Orthonormal Bases, Gram-Schmidt, Reduced QR

Linear Algebra

Josh Engwer

TTU

30 Oct 2015 (revised 14 Nov 2021)

PART I:

Orthogonal Sets

Orthogonal Bases
Orthonormal Bases

Projections onto Subspaces
Converting $[\mathbf{x}]_{\mathcal{E}} \to [\mathbf{x}]_{\widehat{\mathcal{Q}}}$

Orthogonal Sets (Definition & Independence)

Definition

(Orthogonal Set)

Let $(V, \langle\cdot,\cdot\rangle)$ be an inner product space.

Then set $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$ is an **orthogonal set** if: $\begin{cases} \langle \mathbf{q}_i, \mathbf{q}_j \rangle \neq 0 & \text{for } i = j \\ \langle \mathbf{q}_i, \mathbf{q}_j \rangle = 0 & \text{for } i \neq j \end{cases}$

i.e. Every distinct pair of vectors in the set is orthogonal.

Corollary

(Orthogonal Sets are Linearly Independent Corollary – OSLIC)

Given inner product space $(V, \langle\cdot,\cdot\rangle)$ and orthogonal set $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$.
Then, the orthogonal set is linearly independent.

Orthogonal Sets (Definition & Independence)

Definition

(Orthogonal Set)

Let $(V, \langle\cdot,\cdot\rangle)$ be an inner product space.

Then set $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$ is an **orthogonal set** if: $\begin{cases} \langle \mathbf{q}_i, \mathbf{q}_j \rangle & \neq & 0 & \text{for} & i = j \\ \langle \mathbf{q}_i, \mathbf{q}_j \rangle & = & 0 & \text{for} & i \neq j \end{cases}$

i.e. Every distinct pair of vectors in the set is orthogonal.

Corollary

(Orthogonal Sets are Linearly Independent Corollary – OSLIC)

Given inner product space $(V, \langle\cdot,\cdot\rangle)$ and orthogonal set $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$.
Then, the orthogonal set is linearly independent.

PROOF: Let scalars $c_1, c_2, \ldots, c_n \in \mathbb{R}$ and let $c_1\mathbf{q}_1 + c_2\mathbf{q}_2 + \cdots + c_n\mathbf{q}_n = \vec{\mathbf{0}}$.

Let $k \in \{1, 2, \ldots, n\}$. Then, compute c_k by taking inner product with $\mathbf{q}_k$ on both sides of equation:

$$\begin{array}{rccll} & \langle \mathbf{q}_k, c_1\mathbf{q}_1 + \cdots + c_k\mathbf{q}_k + \cdots + c_n\mathbf{q}_n \rangle & = & \langle \mathbf{q}_k, \vec{\mathbf{0}} \rangle & \\ \Longrightarrow & c_1\langle \mathbf{q}_k, \mathbf{q}_1 \rangle + \cdots + c_k\langle \mathbf{q}_k, \mathbf{q}_k \rangle + \cdots + c_n\langle \mathbf{q}_k, \mathbf{q}_n \rangle & = & 0 & \text{(Linearity of IP)} \\ \Longrightarrow & c_1 \cdot 0 + \cdots + c_k \cdot ||\mathbf{q}_k||^2 + \cdots + c_n \cdot 0 & = & 0 & \text{(Defn of Orthogonal Set)} \\ \Longrightarrow & c_k \cdot ||\mathbf{q}_k||^2 & = & 0 & (\text{Since } \langle \mathbf{q}_k, \mathbf{q}_i \rangle = 0) \\ \Longrightarrow & c_k & = & 0 & (\text{Since } \langle \mathbf{q}_k, \mathbf{q}_k \rangle \neq 0) \end{array}$$

Since k was arbitrarily chosen, $c_1 = c_2 = \cdots = c_n = 0 \Longrightarrow$ Set is linearly independent □

Orthogonal & Orthonormal Bases (Definition)

Orthonormal bases are extremely useful in applied mathematics:

Definition

(Orthogonal Basis)

Let $(V, \langle\cdot,\cdot\rangle)$ be a finite-dimensional inner product space.
Then basis $\mathcal{Q} = \{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$ is an **orthogonal basis** for V if:

$$\langle \mathbf{q}_i, \mathbf{q}_j \rangle = 0 \text{ for } i \neq j$$

i.e. Every distinct pair of basis vectors in $\mathcal{Q}$ is orthogonal.

