Review of Some Concepts from Linear Algebra: Part 2

Department of Mathematics

Boise State University

January 13, 2020
The focus of this course is on the vector spaces $\mathbb{R}^m$ and $\mathbb{C}^m$. 

A set of vectors $X$ is called a vector space if for any $x, y, z \in X$ and any scalars $\alpha, \beta \in \mathbb{C}$ the following holds:

1. $x + y \in X$
2. $x + y = y + x$
3. $x + (y + z) = (x + y) + z$
4. There exists a unique zero element $0 \in X$, such that $x + 0 = x$
5. There exists a $-x \in X$ such that $x + (-x) = 0$
6. $\alpha x \in X$
7. $1 \times x = x$
8. $\alpha(x + y) = \alpha x + \alpha y$
9. $(\alpha + \beta)x = \alpha x + \beta x$
10. $\alpha(\beta x) = (\alpha \beta)x$
Vector spaces

A set of vectors $X$ is called a *vector space* if for any $x, y, z \in X$ and any scalars $\alpha, \beta \in \mathbb{C}$ the following holds

1. $x + y \in X$
2. $x + y = y + x$
3. $x + (y + z) = (x + y) + z$
4. There exists a unique zero element $0 \in X$, such that $x + 0 = x$
5. There exists a $-x \in X$ such that $x + (-x) = 0$
6. $\alpha x \in X$
7. $1 x = x$
8. $\alpha(x + y) = \alpha x + \alpha y$
9. $(\alpha + \beta)x = \alpha x + \beta x$
10. $\alpha(\beta x) = (\alpha \beta)x$

The focus of this course is on the vector spaces $\mathbb{R}^m$ and $\mathbb{C}^m$. 
Vector subspaces

A subset of vectors $Y$ of $\mathbb{C}^m$ is called a subspace if for any two vectors $x, y \in Y$ and any scalars $\alpha, \beta \in \mathbb{C}$, the following holds:

$$\alpha x + \beta y \in Y$$

Example: The subset of vectors of $\mathbb{R}^3$ given by

$$x = \begin{bmatrix} t \\ 2t \\ -3t \end{bmatrix}$$

where $t \in \mathbb{R}$, is a subspace of $\mathbb{R}^3$. 
A subset of vectors \( Y \) of \( \mathbb{C}^m \) is called a \textit{subspace} if for any two vectors \( x, y \in Y \) and any scalars \( \alpha, \beta \in \mathbb{C} \), the following holds:

\[
\alpha x + \beta y \in Y
\]

Example: The subset of vectors of \( \mathbb{R}^3 \) given by

\[
x = \begin{bmatrix}
t \\
2t \\
-3t
\end{bmatrix},
\]

where \( t \in \mathbb{R} \), is a subspace of \( \mathbb{R}^3 \).
Linear independence

A set of vectors \( \{x_1, x_2, \ldots, x_n\} \) is \textit{linearly independent} if

\[
\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = 0
\]

only holds when \( \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0 \).

Example: The set of vectors

\[
\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}
\]

is linearly dependent, but the set of vectors

\[
\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}
\]

is not.
A set of vectors \( \{x_1, x_2, \ldots, x_n\} \) is *linearly independent* if

\[
\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = 0
\]

only holds when \( \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0 \).

**Example:** The set of vectors

\[
\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}
\]

is linearly dependent, but the set of vectors

\[
\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \end{bmatrix} \right\}
\]

is not.
The **span** of a set of vectors is the subspace formed from all linear combinations of the vectors.

A **basis** for a subspace $X$ is a set of linearly independent vectors that span $X$.

The **dimension** of a subspace $X$ is the number of vectors in any basis for $X$.

