ACTA MATHEMATICA VIETNAMICA
Volume 34, Number 3, 2009, pp. 335–343

AN ALTERNATING PROJECTIONS ALGORITHM FOR SOLVING LINEAR PROGRAMS

PHAM CANH DUONG AND LE THANH HUE

Abstract. The method of alternating projections was first introduced by Von Neumann in 1933 for finding the projection of a given point onto the intersection of closed subspaces of a Hilbert space. Since its introduction this method has received considerable attention and has found application in many areas of mathematics and physics as well as in other fields of science and engineering. In this paper we show that recent results of the method of alternating projections for general convex feasibility problems may be used to construct a simple algorithm for solving linear programs. This new algorithm is simple to implement, computationally stable and is inherently parallel.

1. Introduction

The *convex feasibility problem* (CFP) is to find a point x^* in the intersection C of finitely many closed, convex sets C_i in a Hilbert space $\mathcal{H}$. This problem appears in various areas of mathematics and physical sciences such as approximation theory, signal and image reconstruction, medical imaging. The problem may be solved by using a simple and powerful algorithm, called the method of *successive orthogonal projections* (SOP), when an iterate sequence is generated by projecting cyclically onto the constraint sets C_i, which was first introduced by Von Neumann, back to 1933, in [13]. Due to its extraordinary utility and broad applicability in many areas, the convex feasibility problem and the method of successive orthogonal projections continue to receive great attention. A large number of research papers with new results in this area are discussed in excellent survey papers by Deutch [7, 8], Baushke et al. [1], Combettes [4], Censor [5, 6]. In this paper we will show that some new results on the behavior of iterate sequences generated by this kind of projection algorithms when applied to the linear feasibility problems may be used to derive a simple and efficient algorithm for solving linear programs approximately. In the next Section 2, we present some basic results on the convex feasibility problems and the related method of successive orthogonal projections. In Section 3, we focus on the behavior of the method of SOP when applied to a system of linear inequalities. Section 4 is devoted to

Received April 14, 2008; in revised form August 19, 2008.
2000 *Mathematics Subject Classification.* 90C05, 90C59, 65K05, 65F10.
Key words and phrases. Linear program, successive orthogonal projections algorithm, linear convergence, convex feasibility problem.

describe our new algorithm for solving linear programs. This algorithm is based upon some new convergence properties of SOP presented in Section 3. Finally, a brief concluding remark and some preliminary numerical results are presented in Section 5.

2. Preliminaries

One of the basic tasks in many computational processes is to find a point satisfying some given set of equations and/or inequalities. An important case where the solution set and the inequalities involved are convex is referred to as *convex feasibility problem* and is formulated as

$$\text{Find } x^* \in C = \bigcap_{i\in I} C_i \tag{2.1}$$

where for each $i \in I = \{1, 2, \ldots, m\}$, C_i is a closed, convex set in a Hilbert space $\mathcal{H}$. In this paper, we will restrict ourself to the case where $\mathcal{H}$ is a finite-dimensional Euclidean space $\mathbb{R}^n$. The successive orthogonal projections (SOP) method is the following. Begin with an arbitrary $x^{(0)}$, let

$$(\forall k \in \mathbb{N}) \quad x^{(k+1)} = P_{(k \bmod m)+1}\, x^{(k)}, \tag{2.2}$$

where $P_i\, x$ denotes the orthogonal projection of x onto the set C_i. In the literature this method is also called the *alternating projections* or *the cyclic projections method*. If problem (2.1) is consistent, i.e. $C \neq \emptyset$, the convergence of sequence (2.2) is established by the following

Theorem 1 (Bregman, [3]). *Suppose $C = \bigcap_{i=1}^m C_i \neq \emptyset$. Then sequence $\{x^{(k)}\}$ converges (weakly) to some point in C.*

If problem (2.1) is inconsistent, that is $C = \emptyset$, the SOP method (2.2) is still employed. Its convergence properties were studied in [8, 11], and recently in [9]. First, let us reindex the iterate sequence $\{x^{(k)}\}$ as follows. Every index $k \geq 0$ can be expressed in the form $k = ms + i$ with $0 \leq i < m$. We then use s and i to index $x^{(k)}$ as

$$x^{(k)} = x_i^{(s)}\,,\ k = ms + (i-1)\,,\ 0 < i \leq m\,. \tag{2.3}$$

So we can write

$$\left\{x^{(k)}\right\}_{k\geq 0} = \left\{\left\{x_i^{(s)} \,|\, i = 1, 2, \ldots, m\right\}\right\}_{s\geq 0}\,.$$

