

GLOBAL BEHAVIOR OF PARALLEL PROJECTION METHOD FOR CERTAIN NONCONVEX FEASIBILITY PROBLEMS

Masao Yamagishi

Isao Yamada

Department of Communications and Computer Engineering
Tokyo Institute of Technology, 2-12-1-S3-60 Ookayama, Meguro-ku, Tokyo 152-8550, Japan
Email: {myamagi, isao}@sp.ce.titech.ac.jp

ABSTRACT

Finding a common point of multiple closed sets in a real Hilbert space has been an important task in a wide range of signal processing. In this paper, we study asymptotic properties of the parallel projection method (PPM) for closed sets satisfying a special feasibility condition, which holds in the context of certain sparse signal processing. Our analysis guarantees that the cluster point set of PPM is exactly the intersection of the closed sets, and the distance to each set along a sequence generated by PPM with arbitrary initial point converges to zero. Moreover, under certain additional assumptions, we prove that the sequence converges to a point in the intersection of the closed sets, while existing analyses gave only local behaviors of PPM.

Index Terms— Nonconvex feasibility problem, parallel projection method, convergence analysis

1. INTRODUCTION

Finding a common point of multiple closed sets in a real Hilbert space is an important task in wide range of signal and image processing [1, 2, 3]. Let \mathcal{H} be a finite dimensional real Hilbert space and $S_i (\subset \mathcal{H})$ ($i = 1, 2, \dots, p$) closed sets satisfying $\bigcap_{i=1}^p S_i \neq \emptyset$. Then, the standard feasibility problem is formulated as

$$\text{Find some point } x_* \in \bigcap_{i=1}^p S_i, \quad (1)$$

where each S_i ($i = 1, 2, \dots, p$) is associated with a certain desired property of a vector x to be estimated. One of the most successful iterative approaches is known as the parallel projection method (PPM) that updates a point $x^{(k)} \in \mathcal{H}$ by using a weighted average of selections $y_i^{(k)}$ ($i = 1, 2, \dots, p$) in the metric projections P_{S_i} of $x^{(k)}$ onto every S_i , i.e.,

$$y_i^{(k)} \in P_{S_i}(x^{(k)}) := \arg \min_{y \in S_i} \|y - x^{(k)}\| \subset \mathcal{H}, \quad (2)$$

where $P_{S_i}(x^{(k)}) \neq \emptyset$ is assumed (Note: this condition holds automatically by Weierstrass's theorem if dimension of \mathcal{H} is finite, i.e., $\dim(\mathcal{H}) < \infty$). For a special case where all S_i s are hyperplanes, Cimmino showed that a sequence generated by the PPM converges to the best approximation, from $x^{(0)}$, to the intersection, i.e., the projection of $x^{(0)}$ onto the intersection in this special case [4] (see also Cimmino's history [5]). If all S_i s are closed affine subspaces (or linear variety), this best approximation property of the PPM holds

This work was supported in part by JSPS Grants-in-Aid (26730128).

PPM:Parallel projection method

The initial point $x^{(0)} \in \mathcal{H}$, convex weights $(w_i)_{i=1}^p$ s.t. $w_i > 0$ and $\sum_{i=1}^p w_i = 1$, and closed sets $S_1, S_2, \dots, S_p \subset \mathcal{H}$ are given.

Repeat

Step 1. Select $y_i^{(k)} \in P_{S_i}(x^{(k)})$ ($i = 1, 2, \dots, p$).

Step 2. Compute $x^{(k+1)} = \sum_{i=1}^p w_i y_i^{(k)}$.

until $x^{(k)}$ converges.

also true. More generally, if all S_i s are assumed to be closed convex sets, convergence of the PPM to a point in the intersection has been guaranteed, see, e.g., [6, 7, 8, 9, 10, 11].

Meanwhile, there are many examples of nonconvex sets associated with useful properties in the recent advance of sparsity-aware signal processing, e.g., the lower level set of ℓ_0 pseudo norm function (see Example 5) and the lower level set of rank function of matrices (see Example 6) (see e.g. [12, 13] for applications of the ℓ_0 pseudo norm and rank function). Sparsity implies that few components of signal are nonzero. The ℓ_0 pseudo norm of a vector is the number of nonzero components of the vector. The rank of a matrix plays also important roles in sparsity-aware signal processing because it is the number of nonzero singular values of the matrix.

