CENG 216 – Numerical Computation

Least Squares - Part II

Mustafa Özuysal
mustafaozuysal@iyte.edu.tr

May 25, 2022

İzmir Institute of Technology

Slide Credits

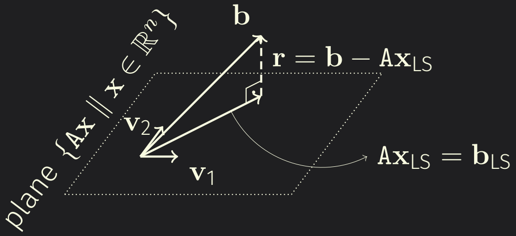
Slides are based on material from the main textbook:

"Numerical Analysis", The new international edition, 2ed, by Timothy Sauer

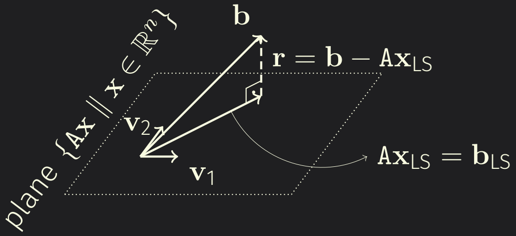
Previously on CENG 216

$\mathtt{A}\mathbf{x} = \mathbf{b}$ with no solutions $\rightarrow$ Equations are inconsistent.

$\mathtt{A}\mathbf{x} = \mathbf{b}$ with no solutions $\rightarrow$ Equations are inconsistent.



$\mathtt{A}\mathbf{x} = \mathbf{b}$ with no solutions $\rightarrow$ Equations are inconsistent.



$$(\mathbf{b} - \mathtt{A}\mathbf{x}_{\mathsf{LS}}) \perp \{\mathtt{A}\mathbf{x} \parallel \mathbf{x} \in \mathbb{R}^n\} \implies \mathtt{A}^\top \mathtt{A}\mathbf{x}_{\mathsf{LS}} = \mathtt{A}^\top \mathbf{b} \quad \text{(Normal Equations)}$$

Input: $\mathtt{A}_{m\times n} = \begin{bmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n\end{bmatrix}$ with linearly independent columns.

```
for j ← 1 to n do
    y = a_j
    for i ← 1 to j − 1
      do
        r_ij = q_i^T a_j
        y = y − r_ij q_i
    end
    r_jj = ||y||
    q_j = y / r_jj
end
```

Output: $\mathtt{Q}_{m\times n}$ and $\mathtt{R}_{n\times n}$ such that

$$\mathtt{A}_{m\times n} = \begin{bmatrix}\mathbf{q}_1 & \mathbf{q}_2 & \dots & \mathbf{q}_n\end{bmatrix}\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn}\end{bmatrix}$$

$$= \mathtt{Q}_{m\times n}\mathtt{R}_{n\times n}$$

Modified Gram-Schmidt

Modified Gram-Schmidt Algorithm

Input: $A_{m\times n} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$ with linearly independent columns.

Classical Gram-Schmidt:

for $j \leftarrow 1$ to n do
  $\mathbf{y} = \mathbf{a}_j$
  for $i \leftarrow 1$ to $j-1$ do
    $r_{ij} = \mathbf{q}_i^\top \mathbf{a}_j$
    $\mathbf{y} = \mathbf{y} - r_{ij}\mathbf{q}_i$
  end
  $r_{jj} = \|\mathbf{y}\|$
  $\mathbf{q}_j = \frac{\mathbf{y}}{r_{jj}}$
end

Modified Gram-Schmidt:

for $j \leftarrow 1$ to n do
  $\mathbf{y} = \mathbf{a}_j$
  for $i \leftarrow 1$ to $j-1$ do
    $r_{ij} = \mathbf{q}_i^\top \mathbf{y}$
    $\mathbf{y} = \mathbf{y} - r_{ij}\mathbf{q}_i$
  end
  $r_{jj} = \|\mathbf{y}\|$
  $\mathbf{q}_j = \frac{\mathbf{y}}{r_{jj}}$
end

Example: Almost Parallel Vectors

Apply classical Gram-Schmidt to $\begin{bmatrix} 1 & 1 & 1 \\ \delta & 0 & 0 \\ 0 & \delta & 0 \\ 0 & 0 & \delta \end{bmatrix}$ with $\delta = 10^{-10}$.

Example: Almost Parallel Vectors

Apply classical Gram-Schmidt to $\begin{bmatrix} 1 & 1 & 1 \\ \delta & 0 & 0 \\ 0 & \delta & 0 \\ 0 & 0 & \delta \end{bmatrix}$ with $\delta = 10^{-10}$.

$$\mathbf{y}_1 = \mathbf{a}_1 = \begin{bmatrix} 1 \\ \delta \\ 0 \\ 0 \end{bmatrix}$$

Example: Almost Parallel Vectors

Apply classical Gram-Schmidt to $\begin{bmatrix} 1 & 1 & 1 \\ \delta & 0 & 0 \\ 0 & \delta & 0 \\ 0 & 0 & \delta \end{bmatrix}$ with $\delta = 10^{-10}$.

$$\mathbf{y}_1 = \mathbf{a}_1 = \begin{bmatrix} 1 \\ \delta \\ 0 \\ 0 \end{bmatrix} \implies \mathbf{q}_1 = \frac{1}{\underbrace{\sqrt{1+\delta^2}}_{\approx 1 \text{ since } \delta^2 = 10^{-20}}} \begin{bmatrix} 1 \\ \delta \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ \delta \\ 0 \\ 0 \end{bmatrix}$$

Example: Almost Parallel Vectors

Apply classical Gram-Schmidt to $\begin{bmatrix} 1 & 1 & 1 \\ \delta & 0 & 0 \\ 0 & \delta & 0 \\ 0 & 0 & \delta \end{bmatrix}$ with $\delta = 10^{-10}$.

$$\mathbf{q}_1 = \begin{bmatrix} 1 \\ \delta \\ 0 \\ 0 \end{bmatrix},$$

$$\mathbf{y}_2 = \begin{bmatrix} 1 \\ 0 \\ \delta \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ \delta \\ 0 \\ 0 \end{bmatrix} \underbrace{\mathbf{q}_1^\top \mathbf{a}_2}_{1} = \begin{bmatrix} 0 \\ -\delta \\ \delta \\ 0 \end{bmatrix}$$

Example: Almost Parallel Vectors

Apply classical Gram-Schmidt to $\begin{bmatrix} 1 & 1 & 1 \\ \delta & 0 & 0 \\ 0 & \delta & 0 \\ 0 & 0 & \delta \end{bmatrix}$ with $\delta = 10^{-10}$.

$$\mathbf{q}_1 = \begin{bmatrix} 1 \\ \delta \\ 0 \\ 0 \end{bmatrix},$$

$$\mathbf{y}_2 = \begin{bmatrix} 1 \\ 0 \\ \delta \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ \delta \\ 0 \\ 0 \end{bmatrix} \underbrace{\mathbf{q}_1^\top \mathbf{a}_2}_{1} = \begin{bmatrix} 0 \\ -\delta \\ \delta \\ 0 \end{bmatrix} \implies \mathbf{q}_2 = \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$$

Example: Almost Parallel Vectors

Apply classical Gram-Schmidt to $\begin{bmatrix} 1 & 1 & 1 \\ \delta & 0 & 0 \\ 0 & \delta & 0 \\ 0 & 0 & \delta \end{bmatrix}$ with $\delta = 10^{-10}$.

$$\mathbf{q}_1 = \begin{bmatrix} 1 \\ \delta \\ 0 \\ 0 \end{bmatrix}, \mathbf{q}_2 = \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix},$$

$$\mathbf{y}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \delta \end{bmatrix} - \begin{bmatrix} 1 \\ \delta \\ 0 \\ 0 \end{bmatrix} \underbrace{\mathbf{q}_1^\top \mathbf{a}_3}_{1} - \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \underbrace{\mathbf{q}_2^\top \mathbf{a}_3}_{0} = \begin{bmatrix} 0 \\ -\delta \\ 0 \\ \delta \end{bmatrix}$$

Example: Almost Parallel Vectors

Apply classical Gram-Schmidt to $\begin{bmatrix} 1 & 1 & 1 \\ \delta & 0 & 0 \\ 0 & \delta & 0 \\ 0 & 0 & \delta \end{bmatrix}$ with $\delta = 10^{-10}$.

