

MULTIPLICATIVE ITERATIVE ALGORITHMS FOR POSITIVE CONSTRAINED RECONSTRUCTIONS IN EMISSION AND TRANSMISSION TOMOGRAPHY

Jun Ma

Department of Statistics, Macquarie University
NSW 2109, Australia

ABSTRACT

This paper introduces a multiplicative iterative (MI) algorithm for image reconstructions in tomography. This algorithm can accommodate objective functions deduced from different probability models for measurements. Poisson and Gaussian (for both emission and transmission scans), or shifted Poisson (for precorrected PET and X-ray CT), are examples of such measurement probability models. This MI algorithm is very easy to implement and respects the positivity constraint. Furthermore, an exact or approximate line search step can be easily incorporated into this algorithm so that the objective functions are guaranteed to increase during the iterations.

Key Words: Multiplicative iterative (MI) algorithm, emission and transmission tomography, line search, EM, ISRA.

I. INTRODUCTION

Multiplicative iterative (MI) algorithms, such as the multiplicative algebraic reconstruction technique (MART) and maximum likelihood EM (ML-EM), are widely used in tomography. They aim to optimize different objective functions, e.g. Shannon's entropy for MART and Poisson based log-likelihood for ML-EM. One distinctive advantage of MI algorithms is that they automatically impose the positivity constraint (i.e. reconstructed image values are not less than zero). In this paper, we propose an MI algorithm for optimizing penalized types of objective functions, such as penalized log-likelihood (PL) or penalized least squares (PLS). This MI scheme can be expressed as a gradient-based algorithm with an ascending direction.

Nowadays, statistical image reconstruction in emission or transmission tomography can be developed based on the specified probability model for measurements y_1, \dots, y_n . For example, for SPECT scans, possible options are Gaussian [1] and Poisson [2] models. Poisson is also proposed for transmission scans [3]. For randoms pre-corrected PET scans, possible measurement models are Gaussian, ordinary Poisson and shifted Poisson [4]. Note that shifted Poisson is also used for modeling X-ray CT measurements [5].

Different algorithms have been developed to maximize their corresponding objective functions. For example, ML-EM [2] is designed for maximizing the likelihood of Poisson

measurements, or the iterative space reconstruction algorithm (ISRA) [1] for maximizing the likelihood of (constant variance) Gaussian measurements. Attractive aspects of both ML-EM and ISRA are that they are very easy to implement and that they respect the positivity constraint. However, if the objective function contains a penalty term then both ML-EM and ISRA become impractical as they involve, in each iteration, non-linear systems that do not possess closed form solutions.

Two types of algorithms exist in tomography that can trivially impose the positivity constraint: iterative coordinate ascent (ICA) [6] and optimization transfer (OT) [7] algorithms. ML-EM and DePierro's modification to EM (MEM) [8] are examples of OT for, respectively, ML and MPL Poisson emission tomography.

Our MI algorithm is competitive with ICA and OT in tomography. We will demonstrate that the MI algorithm automatically enforces the positivity constraint. However, MI by itself may not guarantee monotonicity of the objective function. An attractive feature of MI is that it easily incorporates a line search step so that the positivity constraint is enforced and, at the mean time, the objective function is assured to increase. MI with line search does not incur extra computing cost; its computational burden is tantamount to ML-EM.

Section II first describes the general MI algorithm and then implements it to emission and transmission under various measurement probability models. Simulation results are reported in Section III with concluding remarks given in Section IV.

II. A MULTIPLICATIVE ITERATIVE ALGORITHM FOR OBJECTIVE OPTIMIZATION

II-A. Objective optimization

In what follows, $x_j \geq 0$ denotes the emission activity or attenuation coefficient of pixel j , $j = 1, \dots, p$, and x denotes the p -vector of all x_j . Let the camera measurements be denoted by y_i ($i = 1, \dots, n$) and y be the n -vector for all measurements.

