Problems and Algorithms in Continuous Optimization

Charles L. Byrne
Department of Mathematical Sciences
University of Massachusetts Lowell
Lowell, MA 01854

August 23, 2012

(The most recent version is available as a pdf file at
http://faculty.uml.edu/cbyrne/cbyrne.html)

Contents

Chapter 1

Preliminaries

1.1 Overview

In the next chapter we consider the problem of solving the system of linear equations $Ax = b$, or, more generally, $A^T Ax = A^T b$, along with the iterative Landweber algorithm that provides a solution. In the chapters that follow, we shall study in some detail the many ways this problem can be generalized and the variety of iterative algorithms that extend the Landweber method. In this beginning chapter we present several of the fundamental definitions and results pertaining to matrix algebra that we shall use later.

Occasionally in this book we make reference to [46, 47]. These books are available as pdf documents on my website.

1.2 Matrix Rank and Invertibility

Throughout this book we shall denote by A an I by J matrix with real entries. The extension of definitions and results to the complex case is usually straightforward and we shall omit it without comment. Unless otherwise indicated, the symbol $\|x\|$ will denote the Euclidean norm of $x \in \mathbb{R}^J$, also called the two-norm of x, given by

$$\|x\|_2 = \sqrt{\sum_{j=1}^{J} x_j^2}. \tag{1.1}$$

We shall also have occasion to use the one-norm of x, given by

$$\|x\|_1 = \sum_{j=1}^{J} |x_j|. \tag{1.2}$$

From the definition of Ax, the vector resulting from the multiplication of a vector x by a matrix A, we can easily see that the vector Ax is a linear combination of the columns of A; specifically, we have

$$Ax = \sum_{j=1}^{J} x_j a^j, \tag{1.3}$$

where x_j is the jth entry of the vector x and a^j is the jth column of A. From the definition of matrix multiplication of matrices B and C, the jth column of the product $A = BC$ is a linear combination of the columns of B, with coefficients the entries of the jth column of C.

Definition 1.1 *The* row rank *of A, denoted rr(A), is the maximum number of linearly independent rows of A, and the* column rank *of A, denoted cr(A), is the maximum number of linearly independent columns of A.*

Clearly, the row rank of A equals the column rank of its transpose, A^T.

Proposition 1.1 *The row rank and the column rank of any A are the same.*

Proof: Suppose that the column rank of A is K. Form the I by K matrix B having for its columns K linearly independent columns of A. Then every column of A is a linear combination of the columns of B, which means that there is a K by J matrix C such that $A = BC$. Then $A^T = C^T B^T$, which tells us that each column of A^T is a linear combination of the K columns of C^T. Consequently, there can be at most K linearly independent columns of A^T. So the column rank of A^T, which is the row rank of A, cannot exceed the column rank of A. But since this must be true for all choices of A, the assertion of the proposition follows. ∎

Definition 1.2 *The* rank *of A, denoted r(A), is its column rank.*

Definition 1.3 *A square matrix S is* invertible *if there is a matrix R such that $SR = RS = I$, where I denotes the identity matrix of the appropriate size. Then R is the* inverse *of S and we write $R = S^{-1}$.*

Definition 1.4 *The linear transformation $A : \mathbb{R}^J \to \mathbb{R}^I$ induced by multiplication by A, which we also denote by A, is* onto $\mathbb{R}^I$*, or simply* onto*, if, for every $b \in \mathbb{R}^I$, there is $x \in \mathbb{R}^J$ with $b = Ax$.*

Lemma 1.1 *Let S be any square matrix. The linear transformation S is onto if and only if the linear transformation S^T is onto.*

Proof: Suppose S is I by I and S is onto $\mathbb{R}^I$. Then the columns of S span all of $\mathbb{R}^I$ and so the rank of S is I. Therefore, the rank of S^T is also I, its columns also span all of $\mathbb{R}^I$, and S^T is onto. ∎

Proposition 1.2 *The matrix S invertible if and only if the induced linear transformation S is onto $\mathbb{R}^I$.*

Proof: If S is invertible and b is any vector in $\mathbb{R}^I$, then $x = S^{-1}b$ satisfies $b = Sx$; therefore the transformation is onto $\mathbb{R}^I$. Conversely, if the transformation is onto $\mathbb{R}^I$, then for each i there is a vector r^i such that $Sr^i = \delta^i$, where δ^i denotes the ith column of the identity matrx I. The matrix R whose ith column is r^i then satisfies $SR = I$. Because S^T is also onto, there is a matrix U such that $S^TU = I$ and so $U^TS = I$. Then

$$U^T = U^T(SR) = (U^TS)R = IR = R$$

and $RS = I$. So $R = S^{-1}$. ∎

Corollary 1.1 *The matrix S is invertible if and only if the rank of S is I.*

Proposition 1.3 *Let B be I by K, with rank K. Then B^TB also has rank K and is therefore invertible.*

Proof: If $B^TBu = 0$, then $0 = u^TB^TBu = \|Bu\|_2^2$, so that $Bu = 0$. But since the rank of B is K, the columns of B are linearly independent. Consequently, $u = 0$. It follows that the columns of B^TB must be linearly independent, and its rank is then K. ∎

This brings us to the main result of this chapter.

1.3 The Fundamental Decomposition Theorem

We begin with two definitions.

Definition 1.5 *The* range *of A, denoted $R(A)$, is the span of the columns of A in $\mathbb{R}^I$.*

Definition 1.6 *The* null space *of A^T, denoted $NS(A^T)$, is the set of all $w \in \mathbb{R}^I$ such that $A^Tw = 0$.*

Lemma 1.2 *If $b \in R(A) \cap NS(A^T)$, then $b = 0$.*

Proof: If $b = Ax$, and $A^Tb = A^TAx = 0$, then $x^TA^TAx = \|Ax\|_2^2 = 0$, so $b = Ax = 0$. ∎

Corollary 1.2 *If A is onto, then $NS(A^T) = \{0\}$.*

Proof: Let $w \in NS(A^T)$. Since A is onto, we also have $w \in R(A)$. Therefore $w = 0$. ∎

Once we prove the Fundamental Decomposition Theorem we will be able to prove the converse of this corollary.

Theorem 1.1 The Fundamental Decomposition Theorem *Every* b *in* $\mathbb{R}^I$ *can be written uniquely as*

$$b = Ax + w,$$

for some $x \in \mathbb{R}^J$ *and* $w \in NS(A^T)$.

Proof:
If A^TA is invertible, then we take $w = b - A(A^TA)^{-1}A^Tb$, which is clearly in $NS(A^T)$. We then have $b = Ax + w$, for $x = (A^TA)^{-1}A^Tb$.

Now suppose that A^TA is not invertible, and the rank of A is K. Form the I by K matrix B as previously, so that the rank of B is K, B^TB is invertible, and $R(B) = R(A)$. From the previous paragraph, we know that $b = Bz + v$, where v is in the null space of B^T, and Bz is in $R(B) = R(A)$. Since $A = BC$, we have $A^T = C^TB^T$. It follows that v must be in the null space of A^T. ∎

Remark: Note that w and Ax are unique, but x need not be.

Corollary 1.3 *If* $NS(A^T) = \{0\}$, *then* A *is onto.*

Proof: According to Theorem 1.1, if b is not in $R(A)$, then there is $w \neq 0$ in $NS(A^T)$ and $x \in \mathbb{R}^J$ with $b = Ax + w$. ∎

Corollary 1.4 *The system of linear equations* $A^TAx = A^Tb$ *always has solutions.*

Proof: From Theorem 1.1 we know that $b = Ax + w$, with $A^Tw = 0$. Therefore, $A^Tb = A^TAx$. ∎

When the system $Ax = b$ has no solution, which means A is not onto, we often seek a *least squares solution*, say x_{LS}, which is an exact solution to $A^TAx = A^Tb$, which, as we just saw, always has exact solutions. In fact, we always have $b = Ax_{LS} + w$. If $Ax = b$ does have exact solutions, then $w = 0$, so any least squares solution is an exact solution of $Ax = b$.

When A^TA is invertible, the unique least squares solution is

$$x_{LS} = (A^TA)^{-1}A^Tb. \tag{1.4}$$

When $Ax = b$ has exact solutions, and AA^T is invertible, the unique solution with the smallest two-norm is the *minimum norm solution*, x_{MN}, given by

$$x_{MN} = A^T(AA^T)^{-1}b. \tag{1.5}$$

For large A even calculating A^TA or AA^T is out of the question and we must turn to iterative methods to find the desired solution. The Landweber algorithm, which we take up in the next chapter, is an iterative method for solving these problems.

Chapter 2

Landweber's Algorithm and Beyond

2.1 Landweber's Algorithm

Having chosen x^0 and having calculated x^n, the next iterate given by Landweber's algorithm [100] is

$$x^{n+1} = x^n + \gamma A^T(b - Ax^n), \tag{2.1}$$

where $0 < \gamma < 2/\rho(A^TA)$. The *spectral radius* of a square matrix S, denoted $\rho(S)$, is the largest absolute value of any eigenvector of S. Since $S = A^TA$ is non-negative definite, all its eigenvalues are non-negative and $\rho(A^TA)$ is the largest eigenvalue of A^TA.

Theorem 2.1 *The sequence $\{x^n\}$ obtained using Landweber's algorithm converges to the solution of $A^TAx = A^Tb$ for which $\|x - x^0\|_2$ is minimized.*

Proof: Let $A^TAz = A^Tb$. Then

$$\|z - x^{n+1}\|_2^2 = \|z - x^n\|^2 + 2\langle z - x^n, x^n - x^{n+1}\rangle + \|x^n - x^{n+1}\|_2^2.$$

From

$$\|x^n - x^{n+1}\|_2^2 = \gamma^2\|A^T(Az - Ax^n)\|_2^2$$

and

$$\langle z - x^n, x^n - x^{n+1}\rangle = \gamma\langle z - x^n, A^T(Ax^n - b)\rangle$$

$$= \gamma\langle z - x^n, A^T(Ax^n - Az)\rangle = \gamma\langle Az - Ax^n, Ax^n - Az\rangle = -\gamma\|Az - Ax^n\|_2^2$$

we have

$$\|z - x^n\|_2^2 - \|z - x^{n+1}\|_2^2 = 2\gamma\|Az - Ax^n\|_2^2 - \gamma^2\|A^T(Az - Ax^n)\|_2^2. \tag{2.2}$$

Since

$$\|A^T(Az - Ax^n)\|_2^2 \leq \rho(A^TA)\|Az - Ax^n\|_2^2,$$

we have

$$\|z - x^n\|_2^2 - \|z - x^{n+1}\|_2^2 \geq 2\gamma\|Az - Ax^n\|_2^2 - \gamma^2\rho(A^TA)\|Az - Ax^n\|_2^2.$$

Because γ is chosen so that $0 < \gamma < 2/\rho(A^TA)$, we have

$$2\gamma - \gamma^2\rho(A^TA) > 0.$$

We now know that the sequence $\{\|z - x^n\|_2\}$ is decreasing, so that the sequence $\{\|Az - Ax^n\|_2\}$ converges to zero. The sequence $\{x^n\}$ is then bounded and a subsequence converges to some vector x^*. It follows that $Ax^* = Az$. Replacing the generic z with x^* in the calculations above, we conclude that the sequence $\{\|x^* - x^n\|_2\}$ is decreasing, but must actually converge to zero, since a subsequence converges to zero. So the sequence $\{x^n\}$ converges to a least squares solution of $Ax = b$. We can say more, though.

Since the right side of Equation (2.2) depends on z only through Az, so does the left side. Summing the left side over $n = 0, 1, ...$, we find that $\|z - x^0\|_2^2 - \|z - x^*\|_2^2$ depends only on $Az = Ax^*$, and not on any particular z. Consequently, minimizing $\|z - x^0\|_2$ over all least squares solutions z is equivalent to minimizing $\|z - x^*\|_2$ over all such z; but the solution to the latter problem is obviously $z = x^*$. This concludes the proof of the theorem. ∎

There are several reasons for giving a complete proof of convergence of the Landweber algorithm at this early stage. First, it lets us see where the condition on γ is used. Second, this proof will be used later as a template for proofs of convergence of related algorithms. Finally, Theorem 1.1 will be obtained later as a corollary of more general theorems. Therefore, it is good to see an elementary proof of convergence first.

2.2 Beyond Landweber

Our goal, in this book, is to present a variety of problems and algorithms that emerge from solving $Ax = b$ using Landweber's algorithm. To motivate these generalizations, it is helpful to have several ways to view Landweber's algorithm.

2.2.1 Projection onto Hyperplanes

Whenever we use the least squares solution of $Ax = b$ we should first normalize each equation by dividing both sides by the length of that row of A. Then we have

$$\sum_{j=1}^{J} A_{ij}^2 = 1,$$

for all i. The problem is that the least squares solutions, unlike exact solutions, are dependent on the scaling of each equation. For example, the system $x = 1$ and $x = 2$ has $x_{LS} = 1.5$, while the system $2x = 2$ and $x = 2$ has $x_{LS} = 1.2$. The normalization also simplifies notation. We shall assume, throughout this book, and without further comment, that the system $Ax = b$ has been normalized.

For each i, the hyperplane in $\mathbb{R}^J$ corresponding to the ith equation in $Ax = b$ is

$$H_i := \{x|(Ax)_i = b_i\}. \tag{2.3}$$

For any $z \in \mathbb{R}^J$ the unique member of H_i closest to z, in the two-norm sense, is denoted $P_{H_i}z$, or just P_iz if the context is clear, and is given in closed form by

$$P_iz := z + (b_i - (Az)_i)a^i, \tag{2.4}$$

where a^i denotes the ith column of the matrix A^T. The operator $P_i : \mathbb{R}^J \to \mathbb{R}^J$ is called the *orthogonal projection* onto H_i.

Cimmino's algorithm [64] for finding a least squares solution of $Ax = b$ has the iterative step

$$x^{n+1} = \frac{1}{I}\sum_{i=1}^{I} P_ix^n. \tag{2.5}$$

Cimmino's idea is to find the vectors closest to x^n in each of the hyperplanes, and then to average them to get x^{n+1}. Landweber's algorithm can be written as

$$x^{n+1} = \gamma\sum_{i=1}^{I} P_ix^n. \tag{2.6}$$

Since A is normalized, the traces of A^TA and AA^T are equal to I and so $\rho(A^TA) \leq I$. Therefore, Cimmino's choice of $\gamma = \frac{1}{I}$ is acceptable. Loosely speaking, the larger the γ the larger the iterative step in Landweber's algorithm, and the sooner the algorithm converges.

For large matrices A, calculating the largest eigenvalue of A^TA directly, or even calculating A^TA, is too expensive and time-consuming, so we must estimate $\rho(A^TA)$. Because the spectral radius $\rho(A^TA)$ governs the choice of γ, it is helpful to have a good estimate of $\rho(A^TA)$. Estimating $\rho(A^TA)$ by the trace of A^TA, as in Cimmino's algorithm, is quite conservative in general, especially for large matrices. Many of the matrices A that we encounter in image processing and remote sensing are *sparse*, meaning that most of their entries are zero. As we shall see in a later chapter, there are better estimates of $\rho(A^TA)$ that make use of the sparseness of A.

If, for each i, C_i is a non-empty closed, convex subset of $\mathbb{R}^J$, then the orthogonal projection onto C_i is well defined and we could attempt to find a member of the intersection of the C_i by using these operators instead of the P_i in Landweber or ART. The orthogonal projections onto closed convex sets are particularly nice operators and useful classes of operators are found by generalizing these projection operators.

2.2.2 Using Blocks

Expressing Landweber's algorithm in terms of the orthogonal projections onto the hyperplanes H_i suggests several possible generalizations that we shall consider in more detail in subsequent chapters. Instead of calculating all the orthogonal projections of x^n, we could calculate only a single projection at each step; then we would cycle through each hyperplane in turn. The resulting algorithm, known as Kaczmarz's Algorithm [96], or the Algebraic Reconstruction Technique (ART) [87], is a *sequential* or *row-action* method.

We shall also consider more general *block-iterative* versions of Landweber. We decompose the set $\{i = 1, ..., I\}$ into M blocks; that is, into M (not necessarily disjoint) subsets B_m, $m = 1, ..., M$. For $n = 0, 1, ...$, we set $m = m(n) = n(\bmod M) + 1$. For block-iterative Cimmino, having found x^n, we take x^{n+1} to be the average of the orthogonal projections of x^n onto the hyperplanes H_i for which $i \in B_m$. The ART is a special case of this approach in which each block consists of only a single value of i. Block-iterative Landweber makes use of more general step-length parameters.

2.2.3 Function Minimization

A least squares solution of $Ax = b$ is a minimizer of the function

$$f(x) = \frac{1}{2}\|Ax - b\|_2^2. \tag{2.7}$$

The gradient of $f(x)$ is

$$\nabla f(x) = A^T(Ax - b). \tag{2.8}$$

Therefore, Landweber's algorithm can be written as

$$x^{n+1} = x^n - \gamma\nabla f(x^n); \tag{2.9}$$

this is the form of a gradient descent algorithm with a fixed *step-length* parameter. This suggests studying the properties of the operator

$$Tx = x - \gamma \nabla f(x), \tag{2.10}$$

for any differentiable convex function $f(x)$ to see when the iterative scheme $x^{n+1} = Tx^n$ converges to a fixed point of T. If $Tx^* = x^*$, then $\nabla f(x^*) = 0$ and x^* is a global minimizer of $f(x)$.

2.2.4 Fixed-Point Iteration

As we have just seen, the Landweber algorithm can be viewed as a gradient descent algorithm for the function given in Equation (2.7). The sequence $\{x^n\}$ converges to a vector x^* for which $\nabla f(x^*) = 0$. Therefore $Tx^* = x^*$ for the operator T in Equation (2.10). Often we find that the solution to our problem is a vector x^* that is a fixed point of some operator T; that is, $Tx^* = x^*$. When we search for fixed points of T it is natural to consider an iterative algorithm of the form $x^{n+1} = Tx^n$. Of course, unless T enjoys some special properties, the sequence $\{x^n\}$ need not converge.

Definition 2.1 *An operator* $T : \mathbb{R}^J \to \mathbb{R}^J$ *is* non-expansive, *with respect to a given norm* $\| \cdot \|$ *if, for all* x *and* y, *we have*

$$\|Tx - Ty\| \leq \|x - y\|.$$

You might think that, if T is non-expansive, then the sequence $\{x^{n+1} = Tx^n\}$ will converge, but this is not always the case. Let $T = -I$, the negative of the identity operator, and $x^0 \neq 0$; then we have $x^n = (-1)^n x^0$, which does not converge, even though T has a fixed point, namely $x = 0$. We shall need to consider properties of operators that are stronger than being non-expansive, in order to get convergent iterative sequences.

Later we shall show that the orthogonal projection operator onto C, which we shall denote by P_C, is a *firmly non-expansive* operator. If T is any firmly non-expansive, then the iterative sequence defined by $x^{n+1} = Tx^n$ converges to a fixed point of T, whenever fixed points of T exist. However, the product of two or more firmly non-expansive operators need not be firmly non-expansive. For example, if C_1 and C_2 are two closed convex subsets of $\mathbb{R}^J$, and P_{C_1} and P_{C_2} the associated orthogonal projection operators, the operator $T = P_{C_2} P_{C_1}$ need not be firmly non-expansive. As we shall see, convergence does hold for the class of *averaged* operators, which is a subclass of the operators that are non-expansive in the two-norm, and contains the firmly non-expansive operators. Products of averaged operators are again averaged, making this class of operators nearly ideal for fixed-point iteration.

If $f : \mathbb{R}^J \to \mathbb{R}$ is convex and differentiable, and if

$$\|\nabla f(x) - \nabla f(y)\|_2 \leq L\|x - y\|_2$$

for all x and y, then the gradient of f is said to be L-Lipschitz continuous. In that case the operator

$$Tx = x - \gamma \nabla f(x)$$

is firmly non-expansive, whenever $0 < \gamma < 2/L$.

2.2.5 Alternating Minimization

The Landweber algorithm can also be obtained through the alternating minimization (AM) approach of Csiszár and Tusnády [69]. Let $\Theta : \mathbb{R}^J \times \mathbb{R}^J \to \mathbb{R}$. The AM approach is to minimize $\Theta(x^n, y)$ to get $y = y^n$, and then minimize $\Theta(x, y^n)$ to get $x = x^{n+1}$. To get the Landweber sequence we define $\Theta(x, y)$ by

$$\Theta(x, y) = \gamma\|Ax - b\|_2^2 + \frac{1}{2}\|x - y\|_2^2 - \gamma\|Ax - Ay\|_2^2. \tag{2.11}$$

Suppose that $b = \inf \Theta(x, y) > -\infty$. The main objective of any AM algorithm is to have $\Theta(x^n, y^n) \to b$, although this need not happen without further restrictions on Θ.

2.2.6 Sequential Unconstrained Minimization

We can also obtain the Landweber algorithm through *sequential unconstrained minimization*. At the nth step of a sequential unconstrained minimization approach to minimizing a function $f(x)$, we obtain x^n by minimizing

$$G_n(x) = f(x) + g_n(x), \tag{2.12}$$

where $g_n(x)$ is an appropriately chosen auxiliary function. These auxiliary functions are often selected to impose constraints on the x, but can also be used to permit x^n to be calculated in closed form, as is the case with the Landweber algorithm. To get the Landweber sequence we define

$$G_n(x) = \gamma\|Ax - b\|_2^2 + \frac{1}{2}\|x - x^n\|_2^2 - \gamma\|Ax - Ax^n\|_2^2. \tag{2.13}$$

Suppose that $b := \inf f(x) > -\infty$. The main objective of any sequential unconstrained minimization algorithm is to have $f(x^n) \to b$. In order to achieve this objective, additional conditions have to be placed on the auxiliary functions $g_n(x)$.

The SUMMA class of sequential unconstrained minimization algorithms was defined in [43]. It turns out that this class is quite large and includes many of the popular sequential unconstrained minimization algorithms. To be in the SUMMA class it is required that, for all x,

$$G_n(x) - G_n(x^n) \geq g_{n+1}(x) \geq 0. \tag{2.14}$$

All algorithms in the SUMMA class have $f(x^n) \to b$.

2.2.7 Other Distances

The least squares solution of $Ax = b$ minimizes the distance from Ax to b, with respect to the two-norm. It is sometimes helpful to use other measures of distance, such as the cross-entropy, or Kullback-Leibler [99], distance between non-negative vectors.

For $a >$ and $b > 0$ we define the Kullback-Leibler (KL) distance from a to b by

$$KL(a, b) := a \log \frac{a}{b} + b - a. \tag{2.15}$$

Taking limits, we find that $KL(0, b) = b$ and $KL(a, 0) = +\infty$. We extend the KL distance to non-negative vectors component-wise, so that

$$KL(x, z) := \sum_{j=1}^{J} KL(x_j, z_j). \tag{2.16}$$

Note that $KL(x, z)$ and $KL(z, x)$ are not generally the same. The KL distance is not a metric in the formal sense, but it does have a number of convenient properties, as we shall see later.

Let $b \in \mathbb{R}^I$ be a positive vector, and A an I by J matrix whose entries are non-negative. We can then find exact or approximate non-negative solutions of $Ax = b$ by minimizing either $KL(Ax, b)$ or $KL(b, Ax)$. We shall derive algorithms for solving both of these problems in later chapters.

Both the (square of the) two-norm and the KL distance are *Bregman distances*. For convex differentiable functions $f : \mathbb{R}^J \to \mathbb{R}$ the Bregman distance [12, 19] from x to z is

$$D_f(x, z) := f(x) - f(z) - \langle \nabla f(z), x - z \rangle. \tag{2.17}$$

The square of the two-norm comes from the function $f(x) = \frac{1}{2}\|x\|_2^2$, while the KL distance comes from the function $f(x) = \sum_{j=1}^{J} x_j \log(x_j) - x_j$. In later chapters we shall extend to Bregman distances algorithms using the two-norm and the KL distance.

2.3 Summing Up

We have now seen several different ways to view the problem of solving $Ax = b$ and the Landweber algorithm: using orthogonal projections onto hyperplanes, either simultaneously or sequentially; minimizing a function by gradient descent; iterating to a fixed point of an operator; alternating minimization; and sequential unconstrained minimization. As we shall see in subsequent chapters, each of these approaches leads to other problems and to other algorithms. Several times during our discussion of these topics we shall return to the original problem treated in this chapter, to motivate a definition, or to suggest a proof.

Chapter 3

Block-iterative Landweber

3.1 Block-Iterative Variants of Landweber

Now we consider in some detail solving the problem $Ax = b$ using *block-iterative* versions of Landweber's algorithm (BILW). We decompose the set $\{i = 1, ..., I\}$ into M blocks; that is, into M (not necessarily disjoint) subsets B_m, $m = 1, ..., M$. For $n = 0, 1, ...,$ we set $m = m(n) = n(\text{mod}\, M)+1$. Denote by I_m the cardinality of B_m. Having found x^n, we take x^{n+1} to be the average of the orthogonal projections of x^n onto the hyperplanes H_i for which $i \in B_m$. The ART is a special case of this approach in which each block consists of only a single value of i. Block-iterative versions of Landweber converge to a solution whenever $Ax = b$ has solutions. However, when $Ax = b$ is inconsistent, each subsequence $\{x^{kM+m}\}$ converges to some $x^{*,m}$, but generally, these $x^{*,m}$ are distinct and form what is called a *limit cycle*.

Later we shall study block-iterative versions of other algorithms. The main idea is simply to use, at each step, some, but not all, of the equations in the system to be solved.

3.2 The Algebraic Reconstruction Technique

The *algebraic reconstruction technique* (ART) [87], also known as Kaczmarz's Algorithm [96], is a *sequential* or *row-action* variant of Landweber's algorithm, in which only a single equation in the system is used at each step of the iteration.

The iterative step of the ART is

$$x^{n+1} = x^n + (b_i - (Ax^n)_i)a^i, \tag{3.1}$$

where a^i denotes the ith column of the matrix A^T. We remind the reader that we are assuming that the system has been normalized, so $\|a^i\|_2 = 1$.

Because the ART is a special case of the more general block-iterative Landweber algorithms, convergence of the ART, when $Ax = b$ is consistent, is a consequence of Theorem 3.1 in the following section.

When the system $Ax = b$ has no solution, the ART does not converge to a single vector, but the subsequences $\{x^{kI+i}\}$ converge to vectors $x^{*,i}$ comprising a *limit cycle* [126]. To avoid the limit cycle in the inconsistent case, we can apply the ART to the consistent linear system

$$\begin{bmatrix} A & I \\ 0 & A^T \end{bmatrix} \begin{bmatrix} x \\ w \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix}; \tag{3.2}$$

the resulting x is the least squares solution closest to x^0.

3.3 Convergence of the BILW Algorithm

For $m = 1, 2, ..., M$, denote by A_m the I_m by J matrix obtained from A by removing the ith row for any i not in B_m; similarly form b^m from b.

For $n = 0, 1, ...,$ and $m = n(\operatorname{mod} M) + 1$, let

$$x^{n+1} = x^n + \gamma_m A_m^T (b^m - A_m x^n). \tag{3.3}$$

This is the iterative step of the BILW algorithm. The main convergence theorem for the BILW algorithm is the following.

Theorem 3.1 *Let $Ax = b$ have solutions. For each m, let γ_m be chosen so that*

$$0 < \gamma_m < 2/\rho(A_m^T A_m).$$

For any choice of blocks, the sequence $\{x^n\}$ formed using Equation (3.3) converges to the solution of $Ax = b$ for which $\|x - x^0\|_2$ is minimized.

