EEB313 Mid Project Update

Gabi, Golshan, Samara, Jared

2024-11-14

The original paper, "Investigating the role of biological characteristics on adaptive capacity within alien and rare plant species in China" had several flaws which prevented us from continuing to use its dataset for our project. Namely, the authors attempted to quantify adaptive capacity by simply asking "five experts" to rank categories of plant traits and their levels (e.g., in the category of reproductive mode there were three levels: (a) sexual, (b) asexual, or (c) combined sexual-asexual reproduction) they ranked from "most important: 5" to "least important: 1". Then all of these scores for all plant traits were compiled using a formula that was not fully disclosed in the main article, rather, in the complementary supplementary files. Despite the value of "expert opinions", more robust quantitative methods of determining significant variables should have been utilized. Additionally, it is not a simple feat to quantify adaptive capacity through various "expert-informed" (as they mention) traits as this typically requires more comprehensive methods that include genetic, environmental, and evolutionary variables considered together. Considering these constraints, this dataset would have been incredibly hard to utilize in a unique and useful way.

Instead, we have chosen to analyze the data from a paper by Jiang et al. titled "Restoration of native saltmarshes can reverse arthropod assemblages and trophic interactions changed by a plant invasion". The authors of the paper investigated how arthropod assemblages and trophic interactions changes in estuary salt marshes that are invaded with the non native Spartina alterniflora. They established 15 plots across 5 different plant communities which represents a gradient of stages of Spartina (exotic plant) invasion and removal: OP: Original Phragmites monoculture, TP: Threatened Phragmites monoculture, PS: Phragmites–Spartina mixture, IS: Invasive Spartina monoculture, and RP: Restored Phragmites monoculture. One of the datasets Jiang et al. (2022) provides is a set of raw data which includes four sheets: Arthropod list, Plant traits, Soil traits, and Stable isotopes. To test our hypothesis, we ultimately decided to focus on which traits, whether it be plant or soil traits, is/are most significant in explaining the variation in arthropod community (i.e. Shannon diversity, recall: relative abundance and evenness) across plant communities, and if there are differences in how these traits predict arthropod diversity between plant communities.

We can find the same conclusions of Jiang et al. (2022) in our own way to determine and verify the results. Jiang et al. found that the plant traits most responsible for driving changes in arthropod diversity was aboveground biomass, leaf N, and plant density. They found that the soil trait most responsible for driving changes in arthropod diversity was soil salinity. Here we state:

H0: Soil and plant traits represented by principal component axes (PC1 and PC2) do not significantly predict arthropod diversity across all communities.

H1: Soil and plant traits represented by principal component axes (PC1 and PC2) are strong predictors of arthropod diversity across all communities.

H0.2 Soil and Plant traits represented by principal component axes do not differ in how they predict arthropod diversity across different communities.

H2: Soil and Plant traits represented by principal component axes differ in how they predict arthropod diversity across different communities.

To prepare our data for analysis, we had to reformat the data frame. We started by separating each sheet (plant traits, soil traits, and arthropod list) on the provided Excel file to their own csv with altered simplified

columns. Then, we worked with the arthropod data frame that listed the raw abundances of each arthropod species in every treatment in each of the 5 plant communities. The original data frame had each treatment as columns, and every arthropod species listed in the first column. We transformed the data so the first column would have the treatments and every subsequent column had species as headers. We also removed unnecessary columns (columns that had order, family and trophic group, as well as a column that only had NA values).

Once the data was properly formatted, we could continue with the data analysis:

The following method will be used to create a linear model to determine the strongest predictors of Shannon diversity of all our communities. We intend to compare our results with that of the original authors, in terms of determining which plant and soil traits are primarily responsible for driving differences in arthropod diversity across communities and groups

- We calculated Shannon diversity per treatment for all of the communities (15 treatments for 5 communities = 75 total Shannon diversity values) (we calculated Shannon diversity because the authors chose that diversity metric, we want to keep it consistent so our data is as comparable as possible)
- We then determined the distribution type that best fits the Shannon diversity indices by conducting maximum log-likelihood calculations for a normal, gamma, and inverse gaussian probability density distribution. We used dnorm(), dgamma(), and a hand-made inverse gaussian formula to loop over a range of parameter values.
- Our MLE for normal was -19.28 (when mean = 3.43 and sd = 0.31), the MLE for gamma was -20.56 (when shape = 116.1 and scale = 0.0295), the MLE for inverse gaussian was -21.73 (when mu = 3.4 and lambda = 379.6). Thus, a normal probability density distribution fits our data best. This allows us to use a linear regression model for our future analysis.
- We ran a pairwise correlation test between each of the 11 plant and soil traits.
- We found many of the soil traits to have high correlation values, and thus be linearly correlated. On the other hand, our plant traits did not show high correlations between other traits. We intend to use this data further in our analysis to interpret our results of the linear models to understand the variable aggregations within our PC axes.
- Due to the strong correlations between plant and soil traits, we decided to run a PCA to determine which of our 11 plant and soil traits contributed most strongly to the variation between our treatments to inform model selection, we generated a PCA of all possible predictor variables from the dataset (Soil salinity, soil N, soil C, soil P, soil moisture, soil pH, aboveground plant biomass, leaf C, leaf N, leaf P). We then determined which traits fall on each PC axis using contrib/pc.out function.
- The plant and soil traits that are associated with PC1 are the most responsible (55.64%) for driving variance between treatments, with Soil N content (12.6%), Soil moisture (12.5%) and soil pH (11.4%) being the most significant. For PC2 (10.5% of total variance), the variables, soil C content (20.1%), plant density (16.3%) and soil salinity (15.4%) contributed most strongly.

The analysis we still need to perform are:

- Perform model selection, using the PC axes as fixed effects
- Generate the following models:
- with only PC1
- with both PC1 and PC2,
- with both PC1, PC2 and their interaction
- Select the model with the lowest AIC score

- The fixed effect (PC axis) with the largest coefficient (absolute value) is the most responsible for driving changes in arthropod diversity (Test H1)
- Conduct a group-by-group analysis of each plant community - to see if there are any differences in the drivers of diversity in each community (Test H2)
- Compare results with the original paper

R Code so far:

Loading packages

```
require(tidyverse)
require(deSolve)
require(vegan)
require(car)
require(rgl)
require(ggfortify)
require(fitdistrplus)

options(rgl.useNULL = TRUE)
rgl::setupKnitr(autoprint = TRUE)
```

Loading csv files

```
#Provided original csv files (separated sheets from the master excel file)
# column names were edited for simplicity sake
read_csv("Arthropod_List.csv") -> Arthropod_List
```

```
## Rows: 349 Columns: 79
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (4): Order, Family, Species, Trophic_Group
## dbl (75): OP1, OP2, OP3, OP4, OP5, OP6, OP7, OP8, OP9, OP10, OP11, OP12, OP1...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
read_csv("Plant_Traits.csv") -> Plant_Traits
```

```
## Rows: 75 Columns: 6
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (2): Site, Leaf_N
## dbl (4): Aboveground_Biomass, Plant_Density, Leaf_C, Leaf_P
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
read_csv("Soil_Traits.csv") -> Soil_Traits
```

```
## Rows: 75 Columns: 7
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (1): Site
## dbl (6): Soil_C, Soil_N, Soil_P, Soil_pH, Soil_Salinity, Soil_Water
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

Rearraged Arthropod_List dataframe (each species is a column, each treatment is a row)

```
teatment_cols <- Arthropod_List[,5:79]
species_col <- Arthropod_List[c("Species")]

new_AL <- data.frame(species_col, teatment_cols)

