ÖAW Austrian Academy of Sciences

Doctoral Program
Computational Mathematics
Numerical Analysis and Symbolic Computation

JKU JOHANNES KEPLER UNIVERSITY LINZ

Multigrid Methods for Elliptic Optimal Control Problems with Neumann Boundary Control

Stefan Takacs Walter Zulehner

DK-Report No. 2009-01 10 2009

A–4040 LINZ, ALTENBERGERSTRASSE 69, AUSTRIA

Supported by

Austrian Science Fund (FWF) Upper Austria

FWF Der Wissenschaftsfonds.

LAND OBERÖSTERREICH

DK sponsors:

- **Johannes Kepler University Linz** (JKU)
- **Austrian Science Fund** (FWF)
- **Upper Austria**

Multigrid Methods for Elliptic Optimal Control Problems with Neumann Boundary Control

Stefan Takacs, Walter Zulehner

Abstract In this article we discuss multigrid methods for solving discretized optimality systems for elliptic optimal control problems. We concentrate on a model problem of tracking type with Neumann boundary control. The proposed approach is based on the optimality system, which, for the model problem, leads to a linear system for the state y, the control u and the adjoined state p. An Uzawa type smoother is used for the multigrid method. Moreover, we will compare this approach with standard smoothers, like damped Jacobi iteration applied to the normal equation of the Kuhn-Tucker system. A rigorous multigrid convergence analysis is presented for both smoothers.

1 Formulation of the model problem

We discuss the solution of optimal control problems of tracking type. Let Ω be a bounded convex and polygonal domain in $\mathbb{R}^2$ with boundary $\partial\Omega$. We want to minimize the functional

$$J(y,u) = \frac{1}{2}\|y - y_D\|^2_{L^2(\Omega)} + \frac{\gamma}{2}\|u\|_{L^2(\partial\Omega)}, \tag{1}$$

where y is the state variable and u is the control variable. Here, y_D is given and $\gamma > 0$ is some fixed regularization or cost parameter.

Stefan Takacs
Doctoral Program Computational Mathematics, Johannes Kepler University Linz, Austria, e-mail: stefan.takacs@dk-compmath.jku.at

Walter Zulehner
Institute of Computational Mathematics, Johannes Kepler University Linz, Austria, e-mail: zulehner@numa.uni-linz.ac.at

The work is supported by the Austrian Science Fund (FWF) under grant W1214/DK12.

The minimization is done subject to the following constraint: the state variable fulfills some elliptic boundary value problem (BVP) with Neumann boundary data u. For this paper we restrict ourselves to the simple case of a Laplace-type equation:

$$-\Delta y + y = 0 \text{ in } \Omega \qquad \text{and} \qquad \frac{\partial y}{\partial n} = u \text{ on } \partial\Omega. \tag{2}$$

The functions y and u live in standard Lebesgue and Sobolev spaces:

$$y \in H^1(\Omega) \qquad \text{and} \qquad u \in L^2(\partial\Omega). \tag{3}$$

Observe that for this setting the BVP is uniquely solvable in y for every given control u. The BVP (2) can be written in variational form:

$$(y,p)_{H^1(\Omega)} - (u,p)_{L^2(\partial\Omega)} = 0 \qquad \text{for all } p \in H^1(\Omega).$$

Based on the variational formulation, we can introduce the Lagrange functional

$$\mathscr{L}(y,u,p) = \frac{1}{2}\|y - y_D\|^2_{L^2(\Omega)} + \frac{\gamma}{2}\|u\|^2_{L^2(\partial\Omega)} + (y,p)_{H^1(\Omega)} - (u,p)_{L^2(\partial\Omega)}.$$

Solving the original optimal control problem is equivalent to finding a saddle point of the Lagrange functional which leads to the first order optimality conditions (the Karush-Kuhn-Tucker system), given by: Find $(y,u,p) \in X := H^1(\Omega) \times L^2(\partial\Omega) \times H^1(\Omega)$ such that

$$\begin{array}{rrrl} (y,\widetilde{y})_{L^2(\Omega)} & & + (p,\widetilde{y})_{H^1(\Omega)} & = (y_D,\widetilde{y})_{L^2(\Omega)} \\ & \gamma\,(u,\widetilde{u})_{L^2(\partial\Omega)} & - (p,\widetilde{u})_{L^2(\partial\Omega)} & = 0 \\ (y,\widetilde{p})_{H^1(\Omega)} & - (u,\widetilde{p})_{L^2(\partial\Omega)} & & = 0 \end{array} \tag{4}$$

holds for all $(\widetilde{y},\widetilde{u},\widetilde{p}) \in X$.

