Properties of Orthonormal Vectors

Please refer to the lecture 10A notes online for proofs of the following properties.

(a) **Definition: Orthonormal**

A set of vectors \( \{\vec{v}_1, \ldots, \vec{v}_n\} \) is **orthonormal** if all the vectors are mutually orthogonal to each other and all are of unit length. That is:

- **Orthogonal:** For all pairs of vectors \( \vec{v}_i, \vec{v}_j \) where \( i \neq j \), \( \langle \vec{v}_i, \vec{v}_j \rangle = 0 \). For real vectors, this means \( \vec{v}_i^T \vec{v}_j = 0 \).
- **Normalized:** For all \( i \), \( \|\vec{v}_i\| = 1 \). (This implies that \( \|\vec{v}_i\| = \langle \vec{v}_i, \vec{v}_i \rangle = 1 \).)

(b) Any set of orthogonal (and by extension orthonormal) vectors are linearly independent.

(c) Any set of orthogonal (and orthonormal) vectors \( \{\vec{p}_1, \vec{p}_2, \cdots, \vec{p}_n\} \) form a basis in \( \mathbb{R}^n \).

Let us consider the orthonormal set of vectors \( \{\vec{q}_1, \vec{q}_2, \cdots, \vec{q}_n\} \), which form a basis in \( \mathbb{R}^n \). For any vector \( \vec{v} \) represented in this basis, we have

\[
\vec{v} = \alpha_1 \vec{q}_1 + \alpha_2 \vec{q}_2 + \cdots + \alpha_n \vec{q}_n.
\]

Hence, \( \alpha_i = \vec{q}_i^T \vec{v} \). More compactly, we can write

\[
\alpha = \begin{bmatrix}
\alpha_1 \\
\alpha_2 \\
\vdots \\
\alpha_n \\
\end{bmatrix} = \begin{bmatrix}
\leftarrow & \vec{q}_1^T & \rightarrow \\
\leftarrow & \vec{q}_2^T & \rightarrow \\
\vdots & \ddots & \vdots \\
\leftarrow & \vec{q}_n^T & \rightarrow \\
\end{bmatrix} \begin{bmatrix}
\vec{v} = Q^T \vec{v}
\end{bmatrix}
\]

\( \vec{v}, \alpha, \{\vec{q}_1, \vec{q}_2, \cdots, \vec{q}_n\}, \) and \( Q \) have the following properties.

(d) The vector of projections, \( \alpha \), has the same norm as the original vector \( \vec{v} \).

(e) Given that the columns of \( Q \) are orthonormal, the rows of \( Q \) are also orthonormal.

(f) Given a set of orthonormal vectors \( \{\vec{q}_1, \vec{q}_2, \cdots, \vec{q}_n\} \), for any \( r \leq n \), we have

\[
\|\vec{q}_1\|^2 + \|\vec{q}_2\|^2 + \cdots + \|\vec{q}_r\|^2 = \sum_{i=1}^{r} \sum_{j=1}^{n} q_{ji}^2 = r.
\]

Here, \( q_{ji} \) is the \( j^{th} \) element of \( \vec{q}_i \).
Properties of Real Symmetric Matrices

Let \( T \) be a symmetric matrix on \( \mathbb{R}^{n \times n} \). Then,

(a) The eigenvalues of \( T \) are real.
(b) A set of real eigenvectors \( \{ \vec{p}_1, \vec{p}_2, \ldots, \vec{p}_n \} \) of \( T \) can be found.
(c) The eigenvectors \( \{ \vec{p}_1, \vec{p}_2, \ldots, \vec{p}_n \} \) of \( T \) are orthogonal.
(d) A real orthonormal set of eigenvectors of \( T \) can be found by normalizing each vector in the set \( \{ \vec{p}_1, \vec{p}_2, \ldots, \vec{p}_n \} \).

That is, there exists \( n \) real eigenvalues and \( n \) real linearly independent eigenvectors of \( T \) that form a basis for \( \mathbb{R}^n \). Furthermore, these eigenvectors can be normalized to make an orthonormal basis.

