Journal of Theoretical
and Applied Mechanics
Sofia 1993, Year XXIV No 1

AUGMENTED LAGRANGIANS FOR GENERAL OPTIMAL CONTROL PROBLEMS†

S. GROZDEV

Abstract. The paper is devoted to the Multiplier Method (known as Penalty Shifting Method or Augmented Lagrangian Method) independently introduced by Hestenes [1] and Powell [2]. It is considered an optimal control problem in a general form including state equations of elliptic, parabolic or hyperbolic types with final or distributed observations. An augmented Lagrangian is introduced and it is shown that its unconstrained minimization is equivalent to finding the saddle points of the ordinary Lagrangian function, which corresponds to the primal constrained problem.

Introduction. To state the idea of the Multiplier method we consider the equality constrained nonlinear program: find $\min f(x)$, subjected to $g(x) = 0$, where $f:\mathbb{R}^n \to \mathbb{R}$ and $g:\mathbb{R}^n \to \mathbb{R}^m$, $m < n$ are functions of class C^k with a suitable k. The classical Lagrangian function associated with the nonlinear program above is $l(x,\lambda) = f(x) + \lambda^T g(x)$, λ being the dual variable. The standard primal and dual problems corresponding to the nonlinear program are formulated in terms of the Lagrangian as: find $\inf_x \sup_\lambda l(x,\lambda)$ (or find $\sup_\lambda \inf_x l(x,\lambda)$), which is equivalent to the nonlinear program problem. If the pair (x,λ) is a saddle point of $l(x,\lambda)$, then x is a solution of the nonlinear program problem with λ being the multiplier vector. However, the converse is not always true. For example, in nonconvex problems the dual problem may not have a solution or even if it has, a duality gap appears usually. The idea of the Multiplier Method is to remove the gap by passing to an augmented Lagrangian, which

† This work was supported by the National Foundation for Scientific Investigations (project N 52/1991).

involves for instance quadratic penalty - like terms. Thus, a primal-dual equivalence, in the nonconvex case can be achieved by a corresponding parametrized penalty term. In particular, the idea of the Multiplier Method is to prove that a local solution of the nonlinear program problem and its multiplier vector λ form a saddle point of the augmented Lagrangian. The practical use of the Multiplier Method has been of great importance. Various computational algorithms, which alternately update x and λ to find an approximate saddle point of the augmented Lagrangian have been solved successfully and usually this has been achieved by proving a corresponding equivalence between the saddle points of both the ordinary Lagrangian function and the augmented Lagrangian.

The present investigation is inspired by the papers of G. Di Pillo and L. Grippo [3,4], the second one of which considers boundary and distributed control problems in a bounded open set $\Omega \in \mathbb{R}^n$ and a Sobolev space $H^1(\Omega)$. It turns out that the main ideas of the cited papers are well working in abstract Hilbert spaces too. Thus, the present paper is devoted to an optimal control problem in such spaces. However, the aim is not to solve the optimal control problem itself, but to introduce an augmented Lagrangian, which corresponds to it and to show equivalence between the saddle points of that Lagrangian and the ordinary Lagrangian function.

Problem formulation. If H is a Hilbert space, then $(x,y)_H$, with $x,y \in H$, denotes its inner product and $||x||_H = \sqrt{(x,x)}$ its norm correspondingly. Let Y be a finite – or infinite – dimensional real Hilbert space and A be a positive – definite, self – adjoint, bounded operator on Y. Then, $(Ax,y) = (x,Ay)$ and $m||x||^2 \leq (Ax,x) \leq M||x||^2$ with suitably chosen positive constants m and M.

Let U be a second real Hilbert space and let $B \in \mathcal{L}(U,Y)$. We consider the following general state equation:

$$Ay = f + Bu, \tag{1}$$

where $u \in U$ is the control and $f \in Y$ is given.