Definition

(Orthonormal Basis)

Let $(V, \langle\cdot,\cdot\rangle)$ be a finite-dimensional inner product space.
Then basis $\widehat{\mathcal{Q}} = \{\widehat{\mathbf{q}}_1, \widehat{\mathbf{q}}_2, \ldots, \widehat{\mathbf{q}}_n\}$ is an **orthonormal basis** for V if:

$$\langle \widehat{\mathbf{q}}_i, \widehat{\mathbf{q}}_j \rangle = \delta_{ij} \iff \begin{cases} \langle \widehat{\mathbf{q}}_i, \widehat{\mathbf{q}}_j \rangle = 0 \text{ if } i \neq j \\ \langle \widehat{\mathbf{q}}_i, \widehat{\mathbf{q}}_j \rangle = 1 \text{ if } i = j \end{cases}$$

i.e. Every distinct pair of unit basis vectors in $\widehat{\mathcal{Q}}$ is orthogonal.

Orthogonal & Orthonormal Bases (Examples)

Example of orthogonal basis for $\mathbb{R}^2$: $\mathcal{Q} = \left\{ \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\} \equiv \{\mathbf{q}_1, \mathbf{q}_2\}$

$$\text{since } \langle \mathbf{q}_1, \mathbf{q}_2 \rangle = \mathbf{q}_1^T \mathbf{q}_2 = (2)(1) + (-1)(2) = 0 \implies \mathbf{q}_1 \perp \mathbf{q}_2$$

However, $\mathcal{Q}$ is not an orthonormal basis for $\mathbb{R}^2$ since:

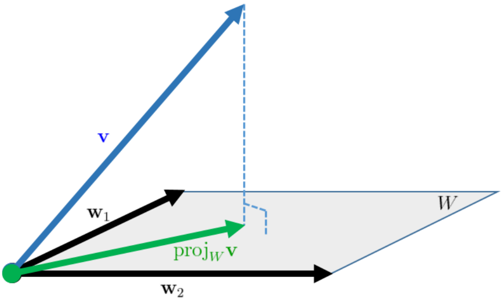
$$\langle \mathbf{q}_1, \mathbf{q}_1 \rangle = \mathbf{q}_1^T \mathbf{q}_1 = (2)(2) + (-1)(-1) = 5 \neq 1$$

and

$$\langle \mathbf{q}_2, \mathbf{q}_2 \rangle = \mathbf{q}_2^T \mathbf{q}_2 = (1)(1) + (2)(2) = 5 \neq 1$$

Example of orthonormal basis: $\widehat{\mathcal{Q}} = \left\{ \begin{bmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix}, \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} \right\} \equiv \{\widehat{\mathbf{q}}_1, \widehat{\mathbf{q}}_2\}$

(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 $\mathcal{Q} = \{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_p\}$ *be an orthogonal basis for subspace* W *of* $\mathbb{R}^n$*.* $(p \leq n)$

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

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

Converting $[\mathbf{x}]_{\mathcal{E}} \to [\mathbf{x}]_{\widehat{\mathcal{Q}}}$ (Procedure)

Coordinate vectors w.r.t orthonormal bases can be found w/o Gauss-Jordan:

Proposition

(Converting $[\mathbf{x}]_{\mathcal{E}} \to [\mathbf{x}]_{\widehat{\mathcal{Q}}}$)

Let $(V, \langle\cdot,\cdot\rangle)$ be a finite-dimensional inner product space.

Let $\widehat{\mathcal{Q}} = \{\widehat{\mathbf{q}}_1, \widehat{\mathbf{q}}_2, \ldots, \widehat{\mathbf{q}}_n\}$ be an ordered orthonormal basis for V.

Let $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ be the ordered standard basis for V.

