AMS526: Numerical Analysis I (Numerical Linear Algebra)

Lecture 3: Orthogonal Vectors and Matrices & Vector Norms

Xiangmin Jiao

SUNY Stony Brook

September 9, 2008

Outline

Components of Vector

- Given an orthonormal set $\{\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_m\}$ forming a basis of $\mathbb{C}^m$, vector $\boldsymbol{v}$ can be decomposed into orthogonal components as $\boldsymbol{v} = \sum_{i=1}^{m} (\boldsymbol{q}_i^* \boldsymbol{v}) \boldsymbol{q}_i$

Components of Vector

- Given an orthonormal set $\{\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_m\}$ forming a basis of $\mathbb{C}^m$, vector $\boldsymbol{v}$ can be decomposed into orthogonal components as $\boldsymbol{v} = \sum_{i=1}^{m} (\boldsymbol{q}_i^* \boldsymbol{v}) \boldsymbol{q}_i$
- Another way to express the condition is$\boldsymbol{v} = \sum_{i=1}^{m} (\boldsymbol{q}_i \boldsymbol{q}_i^*) \boldsymbol{v}$.
- $\boldsymbol{q}_i \boldsymbol{q}_i^*$ is an *orthogonal projection matrix*. Note that it is NOT an orthogonal matrix.

Components of Vector

- Given an orthonormal set $\{\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_m\}$ forming a basis of $\mathbb{C}^m$, vector $\boldsymbol{v}$ can be decomposed into orthogonal components as $\boldsymbol{v} = \sum_{i=1}^{m} (\boldsymbol{q}_i^* \boldsymbol{v}) \boldsymbol{q}_i$

- Another way to express the condition is$\boldsymbol{v} = \sum_{i=1}^{m} (\boldsymbol{q}_i \boldsymbol{q}_i^*) \boldsymbol{v}$.
- $\boldsymbol{q}_i \boldsymbol{q}_i^*$ is an *orthogonal projection matrix*. Note that it is NOT an orthogonal matrix.

- More generally, given an orthonormal set $\{\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_n\}$ with $n \leq m$, we have

$$\boldsymbol{v} = \boldsymbol{r} + \sum_{i=1}^{n} (\boldsymbol{q}_i^* \boldsymbol{v}) \boldsymbol{q}_i = \boldsymbol{r} + \sum_{i=1}^{n} (\boldsymbol{q}_i \boldsymbol{q}_i^*) \boldsymbol{v} \text{ and } \boldsymbol{r}^* \boldsymbol{q}_i = 0, 1 \leq i \leq n$$

- Let $\boldsymbol{Q}$ be composed of column vectors $\{\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_n\}$. $\boldsymbol{Q}\boldsymbol{Q}^* = \sum_{i=1}^{n} (\boldsymbol{q}_i \boldsymbol{q}_i^*)$ is an orthogonal projection matrix.

Unitary Matrices

Definition

A matrix is *unitary* if $\boldsymbol{Q}^* = \boldsymbol{Q}^{-1}$, i.e., if $\boldsymbol{Q}^*\boldsymbol{Q} = \boldsymbol{Q}\boldsymbol{Q}^* = \boldsymbol{I}$.

- In the real case, we say the matrix is *orthogonal*. Its column vectors are *orthonormal*.
- In other words, $\boldsymbol{q}_i\boldsymbol{q}_j = \delta_{ij}$, the *Kronecker delta*.

Unitary Matrices

Definition

A matrix is *unitary* if $\boldsymbol{Q}^* = \boldsymbol{Q}^{-1}$, i.e., if $\boldsymbol{Q}^*\boldsymbol{Q} = \boldsymbol{Q}\boldsymbol{Q}^* = \boldsymbol{I}$.

- In the real case, we say the matrix is *orthogonal*. Its column vectors are *orthonormal*.
- In other words, $\boldsymbol{q}_i \boldsymbol{q}_j = \delta_{ij}$, the *Kronecker delta*.

Question: What is the geometric meaning of multiplication by a unitary matrix?

Unitary Matrices

Definition

A matrix is *unitary* if $\boldsymbol{Q}^* = \boldsymbol{Q}^{-1}$, i.e., if $\boldsymbol{Q}^*\boldsymbol{Q} = \boldsymbol{Q}\boldsymbol{Q}^* = \boldsymbol{I}$.

- In the real case, we say the matrix is *orthogonal*. Its column vectors are *orthonormal*.
- In other words, $\boldsymbol{q}_i\boldsymbol{q}_j = \delta_{ij}$, the *Kronecker delta*.

Question: What is the geometric meaning of multiplication by a unitary matrix?
Answer: It preserves angles and Euclidean length. In the real case, multiplication by an orthogonal matrix $\boldsymbol{Q}$ is a rotation (if $\det(\boldsymbol{Q}) = 1$) or reflection (if $\det(\boldsymbol{Q}) = -1$).

