The Q-R Algorithm of Kublanovskaya & Francis

Michael A. Epton

October 19, 2013

Summary of Results

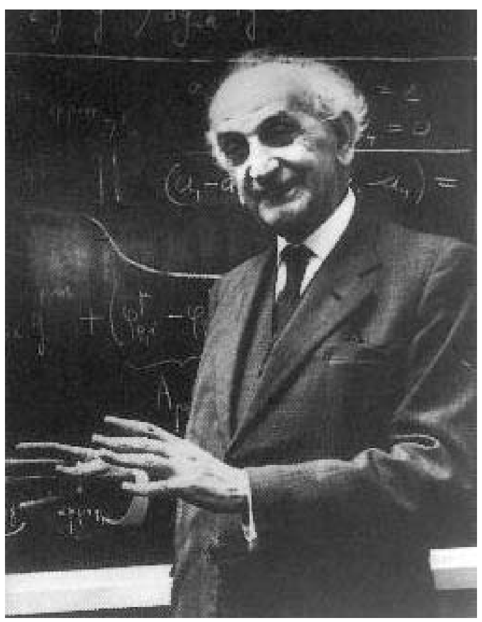
- We consider the eigenvalue problem: Find $\mathbf{x} \in \mathbf{C}^{\nu}$, $\lambda \in \mathbf{C}$ such that $A\mathbf{x} = \lambda\mathbf{x}$.
- Formulate as a *root finding problem* in $\nu + 1$ unknowns: $F(X) = 0$.
- Apply Newton's Method. Obtain formula for the Newton **decrement**: $\Delta = [\delta\mathbf{x}_n, \delta\lambda_n]$
- Key facts: Application of Newton's Method plus insight reveals the algorithmic structure:

  $F(X) = 0 \Rightarrow$ Newton's method $\Rightarrow$ Wielandt Inverse Iteration $\Rightarrow$ Q-R algorithm
- More: If the target λ is simple and known, next iteration determines $\mathbf{x}$ *exactly*
- Motivates getting the *best possible estimate* of the eigenvalue
- This leads us to consider *Inverse Taylor Series Iteration* of degree p
- Newton's Method is the $p = 1$ special case of this family
- We will obtain the $p = 2$ Inverse Taylor Series update to $\mathbf{x}$ and λ
- We will show that the Newton update to λ_n has error $O(\Delta^3)$ in the Hermitian case
  This implies that the Q-R algorithm is cubically convergent in the Hermitian case
- Obtain *all* the terms of Inverse Taylor Series by reversion of series

Heroes of our tale








Root Finding Formulation of Eigenvalue Problem

- Given matrix $A \in \mathbf{C}^{\nu\times\nu}$, find $\mathbf{x} \in \mathbf{C}^{\nu}$, $\lambda \in \mathbf{C}$ such that $A\mathbf{x} = \lambda\mathbf{x}$
- Formulate as a root finding problem in $\nu+1$ variables: $F(X) = 0$ where:

$$X = \begin{bmatrix} \mathbf{x} \\ \lambda \end{bmatrix} \qquad F(X) \stackrel{\text{def}}{=} \begin{bmatrix} (\lambda I - A)\mathbf{x} \\ \mathbf{z}^*\mathbf{x} - 1 \end{bmatrix}$$

  The choice of $\mathbf{z}$ is rather arbitrary. We will change it every iteration

- Last condition is *normalization*: $\mathbf{z}^*\mathbf{x} = 1$. The condition $\mathbf{x}^*\mathbf{x} = 1$ would not work!
- It would lead to an $F(X)$ that is *not analytic* in X.
- In fact we *will* renormalize $\mathbf{x}_n$ each iteration. Also we will set $\mathbf{z} = \mathbf{x}_n$.

