Accurate Computations of Positive Polar Factors of Graded Matrices

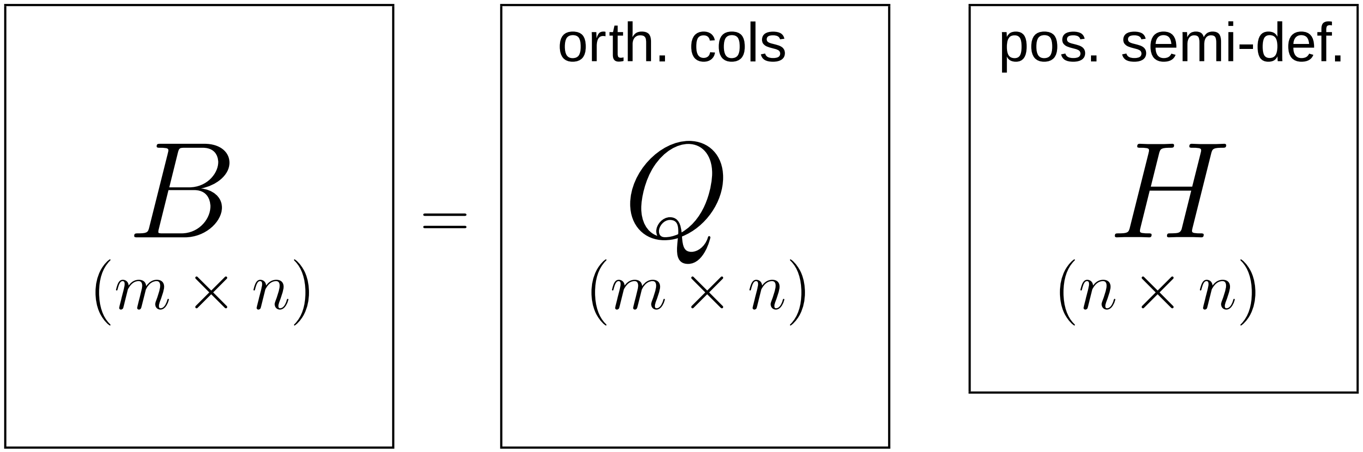
Ren-Cang Li
rcli@ms.uky.edu

http://www.ms.uky.edu/∼rcli

OUTLINE

- ▷ Polar Decomposition
- Existing Perturbation Bounds
- Relative Perturbation Bound for Positive Polar Factors
- Accurate Computations
- Extension to the Matrix Square Root
- Concluding remarks

Polar Decomposition



- Polar decomposition via SVD $B = U_{m \times n} \Sigma V^*$:

$$Q = UV^*, \quad H = V \Sigma V^*.$$

- H *always* unique. Q unique if $\operatorname{rank}(B) = n$.
- If $\operatorname{rank}(B) = n$, then $H = (B^* B)^{1/2}$ and $Q = B(B^* B)^{-1/2}$.

OUTLINE

- Polar Decomposition
- ▷ Existing Perturbation Bounds
- Relative Perturbation Bound for Positive Polar Factors
- Accurate Computations
- Extension to the Matrix Square Root
- Concluding remarks

Existing Perturbation Bounds – Q (absolute pert.)

$B = QH$ perturbed to $\widetilde{B} = \widetilde{Q}\widetilde{H}$.

- (*Mathias, 1993*) If $B \in \mathbb{R}^{n\times n}$, $\text{rank}\,(B) = n$, $\|\Delta B\|_2 < \sigma_n$, then

$$|||\Delta Q||| \le -\frac{2\,|||\Delta B|||}{|||\Delta B|||_2} \ln\left(1 - \frac{|||\Delta B|||_2}{\sigma_n + \sigma_{n-1}}\right) \approx \frac{2}{\sigma_n + \sigma_{n-1}}\,|||\Delta B|||\,,$$

  where $|||\cdot|||_2$ sum of first two largest singular values.

- (*Li, 1995*) If $B \in \mathbb{C}^{n\times n}$ and $\text{rank}\,(B) = n$, then for all unitarily invariant norm $|||\Delta Q||| \le \dfrac{2}{\sigma_n + \widetilde{\sigma}_n}\,|||\Delta B|||\,.$

- (*W. Li and W. Sun, 2002*) If $B \in \mathbb{R}^{n\times n}$, $\text{rank}\,(B) = n$, and $\|\Delta B\|_2 < \sigma_n + \widetilde{\sigma}_n$, then $\|\Delta Q\|_{\mathsf{F}} \le \dfrac{4}{\sigma_{n-1} + \sigma_n + \widetilde{\sigma}_{n-1} + \widetilde{\sigma}_n}\|\Delta B\|_{\mathsf{F}}$.

