QR factorization by Gram-Schmidt orthogonalization process

Rafikul Alam
Department of Mathematics
Indian Institute of Technology Guwahati
Guwahati - 781039, INDIA

Outline

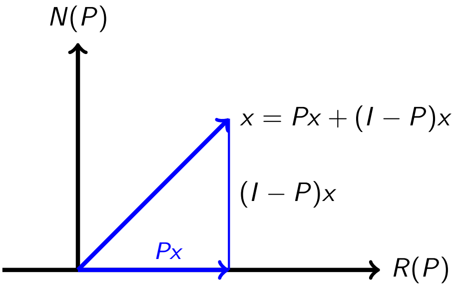
- Classical Gram-Schmidt orthogonalization scheme (CGS)
- Modified Gram-Schmidt ortrhogonalization scheme (MGS)
- QR factorization by Gram-Schmidt orthogonalization

Orthogonal decomposition

If $P \in \mathbb{C}^{n\times n}$ is an orthogonal projection then $\mathbb{C}^n = R(P) \oplus N(P)$ and $R(P) \perp N(P)$.

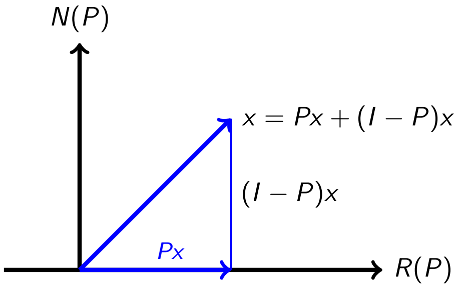
Orthogonal decomposition

If $P \in \mathbb{C}^{n\times n}$ is an orthogonal projection then $\mathbb{C}^n = R(P) \oplus N(P)$ and $R(P) \perp N(P)$.



Orthogonal decomposition

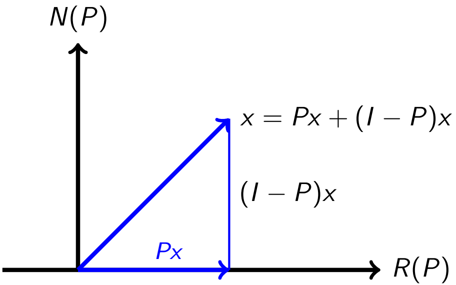
If $P \in \mathbb{C}^{n\times n}$ is an orthogonal projection then $\mathbb{C}^n = R(P) \oplus N(P)$ and $R(P) \perp N(P)$.



Thus, if $P := uu^*$ with $u^*u = 1$ then $Px = uu^*x = \langle x, u\rangle u$ and $(I - P)x \perp u$.

Orthogonal decomposition

If $P \in \mathbb{C}^{n\times n}$ is an orthogonal projection then $\mathbb{C}^n = R(P) \oplus N(P)$ and $R(P) \perp N(P)$.



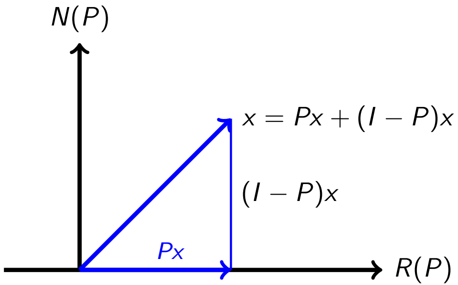
As (I-P)x belongs to N(P) and u belongs to R(P), and they are perpendicular,

Thus, if $P := uu^*$ with $u^*u = 1$ then $Px = uu^*x = \langle x, u\rangle u$ and $(I - P)x \perp u$.

More generally, if $U := \begin{bmatrix} u_1 & \cdots & u_m \end{bmatrix}$ is an isometry then $P := UU^*$ is an orthogonal projection

Orthogonal decomposition

If $P \in \mathbb{C}^{n\times n}$ is an orthogonal projection then $\mathbb{C}^n = R(P) \oplus N(P)$ and $R(P) \perp N(P)$.



Thus, if $P := uu^*$ with $u^*u = 1$ then $Px = uu^*x = \langle x, u\rangle u$ and $(I-P)x \perp u$.

More generally, if $U := \begin{bmatrix} u_1 & \cdots & u_m \end{bmatrix}$ is an isometry then $P := UU^*$ is an orthogonal projection such that $R(P) = R(U)$ and $(I-P)x \perp \{u_1, \ldots, u_m\}$ for any $x \in \mathbb{C}^n$.

Classical Gram-Schmidt scheme (CGS)

Let $x_1, \ldots, x_m$ be linearly independent vectors in $\mathbb{C}^n$.

Classical Gram-Schmidt scheme (CGS)

Let $x_1, \ldots, x_m$ be linearly independent vectors in $\mathbb{C}^n$.

Define $u_1 := x_1/\|x_1\|_2$. Then $\{u_1\}$ is an ONS (orthonormal set) and $(I - u_1 u_1^*)x_2 \perp u_1$.

Classical Gram-Schmidt scheme (CGS)

Let $x_1, \ldots, x_m$ be linearly independent vectors in $\mathbb{C}^n$.

Define $u_1 := x_1/\|x_1\|_2$. Then $\{u_1\}$ is an ONS (orthonormal set) and $(I - u_1 u_1^*)x_2 \perp u_1$.

Define $u_2 := \dfrac{(I - u_1 u_1^*)x_2}{\|(I - u_1 u_1^*)x_2\|_2}$. Then $\{u_1, u_2\}$ is an ONS, $\mathrm{span}(u_1, u_2) = \mathrm{span}(x_1, x_2)$ and

Classical Gram-Schmidt scheme (CGS)

Let $x_1, \ldots, x_m$ be linearly independent vectors in $\mathbb{C}^n$.

Define $u_1 := x_1/\|x_1\|_2$. Then $\{u_1\}$ is an ONS (orthonormal set) and $(I - u_1 u_1^*)x_2 \perp u_1$.

Define $u_2 := \dfrac{(I - u_1 u_1^*)x_2}{\|(I - u_1 u_1^*)x_2\|_2}$. Then $\{u_1, u_2\}$ is an ONS, $\mathrm{span}(u_1, u_2) = \mathrm{span}(x_1, x_2)$ and $(I - \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} u_1 & u_2 \end{bmatrix}^*)x_3 \perp \{u_1, u_2\}$.

