General solution of full row rank linear systems of equations using a new compression ABS model

Leila Asadbeigi¹ · Mahmoud Paripour² · Esmail Babolian³ · Shahn timer Javadi³

Received: 17 January 2017 / Accepted: 15 September 2017 / Published online: 4 October 2017
© The Author(s) 2017. This article is an open access publication

Abstract This paper describes two new ABS algorithms based on two-step ABS methods for solving general solution of full row rank linear systems of equations. For both of our works, the i th iteration solves the first $2i$ equations, but for the second algorithm, we compress the space. We investigate the numerical stability of our models and a class of methods having optimal stability is defined. The condition for the residual perturbation to be minimal is also given. Computational complexity and numerical results demonstrate that, for our new methods, we need less work than corresponding two-step ABS algorithms and Huang's method. The number of multiplications for these new schemes is half of the storage needed by the Gaussian elimination method. Our new version ABS algorithms are computationally better than the classical Gaussian elimination method, having the same arithmetic cost, but using less memory and no pivoting is necessary. The compression ABS model economizes the space more than the first

algorithm, via deleting the zero rows of the Abaffian matrix by two in every iteration.

Keywords ABS methods · Two-step ABS algorithms · New compression ABS model · Numerical stability · Optimally stable method

Introduction

ABS methods have been used broadly for solving linear and nonlinear systems of equations comprising large number of constraints and variables. In addition, ABS methods provide a unification of the field of finitely terminating methods for the solution of linear systems of equations. ABS methods were introduced by Abaffy, Broyden, and Spedicato initially for solving a determined or underdetermined linear system and later extended for linear least squares, nonlinear equations, optimization problems, Diophantine equations, and fuzzy linear systems [1, 11, 13]. These extended ABS algorithms offer some new approaches that are better than classical ones under several respects. In addition, extensive computational experience has shown that now, since two new equations are hat ABS methods are implementable in a stable way, being often more accurate than the corresponding initial algorithm. ABS methods can be more effective than some of the other traditional methods. See more about ABS algorithms in [15–18]. A review of ABS algorithms is observed in [19]. The basic ABS algorithm works on a system of the form:

$$Ax = b, \quad (1)$$

where $A = [a_1, \dots, a_m]^T$, $a_i \in \mathbb{R}^n$, $1 \leq i \leq m$, $x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $m \leq n$.

✉ Mahmoud Paripour
paripour@hut.ac.ir

Leila Asadbeigi
lbeigi14@yahoo.com

Esmail Babolian
babolian@khu.ac.ir

Shahn timer Javadi
javadi@khu.ac.ir

¹ Department of Mathematics, College of Basic Sciences, Hamedan Branch, Islamic Azad University, Hamedan, Iran

² Department of Computer Engineering and Information Technology, Hamedan University of Technology, Hamedan 65165-13418, Iran

³ Faculty of Mathematical Sciences and Computer, Kharazmi University, 50 Taleghani Avenue, Tehran 1561836314, Iran

There is another notation for (1) as the following form:

$$A^T x = b, \quad (2)$$

where $A^T = [a_1, \dots, a_m]^T$, $a_i \in \mathbb{R}^n$, $1 \leq i \leq m$, $x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $m \leq n$.

Notice that systems (1) and (2) are equivalent. Matrix computations are presented in [14]. The basic ABS methods determine solution of the above systems or signify lack of its existence in at most m iterates. Amini et al. proposed two-step ABS algorithms for the general solution of full row rank linear systems of equations [5, 6]. The purpose of our paper is to present two new extended two-step ABS models and study on analysis of error propagation for them. The structure of this paper is organized as follows: Sect. 2 is devoted to the construction of a new two-step ABS model for solving general solution of full row rank linear systems of equations. Rank reducing process is done in two phases, for per iterate. The first phase helps us to have a solution for the i th iteration and the next phase leads to compute general solution of that iteration. In addition, we state and prove related theorems. We present our first two-step ABS algorithm for the general solution of full row rank linear systems of equations in this section. In Sect. 3, to compress required space, we present our second algorithm, such that the number of rows of the Abaffian matrix is reduced by two for per iterate. Thus, considering matrix multiplication, related parameters are selected from proper dimensions. In Sect. 4, we investigate the stability by the backward error analysis techniques due to Broyden's backward error analysis and Galantai's method. The errors in the steps are determined in the terms of projectors constructable from the conjugate directions. A class of methods having optimal stability characteristics is defined. Computational complexity and numerical results are discussed in Sect. 5, in detail. In Sect. 6, we summarize our achievements.

A new class of two-step ABS model

The basic ABS algorithm starts with an initial vector $x_0 \in \mathbb{R}^n$ and a nonsingular matrix $H_0 \in \mathbb{R}^{n \times n}$ (Spedicato's parameter). Given that x_i is a solution of the first i equations, the ABS algorithm computes x_i as the solution of the first $i + 1$ equations as the following steps [1]:

1. Determine z_i (Broyden's parameter), so that $z_i^T H_i a_i \neq 0$ and set $p_i = H_i^T z_i$.
2. Update the solution by $x_{i+1} = x_i + \alpha_i p_i$, where the stepsize α_i is given by $\alpha_i = \frac{b_i - a_i^T x_i}{a_i^T p_i}$.

3. Update the Abaffian matrix H_i by $H_{i+1} = H_i - \frac{H_i a_i w_i^T H_i}{w_i^T H_i a_i}$, where $w_i \in \mathbb{R}^n$ (Abaffy's parameter) is chosen, so that $w_i^T H_i a_i \neq 0$.

Here, we are motivated to study on a method that satisfies two new equations at a time. We consider the system (1) under the assumption that A is full rank in row, i.e., $\text{rank}(A) = m$ and $m \leq n$. Suppose that $m = 2l$ (if m is odd, we can add a trivial equation to the system). Take $A^{2i} = [a_1, \dots, a_{2i}]^T$, $b^{2i} = [b_1, \dots, b_{2i}]^T$ and $r_j(x) = a_j^T x - b_j$ ($j = 1, \dots, m$). Assume that we are at the i th iteration and x_i satisfies $A^{2i} x = b^{2i}$. We determine $H_i \in \mathbb{R}^{n \times n}$, $z_i \in \mathbb{R}^n$ and $\lambda_i \in \mathbb{R}$, so that $x_i = x_{i-1} - \lambda_i H_i^T z_i$ is a solution of the first $2i$ equations of the system (1), which is $A^{2i} x_i = b^{2i}$. As a result, we have $r_j(x_i) = 0$, $j = 1, \dots, 2i$. Thus, for $j = 2i - 1$ and $j = 2i$, we have

$$\begin{cases} a_{2i-1}^T (x_{i-1} - \lambda_i H_i^T z_i) - b_{2i-1} = 0, \\ a_{2i}^T (x_{i-1} - \lambda_i H_i^T z_i) - b_{2i} = 0, \end{cases}$$

or equivalently

$$\begin{cases} \lambda_i (H_i a_{2i-1})^T z_i = r_{2i-1}(x_{i-1}), \\ \lambda_i (H_i a_{2i})^T z_i = r_{2i}(x_{i-1}). \end{cases}$$

Suppose that $r_{2i-1}(x_{i-1}) \neq 0$ and $r_{2i}(x_{i-1}) \neq 0$. Then, λ_i must be nonzero and the above systems are compatible if and only if we take

$$\lambda_i = \frac{\bar{r}_{2i-1}(x_{i-1})}{(H_i a_{2i-1})^T z_i} = \frac{\bar{r}_{2i}(x_{i-1})}{(H_i a_{2i})^T z_i}, \quad (3)$$

where $\bar{r}_{2i-1}(x_{i-1}) = \bar{r}_{2i}(x_{i-1}) = r_{2i-1}(x_{i-1})r_{2i}(x_{i-1})$. There are various ways to satisfy (3). We consider the following model:

1. Choose an appropriate update for H_i , so that $H_i a_{2i-1} = H_i a_{2i} \neq 0$,
2. Select a vector z_i , so that $z_i^T H_i a_{2i} \neq 0$.

