Finding Eigenvalues using Automatic Differentiation

Alex Hirzel, Peter Solfest, and Ryan Bruner

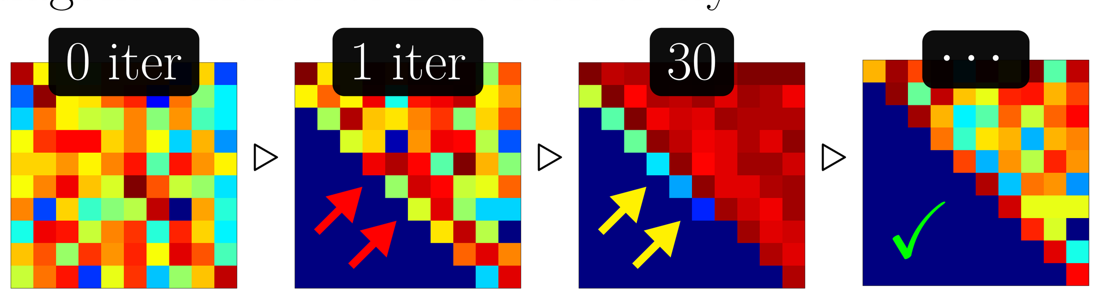
Michigan Technological University

Abstract

Eigenvalues are ubiquitous in all branches of science and engineering. The dominant eigenvalue algorithm reduces input matrices to almost diagonal and then creates and chases bulges using parameters called shifts. Good parameter choices can allow shrinking of the original problem allowing for a more aggressive deflation on the subproblem. Such deflations greatly improve performance. We apply automatic differentiation tools to achieve frequent deflations.

Existing Techniques

A general matrix is first reduced to Hessenberg form, then the **QR algorithm** is used to reduce the subdiagonal entries to zero iteratively:

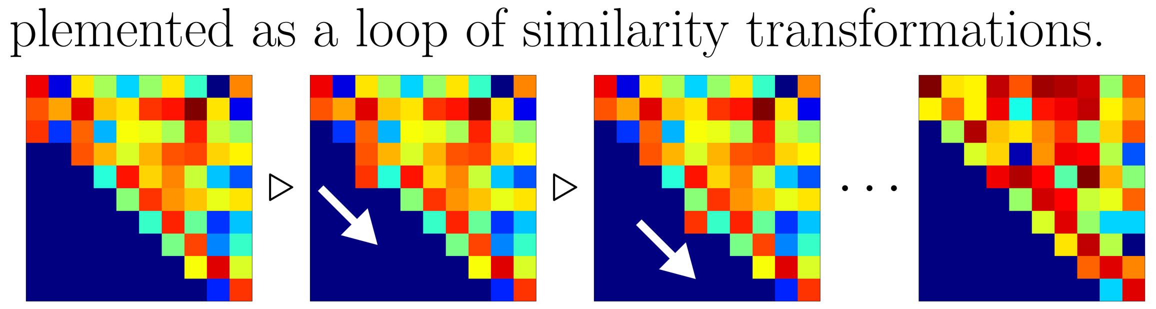


Since the resulting matrix is strictly upper triangular, **the eigenvalues are the diagonal entries**. The QR algorithm consists of repeated iterations of:

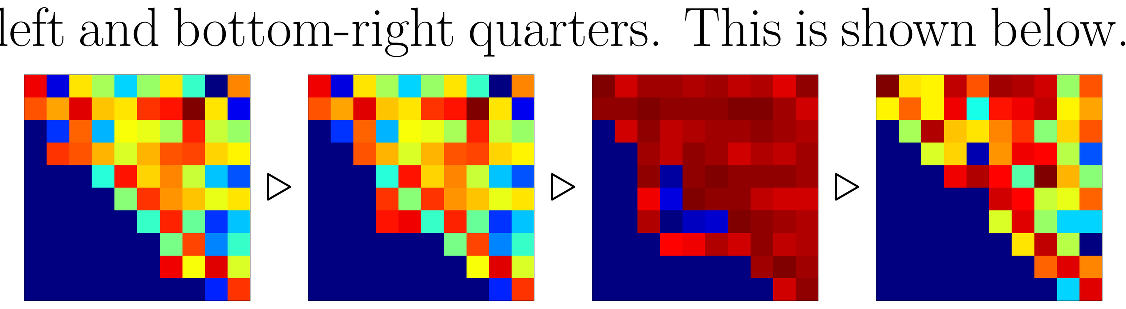
$$A_n \to Q_n R_n \tag{1}$$

$$A_{n+1} \leftarrow R_n Q_n = Q_n^{-1} A_n Q_n \tag{2}$$

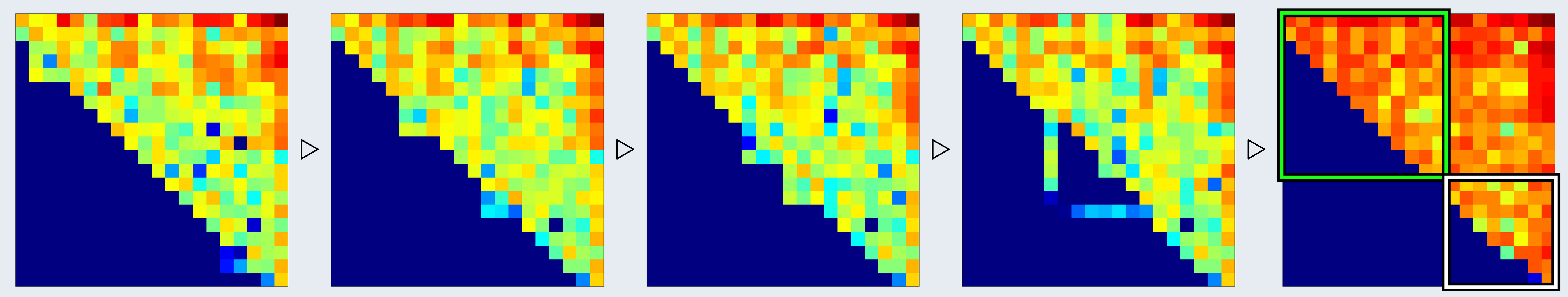
As written, (1) and (2) require multiplication of large matrices. For this reason, the QR algorithm is implemented as a loop of similarity transformations.



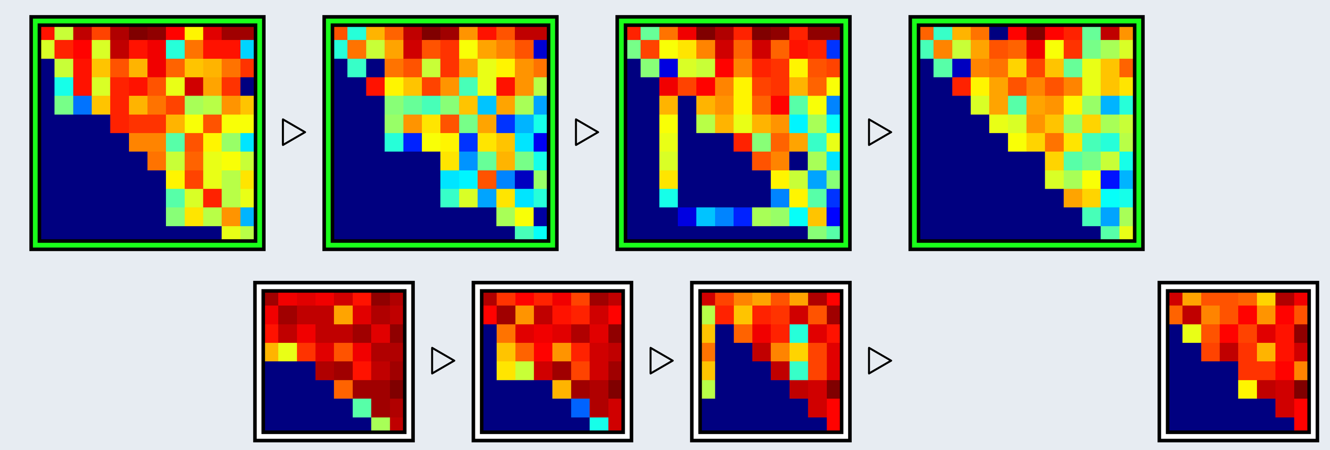
A *bulge* is repeatedly **chased down the diagonal**. To speed convergence [3], the bulge is seeded with one or more *shifts*, which represent **parameter choices**. A second bulge can be started from the bottom, and the two can meet in the middle. A **Schur decomposition is used to form *spikes*** around this bulge. Under the right conditions, the problem can be subdivided recursively into the top-left and bottom-right quarters. This is shown below.



