HE Plot Examples

Michael Friendly

Using heplots version 0.9-12 and candisc version 0.5-21; Date: 2012-05-16

Abstract

This vignette provides some worked examples of the analysis of multivariate linear models (MvLM s) with graphical methods for visualizing results using the heplots package and the candisc package. The emphasis here is on using these methods in R, and understanding how they help reveal aspects of these models that might not be apparent from other graphical displays. No attempt is made to describe the theory of MvLM s or the statistical details behind HE plots and their reduced-rank canonical cousins. For that, see Fox et al. (2009); Friendly (2007, 2006).

Contents

1 MANOVA Designs

1.1 Plastic film data

An experiment was conducted to determine the optimum conditions for extruding plastic film. Three responses, `tear` resistance, film `gloss` and film `opacity` were measured in relation to two factors, `rate` of extrusion and amount of an `additive`, both of these being set to two values, High and Low. The design is thus a 2×2 MANOVA, with $n = 5$ per cell. This example illustrates 2D and 3D HE plots, the difference between "effect" scaling and "evidence" (significance) scaling, and visualizing composite linear hypotheses.

We begin with an overall MANOVA for the two-way MANOVA model. Because each effect has 1 df, all of the multivariate statistics are equivalent, but we specify `test.statistic="Roy"` because Roy's test has a natural visual interpretation in HE plots.

```
> plastic.mod <- lm(cbind(tear, gloss, opacity) ~ rate*additive, data=Plastic)
> Anova(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     1.619     7.55      3     14  0.003 **
additive       1     0.912     4.26      3     14  0.025 *
rate:additive  1     0.287     1.34      3     14  0.302
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

For the three responses jointly, the main effects of `rate` and `additive` are significant, while their interaction is not. In some approaches to testing effects in multivariate linear models (MvLM), significant multivariate tests are often followed by univariate tests on each of the responses separately to determine which responses contribute to each significant effect. In R, these analyses are most convieniently performed using the `update()` method for the `mlm` object `plastic.mod`.

```
> Anova(update(plastic.mod, tear ~ .))
```

```
Anova Table (Type II tests)

Response: tear
              Sum Sq Df F value Pr(>F)
rate            1.74  1    15.8 0.0011 **
additive        0.76  1     6.9 0.0183 *
rate:additive   0.00  1     0.0 0.9471
Residuals       1.76 16
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> Anova(update(plastic.mod, gloss ~ .))
```

```
Anova Table (Type II tests)

Response: gloss
              Sum Sq Df F value Pr(>F)
rate           1.300  1    7.92  0.012 *
additive       0.612  1    3.73  0.071 .
rate:additive  0.544  1    3.32  0.087 .
Residuals      2.628 16
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> Anova(update(plastic.mod, opacity ~ .))
```

```
Anova Table (Type II tests)

Response: opacity
              Sum Sq Df F value Pr(>F)
rate             0.4  1    0.10   0.75
additive         4.9  1    1.21   0.29
rate:additive    4.0  1    0.98   0.34
Residuals       64.9 16
```

The results above show significant main effects for `tear`, a significant main effect of `rate` for `gloss`, and no significant effects for `opacity`, but they don't shed light on the *nature* of these effects. Traditional univariate plots of the means for each variable separately are useful, but they don't allow visualization of the *relations* among the response variables.

We can visualize these effects for pairs of variables in an HE plot, showing the "size" and orientation of hypothesis variation ($\boldsymbol{H}$) in relation to error variation ($\boldsymbol{E}$) as ellipsoids. When, as here, the model terms have 1 degree of freedom, the $\boldsymbol{H}$ ellipsoids degenerate to a line.

```
> # Compare evidence and effect scaling
> colors = c("red", "darkblue", "darkgreen", "brown")
> heplot(plastic.mod, size="evidence", col=colors, cex=1.25)
> heplot(plastic.mod, size="effect", add=TRUE, lwd=4, term.labels=FALSE, col=colors)
```

With effect scaling, both the $\boldsymbol{H}$ and $\boldsymbol{E}$ sums of squares and products matrices are both divided by the error df, giving multivariate analogs of univariate measures of effect size, e.g., $(\bar{y}_1 - \bar{y}_2)/s$. With significance scaling, the $\boldsymbol{H}$ ellipse is further divided by λ_α, the critical value of Roy's largest root statistic. This scaling has the property that an $\boldsymbol{H}$ ellipse will protrude somewhere outside the $\boldsymbol{E}$ ellipse *iff* the multivariate test is significant at level α. Figure 1 shows both scalings, using a thinner line for significance scaling. Note that the (degenerate) ellipse for `additive` is significant, but does not protrude outside the $\boldsymbol{E}$ ellipse in this view. All that is guarranteed is that it will protrude somewhere in the 3D space of the responses.

By design, means for the levels of interaction terms are not shown in the HE plot, because doing so in general can lead to messy displays. We can add them here for the term `rate:additive` as follows:

```
> ## add interaction means
> intMeans <- termMeans(plastic.mod, 'rate:additive', abbrev.levels=2)
> #rownames(intMeans) <- apply(expand.grid(c('Lo','Hi'), c('Lo', 'Hi')), 1, paste, collapse=':')
> points(intMeans[,1], intMeans[,2], pch=18, cex=1.2, col="brown")
> text(intMeans[,1], intMeans[,2], rownames(intMeans), adj=c(0.5,1), col="brown")
> lines(intMeans[c(1,3),1], intMeans[c(1,3),2], col="brown")
> lines(intMeans[c(2,4),1], intMeans[c(2,4),2], col="brown")
```

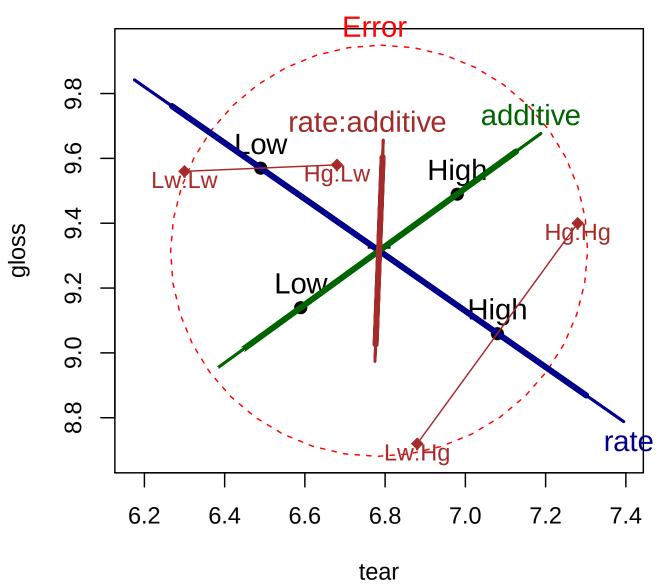

Figure 1: HE plot for effects on `tear` and `gloss` according to the factors `rate`, `additive` and their interaction, `rate:additive`. The thicker lines show effect size scaling, the thinner lines show significance scaling.

The factor means in this plot (Figure 1) have a simple interpretation: The high `rate` level yields greater `tear` resistance but lower `gloss` than the low level. The high `additive` amount produces greater `tear` resistance and greater `gloss`.

The `rate:additive` interaction is not significant overall, though it approaches significance for `gloss`. The cell means for the combinations of `rate` and `additive` shown in this figure suggest an explanation, for tutorial purposes: with the low level of `rate`, there is little difference in `gloss` for the levels of `additive`. At the high level of `rate`, there is a larger difference in `gloss`. The $\boldsymbol{H}$ ellipse for the interaction of `rate:additive` therefore "points" in the direction of `gloss` indicating that this variable contributes to the interaction in the multivariate tests.

In some MANOVA models, it is of interest to test sub-hypotheses of a given main effect or interaction, or conversely to test composite hypotheses that pool together certain effects to test them jointly. All of these tests (and, indeed, the tests of terms in a given model) are carried out as tests of general linear hypotheses in the MvLM.

In this example, it might be useful to test two composite hypotheses: one corresponding to both main effects jointly, and another corresponding to no difference among the means of the four groups (equivalent to a joint test for the overall model). These tests are specified in terms of subsets or linear combinations of the model parameters.