Unfortunately, if at least one of S_1, S_2, \dots, S_p is not convex, convergence analysis of the PPM regardless of the initial point has not been established so far. Although nonconvex projection methods to solve (1) has been investigated and guaranteed their local convergences [14, 15, 16, 17, 18, 19], their analyses require impractical initial points. Certain global properties of a nonconvex projection method were investigated in the case where specific nonconvex sets in [20]. However, it does not establish convergence of the sequence generated by the method. Consequently, establishing convergence analysis regardless of initial point of the PPM is required from the perspective of the signal processing application of the PPM.

In this paper, to establish global convergence of PPM to certain nonconvex feasibility problems, e.g., in sparsity-aware signal processing, we study asymptotic properties of PPM for closed sets S_1, S_2, \dots, S_p satisfying the following special feasibility condition

$$\bigcap_{i=1}^p \text{Inc}(S_i) \neq \emptyset, \quad (3)$$

where $\text{Inc}(S)$ is defined for a nonempty closed set $S \subset \mathcal{H}$ by

$$\text{Inc}(S) := \bigcap_{\substack{x \in \mathcal{H} \\ y \in P_S(x)}} \left\{ z \in S \mid \langle x - y, z - y \rangle \leq 0 \right\}. \quad (4)$$

The closed convexity of $\text{Inc}(S)$ is guaranteed because it is the intersection of closed half spaces. Fortunately, we can verify easily that $\text{Inc}(S)$ is nonempty for important closed sets appearing sparsity-aware signal processing (see Sect. 4 and Sect. 5).

Thanks to the special feasibility condition (3), we present two global asymptotic properties of PPM:

1. the distance to each S_i ($i = 1, 2, \dots, p$) converges to zero (Theorem 1),
2. PPM under the condition (3) never falls into any trap outside the intersection $\bigcap_{i=1}^p S_i$ because the set of all possible cluster points of PPM¹, say $\mathfrak{C}_{\text{PPM}}$, is exactly the intersection $\bigcap_{i=1}^p S_i$ (Theorem 2), where $z_* \in \mathcal{H}$ is called a cluster point of PPM if

$$(\exists(x^{(k)})_{k=0}^{\infty} \subset \mathcal{H} \text{ by PPM})(\exists \iota : \mathbb{N} \rightarrow \mathbb{N} : \text{strictly increasing}) \\ \lim_{k \rightarrow \infty} x^{(\iota(k))} = z_*.$$

Theorem 1 and Theorem 2 guarantee that the sequence generated by PPM has a point belonging to the region where we can utilize techniques in local convergence analysis. Therefore, we can prove that the sequence generated by PPM, converges to a point in $\bigcap_{i=1}^p S_i$, under certain additional assumptions to guarantee local convergences (Theorem 3 and Theorem 4). Consequently, we succeed in establishing global convergence of the sequence generated by PPM under the special feasibility condition (3) as well as additional assumptions to guarantee local convergences. All the proofs of Theorems 1–4 are omitted due to space limitation. Finally, we conduct a numerical example to investigate numerically convergence behavior of PPM, which demonstrates that the sequence generated by PPM converges to a solution of problem (1) rapidly.

2. PRELIMINARIES

Let \mathbb{N} and \mathbb{R} be the sets of all nonnegative integers and all real numbers, respectively, and let \mathcal{H} a finite dimensional real Hilbert space equipped with an inner product $\langle \cdot, \cdot \rangle$ and its induced norm $\| \cdot \|$. Denote by $2^{\mathcal{H}}$ the collection of all subsets of \mathcal{H} .

Let $T : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ be a set-valued mapping. If T maps $x \in \mathcal{H}$ to a singleton $\{y\} \in 2^{\mathcal{H}}$, we write $y = T(x)$. For a nonempty closed set $S \subset \mathcal{H}$, the metric projection P_S onto S satisfies that $x = P_S(x)$ for any $x \in S$.

3. ASYMPTOTIC PROPERTIES OF PPM

3.1. Basic Convergences

We present the following two theorems as part of our main results.