$$\mathbf{q}_1 = \begin{bmatrix} 1 \\ \delta \\ 0 \\ 0 \end{bmatrix}, \mathbf{q}_2 = \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix},$$

$$\mathbf{y}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \delta \end{bmatrix} - \begin{bmatrix} 1 \\ \delta \\ 0 \\ 0 \end{bmatrix} \underbrace{\mathbf{q}_1^\top \mathbf{a}_3}_{1} - \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \underbrace{\mathbf{q}_2^\top \mathbf{a}_3}_{0} = \begin{bmatrix} 0 \\ -\delta \\ 0 \\ \delta \end{bmatrix} \implies \mathbf{q}_3 = \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

Example: Almost Parallel Vectors

Apply classical Gram-Schmidt to $\begin{bmatrix} 1 & 1 & 1 \\ \delta & 0 & 0 \\ 0 & \delta & 0 \\ 0 & 0 & \delta \end{bmatrix}$ with $\delta = 10^{-10}$.

$$\mathbf{q}_1 = \begin{bmatrix} 1 \\ \delta \\ 0 \\ 0 \end{bmatrix}, \mathbf{q}_2 = \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \mathbf{q}_3 = \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix},$$

Example: Almost Parallel Vectors

Apply classical Gram-Schmidt to $\begin{bmatrix} 1 & 1 & 1 \\ \delta & 0 & 0 \\ 0 & \delta & 0 \\ 0 & 0 & \delta \end{bmatrix}$ with $\delta = 10^{-10}$.

$$\mathbf{q}_1 = \begin{bmatrix} 1 \\ \delta \\ 0 \\ 0 \end{bmatrix}, \mathbf{q}_2 = \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \mathbf{q}_3 = \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix},$$

$$\implies \mathbf{q}_2^\top \mathbf{q}_3 = \begin{bmatrix} 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix} \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \frac{1}{2} \quad \text{(Not orthogonal!)}$$

Example: Almost Parallel Vectors

Apply modified Gram-Schmidt to $\begin{bmatrix} 1 & 1 & 1 \\ \delta & 0 & 0 \\ 0 & \delta & 0 \\ 0 & 0 & \delta \end{bmatrix}$ with $\delta = 10^{-10}$.

Apply modified Gram-Schmidt to $\begin{bmatrix} 1 & 1 & 1 \\ \delta & 0 & 0 \\ 0 & \delta & 0 \\ 0 & 0 & \delta \end{bmatrix}$ with $\delta = 10^{-10}$.

$$\mathbf{q}_1 = \begin{bmatrix} 1 \\ \delta \\ 0 \\ 0 \end{bmatrix}, \mathbf{q}_2 = \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$$ (Same as before),

$$\mathbf{y} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \delta \end{bmatrix} - \begin{bmatrix} 1 \\ \delta \\ 0 \\ 0 \end{bmatrix} \underbrace{\mathbf{q}_1^\top \mathbf{a}_3}_{1} = \begin{bmatrix} 0 \\ -\delta \\ 0 \\ \delta \end{bmatrix},$$

Example: Almost Parallel Vectors

Apply modified Gram-Schmidt to $\begin{bmatrix} 1 & 1 & 1 \\ \delta & 0 & 0 \\ 0 & \delta & 0 \\ 0 & 0 & \delta \end{bmatrix}$ with $\delta = 10^{-10}$.

$$\mathbf{q}_1 = \begin{bmatrix} 1 \\ \delta \\ 0 \\ 0 \end{bmatrix}, \mathbf{q}_2 = \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \text{ (Same as before),}$$

$$\mathbf{y} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \delta \end{bmatrix} - \begin{bmatrix} 1 \\ \delta \\ 0 \\ 0 \end{bmatrix} \underbrace{\mathbf{q}_1^\top \mathbf{a}_3}_{1} = \begin{bmatrix} 0 \\ -\delta \\ 0 \\ \delta \end{bmatrix}, \quad \mathbf{y}_3 = \mathbf{y} - \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \underbrace{\mathbf{q}_2^\top \mathbf{y}}_{\frac{\delta}{\sqrt{2}}} = \begin{bmatrix} 0 \\ -\frac{\delta}{2} \\ -\frac{\delta}{2} \\ \delta \end{bmatrix} \implies \mathbf{q}_3 = \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{6}} \end{bmatrix}$$

Example: Almost Parallel Vectors

Apply modified Gram-Schmidt to $\begin{bmatrix} 1 & 1 & 1 \\ \delta & 0 & 0 \\ 0 & \delta & 0 \\ 0 & 0 & \delta \end{bmatrix}$ with $\delta = 10^{-10}$.

$$\mathbf{q}_1 = \begin{bmatrix} 1 \\ \delta \\ 0 \\ 0 \end{bmatrix}, \mathbf{q}_2 = \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \mathbf{q}_3 = \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{6}} \end{bmatrix},$$

Example: Almost Parallel Vectors

Apply modified Gram-Schmidt to $\begin{bmatrix} 1 & 1 & 1 \\ \delta & 0 & 0 \\ 0 & \delta & 0 \\ 0 & 0 & \delta \end{bmatrix}$ with $\delta = 10^{-10}$.

$$\mathbf{q}_1 = \begin{bmatrix} 1 \\ \delta \\ 0 \\ 0 \end{bmatrix}, \mathbf{q}_2 = \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \mathbf{q}_3 = \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{6}} \end{bmatrix},$$

$$\implies \mathbf{q}_2^\top \mathbf{q}_3 = \begin{bmatrix} 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix} \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{6}} \end{bmatrix} = 0 \quad \text{(Orthogonal)}$$

The Full QR Factorization

Given the $n < m$ columns of an $m \times n$ matrix $\mathtt{A}$, add $m - n$ extra independent vectors to $\{\mathbf{a}_j\}$ and compute $\mathbf{q}_1$ to $\mathbf{q}_m$.

$$\underbrace{\begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix}}_{\mathtt{A}_{m\times n}} = \underbrace{\begin{bmatrix} \mathbf{q}_1 & \cdots & \mathbf{q}_n & \mathbf{q}_{n+1} & \cdots & \mathbf{q}_m \end{bmatrix}}_{\mathtt{Q}_{m\times m}} \underbrace{\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}}_{\mathtt{R}_{m\times n}}$$

Example

Find the full QR factorization of $\mathtt{A} = \begin{bmatrix} 1 & -4 \\ 2 & 3 \\ 2 & 2 \end{bmatrix}$.

Find the full QR factorization of $\mathtt{A} = \begin{bmatrix} 1 & -4 \\ 2 & 3 \\ 2 & 2 \end{bmatrix}$.

Previously, we computed $\mathbf{q}_1 = \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{bmatrix}$ and $\mathbf{q}_2 = \begin{bmatrix} -\frac{14}{15} \\ \frac{1}{3} \\ \frac{2}{15} \end{bmatrix}$ from $\mathbf{a}_1$ and $\mathbf{a}_2$. Add $\mathbf{a}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.

Example

Find the full QR factorization of $\mathtt{A} = \begin{bmatrix} 1 & -4 \\ 2 & 3 \\ 2 & 2 \end{bmatrix}$.

