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?

$$c_i \cdot c_j = 0 \quad i \neq j$$

$$c_i \, c_i = 1 = \|c_i\|_2^2$$

- What is QQ^T?

$$Q^T Q = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & 1 \end{bmatrix} = I$$

Orthogonal Transformations Preserve the 2-Norm

We can show that $\|Qv\|_2^2 = \|v\|_2^2$

$$\|Qv\|_2^2 = \underbrace{(Qv)^T}_{\text{row vector}} \underbrace{Qv}_{\text{column vector}}$$

$$= v^T Q^T Q v = v^T v = \|v\|_2^2$$

Orthogonal Transformations Preserve the 2-Norm

Orthogonal Transformations Preserve the 2-Norm

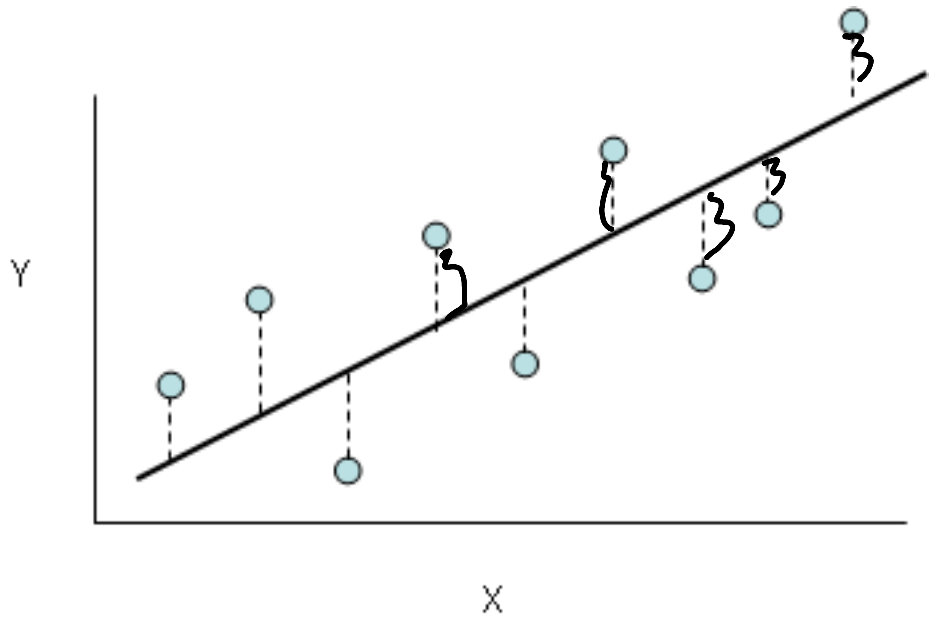
Orthogonal Transformations Preserve the 2-Norm

- What does this mean in terms of amplifying error?

