QR Factorization

CPSC 406 – Computational Optimization

Overview

- Orthogonality
- QR properties
- Solution of linear least-squares problems

Orthogonal vectors

Two vectors x and y in $\mathbb{R}^n$

- recall cosine identity

$$x^\mathsf{T} y = \|x\|_2 \|y\|_2 \cos\theta$$

- x and y in $\mathbb{R}^n$ are **orthogonal** if

$$x^\mathsf{T} y = 0 \quad (\cos\theta = 0)$$

- x and y are **orthonormal** if

$$x^\mathsf{T} y = 0, \quad x^\mathsf{T} x = 1, \quad y^\mathsf{T} y = 1$$

- a set of orthogonal n-vectors $\{q_1, \dots, q_m\}$ are linearly independent
  - if $m = n$ then it's a basis for $\mathbb{R}^n$

Orthogonal matrices

An $n \times r$ matrix Q is **orthonormal** if its columns are pairwise orthonormal:

$$Q = [q_1 \mid \cdots \mid q_r], \qquad Q^\top Q = \begin{bmatrix} q_1^\top q_1 & q_2^\top q_1 & \cdots & q_r^\top q_1 \\ q_1^\top q_2 & q_2^\top q_2 & \cdots & q_r^\top q_2 \\ \vdots & \vdots & \ddots & \vdots \\ q_1^\top q_r & q_2^\top q_r & \cdots & q_r^\top q_r \end{bmatrix} = I_r$$

- if $r = n$ (ie, Q is square) then Q is **orthogonal**

$$Q^{-1} = Q^\top \quad \text{and} \quad Q^\top Q = QQ^\top = I_n$$

- **orthogonal** transformations preserve lengths and angles

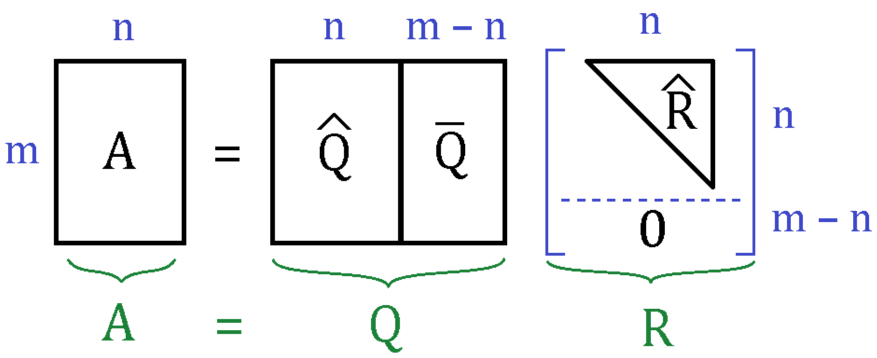
$$\|x\|_2 = \|Qx\|_2 \quad \text{and} \quad x^\top y = x^\top Q^\top Q y = (Qx)^\top (Qy) \quad \text{and} \quad \det(Q) = \pm 1$$

Question: Orthogonal matrices

Let Q be an $n \times n$ orthogonal matrix (so $Q^T Q = I_n$). Which of the following statements is **always** true for any vectors $x, y \in \mathbb{R}^n$?

a. $\|Qx\| = \|x\|$, but $\|Qy\| \neq \|y\|$
b. $(Qx)^\top (Qy) = x^\top y$
c. $\det(Q) = +1$
d. $\|Qx\| \neq \|x\|$, and angles are distored by Q?

QR Factorization



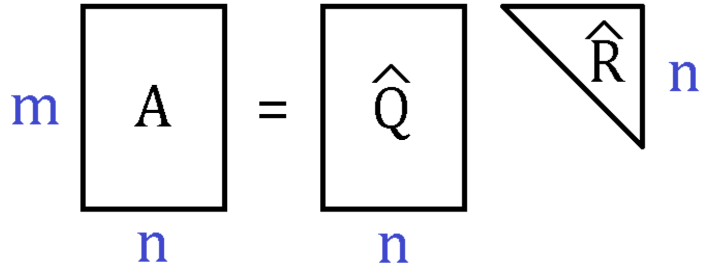
where

1. Q is orthogonal ($Q^\intercal Q = QQ^\intercal = I_m$)
2. $\hat{R}$ is upper triangular ($\hat{R}_{ij} = 0$ for $i > j$)
3. $\mathbf{range}(\hat{Q}) = \mathbf{range}(A)$
4. $\mathbf{range}(\bar{Q}) = \mathbf{range}(A)^\perp \equiv \mathbf{null}(A^\intercal)$

```julia
using LinearAlgebra
Q, R = qr(A)
```

Reduced QR Factorization

For A $m \times n$ with $m \geq n$, full rank



taking column by column, with $\hat{Q} = [q_1 \mid \cdots \mid q_n]$

$$
\begin{aligned}
a_1 &= r_{11}q_1 \\
a_2 &= r_{12}q_1 + r_{22}q_2 \\
a_3 &= r_{13}q_1 + r_{23}q_2 + r_{33}q_3 \\
&\vdots \\
a_n &= r_{1n}q_1 + r_{2n}q_2 + \cdots + r_{nn}q_n
\end{aligned}
$$

Question: Columns of Q

Let A be an $m \times n$ matrix ($m \geq n$) of full column rank, and suppose $A = QR$ is its reduced QR factorization. The columns of Q form an orthonormal basis for which of the following subspaces?

