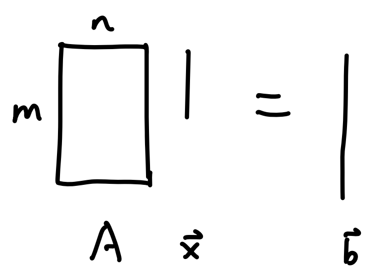
# 5.3 Least-squares problems

Complex vector space $\mathbb{C}^n$:

| Real | Complex |
|---|---|
| $\langle \vec{x}, \vec{y} \rangle = \sum_{i=1}^{n} x_i y_i$ | $\langle \vec{x}, \vec{y} \rangle = \sum_{i=1}^{n} x_i \overline{y_i}$ |
| $\|\vec{x}\|^2 = \sum_{i=1}^{n} x_i^2$ | $\|\vec{x}\|^2 = \sum_{i=1}^{n} \lvert x_i \rvert^2$ |
| Symmetric : $A^T = A$ | Hermitian : $A^* = A$ |
| $\langle \vec{x}, A\vec{y} \rangle = \langle A\vec{x}, \vec{y} \rangle$ | $(A^*)_{ij} = \overline{A_{ji}}$ |
| orthogonal : $A^T A = A A^T = I$ | Unitary : $A^* A = A A^* = I$ |

---

Let $A$ be an $m \times n$ (possibly complex) matrix, $\vec{b} \in \mathbb{C}^m$. We want to find $\vec{x} \in \mathbb{C}^n$ s.t. $\| A\vec{x} - \vec{b} \|$ is minimized.




**Lemma** $\vec{x}$ solves the least squares problem $\Longleftrightarrow$ $\underbrace{A^* A \vec{x} = A^* \vec{b}}_{\text{the normal equation}}$

**Proof** ① If $A^* A \vec{x} = A^* \vec{b}$ then

$$\langle \vec{x}, \vec{y} \rangle = \overline{\langle \vec{y}, \vec{x} \rangle}$$

$$\langle \vec{x}, A\vec{y} \rangle = \langle A^*\vec{x}, \vec{y} \rangle$$

$$\| A(\vec{x} + \vec{y}) - \vec{b} \|^2 = \| (A\vec{x} - \vec{b}) + A\vec{y} \|^2$$

$$= \langle (A\vec{x} - \vec{b}) + A\vec{y}, (A\vec{x} - \vec{b}) + A\vec{y} \rangle$$

$$= \| A\vec{x} - \vec{b} \|^2 + \| A\vec{y} \|^2 + 2\mathrm{Re}\langle A\vec{x} - \vec{b}, A\vec{y} \rangle$$

$$= \| A\vec{x} - \vec{b} \|^2 + \| A\vec{y} \|^2 + 2\mathrm{Re}\langle \underbrace{A^*(A\vec{x} - \vec{b})}_{= \vec{0}}, \vec{y} \rangle$$

$$= \|A\vec{x} - \vec{b}\|^2 + \|A\vec{y}\|^2 \geq \|A\vec{x} - \vec{b}\|^2$$

② If $\vec{x}$ solves least squares, suppose $\vec{z} := A^*(A\vec{x} - \vec{b}) \neq \vec{0}$

$$\|A(\vec{x}+\vec{y}) - \vec{b}\|^2 = \|A\vec{x} - \vec{b}\|^2 + \|A\vec{y}\|^2 + 2\mathrm{Re}\langle \vec{z}, \vec{y}\rangle$$

Take $\vec{y} = -\varepsilon\vec{z}$ w/ $\varepsilon > 0$ small. Then

$$\|A(\vec{x}+\vec{y}) - \vec{b}\|^2 = \|A\vec{x} - \vec{b}\|^2 + \varepsilon^2\|A\vec{z}\|^2 - 2\varepsilon \underbrace{\|\vec{z}\|^2}_{>0}$$

$\Rightarrow$ when $\varepsilon$ is sufficiently small, we get

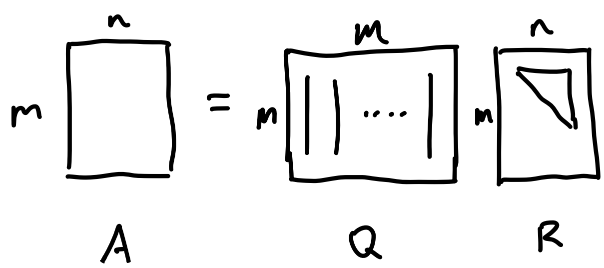
$\|A(\vec{x}+\vec{y}) - \vec{b}\|^2 < \|A\vec{x} - \vec{b}\|^2$ contradiction.

## QR factorization

**Def** A QR-factorization of an $m \times n$ matrix $A$ is

$$A = QR \qquad \Leftrightarrow \qquad Q^*A = R$$

where $Q$ is $m \times m$ unitary and $R$ is $m \times n$ upper-triangular.




