Tutorial 13

SVD

Technion, CS department
SIPC 236327, Winter 2015-16
The Singular Value Decomposition (SVD)

- Any $m \times n$ matrix $\Gamma$ can be decomposed to the form:

\[
\Gamma = U \Lambda^{1/2} V^T
\]

where

- $U$ is an $m \times m$ orthogonal matrix
- $V$ is an $n \times n$ orthogonal matrix

and the diagonal $m \times n$ matrix:

\[
\Lambda^{1/2} = \begin{bmatrix}
\lambda_1^{1/2} & \cdots & 0 \\
0 & \ddots & 0 \\
0 & \cdots & \lambda_r^{1/2}
\end{bmatrix}
\]

where

- $r$ is the rank of $\Gamma$
- and $\lambda_i^{1/2} > 0$
SVD: \( \Gamma = U \Lambda^{1/2} V^T \)

- \( U \) is the orthogonal eigenvector set of \( \Gamma \Gamma^T \)
- \( V \) is the orthogonal eigenvector set of \( \Gamma^T \Gamma \)

- Show that \( \Gamma \Gamma^T \) and \( \Gamma^T \Gamma \) have the same eigenvalues.

We know that \( \Gamma \Gamma^T = U \Lambda_1 U^T \), specifically \( \Gamma \Gamma^T u_i = \lambda_i u_i \)

Multiplying by \( \Gamma^T \) gives

\[ \Gamma^T \Gamma \Gamma^T u_i = \lambda_i \Gamma^T u_i \]

Defining \( \tilde{v}_i = \Gamma^T u_i \) better shows that \( \Gamma^T \Gamma \tilde{v}_i = \lambda_i \tilde{v}_i \)

\[ \Rightarrow \lambda_i \text{ is also an eigenvalue of } \Gamma^T \Gamma. \]
SVD: \( \Gamma = U \Lambda^{1/2} V^T \)

- We've got that \( V \) can be determined by \( U \):

\( \lambda_i \) is also an eigenvalue of \( \Gamma^T \Gamma \)

with a corresponding eigenvector \( \tilde{v}_i = \Gamma^T u_i \).

Normalizing \( \tilde{v}_i \) will give the corresponding member of \( V \):

\[
\begin{align*}
v_i &= \frac{\tilde{v}_i}{\|\tilde{v}_i\|_2} \\
\|\tilde{v}_i\|_2^2 &= (\Gamma^T u_i)^T (\Gamma^T u_i) = u_i^T \Gamma \Gamma^T u_i = u_i^T \lambda_i u_i = \lambda_i \|u_i\|_2^2 = \lambda_i \\
\Rightarrow \quad v_i &= \frac{1}{\lambda_i^{1/2}} \Gamma^T u_i
\end{align*}
\]
SVD: The Compact Form

- There are \( r \) nonzeros on the diagonal of

\[
\begin{bmatrix}
\lambda_1^{1/2} & & 0 \\
& \ddots & \\
0 & & \lambda_r^{1/2}
\end{bmatrix}
\]

Hence, we can write the following compact form:

\[
\Gamma = U_r \Lambda_r^{1/2} V_r^T
\]

where

\[U_r = [u_1, \ldots, u_r]\] is an \( m \times r \) matrix

\[V_r = [v_1, \ldots, v_r]\] is an \( n \times r \) matrix

\[
\Lambda_r^{1/2} = diag(\frac{1}{\lambda_1^2}, \ldots, \frac{1}{\lambda_r^2})
\]
SVD: The Outer-Product Form

The compact form can also be written as

$$\Gamma = \sum_{i=1}^{r} \lambda_i^{1/2} u_i v_i^T$$

Where $u_i$ and $v_i$ are the $i^{th}$ columns of $U$ and $V$, respectively.

Each term in the above summation is a rank-one matrix. Hence, a $k$-rank approximation of $\Gamma$ is given by

$$\hat{\Gamma}_k = \sum_{i=1}^{k} \lambda_i^{1/2} u_i v_i^T$$

$k \leq r$

This is the best approximation in the squared-error sense for a given $\Gamma$. 
SVD: The Image-Transform Perspective

Given a 2D image-matrix $X$ (of $n \times n$ size), the SVD can be utilized as a 2D-transform:

$$\Lambda^{1/2} = U^TXV$$

**Properties:**

- The transform is **image-dependent**.
- The transformed-signal is a diagonal matrix
  - With $r \leq n$ nonzero values.