Theorem 1 (Distance to each set). *Suppose that S_i ($i = 1, 2, \dots, p$) in (1) satisfies the condition (3). Let $(x^{(k)})_{k=0}^{\infty}$ be a sequence gen-*

¹In this paper, the notion of the cluster points is defined for PPM, though it is in general for a sequence: for a sequence $(x^{(k)})_{k=0}^{\infty} \subset \mathcal{H}$, the set of cluster points is defined by

$$\mathfrak{C}((x^{(k)})_{k=1}^{\infty}) := \{x \in \mathcal{H} \mid \exists \iota : \mathbb{N} \rightarrow \mathbb{N} : \text{strictly increasing,} \\ \text{s.t. } x^{(\iota(k))} \rightarrow x \ (k \rightarrow \infty)\}.$$

erated by PPM. Then we have

$$\|x^{(k+1)} - z\|^2 \leq \|x^{(k)} - z\|^2 - \sum_{i=1}^p w_i \|x^{(k)} - y_i^{(k)}\|^2$$

$$\text{for all } z \in \bigcap_{i=1}^p \text{Inc}(S_i) \text{ and for all } k \in \mathbb{N}, \quad (5)$$

and

$$\lim_{k \rightarrow \infty} \|x^{(k)} - P_{S_i}(x^{(k)})\| = 0 \quad (i = 1, 2, \dots, p). \quad (6)$$

Theorem 2 (Set of all possible cluster points). *Suppose that S_i ($i = 1, 2, \dots, p$) in (1) satisfies the condition (3). Then,*

$$\mathfrak{C}_{\text{PPM}} = \bigcap_{i=1}^p S_i. \quad (7)$$

3.2. Global Convergence

Theorem 3 and Theorem 4 below show that additional conditions for S_1, S_2, \dots, S_p guarantee the global convergence of sequences generated by PPM, though Theorem 1 and Theorem 2 do not.

Theorem 3 (Finite union of closed convex sets). *Suppose that S_i ($i = 1, 2, \dots, p$) in (1) satisfies the condition (3). In addition, we assume that each S_i can be expressed as a finite union of closed convex sets $(C_{i,l})_{l=1}^q$, i.e., $S_i = \bigcup_{l=1}^q C_{i,l}$. Let $(x^{(k)})_{k=0}^{\infty}$ be a sequence generated by PPM with initial $x^{(0)} \in \mathcal{H}$. Then $(x^{(k)})_{k=0}^{\infty}$ converges to a point in $\bigcap_{i=1}^p S_i$.*

Theorem 4 (Closed semi-algebraic sets). *Let $\mathcal{H} = \mathbb{R}^n$. Suppose that S_i ($i = 1, 2, \dots, p$) in (1) satisfies the condition (3). In addition, we assume that each S_i is semi-algebraic.² Let $(x^{(k)})_{k=0}^{\infty}$ be a sequence generated by PPM with initial $x^{(0)} \in \mathcal{H}$. Then $(x^{(k)})_{k=0}^{\infty}$ converges to a point in $\bigcap_{i=1}^p S_i$.*

Remark 1: (Idea behind proofs of Theorems 3 and 4) *Thanks to the special feasibility condition (3), Theorem 1 and Theorem 2 hold true, which implies that any sequence generated by PPM becomes bounded and has a cluster point belonging to $\bigcap_{i=1}^p S_i$. In other words, the sequence must have a point sufficiently close to $\bigcap_{i=1}^p S_i$. Consequently, convergence of the sequence can be proven by exploiting techniques in local convergence analysis of projection methods established in existing works, e.g., [29, Theorem 3.4] and [19, Theorem 1 and Theorem 3].*

(Another condition for convergence of sequences) *Although the following condition also guarantees convergence of the sequence generated by PPM, it is not verified easily in general: if S_i ($i = 1, 2, \dots, p$) in (1) satisfies*

$$\text{int} \left(\bigcap_{i=1}^p \text{Inc}(S_i) \right) \neq \emptyset,$$

any sequence generated by PPM converges to a point in $\bigcap_{i=1}^p S_i$.

²For any set $S \subset \mathcal{H} = \mathbb{R}^n$, S is called semi-algebraic [21] if there exists a finite number of real polynomial $P_{ij}, Q_{ij} : \mathcal{H} \rightarrow \mathbb{R}$ such that

$$S = \bigcup_{i \in I} \bigcap_{j \in J} \{x \in \mathcal{H} \mid P_{ij}(x) = 0, Q_{ij}(x) < 0\}.$$

Semi-algebraic nature (or more generally, definable in the o-minimal structure [22, 23]) is exploited to show local convergences of iterative methods including projection methods [18, 19, 24, 25] (see also [26, 27, 28]).

4. ON THE SPECIAL FEASIBILITY CONDITION

We investigate examples of a closed set S such that $\text{Inc}(S) \neq \emptyset$. Examples 1–3 reveal relatively general properties of $\text{Inc}(S)$.

