Review on POCS Algorithms for Image Reconstruction

Ming Jiang
School of Mathematics, Peking University, Beijing 100871, China

Zhaotian Zhang
Department of Information Science, National Nature Science Foundation of China
Beijing 100085, China

Abstract

We provide a brief review on recent advances in POCS algorithms for image reconstruction with a preference towards practical applications and report relevant convergence results in various formulations and topologies for the consistent and inconsistent cases. Mathematical details are minimized for a wide readership.

摘要

本文介绍了近年来关于图象重建的凸集投影算法的进展。我们主要讨论了基于正交投影和广义投影算法。对基于正交投影算法，我们讨论了一般形式的加权松弛迭代格式，这包含了分块迭代格式和同时迭代格式。文中以三个定理形式介绍有关算法的收敛性结果，其中包含了相容和不相容条件和弱强收敛下的结果。对基于广义投影算法，我们介绍了有关的基本概念和例子、基本算法的收敛性结果，并介绍了最近关于引入松弛系数的研究工作及其在 CT 图象重建中的应用。由于这种方法在图象重建和其它反问题（包括不完全数据问题）中的广泛应用前景，本综述的目的在于提请有关读者关注这一重要的方法及应用。

Purpose：Review of recent advances in POCS algorithms for image reconstruction.

Method: POCS based iterative algorithms.

Result: Various recent results and applications are presented.

Conclusion: The powerful algorithmic formulation of the POCS scheme will find various important applications in image reconstruction

Key words: POCS, Iterative Algorithm, Convergence, Image reconstruction.

Submitted to *CT Theory and Applications*

§ 1. Introduction

Image reconstruction arises in many fields of applications. Imaging systems, modeled originally in continuous analytic functional equations, can usually be discretized into linear systems of equations

$$Ax = b \tag{1}$$

where the observed data $b = \left(b^1, \cdots, b^M\right) \in R^M$, the original image $x = \left(x_1, \cdots, x_M\right) \in R^N$, $A = \left(A_{ij}\right)$ a non-zero $M \times N$ matrix. The problem is to reconstruct the image x from the observed data b. A solution is not feasible by a conventional method with complicated computation involving the matrix A, because of the ill-posedness of the problem, noisy data b, unstructured imaging matrix A, and huge data dimensions in practice. Instead, iterative methods, such as the algebraic reconstruction technique (ART), etc., are developed for efficient image reconstruction.

The convex set theoretic image recovery problem is to produce an estimate of the original image in the intersection of a family of closed and convex sets $\{C_i\}_{1 \le i \le m}$ in a real Hilbert space H, which represents the *a priori* knowledge about the problem, such as non-negativity and bound constraints *etc*., and information from the data. The reconstruction problem is then reduced to the convex feasibility problem (CFP):

$$\text{Find } x^* \in C = \bigcap_{1 \le i \le m} C_i \tag{2}$$

§ 2. Orthogonal POCS Algorithms

The mostly used method to solve (2) is called *projections onto convex sets* (POCS), which generates an image as the limit of a sequence $\{x_n\}$ of projections onto the sets, i.e.,

$$x_{n+1} = P_{C_{[n]}}(x_n) \tag{3}$$

where $[n] = n \operatorname{mod}(m) + 1$, P_{C_i} is the orthogonal projection from H onto C_i for $1 \le i \le m$. When C_i is the hyperplane $C_i = \left\{x \in H : \left\langle x, a^i \right\rangle = b^i \right\}$ for $1 \le i \le m = M$, where a^i is the i-th row of A, the resultant algorithm is the ART. Detailed accounts of POCS can be found in [3]. The POCS algorithm can be put in a more general form [1], to accommodate various relaxation, weighting and block-iterative schemes. Let

$$R_i^{(n)} = \mathrm{Id} + \lambda_i^{(n)}(P_{C_i} - \mathrm{Id}), \tag{4}$$

where $\lambda_i^{(n)} \in [0,2]$ is a relaxation parameter and Id is the identity map. A relaxation strategy is called under-relaxation if $\lambda_i^{(n)} \in [0,1]$, over-relaxation if $\lambda_i^{(n)} \in [1,2]$, and unrelaxed if $\lambda_i^{(n)} = 1$, respectively. Let

