THE PUBLISHING HOUSE
OF THE ROMANIAN ACADEMY

PROCEEDINGS OF THE ROMANIAN ACADEMY, Series A,
Volume 9, Number 1/2008, pp. 000–000

ON CIMMINO'S REFLECTION ALGORITHM

Constantin POPA

"Ovidius" University of Constanta, Romania, E-mail: cpopa@univ-ovidius.ro

Gianfranco Cimmino presented for the first time his famous reflection algorithm in 1938. He proved that, if the rank of the problem matrix is greater than one then, for any initial approximation the sequence generated by his algorithm converges to a solution of the normal equation associated to a perturbed (diagonally scaled) least-squares problem. Started from this result we construct a first extension of Cimmino's method which generates sequences of approximations of least-squares solutions for general inconsistent problems. Using some results of H. Keller (1965) and D. Young (1972), we show that the set of limit points of this extension completely characterizes the set of least-squares solutions. The second extension of Cimmino's algorithm is obtained by starting from a previous result of the author, concerning Jacobi's simultaneous projections algorithm. In this sense, we prove that a particular case of Cimmino's method can be considered as a particular case of Jacobi's method and that its limit points also completely characterize the least-squares solutions of the initial problem.

Key words: Cimmino's reflection algorithm, Jacobi's projections algorithm, inconsistent least-squares problems

1. THE ORIGINAL CIMMINO'S ALGORITHM

One year after S. Kaczmarz presented for the first time his famous successive projections algorithm in [6], G. Cimmino proposed in [3] his simultaneous reflections method. This uses instead of the orthogonal projections P_i on the hyperplanes generated by the system equations,

$$P_i = x - \frac{\langle x, a_i \rangle - b_i}{\|a_i\|^2} a_i, \tag{1}$$

orthogonal reflections S_i, given by

$$S_i(x) = x - 2\frac{\langle x, a_i \rangle - b_i}{\|a_i\|^2} a_i, \tag{2}$$

(we denoted by a_i, b_i the i-th row of the $m \times n$ system matrix A and the i-th component of the right hand side $b \in \boldsymbol{R}^m$ and by $\langle \cdot, \cdot \rangle$, $\| \cdot \|$, the Euclidean scalar product and the associated norm on some space $\boldsymbol{R}^q$). Cimmino's reflections S_i are simultaneously applied to an approximation $x^k \in \boldsymbol{R}^m$ and, a convex combination of them defines the next one x^{k+1} by

$$x^{k+1} = \frac{1}{\omega}\sum_{i=1}^{m} \omega_i S_i(x^k) = x^k + \frac{1}{\omega}\sum_{i=1}^{m} \omega_i (S_i(x^k) - x^k) = \\ x^k - \frac{2}{\omega}\sum_{i=1}^{m} \omega_i \frac{\langle x^k, a_i \rangle - b_i}{\|a_i\|^2} a_i, \tag{3}$$

where

$$\omega_i > 0 \,,\, \omega = \sum_{i=1}^{m} \omega_i \,. \tag{4}$$

In [3] Cimmino proved the following main convergence result concerning the algorithm (3)-(4).

Theorem 1. Let *A* and *b* be such that

$$a_i \neq 0 \,,\, \forall i = 1,\ldots,m \,,\, rank(A) \geq 2 \,, \tag{5}$$

and the system

$$Ax = b \,. \tag{6}$$

is consistent. Then, if the weights ω_i are as in (4), for any $x^0 \in \boldsymbol{R}^n$ the sequence $(x^k)_{k \geq 0}$ generated by (3) converges to a solution of (6).

Remark 1. The fact that the limit points of Cimmino's sequence (xk)k ≥ 0 from (4) (with respect to the initial approximation x0) "cover" the set S(A; b) of all solutions of (6) will be proved in the next section.

The "unpleasant" aspect of the above Cimmino's algorithm is that the sequence (xk)k ≥ 0 approximates a solution x only in the consistent case for (6). But unfortunately, "real world problems" usually give rise to inconsistent systems of the form (6), (see e.g. [5], [8], [9]), which must be reformulated as "linear least-squares problems": find x ∈ Rn such that

$$\|Ax - b\| = min\{\|Az - b\|, z \in R^n\} \tag{7}$$

In this general case, Cimmino's algorithm or some of its extensions still converge, but to solutions of "weighted" formulations of (7) (see e.g. [1], [2]). This is why we decided to analyse in this paper some possibilities to "extend" or "adapt" Cimmino's original method (3)-(4) to the more general problem (7), such that the sequence of approximations generated in this way, still converges to one of its solutions (similar with the results obtained by one of the authors in [10], but with respect to Kaczmarz-like projections algorithms). Moreover, we were also interested in the possibility of characterizing with these limit points, the set of all solutions of (7), denoted by *LSS*(*A*;*b*) (see also [11] and Remark 1 before). These versions of Cimmino's algorithm will be described in the next two sections of the paper.

