김일성종합대학학보 (수학) 주체107(2018)년 제64권 제1호

선형화된 분할브레그만방법에 대한 연구

김종철, 리금혁

우리는 선형거꿀문제를 풀기 위한 선형화된 분할브레그만방법에 대하여 연구를 진행하였다.

분할브레그만방법은 l_1 －정규화된 문제를 푸는데 성공적으로 적용되여왔고 수렴성도 증명되여있다. 그런데 이 방법에서는 일부 관측행렬에 대해서는 매 반복에서 대규모적이고 비구조적인 선형련립방정식이 나타나고 이것을 정확히 푸는것은 불가능하다. 그런데 수렴성은 이 련립방정식이 정확히 풀려진다는 가정하에서 증명되였다.

론문에서는 이 결함을 극복하기 위해 선형거꿀문제를 풀기 위한 종전의 분할브레그만방법과 선형화된 브레그만방법을 결합하여 새로운 선형화된 분할브레그만방법을 제기하고 수렴성을 증명하였다. 그리고 그라디엔트하강분할브레그만방법을 제기하고 그것이 선형화된 분할브레그만방법의 특수경우로 된다는것을 밝혔다.

1. 선행연구결과와 문제설정

선형거꿀문제는 보통 다음과 같이 정의된다.

$$g = Pf + \varepsilon$$

여기서 $g \in \mathbf{R}^l$ 은 알려진 벡토르이고 $f \in \mathbf{R}^n$ 은 미지벡토르이며($l \le n$ 이다.) $P \in \mathbf{R}^{l \times n}$ 은 주어진 사영연산자 또는 측정행렬이고 ε 은 측정오차벡토르이다.

론문에서는 측정오차가 없다고 보고 이때 선형거꿀문제의 풀이가 어떤 선형변환하에서 성글다는것이 사전에 알려졌다는 가정하에서 다음의 문제를 고찰한다.

$$\min_f \| Df \|_1, \ Pf = g \tag{1}$$

여기서 $D \in \mathbf{R}^{l \times n}$ ($l \ge n$)은 $\mathbf{R}^n$ 에서 $\mathbf{R}^l$ 에로 작용하는 분해연산자이다.

재구성모형문제 (1)을 효과적으로 풀기 위해 분할브레그만방법[2]이 제기되였고 그 수렴성이 증명[1]되였다. 분할브레그만방법은 실례로 전변동(TV)잡음제거(이 경우에 P 는 항등행렬이다.), 자기공명화상재구성(이 경우에 P 는 푸리에변환행렬의 행들의 부분모임으로 이루어진다.) 등에서와 같이 행렬 P 가 일정한 특수한 구조를 가지는 문제들에 대해서만 효과적으로 문제 (1)을 풀수 있다. CT화상재구성에서와 같이 일부 경우에는 $P^{\mathrm{T}}P$ 의 거꿀행렬을 고속으로 계산할수 있는 P 의 특수한 구조가 알려지지 않으므로 방법의 매 반복에서 최속하강법이나 공액경사법 또는 가우스－자이델방법, 뉴톤법 등을 리용하여 대규모부분문제를 풀어야 한다는것을 의미한다.

그런데 선행연구[1]에서 분할브레그만방법의 수렴성증명은 매 반복에서 부분문제를 정확히 푼다는 전제하에서 진행되였다.

론문에서는 이에 기초하여 임의의 관측행렬에 대해서도 문제 (1)을 효과적으로 푸는데 적합한 선형화된 분할브레그만(LSB)방법을 유도하고 수렴성을 증명한다.

소제목 2에서는 브레그만방법과 선형화된 브레그만방법, 분할브레그만방법을 간단히 소개하며 소제목 3에서는 선형화된 분할브레그만방법을 유도하고 그 수렴성을 증명한다.

2. 브레그만방법과 분할브레그만방법, 선형화된 브레그만방법

1) 브레그만방법

$J(f)=\| Df\|_1$로 표시하자. 최소화문제 (1)을 풀기 위한 브레그만방법[4]은 선행연구[3, 6]에서와 같이 다음의 단순화된 반복으로 쓸수 있다.

$$f^0=0,\ b^0=0$$

$$f^{k+1}=\arg\min_f J(f)+\frac{\lambda}{2}\| Pf-g+b^k\|_2^2 \tag{2}$$

$$b^{k+1}=b^k+(Pf^{k+1}-g) \tag{3}$$

브레그만방법의 수렴성[5]이 증명되였다.

