Deflations Preserving Relative Accuracy

W. Kahan, Prof. Emeritus
Mathematics Dept., and E.E. & Computer Science Dept. #1776
University of California, Berkeley CA 94720-1776
[wkahan ατ eecs d0t berkeley d0t edu]

Abstract

Deflation turns a matrix eigenproblem into two of smaller dimensions by annihilating a block of off-diagonal elements. When does deflation perturb at worst the last significant digit or two of each of an Hermitian matrix's eigenvalues no matter how widely their magnitudes spread? We seek practicable answers to this question, particularly for tridiagonals, analogous to answers for bidiagonals' singular values found by Ren-Cang Li in 1994. How deflation affects singular vectors and eigenvectors is assessed too, as is the exploitation of spectral gaps when known.

Prepared for IWASEP IX in Napa, Calif., 4 - 7 June 2012, and for the 12 Sept. 2012 Scientific Computation Seminar at U.C. Berkeley.

This work has been influenced by Jim Demmel, Ming Gu, Ren-Cang Li and Beresford Parlett; but any errors and oversights are mine alone.

This is posted on my web page at www.eecs.berkeley.edu/~wkahan/4June12.pdf .
Proofs and details are posted at www.eecs.berkeley.edu/~wkahan/ma221/Deflate.pdf .

Introduction

Hermitian $H := H' := \begin{bmatrix} M & B \\ B' & W \end{bmatrix}$ and $Y := Y' := \begin{bmatrix} M & O \\ O' & W \end{bmatrix}$ have ordered *Spectra* respectively

$$\mathbb{E}(H) = \{\ \theta_1 \ge \theta_2 \ge \dots \ge \theta_n\ \} \quad \text{and} \quad \mathbb{E}(Y) = \{\ \eta_1 \ge \eta_2 \ge \dots \ge \eta_n\ \} = \mathbb{E}(M) \uplus \mathbb{E}(W)$$

wherein $\mathbb{E}(M) = \{\ \mu_1 \ge \mu_2 \ge \dots \ge \mu_m\ \}$ and $\mathbb{E}(W) = \{\ \omega_1 \ge \omega_2 \ge \dots \ge \omega_{n-m}\ \}$.

Here union $\uplus$ is the union of *Multisets* because some eigenvalues η_j may be repeated.

Y comes from H via ***Deflation***. Every Absolute Error $|\theta_j - \eta_j| \le ||B||$.

What about Relative Errors $\log(\theta_j/\eta_j)$?

Triangular $S := \begin{bmatrix} D & E \\ O' & F \end{bmatrix}$ and $Z := \begin{bmatrix} D & O \\ O' & F \end{bmatrix}$ have nonnegative singular value sets respectively

$$\mathbb{S}(S) = \{\ \sigma_1 \ge \sigma_2 \ge \dots \ge \sigma_n\ \} \quad \text{and} \quad \mathbb{S}(Z) = \{\ \zeta_1 \ge \zeta_2 \ge \dots \ge \zeta_n\ \} = \mathbb{S}(D) \uplus \mathbb{S}(F)$$

wherein $\mathbb{S}(D) = \{\ \delta_1 \ge \delta_2 \ge \dots \ge \delta_m\ \}$ and $\mathbb{S}(F) = \{\ \phi_1 \ge \phi_2 \ge \dots \ge \phi_{n-m}\ \}$.

Z comes from S via Deflation. Every Absolute Error $|\sigma_j - \zeta_j| \le ||E||$.

What about Relative Errors $\log(\sigma_j/\zeta_j)$?

We seek practicable realistic bounds for relative errors regardless of `values' spreads.

Tools:

A Tiny Tolerance $0 < \tau << 1$. Disregard τ^2.

We need not distinguish among $\tau \approx 1 - e^{-\tau} \approx -\log(1 - \tau) \approx \tau/(1 \pm \tau) \approx \dots$,

nor among inequalities like $\tau > |\log(\theta/\eta)|$, $\tau > |(\theta - \eta)/\theta|$, $\tau > |(\theta - \eta)/\eta|$, …

DEFINE: A **Permissible Deflation** induces `values' relative errors below threshold τ .

A. Ostrowski's now Classical Inequalities

If $Y = C^{-1}{}' \cdot H \cdot C^{-1}$ then $1/\|C^{-1}\|^2 \le \theta_j/\eta_j \le \|C\|^2$ for every j (except $0/0 := 1$).

If $Z = S \cdot C^{-1}$ or if $Z = C^{-1} \cdot S$, then $1/\|C^{-1}\| \le \sigma_j/\zeta_j \le \|C\|$ for every j (but $0/0 := 1$).

A typical choice $C^{\pm 1} := \begin{bmatrix} I & \pm U \\ O' & I \end{bmatrix}$, wherein U may be rectangular, has

$$\|C^{\pm 1}\| = \left\| \begin{bmatrix} 1 & \|U\| \\ 0 & 1 \end{bmatrix} \right\| = \|U\|/2 + \sqrt{(1 + \|U\|^2/4)} = \exp\left(\operatorname{arcsinh}(\|U\|/2)\right).$$

Why? Go to svd(U) .

Streamlined Derivation of Ren-Cang Li's Bounds [1994]

$$S := \begin{bmatrix} D & E \\ O' & F \end{bmatrix}, \quad \mathbb{S}(S) = \{\, \sigma_1 \ge \sigma_2 \ge \dots \ge \sigma_n \,\}, \quad Z := \begin{bmatrix} D & O \\ O' & F \end{bmatrix}, \quad \mathbb{S}(Z) = \{\, \zeta_1 \ge \zeta_2 \ge \dots \ge \zeta_n \,\}.$$

Choose $C := \begin{bmatrix} I & D^{-1}\cdot E \\ O' & I \end{bmatrix}$ to get $Z = S\cdot C^{-1}$, $\quad \|C^{\pm 1}\| = \exp\big(\operatorname{arcsinh}(\|D^{-1}\cdot E\|/2)\big)$.

$\Rightarrow$ every rel. error $|\log(\sigma_j/\zeta_j)| < \|D^{-1}\cdot E\|/2$. Similarly every $|\log(\sigma_j/\zeta_j)| < \|E\cdot F^{-1}\|/2$.

Conclusion: If ***either*** $\|D^{-1}\cdot E\| < 2\tau$ ***or*** $\|E\cdot F^{-1}\| < 2\tau$ then every $|\log(\sigma_j/\zeta_j)| < \tau$.

