Outline

1. Introduction
   - Recap

2. Fundamental Concepts
   - Adjoint / Hermitian
   - Inner Products, Matrix Properties, Orthogonality
   - Unitary Matrices, Vector Norms, Matrix Norms

3. Next...
   - Looking Ahead
A quick review / crash course in basic linear algebra:

- Vectors: Transpose, Addition & Subtraction
- Matrix-Vector Product
- The Vandermonde Matrix ... and Linear Least Squares Problems
- Matrix-Matrix Product
- The Transpose of a Matrix \( A^T \)
- The Range and Nullspace of a Matrix \( A \)
- The Rank of a Matrix \( A_{m \times n} \)
- The Inverse of a Matrix \( A \)
The **Adjoint** a.k.a **Hermitian** (Transpose, or Conjugate) of a matrix $A \in \mathbb{C}^{m \times n}$...

For a scalar $z \in \mathbb{C}$, $z = a + bi$, the **complex conjugate** $\bar{z}$, or $z^*$ is obtained by negating the imaginary part, i.e. $z^* = a - bi$.

Note that if $z \in \mathbb{R}$, then $z^* = z$.

For a matrix $A \in \mathbb{C}^{m \times n}$, the Hermitian Conjugate $A^* \in \mathbb{C}^{n \times m}$ is the matrix

\[
A = \begin{bmatrix}
    a_{11} & a_{12} \\
    a_{21} & a_{22} \\
    a_{31} & a_{32} \\
    a_{41} & a_{42}
\end{bmatrix} \implies A^* = \begin{bmatrix}
    a_{11}^* & a_{21}^* & a_{31}^* & a_{41}^* \\
    a_{12}^* & a_{22}^* & a_{32}^* & a_{42}^*
\end{bmatrix}
\]
The Hermitian Conjugate

If \( A = A^* \), the matrix \( A \) is said to be **Hermitian**.

Note that a **Hermitian matrix must be square**.

In the case that \( A \) is real-valued, *i.e.* \( A \in \mathbb{R}^{m \times n} \), then \( A = A^* = A^T \) (the Hermitian conjugate equals the **transpose**).

If \( A = A^T \), the matrix \( A \) is said to be **Symmetric**.

Our book (*Trefethen-Bau*) tends to state results and theorems in terms of complex vectors and matrices, and hence use the Hermitian conjugate, *i.e.* \( \bar{x}^* \) is a row-vector.

If this is disturbing to you, just imagine that all quantities are real, and that \( * \equiv ^T \).

**The advantage of this approach is that we never have to wonder if a result (stated for the real case) extends to the complex case.**
The inner product, denoted $\langle \bar{x}, \bar{y} \rangle$, of two column vectors $\bar{x}, \bar{y} \in \mathbb{C}^m$ is defined

$$\langle \bar{x}, \bar{y} \rangle = \bar{x}^* \bar{y} = \sum_{i=1}^{m} x_i^* y_i$$

note that the inner product is a scalar quantity.

The Euclidean length, $\|\bar{x}\|$, of $\bar{x} \in \mathbb{C}^m$ is defined

$$\|\bar{x}\| = \sqrt{\langle \bar{x}, \bar{x} \rangle} = \sqrt{\bar{x}^* \bar{x}} = \sqrt{\sum_{i=1}^{m} |x_i|^2}$$
The inner product can also be written

$$\langle \bar{x}, \bar{y} \rangle = \bar{x}^* \bar{y} = \|\bar{x}\| \cdot \|\bar{y}\| \cdot \cos(\alpha)$$

where $\alpha$ is the angle between $\bar{x}$ and $\bar{y}$.
The inner product is **bilinear**, i.e. it is linear in each vector separately:

