Eigenvalue eigenvector

Instructor: Dr. Avijit Pal

Linear algebra- II (IC152)

- Let V be a finite dimensional inner product space. Suppose $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ is a linearly independent subset of V.
- Then the Gram-Schmidt orthogonalisation process uses the vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ to construct new vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ such that

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0 \text{ for } i \neq j, \|\mathbf{v}_i\| = 1$$

$$\text{and}$$

$$\text{Span } \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_i\} = \text{Span } \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_i\} \text{ for } i = 1, 2, \ldots, n.$$

- This process proceeds with the following idea.

Gram-Schmidt orthogonalisation process cont.

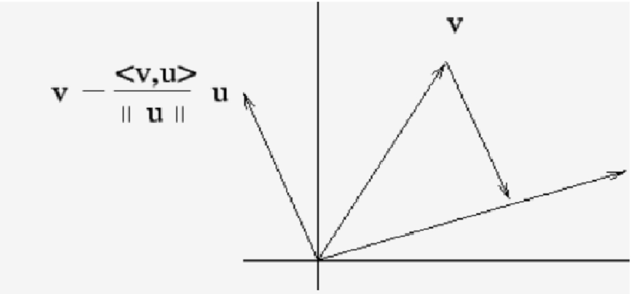

Figure 5.1: Gram-Schmidt Process

- Suppose we are given two vectors $\mathbf{u}$ and $\mathbf{v}$ in a plane.
- If we want to get vectors $\mathbf{z}$ and $\mathbf{y}$ such that $\mathbf{z}$ is a unit vector in the direction of $\mathbf{u}$ and $\mathbf{y}$ is a unit vector perpendicular to $\mathbf{z}$, then they can be obtained in the following way:

- Take the first vector $\mathbf{z} = \dfrac{\mathbf{u}}{\|\mathbf{u}\|}$.
- Let θ be the angle between the vectors $\mathbf{u}$ and $\mathbf{v}$.
- Then $\cos(\theta) = \dfrac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|u\| \ \|v\|}$.
- Defined $\alpha = \|\mathbf{v}\| \ \cos(\theta) = \dfrac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|} = \langle \mathbf{z}, \mathbf{v} \rangle$.
- Then $\mathbf{w} = \mathbf{v} - \alpha \, \mathbf{z}$ is a vector perpendicular to the unit vector $\mathbf{z}$, as we have removed the component of $\mathbf{z}$ from $\mathbf{v}$.
- So, the vectors that we are interested in are $\mathbf{z}$ and $\mathbf{y} = \dfrac{\mathbf{w}}{\|\mathbf{w}\|}$.
- This idea is used to give the Gram-Schmidt Orthogonalisation process which we now describe.

Theorem (Gram-Schmidt Orthogonalisation Process)

Let V be an inner product space. Suppose $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ is a set of linearly independent vectors of V. Then there exists a set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ of vectors of V satisfying the following:

- $\|\mathbf{v}_i\| = 1$ *for* $1 \leq i \leq n$,
- $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ *for* $1 \leq i,j \leq n,\ i \neq j$ *and*
- $L(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_i) = L(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_i)$ *for* $1 \leq i \leq n$.

Outline of the proof

- We successively define the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ as follows.

$$\mathbf{v}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}.$$

- First, we calculate

$$\mathbf{w}_2 = \mathbf{u}_2 - \langle \mathbf{u}_2, \mathbf{v}_1 \rangle \mathbf{v}_1,$$

  and let $\mathbf{v}_2 = \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|}$. In this process we get $\{v_1, v_2\}$.

- To obtain w_3, we need to calculate

$$\mathbf{w}_3 = \mathbf{u}_3 - \langle \mathbf{u}_3, \mathbf{v}_1 \rangle \mathbf{v}_1 - \langle \mathbf{u}_3, \mathbf{v}_2 \rangle \mathbf{v}_2,$$

  and let $\mathbf{v}_3 = \frac{\mathbf{w}_3}{\|\mathbf{w}_3\|}$. In this way we also get $\{v_1, v_2, v_3\}$.

- In general, if $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \ldots, \mathbf{v}_{i-1}$ are already obtained, we compute w_i

$$\mathbf{w}_i = \mathbf{u}_i - \langle \mathbf{u}_i, \mathbf{v}_1 \rangle \mathbf{v}_1 - \cdots - \langle \mathbf{u}_i, \mathbf{v}_{i-1} \rangle \mathbf{v}_{i-1}, \tag{1}$$

Outline of the proof cont.

- We define $\mathbf{v}_i = \dfrac{\mathbf{w}_i}{\|\mathbf{w}_i\|}$.
- We prove the theorem by induction on n, the number of linearly independent vectors.
- For $n = 1$, we have

$$\mathbf{v}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}.$$

- Since $\mathbf{u}_1 \neq \mathbf{0}$, $\mathbf{v}_1 \neq \mathbf{0}$ and

$$\|\mathbf{v}_1\|^2 = \langle \mathbf{v}_1, \mathbf{v}_1 \rangle = \langle \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}, \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} \rangle = \frac{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle}{\|\mathbf{u}_1\|^2} = 1.$$

- Hence, the result holds for $n = 1$.
- Let the result hold for all $k \leq n - 1$. That is, suppose we are given any set of k, $1 \leq k \leq n - 1$ linearly independent vectors $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ of V.

Outline of the proof cont.

