EECS 16B Designing Information Devices and Systems II
Spring 2018 J. Roychowdhury and M. Maharbiz **Discussion 9A**

1 Inner and Outer Products

The **inner product** $\langle \cdot, \cdot \rangle$ on a vector space V over $\mathbb{R}$ is a function that takes in two vectors and outputs a scalar, such that $\langle \cdot, \cdot \rangle$ is symmetric, linear, and positive-definite.

- Symmetric: $\langle \vec{u}, \vec{v} \rangle = \langle \vec{v}, \vec{u} \rangle$
- Linear: $\langle c\vec{u}, \vec{v} + \vec{w} \rangle = c\langle \vec{u}, \vec{v} \rangle + c\langle \vec{u}, \vec{w} \rangle$
- Positive-definite: $\langle \vec{u}, \vec{u} \rangle \geq 0$ with $\langle \vec{u}, \vec{u} \rangle = 0$ if and only if $\vec{u} = \vec{0}$

For two vectors, $\vec{u}, \vec{v} \in \mathbb{R}^n$, we usually define their inner product $\langle \vec{u}, \vec{v} \rangle$ to be $\langle \vec{u}, \vec{v} \rangle = \vec{u}^T \vec{v}$.

We define the **norm**, or the magnitude, of a vector $\vec{v}$ to be $\|\vec{v}\| = \sqrt{\langle \vec{v}, \vec{v} \rangle} = \sqrt{\vec{v}^T \vec{v}}$. For any non-zero vector, we can *normalize*, i.e., set its magnitude to 1 while preserving its direction, by dividing the vector by its norm $\frac{\vec{v}}{\|\vec{v}\|}$.

The **angle** θ between two vectors $\vec{u}$ and $\vec{v}$ is given by the equation $\langle \vec{u}, \vec{v} \rangle = \|\vec{u}\|\|\vec{v}\| \cos\theta$. If $\vec{u}$ and $\vec{v}$ have a magnitude of 1, then this equation simplifies to $\cos\theta = \langle \vec{u}, \vec{v} \rangle$.

The **outer product** of two vectors $\vec{u}, \vec{v} \in \mathbb{R}^n$ is defined as $\vec{u}\vec{v}^T$.

$$\vec{u}\vec{v}^T = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} = \begin{bmatrix} u_1 v_1 & u_1 v_2 & \cdots & u_1 v_n \\ u_2 v_1 & u_2 v_2 & \cdots & u_2 v_n \\ \vdots & \vdots & \ddots & \vdots \\ u_n v_1 & u_n v_2 & \cdots & u_n v_n \end{bmatrix} = \begin{bmatrix} u_1 \vec{v}^T \\ u_2 \vec{v}^T \\ \vdots \\ u_n \vec{v}^T \end{bmatrix} = \begin{bmatrix} v_1 \vec{u} & v_2 \vec{u} & \cdots & v_n \vec{u} \end{bmatrix}$$

Notice that the outer product, unlike the inner product, is an $n \times n$ matrix of rank 1 because each row is a constant multiple of $\vec{v}^T$ (or equivalently, each column is a constant multiple of $\vec{u}$).

2 Orthogonality and Orthonormality

We know that the angle between two vectors is given by this equation $\langle \vec{u}, \vec{v} \rangle = \|\vec{u}\|\|\vec{v}\| \cos\theta$. Notice that if $\theta = \pm 90°$, the right hand side is 0.

Therefore, we define two vectors $\vec{u}$ and $\vec{v}$ to be **orthogonal** to each other if $\langle \vec{u}, \vec{v} \rangle = \vec{u}^T \vec{v} = 0$. A set of vectors is orthogonal if any two vectors in this set are orthogonal to each other.

Furthermore, we define two vectors $\vec{u}$ and $\vec{v}$ to be **orthonormal** to each other if they are orthogonal to each other and their norms are 1. A set of vectors is orthonormal if any two vectors in this set are orthogonal to each other and every vector has a norm of 1. In fact, for any two vectors $\vec{u}$ and $\vec{v}$ in an orthonormal set,

$$\langle \vec{u}, \vec{v} \rangle = \vec{u}^T \vec{v} = \begin{cases} 1, & \text{if } \vec{u} = \vec{v} \\ 0, & \text{otherwise} \end{cases}.$$

A **unitary** matrix is a square matrix whose columns are orthonormal.

$$U = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_n \end{bmatrix}, \qquad \vec{u}_i^T \vec{u}_j = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{otherwise} \end{cases}$$

Note that $U^T U = U U^T = I$, so the inverse of a unitary matrix is its transpose $U^{-1} = U^T$.

