Methods for Inverse Problems: VIII. Adaptivity

Barbara Kaltenbacher, University of Klagenfurt, Austria

overview

1. Motivation: Parameter Identification in PDEs
2. refinement/coarsening based on predicted misfit reduction
3. goal oriented error estimators

Motivation: Parameter Identification in PDEs

- instability: sufficiently high precision (amplification of numerical errors)
- computational effort:
  - large scale problem: each regularized inversion involves several PDE solves
  - repeated solution of regularized problem to determine regularization parameter

Motivation: Parameter Identification in PDEs

- instability: sufficiently high precision (amplification of numerical errors)
- computational effort:
  - large scale problem: each regularized inversion involves several PDE solves
  - repeated solution of regularized problem to determine regularization parameter

Example $-\Delta u = q$:

refine grid for u and q:
- at jumps or large gradients or
- at locations with large error contribution

Motivation: Parameter Identification in PDEs

- instability: sufficiently high precision (amplification of numerical errors)
- computational effort:
  - large scale problem: each regularized inversion involves several PDE solves
  - repeated solution of regularized problem to determine regularization parameter

Example $-\Delta u = q$:

refine grid for u and q:
- at jumps or large gradients or
- at locations with large error contribution

$\rightarrow$ location of large gradients / large errors *a priori unknown*

Motivation: Parameter Identification in PDEs

- instability: sufficiently high precision (amplification of numerical errors)
- computational effort:
  - large scale problem: each regularized inversion involves several PDE solves
  - repeated solution of regularized problem to determine regularization parameter

Example $-\Delta u = q$:

refine grid for u and q:
- at jumps or large gradients or
- at locations with large error contribution

$\rightarrow$ location of large gradients / large errors *a priori unknown*

$\rightarrow$ general strategy for mesh generation possibly *separately for q and u* (example $-\nabla q(u) \nabla u) = f$)

Motivation: Parameter Identification in PDEs

- instability: sufficiently high precision (amplification of numerical errors)
- computational effort:
  - large scale problem: each regularized inversion involves several PDE solves
  - repeated solution of regularized problem to determine regularization parameter

Example $-\Delta u = q$:
refine grid for u and q: • at jumps or large gradients or
• at locations with large error contribution

$\rightarrow$ location of large gradients / large errors *a priori unknown*

$\rightarrow$ general strategy for mesh generation possibly
separately for q and u (example $-\nabla q(u)\nabla u) = f$)

instability $\Rightarrow$ regularization necessary !

Regularization

- unstable operator equation: $F(q) = g$ with $F : q \mapsto u$ or $C(u)$
- solution $q = F^{-1}(g)$ does not depend continuously on g
  i.e., $(\forall (g_n), g_n \to g \;\not\Rightarrow\; q_n := F^{-1}(g_n) \to F^{-1}(g))$
- only noisy data $g^\delta \approx g$ available: $\|g^\delta - g\| \leq \delta$
- making $\|F(q) - g^\delta\|$ small $\not\Rightarrow$ good result for q!
- regularization means approaching solution along stable path:
  given $(g_n), g_n \to g$ construct $q_n := R_{\alpha_n}(g_n)$ such that
  $q_n = R_{\alpha_n}(g_n) \to F^{-1}(g)$
- *regularization method*: family $(R_\alpha)_{\alpha>0}$ with parameter choice $\alpha = \alpha(g^\delta, \delta)$
  such that worst case convergence as $\delta \to 0$:

$$\sup_{\|g^\delta - g\| \leq \delta} \|R_{\alpha(g^\delta,\delta)}(g^\delta) - F^{-1}(g)\| \to 0 \text{ as } \delta \to 0$$

Motivation: Parameter Identification in PDEs

- instability: sufficiently high precision (amplification of numerical errors)
- computational effort:
  - large scale problem: each regularized inversion involves several PDE solves
  - repeated solution of regularized problem to determine regularization parameter

Example $-\Delta u = q$:

refine grid for u and q:
- at jumps or large gradients or
- at locations with large error contribution

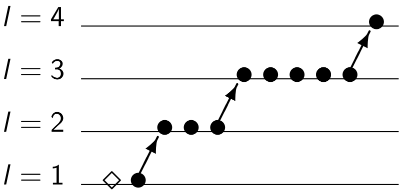
$\rightarrow$ location of large gradients / large errors *a priori unknown*

$\rightarrow$ general strategy for mesh generation possibly *separately for q and u* (example $-\nabla q(u)\nabla u) = f$)

computational effort $\Rightarrow$ efficient numerical strategies necessary !

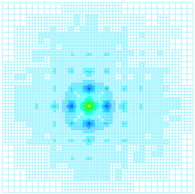
Efficient Methods for PDEs

multilevel iteration:



start with coarse discretization
refine successively

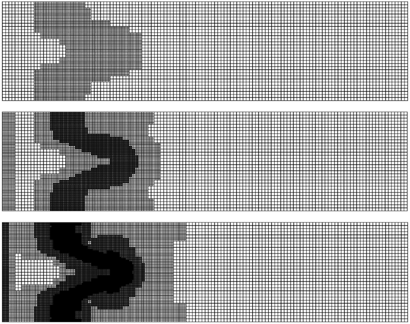
adaptive discretization:



coarse discretization where possible
fine grid only where necessary

Efficient Methods for PDEs

combined multilevel adaptive strategy:



courtesy to [R.Becker&M.Braack&B.Vexler, App.Num.Math., 2005]

start on coarse grid

sucessive adaptive refinement

Some Ideas on Adaptivity for Inverse Problems

- Haber&Heldmann&Ascher'07: Tikhonov with BV type regularization:
  Refine for u to compute residual term sufficiently precisely;
  Refine for q to compute regularization term sufficiently precisely
- Neubauer'03, '06, '07: moving mesh regularization, adaptive grid regularization: Tikhonov with BV type regularization:
  Refine where q has jumps or large gradients
- Borcea&Druskin'02: optimal finite difference grids (a priori refinement): *Refine close to measurements*
- Chavent&Bissell'98, Ben Ameur&Chavent&Jaffré'02, BK&Ben Ameur'02: refinement and coarsening indicators
- Becker&Vexler'04, Griesbaum&BK&Vexler'07, Bangerth'08, BK&Vexler'09: goal oriented error estimators
- ...

