

Visualizing Linear Models: An R Bag of Tricks Session 3: Examples & Extensions

Michael Friendly
SCS Short Course
Oct. 22, 29, Nov. 5, 2020
<https://friendly.github.io/VisMLM-course/>

Today's topics

- Manova examples
 - Distinguishing among psychiatric groups
 - Robust MLMs
- Multivariate regression
 - PA tests & ability
 - Canonical correlation
 - MANCOVA & homogeneity of regression
- Homogeneity of variance

2

Ex: Neuro- & Social-Cognitive measures in psychiatric groups

- A study by Leah Hartman @York examined whether patients classified as 'schizophrenic' or 'schizoaffective' (on DSM-IV) could be distinguished from a normal, control sample on standardized tests in the following domains:
 - **Neuro-Cognitive**: processing speed, attention, verbal learning, visual learning, problem solving
 - **Social-cognitive**: managing emotions, theory of mind, externalizing bias, personalizing bias
- Research questions → MANOVA contrasts
 - Analyze neuro-cog (NC) and social-cog (SC) separately
 - Do the two psychiatric groups differ from the controls?
 - Do the psychiatric groups differ from each other?

See: Friendly & Sigal (2017), Graphical Methods for Multivariate Linear Models in Psychological Research: An R Tutorial
The Quantitative Methods for Psychology, 13, 20-45, <http://dx.doi.org/10.20982/tamp.13.1.p020>

3

Neuro-cognitive measures

```
library(heplots)
library(candisc)
data(NeuroCog, package="heplots")

# fit the MANOVA model, test hypotheses
NC.mlm <- lm(cbind(Speed, Attention, Memory, Verbal, Visual, ProbSolv) ~ Dx,
             data=NeuroCog)
Anova(NC.mlm)

Type II MANOVA Tests: Pillai test statistic
Df test stat approx F num Df den Df Pr(>F)
Dx 2 0.2992 6.8902 12 470 1.562e-11 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

So, the groups differ. But how?

What about the research hypotheses?

```
> contrasts(NeuroCog$Dx)
      [,1] [,2]
Schizophrenia -0.5 1
Schizoaffective -0.5 -1
Control 1.0 0
```

4

Neuro-cognitive measures

A simple result: Control \neq (Schizophrenia \approx Schizoaffective)

```
> print(linearHypothesis(NC.mlm, "Dx1"), SSP=FALSE)
```

Multivariate Tests:

	Df	test stat	approx F	num Df	den Df	Pr(>F)
Pillai	1	0.289	15.9	6	234	2.8e-15 ***
Wilks	1	0.711	15.9	6	234	2.8e-15 ***
Hotelling-Lawley	1	0.407	15.9	6	234	2.8e-15 ***
Roy	1	0.407	15.9	6	234	2.8e-15 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```
> print(linearHypothesis(NC.mlm, "Dx2"), SSP=FALSE)
```

Multivariate Tests:

	Df	test stat	approx F	num Df	den Df	Pr(>F)
Pillai	1	0.006	0.249	6	234	0.96
Wilks	1	0.994	0.249	6	234	0.96
Hotelling-Lawley	1	0.006	0.249	6	234	0.96
Roy	1	0.006	0.249	6	234	0.96