GIVEN: Vector $\mathbf{x} \in V$ in standard basis coordinates: $\mathbf{x} = [\mathbf{x}]_{\mathcal{E}}$

TASK: Write vector $\mathbf{x}$ in non-std orthonormal basis $\widehat{\mathcal{Q}}$-coordinates: $[\mathbf{x}]_{\widehat{\mathcal{Q}}}$

$$\text{(1)} \quad [\mathbf{x}]_{\widehat{\mathcal{Q}}} = \mathit{proj}_{span(\widehat{\mathcal{Q}})}[\mathbf{x}]_{\mathcal{E}} = \mathit{proj}_{\widehat{\mathbf{q}}_1}\mathbf{x} + \mathit{proj}_{\widehat{\mathbf{q}}_2}\mathbf{x} + \cdots + \mathit{proj}_{\widehat{\mathbf{q}}_n}\mathbf{x} = \begin{bmatrix} \langle \mathbf{x}, \widehat{\mathbf{q}}_1 \rangle \\ \langle \mathbf{x}, \widehat{\mathbf{q}}_2 \rangle \\ \vdots \\ \langle \mathbf{x}, \widehat{\mathbf{q}}_n \rangle \end{bmatrix}$$

Converting $[\mathbf{x}]_{\mathcal{E}} \to [\mathbf{x}]_{\widehat{\mathcal{Q}}}$ (Example)

WEX 5-3-1: Let orthonormal basis $\widehat{\mathcal{Q}} = \left\{ \begin{bmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix}, \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} \right\} \equiv \{\widehat{\mathbf{q}}_1, \widehat{\mathbf{q}}_2\}$.

Find $[\mathbf{x}]_{\widehat{\mathcal{Q}}}$ if $[\mathbf{x}]_{\mathcal{E}} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$.

$$\langle \mathbf{x}, \widehat{\mathbf{q}}_1 \rangle = \left\langle \begin{bmatrix} 3 \\ -2 \end{bmatrix}, \begin{bmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix} \right\rangle = (3)\left(\frac{2}{\sqrt{5}}\right) + (-2)\left(-\frac{1}{\sqrt{5}}\right) = \frac{8}{\sqrt{5}}$$

$$\langle \mathbf{x}, \widehat{\mathbf{q}}_2 \rangle = \left\langle \begin{bmatrix} 3 \\ -2 \end{bmatrix}, \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} \right\rangle = (3)\left(\frac{1}{\sqrt{5}}\right) + (-2)\left(\frac{2}{\sqrt{5}}\right) = -\frac{1}{\sqrt{5}}$$

$\therefore$ $\boxed{[\mathbf{x}]_{\widehat{\mathcal{Q}}} = \begin{bmatrix} 8/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix}}$ OR $\boxed{[\mathbf{x}]_{\widehat{\mathcal{Q}}} = \left(\frac{8}{\sqrt{5}}, \frac{-1}{\sqrt{5}}\right)^T}$

PART II:

Classical Gram-Schmidt Orthonormalization w/ late normalization (CGS-LN)
Classical Gram-Schmidt Orthonormalization w/ early normalization (CGS-EN)

(Classical) Gram-Schmidt Orthonormalization

(Classical) Gram[†]-Schmidt[‡] is a procedure to produce an orthonormal basis:

Proposition

(Classical Gram-Schmidt w/ Late Normalization Procedure – CGS-LN)

GIVEN: *Induced-norm inner product space* $(V, \langle\cdot,\cdot\rangle, ||\cdot||)$ *& Basis* $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\} \subseteq V$.
TASK: *Find an orthonormal basis* $\widehat{\mathcal{Q}} = \{\widehat{\mathbf{q}}_1, \widehat{\mathbf{q}}_2, \ldots, \widehat{\mathbf{q}}_n\}$ *for span*$(\mathcal{B})$.

(1) Find orthogonal basis $\mathcal{Q}$ *as follows:* $\left(\mathcal{Q}_k := \{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_k\};\ k = 1, \cdots, n-1\right)$

$$
\begin{array}{rcllcl}
\mathbf{q}_1 & := & \mathbf{v}_1 & = & \mathbf{v}_1 \\
\mathbf{q}_2 & := & \mathbf{v}_2 - \mathit{proj}_{span(\mathcal{Q}_1)}\mathbf{v}_2 & = & \mathbf{v}_2 - \mathit{proj}_{\mathbf{q}_1}\mathbf{v}_2 \\
\mathbf{q}_3 & := & \mathbf{v}_3 - \mathit{proj}_{span(\mathcal{Q}_2)}\mathbf{v}_3 & = & \mathbf{v}_3 - \mathit{proj}_{\mathbf{q}_1}\mathbf{v}_3 - \mathit{proj}_{\mathbf{q}_2}\mathbf{v}_3 \\
 & \vdots & \vdots & \vdots & \ddots \\
\mathbf{q}_n & := & \mathbf{v}_n - \mathit{proj}_{span(\mathcal{Q}_{n-1})}\mathbf{v}_n & = & \mathbf{v}_n - \mathit{proj}_{\mathbf{q}_1}\mathbf{v}_n - \mathit{proj}_{\mathbf{q}_2}\mathbf{v}_n - \cdots - \mathit{proj}_{\mathbf{q}_{n-1}}\mathbf{v}_n
\end{array}
$$