```
> plastic.mod


Call:
lm(formula = cbind(tear, gloss, opacity) ~ rate * additive, data = Plastic)

Coefficients:
                       tear   gloss  opacity
(Intercept)             6.30   9.56   3.74
rateHigh                0.58  -0.84  -0.60
additiveHigh            0.38   0.02   0.10
rateHigh:additiveHigh   0.02   0.66   1.78
```

Thus, for example, the joint test of both main effects tests the parameters `rateHigh` and `additiveHigh`.

```
> print(linearHypothesis(plastic.mod, c("rateHigh", "additiveHigh"), title="Main effects"), SSP=FALSE)


Multivariate Tests: Main effects
                 Df test stat approx F num Df den Df   Pr(>F)
Pillai            2   0.71161   2.7616      6     30 0.029394 *
Wilks             2   0.37410   2.9632      6     28 0.022839 *
Hotelling-Lawley  2   1.44400   3.1287      6     26 0.019176 *
Roy               2   1.26253   6.3127      3     15 0.005542 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


> print(linearHypothesis(plastic.mod, c("rateHigh", "additiveHigh", "rateHigh:additiveHigh"), title="Groups"), S


Multivariate Tests: Groups
                 Df test stat approx F num Df den Df   Pr(>F)
Pillai            3   1.14560   3.2948      9 48.000 0.003350 **
Wilks             3   0.17802   3.9252      9 34.223 0.001663 **
Hotelling-Lawley  3   2.81752   3.9654      9 38.000 0.001245 **
Roy               3   1.86960   9.9712      3 16.000 0.000603 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Correspondingly, we can display these tests in the HE plot by specifying these tests in the `hypothesis` argument to `heplot`(), as shown in Figure 2.

Finally, a 3D HE plot can be produced with `heplot3d`(), giving Figure 3. This plot was rotated interactively to a view that shows both main effects protruding outside the error ellipsoid.

```
> colors = c("pink", "darkblue", "darkgreen", "brown")
> heplot3d(plastic.mod, col=colors)
```

```
> heplot(plastic.mod, hypotheses=list("Group" =
         c("rateHigh", "additiveHigh", "rateHigh:additiveHigh ")),
         col=c(colors, "purple"),
         lwd=c(2, 3, 3, 3, 2), cex=1.25)
> heplot(plastic.mod, hypotheses=list("Main effects" =
         c("rateHigh", "additiveHigh")), add=TRUE,
         col=c(colors, "darkgreen"), cex=1.25)
```

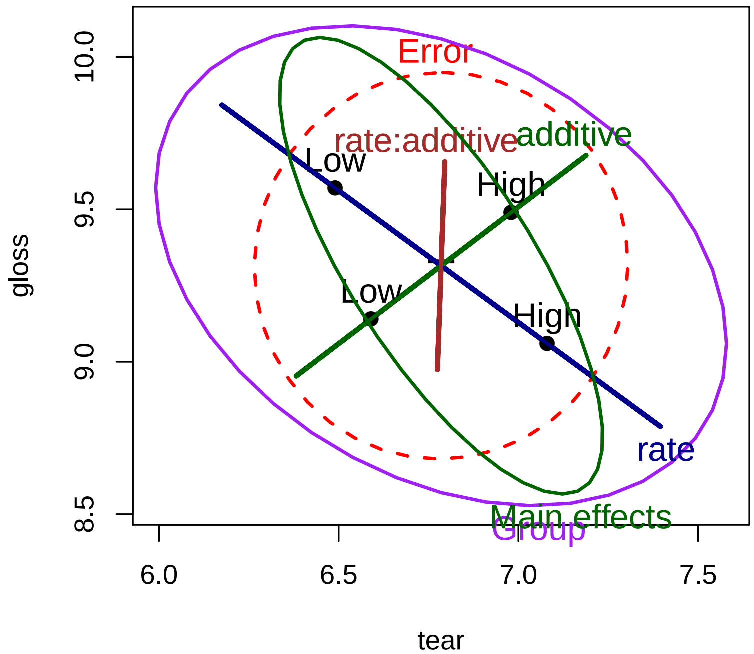

Figure 2: HE plot for `tear` and `gloss`, supplemented with ellipses representing the joint tests of main effects and all group differences

1.2 Effects of physical attractiveness on mock jury decisions

In a social psychology study of influences on jury decisions by Plaster (1989), male participants (prison inmates) were shown a picture of one of three young women. Pilot work had indicated that one woman was beautiful, another of average physical attractiveness, and the third unattractive. Participants rated the woman they saw on each of twelve attributes on scales of 1–9. These measures were used to check on the manipulation of "attractiveness" by the photo.

