AMS526: Numerical Analysis I (Numerical Linear Algebra)

Lecture 2: Matrix-Vector Multiplication Cont'd & Orthogonal Vectors and Matrices

Xiangmin Jiao

SUNY Stony Brook

September 4, 2008

Outline

Review of Matrix-Vector Multiplication

- In $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}$, $\boldsymbol{b}$ is a *linear combination* of column vectors of $\boldsymbol{A}$
- In $\boldsymbol{B} = \boldsymbol{A}\boldsymbol{C}$, each column of $\boldsymbol{B}$ is a linear combination of column vectors of $\boldsymbol{A}$
- Range of matrix $\boldsymbol{A}$ is the column space of $\boldsymbol{A}$, i.e., the space spanned by its column vectors
- Null space of matrix $\boldsymbol{A}$ is the set of vectors $\boldsymbol{x}$ that satisfy $\boldsymbol{A}\boldsymbol{x} = 0$.
- Rank of a matrix is the dimension of its column space
- Note: null($\boldsymbol{A}$) and range($\boldsymbol{A}$) are not complementary

Perspective: Vector Space

A useful way in understanding matrix operations is to think in terms of vector spaces

- Vector space spanned by a set of vectors is composed of linear combinations of these vectors
  - It is closed under addition and scalar multiplication
  - $\mathbf{0}$ is always a member of a subspace
  - Space spanned by m-vectors is subspace of $\mathbb{C}^m$
- If S_1 and S_2 are two subspaces, then $S_1 \cap S_2$ is a subspace, so is $S_1 + S_2$, the space of sum of vectors from S_1 and S_2.
  - Note that $S_1 + S_2$ is different from $S_1 \cup S_2$
- Two subspaces S_1 and S_2 of $\mathbb{C}^m$ are *complementary subspaces* of each other if $S_1 + S_2 = \mathbb{C}^m$ and $S_1 \cap S_2 = \{\mathbf{0}\}$.
  - In other words, $\dim(S_1) + \dim(S_2) = m$ and $S_1 \cap S_2 = \{\mathbf{0}\}$

Transpose and Adjoint

- *Transpose* of $\boldsymbol{A}$, denoted by $\boldsymbol{A}^T$, is the matrix $\boldsymbol{B}$ with $b_{ij} = a_{ji}$
- *Adjoint* or *Hermitian conjugate*, denoted by $\boldsymbol{A}^*$ of $\boldsymbol{A}^H$, is the matrix $\boldsymbol{B}$ with $b_{ij} = \bar{a}_{ji}$
- Note that, $(\boldsymbol{AB})^T = \boldsymbol{B}^T\boldsymbol{A}^T$ and $(\boldsymbol{AB})^* = \boldsymbol{B}^*\boldsymbol{A}^*$
- A matrix $\boldsymbol{A}$ is *symmetric* if $\boldsymbol{A} = \boldsymbol{A}^T$ (i.e., $a_{ij} = a_{ji}$). It is *Hermitian* if $\boldsymbol{A} = \boldsymbol{A}^*$ (i.e., $a_{ij} = \bar{a}_{ji}$)
- For $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, null($\boldsymbol{A}$) and range($\boldsymbol{A}^T$) are complementary subspaces. In addition, null($\boldsymbol{A}$) and range($\boldsymbol{A}^T$) are *orthogonal* to each other (to be explained later)
- For $\boldsymbol{A} \in \mathbb{C}^{m\times n}$, null($\boldsymbol{A}$) and range($\boldsymbol{A}^*$) are complementary subspaces.

Full Rank

Definition

A matrix has *full rank* if it has the maximal possible rank, i.e., $\min\{m, n\}$.

Full Rank

Definition

A matrix has *full rank* if it has the maximal possible rank, i.e., $\min\{m, n\}$.

Theorem

A matrix $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ with $m \geq n$ has full rank if and only if it maps no two distinct vectors to the same vector.

In other word, the linear mapping defined by $\boldsymbol{Ax}$ for $\boldsymbol{x} \in \boldsymbol{C}^n$ is one-to-one.

Full Rank

Definition

A matrix has *full rank* if it has the maximal possible rank, i.e., $\min\{m, n\}$.

Theorem

A matrix $\boldsymbol{A} \in \mathbb{C}^{m \times n}$ *with* $m \geq n$ *has full rank if and only if it maps no two distinct vectors to the same vector.*

In other word, the linear mapping defined by $\boldsymbol{Ax}$ for $\boldsymbol{x} \in \boldsymbol{C}^n$ is one-to-one.

Proof.

