- Modified Gram-Schmidt as triangular orthogonalization.
- Householder Triangularization.

Gram-Schmidt

- Consider classical Gram-Schmidt as a sequence of formulas:

$$q_1 = \frac{P_1 a_1}{||P_1 a_1||}, \quad q_2 = \frac{P_2 a_2}{||P_2 a_2||}, \quad \ldots, \quad q_n = \frac{P_n a_n}{||P_n a_n||}$$

where P_j denotes an orthogonal projector

- Consider classical Gram-Schmidt as a sequence of formulas:

$$q_1 = \frac{P_1 a_1}{||P_1 a_1||}, \quad q_2 = \frac{P_2 a_2}{||P_2 a_2||}, \quad \ldots, \quad q_n = \frac{P_n a_n}{||P_n a_n||}$$

  where P_j denotes an orthogonal projector
- This projector can be represented explicitly. Let Q_{j-1} denote the $m \times (j-1)$ matrix containing the first $(j-1)$ columns of Q:

$$Q_{j-1} = \left[\begin{array}{c|c|c|c} & & & \\ q_1 & q_2 & \ldots & q_{j-1} \\ & & & \end{array} \right]$$

- Consider classical Gram-Schmidt as a sequence of formulas:

$$q_1 = \frac{P_1 a_1}{||P_1 a_1||}, \quad q_2 = \frac{P_2 a_2}{||P_2 a_2||}, \quad \dots, \quad q_n = \frac{P_n a_n}{||P_n a_n||}$$

  where P_j denotes an orthogonal projector
- This projector can be represented explicitly. Let Q_{j-1} denote the $m \times (j-1)$ matrix containing the first $(j-1)$ columns of Q:

$$Q_{j-1} = \left[\begin{array}{c|c|c|c} \\ q_1 & q_2 & \dots & q_{j-1} \\ \\ \end{array} \right]$$

- Then P_j is given by:

$$P_j = I - Q_{j-1} Q_{j-1}^*$$

Modified Gram-Schmidt

- In contrast, if one thinks of the modified Gram-Schmidt algorithm, then computes the same result by a sequence of of $(j-1)$ projections.

Modified Gram-Schmidt

- In contrast, if one thinks of the modified Gram-Schmidt algorithm, then computes the same result by a sequence of of $(j-1)$ projections.
- If $P_{\perp q} = I - qq*$, then in modified Gram-Schmidt,

$$P_j = P_{\perp q_{j-1}} \cdots P_{\perp q_2} P_{\perp q_1}$$

Modified Gram-Schmidt

- In contrast, if one thinks of the modified Gram-Schmidt algorithm, then computes the same result by a sequence of of $(j-1)$ projections.
- If $P_{\perp q} = I - qq*$, then in modified Gram-Schmidt,

$$P_j = P_{\perp q_{j-1}} \cdots P_{\perp q_2} P_{\perp q_1}$$

- Each outer step of the algorithm can be interpreted as a right multiplication by a square upper-triangular matrix.

- In contrast, if one thinks of the modified Gram-Schmidt algorithm, then computes the same result by a sequence of of $(j-1)$ projections.
- If $P_{\perp q} = I - qq*$, then in modified Gram-Schmidt,

$$P_j = P_{\perp q_{j-1}} \cdots P_{\perp q_2} P_{\perp q_1}$$

- Each outer step of the algorithm can be interpreted as a right multiplication by a square upper-triangular matrix.
- For example, beginning with A, the first iteration multiplies the first column a_1 with $\frac{1}{r_{11}}$ and then subtracts r_{1j} times the result from each of the remaining columns a_j.

