Acta Mathematicae Applicatae Sinica, English Series
Vol. 30, No. 1 (2014) 179–186
DOI: 10.1007/s10255-014-0276-0
http://www.ApplMath.com.cn & www.SpringerLink.com

Bregman Iterative Model Using the G-norm

Yu-ying SHI[1], Xiao-zhong YANG[1], Yong-gui ZHU[3]

[1]Department of Mathematics and Physics, North China Electric Power University, Beijing 102206, China
(E-mail: yyshi@amss.ac.cn, xzy1018@hotmail.com)
[2]School of Science, Communication University of China, Beijing, 100024 China (E-mail: ygzhu@cuc.edu.cn)

Abstract In this paper, we analyze the Bregman iterative model using the G-norm. Firstly, we show the convergence of the iterative model. Secondly, using the source condition and the symmetric Bregman distance, we consider the error estimations between the iterates and the exact image both in the case of clean and noisy data. The results show that the Bregman iterative model using the G-norm has the similar good properties as the Bregman iterative model using the L^2-norm.

Keywords Bregman distance, image restoration, total variation, error estimation, iterative regularization
2000 MR Subject Classification 35Q80

1 Introduction

The restoration problem is modeled by $f = Ku + v$, here, K is blurring operator (usually bounded linear operator), u is cartoon part and v is texture and/or noise. The total variation (TV) denoising models are based on a variational problem with constraints using the TV norm ($TV(u) = \int_\Omega |\nabla u| dx$) as a nonlinear nondifferentiable functional. The popular ROF (Rudin-Osher-Fatemi) (i.e., TV-L^2) model[6,11,12] is one of the most famous PDE-based image denoising models in image processing. The TV-L^1 model[1,5,7] has also been used for denoising and cartoon-texture decomposition. Buades summarized different models based on total variation in [2].

Y. Meyer[6] pointed out the crucial role played by a certain functional Banach space, called the space of texture and denoted by G. The TV-G model is:

$$\min_{u\in BV(\Omega)} TV(u) + \frac{\lambda}{2}||f - Ku||_G, \qquad \lambda > 0, \tag{1.1}$$

where $BV(\Omega)$ denotes the space of functions with bounded variation on Ω and $\lambda > 0$ is the regularization parameter that determines the balance between goodness fit to the original image and the amount of regularization done to the original image f in order to produce the approximation u. Because the G-norm is difficult to compute using the usual Euler-Lagrange equation in numerical experiments, the approximation is proposed by Osher, Solé and Vese(OSV)[10]:

$$\min_{u\in BV(\Omega)} TV(u) + \frac{\lambda}{2}||\nabla\triangle^{-1}(f - Ku)||^2, \qquad \lambda > 0, \tag{1.2}$$

where the G-norm used in the TV-G model (1.1) is replaced by an $(H^{-1})^2$ fitting term. The flexibility of the Bregman distances is attractive for achieving certain imaging tasks such as preservation of edges[3].

Manuscript received March 23, 2009. Revised March 28, 2011.
Supported by the National Natural Science Foundation of China (No. 10726035, 10771065, 10801049) and Foundation of North China Electric Power University.

To improve the restoration effect, S. Osher et al.[8] proposed an iterative regularization procedure based on Bregman distance

$$D^p(u,v) \equiv J(u) - J(v) - \langle p, u - v\rangle$$

($\langle\cdot,\cdot\rangle$ denotes the usual duality product) as follows:

Algorithm. *Initializing u_0 and p_0, and for $k = 1, 2, \cdots$,*

$$u_k = \arg\min_u \big\{Q_k(u) := H(u,f) + D^{p_{k-1}}(u, u_{k-1})\big\}, \tag{1.3}$$

where $J(\cdot)$ and $H(\cdot,f)$ are convex non-negative regularization functionals and p_k is the subgradient of $J(u_k)$ with $\partial J(v) = \big\{p : D^p(u,v) \geq 0,\ \forall\, u \in BV(\Omega)\big\}$.

The iterative regularization model is called as ITV-L^2 model[8] with $J(u) = TV(u),\ H(u,f) = \frac{\lambda}{2}||f - Ku||^2$. Similar as the OSV model (1.2), the ITV-G model is with the same $J(u)$ and $H(u,f) = \frac{\lambda}{2}||\nabla\triangle^{-1}(f - Ku)||^2$.

