Math 221, Section 5.2 and 5.3, Part 3

(1) More on orthogonal complements
(2) Orthogonal projection matrix $U(U^T U)^{-1} U^T$
(3) Orthogonal matrices
(4) (optional) About the choice of basis (conceptual)
Review: orthogonal complement

Recall the concept of orthogonal complement:

- Given a subspace $W$ of $\mathbb{R}^n$
- the **orthogonal complement** $W^\perp$ of $W$ is the set of all vectors in $\mathbb{R}^n$ that are orthogonal to $W$.

Therefore, every vector in $W$ is orthogonal to every vector in $W^\perp$.

Note (Sec 5.1, Ex 31) The only vector in both $W$ and $W^\perp$ is $\mathbf{0}$. 
Orthogonal complement of a column space

Recall: given an \( m \times n \) matrix \( A \),

- The **column space** \( \text{Col} \ A \) is the subspace of \( \mathbb{R}^m \) spanned by the columns of \( A \).
- The **null space** \( \text{Nul} \ A \) is the subspace of solutions of \( A \mathbf{x} = 0 \) in \( \mathbb{R}^n \).
- So null space \( \text{Nul} \ A^T \) is a subspace in \( \mathbb{R}^m \).

If \( W \) is a column space \( \text{Col} \ A \) for a matrix \( A \), then

\[
W^\perp = (\text{Col} \ A)^\perp = \text{Nul} \ A^T.
\]

Reason: The orthogonal complements of \( \text{Col} \ A \) consists of \( \mathbf{x} \) such that

\[
\mathbf{u}_1 \cdot \mathbf{x} = \mathbf{u}_2 \cdot \mathbf{x} = 0,
\]

or

\[
\begin{bmatrix}
\mathbf{u}_1 \cdot \mathbf{x} \\
\mathbf{u}_2 \cdot \mathbf{x}
\end{bmatrix} =
\begin{bmatrix}
0 \\
0
\end{bmatrix} = \mathbf{0}.
\]

The left hand side can be rewritten as

\[
\begin{bmatrix}
\mathbf{u}_1^T \mathbf{x} \\
\mathbf{u}_2^T \mathbf{x}
\end{bmatrix} =
\begin{bmatrix}
\mathbf{u}_1^T \\
\mathbf{u}_2^T
\end{bmatrix} \mathbf{x} = [\mathbf{u}_1, \mathbf{u}_2]^T \mathbf{x} = A^T \mathbf{x}.
\]

Hence the orthogonal complement is \( \text{Nul} \ A^T \).
Orthogonal complement of a column space

For example, use Sec 5.3, Ex 15 (p.5 above) again,

\[ A = [u_1, u_2] = \begin{bmatrix} -3 & -3 \\ -5 & 2 \\ 1 & 1 \end{bmatrix}. \]

As an exercise, \( A^T = \begin{bmatrix} -3 & -5 & 1 \\ -3 & 2 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 3 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \text{RREF} \]

\[ 3x - z = 0 \quad \Rightarrow \quad \text{Nul} A^T = \text{Span} \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}. \]

Then \( v = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} \) is orthogonal to both \( u_1 = \begin{bmatrix} -3 \\ -5 \\ 1 \end{bmatrix}, u_2 = \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix}. \)

Conclusion: If \( W = \text{Span}\{u_1, u_2\} \), then

\[ W^\perp \text{ is the solution space of } A^T x = 0. \]
The orthogonal projection matrix \( U(U^T U)^{-1} U^T \)

For the orthogonal projection onto a plane \( W = \text{Span}\{u_1, u_2\} \), the matrix is

\[
\text{Proj}_W = U(U^T U)^{-1} U^T
\]

where \( U \) is the matrix \([u_1, u_2]\).

For example, Sec 5.3.

