Cybersecurity Assignment 2

Mayank Sharma, 160392

August, 31, 2019

1

To Prove:

$$\langle \mathbf{v}, \mathbf{w} \rangle = \langle L\mathbf{v}, L\mathbf{w} \rangle$$

We're given that $\mathbf{L}$ is an isometry and that $\mathbf{L} : V \to W$ is a linear transformation, and we need to show that that the above relation holds (i.e. it preseves the inner products)

Now, we know that $\langle \mathbf{x}, \mathbf{x} \rangle = \|\mathbf{x}\|^2$ and that $\|\mathbf{x}+\mathbf{y}\|_2^2 - \|\mathbf{x}-\mathbf{y}\|_2^2 = 4\langle \mathbf{x}, \mathbf{x} \rangle$
So, we have,

$$\|L\mathbf{v} = L\mathbf{w}\|_2^2 - \|L\mathbf{v} - L\mathbf{w}\|_2^2 = 4\langle \mathbf{v}, \mathbf{w} \rangle \tag{1}$$

Since $\mathbf{L}$ is an isometry,

$$\implies \|\mathbf{v}+\mathbf{w}\|^2 - \|\mathbf{v}-\mathbf{w}\|^2 = 4\langle \mathbf{v}, \mathbf{w} \rangle$$
$$\implies \langle \mathbf{v}+\mathbf{w}, \mathbf{v}+\mathbf{w} \rangle - \langle \mathbf{v}-\mathbf{w}, \mathbf{v}-\mathbf{w} \rangle = 4\langle L\mathbf{v}, L\mathbf{w} \rangle$$

Now, from the additivity of inner products, we get that

$$\implies \langle \mathbf{v}+\mathbf{w}, \mathbf{v}+\mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{v} \rangle + \langle \mathbf{w}, \mathbf{w} \rangle + 2\langle \mathbf{v}, \mathbf{w} \rangle \tag{2}$$
$$\implies \langle \mathbf{v}-\mathbf{w}, \mathbf{v}-\mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{v} \rangle + \langle \mathbf{w}, \mathbf{w} \rangle - 2\langle \mathbf{v}, \mathbf{w} \rangle \tag{3}$$

So, using (2) and (3) in (1), we get that

$$\implies (2\langle \mathbf{v}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle + \langle \mathbf{w}, \mathbf{w} \rangle) - (\langle L\mathbf{v}, L\mathbf{v} \rangle + \langle \mathbf{w}, \mathbf{w} \rangle 2\langle \mathbf{v}, \mathbf{w} \rangle) = 4\langle L\mathbf{v}, L\mathbf{w} \rangle$$
$$\implies 4\langle \mathbf{v}, \mathbf{w} \rangle = 4\langle L\mathbf{v}, L\mathbf{w} \rangle$$
$$\implies \langle \mathbf{v}, \mathbf{w} \rangle = \langle L\mathbf{v}, L\mathbf{w} \rangle$$

Therefore, isometry $\mathbf{L}$ preseves the inner product of two vectors.

Geometrically this result means that an operation of an isometry $\mathbf{L}$ on two vectors $\mathbf{v}, \mathbf{w}$ does not change the magnitude of norms between the two vectors. It's like moving a whole set of vectors by some distance without changing the distances between the vectors (i.e. their relative distance doesn't change).

2

We're given that that $L : V \to W$ is a linear transformation between two normed vector spaces. We need to show that $Row(L) = (Kernel(L))^{\perp}$. The $kernel(L)$ is defined as

$$kernel(L) = \{\mathbf{v} \in V : L\mathbf{v} = 0\}$$

whereas the rowspace of L is the collection of linearly independent vectors that the rows of L. Consider the span of liearly independent rows of L (Here, $\mathbf{r}_i^T$'s are the linearly independent rows of L and c_i is any constant):

$$\mathbf{x} = \sum_{i=0}^{m} c_i \mathbf{r}_i$$

$$L\mathbf{v} = \begin{bmatrix} \mathbf{r}_1^T \\ \mathbf{r}_2^T \\ \cdots \\ \mathbf{r}_m^T \end{bmatrix} \mathbf{v} = \begin{bmatrix} \mathbf{r}_1^T.\mathbf{v} \\ \mathbf{r}_2^T.\mathbf{v} \\ \cdots \\ \mathbf{r}_m^T.\mathbf{v} \end{bmatrix} = \mathbf{0}$$

This means the inner product of $\mathbf{v}$ with each linearly independent row vector of L is 0, i.e. $\mathbf{v}$ is orthogonal to every row vector of L. Hence, $\mathbf{v}$ is orthogonal to any vector $\mathbf{x} \in Row(L)$.

