

扰动因素影响下对修正的 OS-EM 算法的收敛分析

于文杰, 王金平*

(宁波大学 理学院, 浙江 宁波 315211)

摘要: 在扰动因素影响下, 通过引入松弛参数对 Ordered-Subsets Expectation-Maximization (OS-EM) 算法进行修正. 基于修正迭代的 L^1 收敛性, 我们通过修改相应停止规则中的参数得到了一些主要的单调性结果, 进而证明了迭代的 L^2 收敛性.

关键词: 不适定性方程; OS-EM 迭代; 收敛性; Kullback-Leiber 距离

中图分类号: O177.6; O175.5 文献标志码: A 文章编号: 1001-5132 (2017) 02-0117-04

1997 年, Dempster 等^[1]引进 EM(expectation-maximization) 算法, 在处理可逆和不适定性问题时, 这种方法非常有用. 然而, 该算法收敛速率慢, 需要大量的计算才能得到令人满意的图像. 为此, 研究人员做了大量的工作来加速 EM 算法的收敛速率. 1994 年, Hudson 等^[2]提出了 OS-EM 算法, 有效加速了图像的重建进程. 2009 年, Haltmeier 等^[3]将 OS-EM 拓展到了无限维空间里, 并且提出了一种修正的 OS-EM 算法. 在处理可逆问题时, 噪音数据的边界通常在 L^2 范数上给出, 因此本文主要探究了修正后算法在 L^2 上的收敛性.

1 预备知识

我们从下面的算子方程开始研究,

$$Hf = g, \quad (1)$$

其中, $H: L^1(\Omega) \rightarrow L^1(\Sigma)$ 是第一类 Fredholm 积分算子, 且满足

$$Hf(x) = \int_{\Omega} h(x, y)f(y)dy, x \in \Sigma. \quad (2)$$

为求解该方程, 可将 EM 算法写成如下形式:

$$f_{n+1}(y) = f_n(y)H^*\left(\frac{g}{Hf_n}\right)(y) = f_n(y)\int_{\Sigma}\frac{h(x,y)g(x)}{Hf_n(x)}dx, \quad (3)$$

其中, H^* 是 H 的伴随算子.

通过寻求下面函数的最小值可以获得算子方程(1)的解:

$$l(f) := \int_{\Sigma}\left[g(x)\ln\frac{g(x)}{(Hf)(x)} - g(x) + (Hf)(x)\right]dx = D(g, Hf),$$

其中, $D(\phi, \psi)$ 是 KL 距离, 定义如下:

$$D(\phi, \psi) := \int\left[\phi(t)\ln\frac{\phi(t)}{\psi(t)} - \phi(t) + \psi(t)\right]dt, \quad (4)$$

其中, $\phi, \psi \in L^1(\Omega)$ 或 $\phi, \psi \in L^1(\Sigma)$. 注意到对于任意的 (ϕ, ψ) 有 $D(\phi, \psi) \geq 0$, 并且 $D(\phi, \psi) = 0$ 当且仅当 $\phi = \psi$ 时成立.

因此, $l(f)$ 的最小值满足以下关系式:

$$H^*\left(\frac{g}{Hf}\right) = H^*1. \quad (5)$$

假设 $(H^*1)(y) = 1$ 成立, 通过解下面方程我们可以获得(5)式的解:

$$fH^*\left(\frac{g}{Hf}\right) = f. \quad (6)$$

将有序子集的期望最大化(OS-EM)算法推广到无限维空间中时, 我们要将投影数据 g 划分成 N 个子集, 记作 $g_j := g|_{\Sigma_j}$, 在这里 $\Sigma_j \subset \Sigma$ 可以相交, 也可以不相交, 且满足 $\Sigma = \Sigma_0 \cup \dots \cup \Sigma_{N-1}$, 则可以将(1)式分解成如下的第一类算子方程系统:

$$H_j f = g_j, \quad j = 0, \dots, N-1, \quad (7)$$

并且将 H_j 定义为:

$$(H_j f)(x) := \int_{\Omega} h_j(x, y) f(y) dy, \quad (8)$$

其中, $h_j := h|_{\Omega \times \Sigma_j}$. 容易得出, 当 f 是系统(1)的解时, 它也可以解系统(7).

