Lab. notes for MANOVA

Paul Hewson

October 1, 2007

1 Manova practical

The first exercise is based on an example in Johnson and Wichern (2002), page 344 exercise 6.24, although we seem to have a little more data than they do (we have five time periods, including -200 BCE and 150 CE). This is rather a famous data set mentioned in a large number of multivariate text books. Essentially, we have data on Maximum Breadth, Basal Height, Basal Length and Nasal Height for skulls recovered from a number of different time periods. Our interest is in whether the skulls have changed shape and size over time (Johnson and Wichern suggest this may be due to interbreeding with other populations.

Data have been placed in the portal.

First, read in the data:

```
> es <- read.csv("egyptian-skulls.csv")
```

You should be able to do scatterplots of each group by now! Look at some of code for the cluster analysis if you're not sure. What other eda would be useful for these data?

One slight annoyance is that we have to have our response data as a matrix, and our group memberships as a factor

```
> Y <- as.matrix(es[,-5])
> Year <- as.factor(es$Year)
```

More information on the manova command can be found on the helpfile in **R**. This happens to be quite a simple model to fit:

```
> es.manova <- manova(Y ~ Year)
```

And you can get the model fit simply by using `summary()` (with an optional call to `test` to give you one of the four fit criteria. i.e. `test = "Pillai"` or `test = "Wilks"` or `test = "Hotelling-Lawley"` or `test = "Roy"`. What is the default test statistic? Why do you think that one has been chosen?

```
> summary(es.manova, test = "Wilks")
```

You can also consider the various univariate ANOVAs as follows:

```
> summary.aov(es.manova)
```

- Why didn't we just do a whole set of univariate ANOVA?

2   A two way MANOVA

The data on plastic film extrusion is rather famous (and is actually in the R helpfile for summary.manova) and is worked as an example in Johnson and Wichern (page 312, Example 6.11). You can load the data from the Rex.csv file in the portal.

```
> rex <- read.csv("Rex.csv")
```

Again, you may find it useful to conduct a simple eda with these data before you carry out with the analysis.

Do note this time that we will wrap all the Y variables in an `as.matrix` wrapper. For the first call, we have fitted an interaction term between extrusion rate and additive levels:

```
> rex.manova <- manova(as.matrix(rex[,c(1:3)]) ~ rex$rate*rex$additive)
> summary(rex.manova)
```

Look at the diagnostics using Wilk's lambda. Do you need an interaction term?

You may find it useful to compare the results of this analysis with the one worked through in Johnson and Wichern (2002).