AMS526: Numerical Analysis I
(Numerical Linear Algebra)
Lecture 1: Course Overview & Matrix-Vector Multiplication

Xiangmin Jiao
SUNY Stony Brook
Outline

1 Course Overview

2 Matrix-Vector Multiplication
Course Description

- What is numerical linear algebra?
  - Solving linear algebra problems using efficient algorithms on computers


- Prerequisite/Co-requisite:
  - AMS 510 (Analytical Methods for Applied Mathematics and Statistics)
  - Programming in C (AMS 595, Fundamentals of Computing)
  - This must NOT be your first course in linear algebra, or you will get lost


Why Learn Numerical Linear Algebra?

- Numerical linear algebra is foundation of scientific computations
- Many problems ultimately reduce to linear algebra concepts or algorithms, either analytical or computational
- Examples: Finite-element analysis, data fitting, PageRank (Google)

Focus: Fundamental concepts, algorithms, and programming
Course Outline

- Matrix fundamentals, multiplications, orthogonality, norms, and SVD (2.5 weeks).
- QR factorization, projectors, Gram-Schmidt algorithm, Householder triangulation (2 weeks).
- Conditioning and stability (1.5 weeks).
- Solution of linear system of equations, Gaussian elimination, pivoting, LU factorization, Cholesky factorization (2 weeks).
- Eigenvalue problems, Hessenberg tridiagonalization, Rayleigh quotient, power method, inverse power method and shifting (2.5 weeks).
- Iterative methods, Lanczos iteration (1 weeks).
- Linear algebra software packages (1 weeks).
- Course webpage:
  http://www.ams.sunysb.edu/~jiao/teaching/ams526_fall11

Note: Course schedule online is tentative and is subject to change.
Course Policy

- **Assignments (written or programming)**
  - Assignments are due in class one to two weeks after assigned
  - You can discuss course materials and homework problems with others, but you must write your answers completely independently
  - Do NOT copy solutions from any source. Do NOT share your solutions to others

- **Exams and tests**
  - All exams are closed-book
  - However, one-page cheat sheet is allowed

- **Grading**
  - Assignments: 30%
  - Two tests: 40%
  - Final exam: 30%
Outline

1. Course Overview

2. Matrix-Vector Multiplication
Definition

- Matrix-vector product \( \mathbf{b} = \mathbf{A}\mathbf{x} \)

\[
\mathbf{b}_i = \sum_{j=1}^{n} a_{ij} x_j
\]

- All entries belong to \( \mathbb{C} \), the field of complex numbers. The space of \( m \)-vectors is \( \mathbb{C}^m \), and the space of \( m \times n \) matrices is \( \mathbb{C}^{m \times n} \).

- The map \( \mathbf{x} \mapsto \mathbf{A}\mathbf{x} \) is linear, which means that for any \( \mathbf{x}, \mathbf{y} \in \mathbb{C}^n \) and any \( \alpha \in \mathbb{C} \)

\[
\begin{align*}
\mathbf{A}(\mathbf{x} + \mathbf{y}) &= \mathbf{A}\mathbf{x} + \mathbf{A}\mathbf{y}, \\
\mathbf{A}(\alpha \mathbf{x}) &= \alpha \mathbf{A}\mathbf{x}.
\end{align*}
\]

- In lecture notes, I use boldface UPPERCASE letters (such as \( \mathbf{A} \)) for matrices and boldface lowercase letters (such as \( \mathbf{x} \)) for vectors.
Pseudo-Code for Matrix-Vector Product

Pseudo-code for $b = Ax$

\[
\text{for } i = 1 \text{ to } m \text{ do}
\]
\[
b(i) = 0;
\]
\[
\text{for } j = 1 \text{ to } n \text{ do}
\]
\[
b(i) = b(i) + A(i,j) \times x(j);
\]
\[
\text{end for}
\]
\[
\text{end for}
\]

