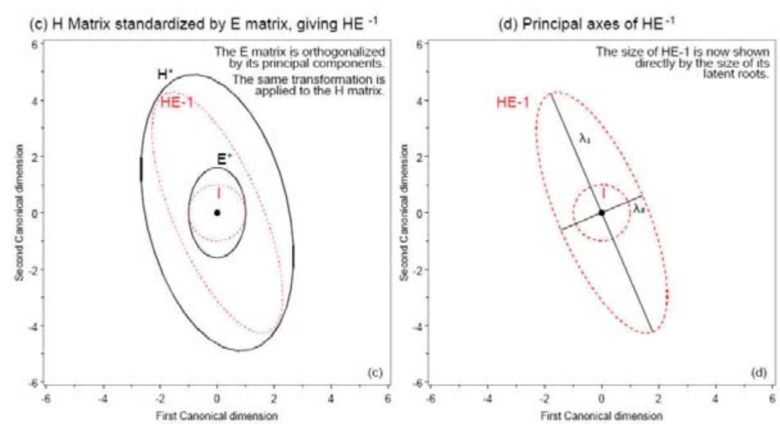



Ideas behind multivariate tests: (a) Data ellipses; (b) ***H*** and ***E*** matrices

- ***H*** ellipse: data ellipse for fitted values, $\hat{\boldsymbol{y}}_{ij} = \bar{\boldsymbol{y}}_j$.
- ***E*** ellipse: data ellipse of residuals, $\hat{\boldsymbol{y}}_{ij} - \bar{\boldsymbol{y}}_j$.



# HE plots: visualizing multivariate tests




Ideas behind multivariate tests: latent roots & vectors of $\boldsymbol{HE}^{-1}$

- $\lambda_i, i = 1, \dots df_h$ show size(s) of ***H*** relative to ***E***.
- latent vectors show canonical directions of maximal difference.

# Example: Dog food– one way design

```
%include data(dogfood);
proc glm data=dogfood order=data;
   class formula;
   model start amount  =  formula / ss3;
                        /*     old new M  A  */
   contrast 'Equality of Groups'  formula 1  0  0 -1,
                                  formula 1  0 -1  0,
                                  formula 1 -1  0  0;

   contrast 'Ours vs. Theirs'     formula 1  1 -1 -1;
   contrast 'Old - New'           formula 1 -1  0  0;
   contrast 'Major vs. Alps'      formula 0  0  1 -1;
   manova h=formula;
   title2 'MANOVA for equality of means';
run;
```

Univariate tests

3 df test = overall test

1 df contrasts

MANOVA tests



## Overall multivariate test:

Note: this is a case where multivariate tests differ. Why??

Characteristic Roots and Vectors of: E Inverse * H, where
H = Type III SSCP Matrix for formula
E = Error SSCP Matrix

| Characteristic Root | Percent | Characteristic Vector V'EV=1 start | amount |
|---|---|---|---|
| 2.03961854 | 98.47 | -0.10279413 | 0.04639418 |
| 0.03174562 | 1.53 | 0.16973304 | 0.02111246 |

MANOVA Test Criteria and F Approximations for
the Hypothesis of No Overall formula Effect

S=2 M=0 N=4.5

| Statistic | Value | F Value | Num DF | Den DF | Pr > F | |
|---|---|---|---|---|---|---|
| Wilks' Lambda | 0.318866 | 2.83 | 6 | 22 | 0.0341 | ✓ |
| Pillai's Trace | 0.701780 | 2.16 | 6 | 24 | 0.0829 | ✗ |
| Hotelling-Lawley Trace | 2.071364 | 3.67 | 6 | 13.032 | 0.0234 | ✓ |
| Roy's Greatest Root | 2.039619 | 8.16 | 3 | 12 | 0.0031 | ✓✓ |



## Multivariate tests of contrasts:

Hypothesis of No Overall Ours vs. Theirs Effect
S=1 M=0 N=4.5

| Statistic | Value | F Value | Num DF | Den DF | Pr > F | |
|---|---|---|---|---|---|---|
| Wilks' Lambda | 0.374715 | 9.18 | 2 | 11 | 0.0045 | ✓ |
| Pillai's Trace | 0.625285 | 9.18 | 2 | 11 | 0.0045 | |
| Hotelling-Lawley Trace | 1.668694 | 9.18 | 2 | 11 | 0.0045 | |
| Roy's Greatest Root | 1.668694 | 9.18 | 2 | 11 | 0.0045 | |

