J Optim Theory Appl (2007) 133: 413–416
DOI 10.1007/s10957-007-9238-5

On the Surrogate Gradient Algorithm for Lagrangian Relaxation

T. Sun · Q.C. Zhao · P.B. Luh

Published online: 7 June 2007

Abstract When applied to large-scale separable optimization problems, the recently developed surrogate subgradient method for Lagrangian relaxation (Zhao et al.: J. Optim. Theory Appl. 100, 699–712, 1999) does not need to solve optimally all the subproblems to update the multipliers, as the traditional subgradient method requires. Based on it, the penalty surrogate subgradient algorithm was further developed to address the homogenous solution issue (Guan et al.: J. Optim. Theory Appl. 113, 65–82, 2002; Zhai et al.: IEEE Trans. Power Syst. 17, 1250–1257, 2002). There were flaws in the proofs of Zhao et al., Guan et al., and Zhai et al.: for problems with inequality constraints, projection is necessary to keep the multipliers nonnegative; however, the effects of projection were not properly considered. This note corrects the flaw, completes the proofs, and asserts the correctness of the methods.

Keywords Lagrangian relaxation · Surrogate subgradient · Inequality constraints

Communicated by W.B. Gong.

This work is supported by the NSFC Grant Nos. 60274011, 60574067, the NCET program (No. NCET-04-0094) of China. The third author was supported in part by US National Science Foundation under Grants ECS-0323685 and DMI-0423607.

T. Sun · Q.C. Zhao (✉) · P.B. Luh
Center for Intelligent and Networked Systems, Department of Automation, Tsinghua University, Beijing 100084, China
e-mail: zhaoqc@tsinghua.edu.cn

P.B. Luh
Department of Electrical and Computer Engineering, University of Connecticut, Storrs, Connecticut, CT 06269-2157, USA

1 Optimization Problem and Lagrangian Multipliers

1.1 Optimization Problem

The separable integer programming problem in [1] can be described as follows (problem descriptions in [2, 3] are a little different, but that does not affect the following analysis):

$$\min_x J_{IP} \equiv \sum_{i=1}^{I} J_i(x_i), \quad \text{s.t.} \quad Ax \le b, \qquad x_i \in Z^{ni}, \quad i = 1, 2, \ldots, I, \tag{1}$$

where $x = [x_1, \ldots, x_I]^{\mathrm{T}}$ is an $N \times 1$ decision variable with $N = \sum_{i=1}^{I} n_i$ and Z^{ni} is a set of integers. Furthermore, A is an $M \times N$ matrix, b is an $M \times 1$ vector and the objectives $\{J_i(x_i)\}$ are possibly nonlinear functions. The Lagrangian relaxation (LR) of (1) is given by

$$\min_x \left[\sum_{i=1}^{I} J_i(x_i) + \lambda^{\mathrm{T}}(Ax - b) \right], \quad \text{s.t.} \quad x_i \in Z^{ni}. \tag{2}$$

Here, λ is an $M \times 1$ vector of Lagrangian multipliers.

1.2 Projection for Updating the Multipliers

For LR, the multipliers are nonnegative when relaxing the inequality constraints [1–4]. Note there is no such requirement for equality constraints. Thus, the projection is necessary when updating the multipliers during the iterations, i.e.,

$$\hat{\lambda}^{k+1} \equiv \lambda^k + s^k g^k, \tag{3}$$

$$\lambda^{k+1} = [\hat{\lambda}^{k+1}]^+. \tag{4}$$

Here, at iteration k, s^k is the stepsize, g^k is the subgradient or surrogate subgradient given by $g^k = Ax^k - b$, and x^k is the argument of (2). The projection within (4) is obtained as

$$\lambda_m^{k+1} = \hat{\lambda}_m^{k+1}, \quad \text{if } \hat{\lambda}_m^{k+1} \ge 0, \tag{5a}$$

$$\lambda_m^{k+1} = 0, \qquad \text{else}, \tag{5b}$$

where $m = 1, \ldots, M$ is the element index of λ^{k+1} and $\hat{\lambda}^{k+1} \equiv \lambda^k + s^k g^k$.

The above shows that the rigorous expression of (7) and (25) in [1] should have the form (4) and be updated according to (5). The same is required for μ in (30) of [2] and (25) of [3].

2 Completion of the Proofs Affected by Projection

The projection in (4) invalidates some of the steps within the proofs in [1–3]. Some of them can be corrected by applying the nonexpansive property of projection and will be presented in Sect. 2.1 below. Others are more involved and will be presented in Sect. 2.2.

