A Posteriori Estimates for Cost Functionals of Optimal Control Problems

Alexandra Gaevskaya,[1] Ronald H.W. Hoppe[1,2] and Sergey Repin[3]

[1] Institute of Mathematics, Universität Augsburg, D-86159 Augsburg, Germany
gaevskaya@math.uni-augsburg.de
[2] Department of Mathematics, University of Houston, Houston, TX 77204-3008, USA
rohop@math.uh.edu
[3] St. Petersburg Division of the Steklov Mathematical Institute, Russian Academy of Sciences, 191011 St. Petersburg, Russia
repin@pdmi.ras.ru

1 Introduction

A posteriori analysis has become an inherent part of numerical mathematics. Methods of a posteriori error estimation for finite element approximations were actively developed in the last two decades (see, e.g., [1, 2, 3, 12] and the references therein). For problems in the theory of optimization, these methods started receiving attention much later. In particular, for optimal control problems governed by PDEs the literature on this matter is rather scarce. In this work, we present a new approach to a class of optimal control problems associated with elliptic type partial differential equations. In the framework of this approach, we obtain directly computable upper bounds for the cost functionals of the respective optimal control problems.

Let $\Omega \in \mathbb{R}^n$ be a Lipschitz domain with boundary $\Gamma := \partial\Omega$.

Problem 1. Given $\psi \in L_\infty(\Omega)$, $y^d \in L_2(\Omega)$, $u^d \in L_2(\Omega)$, $f \in L_2(\Omega)$, and $a \in \mathbb{R}_+$, consider the distributed control problem

$$\text{minimize } J(y(v),v) := \frac{1}{2}\,\|y-y^d\|^2 + \frac{a}{2}\,\|v-u^d\|^2 \tag{1a}$$

$$\text{over } (y,v) \in H_0^1(\Omega) \times L^2(\Omega)\,,$$

$$\text{subject to } -\Delta\, y = v + f \quad \text{a.e. in } \Omega\,, \tag{1b}$$

$$v \in K := \{v \in L^2(\Omega) \mid v \le \psi \text{ a.e. in } \Omega\}\,. \tag{1c}$$

The function y^d is given and presents the desired shape of the state function y, whereas u^d presents the desired control. It is well-known that under the above assumptions Problem 1 has a unique solution (see, e.g. [9]).

There exist many different approaches to optimal control problems of this type. The numerical solution of optimal control problems is usually based on applying specific iterative schemes to the system of optimality conditions, e.g., active set strategies or interior point methods (cf., e.g., [6, 7] and the references therein). Adaptive techniques for optimal control problems governed by PDEs are presented in [4] and [8].

In this work, we follow another approach which is based on so-called functional type a posteriori error estimates. To explain the meaning of such estimates, as a model problem we consider Poisson's equation with homogeneous boundary conditions

$$-\Delta y = v + f \qquad \text{in} \quad \Omega \,, \tag{2a}$$

$$y = 0 \qquad \text{on} \quad \Gamma \,, \tag{2b}$$

which describes the dependence between the control and the state in the optimal control problem (1a)-(1c). Let $\widetilde{y}$ be any function from the admissible set $Y := H_0^1(\Omega)$ which we view as an approximation of the solution of the elliptic problem (2a)-(2b). It was shown (see, e.g., [10] and [11]) that the error of the approximation $\widetilde{y}$ satisfies the following estimate:

$$\|\nabla(y(v) - \widetilde{y})\| \le \|\tau - \nabla\widetilde{y}\| + C_\Omega \|\mathrm{div}\tau + v + f\| \,. \tag{3}$$

Here, C_Ω is the constant in the Friedrichs inequality

$$\|w\| \;\le\; C_\Omega \;\|\nabla w\| \,, \quad w \in H_0^1(\Omega) \tag{4}$$

for the domain Ω and τ is an arbitrary function from the functional class $\Sigma := H_{\mathrm{div}}(\Omega, \mathbb{R}^n)$. Mathematical justifications of functional type a posteriori estimates and their analysis can be found in the above cited literature. Below, we recall the main properties of such estimates:

- For any approximation $\widetilde{y} \in Y$, the right–hand side of (3) gives an upper bound of the error in the natural energy norm of the problem considered;
- Its value is equal to zero if and only if $\widetilde{y}$ coincides with $y(v)$ and $\tau = \nabla y(v)$;
- The estimate is consistent in the sense that its value tends to zero for any sequences $\{\widetilde{y}_k\}$ and $\{\tau_k\}$, converging to the exact solution y and its gradient ∇y, respectively;
- The estimate is exact in the sense that there exists a function τ such that equality holds true;
- The estimate does not depend on the mesh parameters and only contains a global constant.

