Lecture 4
The QR Factorization

MIT 18.335J / 6.337J

Introduction to Numerical Methods

Per-Olof Persson

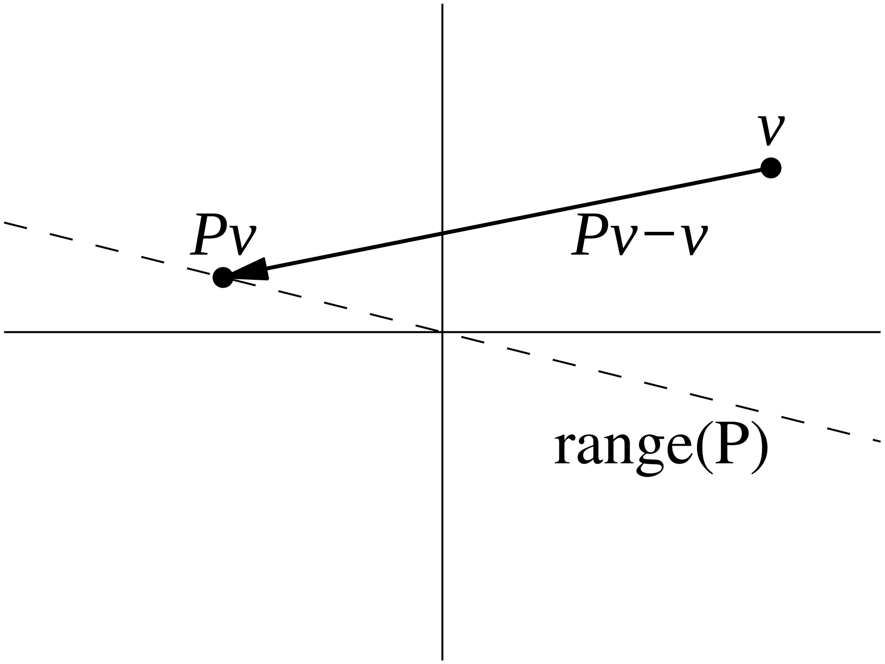
September 19, 2006

Projectors

- A *projector* is a square matrix P that satisfies

$$P^2 = P$$

- Not necessarily an *orthogonal projector* (more later)
- If $v \in \mathrm{range}(P)$, then $Pv = v$
  - Since with $v = Px$,
    $Pv = P^2x = Px = v$
- Projection along the line
  $Pv - v \in \mathrm{null}(P)$
  - Since $P(Pv - v) =$
    $P^2v - Pv = 0$



Complementary Projectors

- The matrix $I - P$ is the *complementary projector* to P
- $I - P$ projects on the nullspace of P:
  - If $Pv = 0$, then $(I - P)v = v$, so $\mathrm{null}(P) \subseteq \mathrm{range}(I - P)$
  - But for any v, $(I - P)v = v - Pv \in \mathrm{null}(P)$, so $\mathrm{range}(I - P) \subseteq \mathrm{null}(P)$
  - Therefore