It persists with $0/0 := 1$ even if some $\zeta_j = 0$ so long as either $\|D^{-1}\cdot E\|$ or $\|E\cdot F^{-1}\|$ exists.

Virtue: $\|D^{-1}\cdot E\| \le \|D^{-1}\|\cdot\|E\| \le \|Z^{-1}\|\cdot\|E\|$ derived from absolute error-bounds. And sometimes $\|D^{-1}\cdot E\| \ll \|D^{-1}\|\cdot\|E\|$ and costs a lot less to compute, as happens in Parlett's fast ***dqds*** process to compute a bidiagonal's singular values.

Example:

Let n-by-n $S := \mathrm{bidiag}\begin{bmatrix} & s & & s & & \dots & & s & & s & & e & \\ 1 & & 1 & & \dots & & \dots & & 1 & & 1 & & f \end{bmatrix} = \begin{bmatrix} D & \mathbf{e} \\ \mathbf{o}' & f \end{bmatrix}$ in which the pair $\begin{bmatrix} s \\ 1 \end{bmatrix}$ is missing from only the first and last columns, and $s > f >> 1 > e > 0$.

When is $\mathbf{e}$ is so small that replacing it by $\mathbf{o}$ deflates S with no relative error worse than τ in a singular value?

The least singular value σ_n of S is very near the least singular value $1/\|D^{-1}\|$ of D :

$$\sigma_n \approx (s^2 - 1)/\sqrt{(s^{2n} - n\cdot s^2 + n - 1)} \qquad \text{for } s > 3 \text{ and } n > 3 .$$

The largest singular values of S are not far from those of D :

$$\sigma_1 \approx s + 1 \qquad \text{for } s > 3 \text{ and } n > 3 .$$

This puts f amidst $\mathbb{S}(D)$, so no *spectral gap* (*cf*. p. 14) is available compared with which to deem e^2 negligible.

Yet R-C. Li's criterion implies that $\mathbf{e}$ is negligible if $e < 2\tau\cdot f$ although this e can exceed σ_n hugely.

No other relative-accuracy-preserving criterion I know would permit this example to be so deflated.

What does a Permissible Deflation do to *Some* of the Singular Vectors?

$$S := \begin{bmatrix} D & E \\ O' & F \end{bmatrix}, \quad \mathbb{S}(S) = \{\, \sigma_1 \ge \sigma_2 \ge \dots \ge \sigma_n \,\}, \quad Z := \begin{bmatrix} D & O \\ O' & F \end{bmatrix}, \quad \mathbb{S}(Z) = \{\, \zeta_1 \ge \zeta_2 \ge \dots \ge \zeta_n \,\}.$$

Let F have singular value ϕ and normalized singular row-vectors $\mathbf{u}'$ and $\mathbf{v}'$ satisfying

$$\|\mathbf{u}\| = \|\mathbf{v}\| = 1\,, \quad F\cdot\mathbf{u} = \phi\cdot\mathbf{v} \quad \text{and} \quad \mathbf{v}'\cdot F = \phi\cdot\mathbf{u}'\,.$$

After deflating S to Z, we accept ϕ in $\mathbb{S}(Z)$ as a computed approximation to some σ in $\mathbb{S}(S)$, and accept Z's corresponding singular row-vectors $[\mathbf{o}', \mathbf{u}']$ and $[\mathbf{o}', \mathbf{v}']$ as computed approximations to singular vectors of S. These vector's residuals are

$$\mathbf{r} := S\cdot\begin{bmatrix} \mathbf{o} \\ \mathbf{u} \end{bmatrix} - \phi\cdot\begin{bmatrix} \mathbf{o} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} E\cdot\mathbf{u} \\ \mathbf{o} \end{bmatrix} \quad \text{and} \quad [\mathbf{o}', \mathbf{v}']\cdot S - \phi\cdot[\mathbf{o}', \mathbf{u}'] = \mathbf{o}'\,.$$

$$\|\mathbf{r}\| = \|E\cdot\mathbf{u}\| = \phi\cdot\|E\cdot F^{-1}\cdot\mathbf{v}\| \le \phi\cdot\|E\cdot F^{-1}\| < 2\tau\cdot\phi \quad \text{when} \quad \|E\cdot F^{-1}\| < 2\tau\,.$$

(***one*** of R-C. Li's deflation criteria)

This *Relatively* (relative to ϕ) tiny residual $\mathbf{r}$ figures in the angles between the desired singular vectors of S and their approximations $[\mathbf{o}', \mathbf{u}']$ and $[\mathbf{o}', \mathbf{v}']$ from Z :

Roughly, $\quad \text{angles} \le \|\mathbf{r}\|/(\text{absolute gap}) = (\,\|\mathbf{r}\|/\phi\,)/(\text{relative gap}) < 2\tau/(\text{relative gap})$

wherein $\quad \text{absolute gap} := \min\{|\zeta_j - \phi| \text{ over } \zeta_j \ne \phi\}\,, \quad \text{relative gap} := (\text{absolute gap})/\phi\,.$

It all generalizes from simple ϕ to clustered with invariant subspaces. *Reassuring* so far?

What does a Permissible Deflation do to *the Rest* of the Singular Vectors?

$$S := \begin{bmatrix} D & E \\ O' & F \end{bmatrix}, \quad \mathbb{S}(S) = \{ \sigma_1 \ge \sigma_2 \ge \dots \ge \sigma_n \}, \quad Z := \begin{bmatrix} D & O \\ O' & F \end{bmatrix}, \quad \mathbb{S}(Z) = \{ \zeta_1 \ge \zeta_2 \ge \dots \ge \zeta_n \}.$$

Let D have singular value δ and normalized singular row-vectors $\mathbf{x}'$ and $\mathbf{y}'$ satisfying

$$\|\mathbf{x}\| = \|\mathbf{y}\| = 1, \quad D\cdot\mathbf{x} = \delta\cdot\mathbf{y} \quad \text{and} \quad \mathbf{y}'\cdot D = \delta\cdot\mathbf{x}' .$$

After deflating S to Z, we accept δ in $\mathbb{S}(Z)$ as a computed approximation to some σ in $\mathbb{S}(S)$, and accept Z's corresponding singular row-vectors $[\mathbf{x}', \mathbf{o}']$ and $[\mathbf{y}', \mathbf{o}']$ as computed approximations to singular vectors of S. These vector's residuals are

$$S\cdot\begin{bmatrix} \mathbf{x} \\ \mathbf{o} \end{bmatrix} - \delta\cdot\begin{bmatrix} \mathbf{y} \\ \mathbf{o} \end{bmatrix} = \mathbf{o} \quad \text{and} \quad \mathbf{r}' := [\mathbf{y}', \mathbf{o}']\cdot S - \delta\cdot[\mathbf{x}', \mathbf{o}'] = [\mathbf{o}', \mathbf{y}'\cdot E] ;$$

$$\|\mathbf{r}'\|/\delta = \|\mathbf{y}'\cdot E\|/\delta = \|\mathbf{x}'\cdot D^{-1}\cdot E\| .$$ Must this be small?

