Working Paper

A Note on the Twice Differentiable Cubic Augmented Lagrangian

Krzysztof C. Kiwiel

WP-94-12
March 1994

IIASA International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria
Telephone: +43 2236 715210 □ Telex: 079 137 iiasa a □ Telefax: +43 2236 71313

A Note on the Twice Differentiable Cubic Augmented Lagrangian

Krzysztof C. Kiwiel

WP-94-12
March 1994

IIASA International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria
Telephone: +43 2236 715210 □ Telex: 079 137 iiasa a □ Telefax: +43 2236 71313

A note on the twice differentiable cubic augmented Lagrangian*

Krzysztof C. Kiwiel†

March 11, 1994

Abstract

Rockafellar's quadratic augmented Lagrangian for inequality constrained minimization is not twice differentiable. To eliminate this drawback, several quite complicated Lagrangians have been proposed. We exhibit a simple cubic Lagrangian that is twice differentiable. It stems from the recent work of Eckstein and Teboulle on Bregman-related Lagrangians.

Key words. Convex programming, augmented Lagrangians, multiplier methods, proximal methods, Bregman functions.

MSC Subject Classification. Primary: 65K05. Secondary: 90C25.

1 Introduction

The purpose of this note is to call attention to a simple modified Lagrangian for the convex program

$$\text{minimize} \quad f_0(x) \quad \text{over all } x \in C \text{ satisfying} \quad f_i(x) \le 0, \quad i = 1\colon m, \tag{1}$$

where C is a nonempty closed convex subset of $\mathbb{R}^n$ and $f_i : C \to \mathbb{R}$ is a closed convex function for $i = 0, 1, \ldots, m$. The *quadratic augmented Lagrangian* of Rockafellar [Roc73] is

$$L_c^{(2)}(x,y) = f_0(x) + \frac{1}{2c}\sum_{i=1}^{m}\{[y_i + cf_i(x)]_+^2 - (y_i)^2\} \tag{2}$$

for $x \in C$ and $y \in \mathbb{R}^m$, where c is a positive number and $[\cdot]_+ = \max\{\cdot, 0\}$. The corresponding *multiplier method* [Roc76] generates sequences $\{x^k\} \subset C$ and $\{y^k\} \subset \mathbb{R}_+^m$, which should converge to the solution and Lagrange multiplier of (1) respectively, via the recursion

$$x^k \in \operatorname{Arg\,min}_{x \in C} L_{c_k}^{(2)}(x, y^k), \tag{3a}$$

$$y_i^{k+1} = [y_i^k + c_k f_i(x^k)]_+, \quad i = 1\colon m, \tag{3b}$$

where $\{c_k\}$ is a nondecreasing sequence of positive numbers (or $\inf_k c_k > 0$ [Eck93]).

*This research was supported by the State Committe for Scientific Research under Grant 8S50502206 and by the International Institute for Applied Systems Analysis, Laxenburg, Austria.

†Systems Research Institute, Newelska 6, 01–447 Warsaw, Poland (kiwiel@ibspan.waw.pl)

Even if all f_i are twice differentiable on C, the Lagrangian $L^{(2)}_{c_k}(\cdot, y^k)$ is differentiable only once. This may create difficulties for methods used to find x^k in (3a) [Ber82, GoT89, KoB76, Man75, TsB93]. Other twice differentiable Lagrangians are either quite complicated [Ber82, GoT74, GoT89, KoB76], or nonconcave with respect to y [Man75], or difficult to analyze [TsB93]. In the next section we exhibit a simple twice differentiable Lagrangian. It is derived from the recent work of [Eck93, Teb92] on Bregman-related Lagrangians.

2 The cubic Lagrangian

Consider using the *cubic augmented Lagrangian*

$$L^{(3)}_c(x,y) = f_0(x) + \frac{1}{3c}\sum_{i=1}^m \{[\mathrm{sign}(y_i)|y_i|^{1/2} + cg_i(x)]^3_+ - |y_i|^{3/2}\} \tag{4}$$

in the method

$$x^k \in \mathrm{Arg\,min}_{x\in C}\, L^{(3)}_{c_k}(x, y^k), \tag{5a}$$

$$y_i^{k+1} = [|y_i^k|^{1/2} + c_k f_i(x^k)]^2_+, \quad i = 1\colon m, \tag{5b}$$

with $y^1 \geq 0$. Clearly, $L^{(3)}_{c_k}(\cdot, y^k)$ is continuously twice differentiable on C if so is each f_i.

Letting

$$p(t;\mu) = \frac{1}{3}\{[\mathrm{sign}(\mu)|\mu|^{1/2} + t]^3_+ - |\mu|^{3/2}\} \quad \text{for } t, \mu \in \mathbb{R},$$

we have $L^{(3)}_c(x,y) = f_0(x) + \sum_{i=1}^m p[cg_i(x), y_i]/c$, $y_i^{k+1} = \nabla_t p[c_k g_i(x^k); y_i^k]$, $i = 1\colon m$. Since p belongs to the class of penalty functions denoted by P_I in [Ber82, p. 305] and by P in [KoB76], these references contain results on global convergence of the method (5), including possible inexact minimization in (5a).