1st approach:

refinement/coarsening based on predicted misfit reduction

Identification of a Distributed Parameter:

Groundwater modelling

$$s\frac{\partial u}{\partial t} - div\,(q\ grad\ u) = f\ \ in\ \Omega \subseteq \mathbb{R}^2$$

with initial and boundary conditions

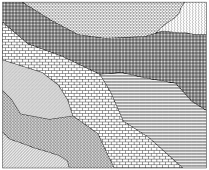
u … hydraulic potential (ground water level),
$s(x, y)$ … storage coefficients,
$q(x, y)$ … hydraulic transmissivity,
$f(x, y, t)$ … source term,

space and time discretization (time step Δt, mesh size h).

Parameter Identification

$$s\frac{\partial u}{\partial t} - \text{div}\,(q\ \text{grad}\ u) = f \text{ in } \Omega$$

Reconstruction of the transmissivity q (pcw. const.) from measurements of u.

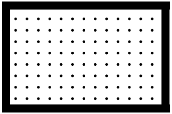


Find zonation and values of q such that

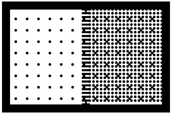
$$J(q) := \|u(q) - u^{obs}\|^2 = \text{min!}$$

[Ben Ameur&Chavent&Jaffré'02], [Chavent&Bissell'98], [BK&Ben Ameur'02]

Refinement Indicators



1 zone $\rightarrow$ 2 zones



$q^* :=$ min of $J(q)$ solves

$$\begin{cases} \min J\left(\begin{pmatrix} q_1 \\ q_2 \end{pmatrix}\right) & \text{s.t.} \\ d^T \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} = q_1 - q_2 = B & =: 0 \end{cases}$$

$\begin{pmatrix} q_1^* \\ q_2^* \end{pmatrix} :=$ min of $J\left(\begin{pmatrix} q_1 \\ q_2 \end{pmatrix}\right)$ solves

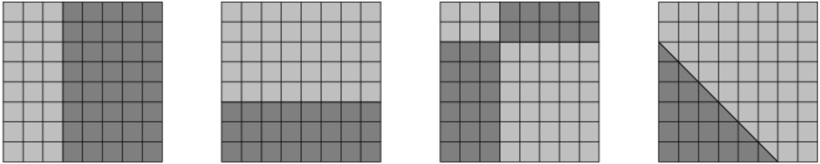
$$\begin{cases} \min J\left(\begin{pmatrix} q_1 \\ q_2 \end{pmatrix}\right) & \text{s.t.} \\ d^T \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} = q_1 - q_2 = B & =: q_1^* - q_2^* \end{cases}$$

$$\frac{\partial}{\partial B} J\left(\begin{pmatrix} q_1^B \\ q_2^B \end{pmatrix}\right) = \lambda^B \;\Rightarrow\; J\left(\begin{pmatrix} q_1^* \\ q_2^* \end{pmatrix}\right) \approx J(q^*) + \lambda^0 (q_1^* - q_2^*)$$

$|\lambda^0|$ large $\Rightarrow$ large possible reduction of data misfit J_{opt}^B

$\lambda^0 = (1/d^T d) d^T \nabla J(q^*)$ (negligible computational effort)

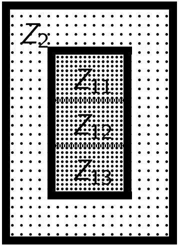
Compute all refinement indicators for zonations generated systematically by families of vertical, horizontal, checkerboard and oblique cuts.



(a) (b) (c) (d)

Mark those cuts that yield largest refinement indicators $|\lambda^0|$

Coarsening Indicators



$$\begin{pmatrix} q_1^* \\ q_2^* \end{pmatrix} := \text{solution of } \min J\left(\begin{pmatrix} q_1 \\ q_2 \end{pmatrix}\right)$$

$$\text{solves } \begin{cases} \min J\left(\begin{pmatrix} q_{11} \\ q_{12} \\ q_{13} \\ q_2 \end{pmatrix}\right) \text{ s.t.} \\ q_{11} - q_2 = B_1 \\ q_{12} - q_2 = B_2 \\ q_{13} - q_2 = B_3 \end{cases}$$

with $B_i := B^* := q_1^* - q_2^*$

$$\underbrace{J\left(\begin{pmatrix} q_{11}^B \\ q_{12}^B \\ q_{13}^B \\ q_2^B \end{pmatrix}\right)\Big|_{B_i=0, B_j=B^*, j\neq i}}_{\text{optimum if } q_{1i} \text{ is aggregated with } q_2} \approx J\left(\begin{pmatrix} q_1^* \\ q_2^* \end{pmatrix}\right) - \underbrace{\lambda_i^{B^*} B^*}_{\text{coarsening indicator}}$$

