05: The Gram-Schmidt Process

Chapter Goal: To learn a practical and important algorithm, the **Gram-Schmidt Process**. This is the answer to the question, "What if our basis isn't 'nice'? Can we make it 'nice'?"

---

1. The Problem: A "Messy" Basis

- **Situation:** We are given a set of linearly independent basis vectors $\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \dots\}$ that can define a space.
- **The Problem:** These vectors are "messy".
  - They are **not mutually orthogonal** (not perpendicular).
  - They are **not unit vectors** (their lengths are not 1).
- **Goal:** To create a new **orthonormal basis** $\{\vec{e}_1, \vec{e}_2, \vec{e}_3, \dots\}$ from the messy `v` basis. This new basis must span the exact same space.

---

2. The Core Idea: "Cleaning" Vectors One by One

The Gram-Schmidt process works by building the new orthonormal basis **iteratively**, one vector at a time. The idea is: "take the next messy vector, subtract the parts we don't like, and then clean up what's left."

---

3. The Step-by-Step Process (The Gram-Schmidt Recipe)

Step 1: Create $\vec{e}_1$ (The First Orthonormal Basis Vector)

1. **Choose:** Take the first vector from the messy set, $\vec{v}_1$, as our starting point.
2. **Direction:** Its direction is fine. We will make this our first fundamental direction.
3. **Clean Up the Length:** The only problem with $\vec{v}_1$ is its length might not be 1. So, we **normalize** it.

$$\vec{e}_1 = \frac{\vec{v}_1}{|\vec{v}_1|}$$

   We now have $\vec{e}_1$, our first orthonormal basis vector.

Step 2: Create $\vec{e}_2$ (The Second Orthonormal Basis Vector)

1. **Choose:** Take the next messy vector, $\vec{v}_2$.

2. **Analyze $\vec{v}_2$:** We can imagine $\vec{v}_2$ as having two components:
    - A part that is **parallel** to $\vec{e}_1$ (this is the "shadow" of $\vec{v}_2$ on $\vec{e}_1$). This is the "dirty" or "unwanted" part because it's not orthogonal to $\vec{e}_1$.
    - A part that is **perpendicular** to $\vec{e}_1$. This is the "clean" and "original" part of $\vec{v}_2$ that we want to keep.
3. **Subtract the "Dirty" Part:**
    - The "dirty" part (parallel to $\vec{e}_1$) is the vector projection of $\vec{v}_2$ onto $\vec{e}_1$.
    - Projection Formula (since $\vec{e}_1$ is a unit vector): $\mathrm{proj}_{\vec{e}_1}(\vec{v}_2) = (\vec{v}_2 \cdot \vec{e}_1)\vec{e}_1$.
    - Now, we "clean" $\vec{v}_2$ by subtracting this dirty part from it:

    $$\vec{u}_2 = \vec{v}_2 - (\vec{v}_2 \cdot \vec{e}_1)\vec{e}_1$$

    - The resulting vector, $\vec{u}_2$, is now **guaranteed to be orthogonal** to $\vec{e}_1$.
4. **Clean Up the Length:** $\vec{u}_2$ has the correct direction, but its length is probably not 1. So, we normalize it.

    $$\vec{e}_2 = \frac{\vec{u}_2}{|\vec{u}_2|}$$

    We now have $\vec{e}_2$, which has a length of 1 and is perpendicular to $\vec{e}_1$.

Step 3: Create $\vec{e}_3$ (and so on...)

1. **Choose:** Take $\vec{v}_3$.
2. **Subtract the "Dirty" Parts:** $\vec{v}_3$ has a component parallel to $\vec{e}_1$ AND a component parallel to $\vec{e}_2$. We must subtract both.
    - Subtract the shadow on $\vec{e}_1$: $(\vec{v}_3 \cdot \vec{e}_1)\vec{e}_1$.
    - Subtract the shadow on $\vec{e}_2$: $(\vec{v}_3 \cdot \vec{e}_2)\vec{e}_2$.

    $$\vec{u}_3 = \vec{v}_3 - (\vec{v}_3 \cdot \vec{e}_1)\vec{e}_1 - (\vec{v}_3 \cdot \vec{e}_2)\vec{e}_2$$

    - $\vec{u}_3$ is now guaranteed to be orthogonal to both $\vec{e}_1$ and $\vec{e}_2$.
3. **Clean Up the Length:**

    $$\vec{e}_3 = \frac{\vec{u}_3}{|\vec{u}_3|}$$

Repeat this process for all remaining `v` vectors.

---

4. Conclusion

- The Gram-Schmidt process is a **constructive algorithm** that transforms any set of basis vectors into a "nice" orthonormal basis.

- This is extremely useful because orthonormal bases make many calculations (inverses, projections, change of basis) significantly easier.
- It is a fundamental tool in many areas of applied mathematics and machine learning.

---

5. Worked Example: Orthonormalizing a Basis

Problem:
We are given a "messy" basis `B` in $\mathbb{R}^2$:

- $\vec{v}_1 = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$
- $\vec{v}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

These vectors are linearly independent (not co-linear), so they form a valid basis. However, they are not orthogonal ($\vec{v}_1 \cdot \vec{v}_2 = 3$) and $\vec{v}_1$ is not a unit vector.