Classical Gram-Schmidt scheme (CGS)

Let $x_1, \ldots, x_m$ be linearly independent vectors in $\mathbb{C}^n$.

Define $u_1 := x_1/\|x_1\|_2$. Then $\{u_1\}$ is an ONS (orthonormal set) and $(I - u_1 u_1^*)x_2 \perp u_1$.

Define $u_2 := \dfrac{(I - u_1 u_1^*)x_2}{\|(I - u_1 u_1^*)x_2\|_2}$. Then $\{u_1, u_2\}$ is an ONS, $\mathrm{span}(u_1, u_2) = \mathrm{span}(x_1, x_2)$ and $(I - \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} u_1 & u_2 \end{bmatrix}^*)x_3 \perp \{u_1, u_2\}$.

Define $u_3 := \dfrac{(I - \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} u_1 & u_2 \end{bmatrix}^*)x_3}{\|(I - \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} u_1 & u_2 \end{bmatrix}^*)x_3\|_2}$. Then $\{u_1, u_2, u_3\}$ is an ONS and $\mathrm{span}(u_1, u_2, u_3) = \mathrm{span}(x_1, x_2, x_3)$.

Classical Gram-Schmidt scheme (CGS)

Let $x_1, \ldots, x_m$ be linearly independent vectors in $\mathbb{C}^n$.

Define $u_1 := x_1/\|x_1\|_2$. Then $\{u_1\}$ is an ONS (orthonormal set) and $(I - u_1 u_1^*)x_2 \perp u_1$.

Define $u_2 := \dfrac{(I - u_1 u_1^*)x_2}{\|(I - u_1 u_1^*)x_2\|_2}$. Then $\{u_1, u_2\}$ is an ONS, $\mathrm{span}(u_1, u_2) = \mathrm{span}(x_1, x_2)$ and $(I - \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} u_1 & u_2 \end{bmatrix}^*)x_3 \perp \{u_1, u_2\}$.

Define $u_3 := \dfrac{(I - \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} u_1 & u_2 \end{bmatrix}^*)x_3}{\|(I - \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} u_1 & u_2 \end{bmatrix}^*)x_3\|_2}$. Then $\{u_1, u_2, u_3\}$ is an ONS and $\mathrm{span}(u_1, u_2, u_3) = \mathrm{span}(x_1, x_2, x_3)$.

This yields the classical Gram-Schmidt orthonormal process: Define

$$
\begin{aligned}
u_1 &:= \frac{x_1}{\|x_1\|_2}, \\
u_j &:= \frac{x_j - \begin{bmatrix} u_1 & \cdots & u_{j-1} \end{bmatrix} \begin{bmatrix} u_1 & \cdots & u_{j-1} \end{bmatrix}^* x_j}{\|x_j - \begin{bmatrix} u_1 & \cdots & u_{j-1} \end{bmatrix} \begin{bmatrix} u_1 & \cdots & u_{j-1} \end{bmatrix}^* x_j\|_2}, \quad j = 2 : m.
\end{aligned}
$$

Classical Gram-Schmidt scheme (CGS)

Let $x_1, \ldots, x_m$ be linearly independent vectors in $\mathbb{C}^n$.

Define $u_1 := x_1/\|x_1\|_2$. Then $\{u_1\}$ is an ONS (orthonormal set) and $(I - u_1u_1^*)x_2 \perp u_1$.

Define $u_2 := \dfrac{(I - u_1u_1^*)x_2}{\|(I - u_1u_1^*)x_2\|_2}$. Then $\{u_1, u_2\}$ is an ONS, $\underline{\mathrm{span}(u_1, u_2) = \mathrm{span}(x_1, x_2)}$ and $(I - \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} u_1 & u_2 \end{bmatrix}^*)x_3 \perp \{u_1, u_2\}$.

? HOW

Does this mean (I - [u1 u2] [u1 u2]*)x3 is orthogonal to both u1 and u2?, => YES

Define $u_3 := \dfrac{(I - \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} u_1 & u_2 \end{bmatrix}^*)x_3}{\|(I - \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} u_1 & u_2 \end{bmatrix}^*)x_3\|_2}$. Then $\{u_1, u_2, u_3\}$ is an ONS and $\mathrm{span}(u_1, u_2, u_3) = \mathrm{span}(x_1, x_2, x_3)$.

This yields the classical Gram-Schmidt orthonormal process: Define

$$
\begin{aligned}
u_1 &:= \frac{x_1}{\|x_1\|_2}, \\
u_j &:= \frac{x_j - \begin{bmatrix} u_1 & \cdots & u_{j-1} \end{bmatrix} \begin{bmatrix} u_1 & \cdots & u_{j-1} \end{bmatrix}^* x_j}{\|x_j - \begin{bmatrix} u_1 & \cdots & u_{j-1} \end{bmatrix} \begin{bmatrix} u_1 & \cdots & u_{j-1} \end{bmatrix}^* x_j\|_2}, \quad j = 2 : m.
\end{aligned}
$$

Then $\{u_1, \ldots, u_m\}$ is an ONS and $\mathrm{span}(u_1, \ldots, u_j) = \mathrm{span}(x_1, \ldots, x_j)$ for $j = 1 : m$.

Orthogonalization

Let $x_1, \ldots, x_m$ be linearly indepedent vectors in $\mathbb{C}^n$. Then there exist orthonormal vectors $v_1, \ldots, v_m$ in $\mathbb{C}^n$ such that

$$\mathrm{span}(x_1, \ldots, x_j) = \mathrm{span}(v_1, \ldots, v_j) \ \text{for} \ \ j = 1 : m. \tag{1}$$

Orthogonalization

Let $x_1, \ldots, x_m$ be linearly indepedent vectors in $\mathbb{C}^n$. Then there exist orthonormal vectors $v_1, \ldots, v_m$ in $\mathbb{C}^n$ such that

$$\mathrm{span}(x_1, \ldots, x_j) = \mathrm{span}(v_1, \ldots, v_j) \text{ for } j = 1 : m. \tag{1}$$

The condition in (1) is a QR factorization in disguise. This can be seen as follows.

