Householder Triangularization

Note Title

Before discussing Householder triangularization, we need to interpret:

★ <u>MGS as Triangular Orthogonalization</u>

Recall the modified Gram-Schmidt (MGS) algorithm.

- Initial set up : $v_j^{(1)} = a_j \quad 1 \le j \le n$
- The first step in the outer for loop:

$$\underbrace{[v_1^{(1)}\ v_2^{(1)} \cdots v_n^{(1)}]}_{=A} \underbrace{\left[\begin{array}{c|ccc} \frac{1}{r_{11}} & \frac{-r_{12}}{r_{11}} & \cdots & \frac{-r_{1n}}{r_{11}} \\ \hline 0 & & & \\ \vdots & & I_{n-1 \times n-1} & \\ 0 & & & \end{array}\right]}_{\text{call this } R_1} = \underbrace{[q_1\ v_2^{(2)} \cdots v_n^{(2)}]}_{= AR_1}$$

- The second step:

$$[q_1\ v_2^{(2)} \cdots v_n^{(2)}] \underbrace{\left[\begin{array}{cc|ccc} 1 & 0 & 0 & \cdots & 0 \\ 0 & \frac{1}{r_{22}} & \frac{-r_{23}}{r_{22}} & \cdots & \frac{-r_{2n}}{r_{22}} \\ \hline \multicolumn{2}{c|}{0} & & I_{n-2 \times n-2} & \end{array}\right]}_{\text{call this } R_2} = \underbrace{[q_1\ q_2\ v_3^{(3)} \cdots v_n^{(3)}]}_{= AR_1R_2}$$

$\vdots$

In the end, we get

$$A \underbrace{R_1 R_2 \cdots R_n}_{\text{call this } \hat{R}^{-1}} = \hat{Q} = [q_1 \cdots q_n]$$

Then $\quad A = \hat{Q}\hat{R}$!

Hence, we can view

MGS = triangular orthogonalization

meaning: applying triangular operations to reduce to orthonormal col' vectors.

<u>Note</u>: <u>In practice</u>, we do not form matrices R_i, $i = 1, \cdots, n$.
These are used to interpret the meaning of the MGS algorithm.

☆ <u>Householder Triangularization</u>

= orthogonal triangularization!

Instead of triangular orthogonalization.

$$\text{MGS}: \quad A \underbrace{R_1 R_2 \cdots R_n}_{= \hat{R}^{-1}} = \hat{Q} \qquad \text{reduced QR}$$

$$\text{Householder}: \quad \underbrace{Q_n Q_{n-1} \cdots Q_2 Q_1}_{= Q^T} A = R \qquad \text{full QR}$$

In essence, it's a triangularization by introducing zeros (Householder, 1958)

$$\underset{A}{\begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \end{bmatrix}} \xrightarrow{Q_1} \underset{Q_1 A}{\begin{bmatrix} \triangle & \triangle & \triangle \\ 0 & \triangle & \triangle \\ 0 & \triangle & \triangle \\ 0 & \triangle & \triangle \\ 0 & \triangle & \triangle \end{bmatrix}} \xrightarrow{Q_2} \underset{Q_2 Q_1 A}{\begin{bmatrix} \triangle & \triangle & \triangle \\ 0 & + & + \\ 0 & 0 & + \\ 0 & 0 & + \\ 0 & 0 & + \end{bmatrix}} \xrightarrow{Q_3} \underset{Q_3 Q_2 Q_1 A}{\begin{bmatrix} \triangle & \triangle & \triangle \\ 0 & + & + \\ 0 & 0 & \triangledown \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}$$

How to construct Q_k ?

$\Rightarrow$ **Householder Reflector**

I_{k-1} is a short notation for $I_{k-1 \times k-1}$

$$Q_k = \begin{array}{c} \left[\begin{array}{cc} I_{k-1} & 0 \\ 0 & F \end{array}\right] \begin{array}{l} \}\, k-1 \\ \}\, m-(k-1) \end{array} \\ \begin{array}{cc} \underbrace{\quad}_{k-1} & \underbrace{\quad}_{m-(k-1)} \end{array} \end{array}$$

$\underbrace{F}_{\text{Householder Reflector}} \in \mathbb{R}^{(m-k+1) \times (m-k+1)}$, $F^T F = F F^T = I_{m-k+1}$ i.e., ortho. mat.

