Singular-Value Decomposition with Applications

Anil Maheshwari

anil@scs.carleton.ca
School of Computer Science
Carleton University
Canada
Outline

Matrices - Eigenvalues & Eigenvectors

Singular Value Decomposition

Low Rank Approximations

An Application

Correctness
Matrices - Eigenvalues & Eigenvectors
Given an $n \times n$ matrix $A$.
A non-zero vector $v$ is an eigenvector of $A$, if $Av = \lambda v$ for some scalar $\lambda$. $\lambda$ is the eigenvalue corresponding to vector $v$.

**Example**

Let $A = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}$

Observe that

$$\begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = 5 \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$
and

$$\begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Thus, $\lambda_1 = 5$ and $\lambda_2 = 1$ are the eigenvalues of $A$.
Corresponding eigenvectors are $v_1 = [1, 3]$ and $v_2 = [1, -1]$, as $Av_1 = \lambda_1 v_1$ and $Av_2 = \lambda_2 v_2$. 
Example

Consider symmetric matrix \( S = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \).

Its eigenvalues are \( \lambda_1 = 4 \) and \( \lambda_2 = 2 \) and the corresponding eigenvectors are \( q_1 = (1/\sqrt{2}, 1/\sqrt{2}) \) and \( q_2 = (1/\sqrt{2}, -1/\sqrt{2}) \), respectively.

Note that eigenvalues are real and the eigenvectors are orthonormal.

\[
S = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}
\]

Eigenvalues of Symmetric Matrices

All the eigenvalues of a real symmetric matrix \( S \) are real. Moreover, all components of the eigenvectors of a real symmetric matrix \( S \) are real.
Property
Any pair of eigenvectors of a real symmetric matrix $S$ corresponding to two different eigenvalues are orthogonal.
Proof: Let $q_1$ and $q_2$ be two eigenvectors corresponding to $\lambda_1 \neq \lambda_2$, respectively. Thus, $Sq_1 = \lambda_1 q_1$ and $Sq_2 = \lambda_2 q_2$. Since $S$ is symmetric, $q_1^T S = \lambda_1 q_1^T$. Multiply by $q_2$ on the right and we obtain $\lambda_1 q_1^T q_2 = q_1^T Sq_2 = q_1^T \lambda_2 q_2$. Since $\lambda_1 \neq \lambda_2$ and $\lambda_1 q_1^T q_2 = q_1^T \lambda_2 q_2$, this implies that $q_1^T q_2 = 0$ and thus the eigenvectors $q_1$ and $q_2$ are orthogonal.
Symmetric Matrices (contd.)

Symmetric matrices with distinct eigenvalues

Let $S$ be a $n \times n$ symmetric matrix with $n$ distinct eigenvalues and let $q_1, \ldots, q_n$ be the corresponding orthonormal eigenvectors. Let $Q$ be the $n \times n$ matrix consisting of $q_1, \ldots, q_n$ as its columns. Then

$$S = Q\Lambda Q^{-1} = Q\Lambda Q^T.$$  Furthermore, $S = \lambda_1 q_1 q_1^T + \lambda_2 q_2 q_2^T + \cdots + \lambda_n q_n q_n^T$

$$S = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} = 4 \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} + 2 \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$$
Symmetric matrix $S$ is *positive definite* if all its eigenvalues $> 0$. It is *positive semi-definite* if all the eigenvalues are $\geq 0$.

**An Alternate Characterization**

Let $S$ be a $n \times n$ real symmetric matrix. For all non-zero vectors $x \in \mathbb{R}^n$, if $x^T S x > 0$ holds, then all the eigenvalues of $S$ are $> 0$. 