- If $A$ has QR-factorization, then least squares problem can be solved

$$A^*A\vec{x} = A^*\vec{b}$$

$$R^*Q^*QR\vec{x} = R^*Q^*\vec{b}$$

$$R^*R\vec{x} = R^*Q^*\vec{b}$$

↳ upper-triangular

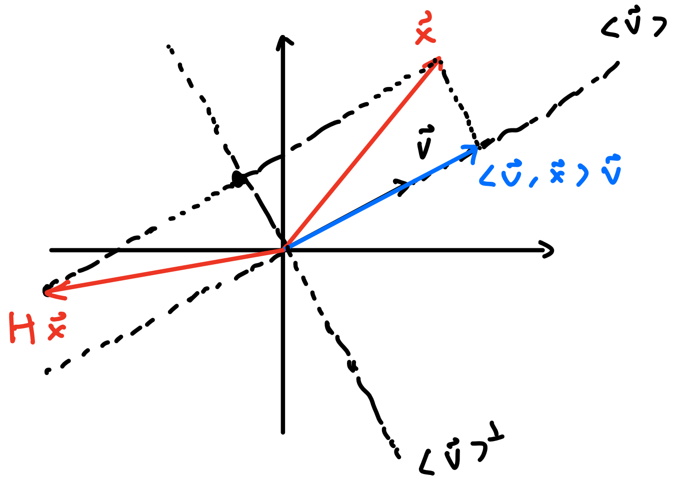
↳ lower-triangular.

We obtain QR-factorization by building $Q^*$ as a product of <u>Householder transformations</u>, i.e., a matrix of the form

$$\begin{pmatrix} I_k & 0 \\ 0 & I_{m-k} - 2\vec{v}\vec{v}^* \end{pmatrix} \quad w/\ \vec{v} \in \mathbb{C}^{m-k},\ \|\vec{v}\| = 1$$

- Geometrically, a Householder transf. is a reflection about $\langle \vec{v} \rangle^\perp$

$H = I - 2\vec{v}\vec{v}^T$ in $\mathbb{R}^2$




$$H\vec{x} = \vec{x} - 2\vec{v}\vec{v}^T\vec{x}$$
$$= \vec{x} - 2\langle \vec{v}, \vec{x} \rangle \vec{v}$$

---

Starting from $A = A^{(1)}$

Step 1 : find $\vec{v}^{(1)}$ s.t.

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\underbrace{\left( I_m - 2\vec{v}^{(1)}\vec{v}^{(1)*} \right)}_{U^{(1)}} \underbrace{\begin{pmatrix} | & & | \\ | & \cdots & | \\ | & & | \end{pmatrix}}_{A^{(1)}} \begin{matrix} n \\ m \end{matrix} = \underbrace{\begin{pmatrix} * & | & \cdots & | \\ 0 & | & & | \\ \vdots & | & & | \\ 0 & | & & | \end{pmatrix}}_{A^{(2)}}$$

Denote $\vec{x}$ as the first column of $A^{(1)}$. We need

$$(I_m - 2\vec{v}\vec{v}^*)\vec{x} = \alpha \vec{e}_1$$
$$\vec{x} - 2\langle \vec{x}, \vec{v} \rangle \vec{v} = \alpha \vec{e}_1$$
$$\vec{x} - \alpha \vec{e}_1 = 2\langle \vec{x}, \vec{v} \rangle \vec{v}$$

$|\alpha| = \|\vec{x}\|$

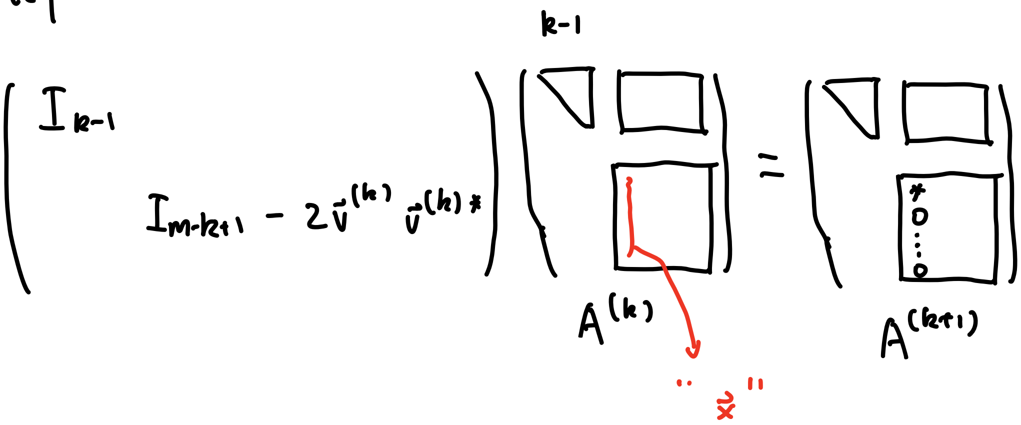
$\alpha = -e^{i\phi}\|\vec{x}\|$

$\hookrightarrow$ choose $\phi = \operatorname{Arg} x_1$ to avoid subtracting close numbers in $x_1 - \alpha$.

$$\Rightarrow \quad \vec{v} = \frac{\vec{x} - \alpha \vec{e}_1}{\| \vec{x} - \alpha \vec{e}_1 \|} \qquad (\text{since } \|\vec{v}\| = 1)$$

$$\vdots$$

Step $k$ :




After $n-1$ steps

$$\underbrace{U^{(n-1)} \cdots U^{(2)} U^{(1)}}_{Q^*} A = A^{(n)} := R$$

$$Q = U^{(1)} U^{(2)} \cdots U^{(n-1)}$$

Total cost of QR-factorization for full matrix : $O(mn^2)$ (for $m \geq n$)