Proposed Technique: Inducing Recursive Middle Deflation



Subdivision is a **HUGE** performance win.

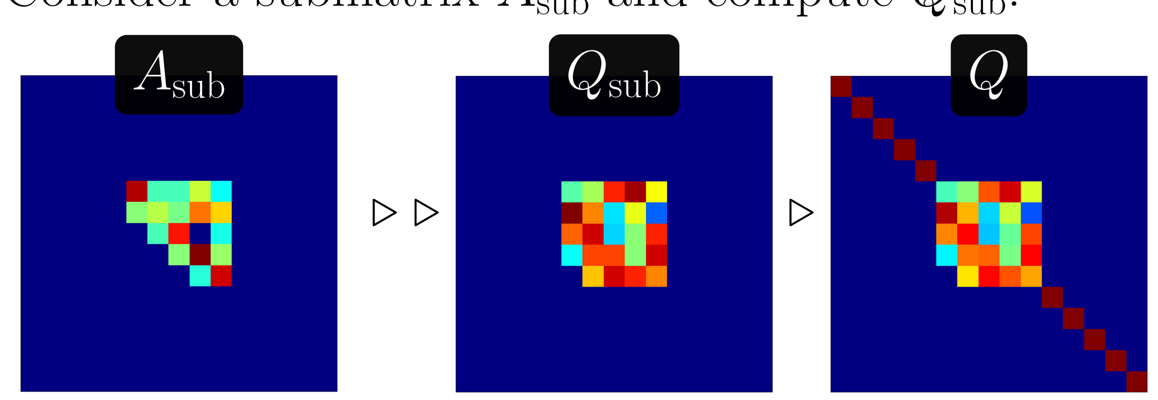


Schur Decomposition

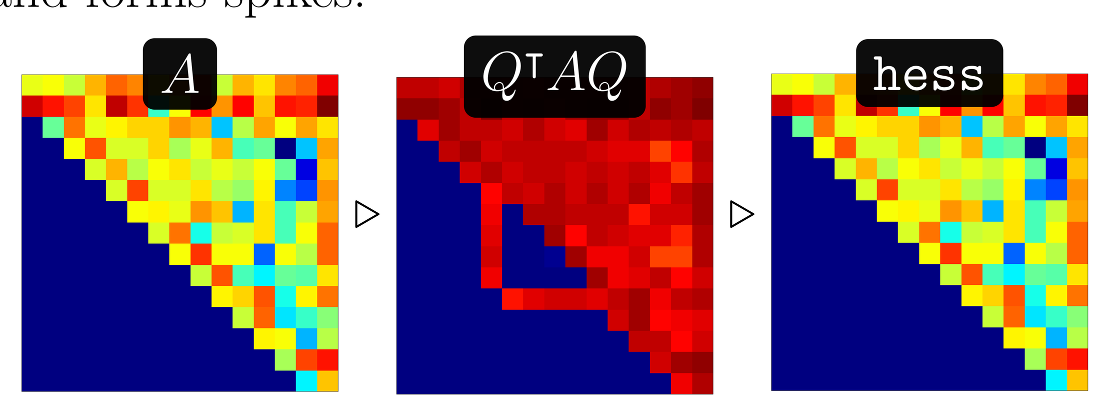
The Schur decomposition allows us to create spikes. For an input matrix A, it forms symmetric Q and diagonal T such that **Q diagonalizes A into T**:

$$A = Q^{\intercal} T Q \tag{3}$$

Consider a submatrix A_{sub} and compute Q_{sub}:



We can then create Q by embedding Q_{sub} into the identity. Applying this to A creates zeros on the subdiagonal in the neighborhood of the decomposition, and forms spikes.