5

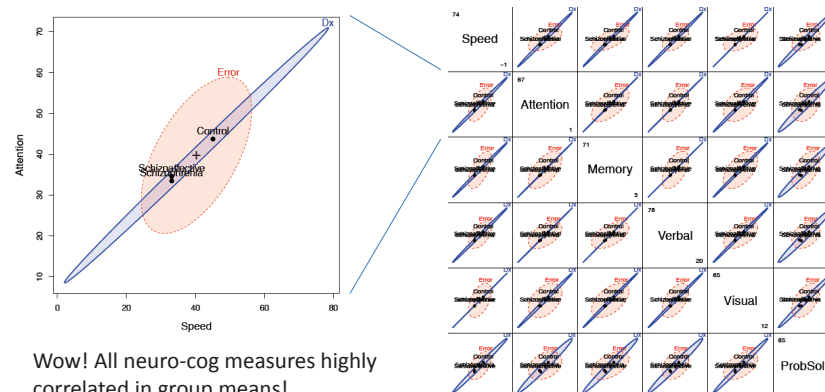
Visualize me: in data space

Bivariate view for any 2 responses:

```
heplot(NC.mlm, var=1:2, ...)
```

HE plot matrix: for all responses

```
pairs(NC.mlm, ...)
```

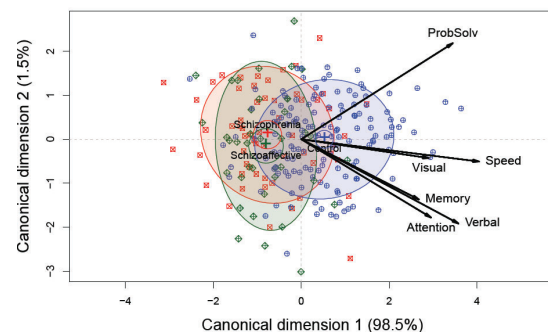


Wow! All neuro-cog measures highly correlated in group means!
Only 1 dim. of H variation

6

Visualize me: in canonical space

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



Very simple interpretation

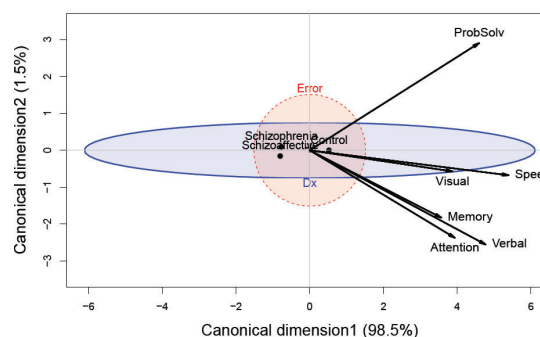
Can1: normal vs. others
All vars highly correlated;

Can2: only 1.5%, NS; but perhaps suggestive

7

Visualize me: canonical HE plots

- Canonical HE plot is just the HE plot of canonical scores, (z_1, z_2) in 2D, or z_1, z_2, z_3 in 3D.
- As in biplot, we add vectors to show relations of the y_i response variables to the canonical variates.
- variable vectors here are **structure coefficients** = correlations of variables with canonical scores.



The multivariate "juicer"

Shows just group means, H ellipse & E ellipse

Variable vectors offer interpretation of Can dimensions.

8

Social cognitive measures

- These measures deal with the person's perception and cognitive processing of emotions of others
 - Scales: managing emotions, theory of mind, externalizing bias, personalizing bias
- Questions:
 - Do these differentiate normal from patient groups?
 - Can they distinguish between schizophrenic & schizoaffective
 - If so, this could be a major finding.

9

Social cognitive measures

```
> data(SocialCog, package="heplots")
> 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 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Test contrasts: Dx1 = Normal vs. Patient; Dx2 = Schizo vs. Schizoaffective

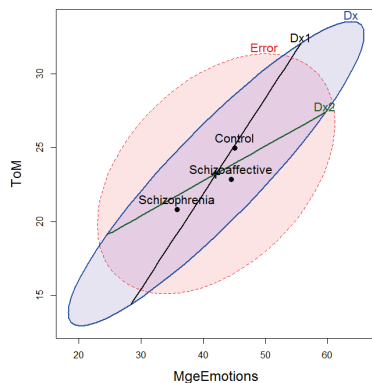
```
> print(linearHypothesis(SC.mlm, "Dx1"), SSP=FALSE)
Multivariate Tests:
Pillai      Df test stat approx F num Df den Df Pr(>F)
1      0.1355      5.212      4    133 0.000624 ***

> print(linearHypothesis(SC.mlm, "Dx2"), SSP=FALSE)
Multivariate Tests:
Pillai      Df test stat approx F num Df den Df Pr(>F)
1      0.0697      2.493      4    133 0.0461 *
```

10

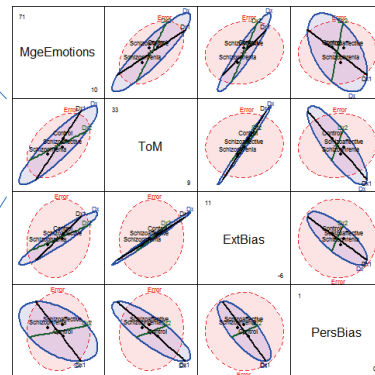
Visualize me: data space

```
heplot(SC.mlm,
        hypotheses=list("Dx1", "Dx2"),...)
```



The groups are ordered Schizo < ScAffective < Control on these measures