Let $\lambda_i$ be an eigenvalue of $S$ and its corresponding unit eigenvector is $q_i$. Note that $q_i^T q_i = 1$. Since $S$ is symmetric, we know that $\lambda_i$ is real. Now we have, $\lambda_i = \lambda_i q_i^T q_i = q_i^T \lambda_i q_i = q_i^T S q_i$. But $q_i^T S q_i > 0$, hence $\lambda_i > 0$. 
Diagonalization Summary

Square Matrices:
A be an $n \times n$ matrix with distinct eigenvalues.
$X_{n \times n} = \text{Matrix of eigenvectors of } A$

$AX = X\Lambda$, $A = X\Lambda X^{-1}$, $\Lambda = X^{-1}\Lambda X$

Symmetric Matrices:
$S$ be an $n \times n$ symmetric matrix with distinct eigenvalues.
$Q_{n \times n} = \text{Matrix of } n\text{-orthonormal eigenvectors of } S$

$S = Q\Lambda Q^T$

What if $A$ is a rectangular matrix of dimensions $m \times n$?
Singular Value Decomposition
Let $A$ be a $m \times n$ matrix of rank $r$ with real entries.

We can find orthonormal vectors in $\mathbb{R}^n$ such that their product with $A$ results in a scaled copy of orthonormal vectors in $\mathbb{R}^m$.

Formally, we can find

1. Orthonormal vectors $v_1, \ldots, v_r \in \mathbb{R}^n$
2. Orthonormal vectors $u_1, \ldots, u_r \in \mathbb{R}^m$
3. Real numbers $\sigma_1, \ldots, \sigma_r \in \mathbb{R}$
4. For $i = 1, \ldots, r$: $Av_i = \sigma_i u_i$
5. $AV = U\Sigma$, i.e.,

$$
A \begin{bmatrix}
  v_1 & \ldots & v_r
\end{bmatrix} =
\begin{bmatrix}
  u_1 & \ldots & u_r
\end{bmatrix}
\begin{bmatrix}
  \sigma_1 \\
  \vdots \\
  \sigma_r
\end{bmatrix}
$$

6. $A = U\Sigma V^T$
An Example: $AV = U\Sigma$

\[
\begin{bmatrix}
1 & 5 \\
0 & 3 \\
1 & 4 \\
4 & 0 \\
5 & 1
\end{bmatrix}
\begin{bmatrix}
.60 & -.8 \\
.8 & .6
\end{bmatrix}
= \begin{bmatrix}
.58 & .39 \\
.31 & .30 \\
.48 & .28 \\
.30 & -.56 \\
.48 & -.59
\end{bmatrix}
\begin{bmatrix}
7.8 & 0 \\
0 & 5.7
\end{bmatrix}
\]

Alternatively, $A = U\Sigma V^T$

\[
\begin{bmatrix}
1 & 5 \\
0 & 3 \\
1 & 4 \\
4 & 0 \\
5 & 1
\end{bmatrix}
\begin{bmatrix}
.58 & .39 \\
.31 & .30 \\
.48 & .28 \\
.30 & -.56 \\
.48 & -.59
\end{bmatrix}
= \begin{bmatrix}
7.8 & 0 \\
0 & 5.7
\end{bmatrix}
\begin{bmatrix}
.60 & .8 \\
-.8 & .6
\end{bmatrix}
\]

Play around with the SVD command in Wolfram Alpha for some matrices.
**Symmetric and Positive semi-definite**

Let $A$ be $m \times n$ matrix, where $m \geq n$. The matrix $A^T A$ is symmetric and positive semi-definite.

**Proof:**

**Symmetric:** $(A^T A)^T = A^T (A^T)^T = A^T A$

**Positive semi-definite:** Take any non-zero vector $x \in \mathbb{R}^n$

$$x^T (A^T A)x = (x^T A^T)(Ax) = (Ax)^T (Ax) = \|Ax\|^2 \geq 0$$
Matrix $A^T A$ (contd.)

$A^T A$ is a symmetric matrix of dimension $n \times n$. Eigenvalues of $A^T A$ are non-negative and the corresponding eigenvectors are orthonormal.

