QR Factorization

Note Title 2024-10-30

$$\min_{x\in\mathbb{R}^n} \|Ax-y\| \quad \text{for given } A, y \qquad x = A^+ y = (A^TA)^{-1}A^T y$$

We see a factorization similar to LU factorization

__Problem:__ $x \in \mathbb{R}^m$ given, find H orthogonal $(H^TH = HH^T = I)$ such that $Hx = \begin{bmatrix} s \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = e_1 s$ for some $s \in \mathbb{R}$

__Observation:__ since $\|Hx\| = \|x\|$, one must take $s = \pm\|x\|$

A solution is given by __Householder reflectors__

H is a __Householder reflector__ if it has the form

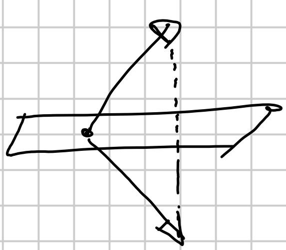
$$H = I - 2\frac{1}{v^Tv} vv^T \quad \text{for some vector } v \in \mathbb{R}^m$$

__Lemma:__ H is orthogonal and symmetric

$$H^T = \left(I - 2\frac{1}{v^Tv}vv^T\right) = I^T - \frac{2}{v^Tv}(vv^T)^T = I - \frac{2}{v^Tv}vv^T = H$$

$$H^TH = H^2 = \left(I - \frac{2}{v^Tv}vv^T\right)\left(I - \frac{2}{v^Tv}vv^T\right) = I - \frac{2}{v^Tv}vv^T - \frac{2}{v^Tv}vv^T$$

$$+ \frac{4}{(v^Tv)^2}v\cancel{(v^Tv)}v^T = I$$



H is a reflection (mirror simmetry) w.r.t. the plane normal to v

$$H = I - \frac{2}{v^T v} v v^T = I - \frac{2}{\|v\|^2} v v^T = I - 2uu^T \quad \text{with} \quad u = \frac{1}{\|v\|} \cdot v$$

<u>Obs</u>: Given $H = I - 2uu^T$ and $x \in \mathbb{R}^m$, one can compute the product $H \cdot x$ in time $O(m)$ rather than $O(m^2)$

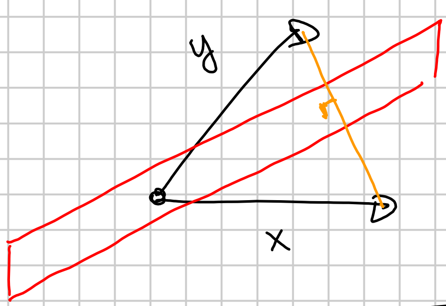
$$Hx = (I - 2uu^T)x = x - 2u(u^T x)$$

1. compute $u^T x = \alpha$ $\quad 2m + O(1)$
2. Compute $x - (2\alpha)u$ $\quad 2m + O(1)$

<u>Lemma</u>: given two vectors $x, y \in \mathbb{R}^m$, with $\|x\| = \|y\|$ the Householder reflector $H = I - \frac{2}{v^T v} v v^T$ constructed with $v = x - y$ is such that $Hx = y$.

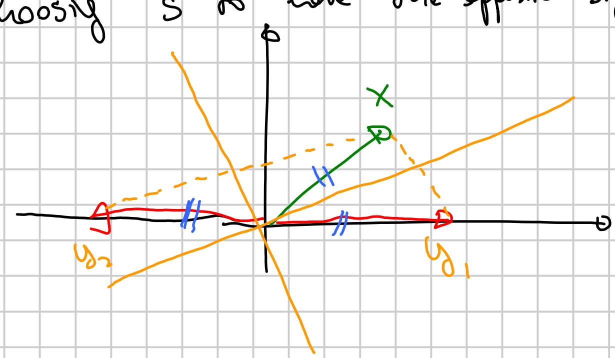
To find H such that

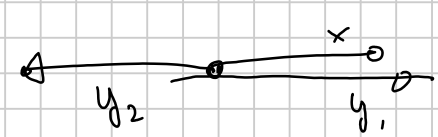
$Hx = \begin{bmatrix} s \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ $\quad s = \pm\|x\|$, we just take $y = \begin{bmatrix} s \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ in the Lemma.



$$v = x - \begin{bmatrix} s \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} - \begin{bmatrix} s \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} x_1 - s \\ x_2 \\ x_3 \\ \vdots \\ x_m \end{bmatrix}$$

By choosing s to have the opposite sign as x_1, we ensure no numerical cancellation





Theorem: for every $A \in \mathbb{R}^{m\times n}$, there exists $Q \in \mathbb{R}^{m\times m}$ orthogonal, $R \in \mathbb{R}^{m\times n}$ upper triangular such that $A = QR$.

