A Posteriori Estimates for Nonlinear Boundary Value Problems

S. Repin

S. Petersburg Department of V.A. Steklov Institute of Mathematics

A fundamental question, which must be answered before

A fundamental question, which must be answered before computations,

A fundamental question, which must be answered before

computations,

mesh refinement,

A fundamental question, which must be answered before

computations,

mesh refinement,

error indicators, estimators, etc.

A fundamental question, which must be answered before
computations,
mesh refinement,
error indicators, estimators, etc.
is as follows:

What the words "accurate numerical solution" mean?

A fundamental question, which must be answered before

computations,

mesh refinement,

error indicators, estimators, etc.

is as follows:

What the words "accurate numerical solution" mean?

or

Which error measure(s) is adequate to the problem studied?

For linear PDEs at least one possible answer is obvious:

the Energy Norm is a right measure.

For linear PDEs at least one possible answer is obvious:

the Energy Norm is a right measure.

However, for many nonlinear models it is still an open question.

Below we discuss possible answers with the paradigm of one class of nonlinear PDE's.

The plan

- Concise introduction (history of a posteriori analysis, new trends, and unsolved problems)

The plan

- Concise introduction (history of a posteriori analysis, new trends, and unsolved problems)
- A posteriori for linear PDEs (key mathematical principles and results)

The plan

- Concise introduction (history of a posteriori analysis, new trends, and unsolved problems)
- A posteriori for linear PDEs (key mathematical principles and results)
- A class of well posed variational nonlinear problems

The plan

- Concise introduction (history of a posteriori analysis, new trends, and unsolved problems)
- A posteriori for linear PDEs (key mathematical principles and results)
- A class of well posed variational nonlinear problems
- Comments on error control of some ill posed variational problems

The plan

- Concise introduction (history of a posteriori analysis, new trends, and unsolved problems)
- A posteriori for linear PDEs (key mathematical principles and results)
- A class of well posed variational nonlinear problems
- Comments on error control of some ill posed variational problems

Starting from mid **80'**, numerical analysis of PDEs is based on the **Adaptive Modeling** conception, which cannot be realized without indication of a posteriori errors.

Error indicators use different motivations and have different forms:

Starting from mid **80'**, numerical analysis of PDEs is based on the **Adaptive Modeling** conception, which cannot be realized without indication of a posteriori errors.

Error indicators use different motivations and have different forms:
Group A. weighted sums of local residuals and interelement jumps:
I. Babuska and W. Rheinboldt +P. Clement interpolation,
further developed by
T. Oden, R. Nochetto, R. Verfurth, M. Ainsworth, T. Stroboulis, C. Carstensen, R. Hoppe......................

Starting from mid **80'**, numerical analysis of PDEs is based on the **Adaptive Modeling** conception, which cannot be realized without indication of a posteriori errors.

Error indicators use different motivations and have different forms:
Group A. weighted sums of local residuals and interelement jumps:
I. Babuska and W. Rheinboldt +P. Clement interpolation,
further developed by
T. Oden, R. Nochetto, R. Verfurth, M. Ainsworth, T. Stroboulis, C. Carstensen, R. Hoppe......................

Group B. Post processing (averaging)
O. C. Zienkiewicz and J. Z. Zhu, C. Carstensen, J. Wang, X. Ye........,
mathematical justifications are based on superconvergence (first publications A. Oganesjan and L. Rukhovetz, M. Zlamal, J. Bramble and A. Schatz.)

Group C. Indicators obtained with the help of adjoint problems

dual-weighted residual method
and estimates for goal-oriented quantities

R. Rannacher, C. Johnson, E. Suli, T. Oden, S. Prhudome,...

Group C. Indicators obtained with the help of adjoint problems
dual-weighted residual method
and estimates for goal-oriented quantities
R. Rannacher, C. Johnson, E. Suli, T. Oden, S. Prhudome,...

Group D. Advanced versions of the Runge's error indicator
Hierarchically based indicators
R. Bank, C. Schwab, ...

Error indicators generate suitable adaptations of meshes (or basic functions) but in general they do not provide guaranteed bounds of errors.

Fully Reliable Error Control

is a new direction (about 15-20 years). It is based on much stronger tools of error control, which satisfy the following conditions:

Fully Reliable Error Control

is a new direction (about 15-20 years). It is based on much stronger tools of error control, which satisfy the following conditions:

- For a concrete solution the estimate must give a guaranteed and realistic estimate of the error.

Fully Reliable Error Control

is a new direction (about 15-20 years). It is based on much stronger tools of error control, which satisfy the following conditions:

- For a concrete solution the estimate must give a guaranteed and realistic estimate of the error.
- Estimates must be fully computable, CPU time required for the computation should be taken into account as a substantial parameter.