The function τ in the expression of the error majorant (3) serves as an image of the exact flux $\nabla y(v)$. It is easy to observe that two terms of the majorant represent the respective errors in the *constitutive relation* $\tau = \nabla y(v)$ and in the *equilibrium equation* $\mathrm{div}\tau + v + f = 0$.

In this paper, we apply this estimate in order to reformulate the original optimal control problem. As a result, we obtain a directly computable and *guaranteed* majorant for the cost functional. Besides, we prove that the sequences of approximate state and control functions, computed by the minimization of the majorant, converge to the exact state and control functions.

2 Majorants for the cost functional

One of the major difficulties in (1a)-(1c) is that the state and control functions must satisfy the equality constraint presented by the boundary-value problem for an elliptic PDE.

Let $v \in K$ and $y \in Y$ be two functions related by the differential equation (1b). For this pair, the cost functional is as follows:

$$J(y(v), v) \ := \ \frac{1}{2}\|y - y^d\|^2 \ + \ \frac{a}{2}\|v - u^d\|^2 \ .$$

Let $\widetilde{y} \in Y$ be some approximation of y so that we may include it in the first term of the cost functional. By the triangle and Friedrichs inequalities, we obtain the estimate

$$J(y(v), v) \leq \ \frac{1}{2}\left(\|\widetilde{y} - y^d\| + C_\Omega\|\nabla(y - \widetilde{y})\|\right)^2 \ + \ \frac{a}{2}\|v - u^d\|^2 \ . \tag{5}$$

Now, using the error majorant (3) we can estimate the weak norm of the error and substitute it to the estimate of the cost functional (5). By this procedure, we exclude the explicit entry of the exact solution y of (2a)-(2b) from our estimate and arrive at the relation

$$J(y(v), v) \leq \ \frac{1}{2}\left(\|\widetilde{y} - y^d\| + C_\Omega\|\nabla\widetilde{y} - \tau\| + C_\Omega^2\|\mathrm{div}\tau + v + f\|\right)^2 + \frac{a}{2}\|v - u^d\|^2 \ .$$

However, from a computational point of view it is desirable to reformulate this estimate such that the right–hand side is given by a quadratic functional. For this purpose, we introduce parameters α, $\beta > 0$ and obtain the following upper bound (hereafter called *the majorant*):

$$J(y(v), v) \leq \ J^{\oplus}(\alpha, \beta; \widetilde{y}, \tau, v) \ , \qquad \forall v \in K \ . \tag{6}$$

Here,

$$J^{\oplus}(\alpha, \beta; \widetilde{y}, \tau, v) := \frac{1+\alpha}{2}\|\widetilde{y} - y^d\|^2 + \frac{(1+\alpha)(1+\beta)}{2\alpha} C_\Omega^2\|\tau - \nabla\widetilde{y}\|^2 + \tag{7}$$
$$+ \frac{(1+\alpha)(1+\beta)}{2\alpha\beta} C_\Omega^4\|\mathrm{div}\tau + v + f\|^2 + \frac{a}{2}\|v - u^d\|^2 \ ,$$

where $\widetilde{y} \in Y$ and τ is an arbitrary function in Σ.

Remark 1. A similar upper estimate can be derived for the optimal control problem with the cost functional

$$\boldsymbol{J}(y,v) = \frac{1}{2}\|\nabla y - \sigma^d\|^2 + \frac{a}{2}\|u - u^d\|^2 \ ,$$

where the vector-valued function σ^d is given and presents the desired gradient of the state function.

Let us consider the majorant as a functional that generates a new minimization problem

Problem 1^***.** Given $\psi \in L_\infty(\Omega)$, $y^d \in L_2(\Omega)$, $u^d \in L_2(\Omega)$, $f \in L_2(\Omega)$, and $a \in \mathbb{R}_+$,

$$\text{minimize } J^\oplus(\alpha,\beta;\widetilde{y},\tau,v) \tag{8a}$$
$$\text{over } v \in K,\ \widetilde{y} \in Y,\ \tau \in \Sigma,\ \alpha,\beta \in \mathbb{R}_+ \ ,$$
$$K \ := \ \{v \in L^2(\Omega) \mid v \le \psi \text{ a.e. in } \Omega\} \ . \tag{8b}$$

We see that in this problem the differential equation (which in (1a)-(1c) defines the respective admissible set) does not appear explicitly. In (8a)-(8b), the functions τ, $\widetilde{y}$ and v act as independent variables. In the next section, we present properties of the majorant (7) and show that Problem 1^* and Problem 1 have one and the same exact lower bound attained on the same state and control functions.