M. Burger et al.[3] considered the error estimations for the TV-L^2 model and the ITV-L^2 model based on the generalized Bregman distances. Motivated by the analysis in [3,4], we first consider the convergence theorem for the ITV-G model, then the error estimations with noise or not will be shown.

The rest of the paper is organized as follows. In Section 2, we give the convergence analysis of the ITV-G model. In Section 3, we consider the error estimations between the iterates and the exact image both in the case of clean and noisy data, and derive the convergence rate. Some concluding remarks are given in Section 4.

2 Convergence Analysis

Now we give the well-definedness of the **Algorithm** in the above section for the ITV-G model.

Theorem 2.1. *Assume $J(u) = TV(u)$, $H(u,f) = \frac{\lambda}{2}||\nabla\triangle^{-1}(f - Ku)||^2$ and let $u_0 = 0$, $p_0 = 0$ for the iterates (1.3). Then for each $k \in N$ there exists a minimizer u_k of $Q_k(u)$, $p_k \in \partial J(u_k)$ and $q_k \in \partial_u H(u_k, f) = \lambda K^*\triangle^{-1}(f - Ku_k)$ such that $p_k + q_k = p_{k-1}$. If K has no nullspace, then the minimizer u_k is unique.*

Proof. We prove the above result by induction. For $k = 1$, we have $u_1 = \arg\min_u\{Q_1(u) = J(u) + H(u,f)\}$ and the existence of minimizers as well as the optimality condition $p_1 + q_1 = p_0 = 0$ is well-known. Moreover, with $r_1 = \lambda\triangle^{-1}(Ku_1 - f)$ we have $p_1 = K^*r_1$.

Now we proceed from $k-1$ to k, and assume that

$$p_{k-1} = K^*r_{k-1}. \tag{2.1}$$

Under the above assumption, the functional

$$Q_k(u) : u \to J(u) - J(u_{k-1}) + H(u,f) - \langle p_{k-1}, u - u_{k-1}\rangle$$

is weak-* lower semicontinuous and it is bounded below by $H(u,f)$ due to the properties of the subgradients. Moreover, we can estimate by (2.1)

$$\begin{aligned} Q_k(u) =& J(u) - J(u_{k-1}) - \langle r_{k-1}, Ku - Ku_{k-1}\rangle + \lambda||\nabla\triangle^{-1}(Ku - f)||^2 \\ =& J(u) - J(u_{k-1}) - \langle r_{k-1}, Ku - f\rangle - \langle r_{k-1}, f - Ku_{k-1}\rangle \\ &+ \lambda||\nabla\triangle^{-1}(Ku - f)||^2. \end{aligned} \tag{2.2}$$

Since

$$\langle r_{k-1}, Ku - f\rangle = -\langle\nabla r_{k-1}, \nabla\triangle^{-1}(Ku - f)\rangle, \tag{2.3}$$

we have

$$
\begin{aligned}
&-\langle r_{k-1}, Ku-f\rangle+\lambda||\nabla\triangle^{-1}(Ku-f)||^2\\
=&\lambda||\nabla\triangle^{-1}(Ku-f)+\frac{1}{2\lambda}\nabla r_{k-1}||^2-\frac{1}{4\lambda}||\nabla r_{k-1}||^2\\
\geq&-\frac{1}{4\lambda}||\nabla r_{k-1}||^2.
\end{aligned}
\tag{2.4}
$$

By (2.2) and (2.4), we get

$$
Q_k(u)\geq J(u)-J(u_{k-1})-\langle r_{k-1}, f-Ku_{k-1}\rangle-\frac{1}{4\lambda}||\nabla r_{k-1}||^2. \tag{2.5}
$$

Since only the first term on the right-hand side of this inequality is not constant, boundedness of $Q_k(u)$ implies of boundedness of $J(u)$. This shows that the level sets of $Q_k(u)$ are bounded in the norm of $BV(\Omega)$, and therefore they are weak*-compact. Hence, there exists a minimizer of $Q_k(u)$ due to the fundamental theorem of optimization. Moreover, if K has no nullspace, the strict convexity of $H(\cdot, f)$ and convexity of the other terms imply the strict convexity of $Q_k(u)$, and therefore the minimizer is unique. Since $p_{k-1}\in\partial J(u_k)+\partial_u H(u_k, f)$, which yields the existence of $p_k\in\partial J(u_k)$ and $q_k=\partial_u H(u_k, f)=\lambda K^*\triangle^{-1}(f-Ku_k)$ satisfying $p_{k-1}=p_k+q_k$. □

We recall below several intermediate results as well as some of the main results shown in [8].