Then the participants were told that the person in the photo had committed a Crime, and asked to rate the seriousness of the crime and recommend a prison sentence, in Years. The data are contained in the data frame `MockJury`.[1]

```
> str(MockJury)
```

[1]The data were made available courtesy of Karl Wuensch, from `http://core.ecu.edu/psyc/wuenschk/StatData/PLASTER.dat`

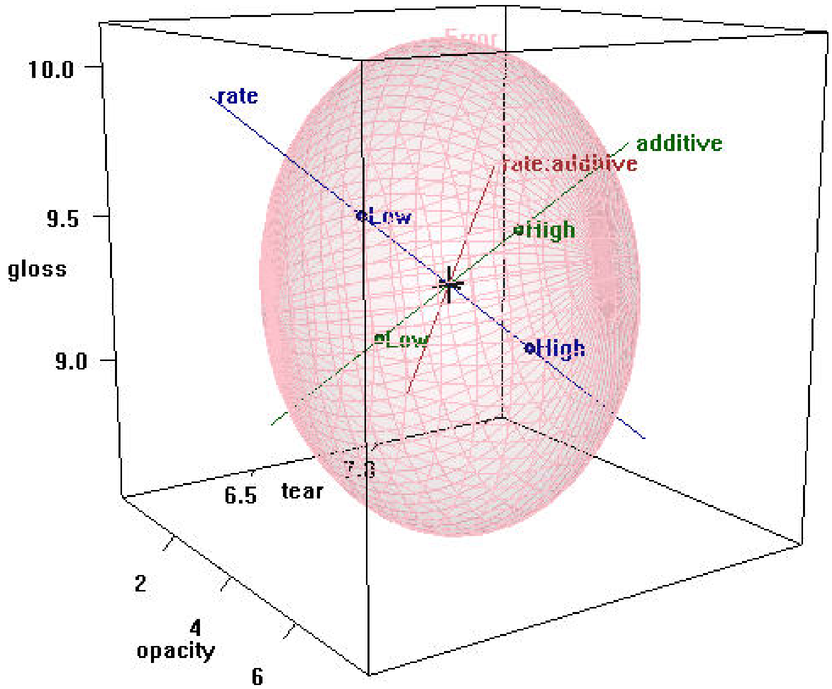

Figure 3: 3D HE plot for the plastic film data

```
'data.frame':        114 obs. of  17 variables:
 $ Attr         : Factor w/ 3 levels "Beautiful","Average",..: 1 1 1 1 1 1 1 1 1 1 ...
 $ Crime        : Factor w/ 2 levels "Burglary","Swindle": 1 1 1 1 1 1 1 1 1 1 ...
 $ Years        : int  10 3 5 1 7 7 3 7 2 3 ...
 $ Serious      : int  8 8 5 3 9 9 4 4 5 2 ...
 $ exciting     : int  6 9 3 3 1 1 5 4 4 6 ...
 $ calm         : int  9 5 4 6 1 5 6 9 8 8 ...
 $ independent  : int  9 9 6 9 5 7 7 2 8 7 ...
 $ sincere      : int  8 3 3 8 1 5 6 9 7 5 ...
 $ warm         : int  5 5 6 8 8 8 7 6 1 7 ...
 $ phyattr      : int  9 9 7 9 8 8 8 5 9 8 ...
 $ sociable     : int  9 9 4 9 9 9 7 2 1 9 ...
 $ kind         : int  9 4 2 9 4 5 5 9 5 7 ...
 $ intelligent  : int  6 9 4 9 7 8 7 9 9 9 ...
 $ strong       : int  9 5 5 9 9 9 5 2 7 5 ...
 $ sophisticated: int  9 5 4 9 9 9 6 2 7 6 ...
 $ happy        : int  5 5 5 9 8 9 5 2 6 8 ...
 $ ownPA        : int  9 7 5 9 7 9 6 5 3 6 ...
```

Sample sizes were roughly balanced for the independent variables in the three conditions of the attractiveness of the photo, and the combinations of this with `Crime`:

```
> table(MockJury$Attr)
```

```
   Beautiful       Average Unattractive 
          39            38           37 
```

```
> table(MockJury$Attr, MockJury$Crime)
```

```
               Burglary Swindle
  Beautiful          21      18
  Average            18      20
  Unattractive       20      17
```

The main questions of interest were: (a) Does attractiveness of the "defendent" influence the sentence or perceived seriousness of the crime? (b) Does attractiveness interact with the nature of the crime?

But first, we try to assess the ratings of the photos in relation to the presumed categories of the independent variable `Attr`. The questions here are (a) do the ratings of the photos on physical attractiveness (`phyattr`) confirm the original classification? (b) how do other ratings differentiate the photos? To keep things simple, we consider ony a few of the other ratings in a one-way MANOVA.

```
> (jury.mod1 <- lm( cbind(phyattr, happy, independent, sophisticated) ~ Attr, data=MockJury))
```

```
Call:
lm(formula = cbind(phyattr, happy, independent, sophisticated) ~
    Attr, data = MockJury)

Coefficients:
                  phyattr  happy   independent  sophisticated
(Intercept)        8.282    5.359   6.410        6.077
AttrAverage       -4.808    0.430   0.537       -1.340
AttrUnattractive  -5.390   -1.359  -1.410       -1.753
```

```
> Anova(jury.mod1, test="Roy")
```

```
Type II MANOVA Tests: Roy test statistic
     Df test stat approx F num Df den Df Pr(>F)
Attr  2      1.77     48.2      4    109 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Note that Beautiful is the baseline category of `Attr`, so the intercept term gives the means for this level. We see that the means are significantly different on all four variables collectively, by a joint multivariate test. A traditional analysis might follow up with univariate ANOVAs for each measure separately.

As an aid to interpretation of the MANOVA results We can examine the test of `Attr` in this model with an HE plot for pairs of variables, e.g., for `phyattr` and `happy` (Figure 4). The means in this plot show that Beautiful is rated higher on physical attractiveness than the other two photos, while Unattractive is rated less happy than the other two. Comparing the sizes of the ellipses, differences among group means on physical attractiveness contributes more to significance than do ratings on happy.

```
> heplot(jury.mod1, main="HE plot for manipulation check")
```

The HE plot for all pairs of variables (Figure 5) shows that the means for `happy` and `independent` are highly correlated, as are the means for `phyattr` and `sophisticated`. In most of these pairwise plots, the means form a triangle rather than a line, suggesting that these attributes are indeed measuring different aspects of the photos.

With 3 groups and 4 variables, the $\boldsymbol{H}$ ellipsoid has only $s = \min(df_h, p) = 2$ dimensions. `candisc`() carries out a canonical discriminant analysis for the MvLM and returns an object that can be used to show an HE plot in the space of the canonical dimensions. This is plotted in Figure 6.