($\Rightarrow$) Column vectors of $\boldsymbol{A}$ forms a basis of range($\boldsymbol{A}$), so every $\boldsymbol{b} \in$ range($\boldsymbol{A}$) has a unique linear expansion in terms of the columns of $\boldsymbol{A}$.
($\Leftarrow$) If $\boldsymbol{A}$ does not have full rank, then its column vectors are linear dependent, so its vectors do not have a unique linear combination. $\square$

Full Rank

Definition

A matrix has *full rank* if it has the maximal possible rank, i.e., $\min\{m, n\}$.

Theorem

A matrix $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ with $m \geq n$ has full rank if and only if it maps no two distinct vectors to the same vector.

In other word, the linear mapping defined by $\boldsymbol{Ax}$ for $\boldsymbol{x} \in \boldsymbol{C}^n$ is one-to-one.

Proof.

($\Rightarrow$) Column vectors of $\boldsymbol{A}$ forms a basis of range($\boldsymbol{A}$), so every $\boldsymbol{b} \in$ range($\boldsymbol{A}$) has a unique linear expansion in terms of the columns of $\boldsymbol{A}$.
($\Leftarrow$) If $\boldsymbol{A}$ does not have full rank, then its column vectors are linear dependent, so its vectors do not have a unique linear combination. □

Definition

A *nonsingular* or *invertible* matrix is a square matrix of full rank.

Inverse

Definition

Given a nonsingular matrix $\boldsymbol{A}$, its *inverse* is written as $\boldsymbol{A}^{-1}$, and $\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}$.

- Note that $(\boldsymbol{A}\boldsymbol{B})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$.
- $(\boldsymbol{A}^{-1})^* = (\boldsymbol{A}^*)^{-1}$, and we use $\boldsymbol{A}^{-*}$ as a shorthand for it

Inverse

Definition

Given a nonsingular matrix $\boldsymbol{A}$, its *inverse* is written as $\boldsymbol{A}^{-1}$, and $\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}$.

- Note that $(\boldsymbol{A}\boldsymbol{B})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$.
- $(\boldsymbol{A}^{-1})^* = (\boldsymbol{A}^*)^{-1}$, and we use $\boldsymbol{A}^{-*}$ as a shorthand for it

Theorem

The following conditions are equivalent:
(a) A has an inverse $\boldsymbol{A}^{-1}$,
(b) rank$(\boldsymbol{A})$ *is* m,
(c) range$(\boldsymbol{A})$ *is* $\mathbb{C}^m$,
(d) null$(\boldsymbol{A})$ *is* $\{\mathbf{0}\}$,
(e) 0 is not an eigenvalue of $\boldsymbol{A}$,
(f) 0 is not a singular value of $\boldsymbol{A}$,
(g) det$(\boldsymbol{A}) \neq 0$.

Matrix Inverse Times a Vector

- When writing $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$, it means $\boldsymbol{x}$ is the solution of $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$
- In other words, $\boldsymbol{A}^{-1}\boldsymbol{b}$ is the vector of coefficients of the expansion of $\boldsymbol{b}$ in the basis of columns of $\boldsymbol{A}$
- Multiplication by $\boldsymbol{A}$ and $\boldsymbol{A}^{-1}$ are *change of basis* operations

Rank-1 Matrices

- Full-rank matrices are important.
- Another interesting space case is rank-1 matrices.
- A matrix $\boldsymbol{A}$ is rank if it can be written as $\boldsymbol{A} = \boldsymbol{u}\boldsymbol{v}^*$, where $\boldsymbol{u}$ and $\boldsymbol{v}$ are nonzero vectors
- $\boldsymbol{u}\boldsymbol{v}^*$ is called the *outer product* of the two vectors, as opposed to the inner product $\boldsymbol{u}^*\boldsymbol{v}$.

Outline

Inner Product

- *Inner product* (*dot product*) of two column vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{C}^m$ is $\boldsymbol{u}^* \boldsymbol{v} = \sum_{i=1}^{m} \bar{u}_i v_i$
- In contrast, *outer product* of $\boldsymbol{u}$ and $\boldsymbol{v}$ is $\boldsymbol{u}\boldsymbol{v}^*$
- Note that *cross product* is different

Inner Product

- *Inner product* (*dot product*) of two column vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{C}^m$ is $\boldsymbol{u}^* \boldsymbol{v} = \sum_{i=1}^{m} \bar{u}_i v_i$
- In contrast, *outer product* of $\boldsymbol{u}$ and $\boldsymbol{v}$ is $\boldsymbol{u}\boldsymbol{v}^*$
- Note that *cross product* is different

- *Euclidean length* of $\boldsymbol{u}$ is the square root of the inner product of $\boldsymbol{u}$ with itself, i.e., $\sqrt{\boldsymbol{u}^* \boldsymbol{u}}$

