1. **Householder reflector**
   - Let $v \in \mathbb{C}^n$ and $v \neq 0$. Then the matrix
     \[ H = I - 2 \frac{vv^*}{v^*v}, \]
     is called a *Householder reflector*.
   - Geometric interpretation
Exercise: Prove that any Householder reflector $H$ satisfies the following properties:
(1) It is hermitian: $H = H^*$
(2) It is unitary: $H^* = H^{-1}$
(3) It is involutary: $H^2 = I$

Exercise: What are the eigenvalues, the determinant, and the singular values of a Householder reflector $H$?
Hint: by definition, consider eigenvalues and eigenvectors of $H$.

Exercise: Prove that $I - \frac{vv^*}{v^*v}$ is the orthogonal projector which projects $\mathbb{C}^m$ onto the hyperplane span$\{v\}^\perp$. 
Theorem 1

For all \( x, y \in \mathbb{C}^m \) with \( x \neq y \), there exists a Householder reflector \( H \) such that \( Hx = y \) if and only if \( \|x\|_2 = \|y\|_2 \) and \( x^*y \in \mathbb{R} \).

Proof.

“⇒” is easy. “⇐”: let \( v = y - x \), verify \( Hx = y \).

Corollary 2

For all nonzero \( x, y \in \mathbb{C}^m \) with \( x \neq y \), there exists a Householder reflector \( H \) and \( z \in \mathbb{C} \) such that \( Hx = zy \).

Proof.

Let
\[
z = \frac{\|x\|_2}{\|y\|_2} \cdot c,
\]
\[
c = \begin{cases} 
\pm \frac{y^*x}{\|x^*y\|}, & \text{if } x^*y \neq 0, \\
e^{i\theta}, & \text{if } x^*y = 0,
\end{cases}
\]

and \( v = zy - x \). Verify \( Hx = zy \).
2. QR factorization via Householder reflectors

- Householder method: $Q_n \cdots Q_2 Q_1 A = R$ is upper-triangular.

$A \rightarrow Q_1 A \rightarrow Q_2 Q_1 A \rightarrow Q_3 Q_2 Q_1 A$

$\times$ denotes an entry not necessarily zero; "blank" are zeros

- At the $k$th step, the unitary matrix $Q_k$ has the form

$$Q_k = \begin{bmatrix} I_{k-1} & H_k \end{bmatrix}.$$ 

Here $H_k$ is an $(m - k + 1) \times (m - k + 1)$ Householder reflector, which maps an $m - k + 1$-vector to a scalar multiple of $e_1$.

- The full QR factorization: $A = Q_1^* Q_2^* \cdots Q_n^* R = QR$
QR factorization with column pivoting: $AP = QR$. Consider “qr”

2.1. Two possible Householder reflections in real case

Choose the one that moves $x$ the larger distance, i.e.,
\[ v = -\text{sign}(x_1)\|x\|_2e_1 - x, \text{ or } v = \text{sign}(x_1)\|x\|_2e_1 + x \]

Convention: $\text{sign}(x_1) = 1$ if $x_1 = 0$
2.2. Algorithms

**Algorithm: Householder QR factorization**

```plaintext
for k = 1 to n
    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
```

**Algorithm: Implicit calculations of Q*b or Qx**

```plaintext
for k = 1 to n
    b_{k:m} = b_{k:m} - 2v_k(v_k^*b_{k:m})
end
for k = n downto 1
    x_{k:m} = x_{k:m} - 2v_k(v_k^*x_{k:m})
end
```
3. **Givens rotation** (We mainly consider the real case).

