Classical Gram Schmidt Orthonormalisation

Let $\{v_1, \ldots, v_m\}$ be an ordered set of linearly independent vectors in $\mathbb{R}^n$. The Classical Gram Schmidt (CGS) process finds an ordered orthonormal set of vectors $\{q_1, \ldots, q_m\}$ in $\mathbb{R}^n$ such that

$$\operatorname{span}\{v_1, \ldots, v_k\} = \operatorname{span}\{q_1, \ldots, q_k\}, \ k = 1, \ldots, m.$$

Classical Gram Schmidt Orthonormalisation

Let $\{v_1, \ldots, v_m\}$ be an ordered set of linearly independent vectors in $\mathbb{R}^n$. The Classical Gram Schmidt (CGS) process finds an ordered orthonormal set of vectors $\{q_1, \ldots, q_m\}$ in $\mathbb{R}^n$ such that

$$\operatorname{span}\{v_1, \ldots, v_k\} = \operatorname{span}\{q_1, \ldots, q_k\}, \; k = 1, \ldots, m.$$

Classical Gram Schmidt (CGS):

Step 1: $q_1 := v_1 / \|v_1\|_2$.

Step 2: $q_2 := \underbrace{(v_2 - (v_2^T q_1) q_1)}_{=:\hat{q}_2} / \|v_2 - (v_2^T q_1) q_1\|_2$.

Step k: Assuming that $q_1, \ldots, q_{k-1}$ are calculated as above,

$$q_k = \underbrace{\left(v_k - \sum_{i=1}^{k-1} (v_k^T q_i) q_i\right)}_{=:\hat{q}_k} / \|v_k - \Sigma_{i=1}^{k-1} (v_k^T q_i) q_i\|_2.$$

Classical Gram Schmidt Orthonormalisation

Let $\{v_1, \dots, v_m\}$ be an ordered set of linearly independent vectors in $\mathbb{R}^n$. The Classical Gram Schmidt (CGS) process finds an ordered orthonormal set of vectors $\{q_1, \dots, q_m\}$ in $\mathbb{R}^n$ such that

$$\mathrm{span}\{v_1, \dots, v_k\} = \mathrm{span}\{q_1, \dots, q_k\}, \; k = 1, \dots, m.$$

$r_{ik} = v_k^T q_i, \; i = 1, \dots, k-1$

$r_{kk} = \|\hat{q}_k\|_2$

Classical Gram Schmidt (CGS):

Step 1: $q_1 := v_1 / \|v_1\|_2$.

Step 2: $q_2 := \underbrace{(v_2 - (v_2^T q_1) q_1)}_{=:\hat{q}_2} / \|v_2 - (v_2^T q_1) q_1\|_2$.

Step k: Assuming that $q_1, \dots, q_{k-1}$ are calculated as above,

$$q_k = \underbrace{(v_k - \sum_{i=1}^{k-1} \underbrace{(v_k^T q_i)}_{=: r_{ik}} q_i)}_{=:\hat{q}_k} / \underbrace{\|v_k - \Sigma_{i=1}^{k-1} (v_k^T q_i) q_i\|_2}_{=: r_{kk}}.$$

$\Rightarrow v_k - \sum_{i=1}^{k-1} r_{ik} q_i = r_{kk} q_k$

$\Rightarrow v_k = \sum_{i=1}^{k} r_{ik} q_i$

Exercise: Show that CGS applied to the basis $\{e_1 + e_2, e_2, e_2 + e_3\}$ in $\mathbb{R}^3$ produces the ordered orthonormal basis

$$\{(e_1 + e_2)/\sqrt{2}, (e_2 - e_1)/\sqrt{2}, e_3\}.$$

Equivalence of CGS and condensed QR decomposition

CGS $\equiv$ condensed QR

Supose $\{v_1, \ldots, v_m\}$ is an ordered linearly independent subset of $\mathbb{R}^n$ and $\{q_1, \ldots, q_m\}$ is the output of CGS on $\{v_1, \ldots, v_m\}$.