Now, since two new equations are considered in each iterate, we use a rank one update as (4) and another rank one update as (5). Therefore, we have a new rank two update for each iterate. Here, we present H_i and H_{i+1} satisfying the following properties:

$$H_i a_{2i-1} = H_i a_{2i} \neq 0, \quad i = 1, \dots, l, \quad (4)$$

and

$$H_{i+1} a_j = 0, \quad j = 1, \dots, 2i. \quad (5)$$

Now, assume

$$c_j = \begin{cases} a_{2i} - a_j, & j \neq 2i, \\ a_{2i}, & j = 2i. \end{cases} \quad (6)$$

As relations (4)–(6), we will construct (7) and (8), such that

$$H_i c_j = 0, \quad j = 1, \dots, 2i - 1, \quad (7)$$

and

$$H_i c_j = 0, \quad j = 1, \dots, 2i. \quad (8)$$

We compute H_{i+1} from H_i , such that the relations (7) and (8) hold and proceed inductively. We define $H_{i+1} = H_i + g_{2i+1} d_{2i+1}^T$ where $g_{2i+1}, d_{2i+1} \in \mathbb{R}^n$. We need to have

$$\begin{cases} H_{i+1} c_j = 0, & j = 1, \dots, 2i + 1, \\ H_i c_j = 0, & j = 1, \dots, 2i. \end{cases}$$

Therefore, we have

$$\begin{cases} (H_i + g_{2i+1} d_{2i+1}^T) c_j = 0, & j = 1, \dots, 2i + 1, \\ H_i c_j = 0, & j = 1, \dots, 2i. \end{cases}$$

Thus, we must define $g_{2i+1}, d_{2i+1} \in \mathbb{R}^n$ in such a way that

$$H_i c_j + (d_{2i+1}^T c_j) g_{2i+1} = 0, \quad j = 1, \dots, 2i + 1, \quad (9)$$

and $H_i c_j = 0, j = 1, \dots, 2i$. The condition (9) is satisfied for $j \leq 2i - 1$ by the induction hypothesis. By taking $j = 2i + 1$ in (9), we get

$$(d_{2i+1}^T c_{2i+1}) g_{2i+1} = -H_i c_{2i+1}. \quad (10)$$

We consider the choice $g_{2i+1} = -H_i c_{2i+1}$ with $d_{2i+1}^T c_{2i+1} = 1$, which clearly holds in (10). Now, we define $d_{2i+1} = H_i^T w_{2i+1}$ for $w_{2i+1} \in \mathbb{R}^n$, such that

$$w_{2i+1}^T H_i c_{2i+1} = 1. \quad (11)$$

Later, as Theorem 2.3, we will conclude that the above system has solution and H_i is well defined. Therefore, the updating formula for H_i is given by the following:

$$H_{i+1} = H_i - H_i c_{2i+1} w_{2i+1}^T H_i, \quad (12)$$

where w_{2i+1} can be any vector satisfying (11). Now, to satisfy (12) and complete the induction, H_1 should be chosen, so that $H_1 a_1 = H_1 a_2 \neq 0$ or

$$H_1 c_1 = 0. \quad (13)$$

Let H_0 be an arbitrary nonsingular matrix. We obtain H_1 from H_0 using a rank one update. Take $H_1 = H_0 - u_1 v_1^T$, where $u_1, v_1 \in \mathbb{R}^n$ are chosen, so that (13) is satisfied. Therefore, we have $H_0 c_1 - (v_1^T c_1) u_1 = 0$. The previous equation is satisfied if we set $u_1 = H_0 c_1, v_1 = H_0^T w_1$, and we choose $w_1 \in \mathbb{R}^n$ satisfying the next condition:

$$w_1^T H_0 c_1 = 1. \quad (14)$$

Clearly, (14) can be held with a proper choice of $w_1 \in \mathbb{R}^n$, whenever a_1 and a_2 are linearly independent. Thus, we have a rank one update as follows:

$$H_1 = H_0 - H_0 c_1 w_1^T H_0, \quad (15)$$

where w_1 is an arbitrary vector satisfying (14). To compute the general solution and update the second phase for the i th iteration, we introduce a matrix H_i with properties $H_i c_j = H_i a_j = 0, j = 1, \dots, 2i$. Therefore, we define the matrix H_i by a rank one update as the next formula:

$$H_i = H_i - H_i a_{2i} w_{2i}^T H_i = H_i - H_i c_{2i} w_{2i}^T H_i, \quad i = 1, \dots, l. \quad (16)$$

Notice that $w_{2i} \in \mathbb{R}^n$ is an arbitrary vector satisfying the following condition:

$$w_{2i}^T H_i a_{2i} = 1. \quad (w_{2i}^T H_i c_{2i} = 1). \quad (17)$$

Hence, the general solution of the i th iteration is given by $x_{l_i} = x_i - H_i^T s$, where $s \in \mathbb{R}^n$ is arbitrary. Clearly, the general solution for the last iteration is presented by the following:

$$x_{l_i} = x_l - H_l^T s, \quad (18)$$

where $s \in \mathbb{R}^n$ is arbitrary.

Lemma 2.1 *The vectors $a_1, \dots, a_m$ are linearly independent if and only if the vectors $c_1, \dots, c_m$ are linearly independent.*

Therefore, we proved the following theorem.

Theorem 2.2 *Assume that we have $m = 2l$ arbitrary linearly independent vectors $a_1, \dots, a_m \in \mathbb{R}^n$ and an arbitrary nonsingular matrix $H_0 \in \mathbb{R}^{n \times n}$. Let H_1 be generated by (15) and w_1 is satisfying (14), and the sequence of matrices $H_i, i = 2, \dots, l$, be generated by the following:*

$$H_i = H_{i-1} - H_{i-1} c_{2i-1} w_{2i-1}^T H_{i-1}, \quad (19)$$

with $w_{2i-1} \in \mathbb{R}^n$ satisfying the following condition:

$$w_{2i-1}^T H_{i-1} c_{2i-1} = 1. \quad (20)$$

In addition, let the sequence of matrices $H_1, \dots, H_l$ be generated by (16) with $w_{2i} \in \mathbb{R}^n$ satisfying (17). Then, when we are at the i th iteration, the following properties (1)–(4) hold for $i = 1, \dots, l$.

- (1) $H_i a_{2i-1} = H_i a_{2i} \neq 0$.
- (2) $H_i a_j = 0, j = 1, \dots, 2(i-1)$.
- (3) $H_i a_j = 0, j = 1, \dots, 2i$.
- (4) $H_i c_j = 0, j = 1, \dots, 2i$.

we are at the i th iteration.

Theorem 2.3 *Assume that $a_1, \dots, a_m$ are linearly independent vectors in $\mathbb{R}^n$. Let $H_0 \in \mathbb{R}^{n \times n}$ be an arbitrary nonsingular matrix and H_1 be defined by (15), with $w_1 \in \mathbb{R}^n$ satisfying (14), and for $i = 2, \dots, l$, the sequence of*

matrices H_i be generated by (19) with $w_{2i-1} \in \mathbb{R}^n$ satisfying (20). Then, for all i , $1 \leq i \leq l$, and j , $2i \leq j \leq m$, the vectors $H_i a_j$ are nonzero and linearly independent.

Proof We proceed by induction. For $i = 1$, the theorem is true. Since if $\sum_{j=2}^m \alpha_j H_1 a_j = 0$, we have

$$\sum_{j=2}^m \alpha_j (H_0 - H_0 c_1 w_1^T H_0) a_j = 0,$$

$$\sum_{j=2}^m \alpha_j H_0 a_j - \left(\sum_{j=2}^m \alpha_j w_1^T H_0 a_j \right) H_0 c_1 = 0,$$

or

$$\beta_1 H_0 a_1 - (\beta_1 - \alpha_2) H_0 a_2 + \sum_{j=3}^m \alpha_j H_0 a_j = 0,$$

where $\beta_1 = \sum_{j=2}^m \alpha_j w_1^T H_0 a_j$. Now, since $a_1, \dots, a_m$ are nonzero and linearly independent and H_0 is nonsingular, $H_0 a_j$ for $1 \leq j \leq m$ are nonzero and linearly independent.