Hypothesis of No Overall Old - New Effect

| Statistic | Value | F Value | Num DF | Den DF | Pr > F | |
|---|---|---|---|---|---|---|
| Wilks' Lambda | 0.752377 | 1.81 | 2 | 11 | 0.2091 | ✗ |
| Pillai's Trace | 0.247623 | 1.81 | 2 | 11 | 0.2091 | |
| Hotelling-Lawley Trace | 0.329121 | 1.81 | 2 | 11 | 0.2091 | |
| Roy's Greatest Root | 0.329121 | 1.81 | 2 | 11 | 0.2091 | |

Hypothesis of No Overall Major vs. Alps Effect

| Statistic | Value | F Value | Num DF | Den DF | Pr > F | |
|---|---|---|---|---|---|---|
| Wilks' Lambda | 0.931490 | 0.40 | 2 | 11 | 0.6768 | |
| Pillai's Trace | 0.068510 | 0.40 | 2 | 11 | 0.6768 | ✗ |
| Hotelling-Lawley Trace | 0.073549 | 0.40 | 2 | 11 | 0.6768 | |
| Roy's Greatest Root | 0.073549 | 0.40 | 2 | 11 | 0.6768 | |



## Univariate tests of contrasts:

Univariate tests give a bleaker story. Why??

Dependent Variable: start   Time to start eating

| Contrast | DF | Contrast SS | Mean Square | F Value | Pr > F | |
|---|---|---|---|---|---|---|
| Equality of Groups | 3 | 9.68750000 | 3.22916667 | 1.50 | 0.2634 | ✗ |
| Ours vs. Theirs | 1 | 7.56250000 | 7.56250000 | 3.52 | 0.0850 | ✗ |
| Old - New | 1 | 2.00000000 | 2.00000000 | 0.93 | 0.3534 | ✗ |
| Major vs. Alps | 1 | 0.12500000 | 0.12500000 | 0.06 | 0.8134 | ✗ |

Dependent Variable: amount   Amount eaten

| Contrast | DF | Contrast SS | Mean Square | F Value | Pr > F | |
|---|---|---|---|---|---|---|
| Equality of Groups | 3 | 585.6875000 | 195.2291667 | 6.00 | 0.0097 | ✓ |
| Ours vs. Theirs | 1 | 473.0625000 | 473.0625000 | 14.55 | 0.0025 | ✓ |
| Old - New | 1 | 84.5000000 | 84.5000000 | 2.60 | 0.1329 | ✗ |
| Major vs. Alps | 1 | 28.1250000 | 28.1250000 | 0.86 | 0.3707 | ✗ |

NB: These are orthogonal contrasts, so

$SS_{Group} = \sum SS_{contrast}$

e.g., 9.69 = 7.56 + 2.00 + 0.125

# Analysis in R

```
# set up special contrasts for formula
contrasts(dogfood$formula) <- matrix(
  c(-1, -1,  1, 1,
     0,  0, -1, 1,
    -1, 1,  0,  0), nrow=4, ncol=3)
contrasts(dogfood$formula)
```

```
> contrasts(dogfood$formula)
      [,1] [,2] [,3]
ALPS    -1    0   -1
MAJOR   -1    0    1
NEW      1   -1    0
OLD      1    1    0
```

Overall MANOVA:

```
dog.mod <- lm(cbind(start, amount) ~ formula, data=dogfood)
Anova(dog.mod)
```

```
> Anova(dog.mod)
Type II MANOVA Tests: Pillai test statistic
        Df test stat approx F num Df den Df Pr(>F)
formula  3     0.702     2.16      6     24  0.083 .
```

Details from:
summary(Anova(dog.mod))

Test contrasts:

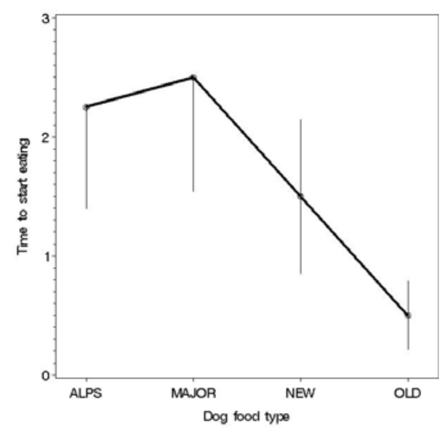
```
# test multivariate contrasts among levels of formula
linearHypothesis(dog.mod, "formula1", title="Ours vs. Theirs")
linearHypothesis(dog.mod, "formula2", title="Old vs. New")
linearHypothesis(dog.mod, "formula3", title="Alps vs. Major")
```

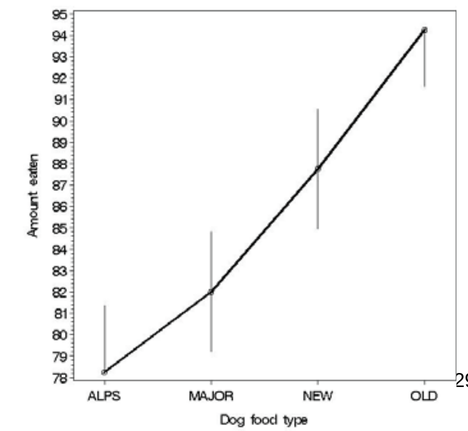


# Visualizing the results

- Univariate plots of means tell a part of the story

```
%meanplot(data=dogfood, var=start amount, class=formula);
```



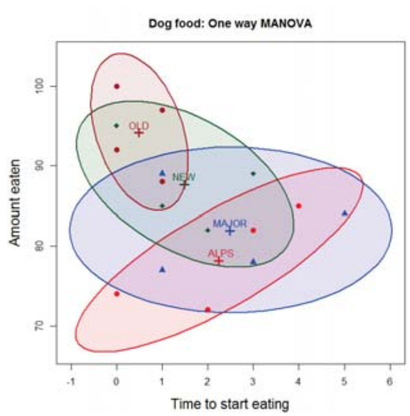



# Visualizing the results: data ellipses

- Data ellipses show between- & within-group variation

```
covEllipses(dogfood[,c("start", "amount")], dogfood$formula)
```

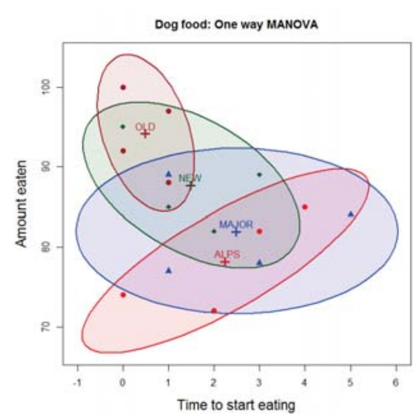
- means of start inversely related to amount
- within-group covariance matrices (Σ) don't look very equal!

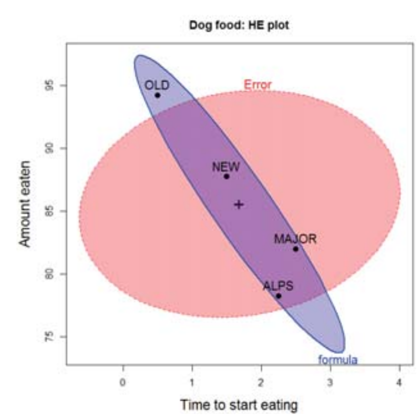





# Visualizing the results: HE plots

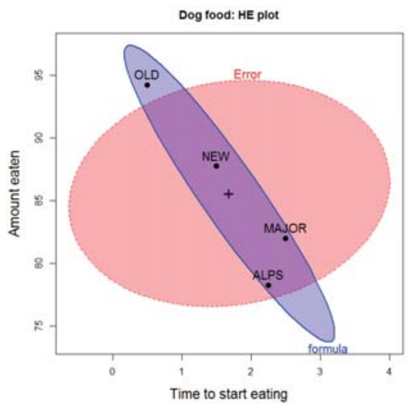
- HE plots show a sufficient visual summary