$$\mathrm{range}(I - P) = \mathrm{null}(P)$$

and

$$\mathrm{null}(I - P) = \mathrm{range}(P)$$

Complementary Subspaces

- For a projector P,

$$\mathrm{null}(I - P) \cap \mathrm{null}(P) = \{0\}$$

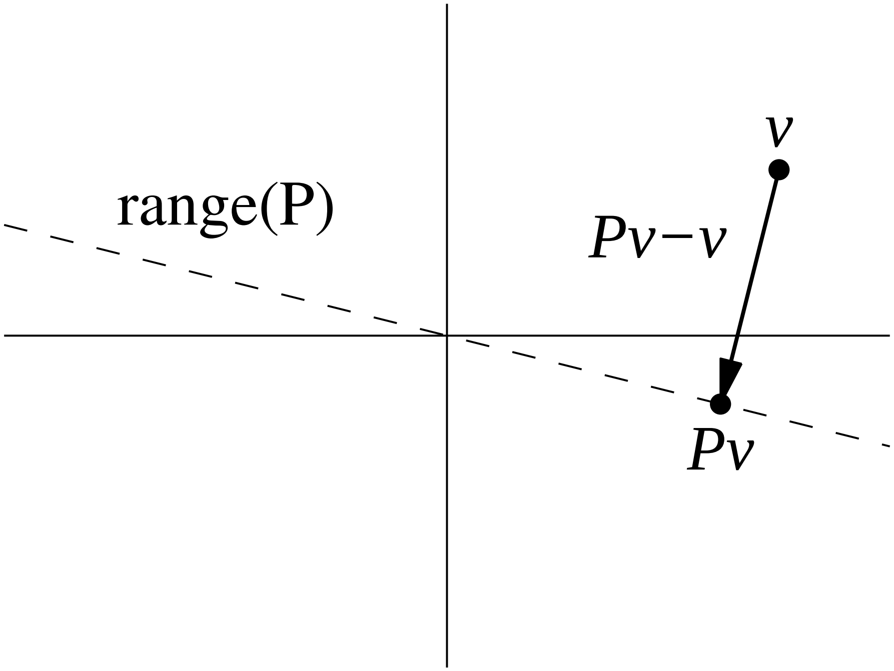
or

$$\mathrm{range}(P) \cap \mathrm{null}(P) = \{0\}$$

- A projector separates $\mathbb{C}^m$ into two spaces S_1, S_2, with $\mathrm{range}(P) = S_1$ and $\mathrm{null}(P) = S_2$
- P is the projector *onto* S_1 *along* S_2

Orthogonal Projectors

- An *orthogonal projector* projects onto S_1 along S_2, with S_1, S_2 orthogonal
- A projector P is orthogonal $\Longleftrightarrow P = P^*$
- *Proof.* Textbook / Black board



Projection with Orthonormal Basis

- Reduced SVD gives projector for orthonormal columns $\hat{Q}$:

$$P = \hat{Q}\hat{Q}^*$$

- Complement $I - \hat{Q}\hat{Q}^*$ also orthogonal, projects onto space orthogonal to $\mathrm{range}(\hat{Q})$
- Special case 1: Rank-1 Orthogonal Projector (gives component in direction q)

$$P_q = qq^*$$

- Special case 2: Rank $m-1$ Orthogonal Projector (eliminates component in direction q)

$$P_{\perp q} = I - qq^*$$

Projection with Arbitrary Basis

- Project v to $y \in \operatorname{range}(A)$. Then

$$y - v \perp \operatorname{range}(A), \text{ or } a_j^*(y - v) = 0, \forall j$$

- Set $y = Ax$:

$$a_j^*(Ax - v) = 0, \forall j \iff A^*(Ax - v) = 0 \iff A^*Ax = A^*v$$

- A^*A is nonsingular, so

$$x = (A^*A)^{-1}A^*v$$

- Finally, we are interested in the projection $y = Ax = A(A^*A)^{-1}A^*v$, giving the orthogonal projector

$$P = A(A^*A)^{-1}A^*$$

The QR Factorization - Main Idea

- Find orthonormal vectors that span the successive spaces spanned by the columns of A:

$$\langle a_1 \rangle \subseteq \langle a_1, a_2 \rangle \subseteq \langle a_1, a_2, a_3 \rangle \subseteq \ldots$$

- This means that (for full rank A),

$$\langle q_1, q_2, \ldots, q_j \rangle = \langle a_1, a_2, \ldots, a_j \rangle \,, \quad \text{for } j = 1, \ldots, n$$

