Orthogonalization with a non-standard inner product and approximate inverse preconditioning

Miro Rozložník

joint work with Jiří Kopal, Alicja Smoktunowicz and Miroslav Tůma

Institute of Computer Science, Czech Academy of Sciences, Prague, Czech Republic

SPOMECH workshop, Ostrava, November 22-24, 2011

Orthogonalization with a non-standard inner product

$A \in \mathcal{R}^{m,m}$ symmetric positive definite, inner product $\langle \cdot, \cdot \rangle_A$
$Z^{(0)} = [z_1^{(0)}, \ldots, z_n^{(0)}] \in \mathcal{R}^{m,n}$, $m \geq n = rank(Z^{(0)})$

A-orthogonal basis of $span(Z^{(0)})$:

$Z = [z_1, \ldots, z_n] \in \mathcal{R}^{m,n}$, $Z^T A Z = I$

$Z^{(0)} = ZU$, $U \in \mathcal{R}^{n,n}$ upper triangular

$$(Z^{(0)})^T A Z^{(0)} = U^T U$$

Non-standard orthogonalization and standard QR factorization

$$A^{1/2}Z^{(0)} = (A^{1/2}Z)U, \;\; Z^TAZ = (A^{1/2}Z)^T(A^{1/2}Z) = I$$

$$\kappa(Z) \leq \kappa^{1/2}(A)$$

$$\kappa(U) = \kappa(A^{1/2}Z^{(0)}) \leq \kappa^{1/2}(A)\kappa(Z^{(0)})$$

$Z^{(0)} = [e_1, \ldots, e_n]$: $Z = \begin{bmatrix} U^{-1} \\ 0 \end{bmatrix} \in \mathcal{R}^{m,n}$ upper triangular

$$\kappa(U) = \kappa(Z) = \kappa^{1/2}(A)$$

Inverse factorization and approximate inverse preconditioning

$$ZZ^T = Z^{(0)}U^{-1}U^{-T}(Z^{(0)})^T = Z^{(0)}[(Z^{(0)})^T A Z^{(0)}]^{-1}(Z^{(0)})^T$$

AZZ^T : orthogonal projector onto $R(AZ^{(0)})$ and orthogonal to $R(Z^{(0)})$

$ZZ^T A$: orthogonal projector onto $R(Z^{(0)})$ and orthogonal to $R(AZ^{(0)})$

$Z^{(0)}$ square and nonsingular: inverse factorization $ZZ^T = A^{-1}$

$Ax = b$, approximate inverse $\bar{Z}\bar{Z}^T \approx A^{-1}$

$$\downarrow$$

$$\bar{Z}^T A \bar{Z} y = \bar{Z}^T b \ , \ \ x = \bar{Z} y, \ \ \|\bar{Z}^T A \bar{Z} - I\| \leq ?$$

$$\bar{U}^T \bar{U} \approx (Z^{(0)})^T A Z^{(0)}, \ \ \|(Z^{(0)})^T A Z^{(0)} - \bar{U}^T \bar{U}\| \leq ?$$

$$\bar{Z}\bar{U} \approx Z^{(0)}, \ \ \|Z^{(0)} - \bar{Z}\bar{U}\| \leq ?$$

Reference approach: Eigenvalue (EIG) based implementation

$$A = V\Lambda V^T, \;\; \Lambda^{1/2} V^T Z^{(0)} = QU, \;\; Z = V\Lambda^{-1/2} Q \text{ (or } Z = Z^{(0)} U^{-1}\text{)}$$

backward stable eigendecomposition + backward stable QR:

$$\|\bar{Z}^T A \bar{Z} - I\| \leq \mathcal{O}(u)\|A\|\|\bar{Z}\|^2$$
$$\|Z^{(0)} - \bar{Z}\bar{U}\| \leq \mathcal{O}(u)\|\bar{Z}\|\|\bar{U}\|$$
$$\|(Z^{(0)})^T A Z^{(0)} - \bar{U}^T \bar{U}\| \leq \mathcal{O}(u)\|A\|^{1/2}\|Z^{(0)}\|\|A^{1/2}Z^{(0)}\|$$

Practical algorithms: classical Gram-Schmidt (CGS) algorithm [Schmidt 1907, 1908], modified Gram-Schmidt (MGS) algorithm [Laplace 1816, Cauchy 1837], Gram-Schmidt with reorthogonalization [Kahan, Parlett, Kaufmann et al...]
Theory for standard inner product $A = I$: [MGS: Björck 1967, Björck and Paige 1992, CGS: Giraud, Langou, R, van den Eshof 2005, Barlow, Langou, Smoktunowicz 2006, CGS2: Giraud, Langou, R, van den Eshof 2005]