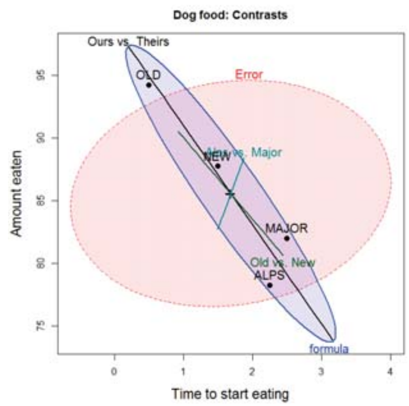


# Visualizing the results: HE plots

- HE plots with contrasts show the breakdown of effects

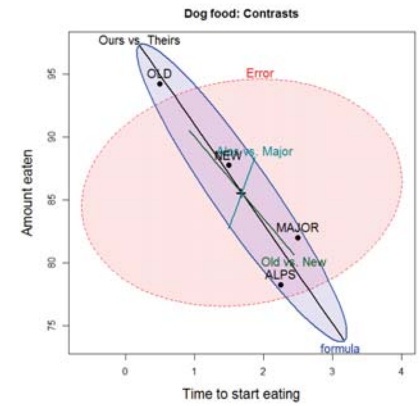





```
linearHypothesis(dog.mod, "formula1", title="Ours vs. Theirs")
linearHypothesis(dog.mod, "formula2", title="Old vs. New")
linearHypothesis(dog.mod, "formula3", title="Alps vs. Major")
heplot(dog.mod, hypotheses=list(
                "Ours vs. Theirs"="formula1",
                "Old vs. New"="formula2",
                "Alps vs. Major"="formula3"))
```

We can easily see that the overall formula effect is largely due to the OursTheirs contrast






# Assumptions: homogeneity of (co)variance

- For univariate *t*-test or ANOVA, we assume equal variance within groups
  - $s^2_1 = s^2_2 = \ldots = s^2_g \rightarrow s^2_{\text{pooled}}$ or MSE
  - Box test or Levine's test often used
  - Visual test: spread-level plot
- Multivariate tests: translates to equality of within-group covariance matrices,
  - $S_1 = S_2 = \ldots = S_g \rightarrow S_{\text{pooled}}$ = E matrix
  - Box test: $H_0 : \Sigma_1 = \Sigma_2 = \ldots = \Sigma_g$

$$V = \frac{\prod | S_i |^{N_i/2}}{| S |_{pooled}^{N/2}} \quad \rightarrow \chi 2 \text{ with } (g-1)p(p+1)/2 \text{ df}$$

  - SAS: proc discrim, **`pool=test`** option
  - R: boxM() in heplots package
  - NB: Box's test *very* susceptible to non-normality
  - Visualize: data ellipses



# Testing homogeneity of (co)variance

**proc discrim** data=dogfood
    **pool=test**;
    class formula;
    var start amount;
    **run**;

```
                  The DISCRIM Procedure
        Test of Homogeneity of Within Covariance Matrices
        Chi-Square        DF     Pr > ChiSq

                  5.689160         9        0.7706

Since the Chi-Square value is not significant at the 0.1 level, a pooled
covariance matrix will be used in the discriminant function.
```

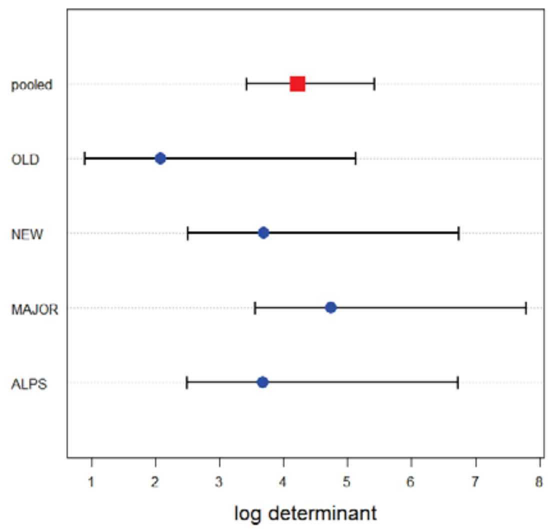
R:

```
dog.mod <- lm(cbind(start, amount) ~ formula, data=dogfood)
boxM(dog.mod)
        Box's M-test for Homogeneity of Covariance Matrices
data:  dogfood[, c("start", "amount")]
Chi-Sq (approx.) = 5.6892, df = 9, p-value = 0.7706
```