2) 분할브레그만방법

식 (2)를 효과적으로 풀기 위한 분할브레그만방법[2]이 제기되였는데 그 방법의 기본내용은 식 (2)에서 비용함수의 l_1항과 l_2항을 분리시키는것이다.

문제 (1)을 풀기 위한 분할브레그만방법에서는 2개의 최량화문제를 부분문제로서 풀어야 한다. 즉

$$f^{k+1}=\arg\min_f \frac{\lambda_1}{2}\| Pf-g+b^k\|_2^2+\frac{\lambda_2}{2}\| d^k-Df-s^k\|_2^2 \tag{4}$$

$$d^{k+1}=\arg\min_d \| d\|_1+\frac{\lambda_2}{2}\| d-Df^{k+1}-s^k\|_2^2 \tag{5}$$

분할브레그만방법의 수렴성[1]은 증명되였다. 수렴성을 담보하기 위해서는 특히 식 (4), (5)가 정확히 풀려야 한다. 식 (5)에서는 최소점 d^{k+1}을 일반화된 수축공식을 리용하여 양적으로 구할수 있다.

그러나 식 (4)에서의 최소화문제는 실례로 1계최량성조건

$$(\lambda P^{\mathrm{T}}P+\mu D^{\mathrm{T}}D)f^{k+1}=\lambda_1 P^{\mathrm{T}}(g-b^k)+\lambda_2 D^{\mathrm{T}}(d^k-s^k) \tag{6}$$

을 리용하여 푼다고 하면 분명히 $\lambda P^{\mathrm{T}}P+\mu D^{\mathrm{T}}D$의 거꿀행렬을 계산하여야 한다. 다시말하여 대규모선형련립방정식을 풀기 위해 여러가지 반복법들이 실행되여야 한다. 실천에서는 식 (4)를 풀기 위해 그라디엔트하강법, 공액경사법 등이 적용될수 있다. 실례로 그라디엔트하강법에 의해

$$f^{k+1}=f^k-\alpha(\lambda_1 P^{\mathrm{T}}(Pf^k-g+b^k)+\lambda_2 D^{\mathrm{T}}(Df^k-d^k+s^k)) \tag{7}$$

을 얻는다. 또한 식 (5)는

$$q_d^{k+1}+\lambda_2(d^{k+1}-Df^{k+1}-s^k)=0,\ q_d^{k+1}\in\partial J(d^{k+1}) \tag{8}$$

과 동등하다. 때문에 분할브레그만방법의 어느 한 걸음에서 그라디엔트하강법을 적용한 그라디엔트하강분할브레그만(GDSB)방법은 다음과 같이 된다.

$$
\begin{aligned}
&f^{k+1}=f^k-\alpha(\lambda_1 P^{\mathrm{T}}(Pf^k-g+b^k)+\lambda_2 D^{\mathrm{T}}(Df^k-d^k+s^k))\\
&q_d^{k+1}+\lambda_2(d^{k+1}-Df^{k+1}-s^k)=0\\
&s^{k+1}=s^k+(Df^{k+1}-d^{k+1})\\
&b^{k+1}=b^k+(Pf^{k+1}-g)
\end{aligned}
\tag{9}
$$

3) 선형화된 브레그만방법

선형화된 브레그만방법[6]은 문제

$$\min_f J(f),\ \text{제한조건}\ Pf=g \tag{10}$$

를 풀기 위해 리용되였다. 선형화된 브레그만방법은 다음과 같다.

$$f^{k+1}=\arg\min_f B_J^{q^k}(f,\ f^k)+\frac{\lambda}{2\alpha}\| f-(f^k-\alpha P^{\mathrm{T}}(Pf^k-g))\|_2^2 \tag{11}$$

$$q^{k+1}=q^k+\frac{\lambda}{\alpha}(f^{k+1}-(f^k-\alpha P^{\mathrm{T}}(Pf^k-g))) \tag{12}$$

여기서 α는 정수이고 매 반복에서 걸음크기로서 생각할수 있다.

