Gram-Schmidt process with a non-standard inner product and its application to the 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

Workshop on Structured Preconditioning and Iterative Methods with Applications,
Tsinghua Sanya International Mathematics Forum, Sanya, March 24-28, 2014

STANDARD INNER PRODUCT: GRAM-SCHMIDT PROCESS AS QR ORTHOGONALIZATION

$$A = (a_1, \ldots, a_n) \in \mathcal{R}^{m,n}, \quad m \geq rank(A) = n$$

orthogonal basis Q of $span(A)$

$Q = (q_1, \ldots, q_n) \in \mathcal{R}^{m,n}$, $Q^T Q = I$

$A = QR$, R upper triangular

$$A^T A = R^T R$$

finite precision arithmetic:

$$\bar{Q} = (\bar{q}_1, \ldots, \bar{q}_n),\ \bar{Q}^T\bar{Q} \neq I_n,\ \|I - \bar{Q}^T\bar{Q}\| \leq ?$$

- **classical** and **modified** Gram-Schmidt are mathematically equivalent, but they have "**different**" numerical properties
- **classical** Gram-Schmidt can be **"quite unstable",** can **"quickly" lose** all semblance of **orthogonality**
- Gram-Schmidt with **reorthogonalization**: **"two-steps are enough"** to preserve the orthogonality to working accuracy

GRAM-SCHMIDT PROCESS VERSUS ROUNDING ERRORS

- **modified** Gram-Schmidt:
  assuming $\mathcal{O}(u)\kappa(A) < 1$
  $\|I - \bar{Q}^T\bar{Q}\| \leq \frac{\mathcal{O}(u)\kappa(A)}{1-\mathcal{O}(u)\kappa(A)}$

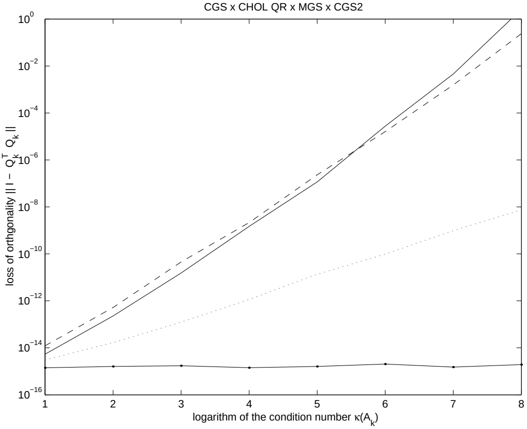
  Björck, 1967, Björck, Paige, 1992

- **classical** Gram-Schmidt:
  assuming $\mathcal{O}(u)\kappa(A) < 1$
  $\|I - \bar{Q}^T\bar{Q}\| \leq \frac{\mathcal{O}(u)\kappa^2(A)}{1-\mathcal{O}(u)\kappa(A)}$

  Giraud, van den Eshof, Langou, R, 2005
  Barlow, Smoktunowicz, Langou, 2006

- classical or modified Gram-Schmidt with **reorthogonalization**:
  assuming $\mathcal{O}(u)\kappa(A) < 1$
  $\|I - \bar{Q}^T\bar{Q}\| \leq \mathcal{O}(u)$

  Giraud, van den Eshof, Langou, R, 2005
  Barlow, Smoktunowicz, 2011



Stewart, "Matrix algorithms" book, p. 284, 1998

ON THE WAY FROM THE STANDARD TO THE NONSTANDARD INNER PRODUCT

- Axel Ruhe. Numerical aspects of. Gram-Schmidt orthogonalization. of vectors. Lin. Alg. and its Appl.,. 52/53 (591-601), 1983.
- T. Ericsson, An analysis of orthogonalization in elliptic norms, to appear.
- M. Gulliksson: Backward error analysis for the constrained and weighted linear least squares problem when using the weighted QR factorization. SIAM J. Matrix Anal. Appl. 16(2), 675-687 (1995)
  M. Gulliksson: On the modified GramSchmidt algorithm for weighted and constrained linear least squares problems. BIT Numer. Math. 35(4), 453-468 (1995)
- S.J. Thomas, R.V.M. Zahar: Efficient orthogonalization in the M-norm. Congr. Numer. 80, 23-32 (1991) 36.
  S.J. Thomas, R.V.M. Zahar, : An analysis of orthogonalization in elliptic norms. Congr. Numer. 86, 193-222 (1992)

