How to find the QR factorization of a matrix A in a way that is numerically stable...

Recall Gram-Schmidt algorithm: (based on column operations to step by step transform the matrix A into the matrix $\tilde{Q}$ which has orthonormal columns).

$$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \in \mathbb{R}^{m\times n} \qquad \tilde{Q} = \begin{bmatrix} q_1 & q_2 & \cdots & q_n \end{bmatrix} \in \mathbb{R}^{m\times n}$$

normalization

$t_1 = a_1$ $\qquad q_1 = \dfrac{t_1}{\|t_1\|_2}$

subtract the part of a_2 that is not perpendicular to a_1

$t_2 = a_2 - (q_1 \cdot a_2) q_1$ $\qquad q_2 = \dfrac{t_2}{\|t_2\|_2}$

$t_3 = a_3 - (q_1 \cdot a_3) q_1 - (q_2 \cdot a_3) q_2$ $\qquad q_3 = \dfrac{t_3}{\|t_3\|_2}$

$\vdots$ $\qquad \vdots$

propagation of all rounding errors from the entire calculation..

$t_n = a_n - (q_1 \cdot a_n) q_1 - \cdots - (q_{n-1} \cdot a_n) q_{n-1}$ $\qquad q_n = \dfrac{t_n}{\|t_n\|_2}$

Note also the algorithm follows a fixed order with not pivoting to reduce rounding error when they are propagated from previous steps..

IDEAS

- Try to reorder the column operations to get larger denominators in the normalization step.
- To nest the projections so the errors in previous steps get projected away. replace

$$t_3 = a_3 - (q_1 \cdot a_3) q_1 - (q_2 \cdot a_3) q_2$$

with

$$t_{3,1} = a_3 - (q_1 \cdot a_3) q_1$$

$$t_{3,2} = t_{3,1} - (q_2 \cdot t_{3,1}) q_2$$

projects off some of the rounding error that happened in $t_{3,1}$.

What happens with all the accumulated rounding error is the vectors $q_1, q_2, \dots, q_n$ that were supposed to be perpendicular to each other, aren't really...

$$\tilde{Q}^T \tilde{Q} = \begin{bmatrix} - & q_1^T & - \\ - & q_2^T & - \\ & \vdots & \\ - & q_n^T & - \end{bmatrix} \begin{bmatrix} q_1 & q_2 & \cdots & q_n \end{bmatrix} = \begin{bmatrix} q_1^T q_1 & q_1^T q_2 & \cdots & q_1^T q_n \\ q_2^T q_1 & q_2^T q_2 & & q_2^T q_n \\ \vdots & & \ddots & \\ q_n^T q_1 & q_n^T q_2 & \cdots & q_n^T q_n \end{bmatrix}$$

$$= \begin{bmatrix} q_1^T q_1 & & 0 \\ & q_2^T q_2 & \\ & & \ddots \\ 0 & & \cdots q_n^T q_n \end{bmatrix} = \begin{bmatrix} 1 & & 0 \\ & 1 & \\ & & \ddots \\ 0 & & 1 \end{bmatrix} = I$$

When solving least squares this cancellation was important for fixing the conditioning.

$$\tilde{R}^T \underbrace{\cancel{\tilde{Q}^T} \cancel{\tilde{Q}}}_{\text{this}} \tilde{R} x = \tilde{R}^T \tilde{Q}^T b$$

$$\cancel{\tilde{R}^T} \tilde{R} x = \cancel{\tilde{R}^T} \tilde{Q}^T b$$

Cancelling the $\tilde{R}^T$ here also...

Thus, it's important for our approximate QR factorization to satisfy $Q^T Q = I$ as good as possible.

Note that $\tilde{R}$ is what we directly calculate with Gram-Schmidt

$t_1 = a_1$ $\qquad q_1 = \dfrac{t_1}{\|t_1\|_2}$

subtract the part of a_2 that is not perpendicular to a_1

$t_2 = a_2 - (q_1 \cdot a_2) q_1$ $\qquad q_2 = \dfrac{t_2}{\|t_2\|_2}$

$t_3 = a_3 - (q_1 \cdot a_3) q_1 - (q_2 \cdot a_3) q_2$ $\qquad q_3 = \dfrac{t_3}{\|t_3\|_2}$

$\vdots$ $\qquad \vdots$

propagation of all rounding errors from the entire calculation..

