Highlights of Paul Willems’ thesis

Beresford Parlett

University of California, Berkeley

September, 2010

LAPACK seminar, UC Berkeley.
Notation for error analysis

Need to bound

$$\sqrt{(1 + \alpha_1)(1 + \alpha_2)(1 + \alpha_3)(1 + \alpha_4)}, |\alpha_i| \leq \varepsilon.$$ 

Get

$$1 + 2\varepsilon + O(\varepsilon^2)$$

Question

*How to avoid $O(\cdot)$ but keep it simple?*
Notation for error analysis

Need to bound

$$\sqrt{(1 + \alpha_1)(1 + \alpha_2)(1 + \alpha_3)(1 + \alpha_4)}, |\alpha_i| \leq \varepsilon.$$ 

Get

$$1 + 2\varepsilon + O(\varepsilon^2)$$

Question

*How to avoid $O(\cdot)$ but keep it simple?*

Higham. $\gamma_n = n\varepsilon/(1 - n\varepsilon)$.

Good for sums and products, but

$$\sqrt{1 + \gamma_4} \neq 1 + \gamma_2$$

Willems. $\varepsilon^{[k]}(n) := n\varepsilon/(1 - kn\varepsilon)$.

Top expression $\leq 1 + \varepsilon^{[2]}(2) = 1 + 2\varepsilon/(1 - 4\varepsilon)$.
In general,

\[(1 + \varepsilon^{[k]}(n))^s = 1 + \varepsilon^{[k/s]}(sn), \quad 0 < s < 1,
\]

\[(1 + \varepsilon^{[k]}(n))^{-1} = 1 + \varepsilon^{[k+1]}(n).\]
Representations of $T$ (without twists)

1. $(T)$ matrix entries \( \{c_i, e_i\} \)

2. $(N) \{d_i, l_i\}$, $T = LDL^*$.
   - **Pro**: defines tiny eigenvalues to high relative accuracy.
   - **Con**: does not always exist. Element growth.

3. $(e) \{d_i, e_i\}$, $e_i = l_i d_i$.

4. $(Z) \{d_i, lld_i\}$, $lld_i = d_i l_i^2 = \text{Schur complements}$.
   \[ c_{i+1} = d_{i+1} + lld_i. \]

Can convert between representations with **no** adds or subtracts. Square Roots for $(Z)$. 
Comparison of Representations

The computation of eigenvectors is organized in the Representation Tree.

Each internal node requires a new representation which defines a specific subset of shifted eigenvalues to high relative accuracy.

Accuracy: \((Z)\) is best. max is 3 ulps for \((Z)\) versus 4 for \((e)\).
Speed: \((e)\), when properly optimized. \((N)\) almost as good.
Conclusion: Always use \((Z)\) but switch to \((N)\) or \((e)\) when a node contains a singleton.
Shift of Origin

\[ L_+ D_+ (L_+)^T = LDL^T - \tau I \]

\[(l^+_i)^2 d^+_i + d^+_{i+1} = l_i^2 d_i + d_{i+1} - \tau \]

\[ d^+_{i+1} = e_i^2 / d_i + d_{i+1} - e_i^2 / d^+_i - \tau \]

Define \( s_i \) by \( s_i - \tau = d^+_i - d_i \) to find

\[ s_{i+1} = e_i^2 (s_i - \tau) / d_i d^+_i, \]

an important recurrence.
a sample error analysis

Algorithm dstqds, unblocked, e-rep

1. \( d_i^+ = d_i + (s_i - \tau) \)
2. \( e_i^+ = e_i \)
3. \( s_{i+1} = e_i^2(s_i - \tau)/d_id_i^+ \)

Keep the computed \( s_i \) sacred.
a sample error analysis

Algorithm dstqds, unblocked, e-rep

1. \( d_i^+ = d_i + (s_i - \tau) \)
2. \( e_i^+ = e_i \)
3. \( s_{i+1} = e_i^2(s_i - \tau)/d_id_i^+ \)

Keep the computed \( s_i \) sacred.

