

Dimension Reduction Principal Components Analysis

Dillon Labonte

February 19, 2025

Table of contents

```
library(tidyverse)
```

```
-- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
v dplyr      1.1.4      v readr      2.1.5
v forcats    1.0.0      v stringr    1.5.1
v ggplot2    3.5.1      v tibble     3.2.1
v lubridate  1.9.3      v tidyr      1.3.1
v purrr      1.0.2
-- 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
```

```
library(tidymodels)
```

```
-- Attaching packages ----- tidymodels 1.2.0 --
v broom       1.0.6      v rsample     1.2.1
v dials       1.3.0      v tune        1.2.1
v infer       1.0.7      v workflows   1.1.4
v modeldata   1.4.0      v workflowsets 1.1.0
v parsnip     1.2.1      v yardstick   1.3.1
v recipes     1.1.0
-- Conflicts ----- tidymodels_conflicts() --
x scales::discard() masks purrr::discard()
```

```
x dplyr::filter() masks stats::filter()
x recipes::fixed() masks stringr::fixed()
x dplyr::lag() masks stats::lag()
x yardstick::spec() masks readr::spec()
x recipes::step() masks stats::step()
* Search for functions across packages at https://www.tidymodels.org/find/
```

```
library(ggforce)
```

Warning: package 'ggforce' was built under R version 4.4.2

```
data <- read_csv("https://raw.githubusercontent.com/agmath/agmath.github.io/master/data/class")
```

Rows: 801 Columns: 2502

-- Column specification -----

Delimiter: ","

chr (1): labels

dbl (2501): gene_16154, gene_9197, gene_15497, gene_4299, gene_340, gene_113...

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.

```
set.seed(434)
data_splits <- initial_split(data, prop = 0.8)

train <- training(data_splits)
test <- testing(data_splits)
```