```
pairs(SC.mlm,
       hypotheses=list("Dx1", "Dx2"),...)
```



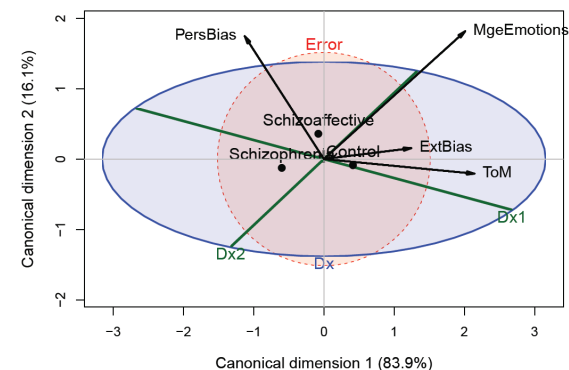
One of these vars is not like the others!

11

Visualize me: canonical space

Contrasts:

Dx1 : Control vs. patients. Controls > patients on MgeEmotions, ExtBias, ToM
Dx2 : Schizo vs. schizoaffective.



Can1: group order

Can2: Schizoaffective vs. others

12

Model checking & remedies

- The MLM assumes residuals are multivariate normal
 - Squared Mahalanobis distances

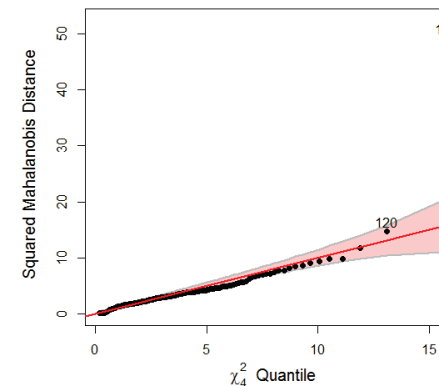
$$D_M^2(\mathbf{y}_i) = (\mathbf{y}_i - \bar{\mathbf{y}})^T \mathbf{S}^{-1} (\mathbf{y}_i - \bar{\mathbf{y}}) \sim \chi_p^2 \text{ with } p \text{ d.f.}$$
 - a quantile – quantile plot of ordered $D_M^2(\mathbf{y}_i)$ vs. quantiles of χ_p^2 should plot as straight line
 - Outliers are readily apparent
 - plots: `heplots::cqplot()`
- Influence plots
 - `mvinfluence::influence()` calculates multivariate analogs of influence measures

13

Social cog: cqplot

`heplots::cqplot()` creates a chi-square QQ plot from a MLM

`cqplot(SC.mlm, id.n=2)`



One observation appears as an extreme outlier.

This was a case w/ ExtBias = -33, but valid range = (-10, +10)

Refitting w/o case 15:
Overall & DX1 tests still OK ☺
Dx2 test: $p=0.074$, now NS ☹

14

Social cog: Influence

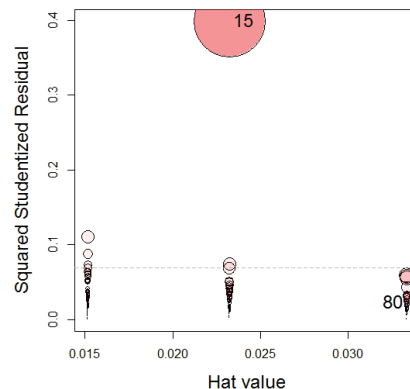
`mvinfluence::influencePlot()` creates a multivariate analog of an influence plot

```
>library(mvinfluence)
>influencePlot(SC.mlm, ...)
```

	H	Q	CookD	L	R
15	0.0233	0.3985	0.4202	0.0238	0.4080
80	0.0333	0.0233	0.0352	0.0345	0.0241

Case 15 stands out as hugely influential

The 3 columns of circles correspond to the 3 groups



15

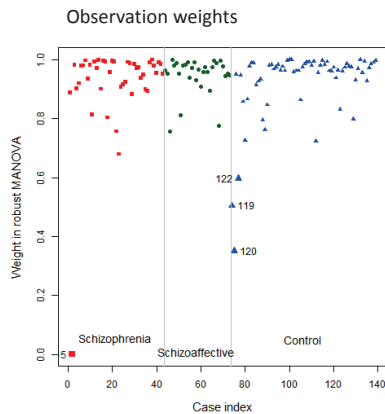
Robust MLMs

- Robust methods for univariate LMs are now well-developed and implemented
 - proper SEs, CIs and hypothesis tests
 - Analogous methods for multivariate LMs are a current hot research topic
 - The `heplots` package now provides `robmlm()` for the fully general MLM (MANOVA, MMReg)
 - Uses simple M-estimator via IRLS
 - Weights: calculated from Mahalanobis D^2 : a robust covariance estimator and weight function, $\psi(D^2)$
- $$D^2 = (\mathbf{Y} - \bar{\mathbf{Y}})^T \mathbf{S}_{\text{robust}}^{-1} (\mathbf{Y} - \bar{\mathbf{Y}}) \sim \chi_p^2$$
- Downside: SEs, p -values only approximate

16

Robust MLMs