Outline

Definition of Norms

- Norm captures "size" of vector or "distance" between vectors
- There are many different measures for "sizes" but a norm must satisfy some requirements:

Definition of Norms

- Norm captures "size" of vector or "distance" between vectors
- There are many different measures for "sizes" but a norm must satisfy some requirements:

Definition

A *norm* is a function $\| \cdot \| : \mathbb{C}^m \to \mathbb{R}$ that assigns a real-valued length to each vector. It must satisfy the following conditions:

$$
\begin{aligned}
&(1)\|\boldsymbol{x}\| \geq 0, \text{ and } \|\boldsymbol{x}\| = 0 \text{ only if } \boldsymbol{x} = \boldsymbol{0}, \\
&(2)\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|, \\
&(3)\|\alpha \boldsymbol{x}\| = |\alpha| \|\boldsymbol{x}\|.
\end{aligned}
$$

Definition of Norms

- Norm captures "size" of vector or "distance" between vectors
- There are many different measures for "sizes" but a norm must satisfy some requirements:

Definition

A *norm* is a function $\|\cdot\| : \mathbb{C}^m \to \mathbb{R}$ that assigns a real-valued length to each vector. It must satisfy the following conditions:

$$
\begin{aligned}
&(1)\|\boldsymbol{x}\| \geq 0, \text{ and } \|\boldsymbol{x}\| = 0 \text{ only if } \boldsymbol{x} = \mathbf{0}, \\
&(2)\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|, \\
&(3)\|\alpha \boldsymbol{x}\| = |\alpha|\|\boldsymbol{x}\|.
\end{aligned}
$$

- An example is Euclidean length (i.e, $\|\boldsymbol{x}\| = \sqrt{\sum_{i=1}^{m} x_i^2}$)

p-norms

- Euclidean length is a special case of p-norms, defined as

$$\|\boldsymbol{x}\|_p = \left(\sum_{i=1}^{m} |x_i|^p\right)^{1/p}$$

for $1 \leq p \leq \infty$

p-norms

- Euclidean length is a special case of p-norms, defined as

$$\|\boldsymbol{x}\|_p = \left(\sum_{i=1}^{m} |x_i|^p \right)^{1/p}$$

  for $1 \leq p \leq \infty$

- Euclidean norm is 2-norm $\|\boldsymbol{x}\|_2$ (i.e., $p = 2$)

p-norms

- Euclidean length is a special case of p-norms, defined as

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^{m} |x_i|^p \right)^{1/p}$$

  for $1 \leq p \leq \infty$

- Euclidean norm is 2-norm $\|\mathbf{x}\|_2$ (i.e., $p = 2$)
- 1-norm: $\|\mathbf{x}\|_1 = \sum_{i=1}^{m} |x_i|$

p-norms

- Euclidean length is a special case of p-norms, defined as

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^{m} |x_i|^p \right)^{1/p}$$

for $1 \le p \le \infty$

- Euclidean norm is 2-norm $\|\mathbf{x}\|_2$ (i.e., $p = 2$)
- 1-norm: $\|\mathbf{x}\|_1 = \sum_{i=1}^{m} |x_i|$
- ∞-norm: $\|\mathbf{x}\|_\infty$. What is its value?

p-norms

- Euclidean length is a special case of p-norms, defined as

$$\|\boldsymbol{x}\|_p = \left(\sum_{i=1}^{m} |x_i|^p \right)^{1/p}$$

  for $1 \leq p \leq \infty$

- Euclidean norm is 2-norm $\|\boldsymbol{x}\|_2$ (i.e., $p = 2$)
- 1-norm: $\|\boldsymbol{x}\|_1 = \sum_{i=1}^{m} |x_i|$
- ∞-norm: $\|\boldsymbol{x}\|_\infty$. What is its value?
  - Answer: $\|\boldsymbol{x}\|_\infty = \max_{1 \leq i \leq m} |x_i|$

p-norms

- Euclidean length is a special case of p-norms, defined as

$$\|\boldsymbol{x}\|_p = \left(\sum_{i=1}^{m} |x_i|^p \right)^{1/p}$$

  for $1 \le p \le \infty$

- Euclidean norm is 2-norm $\|\boldsymbol{x}\|_2$ (i.e., $p = 2$)
- 1-norm: $\|\boldsymbol{x}\|_1 = \sum_{i=1}^{m} |x_i|$
- ∞-norm: $\|\boldsymbol{x}\|_\infty$. What is its value?
  - Answer: $\|\boldsymbol{x}\|_\infty = \max_{1 \le i \le m} |x_i|$
- Why we require $p \ge 1$? What happens if $0 \le p < 1$?

Weighted p-norms

- A generalization of p-norm is *weighted p-norm*, which assigns different weights (priorities) to different components.
  - It is anisotropic instead of isotropic
- Algebraically, $\|\boldsymbol{x}\|_W = \|\boldsymbol{W}\boldsymbol{x}\|$, where $\boldsymbol{W}$ is diagonal matrix with ith diagonal entry $w_i \neq 0$ being weight for ith component
- In other words,

$$\|\boldsymbol{x}\|_W = \left(\sum_{i=1}^{m} |w_i x_i|^p \right)^{1/p}$$

- What happens if we allow $w_i = 0$?

Weighted p-norms

- A generalization of p-norm is *weighted p-norm*, which assigns different weights (priorities) to different components.
  - It is anisotropic instead of isotropic
- Algebraically, $\|\boldsymbol{x}\|_W = \|\boldsymbol{W}\boldsymbol{x}\|$, where $\boldsymbol{W}$ is diagonal matrix with ith diagonal entry $w_i \neq 0$ being weight for ith component
- In other words,

$$\|\boldsymbol{x}\|_W = \left(\sum_{i=1}^{m} |w_i x_i|^p\right)^{1/p}$$

- What happens if we allow $w_i = 0$?
- Can we further generalize it to allow $\boldsymbol{W}$ being arbitrary matrix?
- No. But we can allow $\boldsymbol{W}$ to be arbitrary nonsingular matrix.