Remark 2. It is worth noting that the majorant $J^\oplus(\alpha,\beta;\widetilde{y},\tau,v)$ can be used to find *guaranteed upper bounds* for the cost functional when the minimization problem is solved by known methods. Indeed, since the functions $\widetilde{y}$ and v are arbitrary, we can take them as approximate solutions computed by some optimization procedure and minimize the majorant w.r.t. the function τ and the parameters β and α. The respective value $J^\oplus$ will represent the guaranteed upper bound for the value of the cost functional.

3 Properties of majorants

Theorem 1. *The exact lower bound of the majorant* (7) *coincides with the optimal value of the cost functional of the problem* (1a)-(1c)*, i.e,*

$$\inf_{\substack{\widetilde{y}\in Y,\tau\in\Sigma,\\ v\in K,\alpha,\beta\in\mathbb{R}_+}} J^\oplus(\alpha,\beta;\widetilde{y},\tau,v) = J(y(u),u) \ .$$

The infimum of $J^\oplus$ *is attained for* $v = u$, $\widetilde{y} = y(u)$, $\tau = \nabla y(u)$.

This property means that our transformation of the original problem is mathematically correct in the sense that the new problem is solvable and has the same lower bound as the original one.

Let $\{V_k\}_{k=1}^\infty$, $\{Y_k\}_{k=1}^\infty$ and $\{\Sigma_k\}_{k=1}^\infty$ be sequences of finite-dimensional subspaces that are limit dense in $V := L^2(\Omega)$, Y and Σ, respectively. The discrete

control constraints are given by $K_k := V_k \cap K$. It is not difficult to show that K_k is limit dense in K.
We define the sequence of numbers

$$J_k^{\oplus} := J^{\oplus}(\alpha_k, \beta_k; \widetilde{y}_k, \tau_k, v_k) = \inf_{\substack{\widetilde{y} \in Y_k, \tau \in \Sigma_k, \\ v \in K_k, \alpha, \beta \in \mathbb{R}_+}} J^{\oplus}(\alpha, \beta; \widetilde{y}, \tau, v) \;, \tag{9}$$

which is obtained by solving the problem on sequences of the selected finite–dimensional subspaces.

Theorem 2. *If K_k, Y_k and Σ_k are limit dense in K, Y, and Σ, respectively, then*

(*i*) $J_k^{\oplus} \to J(y(u), u)$ *as* $k \to \infty$;
(*ii*) *the sequence* $\{y(v_s), v_s\}$ *converges to the exact solution of the control problem* $\{y(u), u\}$ *in* $Y \times K$.

The theorem shows that a numerical strategy based upon using the majorant produces sequences of control and state functions which provide a value of the cost functional as close to the value $J(y(u), u)$ as it is required. Moreover, the respective sequences of control and state functions tend to the desired solution of the original problem.

4 Practical implementation

In this section, we briefly discuss the practical implementation of the numerical strategy based on the majorants.

4.1 Discretization of the problem

In the resultes exposed below, we restrict ourselves to the case when the problem is solved by usual finite element approximations on a simplicial mesh which is the same for all functions involved. Let $\mathcal{T}_h(\Omega)$ denote such a shape-regular simplicial triangulation of Ω. For the state function, we use continuous piecewise affine approximations $\widetilde{y}_h \in Y_h$ vanishing on the boundary Γ, whereas for the control $v \in K$ we use piecewise constant approximations $v_h \in K_h$ where K_h is chosen such that $K_h \subset K$. The vector–valued functions $\tau \in \Sigma$ are approximated by piecewise affine functions $\tau_h \in \Sigma_h$.

4.2 Minimization algorithm

To obtain a sharp upper bound of the cost functional, we minimize the majorant $J^{\oplus}(\alpha, \beta; \widetilde{y}_h, \tau_h, v_h)$ over $(\widetilde{y}_h, \tau_h, v_h) \in Y_h \times \Sigma_h \times K_h$ and $\alpha, \beta \in \mathbb{R}^+$.