```
SC.rlm <- robmlm(cbind( MgeEmotions, ToM, ExtBias, PersBias) ~ Dx,
  data=SocialCog)
```



Approx test of Dx2 in robust model

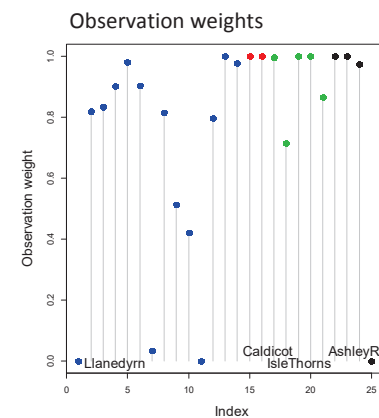
```
> print(linearHypothesis(SC.rlm, "Dx2"), SSP=FALSE)

Multivariate Tests:
Df test stat approx F num Df den Df Pr(>F)
Pillai 1 0.069 2.44 4 132 0.05 *
Wilks 1 0.931 2.44 4 132 0.05 *
```

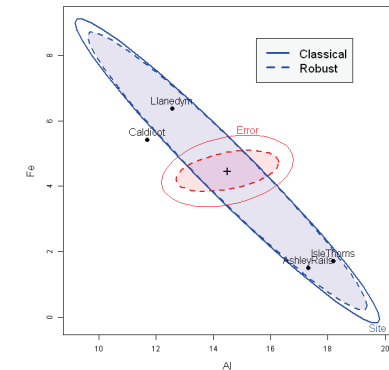
17

Robust MLMs: Pottery data

```
> pottery.mod <- lm(cbind(Al,Fe,Mg,Ca,Na)~Site, data=Pottery)
> pottery.rmod <- robmlm(cbind(Al,Fe,Mg,Ca,Na)~Site, data=Pottery)
```



overlaid HE plots



18

MMRA example: PA tasks & ability

- Rohwer data from Timm (1975)
- How well do paired associate (PA) tasks predict performance on measures of aptitude & achievement in kindergarten children?
 - Samples: 69 children in two groups (schools): 'Lo' | 'Hi' SES
 - Outcomes (Y):
 - Scholastic aptitude test (SAT)
 - Peabody picture vocabulary test (PPVT)
 - Raven progressive matrices (Raven)
 - Predictors (X): Scores (0—40) on PA tasks where the stimuli were:
 - named (n), still (s), named-still (ns), named-action (na), sentence-still (ss)

```
> data("Rohwer", package="heplots")
> car::some(Rohwer, n=5)
  group SES SAT PPVT Raven n s ns na ss
8      1 Lo  8  68   8  0  0 10 19 14
9      1 Lo 49  74  11  0  0  7 16 13
17     1 Lo 19  66  13  7 12 21 35 27
52     2 Hi  38  66  14  0  0  3 16 11
66     2 Hi   8  55  16  4  7 19 20 13
```

Having a group factor makes the analysis more complicated (MANCOVA)

Start with analysis of the Hi SES group

```
> Rohwer2 <- subset(Rohwer, subset=SES=="HI")
```

19

Why not univariate models?

```
rohwer.mod1 <- lm(SAT ~ n + s + ns + na + ss, data = Rohwer2)
rohwer.mod2 <- lm(PPVT ~ n + s + ns + na + ss, data = Rohwer2)
rohwer.mod3 <- lm(Raven ~ n + s + ns + na + ss, data = Rohwer2)
```

```
library(stargazer)
stargazer(rohwer.mod1, rohwer.mod2, rohwer.mod3, type="text", ...)
```

Univariate regression models for Rohwer data

	Dependent variable:		
	SAT	PPVT	Raven
n	3.26*	0.07	0.06
s	3.00	0.37	0.49**
ns	-5.86***	-0.37	-0.16
na	5.67***	1.52*	0.12
ss	-0.62	0.41	-0.12
R2	0.56	0.35	0.31
F Statistic (df = 5; 26)	6.54***	2.85*	2.32

Note: *p<0.05; **p<0.01; ***p<0.001

Results are disappointing

- Only model for SAT highly signif.
- Only a few coefs. signif. ≠0

20

MANOVA tests

