Finite Termination by using the Asymptotic Dual for Dynamic Bundle Method

Paras Bhatnagar
Department of Mathematics,
G.L. Bajaj IT, Gr. Noida, India

Shipra Kaushik
Department of Mathematics,
CMJ University Shillong

Ashok Kumar Vasishtha
Department of Mathematics,
CMJ University Shillong

ABSTRACT

This research paper deals with the question of finite termination of the Algorithm for Dynamic bundle method. For a polyhedral dual function *f* , if the stopping parameter is set to *tol* = 0, and the bundle management is either "no bundle deletion" or "bundle selection", we provide a positive answer for that question.

Keywords

Polyhedral dual function, Dynamic bundle method, minimization problem.

I. INTRODUCTION

We proceed somewhat similarly to Kiwiel.[1], we will show finite termination by using the asymptotic dual results to obtain a contradiction.

Our main assumption is that there exists a finite number *q* of primal points,

$$\{p^1, p^2, ..., p^q\} \subset Q,$$

such that the dual function can be written as

$$f(x) = \max_{i \le q} \left\{ C(p^i) - \left\langle g(p^i), x \right\rangle \right\}, \quad (1)$$

i.e., the dual function is polyhedral.

Although this assumption is made on the dual function, there are many primal conditions that ensure the form (1) for *f* .

Condition (1) implies that at each given x^i there are at most *q* different maximizes p^i as in (4), yielding a finitely generated sub differential

$$\partial f(x^i) = conv\left\{ -g(p^i) : p^i \text{ and } i \le q \right\},$$

Likewise, bundle elements corresponding to past dual evaluations, i.e,

$\left(C_i = C(p^i) : p^i \in Q \right)$ where $i \le q$,

can only take a finite number of different values. This is not the case for aggregate elements, which can take an infinite number of values, simply because they have the expression

$$\left(\hat{C} = \sum_i \alpha_i C(p^i), \hat{\pi} = \sum_i \alpha_i p^i \right),$$

where $\hat{\pi} \in conv\ Q$. This is the underlying reason why we in our next result we cannot handle the "bundle compression" strategy.

Theorem: Suppose the primal problem

$$\begin{cases} \max\limits_{p} \ C(p) \\ p \in Q \subset IR^p \\ g_j(p) \le 0,\ j \in L := \{1, ..., n\}, \end{cases} \quad (2)$$

Satisfies either,

For all $d \ge 0$ $\inf\limits_{p \in Q} \left\langle g(p), d \right\rangle \le 0.$ (3)

or $\sum_{i \le r} \tilde{\alpha}_i g_j(\tilde{p}^i)$ for all $j = 1, ..., n$ (4)

with g affine, conv Q compact, and a dual function of the form (1). Consider Algorithm applied to the minimization problem

$$\min_{x \ge 0} f(x), \text{where } f(x) := \max_{p \in Q} \left\{ C(p) - \sum_{j \in L} g_j(p) x_j \right\} \quad (5)$$

with separation procedure satisfying, $g_l(p) \le 0$. Suppose *tol* = 0 and that Step 5 of the algorithm (Choose a reduced bundle B_{red}; Define $B_{l+1} := B_{red} \bigcup \left\{ (C_l, p^l) \right\}$), always sets the bundle management strategy to be either "no bundle deletion" or "bundle selection". If at null steps $\mu_l = \mu_{k(l)}$, while at serious steps $\mu_{l_k} \le \mu_{\max}$, then the algorithm stops after a finite number of iterations having found a primal convex point $\hat{\pi}^{last}$, solution to conv (1), with x^{last} solving (5).

Proof. Suppose that there is a last serious step $\hat{x}$ followed by infinitely many null steps and let $\hat{\mu}$ denote the corresponding proximal parameter. Note first that, since $\left(C(p^l), p^l \right) \in B_{l+1}$ for any bundle management strategy, having $x^{l+1} = x^l$ implies that $f(x^{l+1}) = f(x^l) = C(p^l) - \left\langle g(p^l), x^{l+1} \right\rangle$. But since by construction $\breve{f} \le f$ and $\breve{f}_{l+1}(x^{l+1}) \ge C(p^l) - \left\langle g(p^l), x^{l+1} \right\rangle$., we conclude that

$\breve{f}_{l+1}(x^{l+1}) = f(x^{l+1})$ and, hence, $\Delta_{l+1} = f(\hat{x}) - f(x^{l+1})$.

By

$$f(x^{l+1}) + \frac{1}{2}\hat{\mu}\left|x^{l+1} - \hat{x}\right|^2 = \breve{f}_{l+1}(x^{l+1}) + \frac{1}{2}\hat{\mu}\left|x^{l+1} - \hat{x}\right|^2$$

$\leq f(\hat{x}) \leq f(x^{l+1})$, so $\hat{x} = x^{l+1}$. In this case the algorithm would eventually stop ($\Delta_{l+1} = 0$, contradicting our starting assumption. Thus, infinite null steps occur only with $x^{l+1} \neq x^l$ For $l \geq last$, consider the following problem :

$$\begin{cases} \min_{r \in R, x \in R^n} r + \frac{1}{2}\hat{\mu}\left|x - \hat{x}\right|^2 \\ r \geq C(p^i) - \left\langle g(p^i), x \right\rangle \; for \; i \in B_l \\ x_{\bar{J}} \geq 0 \; and \; x_{L \setminus \bar{J}} = 0, \end{cases} \qquad (6)$$

and denote its optimal value $O_l := \breve{f}_l(x^l) + \frac{1}{2}\hat{\mu}\left|x^l - \hat{x}\right|^2$. Relation in [HUL93]

$$O_{l+1} \geq O_l + \frac{1}{2}\hat{\mu}\left|x^{l+1} - x^l\right|^2,$$

together with the fact that $x^{l+1} \neq x^l$ imply that the values of O_l are strictly

increasing. The assumption that *Q* is finite implies that B_l contains at most q different pairs (C_i, p^i) . As a result, there is a finite number of different feasible sets in (22) for $l \geq last$, contradicting the fact that the (infinite) values of O_l are strictly increasing.