Bounds for $m \times n$ B ($m > n$) established, too.

Existing Perturbation Bounds – Q (multiplicative pert.)

- (*Li, 1995*) If B and $\widetilde{B} = D_{\mathsf{L}}^* B D_{\mathsf{R}}$ have full column rank, then

$$\begin{aligned}\|\Delta Q\|_{\mathsf{F}} \;\; &\leq \;\; \sqrt{\|I - D_{\mathsf{L}}^{-1}\|_{\mathsf{F}}^2 + \|D_{\mathsf{L}} - I\|_{\mathsf{F}}^2} \\ &\quad + \sqrt{\|I - D_{\mathsf{R}}^{-1}\|_{\mathsf{F}}^2 + \|D_{\mathsf{R}} - I\|_{\mathsf{F}}^2}.\end{aligned}$$

- (*Li, 1997*) (Graded Case) Let $B = GS$ and $\widetilde{B} = \widetilde{G}S$ and assume that G and B have full column rank. If $\|\Delta G\|_2 \|G^\dagger\|_2 < 1$, then

$$\|\Delta Q\|_{\mathsf{F}} \leq \gamma \, \|G^\dagger\|_2 \|\Delta G\|_{\mathsf{F}}, \quad \gamma = \sqrt{1 + \left(1 - \|G^\dagger\|_2 \|\Delta G\|_2\right)^{-2}}.$$

Existing Perturbation Bounds – H always well-conditioned

Kittaneh (1986): $\|H - \widetilde{H}\|_{\mathsf{F}} \leq \sqrt{2}\,\|B - \widetilde{B}\|_{\mathsf{F}}$.

Compared to Q*-factor more extensively studied in the last 20 years*.

H's well-conditionedness also established by *Higham (1986), Barrlund (1989), Kenney and Laub (1991), Chatelin and Gratton (2000)*.

H's well-conditionedness, however, is in the *norm-wise sense*. Relatively? – focus of this talk.

Graded Case – an Example

$$B = GS = \begin{pmatrix} 6 & -2 & 14 & -5 \\ 8 & 5 & -7 & -8 \\ -2 & -11 & 2 & -3 \\ 5 & -8 & -16 & 9 \end{pmatrix} \begin{pmatrix} 10^6 & & & \\ & 10^4 & & \\ & & 10^2 & \\ & & & 1 \end{pmatrix}$$

perturbed to $\widetilde{B} = \widetilde{G}S$ with $\widetilde{G} = G + \Delta G$, $\Delta G = \texttt{10}^{-5} * \texttt{randn}(4)$.

$$H = \begin{pmatrix} 1.1358 \cdot 10^7 & 8.6928 \cdot 10^3 & -4.9320 \cdot 10^2 & -3.7828 \cdot 10^0 \\ 8.6928 \cdot 10^3 & 1.4603 \cdot 10^5 & 3.1908 \cdot 10^2 & -4.4827 \cdot 10^0 \\ -4.9320 \cdot 10^2 & 3.1908 \cdot 10^2 & 2.1691 \cdot 10^3 & -7.7287 \cdot 10^0 \\ -3.7828 \cdot 10^0 & -4.4827 \cdot 10^0 & -7.7287 \cdot 10^0 & 9.2121 \cdot 10^0 \end{pmatrix}$$

Graded Case – an Example (cont'd)

$\|H - \widetilde{H}\|_{\mathsf{F}} \le \sqrt{2}\|\Delta GS\|_{\mathsf{F}} \approx 24.7393$, no assurance to smaller entries!

On the other hand,

$$
\begin{aligned}
\text{(by absolute pert. bound)} \quad \|\Delta Q\|_{\mathsf{F}} &\le 1.90, \\
\text{(by relative pert. bound)} \quad \|\Delta Q\|_{\mathsf{F}} &\le 4.41 \times 10^{-6}.
\end{aligned}
$$

Need a better theory for H-factor!