ORTHOGONALIZATION WITH A NON-STANDARD INNER PRODUCT

$B \in \mathcal{R}^{m,m}$ symmetric positive definite, inner product $\langle \cdot, \cdot \rangle_B$

$A = [a_1, \ldots, a_n] \in \mathcal{R}^{m,n}$, $m \geq n = rank(A)$

B-orthogonal basis of the range of A:

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

$A = ZU$, $U \in \mathcal{R}^{n,n}$ upper triangular

$$A^T B A = U^T U$$

NON-STANDARD ORTHOGONALIZATION AND STANDARD QR

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

$$\kappa(Z) \ll \kappa^{1/2}(B)$$

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

$A = I$: $Z = \begin{bmatrix} U^{-1} \\ 0 \end{bmatrix} \in \mathcal{R}^{m,n}$ upper triangular

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

Inverse factorization and approximate inverse preconditioning

$$ZZ^T = AU^{-1}U^{-T}A^T = A[A^TBA]^{-1}A^T$$

BZZ^T : orthogonal projector onto $R(BA)$ and orthogonal to $R(A)$

ZZ^TB : orthogonal projector onto $R(A)$ and orthogonal to $R(BA)$

$A = I$ square and nonsingular: inverse factorization $ZZ^T = B^{-1}$

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

$$\downarrow$$

$$\bar{Z}^TB\bar{Z}y = \bar{Z}^Tb\ ,\ \ x = \bar{Z}y,\ \ \|\bar{Z}^TB\bar{Z} - I\| \leq ?$$

finite precision arithmetic:

$$\bar{Z} = (\bar{z}_1, \ldots, \bar{z}_n),\ \bar{Z}^TB\bar{Z} \neq I,\ \|I - \bar{Z}^TB\bar{Z}\| \leq ?$$

$$\bar{U}^T\bar{U} \approx A^TBA,\ \ \|A^TBA - \bar{U}^T\bar{U}\| \leq ?$$

$$\bar{Z}\bar{U} \approx A,\ \ \|A - \bar{Z}\bar{U}\| \leq ?$$

REFERENCE AND GRAM-SCHMIDT IMPLEMENTATIONS

$$B = V \Lambda V^T, \ \Lambda^{1/2} V^T A = QU, \ Z = V \Lambda^{-1/2} Q$$

backward stable eigendecomposition + backward stable QR:

$$\|\bar{Z}^T B \bar{Z} - I\| \leq \mathcal{O}(u) \|B\| \|\bar{Z}\|^2$$

$$z_i^{(0)} = a_i, \quad z_i^{(j)} = z_i^{(j-1)} - u_{ji} z_j, \quad j = 1, \dots, i-1$$

$$z_i = z_i^{(i-1)} / u_{ii}, \quad u_{ii} = \|z_i^{(i-1)}\|_B$$

modified Gram-Schmidt $\equiv$ SAINV: $u_{ji} = \langle z_i^{(j-1)}, z_j \rangle_B$
classical Gram-Schmidt: $u_{ji} = \langle a_i, z_j \rangle_B$
AINV algorithm: $u_{ji} = \langle z_i^{(j-1)}, a_j / u_{jj} \rangle_B$

classical (CGS)

for $i = 1, \ldots, n$
  $z_i^{(0)} = a_i$
  for $j = 1, \ldots, i-1$

$$z_i^{(j)} = z_i^{(j-1)} - \langle a_i, z_j \rangle_B z_j$$

  end
  $z_i = z_i^{(i-1)} / \|z_i^{(i-1)}\|_B$
end

modified (MGS)

for $i = 1, \ldots, n$
  $z_i^{(0)} = a_i$
  for $j = 1, \ldots, i-1$

$$z_i^{(j)} = z_i^{(j-1)} - \langle z_i^{(j-1)}, z_j \rangle_B z_j$$

  end
  $z_i = z_i^{(i-1)} / \|z_i^{(i-1)}\|_B$
end

AINV **algorithm**

for $i = 1, \ldots, n$
  $z_i^{(0)} = a_i$
  for $j = 1, \ldots, i-1$

$$z_i^{(j)} = z_i^{(j-1)} - \langle z_i^{(j-1)}, \frac{a_j}{\|z_j^{(j-1)}\|_B} \rangle$$

