International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2015): 78.96 | Impact Factor (2015): 6.391

An Incremental Subgradient Method for Optimization Problems for Bibliometric Mapping of Science

Myong-Hyok Jang[1], Chol-Ryong Ri[2]

[1]Pyongsong Industrial University, Pyongsong, D.P.R. of Korea

[2]Library of Science, Pyongyang, D.P.R. of Korea

Abstract: ***In this paper, we studied a solving method for minimization problem arising in bibliometric mapping of science. We proved that the objective function in this problem is weak-convex function, and then suggested incremental subgradient method to solve the problem and analyzed its convergence.***

Keywords: incremental subgradient method, weak-convex function, bibliometric mapping

1. Introduction

Bibliometrics is the scientific field that quantitatively studies all kinds of bibliographic data such as titles, keywords, authors, and cited references of papers and booklets. Bibliometric mapping of science is a quantitative method for visually representing scientific literatures based on bibliographic data[9].

The general aim of a bibliometric map is to provide an overview of the structure of the scientific literatures belonging to a certain domain or a certain topic.

Let n items to be mapped be 1, 2, … , n. The items may be authors, literatures, or technical terms.

The number of concurrence of items i and j is denoted as c_{ij}, $i, j = \overline{1, n}, (i \neq j)$.

Here for any i, j, $c_{ij} = c_{ji}$ and the total number of concurrence of item i is denoted as $c_i = \sum_{j \neq i} c_{ij}$.

Let s_{ij} be a similarity measure of items i and j.

Relatedness, cosine index, inclusion index and Jaccard index can be used as similarity measures [9].

The quantity $s_{ij} = \dfrac{c_{ij}}{c_i c_j}$ is called relatedness between item i and j. For the definitions of other similarity measures, see [9].

For bibliometric mapping in 2-dimensional plane, let $x_i = (x_{i1}, x_{i2})$ be the vector that represents the coordinate of the i^{th} item on the plane.

Bibliometric mapping is to locate the items on plane so that more similar items may be closer geometrically one another but otherwise items farther one another in 2D plane.

The problem of bibliometric mapping can be modelled as the following optimum problem [9].

$$\min V(x_1, \ldots, x_n) = \sum_{i<j} s_{i,j} \left\| x_i - x_j \right\|^2$$
$$s.t.\ \frac{2}{n(n-1)} \sum_{i<j} \left\| x_i - x_j \right\| = 1 \tag{1}$$

The solution to this problem is obtained from the solution of the following unconditioned optimum problem:

$$\min \sigma(x_1, \ldots, x_n) = \sum_{i<j} s_{i,j} \left\| x_i - x_j \right\|^2 - 2 \sum_{i<j} \left\| x_i - x_j \right\| \to \min \tag{2}$$

Let $x^* = (x_1^*, \ldots, x_n^*)$ be the solution of the problem (2) and assume $c = \dfrac{n(n-1)}{2 \sum_{i<j} \left\| x_i^* - x_j^* \right\|}$ Then $U^* = cx^*$ is the solution of the problem (1).

In this paper, we proposed one solving algorithm for the problem (2).

Generally,

$$\min F(x), \quad x \in X \tag{3}$$

gradient or subgradient descent methods are used frequently in solving nonlinear optimization problems. The general forms are as follows:

$$x^{k+1} = \Pi_X (x^k - \alpha_k g^k), \quad k=0, 1, 2, \ldots \tag{4}$$

Here g^k is gradient or subgradient of the objective function F(x) at the point x^k, and Π_X is the projection on X.

Now let objective function is consist of the sum of several functions. That is,

$$F(x) = \sum_{i=1}^{m} f_i(x) \tag{5}$$

In this case, the gradient of F(x) is the sum of the gradients of all member functions $f_i(x)$. The calculations of the gradients or subgradients of F(x) are very expensive when m is large.

For overcoming this difficulty, several researchers suggested the methods using the gradient or subgradients of individual member function f_i instead of g^k g^k in each iteration of algorithm (4) [3-8].These methods are called incremental gradient or incremental subgradient methods.