Statistical reconstructions in emission and transmission are usually obtained by:

$$\hat{x} = \operatorname{argmax}_{x \geq 0} \Psi(x). \quad (1)$$

The objective function $\Psi(x)$ generally displays the following structure:

$$\Psi(x) = \sum_{i=1}^n l_i(\mu_i|y_i) - hJ(x), \quad (2)$$

where μ_i is a function of x given by

$$\mu_i(x) = \eta_i(x) + r_i = \begin{cases} A_i x + r_i & \text{emission;} \\ b_i e^{-A_i x} + r_i & \text{transmission,} \end{cases} \quad (3)$$

where A_i is the i th row of the system matrix A , b_i the known blank scan counts of the i th detector and r_i the known mean background counts. We assume functions $l_i(\mu_i)$ ($i = 1, \dots, n$) and $J(x)$ are twice differentiable.

The first term of (2) measures data mismatch, which can be depicted by the log-likelihood of y . Thus $l_i(\mu_i|y_i)$ represents the log probability density function (pdf) of y_i . For example, if y_i follows Poisson(μ_i), then $l_i = -\mu_i + y_i \log \mu_i$ [2]; or if y_i follows normal $N(\mu_i, \mu_i)$, then $l_i = -\frac{1}{2}(y_i - \mu_i)^2/\mu_i$ (weighted least squares) [9]. Another example is the randoms precorrected PET scan using the shifted Poisson model for measurements, where $l_i = -(\eta_i + 2r_i) + (y_i + 2r_i) \log(\eta_i + 2r_i)$ [4].

The second term of (2), called the penalty function, quantifies the local smoothness of x . Parameter h (> 0), referred to as the smoothing or regularization parameter, balances two conflicting targets: fidelity of $A\hat{x}$ to y and smoothness of \hat{x} . Function $J(x)$ can take the form of $J(x) = \sum_{j=1}^p \nu_j(C_j x)$, where $C_j x$ meditates a neighborhood operation (such as the first or second order difference) on pixel j , and ν_j is a function for penalty, such as quadratic, Huber or hyperbolic functions. Other $\nu_j(x)$ are also possible.

It is possible that $\Psi(x)$ has multiple local maxima. In this case, MI finds one of the positively constrained local maximum, depending on the starting value of the algorithm.

II-B. Multiplicative iterative (MI) algorithms

We adopt the following notations in this paper. Let $b(z)^+ = \max(0, b(z))$ and $b(z)^- = \min(0, b(z))$, so that

$$b(z) = b(z)^+ + b(z)^-.$$

Besides, $b'(z)$ represents the derivative of b with respect to z and $b'_j(z)$, the derivative of b with respect to the j th element of z (i.e. z_j). $b'(z^{(k)})$ denotes $b'(z)|_{z=z^{(k)}}$ where $z^{(k)}$ is the estimate of z at iteration k .

The Krush-Kuhn-Tucker necessary condition for the positively constrained optimization of $\Psi(x)$ is: $\Psi'_j(x) = 0$ if $x_j > 0$ and $\Psi'_j(x) \leq 0$ if $x_j = 0$, for $j = 1, \dots, p$. Thus, we aim to solve

$$x_j \left(\sum_{i=1}^n l'_i(\mu_i) \frac{\partial \mu_i}{\partial x_j} - hJ'_j(x) \right) = 0, \quad (4)$$

where $\frac{\partial \mu_i}{\partial x_j} = a_{ij}$ for emission and $\frac{\partial \mu_i}{\partial x_j} = -\eta_i a_{ij}$ for transmission tomography.

