Feedback in Iterative Algorithms

Charles Byrne (Charles_Byrne@uml.edu),
Department of Mathematical Sciences,
University of Massachusetts Lowell, Lowell, MA 01854

October 17, 2005

Abstract

When the nonnegative system of linear equations $y = Px$ has no nonnegative exact solutions, block-iterative methods fail to converge, exhibiting, instead, subsequential convergence leading to a limit cycle of two or more distinct vectors. From the vectors of the limit cycle we extract *new data*, replace the data vector y with this new data vector, and restart the iterative process. Success with this approach as applied to the algebraic reconstruction technique (ART) prompts us to apply feedback to other block-iterative methods. We consider here several questions pertaining to this *feedback* approach: Can we show the existence of limit cycles for these other block-iterative algorithms? Does the sequence of successive limit cycles converge to a single vector? If so, which vector is it?

1 Introduction

When the nonnegative system of linear equations $y = Px$ has no nonnegative solutions we say that we are in the *inconsistent case*. In this case the SMART and EMML algorithms still converge, to a nonnegative minimizer of $KL(Px, y)$ and $KL(y, Px)$, respectively. On the other hand, the rescaled block-iterative versions of these algorithms, RBI-SMART and RBI-EMML, do not converge. Instead they exhibit *cyclic subsequential convergence*; for each fixed $n = 1, ..., N$, with N the number of blocks, the subsequence $\{x^{mN+n}\}$ converges to their own limits. These limit vectors then constitute the *limit cycle* (LC). The LC for RBI-SMART is not the same as for RBI-EMML, generally, and the LC varies with the choice of blocks. Our problem is to find a way to calculate the SMART and EMML limit vectors using the RBI methods. More specifically, how can we calculate the SMART and EMML limit vectors from their associated RBI limit cycles? For details concerning these algorithms see [3].

As is often the case with the algorithms based on the KL distance, we can turn to the ART algorithm for guidance. What happens with the ART algorithm in the

inconsistent case is often closely related to what happens with RBI-SMART and RBI-EMML, although proofs for the latter methods are more difficult to obtain. For example, when the system $Ax = b$ has no solution we can prove that ART exhibits cyclic subsequential convergence to a limit cycle. The same behavior is seen with the RBI methods, but no one knows how to prove this. When the system $Ax = b$ has no solution we usually want to calculate the least squares (LS) approximate solution. The problem then is to use the ART to find the LS solution. There are several ways to do this, as discussed in [2, 3]. We would like to be able to borrow some of these methods and apply them to the RBI problem. In this article we focus on one specific method that works for ART and we try to make it work for RBI; it is the *feedback* approach.

2 Feedback in ART

Suppose that the system $Ax = b$ has no solution. We apply the ART and get the limit cycle $\{z^1, z^2, ..., z^M\}$, where M is the number of equations and $z^0 = z^M$. We assume that the rows of A have been normalized so that their lengths are equal to one. Then the ART iterative step gives

$$z_n^m = z_n^{m-1} + A_{mn}(b_m - (Az^{m-1})_m)$$

or

$$z_n^m - z_n^{m-1} = A_{mn}(b_m - (Az^{m-1})_m).$$

Summing over the index m and using $z^0 = z^M$ we obtain zero on the left side, for each n. Consequently $A^T b = A^T c$, where c is the vector with entries $c_m = (Az^{m-1})_m$. It follows that the systems $Ax = b$ and $Ax = c$ have the same LS solutions and that it may help to use both b and c to find the LS solution from the limit cycle. The article [2] contains several results along these lines. One approach is to apply the ART again to the system $Ax = c$, obtaining a new LC and a new candidate for the right side of the system of equations. If we repeat this *feedback* procedure, each time using the LC to define a new right side vector, does it help us find the LS solution? Yes, as Theorem 4 of [2] shows. Our goal in this article is to explore the possibility of using the same sort of feedback in the RBI methods. Some results in this direction are in [2]; we review those now.

3 Feedback in RBI methods

One issue that makes the KL methods more complicated than the ART is the support of the limit vectors, meaning the set of indices j for which the entries of the vector are positive. In [1] it was shown that when the system $y = Px$ has no nonnegative solutions and P has the *full rank property* there is a subset S of $\{j = 1, ..., J\}$ with cardinality at most $I - 1$, such that every nonnegative minimizer of $KL(Px, y)$ has zero for its j-th entry whenever j is not in S. It follows that the minimizer is unique. The same result holds for the EMML, although it has not been proven that the set S is the same as in the SMART case. The same result holds for the vectors of the LC for both RBI-SMART and RBI-EMML.