Inner Product

- *Inner product* (*dot product*) of two column vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{C}^m$ is $\boldsymbol{u}^* \boldsymbol{v} = \sum_{i=1}^{m} \bar{u}_i v_i$
- In contrast, *outer product* of $\boldsymbol{u}$ and $\boldsymbol{v}$ is $\boldsymbol{u}\boldsymbol{v}^*$
- Note that *cross product* is different
- *Euclidean length* of $\boldsymbol{u}$ is the square root of the inner product of $\boldsymbol{u}$ with itself, i.e., $\sqrt{\boldsymbol{u}^* \boldsymbol{u}}$
- Inner product of two unit vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ is the cosine of the angle α between $\boldsymbol{u}$ and $\boldsymbol{v}$, i.e., $\cos \alpha = \frac{\boldsymbol{u}^* \boldsymbol{v}}{\|\boldsymbol{u}\| \|\boldsymbol{v}\|}$

Inner Product

- *Inner product* (*dot product*) of two column vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{C}^m$ is $\boldsymbol{u}^*\boldsymbol{v} = \sum_{i=1}^m \bar{u}_i v_i$
- In contrast, *outer product* of $\boldsymbol{u}$ and $\boldsymbol{v}$ is $\boldsymbol{u}\boldsymbol{v}^*$
- Note that *cross product* is different
- *Euclidean length* of $\boldsymbol{u}$ is the square root of the inner product of $\boldsymbol{u}$ with itself, i.e., $\sqrt{\boldsymbol{u}^*\boldsymbol{u}}$
- Inner product of two unit vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ is the cosine of the angle α between $\boldsymbol{u}$ and $\boldsymbol{v}$, i.e., $\cos\alpha = \frac{\boldsymbol{u}^*\boldsymbol{v}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}$
- Inner product is *bilinear*, in the sense that it is linear in each vertex separately:

$$(\boldsymbol{u}_1 + \boldsymbol{u}_2)^*\boldsymbol{v} = \boldsymbol{u}_1^*\boldsymbol{v} + \boldsymbol{u}_2^*\boldsymbol{v}$$
$$\boldsymbol{u}^*(\boldsymbol{v}_1 + \boldsymbol{v}_2) = \boldsymbol{u}^*\boldsymbol{v}_1 + \boldsymbol{u}^*\boldsymbol{v}_2$$
$$(\alpha\boldsymbol{u})^*(\beta\boldsymbol{v}) = \bar{\alpha}\beta\boldsymbol{u}^*\boldsymbol{v}$$

Orthogonal Vectors

Definition

A pair of vectors are *orthogonal* if $\boldsymbol{x}^*\boldsymbol{y} = 0$.

In other words, the angle between them is 90 degrees

Orthogonal Vectors

Definition

A pair of vectors are *orthogonal* if $\boldsymbol{x}^*\boldsymbol{y} = 0$.

In other words, the angle between them is 90 degrees

Definition

Two sets of vectors X and Y are orthogonal if every $\boldsymbol{x} \in X$ is orthogonal to every $\boldsymbol{y} \in Y$.

Orthogonal Vectors

Definition

A pair of vectors are *orthogonal* if $\boldsymbol{x}^*\boldsymbol{y} = 0$.

In other words, the angle between them is 90 degrees

Definition

Two sets of vectors X and Y are orthogonal if every $\boldsymbol{x} \in X$ is orthogonal to every $\boldsymbol{y} \in Y$.

Definition

A set of nonzero vectors S is *orthogonal* if they are pairwise orthogonal. They are *orthonormal* if it is orthogonal and in addition each vector has unit Euclidean length.

Orthogonal Vectors

Theorem

The vectors in an orthogonal set S are linearly independent.

Orthogonal Vectors

Theorem

The vectors in an orthogonal set S are linearly independent.

Proof.

Prove by contradiction. If a vector can be expressed as linear combination of the other vectors in the set, then it is orthogonal to itself. □

Question: If the column vectors of an $m \times n$ matrix $\boldsymbol{A}$ are orthogonal, what is the rank of $\boldsymbol{A}$?

Orthogonal Vectors

Theorem

The vectors in an orthogonal set S are linearly independent.

Proof.

Prove by contradiction. If a vector can be expressed as linear combination of the other vectors in the set, then it is orthogonal to itself. □

Question: If the column vectors of an $m \times n$ matrix $\boldsymbol{A}$ are orthogonal, what is the rank of $\boldsymbol{A}$?

Answer: $n = \min\{m, n\}$. In other words, $\boldsymbol{A}$ has full rank.