The optimality system can be rewritten as one single variational equation: Find $x \in X$ such that

$$\mathsf{a}(x,\widetilde{x}) = \langle \mathscr{F}, \widetilde{x} \rangle$$

holds for all $\widetilde{x} \in X$, where the bilinear form a and the linear form $\mathscr{F}$ are given by

$$\begin{aligned} \mathsf{a}((y,u,p),(\widetilde{y},\widetilde{u},\widetilde{p})) &:= (y,\widetilde{y})_{L^2(\Omega)} + \gamma\,(u,\widetilde{u})_{L^2(\partial\Omega)} + (p,\widetilde{y})_{H^1(\Omega)} \\ &\quad - (p,\widetilde{u})_{L^2(\partial\Omega)} + (y,\widetilde{p})_{H^1(\Omega)} - (u,\widetilde{p})_{L^2(\partial\Omega)} \\ \langle \mathscr{F}, (\widetilde{y},\widetilde{u},\widetilde{p}) \rangle &:= (y_D,\widetilde{y})_{L^2(\Omega)}. \end{aligned}$$

The optimality system has a natural 2-by-2 block-structure:

$$\begin{array}{ll} a((y,u),(\widetilde{y},\widetilde{u})) + b((\widetilde{y},\widetilde{u}),p) & = (y_D,\widetilde{y})_{L^2(\Omega)} \\ b((y,u),\widetilde{p}) & = 0 \end{array}$$

where

$$a((y,u),(\widetilde{y},\widetilde{u})) := (y,\widetilde{y})_{L^2(\Omega)} + \gamma\,(u,\widetilde{u})_{L^2(\partial\Omega)}$$
$$b((y,u),\widetilde{p}) := (y,\widetilde{p})_{H^1(\Omega)} - (u,\widetilde{p})_{L^2(\partial\Omega)}.$$

Observe that a is a symmetric and bounded bilinear form which is coercive on the kernel of b and b is a bounded bilinear form. Moreover b fulfills the inf-sup-condition

$$\inf_{0\neq\widetilde{p}\in H^1(\Omega)} \sup_{0\neq(y,u)\in H^1(\Omega)\times L^2(\partial\Omega)} \frac{b((y,u),\widetilde{p})}{\|(y,u)\|_{H^1(\Omega)\times L^2(\partial\Omega)}\,\|\widetilde{p}\|_{H^1(\Omega)}} \geq C > 0,$$

which can be seen by plugging in $(y,u) := (\widetilde{p},0)$. Using Brezzi's theorem we obtain:

Lemma 1. *Let $\gamma > 0$ be fixed. The problem* (4) *is well posed in the space X, i.e., there are constants $\underline{C} > 0$ and $\overline{C}$ such that*

$$\underline{C}\,\|x\|_X \leq \sup_{0\neq\widetilde{x}\in X} \frac{\mathsf{a}(x,\widetilde{x})}{\|\widetilde{x}\|_X} \leq \overline{C}\,\|x\|_X$$

holds for all $x \in X$. For every right-hand-side $\mathscr{F} \in X^*$ the problem to find $x \in X$, such that

$$\mathsf{a}(x,\widetilde{x}) = \langle \mathscr{F},\widetilde{x}\rangle$$

holds for all $\widetilde{x} \in X$, has a unique solution.

The discretization is done by standard techniques. For the model problem we use a family of meshes which is obtained based on some coarsest triangular mesh (grid level $k = 0$) and uniform refinement. For $k = 0,1,\ldots$ we denote the size of the largest edge of the triangulation by h_k. Due to the fact that we have uniform refinement $h_k = 2^{-k}h_0$ holds.