Note: Pseudo-code are not real code in that they do not run on any computer. They are human readable, but they should be straightforward to convert into real computer codes in any programming languages.
Linear Combination

- Alternatively, matrix-vector product can be viewed as
  
  \[ b = Ax = \sum_{j=1}^{n} x_j a_j \]

  i.e., \( b \) is a linear combination of column vectors of \( A \)

- Two different views of matrix-vector products:
  1. \( b_i = \sum_{j=1}^{n} a_{ij} x_j \): \( A \) acts on \( x \) to produce \( b \); scalar operations;
  2. \( b = \sum_{j=1}^{n} x_j a_j \): \( x \) acts on \( A \) to produce \( b \); vector operations.

- If \( A \) is \( m \times n \), \( Ax \) can be viewed as a mapping from \( \mathbb{C}^n \) to \( \mathbb{C}^m \)
Matrix-Matrix Multiplication

- If \( A \) is \( l \times m \) and \( C \) is \( m \times n \), then \( B = AC \) is \( l \times n \), with entries defined by

\[
b_{ij} = \sum_{k=1}^{m} a_{ik} c_{kj}.
\]

- Written in columns, we have

\[
b_j = Ac_j = \sum_{k=1}^{m} c_{kj} a_k.
\]

- In other word, each column of \( B \) is a linear combination of the columns of \( A \).
Pseudo-Code for Matrix-Matrix Multiplication

Pseudo-code for \( B = AC \)

\[
\text{for } i = 1 \text{ to } l \text{ do} \\
\quad \text{for } j = 1 \text{ to } n \text{ do} \\
\quad \quad B(i, j) = 0; \\
\quad \quad \text{for } k = 1 \text{ to } m \text{ do} \\
\quad \quad \quad B(i, j) = B(i, j) + A(i, k) \times C(k, j); \\
\quad \quad \text{end for} \\
\quad \text{end for} \\
\text{end for}
\]
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
  - \( 0 \) 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 \( S_1 \cap S_2 = \{0\} \).
  - In other words, \( \dim(S_1) + \dim(S_2) = m \) and \( S_1 \cap S_2 = \{0\} \)
Range and Null Space

Definition
The *range* of a matrix $A$, written as $\text{range}(A)$, is the set of vectors that can be expressed as $Ax$ for some $x$.

Theorem
$\text{range}(A)$ is the space spanned by the columns of $A$. Therefore, the range of $A$ is also called the *column space* of $A$.

Definition
The *null space* of $A \in \mathbb{C}^{m \times n}$, written as $\text{null}(A)$, is the set of vectors $x$ that satisfy $Ax = 0$. Entries of $x \in \text{null}(A)$ give coefficient of $\sum x_i a_i = 0$.

Note: The null space of $A$ is in general not a complementary subspace of $\text{range}(A)$.
Range and Null Space

**Definition**

The *range* of a matrix $A$, written as $\text{range}(A)$, is the set of vectors that can be expressed as $Ax$ for some $x$.

**Theorem**

$\text{range}(A)$ is the space spanned by the columns of $A$.

Therefore, the *range* of $A$ is also called the *column space* of $A$. 
## Range and Null Space

### Definition

The **range** of a matrix $A$, written as $\text{range}(A)$, is the set of vectors that can be expressed as $Ax$ for some $x$.

### Theorem

\[ \text{range}(A) \text{ is the space spanned by the columns of } A. \]

Therefore, the **range** of $A$ is also called the **column space** of $A$.

### Definition

The **null space** of $A \in \mathbb{C}^{m \times n}$, written as $\text{null}(A)$, is the set of vectors $x$ that satisfy $Ax = 0$.

Entries of $x \in \text{null}(A)$ give coefficient of $\sum x_ia_i = 0$.

Note: The null space of $A$ is in general **not** a complementary subspace of $\text{range}(A)$. 
Rank

Definition

The *column rank* of a matrix is the dimension of its column space. The *row rank* is the dimension of the space spanned by its rows.