The numerical results presented below have been obtained using the following algorithm:

Step 1. Initialization. Set $i = 0$, define α^0, β^0, v_h^0, $\widetilde{y}_h^0$.
Step 2. Minimize $J^{\oplus}(\alpha^i, \beta^i; \widetilde{y}_h, \tau_h, v_h)$ over $(\widetilde{y}_h, \tau_h, v_h) \in Y_h \times \Sigma_h \times K_h$.
Step 3. Minimize $J^{\oplus}(\alpha, \beta; \widetilde{y}_h^{i+1}, \tau_h^{i+1}, v_h^{i+1})$ w.r.t. $\beta, \alpha \in \mathbb{R}_+$. Set $i = i + 1$.

Steps 2 and 3 are repeated until

$$\frac{|J_i^{\oplus} - J_{i-1}^{\oplus}|}{J_i^{\oplus}} + \frac{\|v_h^i - v_h^{i-1}\|}{\|v_h^i\|} + \frac{\|\nabla(\widetilde{y}_h^i - \widetilde{y}_h^{i-1})\|}{\|\nabla \widetilde{y}_h^i\|} > \epsilon ,$$

where ϵ is a given tolerance and $J_i^{\oplus} = J^{\oplus}(\alpha^i, \beta^i; \widetilde{y}_h^i, \tau_h^i, v_h^i)$.

5 Numerical experiments

The method described in the previous sections has been numerically tested on a set of various optimal control problems. In all examples, it has been observed that the sequences of computed upper bounds of the cost functionals rapidly converge to the exact lower bound whose value has been computed at high accuracy. Also, it has been observed that the sequences of the state and control functions converge to the exact ones.

Below, we show these results for the model problem in $\Omega = (0, 1)^2$. In this case, $C_\Omega = \frac{1}{\sqrt{2}\pi}$.

The efficiency of the approach is measured by three quantities. The index

$$I = J^{\oplus} / J(y, u)$$

shows the relation between the value of majorant computed for the control function v and the exact lower bound of the cost functional $J(y, u)$. The quantities

$$\eta_y = (\|y - \widetilde{y}\|_{H^1} / \|y\|_{H^1}) * 100\% , \quad \eta_u = (\|v - u\| / \|u\|) * 100\% ,$$

represent the *relative errors* in the state and control functions, respectively.

Example

As an example we take the problem from [6] with the following data:
$a = 0.01$, $\psi(x, y) = 1$, $f(x, y) = 0$, $u^d(x, y) = 0$ and

$$y^d(x, y) = \begin{cases} 200(x - 0.5)^2(1 - y)yx , & x \leq 0.5 , \\ 200(x - 0.5)^2(1 - y)y(x - 1) , & else . \end{cases}$$

The exact solution of this optimal control problem is unknown. Therefore, in order to analyze the efficiency of the method, we have computed a ‘reference solution' using a mesh much finer than those used in the actual computations. For this task, we have used the primal-dual active set strategy (cf., e.g., [5]). The reference value of the cost functional in this case is $J(y,u) = 9.5838 \cdot 10^{-2}$.

The discrete problem has been solved for various uniform meshes with N nodes. Table 1 shows the relative errors in the state and control functions and the index I. In Figure 1, we depict values of the majorant with respect to the minimization time ($N = 1089$). In this figure, the horizontal line shows $J(y(u),u)$ (actual value of the cost functional) whereas the rapidly decaying curve reflects the reduction of the computable upper bound given by the majorant. The desired tolerance $\epsilon = 10^{-4}$ was achieved after $i = 16$ iterations. Approximations $(\widetilde{y}_h, v_h)$ obtained by the algorithm and the reference state and control functions are displayed in Figure 2.

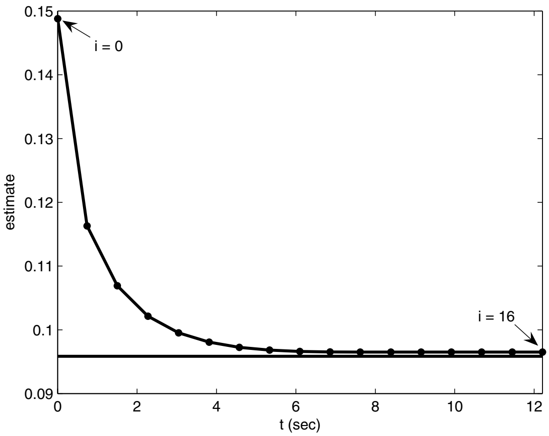

Fig. 1. Reduction of the upper bound of the cost functional w. r. t. CPU time.