The **standard basis** for $\mathbb{R}^n$ or $\mathbb{C}^n$ is given by the \{${e_1, e_2, \ldots, e_n}$\}, where $e \in \mathbb{R}^n$ and

$$(e_i)_j = \begin{cases} 0 & \text{if } i \neq j, \quad i, j = 1, 2, \ldots, n, \\ 1 & \text{if } i = j, \quad i, j = 1, 2, \ldots, n, \end{cases}$$

i.e. $e_i$ contains all zeros except for one 1 in the $i^{th}$ entry.
Span, basis, dimension

- The *span* of a set of vectors is the subspace formed from all linear combinations of the vectors.

- A *basis* for a subspace $X$ is a set of linearly independent vectors that span $X$.

- The *dimension* of a subspace $X$ is the number of vectors in any basis for $X$.

The *standard basis* for $\mathbb{R}^n$ or $\mathbb{C}^n$ is given by the \{${e_1, e_2, \ldots, e_n}$\}, where $e \in \mathbb{R}^n$ and

$$(e_i)_j = \begin{cases} 0 & i \neq j, \\ 1 & i = j, \end{cases} \quad i,j = 1,2,\ldots,n,$$

i.e. $e_i$ contains all zeros except for one 1 in the $i$th entry.
The span of a set of vectors is the subspace formed from all linear combinations of the vectors.

A basis for a subspace $X$ is a set of linearly independent vectors that span $X$.

The dimension of a subspace $X$ is the number of vectors in any basis for $X$.

The standard basis for $\mathbb{R}^n$ or $\mathbb{C}^n$ is given by the $\{e_1, e_2, \ldots, e_n\}$, where $e \in \mathbb{R}^n$ and

$$ (e_i)_j = \begin{cases} 0 & i \neq j, \ i, j = 1, 2, \ldots, n, \\ 1 & i = j \end{cases} $$

i.e. $e_i$ contains all zeros except for one 1 in the $i^{th}$ entry.
Span, basis, dimension

- The *span* of a set of vectors is the subspace formed from all linear combinations of the vectors.
- A *basis* for a subspace $X$ is a set of linearly independent vectors that span $X$.
- The *dimension* of a subspace $X$ is the number of vectors in any basis for $X$.

The *standard basis* for $\mathbb{R}^n$ or $\mathbb{C}^n$ is given by the $\{e_1, e_2, \cdots, e_n\}$, where $e \in \mathbb{R}^n$ and

$$ (e_i)_j = \begin{cases} 0 & i \neq j \text{, } i, j = 1, 2, \ldots, n, \\ 1 & i = j \end{cases} $$

i.e. $e_i$ contains all zeros except for one 1 in the $i^{th}$ entry.
Fundamental subspaces of a matrix

Let $A \in \mathbb{C}^{m \times n}$ then the following are the four fundamental subspaces of $A$:

- **Column space of $A$ (or range of $A$):** the subspace of $\mathbb{C}^m$ formed by the span of the columns of $A$. Denoted as $\mathcal{C}(A)$.

- **Row space of $A$ (or range of $A^T$):** the subspace of $\mathbb{C}^n$ formed by the span of the rows of $A$. Denoted by $\mathcal{C}(A^T)$.

- **Null space of $A$:** The subspace of vectors $x \in \mathbb{C}^n$ that satisfy $Ax = 0$. Denoted by $\mathcal{N}(A)$.

- **Left null space of $A$:** The subspace of vectors $y \in \mathbb{C}^m$ that satisfy $y^TA = 0$. Denoted by $\mathcal{N}(A^T)$. 
Fundamental subspaces of a matrix

Let $A \in \mathbb{C}^{m \times n}$ then the following are the *four fundamental subspaces* of $A$:

- **Column space of $A$ (or range of $A$):** the subspace of $\mathbb{C}^m$ formed by the span of the columns of $A$. Denoted as $\mathcal{C}(A)$.
- **Row space of $A$ (or range of $A^T$):** the subspace of $\mathbb{C}^n$ formed by the span of the rows of $A$. Denoted by $\mathcal{C}(A^T)$.
- **Null space of $A$:** The subspace of vectors $x \in \mathbb{C}^n$ that satisfy $Ax = 0$. Denoted by $\mathcal{N}(A)$.
- **Left null space of $A$:** The subspace of vectors $y \in \mathbb{C}^m$ that satisfy $y^TA = 0$. Denoted by $\mathcal{N}(A^T)$. 
Fundamental subspaces of a matrix