**Question**: Can the column rank and the row rank be different?

**Answer**: No. We will give a proof in future lectures.

We therefore simply say the *rank* of a matrix.

**Question**: Given \( A \in \mathbb{C}^{m \times n} \), what is \( \dim(\text{null}(A)) + \text{rank}(A) \) equal to?

**Answer**: \( n \).
Rank

Definition

The column rank of a matrix is the dimension of its column space. The row rank is the dimension of the space spanned by its rows.

Question: Can the column rank and the row rank be different?
Answer: No. We will give a proof in future lectures.

We therefore simply say the rank of a matrix.

Question: Given $A \in \mathbb{C}^{m \times n}$, what is $\dim(\ker(A)) + \text{rank}(A)$ equal to?
Ran

Definition

The column rank of a matrix is the dimension of its column space. The row rank is the dimension of the space spanned by its rows.

Question: Can the column rank and the row rank be different?  
Answer: No. We will give a proof in future lectures.

- We therefore simply say the rank of a matrix.

Question: Given $A \in \mathbb{C}^{m \times n}$, what is $\dim(\text{null}(A)) + \text{rank}(A)$ equal to?  
Answer: $n$.  

Transpose and Adjoint

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

Definition

A matrix has \textit{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 \( A \in \mathbb{C}^{m \times n} \) 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 \( Ax \) for \( x \in \mathbb{C}^n \) 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 \( A \in \mathbb{C}^{m \times n} \) with \( m \geq n \) has full rank if and only if it maps no two distinct vectors to the same vector.

In other words, the linear mapping defined by \( Ax \) for \( x \in \mathbb{C}^n \) is one-to-one.

Proof.

(\( \Rightarrow \)) Column vectors of \( A \) forms a basis of range(\( A \)), so every \( b \in \text{range}(A) \) has a unique linear expansion in terms of the columns of \( A \).

(\( \Leftarrow \)) If \( A \) does not have full rank, then its column vectors are linear dependent, so its vectors do not have a unique linear combination.
Full Rank

Definition
A matrix has \textit{full rank} if it has the maximal possible rank, i.e., \( \min\{m, n\} \).

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

In other words, the linear mapping defined by \( A x \) for \( x \in \mathbb{C}^n \) is one-to-one.

Proof.
\((\Rightarrow)\) Column vectors of \( A \) forms a basis of range(\( A \)), so every \( b \in \text{range}(A) \) has a unique linear expansion in terms of the columns of \( A \).

\((\Leftarrow)\) If \( A \) does not have full rank, then its column vectors are linear dependent, so its vectors do not have a unique linear combination.

Definition
A \textit{nonsingular} or \textit{invertible} matrix is a square matrix of full rank.
Definition

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

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

Definition

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

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

Theorem

The following conditions are equivalent:

(a) $A$ has an inverse $A^{-1}$,
(b) $\text{rank}(A)$ is $m$,
(c) $\text{range}(A)$ is $\mathbb{C}^m$,
(d) $\text{null}(A)$ is $\{0\}$,
(e) 0 is not an eigenvalue of $A$,
(f) 0 is not a singular value of $A$,
(g) $\det(A) \neq 0$. 
Matrix Inverse Times a Vector

- When writing $x = A^{-1}b$, it means $x$ is the solution of $Ax = b$
- In other words, $A^{-1}b$ is the vector of coefficients of the expansion of $b$ in the basis of columns of $A$
- Multiplying $b$ by $A^{-1}$ is a *change of basis* operations from \{a_1, a_2, \ldots, a_m\} to \{e_1, e_2, \ldots, e_m\}
- Multiplying $A^{-1}b$ by $A$ is a *change of basis* operations from \{e_1, e_2, \ldots, e_m\} to \{a_1, a_2, \ldots, a_m\}
Rank-1 Matrices

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