Outer Products

We can define an outer product between two vectors \( \vec{x} \in \mathbb{R}^n \) and \( \vec{y} \in \mathbb{R}^m \) as follows:

\[
\vec{x} \vec{y}^T = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_{n \times 1} \begin{bmatrix} y_1 & y_2 & \cdots & y_m \end{bmatrix}_{1 \times m} = \begin{bmatrix} x_1 y_1 & x_1 y_2 & \cdots & x_1 y_m \\ x_2 y_1 & x_2 y_2 & \cdots & x_2 y_m \\ \vdots & \vdots & \ddots & \vdots \\ x_n y_1 & x_n y_2 & \cdots & x_n y_m \end{bmatrix}_{n \times m}
\]

Hence, the outer product gives us an \( n \times m \) rank-1 matrix. \textit{Note: Do not confuse the outer product} \( \vec{x} \vec{y}^T \) \textit{with the inner product given by} \( \vec{x}^T \vec{y} \).

We can represent the matrix multiplication as a sum of outer products as follows,

\[
XY^T = \begin{bmatrix} \vec{x}_1 \\ \vec{x}_2 \\ \vdots \\ \vec{x}_r \end{bmatrix}_{n \times r} \begin{bmatrix} \vec{y}_1 \vec{y}_1^T & \vec{y}_2 \vec{y}_2^T & \cdots & \vec{y}_r \vec{y}_r^T \end{bmatrix}_{r \times m} = \vec{x}_1 \vec{y}_1^T + \vec{x}_2 \vec{y}_2^T + \cdots + \vec{x}_r \vec{y}_r^T = \sum_{i=1}^r \vec{x}_i \vec{y}_i^T.
\]

Where \( X \) and \( Y^T \) are any \( n \times r \) and \( r \times m \) matrices.

Singular Value Decomposition

The SVD is a useful way to characterize a matrix. Let \( A \) be a matrix from \( \mathbb{R}^n \) to \( \mathbb{R}^m \) (or \( A \in \mathbb{R}^{m \times n} \)) of rank \( r \). It can be decomposed into a sum of \( r \) rank-1 matrices:

\[
A = \sum_{i=1}^r \sigma_i \vec{u}_i \vec{v}_i^T
\]

where

- \( \vec{u}_1, \ldots, \vec{u}_r \) are orthonormal vectors in \( \mathbb{R}^m \); \( \vec{v}_1, \ldots, \vec{v}_r \) are orthonormal vectors in \( \mathbb{R}^n \).
• The singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ are always real and positive.

We can rewrite the decomposition in the form

$$A = U_1 S V_1^T,$$

where

• $U_1$ is an $[m \times r]$ matrix whose columns consist of $\vec{u}_1, \ldots, \vec{u}_r$ (orthonormal vectors in $\mathbb{R}^m$). Consequently,

$$U_1^T U_1 = I_{r \times r}$$

• $V_1$ is an $[n \times r]$ matrix whose columns consist of $\vec{v}_1, \ldots, \vec{v}_r$ (orthonormal vectors in $\mathbb{R}^n$). Consequently,

$$V_1^T V_1 = I_{r \times r}$$

• $U_1$ characterizes the column space of $A$ and $V_1$ characterizes the row space of $A$.

• $S$ is an $[r \times r]$ matrix whose diagonal entries are the singular values of $A$ arranged in descending order. The singular values are the square roots of the nonzero eigenvalues of $A^T A$ (or, identically, $AA^T$).

The full matrix form of SVD is

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

where $U^T U = I_{m \times m}, V^T V = I_{n \times n}, \Sigma \in \mathbb{R}^{m \times n}$, which contains $S$ and elsewhere zero.

Questions

1. Eigenvectors are Orthogonal

Prove the following: For any symmetric matrix $A$, any two eigenvectors corresponding to distinct eigenvalues of $A$ are orthogonal.

*Hint:* Use the definition of an eigenvalue to show that $\lambda_1 (\vec{v}_1^T \vec{v}_2) = \lambda_2 (\vec{v}_1^T \vec{v}_2)$.

*Answer:* Let $\lambda_1, \lambda_2$ be eigenvalues of $A$ with corresponding eigenvectors $\vec{v}_1, \vec{v}_2$. Since $A$ is symmetric, we have

$$\lambda_1 \vec{v}_1^T \vec{v}_2 = (\lambda_1 \vec{v}_1)^T \vec{v}_2 = (A \vec{v}_1)^T \vec{v}_2 = \vec{v}_1^T A^T \vec{v}_2 = \vec{v}_1^T Av_2 = \vec{v}_1^T (\lambda_2 \vec{v}_2) = \lambda_2 \vec{v}_1^T \vec{v}_2$$

This implies that,

$$(\lambda_1 - \lambda_2) \vec{v}_1^T \vec{v}_2 = 0$$

The only way this equation can be satisfied when $\lambda_1 \neq \lambda_2$ is for $\vec{v}_1^T \vec{v}_2$ to be zero. Therefore $\vec{v}_1$ and $\vec{v}_2$ must be orthogonal.