- The $2 \times 2$ Givens rotation

\[
G = \begin{bmatrix}
\cos \theta & \sin \theta \\
-\sin \theta & \cos \theta
\end{bmatrix}, \quad \cos \theta = \frac{x}{\sqrt{x^2 + y^2}}, \quad \sin \theta = \frac{y}{\sqrt{x^2 + y^2}}
\]

rotates vector $\begin{bmatrix} x \\ y \end{bmatrix}$ in $\mathbb{R}^2$ onto the $x$-axis.
Givens rotation for
\[
\begin{bmatrix}
  c & s \\
  -s & c
\end{bmatrix}
\begin{bmatrix}
  x \\
  y
\end{bmatrix}
= \begin{bmatrix}
  \sqrt{x^2 + y^2} \\
  0
\end{bmatrix}
\]

**Algorithm:** Givens rotation zeroing the 2nd entry

function \([c, s] = \text{givens}(x, y)\)

  if \(y = 0\)
    \(c = 1, \quad s = 0\)
  else
    if \(|y| > |x|\)
      \(\tau = x/y, \quad s = 1/\sqrt{1 + \tau^2}, \quad c = s\tau\)
    else
      \(\tau = y/x, \quad c = 1/\sqrt{1 + \tau^2}, \quad s = c\tau\)
    end
  end

**Exercise:** Design a similar algorithm for a Givens rotation zeroing the 1st entry.
Zeroing a particular entry in a vector using a Givens rotation.

Define the \( m \times m \) Givens rotation \( \mathbf{G}(i, j; \theta) \),

\[
\mathbf{G}(i, j; \theta) = \mathbf{I} + [\mathbf{e}_i \quad \mathbf{e}_j] \begin{bmatrix}
\cos \theta - 1 & \sin \theta \\
-\sin \theta & \cos \theta - 1
\end{bmatrix} \begin{bmatrix}
\mathbf{e}_i^T \\
\mathbf{e}_j^T
\end{bmatrix}
\]

\[
= \begin{bmatrix}
\mathbf{I} \\
\cos \theta & \sin \theta \\
-\sin \theta & \cos \theta
\end{bmatrix}
\begin{bmatrix}
\mathbf{I} \\
\text{row i} \\
\text{row j}
\end{bmatrix}
\]

Exercise: Prove that the matrix \( \mathbf{G}(i, j; \theta) \) is orthogonal.

Creating a sequence of zeros in a vector using Givens rotations

\[
\mathbf{G}_n \mathbf{G}_{n-1} \cdots \mathbf{G}_1 \mathbf{x}
\]
QR factorization via Givens rotations?

**Exercise:** Let \( \mathbf{A} \in \mathbb{R}^{n \times n} \) be given as

\[
\mathbf{A} = \begin{bmatrix}
\alpha_1 & \beta_2 & \beta_3 & \cdots & \beta_n \\
\gamma_2 & \alpha_2 & 0 & \cdots & 0 \\
\gamma_3 & 0 & \alpha_3 & \ddots & \vdots \\
\vdots & \vdots & \ddots & \ddots & 0 \\
\gamma_n & 0 & \cdots & 0 & \alpha_n \\
\end{bmatrix},
\]

\( \alpha_i \neq 0, \ i = 1 : n, \)

\( \beta_i \neq 0, \ i = 2 : n, \)

\( \gamma_i \neq 0, \ i = 2 : n. \)

Describe an algorithm for QR factorization of \( \mathbf{A} \) based on as few Givens rotations as possible.

**Complex case:**

\[
\mathbf{G} = \begin{bmatrix}
c & \bar{s} \\
-s & c
\end{bmatrix}, \quad c \in \mathbb{R}, \quad c^2 + |s|^2 = 1.
\]
4. The least squares problem (LSP)

- LSP: Given $A \in \mathbb{C}^{m \times n}, b \in \mathbb{C}^m$; find $x_{ls} \in \mathbb{C}^n$ such that

$$\|b - Ax_{ls}\|_2 = \min_{x \in \mathbb{C}^n} \|b - Ax\|_2.$$

The least squares solution, $x_{ls}$, maybe not unique. Why?

- Note that the 2-norm corresponds to Euclidean distance.

LSP means we seek a vector $x_{ls} \in \mathbb{C}^n$ such that the vector $Ax_{ls}$ is the closest point in range($A$) to $b$.

The residual, $r_{ls} = b - Ax_{ls}$, is unique. Why?