Example 1 (Closed convex set). *Let $C \subset \mathcal{H}$ be a nonempty closed convex set. Then, $\text{Inc}(C) = C$. This is because the projection onto a nonempty closed convex set C can be characterized by*

$$y = P_C(x) \Leftrightarrow (\forall z \in C) \langle x - y, z - y \rangle \leq 0 \quad (8)$$

for any $x, y \in \mathcal{H}$.

Example 2 (Finite union of closed sets S_i s). *Let $S_i \subset \mathcal{H}$ ($i = 1, 2, \dots, q$) be closed sets such that $\bigcap_{i=1}^q \text{Inc}(S_i) \neq \emptyset$. Then, the set $\bigcup_{i=1}^q S_i$ is closed, and*

$$\text{Inc}\left(\bigcup_{i=1}^q S_i\right) \supset \bigcap_{i=1}^q \text{Inc}(S_i) \neq \emptyset.$$

Example 3 (Nonconvex closed cone). *Let a nonempty closed set $S \subset \mathcal{H}$ be a cone, i.e.,*

$$x \in S \Rightarrow (\forall \alpha \in [0, \infty)) \alpha x \in S.$$

Then, $0 \in \text{Inc}(S)$.

Consequently, Theorems 1–4 are possibly applicable in the case where each S_i ($i = 1, 2, \dots, p$) is a closed convex set or a nonconvex closed cone.

Remark 2: *From comparison between (4) and (8), $\text{Inc}(S) \neq \emptyset$ offers a convex-like inequality for a some point in S , which results in certain global properties of PPM, i.e., Theorem 1 and Theorem 2. Example 3 shows that the condition $\text{Inc}(S) \neq \emptyset$ is not stringent.*

Meanwhile, Examples 4–6 below exemplify useful instances playing important roles in signal processing.

Example 4 (Nonnegativity constraint). *Let $\mathcal{H} = \mathbb{R}^n$ equipped with the standard inner product $\langle \cdot, \cdot \rangle$ and its induced norm $\|\cdot\|$. The nonnegativity constraint is defined by*

$$\mathbb{R}_+^n := \{x = (x_1, \dots, x_n) \in \mathcal{H} \mid x_i \geq 0 \ (\forall i = 1, \dots, n)\}. \quad (9)$$

It is a closed semi-algebraic convex cone, and $\mathbb{R}_+^n = \text{Inc}(\mathbb{R}_+^n)$.

The projection $P_{\mathbb{R}_+^n}$ onto \mathbb{R}_+^n can be computed efficiently: $x^* = P_{\mathbb{R}_+^n}(x)$ if and only if x^* is obtained by replacing all the negative entries of x by zero.

Example 5 (Lower level set of ℓ_0 pseudo norm). *Let $\mathcal{H} = \mathbb{R}^n$ equipped with the standard inner product $\langle \cdot, \cdot \rangle$ and its induced norm $\|\cdot\|$. The ℓ_0 pseudo norm $\|\cdot\|_{\ell_0}$ of $x \in \mathbb{R}^n$ is defined as the number of nonzero components in x . The lower level set of the ℓ_0 pseudo norm is defined by*

$$\mathcal{X}_{\ell_0 \leq a}^n = \{x \in \mathbb{R}^n \mid \|x\|_{\ell_0} \leq a\} \ (a \in \mathbb{N}). \quad (10)$$

Then, $\mathcal{X}_{\ell_0 \leq a}^n$ is a closed semi-algebraic cone³, and $0 \in \text{Inc}(\mathcal{X}_{\ell_0 \leq a}^n)$.

The projection $P_{\mathcal{X}_{\ell_0 \leq a}^n}$ onto $\mathcal{X}_{\ell_0 \leq a}^n$ can be computed efficiently: for any $x \in \mathbb{R}^n$, $x^* \in P_{\mathcal{X}_{\ell_0 \leq a}^n}(x)$ if and only if x^* is obtained by replacing the smallest (in the sense of absolute value) $(n - a)$ entries of x by zero.

Example 6 (Lower level set of rank function of matrices). *Let \mathcal{H} be $\mathbb{R}^{m \times n}$ equipped with the standard inner product $\langle \cdot, \cdot \rangle$ and its induced Frobenius norm $\|\cdot\|_F$. The lower level set of rank function*

$$\mathcal{X}_r^{m \times n} := \{X \in \mathbb{R}^{m \times n} \mid \text{rank}(X) \leq r\} \ (r \in \mathbb{N}) \quad (12)$$

is a closed semi-algebraic cone⁴, and hence $0 \in \text{Inc}(\mathcal{X}_r^{m \times n})$.

