A deep-dive into the QR algorithm

- Basic form of QR:

$$A_{m-1} = Q_m \cdot R_m$$
$$A_m = R_m \cdot Q_m$$

1. Why can we expect convergence at all?

Denote the eigenvectors by v_j and associated eigenvalues as λ_j.

If S is a 1-dim subspace then the sequence

$S, AS, A^2 S, \dots$

converges to the eigenspace $T = \text{span}\{v_1\}$ if $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \dots$.

Correspondingly, if S consists of k vectors, then the sequence

$A, AS, A^2 S, \dots$

converges to the space $T = \text{span}\{v_1, \dots, v_k\}$ if $|\lambda_k| > |\lambda_{k+1}|$.

Now let $Q^{(m)}$ be an othogonal basis of $A^m S$.

We have

$$[Q^{(m)}]^H A Q^{(m)} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

Let $Q^{(m)} = [q_0^{(m)}, q_1^{(m)}, \dots, q_k^{(m)}]$.

Compute $Z^{(m+1)} = A \cdot Q^{(m)}$

and let $Q^{(m+1)} \cdot R = Z^{(m+1)}$

be its QR decomposition.

It follows that also the columns $q_0^{(m)}, \dots, q_j^{(m)}$ converge for each $j \leq k$ against an invariant subspace.

Hence, B_{11} converges to an upper triangular matrix.

For $k = n$ therefor

$$[Q^{(m)}]^H A Q^{(m)} \to R$$

upper triangular

The QR Iteration

Let $A^{(0)} = A$

That's it !!
$$\begin{cases} A^{(m)} = Q^{(m)} \cdot R^{(m)} \\ A^{(m+1)} = R^{(m)} Q^{(m)} \end{cases}$$

Why does this work?

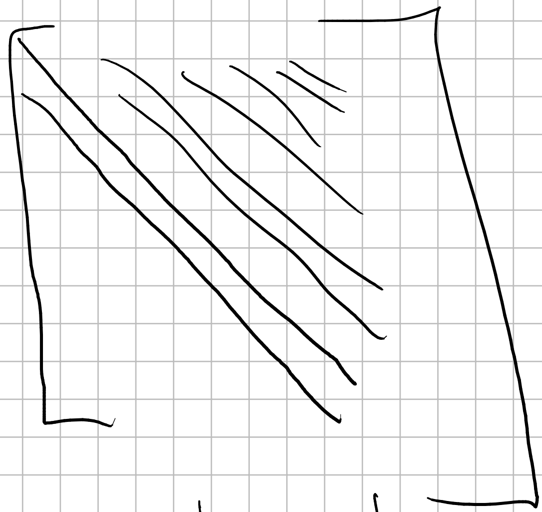
Have
$$A^{(m+1)} = \underbrace{\left[Q^{(m)}\right]^H A^{(m)} Q^{(m)}}_{\text{Looks like the testing from simultaneous iteration.}}$$

With a bit of algebra can show that with $A^{(0)} = \underline{I}$ simultaneous iteration produces the same sequence $A^{(m)}$!!

Speeding up QR.

1.) Transform A to Hessenberg form.

A matrix H is upper Hessenberg if it is upper triangular + first subdiagonal is nonzero

$H =$ 

We can construct H explicitly through eigenvalue preserving transformation

$$H = Q^H A Q$$

in $O(n^3)$ ops.

The QR iteration step

$$H^{(m)} = Q^{(m)} R^{(m)}$$
$$H^{(m+1)} = R^{(m)} \cdot Q^{(m)}$$

preserves Hessenberg form and can be done in $O(n^2)$ ops.

2.) Use shifts

$$A^{(m)} - \sigma I = Q^{(m)} R^{(m)}$$
$$A^{(m+1)} = R^{(m)} Q^{(m)} + \sigma I.$$

How to choose σ?

Can show that for $\sigma = A^{(m)}[n,n]$ [bottom right elem of $A^{(m)}$] the QR iteration acts like a Rayleigh-quotient iteration on the last columns of the Q factors [the exact result is a bit more involved]

$\Rightarrow$ Expect fast convergence in the last column.

$$A^{(m)} = \begin{bmatrix} a_{11}^{(m)} & \ddots & & \\ a_{21}^{(m)} & \ddots & \ddots & \\ & \ddots & \ddots & \ddots \\ 0 & & a_{n,n-1}^{(m)} & a_{nn}^{(m)} \end{bmatrix}$$

Upon convergence $a_{n,n-1}^{(m)} \to 0$.

If sufficiently small, reduce the matrix and only continue with the first $n-1$ rows/cols!