Let $\lambda_1 \geq \ldots \geq \lambda_n$ be eigenvalues of $A^T A$ and let $v_1, \ldots, v_n$ be the corresponding eigenvectors.

$$A^T A v_i = \lambda_i v_i \iff v_i^T A^T A v_i = \lambda_i$$

Define $\sigma_i = ||Av_i|| \implies \sigma_i^2 = ||Av_i||^2 = v_i^T A^T A v_i = \lambda_i$

Hence, $\sigma_i = ||Av_i|| = \sqrt{\lambda_i}$

Consider two cases:

**Full Rank:** Rank of $A^T A$ is $n$.

**Low Rank:** Rank of $A^T A$ is $r < n$. 
Matrix $A^T A$ is Full Rank

Assume, $\sigma_1 \geq \ldots \geq \sigma_n > 0$

($\implies A$ and $A^T A$ has rank $n$)

Define vectors $u_1, \ldots, u_n \in \mathbb{R}^m$ as $u_i = Av_i / \sigma_i$

Orthonormal

The set of vectors $u_i = Av_i / \sigma_i$, for $i = 1, \ldots, n$, are orthonormal.

Proof: $||u_i|| = ||Av_i|| / \sigma_i = \sigma_i / \sigma_i = 1$

Consider the dot product of any two vectors $u_i$ and $u_j$:

$u_i^T u_j = (Av_i / \sigma_i)^T (Av_j / \sigma_j) = \frac{1}{\sigma_i \sigma_j} v_i^T A^T A v_j = \frac{1}{\sigma_i \sigma_j} v_i^T \lambda_j v_j = \frac{\lambda_j}{\sigma_i \sigma_j} v_i^T v_j = 0$

$\blacksquare$

$\Rightarrow Av_i = \sigma_i u_i$ for $i = 1, \ldots, r = n$
Matrix $A^T A$ is Full Rank

Why is $\text{rank}(A) = \text{rank}(A^T A)$?

We will show that $\text{Null} - \text{Space}(A) \subseteq \text{Null} - \text{Space}(A^T A)$ and $\text{Null} - \text{Space}(A^T A) \subseteq \text{Null} - \text{Space}(A)$. This implies that $\text{Null} - \text{Space}(A) = \text{Null} - \text{Space}(A^T A)$ and $\text{rank}(A) = \text{rank}(A^T A) = n - \text{rank}(\text{Null} - \text{Space}(A))$.

Consider a vector $x \in \text{Null} - \text{Space}(A)$.
Then $Ax = \vec{0}$ and $A^T Ax = A^T (Ax) = A^T \vec{0} = \vec{0}$.
$\Rightarrow x \in \text{Null} - \text{Space}(A^T A)$.

Consider a vector $y$ such that $A^T Ay = \vec{0}$.
Then $y^T A^T Ay = \vec{0}$ or $(Ay)^T (Ay) = \vec{0}$.
$\Rightarrow Ay = \vec{0}$ and $y \in \text{Null} - \text{space}(A)$.

$\square$
Suppose $m \geq n$, but $\text{rank}(A) = r < n$.

### Eigenvalues of $A^TA$

The $n - r$ eigenvalues of $A^TA$ are equal to 0.

**Proof:** Consider a basis of the null space of $A$.
Let $x_1, \ldots, x_{n-r}$ be a basis of the null space of $A$.
This implies that $Ax_j = 0$ for $j = 1, \ldots, n - r$.
Now, $A^TAx_j = 0 = 0x_j$.
Thus, 0 is an eigenvalue of $A^TA$ corresponding to each $x_i$’s.
Thus $n - r$ eigenvalues of $A^TA$ are equal to 0.

□
Handling low rank (contd.)

Consider eigenvalues and eigenvectors of \( A^T A \)
Let \( \lambda_1 \geq \ldots \geq \lambda_r > 0 \) and \( \lambda_{r+1} = \ldots = \lambda_n = 0 \)

Let \( v_1, \ldots, v_r \) be the orthonormal vectors corresponding to \( \lambda_1, \ldots, \lambda_r \)

For \( i = 1, \ldots, r \), define \( \sigma_i = \|Av_i\| = \sqrt{\lambda_i} \)
Note that \( \sigma_1 \geq \ldots \sigma_r > 0 \)

For \( i = 1, \ldots, r \), define \( u_i = \frac{1}{\sigma_i} Av_i \)

**SVD for \( A \)**

Vectors \( u_1, \ldots, u_r \) are orthonormal and \( Av_i = \sigma_i u_i \).
Singular Value Decomposition

For a matrix $A$ of dimension $m \times n$, where $m \geq n$, we have

1. $A^T A$ is a symmetric positive semidefinite square matrix of dimension $n \times n$.