Proposition 2.1. *Under the above assumptions, the sequence $H(u_k, f)$ obtained from the iterates (1.3) is monotonically non-increasing, one even has*

$$
H(u_k, f)\leq H(u_k, f)+D^{p_{k-1}}(u_k, u_{k-1})\leq H(u_{k-1}, f). \tag{2.6}
$$

Moreover, let u be such that $J(u)<\infty$, then one has

$$
D^{p_k}(u, u_k)+D^{p_{k-1}}(u_k, u_{k-1})+H(u_k, f)\leq H(u, f)+D^{p_{k-1}}(u, u_{k-1}). \tag{2.7}
$$

Theorem 2.2 (Exact data). *Assume $K\widetilde{u}=f,\ u\in BV(\Omega)$ and $J(\widetilde{u})<\infty$, then*

$$
H(u_k, f)\leq H(\widetilde{u}, f)+\frac{J(\widetilde{u})}{k} \tag{2.8}
$$

and, in particular, u_k is a minimizing sequence.

Moreover, u_k has a weak-convergent subsequence in $BV(\Omega)$, and the limit of each weak*-convergent subsequence is a solution of $Ku=f$. If $\widetilde{u}$ is the unique solution of $Ku=f$, then $u_k\to\widetilde{u}$ in the weak*-topology in $BV(\Omega)$.*

The similar processes of proof are in [8]. Next, we consider the noisy case.

Theorem 2.3 (Noisy data). *Assume $K\widetilde{u}=f,\ u\in BV(\Omega)$,*

$$
||\nabla\triangle^{-1}(f^\delta-f)||\leq\delta \tag{2.9}
$$

and $J(\widetilde{u})<\infty$, then

$$
D^{p_k}(\widetilde{u}, u_k)+\sum_{j=1}^{k}[D^{p_{j-1}}(u_j, u_{j-1})+H(u_j, f^\delta)]\leq\frac{\lambda k\delta^2}{2}+J(\widetilde{u}). \tag{2.10}
$$

Proof. The sequence $H(u_k, f^\delta)$ obtained from the iterations (1.3) is monotonically non-increasing, (2.6) and (2.7) hold when f is replaced by f^δ. Using (2.9), we have

$$D^{p_j}(\widetilde{u}, u_j) + D^{p_{j-1}}(u_j, u_{j-1}) + H(u_j, f^\delta) \leq \frac{\lambda\delta^2}{2} + D^{p_{j-1}}(\widetilde{u}, u_{j-1}), \qquad \forall j \in N. \tag{2.11}$$

Summing (2.11) up from 1 to k, we arrive at (2.10).

We get the similar conclusions as the conclusions of the ITV-L^2 model[8]. It should be noticed that we use $||\nabla\triangle^{-1}(f^\delta - f)|| \leq \delta$ to replace $||f^\delta - f|| \leq \delta$ that is often used in ITV-L^2/TV-L^2 model. □

3 Error Estimation

In the following, we discuss the basic ideas needed for the error estimation. The so-called source condition[3]:

(SC) There exists $\xi \in \partial J(\widetilde{u})$ such that $\xi = K^* q$ for a source element $q \in L^2(\Omega)$.

Since the Bregman distance is not symmetric in general, which can be remedied partly by using the symmetric Bregman distance[3]:

$$D^{\text{symm}}(u_1, u_2) = \langle u_1 - u_2, p_1 - p_2 \rangle = D^{p_1}(u_2, u_1) + D^{p_2}(u_1, u_2), \qquad p_i \in \partial J(u_i). \tag{3.1}$$

The symmetric Bregman distance depends on the specific selection of the subgradients p_i, when the subgradients are not unique.

Now we derive the error estimation between a solution of the equation $Ku = f$ and the iterates u_k produced by the ITV-G model.