AL_long <- new_AL %>%
  pivot_longer(
    cols = (starts_with("OP") | starts_with("RP") | starts_with("TP") |
              starts_with("IS") | starts_with("PS")), # Select treatment columns
    names_to = "Treatment", # Name of new "Treatment" column
    values_to = "Count" #  # Name of new "Count" column
  )
```

```
AL_wide <- AL_long %>%
  pivot_wider(
    names_from = Species, # Create new columns from Species
    values_from = Count   # Fill values from the "Count" column
  )

fixed_arthro <- AL_wide
```

Shannon Diversity Calculations

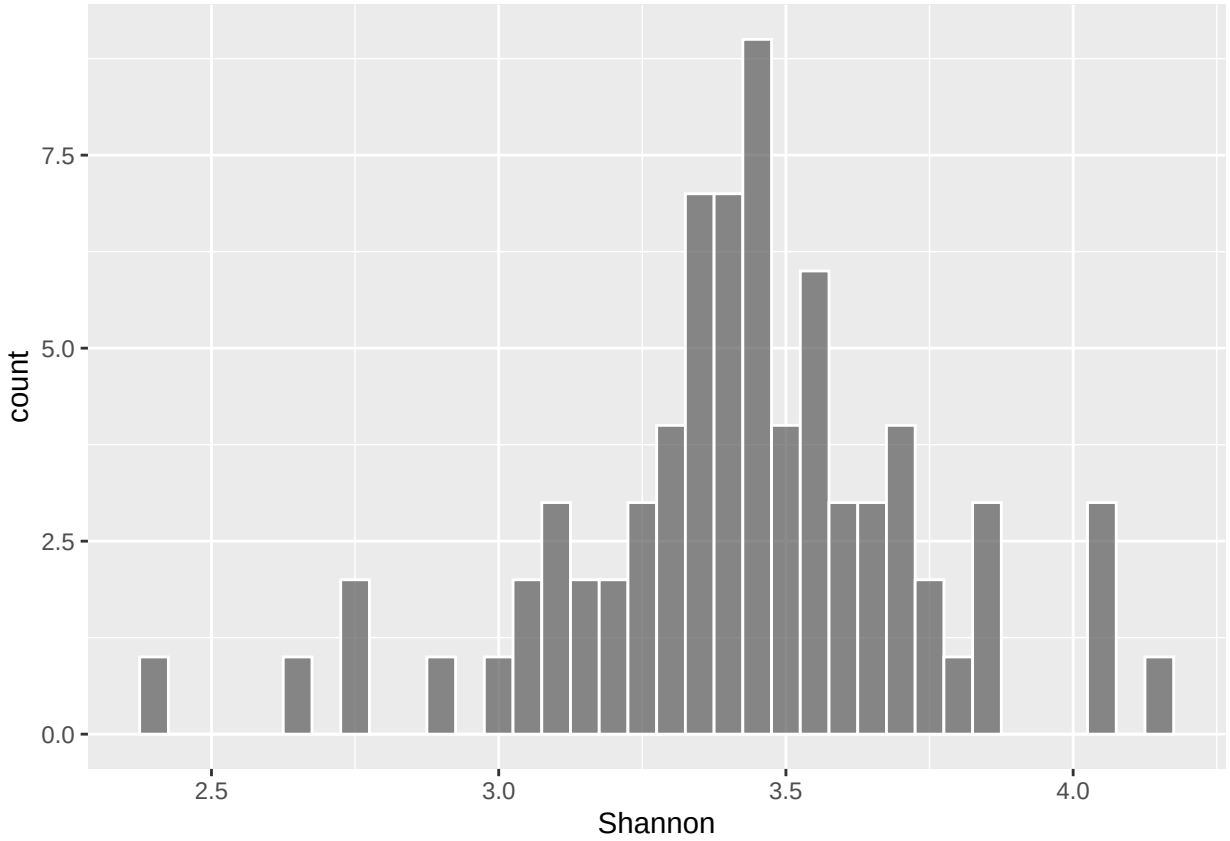
```
#Excluding the first and last column (as they d not contain data)
for_shannon = fixed_arthro[,2:350]

#Shannon diversity
shannon_indices = diversity(for_shannon, index = "shannon")
```

```
#Reformatting with column names specified
final_shannon = data.frame(Treatment = fixed_arthro$Treatment,
                           Shannon = shannon_indices)
```

Histogram of Shannon Diversity

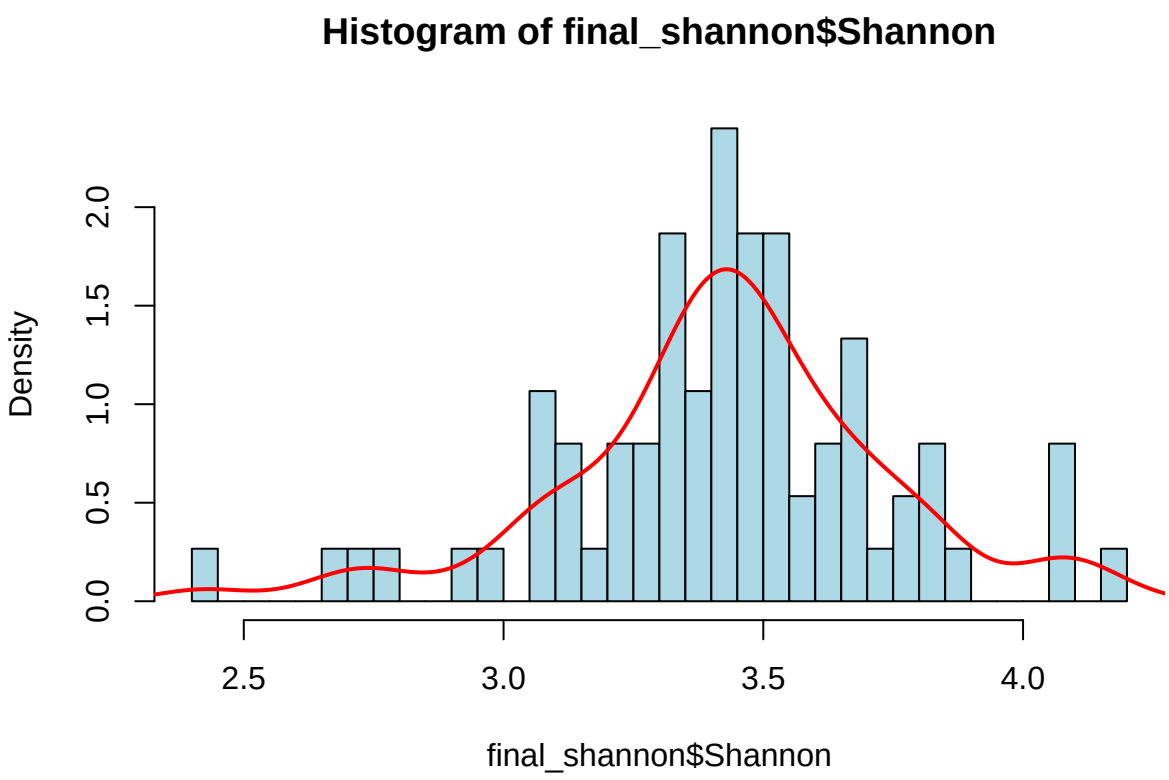
Preliminary to determining distribution type.

```
ggplot(final_shannon, aes(Shannon)) +
  geom_histogram(binwidth = 0.05, alpha = 0.7, colour="white")
```



Another way to histogram... looking normal..

