MGOPT method for semilinear elliptic optimal control problems

Michelle Vallejos*
Institute of Mathematics
University of the Philippines
Diliman, Quezon City, Philippines
michelle.vallejos@up.edu.ph

ABSTRACT

A numerical technique called multigrid for optimization or MGOPT that solves semilinear elliptic optimal control problems is presented. The given problem is discretized by finite difference technique incorporated with a multigrid strategy. To illustrate the MGOPT method, we focus on minimization problems governed by semilinear elliptic differential equations.

1. INTRODUCTION

Among the most efficient tools for the solution of systems of equations arising from the discretization of elliptic optimal control problems are multigrid methods. Some recent results and developments of multigrid include the application to optimal control problems [1, 2, 4], inverse problems [9, 10] and to real- world problems [3]. The purpose of this paper is to formulate a fast numerical technique for solving optimal control problems. We focus on one type of multigrid which is called the multigrid for optimization or MGOPT. We consider the application of this method for solving semilinear elliptic optimal control problems. This work is an extension of [11], where MGOPT is utilized as a solver for linear elliptic optimal control problems. The MGOPT method was first introduced in [6, 8]. In the MGOPT scheme the multigrid solution process represents the outer loop where the control function is considered as the unique dependent variable. The inner loop consists of a classical one-grid optimization scheme.

In the next section, optimal control problems are presented together with the finite difference discretization. This discretization is utilized for the appropriate optimization algorithms. The section ends with the formulation of the multigrid scheme. Numerical experiments follow to demonstrate the ability of MGOPT in solving control-unconstrained and control-constrained semilinear elliptic optimal control problems and a section of conclusion completes this paper.

*Supported by the Natural Sciences Research Institute.

2. PROBLEM DESCRIPTION

The purpose of an optimal control problem is to minimize a cost functional subject to a constraint given by a partial differential equation. In application, take for example a material plate defined over a domain Ω. Let y be the state of the material representing the temperature distribution over the domain Ω. The heat source and additional temperature contributions can be represented by f and $G(y)$, respectively. The idea is to control the state y to be close as possible to a given target function z by applying an additional control function u. This problem can be formulated as a semilinear elliptic optimal control problem

$$\begin{aligned} \min_{u\in U} J(y,u): &= \tfrac{1}{2}\|y-z\|^2_{L^2(\Omega)} + \tfrac{\nu}{2}\|u\|^2_{L^2(\Omega)}, \\ -\Delta y + F(y) - u &= f \ \text{ in } \Omega, \\ y &= 0 \ \text{ on } \partial\Omega, \end{aligned} \tag{1}$$

where $\nu > 0$ is the weight of the cost of the control, $z \in L^2(\Omega)$ is the target function, and $f \in L^2(\Omega)$. For the control-unconstrained case, $U = L^2(\Omega)$ and for the control-constrained case, the control space is a closed convex subset of $L^2(\Omega)$

$$U_{ad} = \{u \in L^2(\Omega) \mid \underline{u} \le u \le \overline{u} \text{ a.e. in } \Omega\}, \tag{2}$$

where $\underline{u}$ and $\overline{u}$ are elements of $L^\infty(\Omega)$. We define the Lagrange functional

$$L(y,u,p) = J(y,u) + \langle -\Delta y + F(y) - u - f, p\rangle_{H_1^{-1},H_1},$$

where p is the Lagrange multiplier. Equating to zero the Fréchet derivatives of L with respect to the triple (y,u,p) results to the first-order necessary optimality conditions for a minimum. We get

$$\begin{aligned} -\Delta y + F(y) - u &= f \quad \text{in } \Omega, \quad y = 0 \ \text{ on } \partial\Omega, \\ -\Delta p + F'(y)p + y &= z \quad \text{in } \Omega, \quad p = 0 \ \text{ on } \partial\Omega, \\ (\nu u - p, v - u) &\ge 0 \quad \text{for all } v \in U_{ad}. \end{aligned} \tag{3}$$

The first equation is called the state equation, the second is the adjoint equation. The inequality condition is called the optimality condition which becomes $\nu u - p = 0$ in Ω for the control-unconstrained case. Equation (3) is called the optimality system which is a characterization of the solution to the given optimization problem (1). The existence of a unique solution to (1) and its characterization are well known. See for example [7].