Previously, we computed $\mathbf{q}_1 = \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{bmatrix}$ and $\mathbf{q}_2 = \begin{bmatrix} -\frac{14}{15} \\ \frac{1}{3} \\ \frac{2}{15} \end{bmatrix}$ from $\mathbf{a}_1$ and $\mathbf{a}_2$. Add $\mathbf{a}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.

$$\mathbf{y}_3 = \mathbf{a}_3 - \mathbf{q}_1 \underbrace{\mathbf{q}_1^\top \mathbf{a}_3}_{\frac{1}{3}} - \mathbf{q}_2 \underbrace{\mathbf{q}_2^\top \mathbf{a}_3}_{-\frac{14}{15}} = \frac{2}{225} \begin{bmatrix} 2 \\ 10 \\ -11 \end{bmatrix}, \quad \mathbf{q}_3 = \begin{bmatrix} \frac{2}{15} \\ \frac{10}{15} \\ -\frac{11}{15} \end{bmatrix}$$

Example

Find the full QR factorization of $\mathtt{A} = \begin{bmatrix} 1 & -4 \\ 2 & 3 \\ 2 & 2 \end{bmatrix}$.

$$\begin{bmatrix} 1 & -4 \\ 2 & 3 \\ 2 & 2 \end{bmatrix} = \frac{1}{15}\begin{bmatrix} 5 & -14 & 2 \\ 10 & 5 & 10 \\ 10 & 2 & -11 \end{bmatrix}\begin{bmatrix} 3 & 2 \\ 0 & 5 \\ 0 & 0 \end{bmatrix} \qquad \text{(Full QR)}$$

Find the full QR factorization of $\mathtt{A} = \begin{bmatrix} 1 & -4 \\ 2 & 3 \\ 2 & 2 \end{bmatrix}$.

$$\begin{aligned}
\begin{bmatrix} 1 & -4 \\ 2 & 3 \\ 2 & 2 \end{bmatrix} &= \frac{1}{15}\begin{bmatrix} 5 & -14 & 2 \\ 10 & 5 & 10 \\ 10 & 2 & -11 \end{bmatrix}\begin{bmatrix} 3 & 2 \\ 0 & 5 \\ 0 & 0 \end{bmatrix} && \text{(Full QR)} \\
&= \frac{1}{15}\begin{bmatrix} 5 & -14 \\ 10 & 5 \\ 10 & 2 \end{bmatrix}\begin{bmatrix} 3 & 2 \\ 0 & 5 \end{bmatrix} && \text{(Reduced QR)}
\end{aligned}$$

Summary of Full QR Factorization with Modified Gram-Schmidt

Input: $\mathtt{A}_{m\times n} = \begin{bmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n\end{bmatrix}$ with linearly independent columns and $m-n$ extra linearly independent vectors.

for $j \leftarrow 1$ **to** m **do**
- $\mathbf{y} = \mathbf{a}_j$
- **for** $i \leftarrow 1$ **to** $j-1$ **do**
  - $r_{ij} = \mathbf{q}_i^\top \mathbf{y}$
  - $\mathbf{y} = \mathbf{y} - r_{ij}\mathbf{q}_i$
- **end**
- $r_{jj} = \|\mathbf{y}\|$
- $\mathbf{q}_j = \frac{\mathbf{y}}{r_{jj}}$

end

Output: $\mathtt{Q}_{m\times m}$ and $\mathtt{R}_{m\times n}$ such that

$$\mathtt{A}_{m\times n} = \begin{bmatrix}\mathbf{q}_1 & \mathbf{q}_2 & \dots & \mathbf{q}_m\end{bmatrix}\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

$$= \mathtt{Q}_{m\times m}\mathtt{R}_{m\times n}$$

Least Squares by QR Factorization

Orthogonal Matrices

Def: A square matrix $\mathtt{Q}$ is orthogonal if $\mathtt{Q}^{-1} = \mathtt{Q}^\top$.

$$\mathtt{Q}^\top \mathtt{Q} = \mathtt{I} \implies \begin{bmatrix} \mathbf{q}_1^\top \\ \mathbf{q}_2^\top \\ \vdots \\ \mathbf{q}_m^\top \end{bmatrix} \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_m \end{bmatrix} = \begin{bmatrix} \mathbf{q}_1^\top \mathbf{q}_1 & \mathbf{q}_1^\top \mathbf{q}_2 & \cdots & \mathbf{q}_1^\top \mathbf{q}_m \\ \mathbf{q}_2^\top \mathbf{q}_1 & \mathbf{q}_2^\top \mathbf{q}_2 & \cdots & \mathbf{q}_2^\top \mathbf{q}_m \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{q}_n^\top \mathbf{q}_1 & \mathbf{q}_n^\top \mathbf{q}_2 & \cdots & \mathbf{q}_n^\top \mathbf{q}_m \end{bmatrix}$$

Def: A square matrix $\mathtt{Q}$ is **orthogonal** if $\mathtt{Q}^{-1} = \mathtt{Q}^\top$.

$$\mathtt{Q}^\top\mathtt{Q} = \mathtt{I} \implies \begin{bmatrix} \mathbf{q}_1^\top \\ \mathbf{q}_2^\top \\ \vdots \\ \mathbf{q}_m^\top \end{bmatrix} \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_m \end{bmatrix} = \begin{bmatrix} \mathbf{q}_1^\top\mathbf{q}_1 & \mathbf{q}_1^\top\mathbf{q}_2 & \cdots & \mathbf{q}_1^\top\mathbf{q}_m \\ \mathbf{q}_2^\top\mathbf{q}_1 & \mathbf{q}_2^\top\mathbf{q}_2 & \cdots & \mathbf{q}_2^\top\mathbf{q}_m \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{q}_n^\top\mathbf{q}_1 & \mathbf{q}_n^\top\mathbf{q}_2 & \cdots & \mathbf{q}_n^\top\mathbf{q}_m \end{bmatrix}$$

$$\implies \mathbf{q}_j^\top\mathbf{q}_j = 1 \text{ and } \mathbf{q}_i^\top\mathbf{q}_j = 0 \quad \forall i \neq j.$$

Def: A square matrix $\mathtt{Q}$ is orthogonal if $\mathtt{Q}^{-1} = \mathtt{Q}^\top$.

$$\mathtt{Q}^\top\mathtt{Q} = \mathtt{I} \implies \begin{bmatrix} \mathbf{q}_1^\top \\ \mathbf{q}_2^\top \\ \vdots \\ \mathbf{q}_m^\top \end{bmatrix} \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_m \end{bmatrix} = \begin{bmatrix} \mathbf{q}_1^\top\mathbf{q}_1 & \mathbf{q}_1^\top\mathbf{q}_2 & \cdots & \mathbf{q}_1^\top\mathbf{q}_m \\ \mathbf{q}_2^\top\mathbf{q}_1 & \mathbf{q}_2^\top\mathbf{q}_2 & \cdots & \mathbf{q}_2^\top\mathbf{q}_m \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{q}_n^\top\mathbf{q}_1 & \mathbf{q}_n^\top\mathbf{q}_2 & \cdots & \mathbf{q}_n^\top\mathbf{q}_m \end{bmatrix}$$

$$\implies \mathbf{q}_j^\top\mathbf{q}_j = 1 \text{ and } \mathbf{q}_i^\top\mathbf{q}_j = 0 \quad \forall i \neq j.$$

A set of vectors satisfying the above conditions is called an orthonormal set of vectors.

Lemma: If $\mathtt{Q}$ is an orthogonal $m \times m$ matrix and $\mathbf{x}$ is an m dimensional vector,

$$\|\mathtt{Q}\mathbf{x}\|^2$$

Lemma: If $\mathtt{Q}$ is an orthogonal $m \times m$ matrix and $\mathbf{x}$ is an m dimensional vector,

$$\|\mathtt{Q}\mathbf{x}\|^2 = (\mathtt{Q}\mathbf{x})^\top \mathtt{Q}\mathbf{x}$$

Lemma: If $\mathtt{Q}$ is an orthogonal $m \times m$ matrix and $\mathbf{x}$ is an m dimensional vector,

$$\|\mathtt{Q}\mathbf{x}\|^2 = (\mathtt{Q}\mathbf{x})^\top \mathtt{Q}\mathbf{x} = \mathbf{x}^\top \mathtt{Q}^\top \mathtt{Q}\mathbf{x}$$

Lemma: If $\mathtt{Q}$ is an orthogonal $m \times m$ matrix and $\mathbf{x}$ is an m dimensional vector,

$$\|\mathtt{Q}\mathbf{x}\|^2 = (\mathtt{Q}\mathbf{x})^\top \mathtt{Q}\mathbf{x} = \mathbf{x}^\top \mathtt{Q}^\top \mathtt{Q}\mathbf{x} = \mathbf{x}^\top \mathbf{x} = \|\mathbf{x}\|^2 .$$

Lemma: If $\mathtt{Q}$ is an orthogonal $m \times m$ matrix and $\mathbf{x}$ is an m dimensional vector,

$$\|\mathtt{Q}\mathbf{x}\|^2 = (\mathtt{Q}\mathbf{x})^\top \mathtt{Q}\mathbf{x} = \mathbf{x}^\top \mathtt{Q}^\top \mathtt{Q}\mathbf{x} = \mathbf{x}^\top \mathbf{x} = \|\mathbf{x}\|^2 .$$

Lemma: If $\mathtt{Q}_1$ and $\mathtt{Q}_2$ are orthogonal then the result of the matrix multiplications $\mathtt{Q}_1\mathtt{Q}_2$ and $\mathtt{Q}_2\mathtt{Q}_1$ are also orthogonal.