The projection $P_{\mathcal{X}_r^{m \times n}}$ onto $\mathcal{X}_r^{m \times n}$ can be computed efficiently: Let $X \in \mathbb{R}^{m \times n}$ and $r \leq \text{rank}(X)$. Then $X_r^* \in P_{\mathcal{X}_r^{m \times n}}(X)$ if and only if X_r^* is of the following form:

$$X_r^* = U \Sigma_r V^\top,$$

where $X = U \Sigma V^\top$ is a singular value decomposition⁵ of X , and $\Sigma_r \in \mathbb{R}^{m \times n}$ is the diagonal matrix obtained from Σ by replacing its $(n - r)$ smallest diagonal elements by zeros.

5. APPLICATION OF THEOREM 4

We give two corollaries of Theorem 4. Since the collection $\{\mathbb{R}_+^{m \times n}, \mathcal{X}_{\ell_0 \leq a}^{m \times n}, \mathcal{X}_r^{m \times n}\}$ (see Examples 4–6) satisfies the requirements⁶ in Theorem 4, we can guarantee the global convergence of PPM applied to nonconvex feasibility problems in terms of $\mathbb{R}_+^{m \times n}$, $\mathcal{X}_{\ell_0 \leq a}^{m \times n}$, and $\mathcal{X}_r^{m \times n}$. This gives us powerful iterative algorithms for signal processing problems.

Let us consider a sparse and low-rank matrix approximation problem. For a given matrix $X_0 \in \mathbb{R}^{m \times n}$, a task to find a sparse and low-rank matrix which approximates X_0 well has several application targets including covariance matrix estimation (see e.g. [32]). For this task, a natural problem formulation is

$$\text{find } X \in \mathcal{X}_{\ell_0 \leq a}^{m \times n} \cap \mathcal{X}_r^{m \times n}. \quad (14)$$

Theorem 4 guarantees that PPM applied to (14) converges to a solution.

Corollary 1. *PPM applied to problem (14) with arbitrarily chosen initial $X^{(0)}$ generates a convergent sequence $(X^{(k)})_{k=0}^\infty$ to a point in $\mathcal{X}_{\ell_0 \leq a}^{m \times n} \cap \mathcal{X}_r^{m \times n}$.*

Note that the limit point of the sequence generated by PPM is empirically close to the initial point $X^{(0)}$. Hence PPM with initial X_0 is expected to be a powerful heuristic algorithm to sparse and low-rank matrix approximation.

³Note that $\mathcal{X}_{\ell_0 \leq a}^n$ is nonconvex if $1 \leq a < n$. A simple example which shows nonconvexity of $\mathcal{X}_{\ell_0 \leq 1}^2$ is

$$\frac{1}{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \notin \mathcal{X}_{\ell_0 \leq 1}^2. \quad (11)$$

Although the first term and the second term in the left hand side in (11) belong to $\mathcal{X}_{\ell_0 \leq 1}^2$, the convex combination of these terms does not belong to $\mathcal{X}_{\ell_0 \leq 1}^2$.

⁴Note that $\mathcal{X}_r^{m \times n}$ is nonconvex if $1 \leq r < \min\{m, n\}$. A simple example which shows nonconvexity of $\mathcal{X}_1^{2 \times 2}$ is

$$\frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \notin \mathcal{X}_1^{2 \times 2}. \quad (13)$$

Although the first term and the second term in the left hand side in (13) belong to $\mathcal{X}_1^{2 \times 2}$, the convex combination of these terms does not belong to $\mathcal{X}_1^{2 \times 2}$.

⁵For a complete discussion of the singular value decomposition, see, e.g., [30, 31].

⁶In fact, we have $\text{Inc}(\mathbb{R}_+^{m \times n}) \cap \text{Inc}(\mathcal{X}_{\ell_0 \leq a}^{m \times n}) \cap \text{Inc}(\mathcal{X}_r^{m \times n}) \ni 0$, as well as $\mathbb{R}_+^{m \times n}$, $\mathcal{X}_{\ell_0 \leq a}^{m \times n}$, and $\mathcal{X}_r^{m \times n}$ are semi-algebraic.