  end
  $z_i = z_i^{(i-1)} / \|z_i^{(i-1)}\|_B$
end

GRAM-SCHMIDT ALGORITHMS WITH COMPLETE REORTHOGONALIZATION

classical (CGS2)

for $i = 1, \ldots, n$
  $z_i^{(0)} = a_i$
  for $k = 1, 2$
  for $j = 1, \ldots, i-1$

$$z_i^{(j)} = z_i^{(j-1)} - \langle a_i, z_j \rangle_B z_j$$

  end
  end
  $z_i = z_i^{(i-1)} / \|z_i^{(i-1)}\|_B$
end

modified (MGS2)

for $i = 1, \ldots, n$
  $z_i^{(0)} = a_i$
  for $k = 1, 2$
  for $j = 1, \ldots, i-1$

$$z_i^{(j)} = z_i^{(j-1)} - \langle z_i^{(j-1)}, z_j \rangle_B z_j$$

  end
  end
  $z_i = z_i^{(i-1)} / \|z_i^{(i-1)}\|_B$
end

CLASSICAL GRAM-SCHMIDT PROVIDES A CHOLESKY FACTOR

Exact arithmetic:

$$u_{j,i} = \langle a_i, z_j \rangle_B = \left\langle a_i, \frac{a_j - \sum_{k=1}^{j-1} u_{k,j} z_k}{u_{j,j}} \right\rangle_B$$

$$= \frac{\langle a_i, a_j \rangle_B - \sum_{k=1}^{j-1} u_{k,i} u_{k,j}}{u_{j,j}}$$

$$A^T B A = U^T U$$

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

$$A^T B A = \bar{U}^T \bar{U} - (\bar{U} - \bar{Z}^T B A)^T \bar{U} + A^T B (A - \bar{Z}\bar{U})$$

THE LOSS OF ORTHOGONALITY

$$A^T BA = \bar{U}^T \bar{U} - (\bar{U} - \bar{Z}^T BA)^T \bar{U} + A^T B(A - \bar{Z}\bar{U})$$
$$A^T BA = \bar{U}^T \bar{Z}^T B\bar{Z}\bar{U} + (A - \bar{Z}\bar{U})^T B\bar{Z}\bar{U} + \bar{U}^T \bar{Z}^T B(A - \bar{Z}\bar{U}) + (A - \bar{Z}\bar{U})^T B(A - \bar{Z}\bar{U})$$
$$\bar{U}^T (I - \bar{Z}^T B\bar{Z})\bar{U} = (\bar{U} - \bar{Z}^T BA)^T \bar{U} - (A - \bar{Z}\bar{U})^T B\bar{Z}\bar{U}$$

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

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

GRAM-SCHMIDT PROCESS WITH REORTHOGONALIZATION (B-CGS2)

$$z_j^{(1)} = a_j - Z_{j-1}u_{1:j-1,j}^{(1)} = (I - Z_{j-1}Z_{j-1}^T B)a_j,$$
$$z_j^{(2)} = z_j^{(1)} - Z_{j-1}u_{1:j-1,j}^{(2)} = (I - Z_{j-1}Z_{j-1}^T B)^2 a_j = z_j^{(1)}$$

$$\|\bar{Z}_{j-1}^T B \left(\frac{\bar{z}_j^{(2)}}{\bar{u}_{j,j}} \right) \| \lessapprox \|I - \bar{Z}_{j-1}^T B \bar{Z}_{j-1}\|^2 \| \frac{\bar{u}_{1:j-1,j}}{\bar{u}_{j,j}} \|$$

$$1/u_{j,j} \leq \sigma_{min}^{-1}(U) = \sigma_{min}^{-1}(B^{1/2}A), \; \|u_{1:j-1,j}\|/u_{j,j} \leq \kappa(B^{1/2}A),$$

finite precision arithmetic:

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

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

LOSS OF B-ORTHOGONALITY IN GRAM-SCHMIDT

modified Gram-Schmidt:

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

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

classical Gram-Schmidt and AINV algorithm:

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

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

classical Gram-Schmidt with reorthogonalization:

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

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

THE LOCAL ERRORS IN A NON-STANDARD INNER PRODUCTS

general positive definite B :

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

diagonal positive (weight matrix) B :