Fully Reliable Error Control

is a new direction (about 15-20 years). It is based on much stronger tools of error control, which satisfy the following conditions:

- For a concrete solution the estimate must give a guaranteed and realistic estimate of the error.
- Estimates must be fully computable, CPU time required for the computation should be taken into account as a substantial parameter.
- The estimate should be applicable to a wide spectrum of approximations, i.e., it should not be based upon special properties of approximations/method (e.g. Galerkin orthogonality);

Fully Reliable Error Control

is a new direction (about 15-20 years). It is based on much stronger tools of error control, which satisfy the following conditions:

- For a concrete solution the estimate must give a guaranteed and realistic estimate of the error.
- Estimates must be fully computable, CPU time required for the computation should be taken into account as a substantial parameter.
- The estimate should be applicable to a wide spectrum of approximations, i.e., it should not be based upon special properties of approximations/method (e.g. Galerkin orthogonality);
- It must not attract extra regularity or other special properties of the exact solution.

Fully Reliable Error Control

is a new direction (about 15-20 years). It is based on much stronger tools of error control, which satisfy the following conditions:

- For a concrete solution the estimate must give a guaranteed and realistic estimate of the error.
- Estimates must be fully computable, CPU time required for the computation should be taken into account as a substantial parameter.
- The estimate should be applicable to a wide spectrum of approximations, i.e., it should not be based upon special properties of approximations/method (e.g. Galerkin orthogonality);
- It must not attract extra regularity or other special properties of the exact solution.

It is clear that above requirements can be satisfied if we consider the problem on the functional level and try to find computable estimates of deviations from the exact solution.

Three key steps, which lead to fully reliable a posteriori estimates for linear PDEs:

1. Helmgholtz type decomposition of vector (tensor) spaces

$$\mathbf{U}(\boldsymbol{\Omega}) := \mathbf{L}^2(\boldsymbol{\Omega}, \mathbb{R}^{\mathbf{d}}) = \mathbf{Q_0}(\boldsymbol{\Omega}) \oplus \mathbf{H}_\nabla(\boldsymbol{\Omega}),$$

where

$$Q_0(\Omega) = \{q \in U \mid \operatorname{div} q = 0\},$$
$$H_\nabla(\Omega) = \{q = \nabla v \mid v \in \overset{\circ}{H}{}^1(\Omega) =: V_0\}.$$

1. Helmgholtz type decomposition of vector (tensor) spaces

$$\mathbf{U}(\boldsymbol{\Omega}) := \mathbf{L}^2(\boldsymbol{\Omega}, \mathbb{R}^{\mathbf{d}}) = \mathbf{Q_0}(\boldsymbol{\Omega}) \oplus \mathbf{H}_\nabla(\boldsymbol{\Omega}),$$

where

$$Q_0(\Omega) = \{q \in U \mid \operatorname{div} q = 0\},$$
$$H_\nabla(\Omega) = \{q = \nabla v \mid v \in \overset{\circ}{H}{}^1(\Omega) =: V_0\}.$$

1. Helmgholtz type decomposition of vector (tensor) spaces

$$\mathbf{U}(\boldsymbol{\Omega}) := \mathbf{L}^{2}(\boldsymbol{\Omega}, \mathbb{R}^{\mathbf{d}}) = \mathbf{Q_0}(\boldsymbol{\Omega}) \oplus \mathbf{H}_{\nabla}(\boldsymbol{\Omega}),$$

where

$$Q_0(\Omega) = \{q \in U \mid \operatorname{div} q = 0\},$$
$$H_{\nabla}(\Omega) = \{q = \nabla v \mid v \in \overset{\circ}{H}{}^{1}(\Omega) =: V_0\}.$$

Yields Prager-Synge (1947) type estimates, e.g. for $\Delta u + f = 0$

$$\|q - \nabla u\|^2 + \|\nabla(\mathbf{v} - \mathbf{u})\|^{\mathbf{2}} = \|\nabla \mathbf{v} - \mathbf{q}\|^{\mathbf{2}} \quad \forall v \in V_0 + u_0,\ q \in Q_f,$$

where

$$Q_f = \{q \in U \mid \operatorname{div} q + f = 0\}.$$

1. Helmgholtz type decomposition of vector (tensor) spaces

$$\mathbf{U}(\mathbf{\Omega}) := \mathbf{L}^2(\mathbf{\Omega}, \mathbb{R}^{\mathbf{d}}) = \mathbf{Q_0}(\mathbf{\Omega}) \oplus \mathbf{H}_\nabla(\mathbf{\Omega}),$$

where

$$Q_0(\Omega) = \{q \in U \mid \operatorname{div} q = 0\},$$
$$H_\nabla(\Omega) = \{q = \nabla v \mid v \in \overset{\circ}{H}{}^1(\Omega) =: V_0\}.$$

Yields Prager-Synge (1947) type estimates, e.g. for $\Delta u + f = 0$

$$\|q - \nabla u\|^2 + \|\nabla(\mathbf{v} - \mathbf{u})\|^2 = \|\nabla \mathbf{v} - \mathbf{q}\|^2 \quad \forall v \in V_0 + u_0,\ q \in Q_f,$$

where

$$Q_f = \{q \in U \mid \operatorname{div} q + f = 0\}.$$

Close ideas were used in the so called "Orthogonal Projection Method":
S. Zaremba (1927), H. Weil (1940), M. Vishik (1947).

