Least Squares, Full QR, Orthogonal Matrices

Linear Algebra

Josh Engwer

TTU

02 Nov 2015 (revised 14 Nov 2021)

PART I:

(Orthogonal) Projections on a Subspace

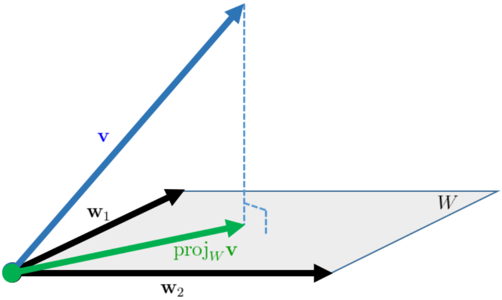
Orthogonal Subspaces

Orthogonal Complements

Direct Sums of Subspaces of $\mathbb{R}^n$

Full QR Factorization via CGS-EN

(Orthogonal) Projection onto a Subspace (Definition)



(Above): In $\mathbb{R}^3$, plane W is subspace spanned by orthogonal vectors $\mathbf{w}_1$ & $\mathbf{w}_2$.

Theorem

(Projection onto a Subspace)

Let $\mathcal{Q} = \{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$ be an orthogonal basis for subspace W of $\mathbb{R}^n$.

Then the ***(orthogonal) projection of vector*** $\mathbf{v} \in \mathbb{R}^n$ ***onto subspace*** W *is:*

$$\textit{proj}_W\mathbf{v} := \textit{proj}_{span(\mathcal{Q})}\mathbf{v} = \textit{proj}_{\mathbf{q}_1}\mathbf{v} + \textit{proj}_{\mathbf{q}_2}\mathbf{v} + \cdots + \textit{proj}_{\mathbf{q}_n}\mathbf{v}$$

Orthogonal Sets & Orthogonal Subspaces (Definition)

Orthogonality generalizes to subsets & subspaces of inner product space $\mathbb{R}^n$:

Definition

(Orthogonal Sets)

Sets $E_1, E_2 \subset \mathbb{R}^n$ are **orthogonal**, denoted $E_1 \perp E_2$, if

$$\mathbf{v}_1^T \mathbf{v}_2 = 0 \qquad \forall \mathbf{v}_1 \in E_1, \forall \mathbf{v}_2 \in E_2$$

i.e. vectors in one set are orthogonal to vectors in the other set.

Definition

(Orthogonal Subspaces)

Subspaces S_1, S_2 of $\mathbb{R}^n$ are **orthogonal**, denoted $S_1 \perp S_2$, if

$$\mathbf{v}_1^T \mathbf{v}_2 = 0 \qquad \forall \mathbf{v}_1 \in S_1, \forall \mathbf{v}_2 \in S_2$$

i.e. vectors in one subspace are orthogonal to vectors in the other subspace.

Orthogonal Subspaces

Unless one of the subspaces contains only the zero vector, there are infinitely many vectors in each subspace to test for orthogonality!!

Fortunately, since inner product space $\mathbb{R}^n$ is finite-dimensional, it suffices to test the basis vectors for the subspaces:

Theorem

(Bases of Orthogonal Subspaces Theorem)

Let subspace $S_1 \subseteq \mathbb{R}^n$ *with basis* $\mathcal{B}_1 = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$.
Let subspace $S_2 \subseteq \mathbb{R}^n$ *with basis* $\mathcal{B}_2 = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$.

Then $S_1 \perp S_2 \iff \mathcal{B}_1 \perp \mathcal{B}_2$

PROOF:

$$S_1 = \text{span}(\mathcal{B}_1) = \{\text{linear combinations } c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_m\mathbf{u}_m = \textstyle\sum_{i=1}^m c_i\mathbf{u}_i\}$$
$$S_2 = \text{span}(\mathcal{B}_2) = \{\text{linear combinations } k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_p\mathbf{v}_p = \textstyle\sum_{j=1}^p k_j\mathbf{v}_j\}$$

$$S_1 \perp S_2 \iff \left(\textstyle\sum_{i=1}^m c_i\mathbf{u}_i\right)^T \left(\textstyle\sum_{j=1}^p k_j\mathbf{v}_j\right) = 0 \overset{T2}{\iff} \left(\textstyle\sum_{i=1}^m c_i\mathbf{u}_i^T\right) \left(\textstyle\sum_{j=1}^p k_j\mathbf{v}_j\right) = 0$$

$$\overset{\text{"FOIL"}}{\iff} \textstyle\sum_{i=1}^m \sum_{j=1}^p c_i k_j \mathbf{u}_i^T\mathbf{v}_j = 0 \iff \mathbf{u}_i^T\mathbf{v}_j = 0 \;\forall i,j \iff \mathcal{B}_1 \perp \mathcal{B}_2$$ QED

Orthogonal Complements (Definition)

Definition

(Orthogonal Complements)

Let W be a subspace of Euclidean inner product space $(\mathbb{R}^n, \langle \cdot, \cdot \rangle_2)$.
Then the **orthogonal complement of** W, denoted $W^\perp$, is

$$W^\perp := \{\mathbf{w}^\perp \in \mathbb{R}^n : \langle \mathbf{w}, \mathbf{w}^\perp \rangle_2 = 0 \ \ \forall \mathbf{w} \in W\}$$

i.e. $W^\perp$ is the set of all vectors in $\mathbb{R}^n$ that are orthogonal to all vectors in W.

Clearly, by definition, $W \perp W^\perp$.

<u>SPECIAL CASES:</u> $(\mathbb{R}^n)^\perp = \{\vec{\mathbf{0}}\}$ AND $\{\vec{\mathbf{0}}\}^\perp = \mathbb{R}^n$

For instance, if $W = \text{span}\left\{\begin{bmatrix} 4 \\ -8 \\ 3 \end{bmatrix}\right\}$, then $W^\perp = \text{span}\left\{\begin{bmatrix} -3 \\ 0 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}\right\}$.

since $\begin{bmatrix} 4 \\ -8 \\ 3 \end{bmatrix}^T \begin{bmatrix} -3 \\ 0 \\ 4 \end{bmatrix} = 0$ AND $\begin{bmatrix} 4 \\ -8 \\ 3 \end{bmatrix}^T \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = 0$

Finding the Orthogonal Complement (Motivation)

Supppose subspace $W = \text{span}(\mathcal{B})$ where basis $\mathcal{B} = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_p\}$.

How to systematically find the orthogonal complement $W^\perp$?

Suppose vector $\mathbf{u} \in W^\perp$. Then:

$$\begin{aligned}
\mathbf{w}^T\mathbf{u} = 0 \quad \forall \mathbf{w} \in W \quad & \text{(Def'n of orth. complement)} \\
\iff \mathbf{w}^T\mathbf{u} = 0 \quad \forall \mathbf{w} \in \mathcal{B} \quad & \text{(Previous Theorem)} \\
\iff \mathbf{w}_1^T\mathbf{u} = 0, \; \cdots, \; \mathbf{w}_p^T\mathbf{u} = 0 \quad &
\end{aligned}$$

$$\iff \underbrace{\begin{bmatrix} \text{——} & \mathbf{w}_1^T & \text{——} \\ & \vdots & \\ \text{——} & \mathbf{w}_p^T & \text{——} \end{bmatrix}}_{A^T} \begin{bmatrix} | \\ \mathbf{u} \\ | \end{bmatrix} = \begin{bmatrix} | \\ \vec{\mathbf{0}} \\ | \end{bmatrix}$$

$$\begin{aligned}
\iff A^T\mathbf{u} = \vec{\mathbf{0}} \quad & \\
\iff \mathbf{u} \in \text{NulSp}(A^T) \quad & \text{(Def'n of Null Space of } A^T\text{)}
\end{aligned}$$

To find $W^\perp$, form columns of A with basis vectors of W, then find $\text{NulSp}(A^T)$.