We now introduce the reduced cost functional

$$\hat{J}(u) = J(y(u), u),$$

together with $\nabla \hat{J}(u) = \nu u - p$ which is the gradient with respect to u. For the MGOPT method, the gradient projection method [5] is utilized as the optimization algorithm. We want to find a solution u of $\min_u \left(\hat{J}(u) - (f,u)\right)$ such that $u \in U_{ad}$, where U_{ad} is given by (2). Define the projection $\mathcal{P}$ onto U_{ad} by

$$\mathcal{P}_{U_{ad}}(u) = \begin{cases} \underline{u} & \text{if } u \leq \underline{u}, \\ u & \text{if } \underline{u} < u < \overline{u}, \\ \overline{u} & \text{if } u \geq \overline{u}. \end{cases}$$

Given the current iterate u^ℓ, the new iterate $u^\ell(\alpha)$ is defined as

$$u^\ell(\alpha) = \mathcal{P}_{U_{ad}}(u^\ell + \alpha d^\ell),$$

where α satisfies the sufficient decrease condition [5]

$$\left[\hat{J}(u^\ell(\alpha)) - (f, u^\ell(\alpha))\right] - \left[\hat{J}(u^\ell) - (f, u^\ell)\right] \leq -\frac{\sigma}{\alpha}\|u^\ell - u^\ell(\alpha)\|^2,$$

for bound constrained problems and d^ℓ is a search direction.

Next we discuss the multigrid procedure. A typical multigrid method uses a sequence of nested discretization grids of increasing fineness

$$\Omega_1 \subset \Omega_2 \subset \ldots \subset \Omega_L = \Omega.$$

Associated to the sequence of grids is a sequence of finite difference spaces

$$V_1 \subset V_2 \subset \ldots \subset V_L = V.$$

This means that at each grid level k, the problem

$$\min_{u_k} \left(\hat{J}_k(u_k) - (f_k, u_k)_k\right) \tag{4}$$

represents a discrete convex optimization problem which is equivalent to solving

$$\nabla \hat{J}_k(u_k) = f_k$$

in Ω_k. The term f_k is introduced to give a recursive formulation, where $f_k = 0$ at the finest resolution $k = L$. We need transfer operators between finer and coarser grids. We define a restriction operator $I_k^{k-1} : V_k \to V_{k-1}$ and a prolongation operator $I_{k-1}^k : V_{k-1} \to V_k$.

One cycle of MGOPT method is presented in the following algorithm.

MGOPT algorithm

Initialize u_k^0. If $k = 1$, solve $\min_{u_k} \left(\hat{J}_k(u_k) - (f_k, u_k)_k\right)$. Else

1. Apply γ_1 iterations of an optimization algorithm to the problem at resolution k.

2. Apply γ cycles of MGOPT to the coarse grid problem

$$\min_{u_{k-1}} \left(\hat{J}_{k-1}(u_{k-1}) - (f_{k-1}, u_{k-1})_{k-1}\right)$$

to obtain u_{k-1}, where

$$\begin{aligned} f_{k-1} &= I_k^{k-1} f_k + \tau_{k-1} \\ \tau_{k-1} &= \nabla \hat{J}_{k-1}(I_k^{k-1} u_k^{\gamma_1}) - I_k^{k-1} \nabla \hat{J}_k(u_k^{\gamma_1}). \end{aligned}$$

3. For a given step length α,

$$u_k^{\gamma_1+1} = u_k^{\gamma_1} + \alpha\, I_{k-1}^k (u_{k-1} - I_k^{k-1} u_k^{\gamma_1}).$$

4. Apply γ_2 iterations of an optimization algorithm to the problem at resolution k.

The parameter γ characterizes the type of multigrid cycle being used. Typical values are $\gamma = 1$ which is called the V-cycle and $\gamma = 2$ is W-cycle.

Now consider the discrete version of the optimality system (3). By the finite difference discretization, $-\Delta_k$ denotes the minus five-point stencil for the Laplacian and hence we have

$$\begin{aligned} -\Delta_k y_k + F(y_k) - u_k &= f_k, \\ -\Delta_k p_k + F'(y_k) p_k + y_k &= z_k, \\ (\nu u_k - p_k, v_k - u_k) &\geq 0. \end{aligned}$$

Let $x \in \Omega_k$ where $x = (ih_k, jh_k)$ and i, j are the indeces of the grid points arranged lexicographically. We first set

$$\begin{aligned} A &= -(y_{i-1,j} + y_{i+1,j} + y_{i,j-1} + y_{i,j+1}) - h^2 f_{i,j} \\ B &= -(p_{i-1,j} + p_{i+1,j} + p_{i,j-1} + p_{i,j+1}) - h^2 z_{i,j}. \end{aligned}$$