- What F does is the following:

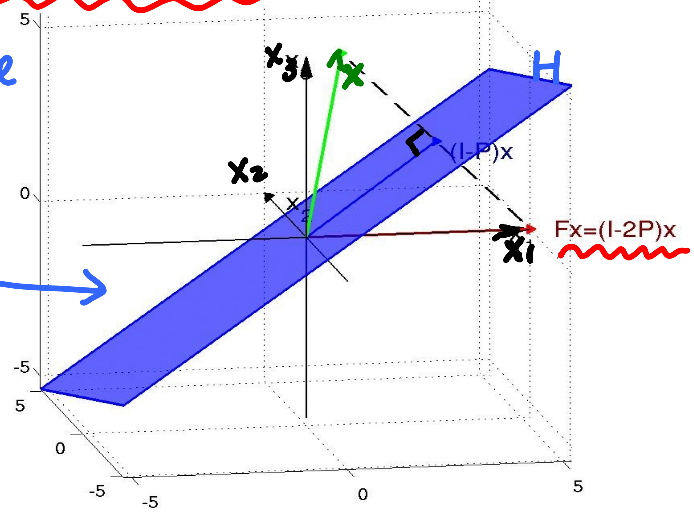
Suppose $\mathbf{x} \in \mathbb{R}^{m-k+1}$, i.e.,

$$\mathbf{x} = \begin{bmatrix} * \\ * \\ * \\ \vdots \\ * \end{bmatrix} \xrightarrow{F} F\mathbf{x} = \begin{bmatrix} \|\mathbf{x}\| \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \|\mathbf{x}\| e_1$$

- Why F is called a reflector?

Define $v = \|\mathbf{x}\| e_1 - \mathbf{x}$

H: Hyperplane normal to v



$P = P_v = \dfrac{v v^T}{v^T v}$: ortho. proj. onto $\langle v \rangle$

But $v \perp H$.

Hence $I - P_v$: ortho. proj. onto H.

Now F can be written as

$$F = I - 2P_v = I - 2\frac{v v^T}{v^T v}$$

Note F is not a projector.

but F is an orthogonal matrix.

<u>why?</u> $(I - 2P_v)^2 = (I-2P_v)(I-2P_v)$

$= I - 2P_v - 2P_v + 4P_v^2$

$= I - 4P_v + 4P_v = I$

$\neq I - 2P_v$ ✓

$(I-2P_v)^T(I-2P_v) = (I^T - 2P_v^T)(I-2P_v)$

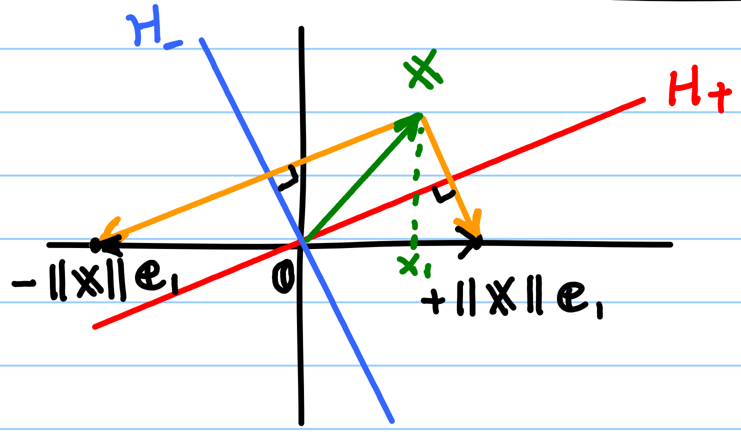
$P_v^T = P_v$ $\longrightarrow$ $= (I - 2P_v)(I-2P_v)$

$= I$ ✓

Similarly $(I-2P_v)(I-2P_v)^T = I$. ✓

☆ <u>Two possible reflectors :</u>

<u>which is better?</u>



Answer: In the above figure, $-\|\mathbf{x}\|\mathbf{e}_1$ is a better choice.

Why? It's better numerically to move $F\mathbf{x}$ as far from $\mathbf{x}$ as possible.

Why? Computing $\mathbf{v}$ involves $\mathbf{v} = \pm\|\mathbf{x}\|\mathbf{e}_1 - \mathbf{x}$. If $\|\mathbf{x}\|\mathbf{e}_1 \approx \mathbf{x}$ then $\|\mathbf{x}\|\mathbf{e}_1 - \mathbf{x}$ loses numerical accuracy called cancellation error.

Hence in general, $\mathbf{v}$ should be chosen as

The first entry of $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \end{bmatrix}$

$$\mathbf{v} = -\operatorname{sgn}(x_1)\|\mathbf{x}\|\mathbf{e}_1 - \mathbf{x}$$

where $\operatorname{sgn}(x_1) := \begin{cases} 1 & \text{if } x_1 \geq 0 \\ -1 & \text{if } x_1 < 0 \end{cases}$

or equivalently,

$$\mathbf{v} = \operatorname{sgn}(x_1)\|\mathbf{x}\|\mathbf{e}_1 + \mathbf{x}.$$

★ **Algorithm (Householder QR fact.)**

for $k = 1:n$

$$\begin{cases} \mathbf{x} = A(k:m, k) \\ \mathbf{v}_k = \operatorname{sgn}(x_1)\|\mathbf{x}\|\mathbf{e}_1 + \mathbf{x} \\ \mathbf{v}_k = \mathbf{v}_k / \|\mathbf{v}_k\| \\ A(k:m, k:n) = A(k:m, k:n) \\ \qquad - 2\mathbf{v}_k(\mathbf{v}_k^T A(k:m, k:n)) \end{cases}$$

$= F A(k:m, k:n)$

Note: In the end, A is replaced by the final R of $A = QR$.