The space of discretized functions $X_k = Y_k \times U_k \times P_k$ is constructed by the Courant element: $Y_k = P_k$ is the set of continuous and piecewise linear functions. U_k is the set of continuous and piecewise linear functions on the boundary.

This setting allows us to show the statement of lemma 1 also if X is replaced by X_k.

Using the standard nodal basis, we can rewrite the optimality system (4) in matrix-vector notation as follows:

$$\underbrace{\left(\begin{array}{cc|c} M_k & 0 & K_k \\ 0 & \gamma M_{\Gamma\Gamma k} & -M_{\Gamma\Omega k}^T \\ \hline K_k & -M_{\Gamma\Omega k} & 0 \end{array}\right)}_{\mathscr{A}_k := \left(\begin{smallmatrix} A_k & B_k^T \\ B_k & 0 \end{smallmatrix}\right) :=} \underbrace{\left(\begin{array}{c} \underline{y}_k \\ \underline{u}_k \\ \underline{p}_k \end{array}\right)}_{\underline{x}_k :=} = \underbrace{\left(\begin{array}{c} \underline{g}_k \\ 0 \\ 0 \end{array}\right)}_{\underline{f}_k :=} \tag{5}$$

with mass matrices M_k, $M_{\Gamma\Omega k}$, $M_{\Gamma\Gamma k}$ and the stiffness matrix K_k. The symbols $\underline{y}_k, \underline{u}_k, \ldots$ denote the coordinate vectors of the corresponding functions $y_k, u_k, \ldots$ with respect to the nodal basis.

2 Multigrid solvers for saddle point problems

Starting from an initial approximation $\underline{x}_k^{(0)}$ one step of the multigrid method for solving the discretized equation (5) on grid level k is given by:

- Apply ν smoothing steps

$$\underline{x}_k^{(0,m)} := \underline{x}_k^{(0,m-1)} + \hat{\mathscr{A}}_k^{-1}(\underline{f}_k - \mathscr{A}_k\, \underline{x}_k^{(0,m-1)}) \qquad \text{for } m = 1,\dots,\nu \tag{6}$$

  with $\underline{x}_k^{(0,0)} := \underline{x}_k^{(0)}$.

- Apply coarse-grid correction
  - Compute defect and restrict to coarser grid
  - Solve problem on coarser grid
  - Prolongate and add result

  If the problem on the coarser grid is solved exactly, then we obtain

$$\underline{x}_k^{(1)} := \underline{x}_k^{(0,\nu)} + I_{k-1}^k \mathscr{A}_{k-1}^{-1} I_k^{k-1}(\underline{f}_k - \mathscr{A}_k\, \underline{x}_k^{(0,\nu)})$$

  for the next iterate (two-grid method).

In practice the problem on grid level $k-1$ is done by applying one (V-cycle) or two (W-cycle) steps of the multigrid method, recursively. On grid level $k=0$ the problem is solved exactly. The convergence of the two-grid method implies the convergence of the W-cycle multigrid method under weak assumptions.

The intergrid-transfer operators I_{k-1}^k and I_k^{k-1} are chosen in a canonical way: we use the canonical embedding for I_{k-1}^k and its adjoined as restriction operator I_k^{k-1}.

The smoother will be specified in the section.

The classical convergence theory of multigrid methods is based on two properties:

- Smoothing property:

$$|||\underline{x}_k^{(0,\nu)} - \underline{x}_k|||_{2,k} \le \eta(\nu)\,|||\underline{x}_k^{(0)} - \underline{x}_k|||_{0,k} \tag{7}$$

  should hold for some function $\eta(\nu)$ independent of k with $\lim_{\nu\to\infty}\eta(\nu) = 0$.
- Approximation property:

$$|||\underline{x}_k^{(1)} - \underline{x}_k|||_{0,k} \le C_A\,|||\underline{x}_k^{(0,\nu)} - \underline{x}_k|||_{2,k} \tag{8}$$

  should hold for some constant $C_A > 0$ independent of k.