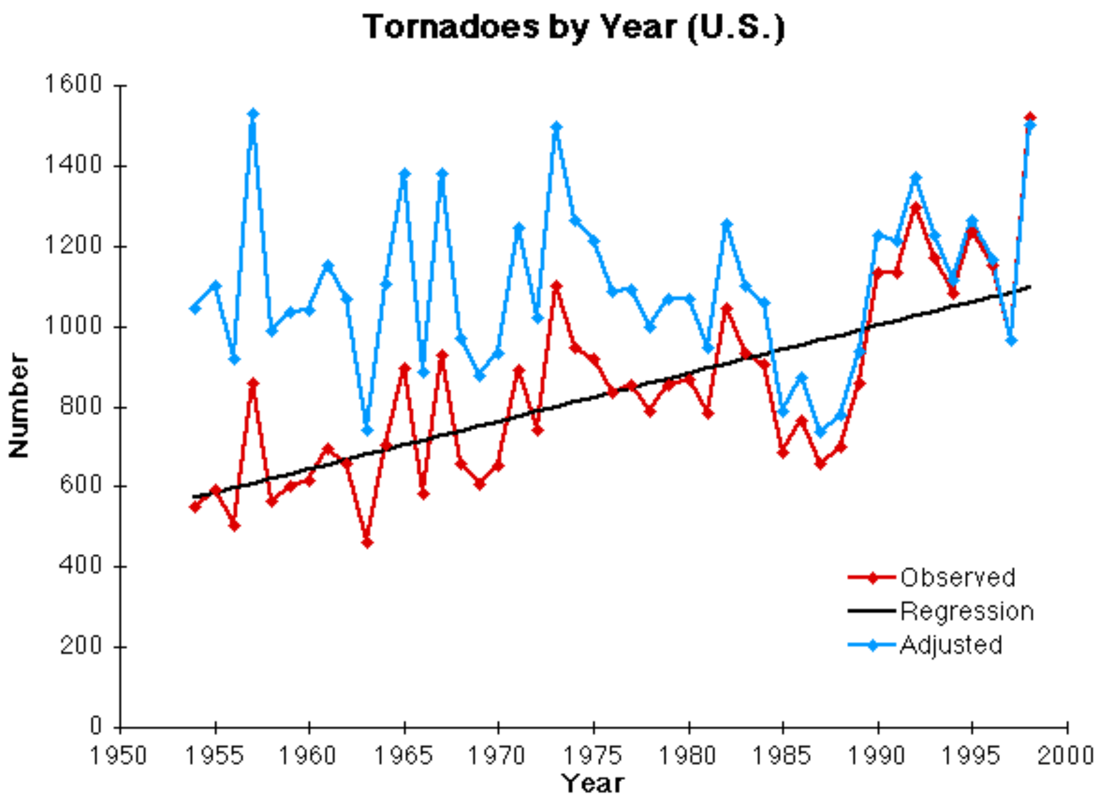
$v = v_{True} + v_e$

Least Squares Problems

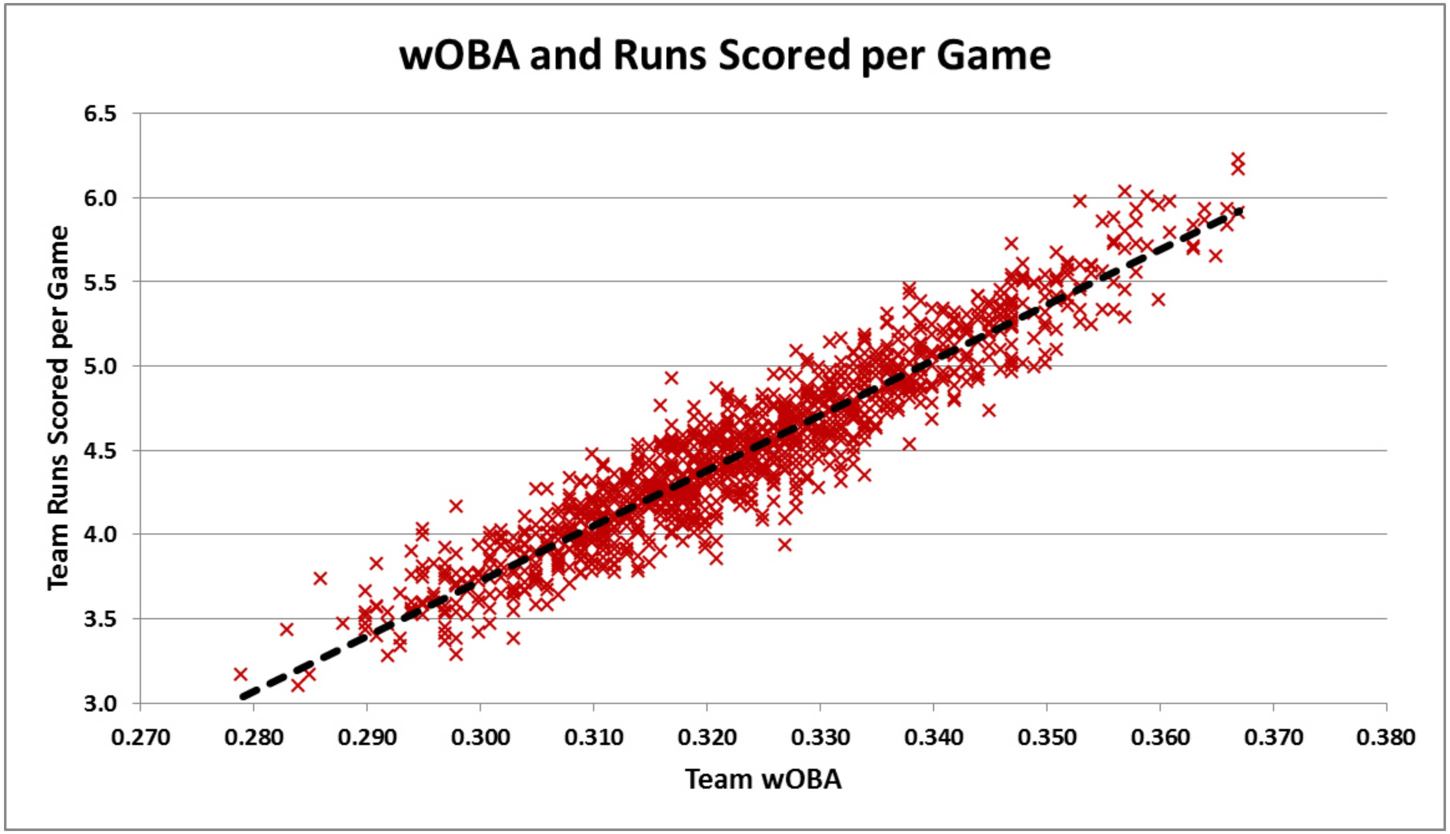
- Lots of interesting problems lack an exact solution….
- Fit a line to a set of points….