Theorem 3.1 (Exact data). *Let $\widetilde{u} \in BV(\Omega)$ be a solution of $Ku = f$ and assume that the source condition (SC) is satisfied. Then*

$$D^{p_k}(\widetilde{u}, u_k) \leq \frac{||\nabla q||^2}{2\lambda k}. \tag{3.2}$$

Proof. Let

$$x_k = \lambda \sum_{j=1}^{k} \triangle^{-1}(Ku_j - f). \tag{3.3}$$

According to Theorem 2.1, $q_k = \partial_u H(u_k, f) = \lambda K^* \triangle^{-1}(f - Ku_k)$ satisfying $p_{k-1} = p_k + q_k$ and $p_0 = 0$. The following equalities are obtained:

$$p_k = -\sum_{j=1}^{k} q_j = -\lambda K^* \sum_{j=1}^{k} \triangle^{-1}(f - Ku_j) = K^* x_k \tag{3.4}$$

and

$$x_{j-1} - x_j = \lambda\triangle^{-1}(f - Ku_j). \tag{3.5}$$

From the definition of symmetric Bregman distance (3.1), we get for any $j \in N$, $0 < j \leq k$,

$$\lambda D^{p_j}(\widetilde{u}, u_j) + \lambda D^{\xi}(u_j, \widetilde{u}) = \lambda\langle p_j - \xi, u_j - \widetilde{u} \rangle. \tag{3.6}$$

The above relation, together with (3.4), (3.5) and (SC), shows

$$\begin{aligned}\lambda D^{p_j}(\widetilde{u}, u_j) + \lambda D^{\xi}(u_j, \widetilde{u}) =& \langle K^* x_j - K^* q, \lambda(u_j - \widetilde{u})\rangle \\ =& \langle x_j - q, \lambda(Ku_j - f)\rangle = \langle x_j - q, \triangle(x_j - x_{j-1})\rangle \\ =& \langle \nabla(x_j - q), \nabla(x_{j-1} - x_j)\rangle. \end{aligned} \tag{3.7}$$

It is obvious that

$$\begin{aligned}&\langle \nabla(x_j - q), \nabla(x_{j-1} - x_j)\rangle \\ =& \langle \nabla(x_j - q), -\nabla(x_j - q) + \nabla(x_{j-1} - q)\rangle \\ =& \frac{1}{2}||\nabla(x_{j-1} - q)||^2 - \frac{1}{2}||\nabla(x_j - q)||^2 - \frac{1}{2}||\nabla x_{j-1} - \nabla x_j||^2 \\ \leq& \frac{1}{2}||\nabla(x_{j-1} - q)||^2 - \frac{1}{2}||\nabla(x_j - q)||^2 - \frac{1}{2}||\lambda \nabla \triangle^{-1}(f - Ku_j)||^2 \\ \leq& \frac{1}{2}||\nabla(x_{j-1} - q)||^2 - \frac{1}{2}||\nabla(x_j - q)||^2. \end{aligned} \tag{3.8}$$

Based on (3.7) and (3.8), we obtain

$$\lambda D^{p_j}(\widetilde{u}, u_j) + \lambda D^{\xi}(u_j, \widetilde{u}) \leq \frac{1}{2}||\nabla(x_{j-1} - q)||^2 - \frac{1}{2}||\nabla(x_j - q)||^2.$$

By summing up the last inequalities from $j = 1$ to k, using the non-negativity of $D^{\xi}(u_j, \widetilde{u})$, and the fact that $D^{p_j}(\widetilde{u}, u_j)$ is non-increasing with respect to j, it follows that

$$D^{p_k}(\widetilde{u}, u_k) \leq \frac{1}{k}\sum_{j=1}^{k} D^{p_j}(\widetilde{u}, u_j) \leq \frac{||\nabla q||^2}{2\lambda k}.$$

□

Suppose that the given noisy data f^{δ} satisfies (2.9). The next result shows that a priori stopping rule $k_*(\delta) \sim \frac{1}{\delta}$ yields semi-convergence of the regularization method.