Hence, $\beta_1 = \alpha_2 = \alpha_3 = \dots = \alpha_m = 0$. Therefore, the vectors $H_1 a_j$, for $2 \leq j \leq m$, are nonzero and linearly independent. Now, we assume that the theorem is true up to $1 \leq i \leq l-1$, and then, we prove it for $i+1$. From (12), for $2i+2 \leq j \leq m$, we have

$$H_{i+1} a_j = H_i a_j - (w_{2i+1}^T H_i a_j) H_i c_{2i+1}. \quad (21)$$

We need to show that the relation

$$\sum_{j=2i+2}^m \alpha_j H_{i+1} a_j = 0, \quad (22)$$

implies that $\alpha_j = 0$, for $2i+2 \leq j \leq m$. Using (21), we can write (22) as follows:

$$\sum_{j=2i+2}^m \alpha_j H_i a_j - \beta'_1 H_i c_{2i+1} = 0,$$

where $\beta'_1 = \sum_{j=2i+2}^m \alpha_j w_{2i+1}^T H_i a_j$. Thus, we have $\sum_{j=2i+2}^m \alpha_j H_i a_j - \beta'_1 H_i (a_{2i+2} - a_{2i+1}) = 0$. As a result

$$\sum_{j=2i+3}^m \alpha_j H_i a_j + \beta'_1 H_i a_{2i+1} - (\beta'_1 - \alpha_{2i+2}) H_i a_{2i+2} = 0. \quad (23)$$

By the induction hypothesis, the vectors $H_i a_j$, for $2i \leq j \leq m$, are nonzero and linearly independent. Now, as the assumption of the induction, we are going to prove that the vectors $H_i a_j$ are nonzero and linearly independent for $2i+1 \leq j \leq m$. Using (16), we have

$$H_i a_j = H_i a_j - H_i a_{2i} w_{2i}^T H_i a_j. \quad (24)$$

We must prove the relation

$$\sum_{j=2i+1}^m \alpha'_j H_i a_j = 0, \quad (25)$$

implies that $\alpha'_j = 0$, for $j \geq 2i+1$. Using (24), we can write (25) as follows:

$$\sum_{j=2i+1}^m \alpha'_j H_i a_j - \left(\sum_{j=2i+1}^m \alpha'_j w_{2i}^T H_i a_j \right) H_i a_{2i} = 0.$$

As the linearly independence of $H_i a_j$, for $2i \leq j \leq m$, we conclude that $\alpha'_j = 0$, for $j \geq 2i+1$. Consequently, for relation (23), we have $\beta'_1 = \alpha_{2i+2} = \alpha_{2i+3} = \dots = \alpha_m = 0$. Hence, the vectors $H_{i+1} a_j$, for $2i+2 \leq j \leq m$ are nonzero and linearly independent. $\square$

Corollary 2.4 Considering the assumptions of Theorem 2.3, the following statement are true.

- (i) When we are at the i th iteration, we have $H_i a_{2i-1} = H_i a_{2i} \neq 0$. In addition, there exists $z_i \in \mathbb{R}^n$ and $w_{2i} \in \mathbb{R}^n$, such that $z_i^T H_i a_{2i} \neq 0$ and $w_{2i}^T H_i a_{2i} \neq 0$.
- (ii) Each of the systems (14) and (20) has solution.
- (iii) H_i , H_l , x_i , and x_l are well defined for $i = 1, \dots, l$.
- (iv) By taking $H_0 = H_{l_0}$ and considering the sequence of matrices $H_{l_0}, \dots, H_{l_{l-1}}$, $0 \leq i \leq l-1$, we have $H_l a_j \neq 0$, $j > i$. Thus, by taking $j = 2i+1$, there exist $z'_j \in \mathbb{R}^n$, such that $z'^T_j H_l a_j \neq 0$. (It is noted again that when we choose $j = 2i$, we are at the i th step. In fact for the i th iteration, we have two choices $j = 2i-1$ and $j = 2i$.)

Theorem 2.5 For the matrices H_i generated by (15) and (19) and the matrices H_l given by (16), we have

$$\begin{aligned} \dim R(H_i) &= n - 2i + 1, \quad 1 \leq i \leq l, \\ \dim N(H_i) &= 2i - 1, \quad 1 \leq i \leq l, \\ \dim R(H_l) &= n - 2i, \quad 1 \leq i \leq l, \\ \dim N(H_l) &= 2i, \quad 1 \leq i \leq l, \\ \dim R(H_l) &= n - m, \\ \dim N(H_l) &= m. \end{aligned}$$

The notations R and N stand for the range and nullspace, respectively.

Theorem 2.6 Consider the matrix $P^i = (\tilde{p}_1, \dots, \tilde{p}_i)$, $i \leq \frac{m}{2}$, where $\tilde{p}_i = (p'_i, p_i)$ and $p_i = H_i^T z_i$ is obtained from the first-phase rank reducing process of the i th iteration. The sequence of $p'_i = H_{l_{i-1}}^T z'_i$ is obtained from $H_0^T z'_1, \dots, H_{l_{i-1}}^T z'_i$. In addition, take $A^i = (a_1, \dots, a_{2i})$, $i \leq m$.

Then, the matrix L^i defined by $L^i = (A^i)^T P^i$ is lower triangular and nonsingular.

Proof As $(L^i)_{j,k} = a_j^T \tilde{p}_k$, Theorem 2.2 implies that $(L^i)_{j,k}$ is zero for $j < k$, and thus, L^i is lower triangular. As Theorem 2.3 and Corollary 2.4, we have $a_j^T \tilde{p}_k \neq 0$, for $j = k$. Consequently L^i is nonsingular. $\square$

Corollary 2.7 If $m = n$, the matrix A can be factorized in the form $A = LS$, with $L = L^n$ and $S = (P^n)^{-1}$.

Remark 2.8 For the matrices H_i generated by (15) and (19), we have $x_i = x_{i-1} - \lambda_i H_i^T z_i$, ($i = 1, \dots, l$), where x_i is a solution of the first $2i$ equations of the system and λ_i will be discussed in Algorithms 1 and 2. The matrix H_0 can be any arbitrary matrix, but to reduce computational complexity, and for stability reasons, we choose H_0 unitary matrix I_n . Numerical stability and computational complexity will be discussed in Sects. 4 and 5, respectively.

Algorithm 1 Assume that $A_{m \times n} = [a_1, a_2, \dots, a_m]^T$ is a full row rank linear matrix, where $m = 2l$ and $Ax = b$ is a compatible full row rank linear system, $x \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$.