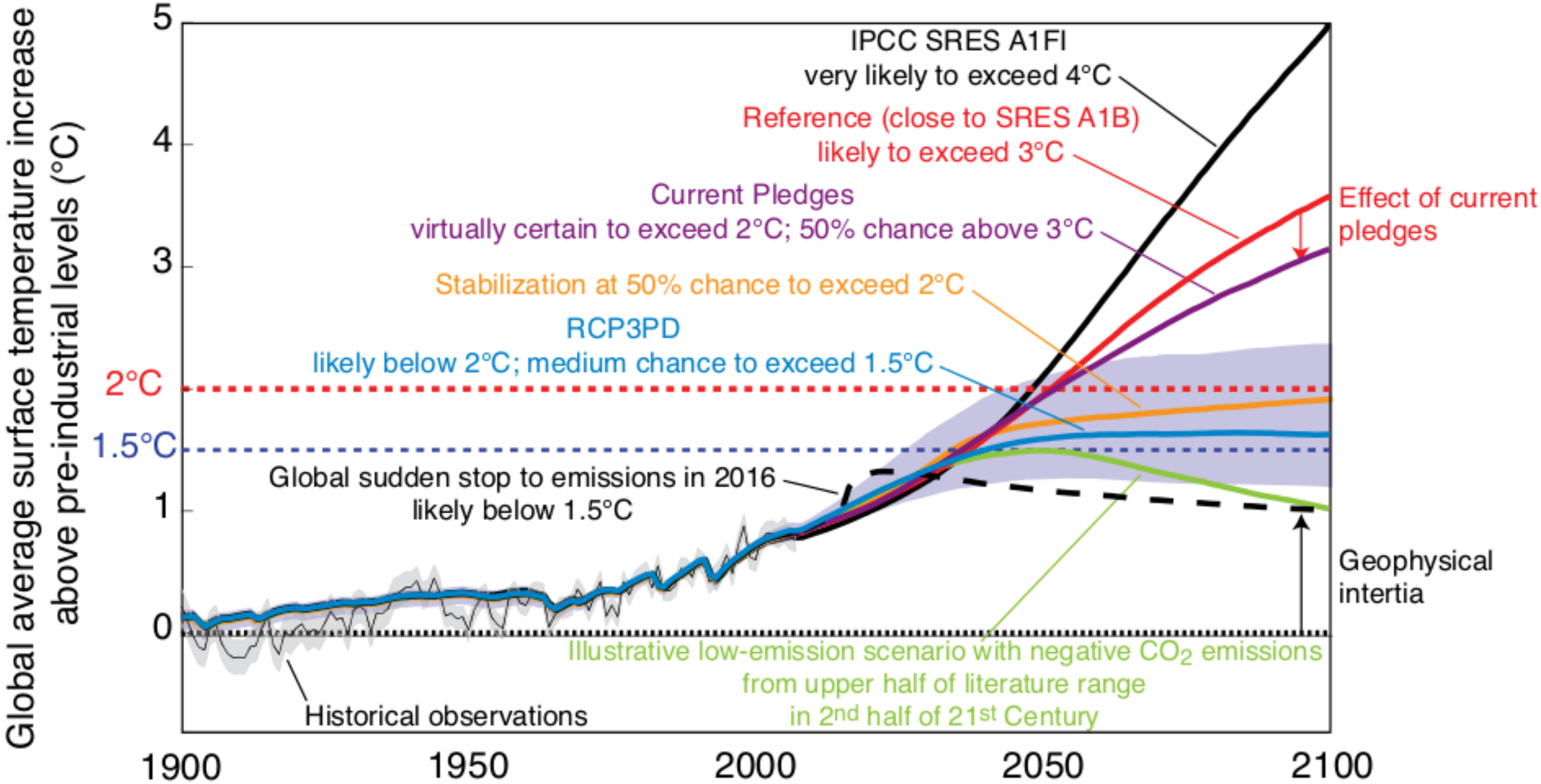
Least Squares Applications



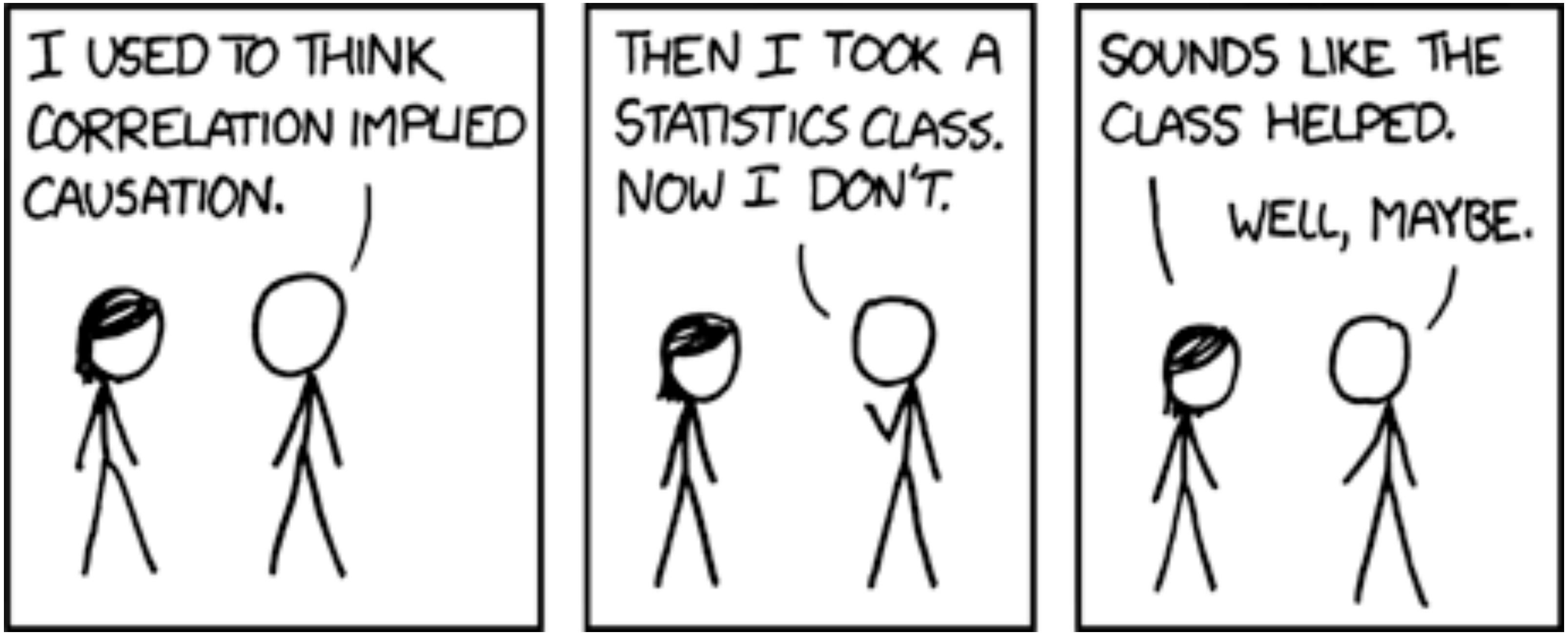
Least Squares Applications



Least Squares Applications



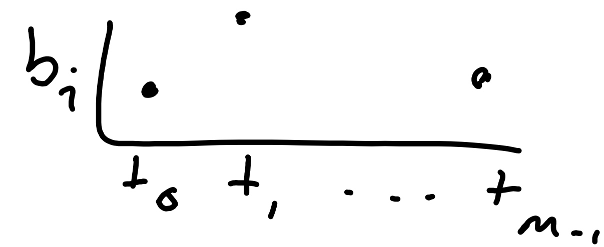
Beware: correlation and causation



Preview: Least Squares as Linear Algebra

Let's fit a line to series of data sampled over times t_0, t_1, t_2…,t_{m-1}

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



