# Finding Eigenvalues using Automatic Differentiation

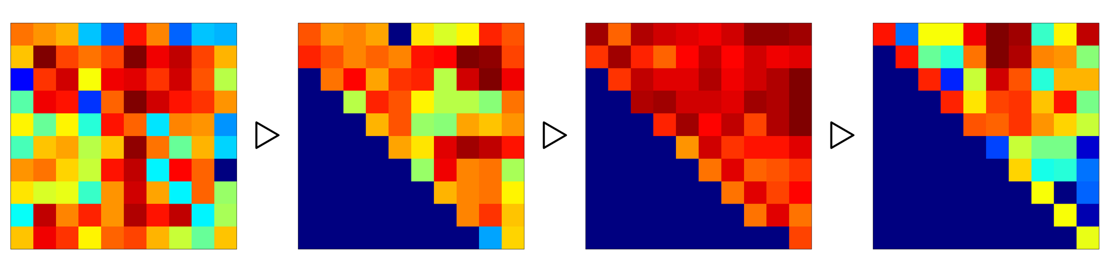
Alex Hirzel, Peter Solfest, and Ryan Bruner

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

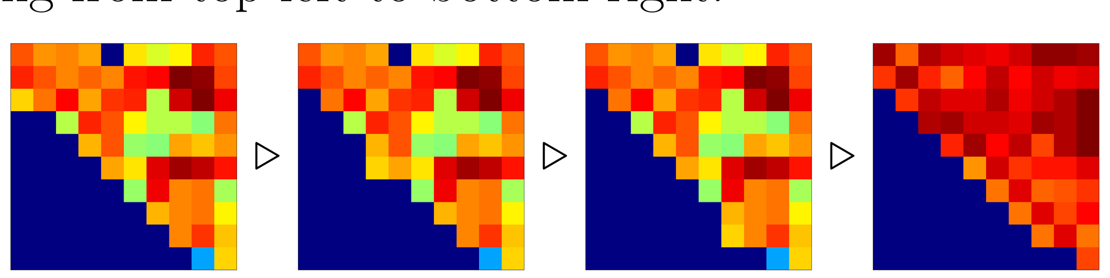


A general matrix is first reduced to Hessenberg form, then the **QR algorithm** is used to reduce the subdiagonal entries to zero (shown above). This exposes eigenvalues on the diagonal. The $QR$ algorithm consists of repeated iterations of:

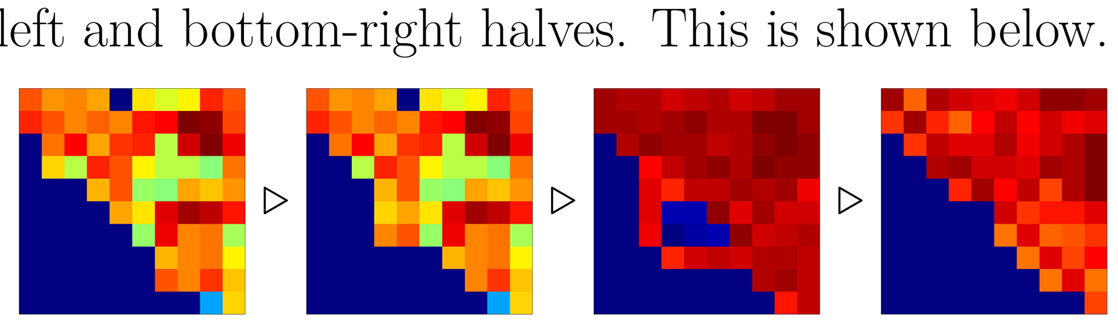
$$A_n \to Q_n \cdot R_n \tag{1}$$

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

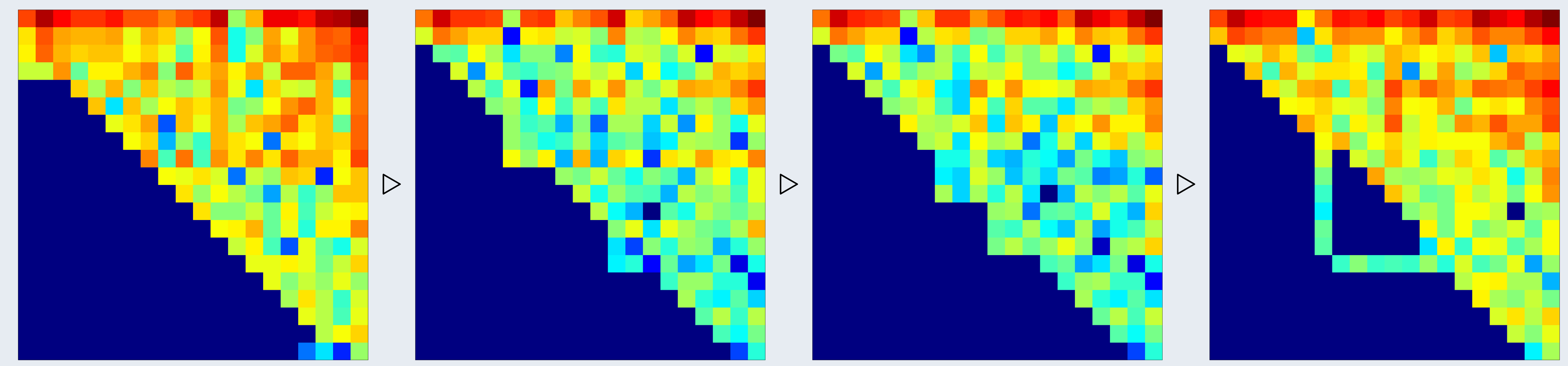
As written, (1) and (2) require matrix multiplies which can be very large. For this reason, the $QR$ algorithm is implemented *implicitly* as a loop of similarity transformations. This results in a *bulge* moving from top-left to bottom-right:



This bulge is seeded with a starting value, a *shift*. A second bulge can be started from the bottom right. When bulges are chased from each side, they meet in the middle. A Schur decomposition is used to form *spikes* around this bulge. When two or more of these *spike values* are sufficiently close to zero, the problem can be subdivided recursively into the top-left and bottom-right halves. This is shown below.



## Proposed Technique



Other text TBD

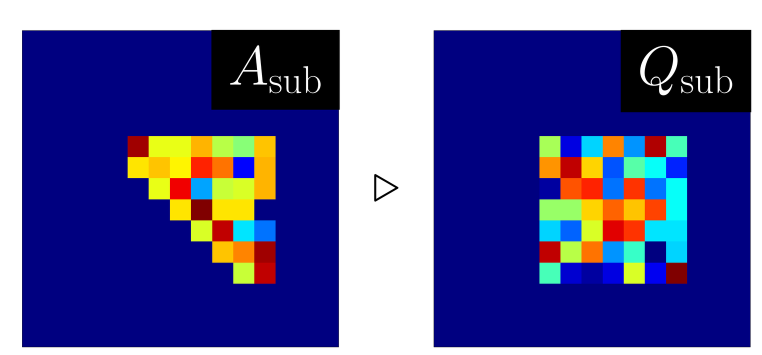
Note that this whole poster is now in English (no Latin :) ).