Orthogonalization

Let $x_1, \ldots, x_m$ be linearly indepedent vectors in $\mathbb{C}^n$. Then there exist orthonormal vectors $v_1, \ldots, v_m$ in $\mathbb{C}^n$ such that

$$\mathrm{span}(x_1, \ldots, x_j) = \mathrm{span}(v_1, \ldots, v_j) \;\text{ for }\; j = 1 : m. \tag{1}$$

The condition in (1) is a QR factorization in disguise. This can be seen as follows. Define $X := \begin{bmatrix} x_1 & \cdots & x_m \end{bmatrix}$ and $Q := \begin{bmatrix} v_1 & \cdots & v_m \end{bmatrix}$.

Orthogonalization

Let $x_1, \ldots, x_m$ be linearly indepedent vectors in $\mathbb{C}^n$. Then there exist orthonormal vectors $v_1, \ldots, v_m$ in $\mathbb{C}^n$ such that

$$\mathrm{span}(x_1, \ldots, x_j) = \mathrm{span}(v_1, \ldots, v_j) \text{ for } j = 1 : m. \tag{1}$$

The condition in (1) is a QR factorization in disguise. This can be seen as follows. Define $X := \begin{bmatrix} x_1 & \cdots & x_m \end{bmatrix}$ and $Q := \begin{bmatrix} v_1 & \cdots & v_m \end{bmatrix}$. Then $\mathrm{rank}(X) = m$ and Q is an isometry.

Orthogonalization

Let $x_1, \ldots, x_m$ be linearly indepedent vectors in $\mathbb{C}^n$. Then there exist orthonormal vectors $v_1, \ldots, v_m$ in $\mathbb{C}^n$ such that

$$\mathrm{span}(x_1, \ldots, x_j) = \mathrm{span}(v_1, \ldots, v_j) \text{ for } j = 1 : m. \tag{1}$$

The condition in (1) is a QR factorization in disguise. This can be seen as follows. Define $X := \begin{bmatrix} x_1 & \cdots & x_m \end{bmatrix}$ and $Q := \begin{bmatrix} v_1 & \cdots & v_m \end{bmatrix}$. Then $\mathrm{rank}(X) = m$ and Q is an isometry.

Fact: The condition in (1) holds $\iff$ there exists a nonsingular upper triangular matrix $R \in \mathbb{C}^{m \times m}$ such that $X = QR$.

Orthogonalization

Let $x_1, \ldots, x_m$ be linearly indepedent vectors in $\mathbb{C}^n$. Then there exist orthonormal vectors $v_1, \ldots, v_m$ in $\mathbb{C}^n$ such that

$$\mathrm{span}(x_1, \ldots, x_j) = \mathrm{span}(v_1, \ldots, v_j) \text{ for } j = 1 : m. \tag{1}$$

The condition in (1) is a QR factorization in disguise. This can be seen as follows. Define $X := \begin{bmatrix} x_1 & \cdots & x_m \end{bmatrix}$ and $Q := \begin{bmatrix} v_1 & \cdots & v_m \end{bmatrix}$. Then $\mathrm{rank}(X) = m$ and Q is an isometry.

Fact: The condition in (1) holds $\iff$ there exists a nonsingular upper triangular matrix $R \in \mathbb{C}^{m\times m}$ such that $X = QR$.

Proof: If (1) holds then $x_j = v_1 r_{1j} + \cdots + v_j r_{jj}$ for some $r_{ij}, i = 1 : j$.

Orthogonalization

Let $x_1, \ldots, x_m$ be linearly indepedent vectors in $\mathbb{C}^n$. Then there exist orthonormal vectors $v_1, \ldots, v_m$ in $\mathbb{C}^n$ such that

$$\mathrm{span}(x_1, \ldots, x_j) = \mathrm{span}(v_1, \ldots, v_j) \text{ for } j = 1 : m. \tag{1}$$

The condition in (1) is a QR factorization in disguise. This can be seen as follows. Define $X := \begin{bmatrix} x_1 & \cdots & x_m \end{bmatrix}$ and $Q := \begin{bmatrix} v_1 & \cdots & v_m \end{bmatrix}$. Then $\mathrm{rank}(X) = m$ and Q is an isometry.

Fact: The condition in (1) holds $\iff$ there exists a nonsingular upper triangular matrix $R \in \mathbb{C}^{m \times m}$ such that $X = QR$.

Proof: If (1) holds then $x_j = v_1 r_{1j} + \cdots + v_j r_{jj}$ for some $r_{ij}, i = 1 : j$. Thus

$$x_j = \begin{bmatrix} v_1 & \cdots & v_j \end{bmatrix} \begin{bmatrix} r_{1j} \\ \vdots \\ r_{jj} \end{bmatrix} \Longrightarrow \begin{bmatrix} x_1 & \cdots & x_m \end{bmatrix} = \begin{bmatrix} v_1 & \cdots & v_m \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ & r_{22} & \cdots & r_{2m} \\ & & \ddots & \vdots \\ & & & r_{mm} \end{bmatrix}$$

Orthogonalization

Let $x_1, \ldots, x_m$ be linearly indepedent vectors in $\mathbb{C}^n$. Then there exist orthonormal vectors $v_1, \ldots, v_m$ in $\mathbb{C}^n$ such that

$$\mathrm{span}(x_1, \ldots, x_j) = \mathrm{span}(v_1, \ldots, v_j) \text{ for } j = 1 : m. \tag{1}$$

The condition in (1) is a QR factorization in disguise. This can be seen as follows. Define $X := \begin{bmatrix} x_1 & \cdots & x_m \end{bmatrix}$ and $Q := \begin{bmatrix} v_1 & \cdots & v_m \end{bmatrix}$. Then $\mathrm{rank}(X) = m$ and Q is an isometry.

Fact: The condition in (1) holds $\iff$ there exists a nonsingular upper triangular matrix $R \in \mathbb{C}^{m \times m}$ such that $X = QR$.

Proof: If (1) holds then $x_j = v_1 r_{1j} + \cdots + v_j r_{jj}$ for some $r_{ij}, i = 1 : j$. Thus

$$x_j = \begin{bmatrix} v_1 & \cdots & v_j \end{bmatrix} \begin{bmatrix} r_{1j} \\ \vdots \\ r_{jj} \end{bmatrix} \Longrightarrow \begin{bmatrix} x_1 & \cdots & x_m \end{bmatrix} = \begin{bmatrix} v_1 & \cdots & v_m \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ & r_{22} & \cdots & r_{2m} \\ & & \ddots & \vdots \\ & & & r_{mm} \end{bmatrix}$$ which

gives $X = QR$. Obviously R is nonsingular (Why?).