Changing variables via $\bar{y}_i = \mathrm{sign}(y_i)|y_i|^{1/2}$, $i = 1\colon m$, we may express $L^{(3)}_c$ as

$$\bar{L}^{(3)}_c(x,\bar{y}) = f_0(x) + \frac{1}{3c}\sum_{i=1}^m \{[\bar{y}_i + cg_i(x)]^3_+ - |\bar{y}_i|^3\}. \tag{6}$$

$\bar{L}^{(3)}_c$ is a Lagrangian of Mangasarian [Man75] (with $\psi(\xi) = |\xi|^3/3c$). $\bar{L}^{(3)}_c(x,\cdot)$ is concave on $\mathbb{R}^m$ if x is feasible in (1) [Man75, Rem. 2.13], and so is $L^{(3)}_c(x,\cdot)$, since $\nabla^2_\mu p(t;\mu) = -1/4|\mu|^{1/2}$ if $\mu < 0$ and $t \leq 0$, or $\mu > 0$ and $\mu^{1/2} + t < 0$, $\nabla^2_\mu p(t;\mu) = -t^2/4\mu^{3/2}$ if $\mu > 0$ and $\mu^{1/2} + t > 0$. If $x \in C$, $L^{(3)}_c(x,\cdot)$ is also concave on $\mathbb{R}^m_+$. In general, neither $L^{(3)}_c(x,\cdot)$ nor $\bar{L}^{(3)}_c(x,\cdot)$ are concave on $\mathbb{R}^m$ if x is infeasible. (In contrast, the concavity of $L^{(2)}_c(x,\cdot)$ for each $x \in C$ (and of other modified Lagrangians [GoT74]) facilitates the development of algorithms; cf. [GoT89, Chaps 3–5].) $\bar{L}^{(3)}_c(x,\cdot)$ is twice differentiable, and so is $L^{(3)}_c(x,\cdot)$, except on the boundary of $\mathbb{R}^m_+$.

As an extension, for an integer $\beta > 1$, consider using the Lagrangian

$$L^{(\beta)}_c(x,y) = f_0(x) + \frac{1}{\beta c}\sum_{i=1}^m \{[\mathrm{sign}(y_i)|y_i|^{1/(\beta-1)} + c_k g_i(x)]^\beta_+ - |y_i|^{\beta/(\beta-1)}\} \tag{7}$$

in the method

$$x^k \in \mathrm{Arg\,min}_{x\in C}\, L^{(\beta)}_{c_k}(x, y^k), \tag{8a}$$

$$y_i^{k+1} = [|y_i^k|^{1/(\beta-1)} + c_k f_i(x^k)]^{\beta-1}_+, \quad i = 1\colon m, \tag{8b}$$

with $\beta = 2$ corresponding to (2)–(3) and $\beta = 3$ to (4)–(5). Note that $L_{c_k}^{(\beta)}(\cdot, y^k)$ is $(\beta - 1)$-times differentiable on C if so is each f_i. Again one may associate $L_{c_k}^{(\beta)}$ with Mangasarian's Lagrangians. Global convergence of the method (8) follows from Theorem 7 of [Eck93], because $L_{c_k}^{(\beta)}$ stems from the Bregman function $h(y) = \sum_{i=1}^{m} |y_i|^{\alpha}/\alpha$ with $\alpha = \beta/(\beta - 1)$; cf. [Eck93, Teb92].

References

[Ber82] D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, New York, 1982.

[Eck93] J. Eckstein, *Nonlinear proximal point algorithms using Bregman functions, with applications to convex programming*, Math. Oper. Res. **18** (1993) 202–226.

[GoT74] E. G. Golshtein and N. V. Tretyakov, *Modified Lagrange functions*, Èkonom. i Mat. Metody **10** (1974) 568–591 (Russian).

[GoT89] ———, *Modified Lagrange Functions; Theory and Optimization Methods*, Nauka, Moscow, 1989 (Russian).

[KoB76] B. W. Kort and D. P. Bertsekas, *Combined primal-dual and penalty methods for convex programming*, SIAM J. Control Optim. **14** (1976) 268–294.

[Man75] O. L. Mangasarian, *Unconstrained Lagrangians in nonlinear programming*, SIAM J. Control **13** (1975) 772–791.

[Roc73] R. T. Rockafellar, *A dual approach to solving nonlinear programming problems by unconstrained optimization*, Math. Programming **5** (1973) 354–373.

[Roc76] ———, *Augmented Lagrangians and applications of the proximal point algorithm in convex programming*, Math. Oper. Res. **1** (1976) 97–116.

[Teb92] M. Teboulle, *Entropic proximal mappings with applications to nonlinear programming*, Math. Oper. Res. **17** (1992) 670–690.

[TsB93] P. Tseng and D. P. Bertsekas, *On the convergence of the exponential multiplier method for convex programming*, Math. Programming **60** (1993) 1–19.