Due to the assumptions, if u is known, then it defines an unique state y, which depends on u continuously. In the simpler case when B is the zero operator, the solution of (1), which is $y = A^{-1}f$, affords a minimum on Y to the function $Fy = \frac{1}{2}(Ay,y)_Y - (f,y)_Y$. Since $F'(y,h) = (Ay - f,h)_Y$, with $h \in Y$, then $F'(y) = Ay - f$. If now K is another positive – definite, bounded and self – adjoint operator on Y, then $< x,y >= (K^{-1}x,y)_Y$, with $x,y \in Y$, is a second inner product, topologically equivalent to the first one. The negative gradient g of F at y relative to this new inner product can be defined by the relation $F'(y,h) = - < g,h >$ for all $h \in Y$. Hence, $g(y) = Kr(y)$, where $r = -F'(y) = f - Ay$ is the negative gradient of F at y relative to the primal inner product. To solve (1), when B is the zero operator, one can use the generalized gradient method, which takes the form $y_{n+1} = y_n + \alpha_n K r_n$ with suitably chosen $\alpha_n > 0$ and $r_n = -F'(y_n)$. The ideal choice of K would be $K = A^{-1}$, but it is impossible, since A^{-1} is not assumed to be known.

In the general case let A be an isomorphism in Y and let F and G be convex and semi – continuous real functions in Y and U, respectively. A further assumption for F is taken from [4]. Let X be arbitrary Hilbert space, $x_d \in X$ and $C \in \mathcal{L}(Y, X)$. Then, $F(y) = \frac{1}{2}||Cy - x_d||_X^2$. Instead of F, an affine function can be taken too.

We consider the following optimal control problem:

(2) $$\text{find } \min I(y,u) = F(y) + G(u), \text{ when } y \text{ and } u \text{ verify (1).}$$

Let Y' be the dual space of Y and λ be the dual variable. Then, the classical Lagrangian function, associated to (2), takes the form:

(3) $$l(y,u,\lambda) = F(y) + G(u) + < \lambda, Ay - f - Bu > .$$

If $F'(y)$ denotes the G–differential of F at y, then we consider the following augmented Lagrangian:

(4) $$L(y,u,\lambda) = l(y,u,\lambda) + \alpha||Ay - f - Bu||_Y^2 + \beta||A^*\lambda - F'y||_{Y'}^2,$$

with suitable constants α and β.

It is studied in the sequel the relationship between in the points minimizing (4) and the saddle points of (3). We note that in the particular case, when $\lambda(y)$ is defined by $A^*\lambda(y) = F'(y)$, the Hestenes [1] augmented Lagrangian is obtained. Using (4), it takes the form $L(y,u,\lambda(y))$. In addition, using the notations form [4], let Ω be a bounded open set in $\mathbb{R}^n$ with boundary Γ, $[O,T]$ a fixed time interval, $Q = \Omega \times (O,T)$ and $\Sigma = \Gamma \times (O,T)$. By $H^1(\Omega)$ it is denoted the Sobolev space:

$$H^1(\Omega) = \left\{\varphi\colon \varphi \in L^2(\Omega),\ \frac{\partial\varphi}{\partial x_i} \in L^2(\Omega)\right\},$$

provided with the norm:

$$||\varphi||_{H^1(\Omega)} = \left(||\varphi||_{L^2(\Omega)}^2 + \sum_{i=1}^{n} \left|\left|\frac{\partial\varphi}{\partial x_i}\right|\right|_{L^2(\Omega)}^2\right)^{1/2},$$

and by $H_0^1(\Omega)$ the subspace of $H^1(\Omega)$:

$$H_0^1(\Omega) = \{\varphi\colon \varphi \in H^1(\Omega), \varphi/\Gamma = 0\}.$$

If we take now $Y = L^2(O,T; H^1(\Omega))$, the second order parabolic operator

$$A = \frac{\partial}{\partial t} + \sum_{i,j=1}^{n} \frac{\partial}{\partial x_i}\left(a_{ij}(x,t)\frac{\partial y}{\partial x_j}\right),$$

with a_{ij} being given functions in Q, such that

$$\sum_{i,j=1}^{n} a_{ij}(x,t)\xi_i\xi_j \geq \alpha \sum_{i=1}^{n} \xi_i^2, \ \alpha > 0, \ \xi_i \in \mathbb{R},$$

and $U = L^2(\Sigma)$, then (1) could be interpreted in a form studied in [4].

