Übung zu MUS1V

Multivariate Statistik (DSE VZ WS19)

Michael Petz, Fh Hagenberg

A05

Find a formula for the number of interactions of degree n with v variables. Check this formula for small numbers, where n goes from 2 to 4 and v from 2 to 4 as well (ie 3x3 = 9 calculations including the associated list of interactions). Example: with 2 variables there is only one interaction of degree 2, namely xy. One of the 3 interactions of degree 4 with 3 variables is x^2yz (interactions come from the context of multivariate regressions, as discussed in the last lecture).

A06

a) Formally deduce the three equations for a quadratic regression of the form $\hat{y} = ax^2+bx+c$ (see lecture for the linear case) and represent the result in matrix notation as well.

b) Formally deduce the three equations for a linear regression with two independent variables of the form $\hat{z} = ax+by+c$ and represent the result in matrix notation as well.

A07

Read the file regr.csv and save it in a dataframe df: `df = read.csv("regr.csv")`

The variables are n (=SerNo), p and e (regressors), as well as s, v and d (dependent variables).

You can calculate the (linear) correlation between p and e with `cor(df$p,df$e)`.

With `cor(df)` you get the pairwise correlations of all variables.

a) Calculate all possible linear correlation coefficients and interpret them.

b) Perform nonlinear multivariate regressions for the 3 dependent variables s, v and d in R. Assume effects of the variables p and e up to degree 2 (=quadratic) and interactions up to degree 3 (=cubic). You get an example for such a regression as follows:

```
Regr = lm(s~p+e+I(p^2)+I(e^2)+I(p*e)+I(p^2*e)+I(p*e^2)-1)
```

Examine the output of `summary(Regr)` with respect to the coefficients of the model and try to find a model with less effects than the one above, which has a similar “power” to explain the dependent variable (e.g. s) as the complete model. If you want to eliminate the linear effect of e and the quadratic effect of p, the call in R is:

```
Regr = lm(s~p+I(e^2)+I(p*e)+I(p^2*e)+I(p*e^2)-1)
```

c) As in a) and b), but with the file regr2.csv

Note: 3 assignments = 3 files to upload, with max 2 points for the first two assignments, and max 3 points for the last one due to being a little „lengthy“.