```
hist(final_shannon$Shannon, probability = TRUE, col = "lightblue", border = "black", breaks = 25)
lines(density(final_shannon$Shannon), col = "red", lwd = 2)  # Adds a density curve
```



This is looks normal, however we need to determine distribution type in order to make the correct analysis and conclusions about our findings...

#Data wrangling ##Group Treatments into Communities

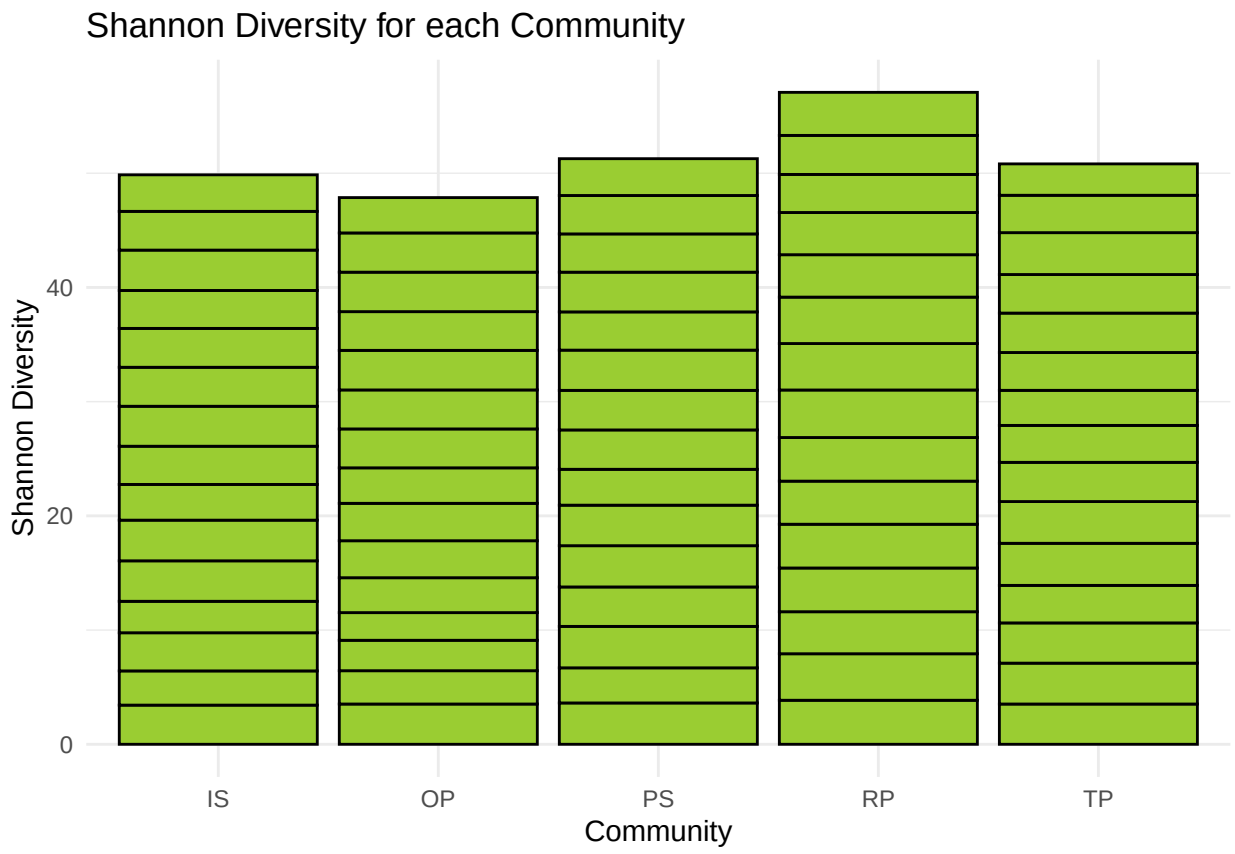
```
#Data is separated by into 75 treatments -> groups with 15 treatments each
#To draw better conclusions, we need to group by group
final_shannon$Community <- ifelse(startsWith(final_shannon$Treatment, "OP"), "OP",
                                  ifelse(startsWith(final_shannon$Treatment, "RP"), "RP",
                                         ifelse(startsWith(final_shannon$Treatment, "TP"), "TP"
                                                ifelse(startsWith(final_shannon$Treatment, "IS"
                                                       ifelse(startsWith(final_shannon$Treatmen
```

Rearranging columns so "Community" is beside "Treatment"

```
final_shannon <- final_shannon[, c("Treatment", "Community", "Shannon")]
```

Visualization of Shannon diversity for each community

```
ggplot(final_shannon, aes(x = Community, y = Shannon)) +
  geom_bar(stat = "identity", fill = "yellowgreen", color = "black") +
  theme_minimal() +
  labs(title = "Shannon Diversity for each Community", x = "Community",
       y = "Shannon Diversity")
```



Add all the predictor variable columns (soil and plant traits) onto final_shannon dataframe. . .

```
final_shannon$soil_C <- Soil_Traits$Soil_C
final_shannon$soil_N <- Soil_Traits$Soil_N
final_shannon$soil_P <- Soil_Traits$Soil_P
final_shannon$soil_pH <- Soil_Traits$Soil_pH
final_shannon$soil_sal <- Soil_Traits$Soil_Salinity
final_shannon$soil_wat <- Soil_Traits$Soil_Water

final_shannon$biomass_above <- Plant_Traits$Aboveground_Biomass
final_shannon$plant_dens <- Plant_Traits$Plant_Density
final_shannon$leaf_N <- Plant_Traits$Leaf_N
final_shannon$leaf_C <- Plant_Traits$Leaf_C
final_shannon$leaf_P <- Plant_Traits$Leaf_P
```

Making sure all the number columns are numeric!!

Note that this also converts 'n/a' cells into NULL – we will have to replace this later with an estimate!

```
final_shannon[, 3:14] <- lapply(final_shannon[, 3:14], as.numeric)
```

```
## Warning in lapply(final_shannon[, 3:14], as.numeric): NAs introduced by
## coercion
```

###Saving the final_shannon dataframe as a csv file for sharing and future edits...

```
write.csv(final_shannon, "final_shannon.csv")
```

##Filtering data for our final dataset OP14 and OP15 have NULL leaf_N values (due to equipment failure) -> here, we replace them with the mean of their community instead...

```
replaced_FS <- final_shannon
```

```
#First filter out the data to involve only the OP community and then
# exclude the 2 rows with "n/a" leaf_N values:

final_shannon %>%
  filter(final_shannon$Community == "OP",
         !is.na(final_shannon$leaf_N)) -> OPonly_noNA_FS
```

```
# Replacing the 2 cells with the mean of the OP'S leaf_N column
replaced_FS <- final_shannon

replaced_FS[replaced_FS$Treatment == "OP14", "leaf_N"] <-
  mean(OPonly_noNA_FS$leaf_N)
replaced_FS[replaced_FS$Treatment == "OP15", "leaf_N"] <-
  mean(OPonly_noNA_FS$leaf_N)
```

###Renaming master data frame to a shorter name bc I will have to use it over and over again in log-likelihood evalutors!

```
data <- replaced_FS
```

Making ranges of parameter combinations for each distribution

MLE: We test for normal, gamma, and inverse gaussian because they are the only distribution types that can reasonably fit our data and are generally accepted in a generalized linear mixed model!