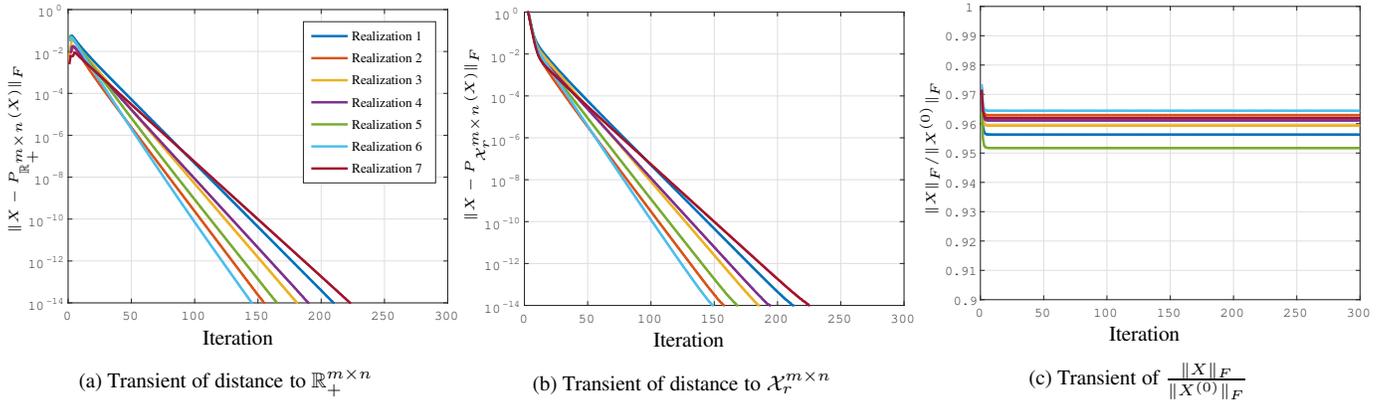


Fig. 1. Convergence behavior of PPM ($m = 20, n = 15, r = 5$).

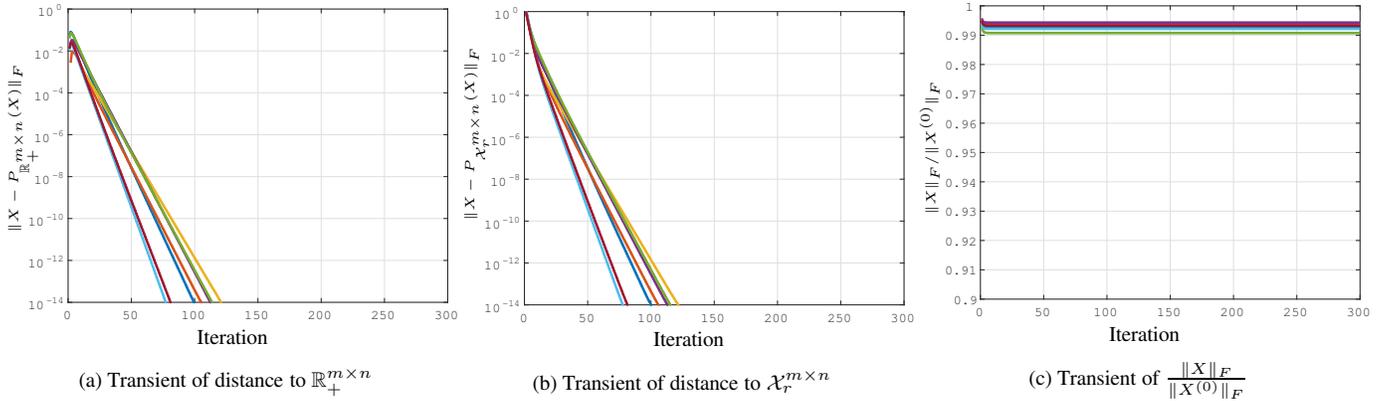


Fig. 2. Convergence behavior of PPM ($m = 20, n = 15, r = 10$).

On the other hand, let us consider a nonnegative and low-rank matrix approximation problem, i.e., for a given matrix $X_0 \in \mathbb{R}_+^{m \times n}$, to find a nonnegative and low-rank matrix which approximates X_0 well. This task relates to feature extraction from a series of images (see e.g. [33, 34]). For this task, a natural problem formulation is

$$\text{find } X \in \mathbb{R}_+^{m \times n} \cap \mathcal{X}_r^{m \times n}. \quad (15)$$

Theorem 4 guarantees that PPM applied to (15) converges to a solution.

Corollary 2. PPM applied to problem (15) with arbitrarily chosen initial $X^{(0)}$ generates a convergent sequence $(X^{(k)})_{k=0}^\infty$ to a point in $\mathbb{R}_+^{m \times n} \cap \mathcal{X}_r^{m \times n}$.