This way of indexing divides the sequence $\{x^{(k)}\}$ into equally sized segments $\left\{x_i^{(s)} \,|\, i = 1, 2, \ldots, m\right\}$, $s \geq 0$, which are called *projection cycles* or, simply *cycles*. It is easily seen that, for every $i \in I$, the subsequences $\{x_i^{(s)}\}_{s\geq 0}$ contains all points from $\{x^{(k)}\}$ which belongs to C_i. Those subsequences are always convergent if system (2.1) is convex.

Theorem 2 (Gubin, Polyak and Raik, [10]). *Let $\{x^{(k)}\}$ be any sequence generated by (2.2). Then there exist points $\{\bar{x}_i\}_{1\leq i\leq m}$ such that $P_1\,\bar{x}_m = \bar{x}_1$ and $P_i\,\bar{x}_{i-1} = \bar{x}_i$ for every $i \in \{2, 3, \ldots, m\}$. Moreover, for every $i \in \{1, 2, \ldots, m\}$, the periodic subsequence $\{x_i^{(s)}\}_{s\geqslant 0}$ converges to such a point $\bar{x}_i \in C_i$.*

The set of limit points $\{\bar{x}_i\}_{1\leq i\leq m}$ is called *limit cycle* of $\{x^{(k)}\}$. Moreover, in the important special cases of (2.1), when all the C_i's are affine spaces or affine half-spaces, the iterate sequences $\{x^{(k)}\}$ and $\{x_i^{(s)}\}$, mentioned in Theorems 1 and 2, are known to be linearly convergent with the rate independent of the starting point (Deutch [8], Bauschke et al. [2]). We recall that a sequence $\{x^{(k)}\}$ is said to *converge linearly* to its limit x^* if there exists a natural number $h > 0$ and a number $0 < c < 1$ such that

$$(\forall k > h) \quad ||x^{(k+1)} - x^{(k)}|| \ \leqslant\ c||x^{(k)} - x^{(k-1)}|| \tag{2.4}$$

An alternative definition is that

$$(\forall k \in \mathbb{N}) \quad ||x^{(k+1)} - x^*|| \ \leqslant\ c||x^{(k)} - x^*|| \tag{2.5}$$

It is interesting to note that the linear convergence of the iterate sequences is proved for more general setting in the inconsistent case of system (2.1)(see Bauschke et al.[2], Fact 5.5.7), in particular, in $\mathbb{R}^n$ it is true for every finite family of closed convex sets $\{C_i\}$ with empty intersection. However, if (2.1) is consistent, it must be assumed to be linear to guarantee the linear convergence of $\{x^{(k)}\}$ ([1], Theorem 5.7). Most known alternating projection type algorithms required some kind of linearity assumption. In the next section, we will focus on the behavior of the method of SOP when applying to the linear feasibility problems.

3. Convergence properties for linear feasibility problem

For the remainder of the paper we shall assume that problem (2.1) is linear, that is all C_i's are affine half-spaces in $\mathbb{R}^n$. If it is consistent then by virtue of Theorem 1 and Theorem 5.7 in [1] the iterate sequence $\{x^{(k)}\}_{k\geq 0}$ converges linearly to some point $x^* \in C = \bigcap_{i\in I} C_i$ with the rate $0 < c < 1$. Applying (2.5) recursively to $x^{(k)}$, $x^{(k-1)}$,...,$x^{(k-m)}$ implies

$$(\forall k \in \mathbb{N}) \quad ||\, x^{(k)} - x^*\, || \ \leqslant\ c^m\, ||\, x^{(k-m)} - x^*\, ||\,.$$

So, setting $\rho = c^m$, for every $i \in I$ we have

$$(\forall s \in \mathbb{N}) \quad ||\, x_i^{(s)} - x^*\, || \ \leqslant\ \rho\, ||\, x_i^{(s-1)} - x^*\, ||\,. \tag{3.1}$$

This means that all subsequences $\{x_i^{(s)}\}_{s\geq 0}$, $i = 1, 2, \ldots, m$, converge linearly to x^* with the rate ρ.