2. Mikhlin's identity for quadratic energy functionals

$$\frac{1}{2}a(u-v,u-v)=J_a(v)-J_a(u),$$

where

$$J_a(v):=\frac{1}{2}a(v,v)-(f,v)$$

and u is a mininimizer.

2. Mikhlin's identity for quadratic energy functionals

$$\frac{1}{2}a(u-v,u-v) = J_a(v) - J_a(u),$$

where

$$J_a(v) := \frac{1}{2}a(v,v) - (f,v)$$

and u is a mininimizer.

Yields "classical"duality/equilibration estimates:
S. Miklin (1962), H. Gaevskii, H. Gröger, K. Zaharias (1974),
P. Mosoalov and P. Myasnikov (1981), D. Kelly (1984),

....

3. Estimates of the distance to the set of equilibrated fields

Lemma. Let y be a vector function in $H(\Omega, \mathrm{div}\,)$, then the distance to the set Q_f can be estimated as follows

$$\inf_{q \in Q_f} \|y - q\|^2 \leq \; C \, \|\mathrm{div}\, y + f\|^2,$$

where C is a computable global constant (e.g., Poincaré-Friedrichs constant)

3. Estimates of the distance to the set of equilibrated fields

Lemma. Let y be a vector function in $H(\Omega, \mathrm{div}\,)$, then the distance to the set Q_f can be estimated as follows

$$\inf_{q \in Q_f} \|y - q\|^2 \leq \; C\, \|\mathrm{div}\, y + f\|^2,$$

where C is a computable global constant (e.g., Poincaré-Friedrichs constant)

Yields a posteriori estimates of the functional type.
First derived in 96-97':
Comput. Meth. Appl. Math. Engrng. 96,
Comptes Rendus. Matematique 97
J. Math. Sci. 97, consequent exposition in *Math. Comput. 2000*.

They provide fully guaranteed and directly computable error bounds for any conforming approximation of a PDE.

General divergent type elliptic problem $\Lambda^* A \Lambda u + \ell = 0$

$A : Y \to Y$ positive definite with the ellipticity constant C_a, $\ell \in V^*$
U Hilbert space with $(\cdot, \cdot)$ and $\|\cdot\|$.

$$\Lambda : V \ \to \ U \qquad \Lambda^* : U \ \to \ V^*$$

$$< \Lambda^* y^*, v >= (y^*, \Lambda v)$$

Let

$$\|w\| \leq C_\Lambda \|\Lambda w\| \quad \forall w \in V,$$

Theorem

Let $\mathcal{V}$ be a Hilbert space s.t. $V \in \mathcal{V} \in V^*$.
For all $v \in V$

$$\|\Lambda(u - v)\| = \inf_{y \in Q} \{\|\Lambda v - y\|_* + C\|\Lambda^* y + \ell\|_{\mathcal{V}}\}$$

where $Q := \{y \in U, \Lambda^* y \in \mathcal{V}\}$ and $C = \frac{C_\Lambda}{C_A}$.

Theorem

Let $\mathcal{V}$ be a Hilbert space s.t. $V \in \mathcal{V} \in V^*$.
For all $v \in V$

$$|\!|\!|\Lambda(u-v)|\!|\!| = \inf_{y \in Q} \{|\!|\!|\Lambda v - y|\!|\!|_* + C\|\Lambda^* y + \ell\|_{\mathcal{V}}\}$$

where $Q := \{y \in U,\ \Lambda^* y \in \mathcal{V}\}$ and $C = \frac{C_\Lambda}{C_A}$.

Corollary 1:

$$M_{\oplus}(v, y) = |\!|\!|\Lambda v - y|\!|\!|_* + C\|\Lambda^* y + \ell\|_{\mathcal{V}}$$

is a computable majorant of the error for any $y \in Q$.
Corollary 2:
Majorant has NO GAP!
In other words, problem is INDEED FULLY CONTROLABLE!

Example. Stationary diffusion model

$$
\begin{aligned}
-\mathrm{div}\, p &= f && \text{in } \Omega, && (1)\\
p &= A\nabla u && \text{in } \Omega, && (2)\\
u &= u_0 && \text{on } \Gamma, && (3)
\end{aligned}
$$

where A is a symmetric matrix satisfying the condition

$$Az \cdot z \geq c_1 |z|^2 \quad \forall z \in \mathbb{R}^d$$

$$|||y|||^2 = \|y\|_A^2 := \int_\Omega Ay \cdot y dx \quad \text{and} \quad |||y|||_*^2 = \|y\|_{A^{-1}}^2 := \int_\Omega A^{-1} y \cdot y dx$$

are the norms equivalent to the natural norm of $Q(\Omega) := L^2(\Omega, \mathbb{R}^d)$ and C_Λ is a constant in the inequality

$$\|w\| \leq C \|\nabla w\|_A \qquad \forall w \in V_0.$$

Since $C = c_1^{-1} C_{F\Omega}$, if $C_{F\Omega}$ (or a computable upper bound of it) is known, then C is easily computable and we arrive at