After separating the positive and negative terms of (4), and moving the negative terms to the other side of the equation, the resulting expression naturally suggests the following MI algorithms:

$$x_j^{(k+1/2)} = x_j^{(k)} \frac{\sum_{i=1}^n l'_i(\mu_i^{(k)})^+ a_{ij} - hJ'_j(x^{(k)})^-}{-\sum_{i=1}^n l'_i(\mu_i^{(k)})^- a_{ij} + hJ'_j(x^{(k)})^+} \quad (5)$$

for emission tomography and

$$x_j^{(k+1/2)} = x_j^{(k)} \frac{-\sum_{i=1}^n l'_i(\mu_i^{(k)})^- \eta_i^{(k)} a_{ij} - hJ'_j(x^{(k)})^-}{\sum_{i=1}^n l'_i(\mu_i^{(k)})^+ \eta_i^{(k)} a_{ij} + hJ'_j(x^{(k)})^+} \quad (6)$$

for transmission tomography. Here $x^{(k+1/2)}$ denotes a temporary update of x from $x^{(k)}$. $x^{(k+1/2)}$ will be further enhanced by a line search step (discussed later) to give $x^{(k+1)}$.

As both the numerator and denominator of (5) or (6) are non-negative, $x^{(k+1/2)} \geq 0$ when $x^{(k)} > 0$.

When $x_j^{(k)} > 0$, $x_j^{(k+1/2)} = 0$ only if $J'_j(x^{(k)}) \geq 0$ together with

$$\begin{cases} \sum_i l'_i(\mu_i^{(k)})^+ a_{ij} = 0 & \text{emission;} \\ \sum_i l'_i(\mu_i^{(k)})^- \eta_i^{(k)} a_{ij} = 0 & \text{transmission.} \end{cases}$$

Once $x_j^{(k)} = 0$ it remains at zero for subsequent iterations. Thus we may remove x_j once it hits zero and continue the iterations with the remaining x 's. For this reason we can assume $x_j^{(k)} > 0$ for the MI algorithm.

Equations (5) or (6) serve as fundamental tools for developing positive constrained algorithms under different options of $l_i(\mu_i)$. For example, if measurements $y_i \sim \text{Poisson}(\mu_i)$ then x is updated by

$$x_j^{(k+1/2)} = x_j^{(k)} \frac{\sum_{i=1}^n a_{ij} y_i / \mu_i^{(k)} - hJ'_j(x^{(k)})^-}{\sum_{i=1}^n a_{ij} + hJ'_j(x^{(k)})^+} \quad (7)$$

for emission tomography, and by

$$x_j^{(k+1/2)} = x_j^{(k)} \frac{\sum_{i=1}^n a_{ij} \eta_i^{(k)} - hJ'_j(x^{(k)})^-}{\sum_{i=1}^n a_{ij} y_i \eta_i^{(k)} / \mu_i^{(k)} + hJ'_j(x^{(k)})^+} \quad (8)$$

for transmission tomography. When $h = 0$ and all $r_i = 0$ (i.e. no penalty and background noise), (7) becomes ML-EM of Shepp and Vardi [2] for emission tomography and (8) reduces to the algorithm of Lange *et al* [10] for transmission tomography. Another example is that when $y_i \sim N(\mu_i, 1/w_i)$, where $w_i > 0$ is independent of x , then (5) and (6) become

$$x_j^{(k+1/2)} = x_j^{(k)} \frac{\sum_{i=1}^n w_i a_{ij} y_i - hJ'_j(x^{(k)})^-}{\sum_{i=1}^n w_i a_{ij} \mu_i^{(k)} + hJ'_j(x^{(k)})^+} \quad (9)$$

and

$$x_j^{(k+1/2)} = x_j^{(k)} \frac{\sum_{i=1}^n w_i a_{ij} \eta_i^{(k)} \mu_i^{(k)} - hJ'_j(x^{(k)})^-}{\sum_{i=1}^n w_i a_{ij} \eta_i^{(k)} y_i + hJ'_j(x^{(k)})^+} \quad (10)$$

respectively. When $h = 0$ and w_i a constant, (9) agrees with ISRA of Titterton [1].

Above examples show that MI demands only projection, backprojection and the first derivative of $J(x)$, and thus is easy to implement in tomography.

