Research Report

KSTS/RR-15/004
March 9, 2015

Block symplectic Gram-Schmidt method

by

Yoichi Matsuo
Takashi Nodera

Yoichi Matsuo
Graduate School of Science and Technology
Keio University

Takashi Nodera
Department of Mathematics
Keio University

Department of Mathematics
Faculty of Science and Technology
Keio University

KSTS/RR-15/004
March 9, 2015

Block symplectic Gram-Schmidt method

Yoichi Matsuo[1] Takashi Nodera[2]

March 3, 2015

Abstract

For large scale linear problems, the symplectic Lanczos method which uses the Symplectic Gram-Schmidt (SGS) method to compute symplectic vectors, is often used. However, previous studies have shown that the selection process of the parameter in the SGS method is flawed, as it results in a partially destroyed J-orthogonality of the J-orthogonal matrix.

In this paper, we have explored a block type SGS and a new condition for the reorthogonalization to maintain J-orthogonality. Applying the block size scheme to this method, we have developed a new procedure for computing symplectic vectors.

KeyWords. symplectic block Gram-Schmidt, optimal block size, J-orthogonality
AMS(MOS) subject classifications. 65F10, 65M12

Contents

1 Introduction **1**

2 Symplectic Gram-Schmidt Method **2**

3 Block Symplectic Gram-Schmidt Method **4**
3.1 Re-orthogonalization 4
3.2 Optimal Block Size 5

4 Numerical Experiments **6**
4.1 Experiment 1 6
4.2 Experiment 2 7

5 Conclusion **8**

References **8**

1 Introduction

The orthogonalization process or the QR factorization by the Gram-Schmidt (GS) method is arguably one of the most important processes in linear algebraic computation and there are numerous studies on this subject [3, 5, 6, 9, 10].

The GS procedure is also used in the Symplectic methods which are structure-preserving methods for solving eigenvalue problems arising from special matrices like the Hamiltonian matrix. In scientific computation, the eigenvalue problem of the Hamiltonian matrix is such an important topic that many studies have been published on this topic [1, 2, 4]. Applications of this symplectic method are, for instance, used to solve the Ricatti equation arising from the control theory. The symplectic methods enable us to compute eigenvalues faster by using the structure of a matrix unlike the QR or Lanczos method. For a given coefficient matrix A, the SGS method computes the symplectic matrix S and triangular matrix R which satisfy $A = SR$. Since this method preserves the important structure of the given matrix, it enables us to comupute eigenvalues more rapidly. According to Van Loan [4], particularly when comparing the Hamiltonian matrix to the QR algorithm in terms of the number of floating-point operations, this method requires only about 25 % storage.

The SR procedure is very similar to the Householder QR algorithm. Salam [8] have proven that this is an equivalent to the modified symplectic Gram-Schmidt (MSGS) method. Another is the SR factorizaton by the classical symplectic Gram-Schmidt (CSGS) and the proposed MSGS method [7]. However, there are less numerical experiments documented on the SGS compared to the GS decomposition for the QR, Arnoldi and Lanczos methods.

In this paper, we have explored the possibility of using the Block Symplectic Gram-Schmidt (BSGS) method by blocking the CSGS method. The Block Gram-Schmidt (BGS) algorithm is a standard generalization of the classical Gram-Schmidt algorithm. A study by Stewart [10] and Matsuo et al. [5] illustrates how the computation time of the QR factorization can be shortened by employing the BGS method. The CSGS method is blocked into the BSGS method. The BSGS method, then enables the computation of the SR factorization more rapidly. Moreover since the optimal block-size m is not consistent when employing the BGS method, it is necessary to determine m. There is no unique m for any matrix X when the BGS method is being used, and m must be determined accurately, through trial and error.

Section 2 is a summary of the SGS method. In section 3, a BSGS method will be proposed. In section 4 and section 5, numerical experiments are given to evaluate the effectiveness of our proposed algorithm, and our conclusion follows.