Triangular Orthogonalization

- This is equivalent to right-multiplication by a matrix R_1:

$$\left[\begin{array}{c|c|c|c} & & & \\ v_1 & v_2 & \dots & v_n \\ & & & \end{array}\right] \begin{bmatrix} \frac{1}{r_{11}} & \frac{-r_{12}}{r_{11}} & \dots & \frac{-r_{1n}}{r_{11}} \\ & 1 & & \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix} = \left[\begin{array}{c|c|c|c} & & & \\ q_1 & v_2^{(2)} & \dots & v_n^{(2)} \\ & & & \end{array}\right]$$

Triangular Orthogonalization

- This is equivalent to right-multiplication by a matrix R_1:

$$\begin{bmatrix} \Bigg| & \Bigg| & & \Bigg| \\ v_1 & v_2 & \dots & v_n \\ \Bigg| & \Bigg| & & \Bigg| \end{bmatrix} \begin{bmatrix} \frac{1}{r_{11}} & \frac{-r_{12}}{r_{11}} & \dots & \frac{-r_{1n}}{r_{11}} \\ & 1 & & \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix} = \begin{bmatrix} \Bigg| & \Bigg| & & \Bigg| \\ q_1 & v_2^{(2)} & \dots & v_n^{(2)} \\ \Bigg| & \Bigg| & & \Bigg| \end{bmatrix}$$

- In general, step i subtracts r_{ij}/r_{ii} times column i of the current A from columns, $j > i$ and replaces column i by $1/r_{ii}$ times itself. This corresponds to multiplication by upper triangular matrix R_i:

$$R_2 = \begin{bmatrix} 1 & & & \\ & \frac{1}{r_{22}} & \frac{-r_{23}}{r_{22}} & \dots \\ & & 1 & \\ & & & \ddots \end{bmatrix} R_3 = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \frac{1}{r_{33}} & \dots \\ & & & \ddots \end{bmatrix}, \dots$$

Triangular Orthogonalization

- At the end of the iteration:

$$A \underbrace{R_1 R_2 \cdots R_n}_{R^{-1}} = Q$$

Triangular Orthogonalization

- At the end of the iteration:

$$A \underbrace{R_1 R_2 \cdots R_n}_{R^{-1}} = Q$$

- This shows that Gram-Schmidt is a method of triangular orthogonalization: It applies triangular operations on the right of a matrix to reduce it to a matrix of orthonormal columns.

Householder triangularization

- Householder method applies a succession of elementary unitary matrices Q_k on the left of A so that the resulting matrix:

$$\underbrace{Q_n \cdots Q_2 Q_1}_{Q*} A = R$$

is upper triangular.

Householder triangularization

- Householder method applies a succession of elementary unitary matrices Q_k on the left of A so that the resulting matrix:

$$\underbrace{Q_n \cdots Q_2 Q_1}_{Q*} A = R$$

  is upper triangular.
- The product, $Q = Q_1^* Q_2^* \cdots Q_n^*$ is unitary too and therefore $A = QR$ is the QR factorization of A.

Householder triangularization

- Householder method applies a succession of elementary unitary matrices Q_k on the left of A so that the resulting matrix:

$$\underbrace{Q_n \cdots Q_2 Q_1}_{Q*} A = R$$

  is upper triangular.
- The product, $Q = Q_1^* Q_2^* \cdots Q_n^*$ is unitary too and therefore $A = QR$ is the QR factorization of A.
-
  - Gram-Schmidt: Triangular Orthogonalization

Householder triangularization

- Householder method applies a succession of elementary unitary matrices Q_k on the left of A so that the resulting matrix:

$$\underbrace{Q_n \cdots Q_2 Q_1}_{Q*} A = R$$

  is upper triangular.
- The product, $Q = Q_1^* Q_2^* \cdots Q_n^*$ is unitary too and therefore $A = QR$ is the QR factorization of A.
-
  - Gram-Schmidt: Triangular Orthogonalization
  - Householder : Orthogonal triangularization

Triangularization by introducing zeroes

- The matrix Q_k are chosen such that it introduces zeros below the diagonal in the kth column, while preserving all the zeroes previously introduced.

