CS 357: Numerical Methods

Lecture 11: QR Decomposition

Eric Shaffer
Why is orthogonality useful?

- Matrices with orthonormal columns can do special things

- $Qv$ preserves the 2-norm of $v$

- Important in Least Squares problems
Orthogonal Transformations Preserve the 2-Norm

- What is true about the columns of an orthonormal matrix $Q$?
- What is $QQ^T$?

\[
\begin{align*}
\langle c_i, c_j \rangle &= 0 \quad \text{if } i \neq j \\
\| c_i \|_2 &= 1
\end{align*}
\]

\[
Q^T Q = \begin{bmatrix}
0 & 1 & \cdots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
0 & \cdots & \cdots & 0 \\
0 & \cdots & \cdots & -1
\end{bmatrix} = I
\]
We can show that $\|Qv\|^2_2 = \|v\|^2_2$

$\|Qv\|^2_2 = (Qv)^TQv$

$= v^TQ^TQv = v^Tv = \|v\|^2_2$
Orthogonal Transformations Preserve the 2-Norm
Orthogonal Transformations Preserve the 2-Norm
Orthogonal Transformations Preserve the 2-Norm

What does this mean in terms of amplifying error?

$$V = V_{true} + V_e$$
Least Squares Problems

- Lots of interesting problems lack an exact solution....
- Fit a line to a set of points....
Least Squares Applications

Tornadoes by Year (U.S.)

- Red: Observed
- Black: Regression
- Blue: Adjusted

Year


Number

0 200 400 600 800 1000 1200 1400 1600
Least Squares Applications

**wOBA and Runs Scored per Game**

- **Y-axis**: Team Runs Scored per Game
- **X-axis**: Team wOBA

The graph shows a strong positive correlation between Team wOBA and Team Runs Scored per Game.
Least Squares Applications

- IPCC SRES A1FI very likely to exceed 4°C
- Reference (close to SRES A1B) likely to exceed 3°C
- Current Pledges virtually certain to exceed 2°C; 50% chance above 3°C
- Stabilization at 50% chance to exceed 2°C
- RCP3PD likely below 2°C; medium chance to exceed 1.5°C

Global average surface temperature increase above pre-industrial levels (°C)

- Global sudden stop to emissions in 2016 likely below 1.5°C
- Illustrative low-emission scenario with negative CO₂ emissions from upper half of literature range in 2nd half of 21st Century

Geophysical intertia

Dates:
- 1900
- 1950
- 2000
- 2050
- 2100

Historical observations
Beware: correlation and causation

I used to think correlation implied causation.

Then I took a statistics class. Now I don’t.

Sounds like the class helped. Well, maybe.
Let's fit a line to series of data sampled over times $t_0, t_1, t_2, \ldots, t_{m-1}$.

The line is given by $f(t) = x_1 t + x_0$.

$$A = \begin{bmatrix} 1 & t_0 \\ 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_{m-1} \end{bmatrix}, \quad X = \begin{bmatrix} x_0 \\ x_1 \end{bmatrix}, \quad b = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \end{bmatrix}$$
So...how does orthogonality relate to least squares?

- The closest fit to the observed data is an orthogonal projection into the column space of a matrix....
- You’ll understand later....

Figure J.1: Geometrical interpretation of orthogonal projection.
Recap: Orthonormal Basis

- A basis is orthonormal if each basis vector:
  - Has unit length
  - Is orthogonal to all other basis vectors.

- Example: (1,0) and (0,1) for 2D Euclidean space
- Can you give another 2D orthonormal basis?

\[
\begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{-\sqrt{2}}{2} \end{pmatrix}
\]
Recap: Orthonormal Basis

For some given vector $\hat{x}$, how do I find coefficients with respect to an ONB?

\[ \hat{x} = (x \cdot b_1)\, \hat{b}_1 + (x \cdot b_2)\, \hat{b}_2 + (x \cdot b_3)\, \hat{b}_3 + \ldots + (x \cdot b_n)\, \hat{b}_n \]

Much easier than finding coefficients by solving a linear system!

Also much cheaper: $O(n^2)$
Can we build a matrix that computes those coefficients for us?

\[ Q = \begin{pmatrix} | & | & | \\ b_1 & \cdots & b_n \\ | & | & | \end{pmatrix} \quad \Rightarrow \quad Q^T \hat{x} = \begin{pmatrix} \hat{b}_1 \cdot \hat{x} \\ \vdots \\ \hat{b}_n \cdot \hat{x} \end{pmatrix} \]

A square matrix whose columns are orthonormal is called orthogonal.
Example

\[ b_1 = \left( \frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right) \]

\[ b_2 = \left( -\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right) \]

\[ Q = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \]

\[ Q^T \begin{bmatrix} \sqrt{2} \\ \sqrt{2} \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \]
What if $Q$ contains a few zero columns instead of orthonormal vectors?

Define $P := QQ^T$

Compute $P\tilde{x}$ for $\tilde{x} = \alpha_1\hat{b}_1 + \ldots + \alpha_n\hat{b}_n$:

$Q^T\tilde{x} = \begin{pmatrix} 
\tilde{x} \cdot \hat{b}_1 \\
\vdots \\
\tilde{x} \cdot \hat{b}_n \\
0 
\end{pmatrix}$
Orthogonal Projection

Define \( P := QQ^T \)

Compute \( P\tilde{x} \) for \( \tilde{x} = \alpha_1 \tilde{b}_1 + \cdots + \alpha_n \tilde{b}_n \):

\[
Q^T\tilde{x} = \begin{pmatrix}
\tilde{x} \cdot \tilde{b}_1 \\
\vdots \\
\tilde{x} \cdot \tilde{b}_n \\
0
\end{pmatrix}
\]

\[
Q(Q^T\tilde{x}) = (\tilde{x} \cdot \tilde{b}_1)\tilde{b}_1 + \cdots + (\tilde{x} \cdot \tilde{b}_k)\tilde{b}_k + 0 \cdot \tilde{b}_{k+1} + \cdots + 0 \cdot \tilde{b}_n
\]

\( \tilde{x} \) projected onto \( \tilde{b}_1, \ldots, \tilde{b}_k \)
Example

\[ q_1 = \frac{a_1}{\| a_1 \|} \]

\[ r = a_2 - (q_1 \cdot a_2) q_1 \]

\[ q_2 = \frac{r}{\| r \|} \]
Gram-Schmidt Orthogonalization

- Given linearly independent $a_1$ and $a_2$
- Find $q_1$ and $q_2$ that are orthonormal and span same space
We can orthogonalize any number of vectors...

```python
for k in range(A.shape[1]):
    avec = A[:, k]
    q = avec
    for j in range(k):
        q = q - np.dot(avec, Q[:, j]) * Q[:, j]
    Q[:, k] = q / la.norm(q)
```
Problems

- Rounding error can destroy orthogonality in the $q_k$ vectors
- Also we need to store A, Q and R separately
  - problematic for large systems
for k in range(A.shape[1]):
    q = A[:, k]
    for j in range(k):
        q = q - np.dot(q, Q[:,j])*Q[:,j]

    Q[:, k] = q/la.norm(q)