Results.

Lemma. *Let X be a Hilbert space, $x_d \in X$ and $C \in \mathcal{L}(Y,X)$. If $F(y) = \frac{1}{2}||Cy - x_d||_X^2$, then it exsists $\delta > 0$, such that*

$$F(y_2) \geq F(y_1) + < F'(y_1), y_2 - y_1 > + \delta||F'(y_1) - F'(y_2)||_{Y'}^2,$$

for $y_1, y_2 \in Y$.

Proof. We have

$$F(y_2) = F(y_1) + < F'(y_1), y_2 - y_1 > + \frac{1}{2}||C(y_2 - y_1)||_X^2,$$

and $F'(y_2) - F'(y_1) = C^*C(y_2 - y_1)$.

Since C is continuous, it exists $\varepsilon > 0$, such that

$$||F'(y_2) - F'(y_1)||_{Y'}^2 \leq \varepsilon||C(y_2 - y_1)||_X^2.$$

Now, the Lemma is verified, if $\delta \leq \frac{\varepsilon}{2}$.

Theorem. *Let $L(y,u,\lambda)$ be the augmented Lagrangian defined by (4). When $\beta > 0$ is fixed, then it exists $\gamma > 0$, such that for $\alpha > \gamma$ the set of all triples (y,u,λ), which minimize $L(y,u,\lambda)$, coincides with the set of the saddle points of $l(y,u,\lambda)$ from (3).*

Proof. Let (y,u,λ) be a saddle point of $l(y,u,\lambda)$ and let $(\tilde{y},\tilde{u})$ be a solution of (2). It is well known that there exists $\tilde{\lambda} \in Y'$, such that $(\tilde{y},\tilde{u},\tilde{\lambda})$ is a saddle point of $l(y,u,\lambda)$ too. We shall use the notation

$$D = L(\tilde{y} + \Delta y, \tilde{u} + \Delta u, \tilde{\lambda} + \Delta\lambda) - L(\tilde{y},\tilde{u},\tilde{\lambda})$$

and prove that $D \geq 0$.

Since $L(\tilde{y},\tilde{u},\tilde{\lambda}) = F(\tilde{y}) + G(\tilde{y})$ and $A(\tilde{y} + \Delta y) - f - B(\tilde{u} + \Delta u) = A\Delta y - B\Delta u$, then

$$D = F(\tilde{y} + \Delta y) - F(\tilde{y}) + G(\tilde{u} + \Delta u) - G(\tilde{u}) - < \tilde{\lambda} + \Delta\lambda, A\Delta y - B\Delta u > +$$
$$+ \alpha||A\Delta y - B\Delta u||_Y^2 + \beta||A^*(\tilde{\lambda} + \Delta\lambda) - F'(\tilde{y} + \Delta y||_{Y'}^2.$$

Using the Lemma and the identity $A^*\tilde{\lambda} = F'(\tilde{y})$, we get

$$F(\tilde{y} + \Delta y) - F(\tilde{y}) - < \tilde{\lambda}, A\Delta y > \geq \delta ||F'(\tilde{y} + \Delta y) - F'(\tilde{y})||^2_{Y'}$$

for a suitable δ and $G(\tilde{u} + \Delta u) - G(\tilde{u}) + < \tilde{\lambda}, B\Delta u > \geq 0$. The simple calculation shows that

$$A^*(\tilde{\lambda} + \Delta\lambda) - F'(\tilde{y} + \Delta y) = A^*\Delta\lambda - F'(\tilde{y} + \Delta y) + F'(\tilde{y}).$$

Thus,

$$D \geq \delta ||F'(\tilde{y} + \Delta y) - F'(\tilde{y})||^2_{Y'} - < \Delta\lambda, A\Delta y - B\Delta u > +$$

$$+\alpha ||A\Delta y - B\Delta u||^2_Y + \beta ||A^*\Delta\lambda - F'(\tilde{y} + \Delta y) + F'(\tilde{y})||^2_{Y'}.$$

