Zyzzyva: The Symmetric Tridiagonal Eigenproblem

Inderjit S. Dhillon
University of Texas at Austin

Bay Area Scientific Computing Days
March 30, 2008

Problem Definition

Given

$$T = \begin{bmatrix} a_1 & b_1 & & & & \\ b_1 & a_2 & b_2 & & & \\ & b_2 & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & a_{n-1} & b_{n-1} \\ & & & & b_{n-1} & a_n \end{bmatrix},$$

Solve

$$Tv = \lambda v$$

for λ and $v \neq 0$.

This is a **central** problem in any symmetric eigenproblem.

Requirements in finite precision :

- $\| T\hat{v}_i - \hat{\lambda}_i \hat{v}_i \| \; = \; O(\varepsilon \| T \|), \quad i = 1, 2, \ldots, n$
- $(\hat{v}_i, \hat{v}_j) \; = \; O(\varepsilon), \quad i \neq j.$

History

- Jacobi [1846]
- Bargmann, Montgomery and von Neumann [1946]
- Lanczos [1950], Givens [1954], Householder [1958]
- Wielandt, Wilkinson [1940/1950s], perhaps earlier
- Francis/Kublanovskaja [1961]
- Golub and Kahan [1965]
- Wilkinson [1965]
- Cuppen [1981], Gu and Eisenstat [1995]
- All algorithms cost $O(n^3)$ in the worst case

Ideal Tridiagonal Eigensolver

DESIRED GOALS:

- **Minimum output complexity** — $O(n)$ per eigenvector
- **Provably numerically accurate**
- Each eigenpair is **independently computable**

Difficulties

- Problem: Given $\hat{\lambda}$, solve

$$(T - \hat{\lambda} I)x \approx 0.$$

- Inverse iteration: Given $\hat{\lambda}$, compute:

$$(T - \hat{\lambda} I)x_{i+1} \; = \; x_i, \quad i = 0, 1, 2, \ldots.$$

  - Costs $O(n)$ per iteration
  - Typically, 1-3 iterations are enough

- **BUT**, inverse iteration only guarantees

$$\| T\hat{v} - \hat{\lambda}\hat{v} \| \; = \; O(\varepsilon \| T \|)$$

Fundamental Limitations

Gap Theorem :

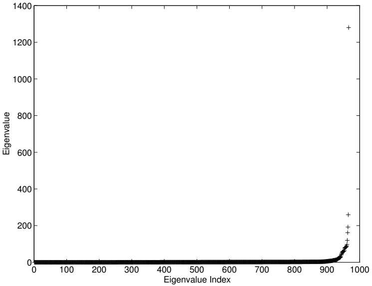
$$\sin \angle(v, \hat{v}) \ \leq \ \frac{\| T\hat{v} - \hat{\lambda}\hat{v} \|}{\mathrm{Gap}(\hat{\lambda})}.$$

$\mathrm{Gap}(\hat{\lambda})$ can be small :

$$\begin{bmatrix} 1 & \varepsilon_1 & & \\ \varepsilon_1 & 1 & \varepsilon_2 & \\ & \varepsilon_2 & 1 & \varepsilon_3 \\ & & \varepsilon_3 & 1 \end{bmatrix}$$

When eigenvalues are close, independently computed eigenvectors **WILL NOT** be mutually orthogonal.

Eigenvalues of Biphenyl Matrix



- LAPACK — one big "cluster" $\lambda_1, \lambda_2, \ldots, \lambda_{939}$.
- Tridiagonal solution takes 80% of Total Time.

Difficulties

- All eigenvalues of T are easily computed in $O(n^2)$ time.
  - QR, Root-free QR (Pal, Walker & Kahan), Bisection + Laguerre
- But eigenvector computation costs $O(n^3)$
  - QR was slow
  - D&C resulted in inaccurate eigenvectors
  - Inverse Iteration expensive when eigenvalues were clustered

Developments at Berkeley

- Ming Gu solves D&C eigenvector problem and comes to Berkeley
- Jeff Rutter writes D&C software to add to LAPACK
- Jim Demmel makes observation:
  - D&C fast when eigenvalues cluster
  - Inverse Iteration fast when eigenvalues are isolated
- Modest GOAL: Combine them to get faster software

Developments

- Twisted Factorizations — Fernando [1995], Godunov et al. [1985], Parlett & Dhillon[1997]
- Compute the appropriate Twisted Factorization :

$$T - \hat{\lambda} I = N_r D_r N_r^T,$$

$$\text{with} \quad N_r = \begin{bmatrix} x & & & & & & & \\ x & x & & & & & & \\ & \cdot & \cdot & & & & & \\ & & x & \textcircled{\gamma_r} & x & & & \\ & & & & \cdot & \cdot & & \\ & & & & & x & x & \\ & & & & & & x & x \\ & & & & & & & x \end{bmatrix}$$

- Solve for z, $N_r D_r N_r^T z = \gamma_r e_r$ ($\Rightarrow N_r^T z = e_r$)
- Solved an open problem posed by Wilkinson (1965).

Initial Ideas

- Handling Clusters
  - Different twists for different eigenvalues in the cluster
- Ming Gu produced 5×5 counter-example:
  - Leftmost $\lambda \approx 1$
  - 3 clustered (not tightly) around 2
  - Righmost $\lambda \approx 3$
- Struggled with it over many sessions on the computer

The Major Breakthrough

- LAPACK test matrices with geometrically distributed eigenvalues
  - One type was breezing through
  - Another type was failing miserably
- High Relative Accuracy
  - Kahan (1966)
  - Demmel & Kahan (1985)
- But tridiagonals do not always determine eigenvalues to high relative accuracy
- REPLACE Tridiagonal by Bidiagonal
- Many dead-ends/false leads
- Representation Tree.......

Software

- Osni Marques
- Christof Vömel's torture tests
- Software is now in LAPACK
- Zyzzyva??

Early References

- Fernando [1995], Godunov et al. [1985], Parlett & Dhillon[1997]
- Strang (2001)
  - Formula for $Diag(T^{-1})$ known in literature on Kalman Filtering

- Birthday Gift for Beresford?