Proposition 3.1. *Let $\widetilde{u} \in BV(\Omega)$ verify $K\widetilde{u} = f$, and assume that the (SC) and (2.9) are satisfied. Moreover, let the stopping index $k_*(\delta)$ be chosen of order $\frac{1}{\delta}$. Then, $\{J(u_{k_*(\delta)})\}_{\delta}$ is bounded and hence, as $\delta \to 0$, there exists a weak*-convergent subsequence $\{u_{k_*(\delta_n)}\}_n$ in $BV(\Omega)$ whose limit is a solution of $Ku = f$. Moreover, if the solution of the equation is unique, then $u_{k_*(\delta)}$ converges in the weak*-topology to the solution as $\delta \to 0$.*

Proof. The proof follows the pattern of the proof of Theorem 2.2 for the exact data case, but it is provided here for the sake of completeness. From inequality (2.10), we have

$$\begin{aligned}J(\widetilde{u}) + \lambda k\delta^2 \geq& \sum_{j=1}^{k} D^{p_{j-1}}(u_j, u_{j-1}) \\ =& J(u_k) - \langle p_{k-1}, u_k - \widetilde{u}\rangle + \sum_{j=1}^{k-1}\langle p_j - p_{j-1}, u_j - \widetilde{u}\rangle, \end{aligned}$$

and by (3.4),

$$J(\widetilde{u}) + \lambda k\delta^2 \geq J(u_k) + \sum_{j=1}^{k-1}\langle q_j, u_k - \widetilde{u}\rangle - \sum_{j=1}^{k-1}\langle q_j, u_j - \widetilde{u}\rangle$$

$$
\begin{aligned}
&= J(u_k) - \lambda \sum_{j=1}^{k-1} \langle \triangle^{-1}(Ku_j - f^\delta), Ku_k - f \rangle \\
&\quad + \lambda \sum_{j=1}^{k-1} \langle \triangle^{-1}(Ku_j - f^\delta), Ku_j - f \rangle \\
&= J(u_k) + \lambda \sum_{j=1}^{k-1} \langle \nabla \triangle^{-1}(Ku_j - f^\delta), \nabla \triangle^{-1}(Ku_k - f) \rangle \\
&\quad - \lambda \sum_{j=1}^{k-1} \langle \nabla \triangle^{-1}(Ku_j - f^\delta), \nabla \triangle^{-1}(Ku_j - f) \rangle \\
&= J(u_k) + \lambda \sum_{j=1}^{k-1} \langle \nabla \triangle^{-1}(Ku_j - f^\delta), \nabla \triangle^{-1}(Ku_k - f^\delta) \rangle \\
&\quad - \lambda \sum_{j=1}^{k-1} \langle \nabla \triangle^{-1}(Ku_j - f^\delta), \nabla \triangle^{-1}(Ku_j - f^\delta) \rangle.
\end{aligned}
$$

Next we use Cauchy-Schwarz inequality, (2.10) and the inequality $ab \leq \frac{a^2}{2} + \frac{b^2}{2}$ to get

$$
\begin{aligned}
J(\widetilde{u}) + \lambda k \delta^2 \geq & J(u_k) - \frac{3\lambda}{2} \sum_{j=1}^{k-1} ||\nabla \triangle^{-1}(Ku_j - f^\delta)||^2 - \frac{k\lambda}{2} ||\nabla \triangle^{-1}(Ku_k - f^\delta)||^2 \\
\geq & J(u_k) - 4J(\widetilde{u}) - 4\lambda k \delta^2.
\end{aligned}
$$

So we have the following estimate

$$
J(u_k) \leq 5J(\widetilde{u}) + 5\lambda k \delta^2, \tag{3.9}
$$

which further implies that the sequence $\{J(u_{k_*(\delta)})\}_\delta$ is bounded for $\delta > 0$ sufficiently small and for $k_*(\delta) \sim \frac{1}{\delta}$. Thus we get the boundedness of $||u_{k_*(\delta)}||_{BV}$ for any $\delta > 0$. On one hand, since $BV(\Omega)$ is provided with a weak*-topology, we conclude that there is a subsequence $\{u_{k_*(\delta_n)}\}_n$ which converges to a point $\bar{u}$ with respect to that topology. Due to the embedding of $(BV(\Omega), w^*)$ into $(L^2(\Omega), || \cdot ||_2)$ for spatial dimension less or equal two, this subsequence converges to $\bar{u}$ in the $L^2(\Omega)$-norm. Because of the continuity of the operator K on $L^2(\Omega)$, thus $\lim\limits_{n\to\infty} Ku_{k_*(\delta_n)} = K\bar{u}$. On the other hand, we derive from (2.10) the inequality

$$
H(u_{k_*(\delta_n)}, f^\delta) \leq \lambda \delta_n^2 + \frac{J(\widetilde{u})}{k_*(\delta_n)}.
$$

By our special choice of $k_*(\delta_n)$, we obtain $\lim\limits_{n\to\infty} Ku_{k_*(\delta_n)} = f$ and then, $K\overline{u} = f$. □

The error estimates for the noisy data case are established below.

