Fast QR Iteration for Companion Matrices

David Bindel

Ming Gu

David Garmire

James Demmel

Shivkumar Chandrasekaran

Motivation: Polynomial root finding

A standard algorithm: QR iteration on a companion matrix

- Robust software exists
- It's normwise backward stable
- It's used in Matlab
- But it takes $O(n^3)$ time and $O(n^2)$ storage

Use structure in QR iterates to get $O(n^2)$ time, $O(n)$ space.

Companion matrices

Given

$$p(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0$$

Define a companion matrix C:

$$C = \begin{bmatrix} -a_{n-1} & -a_{n-2} & -a_{n-3} & \ldots & -a_1 & -a_0 \\ 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & 0 \end{bmatrix}$$

For matrix polynomials, replace a_i by A_i and 1 by I.

Companion matrix structure

Write C as orthonormal + low rank:

$$
\begin{aligned}
C &= P + A \\
P &= \begin{bmatrix} 0^T & 1 \\ I & 0 \end{bmatrix} \\
A &= -e_1 \begin{bmatrix} a_{n-1} & \dots & a_1 & a_0 + 1 \end{bmatrix}
\end{aligned}
$$

This structure is preserved under unitary similarity.

Reminder: CS decomposition

Any block 2-by-2 unitary Z has a CS decomposition:

$$\begin{bmatrix} U_1 & 0 \\ 0 & U_2 \end{bmatrix} \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} \begin{bmatrix} V_1 & 0 \\ 0 & V_2 \end{bmatrix} = \begin{bmatrix} C & S & 0 \\ -S & C & 0 \\ 0 & 0 & I \end{bmatrix}$$

where U and V are also unitary.

Therefore rank$(Z_{21}) = $ rank(Z_{12}).

Off-diagonal rank

Compute a Schur form

$$\begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix} = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix}^H C \begin{bmatrix} Q_1 & Q_2 \end{bmatrix}$$

Then rank$(T_{12}) \leq 2$.

Proof: Write $T = \hat{P} + \hat{A}$, where $\hat{P} := Q^H P Q$ and $\hat{A} := Q^H A Q$.

$$\begin{aligned} \text{rank}(\hat{P}_{21}) &= \text{rank}(-\hat{A}_{21}) = 1 \\ \text{rank}(\hat{T}_{12}) &= \text{rank}(\hat{P}_{12} + \hat{A}_{12}) \\ &\leq \text{rank}(\hat{P}_{12}) + \text{rank}(\hat{A}_{12}) \\ &= 2 \end{aligned}$$

Off-diagonal rank

Similar argument shows:

- Hessenberg forms similar to C have rank(off-diagonal) ≤ 3.
- Real Schur forms similar to C have rank(off-diagonal) ≤ 3.
- Matrix polynomials with coefficient size d have rank$(T_{12}) \leq 2d^2$.

SSS structure

Write Hessenberg matrices with low off-diagonal rank as narrow band matrix + SSS matrix.

$$H = B + \begin{bmatrix} 0 & U_1V_2^T & U_1W_2V_3^T & U_1W_2W_3V_4^T \\ 0 & 0 & U_2V_3^T & U_2W_3V_4^T \\ 0 & 0 & 0 & U_3V_4^T \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Write $H_{ij} = U_iW_{i+1}\dots W_{j-1}V_j$ for $i < j$ where $U_i \in \mathbb{R}^{k\times 3}$, $V_i \in \mathbb{R}^{k\times 3}$, and $W_i \in \mathbb{R}^{3\times 3}$.

Total storage cost: $O(n)$.

SSS structure transformed

$$\hat{H} = \begin{bmatrix} B_{11} & 0 & 0 & 0 \\ \hat{B}_{21} & \hat{B}_{22} & 0 & 0 \\ 0 & \hat{B}_{32} & B_{33} & 0 \\ 0 & 0 & B_{43} & B_{44} \end{bmatrix} + \begin{bmatrix} 0 & U_1 V_2^T & U_1 W_2 \hat{V}_3^T & U_1^T W_2 W_3 V_4^T \\ 0 & 0 & \hat{U}_2 \hat{V}_3^T & \hat{U}_2^T W_3 V_4^T \\ 0 & 0 & 0 & U_3 V_4^T \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Transform second row and column by reflection.
Only affects the band, U_2, and V_2.

What about transformations that cross block edges?
Merge blocks so it never happens – then split them.