Goal:
Use the Gram-Schmidt Process to transform basis `B` into an orthonormal basis `E` = $\{\vec{e}_1, \vec{e}_2\}$.

Step 1: Create $\vec{e}_1$ (The First Orthonormal Basis Vector)

- **Recipe:** Take $\vec{v}_1$ and normalize it (make its length 1).

$$\vec{e}_1 = \frac{\vec{v}_1}{|\vec{v}_1|}$$

1. **Take $\vec{v}_1$:**
   $\vec{v}_1 = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$
2. **Calculate its length, $|\vec{v}_1|$:**
   $|\vec{v}_1| = \sqrt{3^2 + 0^2} = \sqrt{9} = 3$
3. **Divide $\vec{v}_1$ by its length**:

$$\vec{e}_1 = \frac{1}{3}\begin{bmatrix} 3 \\ 0 \end{bmatrix} = \begin{bmatrix} 3/3 \\ 0/3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

- **Result of Step 1:**
  We have found our first orthonormal basis vector:

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

(In this case, it happens to be the standard $\hat{\imath}$, because $\vec{v}_1$ was already on the x-axis).

Jawaban:

Panjang vektor $\vec{a} = |\vec{a}| = \sqrt{a_1{}^2 + a_2{}^2}$

Vektor satuan $\vec{a} = \frac{a}{|\vec{a}|} = \frac{1}{\sqrt{a_1{}^2 + a_2{}^2}}(a_1, a_2)$

Pembahasan:

Vektor merupakan besaran yang punya nilai dan arah. Untuk menentukan panjang vektor kita menggunakan rumus:

$$\text{Panjang vektor } \vec{a} = |\vec{a}| = \sqrt{a_1{}^2 + a_2{}^2}$$

Sedangkan vektor satuan adalah suatu vektor yang panjangnya satu satuan. Untuk menentukan vektor satuan kita menggunakan rumus:

$$\text{Vektor satuan } \vec{a} = \frac{a}{|\vec{a}|} = \frac{1}{\sqrt{a_1{}^2 + a_2{}^2}}(a_1, a_2)$$

Jadi, cara menentukan panjang vektor adalah Panjang vektor $\vec{a} = |\vec{a}| = \sqrt{a_1{}^2 + a_2{}^2}$. Sedangkan untuk menentukan vektor satuan adalah Vektor satuan $\vec{a} = \frac{a}{|\vec{a}|} = \frac{1}{\sqrt{a_1{}^2 + a_2{}^2}}(a_1, a_2)$.

Step 2: Create $\vec{e}_2$ (The Second Orthonormal Basis Vector)

- **Recipe:**
  1. Take $\vec{v}_2$.
  2. "Clean" $\vec{v}_2$ by subtracting its "shadow" on $\vec{e}_1$ to get $\vec{u}_2$.

  $$\vec{u}_2 = \vec{v}_2 - \text{proj}_{\vec{e}_1}(\vec{v}_2) = \vec{v}_2 - (\vec{v}_2 \cdot \vec{e}_1)\vec{e}_1$$

  3. Normalize $\vec{u}_2$ to get $\vec{e}_2$.

  $$\vec{e}_2 = \frac{\vec{u}_2}{|\vec{u}_2|}$$

Let's do the work.

1. **Take $\vec{v}_2$**:
   $\vec{v}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$
2. **Calculate the "dirty" part (the projection of $\vec{v}_2$ onto $\vec{e}_1$)**:
   - First, calculate the dot product $\vec{v}_2 \cdot \vec{e}_1$:

   $$\begin{bmatrix} 1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} = (1 \cdot 1) + (2 \cdot 0) = 1$$

   - The projection vector is $(\vec{v}_2 \cdot \vec{e}_1)\vec{e}_1$:

$$1 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

3. **"Clean" $\vec{v}_2$ to get $\vec{u}_2$:**

$$\vec{u}_2 = \vec{v}_2 - (\text{projection vector})$$

$$\vec{u}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1-1 \\ 2-0 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$

The vector $\vec{u}_2 = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$ is now guaranteed to be orthogonal to $\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. (Check it: $[1,0] \cdot [0,2] = 0$. It works!).

4. **Normalize $\vec{u}_2$ to get $\vec{e}_2$:**
   - Calculate the length of $\vec{u}_2$:
     $|\vec{u}_2| = \sqrt{0^2 + 2^2} = \sqrt{4} = 2.$
   - Divide $\vec{u}_2$ by its length:

$$\vec{e}_2 = \frac{1}{2} \begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 0/2 \\ 2/2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

- **Result of Step 2:**
  We have found our second orthonormal basis vector:

$$\vec{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Final Solution

The new orthonormal basis we constructed from $\{[3,0],[1,2]\}$ is:

$$E = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

In this simple example, the Gram-Schmidt process transformed a "skewed" basis back into the familiar standard basis. For more complex initial bases, the result would be a set of unit vectors that are mutually orthogonal but may not align with the x and y axes.

Tags: #mml-specialization #linear-algebra #gram-schmidt #orthonormal-basis #orthogonalization