Similar to Corollary 1, PPM with initial X_0 is expected to be a powerful heuristic algorithm to nonnegative and low-rank matrix approximation.

Remark 3: Although problems (14) and (15) have a trivial solution, i.e., the zero matrix $O \in \mathbb{R}^{m \times n}$, it is not an interesting solution from the perspective of their signal processing applications. Even though Theorems 1–4 do not guarantee that PPM converges to a nontrivial solution, such a case seems to be quite rare. This is demonstrated in a numerical example in Sect. 6.

6. NUMERICAL EXAMPLE

We conduct a numerical example to investigate convergence behavior of PPM applied to a toy example of the nonconvex feasibility problem (15). Parameters m, n, r are chosen in two scenarios, i.e., $(m, n, r) = (20, 15, 5)$ and $(m, n, r) = (20, 15, 10)$. The weights of PPM are set as $w_1 = w_2 = 1/2$, and the initial point $X^{(0)}$ of PPM is chosen as a realization of the random matrix whose entry is drawn from the uniform distribution over $[0, 1]$.

Figure 1 depicts convergence behavior of PPM in the first scenario, i.e., $(m, n, r) = (20, 15, 5)$. Figures 1(a) and 1(b) illustrate transients of the distance to $\mathbb{R}_+^{m \times n}$ and that to \mathcal{X}_r . As Theorem 1, the distances converge to zero. In addition, they tend to converge rapidly and linearly though such behavior is not guaranteed by our analysis. Figure 1(c) depicts transients of $\frac{\|X\|_F}{\|X^{(0)}\|_F}$. It shows that all the sequences generated by PPM converges to non-zero matrix. Hence their limits provide us non-trivial solutions of problem (15).

Figure 2 depicts convergence behavior of PPM in the second scenario of $(m, n, r) = (20, 15, 10)$. The whole behavior is similar to the first scenario, but convergences of distances in Figures 2(a) and 2(b) are swifter than that of the first scenario. Figures 2(c) shows all realizations provide us non-trivial solutions of problem (15).

These observations demonstrate that the sequence generated by PPM converges rapidly and its limit is a non-trivial solution of problem (15) in this numerical example.

