**Fitting a curve to the 95% quantile of the Salinity/Depth data**

***Background on approach***

To fit a curve to the 95% quantile of the data I created a linear model that models the error structure of the data. A typical linear model has the form:

y = beta0 + beta1\*x1 + beta2\*x2 + ... + betaN\*xN + epsilon

where epsilon represents the unexplained variance or the residuals. Typically these are assumed to be normally distributed with a constant variance. That is

epsilon ~ Norm( mu = 0, sigma2 = constant)

However, it is possible to model changes in the standard deviation as one would model changes in the mean. In such a model, the standard deviation structure of the epsilons is given by something of the form:

epsilon ~ Norm(mu = 0, sigma = gamm0 + gamma1\*x1 + gamma2\*x2 + ... + gammaN\*xN)

In such a model, the variance in the epsilons (residuals) is allowed to change. For more details regarding this kind of model, see Bolker 2008.

Of course, the error model does not have to be linear.

For the Salinity/Depth data set I used the following model:

Depth = beta0 + beta1\*salinity + epsilon

epsilon ~ Norm( mu = 0, sd = beta2 + beta3\*exp(beta4\*salinity)

The betaX estimate are found using maximum likelihood.

I did consider a model where beta2 = 0, but the result is sd -> 0 as salinity -> 30 ppt. That is, the maximum likelihood estimate for beta4 results in values for the standard deviation that are near zero for salinity values beyond about 25 ppt. Since we are interested in the upper boundary of the data as phenomena, I am not especially concerned with trying to interpret beta2 per se. If we were interested in the mechanism creating this boundary, I would be more concerned with understanding why this particular functional form is meaningful.

I considered two versions of the model, one where beta0 and beta1 were fixed (beta0 = beta1 = 0) and one where beta0 and beta1 were estimated. The results are visualized in figure 1.

***Model1 results:***

beta0 = 0

beta1 = 0

beta2 = 0.084

beta3 = 0.085

beta4 = -0.19

-log(L) = -1242.68

***Model2 results:***

beta0 = 0.0058

beta1 = -0.0021

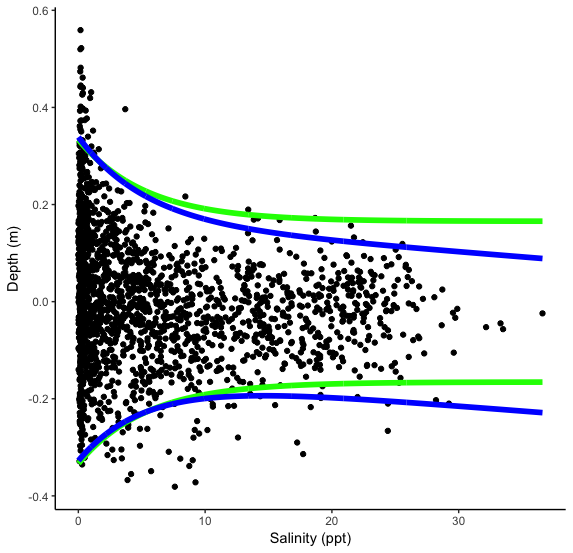
beta2 = 0.081

beta3 = 0.089

beta4 = -0.19

-log(L) = -1279.87

There are other potential approaches but this seemed like a straightforward one.



**Figure 1.** Salinity/Depth distribution of marsh species in coastal Louisiana. Each point represented the mean salinity and depth for a single CRMS station. The green and blue lines represent 95% boundaries for the data. The green line shows the 95% boundary for model1 where beta0 and beta1 were fixed at zero. The blue line shows the 95% boundary for model2 where beta0 and beta1 were estimated.

***Bibliography***

Bolker, B., 2008. Ecological Models and Data In R. Princeton University Press.