Multilevel Refinement and Coarsening Algorithm

[H.Ben Ameur, G.Chavent, J.Jaffré, 2002]

Minimize J on starting zonation
Do until refinement indicators $= 0$
Refinement: compute refinement indicators λ
chooose cuts with largest $|\lambda|$
Coarsening: if chosen cuts yield several sub-zones:
evaluate coarsening indicators
and aggregate zones where possible
Minimize J for each of the retained zonations
and keep those with largest reduction in J

Abstract Setting for Refinement and Coarsening

discretization: $X_N = \text{span}\{\phi_1, \ldots, \phi_N\}$ s.t. $X = \bigcup_{N \in \mathbb{N}} X_N$

misfit minimization

$$\min_{q \in X_N} \|F(q) - g^\delta\|^2 = \min_{\mathbf{a} \in \mathbb{R}^N} \underbrace{\|F(\sum_{i=1}^N a_i \phi_i) - g^\delta\|^2}_{=\mathcal{J}(\mathbf{a})}$$

Abstract Setting for Refinement and Coarsening

discretization: $X_N = \mathrm{span}\{\phi_1, \dots, \phi_N\}$ s.t. $X = \bigcup_{N \in \mathbb{N}} X_N$

misfit minimization

$$\min_{q \in X_N} \|F(q) - g^\delta\|^2 = \min_{\mathbf{a} \in \mathbb{R}^N} \underbrace{\|F(\textstyle\sum_{i=1}^N a_i \phi_i) - g^\delta\|^2}_{=\mathcal{J}(\mathbf{a})}$$

consider misfit minimization on some index set $\mathcal{I} \subseteq \{1, 2, \dots, N\}$:

$$\min_{\mathbf{a} \in \mathbb{R}^{|\mathcal{I}|}} \|F(\textstyle\sum_{i \in \mathcal{I}} a_i \phi_i) - g^\delta\|^2 \quad (P^{\mathcal{I}})$$

$\rightsquigarrow$ solution $\mathbf{a}^{\mathcal{I}}$, $q^{\mathcal{I}}$ with $a_i := 0$ for $i \notin \mathcal{I}$

Abstract Setting for Refinement and Coarsening

discretization: $X_N = \text{span}\{\phi_1, \ldots, \phi_N\}$ s.t. $X = \bigcup_{N \in \mathbb{N}} X_N$

misfit minimization

$$\min_{q \in X_N} \|F(q) - g^\delta\|^2 = \min_{\mathbf{a} \in \mathbb{R}^N} \underbrace{\|F(\textstyle\sum_{i=1}^N a_i \phi_i) - g^\delta\|^2}_{=\mathcal{J}(\mathbf{a})}$$

consider misfit minimization on some index set $\mathcal{I} \subseteq \{1, 2, \ldots, N\}$:

$$\min_{\mathbf{a} \in \mathbb{R}^{|\mathcal{I}|}} \|F(\textstyle\sum_{i \in \mathcal{I}} a_i \phi_i) - g^\delta\|^2 \quad (P^{\mathcal{I}})$$

$\rightsquigarrow$ solution $\mathbf{a}^{\mathcal{I}}$, $q^{\mathcal{I}}$ with $a_i := 0$ for $i \notin \mathcal{I}$ $\rightsquigarrow$ *sparsity*

Abstract Setting for Refinement and Coarsening

discretization: $X_N = \text{span}\{\phi_1, \ldots, \phi_N\}$ s.t. $X = \bigcup_{N \in \mathbb{N}} X_N$

misfit minimization

$$\min_{q \in X_N} \|F(q) - g^\delta\|^2 = \min_{\mathbf{a} \in \mathbb{R}^N} \underbrace{\|F(\textstyle\sum_{i=1}^N a_i \phi_i) - g^\delta\|^2}_{=\mathcal{J}(\mathbf{a})}$$

consider misfit minimization on some index set $\mathcal{I} \subseteq \{1, 2, \ldots, N\}$:

$$\min_{\mathbf{a} \in \mathbb{R}^{|\mathcal{I}|}} \|F(\textstyle\sum_{i \in \mathcal{I}} a_i \phi_i) - g^\delta\|^2 \; (P^{\mathcal{I}})$$

$\rightsquigarrow$ solution $\mathbf{a}^{\mathcal{I}}$, $q^{\mathcal{I}}$ with $a_i := 0$ for $i \notin \mathcal{I}$ $\rightsquigarrow$ *sparsity*

Find index set $\mathcal{I}^\dagger$ and coefficients $a^{\mathcal{I}^\dagger}$ such that
$\|F(\sum_{i \in \mathcal{I}^\dagger} a_i^{\mathcal{I}^\dagger} \phi_i) - g^\delta\|^2 = \min_{\mathbf{a} \in \mathbb{R}^{|\mathcal{I}^\dagger|}} \|F(\sum_{i \in \mathcal{I}^\dagger} a_i^{\mathcal{I}^\dagger} \phi_i) - g^\delta\|^2 =$
$\min_{q \in X_N} \|F(q) - g^\delta\|^2$

Refinement Indicators

current index set $\mathcal{I}^k$ with computed solution $a^{\mathcal{I}^k}$ of $(P^{\mathcal{I}^k})$;

Refinement Indicators

current index set $\mathcal{I}^k$ with computed solution $a^{\mathcal{I}^k}$ of $(P^{\mathcal{I}^k})$;

for some index $\{i_*\} \notin \mathcal{I}^k$ consider constrained minimization prob.