Finding the Orthogonal Complement (Procedure)

How to systematically find the orthogonal complement of a subspace of $\mathbb{R}^n$?

Proposition

(Finding Orthogonal Complement of a Subspace of $\mathbb{R}^n$)

GIVEN: *Subspace W of $\mathbb{R}^n$ such that $W = \text{span}\{\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_p\}$.*
TASK: *Find orthogonal complement $W^\perp$.*

(1) *Form matrix* $A = \begin{bmatrix} | & | & & | \\ \mathbf{w}_1 & \mathbf{w}_2 & \cdots & \mathbf{w}_p \\ | & | & & | \end{bmatrix}$

(2) $W^\perp = \text{NulSp}(A^T) \implies \left[\begin{array}{c|c} A^T & \vec{\mathbf{0}} \end{array}\right] \xrightarrow{Gauss-Jordan} \left[\begin{array}{c|c} \text{RREF}(A^T) & \vec{\mathbf{0}} \end{array}\right]$

Direct Sums of Subspaces of $\mathbb{R}^n$ (Definition)

Definition

(Direct Sum)

Let W_1, W_2 be two subspaces of $\mathbb{R}^n$.

Then $\mathbb{R}^n$ is the **direct sum** of W_1 & W_2, written $\mathbb{R}^n = W_1 \oplus W_2$, if

$\mathbf{v} \in \mathbb{R}^n$ can be uniquely written as $\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_2$, where $\mathbf{w}_1 \in W_1$ & $\mathbf{w}_2 \in W_2$.

i.e. each vector in $\mathbb{R}^n$ can be uniquely written as a sum of a vector from W_1 and a vector from W_2.

e.g. Let $W_1 = \text{span}\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$ and $W_2 = \text{span}\left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \right\}$ and $\mathbf{v} \in \mathbb{R}^3$. Then:

$$\underbrace{\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}}_{\mathbf{v}} = \underbrace{\begin{bmatrix} 0 \\ -\frac{1}{2}v_1 + v_2 - \frac{1}{2}v_3 \\ 0 \end{bmatrix} + \begin{bmatrix} \frac{1}{2}v_1 + \frac{1}{2}v_3 \\ \frac{1}{2}v_1 + \frac{1}{2}v_3 \\ \frac{1}{2}v_1 + \frac{1}{2}v_3 \end{bmatrix}}_{\mathbf{w}_1} + \underbrace{\begin{bmatrix} \frac{1}{2}v_1 - \frac{1}{2}v_3 \\ 0 \\ -\frac{1}{2}v_1 + \frac{1}{2}v_3 \end{bmatrix}}_{\mathbf{w}_2}$$

$$= \underbrace{\begin{bmatrix} \frac{1}{2}v_1 + \frac{1}{2}v_3 \\ v_2 \\ \frac{1}{2}v_1 + \frac{1}{2}v_3 \end{bmatrix}}_{\mathbf{w}_1} + \underbrace{\begin{bmatrix} \frac{1}{2}v_1 - \frac{1}{2}v_3 \\ 0 \\ -\frac{1}{2}v_1 + \frac{1}{2}v_3 \end{bmatrix}}_{\mathbf{w}_2} \implies \mathbb{R}^3 = W_1 \oplus W_2$$

Properties of Orthogonal Complements

Theorem

(Properties of Orthogonal Complements)

Let W be a subspace of Euclidean induced-norm inner product space $(\mathbb{R}^n, \langle \cdot, \cdot \rangle_2, || \cdot ||_2)$. Then:

(i) $W^\perp$ *is also a subspace of* $\mathbb{R}^n$

(ii) $W \cap W^\perp = \{\vec{\mathbf{0}}\}$

(iii) $\mathbb{R}^n = W \oplus W^\perp$

(iv) $\mathit{dim}(\mathbb{R}^n) = \mathit{dim}(W) + \mathit{dim}(W^\perp)$

(v) $(W^\perp)^\perp = W$

Properties of Orthogonal Complements

Theorem

(Properties of Orthogonal Complements)

Let W be a subspace of Euclidean induced-norm inner product space $(\mathbb{R}^n, \langle\cdot,\cdot\rangle_2, ||\cdot||_2)$. Then:

(i) $W^\perp$ *is also a subspace of* $\mathbb{R}^n$

(ii) $W \cap W^\perp = \{\vec{\mathbf{0}}\}$

(iii) $\mathbb{R}^n = W \oplus W^\perp$

(iv) $dim(\mathbb{R}^n) = dim(W) + dim(W^\perp)$

(v) $(W^\perp)^\perp = W$

PROOF:

(i) Let vectors $\mathbf{w}^\perp, \mathbf{w}_1^\perp, \mathbf{w}_2^\perp \in W^\perp$ and scalar $\alpha \in \mathbb{R}$.
Then, since $W^\perp$ is the orthogonal complement (OC) of W:

$$\begin{array}{ccccc} \langle \mathbf{w}, \mathbf{w}_1^\perp + \mathbf{w}_2^\perp\rangle_2 & \overset{IPS5}{=} & \langle \mathbf{w}, \mathbf{w}_1^\perp\rangle_2 + \langle \mathbf{w}, \mathbf{w}_2^\perp\rangle_2 \overset{OC}{=} 0 + 0 = 0 \ \ \forall \mathbf{w} \in W & \Longrightarrow & \mathbf{w}_1^\perp + \mathbf{w}_2^\perp \in W^\perp \\ \langle \alpha\mathbf{w}^\perp, \mathbf{w}\rangle_2 & \overset{IPS4}{=} & \alpha\langle \mathbf{w}^\perp, \mathbf{w}\rangle_2 \overset{OC}{=} \alpha \cdot 0 = 0 \ \ \forall \mathbf{w} \in W & \Longrightarrow & \alpha\mathbf{w}^\perp \in W^\perp \end{array}$$

Hence:
$$\begin{array}{rcr} \mathbf{w}_1^\perp, \mathbf{w}_2^\perp \in W^\perp & \Longrightarrow & \mathbf{w}_1^\perp + \mathbf{w}_2^\perp \in W^\perp \\ \mathbf{w}^\perp \in W^\perp & \Longrightarrow & \alpha\mathbf{w}^\perp \in W^\perp \end{array}$$

$\therefore$ $W^\perp$ is also a subspace of $\mathbb{R}^n$. $\square$

Properties of Orthogonal Complements

Theorem

(Properties of Orthogonal Complements)

Let W be a subspace of Euclidean induced-norm inner product space $(\mathbb{R}^n, \langle\cdot,\cdot\rangle_2, ||\cdot||_2)$. Then:

(i)	$W^\perp$ *is also a subspace of* $\mathbb{R}^n$
(ii)	$W \cap W^\perp = \{\vec{\mathbf{0}}\}$
(iii)	$\mathbb{R}^n = W \oplus W^\perp$
(iv)	$\mathit{dim}(\mathbb{R}^n) = \mathit{dim}(W) + \mathit{dim}(W^\perp)$
(v)	$(W^\perp)^\perp = W$

PROOF:

(ii) Let $\mathbf{w} \in W \cap W^\perp$. Then, since $W \perp W^\perp$:

$$\langle \mathbf{w}, \mathbf{w}\rangle_2 = 0 \overset{DP5}{\Longrightarrow} ||\mathbf{w}||_2^2 = 0 \implies ||\mathbf{w}||_2 = 0 \implies \mathbf{w} = \vec{\mathbf{0}} \quad \square$$

Properties of Orthogonal Complements

Theorem

(Properties of Orthogonal Complements)

Let W be a subspace of Euclidean induced-norm inner product space $(\mathbb{R}^n, \langle\cdot,\cdot\rangle_2, ||\cdot||_2)$. Then:

(i) $W^\perp$ *is also a subspace of* $\mathbb{R}^n$

(ii) $W \cap W^\perp = \{\vec{\mathbf{0}}\}$

(iii) $\mathbb{R}^n = W \oplus W^\perp$

(iv) $dim(\mathbb{R}^n) = dim(W) + dim(W^\perp)$

(v) $(W^\perp)^\perp = W$

PROOF:

(iii)-(v): Too long & tedious. See textbook if interested.