The values A and B are considered constant during the update of the variables at ij. Then

$$\begin{aligned} A + 4y_{i,j} + h^2 F(y_{i,j}) - h^2 u_{i,j} &= 0, \\ B + 4p_{i,j} + h^2 F'(y_{i,j}) p_{i,j} + h^2 y_{i,j} &= 0, \\ (\nu u_{i,j} - p_{i,j}, v_{i,j} - u_{i,j}) &\geq 0. \end{aligned}$$

We can easily compute the updates for the variables $y_{i,j}$ and $p_{i,j}$ by using a Newton method and hence we obtain an update for $u_{i,j}$. In the presence of constraints, a new value for $u_{i,j}$ is obtained by its projection onto the admissible set.

3. NUMERICAL RESULTS

We now present the numerical experiments using multigrid with finite difference method. For the results of the experiments, we use $\gamma_1 = \gamma_2 = 2$ pre- and post- optimization steps. This means that one multigrid cycle uses $\gamma_1 + \gamma_2 = 4$ iterations of the optimization algorithm on the finest level.

We discuss the semilinear elliptic optimal control problem

$$\begin{aligned} \min_{u \in U} J(y, u) &:= \tfrac{1}{2}\|y - z\|_{L^2}^2 + \tfrac{\nu}{2}\|u\|_{L^2}^2, \\ -\Delta y + F(y) - u &= f \quad \text{in} \quad \Omega, \\ y &= 0 \quad \text{on} \quad \partial\Omega. \end{aligned}$$

The domain is the unit square $\Omega = (0,1) \times (0,1)$ with f,$z \in L^2(\Omega)$ given by

$$\begin{aligned} f(x_1, x_2) &= 0, \\ z(x_1, x_2) &= \sin(2\pi x_1)\sin(2\pi x_2). \end{aligned}$$

The target function z is shown in Figure 1.

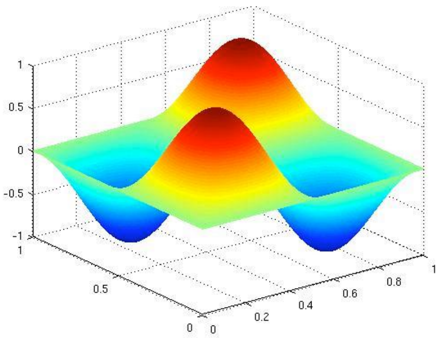

Figure 1: The target function z.

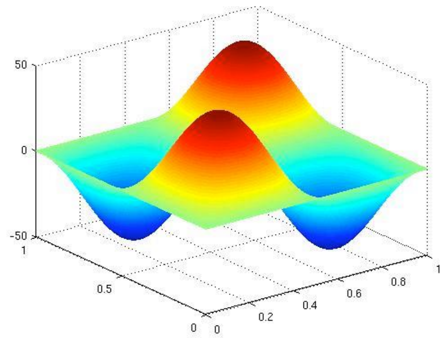

Figure 2: The control function u for the control-unconstrained case of Run 2.

3.1 Run 1

In the semilinear partial differential equation, let $F(y) = y$. The numerical results for the control-unconstrained case are shown in Table 1. In this case, the CPU time in seconds are noted until the L^2-norm of the state and adjoint residuals, $\|r(y)\|_{L^2}$ and $\|r(p)\|_{L^2}$, satisfy a stopping tolerance of $tol = 10^{-11}$. For the control-constrained case, let $\underline{u}(x_1, x_2) = -40$ and $\overline{u}(x_1, x_2) = 40$. The results are reported in Table 2. The CPU time in seconds are noted until the L^2-norm of the difference $\|u^\ell - u^\ell(1)\|_{L^2}$ satisfies a stopping tolerance of $tol = 10^{-5}$. The CPU time approximately increase as a factor of four by halving the mesh size. This shows an almost optimal computational complexity of the MGOPT approach.

Table 1: Numerical results for the control-unconstrained case of Run 1.

ν	mesh	$\|r(y)\|_{L^2}$	$\|r(p)\|_{L^2}$	time (sec)
	33×33	1.451e-13	1.161e-15	0.1
	65×65	2.976e-13	2.230e-15	0.5
10^{-4}	129×129	6.288e-13	6.272e-15	2.4
	257×257	5.057e-12	4.004e-14	10.6
	513×513	1.687e-11	1.368e-13	46.3

Table 2: Numerical results for the control-constrained case of Run 1.