Orthogonalization

Let $x_1, \ldots, x_m$ be linearly indepedent vectors in $\mathbb{C}^n$. Then there exist orthonormal vectors $v_1, \ldots, v_m$ in $\mathbb{C}^n$ such that

$$\mathrm{span}(x_1, \ldots, x_j) = \mathrm{span}(v_1, \ldots, v_j) \text{ for } j = 1 : m. \tag{1}$$

The condition in (1) is a QR factorization in disguise. This can be seen as follows. Define $X := \begin{bmatrix} x_1 & \cdots & x_m \end{bmatrix}$ and $Q := \begin{bmatrix} v_1 & \cdots & v_m \end{bmatrix}$. Then $\mathrm{rank}(X) = m$ and Q is an isometry.

Fact: The condition in (1) holds $\iff$ there exists a nonsingular upper triangular matrix $R \in \mathbb{C}^{m\times m}$ such that $X = QR$.

Proof: If (1) holds then $x_j = v_1 r_{1j} + \cdots + v_j r_{jj}$ for some $r_{ij}, i = 1 : j$. Thus

$$x_j = \begin{bmatrix} v_1 & \cdots & v_j \end{bmatrix} \begin{bmatrix} r_{1j} \\ \vdots \\ r_{jj} \end{bmatrix} \Longrightarrow \begin{bmatrix} x_1 & \cdots & x_m \end{bmatrix} = \begin{bmatrix} v_1 & \cdots & v_m \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ & r_{22} & \cdots & r_{2m} \\ & & \ddots & \vdots \\ & & & r_{mm} \end{bmatrix}$$ which

gives $X = QR$. Obviously R is nonsingular (Why?). The converse part is immediate as $Q = XR^{-1}$. ■

Assume it is singular and contradict.

Classical Gram-Schmidt scheme (CGS)

Equating the columns we have

$$
\begin{aligned}
x_1 &= v_1 r_{11} \Rightarrow r_{11} = \|x_1\|_2 \text{ and } v_1 := x_1/r_{11} \\
x_2 &= v_1 r_{12} + v_2 r_{22} \Rightarrow r_{12} = \langle x_2, v_1 \rangle, r_{22} := \|x_2 - \langle x_2, v_1 \rangle v_1\|_2
\end{aligned}
$$

and $v_2 := (x_2 - \langle x_2, v_1 \rangle v_1)/r_{22}$, where $\langle x, y \rangle = y^* x$.

Classical Gram-Schmidt scheme (CGS)

Equating the columns we have

$$\begin{aligned} x_1 &= v_1 r_{11} \Rightarrow r_{11} = \|x_1\|_2 \text{ and } v_1 := x_1/r_{11} \\ x_2 &= v_1 r_{12} + v_2 r_{22} \Rightarrow r_{12} = \langle x_2,\, v_1\rangle, r_{22} := \|x_2 - \langle x_2,\, v_1\rangle v_1\|_2 \end{aligned}$$

and $v_2 := (x_2 - \langle x_2,\, v_1\rangle v_1)/r_{22}$, where $\langle x,\, y\rangle = y^* x$.

The classical Gram-Schmidt method constructs orthonormal vectors $v_1, \ldots, v_n$ as follows. Define

$$\begin{aligned} v_1 &:= \frac{x_1}{\|x_1\|_2}, \\ v_j &:= \frac{x_j - v_1 v_1^* x_j - \cdots - v_{j-1} v_{j-1}^* x_j}{\|x_j - v_1 v_1^* x_j - \cdots - v_{j-1} v_{j-1}^* x_j\|_2},\ j = 2 : n. \end{aligned}$$

Classical Gram-Schmidt scheme (CGS)

Equating the columns we have

$$\begin{aligned} x_1 &= v_1 r_{11} \Rightarrow r_{11} = \|x_1\|_2 \text{ and } v_1 := x_1/r_{11} \\ x_2 &= v_1 r_{12} + v_2 r_{22} \Rightarrow r_{12} = \langle x_2,\, v_1\rangle,\, r_{22} := \|x_2 - \langle x_2,\, v_1\rangle v_1\|_2 \end{aligned}$$

and $v_2 := (x_2 - \langle x_2,\, v_1\rangle v_1)/r_{22}$, where $\langle x,\, y\rangle = y^* x$.

The classical Gram-Schmidt method constructs orthonormal vectors $v_1, \ldots, v_n$ as follows. Define

$$\begin{aligned} v_1 &:= \frac{x_1}{\|x_1\|_2}, \\ v_j &:= \frac{x_j - v_1 v_1^* x_j - \cdots - v_{j-1} v_{j-1}^* x_j}{\|x_j - v_1 v_1^* x_j - \cdots - v_{j-1} v_{j-1}^* x_j\|_2},\; j = 2 : n. \end{aligned}$$

Then $v_1, \ldots, v_n$ are orthonormal, that is, $v_i^* v_j = 0$ for $i \neq j$ and $\|v_j\|_2 = 1$ for $j = 1 : n$. Note that $\|x_j - v_1 v_1^* x_j - \cdots - v_{j-1} v_{j-1}^* x_j\|_2 \neq 0$ as the vectors $x_1, \ldots, x_j$ are linearly independent.

Proof?