## Visualizing Box's M test

```
plot( boxM(dog.mod))
```

In this problem, sample size is too small to detect the differences in the $\mathbf{S}_i$





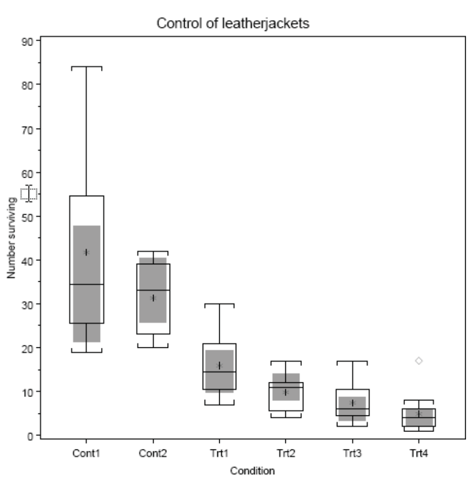

## Spread-level plot: diagnose heterogeneity

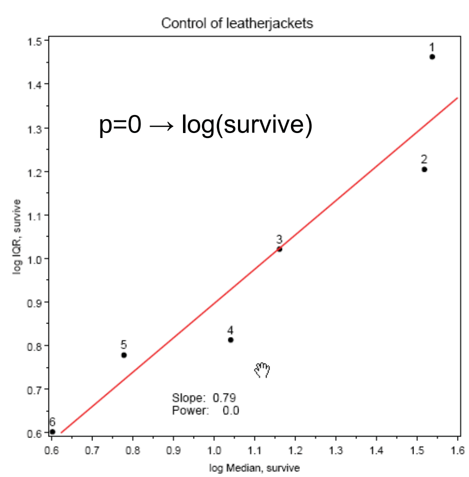
- If within-group variance ~ mean, a plot of log(spread) vs log(level) can determine a transformation, $y \rightarrow y^p$ to make variances approx. equal
  - SAS: sprdplot macro
  - R: car::spreadLevelPlot()
- Plot $\log_{10}$(IQR) vs $\log_{10}$(median)
- If linear with slope=b, → power = 1-b
- There is no multivariate version yet – use on each response variable.



**Ex:** Two controls & 4 insecticides applied to bugs; response=# surviving

```
*-- Boxplot of raw data;
%boxplot(data=control,
     var=survive, varlab=Number surviving,
     class=cond, classlab=Condition, cnotch=graya0);
*-- Spread-level plot: creates the new variable LSURVIVE;
%sprdplot(data=control, var=survive, class=cond, prefix=L);
```



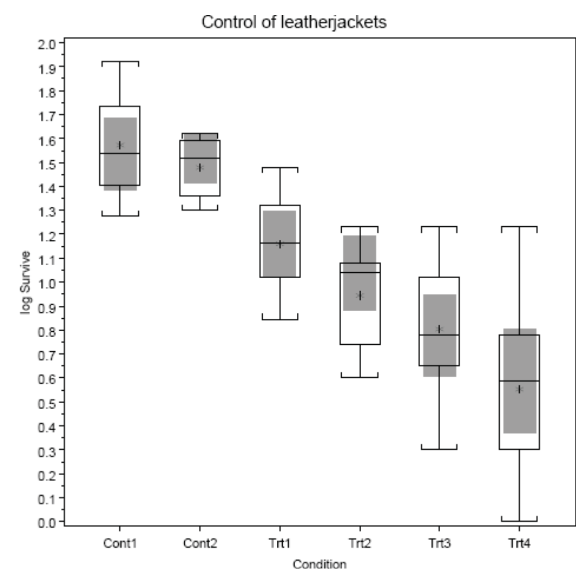



```
*-- Boxplot of transformed data;
%boxplot(data=control, var=lsurvive, class=cond,
       varlab=log Survive, classlab=Condition, cnotch=graya0);
```

log(survive): variances more nearly equal

because the response is a frequency (# surviving),

y → √y is also an option



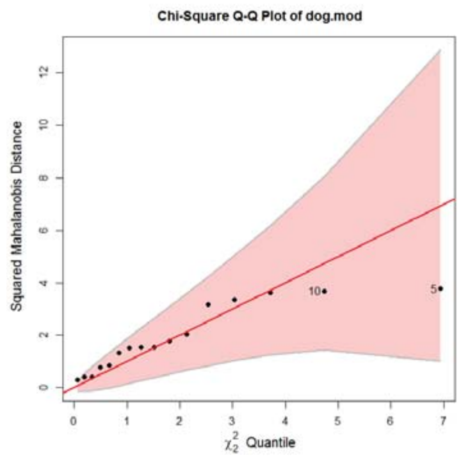
# Assumptions: Normality

- We assume residuals, **E** = **Y-XB** are multivariate normal
- Easiest to check with a $\chi^2$ QQ plot

```
library(heplots)
cqplot(dog.mod, id.n=2)
```