Triangularization by introducing zeroes

- The matrix Q_k are chosen such that it introduces zeros below the diagonal in the kth column, while preserving all the zeroes previously introduced.
- For example, in the 5×3 case, the zeroes are introduced in the following way:

$$\begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix} \xrightarrow{Q_1} \begin{bmatrix} \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} \\ \mathbf{0} & \boldsymbol{\times} & \boldsymbol{\times} \\ \mathbf{0} & \boldsymbol{\times} & \boldsymbol{\times} \\ \mathbf{0} & \boldsymbol{\times} & \boldsymbol{\times} \\ \mathbf{0} & \boldsymbol{\times} & \boldsymbol{\times} \end{bmatrix} \xrightarrow{Q_2} \begin{bmatrix} \times & \times & \times \\ & \boldsymbol{\times} & \boldsymbol{\times} \\ & \mathbf{0} & \boldsymbol{\times} \\ & \mathbf{0} & \boldsymbol{\times} \\ & \mathbf{0} & \boldsymbol{\times} \end{bmatrix} \xrightarrow{Q_3} \begin{bmatrix} \times & \times & \times \\ & \times & \times \\ & & \boldsymbol{\times} \\ & & \mathbf{0} \\ & & \mathbf{0} \end{bmatrix}$$

- How do we construct the reflectors that will do this?

Householder Reflectors

- How do we construct the reflectors that will do this?
- In general, each Q_k is chosen to be a unitary matrix:

$$Q_k = \begin{bmatrix} I & 0 \\ 0 & F \end{bmatrix}$$

where I is a $k \times k$ identity matrix and F is an $(m-k+1) \times (m-k+1)$ unitary matrix.

Householder Reflectors

- How do we construct the reflectors that will do this?
- In general, each Q_k is chosen to be a unitary matrix:

$$Q_k = \begin{bmatrix} I & 0 \\ 0 & F \end{bmatrix}$$

  where I is a $k \times k$ identity matrix and F is an $(m-k+1) \times (m-k+1)$ unitary matrix.
- Multiplication by F should introduce zeroes in the kth column.

Householder reflector

- Suppose at the beginning of step k, the entries $k, \ldots, m$ of the kth column are given by vector $x \in \mathbb{C}^{m-k+1}$

Householder reflector

- Suppose at the beginning of step k, the entries $k, \ldots, m$ of the kth column are given by vector $x \in \mathbb{C}^{m-k+1}$
- To introduce zeroes the Householder reflector, F should have the following effect:

$$x = \begin{bmatrix} \times \\ \times \\ \times \\ \vdots \\ \times \end{bmatrix} \xrightarrow{F} Fx = \begin{bmatrix} ||x|| \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Householder reflector

- Suppose at the beginning of step k, the entries $k, \ldots, m$ of the kth column are given by vector $x \in \mathbb{C}^{m-k+1}$
- To introduce zeroes the Householder reflector, F should have the following effect:

$$x = \begin{bmatrix} \times \\ \times \\ \times \\ \vdots \\ \times \end{bmatrix} \xrightarrow{F} Fx = \begin{bmatrix} ||x|| \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

- The reflector F will reflect the space $\mathbb{C}^{m-k+1}$ across the hyperplane H orthogonal to $v = ||x||e_1 - x$

Householder reflector

- Suppose at the beginning of step k, the entries $k, \ldots, m$ of the kth column are given by vector $x \in \mathbb{C}^{m-k+1}$
- To introduce zeroes the Householder reflector, F should have the following effect:

$$x = \begin{bmatrix} \times \\ \times \\ \times \\ \vdots \\ \times \end{bmatrix} \xrightarrow{F} Fx = \begin{bmatrix} ||x|| \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

- The reflector F will reflect the space $\mathbb{C}^{m-k+1}$ across the hyperplane H orthogonal to $v = ||x||e_1 - x$
- The matrix F is:

$$F = I - 2\frac{vv^*}{v^*v}$$

- Given a non-zero p-vector $y = (y_1, y_2, \ldots, y_p)$ define:

$$w = \begin{bmatrix} y_1 + \mathsf{sign}(y_1)||y|| \\ y_2 \\ \vdots \\ y_p \end{bmatrix}, \qquad v = \frac{1}{||w||} w$$

Householder Reflector

- Given a non-zero p-vector $y = (y_1, y_2, \ldots, y_p)$ define:

$$w = \begin{bmatrix} y_1 + \mathsf{sign}(y_1)||y|| \\ y_2 \\ \vdots \\ y_p \end{bmatrix}, \qquad v = \frac{1}{||w||} w$$

