

Orthogonal Projection Review

©Prakash Ishwar

Department of Electrical and Computer Engineering
Boston University

September 7, 2015

Inner Product

- Let $\mathbf{u} = (u_1, \dots, u_n)^\top$ and $\mathbf{v} = (v_1, \dots, v_n)^\top$ be two vectors in n -dimensional real Euclidean space \mathbb{R}^n
- The **inner product** or dot product of \mathbf{u} and \mathbf{v} , denoted by $\langle \mathbf{u}, \mathbf{v} \rangle$, is defined by:

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{v}^\top \mathbf{u} = \sum_{i=1}^n u_i v_i$$

- Properties:
 - ① **Positivity:** $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$ for all \mathbf{v} .
 - ② **Definiteness:** $\langle \mathbf{v}, \mathbf{v} \rangle = 0 \Rightarrow \mathbf{v} = \mathbf{0}$.
 - ③ **Additivity:** $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$ for all $\mathbf{u}, \mathbf{v}, \mathbf{w}$.
 - ④ **Homogeneity:** $\langle \lambda \mathbf{u}, \mathbf{v} \rangle = \lambda \langle \mathbf{u}, \mathbf{v} \rangle$ for all $\mathbf{u}, \mathbf{v}, \lambda$.
 - ⑤ **Symmetry:** $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$ for all \mathbf{u}, \mathbf{v} .
- **Note-1:** More generally, the complex inner product (not the focus here) is defined by $\mathbf{v}^H \mathbf{u}$, where H denotes conjugate-transpose.
- **Note-2:** Any function mapping a pair of vectors to scalars satisfying the above properties is called a real inner product.

Norm

- The Euclidean **norm**, 2-norm, ℓ_2 -norm, or simply length of a vector $\mathbf{v} \in \mathbb{R}^n$, denoted by $\|\mathbf{v}\|_2$ or simply $\|\mathbf{v}\|$, is defined by:

$$\|\mathbf{v}\| = +\sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} = +\sqrt{\sum_{i=1}^n |v_i|^2}$$

- Properties:
 - Triangle inequality or subadditivity:** $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$ for all \mathbf{u}, \mathbf{v} .
 - Absolute homogeneity:** $\|\lambda \mathbf{v}\| = |\lambda| \cdot \|\mathbf{v}\|$ for all \mathbf{v}, λ .
 - Zero vector:** $\|\mathbf{v}\| = 0 \Rightarrow \mathbf{v} = \mathbf{0}$.
- In fact, any function mapping vectors to scalars satisfying the above properties is called a norm, e.g., the ℓ_p -norm:
 $\|\mathbf{v}\|_p = \left(\sum_{i=1}^n |v_i|^p \right)^{1/p}$ for any $p \geq 1$.

Orthogonality

- Two vectors are **orthogonal** if their inner product is zero. Thus \mathbf{u}, \mathbf{v} are orthogonal if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$ and we write $\mathbf{u} \perp \mathbf{v}$.
- A set of vectors is called orthogonal if any two of them are orthogonal: $\mathbf{u}_1, \dots, \mathbf{u}_n$ is orthogonal if $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0$ for all $i \neq j$.
- A set of vectors is called **orthonormal** if it is orthogonal and all vectors have unit norm, i.e., $\|\mathbf{u}_i\| = 1$ for all i .
- $\|\mathbf{u} + \mathbf{v}\|^2 = \langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\langle \mathbf{u}, \mathbf{v} \rangle$
- **Baudhayana-Pythagoras theorem:** If $\mathbf{u} \perp \mathbf{v}$ then

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2.$$

Orthogonality

- Let $\mathbf{u}_1, \dots, \mathbf{u}_n$ be orthonormal. If $\mathbf{v} = \sum_{i=1}^n \alpha_i \mathbf{u}_i$ and $\mathbf{w} = \sum_{i=1}^n \beta_i \mathbf{u}_i$, then

$$\begin{aligned}\langle \mathbf{v}, \mathbf{w} \rangle &= \sum_i \sum_j \alpha_i \beta_j \langle \mathbf{u}_i, \mathbf{u}_j \rangle \\ &= \sum_{i=1}^n \alpha_i \beta_i \\ \|\mathbf{v}\|^2 &= \sum_{i=1}^n |\alpha_i|^2\end{aligned}$$

- Orthonormal vectors are linearly independent: Let $\mathbf{u}_1, \dots, \mathbf{u}_n$ be orthonormal. If $\mathbf{w} = \sum_{i=1}^n \alpha_i \mathbf{u}_i = \mathbf{0}$, then

$$0 = \|\mathbf{w}\|^2 = \sum_{i=1}^n |\alpha_i|^2 \Rightarrow \alpha_i = 0 \text{ for all } i.$$

