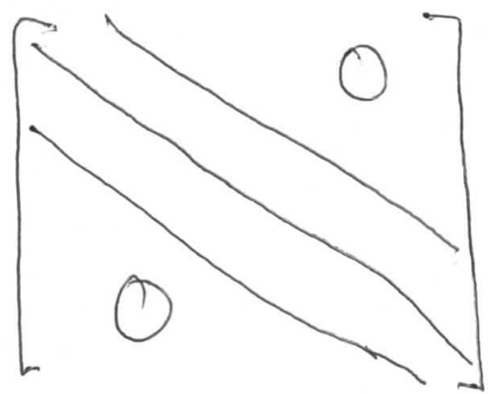
The QR algorithm

A matrix A is said to be <u>tridiagonal</u> iff $a_{ij} = 0$ for $|i-j| > 1$.



Let $\underline{e}$ ~~the~~ be a unit vector in $\mathbb{R}^n$.

The matrix $(I - 2\underline{e}\,\underline{e}^T)$ is

- An orthogonal matrix (why?)
- A symmetric matrix (why?)
- The matrix representing the reflection in a plane orthogonal to $\underline{e}$. (why?)

given a vector $\underline{a} \in \mathbb{R}^n$ let $\underline{e}_a$ be given by $(\underline{a} - \|\underline{a}\|_2 \underline{e}_1) / \|\underline{a} - \|\underline{a}\|_2 \underline{e}_1\|$ (a unit vector) where $\underline{e}_1$ denotes the first canonical basis vector. Then

$$R_a = (I - 2\, \underline{e}_a \underline{e}_a^T)$$

satisfies

$$R_a \underline{a} = \|\underline{a}\|_2 \underline{e}_1$$

Pf:

$$\underline{e}_a \underline{e}_a^T \underline{a} = \frac{(\underline{a} - \|\underline{a}\|_2 \underline{e}_1)(\underline{a} - \|\underline{a}\|_2 \underline{e}_1)^T \underline{a}}{\langle \underline{a} - \|\underline{a}\|_2 \underline{e}_1 ,\ \underline{a} - \|\underline{a}\|_2 \underline{e}_1 \rangle}$$

$$= \frac{(\underline{a} - \|\underline{a}\|_2 \underline{e}_1)(\|\underline{a}\|_2^2 - a_1 \|\underline{a}\|_2)}{\|\underline{a}\|_2^2 - 2\|\underline{a}\|_2 a_1 + \|\underline{a}\|_2^2}$$

$$= \frac{1}{2}(\underline{a} - \|\underline{a}\|_2 \underline{e}_1)$$

and thus

$$R_{\underline{a}}\,\underline{a} = \underline{a} - (\underline{a} - \|\underline{a}\|_2\,\underline{e}_1)$$

$$= \|\underline{a}\|_2\,\underline{e}_1 \qquad \square$$

Let A be an $\underline{n \times n}$ real $\underline{\text{symmetric}}$ matrix

$$A = \left[\begin{array}{c|ccc} a_{11} & a_{12} & \cdots & a_{1n} \\ \hline & & & \\ \underline{a}' & & A' & \\ & & & \end{array}\right]$$

where $\underline{a}' \in \mathbb{R}^{n-1}$ & A' is an $(n-1) \times (n-1)$ matrix

let $R_{\underline{a}'}$ be the $(n-1)\times(n-1)$ reflection matrix from before with

$$R_{\underline{a}'}\,\underline{a}' = \|\underline{a}'\|_2\,\underline{e}_1$$

Then

$$\begin{bmatrix} 1 & 0 \; \cdots \; 0 \\ 0 & \\ \vdots & R_{a'} \\ 0 & \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \; \cdots \; a_{1m} \\ \underline{a}' & A' \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & a_{12} \; \cdots \; a_{1m} \\ R_{a'}\underline{a}' & R_{a'}A' \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & a_{12} \; \cdots \; a_{1m} \\ \|\underline{a}'\|_2 & \\ 0 & \\ \vdots & R_{a'}A' \\ 0 & \end{bmatrix} \quad \text{and therefore}$$

$$\begin{bmatrix} 1 & 0 \; \cdots \; 0 \\ 0 & \\ \vdots & R_{a'} \\ 0 & \end{bmatrix} A \begin{bmatrix} 1 & 0 \; \cdots \; 0 \\ 0 & \\ \vdots & R_{a'}^T \\ 0 & \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & \|\underline{a}'\|_2 \; 0 \cdots 0 \\ \|\underline{a}'\|_2 & \begin{matrix} b_{11} & \underline{b}'^T \\ \underline{b}' & B' \end{matrix} \\ 0 & \\ \vdots & \\ 0 & \end{bmatrix} \quad \text{why?}$$