2. Rank of $A$ is $n$: $\lambda_1 \geq \ldots \geq \lambda_n > 0$ are eigenvalues of $A^T A$ and $v_1, \ldots, v_n$ the corresponding orthonormal eigenvectors. The vectors $u_i = A v_i / \sigma_i$, for $i = 1, \ldots, n$, are orthonormal, where $\sigma_i = \sqrt{\lambda_i}$.

3. Rank of $A$ is $r < n$: $\lambda_1 \geq \ldots \geq \lambda_r > 0$ are non-zero eigenvalues of $A^T A$ and $v_1, \ldots, v_r$ the corresponding orthonormal eigenvectors. The vectors $u_i = A v_i / \sigma_i$, for $i = 1, \ldots, r$, are orthonormal, where $\sigma_i = \sqrt{\lambda_i}$.

4. $AV = U \Sigma$, where $V$ is $n \times r$ matrix consisting of orthonormal eigenvectors of $A^T A$ corresponding to non-zero eigenvalues of $A^T A$, $U$ is $m \times r$ matrix of orthonormal vectors given by $u_i = A v_i / \sigma_i$ for non-zero $\sigma_i$, and $\Sigma$ is $r \times r$ diagonal matrix.

5. $A V V^T = A = U \Sigma V^T \leftarrow$ Singular-Value Decomposition of $A$. 

16
Matrices $A^T A$ and $AA^T$

We have $A = U \Sigma V^T$.

$$A^T A = (U \Sigma V^T)^T (U \Sigma V^T) = (V \Sigma U^T)(U \Sigma V^T) = V \Sigma (U^T U) \Sigma V^T = V \Sigma^2 V^T$$

**Matrix $A^T A$**

$A^T A$ is square symmetric matrix and it is expressed in the diagonalized form $A^T A = V \Sigma^2 V^T$. Thus, $\sigma_i^2$'s are its eigenvalues and $V$ is its eigenvectors matrix.

Similarly, consider $AA^T$ and we obtain that

$$AA^T = (U \Sigma V^T)(U \Sigma V^T)^T = U \Sigma V^T V \Sigma U^T = U \Sigma^2 U^T.$$ 

**Matrix $AA^T$**

$AA^T$ is square symmetric matrix and it is expressed in the diagonalized form $AA^T = U \Sigma^2 U^T$. Thus $U$ is the eigenvector matrix for the symmetric matrix $AA^T$ with the same eigenvalues as $A^T A$. 
Singular Value Decomposition - Summary

- Let $A$ be a $m \times n$ matrix of real numbers of rank $r$

- $A_{m \times n} = U_{m \times r} \Sigma_{r \times r} V_{r \times n}^T$, where

$U$ is a orthonormal $m \times r$ matrix
$V$ is a orthonormal $n \times r$ matrix
$\Sigma$ is an $r \times r$ diagonal matrix and its $(i, i)$-th entry is $\sigma_i$ for $i = 1, \ldots, r$

- Note that $\sigma_1 \geq \sigma_2 \geq \ldots \sigma_r > 0$ and $\sigma_i = \sqrt{\lambda_i}$ where $\lambda_i$ are the eigenvalues of $A^T A$

- The set of orthonormal vectors $v_1, \ldots, v_r$ and $u_1, \ldots, u_r$ are eigenvectors of $A^T A$ and $AA^T$, respectively. The vectors $v$’s and $u$’s satisfy the equation $Av_i = \sigma_i u_i$, for $i = 1, \ldots, r$

- Alternatively, we can express $A$ as the sum of the product of rank 1 matrices

$$A = \sum_{i=1}^{r} \sigma_i u_i v_i^T = \sigma_1 u_1 v_1^T + \ldots + \sigma_r u_r v_r^T$$
Low Rank Approximations
Let $A_{m \times n}$ be the **Utility Matrix**, where $m = 10^8$ users and $n = 10^5$ items.