Q has not constructed explicitly.

This is OK since QR fact. is usually used as an intermediate process for solving some other problem, e.g., $A\mathbb{x} = \mathbb{b}$ or $\min. \|\mathbb{b} - A\mathbb{x}\|^2$

as we showed before, this leads to $\underline{R\mathbb{x} = Q^T\mathbb{b}}$

Hence, as long as $Q^T\mathbb{b}$ is computed we often do not need Q itself.

Algorithm (Implicit calculation of $Q^T\mathbb{b}$)

for $k = 1 : n$

$$\mathbb{b}(k:m) = \mathbb{b}(k:m) - 2v_k(v_k^T\mathbb{b}(k:m))$$

This should be included in the Householder QR fact. algorithm in the previous page.

Or, just store v_k, $k = 1:n$.

Do not store $Q_k = I - 2v_kv_k^T$

No need to store these.

v_k, $k = 1:n$ suffice.

<u>Example</u> Let's consider the familiar example matrix.

$$A = \begin{bmatrix} 1 & 1 & 1 \\ \varepsilon & 0 & 0 \\ 0 & \varepsilon & 0 \\ 0 & 0 & \varepsilon \end{bmatrix}.$$

ε : small s.t. ε^2 can be ignored.

I'm not going through the detailed computations by the Householder reflectors here (it would take several pages to do this by hand). But in the end, you get:

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1/\sqrt{3} & \sqrt{2/3} \\ 0 & 0 & \sqrt{2/3} & 1/\sqrt{3} \end{bmatrix}}_{Q_3} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{Q_2} \underbrace{\begin{bmatrix} -1 & -\varepsilon & 0 & 0 \\ -\varepsilon & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{Q_1} A$$

$$= \underbrace{\begin{bmatrix} -1 & -1 & -1 \\ 0 & \sqrt{2}\varepsilon & \varepsilon/\sqrt{2} \\ 0 & 0 & \left(1/\sqrt{3} + \sqrt{2/3}\right)\varepsilon \\ 0 & 0 & 0 \end{bmatrix}}_{R}$$

If we want $A = QR$,

then $\underbrace{Q_3 Q_2 Q_1}_{Q^T} A = R$

note each Q_i ortho mat.

So, $$Q = (Q_3 Q_2 Q_1)^T$$
$$= Q_1^T Q_2^T Q_3^T$$
$$= Q_1 Q_2 Q_3$$

In this example, we have

$$Q = \begin{bmatrix} -1 & \varepsilon/\sqrt{2} & \varepsilon/\sqrt{6} & -\varepsilon/\sqrt{3} \\ -\varepsilon & -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 0 & \sqrt{2/3} & 1/\sqrt{3} \end{bmatrix}$$

Now let's check $Q^T Q$

$$Q^T Q = \begin{bmatrix} 1+\varepsilon^2 & 0 & 0 & 0 \\ 0 & 1+\varepsilon^2/2 & \varepsilon^2/2\sqrt{3} & -\varepsilon^2/\sqrt{6} \\ 0 & \varepsilon^2/2\sqrt{3} & 1+\varepsilon^2/6 & -\varepsilon^2/3\sqrt{2} \\ 0 & -\varepsilon^2/\sqrt{6} & -\varepsilon^2/3\sqrt{2} & 1+\varepsilon^2/3 \end{bmatrix}$$

no loss of orthogonality!

$$\approx I_{4\times 4} \; !$$

This is a big difference from CGS and even MGS!

☆ Givens Rotations

Define

$$G(i,j,\theta) := \begin{array}{cc} & \begin{array}{ccccccc} & & i & & j & & \end{array} \\ \begin{array}{c} \\ \\ i \\ \\ j \\ \\ \\ \end{array} & \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & \cos\theta & & -\sin\theta & & \\ & & & \ddots & & & \\ & & \sin\theta & & \cos\theta & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \end{array}$$

↑ all the other entries are 0.

We can choose θ s.t.

$$G(i,j,\theta) \begin{bmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_j \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ \sqrt{x_i^2 + x_j^2} \\ \vdots \\ 0 \\ \vdots \\ x_m \end{bmatrix}$$

i.e., $\tan\theta = -x_j / x_i$

So, the Householder reflection can be written as successive applications of the Givens rotations:

$$Q_k = G(k, k+1, \theta_k)\, G(k+1, k+2, \theta_{k+1}) \cdots G(m-1, m, \theta_{m-1})$$

$= m-k$ rotations!