식 (11)을 다음의 동등한 형식을 리용하여 푼다.

$$f^{k+1}=\arg\min_f J(f)+\frac{\lambda}{2\alpha}\left\| f-\left(f^k+\alpha\left(\frac{1}{\lambda}q^k-P^{\mathrm{T}}(Pf^k-g)\right)\right)\right\|_2^2 \tag{13}$$

$J(f)=\| f\|_1$이라고 가정하면 수축연산자를 리용하여 식 (13)의 양적인 풀이를 얻을수 있다. 그러나 $J(f)=\| Df\|_1$이라고 가정하면 $\| Df\|_1$의 특수한 형식과 미분불가능성의 결과로서 식 (13)으로부터 f^{k+1}의 닫긴공식을 얻을수 없다.

3. 선형화된 분할브레그만방법

1) 선형화된 분할브레그만방법의 유도

$J(f)=\| Df\|_1$일 때 분할브레그만방법의 제한성을 극복하기 위해 즉 식 (4)에 있는 문제를 풀기 위해 선형화된 브레그만방법에서의 선형화수법을 리용하자.

$$
\begin{aligned}
f^{k+1}=\arg\min_f \frac{\lambda_1}{2\beta_1}\| f-(f^k-\beta_1 P^{\mathrm{T}}(Pf^k-g+b^k))\|_2^2+\\
+\frac{\lambda_2}{2\beta_2}\| f-(f^k-\beta_2 D^{\mathrm{T}}(Df^k-d^k+s^k))\|_2^2
\end{aligned}
\tag{14}
$$

이제 $\omega_1=\lambda_1/\beta_1,\ \omega_2=\lambda_2/\beta_2$라고 놓으면

$$f^{k+1}=f^k-\frac{\lambda_1 P^{\mathrm{T}}(Pf^k-g+b^k)+\lambda_2 D^{\mathrm{T}}(Df^k-d^k+s^k)}{\omega_1+\omega_2}$$

이 얻어진다.

이로부터 문제 (1)을 풀기 위한 선형화된 분할브레그만방법이 얻어진다.

걸음 1 초기점 $f^0=0,\ d^0=0,\ s^0=0,\ b^0=0$ 과 $0<\varepsilon<<1$ 그리고 옹근수 $K>0$ 이 주어졌다고 하자. $k:=0$ 으로 놓는다.

걸음 2 f 를 갱신한다.

$$f^{k+1}=f^k-\frac{\lambda_1 P^{\mathrm{T}}(Pf^k-g+b^k)+\lambda_2 D^{\mathrm{T}}(Df^k-d^k+s^k)}{\omega_1+\omega_2} \tag{15}$$

$r_k=\| f^{k+1}-f^k\|_2$ 를 계산한다. $r_k<\varepsilon$ 혹은 $k+1>K$ 라면 반복을 중지하고 그렇지 않으면 다음걸음으로 간다.

걸음 3 d 를 갱신한다.

$$q^{k+1}+\lambda_2(d^{k+1}-Df^{k+1}-s^k)=0,\ q^{k+1}\in\partial\| d^{k+1}\|_1 \tag{16}$$

또는

$$d^{k+1}=\max\left(h^k-\frac{1}{\lambda_2},\ 0\right)\frac{Df^{k+1}+s^k}{h^k}$$

여기서 $h^k=\|\nabla f^{k+1}+s^k\|_2$ 이다.

걸음 4 s 를 갱신한다.

$$s^{k+1}=s^k+(Df^{k+1}-d^{k+1}) \tag{17}$$

걸음 5 b 를 갱신한다.

$$b^{k+1}=b^k+(Pf^{k+1}-g) \tag{18}$$

걸음 6 $k:=k+1$ 로 놓고 걸음 2에로 되돌아간다.

2) 선형화된 분할브레그만방법의 수렴성

다음의 정리는 적당한 조건하에서 선형화된 분할브레그만방법의 수렴성을 준다.

정리 $J(f)=\| Df\|_1$ 이라고 하자. 문제 (1)의 적어도 하나의 풀이 f^* 이 존재한다고 가정하고 $\lambda_1>0,\ \lambda_2>0,\ \beta_1>0,\ \beta_2>0$ 이며

$$I-\frac{\lambda_1}{\omega_1+\omega_2}P^{\mathrm{T}}P-\frac{\lambda_2}{\omega_1+\omega_2}D^{\mathrm{T}}D$$

가 정의정값이라고 하자. 여기서 $\omega_1=\lambda_1/\beta_1,\ \omega_2=\lambda_2/\beta_2$ 이다.