**SVD of** $A = U\Sigma V^T$

Let $r$ of $\sigma_i$s are $> 0$

Let $\sigma_1 \geq \ldots \geq \sigma_r > 0$

$A$ can be expressed as $A = \sum_{i=1}^{r} \sigma_i u_i v_i^T = \sigma_1 u_1 v_1^T + \ldots + \sigma_r u_r v_r^T$

Total space required to store $A$ is $r(m + n + 1)$. If rank of $A$ is small, it is better to store $u_1, \ldots, u_r, v_1, \ldots, v_r, \sigma_1, \ldots, \sigma_r$, rather than whole of $A$. 
Energy of $A = U\Sigma V^T$ is given by $E = \sum_{i=1}^{r} \sigma_i^2$

Define $E' = 0.99E$, and let $j \leq r$ be the maximum index such that $\sum_{i=1}^{j} \sigma_i^2 \leq E'$

Approximate $A$ by $\sum_{i=1}^{j} \sigma_i u_i v_i^T$

How many cells we need to store in this representation?

1. First $j$ columns of $U$,
2. $j$ diagonal entries of $\Sigma$, and
3. $j$ rows of $V^T$.

Total Space = $j^2 + j(m + n)$ cells
For our example, dimension of $A_{m \times n}$ are $m = 10^8$ users and $n = 10^5$ items.

If $j = 20$, then we need to store

$$j^2 + j(m + n) = 20^2 + 20 \times (10^8 + 10^5) \approx 5,005,000 \text{ cells}$$

This number is only .02% of $10^{13}$.
Let SVD of $A$ be

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{30} & -1/\sqrt{5} \\ 1/\sqrt{30} & 2/\sqrt{5} \\ 5/\sqrt{30} & 0 \end{bmatrix} \begin{bmatrix} \sqrt{6} & 0 \\ 0 & 1 \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$$

In terms of Rank 1 Components:

$$A = \sqrt{6} \begin{bmatrix} 2/\sqrt{30} \\ 1/\sqrt{30} \\ 5/\sqrt{30} \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}^T + \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}^T$$

Energy of $A$: $\mathcal{E}(A) = \sqrt{6^2} + 1^2 = 7$

Possible $\frac{6}{7}$-Energy approximation of $A$ is given by

$$A \approx \sqrt{6} \begin{bmatrix} 2/\sqrt{30} \\ 1/\sqrt{30} \\ 5/\sqrt{30} \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} \end{bmatrix}^T$$
An Application
Interpreting $U$, $\Sigma$, and $V$

Utility Matrix $M$ as SVD $M = U\Sigma V^T$

$$
\begin{bmatrix}
1 & 1 & 1 & 0 & 0 \\
3 & 3 & 3 & 0 & 0 \\
4 & 4 & 4 & 0 & 0 \\
5 & 5 & 5 & 0 & 0 \\
0 & 2 & 0 & 4 & 4 \\
0 & 0 & 0 & 5 & 5 \\
0 & 1 & 0 & 2 & 2 \\
\end{bmatrix}
= 
\begin{bmatrix}
.13 & -.02 & .01 \\
.41 & -.07 & .03 \\
.55 & -.1 & .04 \\
.68 & -.11 & .05 \\
.15 & .59 & -.65 \\
.07 & .73 & .67 \\
.07 & .29 & -.32 \\
\end{bmatrix}
\begin{bmatrix}
12.5 & 0 & 0 \\
0 & 9.5 & 0 \\
0 & 0 & 1.35 \\
\end{bmatrix}
\begin{bmatrix}
.56 & .59 & .56 & .09 & .09 \\
-.12 & .02 & -.12 & .69 & .69 \\
.40 & -.8 & .40 & .09 & .09 \\
\end{bmatrix}
$$

1. 3 concepts ($= rank$)
2. $U$ maps users to concepts
3. $V$ maps items to concepts
4. $\Sigma$ gives strength of each concept
Rank-2 Approximation

\[
\begin{bmatrix}
0.13 & -0.02 \\
0.41 & -0.07 \\
0.55 & -0.1 \\
0.68 & -0.11 \\
0.15 & 0.59 \\
0.07 & 0.73 \\
0.07 & 0.29 \\
\end{bmatrix}
\begin{bmatrix}
12.5 & 0 \\
0 & 9.5 \\
\end{bmatrix}
\begin{bmatrix}
0.56 & 0.59 & 0.56 & 0.09 & 0.09 \\
-0.12 & 0.02 & -0.12 & 0.69 & 0.69 \\
\end{bmatrix}
\]

\[\text{% Loss in Energy} = \frac{1.35^2}{12.5^2+9.5^2+1.35^2} < 1\%\]
Mapping Users to Concept Space