$$A^{(n)} = \sum_{i=1}^{m} \omega_i^{(n)} R_i^{(n)} \tag{5}$$

where $\{\omega_i^{(n)}\} \subset [0,1]$ is a weight, i.e., $\sum_{i=1}^{m} \omega_i^{(n)} = 1$. The general form of POCS is defined as

$$x_0 \in H \text{ arbitrary}, \quad x_{n+1} = A^{(n)}(x_n)\,. \tag{6}$$

The set of active indices is defined to be $I^{(n)} = \{i \in \{1,2,\cdots,m\} : \omega_i^{(n)} > 0\}$. We say that an index i is active at iteration n or n is activated for i if $\omega_i^{(n)} > 0$. The way that indices are chosen to be active is called a control. The control used in (3) is called cyclic control. If there is an integer $p > 0$ such that

$$i \in I^{(n)} \bigcup I^{(n+1)} \bigcup \cdots \bigcup I^{(n+p)}, \text{ for every index } i \text{ and } n \geq 0, \tag{7}$$

the control is called intermittent or p - intermittent. Cyclic control is a special case of intermittent control. Although mathematically intriguing, controls that are different from the intermittent control seems to be of little use for applications [1].

The primary question with any iterative scheme is its convergence behavior, which characterizes the computational stability of the scheme, hence the reconstructed image quality. Extensive work has been done on the convergence of POCS algorithms [1,3]. We list two general results in the following. To facilitate the presentation, let

$$\mu_i^{(n)} = \omega_i^{(n)} \lambda_i^{(n)} (2 - \sum_{j=1}^{m} \omega_j^{(n)} \lambda_j^{(n)}). \tag{8}$$

The following theorem guarantees the weak convergence of the sequence $\{x_n\}$.

Theorem 1.([1]) If the intersection set C *is not empty and* $\underline{\lim}_{n:n \text{ active for } i}\, \mu_i^{(n)} > 0$ *for every index* i *, then the sequence* $\{x_n\}$ *converges weakly to some point* $x \in C$ *.*

In some applications, the reconstruction problem arises in infinite dimension. In such problems, a strong convergence result for analytic functional iterative algorithms is often much more desirable than weak convergence. The importance of the strong convergence is also relevant when the algorithm is implemented in a finite dimensional setting through discretization of its original infinite dimensional analytic model. The performance of the algorithm is closely related to the discretized counterpart and its convergence behavior. The following is a result that guarantees strong or norm convergence.

Theorem 2.([1]) If the intersection set C *is not empty and there is some* $\varepsilon > 0$ *such that* $0 < \varepsilon \leq \lambda_i^{(n)} \leq 2 - \varepsilon$ *for all* n *and index* i *active at* n *. If the interior of* C *is not empty* *or H is finite dimensional , then the sequence* $\{x_n\}$ *converges strongly to some point* $x \in H$ *. If* $\sum_n \mu_i^{(n)} = +\infty$ *for every index* i *, then* $x \in C$ *.*

The above result is established under the assumption that the intersection set C is not empty, i.e., the constraints C_i are consistent. In the inconsistent case, if one of the sets is bounded, there exists points $\{y_i\}_{1\le i\le m}$ such that $P_1(y_m) = y_1$, $P_i(y_{i-1}) = y_i$, for $2 \le i \le m$. Moreover, the cyclic subsequence $\{x_{mn+i}\}_{n\ge 0}$ converges to $y_i \in C_i$, c.f., [4] and the references therein. The inconsistent case was recently studied in [4,6]. In [4], the exact feasibility problem (2) was replaced by the weighted least-squares feasibility problem

$$\text{Minimize}\, \Phi(x) = \sum_{i=1}^{m} \omega_i d(x, C_i). \tag{9}$$

Theorem 3.([4]) Let $\{\omega_i\}_{1\le i\le m}$ *be strictly convex weights, i.e.,* $\sum_{i=1}^{m} \omega_i = 1$ *and* $\omega_i > 0$ *for* $1 \le i \le m$ *. Assume that one of the sets* $\{C_i\}_{1\le i\le m}$ *is bounded. Then, for any* $x_0 \in H$ *, every sequence of iterates* $\{x_n\}$ *generated by*

$$x_{n+1} = x_n + \lambda^{(n)}\left(\sum_{i=1}^{m} \omega_i P_{C_i}(x_n) - x_n\right), \tag{10}$$

converges weakly to a minimizer of $\Phi(x)$ *, where* $0 < \varepsilon \le \lambda^{(n)} \le 2 - \varepsilon$ *for some* $\varepsilon > 0$ *.*

