



## Weaker hypotheses for the general projection algorithm with corrections

Alexandru Bobe, Aurelian Nicola, Constantin Popa

### Abstract

In an earlier paper [J. of Appl. Math. and Informatics, 29(3-4)(2011), 697-712] we proposed a general projection-type algorithm with corrections and proved its convergence under a set of special assumptions. In this paper we prove convergence of this algorithm under a much weaker set of assumptions. This new framework gives us the possibility to obtain as a particular case of our method the two-step algorithm analysed in [BIT, 38(2)(1998), 275-282].

### 1 Introduction

We will consider the problem: find  $x \in \mathbb{R}^n$  such that

$$\|Ax - b\| = \min\{\|Az - b\|, z \in \mathbb{R}^n\}, \quad (1)$$

where  $\|\cdot\|$  denotes the Euclidean norm ( $\langle \cdot, \cdot \rangle$  will be the Euclidean scalar product). Concerning the matrix involved in (1) we will suppose throughout the paper that it has nonzero rows  $A_i$  and columns  $A^j$ , i.e.,

$$A_i \neq 0, i = 1, \dots, m, \quad A^j \neq 0, j = 1, \dots, n. \quad (2)$$

These assumptions are not essential restrictions of the generality of the problem (1) because, if  $A$  has null rows and/or columns, it can be easily proved that

---

Key Words: inconsistent least-squares problems, projection algorithms, correction  
2010 Mathematics Subject Classification: AMS: 65F10, 65F20  
Received: October, 2014.  
Revised: October, 2014.  
Accepted: October, 2014.

they can be eliminated without affecting its set of classical or least squares solutions. We first introduce some notations: the spectrum and spectral radius of a square matrix will be denoted by  $\sigma(B)$  and  $\rho(B)$ , respectively. By  $A^T$ ,  $\mathcal{N}(A)$ ,  $\mathcal{R}(A)$  we will denote the transpose, null space and range of  $A$ .  $P_S(x)$  will be the orthogonal (Euclidean) projection onto a vector subspace  $S$  of some  $\mathbb{R}^q$ .  $S(A; b)$ ,  $LSS(A; b)$ ,  $x_{LS}$  will stand for the set of classical and least squares solution of (1), respectively, and the (unique) minimal norm solution. In the consistent case for (1) we have  $S(A; b) = LSS(A; b)$ . In the general case the following properties are known (see [5], Chapter 1)

$$x_{LS} \perp \mathcal{N}(A), \quad b = b_A + b_A^*, \quad \text{with } b_A = P_{\mathcal{R}(A)}(b), \quad b_A^* = P_{\mathcal{N}(A^T)}(b), \quad (3)$$

$$LSS(A; b) = x_{LS} + \mathcal{N}(A) \quad \text{and} \quad x \in LSS(A; b) \Leftrightarrow Ax = b_A, \quad (4)$$

$$S(A; b) = x_{LS} + \mathcal{N}(A) \quad \text{and} \quad x \in S(A; b) \Leftrightarrow Ax = b. \quad (5)$$

Let  $Q : n \times n$  and  $R : n \times m$  be real matrices satisfying the (main) assumptions

$$I - Q = RA, \quad (6)$$

and

$$\text{if } \tilde{Q} = QP_{\mathcal{R}(A^T)}, \quad \text{then } \|\tilde{Q}\| < 1, \quad (7)$$

where  $\|\tilde{Q}\|$  denotes the spectral norm of the matrix  $\tilde{Q}$ . In the paper [4] we proposed the following algorithm (a standard example of such an algorithm is Kaczmarz's projection method; see Algorithm 5.4.3 in [1]).

**Algorithm General Projections with Corrections.**

*Initialization:*  $x^0 \in \mathbb{R}^n$  is arbitrary.

*Iterative step:*

$$x^{k+1} = Qx^k + Rb + v^k, \quad (8)$$

where

$$v^k \in \mathcal{R}(A^T), \quad \forall k \geq 0. \quad (9)$$

We introduced in [4] the additional assumptions:

$$\forall y \in \mathbb{R}^m, \quad Ry \in \mathcal{R}(A^T), \quad (10)$$

and, for  $\varepsilon^k$  defined by

$$\varepsilon^k := v^k + Rb_A^*, \quad (11)$$

we assumed that there exist constants  $c > 0$  and  $\delta \in [0, 1)$  such that

$$\|\varepsilon^k\| \leq c\delta^k, \quad \forall k \geq 0. \quad (12)$$

Then, the following result was proved in [4].