Application of Newton's Method to Solving $F(X)=0$

- Apply Newton's Method to find $\Delta = [\delta\mathbf{x}_n, \delta\lambda_n]$ s.t. $\mathbf{x}_{n+1} = \mathbf{x}_n - \delta\mathbf{x}_n$, $\lambda_{n+1} = \lambda_n - \delta\lambda_n$:

$$(F_X)_{X_n}\,\Delta = F(X_n) \quad \Longrightarrow \quad \begin{bmatrix} \lambda_n I - A & \mathbf{x}_n \\ \mathbf{z}^* & 0 \end{bmatrix} \begin{bmatrix} \delta\mathbf{x}_n \\ \delta\lambda_n \end{bmatrix} = \begin{bmatrix} (\lambda_n I - A)\mathbf{x}_n \\ \mathbf{z}^*\mathbf{x}_n - 1 \end{bmatrix}$$

- The Newton decrement relation can be recast as

$$\begin{bmatrix} \lambda_n I - A & \mathbf{x}_n \\ \mathbf{z}^* & 0 \end{bmatrix} \begin{bmatrix} \delta\mathbf{x}_n - \mathbf{x}_n \\ \delta\lambda_n \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

- The inverse of the coefficient matrix is:

$$\begin{bmatrix} \lambda_n I - A & \mathbf{x}_n \\ \mathbf{z}^* & 0 \end{bmatrix}^{-1} = \begin{bmatrix} E/\gamma & \mathbf{g} \\ \mathbf{w}^* & -\gamma \end{bmatrix}, \qquad \begin{array}{lll} \gamma = 1\big/\left(\mathbf{z}^*(\lambda_n I - A)^{-1}\mathbf{x}_n\right), & & G = \gamma(\lambda_n I - A)^{-1} \\ \mathbf{w}^* = \mathbf{z}^* G, & \mathbf{g} = G\mathbf{x}_n, & E = G - \mathbf{g}\mathbf{w}^* \end{array}$$

- Using the matrix inverse formula we find

$$\begin{bmatrix} \delta\mathbf{x}_n - \mathbf{x}_n \\ \delta\lambda_n \end{bmatrix} = \begin{bmatrix} -\mathbf{g} \\ \gamma \end{bmatrix} \quad \Longrightarrow \quad \mathbf{x}_{n+1} = \mathbf{x}_n - \delta\mathbf{x}_n = \mathbf{g} = G\mathbf{x}_n, \qquad \delta\lambda_n = \gamma$$

Newton's Method, Inverse Iteration and the Q-R Algorithm

- We have obtained the Newton iterate $\mathbf{x}_{n+1}$ and eigenvalue decrement $\delta\lambda_n$

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \delta\mathbf{x}_n = \mathbf{g} = G\mathbf{x}_n, \qquad \delta\lambda_n = \gamma \;=\; 1\Big/\left(\mathbf{z}^*(\lambda_n I - A)^{-1}\mathbf{x}_n\right), \qquad G \;=\; \gamma(\lambda_n I - A)^{-1}$$

- Focusing on $\mathbf{x}_{n+1}$, observe its relation to Wielandt Inverse Iteration:

$$\mathbf{x}_{n+1} \;=\; G\mathbf{x}_n \;=\; \delta\lambda_n\,(\lambda_n I - A)^{-1}\,\mathbf{x}_n$$

  - G is the Inverse Iteration matrix. $\mathbf{w}^* = \mathbf{z}^* G$ estimates a *left* eigenvector by inverse iteration

- The L-Q variant of the Q-R algorithm is based on the following setup:
  - The matrix A has been reduced to lower Hessenberg form
  - We obtain the L-Q factorization: $\lambda_n I - A \;=\; LQ$, L lower triangular, Q unitary
  - Q is consequently lower Hessenberg: $Q = L^{-1}(\lambda_n I - A)$
  - $\mathbf{x}_n = \mathbf{e}_\nu$ (last natural unit vector) and $(\mathbf{z}^*)_\nu = 1$. This leads to the following development

$$\begin{aligned} \mathbf{x}_{n+1} \;&=\; \delta\lambda_n(\lambda_n I - A)^{-1}\mathbf{x}_n \;=\; \delta\lambda_n(LQ)^{-1}\mathbf{e}_\nu \;=\; \delta\lambda_n Q^{-1}L^{-1}\mathbf{e}_\nu \\ \Longrightarrow \qquad \mathbf{x}'_{n+1} \;&=\; Q\mathbf{x}_{n+1} \;=\; \delta\lambda_n L^{-1}\mathbf{e}_\nu \;=\; (\delta\lambda_n/\ell_{\nu,\nu})\,\mathbf{e}_\nu \end{aligned}$$

  - Absolutely crucial: $\mathbf{x}'_{n+1}$ is *parallel* to $\mathbf{e}_\nu$: This is half the magic of Q-R!

