

# Testing different strategies to compute a representative cluster profile

Code ▾

Hide

```
library(tidyverse)
library(vegan)
```

## C1 cluster:

Hide

```
true_values <- c(-1.0102, -0.9091, -0.8081, 0.8081, 0.9091, 1.0102)
names(true_values) <- c("t1", "t2", "t3", "t4", "t5", "t6")
df <- read.table("clust_c1.tsv", header = TRUE, row.names = 1)
df_scaled <- t(scale(t(df), center = TRUE, scale = TRUE))
pca <- rda(df_scaled)
eigengenes <- scores(pca)$species[,"PC1"]
means <- apply(df_scaled, 2, mean)
trimmed_means <- apply(df_scaled, 2, mean, trim = 0.25)
medians <- apply(df_scaled, 2, median)
sum(abs(eigengenes - true_values))
```

```
[1] 5.869941
```

Hide

```
sum(abs(means - true_values))
```

```
[1] 0.7849405
```

Hide

```
sum(abs(trimmed_means - true_values))
```

```
[1] 0.6025752
```

Hide

```
sum(abs(medians - true_values))
```

```
[1] 0.4983377
```

## C2 cluster:

Hide

```
true_values <- c(1.0102, 0.9091, 0.8081, -0.8081, -0.9091, -1.0102)
names(true_values) <- c("t1", "t2", "t3", "t4", "t5", "t6")
df <- read.table("clust_c2.tsv", header = TRUE, row.names = 1)
df_scaled <- t(scale(t(df), center = TRUE, scale = TRUE))
pca <- rda(df_scaled)
eigengenes <- scores(pca)$species[, "PC1"]
means <- apply(df_scaled, 2, mean)
trimmed_means <- apply(df_scaled, 2, mean, trim = 0.25)
medians <- apply(df_scaled, 2, median)
sum(abs(eigengenes - true_values))
```

```
[1] 5.346998
```

[Hide](#)

```
sum(abs(means - true_values))
```

```
[1] 0.7891589
```

[Hide](#)

```
sum(abs(trimmed_means - true_values))
```

```
[1] 0.649756
```

[Hide](#)

```
sum(abs(medians - true_values))
```

```
[1] 0.5606012
```

## C3 cluster:

[Hide](#)

```
true_values <- c(-0.9895, -0.2234, 1.2129, 1.2129, -0.2234, -0.9895)
names(true_values) <- c("t1", "t2", "t3", "t4", "t5", "t6")
df <- read.table("clust_c3.tsv", header = TRUE, row.names = 1)
df_scaled <- t(scale(t(df), center = TRUE, scale = TRUE))
pca <- rda(df_scaled)
eigengenes <- scores(pca)$species[, "PC1"]
means <- apply(df_scaled, 2, mean)
trimmed_means <- apply(df_scaled, 2, mean, trim = 0.25)
medians <- apply(df_scaled, 2, median)
sum(abs(eigengenes - true_values))
```

```
[1] 6.209301
```

[Hide](#)

```
sum(abs(means - true_values))
```

```
[1] 0.7098361
```

[Hide](#)

```
sum(abs(trimmed_means - true_values))
```

```
[1] 0.5428989
```

[Hide](#)

```
sum(abs(medians - true_values))
```

```
[1] 0.4522004
```