A Robust Multigrid Method for an Elliptic Optimal Control Problem

René Simon

University Duisburg - Essen

Workshop on PDE Constrained Optimization of Certain and Uncertain Processes

Trier, June 3-5, 2009

Acknowledgement

joint work with:

Joachim Schöberl (RWTH Aachen, Germany)

Walter Zulehner (JKU Linz, Austria)

Outline

An elliptic optimal control problem

Minimize the cost functional

$$J(y, u) = \frac{1}{2} \int_\Omega |y - y_d|^2 \, dx + \frac{\gamma}{2} \int_\Omega |u|^2 \, dx$$

subject to the state equation for y with distributed control u

$$\begin{aligned} L\,y &= u && \text{in } \Omega, \\ B\,y &= 0 && \text{on } \Gamma, \end{aligned}$$

where y_d is the desired state.

L is an elliptic differential operator, B a boundary operator.

Model problem:

$$L\,y = -\Delta y + y, \quad B\,y = \frac{\partial y}{\partial n}.$$

An elliptic optimal control problem

Variational form of the state equation:

For $u \in L^2(\Omega)$, find $y \in Y \subset H^1(\Omega)$ such that

$$a(y, z) = (u, z)_{L^2(\Omega)} \quad \text{for all } z \in Y$$

where a is symmetric, bounded and coercive on Y.

Model problem: $Y = H^1(\Omega)$ and

$$a(y, z) = \int_\Omega \left[\nabla y \cdot \nabla z + y\, z \right]\, dx$$

An elliptic optimal control problem

Lagrangian function: $p \in Q = Y$

$$L(y, u, p) = J(y, u) + a(y, p) - (u, p)_{L^2(\Omega)}$$
$$= \frac{1}{2} \|y - y_d\|_{L^2(\Omega)}^2 + \frac{\gamma}{2} \|u\|_{L^2(\Omega)}^2 + a(y, p) - (u, p)_{L^2(\Omega)}$$

Optimality system (KKT conditions):

$$\langle D_y L(y, u, p), z \rangle = 0 \quad \text{for all } z \in Y$$
$$\langle D_u L(y, u, p), v \rangle = 0 \quad \text{for all } v \in L^2(\Omega)$$
$$\langle D_p L(y, u, p), q \rangle = 0 \quad \text{for all } q \in Y$$

i.e.:

$$(y - y_d, z)_{L^2(\Omega)} + a(z, p) = 0 \quad \text{for all } z \in Y$$
$$\gamma (u, v)_{L^2(\Omega)} - (v, p)_{L^2(\Omega)} = 0 \quad \text{for all } v \in L^2(\Omega)$$
$$a(y, q) - (u, q)_{L^2(\Omega)} = 0 \quad \text{for all } q \in Y$$

An elliptic optimal control problem

Second equation:

$$u = \gamma^{-1} p$$

Reduced problem:

Find $y \in Y$ and $p \in Q$ such that

$$
\begin{aligned}
(y, z)_{L^2(\Omega)} + a(z, p) &= (y_d, z)_{L^2(\Omega)} && \text{for all } z \in Y, \\
a(y, q) \quad - \gamma^{-1} (p, q)_{L^2(\Omega)} &= 0 && \text{for all } q \in Q.
\end{aligned}
$$

Find $(y, p) \in Y \times Q$ such that

$$\mathcal{B}((y, p), (z, q)) = (y_d, z)_{L^2(\Omega)} \quad \text{for all } (z, q) \in Y \times Q$$

with

$$\mathcal{B}((y, p), (z, q)) = (y, z)_{L^2(\Omega)} + a(z, p) + a(y, q) - \gamma^{-1} (p, q)_{L^2(\Omega)}.$$