Now, given a tolerance $\varepsilon > 0$, let's denote by p the number of cycles necessary to ensure that for every $i \in I$ the points $x_i^{(s>p)}$ are close to x^* by ε, i.e.

$$(\forall i \in I)\,(\forall s \in \mathbb{N},\, s > p) \quad ||\, x_i^{(s)} - x^* \,|| \;<\; \varepsilon\,. \tag{3.2}$$

Such a number p can be estimated as follows. From (3.1) we deduce

$$(\forall s \in \mathbb{N}) \quad ||\, x_i^{(s)} - x^* \,|| \;\leqslant\; \rho^s \,||\, x_i^{(0)} - x^* \,||\,.$$

Thus p can be determined from the following inequality

$$\rho^p \,||\, x_i^{(0)} - x^* \,|| \;<\; \varepsilon.$$

Setting $\varepsilon' = \varepsilon \big/ ||x_i^{(0)} - x^*||$ and noting that $\rho < 1$ gives

$$p \;\geqslant\; \frac{\ln(\varepsilon')}{\ln(\rho)}\,.$$

Consequently, the number of cycles necessary to guarantee (3.2) can be taken as

$$p \;=\; \left[\frac{\ln(\varepsilon')}{\ln(\rho)}\right] + 1\,, \tag{3.3}$$

where [ ] denotes the integer-part function.

Next, we will examine in more details the behavior of sequences $\{x_i^{(s)}\}_{s\geq 0}$ near their limits. Let us define

$$(\forall s \in \mathbb{N})\;(\forall i \in I) \quad h_i^{(s)} \;\triangleq\; ||x_{(i+1) \bmod m}^{(s)} - x_i^{(s)}||;$$
$$(\forall s \in \mathbb{N})\;(\forall i \in I) \quad d_i^{(s)} \;\triangleq\; (x_i^{(s)} - x_i^{(s-1)}).$$

Proposition 1. *Suppose (2.1) is linear and consistent. Then:*

(i) $(\forall i \in I) \quad \lim\limits_{s\to\infty} h_i^{(s)} \;=\; 0.$

(ii) $(\forall i \in I) \quad \lim\limits_{s\to\infty} \frac{h_i^{(s)}}{||d_i^{(s)}||} \;=\; a_i \;<\; +\infty.$

Proof. (i) is obvious by Theorem 1 and the definition of linear convergence. (ii): First, we show that for all $i \in I$, $\{h_i^{(s)}\}_{s\geqslant 0}$ converges linearly to 0 with the same rate as $\{x_i^{(s)}\}$. Indeed, by Theorem 1, $\{x^{(k)}\}$ and, consequently, all subsequences $\{x_i^{(s)}\}$,$i \in I$, converge linearly to some point $x^* \in \bigcap_{i\in I} C_i$ with the convergence rate $0 < c < 1$. By the definition of $h_i^{(s)}$, the triangle inequality implies

$$h_i^{(s)} \;\leqslant\; ||x_i^{(s)} - x^*|| \;+\; ||x_{i+1}^{(s)} - x^*||.$$

By the definition of linear convergence, it follows that

$$h_i^{(s)} \;\leqslant\; c^s.(||x_i^{(0)} - x^*|| \;+\; ||x_{i+1}^{(0)} - x^*||). \tag{3.4}$$

Since the amount in the bracket is a positive constant, (3.4) means that for all $i \in I$, $\{h_i^{(s)}\}_{s\geqslant 0}$ converges linearly to 0 with the rate c.