A simple, yet helpful, example to refer to as we proceed is the following.

$$P = \begin{bmatrix} 1 & .5 \\ 0 & .5 \end{bmatrix}, y = \begin{bmatrix} .5 \\ 1 \end{bmatrix}.$$

There is no nonnegative solution to this system of equations and the support set S for SMART, EMML and the RBI methods is $S = \{j = 2\}$.

3.1 The RBI-SMART

Our analysis of the SMART and EMML methods has shown that the theory for SMART is somewhat nicer than that for EMML and the resulting theorems for SMART are a bit stronger. The same is true for RBI-SMART, compared to RBI-EMML. For that reason we begin with RBI-SMART.

Recall that the iterative step for RBI-SMART is

$$x_j^{k+1} = x_j^k \exp(m_n^{-1} s_j^{-1} \sum_{i \in B_n} P_{ij} \log(y_i/(Px^k)_i)),$$

where $n = k(\text{mod}\, N) + 1$, $s_j = \sum_{i=1}^I P_{ij}$, $s_{nj} = \sum_{i \in B_n} P_{ij}$ and $m_n = \max\{s_{nj}/s_j,\, j = 1, ..., J\}$.

For each n let

$G_n(x, z) =$

$$\sum_{j=1}^J s_j KL(x_j, z_j) - m_n^{-1} \sum_{i \in B_n} KL((Px)_i, (Pz)_i) + m_n^{-1} \sum_{i \in B_n} KL((Px)_i, y_i).$$

Exercise 1: Show that

$$\sum_{j=1}^J s_j KL(x_j, z_j) - m_n^{-1} \sum_{i \in B_n} KL((Px)_i, (Pz)_i) \geq 0,$$

so that $G_n(x,z) \geq 0$.

Exercise 2: Show that

$$G_n(x,z) = G_n(z',z) + \sum_{j=1}^{J} s_j KL(x_j, z'_j),$$

where

$$z'_j = z_j \exp(m_n^{-1} s_j^{-1} \sum_{i \in B_n} P_{ij} \log(y_i/(Pz)_i).$$

We assume that there are no nonnegative solutions to the nonnegative system $y = Px$. We apply the RBI-SMART and get the limit cycle $\{z^1, ..., z^N\}$, where N is the number of blocks. We also let $z^0 = z^N$ and for each i let $w_i = (Pz^{n-1})_i$ where $i \in B_n$, the n-th block. Prompted by what we learned concerning the ART, we ask if the nonnegative minimizers of $KL(Px, y)$ and $KL(Px, w)$ are the same. This would be the correct question to ask if we were using the slower unrescaled block-iterative SMART, in which the m_n are replaced by one. For the rescaled case it turns out that the proper question to ask is: Are the nonnegative minimizers of the functions

$$\sum_{n=1}^{N} m_n^{-1} \sum_{i \in B_n} KL((Px)_i, y_i)$$

and

$$\sum_{n=1}^{N} m_n^{-1} \sum_{i \in B_n} KL((Px)_i, w_i)$$

the same? The answer is "Yes, probably." The difficulty has to do with the support of these minimizers; specifically: Are the supports of both minimizers the same as the support of the LC vectors? If so, then we can prove that the two minimizers are identical. This is our motivation for the feedback approach.

The *feedback* approach is the following: beginning with $y^0 = y$ we apply the RBI-SMART and obtain the LC, from which we extract the vector w, which we also call w^0. We then let $y^1 = w^0$ and apply the RBI-SMART to the system $y^1 = Px$. From the resulting LC we extract $w^1 = y^2$, and so on. In this way we obtain an infinite sequence of *data vectors* $\{y^k\}$. We denote by $\{z^{k,1}, ..., z^{k,N}\}$ the LC we obtain from the system $y^k = Px$, so that

$$y_i^{k+1} = (Pz^{k,n})_i, \text{ for } i \in B_n.$$

One issue we must confront is how we use the support sets. At the first step of feedback we apply RBI-SMART to the system $y = y^0 = Px$, beginning with a positive vector

x^0. The resulting limit cycle vectors are supported on a set S^0 with cardinality less than I. At the next step we apply the RBI-SMART to the system $y^1 = Px$. Should we begin with a positive vector (not necessarily the same x^0 as before) or should our starting vector be supported on S^0?

Exercise 3: Show that the RBI-SMART sequence $\{x^k\}$ is bounded.

Hints: For each j let $M_j = \max\{y_i/P_{ij}\,, |P_{ij} > 0\}$ and let $C_j = \max\{x_j^0, M_j\}$. Show that $x_j^k \leq C_j$ for all k.