```
norm_combos <- expand.grid(mean = seq(2.4, 4.2, 0.01),
                           sd = seq(0.15, 0.7, 0.01))

gamma_combos <- expand.grid(shape = seq(100, 150, 0.1),
                            scale = seq(0.02, 0.04, 0.0005))

invgaus_combos <- expand.grid(mu = seq(2, 6, 0.1),
                              lambda = seq(200, 600, 0.1))
```

#Log-Likelihood evalutions

```r
norm_loglik <- function(x = data$Shannon, mean, sd){
  return (sum(log(dnorm(x = x, mean = mean, sd = sd))))
}
```

```r
gamma_loglik <- function(x = data$Shannon, shape, scale){
  return (sum(log(dgamma(x = x, shape = shape, scale = scale))))
}
```

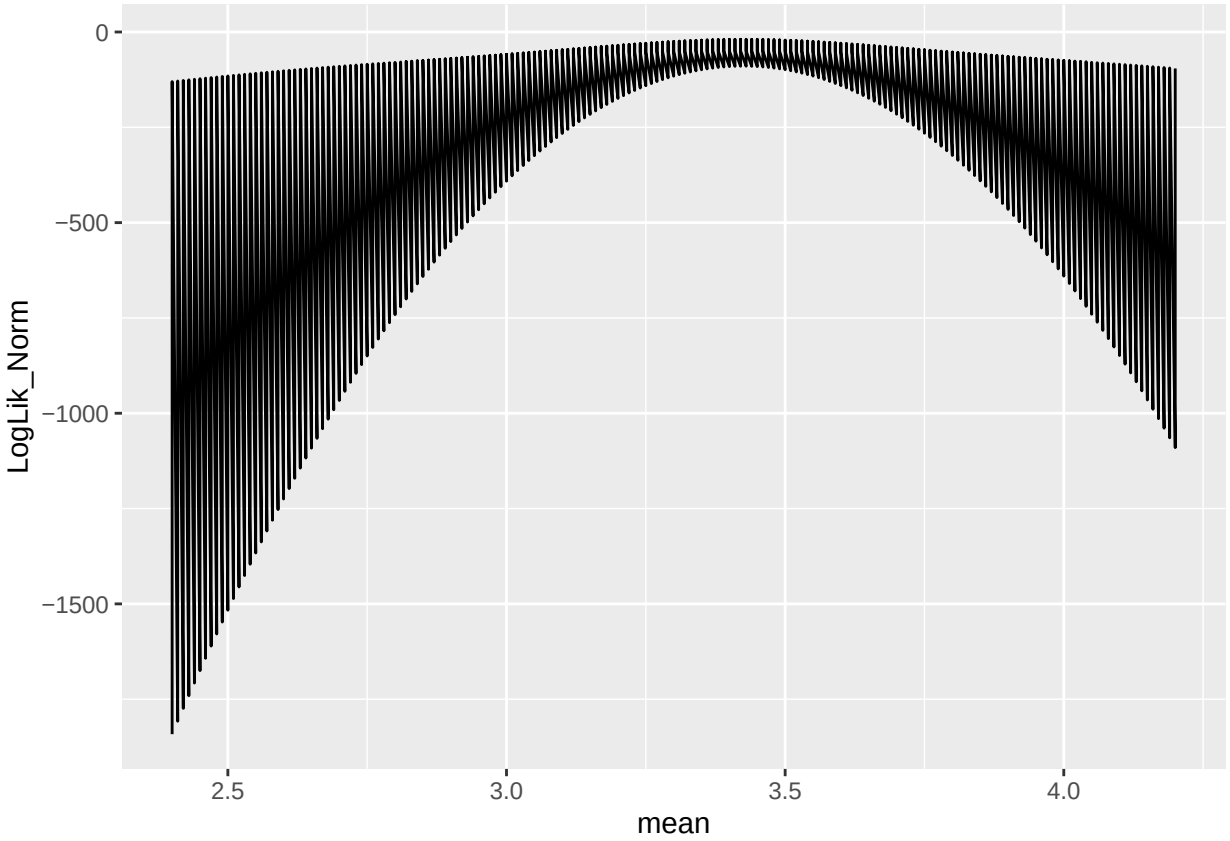
```r
invgaus_loglik <- function(x = data$Shannon, mu, lambda){
  return (sum(log(
    sqrt(lambda/(2*pi*x^3)) * exp((-lambda*(x-mu)^2)/(2*x*mu^2))
    )))
}
```

```r
LogLik_norm <- c() # empty log-likelihoods vector for NORMAL

for (i in 1:nrow(norm_combos)){
  LogLik_norm[i] <- norm_loglik(mean = norm_combos$mean[i],
                                sd = norm_combos$sd[i])
}

LLs_norm <- data.frame(LogLik_Norm = LogLik_norm,
                       mean = norm_combos$mean,
                       sd = norm_combos$sd)
```

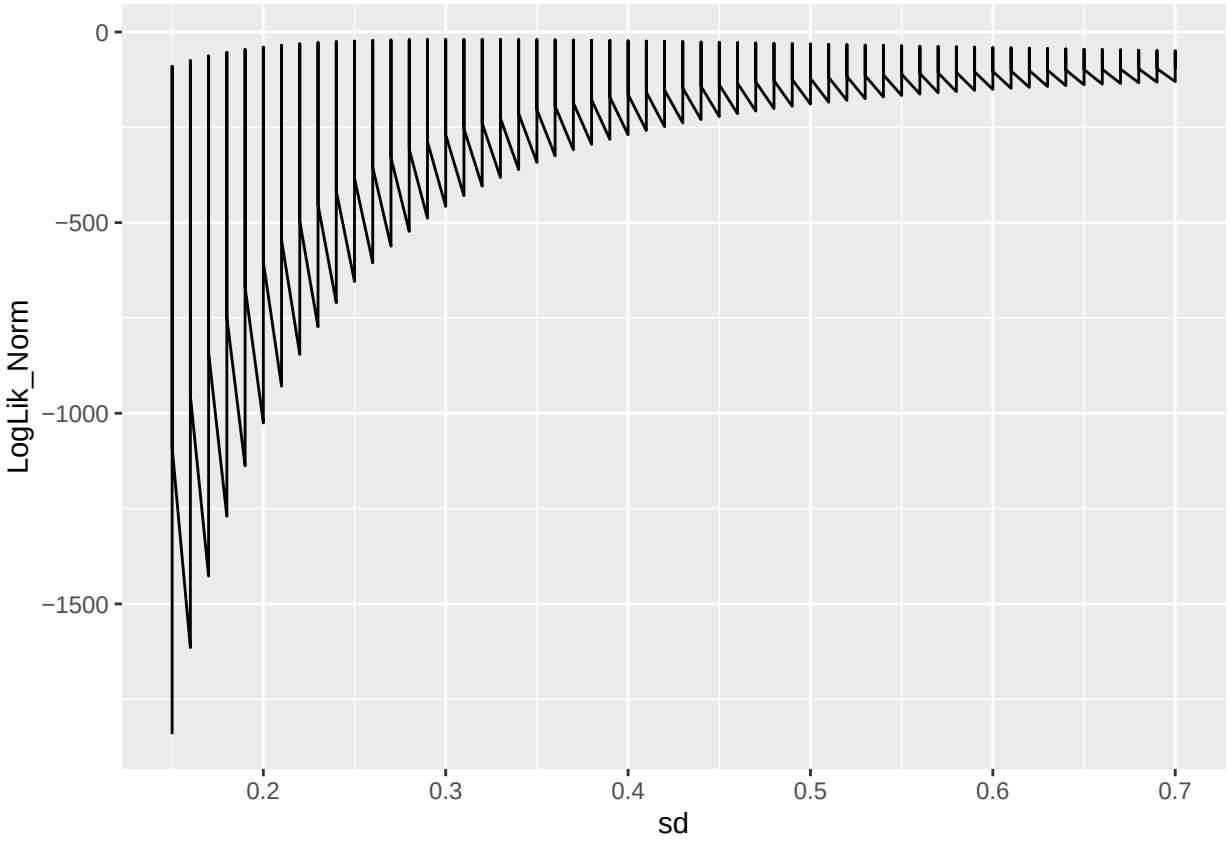
```r
LogLik_gamma <- c() # empty log-likelihoods vector for GAMMA

for (i in 1:nrow(gamma_combos)){
  LogLik_gamma[i] <- gamma_loglik(shape = gamma_combos$shape[i],
                                  scale = gamma_combos$scale[i])
}