Theorem 3.2 (Noisy data). *Let $\widetilde{u} \in BV(\Omega)$ verify $K\widetilde{u} = f$, and assume that the (SC) and (2.9) are satisfied. Then, the following estimate holds:*

$$
D^{p_k}(\widetilde{u}, u_k) \leq \frac{||\nabla q||^2}{2\lambda k} + \delta ||\nabla q|| + \lambda \delta^2 k, \qquad \forall\, k \in N. \tag{3.10}
$$

Moreover, if a prior choice $k_*(\delta) \sim \frac{1}{\delta}$ is made, then the following convergence rate is obtained

$$
D^{p_{k_*(\delta)}}(\widetilde{u}, u_{k_*(\delta)}) = O(\delta).
$$

Proof. Let

$$x_k = \lambda \sum_{j=1}^{k} \triangle^{-1}(Ku_j - f^\delta). \tag{3.11}$$

From (3.4) and (SC), we get for any positive $j \in N,\ 0 < j \leq k$,

$$\begin{aligned}\lambda D^{p_j}(\widetilde{u}, u_j) + \lambda D^{\xi}(u_j, \widetilde{u}) &= \lambda \langle p_j - \xi, u_j - \widetilde{u} \rangle \\ &= \langle K^* x_j - K^* q, \lambda(u_j - \widetilde{u}) \rangle \\ &= \langle x_j - q, \lambda(K u_j - f) \rangle \\ &= \langle x_j - q, \lambda(K u_j - f^\delta) \rangle + \lambda \langle x_j - q, f^\delta - f \rangle.\end{aligned}$$

The above relation, together with (3.11), leads to

$$\begin{aligned}&\lambda D^{p_j}(\widetilde{u}, u_j) + \lambda D^{\xi}(u_j, \widetilde{u}) \\ =&\langle x_j - q, \triangle(x_j - x_{j-1}) \rangle - \lambda \langle \nabla(x_j - q), \nabla \triangle^{-1}(f^\delta - f) \rangle \\ =&\langle \nabla(x_j - q), \nabla(x_{j-1} - x_j) \rangle - \lambda \langle \nabla(x_j - q), \nabla \triangle^{-1}(f^\delta - f) \rangle \\ \leq&\langle \nabla(x_j - q), \nabla(x_{j-1} - x_j) \rangle + \lambda\delta ||\nabla(x_j - q)||,\end{aligned} \tag{3.12}$$

where we used (2.9) in the last equality. Thus we have

$$\langle \nabla(x_j - q), \nabla(x_j - x_{j-1}) \rangle \leq \lambda\delta ||\nabla(x_j - q)||, \tag{3.13}$$

that is, $||\nabla(x_j - q)|| \leq ||\nabla(x_{j-1} - q)|| + \lambda\delta$. By the method of induction, it follows that

$$||\nabla(x_j - q)|| \leq ||\nabla q|| + \lambda\delta j. \tag{3.14}$$

Combining (3.8) and (3.12)–(3.14), we obtain

$$\lambda D^{p_j}(\widetilde{u}, u_j) \leq \frac{1}{2} ||\nabla(x_{j-1} - q)||^2 - \frac{1}{2} ||\nabla(x_j - q)||^2 + \lambda\delta ||\nabla q|| + \lambda^2 \delta^2 j.$$

Fix a positive $k \in N$ and sum the last inequalities up from 1 to k, we get

$$\lambda \sum_{j=1}^{k} D^{p_j}(\widetilde{u}, u_j) \leq \frac{||\nabla q||^2}{2} + \lambda\delta k ||\nabla q|| + \lambda^2 \delta^2 \frac{k(k+1)}{2} \tag{3.15}$$

and thus,

$$\sum_{j=1}^{k} D^{p_j}(\widetilde{u}, u_j) \leq \frac{||\nabla q||^2}{2\lambda} + \delta k ||\nabla q|| + \lambda \delta^2 \frac{k(k+1)}{2}. \tag{3.16}$$