- Let $x_0 \in \mathbb{R}^n$ be an arbitrary vector and choose $H_0 = I_n \in \mathbb{R}^{n \times n}$ (unitary matrix). Set $i = 1$.
- Compute $r_1(x_0) = a_1^T x_0 - b_1$ and $r_2(x_0) = a_2^T x_0 - b_2$.
 - If $r_1(x_0)r_2(x_0) \neq 0$, we let

$$\begin{cases} a_1 = r_2(x_0)a_1, & a_2 = r_1(x_0)a_2, \\ b_1 = r_2(x_0)b_1, & b_2 = r_1(x_0)b_2. \end{cases}$$
 - If $r_1(x_0)r_2(x_0) = 0$ and one of the residual values is nonzero, without loss of generality, we assume that $r_2(x_0) \neq 0$ and we let

$$\begin{cases} a_1 = a_2 + a_1, & a_2 = a_2, \\ b_1 = b_2 + b_1, & b_2 = b_2. \end{cases}$$
 - If $r_1(x_0)r_2(x_0) = 0$, and both of the residual values are zero, x_1 will be x_0 and go to (5).
- Take $c_1 = a_2 - a_1$.
 - Select $w_1 \in \mathbb{R}^n$, such that $w_1^T H_0 c_1 = 1$ and compute $H_1 = H_0 - H_0 c_1 w_1^T H_0$.
 - Select $w_2 \in \mathbb{R}^n$, such that $w_2^T H_1 a_2 = 1$ and compute $H_l = H_1 - H_1 a_2 w_2^T H_1$.
 - Select $z_1 \in \mathbb{R}^n$, so that $z_1^T H_1 a_2 \neq 0$ and compute $\lambda_1 = \frac{r_1(x_0)r_2(x_0)}{z_1^T H_1 a_2}$, if $r_1(x_0)r_2(x_0) \neq 0$, $\lambda_1 = \frac{r_2(x_0)}{z_1^T H_1 a_2}$, if $r_1(x_0) = 0$ and $r_2(x_0) \neq 0$.
- Take $x_1 = x_0 - \lambda_1 H_1^T z_1$.
- Set $i = 2$ and go to 6.

- While $i \leq \frac{m}{2}$, do steps 6(a)–8(b).
 - Compute $r_{2i-1}(x_{i-1}) = a_{2i-1}^T x_{i-1} - b_{2i-1}$ and $r_{2i}(x_{i-1}) = a_{2i}^T x_{i-1} - b_{2i}$.
 - If $r_{2i-1}(x_{i-1})r_{2i}(x_{i-1}) \neq 0$, then we let

$$\begin{cases} a_{2i-1} = r_{2i}(x_{i-1})a_{2i-1}, & a_{2i} = r_{2i-1}(x_{i-1})a_{2i}, \\ b_{2i-1} = r_{2i}(x_{i-1})b_{2i-1}, & b_{2i} = r_{2i-1}(x_{i-1})b_{2i}. \end{cases}$$
 - If $r_{2i-1}(x_{i-1})r_{2i}(x_{i-1}) = 0$ and one of the residual values is nonzero, without loss of generality, we assume that $r_{2i}(x_{i-1}) \neq 0$ and we let

$$\begin{cases} a_{2i-1} = a_{2i} + a_{2i-1}, & a_{2i} = a_{2i}, \\ b_{2i-1} = b_{2i} + b_{2i-1}, & b_{2i} = b_{2i}. \end{cases}$$
 - If $r_{2i-1}(x_{i-1})r_{2i}(x_{i-1}) = 0$ and both the residual values are zero, x_i will be x_{i-1} and go to 8(b).
- Take $c_{2i-1} = a_{2i} - a_{2i-1}$.
 - Select $w_{2i-1} \in \mathbb{R}^n$, such that $w_{2i-1}^T H_{l-1} c_{2i-1} = 1$, and compute $H_i = H_{l-1} - H_{l-1} c_{2i-1} w_{2i-1}^T H_{l-1}$.
 - Select $w_{2i} \in \mathbb{R}^n$, such that $w_{2i}^T H_i a_{2i} = 1$ and compute $H_l = H_i - H_i a_{2i} w_{2i}^T H_i$.
 - Select $z_i \in \mathbb{R}^n$, so that $z_i^T H_i a_{2i} \neq 0$, and compute $\lambda_i = \frac{r_{2i-1}(x_{i-1})r_{2i}(x_{i-1})}{z_i^T H_i a_{2i}}$, if $r_{2i-1}(x_{i-1})r_{2i}(x_{i-1}) \neq 0$, $\lambda_i = \frac{r_{2i}(x_{i-1})}{z_i^T H_i a_{2i}}$, if $r_{2i-1}(x_{i-1}) = 0$ and $r_{2i}(x_{i-1}) \neq 0$.
- Take $x_i = x_{i-1} - \lambda_i H_i^T z_i$.
 - Set $i = i + 1$. Endwhile.
- Stop (x_l is a solution of the system). From (18), we can compute general solution of the system after the final iterate by $x_l = x_l - H_l^T s$ where $s \in \mathbb{R}^n$ is arbitrary.

Remark 2.9 To reduce computational complexity of Algorithm 1, we propose the following tactics:

- Taking x_0 is proper in step (1) as Algorithm 1.
- To compute $w_1 \in \mathbb{R}^n$, we take $t_j = H_0 c_1$, and then, we define w_1 as follows:

$$w_1 = \begin{cases} \frac{1}{\text{sign}(t_{j_M})t_{j_M}}, & i = j_M, \\ 0, & i \neq j_M. \end{cases}$$

- Then, we take $t_{j_M} = \max\{|t_j| : j \in \{1, \dots, n\}, \text{such that } w_1^T t_j = 1\}$. To determine the other $w_{2i-1} \in \mathbb{R}^n$, $i = 2, \dots, l$, we let $t'_j = H_{l-1} c_{2i-1}$, and then, we continue by the similar way.
- We can choose $w_{2i} = z_i \in \mathbb{R}^n$, $i = 1, \dots, l$, by defining the next parameters $d_j = H_i a_{2i}$, and

$$w_{2i} = z_i = \begin{cases} \frac{1}{\text{sign}(d_{j_M})d_{j_M}}, & i = j_M, \\ 0, & i \neq j_M. \end{cases}$$

Now, we take $d_{j_M} = \max\{|d_j| : j \in \{1, \dots, n\}\}$, such that $w_{2i}^T d_j = z_i^T d_j = 1$.

Now, assume that the vectors $a_1, \dots, a_m$ are linearly independent. According to Theorem 2.5, we have $N(H_i) = 2i - 1$ and $N(H_{l_i}) = 2i$. Therefore, $2i - 1$ rows of matrix H_i are depend on the other rows of H_i and $2i$ rows of the matrix H_{l_i} are depend on the other rows of H_{l_i} . As we have defined the matrix H_i and H_{l_i} in such a way that exactly $2i - 1$ rows of H_i and $2i$ rows of matrix H_{l_i} are the zero vectors, we can delete these zero vectors using appropriate operator. Next, we will show how to reduced the number of rows of the Abaffian matrix by two in every iterate and economize the space needed for our new model.

New compression two-step ABS algorithm

Algorithm 2 Assume that $A_{m \times n} = [a_1, a_2, \dots, a_m]^T$ is a full row rank linear matrix, where $m = 2l$ and $Ax = b$ is a compatible full row rank linear system, $x \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$.

1. Let $x_0 \in \mathbb{R}^n$ be an arbitrary vector and take $H_0 = I_n$. Set $i = 1$.
2. (a) Compute $r_1(x_0) = a_1^T x_0 - b_1$ and $r_2(x_0) = a_2^T x_0 - b_2$.
 (b) If $r_1(x_0)r_2(x_0) \neq 0$, we let

$$\begin{cases} a_1 = r_2(x_0)a_1, & a_2 = r_1(x_0)a_2, \\ b_1 = r_2(x_0)b_1, & b_2 = r_1(x_0)b_2. \end{cases}$$
 (c) If $r_1(x_0)r_2(x_0) = 0$ and one of the residual values is nonzero, without loss of generality, we assume that $r_2(x_0) \neq 0$ and we let

$$\begin{cases} a_1 = a_2 + a_1, & a_2 = a_2, \\ b_1 = b_2 + b_1, & b_2 = b_2. \end{cases}$$
 (d) If $r_1(x_0)r_2(x_0) = 0$, and both the residual values are zero, x_1 will be x_0 and go to (5).
3. (a) Take $c_1 = a_2 - a_1$.
 (b) Select $w_1 \in \mathbb{R}^n$, such that $w_1^T H_0 c_1 = 1$ and compute $H_1 = D(H_0 - H_0 c_1 w_1^T H_0)$. (We define D as an operator that delete the zero rows of a matrix).
 (c) Select $w_2 \in \mathbb{R}^{n-1}$, such that $w_2^T H_1 a_2 = 1$ and compute $H_{l_1} = D(H_1 - H_1 a_2 w_2^T H_1)$.