$$\min_{\mathbf{a}\in\mathbb{R}^{|\mathcal{I}^k|+1}} \underbrace{\|F(\textstyle\sum_{i\in\mathcal{I}^k\cup\{i_*\}} a_i\phi_i) - g^\delta\|^2}_{=\mathcal{J}(\mathbf{a})} \quad \text{s.t. } a_{i_*} = \beta \qquad (P_\beta^{\mathcal{I}^k,\ i_*})$$

$\rightsquigarrow$ solution $\mathbf{a}_\beta$ with $a_i := 0$ for $i \notin \mathcal{I}^k \cup \{i_*\}$;

Refinement Indicators

current index set $\mathcal{I}^k$ with computed solution $a^{\mathcal{I}^k}$ of $(P^{\mathcal{I}^k})$;

for some index $\{i_*\} \notin \mathcal{I}^k$ consider constrained minimization prob.

$$\min_{\mathbf{a}\in\mathbb{R}^{|\mathcal{I}^k|+1}} \underbrace{\|F(\textstyle\sum_{i\in\mathcal{I}^k\cup\{i_*\}} a_i\phi_i) - g^\delta\|^2}_{=\mathcal{J}(\mathbf{a})} \quad \text{s.t. } a_{i_*} = \beta \qquad (P_\beta^{\mathcal{I}^k,\ i_*})$$

$\rightsquigarrow$ solution $\mathbf{a}_\beta$ with $a_i := 0$ for $i \notin \mathcal{I}^k \cup \{i_*\}$; note: $\mathbf{a}_{\beta=0} = \mathbf{a}^{\mathcal{I}^k}$ solves $(P^{\mathcal{I}^k})$

Refinement Indicators

current index set $\mathcal{I}^k$ with computed solution $a^{\mathcal{I}^k}$ of $(P^{\mathcal{I}^k})$;

for some index $\{i_*\} \notin \mathcal{I}^k$ consider constrained minimization prob.

$$\min_{\mathbf{a}\in\mathbb{R}^{|\mathcal{I}^k|+1}} \underbrace{\|F(\sum_{i\in\mathcal{I}^k\cup\{i_*\}} a_i\phi_i) - g^\delta\|^2}_{=\mathcal{J}(\mathbf{a})} \quad \text{s.t. } a_{i_*} = \beta \qquad (P_\beta^{\mathcal{I}^k,\ i_*})$$

$\rightsquigarrow$ solution $\mathbf{a}_\beta$ with $a_i := 0$ for $i \notin \mathcal{I}^k \cup \{i_*\}$; note: $\mathbf{a}_{\beta=0} = \mathbf{a}^{\mathcal{I}^k}$ solves $(P^{\mathcal{I}^k})$

Lagrange function $\quad \mathcal{L}(\mathbf{a},\lambda) = \mathcal{J}(\mathbf{a}) + \lambda(\beta - a_{i_*})$

Refinement Indicators

current index set $\mathcal{I}^k$ with computed solution $a^{\mathcal{I}^k}$ of $(P^{\mathcal{I}^k})$;

for some index $\{i_*\} \notin \mathcal{I}^k$ consider constrained minimization prob.

$$\min_{\mathbf{a}\in\mathbb{R}^{|\mathcal{I}^k|+1}} \underbrace{\|F(\textstyle\sum_{i\in\mathcal{I}^k\cup\{i_*\}} a_i\phi_i) - g^\delta\|^2}_{=\mathcal{J}(\mathbf{a})} \quad \text{s.t. } a_{i_*} = \beta \qquad (P_\beta^{\mathcal{I}^k,\ i_*})$$

$\rightsquigarrow$ solution $\mathbf{a}_\beta$ with $a_i := 0$ for $i \notin \mathcal{I}^k \cup \{i_*\}$; note: $\mathbf{a}_{\beta=0} = \mathbf{a}^{\mathcal{I}^k}$ solves $(P^{\mathcal{I}^k})$

Lagrange function $\quad \mathcal{L}(\mathbf{a},\lambda) = \mathcal{J}(\mathbf{a}) + \lambda(\beta - a_{i_*})$

necessary optimality conditions: $0 = \frac{\partial\mathcal{L}}{\partial a_{i_*}}(\mathbf{a}_\beta,\lambda_\beta) = \frac{\partial\mathcal{J}}{\partial a_{i_*}}(\mathbf{a}_\beta) - \lambda_\beta \quad (*)$

Refinement Indicators

current index set $\mathcal{I}^k$ with computed solution $a^{\mathcal{I}^k}$ of $(P^{\mathcal{I}^k})$;

for some index $\{i_*\} \notin \mathcal{I}^k$ consider constrained minimization prob.

$$\min_{\mathbf{a}\in\mathbb{R}^{|\mathcal{I}^k|+1}} \underbrace{\|F(\textstyle\sum_{i\in\mathcal{I}^k\cup\{i_*\}} a_i\phi_i) - g^\delta\|^2}_{=\mathcal{J}(\mathbf{a})} \quad \text{s.t. } a_{i_*} = \beta \qquad (P_\beta^{\mathcal{I}^k,\ i_*})$$

$\rightsquigarrow$ solution $\mathbf{a}_\beta$ with $a_i := 0$ for $i \notin \mathcal{I}^k \cup \{i_*\}$; note: $\mathbf{a}_{\beta=0} = \mathbf{a}^{\mathcal{I}^k}$ solves $(P^{\mathcal{I}^k})$