In the paper[2] author proved the convergence of the incremental subgradient algorithm for minimizing the non-differentiable convex functions. In the paper[6] author proved the convergence of the incremental subgradient algorithm in the problem of convex programming method, Kiwiel[5] proved the convergence of the incremental proximal subgradient method in convex programming problems and Blatt, Hero and Gauchman analyzed the convergence of the method using constant step length in unconditioned minimization problem of convex function. [3]. The authors of [4] suggested ε-subgradient methods and described unified method which analyses the convergence of the incremental subgradient methods.

Summarizing the results in previous literatures, the convergence of incremental gradient or subgradient methods was proved only in convex programming problems.

In this paper, we are going to suggest the incremental subgradient method for minimizing the weak-convex functions, one type of non-convex function. And then we analyse the convergence of it.

The objective function of optimization problem (2) for bibliometric mapping of science has the form of sum of several functions just like (5).

We show the main results of the paper in Section 2.

2. Incremental Subgradient Method for Minimizing Weak-Convex Functions

Definition ([10]): Suppose F(x): $R^n \to R^1$.
The function $F(x)$ is called weak-convex function in R^n if for any vector $x \in R^n$ there exists a nonempty set G(x) of vectors $\widehat{F}_x$ satisfying the following condition;

for any $Z \in R^n$ and $\widehat{F}_x(x) \in G(x)$,

$$F(z)-F(x) \geq (\widehat{F}_x(x), z-x) + r(z,x) \quad (6),$$

where $r(x,y)\|x-y\|^{-1}$ converges to 0 uniformly under $y \to x$ in any closed bounded set.

Any vector $\widehat{F}_x \in G(x)$ is called subgradient of weak-convex function F at x and $G(x)$ is called subdifferetial. Differentiable functions and convex functions are weak-convex functions.

The objective function $\sigma(x_1,\ldots,x_n)$ of the problem (2) is neither a convex function nor a differentiable function. But we can prove that it is a weak-convex function.

Theorem 1. The function $\sigma(x_1,\ldots,x_n)$ is a weak-convex function in R^{2n}.

Proof. That will be sufficient if author proves that the function $\|x_i - x_j\|$ is a weak-convex. The function $\|x_i - x_j\|$ is differentiable at the remaining points except for the points where $x_i \neq x_j$; so it satisfies the weak-convexity condition.

In the case of $x_i = x_j$ let G(x)={0} and $r(z,x)=0$, then the weak- convexity condition is clearly satisfied at x.

From now on we consider the problem (3) with the objective function (5), where we arising X is a convex closed set and each $f_i(x)$ is a weak-convex function.

We consider the following algorithm proposed in [2]:

$$x^{k+1} = z_{m,k} \quad (7a)$$

$$z_{i,k} = \prod\nolimits_x (z_{i-1,k} - \alpha_k g_{i,k}) \quad , g_{i,k} \in \partial f_i(z_{i-1,k})$$

$$i = 1,\cdots,m$$

$$\leq \|z_{i-1,k} - y\|^2 - 2\alpha_k g_{i,k}^T (z_{i-1,k} - y) + \alpha_k^2 C_i^2 \quad (7b)$$

$$z_{0,k} = x^k \quad (7c)$$

Here $\partial f_i(z_{i-1,k})$ is subdifferential at the point $z_{i-1,k}$ of function f_i.
Each step is a subgradient projection iteration for one component function f_i, and there is one step per component function. Thus, an iteration can be viewed as a cycle of m subiterations. If x^k is the vector obtained after k cycles, the vector x^{k+1} is $x^{k+1} = z_{m,k}$ obtained after one further cycle.

The data obtained by (7b) are referred to as the subiterations of the k^{th} cycle.

Assumption 1.

$$\exists C_1, C_2, \ldots, C_m, \quad \|g\| \leq C_i, \forall g \in \partial f_i(x^k) \cup \partial f_i(z_{i-1,k})$$

$$i = \overline{1,m}, k = 0,1,\ldots$$

Since each component f_i is taken as real-value and weak-convex over the entire space R^n, the partial differential $\partial f_i(x)$ is nonempty and compact for all x and i. If the set X is compact or the sequences $\{z_{i,k}\}$ are bounded, then Assumption 1 is satisfied because the set

$\cup_{x\in B}\partial f_i(x)$ is bounded for any bounded set B.

Assumption 2.