Let $\varepsilon > 0$ be arbitrary. Then,

$$- < \lambda, A\Delta y - B\Delta u > \geq -\frac{\varepsilon}{2} ||\Delta\lambda||^2_{Y'} - \frac{1}{2\varepsilon} ||A\Delta y - B\Delta u||^2_Y.$$

We get

$$D \geq \delta ||F'(\tilde{y} + \Delta y) - F'(\tilde{y})||^2_{Y'} - \frac{\varepsilon}{2} ||\Delta\lambda||^2_{Y'} +$$

$$+(\alpha - \frac{1}{2\varepsilon}) ||A\Delta y - B\Delta u||^2_Y + \beta ||A^*\Delta\lambda - F'(\tilde{y} + \Delta y) + F'(\tilde{y})||^2_{Y'}.$$

Since A^* is isomorphism, then there exists $k > 0$, such that

$$||\Delta\lambda||^2_{Y'} \leq k ||A^*\Delta\lambda||^2_{Y'}.$$

We deduce that for all $\varepsilon > 0$

$$-\frac{\varepsilon}{2} ||\Delta\lambda||^2_{Y'} \geq -\frac{k\varepsilon}{2} ||A^*\Delta\lambda||^2_{Y'} \geq$$

$$\geq -k\varepsilon(||A^*\Delta\lambda - F'(\tilde{y} + \Delta y) + F'(\tilde{y})||^2_{Y'} + ||F'(\tilde{y} + \Delta y) - F'(\tilde{y})||^2_{Y'}).$$

At last we get

$$D \geq (\delta - k\varepsilon) ||F'(\tilde{y} + \Delta y) - F'(\tilde{y})||^2_{Y'} + (\alpha - \frac{1}{2\varepsilon}) ||A\Delta y - B\Delta u||^2_Y +$$

$$+(\beta - k\varepsilon) ||A^*\Delta\lambda - F'(\tilde{y} + \Delta y) + F'(\tilde{y})||^2_{Y'}.$$

It is sufficient to chose $0 < \varepsilon < \min\left(\frac{\delta}{k}, \frac{\beta}{k}\right)$ and $\gamma = \frac{1}{2\varepsilon}$.

Let us assume that $D = 0$.Now we have

$$A((\tilde{y} + \Delta y) - f = B(\tilde{u} + \Delta\tilde{u})$$

and

$$A^*(\tilde{\lambda} + \Delta\lambda) = F'(\tilde{y} + \Delta y),$$

which means that the state equation and the dual one are satisfied simultaneonsly. At the same time

$$D = F'(\tilde{y} + \Delta y) - F'(\tilde{y}) + G(\tilde{u} + \Delta u) - G(\tilde{u}),$$

which implies $I(\tilde{y}+\Delta y, \tilde{u}+\Delta u) = I(\tilde{y}, \tilde{u})$. Thus, the couple $(\tilde{y}+\Delta y, \tilde{u}+\Delta u))$ turns out to be a solution of the control problem, while $\tilde{\lambda} + \Delta\lambda$ the dual state. The conclusion is, that $(\tilde{y} + \Delta y, \tilde{u} + \Delta u, \tilde{\lambda} + \Delta\lambda)$ is a saddle point of $l(y, u, \lambda)$.

REFERENCES

[1] M. R. HESTENES. Multiplier and Gradient Methods. *Journal of Optimization Theory and Applications*, **4**, No 5, 1969, 303-320.

[2] M. J. D. POWELL. A Method for Nonlinear Constraints in Minimization Problems. In: Optimization, Academic Press, N. Y., 1969, 283-298.

[3] G. DI PILLO and L. GRIPPO. A New Class of Augmented Lagrangians in Nonlinear Programming. *SIAM Journal of Control and Optimization*, **17**, No 5, 1979, 618-628.

[4] G. DI PILLO and L. GRIPPO. The Multiplier Method for Optimal Control Problems of Parabolic Systems. *Applied Mathematics and Optimization*, **5**, 1979, 253-269.

Received on April 2, 1993