Before we turn to the general POCS algorithms based on non-orthogonal projection, we would like to mention several important aspects about the orthogonal projection based POCS algorithms, not covered in this review due to space limit: (a) the strong convergence was extensively studied in [1] based on the regularity of the convex sets $\{C_i\}_{1\le i\le m}$; Recent progress is to establish weak-to-strong convergence by modifying the projections at each projections [2]. (b) the convergence rate of POCS was studied in [1] based on the linear regularity of the convex sets $\{C_i\}_{1\le i\le m}$, where it was proved that a linear convergence rate of the POCS algorithm could be achieved; (c) it is possible to extend the relaxation range beyond the interval $[0,2]$ and improve the convergence rate of the algorithm [5]. Interested readers may refer to the references for details.

§ 2. General Projections and POCS

The POCS algorithms in §1 utilizes the orthogonal projection induced from the underlying Hilbert space. The generalized distances and projections were first proposed by Bregman and has been studied extensively in the past decades [3].

Let $H = R^N$ and S be a nonempty, open, convex set, such that the closure $\overline{S} \subset \Lambda$, where $\Lambda \subset H$ is the domain of a function $f : \Lambda \to R$. Assume that $f \in C^1(S)$. From f, construct the function $D_f(\cdot,\cdot) : \overline{S} \times S \to R$ by

$$D_f(x, y) = f(x) - f(y) - \langle \nabla f(y), x - y \rangle . \tag{11}$$

This function is called a *generalized distance function* or D-function. The partial level sets of D_f is defined as $L^f(x, \alpha) = \{y \in S : D_f(x, y) \le \alpha\}$. The original definition of Bregman functions was recently simplified in [7].

Definition. (Bregman Functions) Under the above assumptions, the function f *is called a Bregman function if the following conditions hold: (a)* f *is continuous and strictly convex on* $\overline{S}$*; (b) for every* $\alpha \in R$ *and* $x \in R^N$*,* $L^f(x, \alpha)$ *is bounded; (c) if* $y_n \in S$*, for all* $n \ge 0$*, and* $\lim y_n = y^*$*, then* $\lim D_f(y^*, y_n) = 0$*; (d) if* $\{x_n\} \subset S$ *is a convergent sequence with limit* $x^* \in \partial S$*, then* $\lim \langle \nabla f(x_n), x^* - x_n \rangle = 0$.

We denote the family of Bregman functions by $B(S)$, refer to the set S as the zone of the function f. It can be proved that $D_f(x, y) \ge 0$ for $(x, y) \in \overline{S} \times S$ and $D_f(x, y) = 0$ if and only if $x = y$. Hence, $D_f(x, y)$ may be used to measure the distance between x and y in some generalized sense. We use the generalized distance $D_f(x, y)$ to define generalized projection. It can be proved that for any closed convex set $\Omega \subset R^N$ such that the intersection $\Omega \cap \overline{S}$ is not empty, and for every $y \in S$, there exists a unique $x^* \in \Omega \cap \overline{S}$ such that

$$D_f(x^*, y) = \min\{D_f(x, y) : x \in \Omega \cap \overline{S}\} . \tag{12}$$

x^* is called the generalized projection of the point y onto the set Ω and denoted by $P_\Omega(y)$.

Two typical examples of Bregman functions are: (a) the function $f(x) = \| x \|^2 / 2$, with $\Lambda = S = R^N$, and $D_f(x, y)$ is then $D_f(x, y) = \| x - y \|^2 / 2$. The general projection in this case is the ordinary orthogonal projection. (b) the entropy function $f(x) = -\sum_j x_j \log x_j$, with

$\Lambda = R_{+}^{n}$ and $S = R_{++}^{n}$. The distance function $D_f(x, y)$ is then the Kullback-Leibler entropy,

$D_f(x, y) = \sum_j [x_j \log \frac{x_j}{y_j} + y_j - x_j]$.

