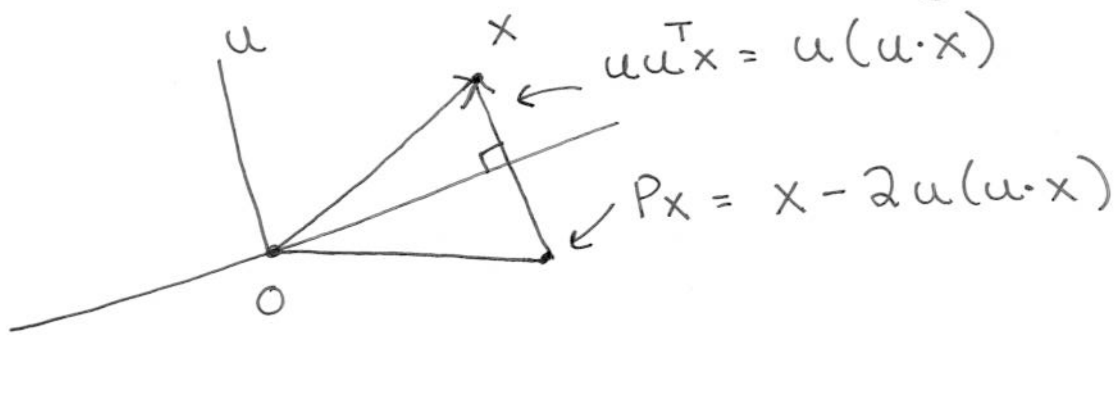
QR Decomposition in Practice.

We have seen that the Gram-Schmidt algorithm can be unstable when the input vectors are close to being linearly dependent.

How can we find another method?

Definition. A Householder reflection is a matrix of the form $P = I - 2uu^T$ where u is a vector of length 1.



Lemma. A Householder reflection is a symmetric orthogonal matrix (and hence $P^2 = I$).

Proof.

$$P^T = (I - 2uu^T)^T$$

$$= I - 2u^{T^T}u^T = I - 2uu^T = P.$$

$$PP^T = (I - 2uu^T)(I - 2uu^T)$$

$$= I - 4uu^T + 4u(u^Tu)u^T$$

$$= I. \qquad \square$$

Now given $\vec{x}$, we want to find a Householder transformation P so that

$Px = ce_1$, or P zeros all but 1st coord.

We can do this by writing

$$Px = x - 2u(u^T x) = ce_1$$

so

$$u = \frac{1}{2(u^T x)}(x - ce_1)$$

Thus u is a linear combination of x and e_1. Now $\|x\|_2 = \|Px\|_2 = \|ce_1\|_2 = |c|$,

so

$$u \parallel x \pm \|x\|_2 e_1$$

parallel to

and

$$u = \frac{x + \|x\|_2 e_1}{\|x + \|x\|_2 e_1\|_2} \quad \text{or} \quad u = \frac{x - \|x\|_2 e_1}{\|x - \|x\|_2 e_1\|_2}.$$

We will choose

this sign avoids cancellation

$$\mathrm{House}(x) := \begin{bmatrix} x_1 + \mathrm{sign}(x_1)\,\|x\|_2 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \Big/ \left\| \begin{bmatrix} x_1 + \mathrm{sign}(x_1)\,\|x\|_2 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \right\|_2$$

this vector

This defines a unit vector associated to x.

We now show how to perform QR via Householder transformations. Suppose A is a general 5×4 matrix

1. Choose P_1 so $P_1 A = \begin{bmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & x & x & x \\ 0 & x & x & x \\ 0 & x & x & x \end{bmatrix}$

2. Let $P_2 = \left[\begin{array}{c|c} 1 & \\ \hline & P_2' \end{array}\right]$ $P_2 P_1 A = \begin{bmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & 0 & x & x \\ 0 & 0 & x & x \\ 0 & 0 & x & x \end{bmatrix}$

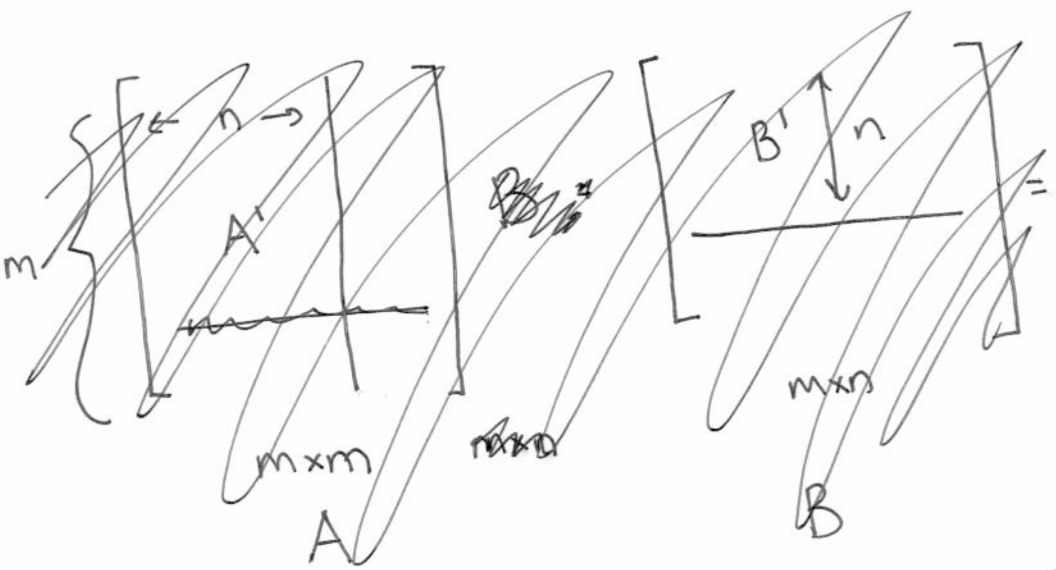
etc

4. $P_4 P_3 P_2 P_1 A = \begin{bmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & 0 & x & x \\ 0 & 0 & 0 & x \\ 0 & 0 & 0 & 0 \end{bmatrix} := \tilde{R}$