1. \((\bar{x}_1 + \bar{x}_2)^* \bar{y} = \bar{x}_1^* \bar{y} + \bar{x}_2^* \bar{y}\)
2. \(\bar{x}^* (\bar{y}_1 + \bar{y}_2) = \bar{x}^* \bar{y}_1 + \bar{x}^* \bar{y}_2\)
3. \((\alpha \bar{x})^* (\beta \bar{y}) = \alpha^* \beta \bar{x}^* \bar{y}\)

where \(\bar{x}, \bar{x}_1, \bar{x}_2, \bar{y}, \bar{y}_1, \bar{y}_2 \in \mathbb{C}^m\), and \(\alpha, \beta \in \mathbb{C}\).
For any two matrices $A$ and $B$, of compatible dimensions, i.e. $A \in \mathbb{C}^{m \times n}$, and $B \in \mathbb{C}^{n \times k}$ the following holds

$$(AB)^* = B^* A^*$$

If the matrices $A$ and $B$ are square, and invertible, the following holds

$$(AB)^{-1} = B^{-1} A^{-1}$$

When necessary, we use the notation $A^{-*}$ for $(A^*)^{-1} \equiv (A^{-1})^*$. 
Orthogonal and Orthonormal Vectors

Two vectors are **orthogonal** if and only if $\langle \bar{x}, \bar{y} \rangle = \bar{x}^* \bar{y} = 0$,

$$0 = \frac{\bar{x}^* \bar{y}}{\|\bar{x}\| \cdot \|\bar{y}\|} = \cos(\alpha) \iff \alpha = \pi/2 + k \cdot \pi.$$

A **set** of non-zero vectors $S$ is **orthogonal** if its elements are pairwise orthogonal, *i.e.*

$$\forall \bar{x}, \bar{y} \in S, \quad \bar{x} \neq \bar{y} \implies \bar{x}^* \bar{y} = 0$$

A **set** of vectors $S$ is **orthonormal** if it is **orthogonal**, and $\forall \bar{x} \in S$, $\|\bar{x}\| = 1$. 
Theorem (Linear Independence)

The vectors in an orthogonal set \( S \) are linearly independent.

**Proof:** If the vectors in \( S \) are not independent, then \( \exists \vec{v}_k \in S \), so that

\[
\vec{v}_k = \sum_{i \neq k} c_i \vec{v}_i.
\]

Since \( \vec{v}_k \neq 0 \), \( \langle \vec{v}_k, \vec{v}_k \rangle > 0 \), now we use the bi-linearity property of inner products, and the orthogonality of \( S \):

\[
0 < \langle \vec{v}_k, \vec{v}_k \rangle = \left\langle \vec{v}_k, \sum_{i \neq k} c_i \vec{v}_i \right\rangle = \sum_{i \neq k} c_i \left\langle \vec{v}_k, \vec{v}_i \right\rangle = 0 - 0, \forall i \neq k
\]

This contradicts the assumption that the vectors are linearly dependent, hence proving the theorem. □
Corollary: Basis for $\mathbb{C}^m$

**Corollary**

If an orthogonal set $S \subseteq \mathbb{C}^m$ contains $m$ vectors, then it is a basis for $\mathbb{C}^m$.

I.e. we can write any vector $\mathbf{v} \in \mathbb{C}^m$ as a unique linear combination:

$$\mathbf{v} = \sum_{i=1}^{m} a_i \mathbf{s}_i,$$

where

$$a_i = \frac{\langle \mathbf{s}_i, \mathbf{v} \rangle}{\| \mathbf{s}_i \|^2}.$$

We can view the computation of $a_i$ as a projection of the vector $\mathbf{v}$ onto the direction $\mathbf{s}_i$.

We can use this in order to decompose arbitrary vectors into orthogonal components...
Suppose we have an orthonormal set of vectors \( \{\bar{q}_1, \bar{q}_2, \ldots, \bar{q}_n\} \), \( \bar{q}_i \in \mathbb{C}^m \), \( n \leq m \).

Now, for any vector \( \bar{v} \in \mathbb{C}^m \), the vector

\[
\bar{r} = \bar{v} - \sum_{i=1}^{n} \langle \bar{q}_i, \bar{v} \rangle \bar{q}_i
\]

is orthogonal to \( \{\bar{q}_1, \bar{q}_2, \ldots, \bar{q}_n\} \):

\[
\langle \bar{q}_k, \bar{r} \rangle = \langle \bar{q}_k, \bar{v} \rangle - \sum_{i=1}^{n} \langle \bar{q}_i, \bar{v} \rangle \langle \bar{q}_k, \bar{q}_i \rangle = 0.
\]

\[
\underbrace{\langle \bar{q}_k, \bar{v} \rangle}_{1} \underbrace{\langle \bar{q}_k, \bar{q}_k \rangle}_{1}
\]
We see that by applying this procedure, we have decomposed the vector $\vec{v}$ into $n + 1$ orthogonal components:

$$\vec{v} = \vec{r} + \sum_{i=1}^{n} \langle \vec{q}_i, \vec{v} \rangle \vec{q}_i$$