PART II:

Fundamental Theorem of Linear Algebra
Pythagorean Theorem for Orthogonal Vectors
Best Approximation Theorem

Full-Rank Least Squares via Normal Equations
Full-Rank Least Squares via Reduced QR
Full-Rank Least Squares via Full QR

Fundamental Subspaces of a Matrix

Theorem

(Fundamental Theorem of Linear Algebra – FTLA)

Let matrix $A \in \mathbb{R}^{m\times n}$ *s.t. rank*$(A) = r$. *Then the fundamental subspaces of* A *are related as so:*

$$\begin{array}{lrcl} (i) & \textit{RowSp}(A) & = & \textit{ColSp}(A^T) \\ (ii) & \textit{ColSp}(A)^\perp & = & \textit{NulSp}(A^T) \\ (iii) & \textit{ColSp}(A^T)^\perp & = & \textit{NulSp}(A) \end{array} \qquad \begin{array}{ll} (iv) & \textit{dim ColSp}(A) = \textit{dim ColSp}(A^T) = r \\ (v) & \mathbb{R}^m = \textit{ColSp}(A) \oplus \textit{NulSp}(A^T) \\ (vi) & \mathbb{R}^n = \textit{ColSp}(A^T) \oplus \textit{NulSp}(A) \end{array}$$

Fundamental Subspaces of a Matrix

Theorem

(Fundamental Theorem of Linear Algebra – FTLA)

Let matrix $A \in \mathbb{R}^{m \times n}$ s.t. rank$(A) = r$. Then the fundamental subspaces of A are related as so:

$$
\begin{array}{llcl}
(i) & \textit{RowSp}(A) & = & \textit{ColSp}(A^T) \\
(ii) & \textit{ColSp}(A)^{\perp} & = & \textit{NulSp}(A^T) \\
(iii) & \textit{ColSp}(A^T)^{\perp} & = & \textit{NulSp}(A)
\end{array}
\qquad
\begin{array}{ll}
(iv) & \textit{dim ColSp}(A) = \textit{dim ColSp}(A^T) = r \\
(v) & \mathbb{R}^m = \textit{ColSp}(A) \oplus \textit{NulSp}(A^T) \\
(vi) & \mathbb{R}^n = \textit{ColSp}(A^T) \oplus \textit{NulSp}(A)
\end{array}
$$

PROOF:

$$
\begin{array}{llcl}
(i)\ \text{RowSp}(A) & := & \text{span}\{\text{Rows of } A\} & \text{(Definition of row space)} \\
 & = & \text{span}\{\text{Columns of } A^T\} & \text{(Definition of a matrix transpose)} \\
 & := & \text{ColSp}(A^T) & \text{(Definition of column space)}
\end{array}
$$

Fundamental Subspaces of a Matrix

Theorem

(Fundamental Theorem of Linear Algebra – FTLA)

Let matrix $A \in \mathbb{R}^{m \times n}$ s.t. rank$(A) = r$. Then the fundamental subspaces of A are related as so:

(i)	*RowSp*(A)	$=$	*ColSp*(A^T)	(iv)	*dim ColSp*$(A) =$ *dim ColSp*$(A^T) = r$
(ii)	*ColSp*$(A)^\perp$	$=$	*NulSp*(A^T)	(v)	$\mathbb{R}^m =$ *ColSp*$(A) \oplus$ *NulSp*(A^T)
(iii)	*ColSp*$(A^T)^\perp$	$=$	*NulSp*(A)	(vi)	$\mathbb{R}^n =$ *ColSp*$(A^T) \oplus$ *NulSp*(A)

PROOF: Below, $\mathbf{a}_1, \cdots, \mathbf{a}_n$ denote the columns of A.

(ii) ColSp$(A)^\perp$ $:= \{\mathbf{v} \in \mathbb{R}^m : \mathbf{a}^T\mathbf{v} = 0 \ \ \forall \mathbf{a} \in \text{ColSp}(A)\}$ (Defn of Orthogonal Complement)

$$= \left\{ \mathbf{v} \in \mathbb{R}^m : \begin{array}{c} \mathbf{a}_1^T\mathbf{v} = 0 \\ \vdots \\ \mathbf{a}_n^T\mathbf{v} = 0 \end{array} \right\}$$ (The columns of A span ColSp(A))

$$= \{\mathbf{v} \in \mathbb{R}^m : A^T\mathbf{v} = \vec{\mathbf{0}}\}$$ (Row-vector view of A^T)

$$:= \text{NulSp}(A^T)$$ (Definition of null space)

Fundamental Subspaces of a Matrix

Theorem

(Fundamental Theorem of Linear Algebra – FTLA)

Let matrix $A \in \mathbb{R}^{m\times n}$ s.t. rank$(A) = r$. Then the fundamental subspaces of A are related as so:

$$\begin{array}{lrcl}
(i) & \text{RowSp}(A) & = & \text{ColSp}(A^T) \\
(ii) & \text{ColSp}(A)^{\perp} & = & \text{NulSp}(A^T) \\
(iii) & \text{ColSp}(A^T)^{\perp} & = & \text{NulSp}(A)
\end{array}
\qquad
\begin{array}{ll}
(iv) & \dim \text{ColSp}(A) = \dim \text{ColSp}(A^T) = r \\
(v) & \mathbb{R}^m = \text{ColSp}(A) \oplus \text{NulSp}(A^T) \\
(vi) & \mathbb{R}^n = \text{ColSp}(A^T) \oplus \text{NulSp}(A)
\end{array}$$

PROOF: Below, $\bar{\mathbf{a}}_1, \cdots, \bar{\mathbf{a}}_n$ denote the rows of A.

$$\begin{array}{rclr}
(iii)\ \text{ColSp}(A^T)^{\perp} & = & \text{RowSp}(A)^{\perp} & \text{(Part } (i)) \\
& := & \{\mathbf{v} \in \mathbb{R}^m : (\bar{\mathbf{a}}^T)^T\mathbf{v} = 0 \ \ \forall \bar{\mathbf{a}} \in \text{RowSp}(A)\} & \text{(Defn of Ortho. Complement)} \\
& = & \left\{ \mathbf{v} \in \mathbb{R}^m : \begin{array}{c} (\bar{\mathbf{a}}_1^T)^T\mathbf{v} = 0 \\ \vdots \\ (\bar{\mathbf{a}}_n^T)^T\mathbf{v} = 0 \end{array} \right\} & \text{(Rows of } A \text{ span RowSp}(A)) \\
& = & \{\mathbf{v} \in \mathbb{R}^m : (A^T)^T\mathbf{v} = \vec{\mathbf{0}}\} & \text{(Row-vector view of } A^T) \\
& = & \{\mathbf{v} \in \mathbb{R}^m : A\mathbf{v} = \vec{\mathbf{0}}\} & \text{(Property of Transpose)} \\
& := & \text{NulSp}(A) & \text{(Definition of null space)}
\end{array}$$