- (d) Select $z_1 \in \mathbb{R}^{n-1}$, so that $z_1^T H_1 a_2 \neq 0$ and compute $\lambda_1 = \frac{r_1(x_0)r_2(x_0)}{z_1^T H_1 a_2}$, if $r_1(x_0)r_2(x_0) \neq 0$, $\lambda_1 = \frac{r_2(x_0)}{z_1^T H_1 a_2}$, if $r_1(x_0) = 0$ and $r_2(x_0) \neq 0$.
4. Take $x_1 = x_0 - \lambda_1 H_1^T z_1$.
5. Set $i = 2$ and go to 6.
6. While $i \leq \frac{m}{2}$, do steps 6(a)–8(b).
 - (a) Compute $r_{2i-1}(x_{i-1}) = a_{2i-1}^T x_{i-1} - b_{2i-1}$ and $r_{2i}(x_{i-1}) = a_{2i}^T x_{i-1} - b_{2i}$.
 - (b) If $r_{2i-1}(x_{i-1})r_{2i}(x_{i-1}) \neq 0$, then we let

$$\begin{cases} a_{2i-1} = r_{2i}(x_{i-1})a_{2i-1}, & a_{2i} = r_{2i-1}(x_{i-1})a_{2i}, \\ b_{2i-1} = r_{2i}(x_{i-1})b_{2i-1}, & b_{2i} = r_{2i-1}(x_{i-1})b_{2i}. \end{cases}$$
 - (c) If $r_{2i-1}(x_{i-1})r_{2i}(x_{i-1}) = 0$ and one of the residual values is nonzero, without loss of generality, we assume that $r_{2i}(x_{i-1}) \neq 0$ and we let

$$\begin{cases} a_{2i-1} = a_{2i} + a_{2i-1}, & a_{2i} = a_{2i}, \\ b_{2i-1} = b_{2i} + b_{2i-1}, & b_{2i} = b_{2i}. \end{cases}$$
 - (d) If $r_{2i-1}(x_{i-1})r_{2i}(x_{i-1}) = 0$ and both of the residual values are zero, x_i will be x_{i-1} and go to 8(b).
7. (a) Take $c_{2i-1} = a_{2i} - a_{2i-1}$.
 (b) Select $w_{2i-1} \in \mathbb{R}^{n-(2i-2)}$, such that $w_{2i-1}^T H_{l_{i-1}} c_{2i-1} = 1$, and compute $H_i = D(H_{l_{i-1}} - H_{l_{i-1}} c_{2i-1} w_{2i-1}^T H_{l_{i-1}})$.
 (c) Select $w_{2i} \in \mathbb{R}^{n-(2i-1)}$, such that $w_{2i}^T H_i a_{2i} = 1$ and compute $H_{l_i} = D(H_i - H_i a_{2i} w_{2i}^T H_i)$.
 (d) Select $z_i \in \mathbb{R}^{n-(2i-1)}$, so that $z_i^T H_i a_{2i} \neq 0$, and compute $\lambda_i = \frac{r_{2i-1}(x_{i-1})r_{2i}(x_{i-1})}{z_i^T H_i a_{2i}}$, if $r_{2i-1}(x_{i-1})r_{2i}(x_{i-1}) \neq 0$, $\lambda_i = \frac{r_{2i}(x_{i-1})}{z_i^T H_i a_{2i}}$, if $r_{2i-1}(x_{i-1}) = 0$ and $r_{2i}(x_{i-1}) \neq 0$.
8. (a) Take $x_i = x_{i-1} - \lambda_i H_i^T z_i$.
 (b) Set $i = i + 1$. Endwhile.
9. Stop (x_l is a solution of the system). General solution of the system is obtained after the final iterate by $x_{l_i} = x_l - H_{l_i}^T s$, where $s \in \mathbb{R}^n$ is arbitrary.

Remark 3.1 We can present the stepsize for Algorithms 1 and 2 as the following form:

$$\lambda_i = \frac{a_{2i-1}^T x_{i-1} - b_{2i-1}}{z_i^T H_i a_{2i}} = \frac{a_{2i}^T x_{i-1} - b_{2i}}{z_i^T H_i a_{2i}} = \frac{\tilde{r}_i}{z_i^T H_i a_{2i}}.$$

Notice that $a_{2i-1}, a_{2i}, b_{2i-1}, b_{2i}$ are updated in steps 2 and 6 and $\tilde{r}_i$ is defined the updated residual value of the i th step. Thus, we have the exact values presented in 3(d) and 7(d)

for the stepsize. Notice that dimension of H_i and z_i is compressed in Algorithm 2.

Remark 3.2 To reduce computational complexity of Algorithm 2, we propose the following tactics.

1. Taking x_0 is proper in step (1) as our new algorithm.
2. To compute $w_1 \in \mathbb{R}^n$, we take $t_j = H_0 c_1$, and then, we define w_1 as follows:

$$w_1 = \begin{cases} \frac{1}{\text{sign}(t_{j_M})t_{j_M}}, & i = j_M, \\ 0, & i \neq j_M. \end{cases}$$

Then, we take $t_{j_M} = \max\{|t_j| : j \in \{1, \dots, n\}, \text{such that } w_1^T t_j = 1\}$. To compute the other $w_{2i-1} \in \mathbb{R}^{n-(2i-2)}$, $i = 2, \dots, l$, we let $t'_j = H_{l-1} c_{2i-1}$, and then, we continue the similar way.

3. We can choose $w_{2i} = z_i \in \mathbb{R}^{n-(2i-1)}$, $i = 1, \dots, l$, by defining the next parameters $d_j = H_i a_{2i}$, and

$$w_{2i} = z_i = \begin{cases} \frac{1}{\text{sign}(d_{j_M})d_{j_M}}, & i = j_M, \\ 0, & i \neq j_M. \end{cases}$$

Now, we take $d_{j_M} = \max\{|d_j| : j \in \{1, \dots, n - (2i-1)\}\}$, such that $w_{2i}^T d_j = z_i^T d_j = 1$.

Notice that by reducing the numbers of rows of the Abaffian matrix by two, for per step, we choose the other parameters as proper dimensions.

Analysis of error propagation in the new versions of two-step ABS models

We investigate the stability of Algorithm 1, for nonsingular system (2), by the backward error analysis technique due to Broyden and Galantai as [8, 9, 12]. Some basic results given by Broyden [9] and Galantai [12] are extended here. Galantai's method has important role in our work. Systems (1) and (2) are equivalent, but for proceeding with Galantai's method, we choose system (2) in all of this section. We prove the analysis of error propagation for nonsingular systems, and then, we will conclude that it is valid for every full row rank system. Galantai studied the numerical stability of the ABS class for linear systems of the form:

$$A^T x = b, \quad (A \in \mathbb{R}^{n \times n})$$

n where the matrix A^T is nonsingular. $P = [p_1, \dots, p_n]$ and $V = [v_1, \dots, v_n]$ are $n \times n$ -type nonsingular matrices with column vectors p_j and v_j ($j = 1, \dots, n$). Denote by I the unit

matrix of $n \times n$ type. The ABS class has the following form:

Algorithm 3

Let $x_0 \in \mathbb{R}^n$ be arbitrary
 For $k = 1, \dots, n$;
 Compute:
 $\alpha_k = \frac{v_k^T (A^T x_{k-1} - b)}{p^T A v_k}$;
 $x_k = x_{k-1} - \alpha_k p_{k-1}$;
 end for;
 where the ABS update algorithm is given by
 Set $H_1 = I$;
 For $k = 1, \dots, n$;
 Compute:
 $p_k = H_k^T z_k$; ($p_k^T A v_k \neq 0$)
 $H_{k+1} = H_k - \frac{H_k A v_k w_k^T H_k}{(w_k^T H_k A v_k)}$; ($w_k^T H_k A v_k \neq 0$)
 end for.