Result: A bulge-chasing pass takes $O(n)$ time vs $O(n^2)$

Splitting blocks

$$\begin{bmatrix} U_1 W_2 \ldots W_{j-1} & 0 \\ \vdots & \vdots \\ U_{j-2} W_{j-1} & 0 \\ U_{j-1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} V_j^T & W_j \\ B_{jj} & U_j \end{bmatrix} \begin{bmatrix} 0 & V_n W_{n-1}^T \ldots W_{j+1}^T \\ \vdots & \vdots \\ 0 & V_{j+2} W_{j+1}^T \\ 0 & V_{j+1} \\ I & 0 \end{bmatrix}^T$$

Splitting blocks

$$\begin{bmatrix} V_j^T & W_j \\ B_{jj} & U_j \end{bmatrix}$$

$$\downarrow$$

$$\begin{bmatrix} V_j^{\alpha T} & W_j^{\alpha} V_j^{\beta T} & W_j^{\alpha} W_j^{\beta} \\ B_{jj}^{\alpha\alpha} & U_j^{\alpha} V_j^{\beta T} & U_j^{\alpha} W_j^{\beta} \\ B_{jj}^{\beta\alpha} & B_{jj}^{\beta\beta} & U_j^{\beta} \end{bmatrix}$$

$$\downarrow$$

Do a pivoted QR decomposition to split U, V, W.

Choose $\begin{bmatrix} W_j^{\alpha} \\ U_j^{\alpha} \end{bmatrix}$ with orthonormal columns.

Balancing

Cannot diagonally scale C and maintain low-rank structure. But we can use geometric scaling.

If C is a companion matrix, so is $|C|$. The Perron vector x of $|C|$ therefore has the form

$$x_i = c\lambda^i$$

Let

$$D = \text{diag}(x_i^{-1}) = \text{diag}(\lambda_i^{-1}).$$

Then DCD^{-1} is optimally balanced in the ∞-norm.

Implementation

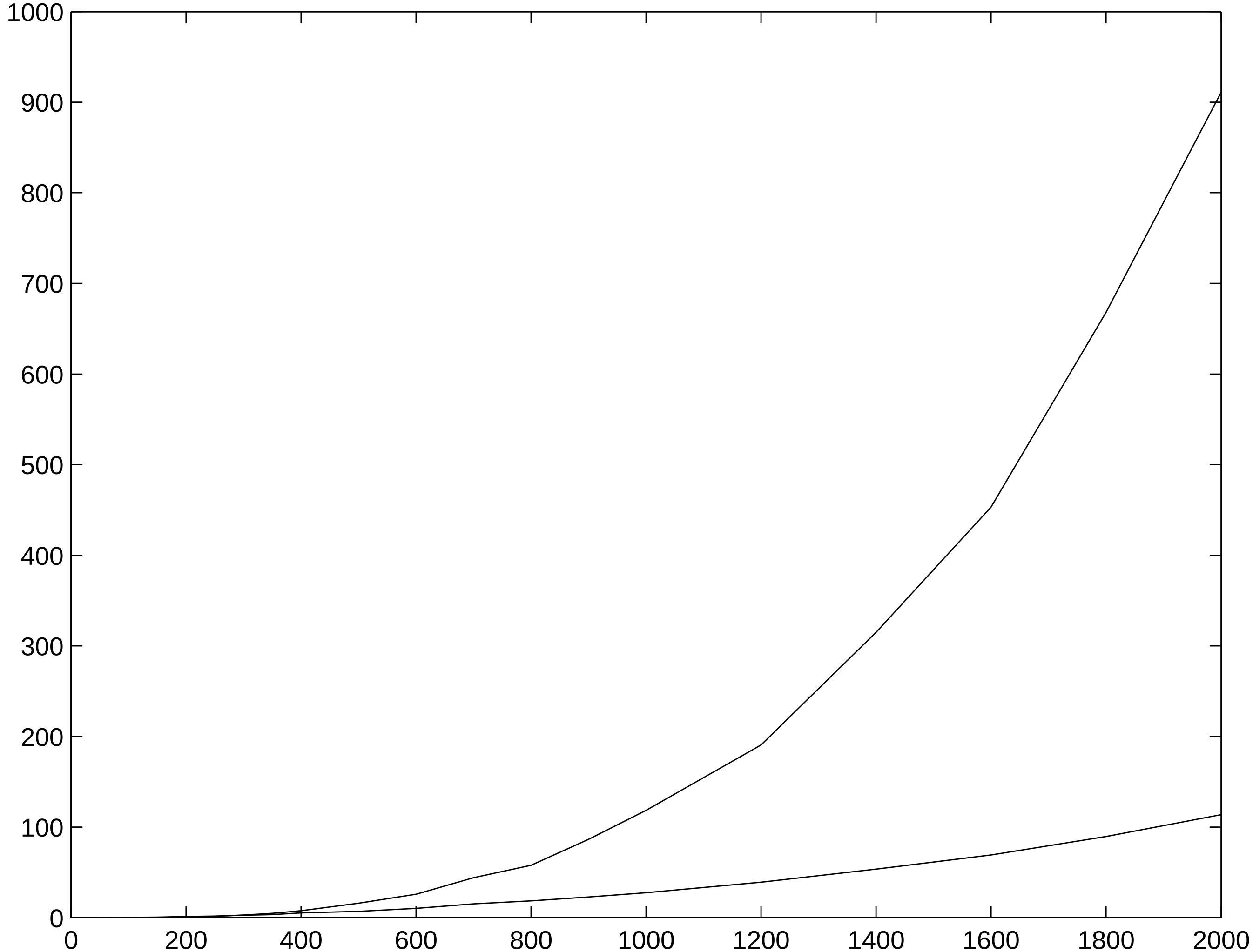
Modified DLAHQR (basic double-shift QR) to use the new data structure.