Exercise 4: Let S be the support of the LC vectors. Show that

$$\sum_{n=1}^{N} m_n^{-1} \sum_{i \in B_n} P_{ij} \log(y_i/w_i) \leq 0 \tag{3.1}$$

for all j, with equality for those $j \in S$. Conclude from this that

$$\sum_{n=1}^{N} m_n^{-1} \sum_{i \in B_n} KL((Px)_i, y_i) - \sum_{n=1}^{N} m_n^{-1} \sum_{i \in B_n} KL((Px)_i, w_i) \geq$$

$$\sum_{n=1}^{N} m_n^{-1} \sum_{i \in B_n} (y_i - w_i),$$

with equality if the support of the vector x lies within the set S.

Hints: For $j \in S$ consider $\log(z_j^n/z_j^{n-1})$ and sum over the index n, using the fact that $z^N = z^0$. For general j assume there is a j for which the inequality does not hold. Show that there is M and $\epsilon > 0$ such that for $m \geq M$

$$\log(x_j^{(m+1)N}/x_j^{mN}) \geq \epsilon.$$

Conclude that the sequence $\{x_j^{mN}\}$ is unbounded.

Exercise 5: Show that

$$\sum_{n=1}^{N} G_n(z^{k,n}, z^{k,n-1}) = \sum_{n=1}^{N} m_n^{-1} \sum_{i \in B_n} (y_i^k - y_i^{k+1}),$$

and conclude that the sequence $\{\sum_{n=1}^{N} m_n^{-1}(\sum_{i \in B_n} y_i^k)\}$ is decreasing and that the sequence $\{\sum_{n=1}^{N} G_n(z^{k,n}, z^{k,n-1})\} \to 0$ as $k \to \infty$.

Hints: Calculate $G_n(z^{k,n}, z^{k,n-1})$ using Exercise 2.

Exercise 6: Show that for all vectors $x \geq 0$ the sequence

$$\{\sum_{n=1}^{N} m_n^{-1} \sum_{i \in B_n} KL((Px)_i, y_i^k)\}$$

is decreasing and the sequence

$$\sum_{n=1}^{N} m_n^{-1} \sum_{i \in B_n} (y_i^k - y_i^{k+1}) \to 0,$$

as $k \to \infty$.

Hints: Calculate

$$\{\sum_{n=1}^{N} m_n^{-1} \sum_{i \in B_n} KL((Px)_i, y_i^k)\} - \{\sum_{n=1}^{N} m_n^{-1} \sum_{i \in B_n} KL((Px)_i, y_i^{k+1})\}$$

and use the previous exercise.

Exercise 7: Extend the boundedness result obtained earlier to conclude that for each fixed n the sequence $\{z^{k,n}\}$ is bounded.

Since the sequence $\{z^{k,0}\}$ is bounded there is a subsequence $\{z^{k_t,0}\}$ converging to a limit vector $z^{*,0}$. Since the sequence $\{z^{k_t,1}\}$ is bounded there is subsequence converging to some vector $z^{*,1}$. Proceeding in this way we find subsequences $\{z^{k_m,n}\}$ converging to $z^{*,n}$ for each fixed n. Our goal is to show that, with certain restrictions on P, $z^{*,n} = z^*$ for each n. We then show that the sequence $\{y^k\}$ converges to Pz^* and that z^* minimizes

$$\sum_{n=1}^{N} m_n^{-1} \sum_{i \in B_n} KL((Px)_i, y_i).$$

It follows from Exercise 5 that

$$\{\sum_{n=1}^{N} G_n(z^{*,n}, z^{*,n-1})\} = 0.$$

Exercise 8: Find suitable restrictions on the matrix P to permit us to conclude from above that $z^{*,n} = z^{*,n-1} = z^*$ for each n.

Exercise 9: Show that the sequence $\{y^k\}$ converges to Pz^*.

Hints: Since the sequence $\{\sum_{n=1}^{N} m_n^{-1} \sum_{i \in B_n} KL((Pz^*)_i, y_i^k)\}$ is decreasing and a subsequence converges to zero, it follows that the whole sequence converges to zero.