Classical Gram-Schmidt scheme (CGS)

Setting $r_{11} := \|x_1\|_2$ and $r_{jj} := \|x_j - v_1 v_1^* x_j - \cdots - v_{j-1} v_{j-1}^* x_j\|_2$, we have

$$\begin{aligned} x_1 &= v_1 r_{11}, \\ x_j &= v_1 v_1^* x_j + \cdots + v_{j-1} v_{j-1}^* x_j + v_j r_{jj}, \quad j = 2 : n. \end{aligned}$$

Setting $r_{kj} := v_k^* x_j$, in matrix notation, we have

Classical Gram-Schmidt scheme (CGS)

Setting $r_{11} := \|x_1\|_2$ and $r_{jj} := \|x_j - v_1 v_1^* x_j - \cdots - v_{j-1} v_{j-1}^* x_j\|_2$, we have

$$\begin{aligned} x_1 &= v_1 r_{11}, \\ x_j &= v_1 v_1^* x_j + \cdots + v_{j-1} v_{j-1}^* x_j + v_j r_{jj}, \quad j = 2 : n. \end{aligned}$$

Setting $r_{kj} := v_k^* x_j$, in matrix notation, we have

$$\begin{aligned} \begin{bmatrix} x_1 & \cdots & x_j \end{bmatrix} &= \begin{bmatrix} v_1 & \cdots & v_j \end{bmatrix} \begin{bmatrix} r_{11} & v_1^* x_2 & \cdots & v_1^* x_j \\ & r_{22} & \cdots & v_2^* x_j \\ & & \ddots & \vdots \\ & & & r_{jj} \end{bmatrix} \\ &= \begin{bmatrix} v_1 & \cdots & v_j \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1j} \\ & r_{22} & \cdots & r_{2j} \\ & & \ddots & \vdots \\ & & & r_{jj} \end{bmatrix} \quad \text{for } \; j = 1 : n. \end{aligned}$$

Classical Gram-Schmidt scheme (CGS)

The CGS computes QR factorization $X = QR$ in which the j-th columns of Q and R are determined at the j-th step.

Classical Gram-Schmidt scheme (CGS)

The CGS computes QR factorization $X = QR$ in which the j-th columns of Q and R are determined at the j-th step.

Algorithm: CGS (Classical Gram-Schmidt algorithm)
Input: Linearly indendent vectors $x_1, \ldots, x_n$
Output: Orthonormal vectors $v_1, \ldots, v_n$

---

for $i = 1\colon n$
  $v_i := x_i$
  for $j = 1\colon i-1$
    $r_{ji} := v_j^* x_i$
    $v_i := v_i - v_j r_{ji}$
  end
$r_{ii} := \|v_i\|_2$
if $r_{ii} = 0$ then quit else
$v_i := v_i / r_{ii}$
end

Example

Consider $x_1 := \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}^\top, x_2 := \begin{bmatrix} 2 & 1 & 0 \end{bmatrix}^\top$ and $\begin{bmatrix} 0 & 1 & 1 \end{bmatrix}^\top$. Then by the Gram-Schmidt process, we have $r_{11} := \|x_1\|_2 = \sqrt{2}$ which gives

$$v_1 := \frac{x_1}{r_{11}} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Example

Consider $x_1 := \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}^\top, x_2 := \begin{bmatrix} 2 & 1 & 0 \end{bmatrix}^\top$ and $\begin{bmatrix} 0 & 1 & 1 \end{bmatrix}^\top$. Then by the Gram-Schmidt process, we have $r_{11} := \|x_1\|_2 = \sqrt{2}$ which gives

$$v_1 := \frac{x_1}{r_{11}} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Next, we have $r_{12} := v_1^\top x_2 = \sqrt{2}$ and

$$\begin{aligned} q_2 &:= x_2 - v_1 v_1^\top x_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} - \frac{\sqrt{2}}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \\ v_2 &:= \frac{q_2}{r_{22}} = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \text{ where } r_{22} := \|q_2\|_2 = \sqrt{3}. \end{aligned}$$

Example

Finally, $r_{13} := v_1^\top x_3 = 1/\sqrt{2}$ and $r_{23} := v_2^\top x_3 = 0$. Hence we have

$$\begin{aligned} q_3 &:= x_3 - v_1 v_1^\top x_3 - v_2 v_2^\top x_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \\ v_3 &:= \frac{q_3}{r_{33}} = \frac{1}{\sqrt{6}}\begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \text{where } r_{33} := \|q_3\|_3 = \frac{\sqrt{6}}{2}. \end{aligned}$$

Setting $A := \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}$ and $Q := \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix}$, we have the QR factorization of A

$$\underbrace{\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} \sqrt{2} & \sqrt{2} & \frac{1}{\sqrt{2}} \\ 0 & \sqrt{3} & 0 \\ 0 & 0 & \frac{\sqrt{6}}{2} \end{bmatrix}}_{R}. \blacksquare$$

Modified Gram-Schmidt Scheme (MGS)

The classical Gram-Schmidt algorithm is known to be numerically unstable. The computed vectors $v_1, \ldots, v_n$ may not be orthogonal.

Modified Gram-Schmidt Scheme (MGS)

The classical Gram-Schmidt algorithm is known to be numerically unstable. The computed vectors $v_1, \ldots, v_n$ may not be orthogonal.

Note that

$$I - v_1 v_1^* - \cdots - v_{j-1} v_{j-1}^* = (I - v_1 v_1^*)(I - v_2 v_2^*) \cdots (I - v_{j-1} v_{j-1}^*).$$

Hence the CGS can be rewritten as

$$\begin{aligned} v_1 &:= \frac{x_1}{\|x_1\|_2}, \\ v_j &:= \frac{(I - v_{j-1} v_{j-1}^*) \cdots (I - v_1 v_1^*) x_j}{\|(I - v_{j-1} v_{j-1}^*) \cdots (I - v_1 v_1^*) x_j\|_2}, \; j = 2 : n. \end{aligned}$$

The modified Gram-Schmidt scheme is obtained by computing v_j incrementally.

Modified Gram-Schmidt Scheme (MGS)

The modified Gram-Schmidt scheme is obtained by computing v_j incrementally as follows

$$
\begin{aligned}
v_1 &:= \frac{x_1}{\|x_1\|_2}, \quad q_k^{(1)} = (I - v_1 v_1^*)x_k, \;\; k = 2:n \\
v_j &:= \frac{q_j^{(j-1)}}{\|q_j^{(j-1)}\|_2}, \; q_k^{(j)} := (I - v_j v_j^*)q_k^{(j-1)}, \; k = j+1:n
\end{aligned}
$$

for $j = 2 : n$.

Modified Gram-Schmidt Scheme (MGS)

The modified Gram-Schmidt scheme is obtained by computing v_j incrementally as follows

$$
\begin{aligned}
v_1 &:= \frac{x_1}{\|x_1\|_2}, \quad q_k^{(1)} = (I - v_1 v_1^*)x_k, \quad k = 2 : n \\
v_j &:= \frac{q_j^{(j-1)}}{\|q_j^{(j-1)}\|_2}, \quad q_k^{(j)} := (I - v_j v_j^*)q_k^{(j-1)}, \quad k = j+1 : n
\end{aligned}
$$

for $j = 2 : n$.