Proof: Let $Az = b$. Then

$$\|z - x^n\|_2^2 - \|z - x^{n+1}\|_2^2$$

$$= 2\gamma_m \langle z - x^n, A_m^T (b^m - A_m x^n)\rangle - \gamma_m^2 \|A_m^T (b^m - A_m x^n)\|_2^2$$

$$= 2\gamma_m \|b^m - A_m x^n\|_2^2 - \gamma_m^2 \|A_m^T (b^m - A_m x^n)\|_2^2.$$

Therefore, we have

$$\|z - x^n\|_2^2 - \|z - x^{n+1}\|_2^2 \geq (2\gamma_m - \gamma_m^2 \rho(A_m^T A_m)) \|b^m - A_m x^n\|_2^2. \tag{3.4}$$

It follows that the sequence $\{\|z - x^n\|_2^2\}$ is decreasing and that the sequence $\{\|b^m - A_m x^n\|_2^2\}$ converges to zero. The sequence $\{x^n\}$ is then bounded;

let $x^* = x^{*,0}$ be any cluster point of the subsequence $\{x^{kM}\}$. Then, for $m = 1, ..., M$, let

$$x^{*,m} = x^{*,m-1} + \gamma_m A_m^T (b^m - A_m x^{*,m-1}).$$

It follows that $x^{*,m} = x^*$ for all m and that $Ax^* = b$. Replacing the arbitrary solution z with x^*, we find that the sequence $\{\|x^* - x^n\|_2^2\}$ is decreasing; but a subsequence converges to zero. Consequently, the sequence $\{\|x^* - x^n\|_2^2\}$ converges to zero. We can therefore conclude that the sequence $\{x^n\}$ converges to a solution, whenever the system $Ax = b$ is consistent. In fact, since we have shown that the difference $\|z-x^n\|_2^2-\|z-x^{n+1}\|_2^2$ is nonnegative and independent of the solution z that we choose, we know that the difference $\|z - x^0\|_2^2 - \|z - x^*\|_2^2$ is also nonnegative and independent of z. It follows that $z = x^*$ is the solution that minimizes $\|z - x^0\|$. ∎

Chapter 4

Convex Feasibility

4.1 The Convex Feasibility Problem

For each $i = 1, ..., I$, let C_i be a non-empty, closed and convex subset of $\mathbb{R}^J$, and C the (possibly empty) intersection of the C_i. The *convex feasibility problem* (CFP) is to find a member of C, if there are any. When C is empty, we find an approximate solution by minimizing the function $F(x)$ given by

$$F(x) = \frac{1}{2I} \sum_{i=1}^{I} \|x - P_{C_i} x\|_2^2. \tag{4.1}$$

In this chapter we show how, by replacing orthogonal projection onto hyperplanes with orthogonal projection onto closed convex sets, the Landweber and ART algorithms can be extended to solve these problems. The algorithms we shall study are the *successive orthogonal projection* (SOP) and *simultaneous orthogonal projection* (SIMOP) methods. We begin with a discussion of the basic properties of orthogonal projection onto convex sets.

4.2 Orthogonal Projection onto Convex Sets

The *Parallelogram Law* is an easy consequence of the definition of the two-norm:

$$||x + y||_2^2 + ||x - y||_2^2 = 2||x||_2^2 + 2||y||_2^2. \tag{4.2}$$

It is important to remember that Cauchy's Inequality and the Parallelogram Law hold only for the two-norm. One consequence of the Parallelogram Law that we shall need is the following: if $x \neq y$ and $\|x\|_2 = \|y\|_2 = d$, then $\|\frac{1}{2}(x + y)\|_2 < d$ (Draw a picture!).

The following proposition is fundamental in the study of convexity and can be found in most books on the subject.

Proposition 4.1 *Given any nonempty closed convex set $C \in \mathbb{R}^J$ and an arbitrary vector x in $\mathbb{R}^J$, there is a unique member $P_C x$ of C closest, in the sense of the two-norm, to x. The vector $P_C x$ is called the* orthogonal (or metric) projection *of x onto C and the operator P_C the* orthogonal projection *onto C.*

Proof: If x is in C, then $P_C x = x$, so assume that x is not in C. Then $d > 0$, where d is the distance from x to C. For each positive integer n, select c^n in C with $||x - c^n||_2 < d + \frac{1}{n}$. Then, since for all n we have

$$\|c^n\|_2 = \|c^n - x + x\|_2 \leq \|c^n - x\|_2 + \|x\|_2 \leq d + \frac{1}{n} + \|x\|_2 < d + 1 + \|x\|_2,$$

the sequence $\{c^n\}$ is bounded; let c^* be any cluster point. It follows easily that $||x - c^*||_2 = d$ and that c^* is in C. If there is any other member c of C with $||x - c||_2 = d$, then, by the Parallelogram Law, we would have $||x - (c^* + c)/2||_2 < d$, which is a contradiction. Therefore, c^* is $P_C x$. ∎

For an arbitrary nonempty closed convex set C in $\mathbb{R}^J$, the orthogonal projection $T = P_C$ is a nonlinear operator, unless, of course, C is a subspace. We may not be able to describe $P_C x$ explicitly, but we do know a useful property of $P_C x$.

Proposition 4.2 *For a given x, a vector z in C is $P_C x$ if and only if*

$$\langle c - z, z - x \rangle \geq 0, \tag{4.3}$$

for all c in the set C.

Proof: Let c be arbitrary in C and α in $(0, 1)$. Then

$$||x - P_C x||_2^2 \leq ||x - (1 - \alpha)P_C x - \alpha c||_2^2 = ||x - P_C x + \alpha(P_C x - c)||_2^2$$

$$= ||x - P_C x||_2^2 - 2\alpha\langle x - P_C x, c - P_C x\rangle + \alpha^2||P_C x - c||_2^2. \tag{4.4}$$

Therefore,

$$-2\alpha\langle x - P_C x, c - P_C x\rangle + \alpha^2||P_C x - c||_2^2 \geq 0, \tag{4.5}$$

so that

$$2\langle x - P_C x, c - P_C x\rangle \leq \alpha||P_C x - c||_2^2. \tag{4.6}$$

Taking the limit, as $\alpha \to 0$, we conclude that

$$\langle c - P_C x, P_C x - x\rangle \geq 0. \tag{4.7}$$

If z is a member of C that also has the property

$$\langle c - z, z - x\rangle \geq 0, \tag{4.8}$$

for all c in C, then we have both

$$\langle z - P_C x, P_C x - x\rangle \geq 0, \tag{4.9}$$

and

$$\langle z - P_C x, x - z\rangle \geq 0. \tag{4.10}$$

Adding on both sides of these two inequalities lead to

$$\langle z - P_C x, P_C x - z\rangle \geq 0. \tag{4.11}$$

But,

$$\langle z - P_C x, P_C x - z\rangle = -||z - P_C x||_2^2, \tag{4.12}$$

so it must be the case that $z = P_C x$. This completes the proof. ∎

Corollary 4.1 *For any x and y in $\mathbb{R}^J$ we have*

$$\langle P_C x - P_C y, x - y\rangle \geq \|P_C x - P_C y\|_2^2. \tag{4.13}$$

Proof: Use Inequality (4.2) to get

$$\langle P_C y - P_C x, P_C x - x\rangle \geq 0, \tag{4.14}$$

and

$$\langle P_C x - P_C y, P_C y - y\rangle \geq 0. \tag{4.15}$$

Add the two inequalities to obtain

$$\langle P_C x - P_C y, x - y\rangle \geq ||P_C x - P_C y||_2^2. \tag{4.16}$$

∎

4.3 The SOP Algorithm

The successive orthogonal projection (SOP) algorithm is a generalization of the ART and has the following iterative step:

$$x^{n+1} = P_{C_i} x^n, \tag{4.17}$$

where $i = n(\text{mod}\, I) + 1$. The main convergence theorem for the SOP is the following:

Theorem 4.1 *If C is not empty, then the sequence $\{x^n\}$ converges to a member of C.*

Proof: Let z be a member of C. Then

$$\|z - x^n\|_2^2 = \|z - P_{C_i}x^n + P_{C_i}x^n - x^n\|_2^2$$

$$= \|z - x^{n+1}\|_2^2 + 2\langle P_{C_i}x^n - x^n, z - P_{C_i}x^n\rangle + \|P_{C_i}x^n - x^n\|_2^2.$$

Since

$$\langle P_{C_i}x^n - x^n, z - P_{C_i}x^n\rangle \geq 0,$$

we have

$$\|z - x^n\|_2^2 - \|z - x^{n+1}\|_2^2 \geq \|x^{n+1} - x^n\|_2^2.$$

Then, for each i and $m = 0, 1, ...,$

$$\|z - x^{mI}\|_2^2 - \|z - x^{(m+1)I}\|_2^2 \geq \sum_{i=1}^{I} \|x^{mI+i} - x^{mI+(i-1)}\|_2^2.$$

Therefore, the sequence $\{\|z-x^{mI}\|_2^2\}$ is decreasing, the sequence $\{\|x^{mI+i} - x^{mI+(i-1)}\|_2^2\}$ converges to zero, the sequence $\{x^{mI}\}$ is bounded, and a subsequence converges to some $x^{*,0}$. With $x^{*,i} = P_{C_i}x^{*,i-1}$ for $i = 1, ..., I$, we find that $x^{*,i} = x^{*,i-1} := x^*$ for all i, and x^* is a member of C. Replacing the generic z with x^*, we conclude that the sequence $\{x^n\}$ converges to x^*. ∎

4.4 The SIMOP Algorithm

The SOP algorithm uses one convex set C_i at each step. The SIMOP algorithm uses all the C_i at each step. To prove convergence of the SIMOP we need to consider the class of firmly non-expansive (fne) operators. Corollary 4.1 says that the operator P_C is fne; we need to know that the convex combination of fne operators is again fne.

It can be shown that, for any non-empty closed convex set C, the function

$$f(x) = \frac{1}{2}\|x - P_C x\|_2^2$$

is convex and differentiable, and the gradient is the operator

$$\nabla f(x) = x - P_C x;$$

for details see [47]. Therefore, the function $F(x)$ in Equation (4.1) is convex and differentiable and

$$\nabla F(x) = \frac{1}{I}\sum_{i=1}^{I}(x - P_{C_i}x) = x - \frac{1}{I}\sum_{i=1}^{I} P_{C_i}x. \tag{4.18}$$

As we shall show, the SIMOP iterative sequence generated by

$$x^{n+1} = x^n - \nabla F(x^n) \tag{4.19}$$

converges to a minimizer of the function $F(x)$ whenever minimizers exist. To prove this we need to investigate properties of the operator $T = I - \nabla F$.

The operator T can be written as

$$Tx = \frac{1}{I}\sum_{i=1}^{I} P_{C_i}x. \tag{4.20}$$

We shall show that T is firmly non-expansive. Since each of the operators P_{C_i} is firmly non-expansive, we need only show that the convex combination of firmly non-expansive operators is again firmly non-expansive.

4.5 Firmly Non-Expansive Operators

Definition 4.1 *An operator $T : \mathbb{R}^J \to \mathbb{R}^J$ is* non-expansive *(ne) with respect to a given norm $\|\cdot\|$ on $\mathbb{R}^J$ if, for all x and y,*

$$\|Tx - Ty\| \leq \|x - y\|. \tag{4.21}$$

Definition 4.2 *An operator $T : \mathbb{R}^J \to \mathbb{R}^J$ is* firmly non-expansive *(fne) if, for all x and y,*

$$\langle Tx - Ty, x - y\rangle - \|Tx - Ty\|_2^2 \geq 0. \tag{4.22}$$

By Cauchy's Inequality, every fne operator is ne in the two-norm. Corollary 4.1 tells us that the operator $T = P_C$ is fne and therefore ne in the two-norm. We show now that the convex combination of fne operators is again fne, so that the operator $T = \frac{1}{I}\sum_{i=1}^{I} P_{C_i}$ is fne.

Proposition 4.3 *An operator $F : \mathbb{R}^J \to \mathbb{R}^J$ is fne if and only if $F = \frac{1}{2}(I + N)$, for some operator N that is ne with respect to the two-norm.*

Proof: Suppose that $F = \frac{1}{2}(I + N)$. We show that F is fne if and only if N is ne in the two-norm. First, we have

$$\langle Fx - Fy, x - y\rangle = \frac{1}{2}\|x - y\|_2^2 + \frac{1}{2}\langle Nx - Ny, x - y\rangle.$$

Also,

$$\|\frac{1}{2}(I+N)x - \frac{1}{2}(I+N)y\|_2^2 = \frac{1}{4}\|x-y\|^2 + \frac{1}{4}\|Nx-Ny\|^2 + \frac{1}{2}\langle Nx-Ny, x-y\rangle.$$

Therefore,

$$\langle Fx - Fy, x - y\rangle \geq \|Fx - Fy\|_2^2$$

if and only if

$$\|Nx - Ny\|_2^2 \leq \|x - y\|_2^2.$$

∎

Corollary 4.2 *If, for each* $m = 1, ..., M$*, the operator* $F_m : \mathbb{R}^J \to \mathbb{R}^J$ *is fne, and* $\alpha_m > 0$*, with* $\sum_{m=1}^M \alpha_m = 1$*, then* $F = \sum_{m=1}^M \alpha_m F_m$ *is fne.*

Corollary 4.3 *The operator* T *defined in Equation (4.20) is firmly non-expansive.*

4.6 Convergence of the SIMOP Algorithm

Let the operator T be given by Equation (4.20). From Equation (4.18), we see that

$$\nabla F(x) = x - Tx. \tag{4.23}$$

It follows that any fixed point of the operator T is a global minimizer of $F(x)$.

Lemma 4.1 *If* $F(z) = 0$ *and* $Tx^* = x^*$*, then* $F(x^*) = 0$*, and* $x^* \in C$*.*

Proof: From $Tx^* = x^*$ it follows that $\nabla F(x^*) = 0$, so that x^* is a global minimizer of $F(x)$. But the minimum of $F(x)$ is zero, since $F(z) = 0$. Therefore, $F(x^*) = 0$, which tells us that $x^* \in C$. ∎

Note that, if C is empty, it is possible for $Tx^* = x^*$ without x^* being in C; consider two parallel lines, with x^* half way between the two lines.

The iterative step of the *simultaneous orthogonal projection* (SIMOP) algorithm is $x^{n+1} = Tx^n$, where T is as defined in Equation (4.20). The main convergence theorem for the SIMOP algorithm is the following.

Theorem 4.2 *Whenever* $F(x)$ *has minimizers, the sequence* $\{x^n\}$ *generated by* $x^{n+1} = Tx^n$ *converges to a fixed point* x^* *of* T*, and* x^* *minimizes* $F(x)$*. If* C *is not empty, then* x^* *is in* C*.*

Proof: Let $Tz = z$. We have

$$\|x^{n+1} - x^n\|_2^2 = \|z - x^n - z + x^{n+1}\|_2^2$$

$$= \|z - x^n\|_2^2 + \|z - x^{n+1}\|_2^2 - 2\langle z - x^n, z - x^{n+1}\rangle.$$

Since

$$\langle z - x^n, z - x^{n+1}\rangle = \langle Tz - Tx^n, z - x^n\rangle,$$

and T is fne, it follows that

$$\langle z - x^n, z - x^{n+1}\rangle \geq \|Tz - Tx^n\|_2^2 = \|z - x^{n+1}\|_2^2.$$

Therefore

$$\|x^{n+1} - x^n\|_2^2 \leq \|z - x^n\|_2^2 + \|z - x^{n+1}\|_2^2 - 2\|z - x^{n+1}\|_2^2,$$

or

$$\|z - x^n\|_2^2 - \|z - x^{n+1}\|_2^2 \geq \|x^{n+1} - x^n\|_2^2.$$

We conclude from this that the sequence $\{\|z - x^n\|\}$ is decreasing, the sequence $\{\|x^{n+1} - x^n\|_2^2\}$ converges to zero, the sequence $\{x^n\}$ is bounded, a subsequence converges to some x^*, and $Tx^* = x^*$. Replacing the generic z with x^*, we find that the sequence $\{\|x^* - x^n\|\}$ is decreasing. But a subsequence converges to zero, so the sequence $\{x^n\}$ converges to x^*, a global minimizer of $F(x)$. If C is not empty, then $x^* \in C$. ∎

Recall that Landweber's algorithm converges to the solution closest to x^0. For the SIMOP algorithm we cannot make the analogous assertion; we cannot claim that x^* is $P_C x^0$. See [47] and the discussion of the HLWB algorithm for further details.

4.7 Block-Iterative Projection Methods

Including both the simultaneous and successive orthogonal projection algorithms are the *block-iterative* projection (BIP) methods [59]. Each step of the BIP algorithm uses a convex combination of some of the operators P_{C_i}. Each of these convex combinations is a fne operator, as we have seen.

As we discussed previously with regard to the BILW algorithm, we decompose the set $\{i = 1, ..., I\}$ into M blocks; that is, into M (not necessarily disjoint) subsets B_m, $m = 1, ..., M$. Denote by I_m the cardinality of B_m. For each m, and each $i \in B_m$, let $\alpha_i^m > 0$, with $\sum_{i \in B_m} \alpha_i^m = 1$. Then let T_m be the operator defined by

$$T_m := \sum_{i \in B_m} \alpha_i^m P_{C_i}. \tag{4.24}$$

The operator T_m is fne.

The iterative step of the BIP algorithm is

$$x^{n+1} = T_m x^n, \tag{4.25}$$

where, for each $n = 0, 1, ...$, we set $m = m(n) = n (\mathrm{mod}\, M) + 1$. The main theorem concerning the BIP is the following.

Theorem 4.3 *The sequence $\{x^n\}$ defined by Equation (4.25) converges to a member of C whenever C is non-empty.*

Proof: We have

$$\|x^{n+1} - x^n\|_2^2 = \|z - x^n - z + x^{n+1}\|_2^2$$

$$= \|z - x^n\|_2^2 + \|z - x^{n+1}\|_2^2 - 2\langle z - x^{n+1}, z - x^n\rangle.$$

From

$$\langle z - x^{n+1}, z - x^n\rangle = \langle T_m z - T_m x^n, z - x^n\rangle$$

and the fact that T_m is fne, we have

$$\langle z - x^{n+1}, z - x^n\rangle \geq \|T_m z - T_m x^n\|_2^2 = \|z - x^{n+1}\|_2^2.$$

Therefore,

$$\|z - x^n\|_2^2 - \|z - x^{n+1}\|_2^2 \geq \|x^{n+1} - x^n\|_2^2.$$

For $k = 0, 1, ...,$ we have

$$\|z - x^{kM}\|_2^2 - \|z - x^{(k+1)M}\|_2^2 \geq \sum_{m=1}^{M} \|x^{kM+m} - x^{kM+(m-1)}\|_2^2.$$

From this inequality we know that the sequence $\{\|z - x^{kM}\|_2^2\}$ is decreasing and non-negative, the sequences $\{\|x^{kM+m} - x^{kM+(m-1)}\|_2^2\}$ converge to zero, the sequence $\{x^{kM}\}$ is bounded, and a subsequence converges to some $x^{*,0}$. With $x^{*,m} = T_m x^{*,m-1}$, we find that $x^{*,m} = x^{*,m-1}$ for all m. The vector $x^* := x^{*,0}$ is then in C. Replacing the generic z with x^*, we have that the sequence $\{x^n\}$ converges to x^*. ∎

A somewhat more general version of the BIP is given in [59].

Chapter 5

SUMMA

5.1 Sequential Unconstrained Optimization

Sequential unconstrained optimization algorithms can be used to minimize a function $f : \mathbb{R}^J \to (-\infty, \infty]$ over a (not necessarily proper) subset C of $\mathbb{R}^J$ [81]. At the nth step of a *sequential unconstrained minimization* method we obtain x^n by minimizing the function

$$G_n(x) = f(x) + g_n(x), \tag{5.1}$$

where the auxiliary function $g_n(x)$ is appropriately chosen. If C is a proper subset of $\mathbb{R}^J$ we may force $g_n(x) = +\infty$ for x not in C, as in the barrier-function methods; then each x^n will lie in C. The objective is then to select the $g_n(x)$ so that the sequence $\{x^n\}$ converges to a solution of the problem, or failing that, at least to have the sequence $\{f(x^n)\}$ converging to the infimum of $f(x)$ over x in C.

5.2 SUMMA

In [43] we presented a particular class of sequential unconstrained minimization methods called SUMMA. As we showed in that paper, this class is broad enough to contain barrier-function methods, proximal minimization methods, and the simultaneous multiplicative algebraic reconstruction technique (SMART). By reformulating the problem, the penalty-function methods can also be shown to be members of the SUMMA class. Any alternating minimization (AM) problem with the five-point property [69] can be reformulated as a SUMMA problem; therefore the *expectation maximization maximum likelihood* (EMML) algorithm for Poisson data, which is such an AM algorithm, must also be a SUMMA algorithm.

For a method to be in the SUMMA class we require that $x^n \in C$ for each n and that each auxiliary function $g_n(x)$ satisfy the inequality

$$0 \leq g_n(x) \leq G_{n-1}(x) - G_{n-1}(x^{n-1}), \tag{5.1}$$

for all x. Note that it follows that $g_n(x^{n-1}) = 0$, for all n.

We assume, throughout this chapter, that the inequality in (5.1) holds for each n. We also assume that $\inf_{x \in C} f(x) = b > -\infty$. The next two results are taken from [43].

Proposition 5.1 *The sequence $\{f(x^n)\}$ is non-increasing and the sequence $\{g_n(x^n)\}$ converges to zero.*

Proof: We have

$$f(x^{n+1}) + g_{n+1}(x^{n+1}) = G_{n+1}(x^{n+1}) \leq G_{n+1}(x^n) = f(x^n). \tag{5.2}$$

∎

Theorem 5.1 *The sequence $\{f(x^n)\}$ converges to b.*

Proof: Suppose that there is $\delta > 0$ such that $f(x^n) \geq b + 2\delta$, for all n. Then there is $z \in C$ such that $f(x^n) \geq f(z) + \delta$, for all n. From the inequality in (5.1) we have

$$g_n(z) - g_{n+1}(z) \geq f(x^n) + g_n(x^n) - f(z) \geq f(x^n) - f(z) \geq \delta, \tag{5.3}$$

for all n. But this cannot happen; the successive differences of a non-increasing sequence of non-negative terms must converge to zero. ∎

5.3 Using SUMMA

Sequential unconstrained minimization algorithms are most commonly used to enforce constraints on the vector variable x. Sometimes, though, they can be used to simplify calculations and to enable us to obtain each x^n in closed form.

For example, we can formulate the Landweber iteration as a SUMMA method and use this approach to prove convergence. For each n let

$$G_n(x) = \frac{1}{2}\|Ax - b\|_2^2 + \frac{1}{2\gamma}\|x - x^{n-1}\|_2^2 - \frac{1}{2}\|Ax - Ax^{n-1}\|_2^2. \tag{5.4}$$

If we require that $0 < \gamma < 1/\rho(A^T A)$, then

$$g_n(x) = \frac{1}{2\gamma}\|x - x^{n-1}\|_2^2 - \frac{1}{2}\|Ax - Ax^{n-1}\|_2^2 \tag{5.5}$$

is non-negative. We also have

$$G_n(x) - G_n(x^n) = \frac{1}{2\gamma}\|x - x^n\|_2^2 \geq g_{n+1}(x), \tag{5.6}$$

so the iteration falls into the SUMMA class. We shall use this idea several times throughout this book.

Let $A^T Az = A^T b$. Then z minimizes

$$f(x) := \frac{1}{2}\|Ax - b\|_2^2.$$

Then we have

$$G_n(z) - G_n(x^n) = f(z) + g_n(z) - f(x^n) - g_n(x^n)$$

$$\leq f(z) + G_{n-1}(z) - G_{n-1}(x^{n-1}) - f(x^n) - g_n(x^n),$$

so that

$$\Big(G_{n-1}(z) - G_{n-1}(x^{n-1})\Big) - \Big(G_n(z) - G_n(x^n)\Big) \geq f(x^n) - f(z) + g_n(x^n) \geq 0.$$

Therefore, the sequence $\{G_n(z) - G_n(x^n)\}$ is decreasing and non-negative, so that the sequences $\{f(x^n) - f(z)\}$ and $\{g_n(x^n)\}$ converge to zero. Since

$$G_n(z) - G_n(x^n) = \frac{1}{2\gamma}\|z - x^n\|_2^2,$$

it follows that the sequence $\{x^n\}$ is bounded, a subsequence converges to some x^*, and $f(z) = f(x^*)$. Replacing z with x^*, we conclude that the sequence $\{x^n\}$ converges to x^*.

Chapter 6

Function Minimization

6.1 Gradient Descent Iteration

As we have seen, the Landweber algorithm minimizes the function $f(x) = \|Ax - b\|^2$, and the SIMOP algorithm minimizes the function $f(x) = F(x)$ defined by Equation (4.1). The iterative step for both algorithms has the gradient descent form

$$x^{n+1} = x^n - \gamma \nabla f(x^n). \tag{6.1}$$

In this chapter we consider the general problem of minimizing a differentiable convex function $f(x)$ using the iteration in Equation (6.1).

6.2 A Convergence Theorem

If $f : \mathbb{R}^J \to \mathbb{R}$ is convex and differentiable, and the gradient operator $Tx = \nabla f(x)$ is L-Lipschitz continuous, that is,

$$\|\nabla f(x) - \nabla f(y)\|_2 \leq L\|x - y\|_2,$$

then the gradient of the function $g(x) = \frac{1}{L} f(x)$ is non-expansive. According to Theorem 16.1, the operator $Fx = \nabla g(x)$ is firmly non-expansive. Therefore, the operator $A := I - \gamma F$ is averaged, for $0 < \gamma < 2/L$, and the iterative sequence $\{A^n x^0\}$ converges to a fixed point of A, and therefore to a global minimizer of $f(x)$, whenever such minimizers exist. In this chapter we prove a slightly weaker form of this convergence theorem, without using the non-trivial Theorem 16.1.

Theorem 6.1 *Let ∇f be L-Lipschitz continuous. The sequence $\{A^n x^0\}$ converges to a fixed point of A, and therefore to a global minimizer of $f(x)$, whenever such minimizers exist, provided that $0 < \gamma < 1/L$.*

Proof: Denote by $D_f(x,z)$ the Bregman distance associated with the function $f(x)$ and given by

$$D_f(x,z) := f(x) - f(z) - \langle \nabla f(z), x-z\rangle. \tag{6.2}$$

For each n we find x^n by minimizing the function

$$G_n(x) = f(x) + \frac{1}{2\gamma}\|x - x^{n-1}\|_2^2 - D_f(x, x^{n-1}). \tag{6.3}$$

It follows then that

$$0 = \nabla f(x^n) + \frac{1}{\gamma}(x^n - x^{n-1}) - \nabla f(x^n) + \gamma \nabla f(x^{n-1})$$

so that

$$x^n = x^{n-1} - \gamma \nabla f(x^{n-1}).$$

The function $G_n(x)$ can be written as

$$G_n(x) = f(x) + g_n(x),$$

for

$$g_n(x) = \frac{1}{2\gamma}\|x - x^{n-1}\|_2^2 - D_f(x, x^{n-1}).$$

From

$$G_n(x) - G_n(x^n) = \frac{1}{2\gamma}\|x - x^n\|_2^2 + f(x) - f(x^n) - \langle \nabla f(x^n), x - x^n\rangle,$$

it follows that

$$G_n(x) - G_n(x^n) \geq \frac{1}{2\gamma}\|x - x^n\|_2^2 \geq g_{n+1}(x).$$

The restriction $0 < \gamma < 1/L$ implies that $g_n(x) \geq 0$. Therefore, this iterative method falls into the SUMMA class.

Let z be a global minimizer of $f(x)$. From the proof of convergence of SUMMA iterations, we may conclude that

$$\Big(G_{n-1}(z) - G_{n-1}(x^{n-1})\Big) - \Big(G_n(z) - G_n(x^n)\Big) \geq f(x^n) - f(z) + g_n(x^n) \geq 0.$$

Therefore, the sequence $\{G_n(z) - G_n(x^n)\}$ is decreasing, and the sequence $\{f(x^n\}$ converges to $f(z)$. From

$$G_n(z) - G_n(x^n) \geq \frac{1}{2\gamma}\|z - x^n\|_2^2,$$

it follows that the sequence $\{x^n\}$ is bounded, that a subsequence converges to some x^*, and that $f(x^*) = f(z)$. Replacing the generic z with x^*, we find that the sequence $\{x^n\}$ converges to x^*. ∎

Chapter 7

Forward-Backward Splitting

7.1 The Forward-Backward Splitting Algorithm

In this chapter we present the *forward-backward splitting* (FBS) algorithm and prove convergence. The FBS algorithm is quite general and contains, as particular cases, every one of the simultaneous iterative algorithms discussed so far in this book.