We have the freedom to choose two norms in (7) and (8). This is done in the following way:

We first introduce the norm $\|\cdot\|_{X_k^-}$ by replacing in $\|\cdot\|_X$ all H^1-norms by L^2-norms scaled by the factor h_k^{-1}:

$$\|(y_k,u_k,p_k)\|^2_{X_k^-} := h_k^{-2}\|y_k\|^2_{L^2(\Omega)} + \|u_k\|^2_{L^2(\partial\Omega)} + h_k^{-2}\|p_k\|^2_{L^2(\Omega)}.$$

This corresponds to a norm for $(\underline{y}_k,\underline{u}_k,\underline{p}_k)$ involving mass matrices. If the mass matrices are replaced by properly scaled identity matrices, we obtain the desired norm $|||\cdot|||_{0,k}$, given by:

$$|||(\underline{y}_k,\underline{u}_k,\underline{p}_k)|||^2_{0,k} := \Bigg(\underbrace{\begin{pmatrix} I & & \\ & h_k I & \\ & & I \end{pmatrix}}_{\mathscr{L}_k :=} \begin{pmatrix} \underline{y}_k \\ \underline{u}_k \\ \underline{p}_k \end{pmatrix}, \begin{pmatrix} \underline{y}_k \\ \underline{u}_k \\ \underline{p}_k \end{pmatrix} \Bigg)_{\ell^2}.$$

According to standard techniques, we choose $|||\cdot|||_{2,k}$ as residual norm corresponding to $|||\cdot|||_{0,k}$, i.e.,

$$|||\underline{x}_k|||_{2,k} := \sup_{\widetilde{\underline{x}}_k \in \mathbb{R}^n} \frac{(\mathscr{A}_k\,\underline{x}_k, \widetilde{\underline{x}}_k)}{|||\widetilde{\underline{x}}_k|||_{0,k}}.$$

3 Construction of Smoothers

Next we construct two simple iterative methods fulfilling the smoothing property (7).

The first kind of smoothers, we want to discuss, are *Uzawa-type smoothers* which have already been successfully applied to distributed control problems (e.g. [6]). These methods can also be extended to Neumann boundary control problems.

We construct the preconditioner $\hat{\mathscr{A}}_k$ in (6) based on the block-LU-factorization $\mathscr{A}_k$: We have

$$\mathscr{A}_k = \begin{pmatrix} A_k & B_k^T \\ B_k & 0 \end{pmatrix} = \begin{pmatrix} A_k & 0 \\ B_k & -S_k \end{pmatrix} \begin{pmatrix} I & A_k^{-1}B_k^T \\ 0 & I \end{pmatrix},$$

where $S_k := B_k A_k^{-1} B_k^T$ is the Schur-compliment. Based on this decomposition we define the preconditioner $\hat{\mathscr{A}}_k$ by replacing A_k and S_k by diagonal matrices $\hat{A}_k$ and $\hat{S}_k$:

$$\hat{\mathscr{A}}_k := \begin{pmatrix} \hat{A}_k & 0 \\ B_k & -\hat{S}_k \end{pmatrix} \begin{pmatrix} I & \hat{A}_k^{-1}B_k^T \\ 0 & I \end{pmatrix}.$$

Then $\hat{\mathscr{A}}_k$ can be inverted easily.

The main issue is how to choose the matrices $\hat{A}_k$ and $\hat{S}_k$. Normally, one would expect to choose $\hat{A}_k$ as the diagonal part of A_k. Instead we propose to choose for $\hat{A}_k$ the left-upper block of $\mathscr{L}_k$. For $\hat{S}_k$ we take the corresponding inexact Schur-compliment. This leads to

$$\hat{A}_k := \frac{1}{\omega}\begin{pmatrix} I & & \\ & h_k I & \\ & & I \end{pmatrix} \qquad \text{and} \qquad \hat{S}_k := \frac{1}{\sigma}\,\mathrm{diag}(B_k \hat{A}_k^{-1} B_k^T)$$

with additional damping parameters ω and σ which are chosen independent of k and such that

$$\hat{A}_k \geq A_k \qquad \text{and} \qquad \hat{S}_k \geq S_k \tag{9}$$

holds. This is possible, as we can choose ω and σ equal to the reciprocal of the number of non-zero entries of A_k or S_k, respectively, which are bounded.