(2) Find orthonormal basis $\widehat{\mathcal{Q}}$ *from* $\mathcal{Q}$ *by normalizing each basis vector:*

$$\widehat{\mathbf{q}}_1 = \frac{\mathbf{q}_1}{||\mathbf{q}_1||}, \qquad \widehat{\mathbf{q}}_2 = \frac{\mathbf{q}_2}{||\mathbf{q}_2||}, \qquad \cdots, \qquad \widehat{\mathbf{q}}_n = \frac{\mathbf{q}_n}{||\mathbf{q}_n||}$$

[†] J. Gram, *Om Raekkenudviklinger bestemte ved Hjaelp af de mindste Kvadraters Methode*, 1879.
[‡] E. Schmidt, "Zur Theorie der linearen und nichtlinearen Integralgleichungen", *M. Ann.*, **63** (1907).

(Classical) Gram-Schmidt Orthonormalization

(Classical) Gram[†]**-Schmidt**[‡] is a procedure to produce an orthonormal basis:

Proposition

(Classical Gram-Schmidt w/ Early Normalization Procedure – CGS-EN)

GIVEN: Induced-norm inner product space $(V, \langle\cdot,\cdot\rangle, ||\cdot||)$ *& Basis* $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\} \subseteq V$.
TASK: Find an orthonormal basis $\widehat{\mathcal{Q}} = \{\widehat{\mathbf{q}}_1, \widehat{\mathbf{q}}_2, \ldots, \widehat{\mathbf{q}}_n\}$ *for span*$(\mathcal{B})$.

(1) Find orthonormal basis $\widehat{\mathcal{Q}}$ *as follows:* $\left(\widehat{\mathcal{Q}}_k := \{\widehat{\mathbf{q}}_1, \widehat{\mathbf{q}}_2, \ldots, \widehat{\mathbf{q}}_k\};\ k = 1, \cdots, n-1\right)$

$$\begin{array}{lllll}
\mathbf{q}_1 & := & \mathbf{v}_1 & = & \mathbf{v}_1 & ;\ \widehat{\mathbf{q}}_1 := \mathbf{q}_1/||\mathbf{q}_1|| \\
\mathbf{q}_2 & := & \mathbf{v}_2 - \mathit{proj}_{span(\widehat{\mathcal{Q}}_1)}\mathbf{v}_2 & = & \mathbf{v}_2 - \mathit{proj}_{\widehat{\mathbf{q}}_1}\mathbf{v}_2 & ;\ \widehat{\mathbf{q}}_2 := \mathbf{q}_2/||\mathbf{q}_2|| \\
\mathbf{q}_3 & := & \mathbf{v}_3 - \mathit{proj}_{span(\widehat{\mathcal{Q}}_2)}\mathbf{v}_3 & = & \mathbf{v}_3 - \mathit{proj}_{\widehat{\mathbf{q}}_1}\mathbf{v}_3 - \mathit{proj}_{\widehat{\mathbf{q}}_2}\mathbf{v}_3 & ;\ \widehat{\mathbf{q}}_3 := \mathbf{q}_3/||\mathbf{q}_3|| \\
 & \vdots & \vdots & \vdots & \ddots & \vdots \\
\mathbf{q}_n & := & \mathbf{v}_n - \mathit{proj}_{span(\widehat{\mathcal{Q}}_{n-1})}\mathbf{v}_n & = & \mathbf{v}_n - \mathit{proj}_{\widehat{\mathbf{q}}_1}\mathbf{v}_n - \cdots - \mathit{proj}_{\widehat{\mathbf{q}}_{n-1}}\mathbf{v}_n & ;\ \widehat{\mathbf{q}}_n := \mathbf{q}_n/||\mathbf{q}_n||
\end{array}$$

[†] J. Gram, *Om Raekkenudviklinger bestemte ved Hjaelp af de mindste Kvadraters Methode*, 1879.
[‡] E. Schmidt, "Zur Theorie der linearen und nichtlinearen Integralgleichungen", *M. Ann.*, **63** (1907).