15. Let \( y = \begin{bmatrix} 5 \\ -9 \\ 5 \end{bmatrix} \), \( u_1 = \begin{bmatrix} -3 \\ -5 \\ 1 \end{bmatrix} \), \( u_2 = \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix} \).

\[
\begin{align*}
\textbf{U} &= \begin{bmatrix}
5 \\
-9 \\
5
\end{bmatrix}, \\
\textbf{U}^T \textbf{U} &= \begin{bmatrix}
5 & -9 & 5
\end{bmatrix} \begin{bmatrix}
5 \\
-9 \\
5
\end{bmatrix} = \begin{bmatrix} 0 & 0 \end{bmatrix},
\end{align*}
\]
\[
\text{so } (U^T U)^{-1} = \begin{bmatrix}
\end{bmatrix}^{-1} = \begin{bmatrix}
\end{bmatrix},
\]
\[
\text{Proj}_W = U(U^T U)^{-1} U^T
\]
\[
= \begin{bmatrix}
\end{bmatrix} \begin{bmatrix}
\end{bmatrix} \begin{bmatrix}
\end{bmatrix} = \begin{bmatrix}
\frac{9}{10} & 0 & -\frac{3}{10} \\
0 & 1 & 0 \\
-\frac{3}{10} & 0 & \frac{1}{10}
\end{bmatrix}.
\]

We check
\[
\text{Proj}_W y = \begin{bmatrix}
\frac{9}{10} & 0 & -\frac{3}{10} \\
0 & 1 & 0 \\
-\frac{3}{10} & 0 & \frac{1}{10}
\end{bmatrix} \begin{bmatrix}
5 \\
-9 \\
5
\end{bmatrix} = \begin{bmatrix}
\end{bmatrix} = \hat{y}.
\]
Note that $U^T U$ is a diagonal matrix with positive diagonal entries:

$$U^T U = [u_1, u_2]^T [u_1, u_2] = \begin{bmatrix} u_1^T \\ u_2^T \end{bmatrix} [u_1, u_2] =$$

$$= \begin{bmatrix} u_1^T u_1 & u_1^T u_2 \\ u_2^T u_1 & u_2^T u_2 \end{bmatrix} = \begin{bmatrix} u_1 \cdot u_1 & u_1 \cdot u_2 \\ u_2 \cdot u_1 & u_1 \cdot u_2 \end{bmatrix}$$

$$= \begin{bmatrix} ||u_1||^2 & 0 \\ 0 & ||u_2||^2 \end{bmatrix}.$$

$$\Rightarrow (U^T U)^{-1} = \begin{bmatrix} \frac{1}{||u_1||^2} & 0 \\ 0 & \frac{1}{||u_2||^2} \end{bmatrix}.$$
When the subspace is a line $L = \text{Span}\{u\}$, then

1. $U^T U = u^T u = u \cdot u$ is a scalar
2. and

$$U(U^T U)^{-1} U^T = u(u^T u)^{-1} u^T = \frac{uu^T}{u^T u}$$

which is given last class.
Orthogonal matrices

If the basis \( \{\mathbf{u}_1, \mathbf{u}_2\} \) is an orthonormal basis, then

\[
U^T U = I \quad \text{and} \quad \text{Proj}_W = UU^T
\]

**Theorem 10**

If \( \{\mathbf{u}_1, \ldots, \mathbf{u}_p\} \) is an orthonormal basis for a subspace \( W \) of \( \mathbb{R}^n \), then

\[
\text{proj}_W \mathbf{y} = (\mathbf{y} \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{y} \cdot \mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{y} \cdot \mathbf{u}_p)\mathbf{u}_p \quad (4)
\]

If \( U = [\mathbf{u}_1 \ \mathbf{u}_2 \ \cdots \ \mathbf{u}_p] \), then

\[
\text{proj}_W \mathbf{y} = UU^T \mathbf{y} \quad \text{for all } \mathbf{y} \text{ in } \mathbb{R}^n
\]

**Note:**

- It means that \( U^{-1} = U^T \) only when \( U \) is a square matrix. In general, \( U \) is a rectangular matrix and is not invertible.
- In case if \( U \) is actually a square matrix, then

\[
U^{-1} = U^T \quad \text{and} \quad \text{Proj}_W = UU^T = I.
\]

We call \( U \) an **orthogonal matrix**. (Its columns form an orthonormal basis.)
Orthogonal matrices preserve dot-products

Matrices with *orthonormal* columns preserve dot-products, hence also preserve lengths and angles.

**Theorem 7**

Let $U$ be an $m \times n$ matrix with orthonormal columns, and let $x$ and $y$ be in $\mathbb{R}^n$. Then

- $\|Ux\| = \|x\|$ 
- $(Ux) \cdot (Uy) = x \cdot y$

**Example 6**

Let $U = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix}$ and $x = \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix}$. Notice that $U$ has orthonormal columns and

$$U^T U = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 2/3 & -2/3 & 1/3 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Verify that $\|Ux\| = \|x\|$.