The **inverse transform** can be considered as:

$$\hat{X} = \sum_{i=1}^{r} \lambda_i^{1/2} u_i v_i^T$$
SVD: The Image-Transform Perspective

What is the **representation cost** of the \( k \)-rank approximation?

The approximation is

\[
\hat{X}_k = \sum_{i=1}^{k} \lambda_i^{1/2} u_i v_i^T
\]

- Each basis-image, \( u_i v_i^T \), is representable using \( 2n \) numbers.
- \( k \)-representation layers require \( k \cdot (2n + 1) \) numbers.
Exercise

Let us consider the following 2x2 image patch:

\[ X = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} \]

What is its SVD transform?

\[ XX^T = \begin{bmatrix} 5 & 4 \\ 4 & 4 \end{bmatrix} \]

\[ \det \left( XX^T - \lambda I \right) = 0 \]

\[ (5 - \lambda)(4 - \lambda) - 16 = 0 \]

\[ (8.53 - \lambda)(0.47 - \lambda) = 0 \]

\[ \lambda_1 = 8.53 \quad \lambda_2 = 0.47 \]

\[ \Rightarrow \quad u_1 = \begin{bmatrix} 0.75 \\ 0.66 \end{bmatrix} \text{ and } u_2 = \begin{bmatrix} -0.66 \\ 0.75 \end{bmatrix} \]
Exercise

• Calculation of \( V \):

\[
v_1 = \frac{1}{\sqrt{\lambda_1}} X^T u_1 = \frac{1}{2.92} \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0.75 \\ 0.66 \end{bmatrix} = \begin{bmatrix} 0.26 \\ 0.97 \end{bmatrix}
\]

\[
v_2 = \frac{1}{\sqrt{\lambda_2}} X^T u_2 = \frac{1}{0.68} \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} -0.66 \\ 0.75 \end{bmatrix} = \begin{bmatrix} -0.97 \\ 0.26 \end{bmatrix}
\]

• Therefore, the SVD transform of \( X = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} \) is

\[
\begin{bmatrix} -0.75 & -0.66 \\ 0.66 & -0.75 \end{bmatrix} \begin{bmatrix} -0.75 & 0.97 \\ -2.82 & -0.26 \end{bmatrix} = \begin{bmatrix} 2.92 & 0 \\ 0 & 0.68 \end{bmatrix}
\]
Exercise

• Let us examine the energy compaction:

\[ \varepsilon_{total} = \lambda_1 + \lambda_2 = 8.53 + 0.47 = 9 \]

\[ \varepsilon_1 = \lambda_1 = 8.53 = \frac{\lambda_1}{\lambda_1 + \lambda_2} \varepsilon_{total} \approx 0.95 \varepsilon_{total} \]

\[ \varepsilon_2 = \lambda_2 \approx 0.05 \varepsilon_{total} \]

The SVD of the given image achieves an impressive energy compaction!
* We should recall that it is a costly image-dependent transform.

The corresponding 1-rank approximation is:

\[ \hat{X}_1 = \sqrt{\lambda_1} u_i v_i^T = 2.92 \begin{bmatrix} 0.75 \\ 0.66 \end{bmatrix} \begin{bmatrix} 0.26 & 0.97 \end{bmatrix} = \begin{bmatrix} 0.56 & 2.12 \\ 0.50 & 1.86 \end{bmatrix} \]
Exercise

• We can assume that $X$ is a realization of a first-order Markov process.

• Therefore, the 2D-DCT should be a good alternative:

$$X_{dct} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} X \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 2.5 & -1.5 \\ 0.5 & 0.5 \end{bmatrix}$$

• Let us examine the energy compaction:
  • Energy is spread on all coefficients.
  • The largest coefficient contain 70% of the total energy

$$\varepsilon_1 = 2.5^2 \approx 0.7 \varepsilon_{total}$$

which is 25% less than in the SVD, and accordingly result in an increased error for the (1-term) approximation.