MA110:
Lecture 13

Saurav Bhaumik
Department of Mathematics
IIT Bombay
Spring 2025

Inner Product and Norm

Let $\mathbb{K} := \mathbb{R}$, the set of real numbers, or $\mathbb{K} := \mathbb{C}$, the set of complex numbers. For a scalar $\alpha \in \mathbb{K}$, we denote its conjugate by $\overline{\alpha}$. If $\alpha \in \mathbb{R}$, then of course, $\overline{\alpha} = \alpha$.

Consider column vectors $\mathbf{x} := \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ and $\mathbf{y} := \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$ in $\mathbb{K}^{n \times 1}$.

The conjugate transpose (or the adjoint) $\mathbf{x}^* := \begin{bmatrix} \overline{x}_1 & \cdots & \overline{x}_n \end{bmatrix}$ of $\mathbf{x}$ is a row vector in $\mathbb{K}^{1 \times n}$. The **inner product** of $\mathbf{x}$ with $\mathbf{y}$ is defined by

$$\langle \mathbf{x}, \mathbf{y} \rangle := \mathbf{x}^* \mathbf{y} = \overline{x}_1 y_1 + \cdots + \overline{x}_n y_n.$$

Note: If $\mathbb{K} = \mathbb{R}$, then $\langle \mathbf{x}, \mathbf{y} \rangle$ is just the scalar product of $\boldsymbol{x} := (x_1, \ldots, x_n)$ and $\boldsymbol{y} := (y_1, \ldots, y_n)$ in $\mathbb{R}^n$.

The inner product function $\langle \cdot\,, \cdot \rangle : \mathbb{K}^{n\times 1} \times \mathbb{K}^{n\times 1} \to \mathbb{K}$ has the following crucial properties. For $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{K}^{n\times 1}$ and $\alpha, \beta \in \mathbb{K}$,

1. $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ and $\langle \mathbf{x}, \mathbf{x} \rangle = 0 \iff \mathbf{x} = \mathbf{0}$ (positive definite),
2. $\langle \mathbf{x}, \alpha\mathbf{y} + \beta\mathbf{z} \rangle = \alpha\langle \mathbf{x}, \mathbf{y} \rangle + \beta\langle \mathbf{x}, \mathbf{z} \rangle$ (linear in 2nd variable),
3. $\langle \mathbf{y}, \mathbf{x} \rangle = \overline{\langle \mathbf{x}, \mathbf{y} \rangle}$ (conjugate symmetric).

From the above three crucial properties, conjugate linearity in the 1st variable follows: $\langle \alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z} \rangle = \overline{\alpha}\langle \mathbf{x}, \mathbf{z} \rangle + \overline{\beta}\langle \mathbf{y}, \mathbf{z} \rangle$.

Let $\mathbf{x} := \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}^{\mathsf{T}} \in \mathbb{K}^{n\times 1}$. We define the **norm** of $\mathbf{x}$ by

$$\|\mathbf{x}\| := \langle \mathbf{x}, \mathbf{x} \rangle^{1/2} = (|x_1|^2 + \cdots + |x_n|^2)^{1/2}.$$

For $n = 1$, the norm of $x \in \mathbb{K}$ is the absolute value $|x|$ of x.

Clearly, $\max\{|x_1|, \ldots, |x_m|\} \leq \|\mathbf{x}\| \leq |x_1| + \cdots + |x_m|$.

If $\mathbf{x} \in \mathbb{K}^{n\times 1}$ and $\|\mathbf{x}\| = 1$, then we say that $\mathbf{x}$ is a **unit vector**.

Theorem

Let $\mathbf{x}, \mathbf{y} \in \mathbb{K}^{n\times 1}$. Then
(i) (Schwarz Inequality) $|\langle \mathbf{x}, \mathbf{y}\rangle| \le \|\mathbf{x}\|\|\mathbf{y}\|$.

(ii) (Triangle Inequality) $\|\mathbf{x}+\mathbf{y}\| \le \|\mathbf{x}\| + \|\mathbf{y}\|$.

Proof. Suppose $\mathbf{x} := \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}^{\mathsf{T}}$ and $\mathbf{y} := \begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix}^{\mathsf{T}}$.