```
> jury.can <- candisc(jury.mod1)
> jury.can
```

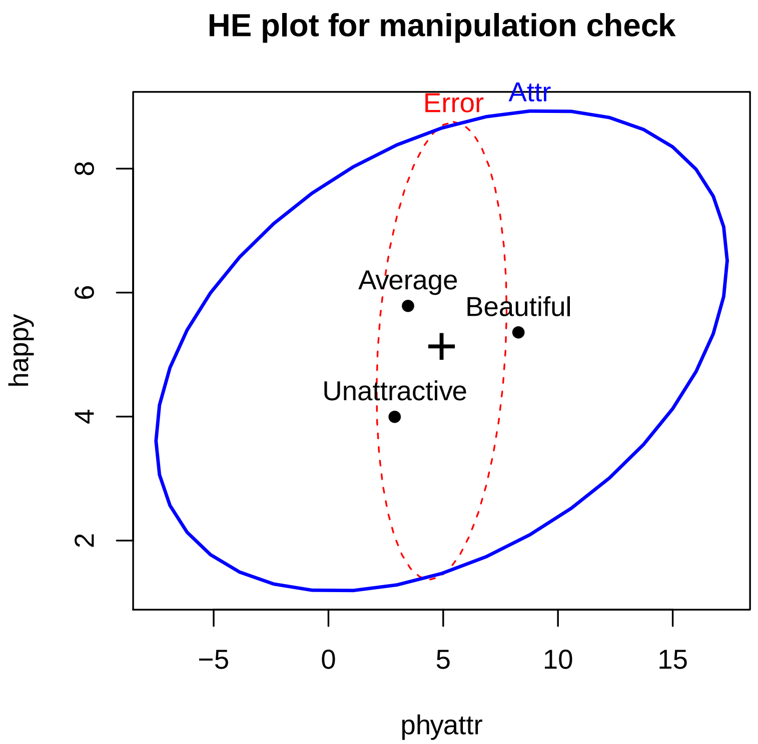

Figure 4: HE plot for ratings of `phyattr` and `happy` according to the classification of photos on `Attr`

```
Canonical Discriminant Analysis for Attr:

  CanRsq Eigenvalue Difference Percent Cumulative
1  0.639      1.767        1.6   91.33       91.3
2  0.144      0.168        1.6    8.67      100.0

Test of H0: The canonical correlations in the
current row and all that follow are zero

  LR test stat approx F num Df den Df Pr(> F)
1        0.309     43.9      4    220 < 2e-16 ***
2        0.856     18.6      1    111 3.5e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

From this we can see that 91% of the variation among group means is accounted for by the first dimension, and this is nearly completely aligned with `phyattr`. The second dimension, accounting for the remaining 9% is determined nearly entirely by ratings on `happy` and `independent`. This display gives a relatively simple account of the results of the MANOVA and the relations of each of the ratings to discrimination among the photos.

Proceeding to the main questions of interest, we carry out a two-way MANOVA of the responses `Years` and `Serious` in relation to the independent variables `Attr` and `Crime`.

```
> # influence of Attr of photo and nature of crime on Serious and Years
> jury.mod2 <- lm( cbind(Serious, Years) ~ Attr * Crime, data=MockJury)
> Anova(jury.mod2, test="Roy")
```

```
> pairs(jury.mod1)
```

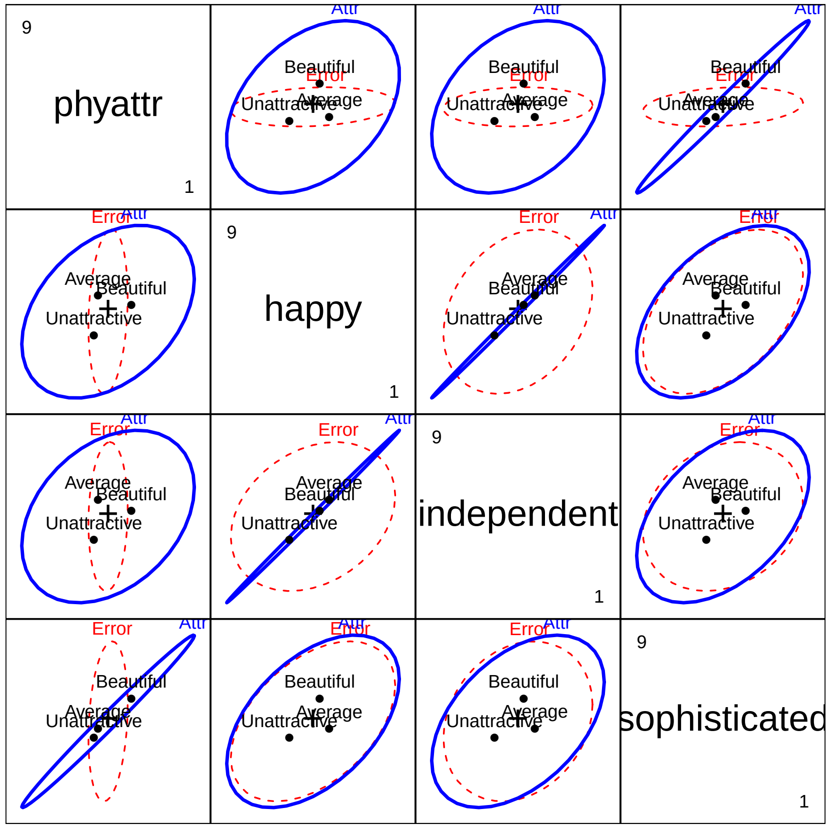

Figure 5: HE plots for all pairs of ratings according to the classification of photos on `Attr`

```
Type II MANOVA Tests: Roy test statistic
            Df test stat approx F num Df den Df Pr(>F)
Attr         2    0.0756     4.08      2    108  0.020 *
Crime        1    0.0047     0.25      2    107  0.778
Attr:Crime   2    0.0501     2.71      2    108  0.071 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

We see that there is a nearly significant interaction between `Attr` and `Crime` and a strong effect of `Attr`.

The HE plot shows that the nearly significant interaction of `Attr:Crime` is mainly in terms of differences among the groups on the response of `Years` of sentence, with very little contribution of `Serious`. We explore this interaction in a bit more detail below. The main effect of `Attr` is also dominated by differences among groups on `Years`.

If we assume that `Years` of sentence is the main outcome of interest, it also makes sense to carry out a step-down test of this variable by itself, controlling for the rating of seriousness (`Serious`) of the crime. The model `jury.mod3` below is equivalent to an ANCOVA for `Years`.