因此, 我们可以通过借助求下方程的最小值来获得系统(7)的解:

$$l_j(f) := \int_{\Sigma} [g_j(x) g_j(x) / ((H_j f)(x)) - g_j(x) + (H_j f)(x)] dx = D(g_j, H_j f), \quad (9)$$

其中, $H_j: L^1(\Omega) \rightarrow L^1(\Sigma)$, 并且 $g_j \in L^1(\Sigma)$.

类比(5)式的处理方式, 若有 $H_j^* 1 = 1, j = 0, \dots, N-1$ 成立, 则 $l_j(f)$ 最小值可通过 $H_j^*(g_j / (H_j f)) = 1$ 求得, 则其相应的方程形式为 $P_j(f) := f H_j^*(g_j / (H_j f)) = f$.

为了求解(7)式, 我们将与 Kaczmarz 类型对应的 OS-EM 算法写成如下形式:

$$f_{n+1} = P_j(f_n) = f_n \int_{\Sigma} \frac{h_j(x, \cdot) g_j(x)}{(H_j f_n)(x)} dx, \quad (10)$$

其中, 指标 $0 \leq j < N$ 与迭代指数 n 有关, 并且 $j = [n] := (n \bmod N)$. 观察可得 $N=1$ 的情形对应于标准的 EM 算法.

假设集合 $\Omega, \Sigma \subset \mathbb{R}^d$ ($d \geq 1$) 是开集并且有界, 令 Δ 表示下面的集合

$$\Delta := \left\{ u \in L^1(\Omega) : u \geq 0, \int_{\Omega} u(t) dt = 1 \right\}, \quad (11)$$

令 $f_0 \in \Delta$ 为(10)式的初始迭代.

为了后续分析, 我们做如下假设:

(A1) 内核函数 $h_j: \Sigma \times \Omega \rightarrow \mathbb{R}, j = 0, \dots, N-1$, 对于任意的 $y \in \Omega$ 有 $\int_{\Sigma} h_j(x, y) dx = 1$ 成立.

(A2) 存在常数 $m, M > 0$, 使得对于任意的 $(x, y) \in \Sigma \times \Omega$, 有 $m \leq h_j(x, y) \leq M$ 成立.

(A3) (7) 式中的精确数据 $g_j \in L^1(\Sigma)$ 满足 $\int_{\Sigma} g_j(x) dx = 1$; 此外, 存在正常数 M' , 使得对于任意 $x \in \Sigma$, 有 $g_j(x) < M'$ 成立.

(A4) (7) 式有一个非负解 $f^* \in \Delta$, 对于任意的 $y \in \Omega$, 它都不会变为 0; 并且 $D(f^*, f_0) < 0$.

由假设(A2)可知算子 $H_j: L^1(\Omega) \rightarrow L^1(\Sigma)$ 是连续的, 事实上, $H_j: L^p(\Omega) \rightarrow L^p(\Sigma)$ 也是连续的.

下面, 我们给出 KL 距离的一些性质, 关于 KL 更多的性质, 可参考文献[4-5].

性质 1 令 ϕ 和 ψ 是两个 L^1 空间上的函数, 那么 (ϕ, ψ) 在 KL 距离的定义域中, 有以下结论成立:

$$(i) \quad \|\psi - \phi\|_{L^1}^2 \leq \left(\frac{2}{3} \|\psi\|_{L^1} + \frac{4}{3} \|\phi\|_{L^1} \right) D(\psi, \phi).$$

(ii) 函数 $(\psi, \phi) \mapsto D(\psi, \phi)$ 是凸的.

(iii) 令 $\{\phi_n\}$ 和 $\{\psi_n\}$ 是 $L^1(\Omega)$ 中的序列, 且其中之一有界. 如果 $\lim_{n \rightarrow \infty} D(\psi_n, \phi_n) = 0$, 则 $\lim_{n \rightarrow \infty} \|\psi_n - \phi_n\|_{L^1} = 0$.

2 主要结果

关于 OS-EM 迭代的收敛性分析已在文献[3]中陈述, 以下我们分析在扰动因素影响下, 修正的 OS-EM 迭代的收敛性.