Let $A \in \mathbb{C}^{m \times n}$ then the following are the four fundamental subspaces of $A$:

- Column space of $A$ (or range of $A$): the subspace of $\mathbb{C}^m$ formed by the span of the columns of $A$. Denoted as $\mathcal{C}(A)$.

- Row space of $A$ (or range of $A^T$): the subspace of $\mathbb{C}^n$ formed by the span of the rows of $A$. Denoted by $\mathcal{C}(A^T)$.

- Null space of $A$: The subspace of vectors $\mathbf{x} \in \mathbb{C}^n$ that satisfy $A\mathbf{x} = \mathbf{0}$. Denoted by $\mathcal{N}(A)$.

- Left null space of $A$: The subspace of vectors $\mathbf{y} \in \mathbb{C}^m$ that satisfy $\mathbf{y}^TA = \mathbf{0}$. Denoted by $\mathcal{N}(A^T)$. 
Let $A \in \mathbb{C}^{m \times n}$ then the following are the *four fundamental subspaces* of $A$:

- **Column space of $A$ (or range of $A$):** the subspace of $\mathbb{C}^m$ formed by the span of the columns of $A$. Denoted as $\mathcal{C}(A)$.
- **Row space of $A$ (or range of $A^T$):** the subspace of $\mathbb{C}^n$ formed by the span of the rows of $A$. Denoted by $\mathcal{C}(A^T)$.
- **Null space of $A$:** The subspace of vectors $x \in \mathbb{C}^n$ that satisfy $Ax = 0$. Denoted by $\mathcal{N}(A)$.
- **Left null space of $A$:** The subspace of vectors $y \in \mathbb{C}^m$ that satisfy $y^TA = 0$. Denoted by $\mathcal{N}(A^T)$.
Fundamental theorem of linear algebra

Part 1: Let $A \in \mathbb{C}^{m \times n}$ then the column space $\mathcal{C}(A)$ and row space $\mathcal{C}(A^T)$ have dimension $r \leq \min(m, n)$ and the null space $\mathcal{N}(A)$ and $\mathcal{N}(A^T)$ have dimension $n - r$ and $m - r$, respectively.

- The dimension $r$ of $\mathcal{C}(A)$ and $\mathcal{C}(A^T)$ is called the rank of $A$.
- If $m \geq n$ and $r = n$ then $A$ is said to be of full column rank.
- If $n \geq m$ and $r = m$ then $A$ is said to be of full row rank.
- If $r < \min(m, n)$ then $A$ is said to be rank deficient.
Fundamental theorem of linear algebra

Part 1: Let \( A \in \mathbb{C}^{m \times n} \) then the column space \( \mathcal{C}(A) \) and row space \( \mathcal{C}(A^T) \) have dimension \( r \leq \min(m, n) \) and the null space \( \mathcal{N}(A) \) and \( \mathcal{N}(A^T) \) have dimension \( n - r \) and \( m - r \), respectively.

- The dimension \( r \) of \( \mathcal{C}(A) \) and \( \mathcal{C}(A^T) \) is called the rank of \( A \).
- If \( m \geq n \) and \( r = n \) then \( A \) is said to be of full column rank.
- If \( n \geq m \) and \( r = m \) then \( A \) is said to be of full row rank.
- If \( r < \min(m, n) \) then \( A \) is said to be rank deficient.
Part 1: Let $A \in \mathbb{C}^{m \times n}$ then the column space $\mathcal{C}(A)$ and row space $\mathcal{C}(A^T)$ have dimension $r \leq \min(m, n)$ and the null space $\mathcal{N}(A)$ and $\mathcal{N}(A^T)$ have dimension $n - r$ and $m - r$, respectively.