```
> # stepdown test (ANCOVA), controlling for Serious
> jury.mod3 <- lm( Years ~ Serious + Attr * Crime, data=MockJury)
> t(coef(jury.mod3))
```

```
     (Intercept) Serious AttrAverage AttrUnattractive CrimeSwindle
[1,]    0.011612 0.83711     0.39586          0.60285     -0.26302
```

```
> opar <- par(xpd=TRUE)
> heplot(jury.can, prefix="Canonical dimension", main="Canonical HE plot")
> par(opar)
```

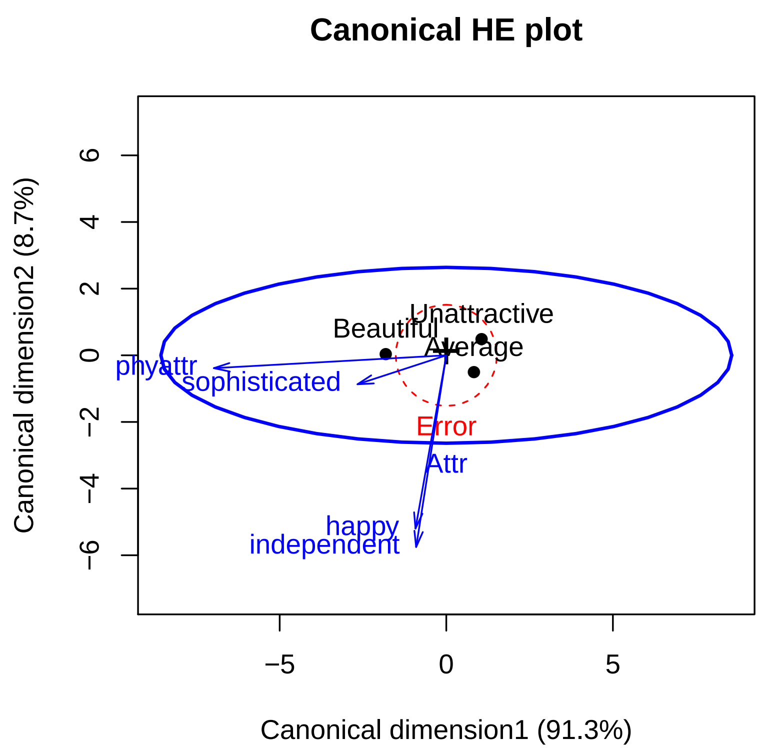

Figure 6: Canonical discriminant HE plot

```
     AttrAverage:CrimeSwindle AttrUnattractive:CrimeSwindle
[1,]                 -0.53701                        2.5123

> Anova(jury.mod3)

Anova Table (Type II tests)

Response: Years
            Sum Sq  Df F value  Pr(>F)
Serious        379   1   41.14 3.9e-09 ***
Attr            74   2    4.02   0.021 *
Crime            4   1    0.43   0.516
Attr:Crime      49   2    2.67   0.074 .
Residuals      987 107
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Thus, even when adjusting for `Serious` rating, there is still a significant main effect of `Attr` of the photo, but also a hint of an interaction of `Attr` with `Crime`. The coefficient for `Serious` indicates that participants awarded 0.84 additional years of sentence for each 1 unit step on the scale of seriousness of crime.

A particularly useful method for visualizing the fitted effects in such univariate response models is provided by the effects package. By default `allEffects`() calculates the predicted

```
> heplot(jury.mod2)
```

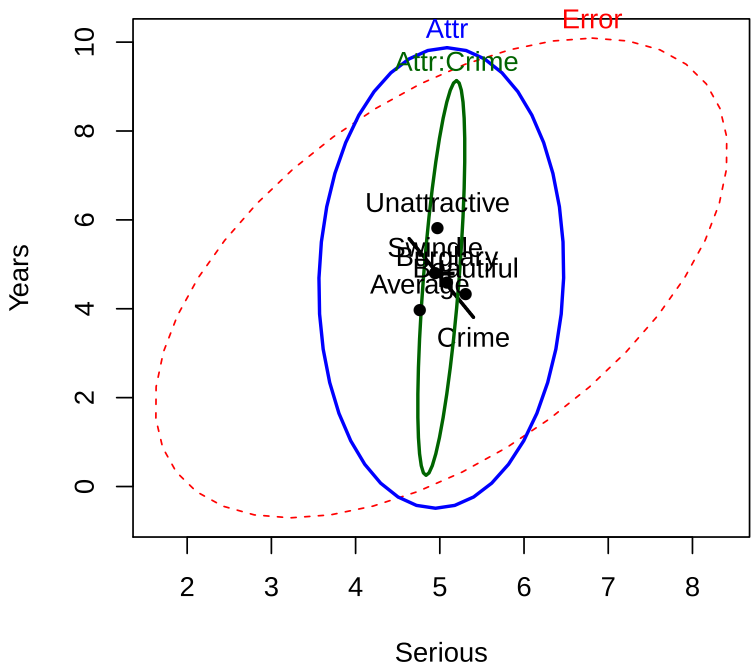

Figure 7: HE plot for the two-way MANOVA for `Years` and `Serious`

values for all high-order terms in a given model, and the `plot` method produces plots of these values for each term. The statements below produce Figure 8.

The effect plot for `Serious` shows the expected linear relation between that variable and `Years`. Of greater interest here is the nature of the possible interaction of `Attr` and `Crime` on `Years` of sentence, controlling for `Serious`. The effect plot shows that for the crime of Swindle, there is a much greater `Years` of sentence awarded to Unattractive defendents.

2 Multivariate Multiple Regression Designs

The ideas behind HE plots extend naturally to multivariate multiple regression (MMRA) and multivariate analysis of covariance (MANCOVA). In MMRA, the $\boldsymbol{X}$ matrix contains only quantitative predictors, while in MANCOVA designs, there is a mixture of factors and quantitative predictors (covariates).

In the MANOVA case, there is often a subtle difference in emphasis: true MANCOVA analyses focus on the differences among groups defined by the factors, adjusting for (or controlling for) the quantitative covariates. Analyses concerned with *homogeneity of regression* focus on quantitative predictors and attempt to test whether the regression relations are the same for all groups defined by the factors.

```
> library(effects)
> jury.eff <- allEffects(jury.mod3)
> plot(jury.eff, ask=FALSE)
```