Gram-Schmidt orthogonalization

$$z_i^{(j)} = z_i^{(j-1)} - \alpha_{ji} z_j, \quad j = 1, \ldots, i-1$$
$$z_i = z_i^{(i-1)} / \alpha_{ii}, \quad \alpha_{ii} = \|z_i^{(i-1)}\|_A$$

modified Gram-Schmidt (MGS) algorithm $\equiv$ SAINV algorithm: $\alpha_{ji} = \langle z_i^{(j-1)}, z_j \rangle_A$
classical Gram-Schmidt (CGS) algorithm: $\alpha_{ji} = \langle z_i^{(0)}, z_j \rangle_A$
AINV algorithm: oblique projections $\alpha_{ji} = \langle z_i^{(j-1)}, z_j^{(0)} / \alpha_{jj} \rangle_A$

Approximate inverse preconditioning: from CGS in AINV to MGS in SAINV; various computation of diagonal entries [Benzi, Meyer, Tůma 1996; Benzi, Cullum, Tůma 2000; Benzi, Tůma 2003; Kharchenko, Kolotilina, Nikishin, Yeremin 2001]

Local errors in the (modified) Gram-Schmidt process

$$z_i^{(j)} = z_i^{(j-1)} - \alpha_{ji}\bar{z}_j$$

$$\alpha_{ji} = \langle z_i^{(j-1)}, \bar{z}_j \rangle_A$$

$$\langle z_i^{(j)}, \bar{z}_j \rangle_A = (1 - \|\bar{z}_j\|_A^2)\langle z_i^{(j-1)}, \bar{z}_j \rangle_A$$

$$z_i^{(j)} = z_i^{(j-1)} - \bar{\alpha}_{ji} z_j$$

$$\bar{\alpha}_{ji} = \mathrm{fl}[\langle z_i^{(j-1)}, z_j \rangle_A]$$

$$\langle z_i^{(j)}, z_j \rangle_A = \left(\mathrm{fl}[\langle z_i^{(j-1)}, z_j \rangle_A] - \langle z_i^{(j-1)}, z_j \rangle_A\right) \|z_j\|_A^2$$

Loss of orthogonality in the Gram-Schmidt algorithms

modified Gram-Schmidt algorithm:

$$\mathcal{O}(u)\kappa(A)\kappa(A^{1/2}Z^{(0)}) < 1$$

$$\|I - \bar{Z}^T A \bar{Z}\| \leq \frac{\mathcal{O}(u)\|A\|^{1/2}\|\bar{Z}\| \max_{i,j} \frac{\|A\|^{1/2}\|\bar{z}_i^{(j)}\|}{\|z_i^{(0)}\|_A} \kappa(A^{1/2}Z^{(0)})}{1 - \mathcal{O}(u)\|A\|^{1/2}\|\bar{Z}\| \max_{i,j} \frac{\|A\|^{1/2}\|\bar{z}_i^{(j)}\|}{\|z_i^{(0)}\|_A} \kappa(A^{1/2}Z^{(0)})}$$

classical Gram-Schmidt and AINV algorithm:

$$\mathcal{O}(u)\kappa(A)\kappa(A^{1/2}Z^{(0)})\kappa(Z^{(0)}) < 1$$

$$\|I - \bar{Z}^T A \bar{Z}\| \leq \frac{\mathcal{O}(u)\|A\|^{1/2}\|\bar{Z}\|\kappa(A^{1/2}Z^{(0)})\kappa^{1/2}(A)\kappa(Z^{(0)})}{1 - \mathcal{O}(u)\|A\|^{1/2}\|\bar{Z}\|\kappa(A^{1/2}Z^{(0)})\kappa^{1/2}(A)\kappa(Z^{(0)})}$$

Classical Gram-Schmidt (CGS2) with reorthogonalization

$$z_i^{(1)} = z_i^{(0)} - \sum_{j=1}^{i-1} \alpha_{ji}^{(1)} z_j, \quad \alpha_{ji}^{(1)} = \langle z_i^{(0)}, z_j \rangle_A$$

$$z_i^{(2)} = z_i^{(1)} - \sum_{j=1}^{i-1} \alpha_{ji}^{(2)} z_j, \quad \alpha_{ji}^{(2)} = \langle z_i^{(1)}, z_j \rangle_A$$

$$z_i = z_i^{(2)} / \alpha_{ii}, \quad \alpha_{ii} = \| z_i^{(2)} \|_A$$

$$\mathcal{O}(u) \kappa^{1/2}(A) \kappa(A^{1/2} Z^{(0)}) < 1$$

$$\| I - \bar{Z}^T A \bar{Z} \| \leq \mathcal{O}(u) \|A\| \|\bar{Z}\| \|\bar{Z}^{(1)}\|$$