Algorithm 3 is finitely terminated in n steps. The matrix V is scaling the system $A^T x = b$. As [20], the pair (P, V) is said to be the A^T -conjugate if $V^T A^T P = L$ is the lower triangular. The ABS algorithm generates all A^T -conjugate directions for suitable choices of parameters [2]. Therefore, the ABS class of methods coincides with those studied by Stewart [20]. Results on the numerical stability of conjugate direction methods are given in [9, 21]. A stability analysis for decent methods is given in [7]. See also [3, 4, 10]. As Corollary 2.7, we conclude that the pair (P, I) is A^T -conjugate for Algorithm 1, considering nonsingular matrix of system (2). Using the basic idea of Broyden's backward error analysis method, we present Algorithm 1 as the following form:

$$X_k = \Psi_{k-1}(X_{k-1}), \quad \left(k = 1, \dots, \frac{n}{2}\right), \quad (26)$$

where X_k is the solution of the problem. Assume that an error ξ_j occurs at the j th step and that is error propagates further. It is also assumed that no other source of error occurs. The exact solution $X_{\frac{n}{2}}$ is given by the following:

$$X_{\frac{n}{2}} = \Psi_{\frac{n}{2}-1} \{ \Psi_{\frac{n}{2}-2} \{ \dots \{ \Psi_j(X_j) \} \dots \} \} = \Omega^{\frac{n}{2}-j}(X_j), \quad (27)$$

while the perturbed solution $X'_{\frac{n}{2}}$ is given by the following:

$$X'_{\frac{n}{2}} = \Omega^{\frac{n}{2}-j}(X_j + \xi_j). \quad (28)$$

If the quantity $\|X_{\frac{n}{2}} - X'_{\frac{n}{2}}\|$ is large, then the algorithm (26) must be very unstable. As Broyden ideas, it must be small for stable algorithms. Consequently, $\|X_{\frac{n}{2}} - X'_{\frac{n}{2}}\|$ is a measure of stability for our new version of algorithm.

Now, we recall some of the important parameters of Algorithm 1 for the system $A^T x = b$, $A \in \mathbb{R}^{n \times n}$, where the matrix A^T is nonsingular considering system (2). Let

$P = [\tilde{p}_1, \dots, \tilde{p}_{\frac{n}{2}}]$, with column vectors defined from Theorem 2.6 as the form $\tilde{p}_i = (p'_i, p_i)$. We know that, $p_i = H_i^T z_i$ is obtained the first-phase rank reducing process of the i th iteration. Thus, P is $n \times n$ -type nonsingular matrix. Here, we take $V = I_n = [v_1, \dots, v_{\frac{n}{2}}]$ with $v_k = e_k = (e_{2k-1}, e_{2k})$, ($k = 1, \dots, \frac{n}{2}$). Therefore, $V = I_n = [e_1, \dots, e_n]$ is an $n \times n$ identify matrix. As Corollary 2.4, we have $p_k^T A e_{2k-1} = p_k^T A e_{2k} \neq 0$. Taking v_k for the stepsize means that we can arbitrary choose one of the values of e_{2k-1} or e_{2k} . Also from Remark 3.1, we have

$$\lambda_k = \frac{e_{2k-1}^T (A^T x_{k-1} - b)}{p_k^T A e_{2k-1}} = \frac{e_{2k}^T (A^T x_{k-1} - b)}{p_k^T A e_{2k}}.$$

We choose $H_0 = I_n$ and we present λ_k, x_k, p_k , and H_k as the following form:

$$\lambda_k = \frac{v_k^T (A^T x_{k-1} - b)}{p_k^T A v_k}, \quad (29)$$

$$x_k = x_{k-1} - \lambda_k p_k, \quad (30)$$

$$p_k = H_k^T z_k, \quad (p_k^T A v_k \neq 0) \quad (31)$$

$$H_k = H_{k-1} - H_{k-1} A (v_{2k} - v_{2k-1}) w_{2k-1}^T H_{k-1}, \quad (32)$$

$$(w_{2k-1}^T H_{k-1} A (v_{2k} - v_{2k-1}) = 1),$$

where $H_k = H_k - H_k A v_{2k} w_{2k}^T H_k$, $w_{2k}^T H_k^T A v_{2k} = 1$. Updating vectors and residual values can be computed as Algorithm 1, for $k = 1, \dots, \frac{n}{2}$. Next, we investigate the stability of the conjugate direction methods of the form (29) and (30) with a special emphasize on the ABS update (31) and (32). It is noted again that for (29)–(32) and the others related parameters, we choose V the unitary matrix and the pair (P, I) is A^T conjugate for a nonsingular matrix as system (2). Considering these points, Algorithm 1 coincides with those studies in [12] for a unitary V and $\frac{n}{2}$ steps. Some basic results of [9, 12, 20] are extended here. Let us introduce the following notation:

$$P_k = \frac{p_k v_k^T A^T}{v_k^T A^T p_k}. \quad (33)$$

It is easily verified that $P_k^2 = P_k$, i.e., P_k is a projector. The matrix P_k is a rank one projector onto the space spanned by p_k along the orthogonal complement of the space spanned by $A v_k$. Therefore, we have $R(P_k) = R(p_k)$ and $N(p_k) = R^\perp(A v_k)$. Note that the projectors $P_1, \dots, P_n$ are called conjugate by Stewart [20], if we have $P_i P_j = 0$ ($i < j$). The following lemma is immediate consequent of Stewart's definition [20].

Lemma 4.1 *Let $P_1, P_2, \dots$ be conjugate projectors. for $i \leq k$ we have:*

$$P_i (I - P_k) (I - P_{k-1}) \cdots (I - P_1) = 0. \quad (34)$$

Proof By conjugacy, for $i < j$, we have

$$P_i (I - P_j) = P_i - P_i P_j = P_i, \quad (35)$$

and since P_i is a projector,

$$P_i (I - P_i) = P_i - P_i^2 = P_i - P_i = 0. \quad (36)$$

Together (35) and (36), we have (34). $\square$

With the notation P_k as (33), the method of the class (29) and (30) has the following form (motivating techniques in [12, 20] and extending them for our model):

$$x_k = (I - P_k) x_{k-1} + d_k, \quad \left(d_k = \frac{p_k v_k^T b}{v_k^T A^T p_k} \right)$$

Denote by x^* the solution of the linear system. By taking $E_k = x^* - x_k$, we have the recursion $E_k = (I - P_k) E_{k-1}$ with the following solution:

$$E_k = (I - P_k) \cdots (I - P_1) E_0.$$

Now, we introduce the notation

$$(I - P_k) \cdots (I - P_j) = Q_{k,j},$$

for $k \geq j$. Now, let us suppose that an error occurs at the $(k-1)$ th step and only at it. The perturbed results of the $(k-1)$ th step are denoted by x'_{k-1} . The perturbed results of further steps are denoted by x'_j ($j = k, \dots, \frac{n}{2}$). Then, we have

$$x'_n = \prod_{j=0}^{\frac{n}{2}-k} (I - P_{\frac{n}{2}-j}) x'_{k-1} + \sum_{j=k}^{\frac{n}{2}} \left(\prod_{t=0}^{\frac{n}{2}-j-1} (I - P_{\frac{n}{2}-t}) \right) d_j,$$

from which it follows that the error occurring in the final iteration:

$$\begin{aligned} x^* - x'_n &= x_n - x'_n = \prod_{j=0}^{\frac{n}{2}-k} (I - P_{\frac{n}{2}-j}) (x_{k-1} - x'_{k-1}) \\ &= Q_{\frac{n}{2},k} (x_{k-1} - x'_{k-1}). \end{aligned}$$