Exercise 10: Use Exercise 4 to obtain conditions that permit us to conclude that the vector z^* is a nonnegative minimizer of the function

$$\sum_{n=1}^{N} m_n^{-1} \sum_{i \in B_n} KL((Px)_i, y_i).$$

3.2 The RBI-EMML

We turn now to the RBI-EMML method, having the iterative step

$$x_j^{k+1} = (1 - m_n^{-1} s_j^{-1} s_{nj}) x_j^k + m_n^{-1} s_j^{-1} x_j^k \sum_{i \in B_n} P_{ij} y_i / (Px^k)_i,$$

with $n = k(\text{mod}\, N)+1$. As we warned earlier, developing the theory for feedback with respect to the RBI-EMML algorithm appears more difficult than in the RBI-SMART case.

Applying the RBI-EMML algorithm to the system of equations $y = Px$ having no nonnegative solution, we obtain the LC $\{z^1, ..., z^N\}$. As before, for each i we let $w_i = (Pz^{n-1})_i$ where $i \in B_n$. There is a subset S of $\{j = 1, ..., J\}$ with cardinality less than I such that for all n we have $z_j^n = 0$ if j is not in S.

The first question that we ask is: Are the nonnegative minimizers of the functions

$$\sum_{n=1}^{N} m_n^{-1} \sum_{i \in B_n} KL(y_i, (Px)_i)$$

and

$$\sum_{n=1}^{N} m_n^{-1} \sum_{i \in B_n} KL(w_i, (Px)_i)$$

the same?

As before, the feedback approach involves setting $y^0 = y$, $w^0 = w = y^1$ and for each k defining $y^{k+1} = w^k$, where w^k is extracted from the limit cycle

$$LC(k) = \{z^{k,1}, ..., z^{k,N} = z^{k,0}\}$$

obtained from the system $y^k = Px$ as $w_i^k = (Pz^{k,n-1})_i$ where n is such that $i \in B_n$. Again, we must confront the issue of how we use the support sets. At the first step of feedback we apply RBI-EMML to the system $y = y^0 = Px$, beginning with a positive vector x^0. The resulting limit cycle vectors are supported on a set S^0 with cardinality less than I. At the next step we apply the RBI-EMML to the system $y^1 = Px$. Should we begin with a positive vector (not necessarily the same x^0 as before) or should our starting vector be supported on S^0? One approach could be to assume first that $J < I$ and that $S = \{j = 1, ..., J\}$ always and then see what can be discovered.

Our conjectures, subject to restrictions involving the support sets, are as follows:
1: The sequence $\{y^k\}$ converges to a limit vector y^∞;
2: The system $y^\infty = Px$ has a nonnegative solution, say x^∞;

3: The LC obtained for each k converge to the singleton x^∞;
4: The vector x^∞ minimizes the function

$$\sum_{n=1}^{N} m_n^{-1} \sum_{i \in B_n} KL(y_i, (Px)_i)$$

over nonnegative x.

Some results concerning feedback for RBI-EMML were presented in [2]. We sketch those results now.

Exercise 11: Show that the quantity

$$\sum_{n=1}^{N} m_n^{-1} \sum_{i \in B_n} y_i^k$$

is the same for $k = 0, 1, ...$

Hints: Show that

$$\sum_{j=1}^{J} s_j \sum_{n=1}^{N} (z_j^{k,n} - z_j^{k,n-1}) = 0$$

and rewrite it in terms of y^k and y^{k+1}.

Exercise 12: Show that there is a constant $B > 0$ such that $z_j^{k,n} \leq B$ for all k, n and j.

Exercise 13: Show that

$$s_j \log(z_j^{k,n-1}/z_j^{k,n}) \leq m_n^{-1} \sum_{i \in B_n} P_{ij} \log(y_i^{k+1}/y_i^k).$$

Hints: Use the convexity of the log function and the fact that the terms $1 - m_n^{-1} s_{nj}$ and $m_n^{-1} P_{ij}$, $i \in B_n$ sum to one.

Exercise 14: Use the previous exercise to prove that the sequence

$$\{\sum_{n=1}^{N} m_n^{-1} \sum_{i \in B_n} KL((Px)_i, y_i^k)\}$$

is decreasing for each nonnegative vector x and the sequence

$$\{\sum_{n=1}^{N} m_n^{-1} \sum_{i \in B_n} P_{ij} \log(y_i^k)\}$$

is increasing.

References

[1] Byrne, C. (1993) "Iterative image reconstruction algorithms based on cross-entropy minimization." *IEEE Transactions on Image Processing* **IP-2**, pp. 96–103.

[2] Byrne, C. (1997) "Convergent block-iterative algorithms for image reconstruction from inconsistent data." *IEEE Transactions on Image Processing* **IP-6**, pp. 1296–1304.

[3] Byrne, C. (2005) *Signal Processing: A Mathematical Approach*, AK Peters, Publ., Wellesley, MA.