How the Q-R Algorithm Works

- We make the following observations:
  - The vector $\mathbf{x}'_{n+1} = Q\mathbf{x}_{n+1}$ is parallel to $\mathbf{e}_\nu$
  - If we change basis using Q and renormalize, $\mathbf{x}'_{n+1}$ can be set equal to $\mathbf{e}_\nu$
  - In the new coordinate system, $A' = QAQ^{-1}$. Consequently:

$$A' = QAQ^{-1} = Q(\lambda_n I - (\lambda_n I - A))Q^{-1} = \lambda_n I - Q(LQ)Q^{-1} = \lambda_n I - QL$$

  - Because Q is lower Hessenberg and L is lower triangular, A' is again lower Hessenberg
- Key features: By *changing basis* using Q, we achieve these useful results:
  - The matrix A' (similar to A) remains lower Hessenberg (tridiagonal if Hermitian)
  - The Newton update to the eigenvector, $\mathbf{x}'_{n+1} = Q\mathbf{x}_{n+1}$, remains parallel to $\mathbf{e}_\nu$
  - Preservation of lower Hessenberg form reduces computational cost by factor of ν.

Condition of the Iteration Matrix at Convergence

- At convergence to $[\mathbf{x}_n, \mu]$ we expect that $A\mathbf{x}_n = \mu\mathbf{x}_n$ and $\mathbf{z}^* A = \mu \mathbf{z}^*$.
  - Put μ in the last position of the Jordan Normal form; assume it is simple
  - This implies $\mathbf{x}_n = V\mathbf{e}_\nu$, and $\mathbf{z}^* = \mathbf{e}_\nu^* V^{-1}$
- Using $V^{-1}AV = J$, we find that $\lambda_n = \mu$, $\mathbf{x}_n = V\mathbf{e}_\nu$ and $\mathbf{z}^* = \mathbf{e}_\nu^* V^{-1}$ imply

$$\begin{bmatrix} \lambda_n I - A & \mathbf{x}_n \\ \mathbf{z}^* & 0 \end{bmatrix} = \begin{bmatrix} \mu V V^{-1} - VJV^{-1} & V\mathbf{e}_\nu \\ \mathbf{e}_\nu^* V^{-1} & 0 \end{bmatrix} = \begin{bmatrix} V & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mu I - J & \mathbf{e}_\nu \\ \mathbf{e}_\nu^* & 0 \end{bmatrix} \begin{bmatrix} V^{-1} & 0 \\ 0 & 1 \end{bmatrix}$$

- For μ simple, the matrix in the middle and its spectral condition number is given

$$\begin{bmatrix} \mu I_{\nu-1} - J_{\nu-1} & \mathbf{0}_{\nu-1} & \mathbf{0}_{\nu-1} \\ \mathbf{0}_{\nu-1}^T & 0 & 1 \\ \mathbf{0}_{\nu-1}^T & 1 & 0 \end{bmatrix} \qquad cond_\rho = \frac{\max\left(1, \ \underset{\lambda_j \neq \mu}{max} \ |\mu - \lambda_j|\right)}{\min\left(1, \ \underset{\lambda_j \neq \mu}{min} \ |\mu - \lambda_j|\right)}$$

- Provided V is well conditioned we conclude these facts about $\mathbf{E}$, $\mathbf{g}$ and $\mathbf{w}$:

$$E = O(\gamma) \qquad \mathbf{g} = O(1) \qquad \mathbf{w}^* = O(1)$$

- For $\gamma = \delta\lambda_n \approx 0$ this implies $G = \mathbf{g}\mathbf{w}^* + E = \mathbf{g}\mathbf{w}^* + O(\delta\lambda_n) \implies G$ approximately rank 1;

Convergence in One Iteration if $\lambda_n = \mu$, a Simple Eigenvalue

- Recast the iteration relation as

$$\begin{bmatrix} A-\lambda_n I & \mathbf{x}_n \\ \mathbf{z}^* & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x}_{n+1} \\ \delta\lambda_n \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix} \qquad \text{put } \lambda_n = \mu \text{ and place in position } \nu$$

- Recalling the Jordan Normal Transformation $A = VJV^{-1}$ we employ V to find

$$\begin{bmatrix} J-\mu I & V^{-1}\mathbf{x}_n \\ \mathbf{z}^* V & 0 \end{bmatrix} \begin{bmatrix} V^{-1}\mathbf{x}_{n+1} \\ \delta\lambda_n \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix}$$

