HE Plot Examples

Michael Friendly

Using heplots version 0.9-4 and candisc version 0.9-4; Date: 2010-09-20

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    1.6188    7.554      3     14 0.00303 **
additive       1    0.9119    4.256      3     14 0.02475 *
rate:additive  1    0.2868    1.339      3     14 0.30178
---
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          1.7405  1  15.787 0.00109 **
additive      0.7605  1   6.898 0.01833 *
rate:additive 0.0005  1   0.005 0.94714
Residuals     1.7640 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          1.3005  1   7.918 0.0125 *
additive      0.6125  1   3.729 0.0714 .
rate:additive 0.5445  1   3.315 0.0874 .
Residuals     2.6280 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            0.42  1   0.104  0.752
additive        4.90  1   1.208  0.288
rate:additive   3.96  1   0.976  0.338
Residuals      64.92 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")
```

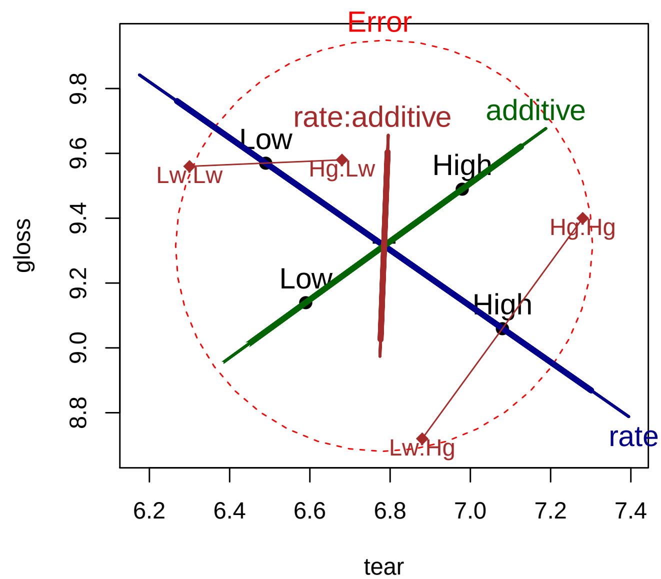

Figure 1: HE plot for effects on `tear` and `gloss` according to the factors `rate`, `additive` and their interaction, `rate:additive`. The thicker lines show effect size scaling, the thinner lines show significance scaling.

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  0.711613  2.76165      6     30 0.0293945 *
Wilks             2  0.374096  2.96317      6     28 0.0228392 *
Hotelling-Lawley  2  1.444000  3.12867      6     26 0.0191755 *
Roy               2  1.262531  6.31266      3     15 0.0055424 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


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


Multivariate Tests: Groups
                 Df test stat approx F num Df  den Df     Pr(>F)
Pillai            3  1.145598  3.29479      9 48.0000 0.00335033 **
Wilks             3  0.178019  3.92517      9 34.2229 0.00166294 **
Hotelling-Lawley  3  2.817516  3.96539      9 38.0000 0.00124500 **
Roy               3  1.869597  9.97118      3 16.0000 0.00060304 ***
---
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)
```

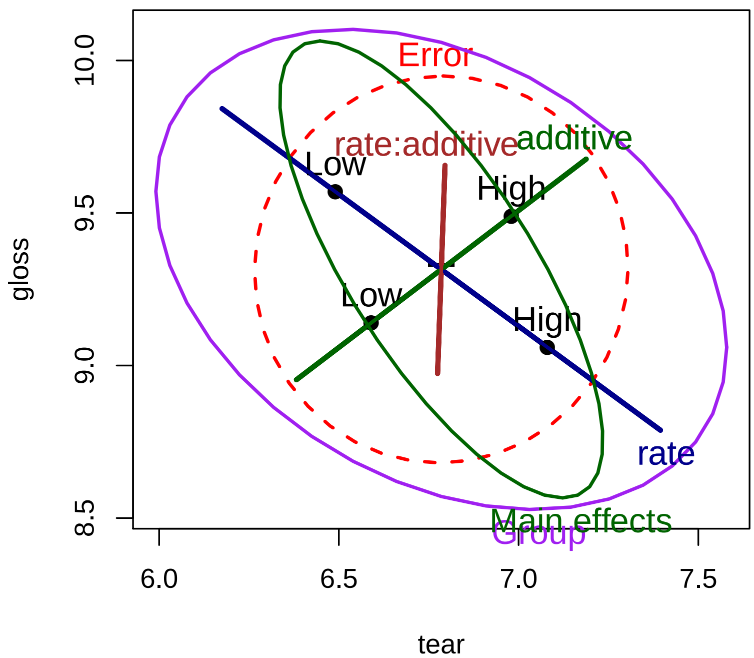

Figure 2: HE plot for `tear` and `gloss`, supplemented with ellipses representing the joint tests of main effects and all group differences