Least Squares by QR Factorization

Given $m \times n$ inconsistent system $\mathtt{A}\mathbf{x} = \mathbf{b}$, find $\mathtt{A} = \mathtt{QR}$, the full QR factorization of $\mathtt{A}$.

$$\|\mathbf{r}\| = \|\mathtt{A}\mathbf{x} - \mathbf{b}\| = \|\mathtt{QR}\mathbf{x} - \mathbf{b}\|$$

Least Squares by QR Factorization

Given $m \times n$ inconsistent system $\mathtt{A}\mathbf{x} = \mathbf{b}$, find $\mathtt{A} = \mathtt{QR}$, the full QR factorization of $\mathtt{A}$.

$$\|\mathbf{r}\| = \|\mathtt{A}\mathbf{x} - \mathbf{b}\| = \|\mathtt{QR}\mathbf{x} - \mathbf{b}\| = \left\|\mathtt{Q}^\top \mathtt{QR}\mathbf{x} - \mathtt{Q}^\top \mathbf{b}\right\|$$

Least Squares by QR Factorization

Given $m \times n$ inconsistent system $\mathtt{A}\mathbf{x} = \mathbf{b}$, find $\mathtt{A} = \mathtt{QR}$, the full QR factorization of $\mathtt{A}$.

$$\|\mathbf{r}\| = \|\mathtt{A}\mathbf{x} - \mathbf{b}\| = \|\mathtt{QR}\mathbf{x} - \mathbf{b}\| = \left\|\mathtt{Q}^\top \mathtt{QR}\mathbf{x} - \mathtt{Q}^\top \mathbf{b}\right\| = \left\|\mathtt{R}\mathbf{x} - \underbrace{\mathtt{Q}^\top \mathbf{b}}_{\mathbf{d}}\right\|$$

Least Squares by QR Factorization

Given $m \times n$ inconsistent system $\mathtt{A}\mathbf{x} = \mathbf{b}$, find $\mathtt{A} = \mathtt{QR}$, the full QR factorization of $\mathtt{A}$.

$$\|\mathbf{r}\| = \|\mathtt{A}\mathbf{x} - \mathbf{b}\| = \|\mathtt{QR}\mathbf{x} - \mathbf{b}\| = \left\|\mathtt{Q}^\top \mathtt{QR}\mathbf{x} - \mathtt{Q}^\top \mathbf{b}\right\| = \left\|\mathtt{R}\mathbf{x} - \underbrace{\mathtt{Q}^\top \mathbf{b}}_{\mathbf{d}}\right\|$$

$$\|\mathtt{A}\mathbf{x} - \mathbf{b}\| = \underbrace{\begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \\ e_{n+1} \\ \vdots \\ e_m \end{bmatrix}}_{\begin{bmatrix} \hat{\mathbf{e}} \\ \mathbf{e}' \end{bmatrix}}$$

Given $m \times n$ inconsistent system $\mathtt{A}\mathbf{x} = \mathbf{b}$, find $\mathtt{A} = \mathtt{QR}$, the full QR factorization of $\mathtt{A}$.

$$\|\mathbf{r}\| = \|\mathtt{A}\mathbf{x} - \mathbf{b}\| = \|\mathtt{QR}\mathbf{x} - \mathbf{b}\| = \left\|\mathtt{Q}^\top \mathtt{QR}\mathbf{x} - \mathtt{Q}^\top \mathbf{b}\right\| = \left\|\mathtt{R}\mathbf{x} - \underbrace{\mathtt{Q}^\top \mathbf{b}}_{\mathbf{d}}\right\|$$

$$\|\mathtt{A}\mathbf{x} - \mathbf{b}\| = \underbrace{\begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \\ e_{n+1} \\ \vdots \\ e_m \end{bmatrix}}_{\begin{bmatrix} \hat{\mathbf{e}} \\ \mathbf{e}' \end{bmatrix}} = \underbrace{\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}}_{\begin{bmatrix} \hat{\mathtt{R}} \\ 0 \end{bmatrix}} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} - \underbrace{\begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \\ d_{n+1} \\ \vdots \\ d_m \end{bmatrix}}_{\begin{bmatrix} \hat{\mathbf{d}} \\ \mathbf{d}' \end{bmatrix}}$$

Given $m \times n$ inconsistent system $\mathtt{A}\mathbf{x} = \mathbf{b}$, find $\mathtt{A} = \mathtt{QR}$, the full QR factorization of $\mathtt{A}$.

$$\|\mathbf{r}\| = \|\mathtt{A}\mathbf{x} - \mathbf{b}\| = \|\mathtt{QR}\mathbf{x} - \mathbf{b}\| = \left\|\mathtt{Q}^\top\mathtt{QR}\mathbf{x} - \mathtt{Q}^\top\mathbf{b}\right\| = \left\|\mathtt{R}\mathbf{x} - \underbrace{\mathtt{Q}^\top\mathbf{b}}_{\mathbf{d}}\right\|$$

$$\|\mathtt{A}\mathbf{x} - \mathbf{b}\| = \underbrace{\begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \\ e_{n+1} \\ \vdots \\ e_m \end{bmatrix}}_{\begin{bmatrix} \hat{\mathbf{e}} \\ \mathbf{e}' \end{bmatrix}} = \underbrace{\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}}_{\begin{bmatrix} \hat{\mathtt{R}} \\ 0 \end{bmatrix}} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} - \underbrace{\begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \\ d_{n+1} \\ \vdots \\ d_m \end{bmatrix}}_{\begin{bmatrix} \hat{\mathbf{d}} \\ \mathbf{d}' \end{bmatrix}} = \underbrace{\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ -d_{n+1} \\ \vdots \\ -d_m \end{bmatrix}}_{\substack{\text{if } \mathbf{x} = \mathbf{x}_{\text{LS}} \\ \hat{\mathtt{R}}\mathbf{x}_{\text{LS}} = \hat{\mathbf{d}}}}$$

Example

Use QR factorization to solve $\begin{bmatrix} 1 & -4 \\ 2 & 3 \\ 2 & 2 \end{bmatrix} \mathbf{x} = \begin{bmatrix} -3 \\ 15 \\ 9 \end{bmatrix}$.

Example

Use QR factorization to solve $\begin{bmatrix} 1 & -4 \\ 2 & 3 \\ 2 & 2 \end{bmatrix} \mathbf{x} = \begin{bmatrix} -3 \\ 15 \\ 9 \end{bmatrix}$.

$$\mathtt{Q}^\top = \frac{1}{15} \begin{bmatrix} 5 & 10 & 10 \\ -14 & 5 & 2 \\ 2 & 10 & -11 \end{bmatrix}, \quad \hat{\mathtt{R}} = \begin{bmatrix} 3 & 2 \\ 0 & 5 \end{bmatrix}$$

Use QR factorization to solve $\begin{bmatrix} 1 & -4 \\ 2 & 3 \\ 2 & 2 \end{bmatrix} \mathbf{x} = \begin{bmatrix} -3 \\ 15 \\ 9 \end{bmatrix}$.

$$\mathtt{Q}^\top = \frac{1}{15} \begin{bmatrix} 5 & 10 & 10 \\ -14 & 5 & 2 \\ 2 & 10 & -11 \end{bmatrix}, \quad \hat{\mathtt{R}} = \begin{bmatrix} 3 & 2 \\ 0 & 5 \end{bmatrix}$$

$$\mathbf{d} = \mathtt{Q}^\top \mathbf{b} = \begin{bmatrix} 15 \\ 9 \\ 3 \end{bmatrix}, \quad \text{Solve } \begin{bmatrix} 3 & 2 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 15 \\ 9 \end{bmatrix} \implies \begin{aligned} x_1 &= 3.8 \\ x_2 &= 1.8 \end{aligned}$$

Example

Use QR factorization to solve $\begin{bmatrix} 1 & -4 \\ 2 & 3 \\ 2 & 2 \end{bmatrix} \mathbf{x} = \begin{bmatrix} -3 \\ 15 \\ 9 \end{bmatrix}$.