**Solution**

$$Ux = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix} \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}$$

$$\|Ux\| = \sqrt{9 + 1 + 1} = \sqrt{11}$$
(optional) About the choice of basis

- The Orthogonal Projection Formula onto a line \( L = \text{Span}\{u\} \),

\[
\hat{y} = \frac{y \cdot u}{u \cdot u} u,
\]

does not depend on the choice of the vector \( u \) on \( L \).

- That means:

  - If you choose another basis vector, say \( L = \text{Span}\{cu\} \),
  - then apply the formula for \( cu \) instead of \( u \)

\[
\frac{y \cdot (cu)}{(cu) \cdot (cu)} (cu) = \frac{y \cdot u}{u \cdot u} u,
\]

because all scalar \( c \) cancel with each other.

- Also, the matrix

\[
\text{Proj}_L = \frac{uu^T}{u^Tu}
\]

does not depend on the choice, because

\[
\frac{(cu)(cu)^T}{(cu)^T(cu)} = \frac{uu^T}{u^Tu},
\]

again because all scalar \( c \) cancel with each other.
(optional) About the choice of basis

For projections onto a subspace in general, the T/F question in Sec 5.3, Ex 21.c:

c. The orthogonal projection $\hat{y}$ of $y$ onto a subspace $W$ can sometimes depend on the orthogonal basis for $W$ used to compute $\hat{y}$.

is false.

▶ The Orthogonal Projection Formula onto a plane

$$\hat{y} = \frac{y \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{y \cdot u_2}{u_2 \cdot u_2} u_2,$$

does not depend on the choice of the orthogonal basis $u_1, u_2$ on $L$.

▶ The matrix

$$\text{Proj}_W = U(U^T U)^{-1} U^T$$

also does not depend on the choice of the basis $u_1, u_2$. 
Example: Sec 5.3. Ex 15

\[ u_1 = \begin{bmatrix} -3 \\ -5 \\ 1 \end{bmatrix}, \quad u_2 = \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix}, \]

is an orthogonal basis for \( W = \text{Span}\{u_1, u_2\} \). Take another basis

\[ a_1 = u_1 + u_2 = \begin{bmatrix} -6 \\ -3 \\ 2 \end{bmatrix}, \quad a_2 = u_1 - u_2 = \begin{bmatrix} 0 \\ -7 \\ 0 \end{bmatrix} \]

Check the following statements.

- \( a_1 \cdot a_2 \neq 0 \), so \( \{a_1, a_2\} \) is not an orthogonal basis,
- Form \( A = [a_1, a_2] \), then

\[ U(U^T U)^{-1} U^T = A(A^T A)^{-1} A^T. \]

We will compute an example next class.
(optional) About the choice of basis. Be careful:

For the orthogonal projection formula

\[
\hat{y} = \frac{y \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{y \cdot u_2}{u_2 \cdot u_2} u_2
\]

- we require the basis \( \{u_1, u_2\} \) to be orthogonal,
- although we are free to choose other basis, as long as it is orthogonal.

But for the orthogonal projection matrix

\[
\text{Proj}_W = U(U^T U)^{-1} U^T
\]

- we don’t require the basis to be orthogonal!
- That means, as long as \( W = \text{Span}\{a_1, a_2\} \), we must have

\[
\text{Proj}_W = A(A^T A)^{-1} A^T
\]

We will use this formula in Sec 5.5.
Appendix: eigenvalues of projection matrices
(Refer to Problem 7 in the final exam of 2015.)

- If \( w \) is in \( W \), then \( \text{Proj}_W w = \)
- If \( x \) is in \( W^\perp \), then \( \text{Proj}_W x = \)
- Therefore, the eigenvalues of \( \text{Proj}_W \) are

\[
\lambda_1 = \quad \text{and} \quad \lambda_2 =
\]

with (both algebraic and geometric) multiplicities equal to the dimensions of \( \) and \( \)

- The eigenspaces are
  - for \( \lambda_1 = \), its eigenspace is
  - for \( \lambda_2 = \), its eigenspace is

- This also implies the square of a projection matrix is itself, that means

\[
(\text{Proj}_W)^2 = \text{Proj}_W.
\]

- (algebraic reason) \( A^2 = PD^2P^{-1} \), what is \( D \)?
- (geometric reason) \( (\text{Proj}_W)^2 y = \text{Proj}_W(\text{Proj}_W y) \)