(i) If $\|\mathbf{x}\| = 0$ or $\|\mathbf{y}\| = 0$, then $\mathbf{x} = \mathbf{0}$ or $\mathbf{y} = \mathbf{0}$. Hence we are done. Now let $\|\mathbf{x}\| \neq 0$ and $\|\mathbf{y}\| \neq 0$. Then

$$\frac{|\overline{x}_j|}{\|\mathbf{x}\|}\frac{|y_j|}{\|\mathbf{y}\|} \le \frac{1}{2}\Big(\frac{|x_j|^2}{\|\mathbf{x}\|^2} + \frac{|y_j|^2}{\|\mathbf{y}\|^2}\Big) \quad \text{for } j = 1, \ldots, n,$$

since $|\overline{\alpha}\beta| = |\alpha|\,|\beta| \le (|\alpha|^2 + |\beta|^2)/2$ for all $\alpha, \beta \in \mathbb{K}$. Hence

$$|\langle \mathbf{x}, \mathbf{y}\rangle| \le \sum_{j=1}^{n} |\overline{x}_j|\,|y_j| \le \frac{\|\mathbf{x}\|\|\mathbf{y}\|}{2}(1+1) = \|\mathbf{x}\|\|\mathbf{y}\|.$$

(ii) Since $\langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle = 2\,\Re\,\langle \mathbf{x}, \mathbf{y} \rangle$, we see that

$$\begin{aligned} \|\mathbf{x}+\mathbf{y}\|^2 &= \langle \mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y} \rangle = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\,\Re\,\langle \mathbf{x}, \mathbf{y} \rangle \\ &\leq \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\,|\langle \mathbf{x}, \mathbf{y} \rangle| \\ &\leq \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\,\|\mathbf{x}\|\|\mathbf{y}\| \quad \text{(by the Schwarz inequality)} \\ &= (\|\mathbf{x}\| + \|\mathbf{y}\|)^2. \end{aligned}$$

Thus $\|\boldsymbol{x}+\boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$. $\square$

We observe that the norm function $\|\cdot\| : \mathbb{K}^{n\times 1} \to \mathbb{K}$ satisfies the following three crucial properties:

(i) $\|\mathbf{x}\| \geq 0$ for all $\mathbf{x} \in \mathbb{K}^{n\times 1}$ and $\|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$,

(ii) $\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\|$ for all $\alpha \in \mathbb{K}$ and $\mathbf{x} \in \mathbb{K}^{n\times 1}$,

(iii) $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{K}^{n\times 1}$.

The properties of the norm function allow us to define the distance between two vectors in $\mathbb{K}^{n\times 1}$. Let $\mathbf{x}, \mathbf{y} \in \mathbb{K}^{n\times 1}$. Then the **distance** between $\mathbf{x}$ and $\mathbf{y}$ is defined by

$$d(\mathbf{x}, \mathbf{y}) := \|\mathbf{x} - \mathbf{y}\|.$$

The distance function $d : \mathbb{K}^{n\times 1} \times \mathbb{K}^{n\times 1} \to \mathbb{K}$ has the following analogous properties.

(i) $d(\mathbf{x}, \mathbf{y}) \geq 0$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{K}^{n\times 1}$, $d(\mathbf{x}, \mathbf{y}) = 0 \iff \mathbf{x} = \mathbf{y}$,

(ii) $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{K}^{n\times 1}$,

(iii) $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$ for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{K}^{n\times 1}$.

The inner product defined earlier allows us to say when two column vectors are perpendicular to each other.

Orthogonality

Let $\mathbf{x}, \mathbf{y} \in \mathbb{K}^{n\times 1}$. We say that $\mathbf{x}$ and $\mathbf{y}$ are **orthogonal** (to each other) if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$, and then we write $\mathbf{x} \perp \mathbf{y}$.

Clearly, $\mathbf{x} \perp \mathbf{x} \iff \|\mathbf{x}\| = \mathbf{0} \iff \mathbf{x} = \mathbf{0}$.

Let E be a subset of $\mathbb{K}^{n\times 1}$, and define

$$E^{\perp} := \{\mathbf{y} \in \mathbb{K}^{n\times 1} : \mathbf{y} \perp \mathbf{x} \text{ for all } \mathbf{x} \in E\}.$$

It is easy to see that $E^{\perp}$ is a subspace of $\mathbb{K}^{n\times 1}$.