OUTLINE

- Polar Decomposition
- Existing Perturbation Bounds
- ▷ Relative Perturbation Bound for Positive Polar Factors
- Accurate Computations
- Extension to the Matrix Square Root
- Concluding remarks

Relative Perturbation Bound for Positive Polar Factors

$B = GS$, $\widetilde{B} = \widetilde{G}S$. Both have full column rank. S is $n \times n$ and nonsingular, not necessarily diagonal.

Theorem. If $\|\Delta G\|_2 \|G^\dagger\|_2 < 1$ then

$$\begin{aligned} \|(\widetilde{H} - H)S^{-1}\|_\mathsf{F} &\leq \sqrt{\|(\Delta G)G^\dagger\|_\mathsf{F}^2 + \left\| I - \left(I + (\Delta G)G^\dagger\right)^{-1} \right\|_\mathsf{F}^2} \, \|G\|_2 + \|\Delta G\|_\mathsf{F} \\ &\leq \left(\sqrt{1 + \frac{1}{(1 - \|G^\dagger\|_2 \|\Delta G\|_2)^2}} \, \|G^\dagger\|_2 \|G\|_2 + 1 \right) \|\Delta G\|_\mathsf{F}. \end{aligned}$$

Example (revisited)

$$\begin{aligned} \|(\widetilde{H} - H)S^{-1}\|_\mathsf{F} &\leq 1.4298 \cdot 10^{-4}, \\ &\leq 2.0756 \cdot 10^{-4}. \end{aligned}$$

And ...

Relative Perturbation Bound – Analysis for Diagonal S

Entry-wise $\|(\widetilde{H} - H)S^{-1}\|_{\mathsf{F}} \leq \varepsilon$ gives

$$|h_{ij} - \widetilde{h}_{ij}| = |h_{ji} - \widetilde{h}_{ji}| \leq \varepsilon \min\{|s_i|, |s_j|\}.$$

Numbers of correct significant decimal digits $-\log_{10}(|h_{ij} - \widetilde{h}_{ij}|/|h_{ij}|)$ are at least

$$-\log_{10}(\varepsilon \min\{|s_i|, |s_j|\}/|h_{ij}|).$$

Use $\varepsilon = 1.4298 \cdot 10^{-4}$. Numbers of correct significant decimal digits are at least, entry-wise,

$$\begin{pmatrix} 4.9 & 3.8 & 4.5 & 4.4 \\ 3.8 & 5.0 & 4.3 & 4.5 \\ 4.5 & 4.3 & 5.2 & 4.7 \\ 4.4 & 4.5 & 4.7 & 4.8 \end{pmatrix},$$

smaller entries contaminated nor more than larger ones.

Relative Perturbation Bound – Proof

$$
\begin{aligned}
\widetilde{Q}\widetilde{H} &= \widetilde{G}S, \\
\widetilde{H}S^{-1} &= \widetilde{Q}^*\widetilde{G} \\
&= \widetilde{Q}^*(G+\Delta G) \\
&= (Q+\Delta Q)^*G + \widetilde{Q}^*\Delta G \\
&= Q^*G + \Delta Q^*G + \widetilde{Q}^*\Delta G \\
&= HS^{-1} + \Delta Q^*G + \widetilde{Q}^*\Delta G.
\end{aligned}
$$

Thus $(\widetilde{H}-H)S^{-1} = \Delta Q^*G + \widetilde{Q}^*\Delta G$, and

$$\|(\widetilde{H}-H)S^{-1}\|_{\mathsf{F}} \le \|\Delta Q^*\|_{\mathsf{F}}\|G\|_2 + \|\Delta G\|_{\mathsf{F}}.$$

Use relative perturbation bound for $\|\Delta Q^*\|_{\mathsf{F}}$ to complete proof.

OUTLINE

- Polar Decomposition
- Existing Perturbation Bounds
- Relative Perturbation Bound for Positive Polar Factors
- ▷ Accurate Computations
- Extension to the Matrix Square Root
- Concluding remarks

Accurate Computations When S Diagonal

Algorithm.

1. Compute SVD: $B \equiv GS = U\Sigma V^*$ by one-sided Jacobi;
2. Set $Q = UV^*$, $W = Q^*G$, and then $H = WS$.

Note: H by $V\Sigma V^*$ does not work!