Twice-is-enough algorithm of Kahan-Parlett, iterated Gram-Schmidt process [Rice 1966, CGS2: Abdelmalek 1971, Daniel,Gragg, Kaufmann, Stewart 1976; Hoffman 1989; Giraud, Langou, R, van den Eshof 2005]

Local errors in the computation of inner products and normalization

general positive definite A :

$$|\mathrm{fl}[\langle \bar{z}_i^{(j-1)}, \bar{z}_j \rangle_A] - \langle \bar{z}_i^{(j-1)}, \bar{z}_j \rangle_A| \leq \mathcal{O}(u) \|A\| \|\bar{z}_i^{(j-1)}\| \|\bar{z}_j\|$$
$$|1 - \|\bar{z}_j\|_A^2| \leq \mathcal{O}(u) \|A\|^{1/2} \|\bar{z}_j\|$$

diagonal (weight matrix) A :

$$|\mathrm{fl}[\langle \bar{z}_i^{(j-1)}, \bar{z}_j \rangle_A] - \langle \bar{z}_i^{(j-1)}, \bar{z}_j \rangle_A| \leq \mathcal{O}(u) \|\bar{z}_i^{(j-1)}\|_A \|\bar{z}_j\|_A$$
$$|1 - \|\bar{z}_j\|_A^2| \leq \mathcal{O}(u)$$

A diagonal similar to orthogonalization with the standard inner product

MGS algorithm:

$$\mathcal{O}(u)\kappa(A^{1/2}Z^{(0)}) < 1$$

$$\|I - \bar{Z}^T A \bar{Z}\| \leq \frac{\mathcal{O}(u)\kappa(A^{1/2}Z^{(0)})}{1 - \mathcal{O}(u)\kappa(A^{1/2}Z^{(0)})}$$

CGS and AINV algorithms:

$$\mathcal{O}(u)\kappa^2(A^{1/2}Z^{(0)}) < 1$$

$$\|I - \bar{Z}^T A \bar{Z}\| \leq \frac{\mathcal{O}(u)\kappa^2(A^{1/2}Z^{(0)})}{1 - \mathcal{O}(u)\kappa^2(A^{1/2}Z^{(0)})}$$

CGS with reorthogonalization:

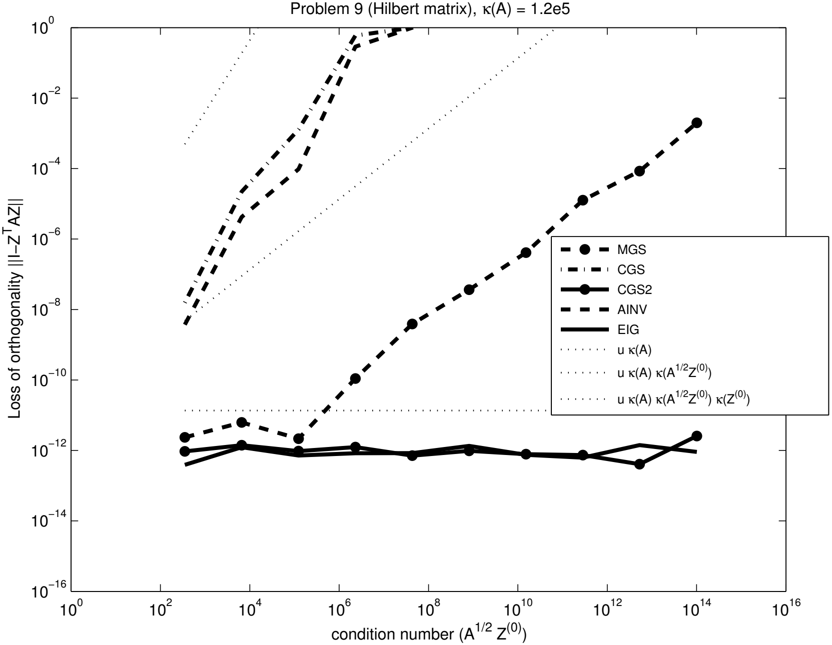
$$\mathcal{O}(u)\kappa(A^{1/2}Z^{(0)}) < 1$$