An elliptic optimal control problem

Inner products in $Y \subset H^1(\Omega)$, $Q \subset H^1(\Omega)$ and $Y \times Q$:

$$(y, z)_Y = (y, z)_{L^2(\Omega)} + \gamma^{1/2} (y, z)_{H^1(\Omega)}$$

$$(p, q)_Q = \gamma^{-1} (p, q)_{L^2(\Omega)} + \gamma^{-1/2} (p, q)_{H^1(\Omega)}$$

$$((y, p), (z, q))_{Y \times Q} = (y, z)_Y + (p, q)_Q$$

Norms in $Y \subset H^1(\Omega)$, $Q \subset H^1(\Omega)$ and $Y \times Q$:

$$\|z\|_Y = (z, z)_Y^{1/2} \sim \|z\|_{L^2(\Omega)} + \gamma^{1/4} \|z\|_{H^1(\Omega)}$$

$$\|q\|_Q = (q, q)_Q^{1/2} \sim \gamma^{-1/2} \|q\|_{L^2(\Omega)} + \gamma^{-1/4} \|q\|_{H^1(\Omega)}$$

$$\|(z, q)\|_{Y \times Q} = (\|z\|_Y^2 + \|q\|_Q^2)^{1/2}$$

An elliptic optimal control problem

Theorem

- $\mathcal{B}$ is bounded on $Y \times Q$:

$$|\mathcal{B}((y,p),(z,q))| \leq C \, \|(y,p)\|_{Y \times Q} \, \|(z,q)\|_{Y \times Q}$$

for all $(y,p), \, (z,q) \in Y \times Q$.

- $\mathcal{B}$ is stable on $X \times Q$:

$$\sup_{0 \neq (z,q) \in Y \times Q} \frac{\mathcal{B}((y,p),(z,q))}{\|(z,q)\|_{Y \times Q}} \geq c \, \|(y,p)\|_{Y \times Q}$$

for all $(y,p) \in Y \times Q$.

- The constants c and C are independent of γ.

An elliptic optimal control problem

Discretization by Galerkin's principle:

$$Y_h = Q_h \subset Y = Q \subset H^1(\Omega)$$

Find $(y_h, p_h) \in Y_h \times Q_h$ such that

$$\mathcal{B}((y_h, p_h), (z_h, q_h)) = (y_d, z_h) \quad \text{for all } (z_h, q_h) \in Y_h \times Q_h.$$

In detail:

$$(y_h, z_h)_{L^2(\Omega)} + a(z_h, p_h) = (y_d, z_h)_{L^2(\Omega)} \quad \text{for all } z_h \in Y_h,$$

$$a(y_h, q_h) - \gamma^{-1} (p_h, q_h)_{L^2(\Omega)} = 0 \quad \text{for all } q_h \in Q_h.$$

Model problem: Courant element (continuous and piecewise P_1 on a triangular subdivision $\mathcal{T}_h$ of Ω)

An elliptic optimal control problem

Basis for $Y_h = Q_h$: $\quad \{\varphi_i : i = 1 \ldots, N_h\}$

$$w_h(x) = \sum_{i=1}^{N_h} w_i \, \varphi_i(x) \quad \longleftrightarrow \quad \underline{w}_h = (w_i)_{i=1,\ldots,N_h}$$

Matrix-vector representations:

$$(y_h, z_h)_{L^2(\Omega)} = (M_h \underline{y}_h, \underline{z}_h)_{\ell^2}, \qquad a(z_h, q_h) = (K_h \underline{z}_h, \underline{q}_h)_{\ell^2},$$
$$(y_d, z_h)_{L^2(\Omega)} = (\underline{f}_h, \underline{z}_h)_{\ell^2}.$$

with the mass matrix M_h and the stiffness matrix K_h.

Linear system in saddle point form:

$$\mathcal{A}_h \begin{pmatrix} \underline{y}_h \\ \underline{p}_h \end{pmatrix} = \begin{pmatrix} \underline{f}_h \\ 0 \end{pmatrix} \quad \text{with} \quad \mathcal{A}_h = \begin{pmatrix} M_h & K_h^T \\ K_h & -\gamma^{-1} M_h \end{pmatrix}$$

An elliptic optimal control problem

Theorem

- $\mathcal{B}$ is bounded on $Y_h \times Q_h$:

$$|\mathcal{B}((y_h, p_h), (z_h, q_h))| \leq C \, \|(y_h, p_h)\|_{Y \times Q} \, \|(z_h, q_h)\|_{Y \times Q}$$

  for all $(y_h, p_h), \ (z_h, q_h) \in Y_h \times Q_h$.

- $\mathcal{B}$ is stable on $X_h \times Q_h$:

$$\sup_{0 \neq (z_h, q_h) \in Y_h \times Q_h} \frac{\mathcal{B}((y_h, p_h), (z_h, q_h))}{\|(z_h, q_h)\|_{Y \times Q}} \geq c \, \|(y_h, p_h)\|_{Y \times Q}$$

  for all $(y_h, p_h) \in Y_h \times Q_h$.