1. \( d_i^+/(1 + \alpha_+) = d_i + (s_i - \tau)(1 + \sigma), \)

\[
\left( \frac{d_i^+}{(1 + \alpha_+)(1 + \sigma)} \right) = \left( \frac{d_i}{1 + \sigma} \right) + (s_i - \tau).
\]

3. \( s_{i+1} = e_i^2(s_i - \tau)(1 + \sigma)(1 + \alpha_s)/d_id_i^+ \)

\( 1 + \alpha_s = (1 + \varepsilon)^4, \quad 3 \text{ mults, } 1 \text{ divide.} \)

\( \tilde{e}_i = e_i \sqrt{\frac{(1 + \alpha_s)(1 + \sigma)}{(1 + \sigma)^2(1 + \alpha_+)}}, \quad \text{cancel } 1 + \sigma, \)

\( = e_i[1 + \varepsilon^{[4]}(3)], \quad \text{Willems notation.} \)
\[ N_k G_k N_k^* \xrightarrow{\text{computed}} N_t^+ G_t^+ (N_t^+)^* \]

Table XX
(Left side)

\[ \tilde{N}_k \tilde{G}_k \tilde{N}_k^* \xrightarrow{\text{exact}} \tilde{N}_t^+ \tilde{G}_t^+ (\tilde{N}_t^+)^* \]

\[ X = I + \delta_k e_k e_k^* + \delta_t e_t e_t^*, \quad \delta_k, \delta_t \equiv \epsilon(1) \]

Figure: Mixed relative error analysis for dtwqds.
Perturbing the shift

Outer Perturbations
$\text{NGN}^* \rightarrow \text{DNGN}^* D, \ D \approx I$

Inner Perturbations
$\text{NGN}^* \rightarrow \text{NDGDN}^*, \ D \approx I$

Ostrowski’s Theorem. $A^* = A$

$$|\lambda(\text{FAF}^*) - \lambda(A)| \leq |\lambda(A)| \cdot ||\text{FF}^* - I||.$$

Outer perturbations make tiny relative changes.

Application. $X \approx I, \ Y = X^{-1/2}$

$$N_+ G_+(N_+)^* = \text{NGN}^* - \tau X$$

$$Y(N_+ G_+(N_+)^*)Y = Y\text{NGN}^* Y - \tau I$$

$G \rightarrow YGY, \ N \rightarrow YNY^{-1}.$
Preserving Tridiagonal Form

Allow $2 \times 2$ blocks in $D$

$$LDL^T$$

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ k_1 & l_2 & 1 & 0 \\ 0 & 0 & l_3 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} d_1 & 1 & 0 & 0 \\ 1 & c_2 & 0 & 0 \\ 0 & 0 & d_3 & 0 \\ 0 & 0 & 0 & d_4 \end{bmatrix}.$$ 

$(3, 1)$ entry: $k_1d_1 + l_2 \cdot 1 + 1 \cdot 0$

Cannot perturb $k_1$ and $l_2$ independently and preserve tridiagonal form
Blocks in LDL

Example. \( T = GK - \alpha l, \quad \alpha = O(\varepsilon). \)

\[
T = \begin{bmatrix}
-\alpha & 1 & 0 & 0 \\
1 & -\alpha & 1 & 0 \\
0 & 1 & -\alpha & \alpha \\
0 & 0 & \alpha & -\alpha \\
\end{bmatrix}, \quad L = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
k_1 & l_2 & 1 & 0 \\
0 & 0 & l_3 & 1 \\
\end{bmatrix}
\]

\[
D = \begin{pmatrix}
d_1 & 1 \\
1 & c_2 \\
\end{pmatrix} \oplus \begin{pmatrix}
d_3 & 0 \\
0 & d_4 \\
\end{pmatrix}.
\]

\[
k_1 = 1/(1 - \alpha^2), \quad l_2 = \alpha/(1 - \alpha^2) \quad l_3 = -(2 - \alpha^2)/(1 - \alpha^2), \quad \Omega = \{2\}.
\]

Willems Representation: \( D, e, \) and \( \Omega = \{i \text{ where a } 2 \times 2 \text{ ends}\} \).

Secondary Data: \( k_i, l_j, \Delta_i = d_i c_{i+1} - e_i^2, inv_D(i) \).

