HW4

Getong Zhong

2023-03-02

4.1

the derived solution shows that the objects arranged in a linear fashion with the objects in the middle being closer to each other and the objects towards the ends being farther apart. This is because the similarity between two objects decreases as the absolute difference between their coordinates (i and j) increases, with a cutoff point of 25. Therefore, the objects towards the middle higher similarities and will be placed closer together in the derived solution, while the objects towards the ends have lower similarities and will be placed farther apart.

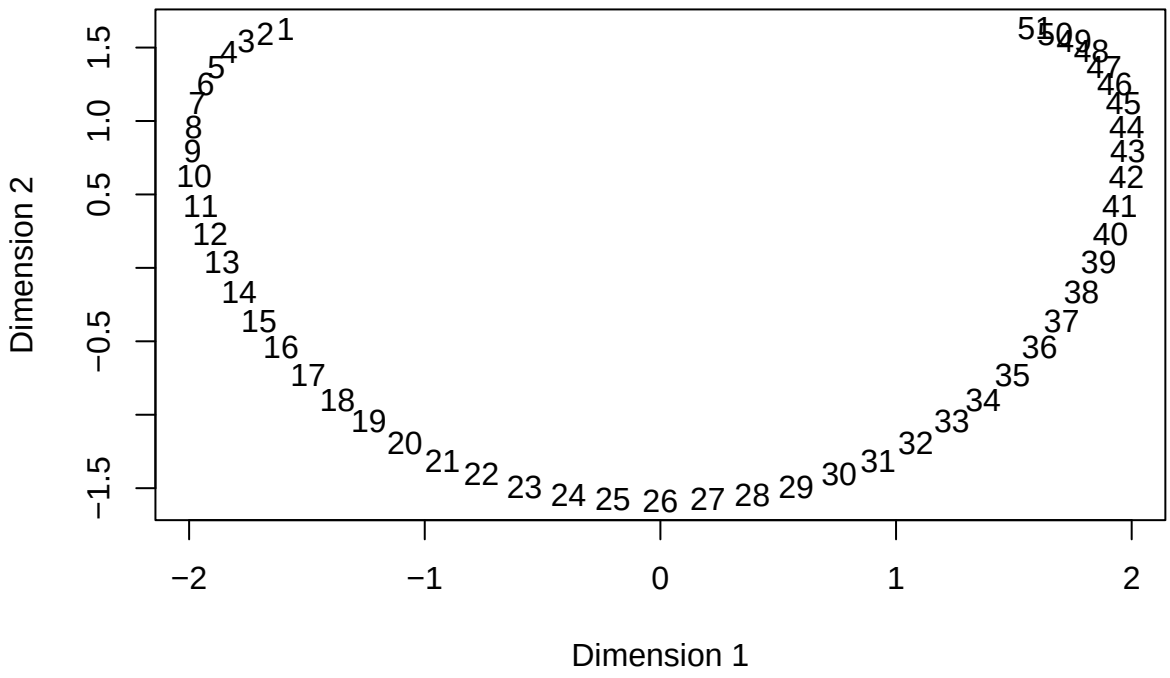
```
s <- matrix(0, nrow = 51, ncol = 51)
for (i in 1:51) {
  for (j in 1:51) {
    if (i == j) {
      s[i,j] <- 9
    } else if (abs(i - j) <= 3) {
      s[i,j] <- 8
    } else if (abs(i - j) <= 6) {
      s[i,j] <- 7
    } else if (abs(i - j) <= 9) {
      s[i,j] <- 6
    } else if (abs(i - j) <= 12) {
      s[i,j] <- 5
    } else if (abs(i - j) <= 15) {
      s[i,j] <- 4
    } else if (abs(i - j) <= 18) {
      s[i,j] <- 3
    } else if (abs(i - j) <= 21) {
      s[i,j] <- 2
    } else if (abs(i - j) <= 24) {
      s[i,j] <- 1
    } else {
      s[i,j] <- 0
    }
  }
}
d <- matrix(0, nrow = 51, ncol = 51)
for (i in 1:51) {
  for (j in 1:51) {
    d[i,j] <- sqrt(s[i,i] + s[j,j] - 2*s[i,j])
  }
}
```

```r
(cmd <- cmdscale(d, k = 2, eig = TRUE))
```

```
## $points
##                [,1]        [,2]
##  [1,] -1.590378e+00  1.63083709
##  [2,] -1.676360e+00  1.59362762
##  [3,] -1.756372e+00  1.54102840
##  [4,] -1.830095e+00  1.47447273
##  [5,] -1.883695e+00  1.36809039
##  [6,] -1.928818e+00  1.25006627
##  [7,] -1.965399e+00  1.12228013
##  [8,] -1.979875e+00  0.96003453
##  [9,] -1.984189e+00  0.79264682
## [10,] -1.978583e+00  0.62219863
## [11,] -1.949812e+00  0.42489157
## [12,] -1.910156e+00  0.23085143
## [13,] -1.860199e+00  0.04201664
## [14,] -1.787037e+00 -0.16487224
## [15,] -1.703305e+00 -0.35927581
## [16,] -1.609936e+00 -0.53979774
## [17,] -1.494368e+00 -0.72955527
## [18,] -1.369582e+00 -0.89800870
## [19,] -1.236845e+00 -1.04470273
## [20,] -1.083914e+00 -1.19299093
## [21,] -9.240808e-01 -1.31271054
## [22,] -7.589079e-01 -1.40469216
## [23,] -5.764206e-01 -1.49283394
## [24,] -3.901724e-01 -1.54764118
## [25,] -2.019561e-01 -1.57147476
## [26,] -7.791538e-16 -1.58897250
## [27,]  2.019561e-01 -1.57147476
## [28,]  3.901724e-01 -1.54764118
## [29,]  5.764206e-01 -1.49283394
## [30,]  7.589079e-01 -1.40469216
## [31,]  9.240808e-01 -1.31271054
## [32,]  1.083914e+00 -1.19299093
## [33,]  1.236845e+00 -1.04470273
## [34,]  1.369582e+00 -0.89800870
## [35,]  1.494368e+00 -0.72955527
## [36,]  1.609936e+00 -0.53979774
## [37,]  1.703305e+00 -0.35927581
## [38,]  1.787037e+00 -0.16487224
## [39,]  1.860199e+00  0.04201664
## [40,]  1.910156e+00  0.23085143
## [41,]  1.949812e+00  0.42489157
## [42,]  1.978583e+00  0.62219863
## [43,]  1.984189e+00  0.79264682
## [44,]  1.979875e+00  0.96003453
## [45,]  1.965399e+00  1.12228013
## [46,]  1.928818e+00  1.25006627
## [47,]  1.883695e+00  1.36809039
## [48,]  1.830095e+00  1.47447273
## [49,]  1.756372e+00  1.54102840
```

```
## [50,]  1.676360e+00  1.59362762
## [51,]  1.590378e+00  1.63083709
##
## $eig
##  [1]  1.260857e+02  6.593531e+01  1.817103e+01  7.820841e+00  7.610276e+00
##  [6]  7.378184e+00  7.018378e+00  5.283354e+00  3.436749e+00  3.096003e+00
## [11]  3.089512e+00  2.697496e+00  1.939351e+00  1.457605e+00  1.427548e+00
## [16]  1.401274e+00  1.341983e+00  1.162903e+00  9.675241e-01  9.495632e-01
## [21]  9.380593e-01  8.766031e-01  7.925229e-01  7.664407e-01  7.041554e-01
## [26]  6.921047e-01  5.994198e-01  5.448786e-01  4.750354e-01  4.738073e-01
## [31]  4.017636e-01  3.996796e-01  3.607046e-01  3.046375e-01  2.892578e-01
## [36]  2.756293e-01  2.664975e-01  2.588366e-01  2.539959e-01  2.509620e-01
## [41]  1.209368e-01  9.981512e-02  6.217249e-15 -9.449958e-02 -1.060236e-01
## [46] -3.067296e-01 -3.298370e-01 -6.738433e-01 -7.105516e-01 -2.002901e+00
## [51] -2.074323e+00
##
## $x
## NULL
##
## $ac
## [1] 0
##
## $GOF
## [1] 0.6744323 0.6896902
```

```
plot(cmd$points, type="n", xlab="Dimension 1", ylab="Dimension 2")
text(cmd$points[,1], cmd$points[,2], labels=1:51)
```