$\forall k = 1, \ldots, m$, $v_k = \sum_{i=1}^{k} r_{ik} q_i \Rightarrow v_k = \underbrace{[q_1 \cdots q_k \; q_{k+1} \cdots q_m]}_{Q} \underbrace{\begin{bmatrix} r_{1k} \\ \vdots \\ r_{kk} \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{R_k \to k^{th} \text{ column of}}$ $\}m-k$

$= Q[R_k]$

$R = \begin{bmatrix} r_{11} & \cdots & r_{1m} \\ & r_{22} \cdots & r_{2m} \\ & & r_{mm} \end{bmatrix}$

$\therefore V = [v_1 \cdots v_m] = [QR_1 \; QR_2 \cdots QR_m]$

$= QR$

CGS $\equiv$ condensed QR

Supose $\{v_1, \ldots, v_m\}$ is an ordered linearly independent subset of $\mathbb{R}^n$ and $\{q_1, \ldots, q_m\}$ is the output of CGS on $\{v_1, \ldots, v_m\}$. Then,

$$\underbrace{[v_1 \cdots v_m]}_{=:V} = \underbrace{[q_1 \cdots q_m]}_{=:Q} \underbrace{\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ & r_{22} & \cdots & r_{2m} \\ & & \ddots & \vdots \\ & & & r_{mm} \end{bmatrix}}_{=:R}$$

where $r_{ij} = v_j^T q_i$ for $j > i$, $r_{jj} = \|\hat{q}_j\|_2$ and $r_{ij} = 0$ otherwise. Clearly, $V = QR$ is a condensed QR decomposition of V.

CGS $\equiv$ condensed QR

Supose $\{v_1, \ldots, v_m\}$ is an ordered linearly independent subset of $\mathbb{R}^n$ and $\{q_1, \ldots, q_m\}$ is the output of CGS on $\{v_1, \ldots, v_m\}$. Then,

$$\underbrace{[v_1 \cdots v_m]}_{=:V} = \underbrace{[q_1 \cdots q_m]}_{=:Q} \underbrace{\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ & r_{22} & \cdots & r_{2m} \\ & & \ddots & \vdots \\ & & & r_{mm} \end{bmatrix}}_{=:R}$$

where $r_{ij} = v_j^T q_i$ for $j > i$, $r_{jj} = \|\hat{q}_j\|_2$ and $r_{ij} = 0$ otherwise. Clearly, $V = QR$ is a condensed QR decomposition of V.

Exercise: Conversely if $V = QR$ be a condensed QR decomposition of $V = [v_1 \cdots v_m] \in \mathbb{R}^{n \times m}$ where $R = [r_{ij}]_{m \times m}$ with $r_{ii} > 0$ for all $i = 1, \ldots, m$, then the columns $q_1, \ldots, q_m$ of Q are equal to those obtained via CGS on the columns of V with

$$r_{ij} = \begin{cases} v_j^T q_i, & i < j, \\ \|\hat{q}_j\|_2, & i = j \\ 0, & \text{otherwise.} \end{cases}$$

Numerical issues associated with CGS and Modified Gram Schmidt (MGS)

Flop count and numerical issues of CGS

For each $k = 1, 2, \ldots, m$, computing,

$\hat{q}_k = v_k - \sum_{i=1}^{k-1}(v_k^T q_i)q_i \longrightarrow 4n(k-1)$ flops

$\|\hat{q}_k\|_2 \longrightarrow 2n$ flops and one square root.

So,

$$q_k = \frac{\hat{q}_k}{\|\hat{q}_k\|_2} \longrightarrow 4nk - n \text{ flops and one square root.}$$

Flop count and numerical issues of CGS

For each $k = 1, 2, \ldots, m$, computing,

$\hat{q}_k = v_k - \sum_{i=1}^{k-1}(v_k^T q_i)q_i \longrightarrow 4n(k-1)$ flops

$\|\hat{q}_k\|_2 \longrightarrow 2n$ flops and one square root.

So,

$q_k = \frac{\hat{q}_k}{\|\hat{q}_k\|_2} \longrightarrow 4nk - n$ flops and one square root.

Therefore the total cost of CGS is

$$\sum_{k=1}^{m} 4nk - n = 2nm^2 + O(nm) + O(m^2)$$

flops.

Flop count and numerical issues of CGS

For each $k = 1, 2, \ldots, m$, computing,

$\hat{q}_k = v_k - \sum_{i=1}^{k-1}(v_k^T q_i)q_i \longrightarrow 4n(k-1)$ flops

$\|\hat{q}_k\|_2 \longrightarrow 2n$ flops and one square root.

So,

$q_k = \frac{\hat{q}_k}{\|\hat{q}_k\|_2} \longrightarrow 4nk - n$ flops and one square root.

Therefore the total cost of CGS is

$$\sum_{k=1}^{m} 4nk - n = 2nm^2 + O(nm) + O(m^2)$$

flops.

The quality of the orthonormalisation is measured by *the departure from orthonormality* $\|I_m - Q^T Q\|_2$ of the computed Q.

It is considered to be good in the presence of rounding error if $\|I_m - Q^T Q\|_2$ is $O(u)$.

CGS is a poor performer in the presence of rounding error

The quality of orthonormalisation in CGS can be poor in the presence of rounding error.