LLs_gamma <- data.frame(LogLik_Gamma = LogLik_gamma,
                       shape = gamma_combos$shape,
                       scale = gamma_combos$scale)
```

```r
LogLik_invgaus <- c() # empty log-likelihoods vector for INVERSE GAUSSIAN

for (i in 1:nrow(invgaus_combos)){
  LogLik_invgaus[i] <- invgaus_loglik(mu = invgaus_combos$mu[i],
                                      lambda = invgaus_combos$lambda[i])
}

LLs_invgaus <- data.frame(LogLik_InvGaus = LogLik_invgaus,
                          mu = invgaus_combos$mu,
                          lambda = invgaus_combos$lambda)
```

```r
ggplot(LLs_norm, aes(x = mean, y = LogLik_Norm)) + geom_line()
```



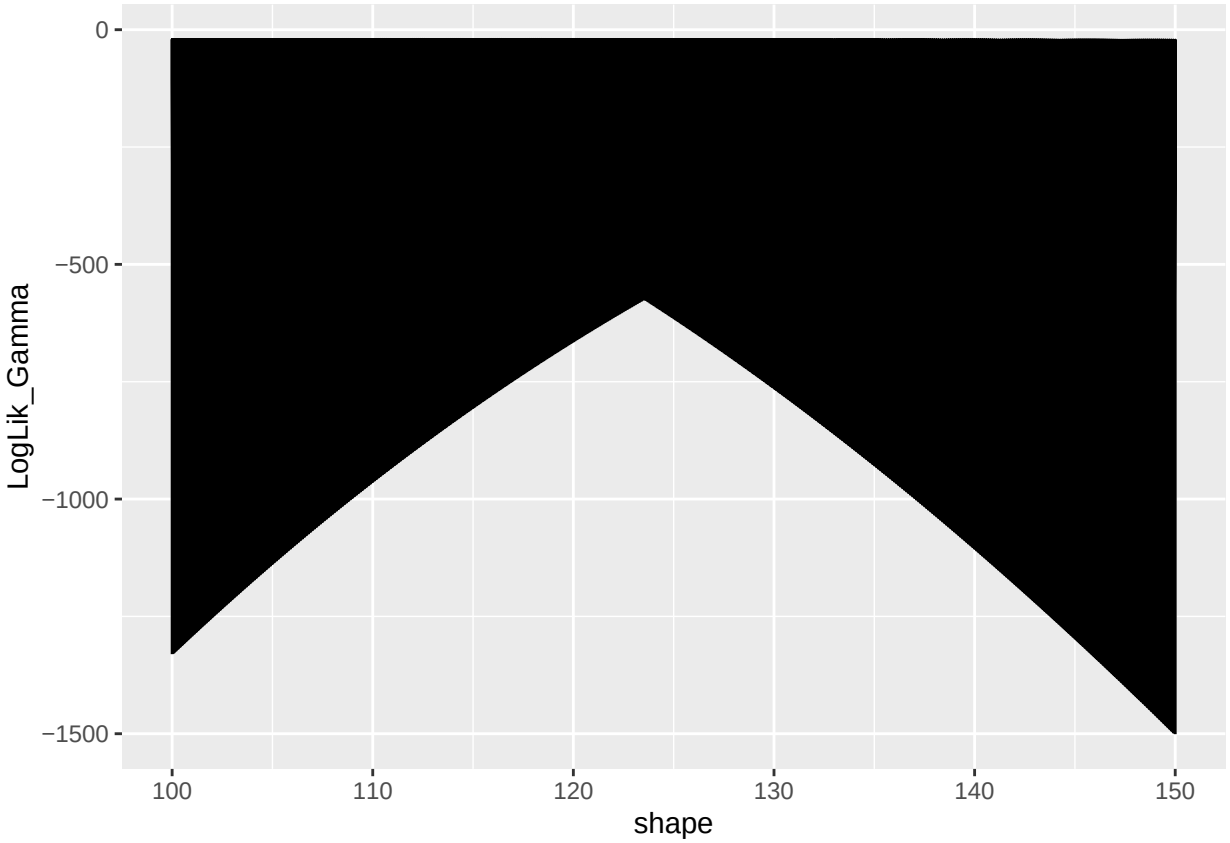

```
ggplot(LLs_norm, aes(x = sd, y = LogLik_Norm)) + geom_line()
```



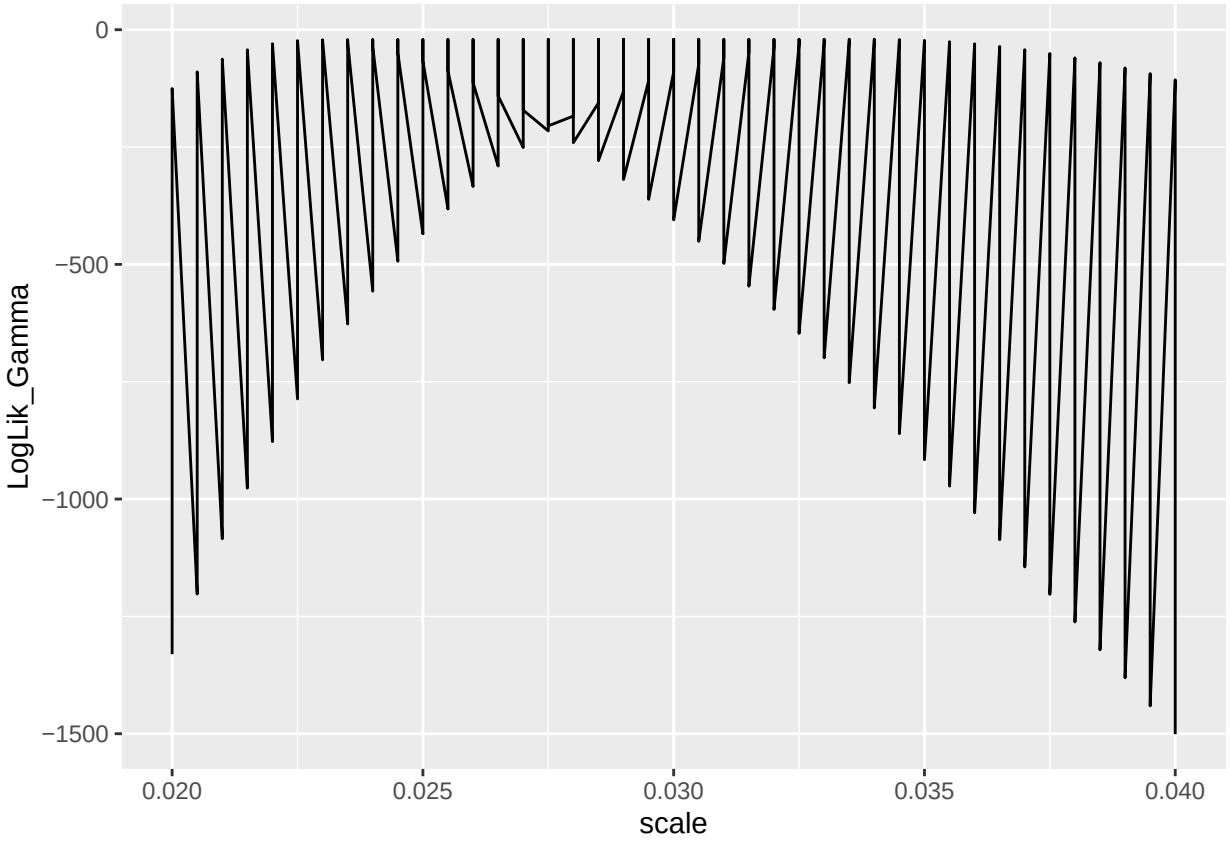