Use the following floating point arithmetic model

$$\text{computed } \alpha \odot \beta \text{ is } (\alpha \odot \beta)(1+\delta), \ |\delta| \le \epsilon_m \text{ for } \odot \in \{+, -, \times, \div\},$$

we have

Theorem. Computed $\widetilde{H}$ satisfies $\|(\widetilde{H} - H)S^{-1}\|_{\mathsf{F}} = \mathcal{O}\left(\epsilon_{\mathsf{m}}\kappa(G)\|G\|_{\mathsf{F}}\right)$.

Accurate Computations – Outline of Proof

Denote computed matrices by the same symbols but with *tildes*.

1. Compute SVD: $B \equiv GS = U\Sigma V^*$ by one-sided Jacobi;

   By Demmel, Gu, Eisenstat, Slapničar, Veselić, and Drmač (1999),

   $$\widetilde{U}\widetilde{\Sigma}\widetilde{V}^* \equiv \widetilde{B} = (I+E)B(I+F),\ \|E\|_{\mathsf{F}} = \mathcal{O}\left(\epsilon_{\mathsf{m}}\kappa(G)\right),\ \|F\|_{\mathsf{F}} = \mathcal{O}\left(\epsilon_{\mathsf{m}}\kappa(G)\right).$$

2. Set $Q = UV^*$, $W = Q^*G$, and then $H = WS$.

   Let $\circ$ denote the entry-wise Hadamard product. Then

   $$\widetilde{Q} = \widetilde{U}\widetilde{V}^* + E_1, \quad \widetilde{W} = \widetilde{Q}^*G + E_2, \quad \widetilde{H} = (\widetilde{W}S) \circ M,$$
   $$\|E_1\|_{\mathsf{F}} = \mathcal{O}\left(\epsilon_{\mathsf{m}}\right), \ \|E_2\|_{\mathsf{F}} = \mathcal{O}\left(\epsilon_{\mathsf{m}}\|G\|_{\mathsf{F}}\right), \quad M_{ij} = 1 + \mathcal{O}\left(\epsilon_{\mathsf{m}}\right)$$

Accurate Computations – Outline of Proof (`cont'd`)

Note $\widetilde{U}\widetilde{V}^* = Q + E_3$ *with* $\|E_3\|_{\mathsf{F}} = \mathcal{O}\left(\|E\|_{\mathsf{F}} + \|F\|_{\mathsf{F}}\right) = \mathcal{O}\left(\epsilon_{\mathsf{m}}\kappa(G)\right)$

$$
\begin{aligned}
\widetilde{H}S^{-1} &= \widetilde{W} \circ M \\
&= (\widetilde{Q}^* G + E_2) \circ M \\
&= [(\widetilde{U}\widetilde{V}^* + E_1)^* G + E_2] \circ M \\
&= [(\widetilde{U}\widetilde{V}^*)^* G] \circ M + (E_1^* G + E_2) \circ M \\
&= (Q^* G) \circ M + (E_3^* G + E_1^* G + E_2) \circ M \\
&= Q^* G + (Q^* G) \circ E_4 + (E_3^* G + E_1^* G + E_2) \circ M \\
&= HS^{-1} + (Q^* G) \circ E_4 + (E_3^* G + E_1^* G + E_2) \circ M,
\end{aligned}
$$

which gives

$$
\|(\widetilde{H} - H)S^{-1}\|_{\mathsf{F}} \leq \|(Q^* G) \circ E_4 + (E_3^* G + E_1^* G + E_2) \circ M\|_{\mathsf{F}} = \mathcal{O}\left(\epsilon_{\mathsf{m}}\kappa(G)\|G\|_{\mathsf{F}}\right).
$$

Accurate Computations – an Example

$n = 10$, $B = GS$, $S = \text{diag}\,(10^3, 10^8, 10^5, 10^4, 1, 10^4, 10^9, 10^8, 10^3, 10^8)$,