7. REFERENCES

- [1] P. L. Combettes, “The foundations of set theoretic estimation,” *Proceedings of the IEEE*, vol. 81, no. 2, pp. 182–208, 1993.
- [2] D. C. Youla and H. Webb, “Image restoration by the method of convex projections: Part 1 theory,” *IEEE Transactions on Medical Imaging*, vol. 1, no. 2, pp. 81–94, 1982.
- [3] H. Stark, Y. Yang, and Y. Yang, *Vector space projections: a numerical approach to signal and image processing, neural nets, and optics*, John Wiley & Sons, Inc., 1998.
- [4] G. Cimmino, *Calcolo approssimato per le soluzioni dei sistemi di equazioni lineari*, La Ricerca Scientifica, 1938.
- [5] M. Benzi, “Gianfranco Cimmino’s contributions to numerical mathematics,” *Atti del Seminario di Analisi Matematica, Dipartimento di Matematica dell’Universita di Bologna. Volume Speciale: Ciclo di Conferenze in Ricordo di Gianfranco Cimmino*, pp. 87–109, 2004.
- [6] A. Auslender, *Méthodes numériques pour la résolution des problèmes d’optimisation avec contraintes*, Faculté des sciences, Université de Grenoble, 1969.
- [7] Y. I. Merzlyakov, “On a relaxation method of solving systems of linear inequalities,” *USSR Computational Mathematics and Mathematical Physics*, vol. 2, no. 3, pp. 504–510, 1963.
- [8] S. Reich, “A limit theorem for projections,” *Linear and Multilinear Algebra*, vol. 13, no. 3, pp. 281–290, 1983.
- [9] P. L. Combettes, “Inconsistent signal feasibility problems: Least-squares solutions in a product space,” *IEEE Transactions on Signal Processing*, vol. 42, no. 11, pp. 2955–2966, 1994.
- [10] H. H. Bauschke and J. M. Borwein, “On projection algorithms for solving convex feasibility problems,” *SIAM Review*, vol. 38, no. 3, pp. 367–426, 1996.
- [11] Y. Censor, “Sequential and parallel projection algorithms for feasibility and optimization,” in *Multispectral Image Processing and Pattern Recognition*. International Society for Optics and Photonics, 2001, pp. 1–9.
- [12] T. Blumensath and M. E. Davies, “Iterative hard thresholding for compressed sensing,” *Applied and Computational Harmonic Analysis*, vol. 27, no. 3, pp. 265–274, 2009.
- [13] J.-F. Cai, E. J. Candès, and Z. Shen, “A singular value thresholding algorithm for matrix completion,” *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1956–1982, 2010.
- [14] P. L. Combettes and H. J. Trussell, “Method of successive projections for finding a common point of sets in metric spaces,” *Journal of Optimization Theory and Applications*, vol. 67, no. 3, pp. 487–507, 1990.
- [15] A. S. Lewis and J. Malick, “Alternating projections on manifolds,” *Mathematics of Operations Research*, vol. 33, no. 1, pp. 216–234, 2008.
- [16] A. S. Lewis, D. Russel. Luke, and J. Malick, “Local linear convergence for alternating and averaged nonconvex projections,” *Foundations of Computational Mathematics*, vol. 9, no. 4, pp. 485–513, 2009.
- [17] H. H. Bauschke, D. R. Luke, H. M. Phan, and X. Wang, “Restricted normal cones and sparsity optimization with affine constraints,” *Foundations of Computational Mathematics*, vol. 14, no. 1, pp. 63–83, 2014.
- [18] D. Drusvyatskiy, A. D. Ioffe, and A. S. Lewis, “Transversality and alternating projections for nonconvex sets,” *Foundations of Computational Mathematics*, vol. 15, no. 6, pp. 1637–1651, 2015.
- [19] D. Noll and A. Rondepierre, “On local convergence of the method of alternating projections,” *Foundations of Computational Mathematics*, vol. 16, no. 2, pp. 425–455, 2016.
- [20] R. Orsi, “Numerical methods for solving inverse eigenvalue problems for nonnegative matrices,” *SIAM Journal on Matrix Analysis and Applications*, vol. 28, no. 1, pp. 190–212, 2006.
- [21] M. Coste, “An introduction to semialgebraic geometry,” *RAAG Notes*, p. 78pp., 2002.
- [22] L. Van den Dries, *Tame topology and o-minimal structures*, vol. 248, Cambridge University Press, 1998.
- [23] M. Coste, “An introduction to o-minimal geometry,” *RAAG Notes*, p. 81pp., 1999.
- [24] J. Bolte, A. Daniilidis, and A. Lewis, “The Lojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems,” *SIAM Journal on Optimization*, vol. 17, no. 4, pp. 1205–1223, 2007.
- [25] H. Attouch, J. Bolte, P. Redont, and A. Soubeyran, “Proximal alternating minimization and projection methods for nonconvex problems: An approach based on the Kurdyka-Lojasiewicz inequality,” *Mathematics of Operations Research*, vol. 35, no. 2, pp. 438–457, 2010.
- [26] Y. Xu and W. Yin, “A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion,” *SIAM Journal on imaging sciences*, vol. 6, no. 3, pp. 1758–1789, 2013.
- [27] E. Chouzenoux, J.-C. Pesquet, and A. Repetti, “A block coordinate variable metric forward–backward algorithm,” *Journal of Global Optimization*, pp. 1–29, 2013.
- [28] P. Ochs, Y. Chen, T. Brox, and T. Pock, “iPiano: Inertial proximal algorithm for nonconvex optimization,” *SIAM Journal on Imaging Sciences*, vol. 7, no. 2, pp. 1388–1419, 2014.
- [29] H. H. Bauschke and D. Noll, “On cluster points of alternating projections,” *Serdica Mathematical Journal*, vol. 39, pp. 355–364, 2013.
- [30] R. A. Horn and C. R. Johnson, *Matrix analysis*, Cambridge University Press, 2012.
- [31] A. Ben-Israel and T. N. E. Greville, *Generalized inverses: theory and applications*, Springer Science & Business Media, 2003.
- [32] E. Richard, P. A. Savalle, and N. Vayatis, “Estimation of simultaneously sparse and low rank matrices,” in *International Conference on Machine Learning (ICML)*, 2012.
- [33] C. Zhang, J. Liu, Q. Tian, C. Xu, H. Lu, and S. Ma, “Image classification by non-negative sparse coding, low-rank and sparse decomposition,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011, pp. 1673–1680.
- [34] G. Zhou, A. Cichocki, and S. Xie, “Fast nonnegative matrix/tensor factorization based on low-rank approximation,” *IEEE Transactions on Signal Processing*, vol. 60, no. 6, pp. 2928–2940, 2012.