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`

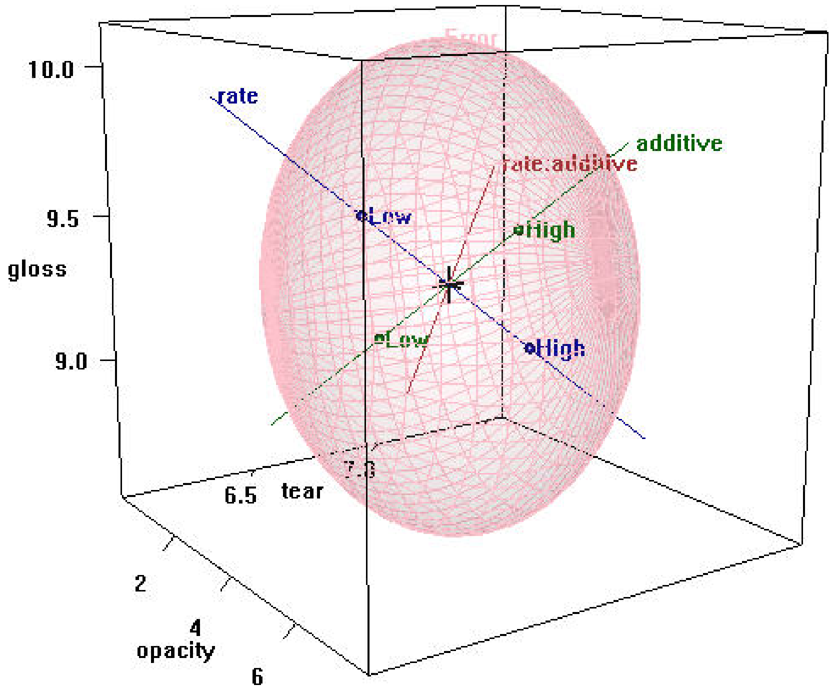

Figure 3: 3D HE plot for the plastic film data

```
'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     1.767    48.16      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
```

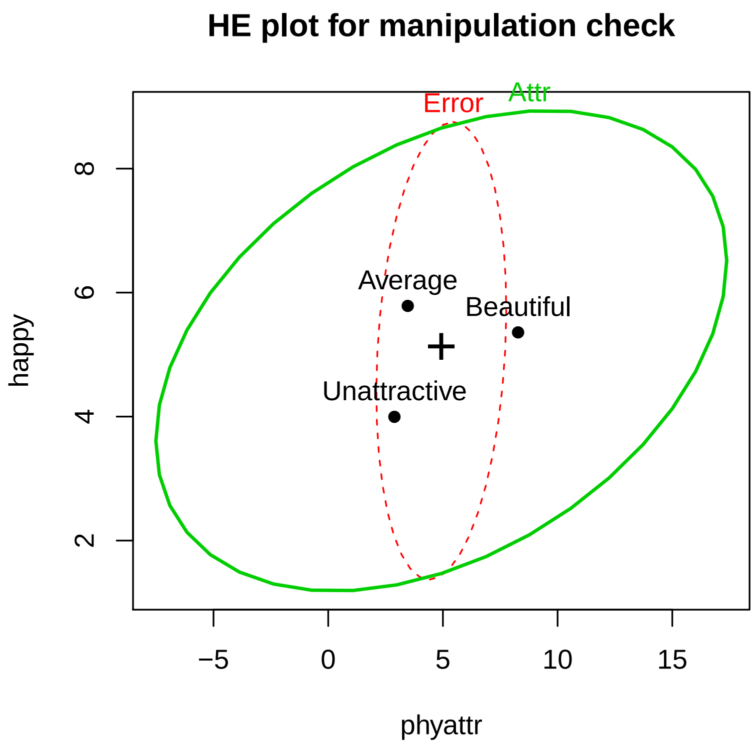

Figure 4: HE plot for ratings of `phyattr` and `happy` according to the classification of photos on `Attr`

```
Canonical Discriminant Analysis for Attr:

  CanRsq Eigenvalue Difference Percent Cumulative
1 0.6386     1.7672      1.600  91.334      91.33
2 0.1436     0.1677      1.600   8.666     100.00

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       0.3095    43.86      4    220  < 2e-16 ***
2       0.8564    18.61      1    111 3.49e-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`.