## Schur Decomposition?

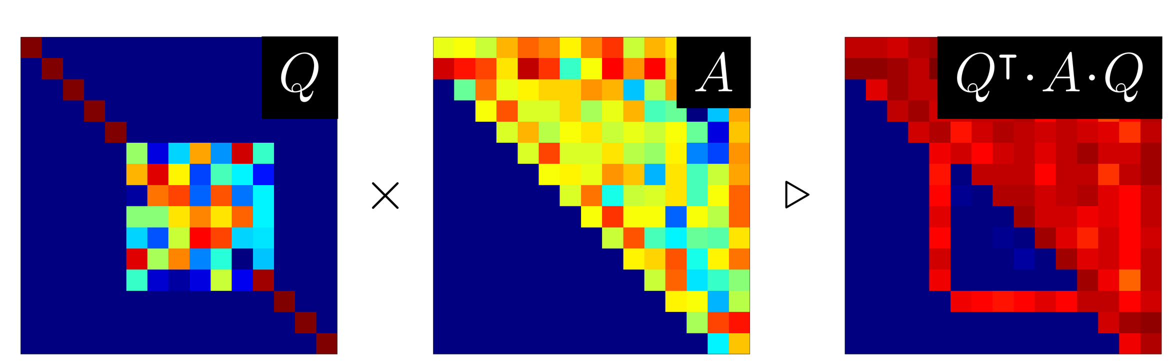
The Schur decomposition is what allows us to create spikes. For an input matrix $A$, it forms $Q$ and $T$ such that $Q$ diagonalizes $T$ into $A$. In symbols,

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

Since $Q$ diagonalizes $A$, we can consider a matrix $A_{\text{sub}}$ and compute $Q_{\text{sub}}$:

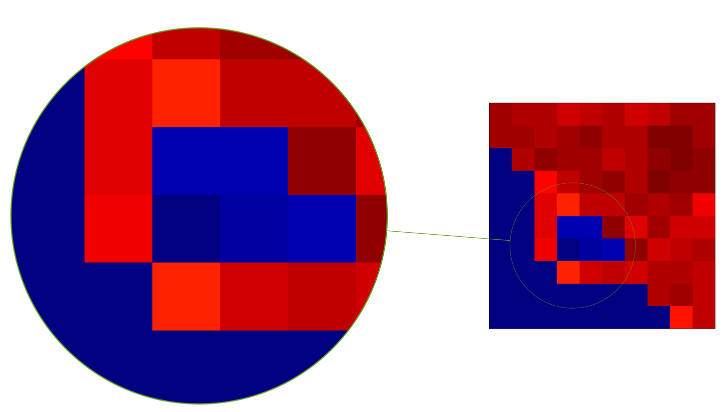



We can then embed $Q_{\text{sub}}$ into $Q$ and apply it to $A$. In the neighborhood of the Schur decomposition, the subdiagonal of $A$ will be zero and the spikes will form.




## Automatic Differentiation?

In practice, middle deflations are improbable because they rely on the values of the spikes. The spike values depend on the shifts fed to the $QR$ algorithm. are generated using a Schur decomposition. They occur when the spike tips become zero due to change or luck. Inducing shifts



We apply **automatic differentiation** (AD) tools to these spikes. AD tools allow *algorithms* to be differentiated in the same sense as algebraic functions. Classical calculus allows us to linearize a function $f(x)$ at $x = a$ as

$$f(x) \approx f(a) + f'(x) \cdot (x - a)$$

which solving for the root ($f(x) = 0$) gives us Newton's method

$$x_{n+1} = x_n - \frac{f(x)}{f'(x)} \tag{4}$$

to find the zeros of $f(x)$. **The technique for finding zeros of the spike values is analogus.** Applying AD tools to the Schur decomposition creates the gradient with respect to the input values (shifts). When the proper shifts are found to induce a zero on the spike, a deflation has been found.

## Conclusion

We successfully apply Automatic Differentiation techniques to induce deflations in the center of matrices. 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 problems. Our preliminary results indicate that a performance gain of XXX can be expected from proper implementations of this technique. This was determined by naively implementing each algorithm in MATLAB and comparing FLOP counts using each algorithm for the same inputs.

## Additional Information

The authors are pursuing publication. For more details on this technique, please contact the authors directly.

## References

[1] G. Anderson, *A procedure for differentiating perfect-foresight-model reduced-from coefficients*, Journal of Economic Dynamics and Control, 11 (1987), pp. 465 – 481.

[2] K. S. Braman, *Toward a Recursive QR Algorithm*, PhD thesis, Lawrence, KS, USA, 2003. AAI3103375.

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

## Acknowledgements

The authors thank **Dr. Allan Struthers** (Mathematics department) for his guidance.

## Contact Information

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

Michigan Tech
Michigan Technological University