An *alternative approach* is to construct smoothers that are based on the *normal equation* $\mathscr{A}_k^* \mathscr{A}_k \underline{x}_k = \mathscr{A}_k^* \underline{f}_k$, where $\mathscr{A}_k^*$ denotes the adjoined of $\mathscr{A}_k$ with respect to the inner product corresponding to the norm $|||\cdot|||_{0,k}$. Using $\mathscr{L}_k$ we can formulate this in standard matrix-vector notation:

$$\mathscr{A}_k^T \mathscr{L}_k^{-1} \mathscr{A}_k \underline{x}_k = \mathscr{A}_k^T \mathscr{L}_k^{-1} \underline{f}_k.$$

We can apply some standard smoother which is applicable to symmetric positive definite problems, like the damped Jacobi iteration:

$$\underline{x}_k^{(0,m)} := \underline{x}_k^{(0,m-1)} + \tau \operatorname{diag}(\mathscr{A}_k^T \mathscr{L}_k^{-1} \mathscr{A}_k)^{-1} \mathscr{A}_k^T \mathscr{L}_k^{-1} (\underline{f}_k - \mathscr{A}_k \underline{x}_k^{(0,m-1)}),$$

where the parameter τ is chosen such that the smallest eigenvalue of the iteration matrix is non-negative.

4 Convergence analysis

A convergence analysis for *distributed control problems* was already done based on approximation and smoothing property, e.g., in [6]. The approximation property was shown following the ideas from [2].

Alternative approaches to obtain convergence results are local mode analysis for distributed control (e.g. [1]) and compactness arguments (e.g. [1] and [3]).

We analyze now the *Neumann boundary control problem* (1) – (3). The smoothing property can be shown in our framework:

Lemma 2 (Smoothing property). *The smoothing property holds for both alternatives of smoothers we discussed in this work with smoothing rate* $\eta(\nu) := \frac{C_S}{\sqrt{\nu}}$ *with some constant* $C_S > 0$ *independent of k, i.e.,*

$$|||\underline{x}_k^{(0,\nu)} - \underline{x}_k|||_{2,k} \leq C_S \frac{1}{\sqrt{\nu}} |||\underline{x}_k^{(0)} - \underline{x}_k|||_{0,k}.$$

The proof for the Uzawa type smoother follows the proof in [6] which is based on [5]. The analysis for the smoothers based on the normal equation uses the fact that the normal equation is symmetric and positive definite. Therefore the analysis can done by standard techniques.

We can show the *approximation property* similar to the proof given in [2] using the following regularity result: Let $f := (f_1, f_2) \in L^2(\Omega) \times L^2(\Omega)$ and $x \in X$ be such that

$$\mathsf{a}(x,\widetilde{x}) = (f_1,\widetilde{y})_{L^2(\Omega)} + (f_2,\widetilde{p})_{L^2(\Omega)}$$

holds for all $\widetilde{x} = (\widetilde{y},\widetilde{u},\widetilde{p}) \in X$. Since Ω is convex, it follows from standard H^2-regularity results for y and p that

$$\|x\|_{H^2(\Omega)\times H^1(\partial\Omega)\times H^2(\Omega)} \leq C \, \|f\|_{L^2(\Omega)\times L^2(\Omega)}$$

holds. Using this result we can show the following lemma.

Lemma 3 (Approximation property). *There is a constant $C_A > 0$ such that*

$$|||\underline{x}_k^{(1)} - \underline{x}_k|||_{0,k} \leq C_A |||\underline{x}_k^{(0,\nu)} - \underline{x}_k|||_{2,k}$$

holds.

The combination of smoothing property (lemma 2) and approximation property (lemma 3) leads to:

Theorem 1 (Convergence of the two-grid-method). *The two-grid-method converges for sufficiently large values of ν:*

$$|||\underline{x}_k^{(1)} - \underline{x}_k|||_{0,k} \leq q \, |||\underline{x}_k^{(0)} - \underline{x}_k|||_{0,k}$$

with convergence rate $q := \frac{C_A C_S}{\sqrt{\nu}} < 1$ independent of grid level k.

