

2.3. Singular Value Decomposition - Geometric interpretation

Lieven Clement

statOmics, Ghent University (<https://statomics.github.io>)

Contents

1	Iris dataset	1
1.1	Subset the data	1
1.2	Center the data	2
2	SVD	2
3	Geometric Interpretation?	2
3.1	Projection of a single data point	3
3.2	Projection of all datapoints: project all rows of X on V2	3

We introduce the geometric interpretation of the svd by using a toy example.

1 Iris dataset

The iris dataset is a dataset on iris flowers.

- Three species (setosa, virginica and versicolor)
- Length and width of Sepal leaves
- Length and width of Petal Leaves

For didactical purposes we will use a subset of the data.

- Virginica Species
- 3 Variables: Sepal Length, Sepal Width, Petal Length
- This allows us to visualise the data in 3D plots
- Illustrate the data compression of the SVD from 3 to two dimensions.

1.1 Subset the data

```
library(tidyverse)
library(plotly)
irisSub <- iris %>%
  filter(Species == "virginica") %>%
  dplyr::select("Sepal.Length", "Sepal.Width", "Petal.Length")
```

1.2 Center the data

```
X <- irisSub %>% scale(scale=FALSE)
```

The data is translated to a mean of $[0, 0, 0]$.

We zoom in and add the original axis in grey in the origin.

2 SVD

1. We adopt the SVD on the centered data

```
irisSvd <- svd(X)
```

2. We extract

- the right singular vectors \mathbf{V} and
- the projections \mathbf{Z}

```
V <- irisSvd$v  
Z <- irisSvd$u %*% diag(irisSvd$d)
```

Note, that

- the SVD is essentially a rotation to a new coordinate system.
- we plotted \mathbf{V}_3 with dots because we will use the SVD for dimension reduction

$$3D \rightarrow 2D$$

Rotate the plot

- Note, that
 - \mathbf{V}_1 points in the direction of the largest variability in the data
 - \mathbf{V}_2 points in a direction orthogonal on \mathbf{V}_1 pointing in the direction of the second largest variability in the data.

3 Geometric Interpretation?

Write the **truncated SVD** as

$$\mathbf{X}_k = \mathbf{U}_k \Delta_k \mathbf{V}_k^T = \mathbf{Z}_k \mathbf{V}_k^T$$

with

$$\mathbf{Z}_k = \mathbf{U}_k \Delta_k$$

an $n \times k$ matrix.

Each of the n rows of \mathbf{Z}_k , say $\mathbf{z}_{k,i}^T$, represents a point in a k -dimensional space.

```
V2 <- V[,1:2]
Z2 <- Z[,1:2]
X2 <- Z2 %*% t(V2)
```

Because of the orthonormality of the singular vectors, we also have

$$\begin{aligned}\mathbf{X}_k \mathbf{V}_k &= \mathbf{Z}_k \mathbf{V}_k^T \mathbf{V}_k \\ \mathbf{X}_k \mathbf{V}_k &= \mathbf{Z}_k.\end{aligned}$$

Thus the matrix \mathbf{V}_k is a **transformation matrix** that may be used to transform \mathbf{X}_k into \mathbf{Z}_k , and \mathbf{Z}_k into \mathbf{X}_k .

More importantly, it can be shown that (thanks to orthonormality of \mathbf{V})

$$\mathbf{X} \mathbf{V}_k = \mathbf{Z}_k.$$

This follows from (w.l.g. $\text{rank}(\mathbf{X})=r$)

$$\begin{aligned}\mathbf{X} \mathbf{V}_k &= \mathbf{U} \mathbf{D} \mathbf{V}^T \mathbf{V}_k = \mathbf{U} \mathbf{D} \begin{pmatrix} \mathbf{v}_1^T \\ \vdots \\ \mathbf{v}_r^T \end{pmatrix} (\mathbf{v}_1 \dots \mathbf{v}_k) \\ &= \mathbf{U} \mathbf{D} \mathbf{V}^T \mathbf{V}_k = \mathbf{U} \mathbf{D} \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} = \mathbf{U}_k \Delta_k = \mathbf{Z}_k\end{aligned}$$

The $p \times k$ matrix \mathbf{V}_k acts as a transformation matrix: transforming n points in a p dimensional space to n points in a k dimensional space.

```
Z2proj <- X %*% V2
range(Z2 - Z2proj)
```

```
## [1] -8.881784e-16 1.082467e-15
```

3.1 Projection of a single data point

- Zoom in to see projection.
- The projection is indicated for the blue point X_{44} to the red point $X_{2,44}$ in the plane spanned by \mathbf{V}_2 .

3.2 Projection of all datapoints: project all rows of \mathbf{X} on \mathbf{V}_2

- Zoom in first look orthonal via direction \mathbf{V}_2 (rotate until text \mathbf{V}_2 is viewed in the origin)
- Zoom in first look orthonal via direction \mathbf{V}_1 (rotate until text \mathbf{V}_1 is viewed in the origin)
- Note, that

- V1 points in the direction of the largest variability in the data
 - V2 points in a direction orthogonal on V1 pointing in the direction of the second largest variability in the data.
- Projection only.
- This clearly shows that the projected points X_2 (X projected on V2) live in a two dimensional space