在处理可逆问题时, 近似测量值 $g_j^{\delta} \in L^2(\Sigma)$, 并且满足:

$$\|g_j - g_j^{\delta}\|_{L^2} \leq \delta_j, j = 0, \dots, N-1. \quad (12)$$

我们致力于寻求下方程的近似解:

$$H_j f = g_j^{\delta}, j = 0, \dots, N-1. \quad (13)$$

而对系统(13)的分析离不开下面的函数:

$$l_j^{\delta}(f) := \int_{\Sigma} [g_j^{\delta}(x) g_j^{\delta}(x) / ((H_j f)(x)) - g_j^{\delta}(x) + (H_j f)(x)] dx = D(g_j^{\delta}, H_j f). \quad (14)$$

我们将修正的 OS-EM 迭代定义为如下形式:

$$f_{n+1}^{\delta} = f_n^{\delta} \omega_n, \quad (15)$$

其中,

$$\omega_n = \begin{cases} \int_{\Sigma} \frac{h_{[n]}(x, \cdot) g_{[n]}^{\delta}(\cdot)}{H_{[n]} f_n^{\delta}(x)} dx =: P_{[n]}^{\delta}(f_n^{\delta}), \\ l_{[n]}^{\delta}(f_n^{\delta}) > \tau \delta_{[n]} \gamma_n, \\ 1, \text{其他.} \end{cases} \quad (16)$$

在(16)式中, 参数的选取满足:

$$\tau > 1, \gamma_n := \left\| \ln \frac{g_{[n]}^{\delta}}{H_{[n]} f_n^{\delta}} \right\|_{L^2}. \quad (17)$$

为后续分析, 介绍以下 3 个引理.

引理 1 如果假设(A1)~(A4)成立, g_j^{δ}, δ_j 由(12)所定义, 且对于 $0 \leq j \leq N-1$ 有 $\delta_j > 0$, 那么对于所有的自然数 n 且满足 $[n] = j$ 有下式成立:

$$l_{[n]}^{\delta}(f_n^{\delta}) - D(g_{[n]}, g_{[n]}^{\delta}) \leq D(f^*, f_n^{\delta}) - D(f^*, f_{n+1}^{\delta}). \quad (18)$$

证明 由假设成立及文献[5]中性质 5.2 可知, 对于任意的 $v, w \in \Delta, 0 \leq j \leq N-1$, 下式成立:

$$D(P_j(w), P_j^{\delta}(v)) \leq D(g_j, g_j^{\delta}) + D(P_j(w), v) - D(P_j(w), w) + l_j(w) - l_j(v).$$

因此, 当取 $j = [n], w = f^*, v = f_n^{\delta}$ 时, 对于满

足条件 $[n] = j$ 的自然数 n , 并且 $P_{[n]}(f^*) = f^*$, 可证得(18)式成立.

引理 2 如果假设(A2)成立, 则对于任意的 $f \in \Delta, x \in \Sigma$, 有 $m \leq (H_j f)(x) \leq M$.

证明 由假设(A2)可知:

$$m \leq h_j(x, y) \leq M. \quad (19)$$

对(19)式不等号两边同乘 $f(y)$ 后, 对其在 $y \in \Omega$ 上积分得到:

$$\begin{aligned} \int_{\Omega} m f(y) dy &\leq \int_{\Omega} h_j(x, y) f(y) dy \leq \int_{\Omega} M f(y) dy, \\ m \int_{\Omega} f(y) dy &\leq \int_{\Omega} h_j(x, y) f(y) dy \leq M \int_{\Omega} f(y) dy, \end{aligned}$$

即 $m \leq (H_j f)(x) \leq M$.

引理 3 如果对于任意的自然数 n , $\{\xi_n\}$ 是有限可测空间上的一个勒贝格可积函数序列, 并且一致有界, 那么当 $\lim_{n \rightarrow \infty} \|\xi_n - \xi\|_1 = 0$ 成立时, 对于任意的 $p \in [1, +\infty)$, 有 $\lim_{n \rightarrow \infty} \|\xi_n - \xi\|_p = 0$.