The QR Factorization - Matrix Form

- In matrix form, $\langle q_1, q_2, \ldots, q_j \rangle = \langle a_1, a_2, \ldots, a_j \rangle$ becomes

$$\left[\begin{array}{c|c|c|c} a_1 & a_2 & \cdots & a_n \end{array} \right] = \left[\begin{array}{c|c|c|c} q_1 & q_2 & \cdots & q_n \end{array} \right] \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ & r_{22} & & \vdots \\ & & \ddots & \vdots \\ & & & r_{nn} \end{bmatrix}$$

or

$$A = \hat{Q}\hat{R}$$

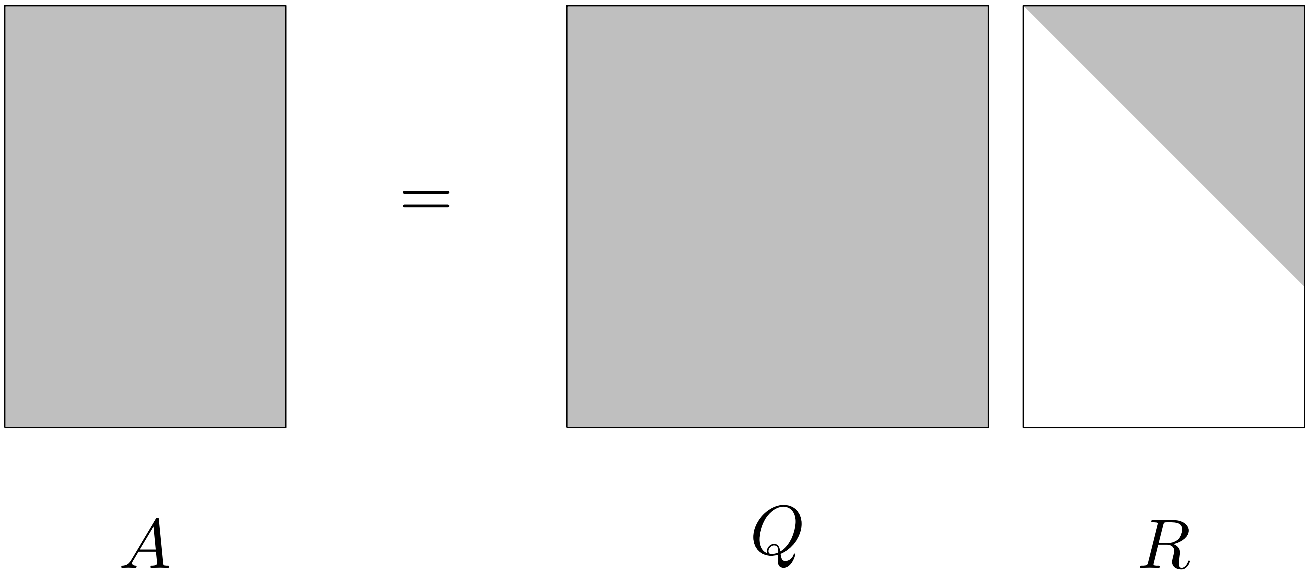
- This is the *reduced QR factorization*
- Add orthogonal extension to $\hat{Q}$ and add rows to $\hat{R}$ to obtain the *full QR factorization*

The Full QR Factorization

- Let A be an $m \times n$ matrix. The full QR factorization of A is the factorization $A = QR$, where