ν	mesh	$\|u^\ell - u^\ell(1)\|_{L^2}$	time (sec)
	33×33	3.31e-05	0.2
	65×65	2.44e-05	0.6
10^{-4}	129×129	4.47e-05	2.7
	257×257	4.33e-05	11.9
	513×513	9.52e-05	54.2

3.2 Run 2

Next we take $F(y) = y^3$. In this experiment, the control function u for the control-unconstrained case and the control-constrained case are shown in Figure 2 and Figure 3, respectively.

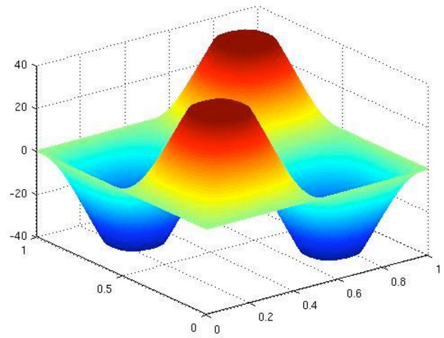

Figure 3: The control function u for the control-constrained case of Run 2.

Here, $\underline{u}(x_1, x_2)$ and $\overline{u}(x_1, x_2)$ are similar as in Run 1. The results for the control-constrained problem are reported in Table 3. Similar to the results of Run 1, the CPU time shows an almost optimal computational complexity of the MGOPT approach.

Table 3: Numerical results for Run 2.

ν	mesh	$\|u^\ell - u^\ell(1)\|_{L^2}$	time (sec)
	33×33	2.467e-05	0.6
	65×65	2.660e-05	1.8
10^{-4}	129×129	9.968e-05	11.4
	257×257	8.441e-05	42.0
	513×513	8.923e-05	221.3

4. CONCLUSIONS

Multigrid optimization scheme with finite difference discretization for solving semilinear elliptic optimal control problems is presented. In particular, the control-unconstrained and control-constrained cases are considered. As an optimization algorithm for the MGOPT method, we consider the gradient projection method. The results of the numerical experiments show that this multigrid strategy provide a multigrid computational efficiency. A topic which can be considered for future research is the appropriate use of other types of optimization algorithms for faster convergence results.

5. REFERENCES

[1] Borzì, A. and Kunisch, K.: A multigrid scheme for elliptic constrained optimal control problems. Comput. Optim. Appl.**31**(3), 309–333 (2005)

[2] Borzì, A. and Schulz, V.: Multigrid methods for PDE optimization. SIAM Review **51**(2), 361–395 (2009)

[3] Dreyer, T. and Maar, S. and Schulz, V.: Multigrid optimization in applications. Journal of Computational and Applied Mathematics **120**(1-2), 67–84 (2000)

[4] Lass, O. and Vallejos, M. and Borzì, A. and Douglas, C.C.: Implementation and analysis of multigrid schemes with finite elements for elliptic optimal control problems. J. Computing **84**(1-2), 27–48 (2009)

[5] Kelley, C.T.: Iterative Methods for Optimization. Kluwer, New York (1987)

[6] Lewis, R.M. and Nash, S.: Model problems for the multigrid optimization of systems governed by differential equations. SIAM J. Sci. Comput. **26**(6), 1811–1837 (2005)

[7] Lions, J.L.: Optimal control of systems governed by partial differential equations. Springer, Berlin (1971)

[8] Nash, S.: A multigrid approach to discretized optimization problems. Optim. Methods Softw. **14**(1-2), 99–116 (2000)

[9] Oh, S. and Bouman, C. and Webb, K.J.: Multigrid tomographic inversion with variable resolution data and image spaces. IEEE Trans. Image Process. **15**(9), 2805–2819 (2006)

[10] Oh, S. and Milstein, A. and Bouman, C. and Webb, K.J.: A general framework for nonlinear multigrid inversion. IEEE Trans. Image Process. **14**(1), 125–140 (2005)

[11] Vallejos, M. and Borzì, A.: Multigrid methods for linear elliptic optimal control problems. Numerical Mathematics and Advanced Applications. Proceedings of ENUMATH 2007, the 7th European Conference on Numerical Mathematics and Advanced Applications, Graz, Austria, September 2007, Springer-Verlag, Heidelberg, 653-660, (2008)