定理 1 假设(A1)~(A4)成立, τ 和 γ_n 在(17)式中被定义, 此外 g_j^δ, δ_j 由(12)式所定义, 且对于 $0 \leq j \leq N-1$ 有 $\delta_j > 0$. 序列 $\{f_n^\delta\}$ 为(15)式和(16)式所定义的迭代, 则对于每一个自然数 n , 下面的结论成立:

$$\begin{aligned} \text{(i)} \quad & l_{[n]}^\delta(f_{n+1}^\delta) \leq l_{[n]}^\delta(f_n^\delta); \\ \text{(ii)} \quad & D(f^*, f_{n+1}^\delta) \leq D(f^*, f_n^\delta); \\ \text{(iii)} \quad & \lim_{n \rightarrow \infty} l_{[n]}^\delta(f_n^\delta) = 0; \\ \text{(iv)} \quad & \lim_{n \rightarrow \infty} D(f_{n+1}^\delta, f_n^\delta) = 0. \end{aligned} \quad (20)$$

证明 (i)由文献[3]引理 3.2(ii)及 $D(\phi, \psi) \geq 0$ 可得结论成立.

(ii)如果 $l_{[n]}^\delta(f_n^\delta) \leq \tau \delta_{[n]} \gamma_n$, 由(16)式可知 $\omega_n = 1$, 即 $f_{n+1}^\delta = f_n^\delta$, 则(20)式得证. 如果 $l_{[n]}^\delta(f_n^\delta) > \tau \delta_{[n]} \gamma_n$, (18)式经过简单的计算可以得到:

$$\begin{aligned} D(f^*, f_n^\delta) - D(f^*, f_{n+1}^\delta) &\geq l_{[n]}^\delta(f_n^\delta) + \\ &\int_{\Sigma} [g_{[n]} - g_{[n]}^\delta] \ln \left(\frac{g_{[n]}^\delta}{H_{[n]} f_n^\delta} \right) dx. \end{aligned}$$

由柯西施瓦兹不等式可知:

$$\begin{aligned} l_{[n]}^\delta(f_n^\delta) + \int_{\Sigma} [g_{[n]} - g_{[n]}^\delta] \ln \left(\frac{g_{[n]}^\delta}{H_{[n]} f_n^\delta} \right) dx &\geq \\ l_{[n]}^\delta(f_n^\delta) - \|g_{[n]} - g_{[n]}^\delta\|_{L^2} \left\| \ln \left(\frac{g_{[n]}^\delta}{H_{[n]} f_n^\delta} \right) \right\|_{L^2} &\geq \end{aligned}$$

$$\tau \delta_{[n]} \gamma_n - \delta_{[n]} \gamma_n \geq (\tau - 1) \delta_{[n]} \gamma_n.$$

由(18)式可知, $D(f^*, f_n^\delta) - D(f^*, f_{n+1}^\delta) \geq (\tau -$

$1) \delta_{[n]} \gamma_n > 0$, 即(20)式得证.

(iii) 由(ii)可知, 存在 $\mu \geq 0$ 使得 $\lim_{n \rightarrow \infty} D(f^*, f_n^\delta) = \mu$, 进而有 $\lim_{n \rightarrow \infty} D(f^*, f_{n+1}^\delta) = \mu$. 由文献[3]引理3.2的(ii)知 $0 \leq \lim_{n \rightarrow \infty} l_{[n]}^\delta(f_n^\delta) \leq \lim_{n \rightarrow \infty} D(f^*, f_n^\delta) - \lim_{n \rightarrow \infty} D(f^*, f_{n+1}^\delta) = \mu - \mu = 0$.

所以 $\lim_{n \rightarrow \infty} l_{[n]}^\delta(f_n^\delta) = 0$.

(iv) 由(i)和(iii)知 $\lim_{n \rightarrow \infty} l_{[n]}^\delta(f_{n+1}^\delta) = 0$, 因此由 $D(f_{n+1}^\delta, f_n^\delta) \leq l_{[n]}^\delta(f_n^\delta) - l_{[n]}^\delta(f_{n+1}^\delta)$, $\lim_{n \rightarrow \infty} l_{[n]}^\delta(f_n^\delta) = 0$ 及 $\lim_{n \rightarrow \infty} l_{[n]}^\delta(f_{n+1}^\delta) = 0$ 可证得(iv).