```
MLE_norm <- LLs_norm %>% subset(LogLik_Norm == max(LogLik_Norm))
MLE_norm
```

```
##      LogLik_Norm mean   sd
## 3000   -19.28128 3.43 0.31
```

```
ggplot(LLs_gamma, aes(x = shape, y = LogLik_Gamma)) + geom_line()
```



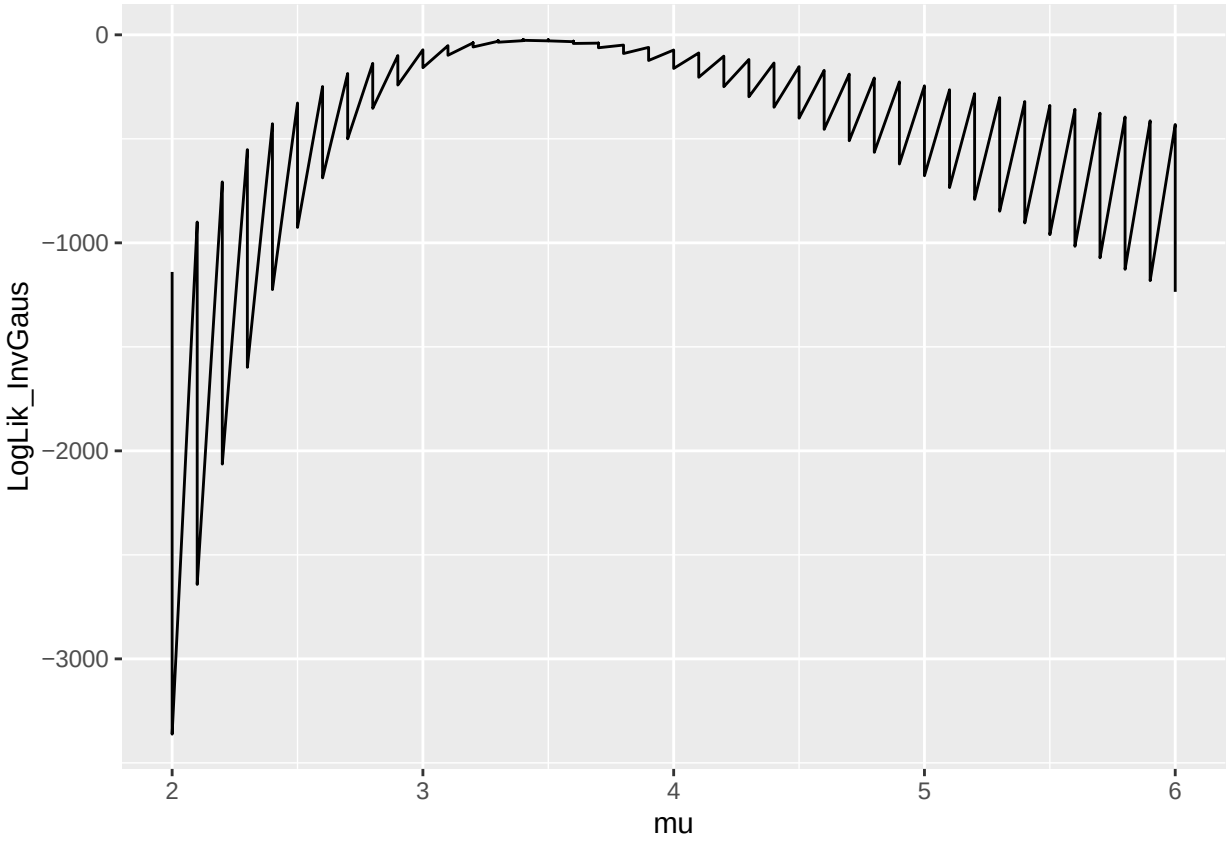

```
ggplot(LLs_gamma, aes(x = scale, y = LogLik_Gamma)) + geom_line()
```



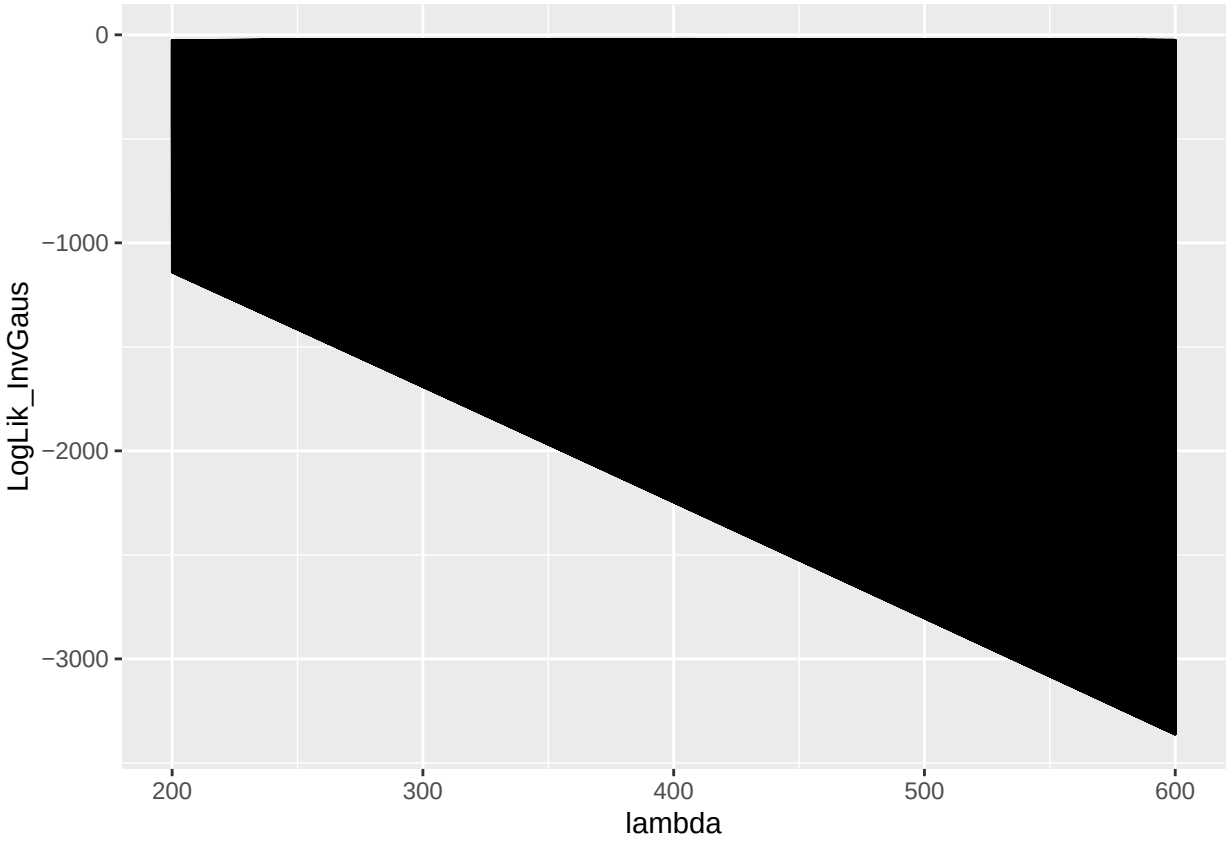

```
MLE_gamma <- LLs_gamma %>% subset(LogLik_Gamma == max(LogLik_Gamma))
MLE_gamma
```