Q is $m \times m$ unitary

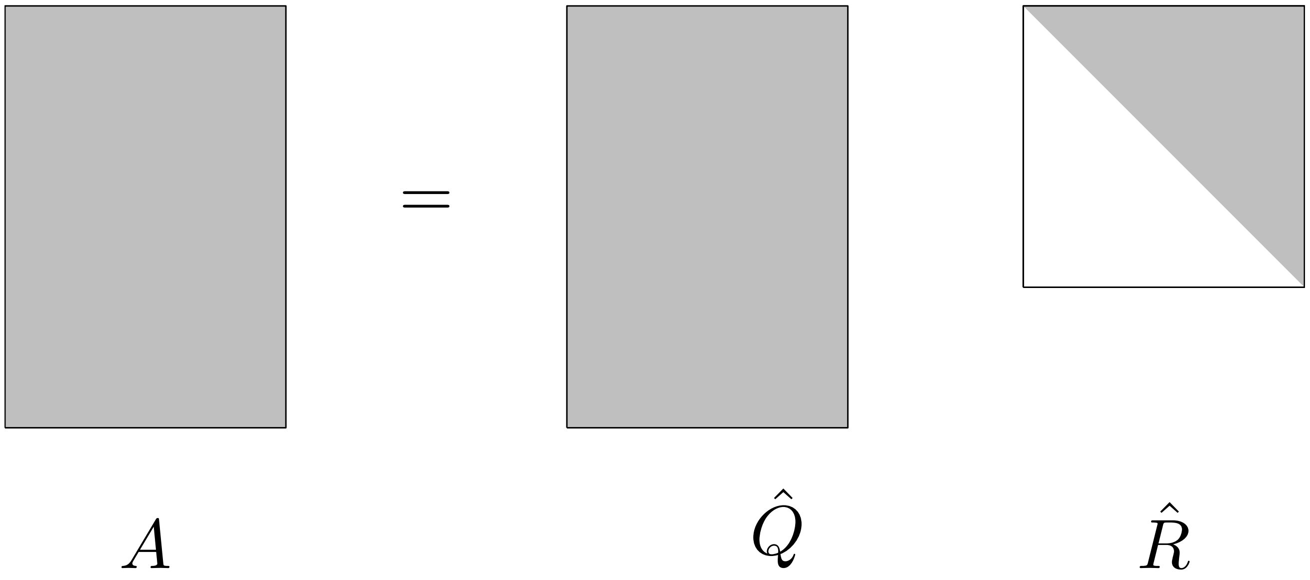
R is $m \times n$ upper-triangular



The Reduced QR Factorization

- A more compact representation is the *Reduced QR Factorization* $A = \hat{Q}\hat{R}$, where (for $m \geq n$)

$$\hat{Q} \text{ is } m \times n \text{ and } \hat{R} \text{ is } m \times n$$



Gram-Schmidt Orthogonalization

- Find new q_j orthogonal to $q_1, \ldots, q_{j-1}$ by subtracting components along previous vectors

$$v_j = a_j - (q_1^* a_j) q_1 - (q_2^* a_j) q_2 - \cdots - (q_{j-1}^* a_j) q_{j-1}$$

- Normalize to get $q_j = v_j / \|v_j\|$
- We then obtain a reduced QR factorization $A = \hat{Q}\hat{R}$, with

$$r_{ij} = q_i^* a_j, \quad (i \neq j)$$

and

$$|r_{jj}| = \|a_j - \sum_{i=1}^{j-1} r_{ij} q_i\|_2$$

Classical Gram-Schmidt

- Straight-forward application of Gram-Schmidt orthogonalization
- Numerically unstable

Algorithm: Classical Gram-Schmidt

```
for $j = 1$ to $n$
    $v_j = a_j$
    for $i = 1$ to $j - 1$
        $r_{ij} = q_i^* a_j$
        $v_j = v_j - r_{ij} q_i$
    $r_{jj} = \|v_j\|_2$
    $q_j = v_j / r_{jj}$
```

Existence and Uniqueness

- Every $A \in \mathbb{C}^{m \times n}$ $(m \geq n)$ has a full QR factorization and a reduced QR factorization
- *Proof.* For full rank A, Gram-Schmidt proves existence of $A = \hat{Q}\hat{R}$. Otherwise, when $v_j = 0$ choose arbitrary vector orthogonal to previous q_i. For full QR, add orthogonal extension to Q and zero rows to R.
- Each $A \in \mathbb{C}^{m \times n}$ $(m \geq n)$ of full rank has unique $A = \hat{Q}\hat{R}$ with $r_{jj} > 0$
- *Proof.* Again Gram-Schmidt, $r_{jj} > 0$ determines the sign