Let $f : \mathbb{R}^J \rightarrow \mathbb{R}$ be convex. For each $z \in \mathbb{R}^J$ the function

$$m_f(z) = \min_x \{f(x) + \frac{1}{2}\|x - z\|_2^2\}$$

is minimized by $x = \mathrm{prox}_f(z)$. Moreau's proximity operator prox_f extends the notion of orthogonal projection onto a closed convex set [108, 109, 110]. We have $x = \mathrm{prox}_f(z)$ if and only if $z - x \in \partial f(x)$, where the set $\partial f(x)$ is the sub-differential of f at x, given by

$$\partial f(x) = \{u | \langle u, y - x\rangle \leq f(y) - f(x), \text{for all}\, y\}.$$

Proximity operators are also firmly non-expansive [66]; indeed, the proximity operator prox_f is the resolvent of the maximal monotone operator $B(x) = \partial f(x)$ and all such resolvent operators are firmly non-expansive [15].

Our objective here is to provide an elementary proof of convergence for the *forward-backward splitting* (FBS) algorithm; a detailed discussion of this algorithm and its history is given by Combettes and Wajs in [66]. Convergence of the version of the FBS algorithm given here involves only the fact the proximity operators are fne. A slightly more general version can be obtained using averaged operators, as we shall see later.

Theorem 7.1 *Let $f : \mathbb{R}^J \to \mathbb{R}$ be convex, with $f = f_1 + f_2$, both convex, f_2 differentiable, and ∇f_2 L-Lipschitz. For $0 < \gamma < \frac{1}{L}$, let*

$$x^n = \text{prox}_{\gamma f_1}\Big(x^{n-1} - \gamma \nabla f_2(x^{n-1})\Big). \tag{7.1}$$

The sequence $\{x^n\}$ converges to a minimizer of the function $f(x)$, whenever such minimizers exist.

Any fixed point of the iteration minimizes the function $f(x)$. Because proximity operators are firmly non-expansive, and therefore averaged, it is a consequence of the Krasnoselskii-Mann Theorem 8.1 [98, 105] for averaged operators that convergence holds for $0 < \gamma < \frac{2}{L}$. The proof given here employs sequential unconstrained minimization and avoids using the non-trivial results that, because the operator $\frac{1}{L}\nabla f_2$ is non-expansive, it is firmly non-expansive, and that the product of averaged operators is averaged.

7.2 Convergence of the FBS algorithm

For each $k = 1, 2, ...$ let

$$G_n(x) = f(x) + \frac{1}{2\gamma}\|x - x^{n-1}\|_2^2 - D_{f_2}(x, x^{n-1}), \tag{7.2}$$

where

$$D_{f_2}(x, x^{n-1}) = f_2(x) - f_2(x^{n-1}) - \langle \nabla f_2(x^{n-1}), x - x^{n-1}\rangle. \tag{7.3}$$

Since $f_2(x)$ is convex, $D_{f_2}(x, y) \geq 0$ for all x and y and is the Bregman distance formed from the function f_2 [12].

Lemma 7.1 *The x^n that minimizes $G_n(x)$ over x is given by Equation (7.1).*

Proof: Since x^n minimizes $G_n(x)$ we know that

$$0 \in \nabla f_2(x^n) + \frac{1}{\gamma}(x^n - x^{n-1}) - \nabla f_2(x^n) + \nabla f_2(x^{n-1}) + \partial f_1(x^n).$$

Therefore,

$$\Big(x^{n-1} - \gamma \nabla f_2(x^{n-1})\Big) - x^n \in \partial \gamma f_1(x^n).$$

Consequently,

$$x^n = \text{prox}_{\gamma f_1}(x^{n-1} - \gamma \nabla f_2(x^{n-1})).$$

∎

The auxiliary function

$$g_n(x) = \frac{1}{2\gamma}\|x - x^{n-1}\|_2^2 - D_{f_2}(x, x^{n-1}) \tag{7.4}$$

can be rewritten as

$$g_n(x) = D_h(x, x^{n-1}), \tag{7.5}$$

where

$$h(x) = \frac{1}{2\gamma}\|x\|_2^2 - f_2(x). \tag{7.6}$$

Therefore, $g_n(x) \geq 0$ whenever $h(x)$ is a convex function.

We know that $h(x)$ is convex if and only if

$$\langle \nabla h(x) - \nabla h(y), x - y \rangle \geq 0, \tag{7.7}$$

for all x and y. This is equivalent to

$$\frac{1}{\gamma}\|x - y\|_2^2 - \langle \nabla f_2(x) - \nabla f_2(y), x - y \rangle \geq 0. \tag{7.8}$$

Since ∇f_2 is L-Lipschitz, the inequality (7.8) holds whenever $0 < \gamma < \frac{1}{L}$.

A relatively simple calculation shows that

$$G_n(x) - G_n(x^n) = \frac{1}{2\gamma}\|x - x^n\|_2^2 +$$

$$\Big(f_1(x) - f_1(x^n) - \langle (x^{n-1} - \gamma\nabla f_2(x^{n-1})) - x^n, x - x^n \rangle\Big). \tag{7.9}$$

Since

$$(x^{n-1} - \gamma\nabla f_2(x^{n-1})) - x^n \in \partial\gamma f_1(x^n),$$

it follows that

$$\Big(f_1(x) - f_1(x^n) - \langle (x^{n-1} - \gamma\nabla f_2(x^{n-1})) - x^n, x - x^n \rangle\Big) \geq 0.$$

Therefore,

$$G_n(x) - G_n(x^n) \geq \frac{1}{2\gamma}\|x - x^n\|_2^2 \geq g_{n+1}(x). \tag{7.10}$$

Therefore, the inequality in (5.1) holds and the iteration fits into the SUMMA class.

Now let $\hat{x}$ minimize $f(x)$ over all x. Then

$$G_n(\hat{x}) - G_n(x^n) = f(\hat{x}) + g_n(\hat{x}) - f(x^n) - g_n(x^n)$$

$$\leq f(\hat{x}) + G_{k-1}(\hat{x}) - G_{k-1}(x^{n-1}) - f(x^n) - g_n(x^n),$$

so that

$$\Big(G_{k-1}(\hat{x}) - G_{k-1}(x^{n-1})\Big) - \Big(G_n(\hat{x}) - G_n(x^n)\Big) \geq f(x^n) - f(\hat{x}) + g_n(x^n) \geq 0.$$

Therefore, the sequence $\{G_n(\hat{x}) - G_n(x^n)\}$ is decreasing and the sequences $\{g_n(x^n)\}$ and $\{f(x^n) - f(\hat{x})\}$ converge to zero.

From

$$G_n(\hat{x}) - G_n(x^n) \geq \frac{1}{2\gamma}\|\hat{x} - x^n\|_2^2,$$

it follows that the sequence $\{x^n\}$ is bounded and that a subsequence converges to some $x^* \in C$ with $f(x^*) = f(\hat{x})$.

Replacing the generic $\hat{x}$ with x^*, we find that $\{G_n(x^*) - G_n(x^n)\}$ is decreasing. By Equation (7.9), it therefore converges to the limit

$$\frac{1}{2\gamma}\|x^* - x^*\|_2^2 + \frac{1}{\gamma}\langle(\text{prox}_{\gamma f_1} - I)(x^* - \gamma\nabla f(x^*)), x^* - \text{prox}_{\gamma f_1}(x^* - \gamma\nabla f(x^*))\rangle = 0.$$

From the inequality in (7.10), we conclude that the sequence $\{\|x^* - x^n\|_2^2\}$ converges to zero, and so $\{x^n\}$ converges to x^*. This completes the proof of the theorem.

7.3 Some Particular Cases of the FBS

As we shall show, the FBS algorithm is quite general and includes, as particular cases, all the simultaneous metods discussed previously in this book.

7.3.1 Projected Gradient Descent

Let C be a non-empty, closed convex subset of $\mathbb{R}^J$ and $f_1(x) = \iota_C(x)$, the function that is $+\infty$ for x not in C and zero for x in C. Then $\iota_C(x)$ is convex, but not differentiable. We have $\text{prox}_{\gamma f_1} = P_C$, the orthogonal projection onto C. The iteration in Equation (7.1) becomes

$$x^n = P_C\Big(x^{n-1} - \gamma\nabla f_2(x^{n-1})\Big). \tag{7.11}$$

This is the iterative step of the *projected gradient descent* (PGD) algorithm. If $C = \mathbb{R}^J$, then the PGD algorithm becomes the *gradient descent* (GD) algorithm. The sequence $\{x^n\}$ converges to a minimizer of f_2 over $x \in C$, whenever such minimizers exist.

7.3.2 The CQ Algorithm

Let A be a real I by J matrix, $C \subseteq \mathbb{R}^J$, and $Q \subseteq \mathbb{R}^I$, both closed convex sets. The *split feasibility problem* (SFP) is to find x in C such that Ax is in Q. The function

$$f_2(x) = \frac{1}{2}\|P_Q Ax - Ax\|_2^2 \tag{7.12}$$

is convex, differentiable and ∇f_2 is L-Lipschitz for $L = \rho(A^T A)$, the spectral radius of $A^T A$. The gradient of f_2 is

$$\nabla f_2(x) = A^T(I - P_Q)Ax. \tag{7.13}$$

We want to minimize the function $f_2(x)$ over x in C, or, equivalently, to minimize the function $f(x) = \iota_C(x) + f_2(x)$. The projected gradient descent algorithm now has the iterative step

$$x^n = P_C\Big(x^{n-1} - \gamma A^T(I - P_Q)Ax^{n-1}\Big); \tag{7.14}$$

this iterative method was called the CQ-algorithm in [38, 39]. The sequence $\{x^n\}$ converges to a solution whenever f_2 has a minimum on the set C.

7.3.3 The Projected Landweber Algorithm

The problem is to minimize the function

$$f_2(x) = \frac{1}{2}\|Ax - b\|_2^2,$$

over $x \in C$. This is a special case of the SFP and we can use the CQ-algorithm, with $Q = \{b\}$. The resulting iteration is the *projected Landweber* (PLW) algorithm; when $C = \mathbb{R}^J$ it becomes the Landweber algorithm.

7.3.4 The SIMOP Algorithm

The function

$$f(x) = \frac{1}{2I}\sum_{i=1}^{I} \|x - P_{C_i}x\|_2^2$$

is convex, differentiable, and its gradient operator is

$$\nabla f(x) = x - \frac{1}{I}\sum_{i=1}^{I} P_{C_i}x.$$

The GD iteration in this case, with $\gamma = 1$, is

$$x^{n+1} = \frac{1}{I}\sum_{i=1}^{I} P_{C_i}x^n,$$

which is the iterative step for the SIMOP algorithm.

Chapter 8

Fixed-Point Methods

8.1 Fixed-Points in Iteration

As we have seen, a number of problems can be solved by formulating the solution to the problem as a fixed-point of a continuous operator T and considering the sequence $\{x^n\}$ generated by the iterative algorithm $x^{n+1} = Tx^n$. If the sequence $\{x^n\}$ converges, then the limit point x^* satisfies $Tx^* = x^*$; that is, x^* is a fixed-point of T. We are concerned, therefore, with conditions on the operator T that guarantee the convergence of the sequence $\{x^n\}$ whenever T has fixed points. It is not enough that T be non-expansive in some norm, as the operator $T = -I$ shows. It is often sufficient that T be firmly non-expansive, but this is overly restrictive in some cases, since the product of fne operators need not be fne. The class of *averaged* operators is a useful class to consider.

8.2 Averaged Operators

The class of *averaged* (av) operators contains the class of fne operators and is contained in the class of ne operators. Averaged operators are closed to finite products. According to the Krasnosel'skii-Mann Theorem 8.1, if $A : \mathbb{R}^J \to \mathbb{R}^J$ is av, then the iterative sequence $\{x^n = A^n x^0\}$ converges to a fixed point of A, whenever fixed points exist.

In a previous chapter, we were able to prove convergence of the quite general FBS algorithm using the fact that prox operators are fne. There we had to assume that the parameter γ satisfies the restrictions $0 < \gamma < 1/L$. Using the full power of averaged operators, we can relax the restrictions to $0 < \gamma < 2/L$.

8.3 Two Useful Identities

The identities in the next two lemmas relate an arbitrary operator T to its complement, $G = I - T$, where I denotes the identity operator. These identities will allow us to transform properties of T into properties of G that may be easier to work with. A simple calculation is all that is needed to establish the following lemma.

Lemma 8.1 *Let T be an arbitrary operator T on $\mathbb{R}^J$ and $G = I - T$. Then*

$$||x-y||_2^2 - ||Tx-Ty||_2^2 = 2(\langle Gx-Gy, x-y\rangle) - ||Gx-Gy||_2^2. \tag{8.1}$$

Lemma 8.2 *Let T be an arbitrary operator T on $\mathbb{R}^J$ and $G = I - T$. Then*

$$\langle Tx-Ty, x-y\rangle - ||Tx-Ty||_2^2 =$$

$$\langle Gx-Gy, x-y\rangle - ||Gx-Gy||_2^2. \tag{8.2}$$

Proof: Use the previous lemma. ∎

The term 'averaged operator' appears in the work of Baillon, Bruck and Reich [15, 4]. There are several ways to define averaged operators. One way is in terms of the complement operator.

Definition 8.1 *An operator G on $\mathbb{R}^J$ is called ν-*inverse strongly monotone *(ν-ism)[86] (also called co-coercive in [65]) if there is $\nu > 0$ such that*

$$\langle Gx-Gy, x-y\rangle \geq \nu||Gx-Gy||_2^2. \tag{8.3}$$

The proof of the following lemma is left to the reader.

Lemma 8.3 *An operator T is ne, with respect to the two-norm, if and only if its complement $G = I - T$ is $\frac{1}{2}$-ism, and T is fne if and only if G is 1-ism, and if and only if G is fne. Also, T is ne if and only if $F = (I+T)/2$ is fne. If G is ν-ism and $\gamma > 0$ then the operator γG is $\frac{\nu}{\gamma}$-ism.*

Definition 8.2 *An operator T is called* averaged *(av) if $G = I - T$ is ν-ism for some $\nu > \frac{1}{2}$. If G is $\frac{1}{2\alpha}$-ism, for some $\alpha \in (0,1)$, then we say that T is α-av.*

It follows that every av operator is ne, with respect to the Euclidean norm, and every fne operator is av.

8.4 Properties of Averaged Operators

The averaged operators are sometimes defined in a different, but equivalent, way, using the following characterization of av operators.

Lemma 8.4 *An operator T is av if and only if, for some operator N that is non-expansive in the two-norm, and $\alpha \in (0,1)$, we have*

$$T = (1-\alpha)I + \alpha N.$$

Consequently, the operator T is av if and only if, for some α in $(0,1)$, the operator

$$N = \frac{1}{\alpha}T - \frac{1-\alpha}{\alpha}I = I - \frac{1}{\alpha}(I-T) = I - \frac{1}{\alpha}G$$

is non-expansive.

Proof: We assume first that there is $\alpha \in (0,1)$ and ne operator N such that $T = (1-\alpha)I + \alpha N$, and so $G = I - T = \alpha(I-N)$. Since N is ne, $I - N$ is $\frac{1}{2}$-ism and $G = \alpha(I-N)$ is $\frac{1}{2\alpha}$-ism. Conversely, assume that G is ν-ism for some $\nu > \frac{1}{2}$. Let $\alpha = \frac{1}{2\nu}$ and write $T = (1-\alpha)I + \alpha N$ for $N = I - \frac{1}{\alpha}G$. Since $I - N = \frac{1}{\alpha}G$, $I - N$ is $\alpha\nu$-ism. Consequently $I - N$ is $\frac{1}{2}$-ism and N is ne. ∎

An averaged operator is easily constructed from a given operator N that is ne in the two-norm by taking a convex combination of N and the identity I. The beauty of the class of av operators is that it contains many operators, such as P_C, that are not originally defined in this way. As we shall see shortly, finite products of averaged operators are again averaged, so the product of finitely many orthogonal projections is av.

We present now the fundamental properties of averaged operators, in preparation for the proof that the class of averaged operators is closed to finite products.

Note that we can establish that a given operator is av by showing that there is an α in the interval $(0,1)$ such that the operator

$$\frac{1}{\alpha}(A - (1-\alpha)I) \tag{8.4}$$

is ne. Using this approach, we can easily show that if T is sc, then T is av.

Lemma 8.5 *Let $T = (1-\alpha)A + \alpha N$ for some $\alpha \in (0,1)$. If A is averaged and N is non-expansive then T is averaged.*

Proof: Let $A = (1-\beta)I + \beta M$ for some $\beta \in (0,1)$ and ne operator M. Let $1 - \gamma = (1-\alpha)(1-\beta)$. Then we have

$$T = (1-\gamma)I + \gamma[(1-\alpha)\beta\gamma^{-1}M + \alpha\gamma^{-1}N]. \tag{8.5}$$

Since the operator $K = (1-\alpha)\beta\gamma^{-1}M + \alpha\gamma^{-1}N$ is easily shown to be ne and the convex combination of two ne operators is again ne, T is averaged. ∎

Corollary 8.1 *If A and B are av and α is in the interval $[0,1]$, then the operator $T = (1-\alpha)A + \alpha B$ formed by taking the convex combination of A and B is av.*

Corollary 8.2 *Let $T = (1-\alpha)F + \alpha N$ for some $\alpha \in (0,1)$. If F is fne and N is ne then T is averaged.*

The orthogonal projection operators P_H onto hyperplanes $H = H(a,\gamma)$ are sometimes used with *relaxation*, which means that P_H is replaced by the operator

$$T = (1-\omega)I + \omega P_H, \tag{8.6}$$

for some ω in the interval $(0,2)$. Clearly, if ω is in the interval $(0,1)$, then T is av, by definition, since P_H is ne. We want to show that, even for ω in the interval $[1,2)$, T is av. To do this, we consider the operator $R_H = 2P_H - I$, which is reflection through H; that is,

$$P_H x = \frac{1}{2}(x + R_H x), \tag{8.7}$$

for each x.

Lemma 8.6 *The operator $R_H = 2P_H - I$ is an isometry; that is,*

$$||R_H x - R_H y||_2 = ||x - y||_2, \tag{8.8}$$

for all x and y, so that R_H is ne.

Lemma 8.7 *For $\omega = 1 + \gamma$ in the interval $[1,2)$, we have*

$$(1-\omega)I + \omega P_H = \alpha I + (1-\alpha)R_H, \tag{8.9}$$

for $\alpha = \frac{1-\gamma}{2}$; therefore, $T = (1-\omega)I + \omega P_H$ is av.

The product of finitely many ne operators is again ne, while the product of finitely many fne operators, even orthogonal projections, need not be fne. It is a helpful fact that the product of finitely many av operators is again av.

If $A = (1-\alpha)I + \alpha N$ is averaged and B is averaged then $T = AB$ has the form $T = (1-\alpha)B + \alpha NB$. Since B is av and NB is ne, it follows from Lemma 8.5 that T is averaged. Summarizing, we have

Proposition 8.1 *If A and B are averaged, then $T = AB$ is averaged.*

Proposition 8.2 *An operator F is firmly non-expansive if and only if $F = \frac{1}{2}(I+N)$, for some operator N that is non-expansive in the two-norm.*

Lemma 8.8 *An operator $F : \mathbb{R}^J \to \mathbb{R}^J$ is fne if and only if $F = \frac{1}{2}(I+N)$, for some operator N that is ne with respect to the two-norm.*

Proof: Suppose that $F = \frac{1}{2}(I + N)$. We show that F is fne if and only if N is ne in the two-norm. First, we have

$$\langle Fx - Fy, x - y\rangle = \frac{1}{2}\|x - y\|_2^2 + \frac{1}{2}\langle Nx - Ny, x - y\rangle.$$

Also,

$$\|\frac{1}{2}(I+N)x - \frac{1}{2}(I+N)y\|_2^2 = \frac{1}{4}\|x-y\|^2 + \frac{1}{4}\|Nx-Ny\|^2 + \frac{1}{2}\langle Nx-Ny, x-y\rangle.$$

Therefore,

$$\langle Fx - Fy, x - y\rangle \geq \|Fx - Fy\|_2^2$$

if and only if

$$\|Nx - Ny\|_2^2 \leq \|x - y\|_2^2.$$

■

8.5 Gradient Operators

Another type of operator that is averaged can be derived from gradient operators.

Definition 8.3 *An operator T on $\mathbb{R}^J$ is* monotone *if*

$$\langle Tx - Ty, x - y\rangle \geq 0, \tag{8.10}$$

for all x and y.

Firmly non-expansive operators on $\mathbb{R}^J$ are monotone operators. Let $g(x) : \mathbb{R}^J \to \mathbb{R}$ be a differentiable convex function and $f(x) = \nabla g(x)$ its gradient. The operator ∇g is also monotone. If ∇g is non-expansive, then ∇g is fne (see [47]) . If, for some $L > 0$, ∇g is L-Lipschitz, for the two-norm, that is,

$$||\nabla g(x) - \nabla g(y)||_2 \leq L||x - y||_2, \tag{8.11}$$

for all x and y, then $\frac{1}{L}\nabla g$ is ne, therefore fne, and the operator $T = I - \gamma\nabla g$ is av, for $0 < \gamma < \frac{2}{L}$. The operators P_C are actually gradient operators; $P_C z = \nabla g(z)$ for

$$g(z) = \frac{1}{2}(\|z\|_2^2 - \|z - P_C z\|_2^2).$$

Note that not all monotone operators are gradient operators.

8.6 The Krasnosel'skii-Mann Theorem

For any operator T that is averaged, convergence of the sequence $\{T^n x^0\}$ to a fixed point of T, whenever fixed points of T exist, is guaranteed by the Krasnosel'skii-Mann (KM) Theorem [98, 105]:

Theorem 8.1 *Let T be α-averaged, for some $\alpha \in (0,1)$. Then the sequence $\{T^n x^0\}$ converges to a fixed point of T, whenever Fix(T) is non-empty.*

Proof: Let z be a fixed point of T. The identity in Equation (8.1) is the key to proving Theorem 8.1.

Using $Tz = z$ and $(I-T)z = 0$ and setting $G = I - T$ we have

$$||z - x^n||_2^2 - ||Tz - x^{n+1}||_2^2 = 2\langle Gz - Gx^n, z - x^n\rangle - ||Gz - Gx^n||_2^2. \tag{8.12}$$

Since, by Lemma 8.4, G is $\frac{1}{2\alpha}$-ism, we have

$$||z - x^n||_2^2 - ||z - x^{n+1}||_2^2 \geq (\frac{1}{\alpha} - 1)||x^n - x^{n+1}||_2^2. \tag{8.13}$$

Consequently the sequence $\{x^n\}$ is bounded, the sequence $\{||z - x^n||_2\}$ is decreasing and the sequence $\{||x^n - x^{n+1}||_2\}$ converges to zero. Let x^* be a cluster point of $\{x^n\}$. Then we have $Tx^* = x^*$, so we may use x^* in place of the arbitrary fixed point z. It follows then that the sequence $\{||x^* - x^n||_2\}$ is decreasing; since a subsequence converges to zero, the entire sequence converges to zero. The proof is complete. ∎

A version of the KM Theorem 8.1, with variable coefficients, appears in Reich's paper [115].

8.7 Norms Derived from Operators

Because it can be difficult to determine the properties of a given operator T with respect to a given norm, we often construct a norm to be compatible with a given operator.

Definition 8.4 *An operator T on X is a* strict contraction *(sc), with respect to a vector norm $||\cdot||$, if there is $r \in (0,1)$ such that*

$$||Tx - Ty|| \leq r||x - y||, \tag{8.14}$$

for all vectors x and y.

Since $\rho(B) \leq ||B||$ for every norm on B induced by a vector norm, B being a strict contraction (sc) implies that $\rho(B) < 1$. When B is Hermitian, the matrix norm of B induced by the Euclidean vector norm is $||B||_2 = \rho(B)$, so if $\rho(B) < 1$, then B is sc with respect to the Euclidean norm.

When B is not Hermitian, it is not as easy to determine if the affine operator T is sc with respect to a given norm. Instead, we often tailor the norm to the operator T. Suppose that B is a diagonalizable matrix, that is, there is a basis for $\mathbb{R}^J$ consisting of eigenvectors of B. Let $\{u^1, ..., u^J\}$ be such a basis, and let $Bu^j = \lambda_j u^j$, for each $j = 1, ..., J$. For each x in $\mathbb{R}^J$, there are unique coefficients a_j so that

$$x = \sum_{j=1}^{J} a_j u^j. \tag{8.15}$$

Then let

$$||x|| = \sum_{j=1}^{J} |a_j|. \tag{8.16}$$

Lemma 8.9 *The expression* $||\cdot||$ *in Equation (8.16) defines a norm on* $\mathbb{R}^J$*. If* $\rho(B) < 1$*, then the affine operator* T *is sc, with respect to this norm.*

It is known that, for any square matrix B and any $\epsilon > 0$, there is a vector norm for which the induced matrix norm satisfies $||B|| \leq \rho(B) + \epsilon$. Therefore, if B is an arbitrary square matrix with $\rho(B) < 1$, there is a vector norm with respect to which B is sc.

8.8 Affine Linear Operators

It may not always be easy to decide if a given operator is averaged. The class of affine linear operators provides an interesting illustration of the problem.

The affine operator $Tx = Bx + d$ will be ne, sc, fne, or av precisely when the linear operator given by multiplication by the matrix B is the same.

8.9 The Hermitian Case

When B is Hermitian, we can determine if B belongs to these classes by examining its eigenvalues λ:

- B is non-expansive if and only if $-1 \leq \lambda \leq 1$, for all λ;
- B is averaged if and only if $-1 < \lambda \leq 1$, for all λ;

- B is a strict contraction if and only if $-1 < \lambda < 1$, for all λ;
- B is firmly non-expansive if and only if $0 \leq \lambda \leq 1$, for all λ.

An operator $T : \mathbb{R}^J \to \mathbb{R}^J$ is *affine* if $Tx = Bx + d$, where B is a linear operator, i.e., a matrix, and d is a fixed vector. Affine linear operators T that arise, for instance, in splitting methods for solving systems of linear equations, generally have non-Hermitian linear part B. Deciding if such operators are fne or av is more difficult. Instead, we can ask if the operator is *paracontractive*, with respect to some norm.

8.10 Paracontractive Operators

By examining the properties of the orthogonal projection operators P_C, we were led to the useful class of averaged operators. The orthogonal projections also belong to another useful class, the paracontractions.

Definition 8.5 *An operator T is called* paracontractive *(pc), with respect to a given norm, if, for every fixed point y of T, we have*

$$||Tx - y|| < ||x - y||, \tag{8.17}$$

unless $Tx = x$.

Paracontractive operators are studied by Censor and Reich in [57].

Proposition 8.3 *The operators $T = P_C$ are paracontractive, with respect to the two-norm.*

Proof: It follows from Cauchy's Inequality that

$$||P_C x - P_C y||_2 \leq ||x - y||_2,$$

with equality if and only if

$$P_C x - P_C y = \alpha(x - y),$$

for some scalar α with $|\alpha| = 1$. But, because

$$0 \leq \langle P_C x - P_C y, x - y\rangle = \alpha||x - y||_2^2,$$

it follows that $\alpha = 1$, and so

$$P_C x - x = P_C y - y.$$

∎

When we ask if a given operator T is pc, we must specify the norm. We often construct the norm specifically for the operator involved. To illustrate, we consider the case of affine operators.