The matrix $Q_{\frac{n}{2},k}$ can be considered as the error matrix. Hence, we have the error bound

$$\|x_n - x'_n\| \leq \|Q_{\frac{n}{2},k}\| \|x_{k-1} - x'_{k-1}\|. \quad (37)$$

A method of the classes (29) and (30) is considered to be optimal in the sense of Broyden [9], if $\|Q_{\frac{n}{2},k}\|$ is minimal for all k . First, we characterize $Q_{\frac{n}{2},k}$. Using the A^T -conjugate property, we have $P_i P_j = 0$ ($i < j$). Therefore, $P_i Q_{\frac{n}{2},k} = 0$ ($k \leq i \leq \frac{n}{2}$) and $(I - P_i) Q_{\frac{n}{2},k} = Q_{\frac{n}{2},k}$ ($k \leq i \leq \frac{n}{2}$), which means that $Q_{\frac{n}{2},k} Q_{\frac{n}{2},k} = Q_{\frac{n}{2},k}$, i.e., $Q_{\frac{n}{2},k}$ is a projector. By observing that $R(I - P_k) = N(P_k)$ and $N(I - P_k) = R(P_k)$, we conclude that

$$R(Q_{\frac{n}{2},k}) = \bigcap_{j=k}^{\frac{n}{2}} N(P_j), \quad N(Q_{\frac{n}{2},k}) = \sum_{j=k}^{\frac{n}{2}} R(P_j).$$

Thus, we have

$$R(Q_{\frac{n}{2},k}) = \bigcap_{j=k}^{\frac{n}{2}} R^\perp(Av_j) = R^\perp(AV^{\frac{n}{2}-k+1}),$$

and

$$N(Q_{\frac{n}{2},k}) = \sum_{j=k}^{\frac{n}{2}} R(P_j) = R(P^{\frac{n}{2}-k+1}) = R^\perp(AV^{k-1}).$$

Here, motivating Householder notations, for any matrix A , we define the matrices $A^{\bar{k}}$, A^k , $A^{\underline{k}}$, and $A^{|k|}$, the submatrices consisting of, respectively, the first $2k$ rows, the first $2k$ columns, the last $2k$ rows, and the last $2k$ columns of A .

The following lemmas show that the symmetric projectors have a minimum property both in the Frobenius and in the spectral norm. See the proof of the following lemmas in [12].

Lemma 4.2 If $A^2 = A$, $A \neq 0$, and $\text{rank}(A) = m$, then $\|A\|_F \geq \sqrt{m}$ and $\|A\| = \sqrt{m}$ if and only if A is symmetric.

Lemma 4.3 If $A^2 = A$ and $A \neq 0$, then $\|A\|_{sp} \geq 1$ and $\|A\| = 1$ if and only if $A = A^T$.

Theorem 4.4 The method of classes (29) and (30) is optimal in the sense of Broyden if and only if (39) holds or equivalently $V^T A^T A V = D$ satisfies for a diagonal matrix D .

Proof A projector P is symmetric if and only if $R(P) = N^\perp(P)$. Hence, $Q_{\frac{n}{2},k}$ is symmetric (and has minimal norm) if and only if

$$R(AV^{k-1}) = R^\perp(AV^{\frac{n}{2}-k+1}). \quad (38)$$

A method is optimal in Broyden's sense if and only if (38) is satisfied for all k . The latter condition is equivalent to

$$Av_i \perp Av_j, \quad (i \neq j). \quad (39)$$

In the matrix formulation, it means that $V^T A^T A V = D$, where D is diagonal. Notice that, since we choose V the unitary matrix, the theorem is true if and only if $A^T A = D$ holds. $\square$

This result was originally obtained by Broyden in [9] in a different way for Algorithm 3, but we extended the proof of Galantai [12] for Algorithm 1.

Theorem 4.5 For the error propagation (26)–(28) and (37), the classes (29) and (30) yield the bound

$$\|x_{\frac{n}{2}} - x'_{\frac{n}{2}}\| \leq k(A) \|x_{k-1} - x'_{k-1}\|,$$

where $k(A)$ is the condition number of matrix A .

Proof Since $Q_{\frac{n}{2},k}$ is a projector onto $R^\perp(AV^{\frac{n}{2}-k+1})$ along $R^\perp(AV^{k-1})$, it can be represented in the form:

$$Q_{\frac{n}{2},k} = (A^{-T} V^{-T})^{k-1} (V^T A^T)^{\overline{k-1}}, \quad (40)$$

As $\|B^k\| \leq \|B\|$ and $\|B^{\bar{k}}\| \leq \|B\|$ are both in the Frobenius and spectral norms, we may bound $Q_{\frac{n}{2},k}$ by $\|Q_{\frac{n}{2},k}\| \leq \|A^{-T} V^{-T}\| \|V^T A^T\| \leq k(A)k(V)$, which is exactly $\|Q_{\frac{n}{2},k}\| \leq k(A)k(V)$. Since we choose V the unitary matrix, $k(V) = 1$ and $\|Q_{\frac{n}{2},k}\| \leq k(A)$. As (37), we conclude that

$$\|x_{\frac{n}{2}} - x'_{\frac{n}{2}}\| \leq k(A) \|x_{k-1} - x'_{k-1}\|.$$

$\square$

Theorem 4.6 The residual error $r'_k = A^T(x_k - x'_k)$ as Algorithm 1 is minimal for all k .

Proof Defining the residual perturbation as $r'_k = A^T(x_k - x'_k)$, we have $r'_{\frac{n}{2}} = A^T Q_{\frac{n}{2},k} A^{-T} r'_k$ for the model (26)–(28) and (37). Using the relation $(AB)^k (CD)^{\bar{k}} = A\{B^k C^{\bar{k}}\}D$ and (40), we have

$$A^T Q_{\frac{n}{2},k} A^{-T} = (V^{-T})^{k-1} (V^T)^{\overline{k-1}} \quad (41)$$

is a projector onto $R((V^{-T})^{k-1})$ along $R((V^{-T})^{\frac{n}{2}-k+1})$. The bound $\|r'_{\frac{n}{2}}\| \leq \|A^T Q_{\frac{n}{2},k} A^{-T}\| \|r'_k\|$ is minimal if and only if

$$R((V^{-T})^{k-1}) = R^\perp((V^{-T})^{\frac{n}{2}-k+1}).$$

The classes of the (29) and (30) can be called optimal for the residual perturbation r'_k if (41) holds for all k . This condition is obviously satisfied if and only if $(V^{-T})^T (V^{-T}) = D^{-1}$ for a suitable diagonal matrix D from which the condition $V^T V = D$ follows. Since we choose V as a unitary matrix, r'_k is minimal for all k . $\square$

Corollary 4.7 Assume that instead of (29) and (30) we have the next recursion

$$x'_k = (I - P_k)(x'_{k-1} + \xi_{k-1}) + d_k, \quad \left(k = 1, \dots, \frac{n}{2}\right),$$

where ξ_{k-1} denote the error which occurred at the $(k-1)$ -th iteration. Then, we have $x_{\frac{n}{2}} - x'_{\frac{n}{2}} = \sum_{k=1}^{\frac{n}{2}} Q_{\frac{n}{2},k} \xi_{k-1}$ from which the bound

$$\|x_{\frac{n}{2}} - x'_{\frac{n}{2}}\| \leq \sum_{k=1}^{\frac{n}{2}} \|Q_{\frac{n}{2},k}\| \|\xi_{k-1}\| \leq k(A) \sum_{k=1}^n \|\xi_{k-1}\|$$

follows. For the optimal method, $k(A)$ is obviously replaced by $\sqrt{m}$ or by 1 depending on the norm chosen.