If $\{\vec{q}_i\}$ is a basis for $\mathbb{C}^m$, then $n = m$ and $\vec{r} = \vec{0}$, i.e.

$$\vec{v} = \sum_{i=1}^{n} \langle \vec{q}_i, \vec{v} \rangle \vec{q}_i = \sum_{i=1}^{n} (\vec{q}_i^* \vec{v}) \vec{q}_i = \sum_{i=1}^{n} \vec{q}_i (\vec{q}_i^* \vec{v}) = \sum_{i=1}^{n} (\vec{q}_i \vec{q}_i^*) \vec{v}$$
\[ \bar{v} = \sum_{i=1}^{n} \langle \bar{q}_i, \bar{v} \rangle \bar{q}_i = \sum_{i=1}^{n} (\bar{q}_i^* \bar{v}) \bar{q}_i = \sum_{i=1}^{n} \bar{q}_i (\bar{q}_i^* \bar{v}) = \sum_{i=1}^{n} (\bar{q}_i \bar{q}_i^*) \bar{v} \]

In the expression \( \bar{v} = \sum_{i=1}^{n} (\bar{q}_i^* \bar{v}) \bar{q}_i \) we view \( \bar{v} \) as a sum of coefficients (circled) times vectors \( \bar{q}_i \), whereas in the equivalent expression \( \bar{v} = \sum_{i=1}^{n} (\bar{q}_i \bar{q}_i^*) \bar{v} \), we view \( \bar{v} \) as a sum of orthogonal projections onto the various directions \( \bar{q}_i \).

We will return to the issue of projection matrices of the form \( \bar{q}_i \bar{q}_i^* \) in a few lectures.
A square matrix $Q \in \mathbb{C}^{m \times m}$ is **unitary** (in the real case “orthogonal”) if

$$Q^* = Q^{-1} \iff Q^* Q = I$$

In terms of the columns, $\bar{q}_i$ of $Q$ this looks like

$$\begin{bmatrix}
\bar{q}_1^* & \bar{q}_2^* & \cdots & \bar{q}_n^*
\end{bmatrix}
\begin{bmatrix}
\bar{q}_1 & \bar{q}_2 & \cdots & \bar{q}_n
\end{bmatrix}
= \begin{bmatrix}
1 & & & \\
& 1 & & \\
& & \ddots & \\
& & & 1
\end{bmatrix}$$

We have $\bar{q}_i^* \bar{q}_j = \delta_{ij}$, the **Kronecker delta**, equal to 1 if-and-only-if $i = j$, and 0 otherwise.
Multiplication by a Unitary Matrix

Since the norm of the columns of a unitary matrix is 1, multiplication by a unitary matrix preserves the Euclidean norm in the following sense:

For a unitary $Q$:

\[(1) \quad \langle Q\bar{x}, Q\bar{y} \rangle = (Q\bar{x})^* (Q\bar{y}) = \bar{x}^* Q^* Q \bar{y} = \bar{x}^* \bar{y} = \langle \bar{x}, \bar{y} \rangle \]
\[(2) \quad \| Q\bar{x} \| = \| \bar{x} \| \]

The invariance of inner products mean that angles between vectors are preserved.

In the real case, multiplication by an orthogonal matrix corresponds to a **rigid rotation** (if $\det(Q) = 1$) or a **reflection** (if $\det(Q) = -1$) of the vector space.
Norms give us the essential notion of size and distance in a vector space — these are our tools for measuring the quality of approximations and convergence in our algorithms.

**Definition (Norm)**

A **norm** is a function $\| \cdot \| : \mathbb{C}^m \rightarrow \mathbb{R}$ that assigns a real-valued (length) to each vector. A norm must satisfy the following three conditions for all vectors $\vec{x}, \vec{y} \in \mathbb{C}^m$, and scalars $\alpha \in \mathbb{C}$,

1. $\| \vec{x} \| \geq 0$, and $\| \vec{x} \| = 0$ only if $\vec{x} = 0$
2. $\| \vec{x} + \vec{y} \| \leq \| \vec{x} \| + \| \vec{y} \|$
3. $\| \alpha \vec{x} \| = |\alpha| \| \vec{x} \|$

(2) is known as the **“triangle inequality.”**
The $p$-norms (sometimes referred to as the $l_p$-norms), parametrized by $p$ are defined by

$$ \| \bar{x} \|_p = \left[ \sum_{i=1}^{m} |x_i|^p \right]^{1/p} $$

As an illustration, the unit sphere $\| \bar{x} \|_p = 1$, $\bar{x} \in \mathbb{R}^2$ is illustrated for some common (and uncommon) $p$-norms, on the following slides.