Consider the utility matrix $M$ and its SVD.

$$M = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 3 & 3 & 3 & 0 & 0 \\ 4 & 4 & 4 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 2 & 0 & 4 & 4 \\ 0 & 0 & 0 & 5 & 5 \\ 0 & 1 & 0 & 2 & 2 \end{bmatrix} \approx \begin{bmatrix} .13 & -.02 \\ .41 & -.07 \\ .55 & -1 \\ .68 & -.11 \\ .15 & .59 \\ .07 & .73 \end{bmatrix} \begin{bmatrix} 12.5 & 0 \\ 0 & 9.5 \end{bmatrix} \begin{bmatrix} .56 & .59 & .56 & .09 & .09 \\ -.12 & .02 & -.12 & .69 & .69 \end{bmatrix}$$

$MV$ gives mapping of each user in concept space:

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 3 & 3 & 3 & 0 & 0 \\ 4 & 4 & 4 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 2 & 0 & 4 & 4 \\ 0 & 0 & 0 & 5 & 5 \\ 0 & 1 & 0 & 2 & 2 \end{bmatrix} \begin{bmatrix} .56 & -.12 \\ .59 & .02 \\ .56 & -.12 \\ .09 & .69 \\ .09 & .69 \end{bmatrix} = \begin{bmatrix} 1.71 & -.22 \\ 5.13 & -.66 \\ 6.84 & -.88 \\ 8.55 & -1.1 \\ 1.9 & 5.56 \\ .9 & 6.9 \\ .96 & 2.78 \end{bmatrix}$$
Mapping Users to Items

Suppose we want to recommend items to a new user $q$ with the following row in the utility matrix $[4 \ 0 \ 0 \ 0 \ 0]$

1. Map $q$ to concept space:

$$qV = \begin{bmatrix} 3 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} .56 & -.12 \\ .59 & .02 \\ .56 & -.12 \\ .09 & .69 \\ .09 & .69 \end{bmatrix} = \begin{bmatrix} 1.68 & -.36 \end{bmatrix}$$

2. Map the vector $qV$ to the Items space by multiplying by $V^T$ as vector $V$ captures the connection between items and concepts.

$$\begin{bmatrix} 1.68 & -.36 \end{bmatrix} \begin{bmatrix} .56 & .59 & .56 & .09 & .09 \\ -.12 & .02 & -.12 & .69 & .69 \end{bmatrix} = \begin{bmatrix} .98 & .98 & .98 & -.1 & -.1 \end{bmatrix}$$
Mapping Users to Items (Contd.)

Suppose we want to recommend items to user $q'$ with the following row in the utility matrix $\begin{bmatrix} 0 & 0 & 0 & 4 & 0 \end{bmatrix}$

1. $q'V = \begin{bmatrix} 0 & 0 & 0 & 4 & 0 \end{bmatrix} \begin{bmatrix} .56 & -.12 \\ .59 & .02 \\ .09 & .69 \\ .09 & .69 \end{bmatrix} = \begin{bmatrix} .36 \\ 2.76 \end{bmatrix}$

2. Map $q'V$ to the Items space by multiplying by $V^T$

$\begin{bmatrix} .36 & 2.76 \end{bmatrix} \begin{bmatrix} .56 & .59 & .56 & .09 & .09 \\ -.12 & .02 & -.12 & .69 & .69 \end{bmatrix} = \begin{bmatrix} -.12 & .26 & -.12 & 1.93 & 1.93 \end{bmatrix}$
Suppose we want to recommend items to user $q''$ with the following row in the utility matrix:

$$q'' = \begin{bmatrix} 0 & 0 & 4 & 4 & 0 \end{bmatrix}$$

1. $q''V = \begin{bmatrix} 0 & 0 & 4 & 4 & 0 \end{bmatrix} \begin{bmatrix} .56 & -.12 \\ .59 & .02 \\ .09 & .69 \\ .09 & .69 \end{bmatrix} = \begin{bmatrix} 2.6 \\ 2.28 \end{bmatrix}$

