Constructing an Orthonormal Basis from a Basis
(The Gram-Schmidt Process)

READING:

[Lay], mostly Section 6.4
Definition 1 (Basis – once more)

Let \( \vec{v}_1, \ldots, \vec{v}_m \in \mathbb{R}^n \). We say \( \vec{v}_1, \ldots, \vec{v}_m \) form a basis of a subspace \( V \) of \( \mathbb{R}^n \) if:

- \( V = \text{span}(\vec{v}_1, \ldots, \vec{v}_m) \), and

- \( \vec{v}_1, \ldots, \vec{v}_m \) are linearly independent.

This means, in particular, that every \( \vec{w} \in V \) can be written uniquely as a linear combination of \( \vec{v}_1, \ldots, \vec{v}_m \):

\[
\vec{w} = c_1 \vec{v}_1 + \cdots + c_m \vec{v}_m
\]

The coefficients \( c_1, \ldots, c_m \) are called the coefficients of \( \vec{w} \) with respect to the basis \( B = \{\vec{v}_1, \ldots, \vec{v}_m\} \).
Definition 2 (Orthogonality, length, unit vectors – once more)

1. Two vectors \( \vec{v}, \vec{w} \in \mathbb{R}^n \) are \textbf{orthogonal} if \( \vec{v} \cdot \vec{w} = 0 \).

2. The \textbf{length} (or \textbf{norm}) of a vector \( \vec{v} \in \mathbb{R}^n \) is \( \|\vec{v}\| = \sqrt{\vec{v} \cdot \vec{v}} \).

3. A vector \( \vec{u} \in \mathbb{R}^n \) is a \textbf{unit vector} if \( \|\vec{u}\| = 1 \).

Definition 3 (Orthonormal vectors – once more)

The vectors \( \vec{u}_1, \ldots, \vec{u}_k \in \mathbb{R}^n \) are called \textbf{orthonormal} if:

1. \( \|\vec{u}_1\| = \cdots = \|\vec{u}_k\| = 1 \),

2. \( \vec{u}_i \cdot \vec{u}_j = 0 \) for all \( 1 \leq i < j \leq k \).

Theorem 4 (Properties of orthonormal vectors – once more)

1. Orthonormal vectors are \textit{linearly independent}.

2. If the vectors \( \vec{u}_1, \ldots, \vec{u}_n \in \mathbb{R}^n \) are orthonormal, they form a \textit{basis} of \( \mathbb{R}^n \).
Theorem 5 (Formula for orthogonal projection — once more )

[Theorem 8, page 348, and Theorem 10, page 351, in [Lay]]

1. Let $V$ be a subspace of $\mathbb{R}^n$ with orthogonal basis $\vec{u}_1, \ldots, \vec{u}_k \in \mathbb{R}^n$. Then, for every $\vec{x} \in \mathbb{R}^n$, we have

$$\text{proj}_V(\vec{x}) = \vec{x}^\parallel = \frac{\vec{u}_1 \cdot \vec{x}}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 + \cdots + \frac{\vec{u}_k \cdot \vec{x}}{\vec{u}_k \cdot \vec{u}_k} \vec{u}_k$$

2. Let $V$ be a subspace of $\mathbb{R}^n$ with orthonormal basis $\vec{u}_1, \ldots, \vec{u}_k \in \mathbb{R}^n$. Then, for every $\vec{x} \in \mathbb{R}^n$, we have

$$\text{proj}_V(\vec{x}) = \vec{x}^\parallel = (\vec{u}_1 \cdot \vec{x}) \vec{u}_1 + \cdots + (\vec{u}_k \cdot \vec{x}) \vec{u}_k$$

Corollary 6 — once more. Let $V = \mathbb{R}^n$ with orthonormal basis $\vec{u}_1, \ldots, \vec{u}_n \in \mathbb{R}^n$. Then, for every $\vec{x} \in \mathbb{R}^n$, we have

$$\vec{x} = (\vec{u}_1 \cdot \vec{x}) \vec{u}_1 + \cdots + (\vec{u}_n \cdot \vec{x}) \vec{u}_n$$
REMINDER:

Let $\vec{v}_1, \vec{v}_2 \in \mathbb{R}^n$ be linearly independent vectors, where $n \geq 2$. Then $\vec{v}_1$ and $\vec{v}_2$ define a subspace $V$ containing both of them. We want an orthonormal basis $\{\vec{u}_1, \vec{u}_2\}$ for the same subspace $V$. One possible approach is to first define:

$$\vec{u}_1 \triangleq \frac{1}{\|\vec{v}_1\|} \vec{v}_1$$

Next, we define $\vec{u}_2$, which we want orthogonal to $\vec{u}_1$ and contained in $V$. As a preliminary computation, find $\vec{v}_2^\perp$ which is orthogonal to both $\vec{v}_1$ and $\vec{u}_1$:

$$\vec{v}_2^\perp = \vec{v}_2 - \vec{v}_2^\| = \vec{v}_2 - (\vec{u}_1 \cdot \vec{v}_2) \vec{u}_1$$

See the book (or Handout 22) for the preceding equation.

After which, we can define $\vec{u}_2$ as:

$$\vec{u}_2 \triangleq \frac{1}{\|\vec{v}_2^\perp\|} \vec{v}_2^\perp$$
QUESTION: How do we transform a basis

\[ B = \{ \vec{v}_1, \ldots, \vec{v}_m \} \]

of subspace \( V \subseteq \mathbb{R}^n \) into an orthonormal basis

\[ A = \{ \vec{u}_1, \ldots, \vec{u}_m \} \]

of the same subspace \( V \).

- Build up your intuition by doing Example 1 and Example 2 in Section 6.4 in [Lay], page 354, first.

- And then read Theorem 11 on page 355.
Theorem 7 (QR factorization)

[Theorem 12, page 357, in [Lay]. Also, read pp. 356-358 in [Lay].]

If $M$ is a $m \times n$ matrix with linearly independent columns $\vec{v}_1, \ldots, \vec{v}_n$, then there is a $m \times n$ matrix $Q$ with orthonormal columns $\vec{u}_1, \ldots, \vec{u}_n$ and an upper triangular $n \times n$ matrix $R$ with positive diagonal entries such that:

1. $\text{span}(\vec{v}_1, \ldots, \vec{v}_n) = \text{span}(\vec{u}_1, \ldots, \vec{u}_n)$,

2. $M$ is uniquely decomposed as $M = QR$,

3. $r_{11} = \|\vec{v}_1\|$ and, for every $2 \leq j \leq n$, $r_{jj} = \|\vec{v}_j^\perp\|$, and

4. for every $1 \leq i < j \leq n$, $r_{ij} = \vec{u}_i \cdot \vec{v}_j$. 