- Then by the inductive assumption, there exists a set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ of vectors satisfying the following:
  1. $\|\mathbf{v}_i\| = 1$ for $1 \leq i \leq k$,
  2. $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ for $1 \leq i \neq j \leq k$, and
  3. $L(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_i) = L(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_i)$ for $1 \leq i \leq k$.
- Now, let us assume that we are given a set of n linearly independent vectors $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ of V.
- Then by the inductive assumption, we already have vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}$ satisfying
  1. $\|\mathbf{v}_i\| = 1$ for $1 \leq i \leq n-1$,
  2. $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ for $1 \leq i \neq j \leq n-1$, and
  3. $L(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_i) = L(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_i)$ for $1 \leq i \leq n-1$.

- Using (1), we define

$$\mathbf{w}_n = \mathbf{u}_n - \langle \mathbf{u}_n, \mathbf{v}_1 \rangle \mathbf{v}_1 - \cdots - \langle \mathbf{u}_n, \mathbf{v}_{n-1} \rangle \mathbf{v}_{n-1}. \tag{2}$$

- We first show that $\mathbf{w}_n \notin L(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1})$. This will also imply that $\mathbf{w}_n \neq \mathbf{0}$ and hence $\mathbf{v}_n = \dfrac{\mathbf{w}_n}{\|\mathbf{w}_n\|}$ is well defined.
- On the contrary, assume that $\mathbf{w}_n \in L(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1})$. Then there exist scalars $\alpha_1, \alpha_2, \ldots, \alpha_{n-1}$ such that

$$\mathbf{w}_n = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_{n-1} \mathbf{v}_{n-1}.$$

- So, by (2)

$$\mathbf{u}_n = \big(\alpha_1 + \langle \mathbf{u}_n, \mathbf{v}_1 \rangle\big) + \cdots + \big(\alpha_n + \langle \mathbf{u}_n, \mathbf{v}_{n-1} \rangle\big)\mathbf{v}_{n-1}.$$

- Thus, by the third induction assumption,

$$\mathbf{u}_n \in L(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}) = L(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{n-1}).$$

- This gives a contradiction to the given assumption that the set of vectors $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ is linear independent.
- So, $\mathbf{w}_n \neq \mathbf{0}$. Define $\mathbf{v}_n = \dfrac{\mathbf{w}_n}{\|\mathbf{w}_n\|}$.
- Then $\|\mathbf{v}_n\| = 1$. Also, it can be easily verified that $\langle \mathbf{v}_n, \mathbf{v}_i \rangle = 0$ for $1 \leq i \leq n-1$.
- Hence, by the principle of mathematical induction, the proof of the theorem is complete.

Example

- Let $\{(1,-1,1,1),(1,0,1,0),(0,1,0,1)\}$ be a linearly independent set in $\mathbb{R}^4(\mathbb{R})$.
- We will find an orthonormal set $\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\}$ such that $L(\ (1,-1,1,1),(1,0,1,0),(0,1,0,1)\ ) = L(\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3)$.
- Let $\mathbf{u}_1 = (1,0,1,0)$. Define $\mathbf{v}_1 = \dfrac{(1,0,1,0)}{\sqrt{2}}$. Let $\mathbf{u}_2 = (0,1,0,1)$.
- Then

$$\mathbf{w}_2 = (0,1,0,1) - \langle (0,1,0,1), \frac{(1,0,1,0)}{\sqrt{2}} \rangle \mathbf{v}_1 = (0,1,0,1).$$

- Hence, $\mathbf{v}_2 = \dfrac{(0,1,0,1)}{\sqrt{2}}$. Let $\mathbf{u}_3 = (1,-1,1,1)$.

- Then

$$\begin{aligned}\mathbf{w}_3 &= (1,-1,1,1) - \langle (1,-1,1,1), \frac{(1,0,1,0)}{\sqrt{2}} \rangle \mathbf{v}_1 \\ &- \langle (1,-1,1,1), \frac{(0,1,0,1)}{\sqrt{2}} \rangle \mathbf{v}_2 \\ &= (0,-1,0,1).\end{aligned}$$

- Also, $\mathbf{v}_3 = \dfrac{(0,-1,0,1)}{\sqrt{2}}$.

Observation: 1

- Let $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ be any basis of a k -dimensional subspace W of $\mathbb{R}^n$.
- Then by Gram-Schmidt orthogonalisation process, we get an orthonormal set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\} \subset \mathbb{R}^n$ with $W = L(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$.
- For $1 \leq i \leq k$,

$$L(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_i) = L(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_i).$$

Oveservation: 2

- Suppose we are given a set of n vectors, $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ of V that are linearly dependent.
- Then there exists a smallest $k,\ 2 \leq k \leq n$ such that

$$L(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k) = L(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{k-1}).$$

- We claim that in this case, $\mathbf{w}_k = \mathbf{0}$.
- Since, we have chosen the smallest k satisfying

$$L(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_i) = L(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{i-1}),$$

  for $2 \leq i \leq n$, the set $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{k-1}\}$ is linearly independent.
- So, by above Theorem, there exists an orthonormal set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{k-1}\}$ such that

$$L(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{k-1}) = L(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{k-1}).$$

- As $\mathbf{u}_k \in L(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{k-1})$, by previous observation

$$\mathbf{u}_k = \langle \mathbf{u}_k, \mathbf{v}_1 \rangle + \cdots + \langle \mathbf{u}_k, \mathbf{v}_{k-1} \rangle \mathbf{v}_{n-1}.$$

- So, by definition of $\mathbf{w}_k,\ \mathbf{w}_k = \mathbf{0}$.
- Therefore, in this case, we can continue with the Gram-Schmidt process by replacing $\mathbf{u}_k$ by $\mathbf{u}_{k+1}$.

Observation:3

- Let S be a countably infinite set of linearly independent vectors. Then one can apply the Gram-Schmidt process to get a countably infinite orthonormal set.

Thank You