PART III

PART III:

Matrix-Vector Products $(A\mathbf{x})$ in terms of column combinations
Matrix-Matrix Products (AB) in terms of column combinations

Reduced QR Factorization via CGS-EN

Matrix-Vector Products $(A\mathbf{x})$ as column combinations

Let $A := \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \equiv \begin{bmatrix} | & | & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \\ | & | & | \end{bmatrix}$ and $\mathbf{x} := \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

Entry View:

$$A\mathbf{x} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{bmatrix}$$

Column View:

$$A\mathbf{x} = \begin{bmatrix} | & | & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \underbrace{x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3}_{\substack{\textit{linear combination of} \\ \textit{the columns of } A}} = \sum_{k=1}^{3} x_k \mathbf{a}_k$$

Matrix-Matrix Products (AB) as column combinations

$$A := \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \equiv \begin{bmatrix} | & | \\ \mathbf{a}_1 & \mathbf{a}_2 \\ | & | \end{bmatrix}, \quad B := \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}, \quad C := \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \end{bmatrix} \equiv \begin{bmatrix} | & | \\ \mathbf{c}_1 & \mathbf{c}_2 \\ | & | \end{bmatrix}$$

Entry View:

$$C := AB = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} (a_{11}b_{11} + a_{12}b_{21}) & (a_{11}b_{12} + a_{12}b_{22}) \\ (a_{21}b_{11} + a_{22}b_{21}) & (a_{21}b_{12} + a_{22}b_{22}) \\ (a_{31}b_{11} + a_{32}b_{21}) & (a_{31}b_{12} + a_{32}b_{22}) \end{bmatrix}$$

Column View:

$$C := AB \implies \begin{cases} c_{11} = a_{11}b_{11} + a_{12}b_{21} \\ c_{21} = a_{21}b_{11} + a_{22}b_{21} \\ c_{31} = a_{31}b_{11} + a_{32}b_{21} \\ \\ c_{12} = a_{11}b_{12} + a_{12}b_{22} \\ c_{22} = a_{21}b_{12} + a_{22}b_{22} \\ c_{32} = a_{31}b_{12} + a_{32}b_{22} \end{cases} \implies \begin{cases} \mathbf{c}_1 = b_{11}\mathbf{a}_1 + b_{21}\mathbf{a}_2 \\ \mathbf{c}_2 = b_{12}\mathbf{a}_1 + b_{22}\mathbf{a}_2 \end{cases}$$

$\therefore$ Each column of C is a linear combination of the columns of A built from a column of B.

Reduced QR Factorization via CGS-EN

Consider 4×3 matrix $A := \begin{bmatrix} | & | & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \\ | & | & | \end{bmatrix}$ such that A has full column rank.

1. Perform CGS-EN $\left(\text{with } \mathbf{v}_1 := \mathbf{a}_1,\ \mathbf{v}_2 := \mathbf{a}_2,\ \mathbf{v}_3 := \mathbf{a}_3;\ \widehat{\mathcal{Q}}_1 := \{\widehat{\mathbf{q}}_1\},\ \widehat{\mathcal{Q}}_2 := \{\widehat{\mathbf{q}}_1, \widehat{\mathbf{q}}_2\}\right)$:

$$\begin{cases} \mathbf{q}_1 & := & \mathbf{a}_1 & = & \mathbf{a}_1 & ; \widehat{\mathbf{q}}_1 := \mathbf{q}_1/||\mathbf{q}_1||_2 \\ \mathbf{q}_2 & := & \mathbf{a}_2 - \text{proj}_{span(\widehat{\mathcal{Q}}_1)}\mathbf{a}_2 & = & \mathbf{a}_2 - (\widehat{\mathbf{q}}_1^T\mathbf{a}_2)\widehat{\mathbf{q}}_1 & ; \widehat{\mathbf{q}}_2 := \mathbf{q}_2/||\mathbf{q}_2||_2 \\ \mathbf{q}_3 & := & \mathbf{a}_3 - \text{proj}_{span(\widehat{\mathcal{Q}}_2)}\mathbf{a}_3 & = & \mathbf{a}_3 - (\widehat{\mathbf{q}}_1^T\mathbf{a}_3)\widehat{\mathbf{q}}_1 - (\widehat{\mathbf{q}}_2^T\mathbf{a}_3)\widehat{\mathbf{q}}_2 & ; \widehat{\mathbf{q}}_3 := \mathbf{q}_3/||\mathbf{q}_3||_2 \end{cases}$$