Since $C = c_1^{-1} C_{F\Omega}$, if $C_{F\Omega}$ (or a computable upper bound of it) is known, then C is easily computable and we arrive at

$$\|\nabla(u-v)\|_A = \\ = \inf_{y \in H(\Omega,\mathrm{div})} \left\{ \int_\Omega \left(A\nabla v \cdot \nabla v + A^{-1} y \cdot y - y \cdot \nabla v \right) dx + \right. \\ \left. + C\|f + \mathrm{div}\, y\| \right\}$$

Since $C = c_1^{-1} C_{F\Omega}$, if $C_{F\Omega}$ (or a computable upper bound of it) is known, then C is easily computable and we arrive at

$$\|\nabla(u-v)\|_A = \\ = \inf_{y\in H(\Omega,\mathrm{div})} \Big\{ \int_\Omega \left(A\nabla v \cdot \nabla v + A^{-1} y \cdot y - y \cdot \nabla v \right) dx + \\ + C\|f + \mathrm{div}\, y\| \Big\}$$

Such type estimates has been derived and tested for reaction-convection-diffusion,

Since $C = c_1^{-1} C_{F\Omega}$, if $C_{F\Omega}$ (or a computable upper bound of it) is known, then C is easily computable and we arrive at

$$\|\nabla(u-v)\|_A = \\ = \inf_{y \in H(\Omega, \mathrm{div})} \left\{ \int_\Omega \left(A\nabla v \cdot \nabla v + A^{-1} y \cdot y - y \cdot \nabla v \right) dx + \right. \\ \left. + C \| f + \mathrm{div}\, y \| \right\}$$

Such type estimates has been derived and tested for
reaction-convection-diffusion,

linear elasticity,

Since $C = c_1^{-1} C_{F\Omega}$, if $C_{F\Omega}$ (or a computable upper bound of it) is known, then C is easily computable and we arrive at

$$\|\nabla(u-v)\|_A = \\ = \inf_{y \in H(\Omega,\mathrm{div})} \left\{ \int_\Omega \left(A\nabla v \cdot \nabla v + A^{-1} y \cdot y - y \cdot \nabla v \right) dx + \right. \\ \left. + C \| f + \mathrm{div}\, y \| \right\}$$

Such type estimates has been derived and tested for
reaction-convection-diffusion,
linear elasticity,
Maxwell,

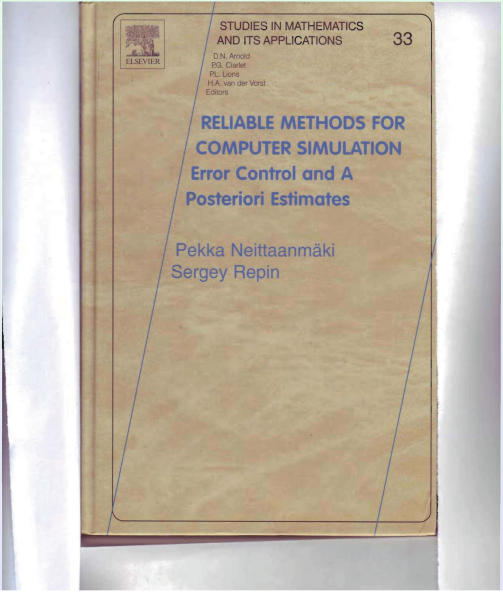
Since $C = c_1^{-1} C_{F\Omega}$, if $C_{F\Omega}$ (or a computable upper bound of it) is known, then C is easily computable and we arrive at

$$\|\nabla(u - v)\|_A = \\ = \inf_{y \in H(\Omega, \mathrm{div})} \left\{ \int_\Omega \left(A\nabla v \cdot \nabla v + A^{-1} y \cdot y - y \cdot \nabla v \right) dx + \right. \\ \left. + C \| f + \mathrm{div}\, y \| \right\}$$

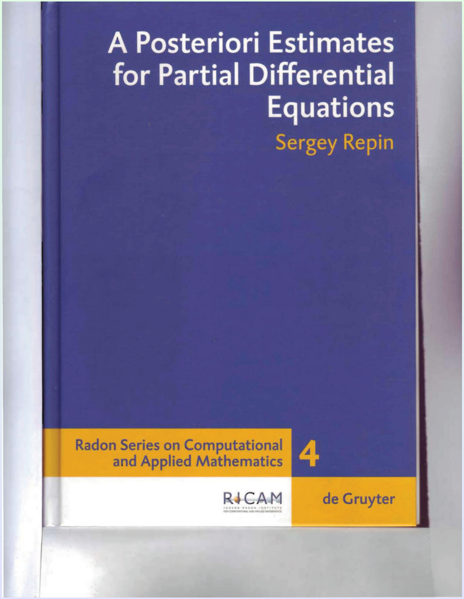
Such type estimates has been derived and tested for
reaction-convection-diffusion,
linear elasticity,
Maxwell,
Stokes, Oseen

....