2 Symplectic Gram-Schmidt Method

The first step is to define matrix $J \in \mathbb{R}^{2n\times 2n}$ with the following equation:

$$J = \begin{bmatrix} 0_n & I_n \\ -I_n & 0_n \end{bmatrix}, \tag{1}$$

where $J^T = J^{-1} = -J$. Then, the J-product is defined for the vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^{2n}$ by the following equation:

$$< \boldsymbol{x}, \boldsymbol{y} >_J = \boldsymbol{x}^T J \boldsymbol{y}. \tag{2}$$

Let M^J be

$$M^J = J^T M^T J, \tag{3}$$

Algorithm 1: Elementary SR Factorization

Data: $A_1 = [\boldsymbol{a}_1, \boldsymbol{a}_2]$

Result: $S_1 = [\boldsymbol{s}_1, \boldsymbol{s}_2], R_1 = [r_{11}, r_{12}, r_{21}, r_{22}]$

begin

  Choose $r_{11} \in \mathbb{R}$, $\boldsymbol{s}_1 = \boldsymbol{a}_1 / r_{11}$;

  Choose $r_{12} \in \mathbb{R}$, $\boldsymbol{y} = \boldsymbol{a}_2 - r_{12}\boldsymbol{s}_1$;

  $r_{22} = \boldsymbol{s}_1^T J \boldsymbol{y}$;

  $\boldsymbol{s}_2 = \boldsymbol{y} / r_{22}$;

end

and let matrix S be symplectic or J-orthogonal when

$$S^J S = J^T S^T J S = I. \tag{4}$$

Next, the elementary symplectic factorization (ESR) , which J-orthogonalizes vectors $X_1 = [\boldsymbol{x}_1, \boldsymbol{x}_2], \boldsymbol{x}_i \in \mathbb{R}^n, i = 1, 2$ into $S_1 = [\boldsymbol{s}_1, \boldsymbol{s}_2]$ by the following equations is introduced:

$$\begin{cases} \boldsymbol{s}_1 = \dfrac{\boldsymbol{x}_1}{r_{11}} \\ \boldsymbol{y} = \boldsymbol{x}_2 - r_{12}\boldsymbol{s}_1 \\ r_{22} = \boldsymbol{s}_1^T J \boldsymbol{y} \\ \boldsymbol{s}_2 = \dfrac{\boldsymbol{y}}{r_{22}}, \end{cases} \tag{5}$$

where r_{11} and r_{12} are arbitrary real values. Note that there are several selections of r_{11}, r_{12}:

- ESR1: $r_{11} = \|\boldsymbol{x}_1\|, \quad r_{12} = 0$
- ESR2: $r_{11} = \|\boldsymbol{x}_1\|, \quad r_{12} = \boldsymbol{s}_1^T \boldsymbol{x}_2$
- ESR3: $r_{11} = \|\boldsymbol{x}_1^T J \boldsymbol{x}_2\|, \quad r_{12} = 0.$

According to Salam [8], the selections of r_{11} and r_{12} is an influential factor in determining the accuracy of the J-orthogonality of the SR factorization and it is reported that the ESR2 method is the most stable, because s_1 becomes orthogonal when set against s_2.

From equation (5), matrix X_1 satisfies the following relation:

$$X_1 = S_1 R_1, \tag{6}$$

where S_1 is the J-orthogonal matrix and R_1 is the upper triangular matrix. The ESR method is illustrated in Algorithm 1.

