Singular Value Decomposition

Sharon Wulff

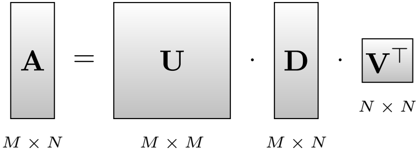
March 7-8, 2013

SVD Exercise 1

- Refresher on class material

SVD Theorem

Let $\mathbf{A}$ be any real M by N matrix, $\mathbf{A} \in \mathbb{R}^{M \times N}$,
then $\mathbf{A}$ can be decomposed as $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^\top$.



- $\mathbf{U}$ is an M by M orthogonal matrix, such that $\mathbf{U}^\top\mathbf{U} = \mathbf{I}_{(M)}$.
- $\mathbf{D}$ is an M by N diagonal matrix
- $\mathbf{V}^\top$ is also an orthogonal matrix, N by N, $\mathbf{V}^\top\mathbf{V} = \mathbf{I}_{(N)}$.

SVD Interpretation

“Users“, “Movies“ and “Concepts“:

- $\mathbf{U}$: Users-to-concept affinity matrix
- $\mathbf{V}$: Movies-to-concept similarity matrix
- $\mathbf{D}$: The diagonal elements of $\mathbf{D}$ represent the “expressiveness“ of each concept in the data.

SVD Example

$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^\top:$$

$$
\begin{array}{c}
\begin{array}{ccccc} \text{Cremators} & \text{Evil spawn} & \text{Fatal justice} & \text{Clerks} & \text{American pie} \end{array} \\
\begin{pmatrix}
5 & 5 & 5 & 0 & 0 \\
4 & 4 & 4 & 0 & 0 \\
5 & 5 & 5 & 0 & 0 \\
3 & 3 & 3 & 0 & 0 \\
0 & 0 & 0 & 4 & 4 \\
0 & 0 & 0 & 5 & 5 \\
0 & 0 & 0 & 4 & 4
\end{pmatrix}
\end{array}
=
\begin{pmatrix}
0.57 & 0 & -0.80 & 0.06 & -0.04 & -0.06 & -0.04 \\
0.46 & 0 & 0.43 & 0.68 & -0.19 & -0.23 & -0.19 \\
0.57 & 0 & 0.37 & -0.70 & -0.08 & -0.11 & -0.08 \\
0.34 & 0 & 0.15 & 0.14 & 0.48 & 0.60 & 0.48 \\
0 & 0.52 & 0 & 0 & -0.71 & 0.35 & 0.28 \\
0 & 0.66 & 0 & 0 & 0.35 & -0.56 & 0.35 \\
0 & 0.52 & 0 & 0 & 0.28 & 0.35 & -0.71
\end{pmatrix}
\times
$$

$$
\begin{pmatrix}
15 & 0 & 0 & 0 & 0 \\
0 & 10.67 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0
\end{pmatrix}
\times
\begin{pmatrix}
0.57 & 0.57 & 0.57 & 0 & 0 \\
0 & 0 & 0 & 0.70 & 0.70 \\
0.81 & -0.40 & -0.40 & 0 & 0 \\
0 & 0.70 & -0.70 & 0 & 0 \\
0 & 0 & 0 & 0.70 & -0.70
\end{pmatrix}
$$

SVD Example

Concepts: **Horror, Comedy**

$\mathbf{U}$: Users-to-concept affinity matrix.

$$
\begin{array}{c}
\begin{array}{ccccc} \text{Cremators} & \text{Evil spawn} & \text{Fatal justice} & \text{Clerks} & \text{American pie} \end{array} \\
\begin{pmatrix} 5 & 5 & 5 & 0 & 0 \\ 4 & 4 & 4 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 3 & 3 & 3 & 0 & 0 \\ 0 & 0 & 0 & 4 & 4 \\ 0 & 0 & 0 & 5 & 5 \\ 0 & 0 & 0 & 4 & 4 \end{pmatrix}
\end{array}
=
\begin{array}{c}
\begin{array}{cc} \textbf{Horror} & \textbf{Comedy} \end{array} \\
\begin{pmatrix} 0.57 & 0 \\ 0.46 & 0 \\ 0.57 & 0 \\ 0.34 & 0 \\ 0 & 0.52 \\ 0 & 0.66 \\ 0 & 0.52 \end{pmatrix}
\end{array}
\times
\begin{pmatrix} 15 & 0 \\ 0 & 10.67 \end{pmatrix}
\times
\begin{pmatrix} 0.57 & 0.57 & 0.57 & 0 & 0 \\ 0 & 0 & 0 & 0.70 & 0.70 \end{pmatrix}
$$

Q: What is the affinity between user1 and horror? 0.57

SVD Example

Concepts: Horror, Comedy

$\mathbf{D}$: Expression level of the different concepts in the data.