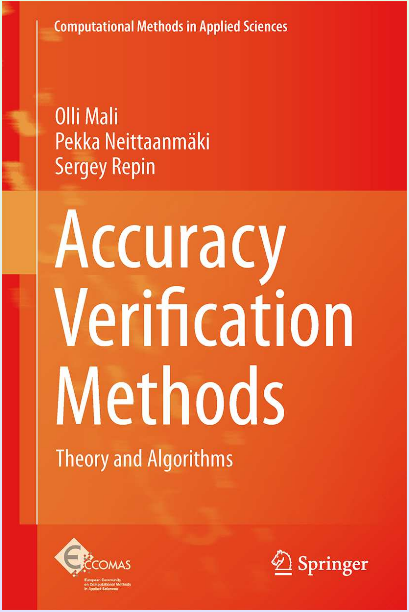
A systematic exposition of the **variational (duality) method**: Elsevier, 2004:



Nonvariational method: Walter de Gruyter, 2008



Numerical and algorithmic aspects: Springer (in press)



Nonlinear Problems

Linear problems are all alike;
every nonlinear problem
is nonlinear in its own way.

A class of nonlinear models

Consider the class of variational problems

$$\inf_{w\in V}\left\{G(\Lambda w)+<\ell,w>\right\} \qquad (A)$$

$G : Y \to \mathbb{R}_+$: convex, continuous, coercive functional vanishing at zero element of Y

A class of nonlinear models

Consider the class of variational problems

$$\inf_{w\in V}\left\{G(\Lambda w)+<\ell,w>\right\} \qquad (A)$$

$G : Y \to \mathbb{R}_+$: convex, continuous, coercive functional vanishing at zero element of Y

$$\begin{array}{cccc} V & V^* & <v^*,v> & \Lambda : V \to Y \\ Y & Y^* & \langle y^*,y\rangle & \Lambda^* : V^* \to Y^* \end{array}$$

$$<\Lambda^* y^*, v> = \langle y^*, \Lambda v\rangle$$

It includes, e.g., α-Laplacian, NonNewtonian fluids, nonlinear models in the theory of solids (e.g., deformation plasticity).

Example. $V = \overset{\circ}{W}{}^{1,\alpha}(\Omega)$, $Y = L^{\alpha}(\Omega, \mathbb{R}^d)$, $Y^* = L^{\alpha'}(\Omega, \mathbb{R}^d)$, $\frac{1}{\alpha} + \frac{1}{\alpha'} = 1$, $\alpha \in (1, +\infty)$

$$\begin{aligned} \Lambda &= \nabla, \quad \Lambda^* = -\mathrm{div}\,, \\ G(y) &= \frac{1}{\alpha}\int_\Omega |y|^\alpha \, dx, \\ J(v) &= \frac{1}{\alpha}\int_\Omega |\nabla v|^\alpha dx - \int_\Omega f v dx. \end{aligned}$$

Euler equation leads to α-Laplacian

$$\mathrm{div}\,|\nabla u|^{\alpha-2}\nabla u + f = 0, \ \mathrm{in}\,\Omega, \quad u = 0 \ \mathrm{on}\ \Gamma.$$

In order to obtain a unified theory, which encompasses linear theory as a special case, we must use special functionals (called compound functionals) instead of norms.

In order to obtain a unified theory, which encompasses linear theory as a special case, we must use special functionals (called compound functionals) instead of norms.

Compound is defined on elements of two complementary spaces X and X^*

$$D_g\,(\xi,\xi^*) := g(\xi) + g^*(\xi^*) - \langle \xi^*, \xi \rangle \geq 0.$$

g^* is the Young-Fenchel conjugate of g.

Important property: $D(\xi,\xi^*) = 0 \;\Leftrightarrow\; \xi^* \subset \partial g(\xi).$

$D_g(\xi, \xi^*)$ can be converted into a norm only in very special cases, where Y is isometrically equivalent to Y^*.

Example. Poisson problem $\Delta u + f = 0$

$V = \overset{\circ}{H}{}^1(\Omega)$, $Y = Y^* = L^2(\Omega, \mathbb{R}^d)$,

$g(\xi) = \frac{1}{2}\|\xi\|^2$, $g^*(\xi^*) = \frac{1}{2}\|\xi^*\|^2$.

Then $D_g(\xi, \xi^*) = \frac{1}{2}\|\xi - \xi^*\|^2$.

Theorem (PS type identity for Problem A)

Let (u, p^*) be the exact solution and exact dual solution, $v \in V$, $q^* \in Y^*_\ell := \{\Lambda^* q^* + \ell = 0\}$. Then

$$\underbrace{D_G(\Lambda v, p^*)}_{\text{measure for } v} + \underbrace{D_G(\Lambda u, q^*)}_{\text{measure for } q^*} = \underbrace{D_G(\Lambda v, q^*)}_{\text{computable}}.$$

Meaning:
Here v is a computed solution and q^* is a computed "flux"("stress"). The left hand side is a certain measure of the distance to (u, p^*). In fact, this measure is a proper one.