```
> jury.mod2 <- lm(cbind(Serious, Years) ~ Attr * Crime, data = MockJury)
> Anova(jury.mod2, test = "Roy")
```

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

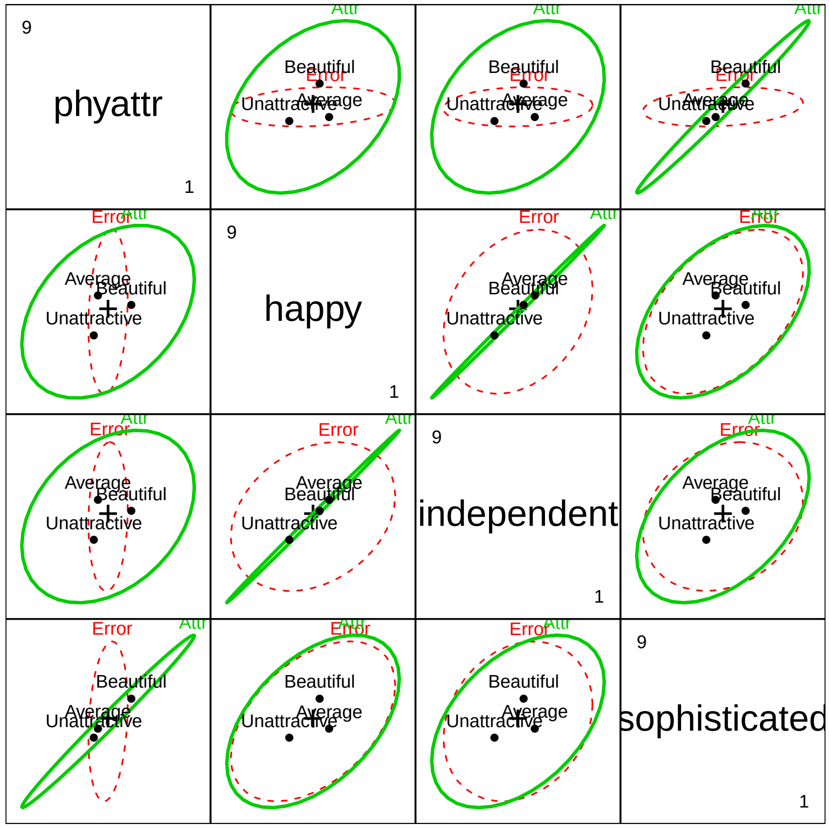

Figure 5: HE plots for all pairs of ratings according to the classification of photos on `Attr`

```
Type II MANOVA Tests: Roy test statistic
            Df test stat approx F num Df den Df Pr(>F)
Attr         2   0.07561    4.083      2    108 0.0195 *
Crime        1   0.00470    0.251      2    107 0.7782
Attr:Crime   2   0.05010    2.706      2    108 0.0714 .
---
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`.

```
> jury.mod3 <- lm(Years ~ Serious + Attr * Crime, data = MockJury)
> t(coef(jury.mod3))
```

```
     (Intercept)  Serious AttrAverage AttrUnattractive CrimeSwindle
[1,]   0.0116122 0.837108    0.395858         0.602846    -0.263018
     AttrAverage:CrimeSwindle AttrUnattractive:CrimeSwindle
[1,]                -0.537006                       2.51226
```