↳ $R_{ij} = 0$ if $i > j$

Idea: work as in LU factorization / Gaussian elimination

$n=1$ case:

$$A = \begin{bmatrix} \\ \\ \end{bmatrix} \qquad HA = \begin{bmatrix} s \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \iff A = H \begin{bmatrix} s \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

already solved!

$$A = \begin{bmatrix} \boxed{\begin{matrix} x \\ x \\ x \\ x \end{matrix}} & \begin{matrix} x & x \\ x & x \\ x & x \\ x & x \end{matrix} \end{bmatrix}$$

$[u_1, s_1]$ = householder_vector$(A(1{:}\text{end}, 1))$ returns a Householder matrix such that

$$A_2 = H_1 A = \begin{bmatrix} s_1 & x & x \\ 0 & x & x \\ 0 & x & x \\ 0 & x & x \end{bmatrix} \qquad Q_1 = H_1$$

$[u_2, s_2]$ = householder_vector$(\underbrace{A2(2{:}\text{end}, 2)}_{(m-1)\times 1})$

$\underbrace{H_2}_{(m-1)\times(m-1)} \underbrace{A_2}_{m\times n} =$ ⚠ Care with dimensions!

$$H_2 = I - 2u_2u_2^T$$

$$Q_2 = \begin{bmatrix} 1 & 0 \cdots 0 \\ \begin{matrix} 0 \\ 0 \\ \vdots \\ 0 \end{matrix} & \boxed{H_2} \end{bmatrix} \in \mathbb{R}^{m\times m}$$

$$A_3\ Q_2 A_2 = \begin{bmatrix} A_2(1, :) \\ \hline H_2 A_2(2{:}\text{end}, :) \end{bmatrix} = \begin{bmatrix} x & x & x \\ 0 & x & x \\ 0 & 0 & x \\ 0 & 0 & x \end{bmatrix}$$

$[u_3, s_3]$ = householder_vector$(A3(3{:}\text{end}, 3))$

$$H_3 = I - 2u_3u_3^T \qquad Q_3 = \left[\begin{array}{c|c} I_2 & 0 \\ \hline 0 & H_3 \end{array}\right]$$

$$Q_k = \left[\begin{array}{c|c} I_{k-1} & 0 \\ \hline 0 & H_k \end{array}\right]$$ is still symmetric + orthogonal.

$$A \to Q_1 A \to Q_2 Q_1 A \to \dots \to \underbrace{Q_n Q_{n-1} \cdots Q_2 Q_1}A = R$$

$$\Updownarrow$$

$$A = \underbrace{Q_1 Q_2 \cdots Q_{n-1} Q_n}_{= Q} R$$

$= Q$. Product of orthogonal matrices $\to$ orthogonal.

Optimization: Householder arithmetic:

$$R = \begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{bmatrix} \qquad Q = \begin{bmatrix} I & 0 \\ 0 & I - 2uu^T \end{bmatrix}$$

R_{22}: $p \times q$

$$Q \cdot R = \begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} - 2u(u^T R_{22}) \end{bmatrix}$$

$p \times q - p \times 1 \; \boxed{1 \times p \;\; p \times q}$

$O(pq)$

$O(pq)$

subtraction: $O(pq)$

With this optimization, we can obtain cost $O(mn)$ per iteration $\to O(mn^2)$ in total for the line $R = Q_k \cdot R$

Problem: if $m \gg n$, we do not want to return a $m \times m$ matrix Q!

Solution 1: use the Householder vectors $u_1, \dots, u_n$ to represent Q:

$$Q = Q_1 Q_2 \cdots Q_n = \left[\, I - 2u_1u_1^T \,\right] \cdot \left[\begin{array}{c|c} 1 & \\ \hline & I - 2u_2u_2^T \end{array}\right] \cdot \dots \cdot \left[\begin{array}{c|c} I_{n-1} & \\ \hline & I - 2u_nu_n^T \end{array}\right]$$

With the u_i's, one can compute for instance Qw given $w \in \mathbb{R}^m$ in time $O(mn)$:

$$\left(\left[\, I - 2u_1u_1^T \,\right] \cdots \left(\left[\begin{array}{c|c} I_{n-1} & \\ \hline & \boxed{I - 2u_nu_n^T} \end{array}\right] w\right)\right)$$

Each factor can be multiplied by a vector (starting from the last) in time $O(m)$.

Solution 2: (thin QR factorization)

$$A = QR = \left[\, \underset{m\times n}{Q_0} \,\middle|\, \underset{m\times(m-n)}{Q_c} \,\right] \cdot \begin{bmatrix} R_0 \\ \hline 0 \end{bmatrix} \begin{matrix} \}\, n\times n \\ \}\, (m-n)\times n \end{matrix} = Q_0R_0 + Q_c \cdot 0 = Q_0R_0$$

Thin QR factorization: for all $A \in \mathbb{R}^{m\times n}$, there exist $Q_0 \in \mathbb{R}^{m\times n}$ with orthonormal columns, $R_0 \in \mathbb{R}^{n\times n}$ upper triangular such that $A = Q_0R_0$

To compute Q_0, we can use Householder arithmetic as in Solution 1 to compute

$$Q_0 = Q \begin{bmatrix} I_n \\ 0 \end{bmatrix} = \left(\left[\, I - 2u_1u_1^T \,\right] \cdot \left(\cdots \cdot \left(\left[\begin{array}{c|c} I_{n-1} & \\ \hline & I - 2u_nu_n^T \end{array}\right] \cdot \begin{bmatrix} I_n \\ 0 \end{bmatrix}\right)\right)\right)$$

Cost: $O(mn^2)$

If we handle Q in one of these ways, QR factorization costs $2mn^2 - \frac{2}{3}n^3 + O(mn)$

If $m \approx n \to \frac{4}{3}n^3$ (twice as much as LU)

If $m \gg n \to 2mn^2$ (linear in the largest dimension)