Consider $l_k \in L'$, $\hat{s}^{l_k} \in \partial\breve{f}_{l_k}(\hat{x}^{k+1})$, $\hat{v}^{l_k} \in N_{J_{l_k}}(\hat{x}^{k+1})$. Since *Q* is finite there is only a finite number of different combinations of J_{l_k} and $Q_{l_k} := \left\{p^i \in Q : i \in B_{l_k}\right\}$.

There exists $\rho > 0$ such that $\left|\hat{s}^{l_k} + \hat{v}^{l_k}\right| < \rho$ implies that $\hat{s}^{l_k} + \hat{v}^{l_k} = 0$. As a result, using(8), $\hat{x}^{k+1} = x^{l_k+1} = \hat{x}^k$. For this value of $\hat{x}^{k+1}$, the descent test in Step 3 of Algorithm (which must hold because l_k gave a serious step) becomes $f(\hat{x}^k) \leq f(\hat{x}^k) - m\Delta_{l_k}$. This inequality is only possible if $\Delta_{l_k} = 0,$, or, $\hat{x}^{k+1} = \hat{x}^k$, if

$$f(\hat{x}^{k+1}) = \breve{f}_{l_k}(\hat{x}^{k+1}). \qquad (7)$$

Let k' be the first index such that $\hat{s}^{l_{k'}} + \hat{v}^{l_{k'}} = 0$, i.e. such that $\hat{x}^{k'+1}$ minimizes $\breve{f}_{l_{k'}}$ on the set $\left\{x \geq 0 : x_{L \setminus J_{l_{k'}}} = 0\right\}$.

For every $l \in L \setminus J_{l_{k'}}$ there is an index $j \in J_{l_{k'}}$ such that

$$\hat{v}_l^{l_{k'}} = -\hat{s}_l^{l^{k'}} = g_l(\hat{\pi}^{l_{k'}}) \leq \beta g_j(\hat{\pi}^{l_{k'}})$$

$$= -\beta\hat{s}_j^{l^{k'}} = \beta\hat{v}_j^{l_{k'}} \leq 0,$$

so $\hat{x}^{k'+1}$ solves the problem $\min_{x \geq 0} \breve{f}_{l_{k'}}(x)$ by Corollary 3.

Since $\breve{f} \leq f$ by construction, we see that

$$\breve{f}_{l_{k'}}(\hat{x}^{k'+1}) \leq f(x^{\infty}), \qquad (8)$$

and by (7), this means that $f(x^{\infty}) \geq f(\hat{x}^{k'+1})$, i.e, the relation is satisfied with equality, because x^{∞} solves (2). Therefore, the algorithm would have $\Delta_{l_{k'}} = 0$ with $I_{l_{k'}} \subseteq J_{l_{k'}}$, and the stopping test would be activated.

II. CONCLUSION

It is worth mentioning that finite termination results in the (static) bundle literature need to modify the descent test in Step 3 by setting the Armijo-like parameter *m* equal to 1. Such requirement is used for example in [1],and [2] to show that there can only be a finite number of serious steps when the function f is polyhedral, with the same assumptions on the bundle management strategy, i.e., either no deletion or bundle selection. Since the static case is covered by the dynamic setting, Theorem extends the known results of finite termination to include the case $m \in (0,1)$.

III. REFERNCES

[1] Kiwiel K.C.. Methods of descent for nondifferentiable optimization. Springer-Verlag, Berlin, 1985.

[2] Kiwiel K.C.. Exact penalty functions in proximal bundle methods for constrained convex nondifferentiable minimization. Math. Programming, 52(2, Ser. B):285–302, 1991.

[3]Kiwiel K.C. Approximations in Proximal Bundle Methods and Decomposition of Convex Programs. Journal of Optimization Theory and Applications, 84:529–548, 1995.

[4] Kiwiel K.C.. A proximal bundle method with approximate subgradient lineariza-tions. SIAM J. Optim., 16:1007–1023, 2006.

[5] Lucena A. and Beasley J. E.. Branch and cut algorithms. In Advances in linear and integer programming, volume 4 of Oxford Lecture Ser. Math. Appl., pages 187–221.Oxford Univ. Press, New York, 1996.

[6] Lemar´echal C., Nemirovskii A., and Nesterov Yu.. New variants of bundle methods. Math. Program., 69:111–148, 1995.

[7] Lucena A.. Steiner problem in graphs: Lagrangean relaxation and cutting-planes. COAL Bulletin, 21(2):2–8, 1992. san and J. Vl.

[8] Luk.cek L.. A bundle-Newton method for nonsmooth unconstrained minimization. Math. Programming, 83(3, Ser. A):373–391, 1998.