Strength of Horror

Strength of Comedy

$$\begin{array}{c} \begin{array}{ccccc} \text{Cremators} & \text{Evil spawn} & \text{Fatal justice} & \text{Clerks} & \text{American pie} \end{array} \\ \begin{pmatrix} 5 & 5 & 5 & 0 & 0 \\ 4 & 4 & 4 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 3 & 3 & 3 & 0 & 0 \\ 0 & 0 & 0 & 4 & 4 \\ 0 & 0 & 0 & 5 & 5 \\ 0 & 0 & 0 & 4 & 4 \end{pmatrix} \end{array} = \begin{pmatrix} 0.57 & 0 \\ 0.46 & 0 \\ 0.57 & 0 \\ 0.34 & 0 \\ 0 & 0.52 \\ 0 & 0.66 \\ 0 & 0.52 \end{pmatrix} \times \begin{pmatrix} 15 & 0 \\ 0 & 10.67 \end{pmatrix} \times \begin{pmatrix} 0.57 & 0.57 & 0.57 & 0 & 0 \\ 0 & 0 & 0 & 0.70 & 0.70 \end{pmatrix}$$

Q: What is the expression of the comedy concept in the data? 10.67

SVD Example

Concepts: Horror, Comedy

$\mathbf{V}$: Movies-to-concept similarity matrix.

$$\begin{array}{ccccc} \text{Cremators} & \text{Evil spawn} & \text{Fatal justice} & \text{Clerks} & \text{American pie} \end{array}$$

$$\begin{pmatrix} 5 & 5 & 5 & 0 & 0 \\ 4 & 4 & 4 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 3 & 3 & 3 & 0 & 0 \\ 0 & 0 & 0 & 4 & 4 \\ 0 & 0 & 0 & 5 & 5 \\ 0 & 0 & 0 & 4 & 4 \end{pmatrix} = \begin{pmatrix} 0.57 & 0 \\ 0.46 & 0 \\ 0.57 & 0 \\ 0.34 & 0 \\ 0 & 0.52 \\ 0 & 0.66 \\ 0 & 0.52 \end{pmatrix} \times \begin{pmatrix} 15 & 0 \\ 0 & 10.67 \end{pmatrix} \times \begin{pmatrix} 0.57 & 0.57 & 0.57 & 0 & 0 \\ 0 & 0 & 0 & 0.70 & 0.70 \end{pmatrix}$$

Q: What is the similarity between Clerks and Horror? 0

What is the similarity between Clerks and Comedy? 0.7

Closest matrix approximation

Let the SVD of $\mathbf{A} \in \mathbb{R}^{M \times N}$ be given by $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^\top$

Define $\mathbf{A}_k$ as

$$\mathbf{A}_k = \sum_{i=1}^{k} d_i \mathbf{u}_i \mathbf{v}_i^\top$$

Where $k < r = \text{Rank}(\mathbf{A})$

Forbenious norm

Def. The Forbenious norm is matrix norm, defined as the square root of the sum of the absolute squares of its elements.
For $\mathbf{A} \in \mathbb{R}^{M \times N}$:

$$\|\mathbf{A}\|_F := \sqrt{\sum_{i=1}^{M} \sum_{j=1}^{N} |A_{i,j}|^2}$$

- The matrix $\mathbf{A}_k$ is also the closets k-rank matrix, under the forbenious norm (why also?).

Forbenious norm

Comparison with the euclidian norm:

- $\|\mathbf{A}\|_2 = d_1$
- $\|\mathbf{A}\|_F^2 = d_1^2 + \ldots + d_r^2$

Forbenious norm

Comparison with the euclidian norm:

- $\|\mathbf{A}\|_2 = d_1$
- $\|\mathbf{A}\|_F^2 = d_1^2 + \ldots + d_r^2$

Therefore:

$$\min_{\text{Rank}(\mathbf{B})=k} \|\mathbf{A} - \mathbf{B}\|_2 = \|\mathbf{A} - \mathbf{A_k}\|_2 = d_{k+1}$$

while

$$\min_{\text{Rank}(\mathbf{B})=k} \|\mathbf{A} - \mathbf{B}\|_F^2 = \|\mathbf{A} - \mathbf{A_k}\|_F^2 = d_{k+1}^2 + \ldots + d_r^2$$

SVD Exercise 1

- Solve the "Pen and Paper" exercise

Assignment - PredictMissingValues

Simple solution:

```
X_pred = PredictMissingValues(X, nil)

  missing_values_indices = find(X == nil);

  % impute missing values
  avg_of_some_sort = func(X, existing values);

  X_pred = X;

  % Fill in imputed values

  X_pred(missing_values_indices) =
  avg_of_some_sort(missing_values_indices);
```

Assignment - PredictMissingValues

```
X_pred = PredictMissingValues(X, nil)

  % Repeat simple solution, obtain a full matrix X_pred

  [U,S,V] = svd(X_pred);

  k = model_selection;

  approximation_by_svd = ...

  X_pred(missing_values_indices) =
  approximation_by_svd(missing_values_indices);
```