One can show that this implies the convergence of the W-cycle multigrid method, see e.g. [4].

5 Numerical results

The numerical tests were done for the unit square. The coarsest mesh (level $k = 0$) was constructed by separating the domain into two congruent triangles. The refinement was done by splitting each triangle into 4 congruent sub-triangles. Table 1 shows the number of iterations and the computing time that were necessary to decrease the norm of the error, $|||\underline{x}_k^{(i)} - \underline{x}_k|||_{0,k}$, by a factor of 10^{-5}.

The parameter γ was set to 1. The parameters ω and σ for the Uzawa-type smoother were chosen such that (9) holds (on a coarse level). For the normal equation method it turned out that it suffices to choose $\tau := 1/4$.

In agreement with the theory we observed that the number of iterations is independent of the grid level. Moreover, the number of iterations decreases as the number of smoothing steps is increased. The iterations that are necessary for both smoothers are in a comparable size of magnitude. The computing time increases linearly with the number of unknowns (optimal convergence). The computing time shows that the Jacobi method applied to the normal equation is slightly better, although one smoothing step of that method is slightly more expensive than one step of the Uzawa like method.

Table 1 Number of iterations and computing time computing time

Level	Number of unknowns	Uzawa-type				Jacobi-type			
		Smoothing steps							
		2		6		2		6	
5	2 306	29	3.2 s	9	2.2 s	21	2.6 s	7	2.1 s
6	8 706	31	11.7 s	10	9.1 s	21	8.9 s	8	7.9 s
7	33 794	31	44.2 s	10	39.4 s	21	33.7 s	8	32.0 s
8	133 122	31	187.1 s	10	160.1 s	21	147.2 s	8	128.9 s
9	528 386	31	751.7 s	11	646.0 s	21	557.0 s	8	517.1 s

6 Conclusion and Further Work

This work shows that the results for the Uzawa smoother and the strategy proposed in [6] for the distributed control problem carries over to the boundary control problem. It was possible to generate comparable results also for the Jacobi-type-smoother applied to the normal equations. The more general approach of the method based on the normal equation will hopefully allow extension of the method for a larger class of optimal control problems. Further work has to be done to find smoothers that are robust in the parameter γ.

References

1. Borzi, A., Kunisch, K., Kwak, D.Y. Accuracy and convergence properties of the finite difference multigrid solution of an optimal control optimality system. *SIAM J. Control Optim*, 41(5):1477 – 1497, 2003.
2. Brenner, S. C. Multigrid methods for parameter dependent problems. *RAIRO, Modélisation Math. Anal. Numér*, 30:265 – 297, 1996.
3. Hackbusch, W. Fast solution of elliptic control problems. *J. Optimization Theory Appl.*, 31:565 – 581, 1980.
4. Hackbusch, W. *Multi-Grid Methods and Applications*. Springer, Berlin, 1985.
5. Schöberl, J., and Zulehner, W. On Schwarz-type smoothers for saddle point problems. *Numer. Math.*, 95:3777 – 399, 2003.
6. Simon, R., and Zulehner, W. On Schwarz-type smoothers for saddle point problems with applications to PDE-constrained optimization problems. *Numer. Math.*, 111:445 – 468, 2009.

Technical Reports of the Doctoral Program "Computational Mathematics"

2009

2009-01 S. Takacs, W. Zulehner: *Multigrid Methods for Elliptic Optimal Control Problems with Neumann Boundary Control* October 2009. Eds.: U. Langer, J. Schicho

Doctoral Program
"Computational Mathematics"

Director:

Prof. Dr. Peter Paule
Research Institute for Symbolic Computation

Deputy Director:

Prof. Dr. Bert Jüttler
Institute of Applied Geometry

Address:

Johannes Kepler University Linz
Doctoral Program "Computational Mathematics"
Altenbergerstr. 69
A-4040 Linz
Austria
Tel.: ++43 732-2468-7174

E-Mail:

office@dk-compmath.jku.at

Homepage:

http://www.dk-compmath.jku.at

Submissions to the DK-Report Series are sent to two members of the Editorial Board who communicate their decision to the Managing Editor.