Fundamental Subspaces of a Matrix

Theorem

(Fundamental Theorem of Linear Algebra – FTLA)

Let matrix $A \in \mathbb{R}^{m \times n}$ *s.t. rank*$(A) = r$. *Then the fundamental subspaces of* A *are related as so:*

$$
\begin{array}{rrcl}
(i) & \textit{RowSp}(A) & = & \textit{ColSp}(A^T) \\
(ii) & \textit{ColSp}(A)^\perp & = & \textit{NulSp}(A^T) \\
(iii) & \textit{ColSp}(A^T)^\perp & = & \textit{NulSp}(A)
\end{array}
\qquad
\begin{array}{rl}
(iv) & \textit{dim ColSp}(A) = \textit{dim ColSp}(A^T) = r \\
(v) & \mathbb{R}^m = \textit{ColSp}(A) \oplus \textit{NulSp}(A^T) \\
(vi) & \mathbb{R}^n = \textit{ColSp}(A^T) \oplus \textit{NulSp}(A)
\end{array}
$$

PROOF:

(iv) Since rank$(A) = r$, via Gauss-Jordan elimination,

RREF(A) has r pivot columns and r pivot rows.

$\therefore$ dim ColSp$(A) =$ (# pivot columns of RREF(A)) $= r$

$\therefore$ dim ColSp$(A^T) \overset{(i)}{=}$ dim RowSp$(A) =$ (# pivot rows of RREF(A)) $= r$

Fundamental Subspaces of a Matrix

Theorem

(Fundamental Theorem of Linear Algebra – FTLA)

Let matrix $A \in \mathbb{R}^{m \times n}$ s.t. rank$(A) = r$. Then the fundamental subspaces of A are related as so:

$$
\begin{array}{clcl}
(i) & \textit{RowSp}(A) & = & \textit{ColSp}(A^T) \\
(ii) & \textit{ColSp}(A)^\perp & = & \textit{NulSp}(A^T) \\
(iii) & \textit{ColSp}(A^T)^\perp & = & \textit{NulSp}(A)
\end{array}
\qquad
\begin{array}{cl}
(iv) & \textit{dim ColSp}(A) = \textit{dim ColSp}(A^T) = r \\
(v) & \mathbb{R}^m = \textit{ColSp}(A) \oplus \textit{NulSp}(A^T) \\
(vi) & \mathbb{R}^n = \textit{ColSp}(A^T) \oplus \textit{NulSp}(A)
\end{array}
$$

PROOF: Let subspace $V \subset \mathbb{R}^m$.

$$
\begin{array}{llcccl}
(v)\ \mathbb{R}^m & = & V & \oplus & V^\perp & \text{(Property of orthogonal complements)} \\
 & = & \text{ColSp}(A) & \oplus & \text{ColSp}(A)^\perp & \text{(Let } V := \text{ColSp}(A)) \\
 & = & \text{ColSp}(A) & \oplus & \text{NulSp}(A^T) & \text{(Part } (ii))
\end{array}
$$

Fundamental Subspaces of a Matrix

Theorem

(Fundamental Theorem of Linear Algebra – FTLA)

Let matrix $A \in \mathbb{R}^{m \times n}$ s.t. rank$(A) = r$. Then the fundamental subspaces of A are related as so:

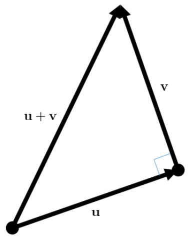
(i)	*RowSp*(A)	$=$	*ColSp*(A^T)	(iv)	*dim ColSp*(A) $=$ *dim ColSp*$(A^T) = r$
(ii)	*ColSp*$(A)^\perp$	$=$	*NulSp*(A^T)	(v)	$\mathbb{R}^m =$ *ColSp*$(A) \oplus$ *NulSp*(A^T)
(iii)	*ColSp*$(A^T)^\perp$	$=$	*NulSp*(A)	(vi)	$\mathbb{R}^n =$ *ColSp*$(A^T) \oplus$ *NulSp*(A)

PROOF: Let subspace $W \subset \mathbb{R}^n$.

$$
\begin{aligned}
(vi)\ \mathbb{R}^n &= W \oplus W^\perp && \text{(Property of orthogonal complements)}\\
&= \text{ColSp}(A^T) \oplus \text{ColSp}(A^T)^\perp && \text{(Let } W := \text{ColSp}(A^T))\\
&= \text{ColSp}(A^T) \oplus \text{NulSp}(A) && \text{(Part } (iii))
\end{aligned}
$$

□

Pythagorean Thm for Orthogonal Vectors (PTFOV)



Theorem

(Pythagorean Theorem for Orthogonal Vectors (PTFOV))

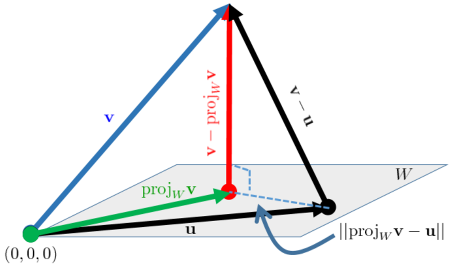
Vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ *are orthogonal* $\iff ||\mathbf{u}+\mathbf{v}||_2^2 = ||\mathbf{u}||_2^2 + ||\mathbf{v}||_2^2$

PROOF: (Recall that $||\cdot||_2$ denotes the Euclidean norm on $\mathbb{R}^n$.)

$$\begin{aligned} ||\mathbf{u}+\mathbf{v}||_2^2 &= (\mathbf{u}+\mathbf{v})^T(\mathbf{u}+\mathbf{v}) = (\mathbf{u}^T+\mathbf{v}^T)(\mathbf{u}+\mathbf{v}) = \mathbf{u}^T\mathbf{u}+\mathbf{u}^T\mathbf{v}+\mathbf{v}^T\mathbf{u}+\mathbf{v}^T\mathbf{v} \\ &= ||\mathbf{u}||_2^2 + \mathbf{u}\cdot\mathbf{v} + \mathbf{v}\cdot\mathbf{u} + ||\mathbf{v}||_2^2 = ||\mathbf{u}||_2^2 + 2(\mathbf{u}\cdot\mathbf{v}) + ||\mathbf{v}||_2^2 \end{aligned}$$

$$\therefore \quad ||\mathbf{u}+\mathbf{v}||_2^2 = ||\mathbf{u}||_2^2 + ||\mathbf{v}||_2^2 \iff 2(\mathbf{u}\cdot\mathbf{v}) = 0 \iff \mathbf{u}\cdot\mathbf{v} = 0 \iff \mathbf{u}\perp\mathbf{v} \qquad \square$$

Best Approximation Theorem



Theorem

(Best Approximation Theorem)

Let W be a subspace of $\mathbb{R}^n$ and $\mathbf{v} \in \mathbb{R}^n$ s.t. $\mathbf{v} \notin W$. Then:

$$||\mathbf{v} - \textit{proj}_W\mathbf{v}||_2 < ||\mathbf{v} - \mathbf{u}||_2 \qquad \forall \mathbf{u} \in S \text{ s.t. } \mathbf{u} \neq \textit{proj}_W\mathbf{v}$$

i.e. the projection of $\mathbf{v}$ onto W is the "closest" vector in W to $\mathbf{v}$ which is not in W.