- The dimension $r$ of $\mathcal{C}(A)$ and $\mathcal{C}(A^T)$ is called the rank of $A$.
- If $m \geq n$ and $r = n$ then $A$ is said to be of full column rank.
- If $n \geq m$ and $r = m$ then $A$ is said to be of full row rank.
- If $r < \min(m, n)$ then $A$ is said to be rank deficient.
Part 2: Let $A \in \mathbb{C}^{m \times n}$ then the column space $\mathcal{C}(A)$ is the \textit{orthogonal complement} of the left null space $\mathcal{N}(A^T)$ and the row space $\mathcal{C}(A^T)$ is the \textit{orthogonal complement} of the null space $\mathcal{N}(A)$.

This means:

- For any $y \in \mathcal{N}(A^T)$, $y^Tz = 0$ for all $z \in \mathcal{C}(A)$.
- For any $x \in \mathcal{N}(A)$, $x^Tw = 0$ for all $w \in \mathcal{C}(A^T)$.
- $\mathbb{C}^m = \mathcal{C}(A) \oplus \mathcal{N}(A^T)$
- $\mathbb{C}^n = \mathcal{C}(A^T) \oplus \mathcal{N}(A)$
Part 2: Let \( A \in \mathbb{C}^{m \times n} \) then the column space \( \mathcal{C}(A) \) is the *orthogonal complement* of the left null space \( \mathcal{N}(A^T) \) and the row space \( \mathcal{C}(A^T) \) is the *orthogonal complement* of the null space \( \mathcal{N}(A) \).

This means:

- For any \( y \in \mathcal{N}(A^T) \), \( y^T z = 0 \) for all \( z \in \mathcal{C}(A) \).
- For any \( x \in \mathcal{N}(A) \), \( x^T w = 0 \) for all \( w \in \mathcal{C}(A^T) \).
- \( \mathbb{C}^m = \mathcal{C}(A) \oplus \mathcal{N}(A^T) \)
- \( \mathbb{C}^n = \mathcal{C}(A^T) \oplus \mathcal{N}(A) \)
Fundamental theorem of linear algebra

Part 2: Let \( A \in \mathbb{C}^{m \times n} \) then the column space \( \mathcal{C}(A) \) is the **orthogonal complement** of the left null space \( \mathcal{N}(A^T) \) and the row space \( \mathcal{C}(A^T) \) is the **orthogonal complement** of the null space \( \mathcal{N}(A) \).

This means:

- For any \( y \in \mathcal{N}(A^T) \), \( y^T z = 0 \) for all \( z \in \mathcal{C}(A) \).
- For any \( x \in \mathcal{N}(A) \), \( x^T w = 0 \) for all \( w \in \mathcal{C}(A^T) \).
- \( \mathbb{C}^m = \mathcal{C}(A) \oplus \mathcal{N}(A^T) \)
- \( \mathbb{C}^n = \mathcal{C}(A^T) \oplus \mathcal{N}(A) \)
Part 3: This is about the *Singular Value Decomposition (SVD)* and we will cover this later.
Invertibility of a matrix

**Theorem:** For a square matrix $A \in \mathbb{C}^{n \times n}$ the following are equivalent:

- $A^{-1}$ exists
- $\text{rank}(A) = n$
- Columns of $A$ form a basis for $\mathbb{C}^n$, i.e. $\mathcal{C}(A) = \mathbb{C}^n$
- Rows of $A$ form a basis for $\mathbb{C}^n$, i.e. $\mathcal{C}(A^T) = \mathbb{C}^n$
- $\mathcal{N}(A) = \{0\}$
Eigenvalues and eigenvectors

Let $A \in \mathbb{C}^{n \times n}$ then we call $\lambda \in \mathbb{C}$ an *eigenvalue* of $A$ if there exists a vector $x \in \mathbb{C}^n$ with $x \neq 0$ such that

$$Ax = \lambda x.$$ 

$x$ is called an *eigenvector* corresponding to the eigenvalue $\lambda$. 
Eigenvalues and eigenvectors

The scalar \( \lambda \in \mathbb{C} \) is an eigenvalue of \( A \in \mathbb{C}^{n \times n} \) if and only if
\[
\det(A - I\lambda) = 0,
\]
where \( I \) is the identity matrix.