Proposition (Pythagoras Theorem)

Let $\mathbf{x}, \mathbf{y} \in \mathbb{K}^{n\times 1}$. If $\mathbf{x} \perp \mathbf{y}$, then $\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$.

Proof.

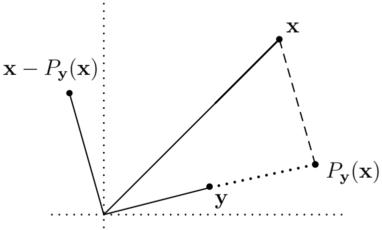
$$\begin{aligned}
\|\mathbf{x}+\mathbf{y}\|^2 &= \langle \mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y} \rangle = \langle \mathbf{x}+\mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{x}+\mathbf{y}, \mathbf{y} \rangle \\
&= \langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle \\
&= \|\mathbf{x}\|^2 + 0 + 0 + \|\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2.
\end{aligned}$$

□

We now introduce an important concept.
Let $\mathbf{y}$ be a nonzero vector in $\mathbb{K}^{n\times 1}$. For $\mathbf{x} \in \mathbb{K}^{n\times 1}$, define

$$P_{\mathbf{y}}(\mathbf{x}) := \frac{\langle \mathbf{y},\, \mathbf{x} \rangle}{\langle \mathbf{y},\, \mathbf{y} \rangle}\, \mathbf{y}.$$

It is called the (perpendicular) **projection** of the vector $\mathbf{x}$ in the direction of the vector $\mathbf{y}$. Note that $P_{\mathbf{y}} : \mathbb{K}^{n\times 1} \to \mathbb{K}^{n\times 1}$ is a linear map and its image space is one dimensional. Also, $P_{\mathbf{y}}(\mathbf{y}) = \mathbf{y}$, so that $(P_{\mathbf{y}})^2 := P_{\mathbf{y}} \circ P_{\mathbf{y}} = P_{\mathbf{y}}$.



Note that $P_{\mathbf{y}}(\mathbf{x})$ is a scalar multiple of $\mathbf{y}$ for every $\mathbf{x} \in \mathbb{K}^{n\times 1}$.

An important property of the projection of a vector in the direction of another (nonzero) vector is the following:

Proposition

Let $\mathbf{y} \in \mathbb{K}^{n\times 1}$ be nonzero. Then for every $\mathbf{x} \in \mathbb{K}^{n\times 1}$,

$$\big(\mathbf{x} - P_{\mathbf{y}}(\mathbf{x})\big) \perp \mathbf{y}.$$

Proof. Let $\mathbf{x} \in \mathbb{K}^{n\times 1}$. The result follows from

$$\langle \mathbf{y},\, \mathbf{x} - P_{\mathbf{y}}(\mathbf{x})\rangle = \langle \mathbf{y},\, \mathbf{x}\rangle - \langle \mathbf{y},\, P_{\mathbf{y}}(\mathbf{x})\rangle = \langle \mathbf{y},\, \mathbf{x}\rangle - \frac{\langle \mathbf{y},\, \mathbf{x}\rangle}{\langle \mathbf{y},\, \mathbf{y}\rangle}\langle \mathbf{y},\, \mathbf{y}\rangle = 0. \quad \square$$

Let E be a subset of $\mathbb{K}^{n\times 1}$. Then E is said to be **orthogonal** if any two (distinct) element of E are orthogonal (to each other), that is, $\mathbf{x} \perp \mathbf{y}$ for all $\mathbf{x}, \mathbf{y}$ in E with $\mathbf{x} \neq \mathbf{y}$.

For example, $E := \{\mathbf{0}, \mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_1 - \mathbf{e}_2\}$ is an orthogonal subset of $\mathbb{K}^{n\times 1}$.

Proposition

Let E be a subset of $\mathbb{K}^{n\times 1}$. If E is orthogonal and if $\mathbf{0} \notin E$, then E is linearly independent.