定理 1 由(ii)可知, 我们可以为修正的 OS-EM 迭代寻找一个停止规则. 当满足

$$k_*^\delta := \min\{mN; f_{mN}^\delta = f_{mN+1}^\delta = \dots = f_{mN+N}^\delta\}, \quad (21)$$

即 $f_{k_*^\delta} = f_{k_*^\delta+1} = \dots = f_{k_*^\delta+N}$ 时, 迭代停止.

定理 2 假设(A1)~(A4)成立, 则对于所有的 $p \in [1, +\infty)$ 和 $j = 0, \dots, N-1$, 有

$$\lim_{\delta \rightarrow 0} \|H_j f_{k_*^\delta}^\delta - g_j\|_{L^p(\Sigma)} = 0. \quad (22)$$

证明 由定理 1(iv)可以得到 $\lim_{\delta \rightarrow 0} D(H_j f_{k_*^\delta}^\delta, g_j) = 0$. 对于 $p=1$ 时的情形, 由性质 1 的(iii)可知(22)式成立. 对于任意的 $1 \leq p < \infty$, 由引理 3 可知结论成立.

定理 3 假设定理 2 中条件成立, 令 $\{\delta^t := (\delta_0^t, \dots, \delta_{N-1}^t)\}$ (t 为自然数)为 $(0, \infty)^N$ 上的一个序列, 且满足 $\lim_{t \rightarrow \infty} \delta_j^t = 0$, 其中 $0 \leq j \leq N-1$. 此外, 令 $\{g^t := (g_0^t, \dots, g_{N-1}^t)\}$ 为扰动数据序列, 且满足 $\|g_j - g_j^t\|_{L^2} \leq \delta_j^t$, $j = 0, \dots, N-1$. 定义 $k_*^t := k_*^{\delta^t} = k_*(\delta^t, g^t)$ 为(21)式所定义的迭代停止指数. 如果 $f_0 \in L^\infty(\Omega)$, 对于 $p \in (1, +\infty)$, 序列 $\{f_{k_*^t}^{\delta^t}\}$ 在 $L^p(\Omega)$ 空间上是有界的, 那么此序列在 $L^p(\Omega)$ 空间里存在一个子列, 且弱收敛于系统(7)的一个解.

证明 首先, 我们证明 $f_n \in L^\infty(\Omega)$.

由于 $f_0 \in \Delta$, 由引理 2 可得 $m \leq H_0 f_0 \leq M$, 又假设(A2)和(A3)成立, 则当 $(x, y) \in \Sigma \times \Omega$ 时, 有

$$\frac{h_0(x, y) g_0(x)}{(H_0 f_0)(x)} \leq \frac{MM'}{m}.$$

因此由 $f_0 \in L^\infty(\Omega)$ 及(10)式可推出 $f_1 \in L^\infty(\Omega)$. 假设 $f_n \in L^\infty(\Omega)$ 成立, 则由假设(A2)和(A3)以及(10)式可知 $f_{n+1} \in L^\infty(\Omega)$. 由归纳法证得 $f_n \in L^\infty(\Omega)$.

对于 $p \in (1, +\infty)$, 由假设可知 $\{f_{k_*^t}^{\delta^t}\}$ 有界. 因此, 存在一个子列 $\{f_{k_*^s}^{\delta^s}\}$ (s 为自然数)弱收敛于 $z \in L^p(\Omega)$. 下面我们分两种情形来证明 z 是系统(13)的解.

情形 1 对于每个 $0 \leq j \leq N-1$, 子列 $\{f_{k_s^s}^{\delta^s}\}$ 中指数 s 都可以表示为 $s = j + mN$, 其中 m 为自然数. 当 $j=0$ 时, 考虑 $\{f_{k_s^s}^{\delta^s}\}$ 的子列 $\{f_{k_{s_i}^s}^{\delta^{s_i}}\} (s_i = m_i N)$. 显然, $[s_i] = 0$. 由定理 2 可知, 在 $L^p(\Sigma)$ 空间里 $H_0 f_{k_{s_i}^s}^{\delta^{s_i}} \rightarrow g_0$. 因为 $H_0 : L^p(\Omega) \rightarrow L^p(\Sigma)$ 是连续的, 所以也是弱连续的, 进而, $H_0 f_{k_s^s}^{\delta^s} \rightarrow H_0 z$. 由弱极限唯一我们得到 $H_0 z = g_0$.