```
> rohwer.mlm <- lm(cbind(SAT, PPVT, Raven) ~ n + s + ns + na + ss, data=rohwer2)
> Anova(rohwer.mlm)
```

```
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
```

Much better!

Multivariate tests:

- pool evidence for all Ys
- take correlations of Ys into account

Can also test overall hypothesis, $H_0: \mathbf{B} = \mathbf{0}$ (all coefs = 0)

```
> print(linearHypothesis(rohwer.mlm,
+ c("n", "s", "ns", "na", "ss")), SSP=FALSE)
```

Multivariate Tests:

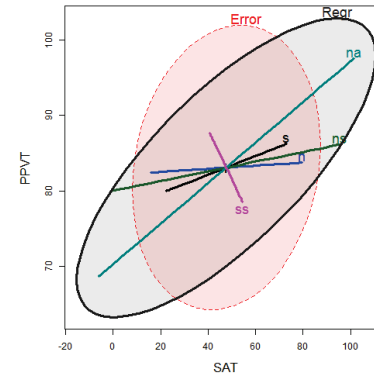
```
Df test stat approx F num Df den Df Pr(>F)
Pillai 5 1.0386 2.753 15 78.00 0.001912 **
Wilks 5 0.2431 2.974 15 66.65 0.001154 **
Hotelling-Lawley 5 2.0615 3.115 15 68.00 0.000697 ***
Roy 5 1.4654 7.620 5 26.00 0.000160 ***
```

Strongly reject H_0 by all criteria

21

Visualize me!

```
cols <- c("red", "blue", "black", "darkgreen", "darkcyan", "magenta", "gray20")
hyp <- list("Regr" = c("n", "s", "ns", "na", "ss")) # Test of B = 0
heplot(rohwer.mlm,
hypoteses = hyp,
fill=TRUE, fill.alpha=0.1, col=cols, lwd=c(1,3))
```



Each predictor gives a 1 df test -> H_i ellipse is a line
 \mathbf{E} here is a 3D ellipsoid ($\text{rank}(\mathbf{E}) = \min(p, q)$)

Interpretation:

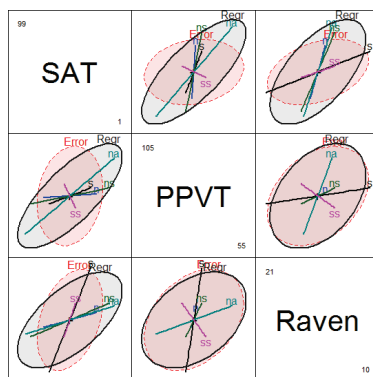
- Any H_i ellipse that protrudes outside \mathbf{E} ellipse is significant by Roy's test
- Length of each H_i line \approx strength of association
- Orientation of each H_i line shows relation of X_i to the two Y_j responses shown.

22

pairs.mlm() plot

View all pairwise HE plots

```
pairs(rohwer.mlm, hypotheses=hyp, col=cols, fill=TRUE, fill.alpha=0.1, ...)
```



Can now understand more subtle aspects

SAT is best predicted overall, but relation with PA tests varies
 The na & ns tasks are strongest for SAT

Raven is weakly predicted

23

Canonical correlations

For quantitative (X, Y) data, canonical correlation analysis is an alternative to MMRA
 It finds the weighted sums of the Y variables most highly correlated with the Xs

```
> X <- rohwer2[, 6:10] # X variables for High SES students
> Y <- rohwer2[, 3:5] # Y variables for High SES students
> (cc <- cancel(X, Y, set.names=c("PA", "Ability")))
```

Canonical correlation analysis of:

with 5 PA variables: n, s, ns, na, ss
 3 Ability variables: SAT, PPVT, Raven

	CanR	CanRSQ	Eigen	percent	cum	scree
1	0.7710	0.5944	1.4654	71.080	71.08	*****
2	0.5465	0.2987	0.4259	20.659	91.74	*****
3	0.3815	0.1455	0.1703	8.261	100.00	***

Test of H_0 : The canonical correlations in the current row and all that follow are zero

	CanR	LR	test stat	approx F	numDF	denDF	Pr(> F)
1	0.77096		0.24307	2.9738	15	66.655	0.001154 **
2	0.54652		0.59926	1.8237	8	50.000	0.094538 .
3	0.38147		0.85448	1.4759	3	26.000	0.244178

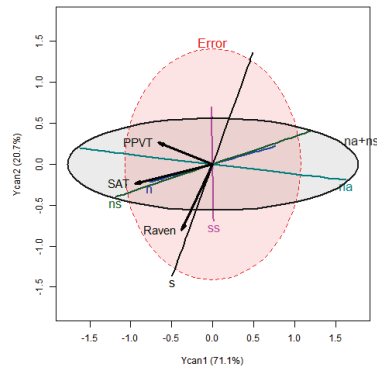
Two dimensions acct for 91.7% of (X,Y) association

Only Can1 is significant

24

Visualize CCA in HE plot

