Orthogonality

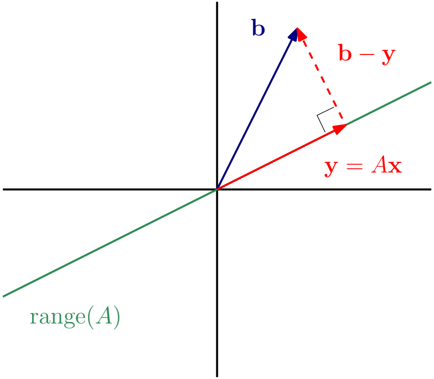
Orthogonality

Normal Equation Revisited

Alternate perspective on the "normal equation":

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

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

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

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

$\vec{u}^T\vec{v} = 0.$

$\perp$

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

Notation. (Kronecker delta function)

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

Example

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

- $\left\{ \overset{\vec{u}}{\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}}, \overset{\vec{v}}{\begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}} \right\}$

- $\vec{u}^T \vec{v} = 1/\sqrt{2} \cdot 1/\sqrt{2} + 1/\sqrt{2} \cdot (-1/\sqrt{2})$

  $= 1/2 - 1/2 = 0$ orthogonal

- $\|\vec{u}\|_2^2 = (1/\sqrt{2})^2 + (1/\sqrt{2})^2 = 1/2 + 1/2 = 1$

- $\|\vec{v}\|_2^2 = (1/\sqrt{2})^2 + (-1/\sqrt{2})^2 = 1/2 + 1/2 = 1$

$\therefore$ Orthonormal!

Matrices with Orthogonal Columns

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

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

$$Q^{\mathrm{T}}Q = \begin{bmatrix} \mathbf{q}_1^{\mathrm{T}} \\ \hline \mathbf{q}_2^{\mathrm{T}} \\ \hline \vdots \\ \hline \mathbf{q}_k^{\mathrm{T}} \end{bmatrix} \left[\begin{array}{c|c|c|c} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_k \end{array}\right] = \begin{bmatrix} \mathbf{q}_1^{\mathrm{T}}\mathbf{q}_1 & \mathbf{q}_1^{\mathrm{T}}\mathbf{q}_2 & \cdots & \mathbf{q}_1^{\mathrm{T}}\mathbf{q}_k \\ \mathbf{q}_2^{\mathrm{T}}\mathbf{q}_1 & \mathbf{q}_2^{\mathrm{T}}\mathbf{q}_2 & \cdots & \mathbf{q}_2^{\mathrm{T}}\mathbf{q}_k \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{q}_k^{\mathrm{T}}\mathbf{q}_1 & \mathbf{q}_k^{\mathrm{T}}\mathbf{q}_2 & \cdots & \mathbf{q}_k^{\mathrm{T}}\mathbf{q}_k \end{bmatrix}.$$

Therefore,

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

because all off-diagonals are of the form

$$\left(\begin{array}{c} \vec{q}_i^{\,T}\vec{q}_j, \quad i \neq j. \\ \parallel \\ 0 \\ \text{by orthogonality} \end{array} \right)$$

Example

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

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

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

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

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

Matrices with Orthonormal Columns

Theorem 2

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

1. $Q^{\mathrm{T}}Q = I \in \mathbb{R}^{k\times k}$; → 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)

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

Orthogonal Matrices

Definition 3

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

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. $Q^{-1} = Q^{\mathrm{T}}$;
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.*

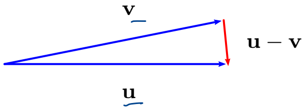
cf) $\|AB\|_2 \le \|A\|_2 \|B\|_2$ in general.

Why Do We Like Orthogonal Vectors?

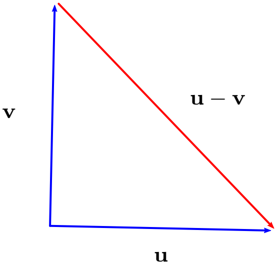
- If $\mathbf{u}$ and $\mathbf{v}$ are orthogonal, then

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

- 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.