Proof. Let $\mathbf{x}_1, \ldots, \mathbf{x}_k$ be distinct vectors in E, and let $\alpha_1, \ldots, \alpha_k$ be scalars such that $\alpha_1\mathbf{x}_1 + \cdots + \alpha_k\mathbf{x}_k = \mathbf{0}$. Fix $j \in \{1, \ldots, k\}$. Since $\langle \mathbf{x}_j, \mathbf{x}_i \rangle = 0$ for all $i \neq j$, we obtain

$$0 = \langle \mathbf{x}_j, \alpha_1\mathbf{x}_1 + \cdots + \alpha_k\mathbf{x}_k \rangle = \sum_{\ell=1}^{k} \alpha_\ell \langle \mathbf{x}_j, \mathbf{x}_\ell \rangle = \alpha_j \langle \mathbf{x}_j, \mathbf{x}_j \rangle.$$

But $\langle \mathbf{x}_j, \mathbf{x}_j \rangle \neq 0$ since $\mathbf{x}_j \neq \mathbf{0}$. Hence $\alpha_j = 0$. $\square$

The converse of the above proposition is not true, that is, a linear linearly independent subset of $\mathbb{K}^{n\times 1}$ need not be orthogonal. For example, the subset $\{\mathbf{e}_1, \mathbf{e}_1 + \mathbf{e}_2\}$ of $\mathbb{K}^{n\times 1}$ is linearly independent, but not orthogonal.

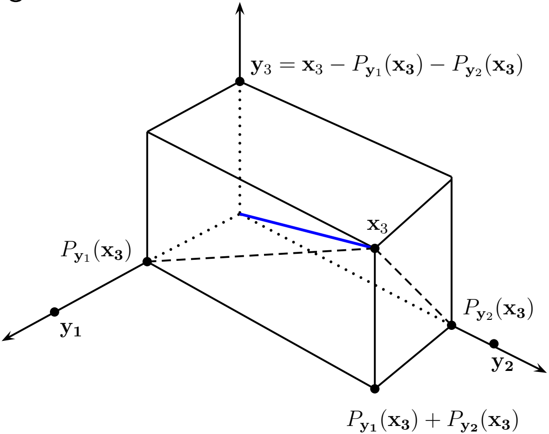
We now address the question: Can we modify a linearly independent set E to construct an orthogonal set, retaining the span of the elements in E at each step of the procedure?

Let E be an ordered linearly independent set of column vectors. Suppose $\mathbf{x}_1$ is the first vector in E. Then $\mathbf{x}_1 \neq 0$. Let $\mathbf{y}_1 := \mathbf{x}_1$. Let $\mathbf{x}_2$ be the second vector in E. If $\mathbf{x}_2$ is not orthogonal to $\mathbf{y}_1$, then it makes sense to subtract from $\mathbf{x}_2$, the projection of $\mathbf{x}_2$ in the direction of $\mathbf{y}_1$, so that $\mathbf{y}_2 := \mathbf{x}_2 - P_{\mathbf{y}_1}(\mathbf{x}_2)$ is orthogonal to $\mathbf{y}_1$. Also, in replacing $\mathbf{x}_2$ by $\mathbf{y}_2$, we do not alter the span of $\{\mathbf{x}_1, \mathbf{x}_2\}$ since $\mathbf{y}_2$ is a linear combination of $\mathbf{x}_2$ and $\mathbf{y}_1$, and $\mathbf{x}_2$ is a linear combination of $\mathbf{y}_2$ and $\mathbf{y}_1$, where $\mathbf{y}_1 = \mathbf{x}_1$.

Let $\mathbf{x}_3$ be the third vector in E. If $\mathbf{x}_3$ is not orthogonal to $\mathbf{y}_1$ and $\mathbf{y}_2$, then we may subtract from $\mathbf{x}_3$, the projections of $\mathbf{x}_3$ in the directions of $\mathbf{y}_1$ and $\mathbf{y}_2$. Then $\mathbf{y}_3 := \mathbf{x}_3 - P_{\mathbf{y}_1}(\mathbf{x}_3) - P_{\mathbf{y}_2}(\mathbf{x}_3)$ is orthogonal to $\mathbf{y}_1$ as well as to $\mathbf{y}_2$. We can see this as follows.

$$\langle \mathbf{y}_1,\, \mathbf{y}_3 \rangle = \langle \mathbf{y}_1,\, \mathbf{x}_3 - P_{\mathbf{y}_1}(\mathbf{x}_3) \rangle - \langle \mathbf{y}_1,\, P_{\mathbf{y}_2}(\mathbf{x}_3) \rangle = 0 - 0 = 0,$$

and similarly $\langle \mathbf{y}_3,\, \mathbf{y}_2 \rangle = 0$. We illustrate these vectors in the following figure.