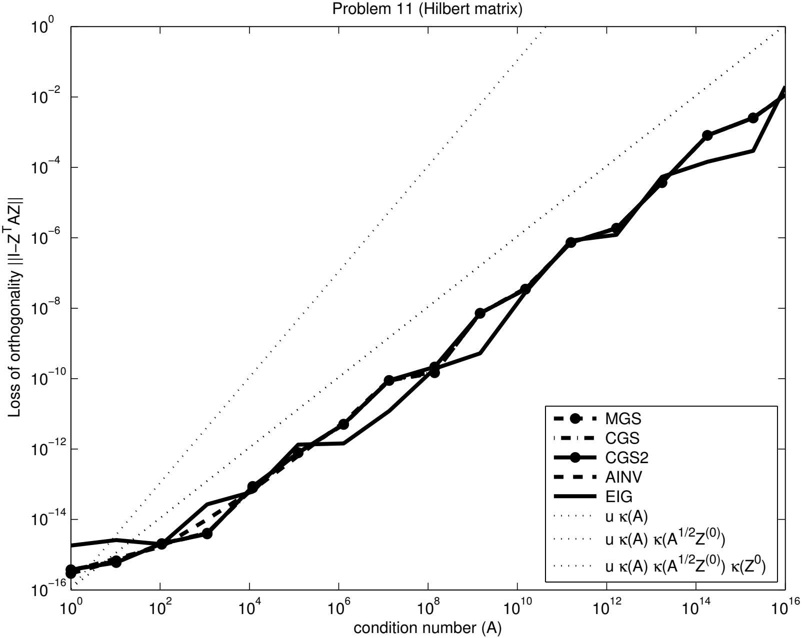
$$\|I - \bar{Z}^T A \bar{Z}\| \leq \mathcal{O}(u)$$

Weighted least squares problem; MGS: [Gulliksson, Wedin 1992, Gulliksson 1995]

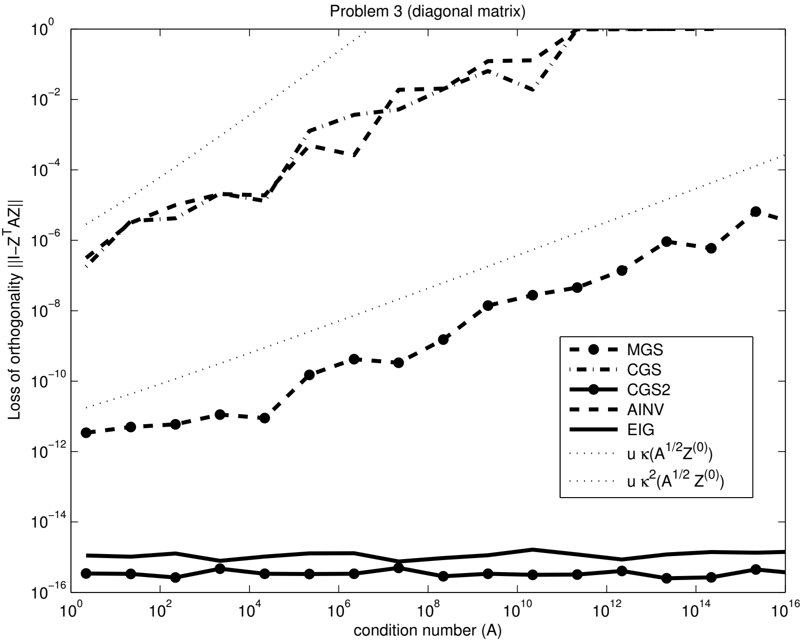
Numerical experiments - extremal cases

1. $\kappa^{1/2}(A) \ll \kappa(A^{1/2}Z^{(0)})$
2. $\kappa(A^{1/2}Z^{(0)}) \ll \kappa^{1/2}(A)$
3. A positive diagonal


Problem 9 (Hilbert matrix), κ(A) = 1.2e5
Loss of orthogonality ||I−Z^T AZ||
condition number (A^{1/2} Z^{(0)})
MGS
CGS
CGS2
AINV
EIG
u κ(A)
u κ(A) κ(A^{1/2}Z^{(0)})
u κ(A) κ(A^{1/2}Z^{(0)}) κ(Z^{(0)})



Problem 11 (Hilbert matrix)
Loss of orthogonality ||I−Z^T AZ||
condition number (A)
MGS
CGS
CGS2
AINV
EIG
u κ(A)
u κ(A) κ(A^{1/2}Z^{(0)})
u κ(A) κ(A^{1/2}Z^{(0)}) κ(Z^0)


Problem 3 (diagonal matrix)
Loss of orthogonality ||I−Z^T AZ||
condition number (A)
MGS
CGS
CGS2
AINV
EIG
u κ(A^{1/2}Z^{(0)})
u κ^2(A^{1/2} Z^{(0)})

Thank you for your attention!!!

Reference: J. Kopal, R, A. Smoktunowicz, and M. Tůma: Rounding error analysis of orthogonalization with a non-standard inner product, submitted 2010.