What if ***only one*** of R-C. Li's deflation criteria, say $\|E\cdot F^{-1}\| < 2\tau$, is satisfied, not the other? Suppose $\|D^{-1}\cdot E\|$ and $\|\mathbf{r}'\|/\delta$ are both huge. It happens with p. 5's example S .

Despite the *Relatively* (relative to δ) ***big*** residual $\mathbf{r}'$, deflation rotates `vectors through *SMALL* angles $< 2\tau/$(relative gap) *ROUGHLY.*

Why?

Singular vectors of S^{-1} are those of S swapped. Residual for S^{-1} is *Relatively* tiny.

What does a Permissible Deflation do to *Some* Eigenvalues?

Recall $H := H' := \begin{bmatrix} M & B \\ B' & W \end{bmatrix}$ and $Y := Y' := \begin{bmatrix} M & O \\ O' & W \end{bmatrix}$ have ordered *Spectra* respectively

$$\mathbb{E}(H) = \{ \theta_1 \ge \theta_2 \ge \ldots \ge \theta_n \} \quad \text{and} \quad \mathbb{E}(Y) = \{ \eta_1 \ge \eta_2 \ge \ldots \ge \eta_n \} = \mathbb{E}(M) \cup \mathbb{E}(W)$$

$$\text{wherein} \quad \mathbb{E}(M) = \{ \mu_1 \ge \mu_2 \ge \ldots \ge \mu_m \} \quad \text{and} \quad \mathbb{E}(W) = \{ \omega_1 \ge \omega_2 \ge \ldots \ge \omega_{n-m} \} .$$

Claim: Some subset of m eigenvalues θ_j in $\mathbb{E}(H)$ are approximated by $\mathbb{E}(M)$ within factors no farther from 1 than are $\exp(\pm 2\cdot\text{arcsinh}(\|M^{-1}\cdot B\|/2))$. Consequently …

The ***Relative*** errors in $\mathbb{E}(M)$ are all smaller than threshold τ whenever $\|M^{-1}\cdot B\| < \tau$.

Useful for tridiagonal M .

Proof: $C := \begin{bmatrix} I & M^{-1}\cdot B \\ O' & I \end{bmatrix}$ makes $C^{-1}{}'\cdot H\cdot C^{-1} = \begin{bmatrix} M & O \\ O' & \overline{W} \end{bmatrix}$ with $\overline{W} := W - B'\cdot M^{-1}\cdot B$.

• • • • •

Claim says nothing about $\mathbb{E}(W)$ and the remaining $n-m$ θ_j's . We could have $W = O$.

Analogous Claim:

The ***Relative*** errors in $\mathbb{E}(W)$ are all smaller than threshold τ whenever $\|B\cdot W^{-1}\| < \tau$.

What if ***both*** $\|M^{-1}\cdot B\| < \tau$ ***and*** $\|B\cdot W^{-1}\| < \tau$? Must $\mathbb{E}(Y)$ approximate ***all*** of $\mathbb{E}(H)$?

What does a Permissible Deflation do to *All* Eigenvalues?

Recall $H := H' := \begin{bmatrix} M & B \\ B' & W \end{bmatrix}$ and $Y := Y' := \begin{bmatrix} M & O \\ O' & W \end{bmatrix}$ have ordered *Spectra* respectively

$$\mathbb{E}(H) = \{ \theta_1 \ge \theta_2 \ge \ldots \ge \theta_n \} \quad \text{and} \quad \mathbb{E}(Y) = \{ \eta_1 \ge \eta_2 \ge \ldots \ge \eta_n \} = \mathbb{E}(M) \uplus \mathbb{E}(W)$$

$$\text{wherein} \quad \mathbb{E}(M) = \{ \mu_1 \ge \mu_2 \ge \ldots \ge \mu_m \} \text{ and } \mathbb{E}(W) = \{ \omega_1 \ge \omega_2 \ge \ldots \ge \omega_{n-m} \} .$$

Claim: If ***both*** $\|M^{-1}\cdot B\| < \tau$ ***and*** $\|B\cdot W^{-1}\| < \tau$, then $\mathbb{E}(Y)$ approximates ***all*** of $\mathbb{E}(H)$; every $|\log(\theta_j/\eta_j)| < \tau + O(\tau^2)$.

cf. R-C. Li's on p.4

Proof:

$C := \begin{bmatrix} I - K & O \\ W^{-1}\cdot B' & I \end{bmatrix}$ will make $C^{-1}{}'\cdot H\cdot C^{-1} = Y$ exactly and $\|C^{\pm 1}\|^2 < 1 + \tau + \tau^2 + O(\tau^4)$ when $(I - K)'\cdot M\cdot(I - K) = M - B\cdot W^{-1}\cdot B'$, which equation has an explicit rapidly-convergent power-series solution K , and

$$\|K\| < \tau^2/2 + O(\tau^4) .$$

v. pp. 5-6 of …/Deflate.pdf

2-by-2 examples $\begin{bmatrix} 1 & \tau \\ \tau & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & \tau \\ \tau & -1 \end{bmatrix}$: • • • Claim is best-possible in the absence of more data about $\mathbb{E}(M)$ and $\mathbb{E}(W)$, but vastly improvable if a known gap separates them.

What does a Permissible Deflation do to Eigenvectors?