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

DIAGONAL CASE IS SIMILAR TO STANDARD CASE

modified Gram-Schmidt:

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

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

classical Gram-Schmidt and AINV algorithm

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

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

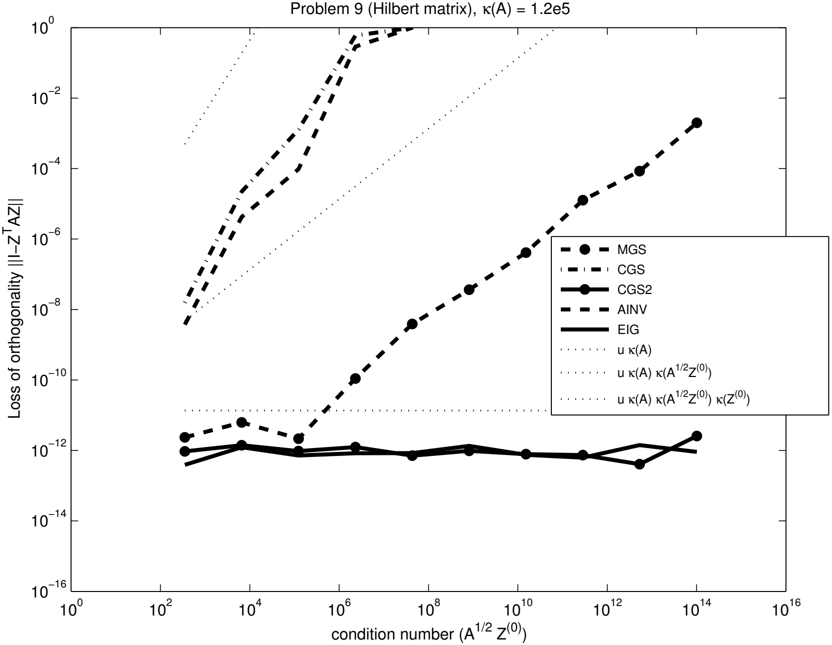
classical Gram-Schmidt with reorthogonalization:

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

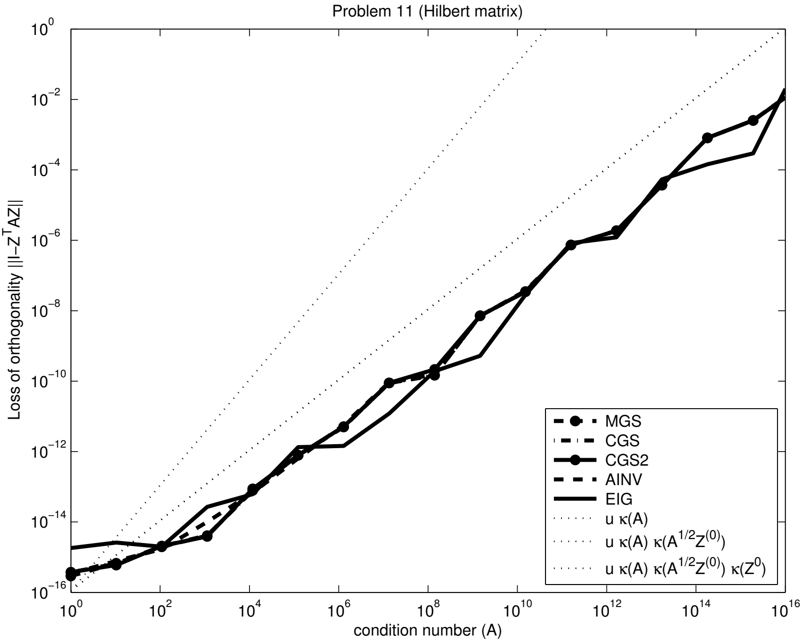
$$\|I - \bar{Z}^T B \bar{Z}\| \leq \mathcal{O}(u)$$