```
cols <- c("red", "blue", "black", "darkgreen", "darkcyan",
          "magenta", "gray20")
heplot(cc, hypotheses=list("na+ns=c('na', 'ns')"),
       fill = TRUE, fill.alpha=0.1, col=cols,
       label.pos = c(3, rep(1,5), .1),
       cex=1.4, var.cex=1.25, var.lwd=3, var.col="black")
```



Residuals are uncorrelated in canonical space

H ellipses for X terms same as in ordinary HE plots – outside E ellipse *iff* signif. by Roy's test

Variable vectors for Ys: correlations with canonical variables Ycan1, Ycan2

- SAT & PPVT: mainly Ycan1
- Raven: more aligned with Ycan2

25

MANCOVA & homogeneity of regression

- With a group variable (SES) can test differences in means (intercepts)
 - `rohwer.mod <- lm(cbind(SAT, PPVT, Raven) ~ SES + n + s + ns + na + ss, data=Rohwer)`
 - This assumes that slopes (B) are the same for both groups (homogeneity of regression)
- Can test for equal slopes by adding interactions of SES with Xs
 - `rohwer.mod1 <- lm(cbind(SAT, PPVT, Raven) ~ SES * (n + s + ns + na + ss))`
- Or, fit separate models for each group

```
rohwer.ses1 <- lm(cbind(SAT, PPVT, Raven) ~ n + s + ns + na + ss,
                  data = Rohwer, subset = SES == "Hi")
rohwer.ses2 <- lm(cbind(SAT, PPVT, Raven) ~ n + s + ns + na + ss,
                  data = Rohwer, subset = SES == "Lo")
```

26

MANCOVA

Fit the MANCOVA model & test hypotheses

```
> 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

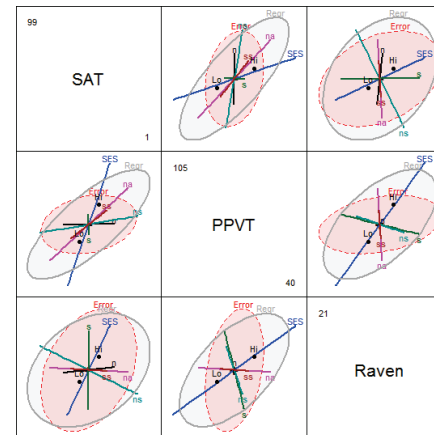
	Df test	stat	approx F	num Df	den Df	Pr(>F)
SES	1	0.379	12.18	3	60	2.5e-06 ***
n	1	0.040	0.84	3	60	0.4773
s	1	0.093	2.04	3	60	0.1173
ns	1	0.193	4.78	3	60	0.0047 **
na	1	0.231	6.02	3	60	0.0012 **
ss	1	0.050	1.05	3	60	0.3770

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

27

Visualize effects

```
pairs(rohwer.mod,
      hypotheses=list("Regr" = c("n", "s", "ns", "na", "ss")),
      fill=TRUE, fill.alpha=0.1)
```



The SES effect is positive for all Y variables
Hi SES group > Lo SES group

28

Fit model with interactions

Fit heterogeneous regression model with SES interactions

```
> rohwer.mod1 <- lm(cbind(SAT, PPVT, Raven) ~ SES * (n + s + ns + na + ss),
+ data=Rohwer)
> Anova(rohwer.mod1)
```

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				

OK, as expected

Hmm ???

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

29

Test interactions

Can test all interactions **simultaneously** with linearHypothesis()

Do I need any interaction terms?