2. Now, solve each equation for $\mathbf{a}_k$ and express each $\mathbf{q}_k$ in terms of $\widehat{\mathbf{q}}_k$:

$$\begin{cases} \mathbf{a}_1 & = & ||\mathbf{q}_1||_2 \cdot \widehat{\mathbf{q}}_1 \\ \mathbf{a}_2 & = & (\widehat{\mathbf{q}}_1^T\mathbf{a}_2) \cdot \widehat{\mathbf{q}}_1 + ||\mathbf{q}_2||_2 \cdot \widehat{\mathbf{q}}_2 \\ \mathbf{a}_3 & = & (\widehat{\mathbf{q}}_1^T\mathbf{a}_3) \cdot \widehat{\mathbf{q}}_1 + (\widehat{\mathbf{q}}_2^T\mathbf{a}_3) \cdot \widehat{\mathbf{q}}_2 + ||\mathbf{q}_3||_2 \cdot \widehat{\mathbf{q}}_3 \end{cases} \equiv \begin{cases} \mathbf{a}_1 & = & r_{11}\widehat{\mathbf{q}}_1 \\ \mathbf{a}_2 & = & r_{12}\widehat{\mathbf{q}}_1 + r_{22}\widehat{\mathbf{q}}_2 \\ \mathbf{a}_3 & = & r_{13}\widehat{\mathbf{q}}_1 + r_{23}\widehat{\mathbf{q}}_2 + r_{33}\widehat{\mathbf{q}}_3 \end{cases}$$

3. Finally, observe that the above system is a linear-combination-of-columns view of a matrix-matrix product:

$$\begin{bmatrix} | & | & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \\ | & | & | \end{bmatrix} = \begin{bmatrix} | & | & | \\ \widehat{\mathbf{q}}_1 & \widehat{\mathbf{q}}_2 & \widehat{\mathbf{q}}_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{bmatrix} \iff A = \hat{Q}\hat{R}$$

Notice that 4×3 matrix $\hat{Q}$ has orthonormal columns.
Notice that 3×3 matrix $\hat{R}$ is upper triangular.

This is known as a **reduced QR factorization** of matrix A.

This factorization applies to tall $(m > n)$ or square $(m = n)$ $m \times n$ matrices with full column rank.

Reduced QR Factorization via CGS-EN

Proposition

(Reduced QR Factorization via CGS-EN)

GIVEN: Tall or square $(m \geq n)$ full column rank matrix $A_{m\times n} := \begin{bmatrix} | & | & & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \\ | & | & & | \end{bmatrix}$.

TASK: Factor $A = \hat{Q}\hat{R}$ where $\hat{Q}_{m\times n}$ has orthonormal columns $\widehat{\mathbf{q}}_k$ and $\hat{R}_{n\times n}$ is upper triangular.

(1) Perform Classical Gram-Schmidt w/ early normalization on the columns of A:
$\left(\widehat{\mathcal{Q}}_k := \{\widehat{\mathbf{q}}_1, \widehat{\mathbf{q}}_2, \ldots, \widehat{\mathbf{q}}_k\};\ k = 1, \cdots, n-1\right)$