$$\mathtt{Q}^\top = \frac{1}{15} \begin{bmatrix} 5 & 10 & 10 \\ -14 & 5 & 2 \\ 2 & 10 & -11 \end{bmatrix}, \quad \hat{\mathtt{R}} = \begin{bmatrix} 3 & 2 \\ 0 & 5 \end{bmatrix}$$

$$\mathbf{d} = \mathtt{Q}^\top \mathbf{b} = \begin{bmatrix} 15 \\ 9 \\ 3 \end{bmatrix}, \quad \text{Solve } \begin{bmatrix} 3 & 2 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 15 \\ 9 \end{bmatrix} \implies \begin{matrix} x_1 = 3.8 \\ x_2 = 1.8 \end{matrix}$$

$$\mathbf{r} = \begin{bmatrix} 1 & -4 \\ 2 & 3 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 3.8 \\ 1.8 \end{bmatrix} - \begin{bmatrix} -3 \\ 15 \\ 9 \end{bmatrix} = \begin{bmatrix} -0.4 \\ -2 \\ 2.2 \end{bmatrix},$$

Use QR factorization to solve $\begin{bmatrix} 1 & -4 \\ 2 & 3 \\ 2 & 2 \end{bmatrix} \mathbf{x} = \begin{bmatrix} -3 \\ 15 \\ 9 \end{bmatrix}$.

$$\mathtt{Q}^\top = \frac{1}{15} \begin{bmatrix} 5 & 10 & 10 \\ -14 & 5 & 2 \\ 2 & 10 & -11 \end{bmatrix}, \quad \hat{\mathtt{R}} = \begin{bmatrix} 3 & 2 \\ 0 & 5 \end{bmatrix}$$

$$\mathbf{d} = \mathtt{Q}^\top \mathbf{b} = \begin{bmatrix} 15 \\ 9 \\ 3 \end{bmatrix}, \quad \text{Solve } \begin{bmatrix} 3 & 2 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 15 \\ 9 \end{bmatrix} \implies \begin{aligned} x_1 &= 3.8 \\ x_2 &= 1.8 \end{aligned}$$

$$\mathbf{r} = \begin{bmatrix} 1 & -4 \\ 2 & 3 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 3.8 \\ 1.8 \end{bmatrix} - \begin{bmatrix} -3 \\ 15 \\ 9 \end{bmatrix} = \begin{bmatrix} -0.4 \\ -2 \\ 2.2 \end{bmatrix},$$

$$\|\mathbf{r}\|^2 = (-0.4)^2 + (-2)^2 + 2.2^2 = 9$$

Use QR factorization to solve $\begin{bmatrix} 1 & -4 \\ 2 & 3 \\ 2 & 2 \end{bmatrix} \mathbf{x} = \begin{bmatrix} -3 \\ 15 \\ 9 \end{bmatrix}$.

$$\mathtt{Q}^\top = \frac{1}{15}\begin{bmatrix} 5 & 10 & 10 \\ -14 & 5 & 2 \\ 2 & 10 & -11 \end{bmatrix}, \quad \hat{\mathtt{R}} = \begin{bmatrix} 3 & 2 \\ 0 & 5 \end{bmatrix}$$

$$\mathbf{d} = \mathtt{Q}^\top \mathbf{b} = \begin{bmatrix} 15 \\ 9 \\ 3 \end{bmatrix}, \quad \text{Solve } \begin{bmatrix} 3 & 2 \\ 0 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 15 \\ 9 \end{bmatrix} \implies \begin{aligned} x_1 &= 3.8 \\ x_2 &= 1.8 \end{aligned}$$

$$\mathbf{r} = \begin{bmatrix} 1 & -4 \\ 2 & 3 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 3.8 \\ 1.8 \end{bmatrix} - \begin{bmatrix} -3 \\ 15 \\ 9 \end{bmatrix} = \begin{bmatrix} -0.4 \\ -2 \\ 2.2 \end{bmatrix},$$

$$\|\mathbf{r}\|^2 = (-0.4)^2 + (-2)^2 + 2.2^2 = 9 = 3^2.$$

Householder Reflectors

An Elimination like QR Factorization Approach

- Even modified Gram-Schmidt is not numerically perfect.

An Elimination like QR Factorization Approach

- Even modified Gram-Schmidt is not numerically perfect.
- We look for new algorithms that works much like LU factorization.

An Elimination like QR Factorization Approach

- Even modified Gram-Schmidt is not numerically perfect.
- We look for new algorithms that works much like LU factorization.
- If we can find a set of orthogonal transformations, such that $\mathtt{Q}_n^\top \cdots \mathtt{Q}_2^\top \mathtt{Q}_1^\top \mathtt{A} = \mathtt{R}$, then we can have $\mathtt{Q} = \mathtt{Q}_1 \mathtt{Q}_2 \cdots \mathtt{Q}_n$.

An Elimination like QR Factorization Approach

- Even modified Gram-Schmidt is not numerically perfect.
- We look for new algorithms that works much like LU factorization.
- If we can find a set of orthogonal transformations, such that $\mathtt{Q}_n^\top \cdots \mathtt{Q}_2^\top \mathtt{Q}_1^\top \mathtt{A} = \mathtt{R}$, then we can have $\mathtt{Q} = \mathtt{Q}_1 \mathtt{Q}_2 \cdots \mathtt{Q}_n$.
- Each orthogonal matrix $\mathtt{Q}_i^\top$ should transform the i^{th} column of $\mathtt{A}$ by creating zeros below the diagonal position so that $\mathtt{R}$ is upper triangular.

An Elimination like QR Factorization Approach

- Even modified Gram-Schmidt is not numerically perfect.
- We look for new algorithms that works much like LU factorization.
- If we can find a set of orthogonal transformations, such that $\mathtt{Q}_n^\top \cdots \mathtt{Q}_2^\top \mathtt{Q}_1^\top \mathtt{A} = \mathtt{R}$, then we can have $\mathtt{Q} = \mathtt{Q}_1 \mathtt{Q}_2 \cdots \mathtt{Q}_n$.
- Each orthogonal matrix $\mathtt{Q}_i^\top$ should transform the i^{th} column of $\mathtt{A}$ by creating zeros below the diagonal position so that $\mathtt{R}$ is upper triangular.
- Orthogonal matrices correspond to either rotations or reflections, so there are two possible algorithms, one using Givens rotations, the other using Householder reflections.

An Elimination like QR Factorization Approach

- Even modified Gram-Schmidt is not numerically perfect.
- We look for new algorithms that works much like LU factorization.
- If we can find a set of orthogonal transformations, such that $\mathtt{Q}_n^\top \cdots \mathtt{Q}_2^\top \mathtt{Q}_1^\top \mathtt{A} = \mathtt{R}$, then we can have $\mathtt{Q} = \mathtt{Q}_1 \mathtt{Q}_2 \cdots \mathtt{Q}_n$.
- Each orthogonal matrix $\mathtt{Q}_i^\top$ should transform the i^{th} column of $\mathtt{A}$ by creating zeros below the diagonal position so that $\mathtt{R}$ is upper triangular.
- Orthogonal matrices correspond to either rotations or reflections, so there are two possible algorithms, one using Givens rotations, the other using Householder reflections.
- We will cover Householder reflections, see the book for Givens rotations in case you are curious.

Look for an orthogonal matrix $\mathtt{H}$ that reflects $\mathbf{x}$ to $\mathbf{w}$ by $\mathtt{H}\mathbf{x} = \mathbf{w}$.

Look for an orthogonal matrix $\mathtt{H}$ that reflects $\mathbf{x}$ to $\mathbf{w}$ by $\mathtt{H}\mathbf{x} = \mathbf{w}$.

Lemma: Assume that $\mathbf{x}$ and $\mathbf{w}$ are of the same length. Then $\mathbf{w} - \mathbf{x}$ and $\mathbf{w} + \mathbf{x}$ are perpendicular.

Look for an orthogonal matrix $\mathtt{H}$ that reflects $\mathbf{x}$ to $\mathbf{w}$ by $\mathtt{H}\mathbf{x} = \mathbf{w}$.