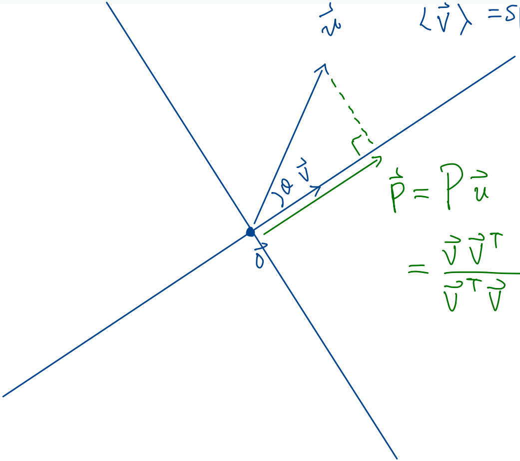


Catastrophic cancellation



Appendix: Projection and Reflection

Given $\vec{u}, \vec{v} \in \mathbb{R}^n$ $\qquad\qquad$ $\|\cdot\|$ for $\|\cdot\|_2$



$\vec{p} = P\vec{u}$

$= \dfrac{\vec{v}\vec{v}^T}{\vec{v}^T\vec{v}} \vec{u}$

$\langle \vec{v} \rangle^{\perp}$ = orthogonal complement of $\langle \vec{v} \rangle$.

- $\|\vec{p}\| = \|\vec{u}\| \cos\theta$

- $\vec{p} = \|\vec{p}\|$ (unit vector in the dir. of $\vec{v}$)

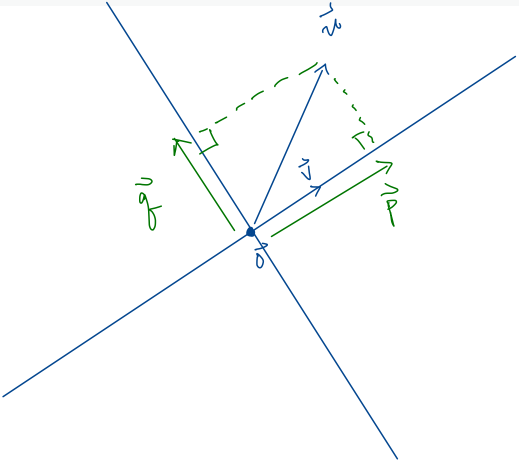
$= \|\vec{p}\| \dfrac{\vec{v}}{\|\vec{v}\|}$

$= \dfrac{\|\vec{u}\| \cos\theta \, \|\vec{v}\|}{\|\vec{v}\| \, \|\vec{v}\|} \vec{v}$

$= \dfrac{\|\vec{u}\| \|\vec{v}\| \cos\theta}{\|\vec{v}\|^2} \vec{v}$

$= \dfrac{\vec{u}^T\vec{v}}{\vec{v}^T\vec{v}} \vec{v} = \vec{v} \dfrac{\vec{v}^T\vec{u}}{\vec{v}^T\vec{v}} = P\vec{u}$

$P = \dfrac{\vec{v}\vec{v}^T}{\vec{v}^T\vec{v}}$



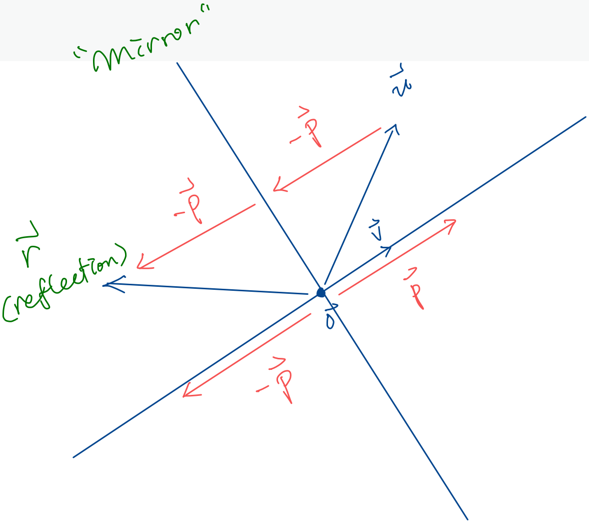
$\langle \vec{v} \rangle = \text{span}\{\vec{v}\}$

$\vec{q} = ?$

$$\vec{p} + \vec{q} = \vec{u}$$

$$\Rightarrow \vec{q} = \vec{u} - \vec{p} = \vec{u} - P\vec{u} = \underbrace{(I - P)}_{= Q}\vec{u}$$

$\langle \vec{v} \rangle^{\perp}$ = orthogonal complement of $\langle \vec{v} \rangle$.



$\langle \vec{v} \rangle = \text{span}\{\vec{v}\}$

$$\vec{r} = \vec{u} - 2\vec{p}$$

$$= \vec{u} - 2P\vec{u}$$

$$= \underbrace{(I - 2P)}_{= R}\,\vec{u}$$

$\langle \vec{v} \rangle^{\perp}$ = orthogonal complement of $\langle \vec{v} \rangle$.

Projection and Reflection Operators

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

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

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

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

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

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

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

Projection and Reflection Operators (cont')

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

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

Then the following matrices carry out geometric transformations

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

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