Since the columns of a unitary matrix are orthonormal vectors, we can interpret these matrices as "rotation" and "reflection" matrices of the standard axes. This also implies that $\|U\vec{v}\| = \|\vec{v}\|$ for any vector $\vec{v}$.

3 Orthonormal Bases

Assume that we are given an orthonormal basis $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$ and that we want to express an arbitrary vector $\vec{v}$ as a linear combination of the orthonormal basis vectors.

$$c_1\vec{u}_1 + c_2\vec{u}_2 + \cdots + c_n\vec{u}_n = \vec{v}$$

Let's first solve for c_1 by left-multiplying both sides with $\vec{u}_1^T$.

$$c_1\vec{u}_1^T\vec{u}_1 + c_2\vec{u}_1^T\vec{u}_2 + \cdots + c_n\vec{u}_1^T\vec{u}_n = \vec{u}_1^T\vec{v}$$

$$c_1\vec{u}_1^T\vec{u}_1 = \vec{u}_1^T\vec{v} \implies c_1 = \frac{\vec{u}_1^T\vec{v}}{\vec{u}_1^T\vec{u}} = \vec{u}_1^T\vec{v}$$

Generalizing this, we can find all c_i by calculating $c_i = \vec{u}_i^T\vec{v}$.

Notice that this is equivalent to finding the scalar projection of $\vec{v}$ onto each of the orthonormal vectors.

$$c_i = \frac{\vec{u}_1^T\vec{v}}{\|\vec{u}_1\|} = \vec{u}_i^T\vec{v}$$

Another method of solving $c_1\vec{u}_1 + c_2\vec{u}_2 + \cdots + c_n\vec{u}_n = \vec{v}$ is to recognize that $U^{-1} = U^T$, where $U = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_n \end{bmatrix}$ and $\vec{c} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$.

$$U\vec{c} = \vec{v} \implies \vec{c} = U^{-1}\vec{v} = U^T\vec{v}$$

This is equivalent to finding the scalar projection of $\vec{v}$ onto each of the column vectors of U.

4 Spectral Theorem

For a real $n \times n$ symmetric matrix A,

(a) All eigenvalues of A are real.

(b) A has n real eigenvectors $\in \mathbb{R}^n$.

(c) A has orthogonal eigenvectors, i.e., $A = V\Lambda V^{-1} = V\Lambda V^T$, where Λ is a diagonal matrix and V is a unitary matrix. We say that A is orthogonally diagonalizable.

Furthermore, if A is of the form $B^T B$ for some arbitrary matrix B, all of its eigenvalues are non-negative, i.e., $\lambda \geq 0$.

1. Real Eigenvalues

Prove the following: All eigenvalues of a real symmetric matrix A are real.

Hint: Use the definition of an eigenvalue to show that $\lambda(\overline{\vec{v}}^T \vec{v}) = \overline{\lambda}(\overline{\vec{v}}^T \vec{v})$.

2. Orthogonal Eigenvectors

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

3. Positive Eigenvalues

Prove the following: For any matrix A, $A^T A$ is symmetric and only has non-negative eigenvalues.

4. Power Iteration

Power iteration is a method for approximating eigenvectors of a matrix A numerically. It's particularly effective when A is very large but very sparse. For example, Google's PageRank algorithm, used to determine the ranking of search results, essentially attempts to perform power iteration on the adjacency matrix of links between all web pages on the internet.

The method starts with any vector $x^{(0)}$ and then iterates the following update:

$$\vec{x}^{(k+1)} = \frac{A\vec{x}^{(k)}}{\|A\vec{x}^{(k)}\|}.$$

Here, $\vec{x}^{(k)}$ denotes the value of $\vec{x}$ in the kth iteration. You will show that this algorithm converges for symmetric A.

(a) Show that if A is a diagonal matrix $\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$ with λ_1 strictly greater than the other λ_i, then the power iteration method converges to $\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, which is the eigenvector corresponding to the largest eigenvalue of A.

(b) Now use the spectral decomposition to show that for any symmetric matrix whose largest eigenvalue is strictly greater than its other eigenvalues, the power iteration method converges to the eigenvector corresponding to the largest eigenvalue.

Contributors:

- Titan Yuan.
- John Maidens.