- Using the fact that $(J)_{\nu,\nu} = \mu = \lambda_n$, this implies

$$\begin{bmatrix} J_{\nu-1}-\mu I_{\nu-1} & \mathbf{0}_{\nu-1} & (V^{-1}\mathbf{x}_n)_{1:\nu-1} \\ \mathbf{0}_{\nu-1}^T & 0 & (V^{-1}\mathbf{x}_n)_\nu \\ (\mathbf{z}^* V)_{1:\nu-1} & (\mathbf{z}^* V)_\nu & 0 \end{bmatrix} \begin{bmatrix} (V^{-1}\mathbf{x}_{n+1})_{1:\nu-1} \\ (V^{-1}\mathbf{x}_{n+1})_\nu \\ \delta\lambda_n \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{\nu-1} \\ 0 \\ 1 \end{bmatrix}$$

- This implies that

eqn ν	$(V^{-1}\mathbf{x}_n)_\nu\, \delta\lambda_n = 0$	$\Longrightarrow$	$\delta\lambda_n = 0$	(no surprise here!)
eqns $1:(\nu-1)$	$(J_{\nu-1}-\mu I_{\nu-1})(V^{-1}\mathbf{x}_{n+1})_{1:\nu-1} = \mathbf{0}_{\nu-1}$	$\Longrightarrow$	$(V^{-1}\mathbf{x}_{n+1})_{1:\nu-1} = \mathbf{0}_{\nu-1}$	
conclude:	$(V^{-1}\mathbf{x}_{n+1}) = \theta\mathbf{e}_\nu$	$\Longrightarrow$	$\mathbf{x}_{n+1} = \theta V\mathbf{e}_\nu$	

- We find $\mathbf{x}_{n+1} \,||\, V\mathbf{e}_\nu$, the last column of V, the eigenvector associated with μ
- **Observation:** Getting λ_n as close as possible to μ is crucial; corresponds to the shift in the Q-R algorithm
- This observation motivated study of the Inverse Taylor Series Method.
- Its stronger convergence enabled proof of cubic convergence for Hermitian problems

The Inverse Taylor Series Method

- Consider determining X^* such that $F(X^*) = 0$.
  - Let G denote the inverse function such that $G(F(X)) = X$. Clearly $X^* = G(0)$
  - Let X_n be our current estimate of X^* and let $Y_n = F(X_n)$; consider the series:

$$G(Y) = G(Y_n) + G_Y(Y - Y_n) + \frac{1}{2} G_{YY}(Y - Y_n)(Y - Y_n) + \frac{1}{6} G_{YYY}(Y - Y_n)(Y - Y_n)(Y - Y_n) + \ldots$$

- Set $Y = 0$ to evaluate X^*. Use $X_n = G(F(X_n)) = G(Y_n)$

$$X^* = G(0) = X_n - G_Y(Y_n) + \frac{1}{2} G_{YY}(Y_n)(Y_n) - \frac{1}{6} G_{YYY}(Y_n)(Y_n)(Y_n) + \ldots$$

  - The Inverse Taylor Series Method of order p retains the first $p+1$ terms.
  - Newton's Method is the special case of $p = 1$. It retains just the first 2 terms.
  - Convergence Analysis is easy: $X_{n+1} - X^* = O(X_n - X^*)^{p+1}$
  - Trivially, for Newton's Method: $X_{n+1} - X^* = O(X_n - X^*)^2$

- Apply this to our $F(X)$ for the eigenvalue problem:

$$F(X) \stackrel{\text{def}}{=} \begin{bmatrix} (\lambda I - A)\mathbf{x} \\ \mathbf{z}^*\mathbf{x} - 1 \end{bmatrix} \qquad (F_X)_{X_n} = \begin{bmatrix} \lambda_n I - A & \mathbf{x}_n \\ \mathbf{z}^* & 0 \end{bmatrix}$$

- The first order term is the Newton decrement:

$$G_Y(Y_n) = F_X^{-1} F(X_n) = \Delta = \begin{bmatrix} \delta \mathbf{x}_n \\ \delta \lambda_n \end{bmatrix} = \begin{bmatrix} \mathbf{x}_n - \mathbf{x}_{n+1} \\ \delta \lambda_n \end{bmatrix} = \begin{bmatrix} \mathbf{x}_n - G\mathbf{x}_n \\ \gamma \end{bmatrix}$$