Hence,

$$Row(L) = (Kernel(L))^{\perp}$$

3

Let us take a $m \times n$ matrix A. Let it's row rank be **r**.
Let it's row span be comprised of linearly independent vectors $x_1, x_2 .... x_r$.
Let's take $v = \sum_{i=1}^{i=r} c_i x_i$.
Now.

$$Av = 0$$
$$\implies A \sum_{i=1}^{i=r} c_i x_i = \sum_{i=1}^{i=r} c_i A x_i = 0$$

Since, $v \in Row(A)$ and $Av = 0$, i.e. v is orthogonal to every row of A and therefore orthogonal to itself, therefore

$$v = 0.$$
$$\implies \sum_{i=1}^{i=r} c_i x_i = 0$$
$$\implies c_i = 0 \quad \forall i$$

Now, since x_is are linearly independent, we get

$$\sum_{i=1}^{i=r} c_i A x_i = 0$$

and

$$c_i = 0 \quad \forall i$$

$\implies$ Ax_is are linearly independent. Also, Ax_is are also vectors of column space of A. Hence, column space of A contains at least r linearly independent vectors, i.e.

$$dim(Col(A)) \geq r$$

$$\implies dim(Col(A)) \geq dim(Row(A))$$

Now,

$$dim(Col(A^T)) \geq dim(Row(A^T))$$
$$\implies dim(Row(A)) \geq dim(Col(A))$$

as row rank of A is the column rank of A^T.
Hence, $dim(Row(A)) = dim(Col(A))$. Q.E.D

4

The SVD is defined as follows:
Let $\mathbf{A}$ be an $(m \times n)$ matrix with $m \geq n$. Then there exist orthogonal matrices $\mathbf{U}_{m\times m}$ and $\mathbf{V}_{n\times n}$ and a diagonal matrix $\mathbf{S}_{(m\times n)} = \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$ with $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_n \geq 0$, such that

$$\mathbf{A} = \mathbf{USV}^T$$

To Prove:
1. $\mathbf{U}$ is a matrix whose columns are the eigen vectors of $\mathbf{AA}^T$ and $\mathbf{V}$ is a matrix whose columns are eigen vectors of $\mathbf{A}^T\mathbf{A}$.
2. $\mathbf{AA}^T$ and $\mathbf{A}^T\mathbf{A}$ are symmetric matrices We firstly prove the second part. Note that,

$$(\mathbf{AA}^T)^T = (\mathbf{A}^T)^T\mathbf{A}^T = \mathbf{AA}^T$$

Hence, $\mathbf{AA}^T$ is a symmetric matrix. Similarly for $\mathbf{A}^T\mathbf{A}$,

$$(\mathbf{A}^T\mathbf{A})^T = \mathbf{A}^T(\mathbf{A}^T)^T = \mathbf{A}^T\mathbf{A}$$

Hence, $\mathbf{A}^T\mathbf{A}$ is also symmetric. We now prove the first part. From the SVD definition we have:

$$\mathbf{S}^T = \mathbf{S} \tag{1}$$
$$\mathbf{UU}^T = \mathbf{U}^T\mathbf{U} = I \tag{2}$$
$$\mathbf{VV}^T = \mathbf{V}^T\mathbf{V} = I \tag{3}$$

Now, using the above three,

$$\begin{aligned}\mathbf{AA}^T &= (\mathbf{USV}^T)(\mathbf{USV}^T)^T \\ &= \mathbf{USV}^T(\mathbf{VSU}^T) \\ &= \mathbf{US}^2\mathbf{U}^T\end{aligned}$$

$$\implies \mathbf{AA}^T\mathbf{U} = \mathbf{US}^2 \tag{4}$$

$$\mathbf{U} = \begin{bmatrix} \mathbf{u}_1, & \mathbf{u}_2, & \ldots & \mathbf{u}_i, & \ldots & \mathbf{u}_m \end{bmatrix}$$
$$\mathbf{S}^2 = \begin{bmatrix} \sigma_1^2, & 0, & \ldots, & 0 \\ \ldots, & \sigma_2^2, & \ldots, & 0 \\ \ldots, & \ldots, & \ldots, & \ldots \\ 0, & 0, & \ldots, & \sigma_m^2 \end{bmatrix}$$

Hence, from (4), we can say that

$$(\mathbf{AA}^T)\mathbf{u}_i = (\sigma_i^2)\mathbf{u}_i$$

Hence, the column vectors of $\mathbf{U}$ are the orthonormal eigen-vectors of matrix $\mathbf{AA}^T$, such that $\mathbf{U}$ is the matrix containing all the eigen vectors of $(\mathbf{AA}^T)$ and $\mathbf{S}^2$ contains all the eigen values.