Recall $H := H' := \begin{bmatrix} M & B \\ B' & W \end{bmatrix}$ and $Y := Y' := \begin{bmatrix} M & O \\ O' & W \end{bmatrix}$ have ordered *Spectra* respectively

$$\mathbb{E}(H) = \{ \theta_1 \ge \theta_2 \ge \dots \ge \theta_n \} \quad \text{and} \quad \mathbb{E}(Y) = \{ \eta_1 \ge \eta_2 \ge \dots \ge \eta_n \} = \mathbb{E}(M) \uplus \mathbb{E}(W)$$
$$\text{wherein} \quad \mathbb{E}(M) = \{ \mu_1 \ge \mu_2 \ge \dots \ge \mu_m \} \ \text{and} \ \mathbb{E}(W) = \{ \omega_1 \ge \omega_2 \ge \dots \ge \omega_{n-m} \} .$$

Choose any $\eta \in \mathbb{E}(Y)$; either $\eta \in \mathbb{E}(M)$ or $\eta \in \mathbb{E}(W)$ or both. Say $\eta = \mu \in \mathbb{E}(M)$; let $\mathbf{u}$ be the normalized eigenvector of M belonging to μ : $M \cdot \mathbf{u} = \mu \cdot \mathbf{u}$ and $\|\mathbf{u}\| = 1$.

Then Y's row-eigenvector $\mathbf{y}' = [\mathbf{u}', \mathbf{o}]$ approximates H's belonging to $\theta \approx \eta$; and residual $\mathbf{r}' = \mathbf{y}' \cdot H - \eta \cdot \mathbf{y}' = [\mathbf{o}', \mathbf{u}' \cdot B]$ has $\|\mathbf{r}'\| = \|\mathbf{u}' \cdot B\| = \|\mu \cdot \mathbf{u}' \cdot M^{-1} \cdot B\| \le |\mu| \cdot \|M^{-1} \cdot B\|$.

This is why $\|\mathbf{r}'\| < \tau \cdot |\eta|$ when $\eta = \mu \in \mathbb{E}(M)$ and $\|M^{-1} \cdot B\| < \tau$.

Similarly $\|\mathbf{r}'\| < \tau \cdot |\eta|$ when $\eta = \omega \in \mathbb{E}(W)$ and $\|B \cdot W^{-1}\| < \tau$.

Thus, eigenvector residuals are ***Relatively*** (relative to eigenvalue η) tiny like $\tau \cdot |\eta|$.

Permissible Deflation rotates eigenvectors of H to those of Y through angles …

Roughly, angles $\le \|\mathbf{r}'\| / (\text{absolute gap}) = (\|\mathbf{r}'\|/|\eta|)/(\text{relative gap}) < 2\tau/(\text{relative gap})$
wherein absolute gap := $\min\{|\eta_j - \eta|$ over $\eta_j \ne \eta\}$, relative gap := (absolute gap)/$|\eta|$.

It all generalizes from simple η to clustered with invariant subspaces. *Reassuring*?

Spectral Gaps figure in Eigenvalues' Quadratic Relative Error-Bounds

Recall $H := H' := \begin{bmatrix} M & B \\ B' & W \end{bmatrix}$ and $Y := Y' := \begin{bmatrix} M & O \\ O' & W \end{bmatrix}$ have ordered *Spectra* respectively

$$\mathbb{E}(H) = \{\ \theta_1 \ge \theta_2 \ge \dots \ge \theta_n\ \} \quad \text{and} \quad \mathbb{E}(Y) = \{\ \eta_1 \ge \eta_2 \ge \dots \ge \eta_n\ \} = \mathbb{E}(M) \underline{\cup} \mathbb{E}(W)$$

$$\text{wherein} \quad \mathbb{E}(M) = \{\ \mu_1 \ge \mu_2 \ge \dots \ge \mu_m\ \} \ \text{ and } \ \mathbb{E}(W) = \{\ \omega_1 \ge \omega_2 \ge \dots \ge \omega_{n-m}\ \}\ .$$

Error-bounds between $\mathbb{E}(H)$ and $\mathbb{E}(Y)$ have been roughly proportional to B so far. When B is small enough, smaller ***Quadratic*** bounds roughly proportional to $B'\cdot B$ may be available provided known ***Gaps*** big enough separate $\mathbb{E}(M)$ from $\mathbb{E}(W)$.

The *Absolute Spectral Gap* $\bar{\gamma}(\eta)$ separates $\eta \in \mathbb{E}(Y)$ from $\mathbb{E}(M)$ or $\mathbb{E}(W)$ thus:

If $\eta \in \mathbb{E}(M)$ then $\bar{\gamma}(\eta) := \min\{\ |\omega - \eta|$ over all $\omega \in \mathbb{E}(W)\ \}$, else
if $\eta \in \mathbb{E}(W)$ then $\bar{\gamma}(\eta) := \min\{\ |\mu - \eta|$ over all $\mu \in \mathbb{E}(M)\ \}$.

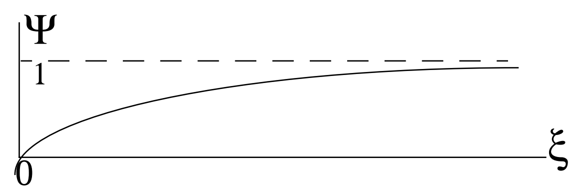
A *Relative Spectral Gap* $\Gamma(\eta)$ separates $\eta \in \mathbb{E}(Y)$ from $\mathbb{E}(M)$ or $\mathbb{E}(W)$ thus:

if $\eta \in \mathbb{E}(M) \cap \mathbb{E}(W)$ then $\Gamma(\eta) := \bar{\gamma}(\eta) = 0$; else $\Gamma(\eta) := \bar{\gamma}(\eta)/|\eta|$.

Define $\quad \Psi(\xi) := \tan\left(\tfrac{1}{2}\arctan(2\xi)\right) = \tanh\left(\tfrac{1}{2}\operatorname{arcsinh}(2\xi)\right) = 2\xi/\left(1 + \sqrt{(1 + 4\xi^2)}\right)$.

Among its properties only these will be needed:

$0 < d\Psi(\xi)/d\xi \le 1$; $\quad \Psi(\xi)/\xi \nearrow 1$ as $\xi \searrow 0$;
$\Psi(\xi) \nearrow 1$ as $\xi \nearrow \infty$.