Computation of 2nd Order Inverse Taylor Series Term

- We focus on the 2nd order update term: $(1/2)G_{YY}(Y_n)(Y_n)$. Consider the development

$$G(F(X))X \quad \Longrightarrow \quad G_Y F_X = I \quad \Longrightarrow \quad G_{YY}F_XF_X + G_YF_{XX} = 0$$

Apply last identity to two copies of Δ $\quad G_{YY}(F_X\Delta)(F_X\Delta) + G_Y(F_{XX}\Delta\Delta) = 0$

Use Newton relation $F_X\Delta = Y_n$, and conclude $\quad G_{YY}Y_nY_n = -G_Y(F_{XX}\Delta\Delta)$

- Doing all the algebra, the second order update term is:

$$\frac{1}{2}G_{YY}(Y_n)(Y_n) \quad = \quad -F_X^{-1}\begin{bmatrix} \delta\lambda_n\delta\mathbf{x}_n \\ 0 \end{bmatrix} \quad = \quad -\begin{bmatrix} E\delta\mathbf{x}_n(\delta\lambda_n/\gamma) \\ \mathbf{w}^*\delta\mathbf{x}_n\delta\lambda_n \end{bmatrix}$$

- Proof that Q-R is cubically convergent for Hermitian matrices
  - Setting $\mathbf{z} = \mathbf{x}_n$, we will show that the quadratic update to λ_n is actually $O(\Delta^3)$

$$\tau_2 \quad = \quad -\mathbf{w}^*\delta\mathbf{x}_n\delta\lambda_n \quad = \quad -\mathbf{z}^*G\delta\mathbf{x}_n\delta\lambda_n \quad = \quad -\mathbf{x}_n^*G\delta\mathbf{x}_n\delta\lambda_n$$

  - Focusing on $\mathbf{x}_n^*G\delta\mathbf{x}_n$:

$$\mathbf{x}_n^*G\delta\mathbf{x}_n \quad = \quad (G^*\mathbf{x}_n)^*\delta\mathbf{x}_n \quad = \quad (G\mathbf{x}_n)^*\delta\mathbf{x}_n \quad = \quad \mathbf{x}_{n+1}^*\delta\mathbf{x}_n \quad = \quad (\mathbf{x}_n - \delta\mathbf{x}_n)^*\delta\mathbf{x}_n \quad = \quad \mathbf{x}_n^*\delta\mathbf{x}_n - \delta\mathbf{x}_n^*\delta\mathbf{x}_n$$

  - But, provided $\mathbf{z}^*\mathbf{x}_n = 1$, the Newton update formula asserts $\mathbf{x}_n^*\delta\mathbf{x}_n = \mathbf{z}^*\delta\mathbf{x}_n = 0$
  - Conclude that $\quad \mathbf{x}_n^*G\delta\mathbf{x}_n = -\delta\mathbf{x}_n^*\delta\mathbf{x}_n \quad \Longrightarrow \quad \tau_2 = |\delta\mathbf{x}_n|^2\delta\lambda_n$ ... third order!

Solving $F(X) = 0$ by series reversion: Setup

- We attack directly the problem of computing $G(Y)$ defined implicity by $G(F(X)) = X$.
- Setting $Y_n = F(X_n)$, expand $G(Y)$ in a series about Y_n:

$$G(Y) \;=\; G(Y_n) + \sum_{r=1}^{\infty} \frac{1}{r!} G_{YY\ldots Y}(Y - Y_n)^r$$

  where the action of the derivative tensor $G_{YY\ldots Y}$ on r copies of V is interpreted as:

$$G_{YY\ldots Y}(V)^r \sim \sum_{j_1=1}^{\nu+1} \sum_{j_2=1}^{\nu+1} \cdots \sum_{j_r=1}^{\nu+1} \frac{\partial^r G^i}{\partial Y^{j_1} \partial Y^{j_2} \ldots \partial Y^{j_r}} V^{j_1} V^{j_2} \ldots V^{j_r}$$