- Assume that $A$ and $b$ are real. Define

$$f(x) := \|b - Ax\|_2^2 = b^T b - x^T A^T b - b^T A x + x^T A^T A x.$$

Then the gradient of $f(x)$ is

$$\nabla f(x) = 2 A^T A x - 2 A^T b.$$
4.1. Example: Polynomial least squares fitting

- Given $m$ distinct points $x_1, \ldots, x_m \in \mathbb{C}$ and data $y_1, \ldots, y_m \in \mathbb{C}$ at these points. Consider a polynomial of degree $n - 1$,

$$p(x) = c_0 + c_1 x + \cdots + c_{n-1} x^{n-1},$$

s.t., $p(x)$ minimizes

$$\sum_{i=1}^{m} |p(x_i) - y_i|^2 = \|y - Ac\|_2^2.$$

$$\begin{bmatrix}
1 & x_1 & x_1^{n-1} \\
1 & x_2 & \cdots & x_2^{n-1} \\
1 & x_3 & \cdots & x_3^{n-1} \\
& \vdots & \ddots & \vdots \\
1 & x_m & \cdots & x_m^{n-1}
\end{bmatrix}
\begin{bmatrix}
c_0 \\
c_1 \\
c_2 \\
\vdots \\
c_{n-1}
\end{bmatrix}
\approx
\begin{bmatrix}
y_1 \\
y_2 \\
y_3 \\
\vdots \\
y_m
\end{bmatrix}$$
4.2. Theory of the least squares problem

**Theorem 3**

Let $\mathbf{P}$ be the orthogonal projector onto $\text{range}(\mathbf{A})$. A vector $\mathbf{x}$ is a least squares solution if and only if $\mathbf{A}\mathbf{x} = \mathbf{P}\mathbf{b}$.

**Proof.**

\[
\|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2^2 = \|\mathbf{P}\mathbf{b} - \mathbf{A}\mathbf{x} + \mathbf{b} - \mathbf{P}\mathbf{b}\|_2^2 = \|\mathbf{P}\mathbf{b} - \mathbf{A}\mathbf{x}\|_2^2 + \|\mathbf{b} - \mathbf{P}\mathbf{b}\|_2^2 \quad \square
\]

**Corollary 4**

A vector $\mathbf{x}$ is a least squares solution if and only if $\mathbf{A}^*\mathbf{A}\mathbf{x} = \mathbf{A}^*\mathbf{b}$, i.e., $\mathbf{A}^*\mathbf{r} = 0$, or $\mathbf{r} \perp \text{range}(\mathbf{A})$, where $\mathbf{r} := \mathbf{b} - \mathbf{A}\mathbf{x}$.

**Proof.**

$\therefore \mathbf{A}^* = \mathbf{A}^*\mathbf{P}$, $\therefore \mathbf{A}^*\mathbf{r} = 0 \Leftrightarrow \mathbf{A}^*(\mathbf{P}\mathbf{b} - \mathbf{A}\mathbf{x}) = 0 \Leftrightarrow \mathbf{A}\mathbf{x} = \mathbf{P}\mathbf{b}$. $\square$

- The system $\mathbf{A}^*\mathbf{A}\mathbf{x} = \mathbf{A}^*\mathbf{b}$ is called the *normal equations*. 

---

**Numerical Linear Algebra**

Lecture 4

October 13, 2020
Corollary 5

The least squares solution $\mathbf{x}$ is unique if and only if $\mathbf{A}$ has full column rank, i.e., $\text{rank}(\mathbf{A}) = n$.

4.3. Geometric interpretation
4.4. Moore-Penrose pseudoinverse solution $A^\dagger b$

- Let $A \in \mathbb{C}^{m \times n}$ have an SVD $A = U_r \Sigma_r V_r^*$. The matrix

$$A^\dagger = V_r \Sigma_r^{-1} U_r^* = \sum_{j=1}^{r} \frac{1}{\sigma_j} v_j u_j^* \in \mathbb{C}^{n \times m},$$

is called the Moore-Penrose pseudoinverse of $A$. If $A$ has full column rank, then $A^\dagger = (A^* A)^{-1} A^*$. (Full row rank case?)

