Dealing with Heterogeniety

This page is a reproducible work for Chapter 4.1 of Zuur et al. (2009). This particular example is useful to practice methods to address heterogeniety (heteroscedasticy) in linear regression models.

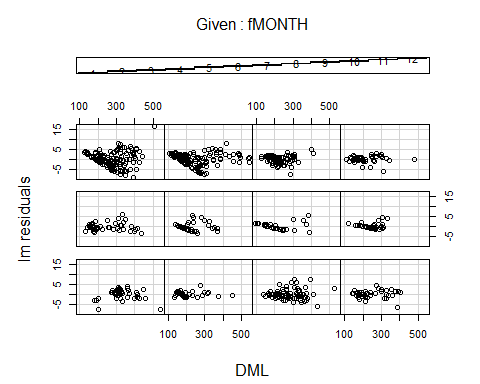
One of the important assumptions in linear regression is “homeoscedasticy (homeogeneity) of variance”, which means that the spread of dataset is same at each X values. (i.e. the spread of reponse is the same for along the range of predictor variable).

One way to deal with heterogeniety is ‘a mean-variance stabilisng’ transformation. Formal test for homeogeniety require normality of the data (e.g. Barlett’s test). Zuur et al (2009) suggest assessing homoegeneity based on ‘graphical inspection’ of residuals (pp20).

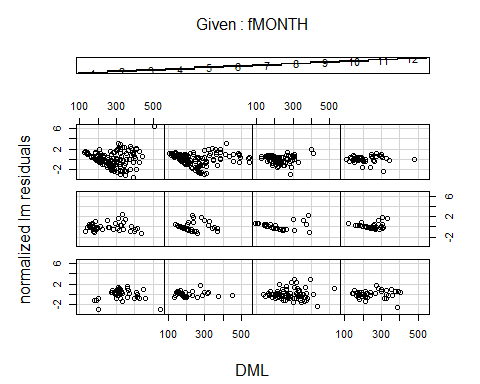
Serious heterogeniety could cause major harm to invalidate the outcome of iinear regression analysis. Ignoring such issue could leave the regression parameters with incorrect standard errors (Zuur et al. 2009), which leads to incorrect distribution of statistics (e.g., F / t statistic is no longer F / t distributed), thus harming statistical significance of the tests.

With extra Mathematical effort, heterogeniety can be incorporated to models and can provide extra biological information (Zuur et al. 2009)

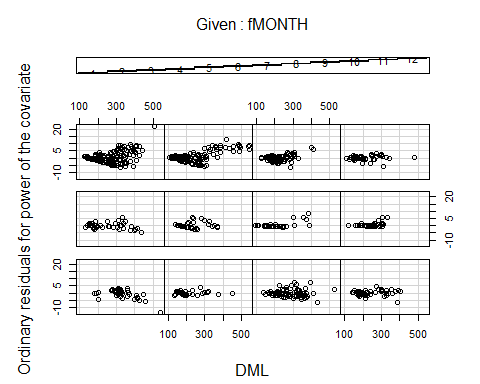
M1<- lm (Testisweight ~DML\*fMONTH, data=Squid)  
E0<-resid(M1)  
coplot(E0~DML|fMONTH,ylab="lm residuals", data=Squid)



M1<-gls(Testisweight ~ DML\*fMONTH, data=Squid)  
E0\_n<-resid(M1, type="normalized")  
coplot(E0\_n~DML|fMONTH,ylab="normalized lm residuals", data=Squid)



### Ordinary residuals  
  
vf4<- varPower(form =~DML| fMONTH)  
M.gls4<-gls(Testisweight ~ DML\*fMONTH, data=Squid, weights=vf4)  
E1<-resid(M.gls4)  
coplot(E1~ DML | fMONTH, ylab ="Ordinary residuals for power of the covariate", data= Squid)



### Ordinary residuals  
E2<-resid(M.gls4,type="normalized")  
coplot(E2~ DML | fMONTH, ylab ="Normalized residuals", data= Squid)