I use 'grep' trick here to find the names of coefficients like 'SES:' containing a ':'

```
> coefs <- rownames(coef(rohwer.mod1)) # store coefficient names in a vector
> print(linearHypothesis(rohwer.mod1, # only test for interaction effects
+ coefs[grep(":", coefs)]), SSP=FALSE)
```

Multivariate Tests:

	Df	test	stat	approx	F	num	Df	den	Df	Pr(>F)
Pillai	5	0.4179	1.845	15	171.0	0.03209	*			
Wilks	5	0.6236	1.894	15	152.2	0.02769	*			
Hotelling-Lawley	5	0.5387	1.927	15	161.0	0.02396	*			
Roy	5	0.3846	4.385	5	57.0	0.00191	**			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Evidence shows that some slopes differ for Hi/Lo SES

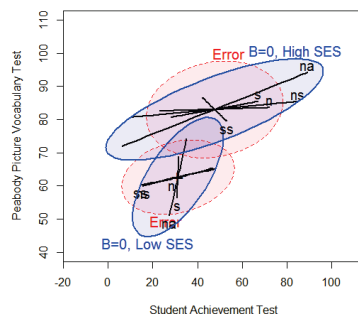
30

Fit separate models

Fitting a model for each group allows all slopes to differ

Also allows within-group covariances to differ

```
rohwer.ses1 <- lm(cbind(SAT, PPVT, Raven) ~ n + s + ns + na + ss,
+ data = Rohwer, subset = SES == "Hi")
rohwer.ses2 <- lm(cbind(SAT, PPVT, Raven) ~ n + s + ns + na + ss,
+ data = Rohwer, subset = SES == "Lo")
```



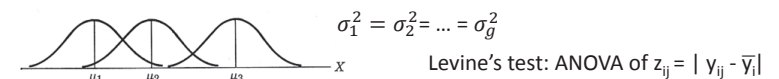
For SAT & PPVT:

- means higher for Hi SES
- within-group covariance larger for Hi SES
- slopes of predictors smaller for Hi SES
→ SAT more important for this group.

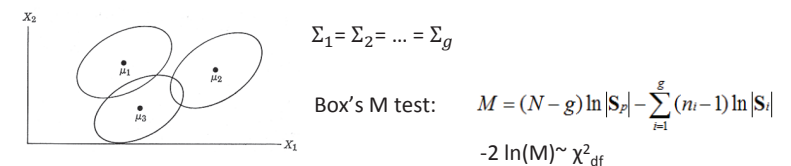
31

Homogeneity of (co)variances

- ANOVA assumes equality of residual variances



- MANOVA: assumes equality of covariance matrices



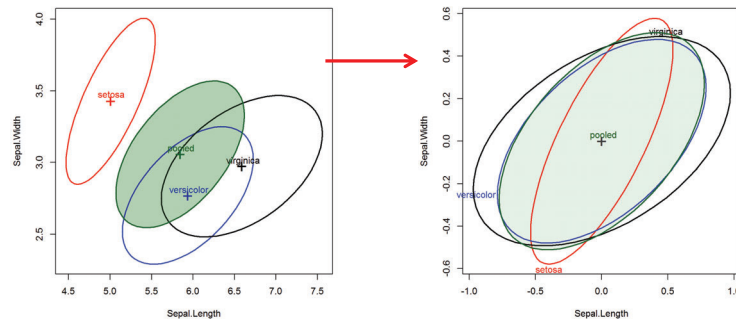
See: <http://www.datavis.ca/papers/EqCov-TAS.pdf>

32

Visualizing covariance matrices

Visualize covariance ellipses in data space
Center to see pure differences in size & shape

```
covEllipses(iris[,1:2], iris$Species, ...)  
covEllipses(iris[,1:2], iris$Species, center=TRUE, ...)
```



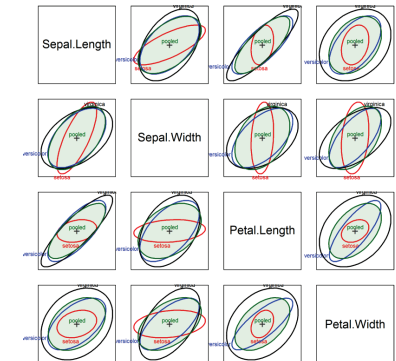
33

View all pairs

```
covEllipses(iris[,1:4], iris$Species,  
center=TRUE, variables=1:4, ...)
```

In all cases, **setosa** stands out as different from the others

- sometimes correlation differs
- sometimes smaller variance(s)