Recall $H := H' := \begin{bmatrix} M & B \\ B' & W \end{bmatrix}$ and $Y := Y' := \begin{bmatrix} M & O \\ O' & W \end{bmatrix}$ have ordered *Spectra* respectively

$$\mathbb{E}(H) = \{\, \theta_1 \ge \theta_2 \ge \ldots \ge \theta_n \,\} \quad \text{and} \quad \mathbb{E}(Y) = \{\, \eta_1 \ge \eta_2 \ge \ldots \ge \eta_n \,\} = \mathbb{E}(M) \uplus \mathbb{E}(W)$$
$$\text{wherein} \quad \mathbb{E}(M) = \{\, \mu_1 \ge \mu_2 \ge \ldots \ge \mu_m \,\} \ \text{ and } \ \mathbb{E}(W) = \{\, \omega_1 \ge \omega_2 \ge \ldots \ge \omega_{n-m} \,\}\,.$$

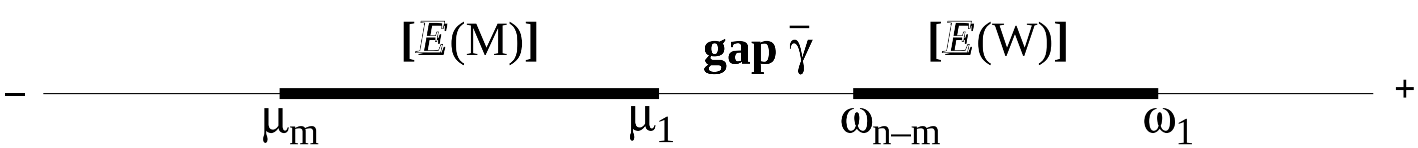
Optimal quadratic absolute error-bounds for eigenvalues from C-K. Li & R-C. Li [2005]:

$$|\theta_j - \eta_j| \le \Psi(\|B\|/\bar{\gamma}(\eta_j))\cdot\|B\| \qquad \mathbb{A}\mathbb{B}$$
$$< \min\{\, \|B\|,\ \|B\|^2/\bar{\gamma}(\eta_j) \,\} \quad \text{when } \|B\| > 0 \text{ and gap } \bar{\gamma}(\eta_j) > 0\,.$$

Quadratic relative error-bounds involving relative gaps $\Gamma(\eta_j)$ come directly from $\mathbb{A}\mathbb{B}$:

$$|\theta_j/\eta_j - 1| \le \Psi(\|B/\eta_j\|/\Gamma(\eta_j))\cdot\|B/\eta_j\|\,. \qquad \mathbb{R}\mathbb{A}\mathbb{B}$$

Those bounds tend to pessimism because they involve $\|B/\eta_j\|$ and are very general, allowing $\mathbb{E}(M)$ and $\mathbb{E}(W)$ to mingle like red and black cards in a shuffled deck. We will impose a **gap** between the smallest interval **[**$\mathbb{E}(W)$**]** containing $\mathbb{E}(W)$, and **[**$\mathbb{E}(M)$**]**:



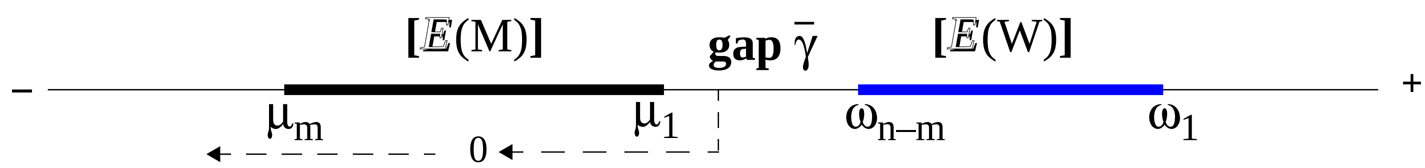
Relative gaps $\Gamma(\eta) := \bar{\gamma}(\eta)/|\eta| \ge \bar{\gamma}/|\eta|$.

Recall $H := H' := \begin{bmatrix} M & B \\ B' & W \end{bmatrix}$ and $Y := Y' := \begin{bmatrix} M & O \\ O' & W \end{bmatrix}$ have ordered *Spectra* respectively

$$\mathbb{E}(H) = \{\, \theta_1 \ge \theta_2 \ge \ldots \ge \theta_n \,\} \quad \text{and} \quad \mathbb{E}(Y) = \{\, \eta_1 \ge \eta_2 \ge \ldots \ge \eta_n \,\} = \mathbb{E}(M) \uplus \mathbb{E}(W)$$

wherein $\mathbb{E}(M) = \{\, \mu_1 \ge \mu_2 \ge \ldots \ge \mu_m \,\}$ and $\mathbb{E}(W) = \{\, \omega_1 \ge \omega_2 \ge \ldots \ge \omega_{n-m} \,\}$.

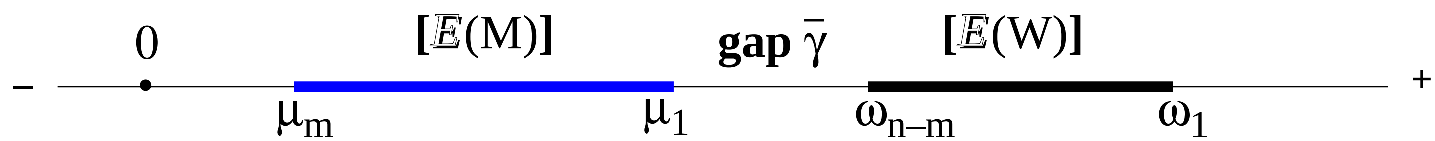
In May 2012 Ren-Cang Li adapted $\mathbb{AB}$'s lengthy proof to get the following usually better bounds for the n–m largest in $\mathbb{E}(H)$ when $\mathbf{[}\mathbb{E}(W)\mathbf{]} > \mathbf{[}\mathbb{E}(M)\mathbf{]}$ and $\mathbf{[}\mathbb{E}(W)\mathbf{]} > 0$:



Relative gaps $\Gamma(\omega_j) = 1 - \mu_1/\omega_j > 0$. Then

$$0 \le \theta_j/\omega_j - 1 \le \Psi(\|B\cdot W^{-1}\|/\Gamma(\omega_j))\cdot\|B\cdot W^{-1}\| \quad \text{for } 1 \le j \le n-m . \qquad \mathbb{RBW}$$

For the m least in $\mathbb{E}(H)$ when $\mathbf{[}\mathbb{E}(W)\mathbf{]} > \mathbf{[}\mathbb{E}(M)\mathbf{]} > 0$ ***and*** $\|M^{-1}\cdot B\| < 1/\sqrt{\left(\left(\frac{\mu_1}{\mu_m}\right)^2 + \frac{\mu_1}{\mu_m}\right)}$:



Relative gaps $\Gamma(\mu_j) = \omega_{n-m}/\mu_j - 1 > 0$. Then

$$0 \le 1 - \theta_{n-m+j}/\mu_j \;<\; \Psi(\|M^{-1}\cdot B\|/\Gamma(\mu_j))\cdot\|M^{-1}\cdot B\| \quad \text{for } 1 \le j \le m . \qquad \mathbb{RMB}$$

Recall $H := H' := \begin{bmatrix} M & B \\ B' & W \end{bmatrix}$ and $Y := Y' := \begin{bmatrix} M & O \\ O' & W \end{bmatrix}$ have ordered *Spectra* respectively

$$\mathbb{E}(H) = \{\ \theta_1 \ge \theta_2 \ge \dots \ge \theta_n\ \} \quad \text{and} \quad \mathbb{E}(Y) = \{\ \eta_1 \ge \eta_2 \ge \dots \ge \eta_n\ \} = \mathbb{E}(M) \,\underline{\cup}\, \mathbb{E}(W)$$

wherein $\mathbb{E}(M) = \{\ \mu_1 \ge \mu_2 \ge \dots \ge \mu_m\ \}$ and $\mathbb{E}(W) = \{\ \omega_1 \ge \omega_2 \ge \dots \ge \omega_{n-m}\ \}$.

0 **$[\mathbb{E}(M)]$** **gap $\bar{\gamma}$** **$[\mathbb{E}(W)]$**

$-$ μ_m μ_1 ω_{n-m} ω_1 $+$

Recall also new bounds $\mathbb{RMB}$ for the m least in $\mathbb{E}(H)$ when $[\mathbb{E}(W)] > [\mathbb{E}(M)] > 0$

and $\|M^{-1}\cdot B\| < 1/\sqrt{((\frac{\mu_1}{\mu_m})^2 + \frac{\mu_1}{\mu_m})}$; then relative gaps $\Gamma(\mu_j) = \omega_{n-m}/\mu_j - 1 > 0$ and

$$0 \le 1 - \theta_{n-m+j}/\mu_j \ < \ \Psi(\|M^{-1}\cdot B\|/\Gamma(\mu_j))\cdot\|M^{-1}\cdot B\| \quad \text{for } 1 \le j \le m\,. \qquad \mathbb{RMB}$$

Is $\mathbb{RMB}$'s extra requirement " $\|M^{-1}\cdot B\| < 1/\sqrt{((\mu_1/\mu_m)^2 + \mu_1/\mu_m)}$ " unavoidable?

See example H_3 on p. 12 of …/Deflate.pdf .

Fortunately the extra requirement is very often satisfied by a permissible deflation's tiny relative error tolerance $\tau > \|M^{-1}\cdot B\|^2$, which amounts to a constraint like $\mu_m/\mu_1 > \sqrt{2\tau}$.

However, a deflation permitted by quadratic relative error-bounds $\mathbb{RMB}$ and $\mathbb{RBW}$ may turn eigenvectors through angles bigger than $\sqrt{\tau}$, thus perhaps spoiling them intolerably.

Spectral Gaps for Singular Values' Quadratic Relative Error-Bounds

Recall $S := \begin{bmatrix} D & E \\ O' & F \end{bmatrix}$, $\mathbb{S}(S) = \{ \sigma_1 \ge \sigma_2 \ge \dots \ge \sigma_n \}$, $Z := \begin{bmatrix} D & O \\ O' & F \end{bmatrix}$, $\mathbb{S}(Z) = \{ \zeta_1 \ge \zeta_2 \ge \dots \ge \zeta_n \}$.

$\mathbb{S}(Z) = \mathbb{S}(D) \uplus \mathbb{S}(F)$ where $\mathbb{S}(D) = \{ \delta_1 \ge \delta_2 \ge \dots \ge \delta_m \}$ and $\mathbb{S}(F) = \{ \phi_1 \ge \phi_2 \ge \dots \ge \phi_{n-m} \}$ are separated by gaps defined in a way now familiar (and now *overloaded*):

The *Absolute Spectral Gap* $\bar{\gamma}(\zeta)$ separates $\zeta \in \mathbb{S}(Z)$ from $\mathbb{S}(D)$ or $\mathbb{S}(F)$ thus:

If $\zeta \in \mathbb{S}(D)$ then $\bar{\gamma}(\zeta) := \min\{ |\phi - \zeta|$ over all $\phi \in \mathbb{S}(F) \}$, else
if $\zeta \in \mathbb{S}(F)$ then $\bar{\gamma}(\zeta) := \min\{ |\delta - \zeta|$ over all $\delta \in \mathbb{S}(D) \}$.

A *Relative Spectral Gap* $\Gamma(\zeta)$ separates $\zeta \in \mathbb{S}(Z)$ from $\mathbb{S}(D)$ or $\mathbb{S}(F)$ thus:

if $\zeta \in \mathbb{S}(D) \cap \mathbb{S}(F)$ then $\Gamma(\zeta) := \bar{\gamma}(\zeta) = 0$; else $\Gamma(\zeta) := \bar{\gamma}(\zeta)/\zeta$.

Li & Li [2005] applied $\mathbb{AB}$ to $\begin{bmatrix} O & D' & O & O \\ D & O & E & O \\ O' & E' & O & F' \\ O' & O' & F & O \end{bmatrix}$ whose eigenvalues are $-\mathbb{S}(S) \uplus \mathbb{S}(S)$ to get

$$|\sigma_j - \zeta_j| \le \Psi(\|E\|/\bar{\gamma}(\zeta_j))\cdot\|E\| \qquad \mathbb{AE}$$
$$< \min\{ \|E\|, \|E\|^2/\bar{\gamma}(\zeta_j) \} \quad \text{when } \|E\| > 0 \text{ and } \bar{\gamma}(\zeta_j) > 0 .$$

Those absolute error-bounds $\mathbb{AE}$ imply promptly these quadratic relative error-bounds:

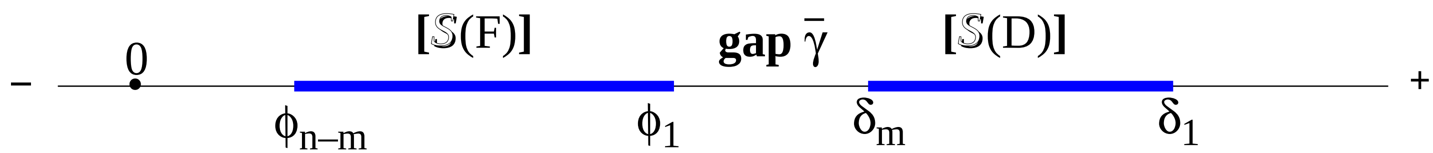
$$|\sigma_j/\zeta_j - 1| \le \Psi(\|E/\zeta_j\|/\Gamma(\zeta_j))\cdot\|E/\zeta_j\| . \qquad \mathbb{RAE}$$

Recall $S := \begin{bmatrix} D & E \\ O' & F \end{bmatrix}$, $\mathbb{S}(S) = \{ \sigma_1 \ge \sigma_2 \ge \ldots \ge \sigma_n \}$, $Z := \begin{bmatrix} D & O \\ O' & F \end{bmatrix}$, $\mathbb{S}(Z) = \{ \zeta_1 \ge \zeta_2 \ge \ldots \ge \zeta_n \}$.