```
##      LogLik_Gamma shape  scale
## 9681    -20.56567 116.1 0.0295
```

```
ggplot(LLs_invgaus, aes(x = mu, y = LogLik_InvGaus)) + geom_line()
```




```
ggplot(LLs_invgaus, aes(x = lambda, y = LogLik_InvGaus)) + geom_line()
```

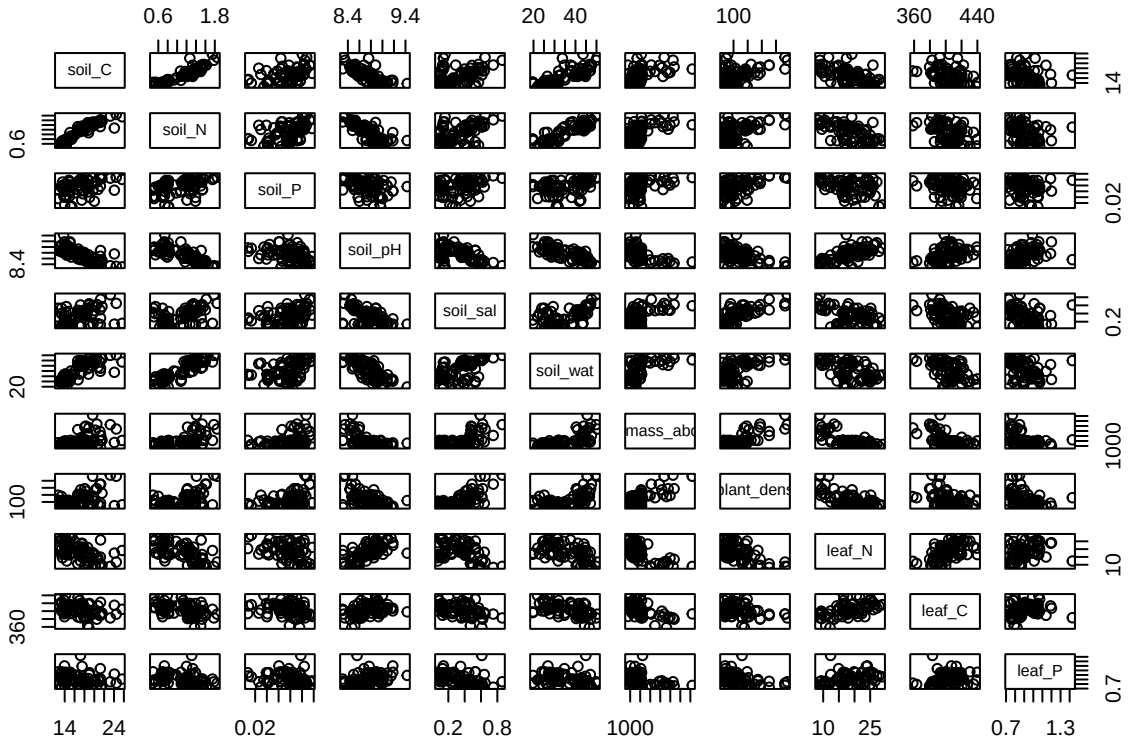



```
MLE_invgaus <- LLs_invgaus %>% subset(LogLik_InvGaus == max(LogLik_InvGaus))
MLE_invgaus
```

```
##       LogLik_InvGaus  mu lambda
## 73651      -21.72646 3.4  379.6
```

Find which varaiables are linearly and positively correlated

```
par(mfrow=c(1,1))

pairs(data[, c("soil_C", "soil_N", "soil_P", "soil_pH", "soil_sal", "soil_wat",
               "biomass_above", "plant_dens", "leaf_N", "leaf_C", "leaf_P")])
```



#Test for co-linearality

```
cor_matrix <- cor(data[, c("soil_C", "soil_N", "soil_P", "soil_pH", "soil_sal",
                           "soil_wat", "biomass_above", "plant_dens", "leaf_N",
                           "leaf_C", "leaf_P")],
                  use = "pairwise.complete.obs")
print(cor_matrix)
```

```
##                   soil_C     soil_N     soil_P    soil_pH   soil_sal   soil_wat
## soil_C         1.0000000  0.9269588  0.4079972 -0.7594546  0.3689419  0.7843768
## soil_N         0.9269588  1.0000000  0.4612699 -0.7713667  0.4504365  0.8769814
## soil_P         0.4079972  0.4612699  1.0000000 -0.3512759  0.3214240  0.3975461
## soil_pH       -0.7594546 -0.7713667 -0.3512759  1.0000000 -0.5082347 -0.7704241
## soil_sal       0.3689419  0.4504365  0.3214240 -0.5082347  1.0000000  0.5721562
## soil_wat       0.7843768  0.8769814  0.3975461 -0.7704241  0.5721562  1.0000000
## biomass_above  0.4958170  0.5719650  0.3556092 -0.5908822  0.5816223  0.5926635
## plant_dens     0.4596174  0.5237192  0.4443572 -0.5009611  0.7244117  0.5890475
## leaf_N        -0.5434564 -0.6102046 -0.4077548  0.6527521 -0.4116900 -0.5335999
## leaf_C        -0.3695966 -0.4709862 -0.2610081  0.3741878 -0.4108897 -0.5415672
## leaf_P        -0.2657054 -0.2912269 -0.1731799  0.3522477 -0.3349665 -0.2966488
##               biomass_above plant_dens     leaf_N     leaf_C     leaf_P
## soil_C            0.4958170  0.4596174 -0.5434564 -0.3695966 -0.2657054
## soil_N            0.5719650  0.5237192 -0.6102046 -0.4709862 -0.2912269
## soil_P            0.3556092  0.4443572 -0.4077548 -0.2610081 -0.1731799
## soil_pH          -0.5908822 -0.5009611  0.6527521  0.3741878  0.3522477
## soil_sal          0.5816223  0.7244117 -0.4116900 -0.4108897 -0.3349665
```

```
## soil_wat          0.5926635  0.5890475 -0.5335999 -0.5415672 -0.2966488
## biomass_above     1.0000000  0.7600965 -0.6732280 -0.5668411 -0.3750552
## plant_dens        0.7600965  1.0000000 -0.6076623 -0.3558787 -0.3951235
## leaf_N           -0.6732280 -0.6076623  1.0000000  0.5349996  0.3001950
## leaf_C           -0.5668411 -0.3558787  0.5349996  1.0000000  0.1715634
## leaf_P           -0.3750552 -0.3951235  0.3001950  0.1715634  1.0000000
```

#PCA

```
pca.out <- prcomp(data[, c("soil_C", "soil_N", "soil_P", "soil_pH", "soil_sal",
                           "soil_wat", "biomass_above", "plant_dens", "leaf_N",
                           "leaf_C", "leaf_P")], scale.=T)
```