Corollary 4.8 Since the achievements of this section are valid for every nonsingular linear system of (2), it is valid for full row rank systems of this class. Algorithm 2 just

Table 1 Number of multiplications required to solve m linear equations in $l = \frac{m}{2}$ iterates, for the main parameters

	The number of multiplications for each iterate
Residual values	$2n$
The products of residual values	1
a_{2i-1}, a_{2i}	$2n$
b_{2i-1}, b_{2i}	2
H_i	$(2n+1)(n-2i+2)$
H_{l_i}	$(2n+1)(n-2i+1)$
x_i	$2n$

deletes the zero rows of the Abaffian matrix as Algorithm 1 and consequently reduces dimensions of the related parameters. Thus, the results of analysis of error propagation are true for the new compression two-step ABS scheme. Notice that systems (1) and (2) are the equivalent. Moreover, the points 2 and 3, as Remarks 2.9 and 3.2, are useful to increase stability of our algorithms.

Computational complexity and numerical results

At first, considering Remark 2.9, we compute the number of multiplications, for Algorithm 1, but for arbitrary non-singular matrix H_0 , as follows (Table 1):

Hence, the total number of multiplications for the l iterates is

$$N = \sum_{i=1}^l [2n+1+2n+2+(2n+1)(n-2i+2) + (2n+1)(n-2i+1)+2n],$$

$$N = 2mn^2 - m^2n + O(mn) - O(m^2) + O(m).$$

The total number for Huang's method is $\frac{3}{2}mn^2 + O(mn)$ multiplications and $3mn^2 - \frac{7}{4}m^2n + \frac{1}{6}m^3 + O(mn) + O(m^2) + O(n^2)$ for the two-step methods presented by Amini et al. [5, 6]. Our computations indicate that our numerical results are better than Amini et al.'s two-step algorithms. In addition, when $n < 2m$, we need cheaper number of multiplications up to the corresponding Huang's algorithm. In addition, when m tends to n , the total number for our two-step method is n^3 multiplications, while it is $\frac{3}{2}n^3$ for Huang's algorithm and $\frac{17}{12}n^3$ for those two-step methods given by Amini et al. [5, 6]. Obviously, when m and n are not too large, the lower order terms have influence on our results.

Remark 5.1 Now, considering Algorithm 1 with Remark 2.9, setting of H_0 the identity matrix I_n , we have the better results. In this case, for the i th iteration, the number of multiplications of the matrix H_i is $(n-2i+2)(4i-1)$ and it is $(n-2i+1)(4i+1)$ multiplications for the matrix H_{l_i} .

Therefore, our algorithm needs $nm^2 - \frac{2}{3}m^3 + O(nm) + O(m^2)$ multiplications. Hence, for $m = n$, $\frac{1}{3}n^3$, multiplications plus lower order terms are needed. Notice that the main storage requirement is the storage of H_{l_i} , which has at most $\frac{1}{4}n^2$ positions. This is half of the storage needed by the Gaussian elimination method. Our new version ABS algorithm is computationally better than the classical Gaussian elimination method, having the same arithmetic cost but using less memory and no pivoting is necessary.

Remark 5.2 The operator D is used to compress and economize our required space, via deleting the zero rows of the Abaffian matrix by two for per iteration. Therefore, it has not any effect on the number of multiplications, and Algorithms 1 and 2 have the same numbers of multiplications. Thus, both of our works are better than Amini et al.'s two-step ABS algorithms and Huang's method. Furthermore, we have good competition with the classical Gaussian elimination method.

Conclusion

This paper presents two new version of ABS models for the general solution of compatible full row rank linear systems of equations in at most $\lceil \frac{m+1}{2} \rceil$ iterates. For the second algorithm, the number of rows of the Abaffian matrix is reduced by two for every iteration and the space is compressed as much as possible. A theory of conjugate projectors is developed and used for the analysis of error propagation. We evaluate the numerical stability of our schemes using Broyden's backward error analysis method and Galantai's technique. In addition, both our models need less computational complexity up to those corresponding two-step ABS algorithms and Huang's method. Our new two-step ABS algorithms is computationally better than the classical Gaussian elimination method, having the same arithmetic cost but using less memory and no pivoting is necessary.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Abaffy, J., Spedicato, E.: *ABS Projection Algorithms: Mathematical Techniques for Linear and Nonlinear Equations*. Prentice-Hall, Inc (1989)
2. Abaffy, J., Galantai, A.: *Conjugate Direction Methods for Linear and Non Linear Systems of Algebraic Equations*, Quaderni DMSIA, 1986–7. Univ. di Bergamo, Bergamo (1986)
3. Abaffy, J., Broyden, C.G., Spedicato, E.: Numerical Performance of the Pseudo Symmetric Algorithm in the ABS Class Versus LU Factorization with Iterative Refinement, Rappporto SOFTMAT 8/83. Univ. di Bergamo, Bergamo (1983)
4. Abaffy, J., Broyden, C.G., Spedicato, E.: A class of direct methods for linear systems. *Numer. Math.* **45**, 361–376 (1984)
5. Amini, K., Mahdavi-Amiri, N., Peyghami, M.R.: ABS-type methods for solving full row rank linear systems using a new rank two update. *Bull. Aust. Math. Soc.* **70**(1), 17–34 (2004)
6. Amini, K., Mahdavi-Amiri, N., Peyghami, M.R.: Extended reduced rank two Abaffian update schemes in the ABS-type methods. *Appl. Math. Comput.* **185**(1), 255–265 (2007)
7. Bollen, J.A.M.: Numerical stability of descent methods for solving linear equations. *Numer. Math.* **43**, 361–377 (1984)
8. Broyden, C.G.: Error propagation in numerical processes. *J. Inst. Math. Appl.* **14**, 131–140 (1974)
9. Broyden, C.G.: On the numerical stability of Huang's and related methods. *J. Optim. Theory Appl.* **47**(401), 401–412 (1985)
10. Egervary, E.: On rank-diminishing operations and their applications to the solution of linear equations. *Z. Angew. Math. Phys.* **11**, 376–386 (1960)
11. Emilio, S., Spedicato, E., Bodon, E., Xia, Z., Mahdavi-Amiri, N.: ABS methods for continous and integer linear equations and optimization. *Cent. Eur. J. Oper. Res.* **18**(1), 73–95 (2010)
12. Galantai, A.: Analysis of error propagation in the ABS class for linear systems. *Ann. Inst. Stat. Math.* **43**(3), 597–603 (1991)
13. Ghanbari, R., Mahdavi-Amiri, N.: New solutions of LR linear systems using ranking functions and ABS algorithms. *Appl. Math. Model.* **34**, 3363–3375 (2010)
14. Golub, G.H., Van Loan, C.F.: *Matrix Computations*. The Johns Hopkins University Press, Baltimore (1983)
15. Jain, S., Mangal, A., Sharma, S.: C-Approach of ABS algorithm for fractional programming problem. *J. Comput. Math. Sci.* **4**(2), 126–134 (2013)
16. Khorramizadeh, M.: A generalization of iteration-free search vectors of ABS methods. *J. Math. Comput. Sci.* **12**, 196–200 (2014)
17. Khorramizadeh, M., Mahdavi-Amiri, N.: An efficient algorithm for sparse null space basis problem using ABS methods. *Numer. Algorithms* **62**(3), 469–485 (2013)
18. Mahdavi-Amiri, N., Golpar-Raboky, E.: Extended rank reduction formulas containing Wedderburn and Abaffy–Broyden–Spedicato rank reducing processes. *Linear Algebra Appl.* **439**(11), 3318–3331 (2013)
19. Spedicato, E., Bodon, E., Popolo, A.D., Mahdavi-Amiri, N.: ABS methods and ABSPACK for linear systems and optimization: a review. *Quart. J. Belgian French Ital. Oper. Res. Soc.* **1**(1), 51–66 (2003)
20. Stewart, G.W.: Conjugate direction methods for solving systems of linear equations. *Numer. Math.* **21**, 285–297 (1973)
21. Wozniakowski, H.: Roundoff error analysis of a new class of conjugate gradient algorithms. *Linear Algebra Appl.* **29**, 507–529 (1980)