$\mathbb{S}(Z) = \mathbb{S}(D) \uplus \mathbb{S}(F)$ where $\mathbb{S}(D) = \{ \delta_1 \ge \delta_2 \ge \ldots \ge \delta_m \}$ and $\mathbb{S}(F) = \{ \phi_1 \ge \phi_2 \ge \ldots \ge \phi_{n-m} \}$ are separated by relative gaps $\Gamma(\zeta_j)$; if they allow $\mathbb{S}(D)$ and $\mathbb{S}(F)$ to mingle, then

$$|\sigma_j/\zeta_j - 1| \le \Psi(\|E/\zeta_j\|/\Gamma(\zeta_j))\cdot\|E/\zeta_j\| . \qquad \mathbb{RAE}$$

To replace $\|E/\zeta_j\|$ by something perhaps smaller and cheaper to compute we relinquish mingling. Assume narrowest containing intervals to be ordered, $[\mathbb{S}(D)] > [\mathbb{S}(F)] \ge 0$, and separated by sufficiently wide relative gaps thus:



If gaps $G_j := 1 - (\phi_1/\delta_j)^2 - \|D^{-1}\cdot E\|^2 > 0$ for $1 \le j \le m$ then

$$0 \le (\sigma_j/\delta_j)^2 - 1 \le \Psi(\|D^{-1}\cdot E\|/G_j)\cdot\|D^{-1}\cdot E\| . \qquad \mathbb{RDE}$$

If gaps $\overline{G}_j := 1 - (\phi_j/\delta_m)^2 - \|D^{-1}\cdot E\|^2 > 0$ for $1 \le j \le n-m$, then (with $0/0 := 1$)

$$0 \le (\phi_j/\sigma_{m+j})^2 - 1 \le \Psi(\|D^{-1}\cdot E\|/\overline{G}_j)\cdot\|D^{-1}\cdot E\| . \qquad \mathbb{REF}$$

The appearance of $\|D^{-1}\cdot E\|$ in both $\mathbb{RDE}$ and $\mathbb{REF}$ is no accident. Compare p. 4.

All four quadratic relative error bounds $\mathbb{RMB}$, $\mathbb{RBW}$, $\mathbb{RDE}$ and $\mathbb{REF}$ look like

“ Relative Error $\le$ $\Psi(\beta/\Gamma)\cdot\beta$ ” in which

β (over)estimates $\|B\cdot W^{-1}\|$ or $\|M^{-1}\cdot B\|$ or $\|D^{-1}\cdot E\|$, and
Γ (under)estimates a relative gap.

Applications of these bounds need not compute the function Ψ because predicate

“ $\Psi(\beta/\Gamma)\cdot\beta < \tau$ ” is equivalent to the simpler “ $\beta^2 < (\tau + \Gamma)\cdot\tau$ ”.

Adequate underestimates Γ of spectral gaps are usually costly to compute unless the matrices in question are dominated enough by their diagonals, as happens during some iterative schemes to compute eigenvalues and singular values.

Formulas that help estimate gaps Γ are tabulated in §8 of …/Deflate.pdf .

CAUTION: Deflations permitted by adequately tiny quadratic relative error-bounds may preserve the accuracy of \`values but spoil the accuracy of \`vectors.

Application to Tests of Computed Eigenvalues' Relative Accuracies

Given: n-by-n $A = A'$, diagonal Λ of m computed eigenvalues, and n-by-m $\overline{Q}$ whose columns are computed *approximately* orthonormal eigenvectors.

Desired: Use Rayleigh-Ritz to tidy up Λ and $\overline{Q}$, and assess Λ's relative accuracy.

Recommended: Accumulate just *Residuals*' scalar products extra-precisely.

Process:

- To tidy up $\overline{Q}$, compute first *Residual* $V := I - \overline{Q}'\cdot\overline{Q}$, then updated $Q := \overline{Q} + \frac{1}{2}\overline{Q}\cdot V$.

  (Now new Q's residual $I - Q'\cdot Q \approx \frac{3}{4}V^2$ should be predictably negligible.)

- To tidy up Λ, compute temporary *Residual* $\overline{R} := A\cdot Q - Q\cdot\Lambda$, then $\overline{\Delta\Lambda} := Q'\cdot\overline{R}$.

  $\overline{\Delta\Lambda} = \overline{\Delta\Lambda}' \pm$ roundoff; clean it up by setting $\Delta\Lambda := \frac{1}{2}(\overline{\Delta\Lambda} + \overline{\Delta\Lambda}')$.

  Compute $M := \Lambda + \Delta\Lambda = M'$ and $R := \overline{R} - Q\cdot\Delta\Lambda \approx A\cdot Q - Q\cdot M \pm$ roundoff.

  (Now $M \approx Q'\cdot A\cdot Q$ should be nearly diagonal, and $Q'\cdot R = O \pm$ roundoff.)

- Relative errors in $\mathbb{E}(M)$ can't exceed $\|R\cdot M^{-1}\|$, which plays rôle of $\|M^{-1}\cdot B\|$ on p. 8.

  (Jacobi's iteration can compute $\mathbb{E}(M)$ quickly.)

- Total work is $O(n\cdot m^2) + \big($*Residual* $\overline{R}$'s $O(m\cdot n^2)$ at worst, much less if A is sparse$\big)$.

Application to Computation of a Bidiagonal's Singular Values by dqds

Given: n-by-n upper bidiagonal S represented by arrays $\{\sqrt{q_j}\}$ and $\{\sqrt{e_j}\}$.