$\textit{proj}_W\mathbf{v}$ is called the ***best approximation*** *to $\mathbf{v}$ in subspace W.*

Theorem

(Best Approximation Theorem)

Let W be a subspace of $\mathbb{R}^n$ and $\mathbf{v} \in \mathbb{R}^n$ s.t. $\mathbf{v} \notin W$. Then:

$$||\mathbf{v} - \textit{proj}_W \mathbf{v}||_2 < ||\mathbf{v} - \mathbf{u}||_2 \qquad \forall \mathbf{u} \in S \text{ s.t. } \mathbf{u} \neq \textit{proj}_W \mathbf{v}$$

i.e. the projection of $\mathbf{v}$ onto W is the "closest" vector in W to $\mathbf{v}$ which is not in W.

*$\textit{proj}_W \mathbf{v}$ is called the **best approximation** to $\mathbf{v}$ in subspace W.*

PROOF: Let $\mathbf{u} \in W$ s.t. $\mathbf{u} \neq \text{proj}_W \mathbf{v}$. Then:

$\mathbf{v} - \mathbf{u} = \mathbf{v} + \vec{\mathbf{0}} - \mathbf{u} = \mathbf{v} + (\text{proj}_W \mathbf{v} - \text{proj}_W \mathbf{v}) - \mathbf{u} = (\mathbf{v} - \text{proj}_W \mathbf{v}) + (\text{proj}_W \mathbf{v} - \mathbf{u})$

Now, $\mathbf{u} \in W$ and $\text{proj}_W \mathbf{v} \in W \implies (\text{proj}_W \mathbf{v} - \mathbf{u}) \in W$

Moreover, $(\mathbf{v} - \text{proj}_W \mathbf{v}) \perp W \implies (\mathbf{v} - \text{proj}_W \mathbf{v}) \perp (\text{proj}_W \mathbf{v} - \mathbf{u})$ Hence:

$$\mathbf{v} - \mathbf{u} = (\mathbf{v} - \text{proj}_W \mathbf{v}) + (\text{proj}_W \mathbf{v} - \mathbf{u}) \stackrel{PTFOV}{\Longrightarrow} ||\mathbf{v} - \mathbf{u}||_2^2 = ||\mathbf{v} - \text{proj}_W \mathbf{v}||_2^2 + ||\text{proj}_W \mathbf{v} - \mathbf{u}||_2^2$$

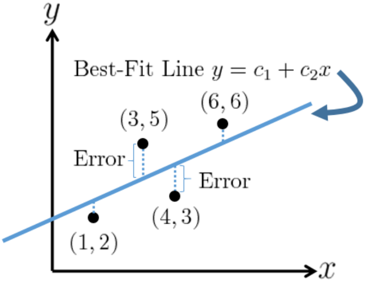
Since $\mathbf{u} \neq \text{proj}_W \mathbf{v}$, $\quad ||\text{proj}_W \mathbf{v} - \mathbf{u}||_2^2 > 0$

$$\implies ||\mathbf{v} - \mathbf{u}||_2^2 = ||\mathbf{v} - \text{proj}_W \mathbf{v}||_2^2 + ||\text{proj}_W \mathbf{v} - \mathbf{u}||_2^2 > ||\mathbf{v} - \text{proj}_W \mathbf{v}||_2^2$$

$$\implies ||\mathbf{v} - \mathbf{u}||_2^2 > ||\mathbf{v} - \text{proj}_W \mathbf{v}||_2^2 \implies ||\mathbf{v} - \mathbf{u}||_2 > ||\mathbf{v} - \text{proj}_W \mathbf{v}||_2 \quad \square$$

The Least-Squares Problem & Solution (Motivation)

Consider fitting a line to a set of points:



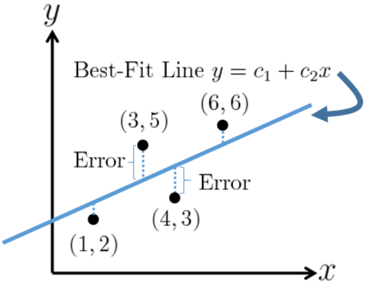
Assume (foolishly) that the line $y = c_1 + c_2 x$ contains all four points. Then:

$$\begin{cases} c_1 + (1)c_2 = 2 \\ c_1 + (3)c_2 = 5 \\ c_1 + (4)c_2 = 3 \\ c_1 + (6)c_2 = 6 \end{cases} \iff \underbrace{\begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 4 \\ 1 & 6 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} 2 \\ 5 \\ 3 \\ 6 \end{bmatrix}}_{\mathbf{b}} \leftarrow$$

Overdetermined Inconsistent Linear System $\therefore$ No Solution

The Least-Squares Problem & Solution (Motivation)

Consider fitting a line to a set of points:



Clearly, there's a best-fit line that minimizes the sum of the errors.
In practice, it's preferred to minimize the sum of the **squares** of the errors.

The overdetermined inconsistent linear system is called the **least-squares problem** & the best-fit line is called the **least-squares solution**.

The Least-Squares Problem & Solution (Definition)

Definition

(Least-Squares Problem & Solution)

Let $A \in \mathbb{R}^{m\times n}$ such that $m > n$ and $\mathbf{b} \notin \text{ColSp}(A)$ such that linear system $A\mathbf{x} = \mathbf{b}$ is inconsistent & overdetermined. Then:

The **least-squares problem** is to find $\mathbf{x} \in \mathbb{R}^n$ s.t. $||\mathbf{b} - A\mathbf{x}||_2^2$ is minimized:

$$\min_{\mathbf{x}\in\mathbb{R}^n} ||\mathbf{b} - A\mathbf{x}||_2^2$$

<u>REMARK:</u> Vector $(\mathbf{b} - A\mathbf{x})$ is called the **residual** of the linear system.

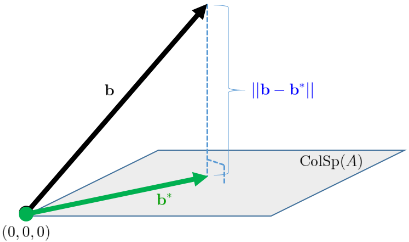
Vector $\mathbf{x}^* \in \mathbb{R}^n$ is a **least-squares solution** to $A\mathbf{x} = \mathbf{b}$ if:

$$\min_{\mathbf{x}\in\mathbb{R}^n} ||\mathbf{b} - A\mathbf{x}||_2^2 = ||\mathbf{b} - A\mathbf{x}^*||_2^2$$

i.e. $||\mathbf{b} - A\mathbf{x}^*||_2^2$ is the minimum square-norm of the residual.

Finding Least-Squares Solution (Derivation)

So how to find $\mathbf{x}^* \in \text{ColSp}(A)$ that minimizes $||\mathbf{b} - A\mathbf{x}||_2^2$??



Let $\mathbf{b}^* = \text{proj}_{\text{ColSp}(A)}\mathbf{b}$ be the best approx. to $\mathbf{b}$ (Best Approx. Thm)

Then $\mathbf{b}^* \in \text{ColSp}(A) \implies \mathbf{b}^* = A\mathbf{x}^*$ (since $A\mathbf{x}^* \in \text{ColSp}(A)$)

Observe that $(\mathbf{b} - \mathbf{b}^*) \perp \text{ColSp}(A) \implies$ Residual $(\mathbf{b} - A\mathbf{x}^*) \perp \text{ColSp}(A)$

$\implies \quad (\mathbf{b} - A\mathbf{x}^*) \in \text{ColSp}(A)^\perp$ (Defn of Orthogonal Complement)

$\implies \quad (\mathbf{b} - A\mathbf{x}^*) \in \text{NulSp}(A^T)$ (Fund. Subspaces of Matrix Thm)

$\implies \quad A^T(\mathbf{b} - A\mathbf{x}^*) = \vec{\mathbf{0}}$ (Defn of Null Space of A^T)

$\implies \quad A^T\mathbf{b} - A^TA\mathbf{x}^* = \vec{\mathbf{0}}$ (Distribute Left-Multiplication by A^T)

$\implies \quad A^TA\mathbf{x}^* = A^T\mathbf{b}$ (**Normal Equations**)

Full-Rank Least-Squares Solution using Normal Eqn's

Proposition

(Full-Rank Least-Squares Procedure using Normal Equations)

GIVEN: $m \times n$ $(m \geq n)$ *linear system* $A\mathbf{x} = \mathbf{b}$, *full column rank* A, $\mathbf{b} \notin$ *ColSp*(A).