Lemma: Assume that $\mathbf{x}$ and $\mathbf{w}$ are of the same length. Then $\mathbf{w} - \mathbf{x}$ and $\mathbf{w} + \mathbf{x}$ are perpendicular.

$$(\mathbf{w} - \mathbf{x})^\top(\mathbf{w} + \mathbf{x}) = \mathbf{w}^\top\mathbf{w} + \mathbf{w}^\top\mathbf{x} - \mathbf{x}^\top\mathbf{w} - \mathbf{x}^\top\mathbf{x} = \|\mathbf{w}\|^2 - \|\mathbf{x}\|^2 = 0.$$

Look for an orthogonal matrix $\mathtt{H}$ that reflects $\mathbf{x}$ to $\mathbf{w}$ by $\mathtt{H}\mathbf{x} = \mathbf{w}$.

Lemma: Assume that $\mathbf{x}$ and $\mathbf{w}$ are of the same length. Then $\mathbf{w} - \mathbf{x}$ and $\mathbf{w} + \mathbf{x}$ are perpendicular.

$$(\mathbf{w} - \mathbf{x})^\top(\mathbf{w} + \mathbf{x}) = \mathbf{w}^\top\mathbf{w} + \mathbf{w}^\top\mathbf{x} - \mathbf{x}^\top\mathbf{w} - \mathbf{x}^\top\mathbf{x} = \|\mathbf{w}\|^2 - \|\mathbf{x}\|^2 = 0.$$

Def: Let $\mathtt{P} = \frac{\mathbf{v}\mathbf{v}^\top}{\mathbf{v}^\top\mathbf{v}}$ then

$$\mathtt{P}^2 = \frac{\mathbf{v}\mathbf{v}^\top\mathbf{v}\mathbf{v}^\top}{(\mathbf{v}^\top\mathbf{v})^2} = \frac{\mathbf{v}^\top\mathbf{v}}{\mathbf{v}^\top\mathbf{v}}\frac{\mathbf{v}\mathbf{v}^\top}{\mathbf{v}^\top\mathbf{v}} = \mathtt{P} = \mathtt{P}^\top.$$

Look for an orthogonal matrix $\mathtt{H}$ that reflects $\mathbf{x}$ to $\mathbf{w}$ by $\mathtt{H}\mathbf{x} = \mathbf{w}$.

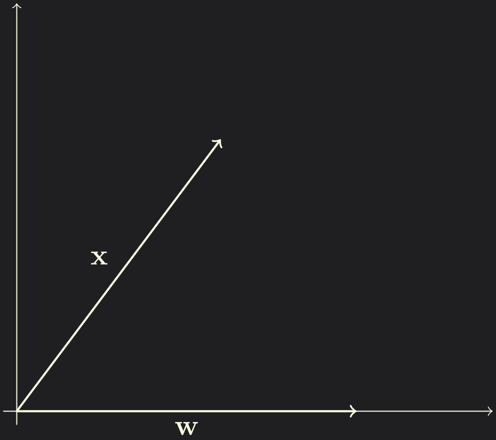
Lemma: Assume that $\mathbf{x}$ and $\mathbf{w}$ are of the same length. Then $\mathbf{w} - \mathbf{x}$ and $\mathbf{w} + \mathbf{x}$ are perpendicular.

$$(\mathbf{w} - \mathbf{x})^\top (\mathbf{w} + \mathbf{x}) = \mathbf{w}^\top \mathbf{w} + \mathbf{w}^\top \mathbf{x} - \mathbf{x}^\top \mathbf{w} - \mathbf{x}^\top \mathbf{x} = \|\mathbf{w}\|^2 - \|\mathbf{x}\|^2 = 0.$$

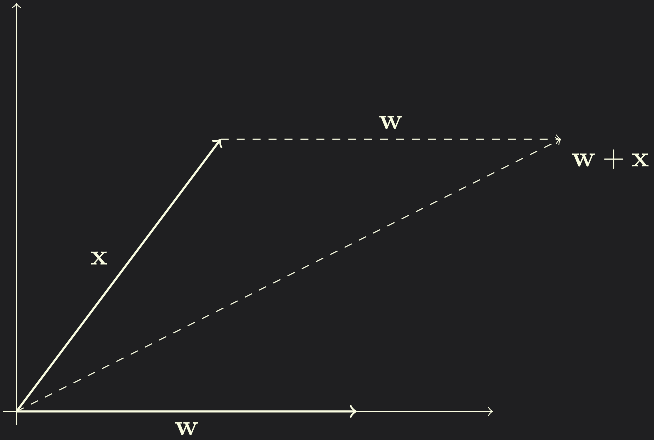
Def: Let $\mathtt{P} = \frac{\mathbf{v}\mathbf{v}^\top}{\mathbf{v}^\top \mathbf{v}}$ then

$$\mathtt{P}^2 = \frac{\mathbf{v}\mathbf{v}^\top \mathbf{v}\mathbf{v}^\top}{(\mathbf{v}^\top \mathbf{v})^2} = \frac{\mathbf{v}^\top \mathbf{v}}{\mathbf{v}^\top \mathbf{v}} \frac{\mathbf{v}\mathbf{v}^\top}{\mathbf{v}^\top \mathbf{v}} = \mathtt{P} = \mathtt{P}^\top.$$

$\mathtt{P}$ is a symmetric projection matrix with $\mathtt{P}\mathbf{v} = \mathbf{v}$ and $\mathtt{P}\mathbf{u}$ is the projection of $\mathbf{u}$ onto $\mathbf{v}$.