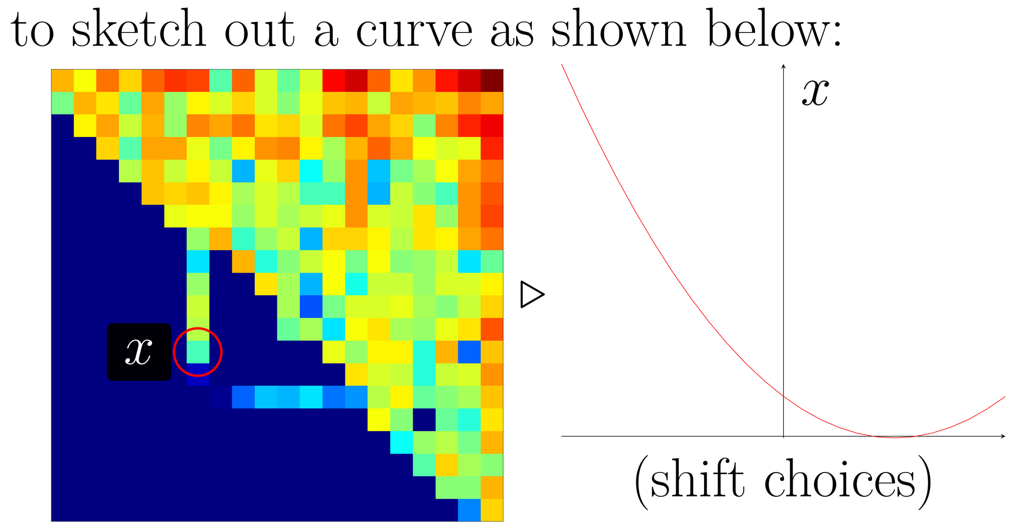
This matrix is then reduced to Hessenberg form again. If half of the values on the spike are near zero, then the Hessenberg form has a zero on the subdiagonal which allows for matrix deflation (see above box).

Achieving Deflation

In practice, middle deflations are improbable because they rely on the spikes having many zeros. The values on the spike depend on the shifts. Since we can control these shifts, the problem of creating a middle deflation can be phrased as

$$\min_{\text{shifts}} \sum |\text{targeted spike values}| \tag{4}$$

Minimizing this implies that the spikes have many zeros, allowing for middle deflation. Selecting a single spike element and tracing its value would cause it to sketch out a curve as shown below:



Many effective methods for minimizing functions rely upon derivative information (e.g. Newton's Method). **Automatic Differentiation** (AD) methods allow for differentiation of matrix operations [1] and the Schur decomposition [2]. Thus by using AD techniques we can utilize existing minimization algorithms to create deflatable matrices.

Conclusions

We have presented an algorithm which aggresively targets middle deflation as opposed to existing algorithms. Furthermore, its recursive nature exposes parralazibility which could lead

With the proper implementation, this will provide a large performance improvement over traditional algorithms in use today. Additionally, this is an enabling technique for extremely large eigenvalue

Future Work

- Incorporate AD methods into minimization algorithms
- Characterize the performance of the algorithm
- Optimize performance
  - Time
  - Memory

GNU Octave.

References

[1] M. Giles, *An extended collection of matrix derivative results for forward and reverse mode algorithmic differentiation*, tech. rep.

[2] A. Struthers, *Automatic differentiation in linear algebra and statistics.*

[3] D. S. Watkins, *The matrix eigenvalue problem: GR and Krylov subspace methods*, vol. 101, SIAM, 2007.

Acknowledgements

The authors thank **Dr. Allan Struthers** (in the Mathematics department) for his guidance and creation of AD tools for the Schur decomposition.

Contact Information

- Email: `ahirzel@mtu.edu`
- Phone: +1 (906) 231 0866

MichiganTech
Michigan Technological University