

Orthogonality

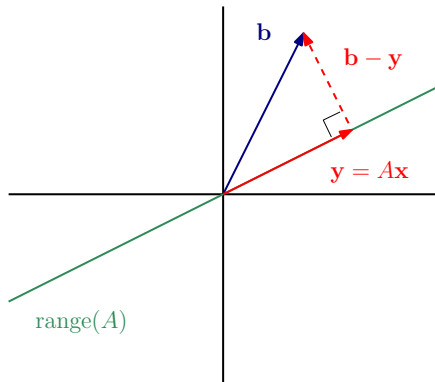
Orthogonality

Normal Equation Revisited

Alternate perspective on the “normal equation”:

$$A^T(\mathbf{b} - A\mathbf{x}) = \mathbf{0} \iff \underbrace{\mathbf{z}^T(\mathbf{b} - A\mathbf{x})}_{\text{residual} = \mathbf{r}} = 0 \quad \text{for all } \mathbf{z} \in \mathcal{R}(A),$$

i.e., \mathbf{x} solves the normal equation if and only if the residual is orthogonal to the range of A .



Orthogonal Vectors

$$\vec{u} \cdot \vec{v} = \|\vec{u}\|_2 \|\vec{v}\|_2 \cos \theta.$$

(or $\vec{u}^T \vec{v}$)

Recall that the angle θ between two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ satisfies

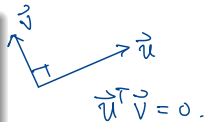
$$\cos(\theta) = \frac{\mathbf{u}^T \mathbf{v}}{\|\mathbf{u}\|_2 \|\mathbf{v}\|_2}.$$

If $\theta = \pm \pi/2$, then

$\cos(\theta) = 0$, so $\vec{u}^T \vec{v} = 0$.

Definition 1

- Two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ are **orthogonal** if $\mathbf{u}^T \mathbf{v} = 0$.
- Vectors $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_k \in \mathbb{R}^n$ are **orthogonal** if $\mathbf{q}_i^T \mathbf{q}_j = 0$ for all $i \neq j$.
- Vectors $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_k \in \mathbb{R}^n$ are **orthonormal** if $\mathbf{q}_i^T \mathbf{q}_j = \delta_{i,j}$.



Notation. (Kronecker delta function)

$$\delta_{i,j} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

\perp $\|\vec{q}_j\|_2 = 1$.

Example

$$\bullet \left\{ \underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}}_{\text{unit vector}}, \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ -5 \end{bmatrix} \right\} : \text{orthogonal but not orthonormal}$$

$$\bullet \left\{ \begin{matrix} \vec{u}_1 & \vec{v}_1 \\ \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} & \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} \end{matrix} \right\}$$

$$\begin{aligned} \bullet \vec{u}^T \vec{v} &= \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} \cdot \left(-\frac{1}{\sqrt{2}}\right) \\ &= \frac{1}{2} - \frac{1}{2} = 0 \quad \text{orthogonal} \end{aligned}$$

$$\bullet \|\vec{u}\|_2^2 = \left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{\sqrt{2}}\right)^2 = \frac{1}{2} + \frac{1}{2} = 1$$

$$\bullet \|\vec{v}\|_2^2 = \left(\frac{1}{\sqrt{2}}\right)^2 + \left(-\frac{1}{\sqrt{2}}\right)^2 = \frac{1}{2} + \frac{1}{2} = 1$$

\therefore Orthonormal!

Matrices with Orthogonal Columns

$$(i,j)\text{-entry of } Q^T Q = \vec{q}_i^T \vec{q}_j$$

Let $Q = [\mathbf{q}_1 \mid \mathbf{q}_2 \mid \cdots \mid \mathbf{q}_k] \in \mathbb{R}^{n \times k}$. Note that

$$Q^T Q = \begin{bmatrix} \mathbf{q}_1^T \\ \mathbf{q}_2^T \\ \vdots \\ \mathbf{q}_k^T \end{bmatrix} \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_k \end{bmatrix} = \begin{bmatrix} \mathbf{q}_1^T \mathbf{q}_1 & \mathbf{q}_1^T \mathbf{q}_2 & \cdots & \mathbf{q}_1^T \mathbf{q}_k \\ \mathbf{q}_2^T \mathbf{q}_1 & \mathbf{q}_2^T \mathbf{q}_2 & \cdots & \mathbf{q}_2^T \mathbf{q}_k \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{q}_k^T \mathbf{q}_1 & \mathbf{q}_k^T \mathbf{q}_2 & \cdots & \mathbf{q}_k^T \mathbf{q}_k \end{bmatrix}.$$

Therefore,

- $\mathbf{q}_1, \dots, \mathbf{q}_k$ are orthogonal. $\iff Q^T Q$ is a $k \times k$ diagonal matrix.
- $\mathbf{q}_1, \dots, \mathbf{q}_k$ are orthonormal. $\iff Q^T Q$ is the $k \times k$ identity matrix.

because all off-diagonals
are of the form
 $\vec{q}_i^T \vec{q}_j, i \neq j.$
0
by orthogonality

Example

$$\cdot \quad \vec{q}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{q}_2 = \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix}, \quad \vec{q}_3 = \begin{bmatrix} 0 \\ 0 \\ -5 \end{bmatrix}.$$

$$Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -5 \end{bmatrix}, \quad Q^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -5 \end{bmatrix} = Q$$

$$Q^T Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 25 \end{bmatrix} \text{ diagonal.}$$

$$\cdot \quad \vec{q}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \quad \vec{q}_2 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}.$$

$$Q = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}, \quad Q^T = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}, \quad Q^T Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \vec{q}_1^T \vec{q}_1 & \vec{q}_1^T \vec{q}_2 \\ \vec{q}_2^T \vec{q}_1 & \vec{q}_2^T \vec{q}_2 \end{bmatrix}$$