- Similarly, the quadratic function $F(X)$ is expanded in a series about X_n

$$F(X) \;=\; F(X_n) + F_X(X - X_n) + \frac{1}{2} F_{XX}(X - X_n)^2$$

- Make the substitutions $G(Y) = X,\; G(Y_n) = X_n,\; F(X) = Y,\; F(X_n) = Y_n$ and obtain

$$X \;=\; X_n + \sum_{r=1}^{\infty} \frac{1}{r!} G_{YY\ldots Y}(Y - Y_n)^r$$

$$Y \;=\; Y_n + F_X(X - X_n) + \frac{1}{2} F_{XX}(X - X_n)^2$$

- Introduce shifted variables $U = X - X_n$, $V = Y - Y_n$ along with linear and multilinear operators $A_1 = F_X, \quad A_2 = (1/2)F_{XX}, \quad B_r = (1/r!)G_{YY\ldots Y}$ and find

$$U \;=\; \sum_{r=1}^{\infty} B_r(V)^r, \qquad V \;=\; A_1 U + A_2(U)^2$$

Solving $F(X) = 0$ by series reversion: Recursion

- Substitute the first of these equations into the second and obtain the identity

$$V = A_1 \left(\sum_{r=1}^{\infty} B_r (V)^r \right) + A_2 \left(\sum_{r=1}^{\infty} B_r (V)^r \right) \left(\sum_{s=1}^{\infty} B_s (V)^s \right)$$

- Matching terms by their degree in V, we obtain the identities

$$\begin{aligned} V &= A_1 B_1 V \\ 0 &= A_1 B_r (V)^r \quad + \quad A_2 \textstyle\sum_{s=1}^{r-1} (B_s V^s)\ (B_{r-s} V^{r-s}) \end{aligned}$$

- Evaluation: We want $U = X - X_n$ corresponding to $Y = 0$, i.e. $V = 0 - Y_n = -Y_n$

$$X - X_n = U = \sum_{r=1}^{\infty} (-1)^r B_r (Y_n)^r$$

- Define $Z_r = B_r (Y_n)^r$ and obtain the recursions

$$Y_n = A_1 Z_1 \qquad\qquad 0 = A_1 Z_r + A_2 \left(\textstyle\sum_{s=1}^{r-1} Z_s Z_{r-s} \right) \quad (r \geq 2)$$

- Solve for Z_r (Δ is the Newton decrement):

$$Z_1 = A_1^{-1} Y_n = F_X^{-1} F(X_n) = \Delta \qquad Z_r = -A_1^{-1} \left(A_2 \textstyle\sum_{s=1}^{r-1} Z_s Z_{r-s} \right) \quad (r \geq 2)$$

Solving $F(X) = 0$ by series reversion: Application

- Partitioning Z_r into vector and scalar parts: $Z_r = [\mathbf{z}_r, \zeta_r]$, we compute the sum in Z_r's definition

$$A_2 \sum_{s=1}^{r-1} Z_s Z_{r-s} = \frac{1}{2} \sum_{s=1}^{r-1} F_{XX} \begin{bmatrix} \mathbf{z}_s \\ \zeta_s \end{bmatrix} \begin{bmatrix} \mathbf{z}_{r-s} \\ \zeta_{r-s} \end{bmatrix} = \frac{1}{2} \sum_{s=1}^{r-1} \begin{bmatrix} \zeta_s \mathbf{z}_{r-s} + \zeta_{r-s} \mathbf{z}_s \\ 0 \end{bmatrix} = \sum_{s=1}^{r-1} \begin{bmatrix} \zeta_s \mathbf{z}_{r-s} \\ 0 \end{bmatrix}$$

- Apply $A_1^{-1} = F_X^{-1}$ to this expression and obtain Z_r

$$\begin{bmatrix} \mathbf{z}_r \\ \zeta_r \end{bmatrix} = Z_r = -A_1^{-1} \sum_{s=1}^{r-1} \begin{bmatrix} \zeta_s \mathbf{z}_{r-s} \\ 0 \end{bmatrix} = - \begin{bmatrix} E/\gamma \\ \mathbf{w}^* \end{bmatrix} \sum_{s=1}^{r-1} \zeta_s \mathbf{z}_{r-s}$$

- Introduce cooefficients $\theta_s = \zeta_s/\gamma$, obtain recursions

$$\begin{bmatrix} \mathbf{z}_r \\ \theta_r \end{bmatrix} = - \begin{bmatrix} E \\ \mathbf{w}^* \end{bmatrix} \sum_{s=1}^{r-1} \theta_s \mathbf{z}_{r-s}$$

- Usually $\mathbf{z}_r = O(\Delta^r)$ and $\theta_r = O(\Delta^{r-1})$. (note: $\theta_1 \equiv 1$)

  But for Hermitian problems, $\theta_r = O(\Delta^r)$ $(r \geq 2)$ because $\mathbf{x}_n^* E = 0$ (superconvergence!)