Note that

$$
v_j = \frac{q_j^{(j-1)}}{\|q_j^{(j-1)}\|_2} = \frac{(I - v_{j-1}v_{j-1}^*)\cdots(I - v_1 v_1^*)x_j}{\|(I - v_{j-1}v_{j-1}^*)\cdots(I - v_1 v_1^*)x_j\|_2}, \quad j = 2 : n.
$$

CGS versus MGS

The tables below illustrate the essential difference between CGS and MGS steps.

Step	CGS Algorithm		
1.	$v_1 := \frac{x_1}{\|x_1\|_2}$	x_2	x_3
2.	v_1	$v_2 := \frac{(I - v_1 v_1^*)x_2}{\|(I - v_1 v_1^*)x_2\|_2}$	x_3
3.	v_1	v_2	$v_3 := \frac{(I - v_1 v_1^* - v_2 v_2^*)x_3}{\|(I - v_1 v_1^* - v_2 v_2^*)x_3\|_2}$

Step	MGS Algorithm		
1.	$v_1 := \frac{x_1}{\|x_1\|_2}$	$(I - v_1 v_1^*)x_2$	$(I - v_1 v_1^*)x_3$
2.	v_1	$v_2 := \frac{(I - v_1 v_1^*)x_2}{\|(I - v_1 v_1^*)x_2\|_2}$	$(I - v_2 v_2^*)(I - v_1 v_1^*)x_3$
3.	v_1	v_2	$v_3 := \frac{(I - v_2 v_2^*)(I - v_1 v_1^*)x_3}{\|(I - v_2 v_2^*)(I - v_1 v_1^*)x_3\|_2}$

Modified Gram-Schmidt Scheme (MGS)

Algorithm: MGS (Modified Gram-Schmidt algorithm)
Input: Linearly indendent vectors $x_1, \ldots, x_n$
Output: Orthonormal vectors $v_1, \ldots, v_n$

---

for $j = 1 \colon n$
  $q_j \colon = x_j$
end
for $i = 1 \colon n$
  $r_{ii} \colon = \|q_i\|_2$
  $v_i \colon = \dfrac{q_i}{r_{ii}}$ (if $r_{ii} \neq 0$ else quit)
  for $j = i + 1 \colon n$
    $r_{ij} \colon = v_i^* q_j$
    $q_j \colon = q_j - r_{ij} v_i$
  end
end

Accuracy

Although CGS and MGS are mathematically the same scheme, however, MGS provides far better results than CGS in finite precision arithmetic.

Accuracy

Although CGS and MGS are mathematically the same scheme, however, MGS provides far better results than CGS in finite precision arithmetic.

If $A = QR$ is computed in finite precision with unit round off $\mathbf{u}$ then $\|\mathtt{Q}' * \mathtt{Q} - \mathtt{eye}(\mathtt{n},\mathtt{n})\|_2$ is a measure of deviation from orthonormality.

Accuracy

Although CGS and MGS are mathematically the same scheme, however, MGS provides far better results than CGS in finite precision arithmetic.

If $A = QR$ is computed in finite precision with unit round off $\mathbf{u}$ then $\|\mathtt{Q}' * \mathtt{Q} - \mathtt{eye}(\mathtt{n}, \mathtt{n})\|_2$ is a measure of deviation from orthonormality.

It can be proved that the isometry Q produced by the MGS algorithm satisfies

$$\|\mathtt{Q}' * \mathtt{Q} - \mathtt{eye}(\mathtt{n}, \mathtt{n})\|_2 \approx \mathbf{u} * \mathtt{cond}(\mathtt{A}),$$

where `cond(A)` is the 2-norm condition number of A.

Accuracy

Although CGS and MGS are mathematically the same scheme, however, MGS provides far better results than CGS in finite precision arithmetic.

If $A = QR$ is computed in finite precision with unit round off $\mathbf{u}$ then $\|\mathtt{Q}' * \mathtt{Q} - \mathtt{eye(n,n)}\|_2$ is a measure of deviation from orthonormality.

It can be proved that the isometry Q produced by the MGS algorithm satisfies

$$\|\mathtt{Q}' * \mathtt{Q} - \mathtt{eye(n,n)}\|_2 \approx \mathbf{u} * \mathtt{cond(A)},$$

where `cond(A)` is the 2-norm condition number of A.

This indicates that although MGS is usually better than CGS in the presence of rounding error, it is unlikely to perform well when $\mathtt{cond(A)}$ is large, that is, when the columns of A are nearly linearly dependent.

Accuracy

Although CGS and MGS are mathematically the same scheme, however, MGS provides far better results than CGS in finite precision arithmetic.

If $A = QR$ is computed in finite precision with unit round off $\mathbf{u}$ then $\|\mathtt{Q}' * \mathtt{Q} - \mathtt{eye}(\mathtt{n},\mathtt{n})\|_2$ is a measure of deviation from orthonormality.

It can be proved that the isometry Q produced by the MGS algorithm satisfies

$$\|\mathtt{Q}' * \mathtt{Q} - \mathtt{eye}(\mathtt{n},\mathtt{n})\|_2 \approx \mathbf{u} * \mathtt{cond}(\mathtt{A}),$$

where `cond(A)` is the 2-norm condition number of A.

This indicates that although MGS is usually better than CGS in the presence of rounding error, it is unlikely to perform well when $\mathtt{cond}(\mathtt{A})$ is large, that is, when the columns of A are nearly linearly dependent.

However if $\mathtt{cond}(\mathtt{A})$ satisfies $1 \ll \mathtt{cond}(\mathtt{A}) \ll 1/\mathbf{u}$, then $\mathbf{u} * \mathtt{cond}(\mathtt{V}) \ll 1$ and in such a case MGS will return a Q which is quite close to being an isometry.

Example

Consider the matrix $A := \begin{bmatrix} 1 & 1 & 1 \\ \epsilon & 0 & 0 \\ 0 & \epsilon & 0 \\ 0 & 0 & \epsilon \end{bmatrix}$. Then $\mathrm{cond}(A) = \dfrac{\sqrt{3+\epsilon^2}}{\epsilon}$.