Simplest linear case:

$\Lambda = \nabla,\ G(y) = \frac{1}{2}\int_\Omega |y|^2 dx,\ <\ell, v>= \int_\Omega fv\, dx$

Simplest linear case:

$\Lambda = \nabla,\ G(y) = \frac{1}{2}\int_\Omega |y|^2 dx,\ <\ell, v>= \int_\Omega fv\, dx$

$D_G(\Lambda v, p^*) = \frac{1}{2}\|\nabla v - p^*\|^2,$

Simplest linear case:

$\Lambda = \nabla,\ G(y) = \frac{1}{2}\int_\Omega |y|^2 dx,\ <\ell, v>= \int_\Omega fv\, dx$

$D_G(\Lambda v, p^*) = \frac{1}{2}\|\nabla v - p^*\|^2,$

$D_G(\Lambda u, q^*) = \frac{1}{2}\|\nabla u - q^*\|^2,$

Simplest linear case:

$\Lambda = \nabla,\ G(y) = \frac{1}{2}\int_\Omega |y|^2 dx,\ <\ell, v> = \int_\Omega fv\, dx$

$D_G(\Lambda v, p^*) = \frac{1}{2}\|\nabla v - p^*\|^2,$

$D_G(\Lambda u, q^*) = \frac{1}{2}\|\nabla u - q^*\|^2,$

$D_G(\Lambda v, q^*) = \frac{1}{2}\|\nabla v - q^*\|^2,$

Simplest linear case:

$\Lambda = \nabla,\ G(y) = \frac{1}{2}\int_\Omega |y|^2 dx,\ <\ell, v>= \int_\Omega fv\, dx$

$D_G(\Lambda v, p^*) = \frac{1}{2}\|\nabla v - p^*\|^2,$

$D_G(\Lambda u, q^*) = \frac{1}{2}\|\nabla u - q^*\|^2,$

$D_G(\Lambda v, q^*) = \frac{1}{2}\|\nabla v - q^*\|^2,$

Theorem results in the PS identity:

$$\|\nabla v - p^*\|^2 + \|\nabla u - q^*\|^2 = \|\nabla v - q^*\|^2.$$

Theorem (Duality gap identity)

For $v \in V$, $q^* \in Y^*_\ell := \{\Lambda^* q^* + \ell = 0\}$, it holds:

$$D_G(\Lambda v, p^*) + D_G(\Lambda u, q^*) = J(v) - I^*(q^*)$$

Corollary:
Set $q^* = p^*$. Then $(I^*(q^*) = J(u)\ !)$

$$D_G(\Lambda v, p^*) = J(v) - J(u)$$

Mikhlin's type identity holds only for the nonlinear measure D_G!

Conclusion:

$$\mu\!\!\mu(v, q^*) := D_G(\Lambda v, p^*) + D_G(\Lambda u, q^*)$$

is an adequate "mixed measure"of the distance between (v, q^*) and (u, p^*) on $V \times Y^*_\ell$.

Motivation: If $\mu\!\!\mu(v - u, q^* - p^*)$ is large, then either

- $J(v) >> J(u)$,
- or $I^*(q^*) << I^*(p^*)$,
- or both.

Conclusion:

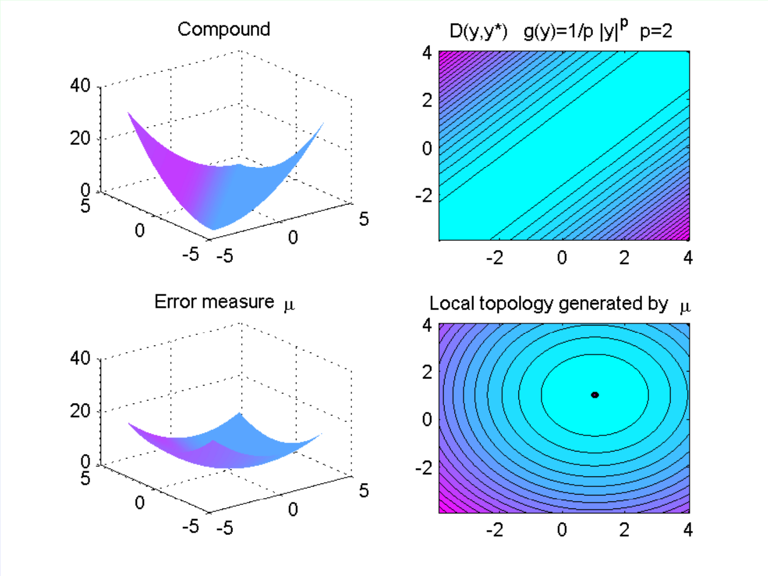
$$\mu\!\!\!\mu(v, q^*) := D_G(\Lambda v, p^*) + D_G(\Lambda u, q^*)$$