$$G = \begin{pmatrix} 4.656 & 7.220 & 3.831 & 2.924 & 7.556 & 7.329 & 4.105 & 2.827 & 6.787 & 7.7860 \\ 4.187 & 2.644 & 5.986 & 9.774 & 7.660 & 7.170 & 1.042 & 2.240 & 5.235 & 6.1470 \\ 7.518 & 6.780 & 5.691 & 1.071 & 8.527 & 2.231 & 0.220 & 7.386 & 9.997 & 4.8770 \\ 3.829 & 7.801 & 2.164 & 1.158 & 1.907 & 4.992 & 7.945 & 9.693 & 7.529 & 2.1330 \\ 3.515 & 1.907 & 9.999 & 0.686 & 2.950 & 2.240 & 5.921 & 9.097 & 5.180 & 3.4390 \\ 0.871 & 4.511 & 4.406 & 3.102 & 4.731 & 1.793 & 3.100 & 9.608 & 6.400 & 6.9290 \\ 8.565 & 8.088 & 9.318 & 3.557 & 6.034 & 9.012 & 1.087 & 8.167 & 1.351 & 4.6020 \\ 1.351 & 5.289 & 3.025 & 3.591 & 1.210 & 2.123 & 8.771 & 4.650 & 8.002 & 9.6420 \\ 2.408 & 2.745 & 3.893 & 4.202 & 5.845 & 3.501 & 0.603 & 2.775 & 0.912 & 9.1680 \\ 0.231 & 2.726 & 6.898 & 9.243 & 3.813 & 5.065 & 0.594 & 8.415 & 5.140 & 1.9450 \end{pmatrix},$$

$\|B\|_{\mathsf{F}} = 3.44 \cdot 10^{10}$, $\kappa(B) = 5.37 \cdot 10^{10}$, $\kappa(G) = 4.60 \cdot 10^{2}$, $\|G\|_{\mathsf{F}} = 5.02 \cdot 10^{1}$.

Accurate Computations – an Example (`cont'd`)

Jacobi: one-sided Jacobi SVD code, courtesy of Z. Drmač.

QR: `sgesvd` from LAPACK.

Exact Q and H by Maple with a 50 decimal digit arithmetic.

	$\|(\widetilde{H} - H)S^{-1}\|_{\mathsf{F}}$	$\|\widetilde{H} - H\|_{\mathsf{F}}$	$\|\widetilde{Q} - Q\|_{\mathsf{F}}$	$\max \|\widetilde{\sigma} - \sigma\|/\sigma$
Jacobi	1.19e-5	8.83e+2	1.75e-6	4.55e-6
QR	8.67e+1	2.47e+3	2.00e+0	2.19e+1

In both cases, norm-wise relative error $\dfrac{\|\widetilde{H} - H\|_{\mathsf{F}}}{\|H\|_{\mathsf{F}}} \approx 10^{-7}$, good enough unless interested in knowing smaller entries.

OUTLINE

- Polar Decomposition
- Existing Perturbation Bounds
- Relative Perturbation Bound for Positive Polar Factors
- Accurate Computations
- ▷ Extension to the Matrix Square Root
- Concluding remarks

Extension to the Matrix Square Root

$A = S^*TS$ perturbed to $\widetilde{A} = S^*\widetilde{T}S$, S ($n \times n$) nonsingular, T positive definite and well-conditioned.

Theorem. Let $\widehat{T} \overset{\text{def}}{=} T^{-1/2}(\widetilde{T} - T)T^{-1/2}$ and $\delta_p = \|\widehat{T}\|_p$ ($p = 2, \mathsf{F}$). If $\delta_2 < 1$, then

$$\begin{aligned} \|(\widetilde{A}^{1/2} - A^{1/2})S^{-1}\|_{\mathsf{F}} &\leq \left(\sqrt{\frac{2-\delta_2}{1-\delta_2}} + 1\right) \|T^{1/2}\|_2 \frac{\delta_{\mathsf{F}}}{1 + \sqrt{1-\delta_2}}, \\ &\approx \frac{\sqrt{2}+1}{2} \|T^{1/2}\|_2 \delta_{\mathsf{F}}. \end{aligned}$$

Existing bound.

$$\|\widetilde{A}^{1/2} - A^{1/2}\|_{\mathsf{F}} \leq \frac{1}{\|\widetilde{A}^{-1/2}\|_2^{-1} + \|A^{-1/2}\|_2^{-1}} \|\widetilde{A} - A\|_{\mathsf{F}}.$$

Extension to the Matrix Square Root – an Example

$A = B^*B = S^*TS$ with

$$B = GS = \begin{pmatrix} 6 & -2 & 14 & -5 \\ 8 & 5 & -7 & -8 \\ -2 & -11 & 2 & -3 \\ 5 & -8 & -16 & 9 \end{pmatrix} \begin{pmatrix} 10^6 & & & \\ & 10^4 & & \\ & & 10^2 & \\ & & & 1 \end{pmatrix}.$$