Now we have

$$\begin{bmatrix} \begin{matrix}1&0\\0&1\end{matrix} & 0 \\ 0 & \underbrace{\boxed{R_{\underline{b}'}}}_{(n-2)\times(n-2)} \end{bmatrix} \begin{bmatrix} a_{11} & \|\underline{a}'\|_2 & \\ \|\underline{a}'\|_2 & b_{11} & \underline{b}'^T \\ 0 & \underline{b}' & \underbrace{\boxed{B'}}_{(n-2)\times(n-2)} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & \|\underline{a}'\|_2 & 0 \\ \|\underline{a}'\|_2 & b_{11} & \underline{b}^T \\ & \|\underline{b}'\|_2 & \\ 0 & 0 & \boxed{R_{\underline{b}'}B'} \end{bmatrix}$$

and so

$$\begin{bmatrix} \begin{matrix}1&0\\0&1\end{matrix} & \\ & \boxed{R_{\underline{b}'}} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \boxed{R_{\underline{a}'}} \end{bmatrix} A \begin{bmatrix} 1 & 0 \\ 0 & R_{\underline{a}'} \end{bmatrix}^T \begin{bmatrix} \begin{matrix}1&0\\0&1\end{matrix} & 0 \\ 0 & R_{\underline{b}'} \end{bmatrix}^T$$

$$= \begin{bmatrix} a_{11} & \|\underline{a}'\|_2 & 0 & \cdots & 0 \\ \|\underline{a}'\|_2 & b_{11} & \|\underline{b}'\|_2 & 0 \cdots & 0 \\ 0 & \|\underline{b}'\|_2 & & & \\ \vdots & 0 & & \boxed{\times} & \\ 0 & \vdots & & & \\ & 0 & & & \end{bmatrix}$$

why?

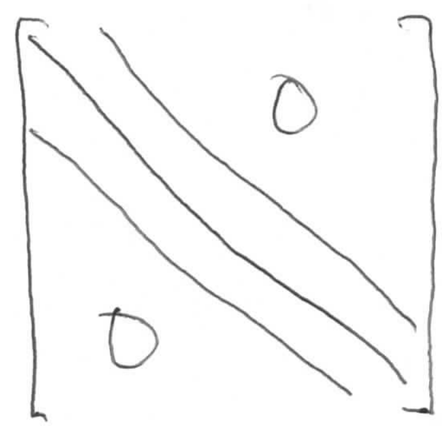
$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & R_a' & \\ 0 & & & \end{bmatrix} \quad \& \quad \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & & & \\ 0 & & & & \\ \vdots & & & R_b' & \\ 0 & & & & \end{bmatrix}$$

are orthogonal and so is the product (why?) in other words

$$QAQ^T = \begin{bmatrix} a_{11} & \|\underline{a}'\|_2 & 0 & \cdots & 0 \\ \|\underline{a}'\|_2 & b_{11} & \|\underline{b}'\|_2 & 0 \cdots & 0 \\ 0 & \|\underline{b}'\|_2 & & & \\ \vdots & 0 & & & \\ & \vdots & & & \\ 0 & 0 & & & \end{bmatrix}$$

by continuing this way we end up with

$QAQ^T =$



<u>tridiagonal</u>

this form of the matrix A is called tridiagonal <u>Hessenberg</u> <u>form</u>.

Once arrived at the Hessenberg tridiagonal form H_0 of the symmetric matrix A, then we write it as

$$H_0 = Q R$$

upper triangular

orthogonal

and then form

$$H_1 := R Q$$

whereupon we continue with the iteration

$$H_k = Q R$$
$$H_{k+1} := R Q$$

as many times as necessary

The matrices H_k will converge to a diagonal matrix, and since all H_k have same eigenvalues as H (and thus as A) the diagonal entries are the eigenvalues of A. Including a shift will make the convergence faster...

(why?)