Auxiliary Quantities: \( s_i = D_+(i, i) - D(i, i) + \tau \) (straightforward)

BUT

\[ s_{i+1} \text{ is much more complicated (9 cases)}. \]

What did I miss?
<table>
<thead>
<tr>
<th>Case</th>
<th>Description</th>
<th>$i \downarrow$</th>
<th>$s_{i+1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>$i, i + 1 \not\in \Omega$ $i, i + 1 \not\in \Omega^+$</td>
<td>$\bullet \bullet$</td>
<td>$e_i^2/d_i - e_i^2/d_i^+$</td>
</tr>
<tr>
<td>S2</td>
<td>$i + 1 \in \Omega$ $i + 1 \in \Omega^+$</td>
<td>$\bullet \circ$</td>
<td>0</td>
</tr>
<tr>
<td>S3</td>
<td>$i \in \Omega$ $i \in \Omega^+$</td>
<td>$\bullet \bullet$</td>
<td>$e_i^2 d_{i-1}/\Delta_{i-1} - e_i^2 d_{i-1}^+/\Delta_{i-1}^+$</td>
</tr>
<tr>
<td>S4</td>
<td>$i + 1 \in \Omega$ $i, i + 1 \not\in \Omega^+$</td>
<td>$\bullet \bullet$</td>
<td>$-e_i^2/d_i^+$</td>
</tr>
<tr>
<td>S5</td>
<td>$i, i + 1 \not\in \Omega$ $i + 1 \in \Omega^+$</td>
<td>$\bullet \circ$</td>
<td>$e_i^2/d_i$</td>
</tr>
<tr>
<td>S6</td>
<td>$i \in \Omega$ $i, i + 1 \not\in \Omega^+$</td>
<td>$\bullet \bullet$</td>
<td>$e_i^2 d_{i-1}/\Delta_{i-1} - e_i^2/d_i^+$</td>
</tr>
<tr>
<td>S7</td>
<td>$i, i + 1 \not\in \Omega$ $i \in \Omega^+$</td>
<td>$\bullet \bullet$</td>
<td>$e_i^2/d_i - e_i^2 d_{i-1}^+/\Delta_{i-1}^+$</td>
</tr>
<tr>
<td>S8</td>
<td>$i \in \Omega$ $i + 1 \in \Omega^+$</td>
<td>$\bullet \circ$</td>
<td>$e_i^2 d_{i-1}/\Delta_{i-1}$</td>
</tr>
<tr>
<td>S9</td>
<td>$i + 1 \in \Omega$ $i \in \Omega^+$</td>
<td>$\bullet \bullet$</td>
<td>$-e_i^2 d_{i-1}^+/\Delta_{i-1}^+$</td>
</tr>
</tbody>
</table>

**Table:** Standard formulae for the next adjustment $s_{i+1}$. 
<table>
<thead>
<tr>
<th></th>
<th>LDL*</th>
<th>L⁺D⁺(L⁺)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>( d_i \sim \tilde{d}_i )</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>( c_i \sim \tilde{c}_i )</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>( d_i^+ \sim \tilde{d}_i^+ )</td>
<td></td>
<td>2</td>
</tr>
</tbody>
</table>

Table: Error bounds to achieve mixed relative stability for Algorithm 4.5, for the concrete parameters \( R = 3, K_\square = 1/8 \), cf. Theorem 4.11 and Figure 4.2. Only first-order bounds are shown, i.e., an entry \( p \) stands for a bound \( pe_\diamond + \mathcal{O}(e_\diamond^2) \).
Recap on MRRR

- Users want orthogonality among eigenvectors.
  Constraint. No Gram-Schmidt (distributive computing)
  George Fann example.
- Extravagant accuracy delivers orthogonality.
- How to achieve $|| Tz - z\lambda || = O(n\epsilon)|\lambda|$?
  Replace $T$ with an RRR (for $\lambda$).
  Not enough. Need $\text{relgap}(\lambda) \geq 10^{-3}$.
- Make relgaps large by shifting origin (to clusters).
  Hence Multiple Representations.
- Organize computation in a Representation Tree.
- Twisted factors permit residual norms proportional to $|\lambda|$.
  Solve $N_k G_k N_k^T z = e_k \gamma_k$. $N_k^T z = e_k$.
  It is $e_k$ that yields $z$ by products only. Can check that $|\gamma_k| = O(n\epsilon)|\lambda|$.
- Differential qd algorithms allow for roundoff between representations.
  Eigenvectors invariant under exact shifts.