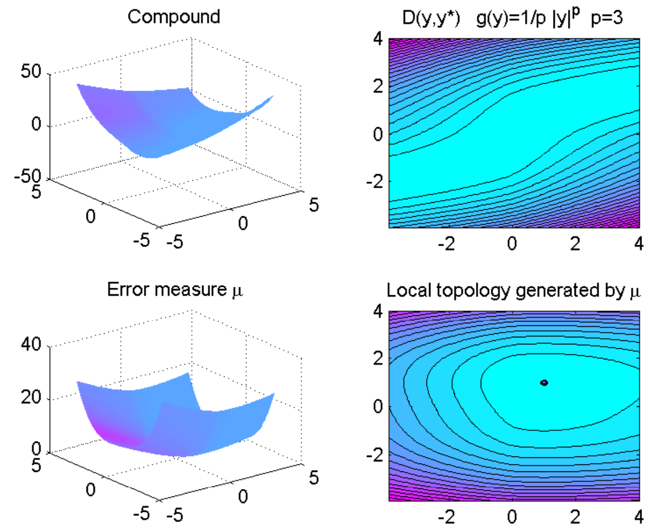
is an adequate "mixed measure"of the distance between (v, q^*) and (u, p^*) on $V \times Y^*_\ell$.

Motivation: If $\mu\!\!\!\mu(v - u, q^* - p^*)$ is large, then either

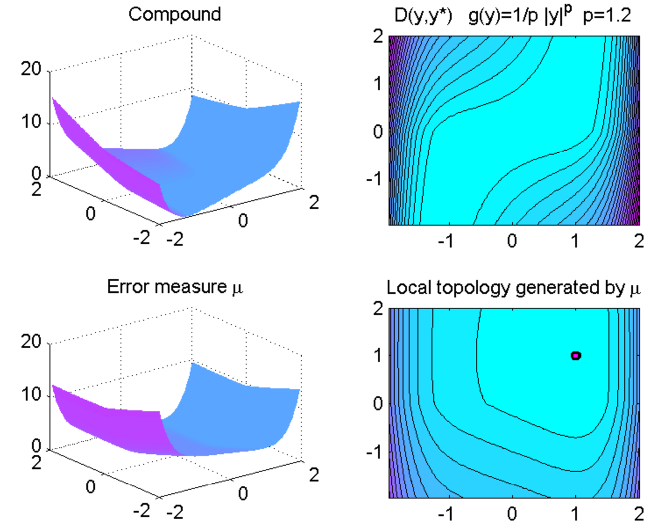
- $J(v) >> J(u)$,
- or $I^*(q^*) << I^*(p^*)$,
- or both.

In general, D_G is not a convex functional, but the measure $\mu\!\!\!\mu(v, p^*)$ generates a convex! topology at the vicinity of the exact pair (u, p^*).


Compound
40
20
0
5
0
-5
-5
0
5
D(y,y*) g(y)=1/p |y|^p p=2
4
2
0
-2
-2
0
2
4
Error measure μ
40
20
0
5
0
-5
-5
0
5
Local topology generated by μ
4
2
0
-2
-2
0
2
4



Compound
50
0
-50
5
0
-5
-5
0
5
D(y,y*) g(y)=1/p |y|^p p=3
4
2
0
-2
-2
0
2
4
Error measure μ
40
20
0
5
0
-5
-5
0
5
Local topology generated by μ
4
2
0
-2
-2
0
2
4


Compound
20
10
0
2
0
-2
-2
0
2
D(y,y*) g(y)=1/p |y|^p p=1.2
2
1
0
-1
-1
0
1
2
Error measure μ
20
10
0
2
0
-2
-2
0
2
Local topology generated by μ
2
1
0
-1
-1
0
1
2

Corollary: primal and dual PS type estimates

The measure $\mu\!\!\mu(v, q^*)$ can be split into two separate measures $\mu\!\!\mu(v)$ and $\mu\!\!\mu(q^*)$, for which we have guaranteed bounds.

$$D_G(\Lambda v, p^*) + D_G(\Lambda u, q^*) = D_G(\Lambda v, q^*)$$

$$\mu\!\!\mu(v) := D_G(\Lambda v, p^*) \leq D_G(\Lambda v, q^*), \ \forall q^* \in Y_\ell^* \qquad (PS\ type\ estimate).$$

Meaning:
$\mu\!\!\mu(v)$ is a convex measure of the distance from Λv to p^* is majorated by computable quantity , $D_G(\Lambda v, q^*)$, where q^* is ANY in Y_ℓ^*.

Dual version:

$$D_G(\Lambda v, p^*) + D_G(\Lambda u, q^*) = D_G(\Lambda v, q^*)$$

$$\mu\!\!\!\mu^*(q^*) := D_G(\Lambda u, q^*) \le D_G(\Lambda v, q^*) \; \forall v \in V \qquad (\textit{dual PS estimate}).$$

Meaning: Convex measure $\mu\!\!\!\mu^*(q^*)$ of the distance from q^* to Λu is majorated by computable quantity , $D_G(\Lambda v, q^*)$, where v is ANY in V.