If you make only 1 diagnostic plot, it should be this!

cqplot() also provides robust versions using MVE and MCD estimates






# Ex: Social cognitive measures in schizophrenia

Three diagnostic groups: Schizophrenic, SchizoAffective, Control

Contrasts: (a) Control vs. others; (b) Schizophrenic vs. SchizoAffective

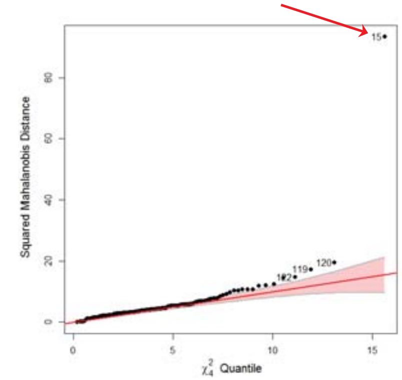
```
library(heplots)
data(SocialCog)
SC.mlm <-  lm(cbind(MgeEmotions,ToM, ExtBias,
             PersBias) ~ Dx, data=SocialCog)
Anova(SC.mlm)
```

```
Type II MANOVA Tests: Pillai test statistic
   Df test stat approx F num Df den Df  Pr(>F)
Dx  2     0.212     3.97      8    268 0.00018 ***
```

```
cqplot(SC.mlm, method="mve", id.n=4)
```

Deleting this obs. changed the *p*-value for the 2nd contrast from 0.045 to 0.074

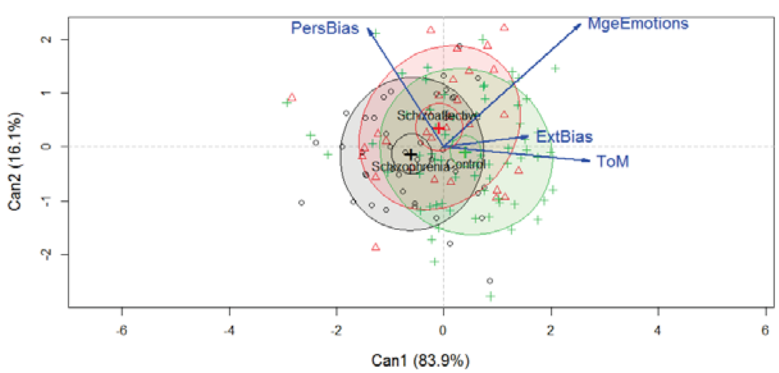
One extreme outlier (data error)






# Canonical HE plots: Low-D views

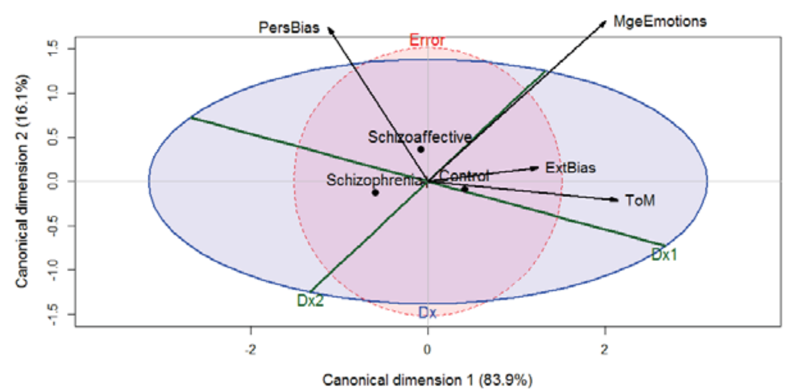
- As with biplot, we can visualize MLM hypothesis variation for *all* responses by projecting $\boldsymbol{H}$ and $\boldsymbol{E}$ into low-rank space.
- Canonical projection: $\boldsymbol{Y}_{n\times p} \mapsto \boldsymbol{Z}_{n\times s} = \boldsymbol{YE}^{-1/2}\boldsymbol{V}$, where $\boldsymbol{V}$ = eigenvectors of $\boldsymbol{HE}^{-1}$.
- This is the view that maximally discriminates among groups, ie max. $\boldsymbol{H}$ wrt $\boldsymbol{E}$ !