```
pca.out
```

```
## Standard deviations (1, .., p=11):
##  [1] 2.4739400 1.0748490 0.9303355 0.8883592 0.8115063 0.7306543 0.5751151
##  [8] 0.4855427 0.4101910 0.3178748 0.2034850
##
## Rotation (n x k) = (11 x 11):
##                      PC1         PC2          PC3          PC4         PC5
## soil_C        -0.3270867  0.44789885  0.166535148 -0.008676304  0.03928796
## soil_N        -0.3547747  0.37618972  0.080336514 -0.012495465  0.03310391
## soil_P        -0.2202713  0.02402190 -0.102260684  0.868617978 -0.28746538
## soil_pH        0.3376054 -0.22697065 -0.188488951  0.129101849 -0.05900525
## soil_sal      -0.2806422 -0.39256180 -0.001680892  0.031298086  0.57534475
## soil_wat      -0.3539508  0.23179513  0.023637151 -0.106935165  0.22817143
## biomass_above -0.3274486 -0.30076393 -0.166191853 -0.113060140 -0.05552193
## plant_dens    -0.3154603 -0.40412756  0.019900979  0.196156607  0.24912045
## leaf_N         0.3142092  0.06465172  0.204688059  0.042696425  0.39640405
## leaf_C         0.2491950  0.08525667  0.592994804  0.361212419  0.30039910
## leaf_P         0.1817902  0.36430053 -0.705485220  0.180643295  0.46436986
##                      PC6          PC7         PC8           PC9        PC10
## soil_C        -0.01522495  0.233006840  0.19895056  0.5003731605 -0.30717955
## soil_N         0.07140457  0.200468481  0.19590211  0.1164541941  0.32749935
## soil_P         0.24960073 -0.095511191 -0.17363341 -0.0004085302  0.01494356
## soil_pH        0.22688892  0.518724636  0.51089232  0.1140463425  0.38564887
## soil_sal       0.24372175 -0.422126461  0.12036129  0.3616190353  0.22135118
## soil_wat       0.25065950  0.094120337 -0.02413831 -0.6709013573  0.29839233
## biomass_above -0.26882044  0.445515232 -0.58930686  0.2239258302  0.28764553
## plant_dens    -0.23672949  0.340168319  0.29150445 -0.2968152659 -0.51071165
## leaf_N         0.54339602  0.353140394 -0.42365379  0.0539462712 -0.26428002
## leaf_C        -0.50049505  0.036650528 -0.03818856 -0.0315869034  0.31180818
## leaf_P        -0.29082673 -0.004844742 -0.05112435  0.0056049776 -0.02770573
##                      PC11
## soil_C         0.47350843
## soil_N        -0.72140762
## soil_P         0.03895873
## soil_pH        0.18850329
## soil_sal       0.05734670
## soil_wat       0.38146521
## biomass_above  0.08899378
## plant_dens    -0.17485820
```

```
## leaf_N        -0.15495914
## leaf_C         0.08220256
## leaf_P        -0.01550201
```

```
summary(pca.out)
```

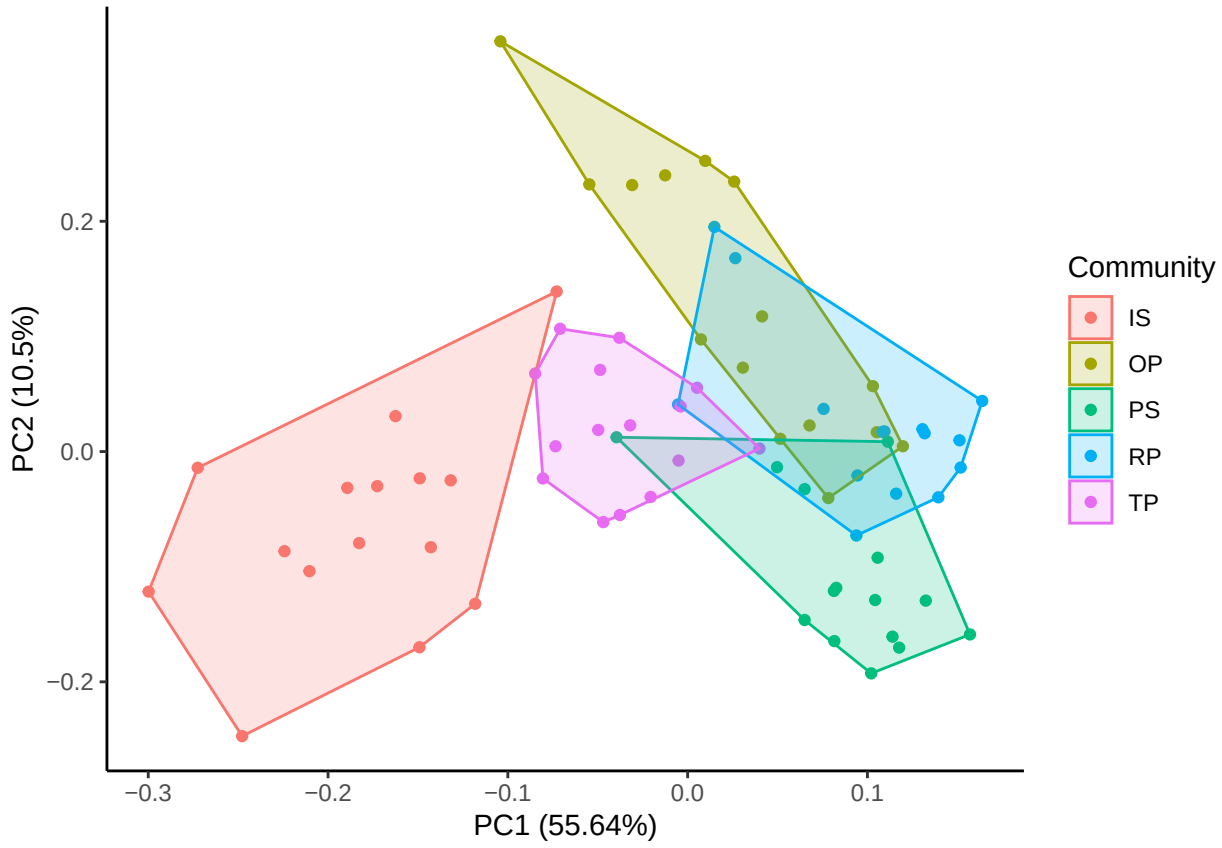
```
## Importance of components:
##                           PC1    PC2     PC3     PC4     PC5     PC6     PC7
## Standard deviation     2.4739 1.0748 0.93034 0.88836 0.81151 0.73065 0.57512
## Proportion of Variance 0.5564 0.1050 0.07868 0.07174 0.05987 0.04853 0.03007
## Cumulative Proportion  0.5564 0.6614 0.74011 0.81185 0.87172 0.92025 0.95032
##                           PC8    PC9    PC10    PC11
## Standard deviation     0.48554 0.4102 0.31787 0.20349
## Proportion of Variance 0.02143 0.0153 0.00919 0.00376
## Cumulative Proportion  0.97175 0.9870 0.99624 1.00000
```

```
head(pca.out$rotation)
```

```
##                 PC1        PC2          PC3          PC4         PC5
## soil_C   -0.3270867  0.4478988  0.166535148 -0.008676304  0.03928796
## soil_N   -0.3547747  0.3761897  0.080336514 -0.012495465  0.03310391
## soil_P   -0.2202713  0.0240219 -0.102260684  0.868617978 -0.28746538
## soil_pH   0.3376054 -0.2269706 -0.188488951  0.129101849 -0.05900525
## soil_sal -0.2806422 -0.3925618 -0.001680892  0.031298086  0.57534475
## soil_wat -0.3539508  0.2317951  0.023637151 -0.106935165  0.22817143
##                 PC6         PC7         PC8           PC9        PC10
## soil_C   -0.01522495  0.23300684  0.19895056  0.5003731605 -0.30717955
## soil_N    0.07140457  0.20046848  0.19590211  0.1164541941  0.32749935
## soil_P    0.24960073 -0.09551119 -0.17363341 -0.0004085302  0.01494356
## soil_pH   0.22688892  0.51872464  0.51089232  0.1140463425  0.38564887
## soil_sal  0.24372175 -0.42212646  0.12036129  0.3616190353  0.22135118
## soil_wat  0.25065950  0.09412034 -0.02413831 -0.6709013573  0.29839233
##                 PC11
## soil_C    0.47350843
## soil_N   -0.72140762
## soil_P    0.03895873
## soil_pH   0.18850329
## soil_sal  0.05734670
## soil_wat  0.38146521
```

Visualized PCA

```
autoplot(pca.out, x=1,y=2,
         data=data,
         colour="Community",
         frame=TRUE) +
  theme_classic()
```



#Things to do after PCA: regression on PC1, PC2, and interaction!! Discriminate!...