**Theorem 1.** *Under the assumptions (6) – (7) and (9) – (12), for any sequence  $(x^k)_{k \geq 0}$  generated by the algorithm (8) we have*

$$P_{\mathcal{N}(A)}(x^k) = P_{\mathcal{N}(A)}(x^0), \forall k \geq 0 \quad (13)$$

and

$$\lim_{k \rightarrow \infty} x^k = P_{\mathcal{N}(A)}(x^0) + x_{LS} \in LSS(A; b). \quad (14)$$

In this paper we will introduce a new set of assumptions weaker than (7) and (12) and we will prove that the algorithm (8) has the same convergence properties. Moreover, we will show that the two step algorithm presented by Elfving in [2] fits into the above mentioned more general hypotheses.

## 2 The new set of weaker assumptions

We will keep the assumptions (6), (10) and (9), but replace (7) with

$$\text{if } \tilde{Q} = QP_{\mathcal{R}(A^T)}, \text{ then } \rho(\tilde{Q}) < 1. \quad (15)$$

and replace (12) with

$$\lim_{k \rightarrow \infty} \varepsilon^k = 0. \quad (16)$$

By a direct application of (6) and (10) we obtain

$$\text{if } x \in \mathcal{N}(A) \text{ then } Qx = x \in \mathcal{N}(A), \quad (17)$$

$$\text{if } x \in \mathcal{R}(A^T) \text{ then } Qx = x - RAx \in \mathcal{R}(A^T). \quad (18)$$

Moreover, the equalities (13) hold as in Lemma 1 of [4].

We can now prove the main result of our paper.

**Theorem 2.** *Under the hypotheses,*

$$\begin{aligned} I - Q &= RA, \\ \text{if } \tilde{Q} &= QP_{\mathcal{R}(A^T)}, \text{ then } \rho(\tilde{Q}) < 1, \\ \forall y \in \mathbb{R}^m, \quad Ry &\in \mathcal{R}(A^T), \\ v^k &\in \mathcal{R}(A^T), \forall k \geq 0, \\ \lim_{k \rightarrow \infty} \varepsilon^k &= 0, \end{aligned}$$

*any sequence  $(x^k)_{k \geq 0}$  generated by the algorithm (8) converges and (14) holds.*

*Proof.* According to (13) we can define the error vector of the iteration (8) as

$$e^k = x^k - (P_{\mathcal{N}(A)}(x^0) + x_{LS}). \quad (19)$$

Moreover, from (15) it follows that  $(I - \tilde{Q})$  is invertible and (see, e.g., [3])

$$(I - \tilde{Q})^{-1} = I + \tilde{Q} + \tilde{Q}^2 + \dots = \sum_{j \geq 0} \tilde{Q}^j. \quad (20)$$

Then, as in [4] we obtain the equality

$$e^{k+1} = \tilde{Q}e^k + \varepsilon^k, \forall k \geq 0, \quad (21)$$

from which we get

$$e^k = \tilde{Q}^k e^0 + \sum_{i=0}^{k-1} \tilde{Q}^{k-1-i} \varepsilon^i, \forall k \geq 1. \quad (22)$$

From (15) (see [3], Lemma 5.6.10) there exist a matrix norm (depending on  $\tilde{Q}$ )  $\|\cdot\|_*$  such that

$$\|\tilde{Q}\|_* < 1. \quad (23)$$

Let  $\|\cdot\|_*$  be a vector norm compatible with the above matrix norm, then

$$\|\tilde{Q}x\|_* \leq \|\tilde{Q}\|_* \|x\|_*, \forall x \in \mathbb{R}^n. \quad (24)$$

But, from (16) we obtain that  $\lim_{k \rightarrow \infty} \|\varepsilon^k\|_* = 0$ , thus (see also (23)) if  $\epsilon > 0$  is arbitrary fixed there exist an  $M > 0$  and an integer  $k_\epsilon^1 \geq 1$  such that

$$\|\varepsilon^k\|_* \leq M, \forall k \geq 1 \quad (25)$$

and

$$\|\tilde{Q}\|_*^k \leq \epsilon, \|\varepsilon^k\|_* < \epsilon, \forall k \geq k_\epsilon^1. \quad (26)$$

Then we define  $k_\epsilon \geq 1$  by

$$k_\epsilon = 2k_\epsilon^1 + 1 \quad (27)$$

and consider  $k = k_\epsilon + \mu$ , with an arbitrary integer  $\mu \geq 0$ . Now, we first take norms, use (26) and split the sum in (22) as

$$\begin{aligned} \|e^k\|_* &\leq \|\tilde{Q}\|_*^k \|e^0\|_* + \sum_{i=0}^{k-1} \|\tilde{Q}\|_*^{k-1-i} \|\varepsilon^i\|_* \\ &\leq \epsilon \|e^0\|_* + \sum_{i=0}^{k_\epsilon^1} \|\tilde{Q}\|_*^{k-1-i} \|\varepsilon^i\|_* + \sum_{i=k_\epsilon^1+1}^{k-1} \|\tilde{Q}\|_*^{k-1-i} \|\varepsilon^i\|_* . \end{aligned} \quad (28)$$