2.1 Nonexpansive Distance of Projection

From the projection theorem in [4], the projection in (4) as given by (5) is nonexpansive, i.e.,

$$\|[v]^+ - [w]^+\| \le \|v - w\|, \quad \forall v, w \in R^M.$$

Because of this, (36) in the proof of Theorem 4.1 [1] should be

$$\|\lambda^* - \lambda^{k+1}\|^2 \le \|\lambda^* - \lambda^k - s^k \tilde{g}^k\|^2 = \|\lambda^* - \lambda^k\|^2 - 2s^k(\lambda^* - \lambda)^{\mathrm{T}}\tilde{g}^k + (s^k)^2\|\tilde{g}^k\|^2.$$

Thus, the theorem still holds since the remaining proof is still valid. The same applies to (39) of [2] and (A4) of [3].

2.2 Completion of the Proofs for the Case with Inequality Constraints

For the surrogate subgradient method, it is important that the surrogate dual is always less than the optimal dual (Proposition 4.1 in [1]). However, the argument within the proof is not correct, since (33) in [1], i.e., (6) below,

$$\tilde{L}^k + (\lambda^{k+1} - \lambda^k)^{\mathrm{T}}(Ax^k - b) = \tilde{L}^k + s^k\|g^k\|^2, \tag{6}$$

may not hold. The reason is that the multiplier λ, used to relax inequality constraints in [1], needs projection as in (4). However, the correction is not straightforward and requires Lemma 2.1 below.

Lemma 2.1 *At the* kth *iteration*,

$$(\lambda^{k+1} - \lambda^k)^{\mathrm{T}} g^k \le s^k\|g^k\|^2,$$

where g^k *is a subgradient or surrogate subgradient as in* [1].

Proof If λ is a multiplier to relax the equality constraints, then there is no projection and

$$(\lambda^{k+1} - \lambda^k)^{\mathrm{T}} g^k = s^k(g^k)^{\mathrm{T}} g^k = s^k\|g^k\|^2.$$

Therefore, (33) in [1] holds. When λ is used to relax the inequality constraints, then

$$(\lambda^{k+1} - \lambda^k)^{\mathrm{T}} g^k - s^k\|g^k\|^2 = (\lambda^{k+1} - \lambda^k - s^k g^k)^{\mathrm{T}} g^k = (\lambda^{k+1} - \hat{\lambda}^{k+1})^{\mathrm{T}} g^k \le 0.$$

To explain the last inequality, let

$$\Delta\lambda \equiv \lambda^{k+1} - \hat{\lambda}^{k+1}.$$

In view of the fact that λ^{k+1} is the projection of $\hat{\lambda}^{k+1}$, the elements of $\Delta\lambda$ are nonnegative for (5). For any element m, $\Delta\lambda_m > 0$ implies $\hat{\lambda}_m^{k+1} < 0$. Furthermore, $g_m^k < 0$, since s^k and λ_m^k are nonnegative in (3). Therefore, $\Delta\lambda^{\mathrm{T}} g^k \le 0$, the last inequality is validated, and Lemma 2.1 is proved. □

From Lemma 2.1, it can be seen that the equality (33) in [1] should be replaced by the following inequality:

$$\tilde{L}^k + (\lambda^{k+1} - \lambda^k)^{\mathrm{T}}(Ax^k - b) \le \tilde{L}^k + s^k \|g^k\|^2.$$

In this way, the remaining part of Proposition 4.1 still holds. Similar corrections apply to the proofs of Theorem 5.1 in [2] and Lemma 2 in [3].

3 Conclusions

Projection is necessary to update the multipliers when Lagrangian relaxation is used for problems with inequality constraints. However, this was not considered adequately in [1–3], causing flaws in the proofs of the surrogate subgradient and related methods. This note corrects the flaws, completes the proofs, and asserts the correctness of the methods for problems with both equality and inequality constraints.

References

1. Zhao, X., Luh, P.B., Wang, J.: Surrogate gradient algorithm for Lagrangian relaxation. J. Optim. Theory Appl. **100**, 699–712 (1999)
2. Guan, X.H., Zhai, Q.Z., Lai, F.: New Lagrangian relaxation based algorithm for resource scheduling with homogeneous subproblems. J. Optim. Theory Appl. **113**, 65–82 (2002)
3. Zhai, Q., Guan, X., Cui, J.: Unit commitment with identical units: successive subproblem solving method based on Lagrangian relaxation. IEEE Trans. Power Syst. **17**, 1250–1257 (2002)
4. Bertsekas, D.P.: Nonlinear Programming. Athena Scientific, Belmont (1999)