- The constants c and C are independent of γ and h.

Multigrid methods

Problem:

$$\mathcal{B}((y_h, p_h), (z_h, q_h)) = \mathcal{F}(z_h, q_h) \quad (z_h, q_h) \in Y_h \times Q_h.$$

Two-grid method: $\quad Y_H \subset Y_h, \quad Q_H \subset Q_h$

$(y_h^{(0)}, p_h^{(0)}) \quad$ starting value

- Smoothing step:

$$(y_h^{(0,k)}, p_h^{(0,k)}) = \mathcal{S}_h (y_h^{(0,k-1)}, p_h^{(0,k-1)}) \quad k = 1, \ldots, m.$$

- Coarse grid correction:

$$\mathcal{R}(z_h, q_h) = \mathcal{F}(z_h, q_h) - \mathcal{B}((y_h^{(0,m)}, p_h^{(0,m)}), (z_h, q_h))$$

$$\mathcal{B}((w_H, s_H), (z_H, q_H)) = \mathcal{R}(z_H, q_H) \quad (z_H, q_H) \in Y_H \times Q_H$$

$$(y_h^{(1)}, p_h^{(1)}) = (y_h^{(0,m)}, p_h^{(0,m)}) + (w_H, s_H)$$

Multigrid methods

Smoothing iteration:

$$(y_h^{(0,k)}, p_h^{(0,k)}) = \mathcal{S}_h(y_h^{(0,k-1)}, p_h^{(0,k-1)})$$

Residuals:

$$\begin{pmatrix} \underline{r}_h \\ \underline{t}_h \end{pmatrix} = \begin{pmatrix} \underline{f}_h \\ 0 \end{pmatrix} - \mathcal{A}_h \begin{pmatrix} \underline{y}_h \\ \underline{p}_h \end{pmatrix}$$

- Collective Gauß-Seidel method (CGS):

  for $i = 1, \ldots, N_h$:

  solve $\quad \begin{pmatrix} M_{ii} & K_{ii} \\ K_{ii} & -\gamma^{-1} M_{ii} \end{pmatrix} \begin{pmatrix} w_i \\ s_i \end{pmatrix} = \begin{pmatrix} r_i \\ t_i \end{pmatrix}$

  update $\quad \underline{y}_h, \underline{p}_h, \underline{r}_h, \underline{t}_h.$

  Vanka-type smoother

Multigrid methods

Borzì & Kunisch & Kwak: optimal control

- Numerical experiments: fast convergence

- Fourier analysis for the two-grid method: robust convergence results with respect to γ.

- Compactness argument: General multigrid convergence results, if the coarse mesh is sufficiently small.

If $K_h^T = K_h$, the original problem

$$\mathcal{A}_h \begin{pmatrix} \underline{y}_h \\ \underline{p}_h \end{pmatrix} = \begin{pmatrix} \underline{f}_h \\ 0 \end{pmatrix} \quad \text{with} \quad \mathcal{A}_h = \begin{pmatrix} M_h & K_h^T \\ K_h & -\gamma^{-1} M_h \end{pmatrix}$$

is equivalent to

$$\tilde{\mathcal{A}}_h \begin{pmatrix} \underline{y}_h \\ \underline{p}_h \end{pmatrix} = \begin{pmatrix} 0 \\ \underline{f}_h \end{pmatrix} \quad \text{with} \quad \tilde{\mathcal{A}}_h = \begin{pmatrix} \gamma\, K_h & -M_h \\ M_h & K_h \end{pmatrix}$$

Multigrid methods

- Symmetric diagonal Uzawa method:

  If $K_h^T = K_h$, the original problem

  $$\mathcal{A}_h \begin{pmatrix} \underline{y}_h \\ \underline{p}_h \end{pmatrix} = \begin{pmatrix} \underline{f}_h \\ 0 \end{pmatrix} \quad \text{with} \quad \mathcal{A}_h = \begin{pmatrix} M_h & K_h^T \\ K_h & -\gamma^{-1} M_h \end{pmatrix}$$

  is equivalent to

  $$\bar{\mathcal{A}}_h \begin{pmatrix} \underline{p}_h \\ \underline{y}_h \end{pmatrix} = \begin{pmatrix} \underline{f}_h \\ 0 \end{pmatrix} \quad \text{with} \quad \bar{\mathcal{A}}_h = \begin{pmatrix} K_h & M_h \\ M_h & -\gamma K_h \end{pmatrix} = \begin{pmatrix} A_h & B_h^T \\ B_h & -C_h \end{pmatrix}$$