Table 1. Index I and relative errors in the state and control.

N	$\eta_y, \%$	$\eta_u, \%$	I
25	67.51	54.39	1.050
81	31.50	25.23	1.029
289	14.59	12.07	1.014
1089	7.55	6.49	1.007
4225	4.67	4.18	1.003
16641	3.65	3.39	1.002

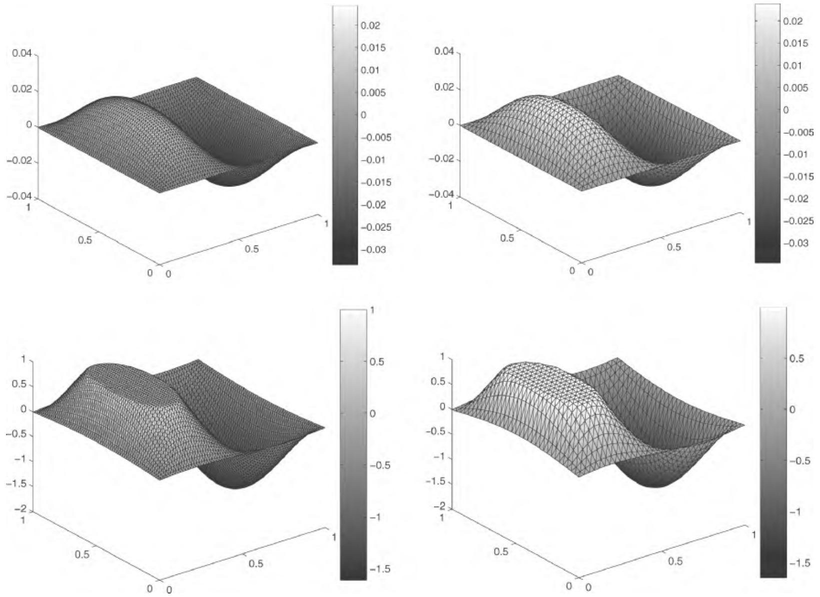

Fig. 2. Exact state y (upper left) and approximate state $\widetilde{y}_h$ (upper right), exact control u (lower left) and approximate control v_h (lower right).

References

1. Ainsworth, M., Oden, J.T.: A posteriori error estimation in finite element analysis. Wiley, New York (2000)
2. Babuška, I., Strouboulis, T.: The finite element method and its reliability, Clarendon Press, Oxford (2001)
3. Bangerth, W., Rannacher, R.: Adaptive finite element methods for differential equations. Birkhäuser, Berlin (2003)
4. Becker, R., Rannacher, R.: An optimal control approach to error estimation and mesh adaptation in finite element methods. In: Acta Numerica (A. Iserles, ed.), 10: 1–102, Cambridge University Press, (2001)
5. Bergounioux, M., Haddou, M, Hintermüller, M., Kunisch, K.: A comparison of a Moreau-Yosida-Based Active Set Strategy and Interior Point Methods for Constrained Optimal Control Problems. SIAM J. Optim., **11**, No. 2, 495–521 (2000)
6. Bergounioux, M., Ito, K., Kunisch, K.: Primal-dual strategy for constrained optimal control problems, SIAM J. Control Optim., **37**, 1176–1194 (1999)
7. Bonnans, J.F., Pola, C., Rebaï, R.: Perturbed path following interior point algorithms, Optim. Methods Softw., 11–12, 183–210 (1999)
8. Li, R., Liu, W., Ma, H., Tang, T.: Adaptive finite element approximation for distributed optimal control problems. SIAM J. Control Optim. **41**, 5, 1321–1349 (2002)

9. Lions, J. L.: Optimal Control of Systems Governed by Partial Differential Equations. Springer, Berlin–Heidelberg–New York, (1971)
10. Repin, S.: A posteriori error estimation for nonlinear variational problems by duality theory, Zapiski Nauchn. Semin. POMI, **243**, 201–214 (1997)
11. Repin, S.: A posteriori error estimation for variational problems with uniformly convex functionals, Math. Comp., **69**, 230, 481–500 (2000)
12. Verfürth, R.: A review of a posteriori error estimation and adaptive mesh-refinement techniques, Wiley and Sons, Teubner, New York, (1996)