Lagrange function $\quad \mathcal{L}(\mathbf{a},\lambda) = \mathcal{J}(\mathbf{a}) + \lambda(\beta - a_{i_*})$

necessary optimality conditions: $0 = \frac{\partial\mathcal{L}}{\partial a_{i_*}}(\mathbf{a}_\beta,\lambda_\beta) = \frac{\partial\mathcal{J}}{\partial a_{i_*}}(\mathbf{a}_\beta) - \lambda_\beta \quad (*)$

Lagrange multipliers = sensitivities: $\frac{d}{d\beta}\mathcal{J}(\mathbf{a}_\beta) = \frac{d}{d\beta}\mathcal{L}(\mathbf{a}_\beta,\lambda_\beta) = \lambda_\beta$

Refinement Indicators

current index set $\mathcal{I}^k$ with computed solution $a^{\mathcal{I}^k}$ of $(P^{\mathcal{I}^k})$;

for some index $\{i_*\} \notin \mathcal{I}^k$ consider constrained minimization prob.

$$\min_{\mathbf{a}\in\mathbb{R}^{|\mathcal{I}^k|+1}} \underbrace{\|F(\textstyle\sum_{i\in\mathcal{I}^k\cup\{i_*\}} a_i\phi_i) - g^\delta\|^2}_{=\mathcal{J}(\mathbf{a})} \quad \text{s.t. } a_{i_*} = \beta \qquad (P_\beta^{\mathcal{I}^k,\ i_*})$$

$\rightsquigarrow$ solution $\mathbf{a}_\beta$ with $a_i := 0$ for $i \notin \mathcal{I}^k \cup \{i_*\}$; note: $\mathbf{a}_{\beta=0} = \mathbf{a}^{\mathcal{I}^k}$ solves $(P^{\mathcal{I}^k})$

Lagrange function $\quad \mathcal{L}(\mathbf{a}, \lambda) = \mathcal{J}(\mathbf{a}) + \lambda(\beta - a_{i_*})$

necessary optimality conditions: $0 = \frac{\partial \mathcal{L}}{\partial a_{i_*}}(\mathbf{a}_\beta, \lambda_\beta) = \frac{\partial \mathcal{J}}{\partial a_{i_*}}(\mathbf{a}_\beta) - \lambda_\beta \quad (*)$

Lagrange multipliers = sensitivities: $\frac{d}{d\beta}\mathcal{J}(\mathbf{a}_\beta) = \frac{d}{d\beta}\mathcal{L}(\mathbf{a}_\beta, \lambda_\beta) = \lambda_\beta$

Taylor expansion $\mathcal{J}(\mathbf{a}_\beta) \approx \mathcal{J}(\mathbf{a}_0) + \frac{d}{d\beta}\mathcal{J}(\mathbf{a}_0)\,\beta = \mathcal{J}(\mathbf{a}^{\mathcal{I}^k}) + \lambda_{\beta=0}\,\beta$

Refinement Indicators

current index set $\mathcal{I}^k$ with computed solution $a^{\mathcal{I}^k}$ of $(P^{\mathcal{I}^k})$;

for some index $\{i_*\} \notin \mathcal{I}^k$ consider constrained minimization prob.

$$\min_{\mathbf{a}\in\mathbb{R}^{|\mathcal{I}^k|+1}} \underbrace{\|F(\textstyle\sum_{i\in\mathcal{I}^k\cup\{i_*\}} a_i\phi_i) - g^\delta\|^2}_{=\mathcal{J}(\mathbf{a})} \quad \text{s.t. } a_{i_*} = \beta \qquad (P_\beta^{\mathcal{I}^k,\ i_*})$$

$\rightsquigarrow$ solution $\mathbf{a}_\beta$ with $a_i := 0$ for $i \notin \mathcal{I}^k \cup \{i_*\}$; note: $\mathbf{a}_{\beta=0} = \mathbf{a}^{\mathcal{I}^k}$ solves $(P^{\mathcal{I}^k})$

Lagrange function $\quad \mathcal{L}(\mathbf{a},\lambda) = \mathcal{J}(\mathbf{a}) + \lambda(\beta - a_{i_*})$

necessary optimality conditions: $0 = \frac{\partial \mathcal{L}}{\partial a_{i_*}}(\mathbf{a}_\beta,\lambda_\beta) = \frac{\partial \mathcal{J}}{\partial a_{i_*}}(\mathbf{a}_\beta) - \lambda_\beta \quad (*)$

Lagrange multipliers = sensitivities: $\frac{d}{d\beta}\mathcal{J}(\mathbf{a}_\beta) = \frac{d}{d\beta}\mathcal{L}(\mathbf{a}_\beta,\lambda_\beta) = \lambda_\beta$

Taylor expansion $\mathcal{J}(\mathbf{a}_\beta) \approx \mathcal{J}(\mathbf{a}_0) + \frac{d}{d\beta}\mathcal{J}(\mathbf{a}_0)\,\beta = \mathcal{J}(\mathbf{a}^{\mathcal{I}^k}) + \lambda_{\beta=0}\,\beta$

$\Rightarrow r^{i_*} := |\lambda_{\beta=0}| \stackrel{(*)}{=} |\frac{\partial \mathcal{J}}{\partial a_{i_*}}(\mathbf{a}^{\mathcal{I}^k})|$ ... refinement indicator

Coarsening Indicators

current index set $\tilde{\mathcal{I}}^k$ with computed solution $a^{\tilde{\mathcal{I}}^k}$ of $(P^{\tilde{\mathcal{I}}^k})$;

for some index $\{l_*\} \in \tilde{\mathcal{I}}^k$ consider constrained minimization probl.