Now we Know

$$QR = A = P_1^T P_2^T P_3^T P_4^T \left(\cancel{P_4 P_3 P_2 P_1}\right) P_4 P_3 P_2 P_1 A$$
$$= P_1^T P_2^T P_3^T P_4^T \tilde{R}.$$

We can't quite conclude that $Q = P_1^T \cdots P_4^T$ since Q is supposed to be $m \times n$ and the P_i are $m \times m$. But note that if we



write

$$P_1^T P_2^T P_3^T P_4^T = P_1 P_2 P_3 P_4 = \begin{array}{cc} Q & \\ \left[\begin{array}{c|c} \uparrow & \tilde{P} \end{array}\right] & \\ m \times n & m \times (m-n) \end{array}$$

and $\tilde{R} = \begin{bmatrix} R \\ 0 \end{bmatrix}$ then $\begin{bmatrix} Q & \tilde{P} \end{bmatrix} \begin{bmatrix} R \\ 0 \end{bmatrix} = \begin{bmatrix} QR \end{bmatrix}$.

so it will suffice to let

$$Q = \text{first } n \ \underline{\text{columns}} \text{ of } P_1 \cdots P_4$$

$$R = \text{first } n \ \underline{\text{rows}} \text{ of } \tilde{R}$$

In general, we have

Algorithm. (QR using Householder)

for $i = 1$ to $\min(m-1, n)$

$u_i = \text{House}(\ A(i:m, i))$

~~lower right~~ ~~minor~~ ~~submatrix~~
last $m-i$ entries
of column i

$P_i' = I - u_i u_i^T$

$A(i:m, i:n) = P_i' A(i:m, i:n)$

end

What about stability?

Lemma. Let P be an exact Householder transformation and $\tilde{P}$ be $fl(P)$. Then

$$fl(\tilde{P}A) = P(A+E), \quad \|E\|_2 = O(\epsilon)\cdot\|A\|_2$$
$$fl(A\tilde{P}) = (A+F)P, \quad \|F\|_2 = O(\epsilon)\,\|A\|_2.$$

This comes from the roundoff error analysis for dot products. Given this lemma, we can show
↳ well, for arbitrary orthogonal matrices

Theorem. For any collection of orthogonal matrices P_i, Q_i we have

$$fl(\tilde{P}_j \cdots \tilde{P}_1 A \tilde{Q}_1 \cdots \tilde{Q}_j) = P_j \cdots P_1 (A+E) Q_1 \cdots Q_j$$

where $\tilde{P}_i = fl(P_i)$, $\tilde{Q}_i = fl(Q_i)$, and $\|E\|_2 = j\,O(\epsilon)\|A\|_2$.

Proof. Suppose $\overline{P}_j := P_j \cdots P_1$, $\overline{Q}_j := Q_1 \cdots Q_j$.
We want to show

$$A_j := fl(\tilde{P}_j A_{j-1} \tilde{Q}_j)$$
$$= \overline{P}_j (A + E_j) \overline{Q}_j$$

~~(where A_j = result of~~ for $\|E_j\|_2 = j\, O(\epsilon) \|A\|_2$.

For $j=0$, there's nothing to show. So suppose the result is true for $j-1$. Then

$$B = fl(\tilde{P}_j A_{j-1})$$
$$= P_j (A_{j-1} + E') \quad \text{by Lemma}$$
$$= P_j (\overline{P}_{j-1} (A + E_{j-1}) \overline{Q}_{j-1} + E') \quad \text{by induction}$$
$$= \overline{P}_j (A + \underbrace{E_{j-1} + \overline{P}_{j-1}^T E' \overline{Q}_{j-1}^T}_{E''}) \overline{Q}_{j-1}$$

we now want to estimate $\| E'' \|$.

We have

$$\|E''\|_2 = \|E_{j-1} + \bar{P}_{j-1}^T E' \bar{Q}_{j-1}^T\|_2$$

$$\le \|E_{j-1}\|_2 + \|E'\|_2$$

$(j-1)\,O(\epsilon)\|A\|_2$ by ~~Lemma~~ induction

$O(\epsilon)\|A\|_2$ by Lemma

$$\le j\,O(\epsilon)\|A\|_2$$

as desired. (Estimating $B\tilde{Q}_j$ is similar.) □

To compare, suppose we had used <u>non</u> orthogonal matrices. Let X be any matrix and $\tilde{X} = fl(X)$.

$$fl(\tilde{X}A) = XA + E = X(A + X^{-1}E)$$
$$= X(A + F)$$

where $\|E\|_2 \le O(\epsilon)\|X\|_2\|A\|_2$, so $\|F\|_2 \le \|X^{-1}\|_2 \cdot \|E_2\|$ or $\|F\|_2 \le O(\epsilon)\kappa_2(X)$.

So if we multiply bey a collection of such transforms, error is magnified by the <u>product</u> of condition numbers, which is <u>1</u> if the matrices are orthogonal.