8.10.1 Linear and Affine Paracontractions

Let the matrix B be diagonalizable and let the columns of V be an eigenvector basis. Then we have $V^{-1}BV = D$, where D is the diagonal matrix having the eigenvalues of B along its diagonal.

Lemma 8.10 *A square matrix B is diagonalizable if all its eigenvalues are distinct.*

Proof: Let B be J by J. Let λ_j be the eigenvalues of B, $Bx^j = \lambda_j x^j$, and $x^j \neq 0$, for $j = 1, ..., J$. Let x^m be the first eigenvector that is in the span of $\{x_j | j = 1, ..., m-1\}$. Then

$$x^m = a_1 x^1 + ...a_{m-1} x^{m-1}, \tag{8.18}$$

for some constants a_j that are not all zero. Multiply both sides by λ_m to get

$$\lambda_m x^m = a_1 \lambda_m x^1 + ...a_{m-1} \lambda_m x^{m-1}. \tag{8.19}$$

From

$$\lambda_m x^m = Ax^m = a_1 \lambda_1 x^1 + ...a_{m-1} \lambda_{m-1} x^{m-1}, \tag{8.20}$$

it follows that

$$a_1(\lambda_m - \lambda_1)x^1 + ... + a_{m-1}(\lambda_m - \lambda_{m-1})x^{m-1} = 0, \tag{8.21}$$

from which we can conclude that some x^n in $\{x^1, ..., x^{m-1}\}$ is in the span of the others. This is a contradiction. ∎

We see from this Lemma that almost all square matrices B are diagonalizable. Indeed, all Hermitian B are diagonalizable. If B has real entries, but is not symmetric, then the eigenvalues of B need not be real, and the eigenvectors of B can have non-real entries. Consequently, we must consider B as a linear operator on $\mathbb{C}^J$, if we are to talk about diagonalizability. For example, consider the real matrix

$$B = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \tag{8.22}$$

Its eigenvalues are $\lambda = i$ and $\lambda = -i$. The corresponding eigenvectors are $(1, i)^T$ and $(1, -i)^T$. The matrix B is then diagonalizable as an operator on $\mathbb{C}^2$, but not as an operator on $\mathbb{R}^2$.

Proposition 8.4 *Let T be an affine linear operator whose linear part B is diagonalizable, and $|\lambda| < 1$ for all eigenvalues λ of B that are not equal to one. Then the operator T is pc, with respect to the norm given by Equation (8.16).*

Proof: The proof is not difficult and we leave it to the reader. ∎

We see from Proposition 8.4 that, for the case of affine operators T whose linear part is not Hermitian, instead of asking if T is av, we can ask if T is pc; since B will almost certainly be diagonalizable, we can answer this question by examining the eigenvalues of B.

Unlike the class of averaged operators, the class of paracontractive operators is not necessarily closed to finite products, unless those factor operators have a common fixed point.

8.10.2 The Elsner-Koltracht-Neumann Theorem

Our interest in paracontractions is due to the Elsner-Koltracht-Neumann (EKN) Theorem [79]:

Theorem 8.2 *Let T be pc with respect to some vector norm. If T has fixed points, then the sequence $\{T^k x^0\}$ converges to a fixed point of T, for all starting vectors x^0.*

We follow the development in [79].

Theorem 8.3 *Suppose that there is a vector norm on $\mathbb{R}^J$, with respect to which each T_i is a pc operator, for $i = 1, ..., I$, and that $F = \cap_{i=1}^I \mathrm{Fix}(T_i)$ is not empty. For $k = 0, 1, ...,$ let $i(k) = k(\mathrm{mod}\, I) + 1$, and $x^{k+1} = T_{i(k)} x^k$. The sequence $\{x^k\}$ converges to a member of F, for every starting vector x^0.*

Proof: Let $y \in F$. Then, for $k = 0, 1, ...,$

$$||x^{k+1} - y|| = ||T_{i(k)} x^k - y|| \leq ||x^k - y||, \tag{8.23}$$

so that the sequence $\{||x^k - y||\}$ is decreasing; let $d \geq 0$ be its limit. Since the sequence $\{x^k\}$ is bounded, we select an arbitrary cluster point, x^*. Then $d = ||x^* - y||$, from which we can conclude that

$$||T_i x^* - y|| = ||x^* - y||, \tag{8.24}$$

and $T_i x^* = x^*$, for $i = 1, ..., I$; therefore, $x^* \in F$. Replacing y, an arbitrary member of F, with x^*, we have that $||x^k - x^*||$ is decreasing. But, a subsequence converges to zero, so the whole sequence must converge to zero. This completes the proof. ∎

Corollary 8.3 *If T is pc with respect to some vector norm, and T has fixed points, then the iterative sequence $\{T^k x^0\}$ converges to a fixed point of T, for every starting vector x^0.*

Corollary 8.4 *If $T = T_I T_{I-1} \cdots T_2 T_1$, and $F = \cap_{i=1}^I \mathrm{Fix}\,(T_i)$ is not empty, then $F = \mathrm{Fix}\,(T)$.*

Proof: The sequence $x^{k+1} = T_{i(k)}x^k$ converges to a member of Fix (T), for every x^0. Select x^0 in F. ∎

Corollary 8.5 *The product T of two or more pc operators T_i, $i = 1, ..., I$ is again a pc operator, if $F = \cap_{i=1}^{I} \text{Fix}\,(T_i)$ is not empty.*

Proof: Suppose that for $T = T_I T_{I-1} \cdots T_2 T_1$, and $y \in F = \text{Fix}\,(T)$, we have

$$||Tx - y|| = ||x - y||. \tag{8.25}$$

Then, since

$$||T_I(T_{I-1} \cdots T_1)x - y|| \leq ||T_{I-1} \cdots T_1 x - y|| \leq ...$$

$$\leq ||T_1 x - y|| \leq ||x - y||, \tag{8.26}$$

it follows that

$$||T_i x - y|| = ||x - y||, \tag{8.27}$$

and $T_i x = x$, for each i. Therefore, $Tx = x$. ∎

Chapter 9

Regularization

9.1 Sensitivity

There are times when we do not want exact solutions of $Ax = b$ or of $A^T Ax = A^T b$. In most remote sensing problems, such as medical imaging from scan data, the vector x represents a digitized version of the image to be reconstructed from data, the matrix A describes the (ideal) relationship between the pixels of the image and the measurements to be taken, and the vector b contains the actual measurement values. Of course, the digitizing amounts to an approximation of the true object of interest, the relationship described by A is never precisely true, and the entries of b are, at the very least, truncated, and usually noisy, versions of the desired measurements. Solving the systems of equations for exact solutions amounts to overlooking the effects of digitization and to placing excessive confidence in the correctness of the description in A and in the accuracy of the measurements in b. The exact solutions of $Ax = b$ or of $A^T Ax = A^T b$ can be quite sensitive to these perturbations, and the resulting reconstructed images can be useless. We need some way to prevent this sensitivity to errors in the model and the measured data; *regularization* is the term used to describe such efforts.

9.2 Where Does Sensitivity Come From?

Consider the least squares solution $x = (A^T A)^{-1} A^T b$. Because A has been normalized, the eigenvalues of $A^T A$, which are all positive here, add up to one. But it is often the case that the largest eigenvalue is relatively much larger than the smallest one; the matrix $A^T A$ is then said to be *ill-conditioned*. For the inverse $(A^T A)^{-1}$ the situation is reversed and the least squares solution can be dominated by the lowest eigenvalue of $A^T A$. This leads to increased sensitivity of x to errors in the model and noise in the data. Increased sensitivity can be detected by calculating the two-norm of the least squares solution; when A^TA is ill-conditioned, the two-norm of the least squares solution is often unreasonably large.

The least squares solution minimizes the function $\|Ax - b\|_2^2$. To avoid sensitivity to noise, we can minimize instead the function

$$f(x) = (1 - \alpha)\|Ax - b\|_2^2 + \alpha\|x\|_2^2, \tag{9.1}$$

where $0 < \alpha < 1$ and is typically near zero. The resulting *regularized* solution, which we shall denote by x_{RLS}, is called the *norm-constrained* least squares solution. It can be written in closed form as

$$x_{RLS} = (A^TA + \epsilon^2 I)^{-1}A^Tb, \tag{9.2}$$

where $\epsilon^2 = \frac{\alpha}{1-\alpha}$.

As we have noted, when A has thousands of rows and columns, calculating A^TA must be avoided. One of the pleasant aspects of both the Landweber algorithm and the ART is that we never need to calculate A^TA. It would appear, however, that we cannot avoid calculating $A^TA + \epsilon^2 I$ if we wish to find x_{RLS}. This is not the case, fortunately, as we shall see.

9.3 Regularizing Landweber's Algorithm

Our goal is to minimize the function in Equation (9.1). Notice that this is equivalent to minimizing the function

$$F(x) = ||Bx - c||_2^2, \tag{9.3}$$

for

$$B = \begin{bmatrix} A \\ \epsilon I \end{bmatrix}, \tag{9.4}$$

and

$$c = \begin{bmatrix} b \\ 0 \end{bmatrix}, \tag{9.5}$$

where 0 denotes the column vector with all entries equal to zero. The Landweber iteration for the problem $Bx = c$ is

$$x^{n+1} = x^n + \gamma B^T(c - Bx^n), \tag{9.6}$$

for $0 < \gamma < 2/\rho(B^TB)$, where $\rho(B^TB)$ is the spectral radius of B^TB. Equation (9.6) can be written as

$$x^{n+1} = (1 - \gamma\epsilon^2)x^n + \gamma A^T(b - Ax^n). \tag{9.7}$$

We turn now to the regularization of the ART.

9.4 Regularizing the ART

We would like to get the regularized solution x_{RLS} by taking advantage of the faster convergence of the ART. Fortunately, there are ways to find x_{RLS}, using only the matrix A and the ART algorithm. We discuss two methods for using ART to obtain regularized solutions of $Ax = b$. The first one is presented in [41], while the second one is due to Eggermont, Herman, and Lent [76].

In our first method we use ART to solve the system of equations given in matrix form by

$$[\,A^T \quad \epsilon I\,] \begin{bmatrix} u \\ v \end{bmatrix} = 0. \tag{9.8}$$

We begin with $u^0 = b$ and $v^0 = 0$. Then, the lower component of the limit vector is $v^\infty = -\epsilon x_{RLS}$, while the upper limit is $u^\infty = b - Ax_{RLS}$.

The method of Eggermont *et al.* is similar. They use ART to solve the system of equations given in matrix form by

$$[\,A \quad \epsilon I\,] \begin{bmatrix} x \\ v \end{bmatrix} = b. \tag{9.9}$$

They begin at $x^0 = 0$ and $v^0 = 0$. Then, the limit vector has for its upper component $x^\infty = x_{RLS}$, and $\epsilon v^\infty = b - Ax_{RLS}$.

Chapter 10

Eigenvalue Bounds

10.1 Introduction and Notation

We are concerned here with iterative methods for solving, at least approximately, the system of I linear equations in J unknowns symbolized by $Ax = b$. In the applications of interest to us, such as medical imaging, both I and J are quite large, making the use of iterative methods the only feasible approach. It is also typical of such applications that the matrix A is sparse, that is, has relatively few non-zero entries. Therefore, iterative methods that exploit this sparseness to accelerate convergence are of special interest to us.

Many of the algorithms considered so far involve a parameter γ that must satisfy certain bounds. In the Landweber algorithm, for example, we require that $0 < \gamma < 2/\rho(A^TA)$. Generally, the larger the γ the larger the increment from x^n to x^{n+1}. If γ is selected well below its upper bound, the algorithm may converge slowly. When A is a large matrix, calculating A^TA is out of the question, so we must estimate $\rho(A^TA)$. Since A is normalized, its eigenvalues sum to the trace of AA^T, which is I. But estimating $\rho(A^TA)$ by I is much too conservative most of the time. We need a better estimate of $\rho(A^TA)$. For notational convenience, we shall often refer to $\rho(A^TA)$ as L in this chapter.

Having a good upper bound for L is important. In the applications of interest, principally medical image processing, the matrix A is large; even calculating A^TA, not to mention computing eigenvalues, is prohibitively expensive. In addition, the matrix A is typically sparse; that is, most of the entries of A are zero. However, A^TA will not be sparse, generally, even when A is. In this section we present upper bounds for L that do not require the calculation of A^TA, and are particularly useful when A is sparse.

10.2 Earlier Work

Many of the concepts we study in computational linear algebra were added to the mathematical toolbox relatively recently, as this area blossomed with the growth of electronic computers. Based on my brief investigations into the history of matrix theory, I believe that the concept of a norm of a matrix was not widely used prior to about 1945. This was recently confirmed when I read the paper [88]; as pointed out there, the use of matrix norms became an important part of numerical linear algebra only after the publication of [131]. Prior to the late 1940's a number of papers were published that established upper bounds on $\rho(A)$ for a general square matrix A. As we now can see, several of these results are immediate consequences of the fact that $\rho(A) \leq \|A\|$, for any induced matrix norm. We give two examples.

For a given N by N matrix A, let

$$C_n = \sum_{m=1}^{N} |A_{mn}|,$$

$$R_m = \sum_{n=1}^{N} |A_{mn}|,$$

and C and R the maxima of C_n and R_m, respectively. We now know that $C = \|A\|_1$, and $R = \|A\|_\infty$, but the earlier authors did not.

In 1930 Browne [13] proved the following theorem.

Theorem 10.1 (Browne) *Let λ be any eigenvalue of A. Then*

$$|\lambda| \leq \frac{1}{2}(C + R).$$

In 1944 Farnell [80] published the following theorems.

Theorem 10.2 (Farnell I) *For any eigenvalue λ of A we have*

$$|\lambda| \leq \sqrt{CR}.$$

Theorem 10.3 (Farnell II) *Let*

$$r_m = \sum_{n=1}^{N} |A_{mn}|^2,$$

and

$$c_m = \sum_{n=1}^{N} |A_{nm}|^2.$$

Then, for any eigenvalue λ of A, we have

$$|\lambda| \leq \sqrt{\sum_{m=1}^{N} \sqrt{r_m c_m}}.$$

In 1946 Brauer [11] proved the following theorem.

Theorem 10.4 (Brauer) *For any eigenvalue λ of A, we have*

$$|\lambda| \leq \min\{C, R\}.$$

Ex. 10.1 *Prove Theorems 10.1, 10.2, and 10.4 using properties of matrix norms. Can you also prove Theorem 10.3 this way?*

Let A be an arbitrary real rectangular matrix. Since the largest singular value of A is the square root of the maximum eigenvalue of the square matrix $S = A^T A$, we could use the inequality

$$\rho(A^T A) = \|A^T A\|_2 \leq \|A^T A\|,$$

for any induced matrix norm, to establish an upper bound for the singular values of A. However, that bound would be in terms of the entries of $A^T A$, not of A itself. In what follows we obtain upper bounds on the singular values of A in terms of the entries of A itself.

Ex. 10.2 *Let A be an arbitrary rectangular matrix. Prove that no singular value of A exceeds $\sqrt{\|A\|_1 \|A\|_\infty}$.*

We see from this exercise that Farnell (I) does generalize to arbitrary rectangular matrices and singular values. Brauer's Theorem 10.4 may suggest that no singular value of a rectangular matrix A exceeds the minimum of $\|A\|_1$ and $\|A\|_\infty$, but this is not true. Consider the matrix A whose SVD is given by

$$A = \begin{bmatrix} 4 & 3 \\ 8 & 6 \\ 8 & 6 \end{bmatrix} = \begin{bmatrix} 1/3 & 2/3 & 2/3 \\ 2/3 & -2/3 & 1/3 \\ 2/3 & 1/3 & -2/3 \end{bmatrix} \begin{bmatrix} 15 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 4/5 & 3/5 \\ 3/5 & -4/5 \end{bmatrix}.$$

The largest singular value of A is 15, $\|A\|_1 = 20$, $\|A\|_\infty = 14$, and we do have

$$15 \leq \sqrt{(20)(14)},$$

but we do not have

$$15 \leq \min\{20, 14\} = 14.$$

10.3 Our Basic Eigenvalue Inequality

In [129] van der Sluis and van der Vorst show that certain rescaling of the matrix A results in none of the eigenvalues of A^TA exceeding one. A modification of their proof leads to upper bounds on the eigenvalues of the original A^TA [45]. For any a in the interval $[0, 2]$ let

$$c_{aj} = c_{aj}(A) = \sum_{i=1}^{I} |A_{ij}|^a,$$

$$r_{ai} = r_{ai}(A) = \sum_{j=1}^{J} |A_{ij}|^{2-a},$$

and c_a and r_a the maxima of the c_{aj} and r_{ai}, respectively. We prove the following theorem.

Theorem 10.5 *For any a in the interval $[0, 2]$, no eigenvalue of the matrix A^TA exceeds the maximum of*

$$\sum_{j=1}^{J} c_{aj}|A_{ij}|^{2-a},$$

over all i, nor the maximum of

$$\sum_{i=1}^{I} r_{ai}|A_{ij}|^a,$$

over all j. Therefore, no eigenvalue of A^TA exceeds c_ar_a.

Proof: Let $A^TAv = \lambda v$, and let $w = Av$. Then we have

$$\|A^Tw\|_2^2 = \lambda\|w\|_2^2.$$

Applying Cauchy's Inequality, we obtain

$$\Big|\sum_{i=1}^{I} \overline{A_{ij}} w_i\Big|^2 \leq \Big(\sum_{i=1}^{I} |A_{ij}|^{a/2}|A_{ij}|^{1-a/2}|w_i|\Big)^2$$

$$\leq \Big(\sum_{i=1}^{I} |A_{ij}|^a\Big)\Big(\sum_{i=1}^{I} |A_{ij}|^{2-a}|w_i|^2\Big).$$

Therefore,

$$\|A^Tw\|_2^2 \leq \sum_{j=1}^{J} \Big(c_{aj}(\sum_{i=1}^{I} |A_{ij}|^{2-a}|w_i|^2)\Big) = \sum_{i=1}^{I} \Big(\sum_{j=1}^{J} c_{aj}|A_{ij}|^{2-a}\Big)|w_i|^2$$

$$\leq \max_i \Big(\sum_{j=1}^J c_{aj} |A_{ij}|^{2-a} \Big) \|w\|^2.$$

The remaining two assertions follow in similar fashion. ∎

As a corollary, we obtain the following eigenvalue inequality.

Corollary 10.1 *For each $i = 1, 2, ..., I$, let*

$$p_i = \sum_{j=1}^J s_j |A_{ij}|^2,$$

and let p be the maximum of the p_i. Then $L \leq p$.

Proof: Take $a = 0$. Then, using the convention that $0^0 = 0$, we have $c_{0j} = s_j$. ∎

Corollary 10.2 *([38]; [128], Th. 4.2) If $\sum_{j=1}^J |A_{ij}|^2 \leq 1$ for each i, then $L \leq s$.*

Proof: For all i we have

$$p_i = \sum_{j=1}^J s_j |A_{ij}|^2 \leq s \sum_{j=1}^J |A_{ij}|^2 \leq s.$$

Therefore,

$$L \leq p \leq s.$$

∎

Corollary 10.3 *Selecting $a = 1$, we have*

$$L = \|A\|_2^2 \leq \|A\|_1 \|A\|_\infty = c_1 r_1.$$

Therefore, the largest singular value of A does not exceed $\sqrt{\|A\|_1 \|A\|_\infty}$.

Corollary 10.4 *Selecting $a = 2$, we have*

$$L = \|A\|_2^2 \leq \|A\|_F^2,$$

where $\|A\|_F$ denotes the Frobenius norm of A.

Corollary 10.5 *Let G be the matrix with entries*

$$G_{ij} = A_{ij} \sqrt{\alpha_i} \sqrt{\beta_j},$$

where

$$\alpha_i \leq \Big(\sum_{j=1}^J s_j \beta_j |A_{ij}|^2 \Big)^{-1},$$

for all i. Then $\rho(G^\dagger G) \leq 1$.

Proof: We have

$$\sum_{j=1}^{J} s_j|G_{ij}|^2 = \alpha_i \sum_{j=1}^{J} s_j\beta_j|A_{ij}|^2 \leq 1,$$

for all i. The result follows from Corollary 10.1. ∎

Corollary 10.6 *If $\sum_{j=1}^{J} s_j|A_{ij}|^2 \leq 1$ for all i, then $L \leq 1$.*

Corollary 10.7 *If $0 < \gamma_i \leq p_i^{-1}$ for all i, then the matrix B with entries $B_{ij} = \sqrt{\gamma_i}A_{ij}$ has $\rho(B^\dagger B) \leq 1$.*

Proof: We have

$$\sum_{j=1}^{J} s_j|B_{ij}|^2 = \gamma_i \sum_{j=1}^{J} s_j|A_{ij}|^2 = \gamma_i p_i \leq 1.$$

Therefore, $\rho(B^\dagger B) \leq 1$, according to the theorem. ∎

Corollary 10.8 *If, for some a in the interval $[0, 2]$, we have*

$$\alpha_i \leq r_{ai}^{-1}, \tag{10.1}$$

for each i, and

$$\beta_j \leq c_{aj}^{-1}, \tag{10.2}$$

for each j, then, for the matrix G with entries

$$G_{ij} = A_{ij}\sqrt{\alpha_i}\sqrt{\beta_j},$$

no eigenvalue of $G^\dagger G$ exceeds one.

Proof: We calculate $c_{aj}(G)$ and $r_{ai}(G)$ and find that

$$c_{aj}(G) \leq \Big(\max_i \alpha_i^{a/2}\Big)\beta_j^{a/2} \sum_{i=1}^{I} |A_{ij}|^a = \Big(\max_i \alpha_i^{a/2}\Big)\beta_j^{a/2} c_{aj}(A),$$

and

$$r_{ai}(G) \leq \Big(\max_j \beta_j^{1-a/2}\Big)\alpha_i^{1-a/2} r_{ai}(A).$$

Therefore, applying the inequalities (10.1) and (10.2), we have

$$c_{aj}(G)r_{ai}(G) \leq 1,$$

for all i and j. Consequently, $\rho(G^\dagger G) \leq 1$. ∎

10.3.1 Another Upper Bound for L

The next theorem ([38]) provides another upper bound for L that is useful when A is sparse. For this theorem we do not assume that A is normalized. As previously, for each i and j, we let $e_{ij} = 1$, if A_{ij} is not zero, and $e_{ij} = 0$, if $A_{ij} = 0$. Let $0 < \nu_i = \sqrt{\sum_{j=1}^J |A_{ij}|^2}$, $\sigma_j = \sum_{i=1}^I e_{ij}\nu_i^2$, and σ be the maximum of the σ_j.

Theorem 10.6 *([38]) No eigenvalue of $A^T A$ exceeds σ.*

Proof: Let $A^T Av = cv$, for some non-zero vector v and scalar c. With $w = Av$, we have

$$w^\dagger AA^T w = cw^\dagger w.$$

Then

$$\Big|\sum_{i=1}^I \overline{A_{ij}} w_i\Big|^2 = \Big|\sum_{i=1}^I \overline{A_{ij}} e_{ij}\nu_i \frac{w_i}{\nu_i}\Big|^2 \leq \Big(\sum_{i=1}^I |A_{ij}|^2 \frac{|w_i|^2}{\nu_i^2}\Big)\Big(\sum_{i=1}^I \nu_i^2 e_{ij}\Big)$$

$$= \Big(\sum_{i=1}^I |A_{ij}|^2 \frac{|w_i|^2}{\nu_i^2}\Big)\sigma_j \leq \sigma\Big(\sum_{i=1}^I |A_{ij}|^2 \frac{|w_i|^2}{\nu_i^2}\Big).$$

Therefore, we have

$$cw^\dagger w = w^\dagger AA^T w = \sum_{j=1}^J \Big|\sum_{i=1}^I \overline{A_{ij}} w_i\Big|^2$$

$$\leq \sigma \sum_{j=1}^J \Big(\sum_{i=1}^I |A_{ij}|^2 \frac{|w_i|^2}{\nu_i^2}\Big) = \sigma \sum_{i=1}^I |w_i|^2 = \sigma w^\dagger w.$$

We conclude that $c \leq \sigma$. ∎

Corollary 10.9 *Let the rows of A have Euclidean length one. Then no eigenvalue of $A^T A$ exceeds the maximum number of non-zero entries in any column of A.*

Proof: We have $\nu_i^2 = \sum_{j=1}^J |A_{ij}|^2 = 1$, for each i, so that $\sigma_j = s_j$ is the number of non-zero entries in the jth column of A, and $\sigma = s$ is the maximum of the σ_j. ∎

Corollary 10.10 *Let ν be the maximum Euclidean length of any row of A and s the maximum number of non-zero entries in any column of A. Then $L \leq \nu^2 s$.*

When the rows of A have length one, it is easy to see that $L \leq I$, so the choice of $\gamma = \frac{1}{I}$ in the Landweber algorithm, which gives Cimmino's algorithm [64], is acceptable, although perhaps much too small.

The proof of Theorem 10.6 is based on results presented by Arnold Lent in informal discussions with Gabor Herman, Yair Censor, Rob Lewitt and me at MIPG in Philadelphia in the late 1990's.

10.4 Eigenvalues and Norms: A Summary

It is helpful, at this point, to summarize the main facts concerning eigenvalues and norms. Throughout this section A will denote an arbitrary matrix, S an arbitrary square matrix, and H an arbitrary Hermitian matrix. We denote by $\|A\|$ an arbitrary induced matrix norm of A.

Here are some of the things we now know:

- 1. $\rho(S^2) = \rho(S)^2$;
- 2. $\rho(S) \leq \|S\|$, for any matrix norm;
- 3. $\rho(H) = \|H\|_2 \leq \|H\|$, for any matrix norm;
- 4. $\|A\|_2^2 = \rho(A^T A) = \|A^T A\|_2 \leq \|A^T A\|$;
- 5. $\|A^T A\|_1 \leq \|A^T\|_1 \|A\|_1 = \|A\|_\infty \|A\|_1$;
- 6. $\|A\|_2^2 \leq \|A\|_1 \|A\|_\infty$;
- 7. $\rho(S) \leq \min\{\|S\|_1, \|S\|_\infty\}$;
- 8. if $\sum_{j=1}^J |A_{ij}|^2 \leq 1$, for all i, then $\|A\|_2^2 \leq s$, where s is the largest number of non-zero entries in any column of A.

Chapter 11

A Tale of Two Algorithms

11.1 Overview

The algorithms discussed in the previous chapters are based on the two-norm. Beginning with the present chapter we extend our study to include other distances. In this chapter the focus is on the cross-entropy, or Kullback-Leibler, distance between non-negative vectors.

Although the EMML and SMART algorithms have quite different histories and are not typically considered together, they are closely related, as we shall see [27, 28]. In this chapter we examine these two algorithms in tandem, following [29]. Forging a link between the EMML and SMART led to a better understanding of both of these algorithms and to new results. The proof of convergence of the SMART in the inconsistent case [27] was based on the analogous proof for the EMML [130], while discovery of the faster version of the EMML, the *rescaled block-iterative* EMML (RBI-EMML) [30] came from studying the analogous block-iterative version of SMART [56]. The proofs we give here are elementary and rely mainly on easily established properties of the cross-entropy or Kullback-Leibler distance. To illustrate this point, many of the proofs are left as exercises for the reader.

11.2 Notation

Let A be an I by J matrix with entries $A_{ij} \geq 0$, such that, for each $j = 1, ..., J$, we have $s_j = \sum_{i=1}^{I} A_{ij} > 0$. Let $b = (b_1, ..., b_I)^T$ with $b_i > 0$ for each i. We shall assume throughout this chapter that $s_j = 1$ for each j. If this is not the case initially, we replace x_j with $x_j s_j$ and A_{ij} with A_{ij}/s_j; the quantities $(Ax)_i$ are unchanged.