$$\min_{\mathbf{a}\in\mathbb{R}^{|\tilde{\mathcal{I}}^k|}} \underbrace{\|F(\textstyle\sum_{i\in\tilde{\mathcal{I}}^k} a_i\phi_i) - g^\delta\|^2}_{=\mathcal{J}(\mathbf{a})} \quad \text{s.t. } a_{l_*} = \gamma \qquad (\tilde{P}_\gamma^{\tilde{\mathcal{I}}^k,\, l_*})$$

$\rightsquigarrow$ solution $\mathbf{a}_\gamma$ with $a_i := 0$ for $i \notin \tilde{\mathcal{I}}^k$; note: $\mathbf{a}_{\gamma_*} = \mathbf{a}^{\tilde{\mathcal{I}}^k}$ with $\gamma_* := a_{l_*}^{\tilde{\mathcal{I}}^k}$ solves $(P^{\tilde{\mathcal{I}}^k})$

Lagrange function $\quad \mathcal{L}(\mathbf{a},\mu) = \mathcal{J}(\mathbf{a}) + \mu(\gamma - a_{l_*})$

necessary optimality conditions: $0 = \frac{\partial\mathcal{L}}{\partial a_{l_*}}(\mathbf{a}_\gamma,\mu_\gamma) = \frac{\partial\mathcal{J}}{\partial a_{l_*}}(\mathbf{a}_\gamma) - \mu_\gamma \quad (*)$

Lagrange multipliers = sensitivities: $\frac{d}{d\gamma}\mathcal{J}(\mathbf{a}_\gamma) = \frac{d}{d\gamma}\mathcal{L}(\mathbf{a}_\gamma,\mu_\gamma) = \mu_\gamma$

Taylor expansion $\mathcal{J}(\mathbf{a}_{\gamma=0}) \approx \mathcal{J}(\mathbf{a}_{\gamma_*}) - \frac{d}{d\gamma}\mathcal{J}(\mathbf{a}_{\gamma_*})\,\gamma_* = \mathcal{J}(\mathbf{a}^{\tilde{\mathcal{I}}^k}) - \mu_{\gamma_*}\gamma_*$

$\Rightarrow c^{l_*} := \mu_{\gamma_*}\gamma_* \overset{(*)}{=} \frac{\partial\mathcal{J}}{\partial a_{l_*}}(\mathbf{a}^{\tilde{\mathcal{I}}^k})\gamma_*$ ... coarsening indicator

Multilevel Refinement and Coarsening Algorithm

$k = 0$: Minimize $\mathcal{J}$ on starting index set $\mathcal{I}^0$ $\rightsquigarrow$ minimal value $\mathcal{J}^0$

Do until refinement indicators $= 0$

Refinement: compute refinement indicators r^{i_*}, $i_* \notin \mathcal{I}^k$

choose index sets $\mathcal{I}^k \cup \{i_*\}$ with largest r^{i_*}

Minimize $\mathcal{J}$ on each of these index sets

and keep $\tilde{\mathcal{I}} := \mathcal{I}^k \cup \{i_*\}$ with largest reduction in $\mathcal{J}$ $\rightsquigarrow$ $\tilde{\mathcal{J}}$

Coarsening (only if $\tilde{\mathcal{J}} < \mathcal{J}^k$): evaluate coarsening indicators c^{l_*}

chooose index sets $\tilde{\mathcal{I}}^k \setminus \{l_*\}$ with largest c^{l_*}

Minimize $\mathcal{J}$ on each of these index sets

and keep $\overline{\mathcal{I}} := \tilde{\mathcal{I}}^k \setminus \{l_*\}$ with largest reduction in $\mathcal{J}$ $\rightsquigarrow$ $\overline{\mathcal{J}}$

If $\overline{\mathcal{J}} \leq \tilde{\mathcal{J}} + \rho(\mathcal{J}^k - \tilde{\mathcal{J}})$ (coarsening does not deteriorate optimal value too much)

set $\mathcal{I}^{k+1} := \overline{\mathcal{I}}$, $\mathcal{J}^{k+1} := \overline{\mathcal{J}}$ (refinement and coarsening)

Else set $\mathcal{I}^{k+1} := \tilde{\mathcal{I}}$, $\mathcal{J}^{k+1} := \tilde{\mathcal{J}}$ (refinement only)

Convergence Proof

For fixed $N < \infty$, Algorithm stops after finitely many steps $k = K$;

$q^K := \sum_{i \in \mathcal{I}^K} a_i^K \phi_i$

- $\mathbf{a}^K$ solves $(P^{\mathcal{I}^K}) \;\Rightarrow\; 0 = \nabla \mathcal{J}(\mathbf{a}^K) \;\Rightarrow$
  $0 = \langle F(q^K) - g^\delta, F'(q^K)\phi_i \rangle \; \forall i \in \mathcal{I}^K$
- refinement indicators vanish $\;\Rightarrow$
  $0 = r^{i_*} = \langle F(q^K) - g^\delta, F'(q^K)\phi_i \rangle \; \forall i \notin \mathcal{I}^K$

$$\Rightarrow \qquad \mathrm{Proj}_{X_N} F'(q^K)^*(F(q^K) - g^\delta) = 0$$

Stability and convergence follow from (existing) results on regularization by discretization

[BK&Offtermatt '09, '10]

Remarks

- more systematic coarsening based on problem specific properties
  (related dofs due to local closeness in groundwater example)
- Lagrange multipliers = gradient components (but we do not carry out gradient steps!): possible improvement by taking into account Hessian information (Newton type)
- Greedy type approach (Burger&Hofinger'04, Denis&Lorenz&Trede'09)
- relation active set strategy $\leftrightarrow$ semismooth Newton (Hintermüller&Ito&Kunisch'03)