# Canonical HE plots: Low-D views

- Canonical HE plot is just the HE plot of canonical scores, $(\boldsymbol{z}_1, \boldsymbol{z}_2)$ in 2D,
- or, $\boldsymbol{z}_1, \boldsymbol{z}_2, \boldsymbol{z}_3$, in 3D.
- As in biplot, we add vectors to show relations of the $\boldsymbol{y}_i$ response variables to the canonical variates.
- variable vectors here are structure coefficients = correlations of variables with canonical scores.

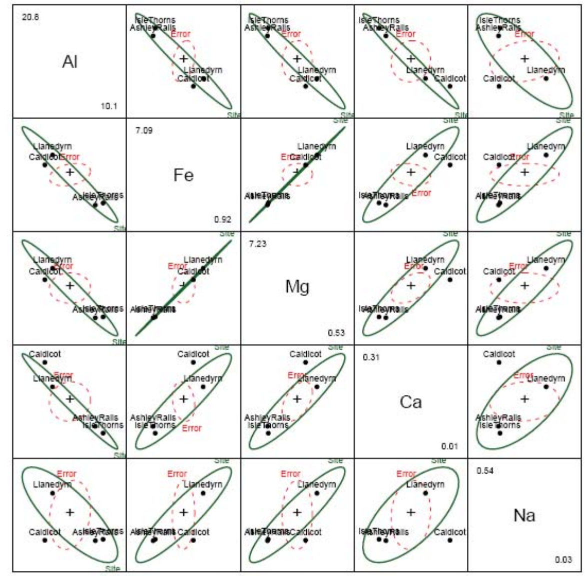


# HE plot example: Romano-British pottery

- Tubb, Parker & Nicholson used atomic absorption spectroscopy to analyze the chemical composition of 26 samples of Romano-British pottery found at four kiln sites in Britain.
    - Sites: Ashley Rails, Caldicot, Isle of Thorns, Llanedryn
    - Variables: aluminum (Al), iron (Fe), magnesium (Mg), calcium (Ca) and sodium (Na)
    - → One-way MANOVA design, 4 groups, 5 responses
- Can the content of Al, Fe, Mg, Ca and Na be used to differentiate the sites?

```
R> library(heplots)
R> pottery.mod <- lm(cbind(Al, Fe, Mg, Ca, Na) ~ Site,
                     data=Pottery)
R> Manova(pottery.mod)
Type II MANOVA Tests: Pillai test statistic
     Df test stat approx F num Df den Df    Pr(>F)
Site  3    1.5539   4.2984     15     60 2.413e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

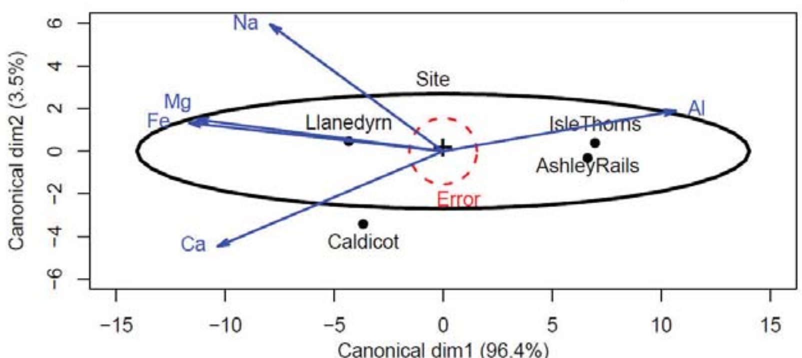
# HE plot matrix: all bivariate views