- \( p_n(\lambda) = \det(A - I\lambda) \) is a polynomial in \( \lambda \) of degree \( n \).
- \( p_n(\lambda) \) is called the characteristic polynomial of \( A \).
- We never use \( p_n(\lambda) \) to numerically compute the eigenvalues of \( A \).
  - We will discuss the correct way to compute the eigenvalues.
Eigenvalues and eigenvectors

The scalar $\lambda \in \mathbb{C}$ is an eigenvalue of $A \in \mathbb{C}^{n \times n}$ if and only if $\det(A - I\lambda) = 0$, where $I$ is the identity matrix.

- $p_n(\lambda) = \det(A - I\lambda)$ is a polynomial in $\lambda$ of degree $n$.
- $p_n(\lambda)$ is called the characteristic polynomial of $A$.
- We never use $p_n(\lambda)$ to numerically compute the eigenvalues of $A$.
  - We will discuss the correct way to compute the eigenvalues.
The scalar \( \lambda \in \mathbb{C} \) is an eigenvalue of \( A \in \mathbb{C}^{n \times n} \) if and only if \( \det(A - I\lambda) = 0 \), where \( I \) is the identity matrix.

- \( p_n(\lambda) = \det(A - I\lambda) \) is a polynomial in \( \lambda \) of degree \( n \).
- \( p_n(\lambda) \) is called the characteristic polynomial of \( A \).
- We never use \( p_n(\lambda) \) to numerically compute the eigenvalues of \( A \).
  - We will discuss the correct way to compute the eigenvalues.
Eigenvalues and eigenvectors

The scalar $\lambda \in \mathbb{C}$ is an eigenvalue of $A \in \mathbb{C}^{n \times n}$ if and only if $\det(A - I\lambda) = 0$, where $I$ is the identity matrix.

- $p_n(\lambda) = \det(A - I\lambda)$ is a polynomial in $\lambda$ of degree $n$.
- $p_n(\lambda)$ is called the characteristic polynomial of $A$.
- **We never use $p_n(\lambda)$ to numerically compute the eigenvalues of $A$.**
  - We will discuss the correct way to compute the eigenvalues.
Vector norms

- A vector norm is a scalar quantity that reflects the “size” of a vector $\mathbf{x}$.
- The norm of a vector $\mathbf{x}$ is denoted as $\|\mathbf{x}\|$.
- There are many ways to define the size of a vector. If $\mathbf{x} \in \mathbb{C}^n$, the three most popular are

\[
\begin{align*}
\text{one-norm:} & \quad \|\mathbf{x}\|_1 = \sum_{k=1}^{n} |x_k|, \\
\text{two-norm:} & \quad \|\mathbf{x}\|_2 = \sqrt{\sum_{k=1}^{n} |x_k|^2}, \\
\text{\(\infty\)-norm:} & \quad \|\mathbf{x}\|_\infty = \max_{1 \leq k \leq n} |x_k|.
\end{align*}
\]
However a vector norm is defined, it must satisfy the following three properties to be called a norm:

1. \( \|x\| \geq 0 \) and \( \|x\| = 0 \) if and only if \( x = 0 \) (i.e. \( x \) contains all zeros as its entries).

2. \( \|\alpha x\| = |\alpha|\|x\| \), for any constant \( \alpha \).

3. \( \|x + y\| \leq \|x\| + \|y\| \), where \( y \in \mathbb{C}^n \). This is called the triangle inequality.
A vector $\mathbf{x}$ is called a **unit vector** if its norm is one, i.e. $\|\mathbf{x}\| = 1$.