Example: Consider the set of vectors $\{v_1, v_2, v_3\}$ where

$$v_1 := \begin{bmatrix} 1 \\ \epsilon \\ 0 \\ 0 \end{bmatrix}, \; v_2 := \begin{bmatrix} 1 \\ 0 \\ \epsilon \\ 0 \end{bmatrix}, \; v_3 := \begin{bmatrix} 1 \\ 0 \\ 0 \\ \epsilon \end{bmatrix},$$

where $\epsilon > 0$ is such that $\epsilon^2 < u$. Perform CGS on the set assuming that $fl(1 + \epsilon^2) = 1$ and there is no other rounding and report the departure from orthonormality.

CGS is a poor performer in the presence of rounding error

The quality of orthonormalisation in CGS can be poor in the presence of rounding error.

Example: Consider the set of vectors $\{v_1, v_2, v_3\}$ where

$$v_1 := \begin{bmatrix} 1 \\ \epsilon \\ 0 \\ 0 \end{bmatrix}, \; v_2 := \begin{bmatrix} 1 \\ 0 \\ \epsilon \\ 0 \end{bmatrix}, \; v_3 := \begin{bmatrix} 1 \\ 0 \\ 0 \\ \epsilon \end{bmatrix},$$

where $\epsilon > 0$ is such that $\epsilon^2 < u$. Perform CGS on the set assuming that $fl(1 + \epsilon^2) = 1$ and there is no other rounding and report the departure from orthonormality.

A modification to the CGS process which is theoretically equivalent to CGS called Modified Gram Schmidt(MGS) is seen to be superior in the presence of rounding error.

Modified Gram Schmidt(MGS)

Let $\{v_1, \ldots, v_m\}$ be a linearly independent subset of $\mathbb{R}^n$.
The first two steps of MGS and CGS are the same. Assume that q_1, q_2 are formed.

Step 3: Let

$$
\begin{aligned}
v_3^{(1)} &:= v_3 - (v_3^T q_1) q_1 \\
v_3^{(2)} &:= v_3^{(1)} - \{\left(v_3^{(1)}\right)^T q_2\} q_2 \\
\tilde{q}_3 &:= v_3^{(2)} / \|v_3^{(2)}\|_2.
\end{aligned}
$$

$v_3^{(2)}$ and $\tilde{q}_3$ are respectively the same as $\hat{q}_3$ and q_3 of CGS in theory.
This is because $\left(v_3^{(1)}\right)^T q_2 = v_3^T q_2$ in theory.

However, the computed q_1 and q_2 are not exactly orthogonal to each other. Also the computed $v_3^{(1)}$ is not exactly orthogonal to q_1. So the computed q_3 and $\tilde{q}_3$ are different.

Modified Gram Schmidt(MGS)

Continuing similarly till $\tilde{q}_1, \tilde{q}_2, \ldots, \tilde{q}_{k-1}$ have been found, let

Step k:

$$
\begin{aligned}
v_k^{(1)} &:= v_k - (v_k^T \tilde{q}_1)\tilde{q}_1 \\
v_k^{(2)} &:= v_k^{(1)} - \{\left(v_k^{(1)}\right)^T \tilde{q}_2\}\tilde{q}_2 \\
&\vdots \\
v_k^{(k-1)} &:= \underbrace{v_k^{(k-2)} - \{\left(v_k^{(k-2)}\right)^T \tilde{q}_{k-1}\}\tilde{q}_{k-1}}_{=\hat{q}_k \text{of CGS in theory}} \\
\tilde{q}_k &:= \underbrace{v_k^{(k-1)}/\|v_k^{(k-1)}\|_2}_{=q_k \text{ in theory}}
\end{aligned}
$$

Thus MGS produces the orthonormal set $\{\tilde{q}_1, \tilde{q}_2, \ldots, \tilde{q}_m\}$.

Modified Gram Schmidt(MGS)

Continuing similarly till $\tilde{q}_1, \tilde{q}_2, \ldots, \tilde{q}_{k-1}$ have been found, let

Step k:

$$
\begin{aligned}
v_k^{(1)} &:= v_k - (v_k^T \tilde{q}_1)\tilde{q}_1 \\
v_k^{(2)} &:= v_k^{(1)} - \{\left(v_k^{(1)}\right)^T \tilde{q}_2\}\tilde{q}_2 \\
&\vdots \\
v_k^{(k-1)} &:= \underbrace{v_k^{(k-2)} - \{\left(v_k^{(k-2)}\right)^T \tilde{q}_{k-1}\}\tilde{q}_{k-1}}_{=\hat{q}_k \text{of CGS in theory}} \\
\tilde{q}_k &:= \underbrace{v_k^{(k-1)}/\|v_k^{(k-1)}\|_2}_{=q_k \text{ in theory}}
\end{aligned}
$$

Thus MGS produces the orthonormal set $\{\tilde{q}_1, \tilde{q}_2, \ldots, \tilde{q}_m\}$.