The classical symplectic Gram-Schmidt (CSGS) algorithm is very similar to the CGS algorithm. Let X be a matrix with $2n \times 2n$. The CSGS method factorizes X into the J-orthogonal matrix S and upper triangular matrix R through the following equation:

$$H_{12} = S^J X \tag{7}$$

$$Y = X - S H_{12} \tag{8}$$

$$Y \to SR \quad \text{(by ESR)}. \tag{9}$$

From equations (7) and (8), it can be seen that the CSGS method is very similar to the CGS method. However in the CSGS method, the vectors are normalized by the ESR instead of by the norm of vectors. Repeating this results in the following relations:

$$A = SR. \tag{10}$$

The J-orthogonalized vectors $S_1, \ldots, S_n, S_i = [\boldsymbol{s}_{2i-1}, \boldsymbol{s}_{2i}]$ in the CSGS method satisfies the following relations:

$$\boldsymbol{s}_{2i-1}^T J \boldsymbol{s}_{2i} = 1, \quad i = 1, \ldots, n, \tag{11}$$

$$S_i^J S_j = \begin{cases} 0, & i \neq j \\ 1, & i = j. \end{cases} \tag{12}$$

The CSGS method, then creates a series of basis vectors:

$$\dim \operatorname{span}\{S_1, \ldots, S_n\} = 2n. \tag{13}$$

3 Block Symplectic Gram-Schmidt Method

In this section, a block type of the CSGS method to speed-up computation of the SR factorization is explored. First, matrix X in equations (7) and (8) are replaced with $X_{\text{block}} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n]$ to create a block type algorithm of the CSGS method. This results in the following equations:

$$H_{12} = S^J X_{\text{block}}, \tag{14}$$

$$Y = X_{\text{block}} - S H_{12}. \tag{15}$$

By using equations (14) and (15), a J-orthogonal matrix X_{block} can be created against the previous J-orthogonalized matrix S. However these steps alone are not enough for creating a J-orthogonalized matrix because the vector in the Y is not J-orthogonalized against other vectors in Y. This makes it necessary to add one more step to create J-orthogonalization in every vector in Y against each other:

$$Y \to SR \quad (\text{ by CSGS}). \tag{16}$$

By relation (16), Y can be J-orthogonalized completely. The block symplectic Gram-Schmidt (BSGS) algorithm is illustrated in Algorithm 2.

3.1 Re-orthogonalization

According to Stewart [10], employing full re-orthogonalization is enough for maintaining the orthogonality of computed vectors when employing the GS method. The condition for re-orthogonalization is as follows:

$$\|\hat{\boldsymbol{y}}\| \geq \frac{1}{2}\|\boldsymbol{x}\|. \tag{17}$$

If an orthogonalized vector $\hat{\boldsymbol{y}}$ does not satisfy this condition, re-orthogonalization is employed.

Algorithm 2: Block Symplectic Gram-Schmidt Algorithm

Data: X=$[X_{\mathrm{block}_1}, \cdots, X_{\mathrm{block}_n}]$
Result: $S = [S_1, \cdots, S_n], R$

begin
  $X_{\mathrm{block}_1} = S_1 R(1:2, 1:2)$;
  for $i = 2 : n$ **do**
    for $j = 1 : i-1$ **do**
      $H_{i,j} = S_j^J X_{\mathrm{block}_i}$;
    end
    $Y_i = X_{\mathrm{block}_i} - \sum_{j=1}^{j=i-1} S_j H_{i,j}$;
    $R(1:2(i-1), 2i-1:2i) = H_{j,i}$;
    $Y_i = S_i \acute{R}$ by CSGS method;
    $R(2i-1:2i, 2i-1:2i) = \acute{R}$;
  end
end

However when running the SGS method, this condition may fail because the norm of J-orthogonalized vectors tends to increase as the SGS steps proceed. This is because the SGS method is unable to normalize computed vector $\hat{\boldsymbol{y}}$ and to address this issue, the ESR method must be utilized instead. The s_{2i-1} vector especially satisfies the following relation:

$$\|\boldsymbol{s}_{2i-1}\| \geq 1. \tag{18}$$

Even though the computed vector $\hat{\boldsymbol{y}}$ lacks J-orthogonality, $\hat{\boldsymbol{y}}$ satisfies condition (17). This makes it possible to propose a new condition for re-orthogonalization:

$$\|\hat{\boldsymbol{y}}\| \leq \frac{1}{2}\|\boldsymbol{x}\|. \tag{19}$$

Through this condition, the norm of the computed vectors can be controlled, and re-orthogonalization can be employed.