- Evaluation of eigenvector/eigenvalue pair: $X = X_n + \sum_{r=1}^{\infty} (-1)^r Z_r$

- Reversion of series not usually a good idea. It works here because $\deg(F) = 2$.

Conclusions

- Formulating the eigenvalue as a root finding problem provides several useful insights
  - — Application of Newton's method produces Wielandt Inverse Iteration
  - — Use of the LQ factorization to solve the system leads to the Q-R algorithm
  - — Quadratic convergence of the Q-R algorithm follows trivially
- Other important insights
  - — Exact eigenvalue $\Longrightarrow$ Convergence on next iteration
  - — Motivates getting best possible eigenvalue estimate: Inverse Taylor Series
  - — Cubic convergence of Q-R in the Hermitian case follows from quadratic Taylor Series
- Getting the best possible eigenvalue estimate is crucial!
  - — Originally developed Inverse Taylor Series Method up through $p = 4$
  - — A web-search turned up notion of Series Reversion
  - — That worked well: All terms are easily computed by a recursion
  - — Consequently: Cost of solving dense eigenvalues problem can be minimized
  - — Also, superconvergence in the Hermitian case persists to all orders
  - — Result: Hermitian at $\sim$ 1.45 iterations/eigenvalue; Non-hermitian at $\sim$ 2.1

Verification of Inversion Formula

- Form the product:

$$\begin{bmatrix} \lambda_n I - A & \mathbf{x}_n \\ \mathbf{z}^* & 0 \end{bmatrix} \begin{bmatrix} E/\gamma & \mathbf{g} \\ \mathbf{w}^* & -\gamma \end{bmatrix} = \begin{bmatrix} (\lambda_n I - A)E/\gamma + \mathbf{x}_n \mathbf{w}^* & (\lambda_n I - A)\mathbf{g} - \gamma \mathbf{x}_n \\ \mathbf{z}^* E/\gamma & \mathbf{z}^* \mathbf{g} \end{bmatrix}$$

with definitions:
$$\begin{aligned} \gamma &= 1 \big/ \left(\mathbf{z}^*(\lambda_n I - A)^{-1}\mathbf{x}_n\right), & G &= \gamma(\lambda_n I - A)^{-1} \\ \mathbf{w}^* &= \mathbf{z}^* G, \quad \mathbf{g} = G\mathbf{x}_n, & E &= G - \mathbf{g}\mathbf{w}^* \end{aligned}$$

- Check the terms of the RHS product. Start with (1,2), then do (1,1).

(1,2) term: $$(\lambda_n I - A)\mathbf{g} - \gamma \mathbf{x}_n = (\lambda_n I - A)\gamma(\lambda_n I - A)^{-1}\mathbf{x}_n - \gamma \mathbf{x}_n = \gamma \mathbf{x}_n - \gamma \mathbf{x}_n = 0$$

(1,1) term: $$\begin{aligned} (\lambda_n I - A)E/\gamma + \mathbf{x}_n \mathbf{w}^* &= \frac{1}{\gamma}\{(\lambda_n I - A)\left[G - \mathbf{g}\mathbf{w}^*\right] + \gamma \mathbf{x}_n \mathbf{w}^*\} \\ &= I + \frac{1}{\gamma}\{-(\lambda_n I - A)\mathbf{g} + \gamma \mathbf{x}_n\}\mathbf{w}^* = I + \frac{1}{\gamma}\{0\}\mathbf{w}^* = I \end{aligned}$$

- By virtue of the Newton update formula, the (2,2) term is given by

$$\mathbf{z}^*\mathbf{g} = \mathbf{z}^* G \mathbf{x}_n = \mathbf{z}^* \mathbf{x}_{n+1} = 1$$

- The (2,1) term can now be easily verified:

$$\mathbf{z}^* E/\gamma = \mathbf{z}^*(G - \mathbf{g}\mathbf{w}^*)/\gamma = \left[\mathbf{w}^* - (\mathbf{z}^*\mathbf{g})\mathbf{w}^*\right]/\gamma = \left[\mathbf{w}^* - \mathbf{w}^*\right]/\gamma = 0$$