$$\overset{A}{\begin{bmatrix} 1 & t_0 \\ 1 & t_1 \\ & \vdots \\ 1 & t_{m-1} \end{bmatrix}} \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} = \begin{bmatrix} b_0 \\ \\ \\ b_{m-1} \end{bmatrix}$$

$f(t_0) = b_0$

So…how does orthogonality relate to least squares?

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

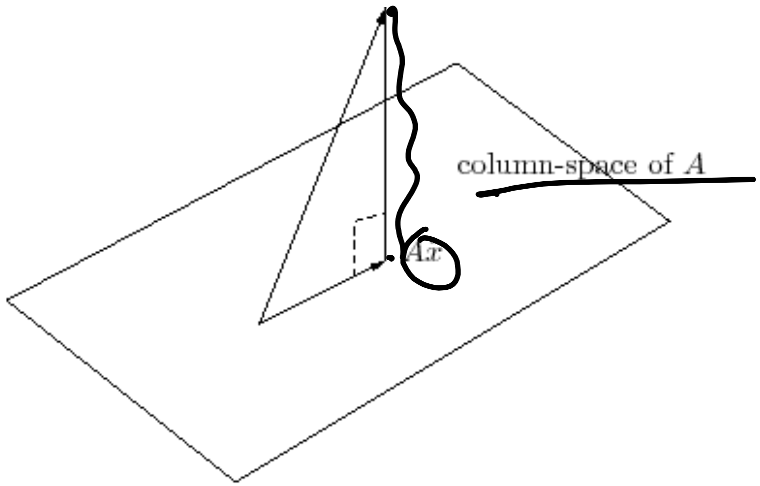


Figure J.1: Geometrical interpretation of orthogonal projection.

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?