11.3 The Kullback-Leibler Distance

For $a > 0$ and $b > 0$, let the cross-entropy or Kullback-Leibler distance from a to b be

$$KL(a,b) = a \log \frac{a}{b} + b - a, \tag{11.1}$$

with $KL(a,0) = +\infty$, and $KL(0,b) = b$. Extend to nonnegative vectors coordinate-wise, so that

$$KL(x,z) = \sum_{j=1}^{J} KL(x_j, z_j). \tag{11.2}$$

Unlike the Euclidean distance, the KL distance is not symmetric; $KL(Ax,b)$ and $KL(b,Ax)$ are distinct, and we can obtain different approximate solutions of $Ax = b$ by minimizing these two distances with respect to non-negative x. Clearly, the KL distance has the property $KL(cx,cz) = cKL(x,z)$ for all positive scalars c.

Ex. 11.1 *Let* $z_+ = \sum_{j=1}^{J} z_j > 0$. *Prove that*

$$KL(x,z) = KL(x_+, z_+) + KL(x, (x_+/z_+)z). \tag{11.3}$$

As we shall see, the KL distance mimics the ordinary Euclidean distance in several ways that make it particularly useful in designing optimization algorithms. The following exercise shows that the KL distance does exhibit some behavior not normally associated with a distance.

Ex. 11.2 *Let* x *be in the interval* $(0,1)$. *Show that*

$$KL(x,1) + KL(1,x^{-1}) < KL(x,x^{-1}).$$

11.4 The Two Algorithms

The algorithms we shall consider are the *expectation maximization maximum likelihood* (EMML) method and the *simultaneous multiplicative algebraic reconstruction technique* (SMART). When $b = Ax$ has nonnegative solutions, both algorithms produce such a solution. In general, the EMML gives a nonnegative minimizer of $KL(b,Ax)$, while the SMART minimizes $KL(Ax,b)$ over nonnegative x.

For both algorithms we begin with an arbitrary positive vector x^0. The iterative step for the EMML method is

$$x_j^{k+1} = (x^k)'_j = x_j^k \sum_{i=1}^{I} A_{ij} \frac{b_i}{(Ax^k)_i}. \tag{11.4}$$

The iterative step for the SMART is

$$x_j^{m+1} = (x^m)''_j = x_j^m \exp\Big(\sum_{i=1}^{I} A_{ij} \log \frac{b_i}{(Ax^m)_i}\Big). \tag{11.5}$$

Note that, to avoid confusion, we use k for the iteration number of the EMML and m for the SMART.

11.5 Background

The *expectation maximization maximum likelihood* (EMML) method we discuss here is actually a special case of a more general approach to likelihood maximization, usually called the EM algorithm [73]; the book by McLachnan and Krishnan [106] is a good source for the history of this more general algorithm.

It was noticed by Rockmore and Macovski [117] that the image reconstruction problems that arise in medical tomography can be formulated as statistical parameter estimation problems. Following up on this idea, Shepp and Vardi [119] suggested the use of the EM algorithm for solving the reconstruction problem in emission tomography. In [101], Lange and Carson presented an EM-type iterative method for transmission tomographic image reconstruction, and pointed out a gap in the convergence proof given in [119] for the emission case. In [130], Vardi, Shepp and Kaufman repaired the earlier proof, relying on techniques due to Csiszár and Tusnády [69]. In [102] Lange, Bahn and Little improved the transmission and emission algorithms, by including regularization to reduce the effects of noise. The question of uniqueness of the solution in the inconsistent case was resolved in [27, 28].

The EMML, as a statistical parameter estimation technique, was not originally thought to be connected to any system of linear equations. In [27], it was shown that the EMML algorithm minimizes the function $f(x) = KL(b, Ax)$, over non-negative vectors x. As in the previous section, b is a vector with positive entries, and A is a matrix with non-negative entries, such that $s_j = \sum_{i=1}^{I} A_{ij} = 1$. Consequently, when the non-negative system of linear equations $Ax = b$ has a non-negative solution, the EMML converges to such a solution.

The EMML has been the subject of much attention in the medical-imaging literature over the past decade. Statisticians like it because it is based on the well-studied principle of likelihood maximization for parameter estimation. Physicists like it because, unlike its competition, filtered back-projection, it permits the inclusion of sophisticated models of the physical situation. Mathematicians like it because it can be derived from iterative optimization theory. Physicians like it because the images are often better than those produced by other means. No method is perfect, however, and the EMML suffers from sensitivity to noise and slow rate of convergence. Research is ongoing to find faster and less sensitive versions of this algorithm.

Another class of iterative algorithms was introduced into medical imaging by Gordon et al. in [87]. These include the *algebraic reconstruction technique* (ART) and its multiplicative version, MART. These methods were derived by viewing image reconstruction as solving systems of linear equations, possibly subject to constraints, such as positivity.

What is usually called the simultaneous multiplicative algebraic reconstruction technique (SMART) was discovered in 1972, independently, by Darroch and Ratcliff [70], working in statistics, and by Schmidlin [118] in medical imaging. The SMART provides another example of alternating minimization having the three- and four-point properties.

Darroch and Ratcliff called their algorithm *generalized iterative scaling*. It was designed to calculate the entropic projection of one probability vector onto a family of probability vectors with a pre-determined marginal distribution. They did not consider the more general problems of finding a non-negative solution of a non-negative system of linear equations $Ax = b$, or of minimizing a function; they did not, therefore, consider what happens in the inconsistent case, in which the system of equations $Ax = b$ has no non-negative solutions. This issue was resolved in [27], where it was shown that the SMART minimizes the function $f(x) = KL(Ax, b)$, over non-negative vectors x. Here b is a vector with positive entries, and A is a matrix with non-negative entries, such that $s_j = \sum_{i=1}^{I} A_{ij} > 0$ for all j.

11.6 The Alternating Minimization Paradigm

For each nonnegative vector x for which $(Ax)_i = \sum_{j=1}^{J} A_{ij}x_j > 0$, let $r(x) = \{r(x)_{ij}\}$ and $q(x) = \{q(x)_{ij}\}$ be the I by J arrays with entries

$$r(x)_{ij} = x_j A_{ij} \frac{b_i}{(Ax)_i}$$

and

$$q(x)_{ij} = x_j A_{ij}.$$

The KL distances

$$KL(r(x), q(z)) = \sum_{i=1}^{I} \sum_{j=1}^{J} KL(r(x)_{ij}, q(z)_{ij})$$

and

$$KL(q(x), r(z)) = \sum_{i=1}^{I} \sum_{j=1}^{J} KL(q(x)_{ij}, r(z)_{ij})$$

will play important roles in the discussion that follows. Note that if there is a nonnegative x with $r(x) = q(x)$ then $b = Ax$.

11.6.1 Some Pythagorean Identities Involving the KL Distance

The iterative algorithms we discuss in this chapter are derived using the principle of *alternating minimization*, according to which the distances $KL(r(x), q(z))$ and $KL(q(x), r(z))$ are minimized, first with respect to the variable x and then with respect to the variable z. Although the KL distance is not Euclidean, and, in particular, not even symmetric, there are analogues of Pythagoras' theorem that play important roles in the convergence proofs.

Ex. 11.3 *Establish the following Pythagorean identities:*

$$KL(r(x), q(z)) = KL(r(z), q(z)) + KL(r(x), r(z)); \tag{11.6}$$

$$KL(r(x), q(z)) = KL(r(x), q(x')) + KL(x', z), \tag{11.7}$$

for

$$x'_j = x_j \sum_{i=1}^{I} A_{ij} \frac{b_i}{(Ax)_i}; \tag{11.8}$$

$$KL(q(x), r(z)) = KL(q(x), r(x)) + KL(x, z) - KL(Ax, Az); \tag{11.9}$$

$$KL(q(x), r(z)) = KL(q(z''), r(z)) + KL(x, z''), \tag{11.10}$$

for

$$z''_j = z_j \exp(\sum_{i=1}^{I} A_{ij} \log \frac{b_i}{(Az)_i}). \tag{11.11}$$

Note that it follows from Equation (11.3) that $KL(x, z) - KL(Ax, Az) \geq 0$.

11.6.2 Convergence of the SMART and EMML

We shall prove convergence of the SMART and EMML algorithms through a series of exercises.

Ex. 11.4 *Show that, for $\{x^k\}$ given by Equation (11.4), $\{KL(b, Ax^k)\}$ is decreasing and $\{KL(x^{k+1}, x^k)\} \to 0$. Show that, for $\{x^m\}$ given by Equation (11.5), $\{KL(Ax^m, b)\}$ is decreasing and $\{KL(x^m, x^{m+1})\} \to 0$. Hint: Use $KL(r(x), q(x)) = KL(b, Ax)$, $KL(q(x), r(x)) = KL(Ax, b)$, and the Pythagorean identities.*

Ex. 11.5 *Show that the EMML sequence $\{x^k\}$ is bounded by showing*

$$\sum_{j=1}^{J} x_j^{k+1} = \sum_{i=1}^{I} b_i.$$

Show that the SMART sequence $\{x^m\}$ is bounded by showing that

$$\sum_{j=1}^{J} x_j^{m+1} \leq \sum_{i=1}^{I} b_i.$$

Ex. 11.6 *Show that $(x^*)' = x^*$ for any cluster point x^* of the EMML sequence $\{x^k\}$ and that $(x^*)'' = x^*$ for any cluster point x^* of the SMART sequence $\{x^m\}$. Hint: Use $\{KL(x^{k+1}, x^k)\} \to 0$ and $\{KL(x^m, x^{m+1})\} \to 0$.*

Ex. 11.7 *Let $\hat{x}$ and $\tilde{x}$ minimize $KL(b, Ax)$ and $KL(Ax, b)$, respectively, over all $x \geq 0$. Then, $(\hat{x})' = \hat{x}$ and $(\tilde{x})'' = \tilde{x}$. Hint: Apply Pythagorean identities to $KL(r(\hat{x}), q(\hat{x}))$ and $KL(q(\tilde{x}), r(\tilde{x}))$.*

Note that, because of convexity properties of the KL distance, even if the minimizers $\hat{x}$ and $\tilde{x}$ are not unique, the vectors $A\hat{x}$ and $A\tilde{x}$ are unique.

Ex. 11.8 *For the EMML sequence $\{x^k\}$ with cluster point x^* and $\hat{x}$ as defined previously, we have the double inequality*

$$KL(\hat{x}, x^k) \geq KL(r(\hat{x}), r(x^k)) \geq KL(\hat{x}, x^{k+1}), \tag{11.12}$$

from which we conclude that the sequence $\{KL(\hat{x}, x^k)\}$ is decreasing and $KL(\hat{x}, x^*) < +\infty$. Hint: For the first inequality calculate $KL(r(\hat{x}), q(x^k))$ in two ways. For the second one, use $(x)'_j = \sum_{i=1}^{I} r(x)_{ij}$ and Equation (11.3).

Ex. 11.9 *Show that, for the SMART sequence* $\{x^m\}$ *with cluster point* x^* *and* $\tilde{x}$ *as defined previously, we have*

$$KL(\tilde{x}, x^m) - KL(\tilde{x}, x^{m+1}) = KL(Ax^{m+1}, b) - KL(A\tilde{x}, b) +$$

$$KL(A\tilde{x}, Ax^m) + KL(x^{m+1}, x^m) - KL(Ax^{m+1}, Ax^m), \qquad (11.13)$$

and so $KL(A\tilde{x}, Ax^*) = 0$*, the sequence* $\{KL(\tilde{x}, x^m)\}$ *is decreasing and* $KL(\tilde{x}, x^*) < +\infty$*. Hint: Expand* $KL(q(\tilde{x}), r(x^m))$ *using the Pythagorean identities.*

Ex. 11.10 *For* x^* *a cluster point of the EMML sequence* $\{x^k\}$ *we have* $KL(b, Ax^*) = KL(b, A\hat{x})$*. Therefore,* x^* *is a nonnegative minimizer of* $KL(b, Ax)$*. Consequently, the sequence* $\{KL(x^*, x^k)\}$ *converges to zero, and so* $\{x^k\} \to x^*$*. Hint: Use the double inequality of Equation (11.12) and* $KL(r(\hat{x}), q(x^*))$*.*

Ex. 11.11 *For* x^* *a cluster point of the SMART sequence* $\{x^m\}$ *we have* $KL(Ax^*, b) = KL(A\tilde{x}, b)$*. Therefore,* x^* *is a nonnegative minimizer of* $KL(Ax, b)$*. Consequently, the sequence* $\{KL(x^*, x^m)\}$ *converges to zero, and so* $\{x^m\} \to x^*$*. Moreover,*

$$KL(\tilde{x}, x^0) \geq KL(x^*, x^0)$$

for all $\tilde{x}$ *as before. Hints: Use Exercise 11.9. For the final assertion use the fact that the difference* $KL(\tilde{x}, x^m) - KL(\tilde{x}, x^{m+1})$ *is independent of the choice of* $\tilde{x}$*, since it depends only on* $Ax^* = A\tilde{x}$*. Now sum over the index* m*.*

Both the EMML and the SMART algorithms are slow to converge. For that reason attention has shifted, in recent years, to *block-iterative* versions of these algorithms.

11.7 Regularization

We discussed previously how the least squares solution can become overly sensitive to noise in the data. The same is true for the SMART and EMML solutions. Both methods benefit from regularization.

11.7.1 Regularizing the SMART

One way to regularize the SMART is to minimize the function

$$f(x) = KL(Ax, b) + \sum_{j=1}^{J} \delta_j KL(x_j, p_j), \tag{11.14}$$

where $p > 0$ is a prior estimate of the desired solution and $\delta_j > 0$.

As we have seen, the iterative step of the SMART is obtained by minimizing the function $KL(q(x), r(x^n))$ over non-negative x, and the limit of the SMART minimizes $KL(Ax, b)$. To obtain x^{n+1} from x^n, we minimize

$$KL(q(x), r(x^n)) + \sum_{j=1}^{J} \delta_j KL(x_j, p_j).$$

There are many penalty functions we could use here, but the one we have chosen permits the minimizing x^{n+1} to be obtained in closed form.

The iterative step of the regularized SMART is as follows:

$$\log x_j^{n+1} = \frac{\delta_j}{\delta_j + s_j} \log p_j +$$

$$\frac{s_j}{\delta_j + s_j}\Big(\log x_j^n + s_j^{-1} \sum_{i=1}^{I} A_{ij} \log \frac{b_i}{(Ax^n)_i}\Big). \tag{11.15}$$

11.7.2 Regularizing the EMML

As we have seen, the iterative step of the EMML is obtained by minimizing the function $KL(r(x^n), q(x))$ over non-negative x, and the limit of the EMML minimizes $KL(b, Ax)$. We can regularize by minimizing

$$KL(b, Ax) + \sum_{j=1}^{J} \delta_j KL(p_j, x_j). \tag{11.16}$$

To obtain x^{n+1} from x^n, we minimize

$$KL(r(x^n), q(x)) + \sum_{j=1}^{J} \delta_j KL(p_j, x_j).$$

Again, there are many penalty functions we could use here, but the one we have chosen permits the minimizing x^{n+1} to be obtained in closed form.

The iterative step of the regularized EMML is as follows:

$$x_j^{n+1} = \frac{\delta_j}{\delta_j + s_j} p_j + \frac{1}{\delta_j + s_j} x_j^n \sum_{i=1}^{I} A_{ij} \Big(\frac{b_i}{(Ax^n)_i}\Big). \tag{11.17}$$

Chapter 12

MART and EMART

12.1 The MART Algorithm

Both the algebraic reconstruction technique (ART) and the *multiplicative algebraic reconstruction technique* (MART) were introduced by Gordon, Bender and Herman [87] as two iterative methods for discrete image reconstruction in transmission tomography. It was noticed somewhat later that the ART is a special case of Kaczmarz's algorithm [96].

Both methods are what are called *row-action* methods, meaning that each step of the iteration uses only a single equation from the system. The MART is limited to non-negative systems for which non-negative solutions are sought. In the under-determined case, both algorithms find the solution closest to the starting vector, in the two-norm or weighted two-norm sense for ART, and in the cross-entropy sense for MART, so both algorithms can be viewed as solving optimization problems. We consider two different versions of the MART.

12.1.1 MART I

The iterative step of the first version of MART, which we call MART I, is the following: for $n = 0, 1, ...,$ and $i = n(\text{mod}\, I) + 1$, let

$$x_j^{n+1} = x_j^n \Big(\frac{b_i}{(Ax^n)_i} \Big)^{A_{ij}/m_i},$$

for $j = 1, ..., J$, where the parameter m_i is defined to be

$$m_i = \max\{A_{ij} | j = 1, ..., J\}.$$

The MART I algorithm converges, in the consistent case, to the non-negative solution of $Ax = b$ for which the KL distance $KL(x, x^0)$ is minimized.

12.1.2 MART II

The iterative step of the second version of MART, which we shall call MART II, is the following: for $n = 0, 1, ...,$ and $i = n(\mathrm{mod}\, I) + 1$, let

$$x_j^{n+1} = x_j^n \Big(\frac{b_i}{(Ax^n)_i} \Big)^{A_{ij}/s_j n_i},$$

for $j = 1, ..., J$, where the parameter n_i is defined to be

$$n_i = \max\{A_{ij} s_j^{-1} | j = 1, ..., J\},$$

and

$$s_j = \sum_{i=1}^{I} A_{ij}.$$

The MART II algorithm converges,in the consistent case, to the non-negative solution of $Ax = b$ for which the KL distance

$$\sum_{j=1}^{J} s_j KL(x_j, x_j^0)$$

is minimized. Just as the Landweber method is a simultaneous cousin of the row-action ART, the MART, not surprisingly, is the row-action cousin of the SMART.

12.2 The EMART Algorithm

When there are non-negative solutions of the non-negative system $Ax = b$, the MART converges faster than the SMART, and to the same solution. The SMART involves exponentiation and a logarithm, and the MART a non-integral power, both of which complicate their calculation. The EMML is considerably simpler in this respect, but, like SMART, converges slowly. We would like to have a row-action variant of the EMML that converges faster than the EMML in the consistent case, but is easier to calculate than the MART. The EM-MART is such an algorithm.

As with the MART, we distinguish two versions, EM-MART I and EM-MART II. When the system $Ax = b$ has non-negative solutions, both EM-MART I and EM-MART II converge to non-negative solutions, but nothing further is known about these solutions. To motivate these algorithms, we rewrite the MART algorithms as follows:

12.2.1 MART I again

The iterative step of MART I can be written as follows: for $n = 0, 1, ...,$ and $i = n(\mathrm{mod}\, I) + 1$, let

$$x_j^{n+1} = x_j^n \exp\Big(\Big(\frac{A_{ij}}{m_i}\Big)\log\Big(\frac{b_i}{(Ax^n)_i}\Big)\Big),$$

or, equivalently, as

$$\log x_j^{n+1} = \Big(1 - \frac{A_{ij}}{m_i}\Big)\log x_j^n + \Big(\frac{A_{ij}}{m_i}\Big)\log\Big(x_j^n \frac{b_i}{(Ax^n)_i}\Big). \tag{12.1}$$

12.2.2 MART II again

Similarly, the iterative step of MART II can be written as follows: for $n = 0, 1, ...,$ and $i = n(\mathrm{mod}\, I) + 1$, let

$$x_j^{n+1} = x_j^n \exp\Big(\Big(\frac{A_{ij}}{s_j n_i}\Big)\log\Big(\frac{b_i}{(Ax^n)_i}\Big)\Big),$$

or, equivalently, as

$$\log x_j^{n+1} = \Big(1 - \frac{A_{ij}}{s_j n_i}\Big)\log x_j^n + \Big(\frac{A_{ij}}{s_j n_i}\Big)\log\Big(x_j^n \frac{b_i}{(Ax^n)_i}\Big). \tag{12.2}$$

We obtain the EM-MART I and EM-MART II simply by removing the logarithms in Equations (12.1) and (12.2), respectively.

12.2.3 EM-MART I

The iterative step of EM-MART I is as follows: for $n = 0, 1, ...,$ and $i = n(\mathrm{mod}\, I) + 1$, let

$$x_j^{n+1} = \Big(1 - \frac{A_{ij}}{m_i}\Big)x_j^n + \Big(\frac{A_{ij}}{m_i}\Big)\Big(x_j^n \frac{b_i}{(Ax^n)_i}\Big). \tag{12.3}$$

12.2.4 EM-MART II

The iterative step of EM-MART II is as follows:

$$x_j^{n+1} = \Big(1 - \frac{A_{ij}}{s_j n_i}\Big)x_j^n + \Big(\frac{A_{ij}}{s_j n_i}\Big)\Big(x_j^n \frac{b_i}{(Ax^n)_i}\Big). \tag{12.4}$$

Convergence of the MART and EMART algorithms follows from the more general convergence theorem for block-iterative algorithms to be discussed in the next chapter.

When the system $Ax = b$ has no non-negative solutions, neither MART nor EMART converge. Instead,as with ART and BILW, they always exhibit sub-sequential convergence to a limit cycle. However, no proof is known that this subsequential convergence necessarily occurs.

Chapter 13

Rescaled Block-Iterative Algorithms

13.1 Ordered-Subset Versions

Those who have used the SMART or the EMML on sizable problems have certainly noticed that they are both slow to converge. An important issue, therefore, is how to accelerate convergence. One popular method is through the use of *block-iterative* (or *ordered subset*) methods. As was the case when we discussed the SMART and the EMML algorithms, A_{ij} is non-negative and b_i is positive, for all i and j.

To illustrate block-iterative methods and to motivate our subsequent discussion we consider now the *ordered subset* EM algorithm (OSEM), which is a popular technique in some areas of medical imaging, as well as an analogous version of SMART, which we shall call here the OSSMART. The OSEM is now used quite frequently in tomographic image reconstruction, where it is acknowledged to produce usable images significantly faster then EMML. From a theoretical perspective both OSEM and OSSMART are incorrect. How to correct them is the subject of much that follows here.

The idea behind the OSEM (OSSMART) is simple: the iteration looks very much like the EMML (SMART), but at each step of the iteration the summations are taken only over the current block. The blocks are processed cyclically.

We begin by decomposing the set $\{i = 1, ..., I\}$ into M (not necessarily disjoint) subsets B_m, $m = 1, ..., M$. For each m let

$$s_{mj} := \sum_{i \in B_m} A_{ij} > 0. \tag{13.1}$$

The OSEM iteration is the following: for $n = 0, 1, ...$ and $m = n(\text{mod}\, M)+$

1, having found x^n let

OSEM:

$$x_j^{n+1} = x_j^n s_{mj}^{-1} \sum_{i\in B_m} A_{ij} \frac{b_i}{(Ax^n)_i}. \tag{13.2}$$

The OSSMART has the following iterative step:

OSSMART

$$x_j^{n+1} = x_j^n \exp\Big(s_{mj}^{-1} \sum_{i\in B_m} A_{ij} \log \frac{b_i}{(Ax^n)_i}\Big). \tag{13.3}$$

In general we do not expect block-iterative algorithms to converge in the inconsistent case, but to exhibit *subsequential convergence* to a *limit cycle*, as happens with MART and EMART. We do, however, want them to converge to a solution in the consistent case; the OSEM and OSSMART fail to do this except when the matrix A and the set of blocks $\{B_m,\, m = 1, ..., M\}$ satisfy the condition known as *subset balance*, which means that the sums s_{mj} depend only on j and not on m. While this may be approximately valid in some special cases, it is overly restrictive, eliminating, for example, almost every set of blocks whose cardinalities are not all the same. When the OSEM does well in practice in medical imaging it is probably because the M is not large and only a few iterations are carried out.

The experience with the OSEM was encouraging, however, and strongly suggested that an equally fast, but mathematically correct, block-iterative version of EMML was to be had; this is the *rescaled block-iterative* EMML (RBI-EMML). Both RBI-EMML and an analogous corrected version of OSSMART, the RBI-SMART, provide fast convergence to a solution in the consistent case, for any choice of blocks.

13.2 The RBI-SMART

We turn next to the block-iterative versions of the SMART, which we shall denote BI-SMART. These methods were known prior to the discovery of RBI-EMML and played an important role in that discovery; the importance of rescaling for acceleration was apparently not appreciated, however.

We start by considering a formulation of BI-SMART that is general enough to include all of the variants we wish to discuss. As we shall see, this formulation is too general and will need to be restricted in certain ways to obtain convergence. Let the iterative step be

$$x_j^{n+1} = x_j^n \exp\Big(\beta_{mj} \sum_{i\in B_m} \alpha_{mi} A_{ij} \log\Big(\frac{b_i}{(Ax^n)_i}\Big)\Big), \tag{13.4}$$

for $j = 1, 2, ..., J$, $m = n(\bmod M) + 1$ and β_{mj} and α_{mi} positive. As we shall see, our convergence proof will require that β_{mj} be separable, that is, $\beta_{mj} = \gamma_j \delta_m$ for each j and m and that

$$\gamma_j \delta_m \sigma_{mj} \leq 1. \tag{13.5}$$

With these conditions satisfied we have the following result.

Theorem 13.1 *Suppose that we are in the consistent case, in which the system $Ax = b$ has non-negative solutions. For any positive vector x^0 and any collection of blocks $\{B_m, m = 1, ..., M\}$ the sequence $\{x^n\}$ given by Equation (13.4) converges to the unique solution of $b = Ax$ for which the weighted cross-entropy $\sum_{j=1}^{J} \gamma_j^{-1} KL(x_j, x_j^0)$ is minimized.*

The inequality in the following lemma is the basis for the convergence proof.

Lemma 13.1 *Let $b = Ax$ for some nonnegative x. Then for $\{x^n\}$ as in Equation (13.4) we have*

$$\sum_{j=1}^{J} \gamma_j^{-1} KL(x_j, x_j^n) - \sum_{j=1}^{J} \gamma_j^{-1} KL(x_j, x_j^{n+1}) \geq \tag{13.6}$$

$$\delta_m \sum_{i \in B_m} \alpha_{mi} KL(b_i, (Ax^n)_i). \tag{13.7}$$

Proof: First note that

$$x_j^{n+1} = x_j^n \exp\Big(\gamma_j \delta_m \sum_{i \in B_m} \alpha_{mi} A_{ij} \log\Big(\frac{b_i}{(Ax^n)_i}\Big)\Big), \tag{13.8}$$

and

$$\exp\Big(\gamma_j \delta_m \sum_{i \in B_m} \alpha_{mi} A_{ij} \log\Big(\frac{b_i}{(Ax^n)_i}\Big)\Big) \tag{13.9}$$

can be written as

$$\exp\Big((1 - \gamma_j \delta_m \sigma_{mj}) \log 1 + \gamma_j \delta_m \sum_{i \in B_m} \alpha_{mi} A_{ij} \log\Big(\frac{b_i}{(Ax^n)_i}\Big)\Big), \tag{13.10}$$

which, by the convexity of the exponential function, is not greater than

$$(1 - \gamma_j \delta_m \sigma_{mj}) + \gamma_j \delta_m \sum_{i \in B_m} \alpha_{mi} A_{ij} \frac{b_i}{(Ax^n)_i}. \tag{13.11}$$

It follows that

$$\sum_{j=1}^{J}\gamma_j^{-1}(x_j^n - x_j^{n+1}) \geq \delta_m \sum_{i\in B_m}\alpha_{mi}((Ax^n)_i - b_i). \tag{13.12}$$

We also have

$$\log(x_j^{n+1}/x_j^n) = \gamma_j\delta_m \sum_{i\in B_m}\alpha_{mi}A_{ij}\log\frac{b_i}{(Ax^n)_i}. \tag{13.13}$$

Therefore

$$\sum_{j=1}^{J}\gamma_j^{-1}KL(x_j, x_j^n) - \sum_{j=1}^{J}\gamma_j^{-1}KL(x_j, x_j^{n+1}) \tag{13.14}$$

$$= \sum_{j=1}^{J}\gamma_j^{-1}(x_j\log(x_j^{n+1}/x_j^n) + x_j^n - x_j^{n+1}) \tag{13.15}$$

$$= \sum_{j=1}^{J}x_j\delta_m \sum_{i\in B_m}\alpha_{mi}A_{ij}\log\frac{b_i}{(Ax^n)_i} + \sum_{j=1}^{J}\gamma_j^{-1}(x_j^n - x_j^{n+1}) \tag{13.16}$$

$$= \delta_m \sum_{i\in B_m}\alpha_{mi}(\sum_{j=1}^{J}x_jA_{ij})\log\frac{b_i}{(Ax^n)_i} + \sum_{j=1}^{J}\gamma_j^{-1}(x_j^n - x_j^{n+1}) \tag{13.17}$$

$$\geq \delta_m\Big(\sum_{i\in B_m}\alpha_{mi}(b_i\log\frac{b_i}{(Ax^n)_i} + (Ax^n)_i - b_i)\Big) = \delta_m \sum_{i\in B_m}\alpha_{mi}KL(b_i, (Ax^n)_i). \tag{13.18}$$

This completes the proof of the lemma. ∎

From the inequality (13.7) we conclude that the sequence

$$\{\sum_{j=1}^{J}\gamma_j^{-1}KL(x_j, x_j^n)\} \tag{13.19}$$

is decreasing, that $\{x^n\}$ is therefore bounded and the sequence

$$\{\sum_{i\in B_m}\alpha_{mi}KL(b_i, (Ax^n)_i)\} \tag{13.20}$$

is converging to zero. Let x^* be any cluster point of the sequence $\{x^n\}$. Then it is not difficult to show that $b = Ax^*$. Replacing x with x^* we have that the sequence $\{\sum_{j=1}^J \gamma_j^{-1} KL(x_j^*, x_j^n)\}$ is decreasing; since a subsequence converges to zero, so does the whole sequence. Therefore x^* is the limit of the sequence $\{x^n\}$. This proves that the algorithm produces a solution of $b = Ax$. To conclude further that the solution is the one for which the quantity $\sum_{j=1}^J \gamma_j^{-1} KL(x_j, x_j^0)$ is minimized requires further work to replace the inequality (13.7) with an equation in which the right side is independent of the particular solution x chosen; see [46] for details.