- Vector w satisfies $||w||^2 = 2(w^* y) = 2||y||(||y|| + |y_1|)$

Householder Reflector

- Given a non-zero p-vector $y = (y_1, y_2, \ldots, y_p)$ define:

$$w = \begin{bmatrix} y_1 + \mathsf{sign}(y_1)||y|| \\ y_2 \\ \vdots \\ y_p \end{bmatrix}, \qquad v = \frac{1}{||w||} w$$

- Vector w satisfies $||w||^2 = 2(w^* y) = 2||y||(||y|| + |y_1|)$
- The reflector $F = I - 2vv^*$ maps y to multiple of $e_1 = (1, 0, \ldots, 0)$:

$$Fy = y - \frac{2(w^* y)}{||w||^2} w = y - w = -\mathsf{sign}(y_1)||y||e_1$$

Householder QR factorization

The following algorithm computes the factor R of a QR factorization of a $m \times n$ matrix A ($m \geq n$), leaving the result in place of A. n reflection vectors, $v_1, v_2, \ldots, v_n$ are stored for later use:

Algorithm 1 Householder QR Factorization

for $k = 1$ to n **do**
  $x = A_{k:m,k}$
  $v_k = \text{sign}(x_1)||x||_2 e_1 + x$
  $v_k = v_k / ||v_k||_2$
  $A_{k:m,k:n} = A_{k:m,k:n} - 2v_k(v_k^* A_{k:m,k:n})$
end for

Applying or forming Q

- Upon completion A has been reduced to upper triangular matrix, but the Q matrix has not been formed.

Applying or forming Q

- Upon completion A has been reduced to upper triangular matrix, but the Q matrix has not been formed.
- This is because explicit construction of Q requires additional work – however we often don't need it explicitly.

Applying or forming Q

- Upon completion A has been reduced to upper triangular matrix, but the Q matrix has not been formed.
- This is because explicit construction of Q requires additional work – however we often don't need it explicitly.
- For solving $Ax = b$, we need to evaluate Q^*b which we do as follows:

Algorithm 6 Implicit calculation of Q^*b

```
for k = 1 to n do
    b_{k:m} = b_{k:m} - 2 v_k (v_k^* b_{k:m})
end for
```

Applying or forming Q

- Upon completion A has been reduced to upper triangular matrix, but the Q matrix has not been formed.
- This is because explicit construction of Q requires additional work – however we often don't need it explicitly.
- For solving $Ax = b$, we need to evaluate Q^*b which we do as follows:

Algorithm 8 Implicit calculation of Q^*b

1: **for** $k = 1$ to n **do**
2: $\quad b_{k:m} = b_{k:m} - 2v_k(v_k^* b_{k:m})$
3: **end for**

- Similarly Qx can also be evaluated:

Algorithm 9 Implicit calculation of Qx

1: **for** $k = n$ down to 1 **do**
2: $\quad x_{k:m} = x_{k:m} - 2v_k(v_k^* x_{k:m})$
3: **end for**

Householder triangularization

- Work for Householder orthogonalization $\sim 2mn^2 - \frac{2}{3}n^3$

- Work for Householder orthogonalization $\sim 2mn^2 - \frac{2}{3}n^3$
- Work for (modified) Gram-Schmidt: $\sim 2mn^2$

- Work for Householder orthogonalization $\sim 2mn^2 - \frac{2}{3}n^3$
- Work for (modified) Gram-Schmidt: $\sim 2mn^2$
- Householder triangularization is numerically more stable than Gram-Schmidt and hence is used for QR factorization.