- A. The **row space** of A
- B. The **column space** of A
- C. The **null space** of A
- D. The **orthogonal complement** of the column space of A

Nonsingular equations with QR

Given $n \times n$ matrix A, full rank, solve

$$Ax = b$$

solve by QR factorization $A = QR$ and $Q^\intercal Q = I$

mathematically

$$x = A^{-1}b = (QR)^{-1}b = R^{-1}Q^{-1}b = R^{-1}Q^\intercal b$$

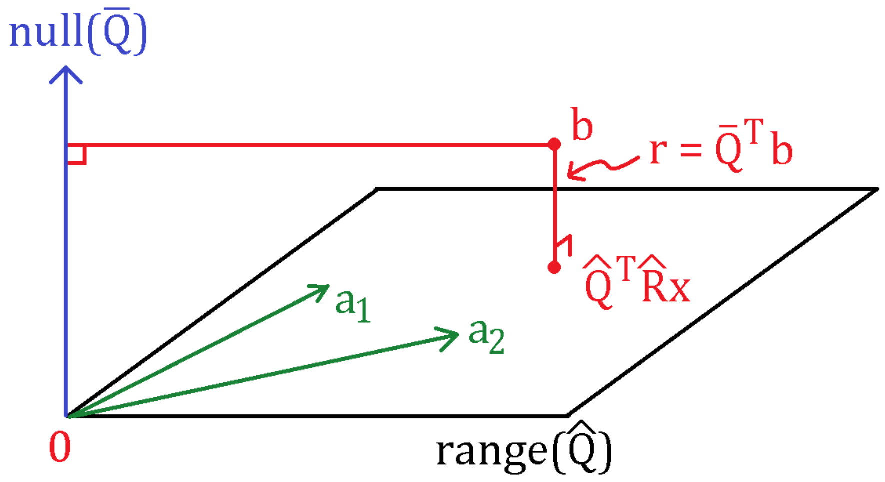
computationally

```julia
using LinearAlgebra
Q, R = qr(A)    # O(n^3) <-- dominant cost
y = Q'b         # O(n^2) <-- matrix-vector multiply
x = R \ y       # O(n^2) <-- triangular solve
```

Geometry of Least-Squares via QR

$$\min_{x\in\mathbb{R}^n} \|Ax-b\|^2, \qquad A = QR$$

$$\begin{aligned}\|Ax-b\|^2 &= (Ax-b)^\intercal(Ax-b)\\ &= (Ax-b)^\intercal QQ^\intercal(Ax-b)\\ &= \|Q^\intercal(Ax-b)\|^2\\ &= \left\|\begin{bmatrix}\hat{R}\\ 0\end{bmatrix}x - \begin{bmatrix}\hat{Q}^\intercal\\ \bar{Q}^\intercal\end{bmatrix}b\right\|^2\\ &= \underbrace{\|\hat{R}x-\hat{Q}^\intercal b\|^2}_{(1)} + \underbrace{\|\bar{Q}^\intercal b\|^2}_{(2)}\end{aligned}$$



where (1) is minimized when $\hat{R}x = \hat{Q}^\intercal b$ and (2) is constant

Question: Least-Squares and Orthogonal Projections

Consider the least-squares problem

$$\min_{x \in \mathbb{R}^n} ||Ax - b||,$$

Let $A = QR$ and $c = Q^T b$. Which of the following best describes the **meaning** of c$?

- A. c is the orthogonal projection of b in the **original space**.
- B. c is the coordinate vector of the projection of b onto the **column space** of Q.
- C. c is orthogonal to every column of Q.
- D. c has no geometric interpretation for the least-squares problem.

Solving Least-Squares via QR

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|^2, \qquad A = QR$$

mathematically

$$\begin{aligned} A^\top A x &= A^\top b \\ R^\top Q^\top Q R x &= R^\top Q^\top b \\ R x &= Q^\top b \\ x &= R^{-1} Q^\top b \end{aligned}$$

computationally

```julia
using LinearAlgebra
F = qr(A)                   # O(n^3) <-- dominant cost
Q, R = Matrix(F.Q), F.R  # extract _thin_ Q, and R
y = Q'b                     # O(n^2) <-- matrix-vector multiply
x = R \ y                   # O(n^2) <-- triangular solve
```

more numerically stable than solving $A^\top A x = A^\top b$ directly

Question: Geometric Interpretation of R

In the factorization $A = QR$, with Q having orthonormal columns, what is the **geometric interpretation** of the triangular matrix R?

a. R is an orthogonal matrix that preserves angles and lengths.

b. R describes how the columns of A can be expressed as linear combinations of the columns of Q, **capturing their coordinates** in the orthonormal basis.

c. R is the null-space basis of A.

d. R is a diagonal matrix containing the singular values of A.

Accuracy of QR vs Normal Equations

For ϵ positive, this matrix has full rank because $\sin^2(\theta) + \cos^2(\theta) = 1$

$$A = \begin{bmatrix} sin^2(\theta_1) & cos^2(\theta_1 + \epsilon) & 1 \\ sin^2(\theta_2) & cos^2(\theta_2 + \epsilon) & 1 \\ \vdots & \vdots & \vdots \\ sin^2(\theta_m) & cos^2(\theta_m + \epsilon) & 1 \end{bmatrix}$$

```julia
using LinearAlgebra
θ = LinRange(0,3,400)
ε = 1e-7
A = @. [sin(θ)^2   cos(θ+ε)^2   θ^0]
xᵉ = [1., 2., 1.]
b  = A*xᵉ

xn = A'A \ A'b                 # Compute xn via normal equations

Q, R = qr(A); Q = Matrix(Q)  # Compute xr via QR
xr = R \ (Q'b)

xb = A \ b                     # Compute xb via backslash

@show xn xr xb;
```