For the first sum in (28), from (25) and (26) we get

$$\begin{aligned}
& \sum_{i=0}^{k_\epsilon^1} \|\tilde{Q}\|_*^{k-1-i} \|\varepsilon^i\|_* \\
& \leq M \sum_{i=0}^{k_\epsilon^1} \|\tilde{Q}\|_*^{k-1-i} = M \left( \|\tilde{Q}\|_*^{k-1} + \dots + \|\tilde{Q}\|_*^{k-1-k_\epsilon^1} \right) \\
& = M \|\tilde{Q}\|_*^{k-1-k_\epsilon^1} \left( \|\tilde{Q}\|_*^{k_\epsilon^1} + \dots + 1 \right) = M \|\tilde{Q}\|_*^{k_\epsilon^1+\mu} \frac{1 - \|\tilde{Q}\|_*^{k_\epsilon^1+1}}{1 - \|\tilde{Q}\|_*} \\
& \leq M \|\tilde{Q}\|_*^{k_\epsilon^1+\mu} \frac{1}{1 - \|\tilde{Q}\|_*} < \frac{M\epsilon}{1 - \|\tilde{Q}\|_*}. \tag{29}
\end{aligned}$$

For the second sum in (28), from the above formula for  $k$ , (27) and (26) we get

$$\begin{aligned}
& \sum_{i=k_\epsilon^1+1}^{k-1} \|\tilde{Q}\|_*^{k-1-i} \|\varepsilon^i\|_* = \|\tilde{Q}\|_*^{k_\epsilon^1+\mu-1} \|\varepsilon^{k_\epsilon^1+1}\|_* \\
& + \|\tilde{Q}\|_*^{k_\epsilon^1+\mu-2} \|\varepsilon^{k_\epsilon^1+2}\|_* + \dots + \|\tilde{Q}\|_* \|\varepsilon^{k_\epsilon^1+\mu-1}\|_* + \|\varepsilon^{k_\epsilon^1+\mu}\|_* \\
& \leq \|\varepsilon^{k_\epsilon^1+1}\|_* \left( \|\tilde{Q}\|_*^{k_\epsilon^1+\mu-1} + \dots + 1 \right) \\
& = \|\varepsilon^{k_\epsilon^1+1}\|_* \frac{1 - \|\tilde{Q}\|_*^{k_\epsilon^1+\mu}}{1 - \|\tilde{Q}\|_*} < \epsilon \frac{1}{1 - \|\tilde{Q}\|_*}. \tag{30}
\end{aligned}$$

From (26) and (28) – (30) we conclude that, for an arbitrary  $\epsilon > 0$ , there exist an integer  $k_\epsilon \geq 1$  (see also (27)), such that for any  $k \geq k_\epsilon$  we have

$$\begin{aligned}
\|e^k\|_* & \leq \epsilon \|e^0\|_* + \sum_{i=0}^{k-1} \|\tilde{Q}\|_*^{k-1-i} \|\varepsilon^i\|_* \\
& \leq \epsilon \|e^0\|_* + \frac{M\epsilon}{1 - \|\tilde{Q}\|_*} + \epsilon \frac{1}{1 - \|\tilde{Q}\|_*} = \epsilon \left( \|e^0\|_* + \frac{M+1}{1 - \|\tilde{Q}\|_*} \right) \tag{31}
\end{aligned}$$

from which (14) follows and the proof is complete.

**Corollary 1.** *Suppose that (6), (15), and (16) hold, and that*

$$x^k \in \mathcal{R}(A^T), \quad \forall k \geq 0. \tag{32}$$

*Then, any sequence  $(x^k)_{k \geq 0}$  generated by the algorithm (8) converges and*

$$\lim_{k \rightarrow \infty} x^k = x_{LS}. \tag{33}$$

*Proof.* We define the error vector  $e^k$  by

$$e^k := x^k - x_{LS}, \quad \forall k \geq 0. \quad (34)$$

Because of (18) and the fact that  $x_{LS} \in \mathcal{R}(A^T)$  we get

$$Qx_{LS} = \tilde{Q}x_{LS}, \quad (35)$$

and

$$\begin{aligned} (I - \tilde{Q})x_{LS} &= x_{LS} - QP_{\mathcal{R}(A^T)}(x_{LS}) \\ &= x_{LS} - Qx_{LS} = (I - Q)x_{LS} = RAx_{LS} = Rb_A, \end{aligned}$$

thus,

$$Rb_A = (I - \tilde{Q})x_{LS} = (I - Q)x_{LS}. \quad (36)$$

Consequently, by using (34), (8), (32) and (36) we successively get ,  $\forall k \geq 0$

$$\begin{aligned} e^{k+1} &= x^{k+1} - x_{LS} = \tilde{Q}x^k + Rb + v^k - x_{LS} \\ &= \tilde{Q}x^k + Rb + v^k - (\tilde{Q}x_{LS} + Rb_A) = \tilde{Q}(x^k - x_{LS}) + v^k + Rb_A^* = \tilde{Q}e^k + \varepsilon^k. \end{aligned} \quad (37)$$