  Preconditioner

  $$\mathcal{P}_h = \begin{pmatrix} \hat{A}_h & 0 \\ B_h & -\hat{S}_h \end{pmatrix} \begin{pmatrix} I & \hat{A}_h^{-1} B_h^T \\ 0 & I \end{pmatrix} = \begin{pmatrix} \hat{A}_h & B_h^T \\ B_h & B_h \hat{A}_h^{-1} B_h^T - \hat{S}_h \end{pmatrix}$$

  with

  $$\hat{A}_h = \frac{1}{\sigma} \operatorname{diag}(A_h), \quad \hat{S}_h = \frac{1}{\tau} \operatorname{diag}(C_h + B_h \hat{A}_h^{-1} B_h)$$

Multigrid methods

Smoothing iteration:

$$\begin{pmatrix} \underline{p}_h^{(k)} \\ \underline{y}_h^{(k)} \end{pmatrix} = \begin{pmatrix} \underline{p}_h^{(k-1)} \\ \underline{y}_h^{(k-1)} \end{pmatrix} + \rho \, \mathcal{P}_h^{-1} \begin{pmatrix} \underline{r}_h^{(k-1)} \\ \underline{t}_h^{(k-1)} \end{pmatrix}$$

with $0 < \rho < 2$

to solve

$$\mathcal{P}_h \begin{pmatrix} \underline{s}_h \\ \underline{w}_h \end{pmatrix} = \begin{pmatrix} \underline{r}_h \\ \underline{t}_h \end{pmatrix}$$

reduces to

$$\hat{A}_h \, \hat{\underline{s}}_h = \underline{r}_h$$
$$\hat{S}_h \, \underline{w}_h = B_h \, \hat{\underline{s}}_h - \underline{t}_h$$
$$\hat{A}_h \, \underline{s}_h = \underline{r}_h - B_h^T \, \underline{w}_h$$

Convergence analysis

For $z \in Y_h$, $q \in Q_h$:

$$\||(z,q)\||_{0,h} \quad \text{and} \quad \||(w,s)\||_{2,h} = \sup_{0 \neq (z,q) \in Y_h \times Q_h} \frac{|\mathcal{B}((w,s),(z,q))|}{\||(z,q)\||_{0,h}}.$$

Approximation property:

$$\||(y_h^{(1)} - y_h, p_h^{(1)} - p_h)\||_{0,h} \leq c_A \||(y_h^{(0,m)} - y_h, p_h^{(0,m)} - p_h)\||_{2,h}$$

Smoothing property:

$$\||(y_h^{(0,m)} - y_h, p_h^{(0,m)} - p_h)\||_{2,h} \leq \eta(m) \||(y_h^{(0)} - y_h, p_h^{(0)} - p_h)\||_{0,h}$$

with

$$\eta(m) \to 0 \quad \text{for } m \to \infty.$$

Convergence analysis

Inner product on $Y \times Q$:

$$
\begin{aligned}
\left((y, p), (z, q)\right)_{Y \times Q} = {} & (y, z)_{L^2(\Omega)} + \gamma^{1/2} \, (y, z)_{H^1(\Omega)} \\
& + \gamma^{-1} \, (p, q)_{L^2(\Omega)} + \gamma^{-1/2} \, (p, q)_{H^1(\Omega)}
\end{aligned}
$$

Mesh-dependent inner product on $Y_h \times Q_h$:

$$
\begin{aligned}
\left((y, p), (z, q)\right)_{0,h} = {} & (y, z)_{L^2(\Omega)} + \gamma^{1/2} \, h^{-2} \, (y, z)_{L^2(\Omega)} \\
& + \gamma^{-1} \, (p, q)_{L^2(\Omega)} + \gamma^{-1/2} \, h^{-2} \, (p, q)_{L^2(\Omega)} \\
= {} & \left(1 + \gamma^{1/2} \, h^{-2}\right) \left[(y, z)_{L^2(\Omega)} + \gamma^{-1} \, (p, q)_{L^2(\Omega)}\right]
\end{aligned}
$$

