NMDS Analysis Tutorial

Introduction

This tutorial demonstrates how to perform a Non-Metric Dimensional Scaling (NMDS) analysis using R. We will use a sample dataset for better understanding.

Setup

First, let's load the necessary packages.

```
library(vegan)
```

```
## Loading required package: permute
```

```
## Loading required package: lattice
```

```
## This is vegan 2.6-4
```

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----------------------- tidyverse 2.0.0 --
## v dplyr     1.1.2     v readr     2.1.4
## v forcats   1.0.0     v stringr   1.5.0
## v ggplot2   3.4.2     v tibble    3.2.1
## v lubridate 1.9.2     v tidyr     1.3.0
## v purrr     1.0.1
```

```
## -- Conflicts --------------------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(RColorBrewer)
```

Sample Data Creation

To keep things simple, we are simulating a sample dataset for the tutorial.

```r
set.seed(123)  # Setting a seed for reproducibility

bsll <- data.frame(
  SampleNo = 1:20,
  Animal = sample(c("A", "B", "C", "D", "E"), 20, replace = TRUE),
  PhaseCategory = sample(c("Baseline", "Probiotics"), 20, replace = TRUE),
  Sp1 = rnorm(20, 10, 5),
  Sp2 = rnorm(20, 15, 4),
  Sp3 = rnorm(20, 12, 3)
)

bsll_env <- bsll %>% select(SampleNo, Animal, PhaseCategory)
bsll_sp <- bsll %>% select(-SampleNo, -Animal, -PhaseCategory)
```

Data Preparation

Next, we prepare the data by converting necessary columns to factors and generating a color palette.

```r
bsll_env$Animal_factor <- as.factor(bsll_env$Animal)
num_sites <- length(unique(bsll_env$Animal_factor))
color_palette <- colorRampPalette(brewer.pal(9, "Set1"))(num_sites)

bsll_env$phase_factor <- as.factor(bsll_env$PhaseCategory)
shape_map <- c(Baseline = 19, Probiotics = 16)  # Circle for Baseline, Square for Probiotics
shape_vec <- shape_map[bsll_env$PhaseCategory]
```

NMDS Analysis

With the data ready, we can now compute the NMDS analysis.

```r
NMDS_result <- metaMDS(bsll_sp, distance = "bray", k = 2)
```

```
## Wisconsin double standardization
## Run 0 stress 0.0260713
## Run 1 stress 0.0260713
## ... Procrustes: rmse 7.589554e-06  max resid 1.99513e-05
## ... Similar to previous best
## Run 2 stress 0.0260713
## ... Procrustes: rmse 3.167506e-06  max resid 1.065786e-05
## ... Similar to previous best
## Run 3 stress 0.0260713
## ... New best solution
## ... Procrustes: rmse 8.223805e-06  max resid 2.467044e-05
## ... Similar to previous best
## Run 4 stress 0.0260713
## ... Procrustes: rmse 9.675616e-06  max resid 2.980061e-05
## ... Similar to previous best
## Run 5 stress 0.0260713
## ... Procrustes: rmse 7.321473e-06  max resid 2.003673e-05
## ... Similar to previous best
## Run 6 stress 0.0260713
```

```
## ... Procrustes: rmse 1.230223e-05   max resid 3.803866e-05
## ... Similar to previous best
## Run 7 stress 0.0260713
## ... Procrustes: rmse 1.878322e-05   max resid 5.841322e-05
## ... Similar to previous best
## Run 8 stress 0.0260713
## ... New best solution
## ... Procrustes: rmse 4.545644e-06   max resid 1.418457e-05
## ... Similar to previous best
## Run 9 stress 0.1648943
## Run 10 stress 0.0260713
## ... Procrustes: rmse 1.583791e-06   max resid 4.702155e-06
## ... Similar to previous best
## Run 11 stress 0.1711949
## Run 12 stress 0.0260713
## ... Procrustes: rmse 4.823734e-06   max resid 1.471969e-05
## ... Similar to previous best
## Run 13 stress 0.2061401
## Run 14 stress 0.0260713
## ... Procrustes: rmse 3.749013e-06   max resid 1.142072e-05
## ... Similar to previous best
## Run 15 stress 0.0260713
## ... Procrustes: rmse 4.360414e-06   max resid 1.334013e-05
## ... Similar to previous best
## Run 16 stress 0.0260713
## ... Procrustes: rmse 6.25119e-06   max resid 1.942665e-05
## ... Similar to previous best
## Run 17 stress 0.0260713
## ... Procrustes: rmse 8.613683e-06   max resid 2.668794e-05
## ... Similar to previous best
## Run 18 stress 0.0260713
## ... Procrustes: rmse 5.264148e-06   max resid 1.533326e-05
## ... Similar to previous best
## Run 19 stress 0.1946148
## Run 20 stress 0.0260713
## ... Procrustes: rmse 6.099683e-06   max resid 1.862275e-05
## ... Similar to previous best
## *** Best solution repeated 9 times
```

Plotting

Finally, let's visualize the NMDS results.

```r
par(mar = c(5.1, 4.1, 4.1, 6))
plot(NMDS_result$points,
     col=color_palette[as.integer(bsll_env$Animal_factor)],
     pch=shape_vec,
     cex=1.2,
     main=paste("NMDS Bacterial Species Analysis\nStress:", round(NMDS_result$stress, 3)),
     xlab= "NMDS Axis 1",
     ylab= "NMDS Axis 2")
ordispider(NMDS_result, groups=bsll_env$PhaseCategory, label = TRUE)
legend('topright',
```

```
    legend=levels(bsll_env$Animal_factor),
    fill=color_palette,
    title="Animal",
    cex=1,
    xpd = TRUE,
    inset = c(-0.10, 0))
```

NMDS Bacterial Species Analysis
Stress: 0.026