Matrices with Orthonormal Columns

Theorem 2

Let $Q = [\mathbf{q}_1 \mid \mathbf{q}_2 \mid \cdots \mid \mathbf{q}_k] \in \mathbb{R}^{n \times k}$ and suppose that $\mathbf{q}_1, \dots, \mathbf{q}_k$ are **orthonormal**. Then

- 1 $Q^T Q = I \in \mathbb{R}^{k \times k}$; \rightarrow explained above
- 2 $\|Q\mathbf{x}\|_2 = \|\mathbf{x}\|_2$ for all $\mathbf{x} \in \mathbb{R}^k$; (norm is preserved)
- 3 $\|Q\|_2 = 1$.

Proof for ②

$$\|Q\vec{x}\|_2^2 = (Q\vec{x})^T (Q\vec{x}) = \vec{x}^T \underbrace{Q^T Q}_{\substack{= \\ I \text{ by ①}}} \vec{x} = \vec{x}^T \vec{x} = \|\vec{x}\|_2^2$$

Orthogonal Matrices

Definition 3

We say that $Q \in \mathbb{R}^{n \times n}$ is an **orthogonal matrix** if $Q^T Q = I \in \mathbb{R}^{n \times n}$.

↪ sq. matrix w/ orthonormal columns.

- A square matrix with orthogonal columns is not, in general, an orthogonal matrix!

Properties of Orthogonal Matrices

$$Q^T Q = I$$

Theorem 4

Let $Q \in \mathbb{R}^{n \times n}$ be orthogonal. Then

- 1 $Q^{-1} = Q^T$;
- 2 Q^T is also an orthogonal matrix;
- 3 $\kappa_2(Q) = 1$;
- 4 For any $A \in \mathbb{R}^{n \times n}$, $\|AQ\|_2 = \|A\|_2$;
- 5 if $P \in \mathbb{R}^{n \times n}$ is another orthogonal matrix, then PQ is also orthogonal.

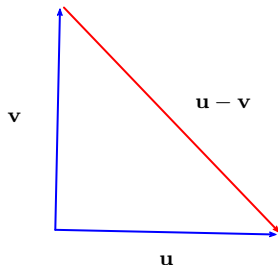
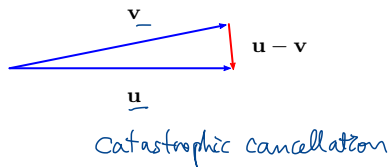
cf) $\|AB\|_2 \leq \|A\|_2 \|B\|_2$ in general.

Why Do We Like Orthogonal Vectors?

- If \mathbf{u} and \mathbf{v} are orthogonal, then

$$\|\mathbf{u} \pm \mathbf{v}\|_2^2 = \|\vec{u}\|_2^2 + \|\vec{v}\|_2^2 \quad (\text{Pythagorean})$$

- Without orthogonality, it is possible that $\|\mathbf{u} - \mathbf{v}\|_2$ is much smaller than $\|\mathbf{u}\|_2$ and $\|\mathbf{v}\|_2$.
- The addition and subtraction of orthogonal vectors are guaranteed to be well-conditioned.



Appendix: Projection and Reflection

Projection and Reflection Operators

Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^m$ be nonzero vectors.

- Projection of \mathbf{u} onto $\langle \mathbf{v} \rangle = \text{span}(\mathbf{v})$:

$$\frac{\mathbf{v}^T \mathbf{u}}{\mathbf{v}^T \mathbf{v}} \mathbf{v} = \underbrace{\left(\frac{\mathbf{v} \mathbf{v}^T}{\mathbf{v}^T \mathbf{v}} \right)}_{=: P} \mathbf{u} =: P \mathbf{u}.$$

- Projection of \mathbf{u} onto $\langle \mathbf{v} \rangle^\perp$, the orthogonal complement of $\langle \mathbf{v} \rangle$:

$$\mathbf{u} - \frac{\mathbf{v}^T \mathbf{u}}{\mathbf{v}^T \mathbf{v}} \mathbf{v} = \left(I - \frac{\mathbf{v} \mathbf{v}^T}{\mathbf{v}^T \mathbf{v}} \right) \mathbf{u} =: (I - P) \mathbf{u}.$$

- Reflection of \mathbf{u} across $\langle \mathbf{v} \rangle^\perp$:

$$\mathbf{u} - 2 \frac{\mathbf{v}^T \mathbf{u}}{\mathbf{v}^T \mathbf{v}} \mathbf{v} = \left(I - 2 \frac{\mathbf{v} \mathbf{v}^T}{\mathbf{v}^T \mathbf{v}} \right) \mathbf{u} =: (I - 2P) \mathbf{u}.$$

Projection and Reflection Operators (cont')

Summary: for given $\mathbf{v} \in \mathbb{R}^m$, a nonzero vector, let

$$P = \frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}} \in \mathbb{R}^{m \times m}.$$

Then the following matrices carry out geometric transformations

- Projection onto $\langle \mathbf{v} \rangle$: P
- Projection onto $\langle \mathbf{v} \rangle^\perp$: $I - P$
- Reflection across $\langle \mathbf{v} \rangle^\perp$: $I - 2P$

Note. If \mathbf{v} were a unit vector, the definition of P simplifies to $P = \mathbf{v}\mathbf{v}^T$.