QR factorization

① Gram-Schmidt (Math 330)

② Householder reflectors (new for this class)

GRAM-SCHMIDT DOES THIS:

Given vectors $v_1, v_2, \ldots, v_n$ that are linearly independent, find vectors $w_1, w_2, \ldots, w_n$ which are orthonormal such that

$$\text{span}\{v_1, \ldots, v_k\} = \text{span}\{w_1, \ldots, w_k\} \text{ for } k = 1, \ldots, n.$$

$\tilde{w}_1 = v_1$ $\qquad$ $w_1 = \dfrac{\tilde{w}_1}{\|\tilde{w}_1\|}$ Euclidean norm i.e. $p=2$.

$\underbrace{(v \cdot x)v}$ same thing projection

$\tilde{w}_2 = v_2 - (w_1 \cdot v_2) w_1$ $\qquad$ $w_2 = \dfrac{\tilde{w}_2}{\|\tilde{w}_2\|}$

note $w_1 \cdot \tilde{w}_2 = 0$

$w_1 \cdot \tilde{w}_2 = w_1 \cdot (v_2 - (w_1 \cdot v_2) w_1)$

$= \cancel{w_1 \cdot v_2} - \cancel{(w_1 \cdot v_2)}(\cancelto{1}{w_1 \cdot w_1}) = 0$

$\tilde{w}_3 = v_3 - (w_1 \cdot v_3) w_1 - (w_2 \cdot v_3) w_2$ $\qquad$ $w_3 = \dfrac{\tilde{w}_3}{\|\tilde{w}_3\|}$

triangle of orange-circled coefficients make R.

⋮

$\tilde{w}_n = v_n - (w_1 \cdot v_n) w_1 - \cdots - (w_{n-1} \cdot v_n) w_{n-1}$ $\qquad$ $w_n = \dfrac{\tilde{w}_n}{\|\tilde{w}_n\|}$

if vectors were not indep. eventually a zero here in the denominator.

Focus is factoring a matrix $A \in \mathbb{R}^{m \times n}$ (cols: n, rows: m)

into QR

orthogonal matrix — Q; upper triangular... — R

so $m \geq n$...

Need the columns of A to be linearly indep.

$$A = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}$$

$$\xrightarrow{\text{Gram-schmidt}} \quad \tilde{Q} = \begin{bmatrix} w_1 & w_2 & \cdots & w_n \end{bmatrix}$$

$$\boxed{\tilde{Q}^T \tilde{Q} = \begin{bmatrix} w_1^T \\ w_2^T \\ \vdots \\ w_n \end{bmatrix} \begin{bmatrix} w_1 & w_2 & \cdots & w_n \end{bmatrix} = I_{n \times n}}$$

Try to factor $A \in \mathbb{R}^{m \times n}$

$$\underset{m \times n}{A} = \underset{m \times n}{\tilde{Q}} \; \underset{n \times n}{\tilde{R}}$$

reduced QR decomposition because Q might not be square

$$\tilde{Q}^T A = \underbrace{\tilde{Q}^T \tilde{Q}}_{I} \tilde{R} = \tilde{R}$$

Let $\tilde{R} = \tilde{Q}^T A$ and then $A = \tilde{Q}\tilde{R}$.

Note reduced $\tilde{Q}\tilde{R}$ works fine for solving least squares problems...

How to extend $\tilde{Q}$ to a square matrix?

just add some more orthonormall vectors to $\tilde{Q}$ untill it's square

$$Q = \Big[\ \underset{m\times n}{\tilde{Q}}\ \Big|\ \underbrace{w_{n+1}\ \Big|\ w_{n+2}\ \Big|\ \cdots\ \Big|\ w_m}_{m-n \text{ more vectors}}\ \Big] \in \mathbb{R}^{m\times m}$$

only requirement is there are orthonormal with each other and with the ones that came from Gram-Schmidt

Now

$Q^TQ = I_{m\times m}$ but also $QQ^T = I_{m\times m}$ $\rightarrow$ $Q^T = Q^{-1}$

$$\underset{m\times n}{A} = \underset{m\times m}{Q}\ \underset{m\times n}{R}$$

$$R = \begin{bmatrix} \tilde{R} \\ \hline 0 \\ \hline 0 \\ \hline \vdots \\ \hline 0 \end{bmatrix} \left.\vphantom{\begin{matrix}0\\0\\\vdots\\0\end{matrix}}\right\} m-n \text{ rows of zeros...}$$