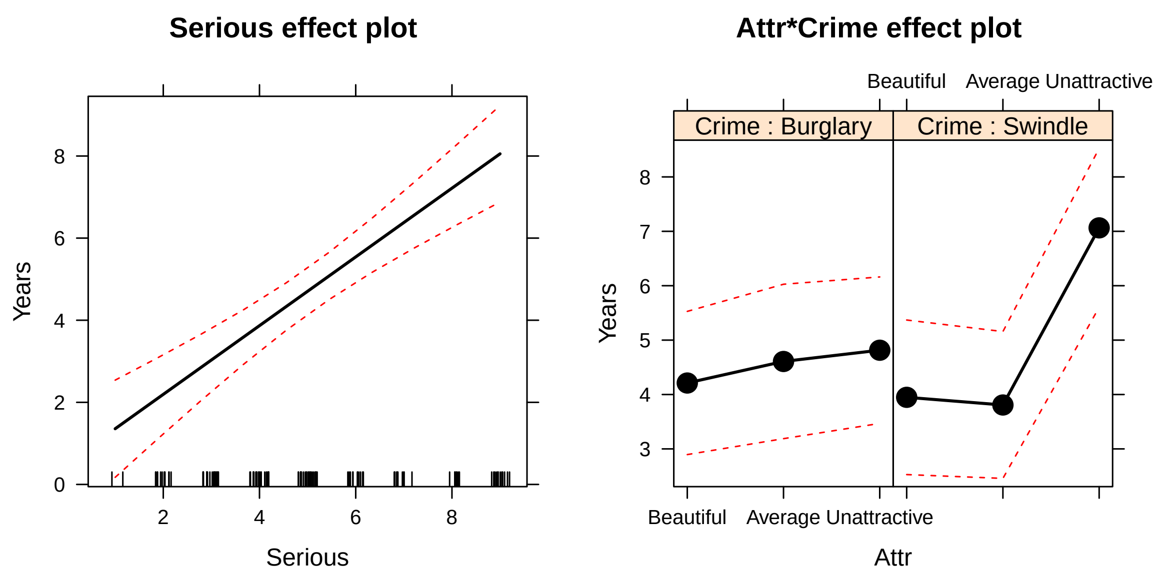

Figure 8: Effect plots for `Serious` and the `Attr * Crime` in the ANCOVA model `jury.mod3`.

2.1 Rohwer data

To illustrate the homogeneity of regression flavor, we use data from a study by Rohwer (given in Timm, 1975: Ex. 4.3, 4.7, and 4.23) on kindergarten children, designed to determine how well a set of paired-associate (PA) tasks predicted performance on the Peabody Picture Vocabulary test (`PPVT`), a student achievement test (`SAT`), and the Raven Progressive matrices test (`Raven`). The PA tasks varied in how the stimuli were presented, and are called *named* (`n`), *still* (`s`), *named still* (`ns`), *named action* (`na`), and *sentence still* (`ss`).

Two groups were tested: a group of $n = 37$ children from a low socioeconomic status (SES) school, and a group of $n = 32$ high SES children from an upper-class, white residential school. The data are in the data frame `Rohwer` in the `heplots` package:

```
> some(Rohwer,n=6)
```

```
   group SES SAT PPVT Raven n  s ns na ss
14     1  Lo  30   55    13 2  1 12 20 17
17     1  Lo  19   66    13 7 12 21 35 27
18     1  Lo  45   54    10 0  6  6 14 16
21     1  Lo  32   48    16 0  7  9 14 18
37     1  Lo  79   54    14 0  6  6 15 14
57     2  Hi  99   94    16 4  6 14 27 19
```

At one extreme, we might be tempted to fit separate regression models for each of the High and Low SES groups. This approach is *not* recommended because it lacks power and does not allow hypotheses about equality of slopes and intercepts to be tested directly.

```
> rohwer.ses1 <- lm(cbind(SAT, PPVT, Raven) ~ n + s + ns + na + ss, data=Rohwer, subset=SES=="Hi")
> Anova(rohwer.ses1)
```

```
Type II MANOVA Tests: Pillai test statistic
   Df test stat approx F num Df den Df Pr(>F)
n   1     0.202     2.02      3     24 0.1376
s   1     0.310     3.59      3     24 0.0284 *
ns  1     0.358     4.46      3     24 0.0126 *
na  1     0.465     6.96      3     24 0.0016 **
ss  1     0.089     0.78      3     24 0.5173
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> rohwer.ses2 <- lm(cbind(SAT, PPVT, Raven) ~ n + s + ns + na + ss, data=Rohwer, subset=SES=="Lo")
> Anova(rohwer.ses2)
```

```
Type II MANOVA Tests: Pillai test statistic
   Df test stat approx F num Df den Df Pr(>F)
n   1    0.0384     0.39      3     29  0.764
s   1    0.1118     1.22      3     29  0.321
ns  1    0.2252     2.81      3     29  0.057 .
na  1    0.2675     3.53      3     29  0.027 *
ss  1    0.1390     1.56      3     29  0.220
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

This allows separate slopes and intercepts for each of the two groups, but it is difficult to compare the coefficients numerically.

```
> coef(rohwer.ses1)
```

```
                  SAT      PPVT     Raven
(Intercept) -28.46747 39.697090 13.243836
n             3.25713  0.067283  0.059347
s             2.99658  0.369984  0.492444
ns           -5.85906 -0.374380 -0.164022
na            5.66622  1.523009  0.118980
ss           -0.62265  0.410157 -0.121156
```

```
> coef(rohwer.ses2)
```

```
                  SAT      PPVT     Raven
(Intercept)  4.151060 33.005769 11.173378
n           -0.608872 -0.080567  0.210995
s           -0.050156 -0.721050  0.064567
ns          -1.732395 -0.298303  0.213584
na           0.494565  1.470418 -0.037318
ss           2.247721  0.323965 -0.052143
```

Nevertheless, we can visualize the results with HE plots, and here we make use of the fact that several HE plots can be overlaid using the option `add=TRUE` as shown in Figure 9.

```
> heplot(rohwer.ses1, ylim=c(40,110),col=c("red", "black"), lwd=2, cex=1.2)
> heplot(rohwer.ses2, add=TRUE, col=c("blue", "black"), grand.mean=TRUE, error.ellipse=TRUE, lwd=2, cex=1.2)
> means <- aggregate(cbind(SAT,PPVT)~SES, data=Rohwer,  mean)
> text(means[,2], means[,3], labels=means[,1], pos=3, cex=2, col=c("red", "blue"))
```

We can readily see the difference in means for the two SES groups (High greater on both variables) and it also appears that the slopes of the predictor ellipses are shallower for the High than the Low group, indicating greater relation with the `SAT` score.

Alternatively (and optimistically), we can fit a MANCOVA model that allows different means for the two SES groups on the responses, but constrains the slopes for the PA covariates to be equal.

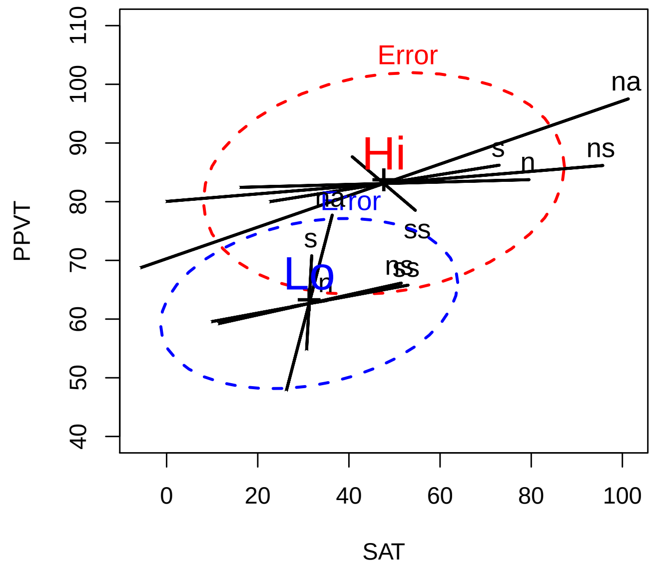

Figure 9: HE plot for `SAT` and `PPVT`, showing the effects for the PA predictors for the High and Low SES groups separately