Iterations in (5) or (6) may fail to assure increments of $\Psi(x)$. To circumvent this problem we introduce a line search step. First, write (5) or (6) as

$$x^{(k+1/2)} = x^{(k)} + S^{(k)}\Psi'(x^{(k)}), \quad (11)$$

where $S = \text{diag}(s_1, \dots, s_p)$, a diagonal matrix with entries $s_j = x_j / (-\sum_{i=1}^n l'_i(\mu_i)^- a_{ij} + hJ'_j(x))$ for emission and $s_j = x_j / (\sum_{i=1}^n l'_i(\mu_i)^+ \eta_i a_{ij} + hJ'_j(x))$ for transmission problems. Let $d^{(k)} = x^{(k+1/2)} - x^{(k)}$.

Since all $s_j^{(k)}$ are positive, $d^{(k)} = S^{(k)}\Psi'(x^{(k)})$ is in an uphill direction. If it occurs that $\Psi(x^{(k+1/2)}) < \Psi(x^{(k)})$ (overshooting step size) then a line search is necessitated, i.e. find $0 < \alpha^{(k)} < 1$ such that

$$\begin{aligned} x^{(k+1)} &= x^{(k)} + \alpha^{(k)}d^{(k)} \\ &= (1 - \alpha^{(k)})x^{(k)} + \alpha^{(k)}x^{(k+1/2)}, \end{aligned} \quad (12)$$

with the property $\Psi(x^{(k+1)}) \geq \Psi(x^{(k)})$, where the equality holds only when $x^{(k+1)} = x^{(k)}$.

From (12), it has $x^{(k+1)} \geq 0$ whenever $0 < \alpha^{(k)} < 1$, i.e. the results of line search automatically meet the positivity constraint. This is a remarkable advantage over other line search methods in tomography, such as the bent line approach [11].

Ideally, $\alpha^{(k)}$ is determined by:

$$\alpha^{(k)} = \arg \max_{\alpha} \Psi(x^{(k)} + \alpha d^{(k)}).$$

Two possible line searches are: exact or approximate line search. For exact line search we first let $F = -E[\Psi''(x)]$, the expected information matrix of y . It can show that, when $\Psi(x)$ is given by (2) with $E[l_i(\mu_i|Y_i)] = 0$ (which is always true when l_i is logarithm of pdf), F is given by: $F = A^T W A + hJ''(x)$, where $W = D V D$. Here D and V are diagonal matrices: $D = I$ for emission tomography and $D = \text{diag}\{\eta_1, \dots, \eta_n\}$ for transmission tomography, while $V = \text{diag}\{-E(\frac{\partial^2 l_1}{\partial \mu_1^2}), \dots, -E(\frac{\partial^2 l_p}{\partial \mu_p^2})\}$. When $\Psi(x)$ is approximated by

$$\begin{aligned} \Psi(x) &\approx \Psi(x^{(k)}) + (x - x^{(k)})^T \Psi'(x^{(k)}) \\ &\quad + \frac{1}{2} (x - x^{(k)})^T F^{(k)} (x - x^{(k)}), \end{aligned} \quad (13)$$

the ‘‘optimal’’ $\alpha^{(k)}$ is:

$$\alpha^{(k)} = \frac{(d^{(k)})^T (S^{(k)})^{-1} d^{(k)}}{(d^{(k)})^T A^T W^{(k)} A d^{(k)} + h(d^{(k)})^T J''(x^{(k)}) d^{(k)}}. \quad (14)$$

Both the numerator and denominator of (14) are simple to compute given $J''(x)$ is not too complicated. The operation $A d^{(k)}$ adds no extra cost as it can be used to update Ax : $Ax^{(k+1)} = Ax^{(k)} + \alpha^{(k)} A d^{(k)}$.

For inexact line search we propose the following backtracking scheme: for a selected ξ check the following Armijo’s condition

$$\Psi(x^{(k)} + \xi d^{(k)}) \geq \Psi(x^{(k)}) + \epsilon \xi d^{(k)} \Psi'(x^{(k)}) \quad (15)$$

where $0 < \epsilon < 1$ is a fixed threshold (for example $\epsilon = 10^{-2}$); if (15) is satisfied then stop, otherwise re-set $\xi = \rho \xi$ (where $\rho < 1$, such as $\rho = 0.8$) and re-evaluate (15). This procedure is continued until a suitable ξ is obtained and then $\alpha^{(k)} = \xi$. Note this inexact line search requires repeated computations of $\Psi(x^{(k)} + \xi d^{(k)})$ at different ξ values.