이때

$$\lim_{k\to\infty}\| Pf^k-g\|_2=0,\ \lim_{k\to\infty}\| Df^k\|_1=\| Df^*\|_1 \tag{19}$$

이 성립한다. 더우기 문제 (1)이 유일한 풀이를 가진다면 다음식이 성립된다.

$$\lim_{k\to\infty}\| f^k-f^*\|_2=0 \tag{20}$$

증명 라그랑쥬벡토르가 w 인 문제 (1)의 라그랑쥬형식은

$$L(f,\ w)=\| Df\|_1+\langle w,\ Pf-g\rangle$$

이다. 가정에 의해 문제 (1)이 적어도 하나의 풀이 f^* 이 존재하므로 쿤－타커조건으로부터 적당한 w^* 이 있어서

$$\begin{cases}0=D^{\mathrm{T}}q^*+P^{\mathrm{T}}w^*\ \ (q^*\in\partial\| d^*\|_1,\ d^*=Df^*)\\ Pf^*=g\end{cases} \tag{21}$$

이다.[2]

$s^* = q^*/\lambda_2$, $b^* = w^*/\lambda_1$ 을 도입하자. 이때 식 (21)로부터

$$\begin{aligned} & f^* = f^* - \frac{\lambda_1 P^{\mathrm{T}}(Pf^* - g + b^*) + \lambda_2 D^{\mathrm{T}}(Df^* - d^* + s^*)}{\omega_1 + \omega_2} \\ & q^* + \lambda_2(d^* - Df^* - s^*) = 0 \\ & s^* = s^* + (Df^* - d^*) \\ & b^* = b^* + (Pf^* - g) \end{aligned} \tag{22}$$

를 얻을수 있다. 이것은 f^*, d^*, s^*, b^*이 식 (15)−(18)의 부동점이라는것을 의미한다.

$$f_e^k := f^k - f^*,\ d_e^k := d^k - d^*,\ s_e^k := s^k - s^*,\ b_e^k := b^k - b^*$$

로 표시하자.

$B = (I - \lambda_1 \alpha P^{\mathrm{T}} P - \lambda_2 \alpha D^{\mathrm{T}} D)^{1/2}$ 을 도입하면 식 (15)−(17)로부터

$$\begin{aligned} & \frac{1}{2\alpha}\sum_{k=0}^{K} \| B(f_e^{k+1} - f_e^k)\|^2 + \sum_{k=0}^{K} < q_e^{k+1},\ d_e^{k+1} > + \frac{\lambda_1}{2}\sum_{k=0}^{K} \| Pf_e^{k+1}\|^2 + \frac{\lambda_2}{2}\sum_{k=0}^{K} \| Df_e^{k+1} - d_e^k\|^2 + \\ & + \frac{1}{2\alpha}\| Bf_e^{K+1}\|^2 + \frac{\lambda_1}{2}\| b_e^{K+1}\|^2 + \frac{\lambda_2}{2}\| d_e^{K+1}\|^2 + \frac{\lambda_2}{2}\| s_e^{K+1}\|^2 = \\ & = \frac{1}{2\alpha}\| Bf_e^0\|^2 + \frac{\lambda_1}{2}\| b_e^0\|^2 + \frac{\lambda_2}{2}\| d_e^0\|^2 + \frac{\lambda_2}{2}\| s_e^0\|^2 < C \end{aligned} \tag{23}$$

를 얻을수 있다. 여기서 K는 임의의 정의 옹근수이고 C는 K에 무관계한 상수이다.

J가 불룩함수이면

$$B_J^q(u,\ v) + B_J^p(v,\ u) = < p - q,\ u - v > \ge 0 \ \ (p \in \partial J(u),\ q \in \partial J(v))$$

이다. 이로부터

$$< q_e^{k+1},\ d_e^{k+1} > = < q^{k+1} - q^*,\ d^{k+1} - d^* > = B_{\|\cdot\|_1}^{q^*}(d^{k+1},\ d^*) + D_{\|\cdot\|_1}^{q^{k+1}}(d^*,\ d^{k+1}) \ge 0$$

이고 식 (23)으로부터 다음식이 성립된다.

$$\lim_{k\to\infty}\| Pf^k - g\| = 0,\ \lim_{k\to\infty}\| B(f^k - f^{k-1})\| = 0,\ \lim_{k\to\infty}\| Df^k - d^{k-1}\| = 0$$

따라서 식 (19)의 첫 식이 증명되였다.