TASK: *Find Least-Squares Solution* $\mathbf{x}^*$ *s.t.* $||\mathbf{b} - A\mathbf{x}||_2^2$ *is minimized.*

(1) *Form* ***normal equations*** *for* $\mathbf{x}^*$: $A^T A\mathbf{x}^* = A^T\mathbf{b}$

(2) *Solve normal equations for* $\mathbf{x}^*$: $\left[\, A^T A \,\middle|\, A^T\mathbf{b} \,\right] \xrightarrow{Gauss-Jordan} \left[\, I \,\middle|\, \mathbf{x}^* \,\right]$

(3) *Minimize square-norm of Residual:* $\min_{\mathbf{x}\in\mathbb{R}^n} ||\mathbf{b} - A\mathbf{x}||_2^2 = ||\mathbf{b} - A\mathbf{x}^*||_2^2$

(4) *Find Projection Matrix onto ColSp*(A): $\bar{P} = A(A^T A)^{-1}A^T$

(5) *Find Best Approximation* $\mathbf{b}^* \in$ *ColSp*(A) *to* $\mathbf{b}$: $\mathbf{b}^* = \bar{P}\mathbf{b} = A\mathbf{x}^*$

Full-Rank Least-Squares Solution using Reduced QR

Proposition

(Full-Rank Least-Squares Procedure using Reduced QR)

GIVEN: $m \times n$ $(m \geq n)$ *linear system* $A\mathbf{x} = \mathbf{b}$, *full column rank* A, $\mathbf{b} \notin$ *ColSp*(A).

TASK: *Find Least-Squares Solution* $\mathbf{x}^*$ *s.t.* $||\mathbf{b} - A\mathbf{x}||_2^2$ *is minimized.*

(1) *Perform Reduced QR Factorization using CGS-EN:* $A = \hat{Q}\hat{R}$
(Recall that with Reduced QR, $\hat{Q}$ *is* $m \times n$ *and* $\hat{R}$ *is* $n \times n$.*)*

(2) *Find Projection Matrix onto ColSp*(A)*:* $\bar{P} = \hat{Q}\hat{Q}^T$

(3) *Find Best Approximation* $\mathbf{b}^* \in$ *ColSp*(A) *to* $\mathbf{b}$*:* $\mathbf{b}^* = \bar{P}\mathbf{b}$

(4) *Minimize square-norm of Residual:* $\min_{\mathbf{x} \in \mathbb{R}^n} ||\mathbf{b} - A\mathbf{x}||_2^2 = ||\mathbf{b} - \bar{P}\mathbf{b}||_2^2$

(5) *Back-solve linear system* $\hat{R}\mathbf{x}^* = \hat{Q}^T\mathbf{b}$ *for* $\mathbf{x}^*$.

Full QR Factorization via CGS-EN

Proposition

(Full QR Factorization via CGS-EN)

GIVEN: Tall or square $(m \geq n)$ full column rank matrix $A_{m\times n}$ with columns $\mathbf{a}_k$.
TASK: Factor $A = QR$ where $Q_{m\times m}$ has orthonormal columns $\widehat{\mathbf{q}}_k$ and $R_{m\times n}$ is upper triangular.

(1) Perform Classical Gram-Schmidt w/ early normalization on the columns of A, $\{\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n\}$:

$$\hat{Q} = \begin{bmatrix} | & | & & | \\ \widehat{\mathbf{q}}_1 & \widehat{\mathbf{q}}_2 & \cdots & \widehat{\mathbf{q}}_n \\ | & | & & | \end{bmatrix}, \qquad \hat{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1n} \\ 0 & r_{22} & r_{23} & \cdots & r_{2n} \\ 0 & 0 & r_{33} & \cdots & r_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & r_{nn} \end{bmatrix}$$

(2) Produce a basis $\{\mathbf{a}_{n+1}, \mathbf{a}_{n+2}, \cdots, \mathbf{a}_m\}$ for orthogonal complement of column space of A:

$$\left[\, A^T \,\middle|\, \vec{\mathbf{0}} \,\right] \xrightarrow{Gauss-Jordan} \left[\, \text{RREF}(A^T) \,\middle|\, \vec{\mathbf{0}} \,\right]$$

(3) Perform CGS-EN on the basis $\{\mathbf{a}_{n+1}, \mathbf{a}_{n+2}, \cdots, \mathbf{a}_m\}$, resulting in $\hat{Q}_r$ matrix.
(4) Form Q by augmenting $\hat{Q}_r$ to $\hat{Q}$, and form R by augmenting zero matrix below $\hat{R}$:

$$Q_{m\times m} := \left[\hat{Q}_{m\times n} \quad \hat{Q}_r\right] = \begin{bmatrix} | & | & & | & | & & | \\ \widehat{\mathbf{q}}_1 & \widehat{\mathbf{q}}_2 & \cdots & \widehat{\mathbf{q}}_n & \widehat{\mathbf{q}}_{n+1} & \cdots & \widehat{\mathbf{q}}_m \\ | & | & & | & | & & | \end{bmatrix}, \qquad R_{m\times n} := \begin{bmatrix} \hat{R}_{n\times n} \\ O \end{bmatrix}$$

Full-Rank Least-Squares Solution using Full QR

Proposition

(Full-Rank Least-Squares Procedure using Full QR)

GIVEN: $m \times n$ $(m \geq n)$ *linear system* $A\mathbf{x} = \mathbf{b}$, *full column rank* A, $\mathbf{b} \notin$ *ColSp*(A).

TASK: *Find Least-Squares Solution* $\mathbf{x}^*$ *s.t.* $||\mathbf{b} - A\mathbf{x}||_2^2$ *is minimized.*

(1) Perform Full QR Factorization using CGS-EN: $A = QR$

$$Q_{m\times m} := \left[\hat{Q}_{m\times n} \quad \hat{Q}_r\right], \qquad R_{m\times n} := \begin{bmatrix} \hat{R}_{n\times n} \\ O \end{bmatrix}$$

(2) Find Projection Matrix onto ColSp(A): $\bar{P} = \hat{Q}\hat{Q}^T$

(3) Find best Approximation $\mathbf{b}^* \in$ ColSp(A) to $\mathbf{b}$: $\mathbf{b}^* = \bar{P}\mathbf{b}$

(4) Find Projection Matrix onto ColSp$(A)^\perp$: $\bar{P}_r = \hat{Q}_r\hat{Q}_r^T$

(5) Minimize square-norm of Residual: $\min_{\mathbf{x}\in\mathbb{R}^n} ||\mathbf{b} - A\mathbf{x}||_2^2 = ||\bar{P}_r\mathbf{b}||_2^2$

(6) Back-solve linear system $\hat{R}\mathbf{x}^* = \hat{Q}^T\mathbf{b}$ *for* $\mathbf{x}^*$.

PART III:

Orthogonal Matrices

Definition

Properties

Determinants

Preservation

Orthogonal Matrices (Definition & Properties)

The square matrix Q produced from the Full QR Factorization is special:

Definition

(Orthogonal Matrix)

A square matrix Q is **orthogonal** if its columns are orthonormal.