This procedure can be continued to yield the famous

Gram-Schmidt Orthogonalization Process (G-S OP)

Let $(\mathbf{x}_1, \ldots, \mathbf{x}_k)$ be an ordered linearly independent set in $\mathbb{K}^{n\times 1}$. Define $\mathbf{y}_1 := \mathbf{x}_1$.
Let $1 \leq j < k$. Suppose we have found $\mathbf{y}_1, \ldots, \mathbf{y}_j$ in $\mathbb{K}^{n\times 1}$ such that the set $\{\mathbf{y}_1, \ldots, \mathbf{y}_j\}$ is orthogonal, and also $\operatorname{span}\{\mathbf{y}_1, \ldots, \mathbf{y}_j\} = \operatorname{span}\{\mathbf{x}_1, \ldots, \mathbf{x}_j\}$. Define

$$\mathbf{y}_{j+1} := \mathbf{x}_{j+1} - P_{\mathbf{y}_1}(\mathbf{x}_{j+1}) - \cdots - P_{\mathbf{y}_j}(\mathbf{x}_{j+1}).$$

Then $\operatorname{span}\{\mathbf{y}_1, \ldots, \mathbf{y}_{j+1}\} = \operatorname{span}\{\mathbf{x}_1, \ldots, \mathbf{x}_{j+1}\}$ since $\mathbf{y}_{j+1} \in \operatorname{span}\{\mathbf{y}_1, \ldots, \mathbf{y}_j, \mathbf{x}_{j+1}\} = \operatorname{span}\{\mathbf{x}_1, \ldots, \mathbf{x}_j, \mathbf{x}_{j+1}\}$ and $\mathbf{x}_{j+1} \in \operatorname{span}\{\mathbf{y}_1, \ldots, \mathbf{y}_j, \mathbf{y}_{j+1}\}$.

To show that the set $\{\mathbf{y}_1, \ldots, \mathbf{y}_{j+1}\}$ is orthogonal, it is enough to show that $\mathbf{y}_{j+1} \in \{\mathbf{y}_1, \ldots, \mathbf{y}_j\}^{\perp}$.

Let $i \in \{1, \ldots, j\}$. Then

$$\begin{aligned}
\langle \mathbf{y}_i, \mathbf{y}_{j+1} \rangle &= \langle \mathbf{y}_i, \mathbf{x}_{j+1} - P_{\mathbf{y}_1}(\mathbf{x}_{j+1}) - \cdots - P_{\mathbf{y}_j}(\mathbf{x}_{j+1}) \rangle \\
&= \langle \mathbf{y}_i, \mathbf{x}_{j+1} \rangle - \langle \mathbf{y}_i, P_{\mathbf{y}_1}(\mathbf{x}_{j+1}) \rangle - \cdots - \langle \mathbf{y}_i, P_{\mathbf{y}_j}(\mathbf{x}_{j+1}) \rangle \\
&= \langle \mathbf{y}_i, \mathbf{x}_{j+1} \rangle - \langle \mathbf{y}_i, P_{\mathbf{y}_i}(\mathbf{x}_{j+1}) \rangle \quad (\text{since } \mathbf{y}_i \perp \mathbf{y}_j,\ i \neq j) \\
&= \langle \mathbf{y}_i, \mathbf{x}_{j+1} - P_{\mathbf{y}_i}(\mathbf{x}_{j+1}) \rangle \\
&= 0 \quad (\text{by the important property of the projection}).
\end{aligned}$$

We remark that since the set $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ is linearly independent, all vectors $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k$ constructed in the G-S OP are nonzero: Clearly, $\mathbf{y}_1 = \mathbf{x}_1 \neq \mathbf{0}$. Also, if $\mathbf{y}_{j+1} = \mathbf{0}$ for some $j \in \{1, \ldots, k-1\}$, then $\mathbf{x}_{j+1}$ would belong to $\text{span}\{\mathbf{y}_1, \ldots, \mathbf{y}_j\} = \text{span}\{\mathbf{x}_1, \ldots, \mathbf{x}_j\}$.

This completes the construction of the G-S OP.