2. THE FIRST EXTENSION

For the construction of our first version of the algorithm (3), we started from a remark of G. Cimmino, made in the original paper [3]. This result (based on some steps from the proof of Theorem 1) can be briefly described as follows.

Corollary 1. If he system (6) is not consistent and *A* verifies (5), for any $x^0 \in \boldsymbol{R}^n$, the sequence $(x^k)_{k \geq 0}$ generated by (3) converges to a solution *x* of the normal equation

$$\left(A^{\Omega}\right)^t A^{\Omega} x = \left(A^{\Omega}\right)^t b^{\Omega} \,, \tag{8}$$

where

$$\Omega = \left(\omega_1, \omega_2, \ldots, \omega_m\right) \in R^m \,, \tag{9}$$

$$A^{\Omega} = D_A^{\Omega} A \,;\; b^{\Omega} = D_A^{\Omega} b \tag{10}$$

and $D_A{}^{\Omega}$ is the diagonal $m \times m$ matrix given by

$$D_A^{\Omega} = diag\left(\frac{\sqrt{\omega_1}}{\|a_1\|}, \ldots, \frac{\sqrt{\omega_m}}{\|a_m\|}\right) \tag{11}$$

Remark 1. If the system (6) is consistent, the above Corollary 1 does not contradict the result in Theorem 1. Indeed, in this case and because D_A^Ω is invertible ($\omega_i > 0$), the (consistent) system (6) is equivalent with the (consistent) system

$$D_A^\Omega Ax = D_A^\Omega b\,, \tag{12}$$

which is equivalent with the normal equation (8). Starting from the above Corollary 1, we obtain the following version of Cimmino's algorithm (3) convergent to solutions of the general least-squares formulation (7).

Proposition 1. Let A be as in (5), $x^0 \in R^n$ an arbitrary initial approximation and the weights ω_i from (4) given by

$$\omega_i = \|a_i\|^2\,,\ i = 1,\dots,m\,. \tag{13}$$

Then, the sequence $(x^k)_{k\geq 0}$ generated by (3)-(4) converges to an element from $LSS(A;b)$.

Proof For ω_i as in (13) the matrix D_A^Ω from (11) is the identity, thus the normal equation (8) is identical with

$$A^t Ax = A^t b\,, \tag{14}$$

i.e. the normal equation associated to the problem (7). Then Corollary 1 applies and the proof is complete.
In the rest of this section we shall prove that the set of all limit points of the algorithm (3)-(4) with the weights choice (13) (which we shall denote by $LPC(A;b)$) coincides with $LSS(A;b)$. For this we shall briefly replay some constructions and results form papers [7] and [12]. Let B and N be $n \times n$ real matrices (with N invertible) and $d \in \boldsymbol{R}^n$ such that the system

$$Bx = d \tag{15}$$

is consistent. For approximating its solutions, we consider the iterative process: $x^0 \in \boldsymbol{R}^n$,

$$x^{k+1} = Tx^k + N^{-1}d,\ k \geq 0\,, \tag{16}$$

with

$$T = I - N^{-1}B\,. \tag{17}$$

The following result is proved in [7].

Theorem 2. Let us suppose that the matrix B is symmetric and nonnegative definite and consider the following splitting of it

$$B = D + M \tag{18}$$

with D symmetric and invertible. Let E be another invertible $n \times n$ matrix and P_E defined by

$$P_E = (E^{-1}D) + (E^{-1}D)^t - B\,. \tag{19}$$

Let also N be given by

$$N = N_E = E^{-1}D\,. \tag{20}$$

Then, for any $x^0 \in \boldsymbol{R}^n$ the method (16)-(17) is convergent to an element from $S(B;d)$ if and only if the matrix P_E is positive definite.
Let now $L(N;d)$ be the set of all limit points of the iteration (16)-(17) (with respect to $x^0 \in \boldsymbol{R}^n$). In the paper [12], the following result is proved.

Proposition 2. If the method (16)-(17) is convergent, then the following equality holds

$$L(N;d) = S(B;d)\,. \tag{21}$$

We are now able to prove the previously announced result.