Next, using the definition of $h_i^{(s)}$ and $d_i^{(s)}$ we can write

$$h_i^{(s)} \;=\; ||(x_{i+1}^{(s-1)} + d_{i+1}^{(s)}) - (x_i^{(s-1)} + d_i^{(s)})||.$$

The triangle inequality implies

$$h_i^{(s)} \;=\; ||(x_{i+1}^{(s-1)} - x_i^{(s-1)}) + (d_{i+1}^{(s)} - d_i^{(s)})|| \;\geqslant\; |h_i^{(s-1)} - ||d_{i+1}^{(s)} - d_i^{(s)}||\,|.$$

Dividing both sides of this inequality by $h_i^{(s-1)} > 0$ gives

$$\frac{h_i^{(s)}}{h_i^{(s-1)}} \;\geqslant\; |1 - \frac{||d_{i+1}^{(s)} - d_i^{(s)}||}{h_i^{(s-1)}}\,|. \tag{3.5}$$

Since $d_{i+1}^{(s)} = P_i d_i^{(s)}$ and the projection operator is non-expansive, it follows that $||d_{i+1}^{(s)}|| \leqslant ||d_i^{(s)}||$. Hence, the second term in the right-hand side of (3.5) can be estimated as

$$\frac{||d_{i+1}^{(s)} - d_i^{(s)}||}{h_i^{(s-1)}} \;\leqslant\; \frac{||d_{i+1}^{(s)}|| + ||d_i^{(s)}||}{h_i^{(s-1)}} \;\leqslant\; \frac{2||d_i^{(s)}||}{h_i^{(s-1)}} \;\leqslant\; \frac{2||d_i^{(s)}||}{h_i^{(s)}}.$$

Now, if we suppose to the contrary, (ii) does not hold, this term would tend to 0 as $s \to \infty$ which makes (3.5) contradict to (3.4) for sufficiently large s. The proof is complete. □

The following proposition is an immediate consequence of Theorem 2 and the fact that $\{x_i^{(s)}\}_{s\geqslant 0}$ are linearly convergent.

Proposition 2. *Suppose* (2.1) *is linear and inconsistent. Then*
$$(\forall i \in I) \quad \lim_{s\to\infty} \frac{h_i^{(s)}}{||d_i^{(s)}||} \;=\; +\infty.$$

Moreover, it is easy to see that the sequence $\{\frac{h_i^{(s)}}{||d_i^{(s)}||}\}$ grows to $+\infty$ as fast as $\{\frac{1}{\rho^s}\}$.

Next, let us define

$$(\forall i \in I)\,(\forall s \in \mathbb{N}) \quad D_i^{(s)} \;\triangleq\; |\frac{h_i^{(s+1)}}{||d_i^{(s+1)}||} \;-\; \frac{h_i^{(s)}}{||d_i^{(s)}||}|.$$

By Propositions 1 and 2 we obtain the following

$$(\forall i \in I) \quad \lim_{s\to\infty} D_i^{(s)} \;=\; 0, \quad \text{if } C \neq \emptyset, \tag{3.6}$$

$$(\forall i \in I) \quad \lim_{s\to\infty} D_i^{(s)} \;=\; \infty, \quad \text{if } C = \emptyset. \tag{3.7}$$

These relationships will be used as stopping rules in the proposed algorithm for solving approximately linear programs.

4. Description of the algorithm

Let's consider a linear program of the form

$$(4.1)\qquad \begin{cases} \min \; < \mathbf{c}, x > \\ s.t. \quad x \in \mathbf{M} \;\triangleq\; \{x \in \mathbb{R}^n \,| \mathbf{A}x \;\leqslant\; \mathbf{b}\}\,. \end{cases}$$

where $\mathbf{A}$ is a $m \times n$-matrix and $\mathbf{b}$ is a vector in $\mathbb{R}^n$. For the simplicity of explanation we assume the following

(1) $\mathbf{A}x \;\leqslant\; \mathbf{b}$ is consistent;
(2) (4.1) has a finite solution x^*.

For every $i \in I$, let's denote by $\mathbf{M}_i$ the affine halfspace defined by the i-th row of the constraint system, i.e.

$$\mathbf{M}_i \;\triangleq\; \{x \in \mathbb{R}^n \mid < \mathbf{A}_i, x > \leqslant\; \mathbf{b}_i\}.$$

We will propose an algorithm for solving (4.1) approximately in the following sense. Given a tolerance $\varepsilon > 0$, it uses the algorithm of SOP (2.2) to find an approximate solution x^*_ε to the system (4.1) that satisfies

$$d(x^*_\varepsilon\,, \mathbf{M}_i) \;<\; \varepsilon \quad (\forall i \in I);$$
$$|\; < \mathbf{c}, x^*_\varepsilon > \;-\; < \mathbf{c}, x^* > \;|\; < \; \varepsilon\,.$$