Note that monotonicity of the sequence $D^{p_j}(\widetilde{u}, u_j)$ is not guaranteed, as in the noise free case. So we employ a monotonicity-like inequality derived from (2.11):

$$D^{p_{j+1}}(\widetilde{u}, u_{j+1}) - D^{p_j}(\widetilde{u}, u_j) \leq \frac{\lambda\delta^2}{2}.$$

Summing up these inequalities up to k yields

$$D^{p_k}(\widetilde{u}, u_k) - D^{p_j}(\widetilde{u}, u_j) \leq \frac{(k-j)\lambda\delta^2}{2}.$$

Summing up again with respect to j up to k implies

$$kD^{p_k}(\widetilde{u},u_k)-\sum_{j=1}^{k}D^{p_j}(\widetilde{u},u_j)\le\frac{\lambda}{2}\Big(k^2-\sum_{j=1}^{k}j\Big)\delta^2,$$

which means

$$kD^{p_k}(\widetilde{u},u_k)\le\sum_{j=1}^{k}D^{p_j}(\widetilde{u},u_j)+\frac{k(k-1)\lambda\delta^2}{4}. \tag{3.17}$$

From (3.16) and (3.17), we have

$$kD^{p_k}(\widetilde{u},u_k)\le\frac{||\nabla q||^2}{2\lambda}+\delta k||\nabla q||+\lambda\delta^2k^2.$$

Thus (3.10) follows immediately. □

4 Concluding Remarks

In this paper, we discuss the iterative regularization method based on the generalized Bregman distance, especially using the G-norm. Firstly, we analyze and show the convergence of the Bregman iterative model (i.e. ITV-G model). Secondly, we consider the error estimations for the ITV-G model based on the source condition and the symmetric Bregman distance. The results illustrate the ITV-G model has the similar good properties with the ITV-L^2 model. It should be noticed that $||f^\delta-f||\le\delta$ (that is often used when analyzing the ITV-L^2 model) is replaced by $||\nabla\triangle^{-1}(f^\delta-f)||\le\delta$ to analyze the ITV-G model.

References

[1] Alliney, S. Digital filters as absolute norm regularizers. *IEEE Trans. Signal Process*, 40: 1548–1562 (1992)

[2] Buades, A., Le, T.M., Morel, J.M., Vese, L.A. Fast cartoon+texture image filters. *IEEE. Trans. Image Process*, 19: 1978–1986 (2010)

[3] Burger, M., Resmerita, E., He. L. Error estimation for Bregman iterations and inverse scale space methods in image restoration. *Computing*, 81: 109–135 (2007)

[4] Burger, M., Osher, S. Convergence rates for convex variational regularization. *Inverse Problems*, 20: 1411–1421 (2004)

[5] Chan, T., Esedoglu, S. Aspects of total regularized L^1 function approximation. *SIAM J. Appl. Math.*, 65: 1817–1837 (2005)

[6] Meyer, Y. Oscillating Patterns in Image Processing and Nonlinear Evolution Equations, AMS, Providence, 2001

[7] Nikolova, M. Minimizers of cost-functions involving nonsmooth data-fidelity terms. *SIAM J. Numer. Anal.*, 40: 965–994 (2002)

[8] Osher, S., Burger, M., Goldfarb, D., Xu, J., Yin, W. An iterative regularization method for total variation based image restoration. *Multiscale Model. Simul.*, 4: 460–489 (2005)

[9] Osher, S., Scherzer., O. G-norm properties of bounded variation regularization. *Comm. Math. Sci.*, 2: 237–254 (2004)

[10] Osher, S., Solé, A., Vese, L. Image decomposition and restoration using total variation minimization and the H^{-1} norm. *Multiscale Modeling and Simulation*, 1: 349–370 (2003)

[11] Rudin, L., Osher, S., Fatemi, E. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60: 259–268 (1992)

[12] Tadmor, E., Nezzar, S., Vese, L. A multiscale image representation using hierarchical (BV,L^2) decompositions. *Multiscale Model. Simul.*, 2: 554–579 (2004)