考虑 $j=1, \dots, N-1$ 的情况, 可以得出 $H_j z = g_j$. 显然, z 是系统(13)的解.

情形 2 至少有一个 $0 \leq j_0 \leq N-1$ 使得序列 $\{f_{k_s^s}^{\delta^s}\}$ 中的指数 s 可以表示为 $s = j_0 + mN$, 其中 m 为自然数. 类似情形 1 的证明可以得出 $H_{j_0} z = g_{j_0}$.

接下来, 对于 $s_i + 1 = (j_0 + 1) + m_i N$, 我们考虑子列 $\{f_{k_{s_i+1}^s}^{\delta^{s_i+1}}\}$. 作为弱收敛序列 $\{f_{k_s^s}^{\delta^s}\}$ 的一个子列, 容易得到在 $L^p(\Omega)$ 上 $f_{k_{s_i+1}^s}^{\delta^{s_i+1}} \rightarrow z$. 由于 $H_{[j_0+1]}$ 是连续的, 所以在 $L^p(\Sigma)$ 空间里 $H_{[j_0+1]} f_{k_{s_i+1}^s}^{\delta^{s_i+1}} \rightarrow H_0 z$. 此外, 由定理 2 可知, 在 $L^p(\Sigma)$ 空间里 $H_{[j_0+1]} f_{k_{s_i+1}^s}^{\delta^{s_i+1}} \rightarrow g_{[j_0+1]}$. 故 $H_{[j_0+1]} z \rightarrow g_{[j_0+1]}$.

对于序列 $\{f_{k_{s_i+2}^s}^{\delta^{s_i+2}}\}, \dots, \{f_{k_{s_i+N-1}^s}^{\delta^{s_i+N-1}}\}$ 运用相同的方法, 对于每个 j , 可以得到 $H_j z \rightarrow g_j$.

综上所述, z 是系统(13)的解.

证毕.

3 总结

本文主要致力于为求解线性不适定方程问题所提出的修正的 OS-EM 算法的研究, 并且更多关注的是修正后算法的正则化性质, 也给出了 KL

距离的单调性以及迭代次数有界的情形下的弱收敛性. 在有扰动因素影响下, 有研究者也提出了一种修正的 OS-EM 迭代, 本文主要探究的是当满足条件 $\|g_j - g_j^\delta\|_{L^2} \leq \delta_j$ 时, 这种修正方法的一些单调性结果以及收敛性.

文中所得到的结果是在一些假设成立的条件下得到的, 因此我们后续将探究这些假设条件被放松时现有结论是否依然成立.

参考文献:

- [1] DEMPSTER A P, LAIRD N M, RUBIN D B. Maximum likelihood from incomplete data via the EM algorithm[J]. Journal of the Royal Statistical Society, 1997, 39(1):1-38.
- [2] HUDSON H M, LARKIN R S. Accelerated image reconstruction using ordered subsets of projection data[J]. IEEE Transactions on Medical Imaging, 1994, 13(4):601-609.
- [3] HALTMEIER M, LEITAO A, RESMERITA E. On the regularization methods of EM-Kaczmarz type[J]. Inverse Problems, 2009, 25(7):1-17.
- [4] RESMERITA E, ANDERSEN R S. Joint additive Kullback-Leibler residual minimization and regularization for linear inverse problems[J]. Mathematical Methods Applicable Scientia, 2007, 30(13):1527-1544.
- [5] RESMERITA E, ENGL H W, IUSEM A N. The expectation-maximization algorithm for ill-posed integral equations: A convergence analysis[J]. Inverse Problems, 2007, 24(5):2575-2588.

Convergence analysis of the modified OS-EM algorithm for noisy data

YU Wen-jie, WANG Jin-ping*

(Faculty of Science, Ningbo University, Ningbo 315211, China)

Abstract: In the noisy data case, we modify the Ordered-Subset Expectation-Maximization (OS-EM) algorithm by introducing a relaxation parameter. Based on the convergence of modified iteration in L^1 norm, we obtained some main monotonic results by modifying the parameters of the corresponding stopping rule, and proved the convergence for this modified iteration in L^2 norm.

Key words: ill-posed equations; OS-EM iteration; convergence; Kullback-Leiber distance

(责任编辑 章践立)