```
R> pairs(pottery.mod)
```

# Canonical HE plot

- Canonical HE plots provide 2D (3D) visual summary of ***H*** vs. ***E*** variation
- Pottery data: $p = 5$ variables, 4 groups $\mapsto df_H = 3$
- Variable vectors: Fe, Mg and Al contribute to distinguishing (Caldicot, Llandryn) from (Isle Thorns, Ashley Rails): 96.4% of mean variation
- Na and Ca contribute an additional 3.5%. End of story!




HEplot movie



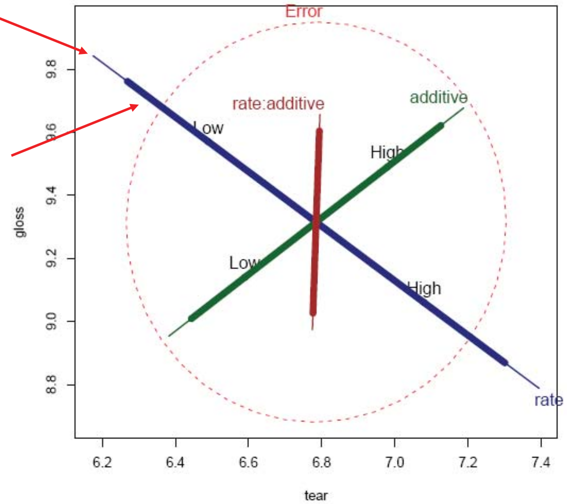
# Two-way MANOVA: Plastic film data

- Data from an experiment to deterimne the optimal conditions for extruding plastic film.
    - Factors: Rate of extrusion (low/high), amount of additive (low/high)
    - Responses: Tear resistance, film gloss, opacity
    - → $2 \times 2$ MANOVA design, 3 responses, $n = 5$ per cell.
- HE plots show main effects, interactions and linear hypotheses in relation to each other

```
R> plastic.mod <- lm(cbind(tear, gloss, opacity) ~
                     rate*additive, data=Plastic)
R> Manova(plastic.mod, test.statistic="Roy")
Type II MANOVA Tests: Roy test statistic
              Df test stat approx F num Df den Df    Pr(>F)
rate           1    1.6188   7.5543      3     14 0.003034 **
additive       1    0.9119   4.2556      3     14 0.024745 *
rate:additive  1    0.2868   1.3385      3     14 0.301782
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

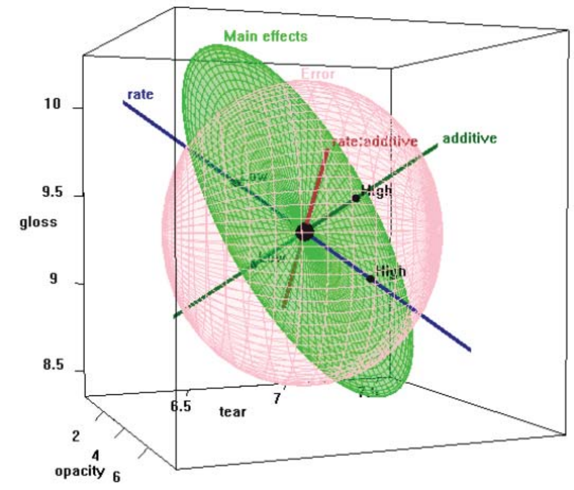
# Two-way MANOVA: Plastic film data

- Main effects & interaction–evidence scaling
- Only effect of `rate` exceeds the ***E*** ellipse. Why?
- Main effects & interaction–effect scaling




# 3D HE plots: Plastic film data

- 3D HE plot shows ellipsoids for ***H*** and ***E*** matrices
- 1 df hypotheses $\mapsto$ lines
- 2 df hypotheses $\mapsto$ ellipses
- `heplot3d` function provides interactive rotation
- This view shows the significant main effects of `rate` and `additive`




# Summary

- MANOVA: Just another GLM
- All tests: $$\mathbf{Y}_{n\times p} = \mathbf{X}_{n\times q}\ \mathbf{B}_{q\times p} + \mathcal{E}_{n\times p}$$ $$H_0 :\ \mathbf{L}\ \mathbf{B}\ \mathbf{M} = \mathbf{0}$$ → SSP matrices for **H** & **E**
- Contrasts: Give **X**, provide interpretable tests
- Test statistics: How big is H relative to E?
- Visualize: HE plots
  - # of large dimensions
  - Relation to response variables
  - Canonical views: space of largest diff$^{ces}$