Comprehension Check: Distance

Q1

1/1 point (graded)

Load the following dataset:

```
library(dslabs)
data("tissue_gene_expression")
```

This dataset includes a matrix `x`:

```
dim(tissue_gene_expression$x)
```

This matrix has the gene expression levels of 500 genes from 189 biological samples representing seven different tissues. The tissue type is stored in `y`:

```
table(tissue_gene_expression$y)
```

Which of the following lines of code computes the Euclidean distance between each observation and stores it in the object `d`?

- `d <- dist(tissue_gene_expression$x, distance='maximum')`
- `d <- dist(tissue_gene_expression)`
- **`d <- dist(tissue_gene_expression$x)`** ✔
- `d <- cor(tissue_gene_expression$x)`

Submit

You have used 1 of 2 attempts

Answers are displayed within the problem

Q2

1/1 point (graded)

Using the dataset from Q1, compare the distances between observations 1 and 2 (both cerebellum), observations 39 and 40 (both colon), and observations 73 and 74 (both endometrium).

Distance-wise, are samples from tissues of the same type closer to each other?

- No, the samples from the same tissue type are not necessarily closer.
- The two colon samples are closest to each other, but the samples from the other two tissues are not.
- The two cerebellum samples are closest to each other, but the samples from the other two tissues are not.
- **Yes, the samples from the same tissue type are closest to each other.** ✔

Explanation

You can calculate the distances using the following code:

```
ind <- c(1, 2, 39, 40, 73, 74)
as.matrix(d)[ind,ind]
```

Submit You have used 2 of 2 attempts

Answers are displayed within the problem

Q3

1/1 point (graded)

Make a plot of all the distances using the `image` function to see if the pattern you observed in Q2 is general.

Which code would correctly make the desired plot?

- `image(d)`
- **`image(as.matrix(d))`** ✔
- `d`
- `image()`

Explanation
When we examine the plot, we do see that the pattern holds and that samples from the same tissue are closest to each other, although there do appear to be some additional close distances between tissue types as well.

Submit You have used 1 of 2 attempts

Answers are displayed within the problem

Ask your questions or make your comments about Distance here! **Remember, one of the best ways to reinforce your own learning is by explaining something to someone else, so we encourage you to answer each other's questions (without giving away the answers, of course).**

Some reminders:

- Search the discussion board before posting to see if someone else has asked the same thing before asking a new question.
- Please be specific in the title and body of your post regarding which question you're asking about to facilitate answering your question.
- Posting snippets of code is okay, but posting full code solutions is not.
- If you do post snippets of code, please format it as code for readability. If you're not sure how to do this, there are instructions in a pinned post in the "general" discussion forum.

Discussion: Distance

Show Discussion

Topic: Section 4: Distance, Knn, Cross-Validation, and Generative Models / 4.1.1: Distance