$\exists C_1(y),$

$$2\sum_{i=1}^{m}[|r_i(z_{i-1,k},y)|+|r_i(z_{i-1,k},x^k)|]\le C_1(y),\forall k=0,1,...$$

Lemma. If Assumption 1 and 2 are satisfied, then $\forall y\in X,\forall k=0,1,2,...,$

$$\|x^{k+1}-y\|^2\le\|x^k-y\|^2-2\alpha_k(f(x^k)-f(y))+\alpha_k^2C^2+\alpha_kC_1(y)$$

where $C=\sum_{i=1}^{m}C_i$ and C_i is as the same as in assumption 1.

Proof. For any $\forall y\in X$, from the property of projection, and weak-convexity of f_i of Assumption 1 and 2,

$$\|z_{i-1,k}-y\|^2=\|P_X(z_{i-1,k}-\alpha_k g_{i,k})-y\|^2\le$$

$$\le\|z_{i-1,k}-\alpha_k g_{i,k}-y\|^2\le\|z_{i-1,k}-y\|^2-$$

$$-2\alpha_k f_i(z_{i-1,k})-f_i(y)-r_i(z_{i-1,k},y)+\alpha_k^2C_i^2,$$

$$,\forall i,k$$

By adding this inequalities in relation to $i=\overline{1,m}$, we lets have, for any $y\in X,\ k=0,1,2,...,$.

$$\|x^{k+1}-y\|^2\le\|x^k-y\|^2-2\alpha_k\sum_{i=1}^{m}(f_i(z_{i-1,k})-$$

$$(f_i(z_{i-1,k})-f_i(y))-2\alpha_k\sum_{i=1}^{m}r_i(z_{i-1,k},y)+\alpha_k^2\sum_{i=1}^{m}C_i^2=$$

$$=\|x^k-y\|^2-2\alpha_k[f(x^k)-f(y)+$$

$$+\sum_{i=1}^{m}(f_i(z_{i-1,k})-f_i(x^k))]-$$

$$-2\alpha_k\sum_{i=1}^{m}r_i(z_{i-1,k},y)+\alpha_k^2\sum_{i=1}^{m}C_i^2$$

$$\|x^k-y\|^2-2\alpha_k(f(x^k)-f(y))+$$

$$2\alpha_k\sum_{i=1}^{m}[g_{i-1,k}^T(z_{i-1,k}-x^k)-r_i(z_{i-1,k},x^k)]-$$

$$-2\alpha_k\sum_{i=1}^{m}r_i(z_{i-1,k},y)+\alpha_k^2\sum_{i=1}^{m}C_i^2\le$$

$$\le\|x^k-y\|^2-2\alpha_k(f(x^k)-f(y))+$$

$$+2\alpha_k\sum_{i=1}^{m}C_i\|z_{i-1,k}-x^k\|-2\alpha_k\sum_{i=1}^{m}[r_i(z_{i-1,k},y)+$$

$$+r_i(z_{i-1,k},x^k)]+\alpha_k^2\sum_{i=1}^{m}C_i^2\le$$

$$\le\|x^k-y\|^2-2\alpha_k(f(x^k)-f(y))+$$

$$+\alpha_k^2[2\sum_{i=2}^{m}C_i(\sum_{j=1}^{i-1}C_j)+\sum_{i=1}^{m}C_i^2]-$$

$$-2\alpha_k\sum_{i=1}^{m}[r_i(z_{i-1,k},y)+r_i(z_{i-1,k},x^k)]\le$$

$$\le\|x^k-y\|^2-2\alpha_k(f(x^k)-f(y))+\alpha_k^2C^2+$$

$$+2\alpha_k\sum_{i=1}^{m}[|r_i(z_{i-1,k},y)|+|r_i(z_{i-1,k},x^k)|]\le$$

$$\le\|x^k-y\|^2-2\alpha_k(f(x^k)-f(y))+\alpha_k^2C^2+\alpha_kC_1(y)$$

where $g_{i-1,k}\in\partial f_i(z_{i-1,k})$. As showing above we used the facts that X is convex closed set and the inequality $\|\Pi_X(a)-\Pi_X(b)\|\le\|a-b\|,\ \forall a,b\in R^n$ holds.

Theorem 2. Let assumption 1 and 2 be satisfied and $\alpha_k\equiv\alpha,\forall k$.