In system (2.1) let's set

$$(4.2)\qquad \begin{aligned} (\forall i \in I) \quad C_i \;&=\; \{x \mid \; < \mathbf{A}_i\,, x > \leqslant\; \mathbf{b}_i\,\}\,; \\ C_{m+1} \;&=\; \{x \mid \; < \mathbf{c}\,, x > \leqslant\; v\,\}\,, \end{aligned}$$

where v is some real number. We will say that (4.2) is ε-consistent if there exists $x \in \mathbb{R}^n$ such that $d(x, \mathbf{M}_i) \;<\; \varepsilon$ for all $i \in I$, where $d(x, \mathbf{M}_i)$ denotes the usual euclidean distance from x to the set $\mathbf{M}_i$.

Given a tolerance ε and an arbitrary point $x^{(0)}$, by Proposition 1 , (3.2), (3.3) and (3.7) one can see that the algorithm of SOP, applied to system (4.2), will give $||d^{(s)}_{m+1}|| \;<\; \varepsilon$ with either $D_{m+1} \;\approx\; 0$ (the system is ε-consistent) or $D_{m+1} \;>\; 1$ (the system is inconsistent), after a finite number p of projection cycles. So for every real number v the algorithm SOP may be used to detect whether (4.2) is ε-consistent or not. This fact is used in the following algorithm.

The Algorithm:

1. Initialization

a. Choose the tolerance $\varepsilon > 0$;
b. Pick an arbitrary starting point $x^{(0)}$;
c. Choose numbers v_0^+ and v_0^- such that system (4.2) is consistent with $v \;=\; v_0^+$ and inconsistent with $v \;=\; v_0^-$;
d. Set $v \;=\; (v_0^+ \;+\; v_0^-)/2$, $k \;=\; 1$.

2. Step k

while $(||d^{(.)}_{m+1}|| \;>\; \varepsilon)$

Apply the algorithm SOP to (4.2);
endwhile
if $(D^{(.)}_{m+1} < 1)$ **then**
(4.2) is ε-consistent.
Set $v^+_{k+1} = (v^+_k + v^-_k)/2$, $v^-_{k+1} = v^-_k$;
else $(D^{(.)}_{m+1} > 1)$
(4.2) is inconsistent.
Set $v^-_{k+1} = (v^+_k + v^-_k)/2$, $v^+_{k+1} = v^+_k$;
endif
if $((v^+_k - v^-_k) < \varepsilon)$ **then goto Stop**;
endif
Set $v_{k+1} = (v^+_{k+1} + v^-_{k+1})/2$, $k \leftarrow k+1$;
goto Step k+1;

3. Stop ($x^*_\varepsilon =$ current $x^{(s)}_{m+1}$; appox. obj. Value = current v_k).

We note that the initial value of v^+_0 and v^-_0, mentioned in the initialization step, are often known beforehand for many practical problems. v^+_0 and v^-_0 may also be obtained using SOP in a similar way as described earlier in step k.

Complexity analysis. Let $\{x^{(k)}\}$ be the sequence of projection points generated by the algorithm SOP at j-th iteration. As explained in Section 2, this sequence consists of a number of projection cycles. Each such a cycle requires m n-dimensional vector-to-vector multiplications which is equivalent to nm floating point multiplications (the number of additions can be ignored). The number of cycles in one iteration, obtained earlier in (3.3), is bounded. So, the complexity of one iteration of the proposed algorithm is of order $O(nm)$. The number of iterations, denoted by q, is estimated using the following inequality

$$\frac{| v^0_+ - v^0_- |}{2^q} \leq \varepsilon.$$

We see that q does not depend on n and m. Therefore, the overall complexity of our algorithm is $O(nm)$.