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

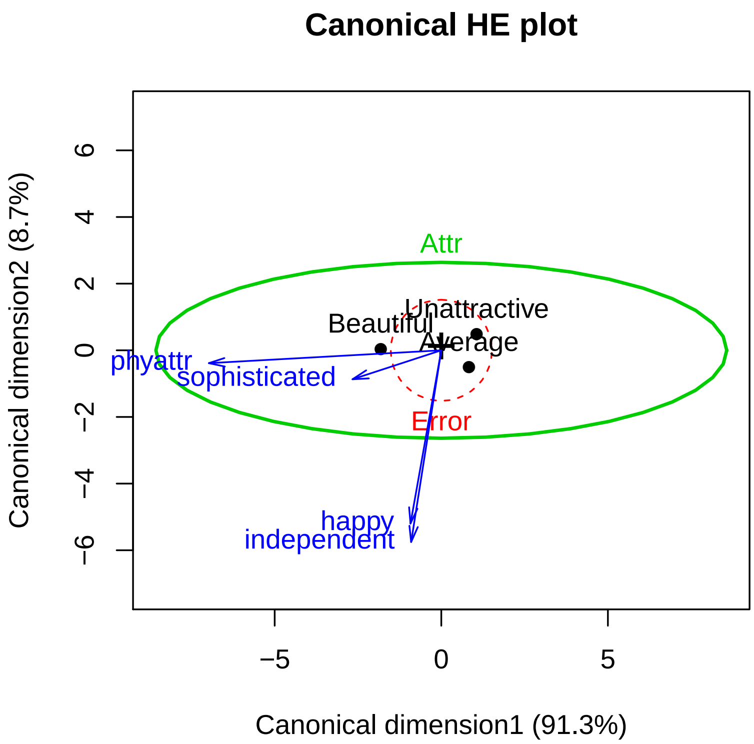

Figure 6: Canonical discriminant HE plot

```
> Anova(jury.mod3)
```

```
Anova Table (Type II tests)

Response: Years
            Sum Sq  Df F value    Pr(>F)    
Serious      379.5   1  41.142 3.94e-09 ***
Attr          74.2   2   4.023   0.0207 *  
Crime          3.9   1   0.425   0.5156    
Attr:Crime    49.3   2   2.672   0.0737 .  
Residuals    986.9 107                     
---
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 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.

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

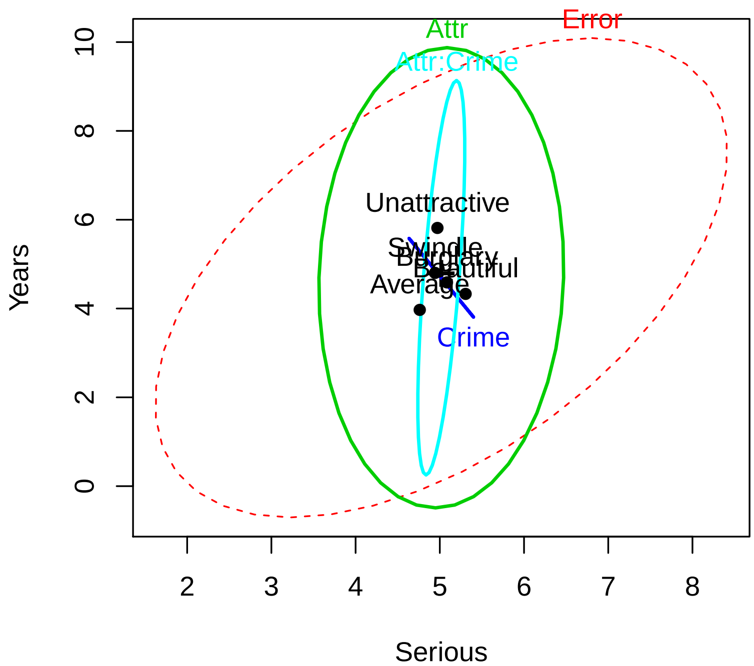

Figure 7: HE plot for the two-way MANOVA for `Years` and `Serious`

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.

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.

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

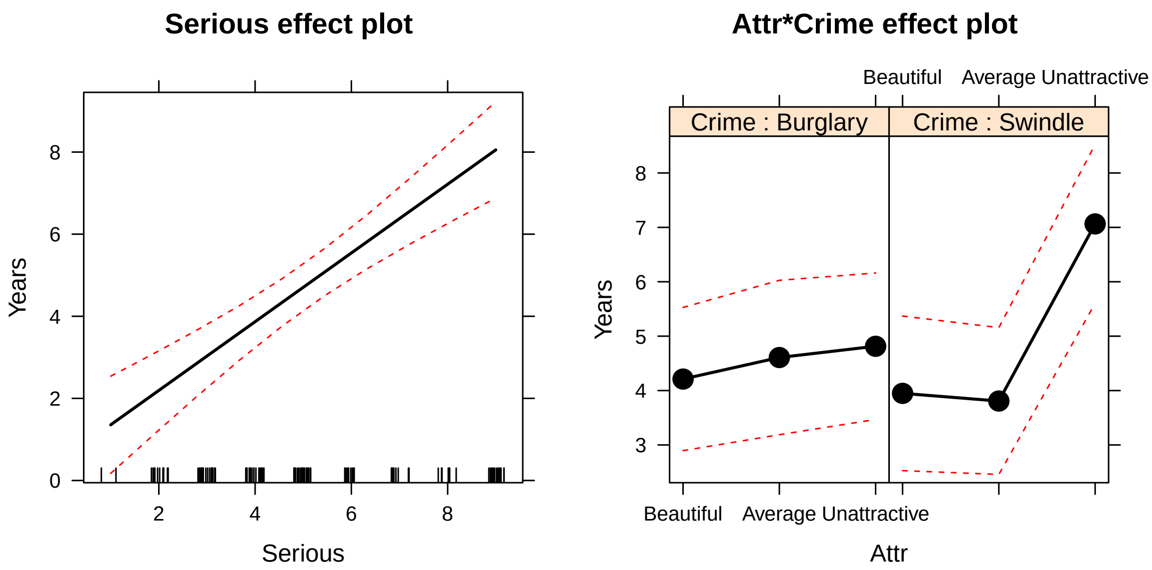

Figure 8: Effect plots for `Serious` and the `Attr * Crime` in the ANCOVA model `jury.mod3`.