Suppose that x^* is the solution of the problem (3), (5), then

$$\liminf_{k\to\infty}f(x^k)\le f^*+\frac{\alpha C^2}{2}+\frac{C_1(x^*)}{2}.$$

Proof. Suppose that the result does not hold. Then there exists an ε>0 so that

$$\liminf_{k\to\infty}f(x^k)>f^*+\frac{\alpha C^2}{2}+\frac{C_1(x^*)}{2}+2\varepsilon\ .$$

Since x^* is the global minimum of this problem

$$\liminf_{k\to\infty}f(x^k)>f(x^*)+\frac{\alpha C^2}{2}+\frac{C_1(x^*)}{2}+2\varepsilon\ (8)$$

On the other hand, as there is number k_0 for all $k\ge k_0$

$$f(x^k)>\liminf_{k\to\infty}f(x^k)-\varepsilon\ \ (9)$$

hold.

By adding inequalities (8) and (9), for all $k\ge k_0$ we have

$$f(x^k)-f(x^*)\ge\frac{\alpha C^2}{2}+\frac{C_1(x^*)}{2}+\varepsilon$$

In above lemma, if x^* is used instead of y, for any $k\ge k_0$,

$$\|x^{k+1}-x^*\|^2\le\|x^k-x^*\|^2-2\alpha(f(x^k)-f(x^*))+$$

$$+\alpha^2C^2+\alpha C_1(x^*)\le\|x^k-x^*\|^2+\alpha(-\alpha C^2-$$

$$-C_1(x^*)-2\varepsilon)++\alpha^2C^2+\alpha C_1(x^*)=$$

$$=\|x^k-x^*\|^2-2\alpha\varepsilon$$

Therefore,

$$\|x^{k+1}-x^*\|^2\le\|x^k-x^*\|^2-2\alpha\varepsilon\le\|x^{k-1}-x^*\|^2-4\alpha\varepsilon\le$$

$$\le...\le\|x^{k_0}-x^*\|^2-2(k+1-k_0)\alpha\varepsilon$$

If k is large enough, then right side of this inequality becomes negative. This is contradiction.

The following theorem shows the result in the case that step length converges to 0.

Theorem 3. Let the assumption 1 and 2 be satisfied . If for step lengths α_k the condition

$$\alpha_k > 0, \lim_{k\to\infty} \alpha_k = 0, \sum_{k=0}^{\infty} \alpha_k = \infty$$

hold, then

$$\liminf_{k\to\infty} f(x^k) \le f^* + \frac{c_1(x^*)}{2} .$$

3. Conclusions

We have analyzed convergence properties of an incremental subgradient method for optimization problems with weak convex objective functions. The methods that employ the constant and the dynamic step size rules are analyzed.

Several incremental gradient, subgradient methods and incremental proximal methods [2] for optimization problems with weak convex objective functions remains open.

References

[1] Bertsekas D. P., A new class of incremental gradient methods for least squares problems, SIAM J. Optim., 7, 913–926 (1997)
[2] Bertsekas D. P., Incremental proximal method for large scale convex optimiza- tion,Math program.,Ser.B,129,163-195(2011)
[3] Blatt D.,Hero A.O.,Gauchman H., A convergent incremental gradient method with a constant stepsize. SIAM J.Optim.18, 29-51 (2008)
[4] Helou E.S., De Pierro A.R., Incremental subgradients for constrained convex optimization, a unified framework and new methods. SIAM J. Optim. 20, 1547–1572 (2009)
[5] Kiwiel K.C., Convergence of approximate and incremental subgradient methods for convex optimization. SIAM J. Optim. 14, 807–840 (2004)
[6] Nedic´A., Bertsekas D.P., Incremental subgradient methods for nondifferentiable optimization. SIAM J. Optim. 12, 109–138 (2001).
[7] Solodov M.V. Incremental gradient algorithms with stepsizes bounded awey from zero. Compt. Optim. Appl. 11, 28-35 (1998).
[8] Tseng P.Incremental gradient(projection) method with momentum term and adaptive stepsize rule, SIAM J.Optim.8, 506-531 (1998).
[9] Van Eck N.J., Methodological advances in bibliometric mapping of science, Springer (2011).
[10] Шор Н. Ж., Методы стохастического программирования. М.Мир (1976).