Orthogonal matrices have some very nice properties:

Theorem

(Properties of Orthogonal Matrices)

Let Q be an $m \times m$ square matrix. Then, the following properties are all equivalent:

(a) *Q is an orthogonal matrix*
(b) *The columns of Q are orthonormal*
(c) $Q^T Q = QQ^T = I$
(d) $Q^{-1} = Q^T$
(e) *Q^T is an orthogonal matrix*
(f) *The rows of Q are orthonormal*

Theorem

(Properties of Orthogonal Matrices)

Let Q be an $m \times m$ square matrix. Then, the following properties are all equivalent:

$$
\begin{array}{ll}
(a) & Q \text{ is an orthogonal matrix} \\
(b) & \text{The columns of } Q \text{ are orthonormal} \\
(c) & Q^T Q = QQ^T = I \\
(d) & Q^{-1} = Q^T \\
(e) & Q^T \text{ is an orthogonal matrix} \\
(f) & \text{The rows of } Q \text{ are orthonormal}
\end{array}
$$

PROOF: $[(a) \iff (b)]$ Follows immediately from the definition of an orthogonal matrix.

Orthogonal Matrices (Properties)

Theorem

(Properties of Orthogonal Matrices)

Let Q be an $m \times m$ square matrix. Then, the following properties are all equivalent:

$$\begin{array}{ll} (a) & Q \text{ is an orthogonal matrix} \\ (b) & \text{The columns of } Q \text{ are orthonormal} \\ (c) & Q^TQ = QQ^T = I \\ (d) & Q^{-1} = Q^T \\ (e) & Q^T \text{ is an orthogonal matrix} \\ (f) & \text{The rows of } Q \text{ are orthonormal} \end{array}$$

PROOF: $[(b) \iff (c)]$ The columns of $Q, \widehat{\mathbf{q}}_1, \cdots, \widehat{\mathbf{q}}_m,$ are orthonormal.

$\iff \quad \widehat{\mathbf{q}}_i^T\widehat{\mathbf{q}}_j = \delta_{ij}$ (definition of orthonormal vectors)

$$\iff \quad Q^TQ = \begin{bmatrix} \text{---} & \widehat{\mathbf{q}}_1^T & \text{---} \\ & \vdots & \\ \text{---} & \widehat{\mathbf{q}}_m^T & \text{---} \end{bmatrix} \begin{bmatrix} | & & | \\ \widehat{\mathbf{q}}_1 & \cdots & \widehat{\mathbf{q}}_m \\ | & & | \end{bmatrix} = \begin{bmatrix} \widehat{\mathbf{q}}_1^T\widehat{\mathbf{q}}_1 & \cdots & \widehat{\mathbf{q}}_1^T\widehat{\mathbf{q}}_m \\ \vdots & \ddots & \vdots \\ \widehat{\mathbf{q}}_m^T\widehat{\mathbf{q}}_1 & \cdots & \widehat{\mathbf{q}}_m^T\widehat{\mathbf{q}}_m \end{bmatrix}$$

$$\iff \quad Q^TQ = \begin{bmatrix} \widehat{\mathbf{q}}_1^T\widehat{\mathbf{q}}_1 & \cdots & \widehat{\mathbf{q}}_1^T\widehat{\mathbf{q}}_m \\ \vdots & \ddots & \vdots \\ \widehat{\mathbf{q}}_m^T\widehat{\mathbf{q}}_1 & \cdots & \widehat{\mathbf{q}}_m^T\widehat{\mathbf{q}}_m \end{bmatrix} = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix} = I \quad (\text{since } \widehat{\mathbf{q}}_i^T\widehat{\mathbf{q}}_j = \delta_{ij})$$

$\iff \quad Q^TQ = I$

Orthogonal Matrices (Properties)

Theorem

(Properties of Orthogonal Matrices)

Let Q be an $m \times m$ square matrix. Then, the following properties are all equivalent:

$$\begin{array}{ll} (a) & Q \text{ is an orthogonal matrix} \\ (b) & \text{The columns of } Q \text{ are orthonormal} \\ (c) & Q^T Q = QQ^T = I \\ (d) & Q^{-1} = Q^T \\ (e) & Q^T \text{ is an orthogonal matrix} \\ (f) & \text{The rows of } Q \text{ are orthonormal} \end{array}$$

PROOF: $[(b) \iff (c)]$ The columns of $Q, \widehat{\mathbf{q}}_1, \cdots, \widehat{\mathbf{q}}_m,$ are orthonormal.

$\iff \quad \widehat{\mathbf{q}}_i^T \widehat{\mathbf{q}}_j = \delta_{ij}$ (definition of orthonormal vectors)

$\iff \quad QQ^T = \widehat{\mathbf{q}}_1 \widehat{\mathbf{q}}_1^T + \cdots + \widehat{\mathbf{q}}_m \widehat{\mathbf{q}}_m^T$ (Outer product expansion of QQ^T)

$\iff \quad QQ^T = \bar{P}_m = I$ (since the columns of Q are an orthonormal basis for $\mathbb{R}^m$... ...and matrix $\bar{P}_m$ projects onto $\mathbb{R}^m$, meaning $\bar{P}_m \mathbf{u} = \mathbf{u} \ \ \forall \mathbf{u} \in \mathbb{R}^m$)

$\iff \quad Q^T Q = QQ^T = I$

Orthogonal Matrices (Properties)

Theorem

(Properties of Orthogonal Matrices)

Let Q be an $m \times m$ square matrix. Then, the following properties are all equivalent:

$$
\begin{array}{ll}
(a) & Q \text{ is an orthogonal matrix} \\
(b) & \text{The columns of } Q \text{ are orthonormal} \\
(c) & Q^TQ = QQ^T = I \\
(d) & Q^{-1} = Q^T \\
(e) & Q^T \text{ is an orthogonal matrix} \\
(f) & \text{The rows of } Q \text{ are orthonormal}
\end{array}
$$

PROOF: $[(c) \iff (d)]$ $Q^TQ = QQ^T = I$.

$\iff$ $Q^{-1} = Q^T$ (definition of inverse of square matrix)

Orthogonal Matrices (Properties)

Theorem

(Properties of Orthogonal Matrices)

Let Q be an $m \times m$ square matrix. Then, the following properties are all equivalent:

(a)	*Q is an orthogonal matrix*
(b)	*The columns of Q are orthonormal*
(c)	$Q^T Q = QQ^T = I$
(d)	$Q^{-1} = Q^T$
(e)	*Q^T is an orthogonal matrix*
(f)	*The rows of Q are orthonormal*

PROOF: $[(d) \iff (e)]$ $Q^{-1} = Q^T$.

$\iff$	$Q^T Q = QQ^T = I$	(definition of inverse of square matrix)
$\iff$	$(Q^T Q)^T = (QQ^T)^T = I^T$	(transpose equation)
$\iff$	$Q^T (Q^T)^T = (Q^T)^T Q^T = I$	(transpose of matrix product and identity matrix)
$\iff$	The columns of Q^T are orthonormal	(since $(b) \iff (c)$)
$\iff$	Q^T is an orthogonal matrix	(definition of orthogonal matrix)

Orthogonal Matrices (Properties)

Theorem

(Properties of Orthogonal Matrices)

Let Q be an $m \times m$ square matrix. Then, the following properties are all equivalent:

$$
\begin{array}{ll}
(a) & Q \text{ is an orthogonal matrix} \\
(b) & \text{The columns of } Q \text{ are orthonormal} \\
(c) & Q^T Q = QQ^T = I \\
(d) & Q^{-1} = Q^T \\
(e) & Q^T \text{ is an orthogonal matrix} \\
(f) & \text{The rows of } Q \text{ are orthonormal}
\end{array}
$$

PROOF: $[(e) \iff (f)]$ Q^T is an orthogonal matrix.