Take $E = 10^{-5} * \texttt{randn}(4)$ and $\Delta T = E + E^* + EE^*$. Then

Existing bound $\qquad \|\widetilde{A}^{1/2} - A^{1/2}\|_{\mathsf{F}} \leq 1.4288\, 10^6,$

New bound $\qquad \|(\widetilde{A}^{1/2} - A^{1/2})S^{-1}\|_{\mathsf{F}} \leq 1.17363\, 10^{-5}.$

Extension to the Matrix Square Root – Outline of Proof

$$
\begin{aligned}
A &= S^*T^{1/2}T^{1/2}S = (GS)^*(GS) \equiv B^*B, \\
\widetilde{A} &= S^*(T+\Delta T)S = S^*T^{1/2}\left[I + T^{-1/2}(\Delta T)T^{-1/2}\right]T^{1/2}S = \widetilde{B}^*\widetilde{B}.
\end{aligned}
$$

$\widetilde{B} = \widetilde{G}S$ with $\widetilde{G} = \left(I + \widehat{T}\right)^{1/2} T^{1/2}$ and $\widehat{T} = T^{-1/2}(\Delta T)T^{-1/2}$.

Polar decompositions $B = QH$, $\widetilde{B} = \widetilde{Q}\widetilde{H} \Longrightarrow A^{1/2} = H$, $\widetilde{A}^{1/2} = \widetilde{H}$.

Previous relative perturbation bound for H-factor applies:

$$
\begin{aligned}
\Delta G &= \widetilde{G} - G = [\left(I + \widehat{T}\right)^{1/2} - I]G, \\
(\Delta G)G^{-1} &= \left(I + \widehat{T}\right)^{1/2} - I, \\
I - \left(I + (\Delta G)G^{-1}\right)^{-1} &= I - \left(I + \widehat{T}\right)^{-1/2}.
\end{aligned}
$$

Extension to the Matrix Square Root – Accurate Computation

Algorithm (for diagonal S).

1. Decompose $T = G^*G$ (e.g., Cholesky decomposition);
2. Compute polar decomposition $GS = QH$ by the previous algorithm, and return $H \approx A^{1/2}$.

Theorem. Computed $\widetilde{H} \approx A^{1/2}$ satisfies
$\|(\widetilde{H} - A^{1/2})S^{-1}\|_\mathsf{F} = \mathcal{O}\left(\epsilon_\mathsf{m}\kappa(T)\|T^{1/2}\|_\mathsf{F}\right)$.

Outline of Proof. Computed $\widetilde{G}$ satisfies $\|\Delta T\|_\mathsf{F} = \mathcal{O}\left(\epsilon_\mathsf{m}\|T\|_\mathsf{F}\right)$ (*J. Sun 1992*). Let $\widehat{A} = S^*\widetilde{G}^*\widetilde{G}S$.

$$\begin{aligned}
\|(\widehat{A}^{1/2} - A^{1/2})S^{-1}\|_\mathsf{F} &= \mathcal{O}\left(\epsilon_\mathsf{m}\|T^{1/2}\|_2\|T^{-1}\|_2\|T\|_\mathsf{F}\right) \approx \mathcal{O}\left(\epsilon_\mathsf{m}\kappa(T)\,\|T^{1/2}\|_2\right), \\
\|(\widetilde{H} - \widehat{A}^{1/2})S^{-1}\|_\mathsf{F} &= \mathcal{O}\left(\epsilon_\mathsf{m}\kappa(\widetilde{G})\|\widetilde{G}\|_\mathsf{F}\right) \approx \mathcal{O}\left(\epsilon_\mathsf{m}\kappa(T^{1/2})\|T^{1/2}\|_\mathsf{F}\right).
\end{aligned}$$

OUTLINE

- Polar Decomposition
- Existing Perturbation Bounds
- Relative Perturbation Bound for Positive Polar Factors
- Accurate Computations
- Extension to the Matrix Square Root
- ▷ Concluding remarks

Concluding remarks

- New perturbation bound measuring the scaled difference in positive polar factors
- In highly graded cases, smaller entries in H contaminated no more than larger ones
- Can compute H as accurately as predicted
- Extensible to matrix square roots.