Similarly, we can consider the matrix $\mathbf{A}^T\mathbf{A}$

$$\begin{aligned}
\mathbf{A}^T\mathbf{A} &= (\mathbf{U}\mathbf{S}\mathbf{V}^T)^T(\mathbf{U}\mathbf{S}\mathbf{V}^T) \\
&= \mathbf{V}\mathbf{S}\mathbf{U}^T(\mathbf{U}\mathbf{S}\mathbf{V}^T) \\
&= \mathbf{V}\mathbf{S}^2\mathbf{V}^T
\end{aligned}$$

$$\implies \mathbf{A}^T\mathbf{A}\mathbf{V} = \mathbf{V}\mathbf{S}^2$$

Hence, column vectors of V are the orthonormal eigenvectors of matrix $\mathbf{A}^T\mathbf{A}$ and the matrix $\mathbf{V}$ contains all the eigenvectors and $\mathbf{S}^2$ contains all the eigen values.

Q.E.D

5

We're given that $\mathbf{u}_i$ is an m-dimensional vector and $\mathbf{v}_i$ is an n-dimensional vector. Now, the product of these two vectors along the common axis (i.e. 1) will be an $m \times n$ matrix (say $\mathbf{A}$).

$$\mathbf{A}_{m\times n} = \sum_i (\sigma_i \mathbf{u}_i)(\mathbf{v}_i^T) \tag{1}$$

Now, from out knowledge about SVD, the matrix $\mathbf{A}$ can also be written as the summation of the outer product of eigen vectors of the matrices $\mathbf{U}$ and $\mathbf{V}$, where these $\mathbf{U}$ and $\mathbf{V}$ matrices satisfy the below relationship.

$$\mathbf{A} = \mathbf{USV}^T$$

such that the matrix $\mathbf{S}$ contains all the singular values of the matrix $\mathbf{A}$. Aoreover, the orthogonal matrices $\mathbf{U}_{m\times m}$ and $\mathbf{V}_{n\times n}$ and the diagonal matrix $\mathbf{S}_{(m\times n)} = \text{diag}(\sigma_1, \ldots, \sigma_n)$ such that $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_n \geq 0$.

The significance of the above decomposition of matrix $\mathbf{A}$ is that we get the singular values in a decreasing way. So, what we can do is in the summation given by (1), we can ignore the terms corresponding to the smaller singular values and take the first k signicant terms.

This allows us to approximate the matrix $\mathbf{A}$ (having a rank r) with another matrix $\mathbf{B}$ having a rank of k (such that $k < r$).

The advantage that this method provides is that we can significantly reduce the computation depending upon the precision of the output that we need. If we need a higher precision, we can ignore a lesser number of terms in the summation given by (1), and vice versa. It's particularly useful in signal processing where there can be many smaller components which don't play a signicant part and can easily be ignored.

6

Given that $\mathbf{S}$ is a symmetric matrix (i.e. $\mathbf{S} = \mathbf{S}^T$) with its eigenvectors $(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n)$ and the corresponding eigen values $(\lambda_1, \lambda_2, ..., \lambda_n)$. Now, by the property of symmetric matrices, these eigenvectors are orthogonal.

To Prove:

$$\mathbf{S} = \sum_{i=1}^{n} \lambda_i \mathbf{v}_i \mathbf{v}_i^T$$

Now, we know that we can decompose $\mathbf{S}$ as $\mathbf{S} = \mathbf{PDP}^{-1}$, where $\mathbf{P}$ is a matrix whose columns are the eigenvectors of $\mathbf{S}$ and $\mathbf{D}$ is a diagonal matrix whose diagonal values are eigenvalues of $\mathbf{S}$.

$$\mathbf{P} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix}$$
$$\mathbf{D} = diag(\ \lambda_1, \quad \lambda_2, \quad \lambda_3\ )$$

Since, $\mathbf{P}$ is an orthogonal matrix because all its column vectors are orthogonal, so we have $\mathbf{P}^{-1} = \mathbf{P}^T$.

$$\mathbf{S} = \mathbf{PDP}^{-1} = \mathbf{PDP}^T$$

$$\implies \mathbf{S} = \begin{bmatrix} \mathbf{v}_1, & \mathbf{v}_2, & \dots, & \mathbf{v}_n \end{bmatrix} \begin{bmatrix} \lambda_1, & 0, & \dots, & 0 \\ \dots, & \lambda_2, & \dots, & 0 \\ \dots, & \dots, & \dots, & \dots \\ 0, & 0, & \dots, & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \dots \\ \mathbf{v}_n^T \end{bmatrix}$$

Once we multiply the above three matrices, we get

$$\mathbf{S} = \sum_{i=1}^{n} \lambda_i \mathbf{v}_i \mathbf{v}_i^T$$

The significance is that this result can be used in reducing the dimensionality of some dataset, which can help us to reduce the number of computations done.

Q.E.D