Proposition 3. In the hypothesis of Proposition 1 the following equality holds

$$LPC(A;b) = LSS(A;b)\,. \tag{22}$$

Proof. Let the $n \times n$ matrix B and $d \in \boldsymbol{R}^n$ be defined by

$$B = A^t A,\ d = A^t b\,. \tag{23}$$

We then have (see e.g. [1])

$$LSS(A;b) = S(B;d) \tag{24}$$

Moreover, Cimmino's algorithm (3)-(4), with ω_i given by (13) becomes an algorithm from the class (16)-(17) if we define the matrices D and E, in the above Theorem 2 by

$$D = I,\ E = \frac{1}{\omega} I\ , \tag{25}$$

with ω given by

$$\omega = \frac{1}{2}\sum_{i=1}^{m} \|a_i\|^2 \tag{26}$$

Then (see (20))

$$N = \omega I \tag{27}$$

and the above Proposition 2, (24) and (27) tell us that

$$LPC(A;b) = L(N;d) = LSS(A;b) \tag{28}$$

if and only if the symmetric matrix $P=P_E$, associated to the above Cimmino's algorithm and given by

$$P_E = 2\omega I - B \tag{29}$$

is positive definite or equivalently

$$\omega > \frac{\rho(B)}{2}\,. \tag{30}$$

But, because B is symmetric and nonnegative definite and

$$trace(B) = \sum_{i=1}^{m} \|a_i\|^2\ , \tag{31}$$

for ω from (26) we always have

$$\omega \geq \frac{\rho(B)}{2}\,. \tag{32}$$

If we would have equality in (32) then, from the properties of the matrix B and (31) it would result that B would have only one nonzero eigenvalue, with (algebraic) multiplicity equal to 1. Then, $rank(A)=1$ which would contradict (5). Then (30) holds and the proof is complete.

3. THE SECOND EXTENSION

We shall start the presentation of this section by observing that the classical Cimmino’s algorithm (3) coincides, in the particular case

$$\omega_i = \gamma > 0,\ \forall i = 1,\ldots,m \tag{32}$$

with Jacobi's simultaneous projections method (see [4], [10])

$$x^{k+1} = x^k - \Omega \sum_{i=1}^{m} \frac{\langle x^k, a_i \rangle - b_i}{\|a_i\|^2} a_i . \tag{33}$$

Indeed, from (3)-(4) and (27) it results that

$$x^{k+1} = x^k - \frac{2}{m} \sum_{i=1}^{m} \frac{\langle x^k, a_i \rangle - b_i}{\|a_i\|^2} a_i , \tag{34}$$

which coincides with (28) for

$$\Omega = \Omega_0 = \frac{2}{m} . \tag{35}$$

The convergence condition for (28) (see again [4], [10]) is

$$0 < \Omega < \frac{2}{\rho(E)} , \tag{36}$$

with

$$E = \sum_{i=1}^{m} \frac{1}{\|a_i\|^2} (a_i a_i^t) . \tag{37}$$

Proposition 4. If the assumptions (5) hold for A, then he number Ω_0 from (35) satisfies (36).
Proof. Because of the symmetry of E, we successively get

$$\rho(E) = \|E\| \le \sum_{i=1}^{m} \frac{1}{\|a_i\|^2} \|a_i a_i^t\| = \sum_{i=1}^{m} \frac{1}{\|a_i\|^2} \|a_i^t a_i\| = m ,$$

which together with (35) gives us

$$\Omega_0 \le \frac{2}{\rho(E)} . \tag{38}$$

If the equality would hold in (38), we would have (see (35))

$$\rho(E) = m . \tag{39}$$

But, because the matrix E from (37) is nonnegative definite and symmetric, we know that $\sigma(E) \subset [\ 0, \infty)$, which means $\rho(E) \in \sigma(E)$. Let $x \in \boldsymbol{R}^n \setminus \{0\}$ be a corresponding eigenvector. We then successively obtain

$$m\|x^2\| = \langle Ex, x \rangle = \sum_{i=1}^{m} \frac{\langle x, a_i \rangle^2}{\|a_i\|^2} \le \sum_{i=1}^{m} \frac{\|x\|^2 \|a_i\|^2}{\|a_i\|^2} = m\|x^2\| . \tag{40}$$