An orthogonal set whose elements are unit vectors is called an **orthonormal set**.

Any orthogonal set whose elements are nonzero vectors can always be turned into an orthonormal set by dividing each element by its own norm.

Thus given an ordered linearly independent set $(\mathbf{x}_1, \ldots, \mathbf{x}_k)$, we can construct an ordered orthogonal set $(\mathbf{y}_1, \ldots, \mathbf{y}_k)$ by the G-S OP, and if we let $\mathbf{u}_j := \mathbf{y}_j / \|\mathbf{y}_j\|$ for $j = 1, \ldots, k$, then $(\mathbf{u}_1, \ldots, \mathbf{u}_k)$ is an ordered orthonormal set such that $\mathrm{span}\{\mathbf{x}_1, \ldots, \mathbf{x}_k\} = \mathrm{span}\{\mathbf{y}_1, \ldots, \mathbf{y}_k\} = \mathrm{span}\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$.

Example

For $j = 1, \ldots, n$, let $\mathbf{x}_j := j(\mathbf{e}_1 + \cdots + \mathbf{e}_j)$. Then $E := (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ is an ordered linearly independent subset of $\mathbb{K}^{n \times 1}$. We claim that the G-S OP gives $\mathbf{y}_j := j\mathbf{e}_j$ for $j = 1, \ldots, n$. Indeed, $\mathbf{y}_1 := \mathbf{x}_1 = \mathbf{e}_1$. Also, assuming that $\mathbf{y}_j = j\mathbf{e}_j$, we see that

$$
\begin{aligned}
\mathbf{y}_{j+1} &= \mathbf{x}_{j+1} - P_{\mathbf{y}_1}(\mathbf{x}_{j+1}) - \cdots - P_{\mathbf{y}_j}(\mathbf{x}_{j+1}) \\
&= (j+1)(\mathbf{e}_1 + \cdots + \mathbf{e}_{j+1}) - (j+1)\mathbf{e}_1 - \cdots - (j+1)\mathbf{e}_j \\
&= (j+1)\mathbf{e}_{j+1}.
\end{aligned}
$$

Hence our claim is justified. Since $\|\mathbf{y}_j\| = j$ for each j, we let $\mathbf{u}_j := \mathbf{y}_j/j$, so that $\mathbf{u}_j = \mathbf{e}_j$ for each $j = 1, \ldots, n$. Clearly, $(\mathbf{u}_1, \ldots, \mathbf{u}_k)$ is an ordered orthonormal set in $\mathbb{K}^{n\times 1}$.

Let V be a subspace of $\mathbb{K}^{n\times 1}$. An **orthonormal basis** for V is a basis for V which is an orthonormal subset of V.

The G-S OP enables us to modify a given basis for a subspace of $\mathbb{K}^{n\times 1}$ to an orthonormal basis for that subspace.

Also, we can expand an orthonormal set in V to a possibly larger orthonormal set in V as follows.

Proposition

Let V be a subspace of $\mathbb{K}^{n\times 1}$, and let $\mathbf{u}_1, \ldots, \mathbf{u}_k$ be an orthonormal set in V. Then there is an orthonormal basis for V which contains $\mathbf{u}_1, \ldots, \mathbf{u}_k$.

Proof. If $\text{span}\{\mathbf{u}_1, \ldots, \mathbf{u}_k\} = V$, then there is nothing to prove.

Now suppose $\text{span}\{\mathbf{u}_1, \ldots, \mathbf{u}_k\} \neq V$. Let $\dim V = r$. Since the set $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\} \neq V$ is linearly independent, there are $\mathbf{y}_{k+1}, \ldots, \mathbf{y}_r$ in V such that $\{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{y}_{k+1}, \ldots, \mathbf{y}_r\}$ is a basis for V. By the G-S OP, we can find $\mathbf{u}_{k+1}, \ldots, \mathbf{u}_r$ in V such that the set $\{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_r\}$ is orthonormal and its span is equal to $\text{span}\{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{y}_{k+1}, \ldots, \mathbf{y}_r\} = V$. $\square$

Corollary

Every nonzero vector subspace V has an orthonormal basis.

Proof. If $\mathbf{0} \neq \mathbf{x}_1 \in V$, then extend $\{\mathbf{x}_1/\|\mathbf{x}_1\|\}$ to an o. n. basis.