If $\epsilon = \sqrt{\mathbf{u}}$ then $\mathbf{u}\,\mathrm{cond}(A) \approx \sqrt{3}\sqrt{\mathbf{u}}$.

Example

Consider the matrix $A := \begin{bmatrix} 1 & 1 & 1 \\ \epsilon & 0 & 0 \\ 0 & \epsilon & 0 \\ 0 & 0 & \epsilon \end{bmatrix}$. Then $\mathrm{cond}(A) = \dfrac{\sqrt{3+\epsilon^2}}{\epsilon}$.

If $\epsilon = \sqrt{\mathbf{u}}$ then $\mathbf{u}\,\mathrm{cond}(A) \approx \sqrt{3}\sqrt{\mathbf{u}}$.

Assume that $\epsilon^2 < \mathbf{u}$. Then, in finite precision arithmetic, Q_{CS} produced by CGS and Q_{MS} produced by MGS are given by

$$Q_{\mathrm{CS}} = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \epsilon & \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \text{ and } Q_{\mathrm{MS}} = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \epsilon & \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} \\ 0 & 0 & \frac{\sqrt{2}}{\sqrt{3}} \end{bmatrix}.$$

Example

Consider the matrix $A := \begin{bmatrix} 1 & 1 & 1 \\ \epsilon & 0 & 0 \\ 0 & \epsilon & 0 \\ 0 & 0 & \epsilon \end{bmatrix}$. Then $\mathrm{cond}(A) = \dfrac{\sqrt{3+\epsilon^2}}{\epsilon}$.

If $\epsilon = \sqrt{\mathbf{u}}$ then $\mathbf{u}\,\mathrm{cond}(A) \approx \sqrt{3}\sqrt{\mathbf{u}}$.

Assume that $\epsilon^2 < \mathbf{u}$. Then, in finite precision arithmetic, Q_{CS} produced by CGS and Q_{MS} produced by MGS are given by

$$Q_{\mathrm{CS}} = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \epsilon & \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \text{ and } Q_{\mathrm{MS}} = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \epsilon & \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} \\ 0 & 0 & \frac{\sqrt{2}}{\sqrt{3}} \end{bmatrix}.$$

Note that $\langle u_2, u_3 \rangle = 1/2$. Hence u_2 and u_3 are far from being orthogonal.

Example

Consider the matrix $A := \begin{bmatrix} 1 & 1 & 1 \\ \epsilon & 0 & 0 \\ 0 & \epsilon & 0 \\ 0 & 0 & \epsilon \end{bmatrix}$. Then $\mathrm{cond}(A) = \dfrac{\sqrt{3+\epsilon^2}}{\epsilon}$.

If $\epsilon = \sqrt{\mathbf{u}}$ then $\mathbf{u}\ \mathrm{cond}(A) \approx \sqrt{3}\sqrt{\mathbf{u}}$.

Assume that $\epsilon^2 < \mathbf{u}$. Then, in finite precision arithmetic, Q_{CS} produced by CGS and Q_{MS} produced by MGS are given by

$$Q_{\mathrm{CS}} = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \epsilon & \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \text{ and } Q_{\mathrm{MS}} = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \epsilon & \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} \\ 0 & 0 & \frac{\sqrt{2}}{\sqrt{3}} \end{bmatrix}.$$

Note that $\langle u_2, u_3 \rangle = 1/2$. Hence u_2 and u_3 are far from being orthogonal. However, we have $|\langle u_1, u_2 \rangle| = |\langle u_1, u_3 \rangle| = \epsilon/\sqrt{2}$.

Example

Consider the matrix $A := \begin{bmatrix} 1 & 1 & 1 \\ \epsilon & 0 & 0 \\ 0 & \epsilon & 0 \\ 0 & 0 & \epsilon \end{bmatrix}$. Then $\mathrm{cond}(A) = \dfrac{\sqrt{3+\epsilon^2}}{\epsilon}$.

If $\epsilon = \sqrt{\mathbf{u}}$ then $\mathbf{u}\, \mathrm{cond}(A) \approx \sqrt{3}\sqrt{\mathbf{u}}$.

Assume that $\epsilon^2 < \mathbf{u}$. Then, in finite precision arithmetic, Q_{CS} produced by CGS and Q_{MS} produced by MGS are given by

$$Q_{\mathrm{CS}} = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \epsilon & \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \text{ and } Q_{\mathrm{MS}} = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \epsilon & \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} \\ 0 & 0 & \frac{\sqrt{2}}{\sqrt{3}} \end{bmatrix}.$$

Note that $\langle u_2, u_3 \rangle = 1/2$. Hence u_2 and u_3 are far from being orthogonal. However, we have $|\langle u_1, u_2 \rangle| = |\langle u_1, u_3 \rangle| = \epsilon/\sqrt{2}$. By contrast, we have $\langle v_2, v_3 \rangle = 0$ so v_2 and v_3 are orthogonal.

Example

Consider the matrix $A := \begin{bmatrix} 1 & 1 & 1 \\ \epsilon & 0 & 0 \\ 0 & \epsilon & 0 \\ 0 & 0 & \epsilon \end{bmatrix}$. Then $\mathrm{cond}(A) = \dfrac{\sqrt{3+\epsilon^2}}{\epsilon}$.

If $\epsilon = \sqrt{\mathbf{u}}$ then $\mathbf{u}\,\mathrm{cond}(A) \approx \sqrt{3}\sqrt{\mathbf{u}}$.

Assume that $\epsilon^2 < \mathbf{u}$. Then, in finite precision arithmetic, Q_{CS} produced by CGS and Q_{MS} produced by MGS are given by

$$Q_{\mathrm{CS}} = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \epsilon & \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \text{ and } Q_{\mathrm{MS}} = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \epsilon & \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} \\ 0 & 0 & \frac{\sqrt{2}}{\sqrt{3}} \end{bmatrix}.$$

Note that $\langle u_2, u_3 \rangle = 1/2$. Hence u_2 and u_3 are far from being orthogonal. However, we have $|\langle u_1, u_2 \rangle| = |\langle u_1, u_3 \rangle| = \epsilon/\sqrt{2}$. By contrast, we have $\langle v_2, v_3 \rangle = 0$ so v_2 and v_3 are orthogonal. Also, we have $|\langle v_1, v_2 \rangle| = \epsilon/\sqrt{2}$ and $|\langle v_1, v_3 \rangle| = \epsilon/\sqrt{6}$.