$t_n = a_n - (q_1 \cdot a_n) q_1 - \cdots - (q_{n-1} \cdot a_n) q_{n-1}$ $\qquad q_n = \dfrac{t_n}{\|t_n\|_2}$

$$\tilde{R} = \begin{bmatrix} \|t_1\|_2 & (q_1 \cdot a_2) & (q_1 \cdot a_3) & \cdots & (q_1 \cdot a_n) \\ & \|t_2\|_2 & (q_2 \cdot a_3) & \cdots & \vdots \\ & & \|t_3\|_2 & & \vdots \\ & 0 & & \ddots & (q_{n-1} \cdot a_n) \\ & & & & \|t_n\|_2 \end{bmatrix}$$

Idea ... Focus on making sure Q satisfies $Q^T Q = I$.

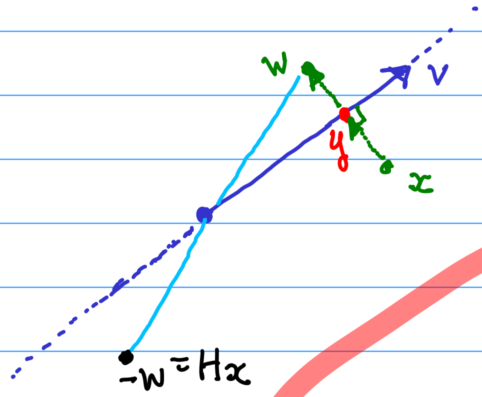
Recall Gram-Schmidt algorithm. (based on column operations. to step by step transform the matrix A into the matrix $\tilde{Q}$ which has orthonormal columns).

Note this gives the reduced $\tilde{Q}\tilde{R}$ factorization

Householder method: (based on reflections that step by step transform the matrix A into R.)

Note this gives the full QR factorization

Householder reflector... reflect a point x about a line given by the unit vector v.



The two vectors $\vec{yw}$ and $\vec{xy}$ are the same

$$\boxed{w-y} = y-x$$

The projection of x on v is y.

$$y = (v\cdot x)v$$

Now solve for w:

linear function of x

$$w = \boxed{w-y} + y = y - x + y = 2y - x = 2(v\cdot x)v - x$$

write a linear function as matrix multiplication..

$$w = 2(v\cdot x)v - x = 2v(\underbrace{v\cdot x}_{v^Tx}) - x = 2v(v^Tx) - x$$

$$= 2v(v^Tx) - x = (2vv^T)x - Ix = (\underbrace{2\overbrace{vv^T}^{\text{outer product}} - I}_{\text{matrix}})x$$

Define $H = -(2vv^T - I) = I - 2vv^T$

this is two reflection... first the reflection about the line given by v follow by reflection through the origin.

Householder reflector... view the minus sign as tradition..

Check that H is an orthogonal matrix (clearly square)

$$H^TH = (I - 2vv^T)^T(I - 2vv^T) = I - 2vv^T - 2vv^T + 4vv^Tvv^T$$

$$I - 4vv^T + 4v\underbrace{v^Tv}_{v\cdot v = 1}v^T = I - \cancel{4vv^T} + \cancel{4vv^T} = I$$

Now find a unit vector v such that

$$HA = H\begin{bmatrix} a_1 & a_2 & \cdots & a_n \\ \vdots & \vdots & & \vdots \end{bmatrix} = \begin{bmatrix} Ha_1 & Ha_2 & \cdots & Ha_n \\ \vdots & \vdots & & \vdots \end{bmatrix} = \begin{bmatrix} c & Ha_2 & \cdots & Ha_n \\ 0 & \vdots & & \vdots \\ 0 & & & \\ \vdots & & & \\ 0 & & & \end{bmatrix}$$

Idea is find v so that $Ha_1 = ce_1$ where $e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$