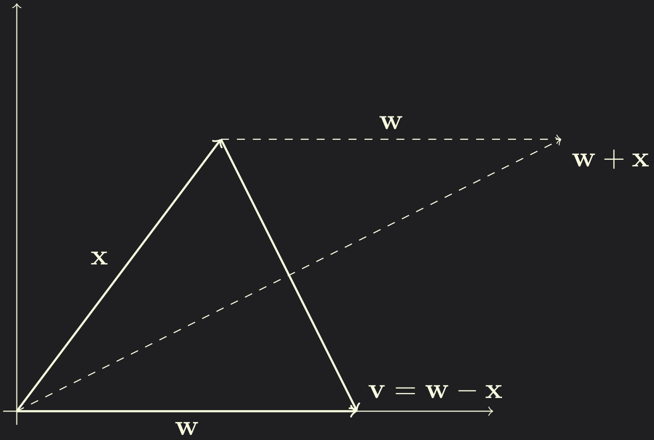


$$\|\mathbf{x}\| = \|\mathbf{w}\|$$



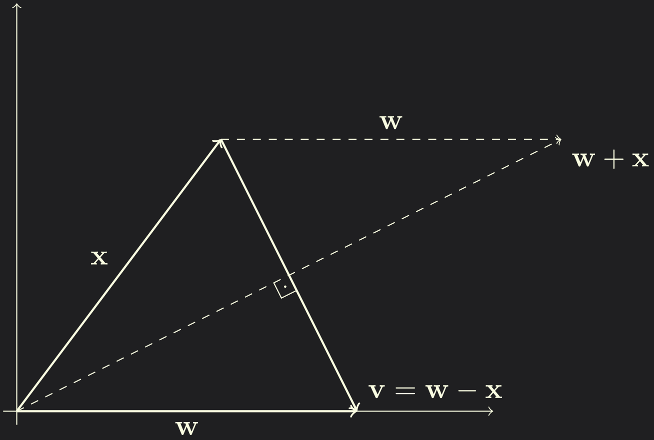
$$\|\mathbf{x}\| = \|\mathbf{w}\|$$

Householder Reflections: The Geometry



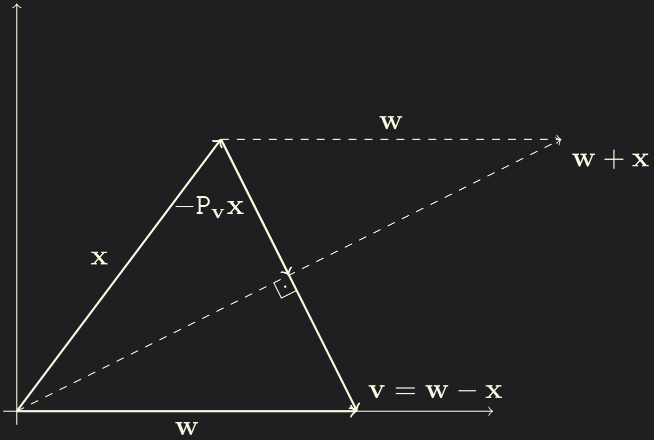
$$\|\mathbf{x}\| = \|\mathbf{w}\|$$

Householder Reflections: The Geometry



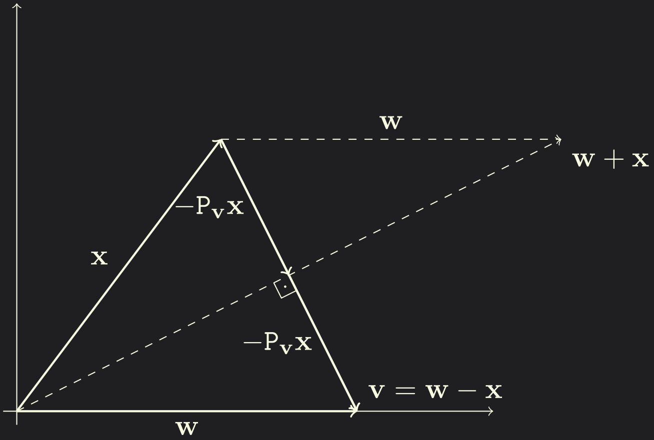
$$\|\mathbf{x}\| = \|\mathbf{w}\|$$

Householder Reflections: The Geometry



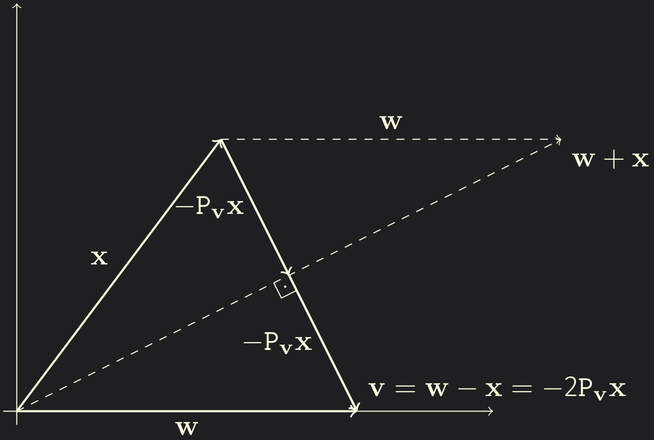
$$\|\mathbf{x}\| = \|\mathbf{w}\|$$

Householder Reflections: The Geometry



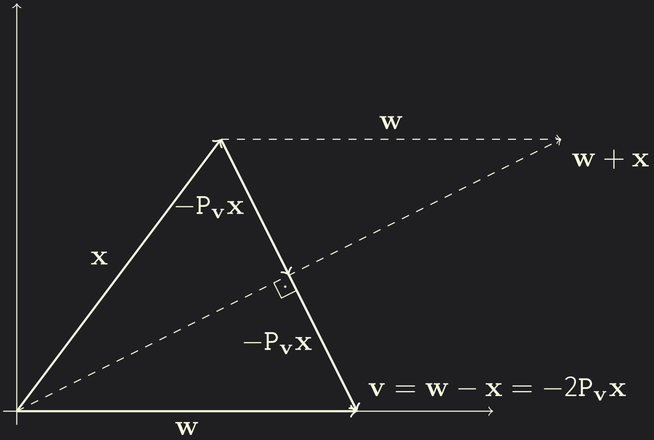
$$\|\mathbf{x}\| = \|\mathbf{w}\|$$

Householder Reflections: The Geometry



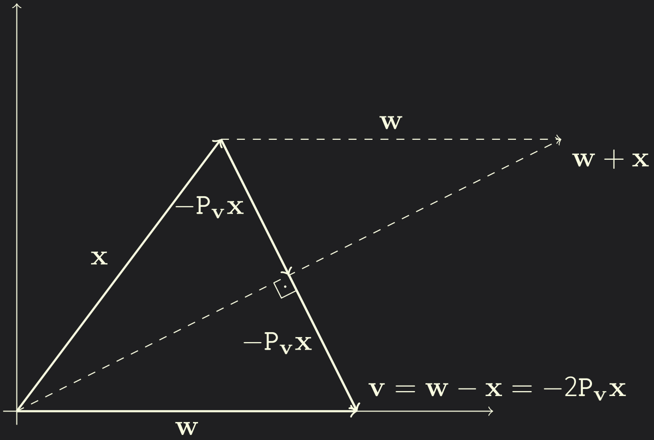
$$\|\mathbf{x}\| = \|\mathbf{w}\|$$

Householder Reflections: The Geometry



$$\|\mathbf{x}\| = \|\mathbf{w}\|, \quad \mathbf{w} = \mathbf{x} - 2\mathrm{P}_{\mathbf{v}}\mathbf{x} = (\mathrm{I} - 2\mathrm{P}_{\mathbf{v}})\mathbf{x} = \mathtt{H}\mathbf{x}$$

Householder Reflections: The Geometry



$$\|\mathbf{x}\| = \|\mathbf{w}\|, \quad \mathbf{w} = \mathbf{x} - 2\mathrm{P}_{\mathbf{v}}\mathbf{x} = (\mathtt{I} - 2\mathrm{P}_{\mathbf{v}})\mathbf{x} = \mathtt{H}\mathbf{x}$$

$$\begin{aligned}\mathtt{H}^\top\mathtt{H} &= (\mathtt{I} - 2\mathrm{P}_{\mathbf{v}})^\top(\mathtt{I} - 2\mathrm{P}_{\mathbf{v}}) = (\mathtt{I} - 2\mathrm{P}_{\mathbf{v}})(\mathtt{I} - 2\mathrm{P}_{\mathbf{v}})\\ &= \mathtt{I} - 2\mathrm{P}_{\mathbf{v}} - 2\mathrm{P}_{\mathbf{v}} + 4\mathrm{P}_{\mathbf{v}}^2 = \mathtt{I}.\end{aligned}$$

Example

Let $\mathbf{x} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$, find $\mathtt{H}$ s.t. $\mathtt{H}\mathbf{x} = \mathbf{w}$.

$$\mathbf{v} = \mathbf{w} - \mathbf{x} = \begin{bmatrix} 5 \\ 0 \end{bmatrix} - \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ -4 \end{bmatrix},$$

Let $\mathbf{x} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$, find $\mathtt{H}$ s.t. $\mathtt{H}\mathbf{x} = \mathbf{w}$.

$$\mathbf{v} = \mathbf{w} - \mathbf{x} = \begin{bmatrix} 5 \\ 0 \end{bmatrix} - \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ -4 \end{bmatrix},$$

$$\mathrm{P}_{\mathbf{v}} = \frac{\mathbf{v}\mathbf{v}^\top}{\mathbf{v}^\top\mathbf{v}} = \frac{\begin{bmatrix} 4 & -8 \\ -8 & 16 \end{bmatrix}}{20} = \begin{bmatrix} 0.2 & -0.4 \\ -0.4 & 0.8 \end{bmatrix},$$

Example

Let $\mathbf{x} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$, find $\mathtt{H}$ s.t. $\mathtt{H}\mathbf{x} = \mathbf{w}$.

$$\mathbf{v} = \mathbf{w} - \mathbf{x} = \begin{bmatrix} 5 \\ 0 \end{bmatrix} - \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ -4 \end{bmatrix},$$

$$\mathtt{P}_{\mathbf{v}} = \frac{\mathbf{v}\mathbf{v}^\top}{\mathbf{v}^\top\mathbf{v}} = \frac{\begin{bmatrix} 4 & -8 \\ -8 & 16 \end{bmatrix}}{20} = \begin{bmatrix} 0.2 & -0.4 \\ -0.4 & 0.8 \end{bmatrix},$$

$$\mathtt{H} = \mathtt{I} - 2\mathtt{P}_{\mathbf{v}} = \begin{bmatrix} 0.6 & 0.8 \\ 0.8 & -0.6 \end{bmatrix},$$

Let $\mathbf{x} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$, find $\mathtt{H}$ s.t. $\mathtt{H}\mathbf{x} = \mathbf{w}$.

$$\mathbf{v} = \mathbf{w} - \mathbf{x} = \begin{bmatrix} 5 \\ 0 \end{bmatrix} - \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ -4 \end{bmatrix},$$

$$\mathrm{P}_{\mathbf{v}} = \frac{\mathbf{v}\mathbf{v}^\top}{\mathbf{v}^\top\mathbf{v}} = \frac{\begin{bmatrix} 4 & -8 \\ -8 & 16 \end{bmatrix}}{20} = \begin{bmatrix} 0.2 & -0.4 \\ -0.4 & 0.8 \end{bmatrix},$$

$$\mathtt{H} = \mathtt{I} - 2\mathrm{P}_{\mathbf{v}} = \begin{bmatrix} 0.6 & 0.8 \\ 0.8 & -0.6 \end{bmatrix},$$

$$\mathtt{H}\mathbf{x} = \begin{bmatrix} 1.8 + 3.2 \\ 2.4 - 2.4 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \end{bmatrix} = \mathbf{w}.$$

QR Factorization by Householder Reflections

Given $\mathtt{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots, \mathbf{a}_n \end{bmatrix}$, take $\mathbf{x}_1 = \mathbf{a}_1$ and $\mathbf{w} = \pm \begin{bmatrix} \|\mathbf{x}_1\| \\ 0 \\ \vdots \\ 0 \end{bmatrix}$.