2. Frobenius Norm

In this problem we will investigate the properties of the Frobenius norm.
Much like the norm of a vector $\vec{x} \in \mathbb{R}^N$ is $||x|| = \sqrt{\sum_{i=1}^{N} x_i^2}$, the Frobenius norm of a matrix $A \in \mathbb{R}^{N \times N}$ is defined as

$$||A||_F = \sqrt{\sum_{i=1}^{N} \sum_{j=1}^{N} |A_{ij}|^2}.$$

Note that matrices have other types of norms as well.

(a) With the above definitions, show that,

$$||A||_F = \sqrt{\text{Tr}\{A^T A\}}.$$

*Note:* The trace of a matrix is the sum of its diagonal entries. For example, let $A \in \mathbb{R}^{N \times N}$, then,

$$\text{Tr}\{A\} = \sum_{i=1}^{N} A_{ii}$$

*Answer:*

$$\text{Tr}\{A^T A\} = \sum_{i=1}^{N} (A^T A)_{ii}$$

$$= \sum_{i=1}^{N} \left( \sum_{j=1}^{N} (A^T)_{ij} A_{ji} \right)$$

$$= \sum_{i=1}^{N} \left( \sum_{j=1}^{N} A_{ji} A_{ji} \right)$$

$$= \sum_{i=1}^{N} \sum_{j=1}^{N} (A_{ji})^2$$

$$= ||A||_F^2$$

(b) Show that if $U$ and $V$ are orthonormal matrices, then

$$||UA||_F = ||AV||_F = ||A||_F.$$

*Answer:*

$$||UA||_F = \sqrt{\text{Tr}\{(UA)^T (UA)\}} = \sqrt{\text{Tr}\{A^T U^T U A\}} = \sqrt{\text{Tr}\{A^T A\}} = ||A||_F$$

To show the second set of equality, we must note that $\text{Tr}\{A^T A\} = \text{Tr}\{AA^T\}$. Hence,

$$||AV||_F = \sqrt{\text{Tr}\{(AV)(AV)^T\}} = \sqrt{\text{Tr}\{AVV^T A^T\}} = \sqrt{\text{Tr}\{AA^T\}} = ||A||_F$$
(c) Show that \( \|A\|_F = \sqrt{\sum_{i=1}^{N} \sigma_i^2} \), where \( \sigma_1, \ldots, \sigma_N \) are the singular values of \( A \).

**Answer:**

\[
\|A\|_F = \|USV^T\|_F = \|SV^T\|_F = \|\Sigma\|_F \\
= \sqrt{\text{Tr}\{\Sigma^T \Sigma\}} = \sqrt{\sum_{i=1}^{N} \sigma_i^2}
\]

3. **SVD Short Questions**

Assume we have the compact form of the SVD of

\[ A = U_1SV_1^T = \sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^T. \]

(a) Compute \( AV_1V_1^T \).

**Answer:**

Recall that \( V_1 \) is an orthogonal matrix, so it has orthonormal columns, giving it the property \( V_1^TV_1 = I \). Hence we can write:

\[ AV_1V_1^T = U_1SV_1^TV_1V_1^T = U_1SV_1^T = A \]

(b) What is the subspace that spans the column space of \( A \)?

**Answer:**

Given a vector \( \vec{x} \), the column space of \( A \) is also the same as the space of all possible \( A\vec{x} \).

\[ A\vec{x} = \sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^T \vec{x} \]

But, \( \vec{v}_i^T \vec{x} \) is a scalar, hence,

\[ A\vec{x} = \sum_{i=1}^{r} (\sigma_i \vec{v}_i^T \vec{x}) \vec{u}_i \]

From that decomposition, we can see that \( A\vec{x} \) is a linear combination of \( \vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r \). Hence the span of columns of \( A \) is the subspace spanned by the columns of \( U_1 \).

4. **(Optional) Symmetric Matrix Properties**

Given the symmetric matrix

\[ T = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \]

find the eigenvalues and eigenvectors of \( T \).

Show that the eigenvalues are real and the eigenvectors are orthonormal.

**Contributors:**

Discussion 10A ©UCB EECS 16B, Spring 2019. All Rights Reserved. This may not be publicly shared without explicit permission.
• Aditya Arun.
• Siddharth Iyer.
• Jane Liang.
• Regina Eckert.