Example

$$A = \begin{bmatrix} -1 & -1 & 1 \\ 1 & 3 & 3 \\ -1 & -1 & 5 \\ 1 & 3 & 7 \end{bmatrix} = Q_1 Q_2 Q_3 \begin{bmatrix} R \\ 0 \end{bmatrix}$$

Example

$$A = \begin{bmatrix} -1 & -1 & 1 \\ 1 & 3 & 3 \\ -1 & -1 & 5 \\ 1 & 3 & 7 \end{bmatrix} = Q_1 Q_2 Q_3 \begin{bmatrix} R \\ 0 \end{bmatrix}$$

We compute reflectors Q_1, Q_2, Q_3 that trangularize A:

$$Q_3 Q_2 Q_1 A = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ 0 & R_{22} & R_{23} \\ 0 & 0 & R_{33} \\ 0 & 0 & 0 \end{bmatrix}$$

Example

Compute the reflector that maps first column of A to multiple of e_1

$$y = \begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix},\ w = y - ||y||e_1 = \begin{bmatrix} -3 \\ 1 \\ -1 \\ 1 \end{bmatrix},\ v_1 = \frac{1}{||w||}w = \frac{1}{2\sqrt{3}} \begin{bmatrix} -3 \\ 1 \\ -1 \\ 1 \end{bmatrix}$$

Example

Compute the reflector that maps first column of A to multiple of e_1

$$y = \begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix}, \; w = y - ||y|| e_1 = \begin{bmatrix} -3 \\ 1 \\ -1 \\ 1 \end{bmatrix}, \; v_1 = \frac{1}{||w||} w = \frac{1}{2\sqrt{3}} \begin{bmatrix} -3 \\ 1 \\ -1 \\ 1 \end{bmatrix}$$

Overwrite A with $I - 2v_1 v_1^*$

$$A := (I - 2v_1 v_1^*)A = \begin{bmatrix} 2 & 4 & 2 \\ 0 & 4/3 & 8/3 \\ 0 & 2/3 & 16/3 \\ 0 & 4/3 & 20/3 \end{bmatrix}$$

Example

Compute the reflector that maps $A_{2:4,2}$ to multiple of e_1

$$y = \begin{bmatrix} 4/3 \\ 2/3 \\ 4/3 \end{bmatrix}, \; w = y + ||y||e_1 = \begin{bmatrix} 10/3 \\ 2/3 \\ 4/3 \end{bmatrix}, \; v_2 = \frac{1}{||w||} w = \frac{1}{\sqrt{30}} \begin{bmatrix} 5 \\ 1 \\ 2 \end{bmatrix}$$

Example

Compute the reflector that maps $A_{2:4,2}$ to multiple of e_1

$$y = \begin{bmatrix} 4/3 \\ 2/3 \\ 4/3 \end{bmatrix}, \ w = y + ||y||e_1 = \begin{bmatrix} 10/3 \\ 2/3 \\ 4/3 \end{bmatrix}, \ v_2 = \frac{1}{||w||} w = \frac{1}{\sqrt{30}} \begin{bmatrix} 5 \\ 1 \\ 2 \end{bmatrix}$$

Overwrite $A_{2:4,2:3}$ with $I - 2v_2v_2^*$

$$A := \begin{bmatrix} 1 & 0 \\ 0 & I - 2v_2v_2^* \end{bmatrix} A = \begin{bmatrix} 2 & 4 & 2 \\ 0 & -2 & -8 \\ 0 & 0 & 16/5 \\ 0 & 0 & 12/5 \end{bmatrix}$$

Example

Compute the reflector that maps $A_{3:4,3}$ to multiple of e_1

$$y = \begin{bmatrix} 16/5 \\ 12/5 \end{bmatrix},\ w = y + ||y||e_1 = \begin{bmatrix} 36/5 \\ 12/5 \end{bmatrix},\ v_3 = \frac{1}{||w||} w = \frac{1}{\sqrt{10}} \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

Example

Compute the reflector that maps $A_{3:4,3}$ to multiple of e_1

$$y = \begin{bmatrix} 16/5 \\ 12/5 \end{bmatrix},\ w = y + ||y||e_1 = \begin{bmatrix} 36/5 \\ 12/5 \end{bmatrix},\ v_3 = \frac{1}{||w||} w = \frac{1}{\sqrt{10}} \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

Overwrite $A_{3:4,3}$ with $I - 2v_3v_3^*$

$$A := \begin{bmatrix} I & 0 \\ 0 & I - 2v_3v_3^* \end{bmatrix} A = \begin{bmatrix} 2 & 4 & 2 \\ 0 & -2 & -8 \\ 0 & 0 & -4 \\ 0 & 0 & 0 \end{bmatrix}$$