- Convert to right proper form.
- Choose shifts (logic from LAPACK).
- Bulge chase top to bottom / convert to left proper form.
- Check for convergence / deflate (logic from LAPACK).

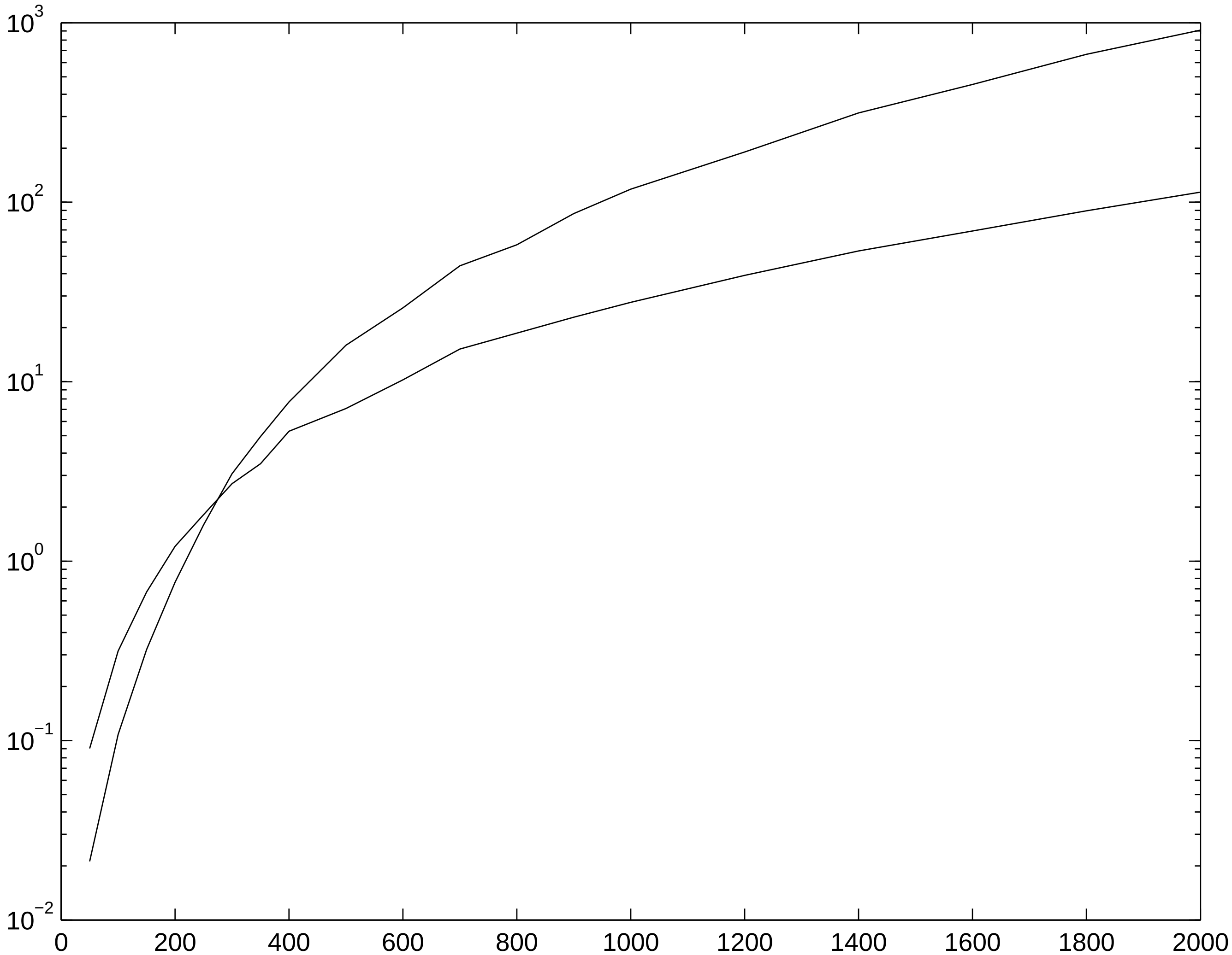
Performance

- Compared against LAPACK DLAHQR (not DHSEQR).
- Compiled using g77 without optimizations.
- LAPACK with standard optimizations, ATLAS BLAS
- Pentium 3 laptop at 700 MHz.
- Polynomials of degree 50-2000.
- For degree 10000, current code takes 41 minutes.

Performance



Performance



Generalizations and Future Work

- Profiling and basic tuning.
- Modified multi-shift DHSEQR code.
- Balancing.
- Polynomial eigenvalue problems.
- Sequences of symmetric + low rank problems.

Conclusions

- Schur({ symmetric, skew, unitary } + low rank) has low off-diagonal rank.
- Can use low-rank structure for fast bulge-chasing.
- Only orthonormal transformations, so backward stable.
- Can re-use most existing QR lore and code.
- Can still do ∞-norm balancing for companion matrices.