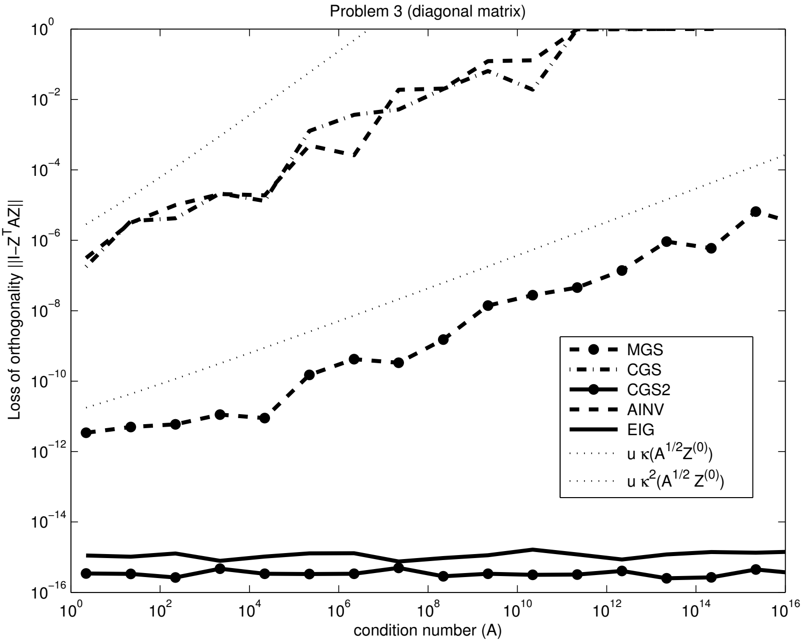
Gulliksson, Wedin 1992, Gulliksson 1995

1. $\kappa^{1/2}(B) \ll \kappa(B^{1/2}A)$
2. $\kappa(B^{1/2}A) \leq \kappa^{1/2}(B)$
3. B 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)})

ON THE WAY FROM THE INNER PRODUCT TO THE BILINEAR FORM

- Symmetric indefinite eigenvalue problems. The bilinear form $\langle x, y\rangle_B = y^T Bx$ can have $\langle x, x\rangle_B < 0$ and $\langle x, x\rangle_B = 0$ for some $x \neq 0$.
- Eigenvectors Q can be chosen such that $Q^T BQ = \Omega$ where $\Omega = \mathrm{diag}(\pm 1)$ is a signature matrix. Isotropic vectors $x^T Bx = 0$.
- Structured eigenvalue problems. The SR factorization. The skew-symmetric bilinear form $\langle x, y\rangle_B = y^T Bx$, where $B^T = -B$. Each vector satisfies $x^T Bx = 0$.

INDEFINITE ORTHOGONALIZATION WITH A SYMMETRIC BILINEAR FORM

$B \in \mathcal{R}^{m,m}$ symmetric indefinite and nonsingular, bilinear form

$$A = (a_1, \ldots, a_n) \in \mathcal{R}^{m,n},\ m \geq n = rank(A)$$

B-orthonormal basis of $span(A)$:

$$Q = (q_1, \ldots, q_n) \in \mathcal{R}^{m,n},\ Q^T BQ = \Omega \in \mathrm{diag}(\pm 1)$$

$A = QR$, $R \in \mathcal{R}^{n,n}$ upper triangular with positive diagonal

if no principal minor of $A^T BA$ vanishes (if $A^T BA$ is strongly nonsingular)

$$A^T BA = R^T \Omega R$$

GRAM-SCHMIDT WITH SKEW-SYMMETRIC BILINEAR FORM: SR FACTORIZATION

$$B = \begin{pmatrix} 0 & I_m \\ -I_m & 0 \end{pmatrix} \in \mathcal{R}^{2m,2m} \text{ skew-symmetric and orthogonal}$$

$$A = [a_1, a_2] \in \mathcal{R}^{2m,2}, \text{ non-isotropic with } a_1^T B a_2 \neq 0$$

$$V = [v_1, v_2] \in \mathcal{R}^{2m,2}, \text{ symplectic } V^B V = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} V^T B V = I_2,$$

$$v_1^T B v_2 = 1$$

$$A = VR,\ R = \begin{pmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{pmatrix} \in \mathcal{R}^{2,2} \text{ upper triangular with } r_{11} r_{22} = a_1^T B a_2$$

Thank you for your attention!!!

References: R, J. Kopal, M. Tuma, A. Smoktunowicz: Numerical stability of orthogonalization methods with a non-standard inner product, BIT Numerical Mathematics (2012) 52:1035-1058.
R, F. Okulicka-Dluzewska, A. Smoktunowicz: Indefinite orthogonalization with rounding errors, submitted 2013.