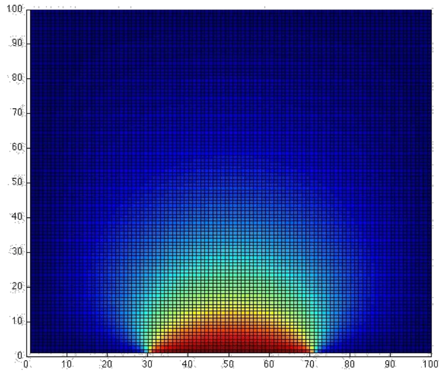


Lecture 06:
The QR Method

CS 111: Intro to Computational Science
Spring 2023

Ziad Matni, Ph.D.
Dept. of Computer Science, UCSB

Administrative

- New homework!
- Lab tomorrow!
- Preliminary slides for this lecture available

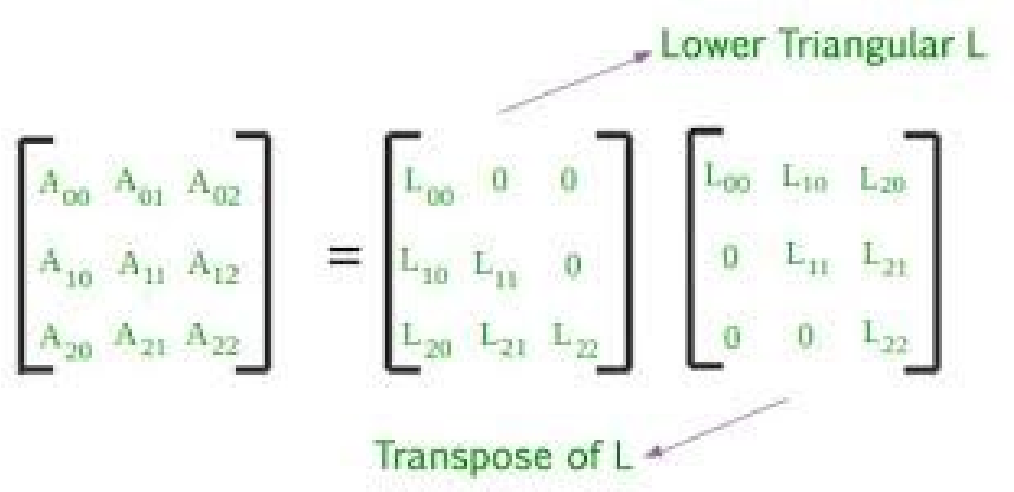
- Quiz 1 Grades are up on Canvas
  - Median: 9/9   Average: 8.8/9   Reaction: Wow!

- Assignment 1 Grades are released on Gradescope
  - Reminder: use LaTeX for answers in your submissions

Reminder: Cholesky Factorization

- Cholesky factorization is a particular form where **R** is a **lower triangular with positive diagonals**

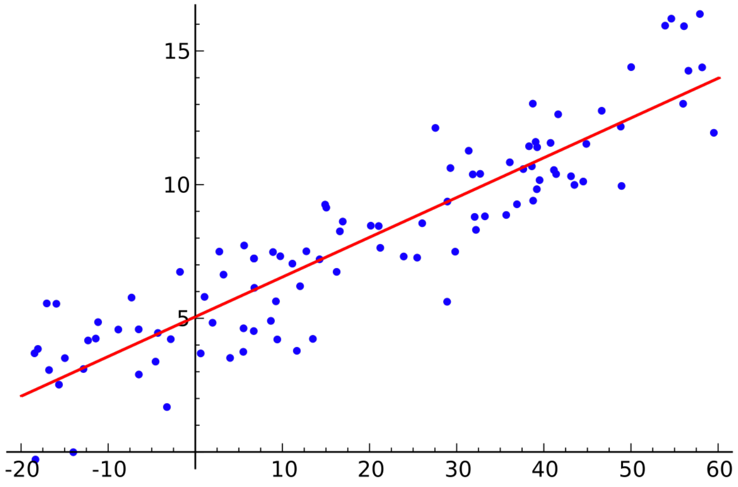
- This method should *ONLY* be used to factorize <u>**SPD matrices**</u>!



QR Factorization

- Often used to solve the **linear least squares problem**
  - An approximation of *fitting linear functions to data*
  - Re: solving statistical problems in **linear regression**

- What is **linear regression**?



QR Factorization

- **A = QR**, where **Q** is an *orthogonal matrix* based on **A**

- ***Condition of use:*** ***A*** *just has to be a* *<u>real</u> <u>square</u> matrix*
  - *i.e. it can be non-symmetrical, unlike Cholesky*

- ***Orthogonal*** matrix ➔ It's columns are "*orthonormal*"
  - Orthogonal vectors ➔ are perpendicular to each other
  - Normal vector ➔ its length is 1
  - **Orthonormal** = both orthogonal and normal

QR Factorization

Characteristic of orthogonal matrices:

- Multiplication of *any* 2 of its columns to each other (using the *dot-product*) will be equal **zero**.