3.2 Optimal Block Size

Through using a blocking procedure, it is possible to compute SR factorization quickly. However, the computation time is dependent on block size m. Moreover since the optimal block-size m is not consistent when employing the BSGS, it is mandatory to determine m. There is no unique m for any matrix X when the BSGS is used, and it is necessary to determine m accurately, through trial and error. The next step is to combine determining optimal block size by estimating computation time from the sample of a one step BSGS and computation cost of the BSGS method. The determination method of block size used is from Matsuo *et al.* [6]. The new method used to estimate optimal block size is illustrated in Algorithm 3.

Algorithm 3: The New Method for Estimating Optimal Block Size

Data: $:X \in \mathbb{R}^{n\times n}$
Result: m
begin
  for $i = 1:5$ **do**
    $m_i := 2^{i-1}$;
    for $j = 0:1$ **do**
      $start := gettimeofday()$;
      Block Symplectic Gram-Schmidt step;
      $end := gettimeofday()$;
      $t_{ij} := end - start$;
    end
    $a = (t_{i0} - t_{i1})/m_i$;
    $b[i] := (1/2)\, n^2 a + t_{i0} - a(h - m)$;
    for *j=5:1* **do**
      $A[ij] = m_i^{j-1}$;
    end
  end
  solve $A\boldsymbol{x} = \boldsymbol{b}$;
  $f(m) := x_1 m^4 + x_2 m^3 + x_3 m^2 + x_4 m + x_5$;
  solve $m := \min_{m\in[0,\frac{1}{2}N]} f(m)$;
end

4 Numerical Experiments

In this section, the BSGS with Algorithm 2 and Algorithm 3, CSGS and modified symplectic Gram-Schmidt (MSGS) method [7] are evaluated. The numerical environment is as follows:

- CPU : Intel(R) Xeon(R) CPU E3-1270 V2 3.50GHz
- Memory : 16GB

4.1 Experiment 1

Firstly, it will be shown how the J-orthogonality of the SGS method is unstable and how our new re-orthogonalization condition is effective in addressing this. These numerical experiments were implemented by MATLAB2013B and the test matrices were Hamiltonian matrices H with sizes 20×20, 40×40, ..., 200×200 respectively with random values. The CGS and the CSGS methods were employed to calculate orthogonality and the J-orthogonality of the computed matrices by the following equations respectively:

$$\|I - Q^T Q\|_2, \quad \|I - S^J S\|_2. \tag{20}$$

Numerical results are shown in Figure 1 and 2. The results suggest that the J-orthogonality of the CSGS method is very unstable. For small size problems, the accuracy of the J-orthogonality is approximately 10^{-10}. However, even for a 200×200 size Hamiltonian

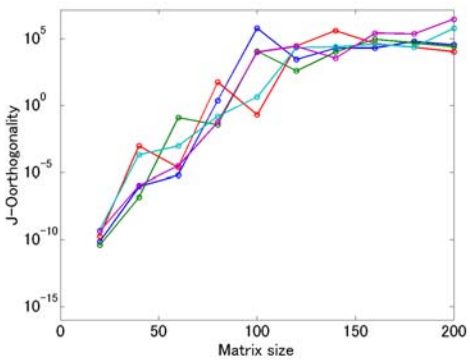

Figure 1: J-Orthogonality of the CSGS Method

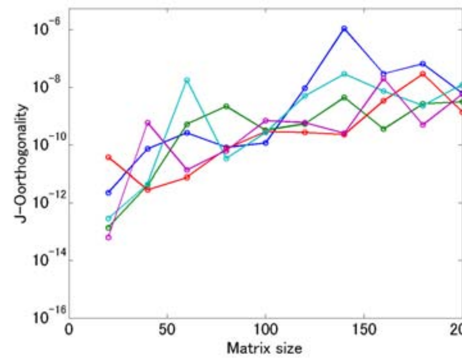

Figure 2: J-Orthogonality of the CSGS with Re-Orthogonalization

Table 1: Ex. 1: CSGS, MSGS, BSGS: Hamiltonian matrix H_1

Method	m	t_m	Accuracy
CSGS	2	0.104	8.70e-06
MSGS	2	0.107	4.65e-06
BSGS	10	0.039	2.80e-06