Unit vectors will be different depending on the norm applied.

Below are several unit vectors in the one, two, and $\infty$ norms for $\mathbf{x} \in \mathbb{R}^2$.

(a) One-norm  
(b) Two-norm  
(c) $\infty$-norm
A matrix norm is a scalar quantity that reflects the “size” of a matrix $A \in \mathbb{C}^{m \times n}$.

The norm of $A$ is denoted as $\|A\|$.

Any matrix norm must satisfy the following four properties:

1. $\|A\| \geq 0$ and $\|A\| = 0$ if and only if $A = 0$ (i.e. $A$ contains all zeros as its entries).
2. $\|\alpha A\| = |\alpha|\|A\|$, for any constant $\alpha$.
3. $\|A + B\| \leq \|A\| + \|B\|$, where $B \in \mathbb{C}^{m \times n}$.
4. $\|AB\| \leq \|A\|\|B\|$, where $B \in \mathbb{C}^{n \times p}$. This is called the submultiplicative inequality.
Matrix norms

Each vector norm induces a matrix norm according to the following definition:

\[ \|A\|_p = \max \frac{\|Ax\|_p}{\|x\|_p} = \max \frac{\|Ax\|_p}{\|x\|_p = 1}, \]

where \( x \in \mathbb{C}^n \) and \( p = 1, 2, \ldots \).

Induced norms describe how the matrix stretches unit vectors with respect to that norm.
Two popular and easy to define induced matrix norms are

One-norm: \[ \|A\|_1 = \max_{1 \leq k \leq n} \sum_{j=1}^{m} |a_{jk}|, \]

\[ \infty \text{-norm}: \|A\|_\infty = \max_{1 \leq j \leq m} \sum_{k=1}^{n} |a_{jk}|. \]

- The one-norm corresponds to the maximum of the one norm of every column.
- The \( \infty \)-norm corresponds to the maximum of the one norm of every row.
Induced matrix norms

The two-norm is also a popular induced matrix norm, but is not easily derived.

**Definition:** Let $B \in \mathbb{C}^{n \times n}$ then the *spectral radius* of $B$ is

$$\rho(B) = \max_{1 \leq j \leq n} |\lambda_j|,$$

where $\lambda_j$ are the $n$ eigenvalues of $B$.

**Theorem:** Let $A \in \mathbb{C}^{m \times n}$ then $\|A\|_2 = \sqrt{\rho(A^*A)}$. 
Geometric illustration of the two-norm

Let $A = \begin{bmatrix} 4 & 1 \\ 1 & 3 \end{bmatrix}$ and consider applying $A$ to all vectors $\mathbf{x}$ such that $\|\mathbf{x}\|_2 = 1$.

(a) Unit vectors w.r.t the two-norm

(b) Transformation of unit vectors after applying $A$

The semi-major axis of the ellipse on the right (marked in black) is the two-norm of $A$. 
The most popular matrix norm that is not an induced norm is the *Frobenius* norm:

$$
\|A\|_F = \sqrt{\sum_{j=1}^{m} \sum_{k=1}^{n} |a_{jk}|^2}.
$$
Important results on matrix norms

The following are some useful inequalities involving matrix norms. Here $A \in \mathbb{R}^{m \times n}$:

- $\rho(A) \leq \|A\|$ for any matrix norm
- $\|Ax\| \leq \|A\|\|x\|
- $\frac{1}{\sqrt{m}} \|A\|_1 \leq \|A\|_2 \leq \sqrt{n} \|A\|_1$
- $\frac{1}{\sqrt{n}} \|A\|_\infty \leq \|A\|_2 \leq \sqrt{m} \|A\|_\infty$
- $\|A\|_2 \leq \|A\|_F \leq \sqrt{n} \|A\|_2$
- $\|A\|_2 \leq \sqrt{\|A\|_1 \|A\|_\infty}$