Remark 1. It is easy to determine the number of multiplications needed by one iteration of Dantzig's original simplex algorithm and its revised form. It is about $(n+1)(m+1)$ for the primal simplex method and is $nm+(m+1)^2$ for the revised simplex algorithm (see for ex. [12]). However, no good estimate for the number of iterations has been found so far. From a large number of experiments it is believed that this number grows linearly with m, the number of constraints, and is some where between $3m$ and $10m$. Comparing this with the complexity result obtained above shows that our algorithm may outperform both of them in finding approximate solutions, especially for systems with large number of constraints. However, it is worthy to note that this algorithm can only provide us with an ε-feasible optimal solution which may not be feasible to the input system. Therefore, if an exact solution is needed the simplex method is likely a better choice.

5. Numerical experiments

The proposed algorithm is simple to implement and requires little system resources. The amount of memory needed is nearly minimal and is used mainly for storing all entries in matrix $\mathbf{A}$ and vectors $\mathbf{b}$ and $\mathbf{c}$. This algorithm uses mainly vector-to-vector multiplications and do not alter the input data during calculation. So it may be directly used for solving sparse systems and, in this case, may be implemented in small, resource-constrained systems. Another advantage of the proposed algorithm is its computational stability. We have done some numerical experiments, using MATLAB, on a number of test problems of small and moderate size. The performance of this algorithm is quite satisfactory, especially for the problems with $m \gg n$. As it is based on a basic form of SOP, it may converge slowly for some set of data. The speed of convergence may be notably improved by using accelerated variants of SOP (see, for ex. in [6, 9] or recently, in [14]).

References

[1] H. H. Bauschke and J. M. Borwein, On projection algorithms for solving convex feasibility problems, *SIAM Rev.* **38** (1996), 367–426.

[2] H. H. Bauschke, J. M. Borwein and A. S. Lewis, *On the method of cyclic projections for convex sets in Hilbert space*, Research report, Simon Fraser University, 1994.

[3] L. M. Bregman, The method of successive projection for finding a common point of convex sets, *Soviet Mathematics - Doklady* **6** (3) (1965), 688-692.

[4] P. L. Combettes, *The Convex Feasibility Problem in Image Recovery*, volume 97 of Advances in Imaging and Electron Physics. Academic Press, New York, 1995.

[5] Y. Censor, Rowaction methods for huge and sparse systems and their applications, *SIAM Rev.* **23** (4) (1981), 444–466.

[6] Y. Censor, On sequential and parallel projection algorithms for feasibility and optimization (Keynote Paper), In: *Visualization and Optimization Techniques,* Y. Censor and M. Ding (Editors), Proceedings of SPIE Volume 4553, SPIE-The Society for Optical Engineering, Bellingham, WA, USA, 1-9, 2001.

[7] F. Deutsch, The method of alternating orthogonal projections, In: *Approximation Theory, Spline Functions and Applications*, (S. P. Singh, editor). The Netherlands: Kluwer, 105 -121, 1992.

[8] F. Deutsch, Rate of convergence of the method of alternating projections, In: *Parametric optimization and approximation* (B. Brosowski and F. Deutsch, editors). International Series of Numerical Mathematics Vol 72. 96 -107. Birkhauser, 1983.

[9] A. Galantai, On the rate of convergence of the alternating projection method in finite dimensional spaces, *J. Math. Anal. Appl.* **310** (1) (2005), 30–44.

[10] L. G. Gubin, B. T. Polyak and E. V. Raik, The method of projections for finding the common point of convex sets, *USSR Computational Mathematics and Mathematical Physics* **7** (1967), 1–24.

[11] J. Mandel, Convergence of the cyclic relaxation method for linear inequalities, *Math. Program.* **30** (1984), 218–228.

[12] E. D. Nering and A. W. Tucker, *Linear Programs and Related Problems*, Academic Press, 1993.

[13] J. Von Neumann, On rings of operators. Reduction theory, *Annals of Math.* **50** (1949), 401–485 (the result of interest first appeared in 1933 in lecture notes).

[14] Thomas, Roman, *A randomized Kaczmarz algorithm with exponential convergence*, eprint arXiv:math/0702226, 2007, submitted.

Department of Software Development and Applied Mathematics
Institut of Mathematics, Vietnamese Academy of Science and Technology
18 Hoang Quoc Viet, Hanoi, Vietnam

E-mail address: pcduong@math.ac.vn

Faculty of Informatic Technology
University of Mining and Geology, Hanoi, Vietnam

E-mail address: huelt63@yahoo.com