$\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$ $\left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$

Recap: Orthonormal Basis

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

$$\vec{x} = \underbrace{(x \cdot b_1)}\ \vec{b}_1 + \underbrace{(x \cdot b_2)}\ \vec{b}_2 + \underbrace{(x \cdot b_3)}\ \vec{b}_3 + \cdots + \underbrace{(x \cdot b_n)}\ \vec{b}_n$$

Much easier than finding coefficients by solving a linear system!

Also much cheaper: $O(n^2)$

Recap: Orthonormal Basis

Can we build a matrix that computes those coefficients for us?

$$Q = \begin{pmatrix} | & & | \\ \vec{b}_1 & \cdots & \vec{b}_n \\ | & & | \end{pmatrix} \rightsquigarrow Q^T \vec{x} = \begin{pmatrix} \vec{b}_1 \cdot \vec{x} \\ \vdots \\ \vec{b}_n \cdot \vec{x} \end{pmatrix}$$

A square matrix whose columns are orthonormal is called orthogonal.

Example

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

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

$$Q = \begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}$$

$$Q^T \begin{bmatrix} \sqrt{2} \\ \sqrt{2} \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

Orthogonal Projection

What if Q contains a few zero columns instead of orthonormal vectors?

$$Q = \left[\begin{array}{ccc|c} | & & | & \\ \vec{b}_1 & \cdots & \vec{b}_k & 0 \\ | & & | & \end{array}\right]$$

orthonormal columns

Define $P := QQ^T$

Compute $P\vec{x}$ for $\vec{x} = \alpha_1 \vec{b}_1 + \cdots + \alpha_n \vec{b}_n$:

$$Q^T \vec{x} = \begin{pmatrix} \vec{x} \cdot \vec{b}_1 \\ \vdots \\ \vec{x} \cdot \vec{b}_k \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Orthogonal Projection

Define $P := QQ^T$

Compute $P\vec{x}$ for $\vec{x} = \alpha_1\vec{b}_1 + \cdots + \alpha_n\vec{b}_n$:

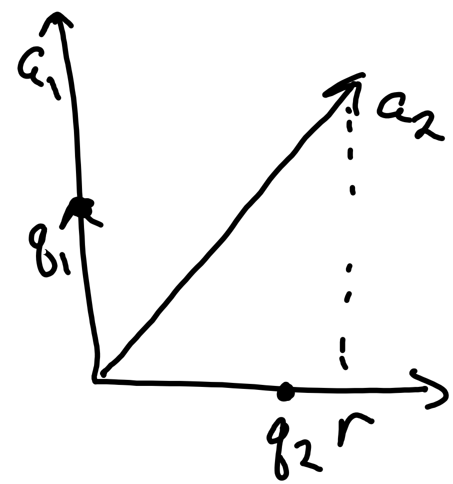
$$Q^T\vec{x} = \begin{pmatrix} \vec{x}\cdot\vec{b}_1 \\ \vdots \\ \vec{x}\cdot\vec{b}_k \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$Q(Q^T\vec{x}) = \underbrace{(x\cdot\vec{b}_1)\vec{b}_1 + \cdots + (\vec{x}\cdot\vec{b}_k)\vec{b}_k}_{\text{x projected onto } \vec{b}_1 \ldots \vec{b}_k} + \underbrace{0\cdot\vec{b}_{k+1} + \cdots + 0\cdot\vec{b}_n}_{\vec{0}}$$

Example

$$q_1 = a_1 / \|a_1\|$$

$$\begin{cases} r = a_2 - (q_1 \cdot a_2)\, q_1 \\ q_2 = r / \|r\| \end{cases}$$



Gram-Schmidt Orthogonalization

- Given linearly independent a1 and a2
- Find q1 and q2 that are orthonormal and span same space

Classical Gram-Schmidt

- We can orthogonalize any number of vectors...

current

K = column

```
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)
```

→ loop over 1 to k

orthogonalizing against column j

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

Modified Gram-Schmidt

```
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)
```