Exercise: Prove that MGS produces an exactly orthonormal set when applied on the set of vectors $\{v_1, v_2, v_3\}$ considered earlier under the same assumptions with respect to rounding.

MGS $\equiv$ Condensed QR

Exercise: *Let* $\{v_1, \ldots, v_m\}$ *be a linearly independent subset of* $\mathbb{R}^n$. *For* $j = 1, \ldots, i-1$, *and* $i = 1, \ldots, m$, *let* $\tilde{q}_i$ *and* $v_i^{(j)}$ *be the vectors obtained via Modified Gram Schmidt.*
Let $\tilde{Q} = [\tilde{q}_1 \cdots \tilde{q}_m]$, $V = [v_1 \cdots v_m]$, *and* $\tilde{R} = [\tilde{r}_{ij}] \in \mathbb{R}^{m \times m}$ *be an upper triangular matrix with* $\tilde{r}_{ik} = \left(v_k^{(i)}\right)^T q_i$ *for* $1 \leq i \leq k-1$ *and* $\tilde{r}_{kk} = \|v_k^{(k-1)}\|_2$ *for* $k = 1, \ldots, m$. *Prove that*

$$V = \tilde{Q}\tilde{R}$$

is theoretically the same condensed QR decomposition of V *as the one via CGS and has the exact same flop count.*

MGS $\equiv$ Condensed QR

Exercise: *Let $\{v_1, \ldots, v_m\}$ be a linearly independent subset of $\mathbb{R}^n$. For $j = 1, \ldots, i-1$, and $i = 1, \ldots, m$, let $\tilde{q}_i$ and $v_i^{(j)}$ be the vectors obtained via Modified Gram Schmidt.*
Let $\tilde{Q} = [\tilde{q}_1 \cdots \tilde{q}_m]$, $V = [v_1 \cdots v_m]$, and $\tilde{R} = [\tilde{r}_{ij}] \in \mathbb{R}^{m\times m}$ be an upper triangular matrix with $\tilde{r}_{ik} = \left(v_k^{(i)}\right)^T q_i$ for $1 \leq i \leq k-1$ and $\tilde{r}_{kk} = \|v_k^{(k-1)}\|_2$ for $k = 1, \ldots, m$. Prove that

$$V = \tilde{Q}\tilde{R}$$

is theoretically the same condensed QR decomposition of V as the one via CGS and has the exact same flop count.

The computed $\tilde{Q}_c$ from MGS satisfies $\|I_m - \tilde{Q}_c^T\tilde{Q}_c\|_2 \approx \kappa_2(V)u$. [Higham, 96], [Björck, 96] So, orthonormalisation is poor if $\kappa_2(V)$ is large.

MGS $\equiv$ Condensed QR

Exercise: *Let $\{v_1, \ldots, v_m\}$ be a linearly independent subset of $\mathbb{R}^n$. For $j = 1, \ldots, i-1$, and $i = 1, \ldots, m$, let $\tilde{q}_i$ and $v_i^{(j)}$ be the vectors obtained via Modified Gram Schmidt.*
Let $\tilde{Q} = [\tilde{q}_1 \cdots \tilde{q}_m]$, $V = [v_1 \cdots v_m]$, and $\tilde{R} = [\tilde{r}_{ij}] \in \mathbb{R}^{m\times m}$ be an upper triangular matrix with $\tilde{r}_{ik} = \left(v_k^{(i)}\right)^T q_i$ for $1 \leq i \leq k-1$ and $\tilde{r}_{kk} = \|v_k^{(k-1)}\|_2$ for $k = 1, \ldots, m$. Prove that

$$V = \tilde{Q}\tilde{R}$$

is theoretically the same condensed QR decomposition of V as the one via CGS and has the exact same flop count.

The computed $\tilde{Q}_c$ from MGS satisfies $\|I_m - \tilde{Q}_c^T \tilde{Q}_c\|_2 \approx \kappa_2(V)u$. [Higham, 96], [Björck, 96] So, orthonormalisation is poor if $\kappa_2(V)$ is large.

Numerically CGS with one more re-orthogonalisation is done. The computed Q_c satisfies $\|I_m - Q_c^T Q_c\|_2 \approx cu$ for some small $c > 0$ if $\kappa_2(V) \ll 1/u$. [Giraud et. al., 2005]