Orthogonal projection onto a subspace

- A **subspace** \mathcal{S} of \mathbb{R}^n is a subset of vectors which is closed under linear combinations.
- The **span** of a set of set of vectors is the smallest subspace which contains it, or equivalently, the set of all possible linear combinations of vectors in the set.
- The **orthogonal projection** of a vector \mathbf{v} onto a subspace \mathcal{S} is the unique vector $\mathbf{w} = \text{Proj}_{\mathcal{S}}(\mathbf{v})$ in \mathcal{S} that is closest to \mathbf{v} , i.e.,
 - ① $\text{Proj}_{\mathcal{S}}(\mathbf{v}) \in \mathcal{S}$ and
 - ② $\|\mathbf{v} - \text{Proj}_{\mathcal{S}}(\mathbf{v})\| \leq \|\mathbf{v} - \mathbf{u}\|$ for all \mathbf{u} in \mathcal{S} .
- **Orthogonality principle:** A vector \mathbf{w} in subspace \mathcal{S} is the orthogonal projection of the vector \mathbf{v} onto \mathcal{S} if, and only if, the error $\mathbf{v} - \mathbf{w}$ is orthogonal to all vectors in \mathcal{S} , i.e.,

$$\mathbf{w} = \text{Proj}_{\mathcal{S}}(\mathbf{v}) \Leftrightarrow \mathbf{w} \in \mathcal{S} \text{ and } \mathbf{v} - \mathbf{w} \perp \mathcal{S},$$

i.e., $\langle \mathbf{v} - \mathbf{w}, \mathbf{u} \rangle = 0$ for all \mathbf{u} in \mathcal{S} .

- $\|\mathbf{v}\|^2 = \|\mathbf{v} - \text{Proj}_{\mathcal{S}}(\mathbf{v})\|^2 + \|\text{Proj}_{\mathcal{S}}(\mathbf{v})\|^2$.

Orthogonal projection onto a subspace

- Let $\mathcal{S} = \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_n)$ where $\mathbf{u}_1, \dots, \mathbf{u}_n$ are linearly independent.
- $\text{Proj}_{\mathcal{S}}(\mathbf{v}) \in \mathcal{S} \Rightarrow \text{Proj}_{\mathcal{S}}(\mathbf{v}) = \sum_{i=1}^n \alpha_i \mathbf{u}_i$ for some scalars $\alpha_1, \dots, \alpha_n$.
- By the orthogonality principle, $\langle \mathbf{v} - \text{Proj}_{\mathcal{S}}(\mathbf{v}), \mathbf{u}_j \rangle = 0$ for all j .
- Thus, $\langle \mathbf{v}, \mathbf{u}_j \rangle = \sum_{i=1}^n \alpha_i \langle \mathbf{u}_i, \mathbf{u}_j \rangle$, $j = 1, \dots, n$, a system of n linear equations in the n unknowns $\alpha_1, \dots, \alpha_n$ with a unique solution:

$$\begin{bmatrix} \langle \mathbf{v}, \mathbf{u}_1 \rangle \\ \vdots \\ \langle \mathbf{v}, \mathbf{u}_n \rangle \end{bmatrix} = \underbrace{\begin{bmatrix} \langle \mathbf{u}_1, \mathbf{u}_1 \rangle & \dots & \langle \mathbf{u}_n, \mathbf{u}_1 \rangle \\ \vdots & \ddots & \vdots \\ \langle \mathbf{u}_1, \mathbf{u}_n \rangle & \dots & \langle \mathbf{u}_n, \mathbf{u}_n \rangle \end{bmatrix}}_{n \times n \text{ invertible Gram matrix}} \underbrace{\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}}_{n \text{ unknowns}}$$

Orthogonal projection onto a subspace

Implications:

- If $\mathcal{S} = \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_n)$ and $\mathbf{u}_1, \dots, \mathbf{u}_n$ are orthonormal, then

$$\text{Proj}_{\mathcal{S}}(\mathbf{v}) = \sum_{i=1}^n \langle \mathbf{v}, \mathbf{u}_i \rangle \mathbf{u}_i$$

- If $\mathcal{S} = \text{Span}(\mathbf{u})$, a one dimensional subspace, then

$$\text{Proj}_{\mathcal{S}}(\mathbf{v}) = \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\|\mathbf{u}\|^2} \mathbf{u}$$

- **Cauchy-Schwartz-Bunyakovski inequality:** $|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \cdot \|\mathbf{v}\|$
with equality, if, and only if, one of \mathbf{u} , \mathbf{v} is a scalar multiple of the other.