Arnoldi

Given

$$A \in \mathbb{R}^{n \times n}, \quad b \in \mathbb{R}, \quad \|b\|_2 = \alpha_0$$

Let's investigate the Gram-Schmidt orthogonalization of the columns of

$$C = [b, \ Ab, \ \ldots, \ A^{m-1}b].$$

The Arnoldi method does this, giving a Hessenberg matrix H such that

$$AV = VH + we_m^t, \quad V^tV = I_m, \quad V^tw = 0.$$

Let's assume that $h_{j+1,j} \neq 0, \ j = 1, 2, \ldots, m-1$ (thus H is *unreduced*). The orthogonal projector onto the column space of C is $V(V^tV)^{-1}V^t = VV^t$, thus

$$VV^tA^kb = A^kb, \quad k \leq m-1,$$

and the Gram-Schmidt QR factorization we wanted is

$$C = VR, \quad \text{with} \quad R \equiv \alpha_0[e_1, \ He_1, \ \ldots, \ H^{m-1}e_1] \quad \text{and} \quad r_{ii} = \alpha_0 \prod_{j=1}^{m-1} h_{j+1,j}.$$

Now

$$HRe_k = Re_{k+1}, \quad k = 1, 2, \ldots m-1,$$

means the companion matrix (rational canonical form) for H is

$$F = R^{-1}HR.$$

So, if q_H is the characteristic polynomial of H, then

$$q_H(x) = x^m - \sum_{i=0}^{m-1} c_i x^i, \quad \text{where } c = R^{-1}HRe_m.$$

Finally, if $p(x) = x^m - \sum_{i=0}^{m-1} a_i x^i$, then $p(A)b = ACe_m - Ca$, and remarkably

$$q_H = \underset{p \in \mathcal{M}onic_m}{\text{argmin}} \ \|p(A)b\|_2.$$

We have just connected a Hessenberg form, a QR factorization, a companion form, and a variational principle all under the Arnoldi umbrella, but can't resist squeezing in a few more players: An unreduced Hessenberg matrix with multiple eigenvalues is not diagonalizable. Thus, if our A is diagonalizable, then its eigenvalues are distinct. Suppose $AX = X\Lambda$ is a diagonalization of A. Then

$$\begin{aligned} C &= [b, \ (X\Lambda X^{-1})b \ , (X\Lambda X^{-1})^2 b, \ldots, \ (X\Lambda X^{-1})^{m-1}b] \\ &= X[X^{-1}b, \ \Lambda X^{-1}b, \ \Lambda^2 X^{-1}b, \ldots, \ \Lambda^{m-1}X^{-1}b] \\ &= X \ \text{diag}(X^{-1}b) \ [e, \ \Lambda e, \ \Lambda^2 e, \ldots, \ \Lambda^{m-1}e] \\ &\equiv YW. \end{aligned}$$

Hopefully you recognize the rightmost factor W as a Vandermonde matrix on the eigenvalues $\lambda_1, \ \lambda_2, \ldots, \ \lambda_n$. Now if we take $m = n$, $F = R^{-1}HR = C^{-1}AC$ is a companion form of A, whose eigenvalues are Λ, while the eigenvectors of A are the columns of $Y = X \ \text{diag}(X^{-1}b) = CW^{-1}$.