Then, as in the proof of Theorem 2 we obtain (22) for  $e^k$  from (34) and then (33), because for  $x^0 \in \mathcal{R}(A^T)$  we have  $P_{\mathcal{N}(A)}(x^0) = 0$ . This completes the proof.

Next we show that the two steps algorithm of Elfving [2] is a special case of our algorithm.

**Algorithm Elfving (ELF).**

*Initialization:*

$$x^0 \in \mathcal{R}(A^T), \quad y^0 = b - Az^0, \quad \text{for some } z^0 \in \mathbb{R}^n. \quad (38)$$

*Iterative step:*

$$y^{k+1} = (I - A\Gamma)y^k, \quad (39)$$

$$x^{k+1} = Qx^k + R(b - y^{k+1}). \quad (40)$$

The matrices  $Q : n \times n$ ,  $R : n \times m$ ,  $\Gamma : m \times n$  satisfy

$$Q + RA = I, \quad (41)$$

and

$$\text{for } w \in \mathcal{R}(A), z = Qz + Rw \text{ if and only if } Az = w, \quad (42)$$

$$\rho((I - A\Gamma)P_{\mathcal{R}(A)}) < 1, \quad (43)$$

$$\rho(QP_{\mathcal{R}(A^T)}) < 1, \quad (44)$$

$$z \in \mathcal{N}(A^T) \implies \Gamma z = 0, \quad (45)$$

$$u \in \mathcal{R}(A) \implies Ru \in \mathcal{R}(A^T). \quad (46)$$

**Proposition 1.** (i) *The algorithm (ELF) is identical with algorithm (8), with the corrections  $v^k$  defined by*

$$v^k := -Ry^{k+1}. \quad (47)$$

(ii) *The assumptions (6), (15), (16) and (32) are satisfied.*

*Proof.* The assumptions (6) and (15) are identical with (41) and (44), respectively.

To show that assumption (32) holds we argue by mathematical induction, as in the proof of Proposition 3 in [4].

Assumption (16) holds because, from (47), and again as in the proof cited before, we get

$$v^k + Rb_A^* = -Ry^{k+1} + Rb_A^* = -R[(I - A\Gamma)P_{\mathcal{R}(A)}]^{k+1}(b_A).$$

According to the hypothesis (43) and Theorem 1, from Chapter 1 in [5], the matrix  $(I - A\Gamma)P_{\mathcal{R}(A)}$  is convergent, i.e.  $\lim_{k \rightarrow \infty} ((I - A\Gamma)P_{\mathcal{R}(A)})^k = 0$ , from which we get (16) and the proof is complete.

**Acknowledgements.** The authors wish to thank to the anonymous referees for their valuable comments that improved the initial version of the paper.

## References

- [1] Censor Y., S. A. Zenios *Parallel Optimization: Theory, Algorithms and Applications*, Oxford Univ. Press, New York, 1997.
- [2] Elfving T., *A stationary iterative pseudoinverse algorithm*, *B I T*, **38(2)** (1998), 275-282.
- [3] Horn R. A., Johnson C. R., *Matrix Analysis*, Cambridge University Press, New York, 1990.
- [4] Nicola A., Popa C. and Rude U., *Projection algorithms with correction*, *Journal of Applied Mathematics and Informatics*, **29 (3-4)** (2011), 697-712.
- [5] Popa C., *Projection Algorithms - Classical Results and Developments. Applications to Image Reconstruction*, Lambert Academic Publishing - AV Akademikerverlag GmbH & Co. KG, Saarbrucken, Germany, 2012.

Alexandru Bobe,  
Ovidius University of Constanta,  
Faculty of Mathematics and Informatics,  
Blvd. Mamaia 124, 900527, Constanta, Romania  
Email: alexb@univ-ovidius.ro

Aurelian Nicola,  
Ovidius University of Constanta,  
Faculty of Mathematics and Informatics,  
Blvd. Mamaia 124, 900527, Constanta, Romania  
Email: anicola@univ-ovidius.ro, aurelian.nicola@gmail.com

Constantin Popa,  
Ovidius University of Constanta,  
Faculty of Mathematics and Informatics,  
Blvd. Mamaia 124, 900527, Constanta, Romania  
Email: cpopa@univ-ovidius.ro, cpopa1956@gmail.com