$\iff$	The columns of Q^T are orthonormal	(definition of orthogonal matrix)
$\iff$	The rows of Q are orthonormal	(definition of transpose of a matrix)

$\therefore$ $(a) \iff (b) \iff (c) \iff (d) \iff (e) \iff (f)$

$\therefore$ The properties are all equivalent. □

Orthogonal Matrices (Determinants)

Corollary

(Orthogonal Matrices & Determinants)

(a) Q *is orthogonal matrix* $\implies$ *det*$(Q) = \pm 1$.
(b) *The converse is not necessarily true:* *det*$(Q) = \pm 1 \not\Longrightarrow Q$ *is orthogonal matrix*

Orthogonal Matrices (Determinants)

Corollary

(Orthogonal Matrices & Determinants)

(a) *Q is orthogonal matrix* $\Longrightarrow$ *det*$(Q) = \pm 1$.
(b) *The converse is not necessarily true: det*$(Q) = \pm 1 \not\Longrightarrow$ *Q is orthogonal matrix*

PROOF:

(a) Let Q be an orthogonal matrix. Then:

$$
\begin{array}{rcll}
 & Q^T Q = I & & \text{(Orthogonal Matrix Property)} \\
\Longrightarrow & \det(Q^T Q) = \det(I) & & \text{(Take determinant on both sides)} \\
\Longrightarrow & \det(Q^T Q) = \det(I) = 1 & & \text{(Determinant of Identity Matrix is One)} \\
\Longrightarrow & \det(Q^T) \cdot \det(Q) = 1 & & \text{(Determinant of Matrix Product)} \\
\Longrightarrow & \det(Q) \cdot \det(Q) = 1 & & \text{(Determinant of Matrix Transpose)} \\
\Longrightarrow & [\det(Q)]^2 = 1 & & \text{(Determinant of Matrix Transpose)} \\
\Longrightarrow & |\det(Q)| = 1 & & \text{(Take Square Roots on both sides)} \\
\Longrightarrow & \det(Q) = \pm 1 & & \text{(Definition of Absolute Value)}
\end{array}
$$

Orthogonal Matrices (Determinants)

Corollary

(Orthogonal Matrices & Determinants)

(a) *Q is orthogonal matrix* $\implies$ *det*$(Q) = \pm 1.$
(b) *The converse is not necessarily true: det*$(Q) = \pm 1 \;\not\Longrightarrow\;$ *Q is orthogonal matrix*

PROOF:

(b) Below are several counterexamples:

$$D := \begin{bmatrix} 0.1 & 0 \\ 0 & 10 \end{bmatrix}, \qquad U := \begin{bmatrix} -2 & 1 & -1 & 1 \\ 0 & 4 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1/8 \end{bmatrix}$$

Then, det$(D) = (0.1)(10) = 1,$ and det$(U) = (-2)(4)(1)(1/8) = -1$ ... but:

$$D^T D = \begin{bmatrix} 0.01 & 0 \\ 0 & 100 \end{bmatrix} \neq \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_{2\times 2} \implies D \text{ is not orthogonal}$$

$$U^T U = \begin{bmatrix} 4 & -2 & 2 & -2 \\ -2 & 17 & -1 & 5 \\ 2 & -1 & 2 & 1 \\ -2 & 5 & 1 & 385/64 \end{bmatrix} \neq \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = I_4 \implies U \text{ is not orthogonal} \qquad \square$$

Orthogonal Matrices (Preservation)

The following theorem is the cornerstone to many stable numerical algorithms involving orthogonal matrices:

Theorem

(Orthogonal Preservation Theorem)

Consider the Euclidean inner product space $(\mathbb{R}^n, \langle\cdot,\cdot\rangle_2)$ *where* $\mathbf{v}, \mathbf{w}, \mathbf{x} \in \mathbb{R}^n$ *and*

$$\begin{array}{lrcl} \textit{Inner product} & \langle \mathbf{v}, \mathbf{w} \rangle_2 & := & \mathbf{v}^T\mathbf{w} \\ \textit{Induced norm} & ||\mathbf{x}||_2 & := & \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle_2} \\ \textit{Induced metric} & d_2(\mathbf{v}, \mathbf{w}) & := & ||\mathbf{v} - \mathbf{w}||_2 \end{array}$$

Then orthogonal matrix $Q \in \mathbb{R}^{n\times n}$ *preserves inner products, norms & metrics:*

$(i)\ \langle Q\mathbf{v}, Q\mathbf{w} \rangle_2 = \langle \mathbf{v}, \mathbf{w} \rangle_2, \qquad (ii)\ ||Q\mathbf{x}||_2 = ||\mathbf{x}||_2, \qquad (iii)\ d_2(Q\mathbf{v}, Q\mathbf{w}) = d_2(\mathbf{v}, \mathbf{w})$

Orthogonal Matrices (Preservation)

Theorem

(Orthogonal Preservation Theorem)

Consider the <u>*Euclidean*</u> *inner product space* $(\mathbb{R}^n, \langle \cdot, \cdot \rangle_2)$ *where* $\mathbf{v}, \mathbf{w}, \mathbf{x} \in \mathbb{R}^n$ *and*

Inner product	$\langle \mathbf{v}, \mathbf{w} \rangle_2$	$:=$	$\mathbf{v}^T \mathbf{w}$
Induced norm	$\|\|\mathbf{x}\|\|_2$	$:=$	$\sqrt{\langle \mathbf{x}, \mathbf{x} \rangle_2}$
Induced metric	$d_2(\mathbf{v}, \mathbf{w})$	$:=$	$\|\|\mathbf{v} - \mathbf{w}\|\|_2$

Then orthogonal matrix $Q \in \mathbb{R}^{n \times n}$ *preserves inner products, norms & metrics:*

$$(i)\ \langle Q\mathbf{v}, Q\mathbf{w} \rangle_2 = \langle \mathbf{v}, \mathbf{w} \rangle_2, \qquad (ii)\ ||Q\mathbf{x}||_2 = ||\mathbf{x}||_2, \qquad (iii)\ d_2(Q\mathbf{v}, Q\mathbf{w}) = d_2(\mathbf{v}, \mathbf{w})$$

<u>PROOF:</u>

$$(i)\ \ \langle Q\mathbf{v}, Q\mathbf{w} \rangle_2 := (Q\mathbf{v})^T (Q\mathbf{w}) \overset{T4}{=} \mathbf{v}^T (Q^T Q)\mathbf{w} \overset{Q}{=} \mathbf{v}^T I \mathbf{w} \overset{I}{=} \mathbf{v}^T \mathbf{w} := \langle \mathbf{v}, \mathbf{w} \rangle_2$$

$$(ii)\ \ ||Q\mathbf{x}||_2^2 = \langle Q\mathbf{x}, Q\mathbf{x} \rangle_2 \overset{(i)}{=} \langle \mathbf{x}, \mathbf{x} \rangle_2 = ||\mathbf{x}||_2^2 \overset{\sqrt{\cdot}}{\Longrightarrow} ||Q\mathbf{x}||_2 = ||\mathbf{x}||_2$$

$$(iii)\ d_2(Q\mathbf{v}, Q\mathbf{w}) := ||Q\mathbf{v} - Q\mathbf{w}||_2 \overset{M3}{=} ||Q(\mathbf{v} - \mathbf{w})||_2 \overset{(ii)}{=} ||\mathbf{v} - \mathbf{w}||_2 = d_2(\mathbf{v}, \mathbf{w}) \qquad \square$$

Fin.