We denote the MI algorithm with exact or inexact line search by MI-EL and MI-IEL respectively. The line search is only required when $\Psi(x^{(k+1/2)}) \leq \Psi(x^{(k)})$. For both MI-EL and MI-IEL we can show that

- 1) If $\Psi(x)$ is strictly concave then, under certain regularity conditions, MI-EL or MI-IEL produces a convergent sequence $\{x^{(k)}\}$ from any initial $x^{(0)} > 0$.
- 2) Assume MI-EL or MI-IEL produced iterations converging to x^* : $x^{(k)} \rightarrow x^*$ when $k \rightarrow \infty$, then the limit x^* satisfies the Krush-Kuhn-Tucker condition.

III. RESULTS

This section reports a simulation study comparing MI-EL with three competitors: (i) the modified EM (MEM) algorithm of De Pierro [8]; (ii) the SPS algorithm of Fessler and Erdoĝan [12] and (iii) the ICA algorithm of Bouman and Sauer [6] where pixel level optimization was performed by Fisher scoring (denoted ICA/FS). We only consider a simulated SPECT system and reconstructions in emission tomography in this paper.

This simulation used a phantom of size 64×64 pixels (so that $p = 64^2$). There were 64 attenuated projections uniformly spaced over 360° , each projection contained 64 measurements (thus $n = 64^2$). Attenuation coefficients were 0.15 /cm (water) within the body and 0.0375 /cm within the two lungs. We did not consider collimator blurring nor scattering in this simulation. The system matrix A (with dimension $64^2 \times 64^2$) was pre-determined by the geometry of pixels, adjusted according to the attenuation coefficients. Poisson noise was added to expected projections $\mu = Ax$ to form the observed measurements vector y . The total projection count was 400,605.

We used a quadratic penalty in the MPL reconstructions: $J(x) = \frac{1}{2} \sum_{j=1}^p (x_j - \text{ave}\{x_k\})^2$, where $k \in \mathcal{N}_j$ with \mathcal{N}_j the neighborhood of pixel j . Two smoothing values were tested: $h = 10^{-3}$ (over smoothing) and $h = 10^{-5}$ (moderate smoothing), and the algorithms were compared corresponding to these smoothing values. All tested algorithms had an equal (uniform) initial start in order to induce a fair comparison.

Fig. 1 provides the plot of log-likelihood (penalized) against iteration numbers for $h = 10^{-3}$ (in (a)) and

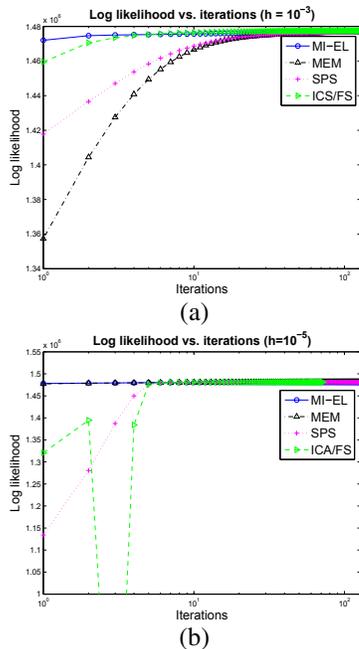


Fig. 1. log-likelihood versus iterations plot: (a) $h = 10^{-3}$ and (b) $h = 10^{-5}$.

$h = 10^{-5}$ (in (b)). The MI-EL performed well in both cases: it offered faster initial speed of convergence than the competitors. ICA/FS oscillated in the first four iterations for $h = 10^{-5}$, however it converged faster once it was stabilized.