We see from the theorem that how we select the γ_j is determined by how we wish to weight the terms in the sum $\sum_{j=1}^J \gamma_j^{-1} KL(x_j, x_j^0)$. In some cases we want to minimize the cross-entropy $KL(x, x^0)$ subject to $b = Ax$; in this case we would select $\gamma_j = 1$. In other cases we may have some prior knowledge as to the relative sizes of the x_j and wish to emphasize the smaller values more; then we may choose γ_j proportional to our prior estimate of the size of x_j. Having selected the γ_j, we see from the inequality (13.7) that convergence will be accelerated if we select δ_m as large as permitted by the condition $\gamma_j \delta_m \sigma_{mj} \leq 1$. This suggests that we take

$$\delta_m = 1/\min\{\sigma_{mj}\gamma_j,\, j = 1, ..., J\}. \tag{13.21}$$

The *rescaled* BI-SMART (RBI-SMART) as presented in [30, 31, 32] uses this choice, but with $\alpha_{mi} = 1$ for each m and i. For each $m = 1, ..., M$ let

$$r_m = \max\{s_{mj} s_j^{-1} | j = 1, ..., J\}. \tag{13.22}$$

The original RBI-SMART is as follows:

Algorithm 13.1 (RBI-SMART) *Let x^0 be an arbitrary positive vector. For $n = 0, 1, ...,$ let $m = n(\mathrm{mod}\, M) + 1$. Then let*

$$x_j^{n+1} = x_j^n \exp\Big(r_m^{-1} s_j^{-1} \sum_{i \in B_m} A_{ij} \log\Big(\frac{b_i}{(Ax^n)_i}\Big)\Big). \tag{13.23}$$

Notice that Equation (13.23) can be written as

$$\log x_j^{n+1} = (1 - r_m^{-1} s_j^{-1} s_{mj}) \log x_j^n + r_m^{-1} s_j^{-1} \sum_{i \in B_m} A_{ij} \log\Big(x_j^n \frac{b_i}{(Ax^n)_i}\Big), \tag{13.24}$$

from which we see that x_j^{n+1} is a weighted geometric mean of x_j^n and the terms

$$(Q_i x^n)_j = x_j^n \Big(\frac{b_i}{(Ax^n)_i}\Big),$$

for $i \in B_m$. This will be helpful in deriving block-iterative versions of the EMML algorithm. The vectors $Q_i(x^n)$ are sometimes called weighted KL projections.

Let's look now at some of the other choices for these parameters that have been considered in the literature.

First, we notice that the OSSMART does not generally satisfy the requirements, since in (13.3) the choices are $\alpha_{mi} = 1$ and $\beta_{mj} = s_{mj}^{-1}$; the only times this is acceptable is if the s_{mj} are separable; that is, $s_{mj} = u_j t_m$ for some u_j and t_m. This is slightly more general than the condition of subset balance and is sufficient for convergence of OSSMART.

In [56] Censor and Segman make the choices $\beta_{mj} = 1$ and $\alpha_{mi} > 0$ such that $\sigma_{mj} \leq 1$ for all m and j. In those cases in which σ_{mj} is much less than 1 for each m and j their iterative scheme is probably excessively relaxed; it is hard to see how one might improve the rate of convergence by altering only the weights α_{mi}, however. Limiting the choice to $\gamma_j \delta_m = 1$ reduces our ability to accelerate this algorithm.

The original SMART uses $M = 1$, $\gamma_j = s_j^{-1}$ and $\alpha_{mi} = \alpha_i = 1$. Clearly the inequality (13.5) is satisfied; in fact it becomes an equality now.

For the row-action version of SMART, the *multiplicative* ART (MART), due to Gordon, Bender and Herman [87], we take $M = I$ and $B_m = B_i = \{i\}$ for $i = 1, ..., I$. The MART has the iterative

$$x_j^{n+1} = x_j^n \Big(\frac{b_i}{(Ax^n)_i} \Big)^{m_i^{-1} A_{ij}}, \tag{13.25}$$

for $j = 1, 2, ..., J$, $i = n(\text{mod}\, I) + 1$ and $m_i > 0$ chosen so that $m_i^{-1} A_{ij} \leq 1$ for all j. The smaller m_i is the faster the convergence, so a good choice is $m_i = \max\{A_{ij}|, j = 1, ..., J\}$. Although this particular choice for m_i is not explicitly mentioned in the various discussions of MART I have seen, it was used in implementations of MART from the beginning [92].

Darroch and Ratcliff included a discussion of a block-iterative version of SMART in their 1972 paper [70]. Close inspection of their version reveals that they require that $s_{mj} = \sum_{i \in B_m} A_{ij} = 1$ for all j. Since this is unlikely to be the case initially, we might try to rescale the equations or unknowns to obtain this condition. However, unless $s_{mj} = \sum_{i \in B_m} A_{ij}$ depends only on j and not on m, which is the *subset balance* property used in [95], we cannot redefine the unknowns in a way that is independent of n.

The MART fails to converge in the inconsistent case. What is always observed, but for which no proof exists, is that, for each fixed $i = 1, 2, ..., I$, as $k \to +\infty$, the MART subsequences $\{x^{kI+i}\}$ converge to separate limit vectors, say $x^{\infty,i}$. This *limit cycle* LC $= \{x^{\infty,i} | i = 1, ..., I\}$ reduces to a single vector whenever there is a nonnegative solution of $b = Ax$. The greater the minimum value of $KL(Ax, b)$ the more distinct from one another the vectors of the limit cycle are. An analogous result is observed for BI-SMART.

13.3 The RBI-EMML

As we did with SMART, we consider now a formulation of BI-EMML that is general enough to include all of the variants we wish to discuss. Once again, the formulation is too general and will need to be restricted in certain ways to obtain convergence. Let the iterative step be

$$x_j^{n+1} = x_j^n(1 - \beta_{mj}\sigma_{mj}) + x_j^n\beta_{mj}\sum_{i\in B_m}\alpha_{mi}A_{ij}\frac{b_i}{(Ax^n)_i}, \tag{13.26}$$

for $j = 1, 2, ..., J$, $m = n(\mathrm{mod}\, M) + 1$ and β_{mj} and α_{mi} positive. As in the case of BI-SMART, our convergence proof will require that β_{mj} be separable, that is,

$$\beta_{mj} = \gamma_j\delta_m \tag{13.27}$$

for each j and m and that the inequality (13.5) hold. With these conditions satisfied we have the following result.

Theorem 13.2 *Suppose that we are in the consistent case. For any positive vector x^0 and any collection of blocks $\{B_m, m = 1, ..., M\}$ the sequence $\{x^n\}$ given by Equation (13.4) converges to a nonnegative solution of $b = Ax$.*

When there are multiple nonnegative solutions of $b = Ax$ the solution obtained by BI-EMML will depend on the starting point x^0, but precisely how it depends on x^0 is an open question. Also, in contrast to the case of BI-SMART, the solution can depend on the particular choice of the blocks. The inequality in the following lemma is the basis for the convergence proof.

Lemma 13.2 *Let $Ax = b$ for some non-negative x. Then, for $\{x^n\}$ as in Equation (13.26), we have*

$$\sum_{j=1}^{J}\gamma_j^{-1}KL(x_j, x_j^n) - \sum_{j=1}^{J}\gamma_j^{-1}KL(x_j, x_j^{n+1}) \geq \tag{13.28}$$

$$\delta_m\sum_{i\in B_m}\alpha_{mi}KL(b_i, (Ax^n)_i). \tag{13.29}$$

Proof: From the iterative step we have

$$x_j^{n+1} = x_j^n(1 - \gamma_j\delta_m\sigma_{mj}) + x_j^n\gamma_j\delta_m\sum_{i\in B_m}\alpha_{mi}A_{ij}\frac{b_i}{(Ax^n)_i} \tag{13.30}$$

$$\log(x_j^{n+1}/x_j^n) = \log\Big((1 - \gamma_j\delta_m\sigma_{mj}) + \gamma_j\delta_m\sum_{i\in B_m}\alpha_{mi}A_{ij}\frac{b_i}{(Ax^n)_i}\Big). \tag{13.31}$$

By the concavity of the logarithm we obtain the inequality

$$\log(x_j^{n+1}/x_j^n) \geq \Big((1-\gamma_j\delta_m\sigma_{mj})\log 1 + \gamma_j\delta_m \sum_{i\in B_m} \alpha_{mi}A_{ij}\log\frac{b_i}{(Ax^n)_i}\Big), \tag{13.32}$$

or

$$\log(x_j^{n+1}/x_j^n) \geq \gamma_j\delta_m \sum_{i\in B_m} \alpha_{mi}A_{ij}\log\frac{b_i}{(Ax^n)_i}. \tag{13.33}$$

Therefore

$$\sum_{j=1}^J \gamma_j^{-1}x_j\log(x_j^{n+1}/x_j^n) \geq \delta_m \sum_{i\in B_m} \alpha_{mi}(\sum_{j=1}^J x_jA_{ij})\log\frac{b_i}{(Ax^n)_i}. \tag{13.34}$$

Note that it is at this step that we used the separability of the β_{mj}. Also

$$\sum_{j=1}^J \gamma_j^{-1}(x_j^{n+1}-x_j^n) = \delta_m \sum_{i\in B_m}((Ax^n)_i - b_i). \tag{13.35}$$

This concludes the proof of the lemma. ∎

From the inequality in (13.29) we conclude, as we did in the BI-SMART case, that the sequence $\{\sum_{j=1}^J \gamma_j^{-1}KL(x_j, x_j^n)\}$ is decreasing, that $\{x^n\}$ is therefore bounded and the sequence $\{\sum_{i\in B_m}\alpha_{mi}KL(b_i,(Ax^n)_i)\}$ is converging to zero. Let x^* be any cluster point of the sequence $\{x^n\}$. Then it is not difficult to show that $b = Ax^*$. Replacing x with x^* we have that the sequence $\{\sum_{j=1}^J \gamma_j^{-1}KL(x_j^*, x_j^n)\}$ is decreasing; since a subsequence converges to zero, so does the whole sequence. Therefore x^* is the limit of the sequence $\{x^n\}$. This proves that the algorithm produces a nonnegative solution of $b = Ax$. So far, we have been unable to replace the inequality in (13.29) with an equation in which the right side is independent of the particular solution x chosen.

Having selected the γ_j, we see from the inequality in (13.29) that convergence will be accelerated if we select δ_m as large as permitted by the condition $\gamma_j\delta_m\sigma_{mj} \leq 1$. This suggests that once again we take

$$\delta_m = 1/\min\{\sigma_{mj}\gamma_j,\, j = 1, ..., J\}. \tag{13.36}$$

The *rescaled* BI-EMML (RBI-EMML) as presented in [30, 31, 32] uses this choice, but with $\alpha_{mi} = 1$ for each m and i. The original motivation for the RBI-EMML came from consideration of Equation (13.24), replacing the geometric means with arithmetic means. This RBI-EMML is as follows:

Algorithm 13.2 (RBI-EMML) *Let x^0 be an arbitrary positive vector. For $n = 0, 1, ...,$ let $m = n(\mathrm{mod}\, M) + 1$. Then let*

$$x_j^{n+1} = (1 - r_m^{-1} s_j^{-1} s_{mj}) x_j^k + r_m^{-1} s_j^{-1} x_j^n \sum_{i \in B_m} (A_{ij} \frac{b_i}{(Ax^n)_i}). \tag{13.37}$$

Let's look now at some of the other choices for these parameters that have been considered in the literature.

First, we notice that the OSEM does not generally satisfy the requirements, since in (13.2) the choices are $\alpha_{mi} = 1$ and $\beta_{mj} = s_{mj}^{-1}$; the only times this is acceptable is if the s_{mj} are separable; that is, $s_{mj} = u_j t_n$ for some u_j and t_n. This is slightly more general than the condition of subset balance and is sufficient for convergence of OSEM.

The original EMML uses $M = 1$, $\gamma_j = s_j^{-1}$ and $\alpha_{mi} = \alpha_i = 1$. Clearly the inequality (13.5) is satisfied; in fact it becomes an equality now.

Notice that the calculations required to perform the BI-SMART are somewhat more complicated than those needed in BI-EMML. Because the MART converges rapidly in most cases there is considerable interest in the row-action version of EMML. It was clear from the outset that using the OSEM in a row-action mode does not work. We see from the formula for BI-EMML that the proper row-action version of EMML, which we call the EM-MART, is the following:

Algorithm 13.3 (EM-MART) *Let x^0 be an arbitrary positive vector and $i = n(\mathrm{mod}\, I) + 1$. Then let*

$$x_j^{n+1} = (1 - \delta_i \gamma_j \alpha_{ii} A_{ij}) x_j^n + \delta_i \gamma_j \alpha_{ii} x_j^n A_{ij} \frac{b_i}{(Ax^n)_i}, \tag{13.38}$$

with

$$\gamma_j \delta_i \alpha_{ii} A_{ij} \leq 1 \tag{13.39}$$

for all i and j.

The optimal choice would seem to be to take $\delta_i \alpha_{ii}$ as large as possible; that is, to select $\delta_i \alpha_{ii} = 1/\max\{\gamma_j A_{ij}, j = 1, ..., J\}$. With this choice the EM-MART is called the *rescaled* EM-MART (REM-MART).

The EM-MART fails to converge in the inconsistent case. What is always observed, but for which no proof exists, is that, for each fixed $i = 1, 2, ..., I$, as $k \to +\infty$, the EM-MART subsequences $\{x^{kI+i}\}$ converge to separate limit vectors, say $x^{\infty,i}$. This *limit cycle* LC $= \{x^{\infty,i} | i = 1, ..., I\}$ reduces to a single vector whenever there is a nonnegative solution of $b = Ax$. The greater the minimum value of $KL(b, Ax)$ the more distinct from one another the vectors of the limit cycle are. An analogous result is observed for BI-EMML.

As we stated earlier, in the consistent case the sequence $\{x^n\}$ generated by the BI-SMART algorithm and given by Equation (13.8) converges to the unique solution of $b = Ax$ for which the distance $\sum_{j=1}^{J} \gamma_j^{-1} KL(x_j, x_j^0)$ is minimized. For details, see [46].

Chapter 14

Alternating Minimization

14.1 Overview

As we have seen, both the EMML and the SMART are best derived as alternating minimization (AM) algorithms. The idea of using the AM framework for EMML is due to Vardi, Shepp and Kaufman [130]. The main reference for alternating minimization is the paper [69] of Csiszár and Tusnády. As the authors of [130] remark, the geometric argument in [69] is "deep, though hard to follow". As we shall see, all AM methods for which the five-point property of [69] holds fall into the SUMMA class (see [49]). Consequently, both the SMART and EMML algorithms are also SUMMA algorithms.

14.2 Alternating Minimization

The alternating minimization (AM) iteration of Csiszár and Tusnády [69] provides a useful framework for the derivation of iterative optimization algorithms. In this section we discuss their five-point property and use it to obtain a somewhat simpler proof of convergence for their AM algorithm.

14.2.1 The AM Framework

Suppose that P and Q are arbitrary non-empty sets and the function $\Theta(p,q)$ satisfies $-\infty < \Theta(p,q) \leq +\infty$, for each $p \in P$ and $q \in Q$. We assume that, for each $p \in P$, there is $q \in Q$ with $\Theta(p,q) < +\infty$. Therefore, $b = \inf_{p\in P,\, q\in Q} \Theta(p,q) < +\infty$. We assume also that $b > -\infty$; in many applications, the function $\Theta(p,q)$ is non-negative, so this additional assumption is unnecessary. We do not always assume there are $\hat{p} \in P$ and $\hat{q} \in Q$ such that $\Theta(\hat{p},\hat{q}) = b$; when we do assume that such a $\hat{p}$ and $\hat{q}$

exist, we will not assume that $\hat{p}$ and $\hat{q}$ are unique with that property. The objective is to generate a sequence $\{(p^n, q^n)\}$ such that $\Theta(p^n, q^n) \to b$.

14.2.2 The AM Iteration

The general AM method proceeds in two steps: we begin with some q^0, and, having found q^n, we

- **1.** minimize $\Theta(p, q^n)$ over $p \in P$ to get $p = p^{n+1}$, and then
- **2.** minimize $\Theta(p^{n+1}, q)$ over $q \in Q$ to get $q = q^{n+1}$.

In certain applications we consider the special case of alternating cross-entropy minimization. In that case, the vectors p and q are non-negative, and the function $\Theta(p, q)$ will have the value $+\infty$ whenever there is an index j such that $p_j > 0$, but $q_j = 0$. It is important for those particular applications that we select q^0 with all positive entries. We therefore assume, for the general case, that we have selected q^0 so that $\Theta(p, q^0)$ is finite for all p.

The sequence $\{\Theta(p^n, q^n)\}$ is decreasing and bounded below by b, since we have

$$\Theta(p^n, q^n) \geq \Theta(p^{n+1}, q^n) \geq \Theta(p^{n+1}, q^{n+1}). \tag{14.1}$$

Therefore, the sequence $\{\Theta(p^n, q^n)\}$ converges to some $B \geq b$. Without additional assumptions, we can say little more.

We know two things:

$$\Theta(p^{n+1}, q^n) - \Theta(p^{n+1}, q^{n+1}) \geq 0, \tag{14.2}$$

and

$$\Theta(p^n, q^n) - \Theta(p^{n+1}, q^n) \geq 0. \tag{14.3}$$

Equation 14.3 can be strengthened to

$$\Theta(p, q^n) - \Theta(p^{n+1}, q^n) \geq 0. \tag{14.4}$$

We need to make these inequalities more precise.

14.2.3 The Five-Point Property for AM

The five-point property is the following: for all $p \in P$ and $q \in Q$ and $n = 1, 2, ...$

The Five-Point Property

$$\Theta(p, q) + \Theta(p, q^{n-1}) \geq \Theta(p, q^n) + \Theta(p^n, q^{n-1}). \tag{14.5}$$

14.2.4 The Main Theorem for AM

We want to find sufficient conditions for the sequence $\{\Theta(p^n, q^n)\}$ to converge to b, that is, for $B = b$. The following is the main result of [69].

Theorem 14.1 *If the five-point property holds then* $B = b$.

Proof: Suppose that $B > b$. Then there are p' and q' such that $B > \Theta(p', q') \geq b$. From the five-point property we have

$$\Theta(p', q^{n-1}) - \Theta(p^n, q^{n-1}) \geq \Theta(p', q^n) - \Theta(p', q'), \tag{14.6}$$

so that

$$\Theta(p', q^{n-1}) - \Theta(p', q^n) \geq \Theta(p^n, q^{n-1}) - \Theta(p', q') \geq 0. \tag{14.7}$$

All the terms being subtracted can be shown to be finite. It follows that the sequence $\{\Theta(p', q^{n-1})\}$ is decreasing, bounded below, and therefore convergent. The right side of Equation (14.7) must therefore converge to zero, which is a contradiction. We conclude that $B = b$ whenever the five-point property holds in AM. ∎

14.2.5 The Three- and Four-Point Properties

In [69] the five-point property is related to two other properties, the three- and four-point properties. This is a bit peculiar for two reasons: first, as we have just seen, the five-point property is sufficient to prove the main theorem; and second, these other properties involve a second function, $\Delta : P \times P \to [0, +\infty]$, with $\Delta(p, p) = 0$ for all $p \in P$. The three- and four-point properties jointly imply the five-point property, but to get the converse, we need to use the five-point property to define this second function; it can be done, however.

The three-point property is the following:

The Three-Point Property

$$\Theta(p, q^n) - \Theta(p^{n+1}, q^n) \geq \Delta(p, p^{n+1}), \tag{14.8}$$

for all p. The four-point property is the following:

The Four-Point Property

$$\Delta(p, p^{n+1}) + \Theta(p, q) \geq \Theta(p, q^{n+1}), \tag{14.9}$$

for all p and q.

It is clear that the three- and four-point properties together imply the five-point property. We show now that the three-point property and the four-point property are implied by the five-point property. For that purpose we need to define a suitable $\Delta(p, \tilde{p})$. For any p and $\tilde{p}$ in P define

$$\Delta(p, \tilde{p}) = \Theta(p, q(\tilde{p})) - \Theta(p, q(p)), \tag{14.10}$$

where $q(p)$ denotes a member of Q satisfying $\Theta(p, q(p)) \leq \Theta(p, q)$, for all q in Q. Clearly, $\Delta(p, \tilde{p}) \geq 0$ and $\Delta(p, p) = 0$. The four-point property holds automatically from this definition, while the three-point property follows from the five-point property. Therefore, it is sufficient to discuss only the five-point property when speaking of the AM method.

In the next two sections we discuss the SMART and EMML algorithms, two important instances of alternating minimization.

14.3 The SMART as AM

In this section we consider the *simultaneous multiplicative algebraic reconstruction technique* (SMART) as an example of AM. Let $\mathcal{X}$ be the set of all $x \geq 0$ for which the vector Ax has only positive entries. For each $x \in \mathcal{X}$, let $t(x)$ and $r(x)$ be the I by J arrays with entries

$$t(x)_{ij} = x_j A_{ij}, \tag{14.11}$$

and

$$r(x)_{ij} = x_j A_{ij} b_i / (Ax)_i. \tag{14.12}$$

We then let

$$\mathcal{R} = \{r = \{r_{ij} \geq 0\} | \sum_{j=1}^{J} r_{ij} = b_i, \text{ for } i = 1, 2, ..., I\}, \tag{14.13}$$

and

$$\mathcal{T} = \{t = t(x) | x \in \mathcal{X}\}. \tag{14.14}$$

The sets $\mathcal{R}$ and $\mathcal{T}$ are convex in the space $\mathbb{R}^{I+J}$.

The function $KL(Ax, b)$ is continuous in the variable x and has bounded level sets, so there is at least one minimizer; call it $\hat{x}$. The vector $A\hat{x}$ is unique, even if the vector $\hat{x}$ is not. For notational convenience we shall assume that $s_j = 1$ for all j. If this is not the case initially, we replace A_{ij} with A_{ij}/s_j and x_j with $x_j s_j$; the product Ax is unchanged.

The Pythagorean identities for the SMART can be written as follows:

$$KL(t(x), r(z)) = KL(t(x), r(x)) + KL(x, z) - KL(Ax, Az); \tag{14.15}$$

and

$$KL(t(x), r(z)) = KL(t(z^*), r(z)) + KL(x, z^*), \tag{14.16}$$

where x and z are arbitrary members of $\mathcal{X}$ and

$$z_j^* = z_j \exp\Big(\sum_{i=1}^{I} A_{ij} \log\Big(\frac{b_i}{(Az)_i}\Big)\Big), \tag{14.17}$$

for each j. Note that

$$KL(Ax, b) = KL(t(x), r(x)), \tag{14.18}$$

and

$$KL(x, z) - KL(Ax, Az) \geq 0. \tag{14.19}$$

To put the SMART algorithm into the framework of alternating minimization, we take the sets $Q = \mathcal{R}$ and $P = \mathcal{T}$ as above and let $p^n = t(x^n)$, and $q^n = r(x^n)$. Generic vectors are $p = t(x)$ for some $x \in \mathcal{X}$ and $q = r(z)$ for some $z \in \mathcal{X}$. Then we set

$$\Theta(p, q) = KL(t(x), r(z)), \tag{14.20}$$

and, for arbitrary $p = t(x)$ and $\tilde{p} = t(\tilde{x})$,

$$\Delta(p, \tilde{p}) = KL(t(x), t(\tilde{x})) = KL(x, \tilde{x}). \tag{14.21}$$

From the Pythagorean identity (14.16) we have

$$KL(t(x), r(x^{n-1})) = KL(t(x^n), r(x^{n-1})) + KL(x, x^n) \tag{14.22}$$

so that

$$\Theta(p, q^{n-1}) = \Theta(p^n, q^{n-1}) + \Delta(p, p^n), \tag{14.23}$$

which is then the three-point property. From

$$KL(t(x), r(x^n)) - KL(t(x), r(x)) =$$

$$KL(x, x^n) - KL(Ax, Ax^n) \leq KL(x, x^n) \tag{14.24}$$

we have

$$\Delta(p, p^n) \geq \Theta(p, q^n) - \Theta(p, q(p)) \geq \Theta(p, q^n) - \Theta(p, q), \tag{14.25}$$

which is the four-point property.

The iterative step of the SMART is then to minimize the function

$$\Theta(p, q^{n-1}) = KL(t(x), r(x^{n-1})) \tag{14.26}$$

to get $x = x^n = (x^{n-1})^*$. Since the SMART is a particular case of AM for which the five-point property holds, we know that

$$\{KL(Ax^n, b)\} \to \inf\{KL(Ax, b)\,|x \geq 0\}. \tag{14.27}$$

As we have seen, using the Pythagorean identities we can show more: the sequence $\{x^n\}$ converges to the non-negative minimizer of the function $KL(Ax, b)$ for which $KL(x, x^0)$ is minimized ([27, 29]) .

14.3.1 Related work of Csiszár

In [68] Csiszár shows that the generalized iterative scaling method of Darroch and Ratcliff can be formulated in terms of successive entropic projection onto the sets $\mathcal{R}$ and $\mathcal{T}$. In other words, he views their method as an alternating projection method, not as alternating minimization. He derives the generalized iterative scaling algorithm in two steps:

- 1. minimize $KL(r(x), t(x^n))$ to get $r(x^n)$; and then
- 2. minimize $KL(t(x), r(x^n))$ to get $t(x^{n+1})$.

Although [68] appeared five years after [69], Csiszár does not reference [69], nor does he mention alternating minimization, instead basing his convergence proof here on his earlier paper [67], which deals with entropic projection. He is able to make this work because the order of the $t(x^n)$ and $r(x)$ does not matter in the first step. Therefore, the generalized iterative scaling, and, more generally, the SMART, is also an alternating projection algorithm, as well.

14.4 The EMML Algorithm as AM

Because $KL(b, Ax)$ is continuous in the variable x and has bounded level sets, there is at least one non-negative minimizer; call it $\hat{x}$. The vector $A\hat{x}$ is unique, even if $\hat{x}$ is not.