From (40) and the Cauchy-Schwarz inequality

$$\langle x, a_i \rangle^2 \le \|x\|^2 \|a_i\|^2 ,\ \forall i = 1, \ldots, m , \tag{41}$$

we obtain the equalities in (41), which means that x is collinear with a_i, $\forall\ i=1,\ldots, m$. But, this would mean that

$$rank(A) \le 1 , \tag{42}$$

which would contradict (5). Thus, strict inequality holds in (38) and the proof is complete.
Let now φ_j, $\Phi(\alpha; \cdot)$ be defined by (see [10])

$$\varphi_j(y) = \frac{\langle y, \alpha_j \rangle - b_i}{\|\alpha_j\|^2} \alpha_j, \; j = 1, \ldots, n, \tag{43}$$

$$\Phi(\alpha; y) = y - \alpha \sum_{j=1}^{n} \varphi_j(y), \; y \in R^m, \tag{44}$$

where $\alpha_j \neq 0$ is the j-th column of A and $\alpha \in \boldsymbol{R}$, $\alpha \neq 0$. Let D be the $n \times n$ the matrix defined by

$$D = \sum_{j=1}^{n} \frac{\alpha_j \alpha_j^t}{\|\alpha_j\|^2}. \tag{45}$$

We then consider the following **C**immino **E**xtended (**CE**, for short) algorithm.

ALGORITHM CE.

Let $y^0 = b$, $x^0 \in \boldsymbol{R}^n$ and x^k an already computed approximation. The next one, x^{k+1} is given by

$$y^{k+1} = \Phi(\alpha; y^k), \tag{46}$$

$$b^{k+1} = b - y^{k+1}, \tag{47}$$

$$x^{k+1} = x^k - \frac{2}{m} \sum_{i=1}^{m} \frac{\langle x^k, a_i \rangle - b_i^{k+1}}{\|a_i\|^2} a_i, \tag{48}$$

where by b_i^{k+1} we denoted the i-th component of the vector b^{k+1} from (47). Then, the following result holds (as in Theorem 6.7 from [10]).

Theorem 3. For any matrix A satisfying (5) and $\alpha_j \neq 0$, j=1,…, n, any vector $b \in \boldsymbol{R}^m$, any initial approximation $x^0 \in \boldsymbol{R}^n$ and any α such that

$$\alpha \in \left(0, \frac{2}{\rho(D)}\right)$$

the sequence $(x^k)_{k \geq 0}$ generated with the algorithm (46)-(48) converges to an element $x \in LSS(A;b)$. Conversely, any element $x \in LSS(A;b)$ can be obtained as the limit point of such a sequence, for an appropriate choice of x^0. Moreover, for x^0 in the range of A^t, the sequence $(x^k)_{k \geq 0}$ converges to the minimal norm solution of the problem (7).

REFERENCES

1. BJÖRK, A., *Numerical methods for least squares problems*, SIAM Philadelphia, 1996.
2. CENSOR, Y., ELFVING, T., *Block-iterative algorithms with diagonally scaled oblique projections for the linear feasibility problem*, SIAM Matrix Anal. and Appl., **24**(2002), 40-58.
3. CIMMINO, G., *Calcolo approssiomatto per le soluzioni dei sistemi di equazioni lineari*, Ric. Sci. progr. tecn. econom. naz., **1**, pp. 326-333, 1938.
4. ELFVING, T., *Block-Iterative Methods for Consistent and Inconsistent Linear Equations*, Numer. Math., **35**, pp. 1-12, 1980.
5. GROETSCH, C.W., *Inverse problems in the mathematical sciences*, Vieweg, Wiesbaden, Germany, 1993
6. KACZMARZ, S., *Angenaherte Auflosung von Systemen lineare Gleichungen*, Bull. Acad. Polonaise Sci. et Lettres., **A**, pp. 355-357, 1937.
7. KELLER, H., *On the solution of singular and semidefinite linear systems by iteration*, SIAM J. Numer. Anal., Ser. B, **2(2)**, pp. 281-290, 1965.
8. MOHR, M., POPA, C., RUEDE, U., *An experimental analysis of a differential inverse problem*, Technical Report **99-1**, IMMD10, FAU Erlangen-Nurnberg, Germany, 1999.
9. MOHR, M., *Comparison of solvers for a bioelectric field problem*, Technical Report **00-1**, IMMD10, FAU Erlangen-Nurnberg, Germany, 2001.

10. POPA, C., Extensions of block-projections methods with relaxation parameters to inconsistent and rank-deficient least-squares *problems*, BIT, **38(1)**, pp. 151-176, 1998.
11. POPA, C., *Characterization of the solutions set of inconsistent least-squares problems by an extended Kaczmarz algorithm*, Korean Journal on Comp. and Appl. Math., **6(1)**, pp. 51-64, 1999.
12. YOUNG, D.M., *On the consistency of linear stationary iterative methods*, SIAM J. Numer. Anal., **9(1)**, pp. 89-96, 1972.

Received December 18, 2007