2nd approach:

goal oriented error estimators

Tikhonov Regularization and the Discrepancy Principle

Parameter identification as a nonlinear operator equation

$$F(q) = g$$

$g^\delta \approx g \ldots$ given data; noise level $\delta \geq \|g^\delta - g\|$
$F \ldots$ forward operator: $F(q) = (C \circ S)(q) = C(u)$ where $u = S(q)$ solves

$$A(q,u)(v) = (f,v) \quad \forall v \in V \quad \ldots \text{ PDE in weak form}$$

Tikhonov regularization:

$$\text{Minimize} \quad j_\alpha(q) = \|F(q) - g^\delta\|^2 + \alpha\|q\|^2 \text{ over } q \in Q\,,$$

Choice of α: *discrepancy principle* (fixed constant $\tau \geq 1$)

$$\|F(q^\delta_{\alpha_*}) - g^\delta\| = \tau\delta$$

Convergence analysis: [Engl& Hanke& Neubauer 1996] and references there

Goal Oriented Error Estimators in PDE Constrained Optimization (I)

[Becker&Kapp&Rannacher'00], [Becker&Rannacher'01], [Becker&Vexler '04, '05]

$$\begin{array}{ll} \text{Minimize} \quad J(q,u) & \text{over } q \in Q\,,\ u \in V \\ \text{under the constraints} & A(q,u)(v) = f(v) \quad \forall v \in V\,, \end{array}$$

Lagrange functional:

$$\mathcal{L}(q,u,z) = J(q,u) + f(z) - A(q,u)(z)\,.$$

First order optimality conditions:

$$\mathcal{L}'(q,u,z)[(p,v,y)] = 0 \quad \forall (p,v,y) \in Q \times V \times V \qquad (1)$$

Discretization $Q_h \subseteq Q$, $V_h \subseteq V \rightsquigarrow$ discretized version of (1).

Estimate discretization error in some *quantity of interest* I:

$$I(q,u) - I(q_h,u_h) \leq \eta$$

Goal Oriented Error Estimators (II)

Auxiliary functional:

$$\mathcal{M}(q,u,z,p,v,y) = I(q,u) + \mathcal{L}'(q,u,z)[(p,v,y)] \quad (q,u,z,p,v,y) \in (Q \times$$

Consider additional equations:

$$\mathcal{M}'(x_h)(dx_h) = 0 \quad \forall dx_h \in X_h = (Q_h \times V_h \times V_h)^2$$

Proposition ([Becker&Vexler, J. Comp. Phys., 2005]:

$$I(q,u) - I(q_h,u_h) = \underbrace{\frac{1}{2}\mathcal{M}'(x_h)(x - \tilde{x}_h)}_{=:\eta} + O(\|x - x_h\|^3) \quad \forall \tilde{x}_h \in X_h \,.$$

error estimator η = sum of local contributions due to q, u, z, p, v, y:

$$\eta = \sum_{i=1}^{N_q} \eta_i^q + \sum_{i=1}^{N_u} \eta_i^u + \sum_{i=1}^{N_z} \eta_i^z + \sum_{i=1}^{N_p} \eta_i^p + \sum_{i=1}^{N_v} \eta_i^v + \sum_{i=1}^{N_y} \eta_i^y$$

$\rightsquigarrow$ local refinement separately for $q \in Q_h$, $u \in V_h$, $z \in V_h$, ...

Choice of Quantity of Interest ?

aim:
recover infinite dim. convergence results for Tikhonov + discr. princ.
in the adaptively discretized setting

challenge: carrying over infinite dimensional results is

…straightforward if we can guarantee smallness of operator norm
$\|F_h - F\|$
$\rightsquigarrow$ *huge number of quantities of interest!*

…not too hard if we can guarantee smallness of
$\|F_h(q^\dagger) - F(q^\dagger)\|$
$\rightsquigarrow$ *large number of quantities of interest!*

…but we only want to guarantee precision of
one or two quantities of interest

Convergence Analysis ⇝ Choice of Quantity of Interest

Proposition [Griesbaum&BK& Vexler'07], [BK& Kirchner&Vexler'10]:

$\alpha_* = \alpha_*(\delta, g^\delta)$ and $Q_h \times V_h \times V_h$ such that for

$$I(q,u) := \|C(u) - g^\delta\|_G^2 = \|F(q) - g^\delta\|_G^2$$

$$\underline{\underline{\tau}}^2 \delta^2 \leq I(q_{h,\alpha_*}^\delta, u_{h,\alpha_*}^\delta) \leq \overline{\overline{\tau}} \delta^2$$

(i) If additionally

$$|I(q_{h,\alpha_*}^\delta, u_{h,\alpha_*}^\delta) - I(q_{\alpha_*}^\delta, u_{\alpha_*}^\delta)| \leq c I(q_{h,\alpha_*}^\delta, u_{h,\alpha_*}^\delta)$$

for some sufficiently small constant $c > 0$ then $q_{\alpha_*}^\delta \longrightarrow q^\dagger$ as $\delta \to 0$.

Optimal rates under source conditions (logarithic/Hölder).