Table 2: Ex. 2: CSGS, MSGS, BSGS: Hamiltonian matrix H_2

Method	m	t_m	Accuracy
CSGS	2	21.24	1.55e-04
MSGS	2	21.40	1.03e-04
BSGS	40	2.50	3.74e-05
BSGS	100	3.21	7.14e-05
BSGS-m	72	2.72	4.48e-05

matrix, the accuracy of the J-orthogonality was unacceptable. We have not identified the reason for this yet, but it is possible that the calculation errors are caused by the ESR increasing the norm of the computed vector.

Contrary to what is illustrated in Figure 2, a new condition of re-orthogonalization was employed, resulting in a much improved accuracy of the J-orthogonality. It can be suggested that the new orthogonalization condition works for the CSGS method.

4.2 Experiment 2

The effectiveness of the BSGS method is illustrated in this section. These numerical experiments were implemented by C language with double precision and the test matrices used were the Hamiltonian matrix $H_1 \in \mathbb{R}^{200\times 200}$ and $H_2 \in \mathbb{R}^{2000\times 2000}$ with a random value. The SR factorization was employed by the BSGS, CSGS and MSGS method and re-orthogonalization was implemented in each procedure. In this table, Accuracy refers to the calculation accuracy of the J-orthogonality by equation (20) and t refers to computation time. BSGS-m refers to the BSGS method with our proposed method for determining block size.

Numerical experiments are illustrated in Table 1 and Table 2. From Table 1, we can see that the BSGS method is the fastest and has the highest accuracy in terms of J-orthogonality. This is because by blocking X, we can calculate the computation with BLAS. Compared to the CSGS method, the accuracy of the BSGS method is significantly better.

From Table 2, we can also see that the BSGS method is the fastest and has the highest accuracy in terms of J-orthogonality. The BSGS method is approximately ten times faster than the CSGS and the MSGS methods. The accuracy of the BSGS method is significantly better than that of the CSGS and MSGS methods. Since H_2 is larger, the SR factorization of matrix H_2 is more unstable than that of matrix H_1. The BSGS-m is not the fastest in Table 2, but the BSGS-m performed only 10 % slower, more or less, than the fastest method.

5 Conclusion

The BSGS method proposed in this paper blocks the CSGS method to speed-up computation, and combines this with determining the optimal block size. This is necessary, because the computation time of the BSGS method changes significantly depending on block size.

In section 4, numerical experiments were shown, as well as the effectiveness of the re-orthogonalization condition of the BSGS method. It was clear that the new condition worked for the SGS method. Our proposed method is much faster and more accurate than either the CSGS and the MSGS methods. And in terms of determining block size this method selects block size automatically.

In future studies, it will be useful to analyze J-orthogonality and study how this proposed method works when dealing with a large scale problem.

References

[1] Ammar, G. and Benner, P.: On Hamiltonian and Symplectic Hessenberg Forms, *Linear Algebra Appl.*, Vol. 149, No. 1, pp. 55–72 (1991). http://dx.doi.org/10.1016/0024-3795(91)90325-Q

[2] Benner, P. and Fassbender, H.: An Implicitly Restarted Symplectic Lanczos Method for the Hamiltonian Eigenvalue Problem, *Linear Algebra Appl.*, Vol. 263, pp. 75–111 (1997). http://dx.doi.org/10.1016/S0024-3795(96)00524-1