With the generalized projection defined in this way, we can formulate unrelaxed POCS algorithms as in [3]. We have the following result.

Theorem 1.([3]) Assume that the intersection set C *is not empty and* $f \in B(S)$ *such that* $C \bigcap \overline{S}$ *is not empty. If for every* $y \in S$, $P_{C_i}(y) \in C_i$, *for every index* i, *then the sequence* $\{x_n\}$ *converges to some point* $x \in C$.

For some time, it is not clear how to introduce relaxation into the POCS algorithms based on the generalized projection for convex sets that are not linear. Recently, Censor and Gabor provided a method to construct under-relaxation for this approach [8]. The practical importance of relaxation is also demonstrated through a CT image reconstruction example. Please refer to [8] for details.

In finishing this review, we would like to draw the readers' attention to the convergence results on another class of algebraic iterative algorithms based on the Landweber scheme, which includes the SART and other recently well designed algebraic reconstruction algorithms as special cases, see [9] for details.

§ 3. Conclusion

We have reviewed recent advances in POCS algorithms and discussed both the orthogonal and generalized projection based algorithms. We have presented the orthogonal projection based algorithms in the form of general weighted relaxation, which includes block iterative and simultaneous versions of the algorithm, and then reported relevant convergence results in three theorems, for the consistent and inconsistent cases, in the weak and strong topologies, respectively. For the generalized projection based algorithms, we have discussed basic concepts and examples, presented the convergence result for the unrelaxed algorithm, and reported recent results in relaxation and its applications in CT image reconstruction. It is certain that the powerful algorithmic formulation of the POCS scheme will find various important applications in image reconstruction.

Acknowledgment

Ming Jiang's work was supported in part by an NKBRSF grant (G1998030606) and an NSFC grant (60272018).

References

[1] H. H. Bauschke and J. M. Borwein. On projection for solving convex feasibility problems. *SIAM Review*, 38:367-426, 1996.

[2] H. H. Bauschke and P. L. Combettes. A weak-to-strong convergence principles for

Fejér-monotone methods in Hilbert spaces. *Mathematics for Operations Research*, 26:248-264, 2001.

[3] Y. Censor and S. A. Zenios. *Parallel Optimization: Theory, Algorithms and Applications*. Oxford University Press, 1997.

[4] P. L. Combettes. Inconsistent signal feasibility problems: least-squares solutions in a product space. *IEEE Trans. Signal Processing*, 42:2955-2966, 1994.

[5] P. L. Combettes. Convex set theoretic image recovery by extrapolated iterations of parallel subgradient projections. *IEEE Trans. Image Processing*, 6:493-506, 1997.

[6] T. Kotzer, N. Cohen and J. Shamir. A projection-based algorithm for consistent and inconsistent constraints. SIAM J. Optim, 7:527-546, 1997.

[7] D. Butnariu, C. Byrne and Y. Censor. Redundant axioms in the definition of Bregman functions. *Journal of Convex Analysis*. To appear.

[8] Y. Censor and G.T. Herman. Block-iterative algorithms with underrelaxed Bregman projections. *SIAM Journal on Optimizations*, 13:283-297, 2002.

[9] Ming Jiang, Ge Wang, *Convergence Studies on Iterative Algorithms for Image Reconstruction*, IEEE Transactions on Medical Imaging, 2003, to appear.

About the authors: Ming Jiang is a professor with the School of Mathematical Science, Peking University. He was a post-doctoral fellow with the Microprocessor Laboratory, International Centre for Theoretical Physics, Italy, from 1996 to 1997; a visiting professor and associate director of the CT/Micro-CT Laboratory, the Department of Radiology, University of Iowa, from 2000 to 2002. His interests are image reconstruction, restoration and analysis.
Zhaotian Zhang is a program director of Department of Information Science, NSFC. His researching interests are signal processing and image processing on CT/CST.