For each $x \in \mathcal{X}$, let $t(x)$ and $r(x)$ be as previously defined. The Pythagorean identities for the EMML algorithm can be written as follows:

$$KL(r(x), t(z)) = KL(r(z), t(z)) + KL(r(x), r(z)); \tag{14.28}$$

and

$$KL(r(x), t(z)) = KL(r(x), t(x')) + KL(x', z), \tag{14.29}$$

where x and z are arbitrary members of $\mathcal{X}$ and the entries of x' are defined by

$$x'_j = x_j \sum_{i=1}^{I} A_{ij} \frac{b_i}{(Ax)_i}, \tag{14.30}$$

for each j. Note that $KL(b, Ax) = KL(r(x), t(x))$.

In the EMML algorithm we minimize the function $KL(r(x^n), t(x))$ to get $x = x^{n+1}$. The EMML iteration begins with a positive vector x^0. Having found the vector x^n, the next vector in the EMML sequence is $x^{n+1} = (x^n)'$, with entries given by

$$x_j^{n+1} = (x^n)'_j = x_j^n \sum_{i=1}^{I} A_{ij} \Big(\frac{b_i}{(Ax^n)_i} \Big) = \sum_{i=1}^{I} r(x^n)_{ij}. \tag{14.31}$$

The sequence $\{x^n\}$ converges to a non-negative minimizer of the function $KL(b, Ax)$.

We put the EMML algorithm into an AM framework using $P = \mathcal{R}$, $Q = \mathcal{T}$, $p = r(x)$, $q = t(z)$, $\Theta(p, q) = KL(r(x), t(z))$, and minimizing $KL(r(x), t(x)) = KL(b, Ax)$. Using the AM notation, we let $q^{n-1} = t(x^{n-1})$, $p^n = r(x^{n-1})$, $p = r(x)$, $\tilde{p} = r(\tilde{x})$, and $q(p) = t(x')$. At the nth step of the EMML algorithm we obtain $p^n = r(x^{n-1})$ by minimizing

$$\Theta(p, q^{n-1}) = KL(r(x), t(x^{n-1})). \tag{14.32}$$

According to the Pythagorean identities (14.28) and (14.29) and Lemma 11.3, we have $x^n = (x^{n-1})'$ and

$$\Theta(p, q^{n-1}) - \Theta(p^n, q^{n-1}) = KL(r(x), r(x^{n-1}))$$

$$\geq KL(x', (x^{n-1})') = KL(x', x^n). \tag{14.33}$$

With $\Delta(p, \tilde{p})$ defined as

$$\Delta(p, \tilde{p}) = KL(r(x), r(\tilde{x})), \tag{14.34}$$

it follows that

$$\Delta(p, p^n) = KL(r(x), r(x^{n-1})), \tag{14.35}$$

so that

$$\Theta(p, q^{n-1}) - \Theta(p^n, q^{n-1}) \geq \Delta(p, p^n), \tag{14.36}$$

which is the three-point property.

We know that

$$KL(r(x), t(x^n)) - KL(r(x), t(x')) = KL(x', x^n) \tag{14.37}$$

and

$$KL(r(x), r(x^{n-1})) \geq KL(x', x^n), \tag{14.38}$$

from which it follows that

$$KL(r(x), r(x^{n-1})) \geq KL(r(x), t(x^n)) - KL(r(x), t(x')); \tag{14.39}$$

this is the four-point property.

14.5 Alternating Bregman Distance Minimization

The general problem of minimizing $\Theta(p, q)$ is simply a minimization of a real-valued function of two variables, $p \in P$ and $q \in Q$. In many cases the function $\Theta(p, q)$ is a distance between p and q, either $\|p - q\|_2^2$ or $KL(p, q)$. In the case of $\Theta(p, q) = \|p - q\|_2^2$, each step of the alternating minimization algorithm involves an orthogonal projection onto a closed convex set; both projections are with respect to the same Euclidean distance function. In the case of cross-entropy minimization, we first project q^n onto the set P by minimizing the distance $KL(p, q^n)$ over all $p \in P$, and then project p^{n+1} onto the set Q by minimizing the distance function $KL(p^{n+1}, q)$. This suggests the possibility of using alternating minimization with respect to more general distance functions. We shall focus on Bregman distances.

14.5.1 Bregman Distances

Let $f : \mathbb{R}^N \rightarrow \mathbb{R}$ be a Bregman function [12, 59, 19], and so $f(x)$ is convex on its domain and differentiable in the interior of its domain. Then, for x in the domain and z in the interior, we define the Bregman distance $D_f(x, z)$ by

$$D_f(x, z) = f(x) - f(z) - \langle \nabla f(z), x - z \rangle. \tag{14.40}$$

For example, the KL distance is a Bregman distance with associated Bregman function

$$f(x) = \sum_{j=1}^{J} x_j \log x_j \; - x_j. \tag{14.41}$$

Suppose now that $f(x)$ is a Bregman function and P and Q are closed convex subsets of the interior of the domain of $f(x)$. Let p^{n+1} minimize $D_f(p, q^n)$ over all $p \in P$. It follows then that

$$\langle \nabla f(p^{n+1}) - \nabla f(q^n), p - p^{n+1} \rangle \geq 0, \tag{14.42}$$

for all $p \in P$. Since

$$D_f(p, q^n) - D_f(p^{n+1}, q^n) =$$

$$D_f(p, p^{n+1}) + \langle \nabla f(p^{n+1}) - \nabla f(q^n), p - p^{n+1} \rangle, \tag{14.43}$$

it follows that the three-point property holds, with

$$\Theta(p, q) = D_f(p, q), \tag{14.44}$$

and

$$\Delta(p, \hat{p}) = D_f(p, \tilde{p}). \tag{14.45}$$

To get the four-point property we need to restrict D_f somewhat; we assume from now on that $D_f(p, q)$ is jointly convex, that is, it is convex in the combined vector variable (p, q) (see [8]). Now we can invoke a lemma due to Eggermont and LaRiccia [78].

14.5.2 The Eggermont-LaRiccia Lemma

Lemma 14.1 *Suppose that the Bregman distance $D_f(p, q)$ is jointly convex. Then it has the four-point property.*

Proof: By joint convexity we have

$$D_f(p, q) - D_f(p^n, q^n) \geq$$

$$\langle \nabla_1 D_f(p^n, q^n), p - p^n \rangle + \langle \nabla_2 D_f(p^n, q^n), q - q^n \rangle,$$

where ∇_1 denotes the gradient with respect to the first vector variable. Since q^n minimizes $D_f(p^n, q)$ over all $q \in Q$, we have

$$\langle \nabla_2 D_f(p^n, q^n), q - q^n \rangle \geq 0,$$

for all q. Also,

$$\langle \nabla_1(p^n, q^n), p - p^n \rangle = \langle \nabla f(p^n) - \nabla f(q^n), p - p^n \rangle.$$

It follows that

$$D_f(p, q^n) - D_f(p, p^n) = D_f(p^n, q^n) + \langle \nabla_1(p^n, q^n), p - p^n \rangle$$

$$\leq D_f(p,q) - \langle \nabla_2 D_f(p^n,q^n), q-q^n\rangle \leq D_f(p,q).$$

Therefore, we have

$$D_f(p,p^n) + D_f(p,q) \geq D_f(p,q^n).$$

This is the four-point property. ∎

We now know that the alternating minimization method works for any Bregman distance that is jointly convex. This includes the Euclidean and the KL distances.

14.6 Minimizing a Proximity Function

We present now an example of alternating Bregman distance minimization taken from [37]. The problem is the *convex feasibility problem* (CFP), to find a member of the intersection $C \subseteq R^J$ of finitely many closed convex sets C_i, $i = 1, ..., I$, or, failing that, to minimize the proximity function

$$F(x) = \sum_{i=1}^{I} D_i(\overleftarrow{P}_i x, x), \tag{14.46}$$

where f_i are Bregman functions for which D_i, the associated Bregman distance, is jointly convex, and $\overleftarrow{P}_i x$ are the *left* Bregman projection of x onto the set C_i, that is, $\overleftarrow{P}_i x \in C_i$ and $D_i(\overleftarrow{P}_i x, x) \leq D_i(z, x)$, for all $z \in C_i$. Because each D_i is jointly convex, the function $F(x)$ is convex.

The problem can be formulated as an alternating minimization, where $P \subseteq R^{IJ}$ is the product set $P = C_1 \times C_2 \times ... \times C_I$. A typical member of P has the form $p = (c^1, c^2, ..., c^I)$, where $c^i \in C_i$, and $Q \subseteq R^{IJ}$ is the *diagonal* subset, meaning that the elements of Q are the I-fold product of a single x; that is $Q = \{d(x) = (x, x, ..., x) \in R^{IJ}\}$. We then take

$$\Theta(p,q) = \sum_{i=1}^{I} D_i(c^i, x), \tag{14.47}$$

and $\Delta(p, \tilde{p}) = \Theta(p, \tilde{p})$.

In [53] a similar iterative algorithm was developed for solving the CFP, using the same sets P and Q, but using alternating projection, rather than alternating minimization. Now it is not necessary that the Bregman distances be jointly convex. Each iteration of their algorithm involves two steps:

- 1. minimize $\sum_{i=1}^{I} D_i(c^i, x^n)$ over $c^i \in C_i$, obtaining $c^i = \overleftarrow{P}_i x^n$, and then
- 2. minimize $\sum_{i=1}^{I} D_i(x, \overleftarrow{P}_i x^n)$.

Because this method is an alternating projection approach, it converges only when the CFP has a solution, whereas the previous alternating minimization method minimizes $F(x)$, even when the CFP has no solution.

14.6.1 Right and Left Projections

Because Bregman distances D_f are not generally symmetric, we can speak of *right* and *left* Bregman projections onto a closed convex set. For any allowable vector x, the *left* Bregman projection of x onto C, if it exists, is the vector $\overleftarrow{P}_C x \in C$ satisfying the inequality $D_f(\overleftarrow{P}_C x, x) \leq D_f(c, x)$, for all $c \in C$. Similarly, the *right* Bregman projection is the vector $\overrightarrow{P}_C x \in C$ satisfying the inequality $D_f(x, \overrightarrow{P}_C x) \leq D_f(x, c)$, for any $c \in C$.

The alternating minimization approach described above to minimize the proximity function

$$F(x) = \sum_{i=1}^{I} D_i(\overleftarrow{P}_i x, x) \tag{14.48}$$

can be viewed as an alternating projection method, but employing both right and left Bregman projections.

Consider the problem of finding a member of the intersection of two closed convex sets C and D. We could proceed as follows: having found x^n, minimize $D_f(x^n, d)$ over all $d \in D$, obtaining $d = \overrightarrow{P}_D x^n$, and then minimize $D_f(c, \overrightarrow{P}_D x^n)$ over all $c \in C$, obtaining $c = x^{n+1} = \overleftarrow{P}_C \overrightarrow{P}_D x^n$. The objective of this algorithm is to minimize $D_f(c, d)$ over all $c \in C$ and $d \in D$; such a minimum may not exist, of course.

In [9] the authors note that the alternating minimization algorithm of [37] involves right and left Bregman projections, which suggests to them iterative methods involving a wider class of operators that they call “Bregman retractions”.

14.7 More Proximity Function Minimization

Proximity function minimization and right and left Bregman projections play a role in a variety of iterative algorithms. We survey several of them in this section.

14.7.1 Cimmino’s Algorithm

Our objective here is to find an exact or approximate solution of the system of I linear equations in J unknowns, written $Ax = b$. For each i let

$$H_i = \{z | (Az)_i = b_i\}, \tag{14.49}$$

and $P_{H_i}x$ be the orthogonal projection of x onto H_i. Then

$$(P_{H_i}x)_j = x_j + \alpha_i A_{ij}(b_i - (Ax)_i), \tag{14.50}$$

where

$$(\alpha_i)^{-1} = \sum_{j=1}^{J} A_{ij}^2. \tag{14.51}$$

Let

$$F(x) = \sum_{i=1}^{I} \|P_{H_i}x - x\|_2^2. \tag{14.52}$$

Using alternating minimization on this proximity function gives Cimmino's algorithm, with the iterative step

$$x_j^{n+1} = x_j^n + \frac{1}{I}\sum_{i=1}^{I} \alpha_i A_{ij}(b_i - (Ax^n)_i). \tag{14.53}$$

14.7.2 Simultaneous Projection for Convex Feasibility

Now we let C_i be any closed convex subsets of $\mathbb{R}^J$ and define $F(x)$ as in the previous section. Again, we apply alternating minimization. The iterative step of the resulting algorithm is

$$x^{n+1} = \frac{1}{I}\sum_{i=1}^{I} P_{C_i}x^n. \tag{14.54}$$

The objective here is to minimize $F(x)$, if there is a minimum.

14.7.3 The Bauschke-Combettes-Noll Problem

In [10] Bauschke, Combettes and Noll consider the following problem: minimize the function

$$\Theta(p,q) = \Lambda(p,q) = \phi(p) + \psi(q) + D_f(p,q), \tag{14.55}$$

where ϕ and ψ are convex on R^J, $D = D_f$ is a Bregman distance, and $P = Q$ is the interior of the domain of f. They assume that

$$b = \inf_{(p,q)} \Lambda(p,q) > -\infty, \tag{14.56}$$

and seek a sequence $\{(p^n, q^n)\}$ such that $\{\Lambda(p^n, q^n)\}$ converges to b. The sequence is obtained by the AM method, as in our previous discussion. They prove that, if the Bregman distance is jointly convex, then $\{\Lambda(p^n, q^n)\} \downarrow b$. In this subsection we obtain this result by showing that $\Lambda(p, q)$ has the five-point property whenever $D = D_f$ is jointly convex. Our proof is loosely based on the proof of the Eggermont-LaRiccia lemma.

The five-point property for $\Lambda(p, q)$ is

$$\Lambda(p, q^{n-1}) - \Lambda(p^n, q^{n-1}) \geq \Lambda(p, q^n) - \Lambda(p, q). \tag{14.57}$$

A simple calculation shows that the inequality in (14.57) is equivalent to

$$\Lambda(p, q) - \Lambda(p^n, q^n) \geq$$

$$D(p, q^n) + D(p^n, q^{n-1}) - D(p, q^{n-1}) - D(p^n, q^n). \tag{14.58}$$

By the joint convexity of $D(p, q)$ and the convexity of ϕ and ψ we have

$$\Lambda(p, q) - \Lambda(p^n, q^n) \geq$$

$$\langle \nabla_p \Lambda(p^n, q^n), p - p^n \rangle + \langle \nabla_q \Lambda(p^n, q^n), q - q^n \rangle, \tag{14.59}$$

where $\nabla_p \Lambda(p^n, q^n)$ denotes the gradient of $\Lambda(p, q)$, with respect to p, evaluated at (p^n, q^n).

Since q^n minimizes $\Lambda(p^n, q)$, it follows that

$$\langle \nabla_q \Lambda(p^n, q^n), q - q^n \rangle = 0, \tag{14.60}$$

for all q. Therefore,

$$\Lambda(p, q) - \Lambda(p^n, q^n) \geq \langle \nabla_p \Lambda(p^n, q^n), p - p^n \rangle. \tag{14.61}$$

We have

$$\langle \nabla_p \Lambda(p^n, q^n), p - p^n \rangle =$$

$$\langle \nabla f(p^n) - \nabla f(q^n), p - p^n \rangle + \langle \nabla \phi(p^n), p - p^n \rangle. \tag{14.62}$$

Since p^n minimizes $\Lambda(p, q^{n-1})$, we have

$$\nabla_p \Lambda(p^n, q^{n-1}) = 0, \tag{14.63}$$

or

$$\nabla \phi(p^n) = \nabla f(q^{n-1}) - \nabla f(p^n), \tag{14.64}$$

so that

$$\langle \nabla_p \Lambda(p^n, q^n), p - p^n \rangle = \langle \nabla f(q^{n-1}) - \nabla f(q^n), p - p^n \rangle \tag{14.65}$$

$$= D(p, q^n) + D(p^n, q^{n-1}) - D(p, q^{n-1}) - D(p^n, q^n). \tag{14.66}$$

Using (14.61) we obtain the inequality in (14.58). This shows that $\Lambda(p, q)$ has the five-point property whenever the Bregman distance $D = D_f$ is jointly convex.

From our previous discussion of AM, we conclude that the sequence $\{\Lambda(p^n, q^n)\}$ converges to b; this is Corollary 4.3 of [10].

In [51] it was shown that, in certain cases, the expectation maximization maximum likelihood (EM) method involves alternating minimization of a function of the form $\Lambda(p, q)$.

If $\psi = 0$, then $\{\Lambda(p^n, q^n)\}$ converges to b, even without the assumption that the distance D_f is jointly convex. In such cases, $\Lambda(p, q)$ has the form of the objective function in proximal minimization and therefore the problem falls into the SUMMA class.

14.8 AM as SUMMA

We show now that the SUMMA class of sequential unconstrained minimization methods includes all the AM methods for which the five-point property holds.

14.8.1 Reformulating AM as SUMMA

For each p in the set P, define $q(p)$ in Q as a member of Q for which $\Theta(p, q(p)) \leq \Theta(p, q)$, for all $q \in Q$. Let $f(p) = \Theta(p, q(p))$.

At the nth step of AM we minimize

$$G_n(p) = \Theta(p, q^{n-1}) = \Theta(p, q(p)) + \Big(\Theta(p, q^{n-1}) - \Theta(p, q(p))\Big) \quad (14.67)$$

to get p^n. With

$$g_n(p) = \Big(\Theta(p, q^{n-1}) - \Theta(p, q(p))\Big) \geq 0, \quad (14.68)$$

we can write

$$G_n(p) = f(p) + g_n(p). \quad (14.69)$$

According to the five-point property, we have

$$G_n(p) - G_n(p^n) \geq \Theta(p, q^n) - \Theta(p, q(p)) = g_{n+1}(p). \quad (14.70)$$

It follows that AM is a member of the SUMMA class.

14.9 SMART and EMML as SUMMA

We have seen that both the SMART and the EMML can be obtained as AM algorithms for which the five-point property holds. Consequently, both SMART and EMML are particular cases of SUMMA.

14.9.1 The SMART as SUMMA

In the case of SMART

$$\Theta(p,q) = KL(t(x), r(z)), \tag{14.71}$$

and

$$f(p) = \Theta(p, q(p)) = KL(t(x), r(x)) = KL(Px, y), \tag{14.72}$$

which is the function of x we seek to minimize over $x \in \mathcal{X}$.

14.9.2 The EMML as SUMMA

In the case of EMML

$$\Theta(p,q) = KL(r(x), t(z)), \tag{14.73}$$

and

$$f(p) = \Theta(p, q(p)) = KL(r(x), t(x')), \tag{14.74}$$

which is not $KL(b, Ax)$. In order to obtain the EMML from an AM formulation having the five-point property, and therefore to show that EMML is in the SUMMA class, we need to view the problem as minimizing not $KL(b, Ax)$ but $f(x) = KL(r(x), t(x'))$. The minima are the same, however, as are the minimizers.

For the EMML we get $x^n = (x^{n-1})'$ by minimizing

$$G_n(x) = KL(r(x), t((x^{n-1})')) = f(x) + g_n(x), \tag{14.75}$$

where

$$g_n(x) = KL(r(x), t((x^{n-1})')) - KL(r(x), t(x')). \tag{14.76}$$

We need to show that

$$G_n(x) - G_n(x^n) \geq g_{n+1}(x). \tag{14.77}$$

From the Pythagorean identities for EMML we have

$$G_n(x) - G_n(x^n) = KL(r(x), r(x^n)), \tag{14.78}$$

and

$$g_{n+1}(x) = KL(x', (x^n)') \leq KL(r(x), r(x^n)), \tag{14.79}$$

which shows the EMML to be a member of the SUMMA class.

Consequently, the sequence $\{KL(b, Ax^n)\}$ converges to the infimum of the function $KL(b, Ax)$ over all $x \in \mathcal{X}$. The infimum is always attained at some $x \geq 0$ in the closure of $\mathcal{X}$ and it can be shown that the sequence $\{x^n\}$ converges to a minimizer of $KL(b, Ax)$ over x in the closure of $\mathcal{X}$ ([27, 29]).

14.10 Conclusion

It was shown previously in [43] that the SUMMA class includes a wide variety of optimization algorithms, including the barrier-function methods, the proximal minimization algorithm of Censor and Zenios [58, 59], the entropic proximal method of Teboulle [127], and the simultaneous multiplicative algebraic reconstruction technique (SMART)[70, 118, 68, 27, 28]. With some reformulation, it also contains the penalty-function methods. We have now shown that the alternating minimization methods of [69] are included in the SUMMA class whenever the five-point property holds. As a consequence, we learn that the EMML algorithm for Poisson mixtures [119, 101, 130, 102, 27, 28] is also a member of the SUMMA class.

Chapter 15

Appendix: Bregman-Legendre Functions

15.1 Chapter Summary

In [7] Bauschke and Borwein show convincingly that the Bregman-Legendre functions provide the proper context for the discussion of Bregman projections onto closed convex sets. The summary here follows closely the discussion given in [7].

15.2 Essential Smoothness and Essential Strict Convexity

Following [116] we say that a closed proper convex function $f : \mathbb{R}^J \to \mathbb{R}$ is *essentially smooth* if int$D$ is not empty, f is differentiable on int$D$ and $x^n \in \text{int}D$, with $x^n \to x \in \text{bd}D$, implies that $||\nabla f(x^n)||_2 \to +\infty$. Here int$D$ and bd$D$ denote the interior and boundary of the set D. A closed proper convex function f is *essentially strictly convex* if f is strictly convex on every convex subset of dom ∂f.

The closed proper convex function f is essentially smooth if and only if the subdifferential $\partial f(x)$ is empty for $x \in \text{bd}D$ and is $\{\nabla f(x)\}$ for $x \in \text{int}D$ (so f is differentiable on int$D$) if and only if the function f^* is essentially strictly convex.

Definition 15.1 *A closed proper convex function f is said to be a* Legendre function *if it is both essentially smooth and essentialy strictly convex.*

So f is Legendre if and only if its conjugate function is Legendre, in which case the gradient operator ∇f is a topological isomorphism with ∇f^* as its inverse. The gradient operator ∇f maps int dom f onto int dom f^*. If int dom $f^* = \mathbb{R}^J$ then the range of ∇f is $\mathbb{R}^J$ and the equation $\nabla f(x) = y$ can be solved for every $y \in \mathbb{R}^J$. In order for int dom $f^* = \mathbb{R}^J$ it is necessary and sufficient that the Legendre function f be *super-coercive*, that is,

$$\lim_{||x||_2 \to +\infty} \frac{f(x)}{||x||_2} = +\infty. \tag{15.1}$$

If the effective domain of f is bounded, then f is super-coercive and its gradient operator is a mapping onto the space $\mathbb{R}^J$.

15.3 Bregman Projections onto Closed Convex Sets

Let f be a closed proper convex function that is differentiable on the nonempty set int$D$. The corresponding *Bregman distance* $D_f(x, z)$ is defined for $x \in \mathbb{R}^J$ and $z \in \text{int}D$ by

$$D_f(x, z) = f(x) - f(z) - \langle \nabla f(z), x - z \rangle. \tag{15.2}$$

Note that $D_f(x, z) \geq 0$ always and that $D_f(x, z) = +\infty$ is possible. If f is essentially strictly convex then $D_f(x, z) = 0$ implies that $x = z$.

Let K be a nonempty closed convex set in $\mathbb{R}^J$, with $K \cap \text{int}D \neq \emptyset$. Pick $z \in \text{int}D$. The *Bregman projection* of z onto K, with respect to f, is

$$P_K^f(z) = \text{argmin}_{x \in K \cap D} D_f(x, z). \tag{15.3}$$

If f is essentially strictly convex, then $P_K^f(z)$ exists. If f is strictly convex on D then $P_K^f(z)$ is unique. If f is Legendre, then $P_K^f(z)$ is uniquely defined and is in int$D$; this last condition is sometimes called *zone consistency*.

Example: Let $J = 2$ and $f(x)$ be the function that is equal to one-half the norm squared on D, the nonnegative quadrant, $+\infty$ elsewhere. Let K be the set $K = \{(x_1, x_2) | x_1 + x_2 = 1\}$. The Bregman projection of $(2, 1)$ onto K is $(1, 0)$, which is not in int$D$. The function f is not essentially smooth, although it is essentially strictly convex. Its conjugate is the function f^* that is equal to one-half the norm squared on D and equal to zero elsewhere; it is essentially smooth, but not essentially strictly convex.

If f is Legendre, then $P_K^f(z)$ is the unique member of $K \cap \text{int}D$ satisfying the inequality

$$\langle \nabla f(P_K^f(z)) - \nabla f(z), P_K^f(z) - c \rangle \geq 0, \tag{15.4}$$

for all $c \in K$. From this we obtain the *Bregman Inequality*:

$$D_f(c,z) \geq D_f(c, P_K^f(z)) + D_f(P_K^f(z), z), \tag{15.5}$$

for all $c \in K$.

15.4 Bregman-Legendre Functions

Following Bauschke and Borwein [7], we say that a Legendre function $f : \mathbb{R}^J \rightarrow \mathbb{R}$ is a *Bregman-Legendre* function if the following properties hold:

B1: for x in D and any $a > 0$ the set $\{z|D_f(x,z) \leq a\}$ is bounded.
B2: if x is in D but not in int$D$, for each positive integer n, y^n is in int$D$ with $y^n \rightarrow y \in$ bd$D$ and if $\{D_f(x, y^n)\}$ remains bounded, then $D_f(y, y^n) \rightarrow 0$, so that $y \in D$.
B3: if x^n and y^n are in int$D$, with $x^n \rightarrow x$ and $y^n \rightarrow y$, where x and y are in D but not in int$D$, and if $D_f(x^n, y^n) \rightarrow 0$ then $x = y$.

Bauschke and Borwein then prove that Bregman's SGP method converges to a member of K provided that one of the following holds: 1) f is Bregman-Legendre; 2) $K \cap \text{int}D \neq \emptyset$ and dom f^* is open; or 3) dom f and dom f^* are both open.

The Bregman functions form a class closely related to the Bregman-Legendre functions. For details see [19].

15.5 Useful Results about Bregman-Legendre Functions

The following results are proved in somewhat more generality in [7].
R1: If $y^n \in$ int dom f and $y^n \rightarrow y \in$ int dom f, then $D_f(y, y^n) \rightarrow 0$.
R2: If x and $y^n \in$ int dom f and $y^n \rightarrow y \in$ bd dom f, then $D_f(x, y^n) \rightarrow +\infty$.
R3: If $x^n \in D$, $x^n \rightarrow x \in D$, $y^n \in$ int D, $y^n \rightarrow y \in D$, $\{x, y\} \cap$ int $D \neq \emptyset$ and $D_f(x^n, y^n) \rightarrow 0$, then $x = y$ and $y \in$ int D.
R4: If x and y are in D, but are not in int D, $y^n \in$ int D, $y^n \rightarrow y$ and $D_f(x, y^n) \rightarrow 0$, then $x = y$.
As a consequence of these results we have the following.
R5: If $\{D_f(x, y^n)\} \rightarrow 0$, for $y^n \in$ int D and $x \in \mathbb{R}^J$, then $\{y^n\} \rightarrow x$.

Proof of R5: Since $\{D_f(x, y^n)\}$ is eventually finite, we have $x \in D$. By Property B1 above it follows that the sequence $\{y^n\}$ is bounded; without loss of generality, we assume that $\{y^n\} \rightarrow y$, for some $y \in \overline{D}$. If x is in int

D, then, by result R2 above, we know that y is also in int D. Applying result R3, with $x^n = x$, for all n, we conclude that $x = y$. If, on the other hand, x is in D, but not in int D, then y is in D, by result R2. There are two cases to consider: 1) y is in int D; 2) y is not in int D. In case 1) we have $D_f(x, y^n) \to D_f(x, y) = 0$, from which it follows that $x = y$. In case 2) we apply result R4 to conclude that $x = y$. ∎

Chapter 16

Appendix: Convex Functions

16.1 Gradient Operators

Definition 16.1 *An operator T on $\mathbb{R}^J$ is called L-*Lipschitz continuous*, with respect to a given norm on $\mathbb{R}^J$, if, for every x and y in $\mathbb{R}^J$, we have*

$$\|Tx - Ty\| \leq L\|x - y\|. \tag{16.1}$$

Clearly, if an operator T is L-Lipschitz continuous, then the operator $N = \frac{1}{L}T$ is non-expansive. We have the following theorem concerning the gradient of a differentiable convex function $h(x)$.