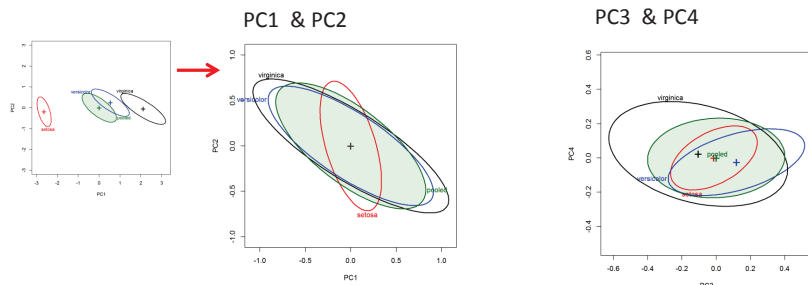


34

Visualize in PCA space

PCA projects the data into an orthogonal space accounting for maximum variance
Covariance ellipses show the differences among groups in this space

Surprisingly, the small dimensions contribute largely to Box's M test.



```
iris.pca <- prcomp(iris[,1:4])  
covEllipses(iris.pca$x, iris$Species, ...)  
covEllipses(iris.pca$x, iris$Species, center=TRUE, ...)
```

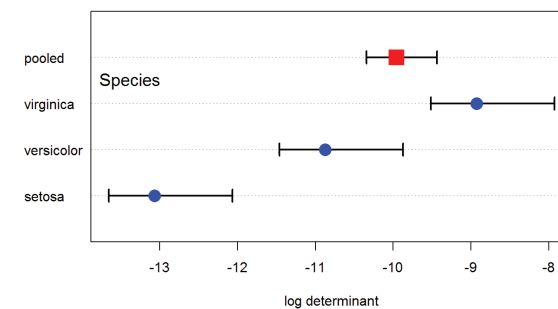
```
covEllipses(iris.pca$x, iris$Species, center=TRUE,  
variables=3:4, ...)
```

35

Visualizing Box's M test

Box's test is based on a comparison of the $\log |S_i|$ relative to $\log |S_p|$: **plot them!**

```
iris.boxm <- boxM(iris[, 1:4], iris[, "Species"])  
plot(iris.boxm, glabel="Species")
```



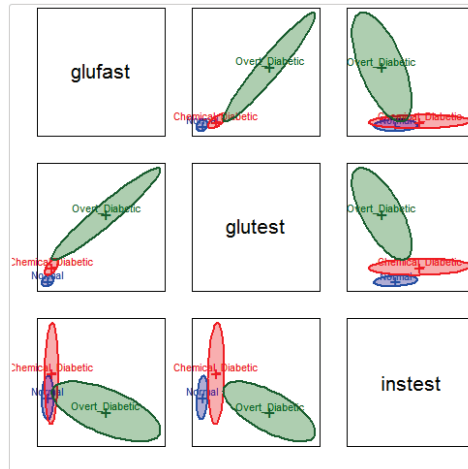
CIs based on an asymptotic CLT \approx distribution of $\ln |S|$ (Cai, Liang, and Zhou 2016)
(Thx: Augustine Wong)

Unsolved: Bootstrap CI

36

Diabetes data

```
covEllipses(Diabetes[,2:5], Diabetes$group, fill=TRUE, pooled=FALSE,
  col=c("blue", "red", "darkgreen"), variables=1:3)
```



Two groups of diabetic patients and a normal group were measured on blood glucose and insulin

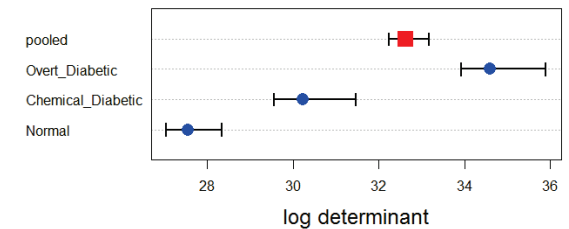
The differences in correlation and variances are dramatic here

37

Box's M test

```
diab.boxm <- boxM(Diabetes[,2:5], Diabetes$group)
diab.boxm
##
## Box's M-test for Homogeneity of Covariance Matrices
##
## data: Diabetes[, 2:5]
## Chi-Sq (approx.) = 383, df = 20, p-value <2e-16
plot(diab.boxm)
```

All pairs differ!



38

Summary

- MANOVA tests of MLMs are easily visualized in HE plots
 - Contrasts among groups can be easily shown
 - Canonical plots show data in 2D/3D space of max. group differences
 - Robust methods can help guard against outliers
- MMRA models
 - Visualize effects of quant. predictors as lines in data space
 - Test & visualize any linear hypothesis
 - Canonical correlations: visualize in 2D/3D of max. (X, Y) correlations
- Homogeneity of covariances
 - Visualize within-group S_i and pooled S_p by data ellipses
 - Visualize Box's M test by simple dot plot of $|S_p|$ and $|S_i|$

39