The 2-norm is the standard Euclidean length function.

The 1-norm is sometimes referred to as the Manhattan/taxicab-distance.
The $p$-norms

Some commonly used $p$-norms

$$\| \bar{x} \|_1 = \sum_{i=1}^{m} |x_i|, \quad \| \bar{x} \|_2 = \left[ \sum_{i=1}^{m} |x_i|^2 \right]^{1/2}, \quad \| \bar{x} \|_{\infty} = \max_{i=1 \ldots m} |x_i|$$
Some exotic $p$-norms

\[
\|\bar{x}\|_4 = \left[ \sum_{i=1}^{m} |x_i|^4 \right]^{1/4}, \quad \|\bar{x}\|_{1/2} = \left[ \sum_{i=1}^{m} |x_i|^{1/2} \right]^2, \quad \|\bar{x}\|_{1/4} = \left[ \sum_{i=1}^{m} |x_i|^{1/4} \right]^4
\]

∃ Movie.
The **weighted** $p$-**norms** $\| \cdot \|_{W,p}$ are derived from the $p$-norms:

$$\| \bar{x} \|_{W,p} = \| W \bar{x} \|_p$$

where $W$ is e.g. a diagonal matrix, in which the $i$th diagonal entry is the weight $w_i \neq 0$:

$$\| \bar{x} \|_{W,p} = \left[ \sum_{i=1}^{m} |w_i x_i|^p \right]^{1/p}$$
Weighted $p$-norms

**Figure**: Visualization of the unit-sphere for the weighted 1-, 2- and $\infty$-norms, where $W = \text{diag}(2, 1)$.

The concept of weighted $p$-norms can be generalized to arbitrary non-singular weight matrices $W$. 
Weighted $p$-norms

Figure: Visualization of the unit-sphere for the weighted 1-, 2- and $\infty$-norms, where $W = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$.

∃ Movie.
Given a vector norms $\| \cdot \|_m$ and $\| \cdot \|_n$ on the domain and range of $A \in \mathbb{C}^{m \times n}$, the induced matrix norm $\|A\|_{(m,n)}$ is

$$\|A\|_{(m,n)} = \sup_{\bar{x} \in \mathbb{C}^n \setminus \{0\}} \left[ \frac{\|A\bar{x}\|_m}{\|\bar{x}\|_n} \right]$$

In any sane application, both $\| \cdot \|_m$ and $\| \cdot \|_n$ will be of the same type, i.e. the $p$-norms (with the same $p$).

Due to the linearity of norms — the third norm-condition — it is sufficient to maximize the matrix norm over $\bar{x} \in \mathbb{C}^n : \|\bar{x}\| = 1$...

Most of the time the norms with $p = 2$ are used. Indeed, if nothing else is specified, this is usually implied.
Illustration: Matrix Norms

\[ A = \begin{bmatrix} 1 & 2 \\ 1/3 & 2 \end{bmatrix} \]

\[ \|A\|_1 = 4 \quad \|A\|_2 \approx 2.9852 \quad \|A\|_\infty = 3 \]
Special Cases: Matrix $p$-norms

If $D$ is a diagonal matrix, then

$$\|D\|_p = \max_{1 \leq i \leq m} |d_i|.$$  

The 1-norm of a matrix is the maximal column-abs-sum:

$$\|A\|_1 = \max_{1 \leq j \leq n} \|\bar{a}_j\|_1$$

The $\infty$-norm of a matrix is the maximal row-abs-sum:

$$\|A\|_\infty = \max_{1 \leq i \leq m} \|\bar{a}_i^*\|_1$$
Next Time

- Finish up the discussion on norms:
  - Inequalities, General matrix norms, The Frobenius norm, Bounds on norms of products of matrices.
  - The Singular Value Decomposition (SVD).