**Theorem 6**

*Let $A \in \mathbb{C}^{m \times n}$ have rank $r < n$ and $b \in \mathbb{C}^m$. Then the vector $A^\dagger b$ is the unique least squares solution with minimal 2-norm.*

**Proof.**

By SVD of $A$, the least squares solutions can be expressed as

$$x_{ls} = A^\dagger b + V_c z, \quad z \in \mathbb{C}^{n-r}.$$

Then the statement follows from $A^\dagger b \perp V_c z$. 

** Numerical Linear Algebra  **

**Lecture 4**

October 13, 2020  16 / 21
4.5. Full column rank LSP solvers

- Normal equations: classical way to solve LSP, best for speed
- QR factorization: “modern classical” method to solve LSP, numerically stable. By

$$A = QR = \begin{bmatrix} Q_n & Q_c \end{bmatrix} \begin{bmatrix} R_n \\ 0 \end{bmatrix},$$

we have

$$\min_{x \in \mathbb{C}^n} \|b - Ax\|_2 = \min_{x \in \mathbb{C}^n} \|b - QRx\|_2 = \min_{x \in \mathbb{C}^n} \|Q^*b - Rx\|_2$$

$$= \min_{x \in \mathbb{C}^n} \left\| \begin{bmatrix} Q_n^*b - R_nx \\ Q_c^*b \end{bmatrix} \right\|_2$$

- SVD, numerically stable, for problems close to rank-deficient. By

$$A = U\Sigma V^* = U_n \Sigma_n V^* = \begin{bmatrix} U_n & U_c \end{bmatrix} \begin{bmatrix} \Sigma_n \\ 0 \end{bmatrix} V^*,$$
we have

\[
\min_{x \in \mathbb{C}^n} \|b - Ax\|_2 = \min_{x \in \mathbb{C}^n} \|b - U\Sigma V^* x\|_2
\]

\[
= \min_{x \in \mathbb{C}^n} \|U^* b - \Sigma V^* x\|_2
\]

\[
= \min_{x \in \mathbb{C}^n} \|\begin{bmatrix} U_n^* b - \Sigma_n V^* x \\ U_c^* b \end{bmatrix}\|_2.
\]

**Exercise:** Given \( A \in \mathbb{C}^{m \times n} \) of full column rank, \( m > n \), \( b \in \mathbb{C}^m \), \( b \notin \text{range}(A) \) and \( QR = [A \ b] \). Show that

\[
\min_{x \in \mathbb{C}^n} \|b - Ax\|_2 = \|R(n + 1, n + 1)\|,
\]

and the least squares solution is given by

\[
x = R(1 : n, 1 : n) \setminus R(1 : n, n + 1).
\]
4.6. Rank-deficient LSP solvers: \( \text{rank}(A) = r < n \leq m \)

- QR factorization with column pivoting:

\[
AP = QR = Q \begin{bmatrix}
R_{11} & R_{12} \\
0 & 0 
\end{bmatrix},
\]

where \( Q \) is unitary, and \( R_{11} \in \mathbb{R}^{r \times r} \) is upper triangular. Assume

\[
Q^* b = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \quad \text{and} \quad P^* x = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.
\]

The general least squares solution is

\[
x_{ls} = P \begin{bmatrix}
R_{11}^{-1}(d_1 - R_{12}y_2) \\
y_2 
\end{bmatrix}, \quad y_2 = \text{arbitrary}.
\]

The case \( y_2 = 0 \) yields the least squares solution with at least \( n - r \) zero components. Consider "\" in MATLAB.
Complete orthogonal factorization (also called UTV factorization)

\[
A = U \begin{bmatrix} R_{11} & 0 \\ 0 & 0 \end{bmatrix} V^*,
\]

where \( U \) and \( V \) are unitary, and \( R_{11} \in \mathbb{R}^{r \times r} \) is upper triangular. Assume

\[
U^* b = \begin{bmatrix} g_1 \\ g_2 \end{bmatrix} \quad \text{and} \quad V^* x = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.
\]

The general least squares solution is

\[
x_{ls} = V \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = V \begin{bmatrix} R_{11}^{-1} g_1 \\ y_2 \end{bmatrix}, \quad y_2 = \text{arbitrary}.
\]

The case \( y_2 = 0 \) yields the minimum norm least squares solution. Consider \texttt{lsqminnorm} in MATLAB.