  - Example: $Q = \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ 1 & 1 \end{bmatrix}$

  - The 2 columns $q_0 = [1, 2, 1]^T$ and $q_1 = [1, -1, 1]^T$ are *orthogonal vectors* because $q_0 \cdot q_1 = 1 - 2 + 1 = \mathbf{0}$

- **Property** of orthonormal matrix **Q**: $Q^T.Q = Q.Q^T = I$ *(identity matrix)*

QR Factorization

`Q,R = npla.qr(A)`

- **A = QR**, where **Q** is an *orthogonal* matrix based on **A**
  - **Property** of orthogonal matrix **Q**: $Q^T.Q = Q.Q^T = I$ *(identity matrix)*

- **R** is an *upper triangular matrix*
  - Property: If $A = Q.R$ ➔ $Q^T.A = \cancelto{I}{Q^T.Q}.R$ (multiply both sides by Q^T)
  - Therefore, **$R = Q^TA$** *(so, if you find **Q** first, then you can find **R**)*

QR Factorization

Demonstration (on blackboard) using the *Gram-Schmidt process*

https://en.wikipedia.org/wiki/Gram%E2%80%93Schmidt_process

It's an orthonormalizing process for a set of vectors (i.e **A**'s columns)

Based on the following characteristic of factor **Q**:

It's columns are *differences of normalized projections* of the columns in **A**

$$A = \begin{array}{c} \begin{bmatrix} 1 & 2 & 4 \\ 0 & 0 & 5 \\ 0 & 3 & 6 \end{bmatrix} \\ \begin{matrix} a & \; b & \; c \end{matrix} \end{array}$$

Find Q, R

dot prod.

<u>Recall</u>: orthogonal vectors $v_1 \cdot v_2 = 0$

$$Q = \begin{bmatrix} \vdots & \vdots & \vdots \\ q_1 & q_2 & q_3 \\ \vdots & \vdots & \vdots \end{bmatrix}$$

q_1, q_2, q_3 are <u>ORTHONORMAL</u>

$$q_1 = \frac{a}{\|a\|} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \times \frac{1}{\sqrt{1^2+0+0}} = \boxed{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}}$$

$q_2 = \dfrac{q_2'}{\|q_2'\|}$ where q_2' is the proj. diff. onto q_1

$$q_2' = b - (b \cdot q_1)\, q_1$$

$$= \begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix} - \Bigg(\underbrace{\begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}}_{2} \Bigg) \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix} - \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$$

<u>So:</u>

$$q_2 = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix} \cdot \frac{1}{\sqrt{0+0+3^2}} \;\; \frac{1}{3} = \boxed{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}$$

$$q_3 = \frac{q_3'}{\|q_3'\|} \quad \text{where } q_3' = c - (c \cdot q_2) q_2 - \underline{(c \cdot q_1) q_1}$$

$$q_3' = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} - \Big(\underbrace{\begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}_{6} \Big) \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - \Big(\underbrace{\begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}}_{4} \Big) \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 6 \end{bmatrix} - \begin{bmatrix} 4 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \\ 0 \end{bmatrix}$$

$$\Rightarrow q_3 = \begin{bmatrix} 0 \\ 5 \\ 0 \end{bmatrix} \cdot \frac{1}{\sqrt{5^2}} = \boxed{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}$$

$$\left.\begin{aligned} q_1 \cdot q_2 &= 0 \\ q_1 \cdot q_3 &= 0 \\ q_2 \cdot q_3 &= 0 \end{aligned}\right\} \text{Proof that they are orthonormal}$$

Since: $R = Q^T A$

$$\boxed{Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}}$$

so

$$\boxed{R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 4 \\ 0 & 0 & 5 \\ 0 & 3 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 4 \\ 0 & 3 & 6 \\ 0 & 0 & 5 \end{bmatrix}}$$

Numerical Stability

A generally desirable property of numerical algorithms

- Consider **f(x) = y** (a *mathematical* definition)
- We calculate it, using some computation, to be **y***
  - **y*** is a ***deviation*** from the "true" solution **y** (it's close in value to **y**, but not exactly the same)
  - This can happen because of round-off errors and/or truncation errors

DEFINITIONS:

- **Forward Error**: $\Delta y = y^* - y$
- **Backward Error**: Smallest Δx such that $f(x + \Delta x) = y^*$
- **Relative Error**: $|\Delta x| / |x|$

We ideally want a small Δx to give us a small Δy

Your TO DOs!

- Turn in your homework on Monday
- Lab Thursday!

</LECTURE>