Example

Now, consider the matrix $A := \begin{bmatrix} 1 & 1 & 1 \\ \epsilon & 0 & 0 \\ 0 & \epsilon & 0 \\ 0 & 0 & \epsilon \end{bmatrix}$ and choose $\epsilon = .5 \times 10^{-8}$.

Computing QR factorization $A = QR$ using CGS, MGS and Householder QR factorization in MATLAB yields

Error	cgs	mgs	qr
$\|Q^*Q - I\|_2$	5.0000e − 01	4.0825e − 09	2.2888e − 16
$\|A - QR\|_2$	1.4904e − 25	1.1293e − 25	2.9772e − 24

The last column shows that Q is an isometry up to working precision showing that qr is backward stable. We have $\mathrm{cond}(A) = 3.5 \times 10^8$.

Example

Now, consider the matrix $A := \begin{bmatrix} 1 & 1 & 1 \\ \epsilon & 0 & 0 \\ 0 & \epsilon & 0 \\ 0 & 0 & \epsilon \end{bmatrix}$ and choose $\epsilon = .5 \times 10^{-8}$.

Computing QR factorization $A = QR$ using CGS, MGS and Householder QR factorization in MATLAB yields

Error	`cgs`	`mgs`	`qr`
$\|Q^*Q - I\|_2$	5.0000e − 01	4.0825e − 09	2.2888e − 16
$\|A - QR\|_2$	1.4904e − 25	1.1293e − 25	2.9772e − 24

The last column shows that Q is an isometry up to working precision showing that `qr` is backward stable. We have $\mathrm{cond}(A) = 3.5 \times 10^8$.

For MGS $\|Q^*Q - I\|_2 \approx \mathrm{cond}(A)\mathbf{u}$. Hence MGS is conditionally backward stable.

Example

Now, consider the matrix $A := \begin{bmatrix} 1 & 1 & 1 \\ \epsilon & 0 & 0 \\ 0 & \epsilon & 0 \\ 0 & 0 & \epsilon \end{bmatrix}$ and choose $\epsilon = .5 \times 10^{-8}$.

Computing QR factorization $A = QR$ using CGS, MGS and Householder QR factorization in MATLAB yields

Error	cgs	mgs	qr
$\|Q^*Q - I\|_2$	5.0000e − 01	4.0825e − 09	2.2888e − 16
$\|A - QR\|_2$	1.4904e − 25	1.1293e − 25	2.9772e − 24

The last column shows that Q is an isometry up to working precision showing that `qr` is backward stable. We have $\mathrm{cond}(A) = 3.5 \times 10^8$.

For MGS $\|Q^*Q - I\|_2 \approx \mathrm{cond}(A)\mathbf{u}$. Hence MGS is conditionally backward stable. But Q computed by CGS is not an isometry and hence CGS is an unstable algorithm. ■

CGS versus MGS in finite precision arithmetic

The CGS and MGS are not equivalent in finite precison arithmetic because the equality

$$I - v_1 v_1^* - \cdots - v_{j-1} v_{j-1}^* = (I - v_1 v_1^*)(I - v_2 v_2^*) \cdots (I - v_{j-1} v_{j-1}^*) \qquad (**)$$

does not hold in finite precision arithmetic.

CGS versus MGS in finite precision arithmetic

The CGS and MGS are not equivalent in finite precison arithmetic because the equality

$$I - v_1 v_1^* - \cdots - v_{j-1} v_{j-1}^* = (I - v_1 v_1^*)(I - v_2 v_2^*) \cdots (I - v_{j-1} v_{j-1}^*) \qquad (**)$$

does not hold in finite precision arithmetic.

Note that $P_j := \begin{bmatrix} v_1 & \cdots & v_j \end{bmatrix} \begin{bmatrix} v_1 & \cdots & v_j \end{bmatrix}^*$ is the orthononal projection on $\mathrm{span}(v_1, \ldots, v_j)$ and $Q_j := v_j v_j^*$ is the orthogonal projection on $\mathrm{span}(v_j)$ for $j = 1 : n$. By (**), we have

$$I - P_j = (I - Q_j)(I - Q_{j-1}) \cdots (I - Q_1).$$

In finite precision arithmetic, P_j may not be an orthogonal projection. Therefore, the computed $v_j = (I - P_{j-1})x_j / \|(I - P_{j-1})x_j\|_2$ may fail to be orthogonal to $v_1, \ldots, v_{j-1}$.

CGS versus MGS in finite precision arithmetic

The CGS and MGS are not equivalent in finite precison arithmetic because the equality

$$I - v_1 v_1^* - \cdots - v_{j-1} v_{j-1}^* = (I - v_1 v_1^*)(I - v_2 v_2^*) \cdots (I - v_{j-1} v_{j-1}^*) \qquad (**)$$

does not hold in finite precision arithmetic.

Note that $P_j := \begin{bmatrix} v_1 & \cdots & v_j \end{bmatrix} \begin{bmatrix} v_1 & \cdots & v_j \end{bmatrix}^*$ is the orthononal projection on $\mathrm{span}(v_1, \ldots, v_j)$ and $Q_j := v_j v_j^*$ is the orthogonal projection on $\mathrm{span}(v_j)$ for $j = 1 : n$. By (**), we have

$$I - P_j = (I - Q_j)(I - Q_{j-1}) \cdots (I - Q_1).$$

In finite precision arithmetic, P_j may not be an orthogonal projection. Therefore, the computed $v_j = (I - P_{j-1})x_j / \|(I - P_{j-1})x_j\|_2$ may fail to be orthogonal to $v_1, \ldots, v_{j-1}$.

By contrast, Q_j is an orthogonal projection in finite precision arithmetic and hence

$$v_j = \frac{(I - Q_{j-1}) \cdots ((I - Q_2)((I - Q_1)x_j))}{\|(I - Q_{j-1}) \cdots ((I - Q_2)((I - Q_1)x_j))\|_2}$$

when computed incrementally as in the case of MGS is likely to be orthogonal to $v_1, \ldots, v_{j-1}$.