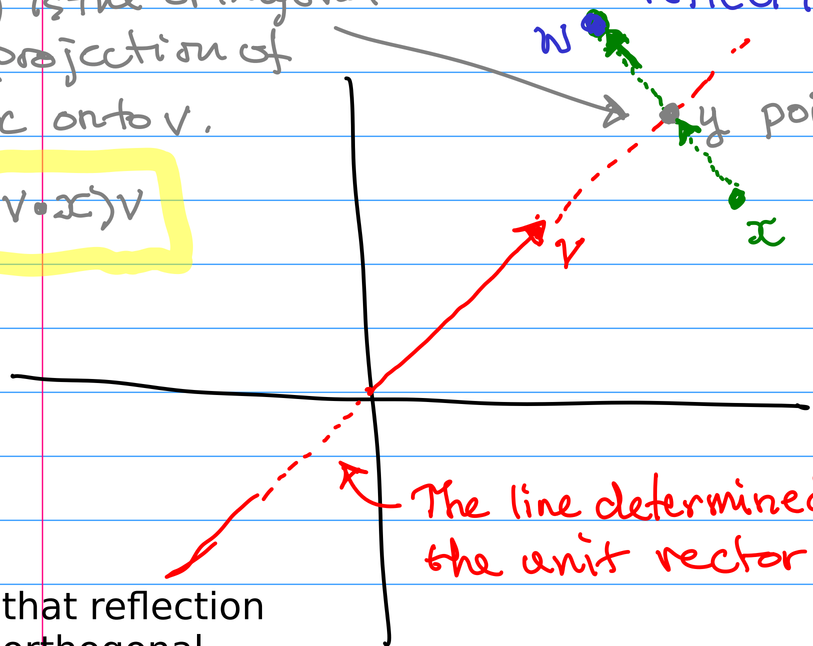
Algorithm to find Q directly...

What is a Householder reflector? H

reflection of w across v.

y is the orthogonal projection of x onto v.

$$\boxed{y = (v \cdot x) v}$$



y point of intersection

x the point to reflect

The line determined by the unit vector v

Note that reflection is an orthogonal tranformation...

$$\begin{aligned} w &= y + (w - y) \\ &= y + (y - x) \\ &= 2y - x \\ &= 2\underbrace{(v \cdot x) v}_{\text{factor out the } x} - x \\ &= 2(v^T x) v - x \\ &= 2v(v^T x) - x \\ &= 2 v v^T x - x \\ &= (2vv^T - I)x \end{aligned}$$

$$w = -Hx$$

Identity

rank 1 matrix

Let $H = I - 2vv^T$

different v...

Idea multiply A by H to make some that is upper triangular one column at a time

$\pm \|a_1\|$

$$HA = H \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} = \begin{bmatrix} c & Ha_2 & \cdots & Ha_n \\ 0 & & & \\ 0 & & & \\ \vdots & & & \\ 0 & & & \end{bmatrix}$$

$$(I - 2vv^T)a_1 = \begin{bmatrix} c\\0\\0\\0\\0\end{bmatrix} = ce_1 \qquad \text{where } e_1 = \begin{bmatrix}1\\0\\0\\\vdots\\0\end{bmatrix}$$

vector in this equation are $v, a_1, e_1 \in \mathbb{R}^m$

$$(I-2vv^T)a_1 = ce_1$$

First solve for c:

$$v^T(I-2vv^T)a_1 = v^T c e_1 = cv_1$$

$$v^Ta_1 - 2\cancelto{1}{v^Tv}\,v^Ta_1 = cv_1$$

$$v^Ta_1 - 2v^Ta_1 = cv_1$$

$$\boxed{-v^Ta_1 = cv_1}$$ substitute

$$a_1^T(I-2vv^T)a_1 = a_1^T c e_1 = ca_{11}$$

$$a_1^Ta_1 - 2(a_1^Tv)(v^Ta_1) = ca_{11}$$

$$\|a_1\|^2 - 2(a_1^Tv)^2 = ca_{11}$$

$$\|a_1\|^2 - 2c^2v_1^2 = ca_{11}$$

these equations are left

$$e_1^T(I-2vv^T)a_1 = e_1^T c e_1 = c$$

$$a_{11} - 2\boxed{e_1^Tv}\,v^Ta_1 = c$$

$$a_{11} - 2v_1v^Ta_1 = c$$

$$a_{11} + 2cv_1^2 = c$$

$$\|a_1\|^2 - \cancel{2c^2v_1^2} = \cancel{ca_{11}}$$

$$\cancel{ca_{11}} + \cancel{2c^2v_1^2} = c^2$$

$$\|a_1\|^2 = c^2$$

$$c = \pm\|a_1\|$$

Solve for v:

For simpliciy continue to write c, even though we know what it is

$$(I - 2vv^T)a_1 = ce_1$$

thats the

$$a_1 - 2v(v^Ta_1) = ce_1$$

$$a_1 - ce_1 = 2v(v^Ta_1)$$

just a number

Since v is a unit vector then

$$v = \frac{a_1 - ce_1}{2v^Ta_1} = \frac{a_1 - ce_1}{\|a_1 - ce_1\|}$$

as big as possible

$$c = \pm\|a_1\|$$

how to choose + or − ?

If $a_{11} > 0$ then take $c = -\|a_1\|$

$a_{11} < 0$ then take $c = +\|a_1\|$

If a_{11} is complex valued then one could check $\|a_1 \pm \|a_1\| e_1\|$ to see which one is bigger or simply compare $\mathrm{Re}\, a_{11}$ as and follow the same choice as in the real case.