Mesh-dependent norm on $Y_h \times Q_h$:

$$
\begin{aligned}
\||(z, q)\||_{0,h} &= \left((z, q), (z, q)\right)_{0,h}^{1/2} \\
&\sim (1 + \gamma^{1/4} \, h^{-1}) \left[\|z\|_{L^2(\Omega)} + \gamma^{-1/2} \, \|q\|_{L^2(\Omega)}\right]
\end{aligned}
$$

Convergence analysis

Assumption: Full elliptic regularity in y of the state equation

If $f \in L^2(\Omega)$, then there exists $y \in H^2(\Omega)$ such that

$$a(y, z) = (f, z)_{L^2(\Omega)} \quad \text{for all } z \in Y$$

and

$$\|y\|_{H^2(\Omega)} \le c \, \|f\|_{L^2(\Omega)}.$$

Model problem: Ω is a convex polygonal domain,
Ω has a smooth boundary.

Theorem (Approximation property)

$$\||(y_h^{(1)} - y_h, p_h^{(1)} - p_h)\||_{0,h} \le c_A \, \||(y_h^{(0,m)} - y_h, p_h^{(0,m)} - p_h)\||_{2,h}$$

Convergence analysis

Lemma

If the parameters σ, τ are $O(1)$ and

$$\hat{A}_h \geq A_h, \quad \hat{S}_h \geq C_h + B_h \hat{A}_h^{-1} B_h^T,$$

then

$$\||(y_h^{(0,m)} - y_h, p_h^{(0,m)} - p_h)\||_{2,h} \leq \eta(m, \gamma, h) \, \||(y_h^{(0)} - y_h, p_h^{(0)} - p_h)\||_{0,h}$$

with

$$\eta(m, \gamma, h) = c_S \, \eta_0(m) \frac{\gamma^{1/2} h^{-2} + \gamma^{-1/2} h^2}{1 + \gamma^{1/2} h^{-2}}$$

where

$$\eta_0(m) = \frac{1}{2^{m-1}} \binom{m-1}{[m/2]} = O(\frac{1}{\sqrt{m}}).$$

Convergence analysis

Theorem (Smoothing property)

If $\gamma \geq c\,h^4$, then

$$\||(y_h^{(0,m)} - y_h, p_h^{(0,m)} - p_h)\||_{2,h} \leq \eta(m)\,\||(y_h^{(0)} - y_h, p_h^{(0)} - p_h)\||_{0,h}$$

with $\eta(m) = O(1/\sqrt{m})$.

Theorem (Two-grid convergence)

If $\gamma \geq c\,h^4$, then there is a constant C independent of γ and h such that

$$\||(y_h^{(1)} - y_h, p_h^{(1)} - p_h)\||_{0,h} \leq \frac{C}{\sqrt{m}}\,\||(y_h^{(0)} - y_h, p_h^{(0)} - p_h)\||_{0,h}$$

Numerical experiments

Geometry and mesh:

- $\Omega = (0, 1)^2$
- initial mesh $\ell = 0$: 2 triangles
- uniform refinement, final mesh $\ell = L$.

Multigrid method:

- W-cycle, V-cycle
- Starting values $\underline{y}_h^{(0)}$ and $\underline{p}_h^{(0)}$ randomly generated.
- stopping rule $r^{(k)} \leq \varepsilon \, r^{(0)}$ with $\varepsilon = 10^{-8}$.
- smoother: symmetric diagonal Uzawa method
  - σ, τ: Simple power iteration on coarse level,
  - $\rho = 1.6$ (motivated by a two-grid Fourier analysis)

Numerical experiments

Dependence on *h* and *m*:

level *L*	*n + m*	smoothing steps					
		1+1		2+2		3+3	
4	2 178	16	0.301	9	0.127	7	0.067
5	8 450	16	0.302	9	0.128	7	0.066
6	33 282	16	0.302	10	0.135	7	0.067
7	132 098	16	0.302	10	0.135	7	0.067
8	526 338	16	0.302	10	0.135	7	0.068

Dependence on γ:

γ	Iter.	conv. rate
1	16	0.302
10^{-2}	16	0.302
10^{-4}	16	0.302
10^{-6}	16	0.302

Concluding remarks

- one-shot multigrid method

- simple tuning of the parameters

- convergence analysis: robust in γ and h

- to be done: other classes of problems