식 (23)으로부터 다음식이 얻어진다.

$$\lim_{k\to\infty}(\| d^k\|_1 - \| d^*\|_1 - < q^*,\ d^k - d^* >) = 0$$

$\lim_{k\to\infty}\| Df^k - d^{k-1}\| = 0$ 이므로 $\lim_{k\to\infty}\| Df^k\| = \lim_{k\to\infty}\| d^{k-1}\| = 0$ 이고 따라서

$$\lim_{k\to\infty}(\| Df^{k+1}\|_1 - \| Df^*\|_1 - < q^*,\ Df^{k+1} - Df^* >) = 0$$

이 성립된다.

한편 $Pf^k \to Pf^* = g\ \ (k \to \infty)$ 이므로

$$\lim_{k\to\infty}(\| Df^{k+1}\|_1 - \| Df^*\|_1) = 0$$

을 얻을수 있다. 이로부터 식 (19)의 두번째 식이 증명되였다.

이제는 문제 (1)이 유일한 풀이를 가진다고 하고 식 (20)을 증명하자.

식 (20)이 성립하지 않는다고 가정하면 어떤 정수 ε에 대하여 $\| f^{k_i} - f^* \|_2 > \varepsilon,\ \forall k_i > 0$ 인 부분렬 $\{f^{k_i}\}$가 존재한다. $\tilde{f}^{k_i}$가 구면 $\{f \mid \| f - f^* \|_2 = \varepsilon\}$과 f^*로부터 f^{k_i}까지의 선분의 사귐점이라고 하자. 이때 $\tilde{f}^{k_i} = tf^* + (1-t) f^{k_i}$ 인 $t \in (0,\ 1)$이 유일하게 존재한다.

$J(f) = \| Df \|_1$이고 $\hat{f} = \arg\min_f \{J(f) \mid \| f - f^* \|_2 = \varepsilon\}$이라고 하자. J가 불룩이고 f^*이 식 (14)의 유일한 풀이이므로

$$J(f^*) < J(\hat{f}) \le J(\tilde{f}^{k_i}) = J(tf^* + (1-t) f^{k_i}) \le tJ(f^*) + (1-t) J(f^{k_i}) < J(f^{k_i})$$

이고 극한을 취하면

$$J(f^*) < J(\hat{f}) \le \lim_{k \to \infty} J(f^{k_i}) = J(f^*)$$

이 얻어진다. 이것은 모순이고 따라서 식 (20)도 증명되였다.(증명끝)

주의 1 선형화된 분할브레그만방법에서 f^{k+1}과 d^{k+1}의 갱신순서를 바꾸어도 같은 론의로 수렴성을 증명할수 있다.

주의 2 정리 1의 증명과정에서 α의 구체적인 정의식은 어디에서도 리용되지 않았으므로 정리는 결국 그라디엔트하강분할브레그만방법의 수렴성정리로도 된다.

참 고 문 헌

[1] J. F. Cai et al.; SIAM Interdisciplinary Journal, 8, 2, 337, 2009.
[2] T. Goldstein, S. Osher; SIAM Journal on Imaging Sciences, 2, 323, 2009.
[3] R. Jia et al.; Appl. Comput. Harmon. Anal., 27, 3, 367, 2009.
[4] S. Osher et al.; Simul., 4, 2, 460, 2005.
[5] S. Osher et al.; Simul., 4, 40, 2005.
[6] W. Yin et al.; SIAM J. Imaging Sci., 1, 1, 143, 2008.

주체106(2017)년 8월 10일 원고접수

A Linearized Split Bregman Method

Kim Jong Chol, Ri Kum Hyok

We propose a linearized split Bregman method combining with the split Bregman method and the linearized Bregman method for solving linear inverse problems, and prove its convergence.

We also present a gradient descent split Bregman method, and prove that it is a special case of the linearized split Bregman method.

Key words: linear inverse problem, split Bregman method, linearized Bregman method