Desired: Compute the squared singular values of S as eigenvalues of a tridiagonal

$$S\cdot S' = \text{tridiag}\left(\begin{bmatrix} & \sqrt{q_2\cdot e_1} & & \sqrt{q_3\cdot e_2} & & \cdots & \sqrt{q_{n-1}\cdot e_{n-2}} & & \sqrt{q_n\cdot e_{n-1}} & \\ q_1+e_1 & & q_2+e_2 & & q_3+e_3 & \cdots & & q_{n-1}+e_{n-1} & & q_n \\ & \sqrt{q_2\cdot e_1} & & \sqrt{q_3\cdot e_2} & & \cdots & \sqrt{q_{n-1}\cdot e_{n-2}} & & \sqrt{q_n\cdot e_{n-1}} & \end{bmatrix}\right)$$

without ever computing the elements of S or $S\cdot S'$ explicitly. *Why not?* p. 5.

Process: Each dqds iteration chooses a *Shift* $ß \ge 0$ and overwrites the current S by the upper bidiagonal Cholesky factor $\overline{S}$ of $S\cdot S' - ß\cdot I = \overline{S}'\cdot\overline{S}$, unless ß is too big. If $\sqrt{ß}$ > (the least singular value of S) choose a smaller ß for another attempt. After a successful attempt, update $\sum ß$ to $\underline{\sum ß} := ß + \sum ß$. Ultimately $ß \to 0$.

Ultimately, iteration drives every $e_j \to 0$ and $\sum ß + q_n \to$ (original least singular value)^2 .

Avoid lethargic convergence by exploiting every Permissible Deflation to set an $e_j \to 0$.

Tests for a tiny e_j add costs to an inner loop that already has in it one division *etc.*:

- R-C. Li's Relative Error test — a multiply and compare
- An absolute error test — a compare
- Ming Gu's absolute error test $\Rightarrow q_n \to 0 \Rightarrow e_{n-1} \to 0$ — a compare (only if $ß = 0$)

To choose ß well needs a test for min{…} too. *Which tests are indispensable?*

Conclusions

• Considering how expensive are worthwhile estimates of spectral gaps, and how rarely deflation is permitted by quadratic error-bounds, what good are they? Perhaps they serve here mostly to explain why the non-quadratic bounds of p. 4 and p. 8 are so often so pessimistic though best-possible without estimates of gaps.

The last word about quadratic error-bounds probably remains to be written.

• Like criteria for terminating an iteration, criteria for deflation have to be chosen by the error-analyst to avoid excessive computation without incurring excessive inaccuracy.

Deflation may be permitted by more than one criterion at each of very many sites; the opportunities are too numerous for all criteria to be tested at all sites. Instead an economical subset must be found.

The quest continues.

Citations in www.eecs.berkeley.edu/~wkahan/ma221/Deflate.pdf

J. Demmel, B. Diament & G. Malajovich [2001] "On the Complexity of Computing Error Bounds" pp. 101-125 in ***FOUNDATIONS OF COMPUTATIONAL MATH.*** **1**. Somewhat speculative.

J.W. Demmel & W. Kahan [1990] "Accurate Singular Values of Bidiagonal Matrices" pp. 873-912 in ***SIAM J. Sci. Stat. Comput.*** **11** #3.

K.V. Fernando & B.N. Parlett [1994] "Accurate singular values and differential qd algorithms" pp. 191-229 in ***Numerische Mathematik*** **67**.

N.J. Higham [1987] "A Survey of Condition Number Estimation for Triangular Matrices" pp. 573-596 in ***SIAM REVIEW*** **29** #4. This actually surveys (over)estimators, of norms of inverses of arbitrary triangular matrices, that cost much less to compute than the inverses do.

Leslie Hogben [2007] ed. ***Handbook of Linear Algebra*** 1504 pp., Chapman & Hall/CRC; a huge encyclopedic survey of facts citing the literature for their proofs.

C.R. Johnson [1989] "A Gersgorin-type lower bound for the smallest singular value" pp. 1-7 in ***Linear Algebra & Its Applications*** **112** : $1/\|C^{-1}\| \geq \max\{\,0,\ \min_j \{\,|c_{jj}| - \sum_{k\neq j} (|c_{kj}| + |c_{jk}|)/2\,\}\,\}$.

W. Kahan [2012'] "A Tutorial Overview of Vector and Matrix Norms" posted on my web page at www.eecs.berkeley.edu/~wkahan/MathH110/NormOvrv.pdf .

Chi-Kwong Li & Roy Mathias (1999) "The Lidskii-Mirsky-Wielandt Theorem — additive and multiplicative versions" pp. 377-413 in ***Numerische Mathematik*** **81**, surveys unitarily invariant matrix norms' relations with perturbed Hermitian matrices' spectra, with elegant proofs.

Chi-Kwong Li & Ren-Cang Li [2005] “A note on eigenvalues of perturbed Hermitian matrices”, pp. 183-190 in ***Linear Algebra and its Applications*** **395**, exploits spectral gaps optimally.

Ren-Cang Li [1994] “On Deflating Bidiagonal Matrices” unpublished note, Mathematics Dept., Univ. of Calif. @ Berkeley, CA 94720. The formulation here was derived as a special case of more general and much more complicated relationships published later in five papers …

R-C. Li [1997] “Relative Perturbation Theory: (III) More Bounds On Eigenvalue Variation” pp. 337—345 in ***Linear Algebra and its Applications*** **266**.

R-C. Li [1998] “Relative Perturbation Theory: (I) Eigenvalue and Singular Value Variations” pp. 956—982 in ***SIAM J. Matrix Anal. Appl.*** **19**.

R-C. Li [1999] “Relative Perturbation Theory: (II) Eigenspace And Singular Space Variations” pp. 471—492 in ***SIAM J. Matrix Anal. Appl.*** **20**.

R-C. Li [2000] “Relative Perturbation Theory: (IV) sin2θ Theorems” pp. 45—60 in ***Linear Algebra and its Applications*** **311**.

R-C. Li [2000] “A Bound On The Solution To A Structured Sylvester Equation With An Application To Relative Perturbation Theory” pp. 471—492 in ***SIAM J. Matrix Anal. Appl***. **21**

S-G. Li, M. Gu & B.N. Parlett [2012] “A Modified DQDS Algorithm” to appear.

B.N. Parlett [1998] ***The Symmetric Eigenvalue Problem*** 426 pp., SIAM, Philadelphia

B.N. Parlett & O.A. Marques [2000] “An Implementation of the dqds Algorithm (Positive Case)” pp. 217-259 in ***Linear Algebra and its Applications*** **309**.