$$\begin{aligned}
\mathbf{q}_1 &:= \mathbf{a}_1 && = \mathbf{a}_1 && ;\ \widehat{\mathbf{q}}_1 := \mathbf{q}_1 / \underbrace{\|\mathbf{q}_1\|_2}_{r_{11}} \\
\mathbf{q}_2 &:= \mathbf{a}_2 - \mathit{proj}_{span(\widehat{\mathcal{Q}}_1)}\mathbf{a}_2 && = \mathbf{a}_2 - \underbrace{(\widehat{\mathbf{q}}_1^T\mathbf{a}_2)}_{r_{12}}\widehat{\mathbf{q}}_1 && ;\ \widehat{\mathbf{q}}_2 := \mathbf{q}_2 / \underbrace{\|\mathbf{q}_2\|_2}_{r_{22}} \\
\mathbf{q}_3 &:= \mathbf{a}_3 - \mathit{proj}_{span(\widehat{\mathcal{Q}}_2)}\mathbf{a}_3 && = \mathbf{a}_3 - \underbrace{(\widehat{\mathbf{q}}_1^T\mathbf{a}_3)}_{r_{13}}\widehat{\mathbf{q}}_1 - \underbrace{(\widehat{\mathbf{q}}_2^T\mathbf{a}_3)}_{r_{23}}\widehat{\mathbf{q}}_2 && ;\ \widehat{\mathbf{q}}_3 := \mathbf{q}_3 / \underbrace{\|\mathbf{q}_3\|_2}_{r_{33}} \\
&\vdots && \quad \ddots && \quad \vdots \\
\mathbf{q}_n &:= \mathbf{a}_n - \mathit{proj}_{span(\widehat{\mathcal{Q}}_{n-1})}\mathbf{a}_n && = \mathbf{a}_n - \underbrace{(\widehat{\mathbf{q}}_1^T\mathbf{a}_n)}_{r_{1n}}\widehat{\mathbf{q}}_1 - \cdots - \underbrace{(\widehat{\mathbf{q}}_{n-1}^T\mathbf{a}_n)}_{r_{n-1,n}}\widehat{\mathbf{q}}_{n-1} && ;\ \widehat{\mathbf{q}}_n := \mathbf{q}_n / \underbrace{\|\mathbf{q}_n\|_2}_{r_{nn}}
\end{aligned}$$

Reduced QR Factorization via CGS-EN

Proposition

(Reduced QR Factorization via CGS-EN)

GIVEN: *Tall or square* $(m \geq n)$ *full column rank matrix* $A_{m\times n} := \begin{bmatrix} | & | & & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \\ | & | & & | \end{bmatrix}$.

TASK: *Factor* $A = \hat{Q}\hat{R}$ *where* $\hat{Q}_{m\times n}$ *has orthonormal columns* $\widehat{\mathbf{q}}_k$ *and* $\hat{R}_{n\times n}$ *is upper triangular.*

(2) Use the $\widehat{\mathbf{q}}_k$ *vectors to build* $\hat{Q}$ *matrix and* r_{ij} *entries to build* $\hat{R}$ *matrix:*

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

Reduced QR Factorization via CGS-EN (Properties)

Proposition

(Properties of Reduced QR Factorization)

<u>*GIVEN:*</u> *Tall or square $(m \geq n)$ full column rank $A_{m\times n}$. Let $A = \hat{Q}\hat{R}$ as Reduced QR via CGS-EN:*

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

Then:

$$\begin{array}{ll} (a) & \hat{Q}^T\hat{Q} = I_{n\times n} \\ (b) & r_{kk} \neq 0 \;\; \forall k \\ (c) & \hat{R} \text{ is invertible} \end{array}$$

<u>PROOF:</u>

$$(a)\ \hat{Q}^T\hat{Q} = \begin{bmatrix} \text{---} & \widehat{\mathbf{q}}_1^T & \text{---} \\ \text{---} & \widehat{\mathbf{q}}_2^T & \text{---} \\ & \vdots & \\ \text{---} & \widehat{\mathbf{q}}_n^T & \text{---} \end{bmatrix} \begin{bmatrix} | & | & & | \\ \widehat{\mathbf{q}}_1 & \widehat{\mathbf{q}}_2 & \cdots & \widehat{\mathbf{q}}_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} \widehat{\mathbf{q}}_1^T\widehat{\mathbf{q}}_1 & \widehat{\mathbf{q}}_1^T\widehat{\mathbf{q}}_2 & \cdots & \widehat{\mathbf{q}}_1^T\widehat{\mathbf{q}}_n \\ \widehat{\mathbf{q}}_2^T\widehat{\mathbf{q}}_1 & \widehat{\mathbf{q}}_2^T\widehat{\mathbf{q}}_2 & \cdots & \widehat{\mathbf{q}}_2^T\widehat{\mathbf{q}}_n \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{\mathbf{q}}_n^T\widehat{\mathbf{q}}_1 & \widehat{\mathbf{q}}_n^T\widehat{\mathbf{q}}_2 & \cdots & \widehat{\mathbf{q}}_n^T\widehat{\mathbf{q}}_n \end{bmatrix}$$

Reduced QR Factorization via CGS-EN (Properties)

Proposition

(Properties of Reduced QR Factorization)