4.2

```
chi_dist <- function(x) {
  row <- rowSums(x)
  col <- colSums(x)
  total <- sum(x)
  exp_vals <- outer(row, col, '*')/total
  row_chi <- apply(x, 1, function(row) sum((row-exp_vals[row!=0,col!=0])^2/exp_vals[row!=0,col!=0]))
  col_chi <- apply(x, 2, function(col) sum((col-exp_vals[row!=0,col!=0])^2/exp_vals[row!=0,col!=0]))
  list(row_chi= as.matrix(row_chi), col_chi = as.matrix(col_chi))
}
```

4.3

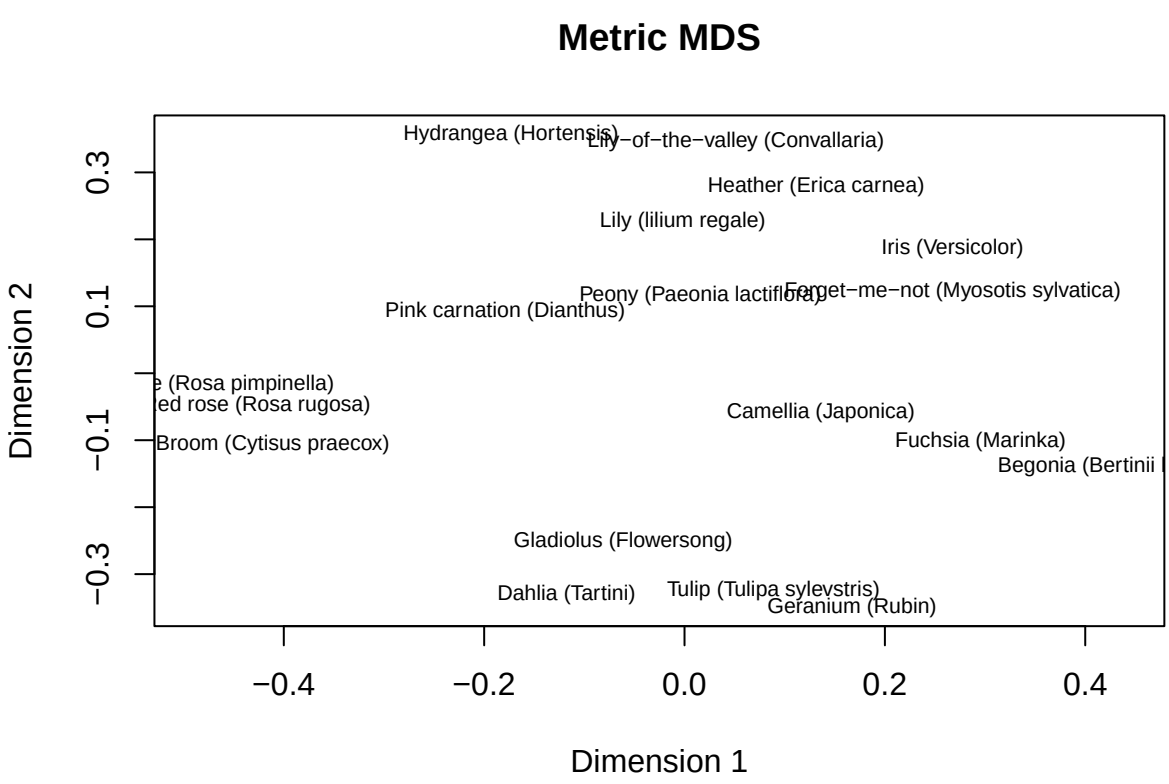
```
library(HSAUR3)
```

```
## Loading required package: tools
```

```
dat <- as.matrix(gardenflowers)
(cmd <- cmdscale(dat, k = 2, eig = TRUE))
```

```
## $points
##                                          [,1]        [,2]
## Begonia (Bertinii bolivienis)       0.441626810 -0.13910199
## Broom (Cytisus praecox)            -0.411002097 -0.10646858
## Camellia (Japonica)                 0.135998713 -0.05878727
## Dahlia (Tartini)                   -0.117861814 -0.32989431
## Forget-me-not (Myosotis sylvatica)  0.267792041  0.12142878
## Fuchsia (Marinka)                   0.295445879 -0.10140928
## Geranium (Rubin)                    0.167256180 -0.34946203
## Gladiolus (Flowersong)             -0.061150521 -0.25197021
## Heather (Erica carnea)              0.131374192  0.27948118
## Hydrangea (Hortensis)              -0.173171040  0.35681051
## Iris (Versicolor)                   0.267620864  0.18687636
## Lily (lilium regale)               -0.001753092  0.22650611
## Lily-of-the-valley (Convallaria)    0.051226880  0.34595178
## Peony (Paeonia lactiflora)          0.015761013  0.11587345
## Pink carnation (Dianthus)          -0.178928108  0.09234363
## Red rose (Rosa rugosa)             -0.427237687 -0.04816267
## Scotch rose (Rosa pimpinella)      -0.491784734 -0.01604355
## Tulip (Tulipa sylevstris)           0.088786521 -0.32397190
##
## $eig
##  [1]  1.173085e+00  8.944353e-01  5.732009e-01  4.843006e-01  2.638516e-01
##  [6]  2.298165e-01  8.383417e-02  6.645861e-02  2.984017e-02 -1.387779e-16
## [11] -2.147959e-02 -4.094094e-02 -4.366468e-02 -8.867401e-02 -1.045334e-01
## [16] -1.275889e-01 -1.714295e-01 -2.045124e-01
##
## $x
## NULL
##
## $ac
## [1] 0
##
## $GOF
## [1] 0.4493002 0.5442529
```

```
plot(cmd$points[,1], cmd$points[,2], xlab = "Dimension 1", ylab = "Dimension 2", main = "Metric MDS", ty
text(cmd$points[,1], cmd$points[,2], labels = row.names(dat), cex = 0.7)
```


Metric MDS
Hydrangea (Hortensis)
Lily−of−the−valley (Convallaria)
Heather (Erica carnea)
Lily (lilium regale)
Iris (Versicolor)
Peony (Paeonia lactiflora)
Forget−me−not (Myosotis sylvatica)
Pink carnation (Dianthus)
e (Rosa pimpinella)
ed rose (Rosa rugosa)
Camellia (Japonica)
Broom (Cytisus praecox)
Fuchsia (Marinka)
Begonia (Bertinii
Gladiolus (Flowersong)
Dahlia (Tartini)
Tulip (Tulipa sylevstris)
Geranium (Rubin)
Dimension 2
0.3
0.1
−0.1
−0.3
−0.4
−0.2
0.0
0.2
0.4
Dimension 1