```
> # MANCOVA, assuming equal slopes
> rohwer.mod <- lm(cbind(SAT, PPVT, Raven) ~ SES + n + s + ns + na + ss,
                   data=Rohwer)
> Anova(rohwer.mod)
```

```
Type II MANOVA Tests: Pillai test statistic
    Df test stat approx F num Df den Df  Pr(>F)
SES  1     0.379    12.18      3     60 2.5e-06 ***
n    1     0.040     0.84      3     60  0.4773
s    1     0.093     2.04      3     60  0.1173
ns   1     0.193     4.78      3     60  0.0047 **
na   1     0.231     6.02      3     60  0.0012 **
ss   1     0.050     1.05      3     60  0.3770
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Note that, although the multivariate tests for two of the covariates (`ns` and `na`) are highly significant, univariate multiple regression tests for the separate responses [from `summary(rohwer.mod)`] are relatively weak. We can also test the global 5 df hypothesis, $\boldsymbol{B} = \mathbf{0}$, that *all* covariates have null effects for all responses as a linear hypothesis (suppressing display of the error and hypothesis SSP matrices),

```
> (covariates <- rownames(coef(rohwer.mod))[-(1:2)])
```

```
[1] "n"  "s"  "ns" "na" "ss"
```

```
> Regr<-linearHypothesis(rohwer.mod, covariates)
> print(Regr, digits=5, SSP=FALSE)
```

```
Multivariate Tests:
                 Df test stat approx F num Df den Df    Pr(>F)    
Pillai            5   0.66579   3.5369     15 186.00 2.309e-05 ***
Wilks             5   0.44179   3.8118     15 166.03 8.275e-06 ***
Hotelling-Lawley  5   1.03094   4.0321     15 176.00 2.787e-06 ***
Roy               5   0.75745   9.3924      5  62.00 1.062e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Then 2D views of the additive MANCOVA model `rohwer.mod` and the overall test for all covariates are produced as follows, giving the plots in Figure 10.

```
> colors <- c("red", "blue", rep("black",5), "#969696")
> heplot(rohwer.mod, col=colors,
        hypotheses=list("Regr" = c("n", "s", "ns", "na", "ss")),
        cex=1.5, lwd=c(2, rep(3,5), 4),
        main="(SAT, PPVT, Raven) ~ SES + n + s + ns + na + ss")


> heplot(rohwer.mod, col=colors,  variables=c(1,3),
        hypotheses=list("Regr" = c("n", "s", "ns", "na", "ss")),
        cex=1.5, lwd=c(2, rep(3,5), 4),
        main="(SAT, PPVT, Raven) ~ SES + n + s + ns + na + ss")
```

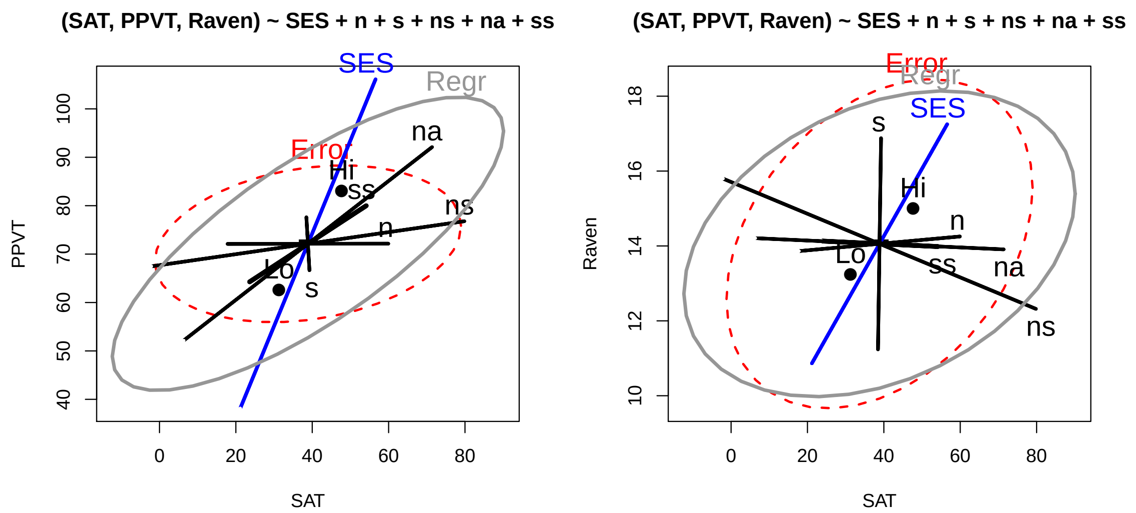

Figure 10: HE plot for `SAT` and `PPVT` (left) and for `SAT` and `Raven` (right) using the MANCOVA model

The positive orientation of the `Regr` ellipses shows that the prediced values for all three responses are positively correlated (more so for `SAT` and `PPVT`). As well, the High SES group is higher on all responses than the Low SES group.

Alternatively, all pairwise plots among these responses could be drawn using the `pairs` function (figure not shown),

```
> pairs(rohwer.mod, col=colors,
        hypotheses=list("Regr" = c("n", "s", "ns", "na", "ss")),
        cex=1.3, lwd=c(2, rep(3,5), 4))
```

or as a 3D plot, using `heplot3d()` as shown in Figure 11.

```
> colors <- c("pink", "blue", rep("black",5), "#969696")
> heplot3d(rohwer.mod, col=colors,
          hypotheses=list("Regr" = c("n", "s", "ns", "na", "ss")))
```

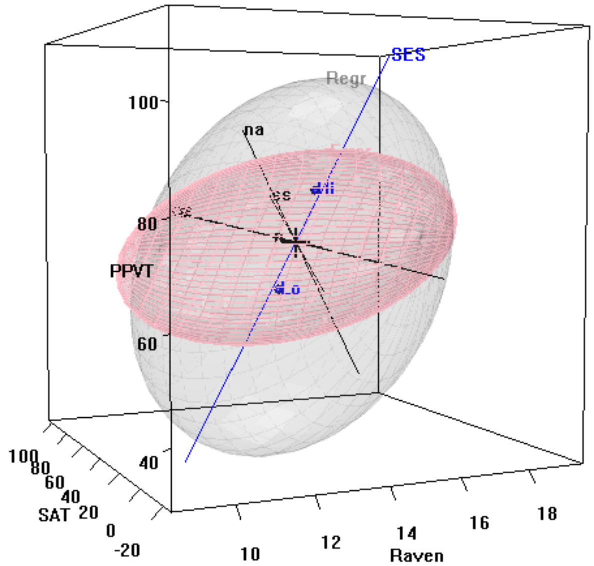

Figure 11: 3D HE plot for the MANCOVA model fit to the Rohwer data

The MANCOVA model, `rohwer.mod`, has relatively simple interpretations (large effect of `SES`, with `ns` and `na` as the major predictors) but the test of relies on the assumption of homogeneity of slopes for the predictors. We can test this as follows, adding interactions of `SES` with each of the covariates:

```
> rohwer.mod2 <- lm(cbind(SAT, PPVT, Raven) ~ SES * (n + s + ns + na + ss),
                    data=Rohwer)
> Anova(rohwer.mod2)
```

```
Type II MANOVA Tests: Pillai test statistic
       Df test stat approx F num Df den Df  Pr(>F)
SES     1     0.391    11.78      3     55 4.5e-06 ***
n       1     0.079     1.57      3     55 0.20638
s       1     0.125     2.62      3     55 0.05952 .
ns      1     0.254     6.25      3     55 0.00100 ***
na      1     0.307     8.11      3     55 0.00015 ***
ss      1     0.060     1.17      3     55 0.32813
SES:n   1     0.072     1.43      3     55 0.24417
SES:s   1     0.099     2.02      3     55 0.12117
SES:ns  1     0.118     2.44      3     55 0.07383 .
SES:na  1     0.148     3.18      3     55 0.03081 *
SES:ss  1     0.057     1.12      3     55 0.35094
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

It appears from the above that there is only weak evidence of unequal slopes from the separate `SES:` terms. The evidence for heterogeneity is stronger, however, when these terms are tested collectively using the `linearHypothesis`() function:

```
> (coefs <- rownames(coef(rohwer.mod2)))
```

```
 [1] "(Intercept)" "SESLo"       "n"           "s"           "ns"
 [6] "na"          "ss"          "SESLo:n"     "SESLo:s"     "SESLo:ns"
[11] "SESLo:na"    "SESLo:ss"
```

```
> print(linearHypothesis(rohwer.mod2, coefs[grep(":", coefs)]), SSP=FALSE)


Multivariate Tests:
                 Df test stat approx F num Df den Df   Pr(>F)
Pillai            5   0.41794   1.8452     15 171.00 0.032086 *
Wilks             5   0.62358   1.8936     15 152.23 0.027695 *
Hotelling-Lawley  5   0.53865   1.9272     15 161.00 0.023962 *
Roy               5   0.38465   4.3850      5  57.00 0.001905 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

This model (`rohwer.mod2`) is similar in spirit to the two models (`rohwer.ses1` and `rohwer.ses2`) fit for the two SES groups separately and show in Figure 9, except that model `rohwer.mod2` assumes a common within-groups error covariance matrix and allows overall tests.

To illustrate model `rohwer.mod2`, we construct an HE plot for `SAT` and `PPVT` shown in Figure 12. To simplify this display, we show the hypothesis ellipses for the overall effects of the PA tests in the baseline high-SES group, and a single combined ellipse for all the `SESLo:` interaction terms that we tested previously, representing differences in slopes between the low and high-SES groups.

Because SES is "treatment-coded" in this model, the ellipse for each covariate represents the hypothesis that the slopes for that covariate are zero in the high-SES baseline category. With this parameterization, the ellipse for `Slopes` represents the joint hypothesis that slopes for the covariates do not differ in the low-SES group.

```
> colors <- c("red", "blue", rep("black",5), "#969696")
> heplot(rohwer.mod2, col=c(colors, "brown"),
        terms=c("SES", "n", "s", "ns", "na", "ss"),
        hypotheses=list("Regr" = c("n", "s", "ns", "na", "ss"),
                        "Slopes" = coefs[grep(":", coefs)]))
```

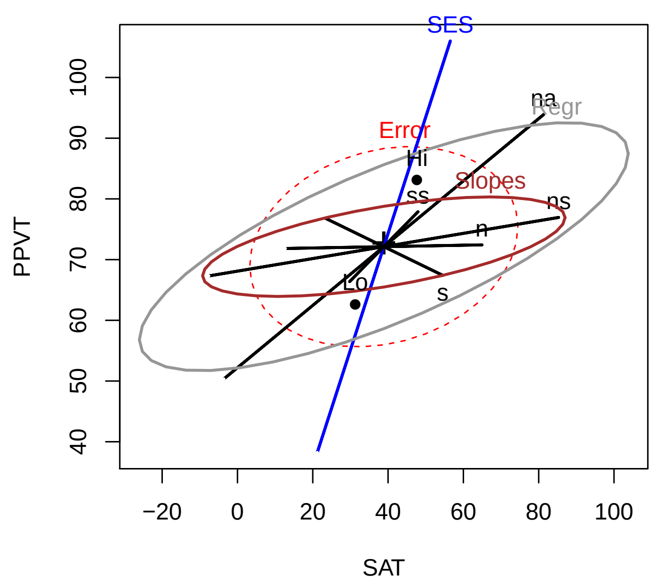

Figure 12: HE plot for `SAT` and `PPVT`, fitting the model `rohwer.mod2` that allows unequal slopes for the covariates.

Comparing Figure 12 for the heterogeneous slopes model with Figure 10 (left) for the homogeneous slopes model, it can be seen that most of the covariates have ellipses of similar size and orientation, reflecting simlar evidence against the respective null hypotheses, as does the effect of `SES`, showing the greater performance of the high-SES group on all response measures. Somewhat more subtle, the error ellipse is noticeably smaller in Figure 12, reflecting the additional variation accounted for by differences in slopes.

References

J. Fox, M. Friendly, and G. Monette. Visualizing hypothesis tests in multivariate linear models: The *heplots* package for R. *Computational Statistics*, 24(2):233–246, 2009. (Published online: 15 May 2008).

M. Friendly. Data ellipses, HE plots and reduced-rank displays for multivariate linear models: SAS software and examples. *Journal of Statistical Software*, 17(6):1–42, 2006.

M. Friendly. HE plots for multivariate general linear models. *Journal of Computational and Graphical Statistics*, 16(2):421–444, 2007.

M. E. Plaster. *The Effect of Defendent Physical Attractiveness on Juridic Decisions Using Felon Inmates as Mock Jurors.* Unpublished master's thesis, East Carolina University, Greenville, NC, 1989.

N. H. Timm. *Multivariate Analysis with Applications in Education and Psychology.* Wadsworth (Brooks/Cole), Belmont, CA, 1975.