[3] Gudula, R. and Michael, S.: Comparison of Different Parallel Modified Gram-Schmidt Algorithms, *Euro-Par 2005, LNCS*, Vol. 3648, pp. 826–836 (2005). http://dx.doi.org/10.1007/11549468_90

[4] Loan, C. V.: A Symplectic Method for Approximating All the Eigenvalues of a Hamiltonian Matrix, *Linear Algebra Appl.*, Vol. 61, pp. 223–251 (1984). http://dx.doi.org/10.1016/0024-3795(84)90034-X

[5] Matsuo, Y. and Nodera, T.: The Optimal Block-Size for the Block Gram-Schmidt Orthogonalization, *J. Sci. Tech.*, Vol. 49, pp. 569–584 (2011). http://vjs.ac.vn/index.php/jst

[6] Matsuo, Y. and Nodera, T.: An Efficient Implementation of the Block Gram–Schmidt Method, *Anziam J.*, Vol. 54, pp. C394–409 (2013). http://dx.doi.org/10.0000/anziamj.v54i0.6327

[7] Salam, A.: On Theoretical and Numerical Aspects of Symplectic Gram-Schmidt-like Algorithms, *Numer. Algo.*, Vol. 39, pp. 437–462 (2005). http://dx.doi.org/10.1007/s11075-005-0963-2

[8] Salam, A. and Al-Aidarous, E.: Equivalence Between Modified Symplectic Gram-Schmidt and Householder SR Algorithms, *BIT Numer. Math.*, Vol. 54, pp. 283–302 (2014). http://dx.doi.org/10.1007/s10543-013-0441-5

[9] Steven, J. L., Ake, B. and Walter, G.: Gram-Schmidt Orthogonalization: 100 years and more, *Numer. Linear Algebra Appl.*, Vol. 20, pp. 492–532 (2013). http://dx.doi.org/10.1002/nla.1839

[10] Stewart, G. W.: Block Gram-Schmidt Orthogonalization, *SIAM J. Sci. Comput.*, Vol. 31, pp. 761–775 (2008). http://dx.doi.org/10.1137/070682563

Author addresses

1. **Yoichi Matsuo**, School of Fundamental Science and Technology, Graduate School of Science and Technology, Keio University, 3-14-1 Hiyoshi, Kohoku, Yokohama, Kanagawa, 223-8522, Japan.
   mailto:matsuo@math.keio.ac.jp

2. **Takashi Nodera**, Department of Mathematics, Faculty of Science and Technology, Keio University, 3-14-1 Hiyoshi, Kohoku, Yokohama, Kanagawa, 223-8522, Japan.
   mailto:nodera@math.keio.ac.jp

Department of Mathematics
Faculty of Science and Technology
Keio University

Research Report

2014

[14/001] A. Larraín-Hubach, Y. Maeda, S. Rosenberg, F. Torres-Ardila,
Equivariant, strong and leading order characteristic classes associated to fibrations,
KSTS/RR-14/001, January 6, 2014

[14/002] Ryoichi Suzuki,
A Clark-Ocone type formula under change of measure for canonical Lévy processes,
KSTS/RR-14/002, March 12, 2014

2015

[15/001] Shiro Ishikawa,
Linguistic interpretation of quantum mechanics: Quantum Language,
KSTS/RR-15/001, January 22, 2015

[15/002] Takuji Arai, Ryoichi Suzuki
Local risk-minimization for Lévy markets,
KSTS/RR-15/002, February 2, 2015

[15/003] Kazuma Teramoto, Takashi Nodera
Lanczos type method for computing PageRank,
KSTS/RR-15/003, March 9, 2015

[15/004] Yoichi Matsuo, Takashi Nodera
Block symplectic Gram-Schmidt method,
KSTS/RR-15/004, March 9, 2015

[15/005] Yuto Yokota, Takashi Nodera
The L-BFGS method for nonlinear GMRES acceleration,
KSTS/RR-15/005, March 9, 2015