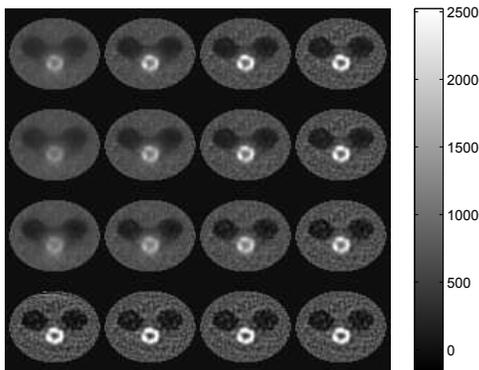


Fig. 2. Reconstructed images by MI-EL (row 1), MEM (row 2), SPS (row 3) and ICA/FS (row 4). Columns correspond to iterations 8, 16, 32 and 64.

Fig. 2 gives reconstructions by the tested algorithms: MI-EL (row 1), MEM (row 2), SPS (row 3) and ICA/FS (row 4). Columns of Fig. 2 correspond to iterations 8, 16, 32 and 64. Clearly, ICA/FS restored high frequency components in the early iterations whereas the other three algorithms only recovered low frequency components. This phenomena is well understood and documented in, for example, [6].

IV. CONCLUSIONS

We conclude that MI is an efficient algorithm for constrained MPL in tomography. Unlike OT or ICA, MI is very simple to derive and implement. It can be conceived as a unifying algorithm for tomography.

The MI algorithm not only enforces the positivity constraint automatically, but also easily enables an exact line search. Our limited simulations demonstrate that MI competes favorably (at least in Poisson emission model) with its competitors, such as MEM, SPS and ICA.

V. REFERENCES

- [1] D. M. Titterton, "On the iterative image space reconstruction algorithm for ECT," *IEEE Trans. Med. Imaging*, vol. 6, pp. 52 – 56, 1987.
- [2] L. A. Shepp and Y. Vardi, "Maximum likelihood estimation for emission tomography," *IEEE Trans. Med. Imaging*, vol. MI-1, pp. 113–121, 1982.
- [3] K. Lange and R. Carson, "EM reconstruction algorithms for emission and transmission tomography," *J. Comp. Assisted Tomography*, vol. 8, pp. 306–316, 1984.
- [4] M. Yavuz and J. A. Fessler, "Statistical image reconstruction methods for randoms-precorrected PET scans," *Medical Image Analysis*, vol. 2, pp. 369–378, 1998.
- [5] B. R. Whiting, "Signal statistics in x-ray computed tomography," in *Proc. SPIE 4682, Medical Imaging 2002: Med Phys.*, pp. 53 – 60, 2002.
- [6] C. A. Bouman and K. Sauer, "A unified approach to statistical tomography using coordinate descent optimization," *IEEE Trans. Med. Imaging*, vol. 5, pp. 480–492, 1996.
- [7] K. Lange, D. R. Hunter, and I. Yang, "Optimization transfer using surrogate objective functions," *J. Comput. Graphical Statist.*, vol. 9, pp. 1–20, 2000.
- [8] De Pierro, "A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography," *IEEE Trans. Med. Imaging*, vol. 14, pp. 132–137, 1995.
- [9] J. M. M. Anderson, B. A. Mair, M. Rao, and C-H. Wu, "Weighted least-squares reconstruction methods for positron emission tomography," *IEEE Trans. Med. Imaging*, vol. 16(2), pp. 159–165, 1997.
- [10] K. Lange, M. Bahn, and R. Little, "A theoretical study of some maximum likelihood algorithms for emission and transmission tomography," *IEEE Trans. Med. Imaging*, vol. 6, pp. 106–114, 1987.
- [11] L. Kaufman, "Implementing and accelerating the EM algorithm for positron emission tomography," *IEEE Trans. Med. Imaging*, vol. 6(1), pp. 37–51, 1987.
- [12] J. A. Fessler and Erdoğan, "A paraboloidal surrogates algorithm for convergent penalized-likelihood emission image reconstruction," *Proc. IEEE Nuc. Sci. Symp. Med. Im. Conf.*, pp. 1132–1135, 1998.