Proposition [Griesbaum&BK& Vexler'07], [BK&Kirchner&Vexler'10]:

$\alpha_* = \alpha_*(\delta, g^\delta)$ and $Q_h \times V_h \times V_h$ such that for

$$I(q,u) := \|C(u) - g^\delta\|_G^2 = \|F(q) - g^\delta\|_G^2$$

$$\underline{\underline{\tau}}^2\delta^2 \leq I(q^\delta_{h,\alpha_*}, u^\delta_{h,\alpha_*}) \leq \overline{\overline{\tau}}\delta^2$$

(ii) If additionally for

$$I_2(q,u) := J_\alpha(q,u)$$

Convergence Analysis $\rightsquigarrow$ Choice of Quantity of Interest

Proposition [Griesbaum&BK& Vexler'07], [BK&Kirchner&Vexler'10]:

$\alpha_* = \alpha_*(\delta, g^\delta)$ and $Q_h \times V_h \times V_h$ such that for

$$I(q,u) := \|C(u) - g^\delta\|_G^2 = \|F(q) - g^\delta\|_G^2$$

$$\underline{\underline{\tau}}^2 \delta^2 \leq I(q^\delta_{h,\alpha_*}, u^\delta_{h,\alpha_*}) \leq \overline{\overline{\tau}} \delta^2$$

(ii) If additionally for

$$I_2(q,u) := J_\alpha(q,u)$$

$$|I_2(q^\delta_{h,\alpha_*}, u^\delta_{h,\alpha_*}) - I_2(q^\delta_{\alpha_*}, u^\delta_{\alpha_*})| \leq \sigma\delta^2$$

for some constant $C > 0$ with $\underline{\underline{\tau}}^2 \geq 1 + \sigma$, then $q^\delta_{h,\alpha_*} \longrightarrow q^\dagger$ as $\delta \to 0$

Proposition [Griesbaum&BK& Vexler'07], [BK&Kirchner&Vexler'10]:

$\alpha_* = \alpha_*(\delta, g^\delta)$ and $Q_h \times V_h \times V_h$ such that for

$$I(q,u) := \|C(u) - g^\delta\|_G^2 = \|F(q) - g^\delta\|_G^2$$

$$\underline{\underline{\tau}}^2\delta^2 \leq I(q^\delta_{h,\alpha_*}, u^\delta_{h,\alpha_*}) \leq \overline{\overline{\tau}}\delta^2$$

(ii) If additionally for

$$I_2(q,u) := J_\alpha(q,u)$$

$$|I_2(q^\delta_{h,\alpha_*}, u^\delta_{h,\alpha_*}) - I_2(q^\delta_{\alpha_*}, u^\delta_{\alpha_*})| \leq \sigma\delta^2$$

for some constant $C > 0$ with $\underline{\underline{\tau}}^2 \geq 1 + \sigma$, then $q^\delta_{h,\alpha_*} \longrightarrow q^\dagger$ as $\delta \to 0$

see also [Neubauer&Scherzer 1990]

J as quantity of interest $\rightsquigarrow$ [Becker&Kapp&Rannacher'00], [Becker&Rannacher'01],

Idea of Proof

error bound $|J_{\alpha_*}(q^\delta_{h,\alpha_*}, u^\delta_{h,\alpha_*}) - J_{\alpha_*}(q^\delta_{\alpha_*}, u^\delta_{\alpha_*})| \leq \sigma\delta^2$ and optimality of $q^\delta_{\alpha_*}, u^\delta_{\alpha_*}$ imply

$$J_{\alpha_*}(q^\delta_{h,\alpha_*}, u^\delta_{h,\alpha_*}) \leq J_{\alpha_*}(q^\delta_{\alpha_*}, u^\delta_{\alpha_*}) + \sigma\delta^2 \leq J_{\alpha_*}(q^\dagger, u^\dagger) + \sigma\delta^2$$

on the other hand, by the discrepancy principle $\underline{\underline{\tau}}^2\delta^2 \leq \|F(q^\delta_{h,\alpha_*}) - g^\delta\|^2 \leq \overline{\overline{\tau}}\delta^2$ and the definition of the cost functional $J_\alpha(q,u) = \|F(q) - g^\delta\|^2 + \alpha\|q\|^2$

$$J_{\alpha_*}(q^\delta_{h,\alpha_*}, u^\delta_{h,\alpha_*}) \geq \underline{\underline{\tau}}^2\delta^2 + \alpha_*\|q^\delta_{h,\alpha_*}\|^2$$

$$J_{\alpha_*}(q^\dagger, u^\dagger) \leq \delta^2 + \alpha_*\|q^\dagger\|^2$$

Combining these estimates and the choice $\underline{\underline{\tau}}^2 > 1 + \sigma$ we get

$$\|q^\delta_{h,\alpha_*}\|^2 \leq \|q^\dagger\|^2 + \frac{1}{\alpha_*}(1 + \sigma - \underline{\underline{\tau}}^2)\delta^2 \leq \|q^\dagger\|^2 .$$

The rest of the proof is standard.

(Also works for stationary points q^δ_{h,α_*} instead of global minimizers.)

Remarks

- goal oriented error estimators allow to control the error in some quantity of interest
- suff. small error in residual norm $i(\frac{1}{\alpha})$ and its derivative $i'(\frac{1}{\alpha})$
  $\Rightarrow$ fast convergence of Newton's method for choosing α_* (discr. princ.)
  $\rightsquigarrow$ coarse grids at the beginning of Newton's method
  $\rightarrow$ save computational effort
- sufficiently small error in residual norm and Tikhonov functional
  $\Rightarrow$ convergence of Tikhonov regularization preserved
- other regularization methods:
  regularization by discretization [BK&Kirchner&Vexler]
  IRGNM [BK&Veljovic]

$\rightarrow$ other regularization parameter choice strategies: e.g., balancing principle