2. Map $q''V$ to the Items space by multiplying by $V^T$

$$\begin{bmatrix} 2.6 & 2.28 \end{bmatrix} \begin{bmatrix} .56 & .59 & .56 & .09 & .09 \\ -.12 & .02 & -.12 & .69 & .69 \end{bmatrix} = \begin{bmatrix} 1.18 & 1.57 & 1.18 & 1.8 & 1.8 \end{bmatrix}$$
Correctness
### Frobenius Norm

Let $A$ be a matrix of real numbers. Its Frobenius Norm $\|A\|_F$ is defined as

$$\|A\|_F = \sqrt{\sum_{i,j} A[i,j]^2}$$

### Frobenius Norm via SVD of $A$

For a rank $r$ matrix $A$ with its singular-value decomposition $A = U\Sigma V^T$, its Frobenius norm is

$$\|A\|_F^2 = \Sigma_{11}^2 + \cdots + \Sigma_{rr}^2.$$
Why is $\|A\|_F^2 = \sum_{11}^2 + \cdots + \sum_{rr}^2$?

Let SVD of $A = PQR$ (Note: $P = U, Q = \Sigma, R = V^T$.)

Now $A_{ij} = \sum_k \sum_l p_{ik} q_{kl} r_{lj}$.

\[
\|A\|_F^2 = \sum_i \sum_j A_{ij}^2
= \sum_i \sum_j (\sum_k \sum_l p_{ik} q_{kl} r_{lj})^2
= \sum_i \sum_j \sum_k \sum_l \sum_m \sum_n p_{ik} q_{kl} r_{lj} p_{im} q_{mn} r_{nj}
\]

Now use the fact that $q_{ab} = 0$ for $a \neq b$ and the dot-product of any two columns of $p$ is 0 due to orthonormality of $P = U$. Similarly, the dot-product of any two rows of $R = V^T$ is 0. This allows us to show

\[
\|A\|_F^2 = \sum_k q_{kk}^2 = \sum_{11}^2 + \cdots + \sum_{rr}^2
\]
Let $A$ and $A'$ be two matrices of real numbers of same dimensions.

**Error in approximating $A$ by $A'$**

The error in approximating $A$ by $A'$ is defined as the Frobenius Norm of

$$||A - A'||_F = \sqrt{\sum_{i,j} (A[i,j] - A'[i,j])^2}$$

Let $A = U \Sigma V^T$ be a $m \times n$ matrix of real numbers of rank $r$. Let $1 \leq r' < r$. Define a $r \times r$ diagonal matrix $\Sigma'$ as follows:

For $i = 1$ to $r'$, $\Sigma'_{ii} = \Sigma_{ii}$ and all other entries of $\Sigma'$ are 0. Let $A' = U \Sigma' V^T$.

**Claim:** $A'$ is the best rank $r' < r$ approximation of $A$, i.e., for any rank $r'$ $m \times n$ matrix $B$, $||A - A'||_F \leq ||A - B||_F$. 


Given \( A = U \Sigma V^T \) and \( A' = U \Sigma' V^T \), \( A - A' = U(\Sigma - \Sigma')V^T \).

Thus, \( \|A - A'\|_F^2 = \Sigma_{r'+1,r'+1}^2 + \cdots + \Sigma_{rr}^2 \).

Note that the elements \( \Sigma_{r'+1,r'+1}, \ldots, \Sigma_{rr} \) were set to 0 in \( \Sigma \) to obtain \( A' \). These are the lowest energy terms in \( A \).

**Best low rank approximation of \( A \)**

For a rank \( r \) matrix \( A \) with its SVD \( A = U \Sigma V^T \), its best rank \( r' < r \) approximation is obtained by the matrix \( A' \) where \( A' = U \Sigma' V^T \) and \( \Sigma' \) is obtained from \( \Sigma \) by setting its \( r - r' \) smallest diagonal entries to 0.