Theorem 16.1 *Let $h(x)$ be convex and differentiable and its derivative, $\nabla h(x)$, non-expansive in the two-norm. Then $\nabla h(x)$ is firmly non-expansive; that is,*

$$\langle \nabla h(x) - \nabla h(y), x - y \rangle \geq ||\nabla h(x) - \nabla h(y)||_2^2. \tag{16.2}$$

Suppose that $g(x) : \mathbb{R}^J \to \mathbb{R}$ is convex and the function $Tx = \nabla g(x)$ is L-Lipschitz. Let $h(x) = \frac{1}{L}g(x)$, so that $Nx = \nabla h(x)$ is a non-expansive operator. Then, according to Theorem 16.1, the operator $\nabla h(x) = \frac{1}{L}\nabla g(x)$ is firmly non-expansive.

The proof of Theorem 16.1 is not trivial. In [86] Golshtein and Tretyakov prove the following theorem, from which Theorem 16.1 follows immediately.

Theorem 16.2 *Let $g : \mathbb{R}^J \to \mathbb{R}$ be convex and differentiable. The following are equivalent:*

- **1)**

$$||\nabla g(x) - \nabla g(y)||_2 \leq ||x - y||_2; \tag{16.3}$$

- **2)**

$$g(x) \geq g(y) + \langle \nabla g(y), x - y \rangle + \frac{1}{2}||\nabla g(x) - \nabla g(y)||_2^2; \tag{16.4}$$

and

- **3)**

$$\langle \nabla g(x) - \nabla g(y), x - y \rangle \geq ||\nabla g(x) - \nabla g(y)||_2^2. \tag{16.5}$$

Proof: The only difficult step in the proof is showing that Inequality (16.3) implies Inequality (16.4). To prove this part, let $x(t) = (1-t)y + tx$, for $0 \leq t \leq 1$. Then

$$g'(x(t)) = \langle \nabla g(x(t)), x - y \rangle, \tag{16.6}$$

so that

$$\int_0^1 \langle \nabla g(x(t)) - \nabla g(y), x - y \rangle dt = g(x) - g(y) - \langle \nabla g(y), x - y \rangle. \tag{16.7}$$

Therefore,

$$g(x) - g(y) - \langle \nabla g(y), x - y \rangle \leq$$

$$\int_0^1 ||\nabla g(x(t)) - \nabla g(y)||_2 ||x(t) - y||_2 dt \tag{16.8}$$

$$\leq \int_0^1 ||x(t) - y||_2^2 dt = \int_0^1 ||t(x - y)||_2^2 dt = \frac{1}{2}||x - y||_2^2, \tag{16.9}$$

according to Inequality (16.3). Therefore,

$$g(x) \leq g(y) + \langle \nabla g(y), x - y \rangle + \frac{1}{2}||x - y||_2^2. \tag{16.10}$$

Now let $x = y - \nabla g(y)$, so that

$$g(y - \nabla g(y)) \leq g(y) + \langle \nabla g(y), \nabla g(y) \rangle + \frac{1}{2}||\nabla g(y)||_2^2. \tag{16.11}$$

Consequently,

$$g(y - \nabla g(y)) \leq g(y) - \frac{1}{2}||\nabla g(y)||_2^2. \tag{16.12}$$

Therefore,

$$\inf g(x) \leq g(y) - \frac{1}{2}||\nabla g(y)||_2^2, \tag{16.13}$$

or

$$g(y) \geq \inf g(x) + \frac{1}{2}||\nabla g(y)||_2^2. \tag{16.14}$$

Now fix y and define the function $h(x)$ by

$$h(x) = g(x) - g(y) - \langle \nabla g(y), x - y \rangle. \tag{16.15}$$

Then $h(x)$ is convex, differentiable, and non-negative,

$$\nabla h(x) = \nabla g(x) - \nabla g(y), \tag{16.16}$$

and $h(y) = 0$, so that $h(x)$ attains its minimum at $x = y$. Applying Inequality (16.14) to the function $h(x)$, with z in the role of x and x in the role of y, we find that

$$\inf h(z) = 0 \leq h(x) - \frac{1}{2}||\nabla h(x)||_2^2. \tag{16.17}$$

From the definition of $h(x)$, it follows that

$$0 \leq g(x) - g(y) - \langle \nabla g(y), x - y \rangle - \frac{1}{2}||\nabla g(x) - \nabla g(y)||_2^2. \tag{16.18}$$

This completes the proof of the implication. ∎

Bibliography

1. Anderson, A. and Kak, A. (1984) “Simultaneous algebraic reconstruction technique (SART): a superior implementation of the ART algorithm.” *Ultrasonic Imaging* **6**, pp. 81–94.

2. Aubin, J.-P., (1993) *Optima and Equilibria: An Introduction to Nonlinear Analysis*, Springer-Verlag.

3. Auslander, A., and Teboulle, M. (2006) “Interior gradient and proximal methods for convex and conic optimization.” *SIAM Journal on Optimization*, **16(3)**, pp. 697–725.

4. Baillon, J.-B., Bruck, R.E., and Reich, S. (1978) “On the asymptotic behavior of nonexpansive mappings and semigroups in Banach spaces.” *Houston Journal of Mathematics*, **4**, pp. 1–9.

5. Bauschke, H., and Borwein, J. (1993) “On the convergence of von Neumann’s alternating projection algorithm for two sets.” *Set-Valued Analysis* **1**, pp. 185–212.

6. Bauschke, H., and Borwein, J. (1996) “On projection algorithms for solving convex feasibility problems.” *SIAM Review*, **38 (3)**, pp. 367–426.

7. Bauschke, H., and Borwein, J. (1997) “Legendre functions and the method of random Bregman projections.” *Journal of Convex Analysis*, **4**, pp. 27–67.

8. Bauschke, H., and Borwein, J. (2001) “Joint and separate convexity of the Bregman distance.” in *Inherently Parallel Algorithms in Feasibility and Optimization and their Applications*, edited by D. Butnariu, Y. Censor and S. Reich, pp. 23–36, Studies in Computational Mathematics 8. Amsterdam: Elsevier Publ.

9. Bauschke, H., and Combettes, P. (2003) “Iterating Bregman retractions.” *SIAM Journal on Optimization*, **13**, pp. 1159–1173.

10. Bauschke, H., Combettes, P., and Noll, D. (2006) "Joint minimization with alternating Bregman proximity operators." *Pacific Journal of Optimization*, **2**, pp. 401–424.

11. Brauer, A. (1946) "Characteristic roots of a matrix." *Duke Mathematics Journal*, **13**, pp. 387–395.

12. Bregman, L.M. (1967) "The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming." *USSR Computational Mathematics and Mathematical Physics* **7**: pp. 200–217.

13. Browne, E. (1930) "The characteristic roots of a matrix." *Bulletin of the American Mathematical Society*, **36**, pp. 705–710.

14. Browne, J. and A. DePierro, A. (1996) "A row-action alternative to the EM algorithm for maximizing likelihoods in emission tomography." *IEEE Trans. Med. Imag.* **15**, pp. 687–699.

15. Bruck, R.E., and Reich, S. (1977) "Nonexpansive projections and resolvents of accretive operators in Banach spaces." *Houston Journal of Mathematics*, **3**, pp. 459–470.

16. Burg, J. (1967) "Maximum entropy spectral analysis." *paper presented at the 37th Annual SEG meeting, Oklahoma City, OK.*

17. Burg, J. (1972) "The relationship between maximum entropy spectra and maximum likelihood spectra." *Geophysics* **37**, pp. 375–376.

18. Burg, J. (1975) *Maximum Entropy Spectral Analysis*, Ph.D. dissertation, Stanford University.

19. Butnariu, D., Byrne, C., and Censor, Y. (2003) "Redundant axioms in the definition of Bregman functions." *Journal of Convex Analysis*, **10**, pp. 245–254.

20. Byrne, C. and Fitzgerald, R. (1979) "A unifying model for spectrum estimation." In *Proceedings of the RADC Workshop on Spectrum Estimation*, Griffiss AFB, Rome, NY, October.

21. Byrne, C. and Fitzgerald, R. (1982) "Reconstruction from partial information, with applications to tomography." *SIAM J. Applied Math.* **42(4)**, pp. 933–940.

22. Byrne, C., Fitzgerald, R., Fiddy, M., Hall, T., and Darling, A. (1983) "Image restoration and resolution enhancement." *J. Opt. Soc. Amer.* **73**, pp. 1481–1487.

23. Byrne, C. and Fitzgerald, R. (1984) "Spectral estimators that extend the maximum entropy and maximum likelihood methods." *SIAM J. Applied Math.* **44(2)**, pp. 425–442.

24. Byrne, C., Levine, B.M., and Dainty, J.C. (1984) "Stable estimation of the probability density function of intensity from photon frequency counts." *JOSA Communications* **1(11)**, pp. 1132–1135.

25. Byrne, C. and Fiddy, M. (1987) "Estimation of continuous object distributions from Fourier magnitude measurements." *JOSA A* **4**, pp. 412–417.

26. Byrne, C. and Fiddy, M. (1988) "Images as power spectra; reconstruction as Wiener filter approximation." *Inverse Problems* **4**, pp. 399–409.

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

28. Byrne, C. (1995) "Erratum and addendum to 'Iterative image reconstruction algorithms based on cross-entropy minimization'." *IEEE Transactions on Image Processing* **IP-4**, pp. 225–226.

29. Byrne, C. (1996) "Iterative reconstruction algorithms based on cross-entropy minimization." in *Image Models (and their Speech Model Cousins)*, S.E. Levinson and L. Shepp, editors, IMA Volumes in Mathematics and its Applications, Volume 80, pp. 1–11. New York: Springer-Verlag.

30. Byrne, C. (1996) "Block-iterative methods for image reconstruction from projections." *IEEE Transactions on Image Processing* **IP-5**, pp. 792–794.

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

32. Byrne, C. (1998) "Accelerating the EMML algorithm and related iterative algorithms by rescaled block-iterative (RBI) methods." *IEEE Transactions on Image Processing* **IP-7**, pp. 100–109.

33. Byrne, C. (1998) "Iterative algorithms for deblurring and deconvolution with constraints." *Inverse Problems*, **14**, pp. 1455–1467.

34. Byrne, C. (2000) "Block-iterative interior point optimization methods for image reconstruction from limited data." *Inverse Problems* **16**, pp. 1405–1419.

35. Byrne, C. (2001) "Bregman-Legendre multi-distance projection algorithms for convex feasibility and optimization." in *Inherently Parallel Algorithms in Feasibility and Optimization and their Applications*, edited by D. Butnariu, Y. Censor and S. Reich, pp. 87-100, Studies in Computational Mathematics 8. Amsterdam: Elsevier Publ.

36. Byrne, C. (2001) "Likelihood maximization for list-mode emission tomographic image reconstruction."*IEEE Transactions on Medical Imaging* **20(10)**, pp. 1084–1092.

37. Byrne, C., and Censor, Y. (2001) "Proximity function minimization using multiple Bregman projections, with applications to split feasibility and Kullback-Leibler distance minimization." *Annals of Operations Research*, **105**, pp. 77–98.

38. Byrne, C. (2002) "Iterative oblique projection onto convex sets and the split feasibility problem."*Inverse Problems* **18**, pp. 441–453.

39. Byrne, C. (2004) "A unified treatment of some iterative algorithms in signal processing and image reconstruction."*Inverse Problems* **20**, pp. 103–120.

40. Byrne, C. (2005) "Choosing parameters in block-iterative or ordered-subset reconstruction algorithms." *IEEE Transactions on Image Processing*, **14 (3)**, pp. 321–327.

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

42. Byrne, C. (2007) *Applied Iterative Methods*, AK Peters, Publ., Wellesley, MA.

43. Byrne, C. (2008) "Sequential unconstrained minimization algorithms for constrained optimization." *Inverse Problems*, **24(1)**, article no. 015013.

44. Byrne, C. (2009) "Block-iterative algorithms." *International Transactions in Operations Research*, **16(4)**, pp. 427–463.

45. Byrne, C. (2009) "Bounds on the largest singular value of a matrix and the convergence of simultaneous and block-iterative algorithms for sparse linear systems." *International Transactions in Operations Research*, **16(4)**, pp. 465–479.

46. Byrne, C. (2009) *Applied and Computational Linear Algebra: A First Course*, available as a pdf file at my web site.

47. Byrne, C. (2011) *A First Course in Optimization*, available as a pdf file at my web site.

48. Byrne, C. "Alternating Minimization and Alternating Projection." *preprint*.

49. Byrne, C. (2012) "Alternating and sequential unconstrained minimization algorithms." submitted.

50. Byrne, C., and Eggermont, P. (2011) "EM Algorithms." in *Handbook of Mathematical Methods in Imaging*, Otmar Scherzer, ed., Springer-Science.

51. Byrne, C., and Eggermont, P. (2012) "EM algorithms." in preparation.

52. Candès, E., Wakin, M., and Boyd, S. (2007) "Enhancing sparsity by reweighted l_1 minimization." preprint available at http://www.acm.caltech.edu/ emmanuel/publications.html .

53. Censor, Y. and Elfving, T. (1994) "A multi-projection algorithm using Bregman projections in a product space." *Numerical Algorithms*, **8**, pp. 221–239.

54. Censor, Y., Bortfeld, T., Martin, B., and Trofimov, A. "A unified approach for inversion problems in intensity-modulated radiation therapy." *Physics in Medicine and Biology* 51 (2006), 2353-2365.

55. Censor, Y., Elfving, T., Kopf, N., and Bortfeld, T. (2005) "The multiple-sets split feasibility problem and its application for inverse problems." *Inverse Problems*, **21** , pp. 2071-2084.

56. Censor, Y. and Segman, J. (1987) "On block-iterative maximization." *J. of Information and Optimization Sciences* **8**, pp. 275–291.

57. Censor, Y., and Reich, S. (1996) "Iterations of paracontractions and firmly nonexpansive operators with applications to feasibility and optimization." *Optimization*, **37**, pp. 323–339.

58. Censor, Y., and Zenios, S.A. (1992) "Proximal minimization algorithm with D-functions." *Journal of Optimization Theory and Applications*, **73(3)**, pp. 451–464.

59. Censor, Y. and Zenios, S.A. (1997) *Parallel Optimization: Theory, Algorithms and Applications*. New York: Oxford University Press.

60. Censor, Y., Gordon, D., and Gordon, R. (2001) "Component averaging: an efficient iterative parallel algorithm for large and sparse unstructured problems." *Parallel Computing*, **27**, pp. 777–808.

61. Censor, Y., Gordon, D., and Gordon, R. (2001) "BICAV: A block-iterative, parallel algorithm for sparse systems with pixel-related weighting." *IEEE Transactions on Medical Imaging*, **20**, pp. 1050–1060.

62. Censor, Y., Eggermont, P.P.B., and Gordon, D. (1983) "Strong underrelaxation in Kaczmarz's method for inconsistent systems." *Numerische Mathematik* **41**, pp. 83–92.

63. Cheney, W., and Goldstein, A. (1959) "Proximity maps for convex sets." *Proc. Amer. Math. Soc.*, **10**, pp. 448–450.

64. Cimmino, G. (1938) "Calcolo approssimato per soluzioni dei sistemi di equazioni lineari." *La Ricerca Scientifica XVI, Series II, Anno IX* **1**, pp. 326–333.

65. Combettes, P. (2000) "Fejér monotonicity in convex optimization." in *Encyclopedia of Optimization*, C.A. Floudas and P. M. Pardalos, editors, Boston: Kluwer Publ.

66. Combettes, P., and Wajs, V. (2005) "Signal recovery by proximal forward-backward splitting." *Multiscale Modeling and Simulation*, **4(4)**, pp. 1168–1200.

67. Csiszár, I. (1975) "I-divergence geometry of probability distributions and minimization problems." *The Annals of Probability* **3(1)**, pp. 146–158.

68. Csiszár, I. (1989) "A geometric interpretation of Darroch and Ratcliff's generalized iterative scaling." *The Annals of Statistics* **17(3)**, pp. 1409–1413.

69. Csiszár, I. and Tusnády, G. (1984) "Information geometry and alternating minimization procedures." *Statistics and Decisions* **Supp. 1**, pp. 205–237.

70. Darroch, J. and Ratcliff, D. (1972) "Generalized iterative scaling for log-linear models." *Annals of Mathematical Statistics* **43**, pp. 1470–1480.

71. De Pierro, A. (1995) "A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography." *IEEE Transactions on Medical Imaging* **14**, pp. 132–137.

72. De Pierro, A., and Yamaguchi, M. (2001) "Fast EM-like methods for maximum 'a posteriori' estimates in emission tomography" *Transactions on Medical Imaging*, **20 (4)**.

73. Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977) "Maximum likelihood from incomplete data via the EM algorithm." *Journal of the Royal Statistical Society, Series B* **37**, pp. 1–38.

74. Dolidze, Z.O. (1982) "Solution of variational inequalities associated with a class of monotone maps."*Ekonomika i Matem. Metody* **18 (5)**, pp. 925–927 (in Russian).

75. Donoho, D. (2006) "Compressed sampling" *IEEE Transactions on Information Theory*, **52 (4)**. (download preprints at http://www.stat.stanford.edu/ donoho/Reports).

76. Eggermont, P.P.B., Herman, G.T., and Lent, A. (1981) "Iterative algorithms for large partitioned linear systems, with applications to image reconstruction. "*Linear Algebra and its Applications* **40**, pp. 37–67.

77. Eggermont, P.P.B., LaRiccia, V.N. (1995) "Smoothed maximum likelihood density estimation for inverse problems." Annals of Statistics **23**, pp. 199–220.

78. Eggermont, P., and LaRiccia, V. (2001) *Maximum Penalized Likelihood Estimation, Volume I: Density Estimation*. New York: Springer.

79. Elsner, L., Koltracht, L., and Neumann, M. (1992) "Convergence of sequential and asynchronous nonlinear paracontractions." *Numerische Mathematik*, **62**, pp. 305–319.

80. Farnell, A.B. (1944) "Limits for the characteristic roots of a matrix." *Bulletin of the American Mathematical Society*, **50**, pp. 789–794.

81. Fiacco, A., and McCormick, G. (1990) *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. Philadelphia, PA: SIAM Classics in Mathematics (reissue).

82. Fiddy, M. (2008) *private communication*.

83. Geman, S., and Geman, D. (1984) "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images."*IEEE Transactions on Pattern Analysis and Machine Intelligence* **PAMI-6**, pp. 721–741.

84. Gerchberg, R. W. (1974) "Super-restoration through error energy reduction."*Optica Acta* **21**, pp. 709–720.

85. Gill, P., Murray, W., Saunders, M., Tomlin, J., and Wright, M. (1986) "On projected Newton barrier methods for linear programming and an equivalence to Karmarkar's projective method." *Mathematical Programming*, **36**, pp. 183–209.

86. Golshtein, E., and Tretyakov, N. (1996) *Modified Lagrangians and Monotone Maps in Optimization*. New York: John Wiley and Sons, Inc.

87. Gordon, R., Bender, R., and Herman, G.T. (1970) "Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and x-ray photography." *J. Theoret. Biol.* **29**, pp. 471–481.

88. Grcar, J. (2011) "John von Neumann's analysis of Gaussian elimination and the origins of modern numerical analysis." *SIAM Review*, **53 (4)**, pp. 607–682.

89. Green, P. (1990) "Bayesian reconstructions from emission tomography data using a modified EM algorithm." *IEEE Transactions on Medical Imaging* **9**, pp. 84–93.

90. Hebert, T. and Leahy, R. (1989) "A generalized EM algorithm for 3-D Bayesian reconstruction from Poisson data using Gibbs priors." *IEEE Transactions on Medical Imaging* **8**, pp. 194–202.

91. Herman, G.T., Censor, Y., Gordon, D., and Lewitt, R. (1985) "Comment (on the paper [130])." *Journal of the American Statistical Association* **80**, pp. 22–25.

92. Herman, G. T. (1999) *private communication.*

93. Holte, S., Schmidlin, P., Linden, A., Rosenqvist, G. and Eriksson, L. (1990) "Iterative image reconstruction for positron emission tomography: a study of convergence and quantitation problems." *IEEE Transactions on Nuclear Science* **37**, pp. 629–635.

94. Hudson, M., Hutton, B., and Larkin, R. (1992) "Accelerated EM reconstruction using ordered subsets." *Journal of Nuclear Medicine*, **33**, p.960.

95. Hudson, H.M. and Larkin, R.S. (1994) "Accelerated image reconstruction using ordered subsets of projection data." *IEEE Transactions on Medical Imaging* **13**, pp. 601–609.

96. Kaczmarz, S. (1937) "Angenäherte Auflösung von Systemen linearer Gleichungen." *Bulletin de l'Academie Polonaise des Sciences et Lettres* **A35**, pp. 355–357.

97. Karmarkar, N. (1984) "A new polynomial-time algorithm for linear programming." *Combinatorica*, **4**, pp. 373–395.

98. Krasnosel'skiĭ, M. (1955) "Two remarks on the method of successive approximations" (in Russian). *Uspekhi Mathematicheskikh Nauk*, **10**, pp. 123–127.

99. Kullback, S. and Leibler, R. (1951) "On information and sufficiency." *Annals of Mathematical Statistics* **22**, pp. 79–86.

100. Landweber, L. (1951) "An iterative formula for Fredholm integral equations of the first kind." *Amer. J. of Math.* **73**, pp. 615–624.

101. Lange, K. and Carson, R. (1984) "EM reconstruction algorithms for emission and transmission tomography." *Journal of Computer Assisted Tomography* **8**, pp. 306–316.

102. Lange, K., Bahn, M. and Little, R. (1987) "A theoretical study of some maximum likelihood algorithms for emission and transmission tomography." *IEEE Trans. Med. Imag.* **MI-6(2)**, pp. 106–114.

103. Leahy, R. and Byrne, C. (2000) "Guest editorial: Recent development in iterative image reconstruction for PET and SPECT." *IEEE Trans. Med. Imag.* **19**, pp. 257–260.

104. Liao, C.-W., Fiddy, M., and Byrne, C. (1997) "Imaging from the zero locations of far-field intensity data." *Journal of the Optical Society of America -A* **14 (12)**, pp. 3155–3161.

105. Mann, W. (1953) "Mean value methods in iteration." *Proceedings of the American Mathematical Society*, **4**, pp. 506–510.

106. McLachlan, G.J. and Krishnan, T. (1997) *The EM Algorithm and Extensions*. New York: John Wiley and Sons, Inc.

107. Meidunas, E. (2001) *Re-scaled Block Iterative Expectation Maximization Maximum Likelihood (RBI-EMML) Abundance Estimation and Sub-pixel Material Identification in Hyperspectral Imagery*, MS thesis, Department of Electrical Engineering, University of Massachusetts Lowell.

108. Moreau, J.-J. (1962) "Fonctions convexes duales et points proximaux dans un espace hilbertien." *C.R. Acad. Sci. Paris Sér. A Math.*, **255**, pp. 2897–2899.

109. Moreau, J.-J. (1963) "Propriétés des applications 'prox'." *C.R. Acad. Sci. Paris Sér. A Math.*, **256**, pp. 1069–1071.

110. Moreau, J.-J. (1965) "Proximité et dualité dans un espace hilbertien." *Bull. Soc. Math. France*, **93**, pp. 273–299.

111. Narayanan, M., Byrne, C. and King, M. (2001) "An interior point iterative maximum-likelihood reconstruction algorithm incorporating upper and lower bounds with application to SPECT transmission imaging." *IEEE Transactions on Medical Imaging* **TMI-20 (4)**, pp. 342–353.

112. Nash, S. and Sofer, A. (1996) *Linear and Nonlinear Programming*. New York: McGraw-Hill.

113. Nesterov, Y., and Nemirovski, A. (1994) *Interior-Point Polynomial Algorithms in Convex Programming.* Philadelphia, PA: SIAM Studies in Applied Mathematics.

114. Papoulis, A. (1975) "A new algorithm in spectral analysis and band-limited extrapolation."*IEEE Transactions on Circuits and Systems* **22**, pp. 735–742.

115. Reich, S. (1979) "Weak convergence theorems for nonexpansive mappings in Banach spaces." *Journal of Mathematical Analysis and Applications*, **67**, pp. 274–276.

116. Rockafellar, R. (1970) *Convex Analysis*. Princeton, NJ: Princeton University Press.

117. Rockmore, A., and Macovski, A. (1976) "A maximum likelihood approach to emission image reconstruction from projections." *IEEE Transactions on Nuclear Science*, **NS-23**, pp. 1428–1432.

118. Schmidlin, P. (1972) "Iterative separation of sections in tomographic scintigrams." *Nucl. Med.* **15(1)**.

119. Shepp, L., and Vardi, Y. (1982) "Maximum likelihood reconstruction for emission tomography." *IEEE Transactions on Medical Imaging*, **MI-1**, pp. 113–122.

120. Shieh, M., Byrne, C., and Fiddy, M. (2006) "Image reconstruction: a unifying model for resolution enhancement and data extrapolation: Tutorial." *Journal of the Optical Society of America, A*, **23(2)**, pp. 258–266.

121. Shieh, M., Byrne, C., Testorf, M., and Fiddy, M. (2006) "Iterative image reconstruction using prior knowledge." *Journal of the Optical Society of America, A*, **23(6)**, pp. 1292–1300.

122. Shieh, M., and Byrne, C. (2006) "Image reconstruction from limited Fourier data." *Journal of the Optical Society of America, A*, **23(11)**, pp. 2732–2736.

123. Silverman, B., Jones, M., Wilson, J., and Nychka, D. (1990) "A smoothed EM approach to indirect estimation problems, with particular reference to stereology and emission tomography (with discussion)." *Journal of the Royal Statistical Society* **B 52**, pp. 271–324.

124. Smith, C. Ray and Grandy, W.T., editors (1985) *Maximum-Entropy and Bayesian Methods in Inverse Problems*. Dordrecht: Reidel Publ.

125. Smith, C. Ray and Erickson, G., editors (1987) *Maximum-Entropy and Bayesian Spectral Analysis and Estimation Problems*. Dordrecht: Reidel Publ.

126. Tanabe, K. (1971) "Projection method for solving a singular system of linear equations and its applications." *Numer. Math.* **17**, pp. 203–214.

127. Teboulle, M. (1992) "Entropic proximal mappings with applications to nonlinear programming." *Mathematics of Operations Research*, **17(3)**, pp. 670–690.

128. van der Sluis, A. (1969) "Condition numbers and equilibration of matrices." *Numer. Math.*, **14**, pp. 14–23.

129. van der Sluis, A., and van der Vorst, H.A. (1990) "SIRT- and CG-type methods for the iterative solution of sparse linear least-squares problems." *Linear Algebra and its Applications*, **130**, pp. 257–302.

130. Vardi, Y., Shepp, L.A. and Kaufman, L. (1985) "A statistical model for positron emission tomography." *Journal of the American Statistical Association* **80**, pp. 8–20.

131. von Neumann, J., and Goldstine, H. H. (1947) "Numerical inverting of matrices of high order." Bulletin of the American Mathematical Society, **53**, pp. 1021–1099.

132. Wernick, M. and Aarsvold, J., editors (2004) *Emission Tomography: The Fundamentals of PET and SPECT*. San Diego: Elsevier Academic Press.

133. Wright, M. (2005) "The interior-point revolution in optimization: history, recent developments, and lasting consequences." *Bulletin (New Series) of the American Mathematical Society*, **42(1)**, pp. 39–56.

134. Yang, Q. (2004) "The relaxed CQ algorithm solving the split feasibility problem." *Inverse Problems*, **20**, pp. 1261–1266.

Index