Given $\mathtt{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots, \mathbf{a}_n \end{bmatrix}$, take $\mathbf{x}_1 = \mathbf{a}_1$ and $\mathbf{w} = \pm \begin{bmatrix} \|\mathbf{x}_1\| \\ 0 \\ \vdots \\ 0 \end{bmatrix}$.

Take the sign to be equal to the opposite sign of the first component of $\mathbf{x}_1$ for numerical stability.

Given $\mathtt{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots, \mathbf{a}_n \end{bmatrix}$, take $\mathbf{x}_1 = \mathbf{a}_1$ and $\mathbf{w} = \pm \begin{bmatrix} \|\mathbf{x}_1\| \\ 0 \\ \vdots \\ 0 \end{bmatrix}$.

Take the sign to be equal to the opposite sign of the first component of $\mathbf{x}_1$ for numerical stability.

$$\mathtt{H}_1\mathtt{A} = \mathtt{H}_1 \begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix} = \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \end{bmatrix}$$

Given $\mathtt{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots, \mathbf{a}_n \end{bmatrix}$, take $\mathbf{x}_1 = \mathbf{a}_1$ and $\mathbf{w} = \pm \begin{bmatrix} \|\mathbf{x}_1\| \\ 0 \\ \vdots \\ 0 \end{bmatrix}$.

Take the sign to be equal to the opposite sign of the first component of $\mathbf{x}_1$ for numerical stability.

$$\mathtt{H}_1\mathtt{A} = \mathtt{H}_1 \begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix} = \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \end{bmatrix}$$

$$\mathtt{H}_2\mathtt{H}_1\mathtt{A} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & \multicolumn{2}{c}{\hat{\mathtt{H}}_2} \\ 0 & & \end{bmatrix}}_{\mathtt{H}_2} \underbrace{\begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \end{bmatrix}}_{\mathtt{H}_1\mathtt{A}} = \underbrace{\begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & 0 & \times \end{bmatrix}}_{\mathtt{R}}$$

Given $\mathtt{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots, \mathbf{a}_n \end{bmatrix}$, take $\mathbf{x}_1 = \mathbf{a}_1$ and $\mathbf{w} = \pm \begin{bmatrix} \|\mathbf{x}_1\| \\ 0 \\ \vdots \\ 0 \end{bmatrix}$.

Take the sign to be equal to the opposite sign of the first component of $\mathbf{x}_1$ for numerical stability.

$$\mathtt{H}_1\mathtt{A} = \mathtt{H}_1 \begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix} = \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \end{bmatrix}$$

$$\mathtt{H}_2\mathtt{H}_1\mathtt{A} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & \multicolumn{2}{c}{\hat{\mathtt{H}}_2} \\ 0 & & \end{bmatrix}}_{\mathtt{H}_2} \underbrace{\begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \end{bmatrix}}_{\mathtt{H}_1\mathtt{A}} = \underbrace{\begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & 0 & \times \end{bmatrix}}_{\mathtt{R}}$$

$$\mathtt{A} = \mathtt{H}_1\mathtt{H}_2\mathtt{R} \quad (\mathtt{H}_i^{-1} = \mathtt{H}_i)$$

Example

Find the QR factorization of $\mathtt{A} = \begin{bmatrix} 1 & -4 \\ 2 & 3 \\ 2 & 2 \end{bmatrix}$.

Example

Find the QR factorization of $\mathtt{A} = \begin{bmatrix} 1 & -4 \\ 2 & 3 \\ 2 & 2 \end{bmatrix}$.

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \mathbf{w}_1 = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 2 \\ -2 \\ -2 \end{bmatrix}, \mathtt{H}_1 = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

Example

Find the QR factorization of $\mathtt{A} = \begin{bmatrix} 1 & -4 \\ 2 & 3 \\ 2 & 2 \end{bmatrix}$.

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \mathbf{w}_1 = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 2 \\ -2 \\ -2 \end{bmatrix}, \mathtt{H}_1 = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

$$\mathtt{H}_1\mathtt{A} = \begin{bmatrix} 3 & 2 \\ 0 & -3 \\ 0 & -4 \end{bmatrix}, \hat{\mathbf{x}}_2 = \begin{bmatrix} -3 \\ -4 \end{bmatrix}, \hat{\mathbf{w}}_2 = \begin{bmatrix} 5 \\ 0 \end{bmatrix}, \mathtt{H}_2 = \begin{bmatrix} -0.6 & -0.8 \\ -0.8 & 0.6 \end{bmatrix}$$

Example

Find the QR factorization of $\mathtt{A} = \begin{bmatrix} 1 & -4 \\ 2 & 3 \\ 2 & 2 \end{bmatrix}$.

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \mathbf{w}_1 = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 2 \\ -2 \\ -2 \end{bmatrix}, \mathtt{H}_1 = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

$$\mathtt{H}_1\mathtt{A} = \begin{bmatrix} 3 & 2 \\ 0 & -3 \\ 0 & -4 \end{bmatrix}, \hat{\mathbf{x}}_2 = \begin{bmatrix} -3 \\ -4 \end{bmatrix}, \hat{\mathbf{w}}_2 = \begin{bmatrix} 5 \\ 0 \end{bmatrix}, \mathtt{H}_2 = \begin{bmatrix} -0.6 & -0.8 \\ -0.8 & 0.6 \end{bmatrix}$$

$$\mathtt{H}_2\mathtt{H}_1\mathtt{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -0.6 & -0.8 \\ 0 & -0.8 & 0.6 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 0 & -3 \\ 0 & -4 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 0 & 5 \\ 0 & 0 \end{bmatrix} = \mathtt{R},$$

Example

Find the QR factorization of $\mathtt{A} = \begin{bmatrix} 1 & -4 \\ 2 & 3 \\ 2 & 2 \end{bmatrix}$.

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \mathbf{w}_1 = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 2 \\ -2 \\ -2 \end{bmatrix}, \mathtt{H}_1 = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

$$\mathtt{H}_1\mathtt{A} = \begin{bmatrix} 3 & 2 \\ 0 & -3 \\ 0 & -4 \end{bmatrix}, \hat{\mathbf{x}}_2 = \begin{bmatrix} -3 \\ -4 \end{bmatrix}, \hat{\mathbf{w}}_2 = \begin{bmatrix} 5 \\ 0 \end{bmatrix}, \mathtt{H}_2 = \begin{bmatrix} -0.6 & -0.8 \\ -0.8 & 0.6 \end{bmatrix}$$

$$\mathtt{H}_2\mathtt{H}_1\mathtt{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -0.6 & -0.8 \\ 0 & -0.8 & 0.6 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 0 & -3 \\ 0 & -4 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 0 & 5 \\ 0 & 0 \end{bmatrix} = \mathtt{R},$$

$$\mathtt{Q} = \mathtt{H}_1\mathtt{H}_2 = \frac{1}{15} \begin{bmatrix} 5 & -14 & -2 \\ 10 & 5 & -10 \\ 10 & 2 & 11 \end{bmatrix}$$

QR factorization is not unique since if $\mathtt{D} = \mathrm{diag}(d_1, \ldots, d_m)$ with $d_i = \pm 1$ then

$$\mathtt{A} = \mathtt{QR} = \underbrace{\mathtt{QD}}_{\mathtt{Q}'}\underbrace{\mathtt{DR}}_{\mathtt{R}'}$$

since

$$\mathtt{D}^2 = \mathtt{I}$$