<u>*GIVEN:*</u> *Tall or square* $(m \geq n)$ *full column rank* $A_{m\times n}$. *Let* $A = \hat{Q}\hat{R}$ *as Reduced* QR *via CGS-EN:*

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

Then:

$$\begin{array}{ll} (a) & \hat{Q}^T\hat{Q} = I_{n\times n} \\ (b) & r_{kk} \neq 0 \ \ \forall k \\ (c) & \hat{R} \text{ is invertible} \end{array}$$

<u>PROOF:</u>

$$(a)\ \hat{Q}^T\hat{Q} = \begin{bmatrix} \widehat{\mathbf{q}}_1^T\widehat{\mathbf{q}}_1 & \widehat{\mathbf{q}}_1^T\widehat{\mathbf{q}}_2 & \cdots & \widehat{\mathbf{q}}_1^T\widehat{\mathbf{q}}_n \\ \widehat{\mathbf{q}}_2^T\widehat{\mathbf{q}}_1 & \widehat{\mathbf{q}}_2^T\widehat{\mathbf{q}}_2 & \cdots & \widehat{\mathbf{q}}_2^T\widehat{\mathbf{q}}_n \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{\mathbf{q}}_n^T\widehat{\mathbf{q}}_1 & \widehat{\mathbf{q}}_n^T\widehat{\mathbf{q}}_2 & \cdots & \widehat{\mathbf{q}}_n^T\widehat{\mathbf{q}}_n \end{bmatrix} \overset{(*)}{=} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = I_{n\times n}$$

$(*)$ $\hat{Q}$ was formed via Reduced QR $\implies$ $\hat{Q}$ has orthonormal columns $\implies$ $\widehat{\mathbf{q}}_i^T\widehat{\mathbf{q}}_j = \delta_{ij}$

Reduced QR Factorization via CGS-EN (Properties)

Proposition

(Properties of Reduced QR Factorization)

GIVEN: Tall or square $(m \geq n)$ full column rank $A_{m\times n}$. Let $A = \hat{Q}\hat{R}$ as Reduced QR via CGS-EN:

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

Then:

$$\begin{array}{ll} (a) & \hat{Q}^T\hat{Q} = I_{n\times n} \\ (b) & r_{kk} \neq 0 \ \ \forall k \\ (c) & \hat{R} \text{ is invertible} \end{array}$$

PROOF:

(b) A has full column rank $\Longrightarrow$ the columns of A, $\{\mathbf{a}_1, \cdots, \mathbf{a}_n\}$, are a basis for ColSp$(A)$

$\Longrightarrow$ the columns of A, $\{\mathbf{a}_1, \cdots, \mathbf{a}_n\}$, are linearly independent

$\Longrightarrow$ per CGS-EN, each $\text{proj}_{span(\widehat{\mathcal{Q}}_{k-1})}\mathbf{a}_k \neq \mathbf{a}_k$

$\Longrightarrow$ per CGS-EN, each $\mathbf{q}_k := \mathbf{a}_k - \text{proj}_{span(\widehat{\mathcal{Q}}_{k-1})}\mathbf{a}_k \neq \vec{\mathbf{0}}$

$\Longrightarrow$ per CGS-EN, each $r_{kk} := ||\mathbf{q}_k||_2 \neq 0$ □

Reduced QR Factorization via CGS-EN (Properties)

Proposition

(Properties of Reduced QR Factorization)

<u>*GIVEN:*</u> *Tall or square $(m \geq n)$ full column rank $A_{m\times n}$. Let $A = \hat{Q}\hat{R}$ as Reduced QR via CGS-EN:*

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

Then:

$$\begin{array}{ll} (a) & \hat{Q}^T\hat{Q} = I_{n\times n} \\ (b) & r_{kk} \neq 0 \ \ \forall k \\ (c) & \hat{R} \text{ is invertible} \end{array}$$

<u>PROOF:</u>

(c) From part (b), $r_{kk} \neq 0 \ \ \forall k$

$$\begin{aligned} &\implies \det(\hat{R}) \overset{(*)}{=} \textstyle\prod_{k=1}^{n} r_{kk} = r_{11}r_{22}\cdots r_{nn} \neq 0 \\ &\implies \det(\hat{R}) \neq 0 \\ &\implies \hat{R} \text{ is invertible} \quad \square \end{aligned}$$

$(*)$ The determinant of an upper triangular square matrix is the product of its main diagonal entries.

Fin

Fin.