Other suitable error measures

Assume that

$$G \text{ is differentiable} \tag{4}$$

and uniformly convex, i.e.,

$$G(\frac{y_1 + y_2}{2}) + \Phi(y_1 - y_2) \leq \frac{1}{2}G(y_1) + \frac{1}{2}G(y_2) \qquad \forall y_1, y_2 \in \tag{5}$$

where $\Phi : Y \to \mathbb{R}^+$ is a forcing functional.
Define

$$\mu\!\!\mu^+(v) := \left\langle G'(\Lambda u) - G'(\Lambda v), \Lambda v - \Lambda u \right\rangle,$$
$$\mu\!\!\mu^-(v) := \Phi(\Lambda(v - u)).$$

Theorem

If G satisfies (1) and (2) then

$$\mu\!\!\mu^-(\mathbf{v}) \leq \mu\!\!\mu(\mathbf{v}) \leq \mu\!\!\mu^+(\mathbf{v}) \qquad \forall \mathbf{v} \in \mathbf{V}.$$

Theorem I guarantees that:

$$\mu\!\!\!\mu(v) = \inf_{q^* \in Y_\ell^*} D(\Lambda v, q^*),$$
$$\mu\!\!\!\mu(q^*) = \inf_{v \in V} D(\Lambda v, q^*),$$

i.e., both estimates have no "gap".

Our goal is to deduce a fully computable majorant with the same property defined on a mich wider set
$Q^* = \{y^* \in Y^*, \Lambda^* y^* \in U\}$ WITHOUT EQUILIBRATION CONDITIONS.
U is a Hilbert space: $V \subset U \subset V^*$.
Simple example: $Q^* = H(\Omega, \mathrm{div}\,)$.

Lemma (Distance to the set of "equilibrated"fields Y^*_ℓ)

Assume that there exists a nonnegative continuous functional $H : V \to \mathbb{R}_+$ such that

$$G(\Lambda w) \geq H(w) \qquad \forall w \in V \tag{6}$$

where $H^* : V^* \to \mathbb{R}_+$ is the Young– Fenchel conjugate to H. Then for any $y^* \in Q^*$, the following estimate holds

$$\inf_{q^* \in Y^*_\ell} G^*(y^* - q^*)\, dx \leq H^*(\mathcal{R}), \tag{7}$$

where $\mathcal{R} : V \to \mathbb{R}$ is a linear functional defined by

$$\langle \mathcal{R}, w \rangle := \langle y^*, \Lambda w \rangle - < \ell, w > .$$

Note: Y^*_ℓ contains y^* such that $\langle \mathcal{R}, w \rangle = 0$ for any $w \in V$.

Example. $V = \overset{\circ}{W}{}^{1,\alpha}$, $G(y) = \frac{1}{\alpha}\|y\|_\alpha^\alpha$, and $G^*(y^*) = \frac{1}{\alpha'}\|y^*\|_{\alpha'}^{\alpha'}$.
Since $\|w\|_\alpha \leq C_F \|\nabla w\|_\alpha$.

$$\frac{1}{\alpha}\int_\Omega |\nabla w|^\alpha \, dx = G(\nabla w) \geq \frac{1}{\alpha C_F^\alpha}\|w\|^\alpha = H(w).$$

If $w^* \in L^{\alpha'}(\Omega)$, then

$$H^*(w^*) = \sup_{w \in V}\left\{\int_\Omega w^* w \, dx - \frac{1}{\alpha C_F^\alpha}\|w\|^\alpha\right\} = \frac{C_F^{\alpha'}}{\alpha'}\|w^*\|_{\alpha'}^{\alpha'}.$$

Thus, if $\operatorname{div} y^* + \ell \in L^{\alpha'}$ then

$$\inf_{q^* \in Y_\ell^*} G^*(q^* - y^*) \leq \frac{C_F^{\alpha'}}{\alpha'}\|\operatorname{div} y^* - \ell\|_{\alpha'}^{\alpha'}.$$

Theorem (General form of the error majorant; *Russ. J. Numer. Anal.* 2012)

For any $v \in V$

$$\mu(v) = \inf_{\substack{y^* \in Y_Z \\ \lambda \in (0,1)}} D_G(\Lambda v, y^*) + H^*\left(\frac{\mathcal{R}}{1-\lambda}\right) + \Re(\lambda, y^*),$$

where

$$\Re(\lambda, y^*) = \lambda G^*\left(\frac{y^*}{\lambda}\right) - G^*(y^*) + \langle y^*, \Lambda v\rangle - < \ell, v >$$

and $\mu(v) = D_G(\Lambda v, p^*)$ is the above defined nonlinear measure.

Example.

$G(\Lambda w) = \frac{1}{\alpha} \int_\Omega |\nabla w|^\alpha \, dx$. $G^*(y^*) = \frac{1}{\alpha'} \int_\Omega |y^*|^{\alpha'} dx$.

$$\begin{aligned} \mu(v) \leq \; & \int_\Omega \left(\frac{1}{\alpha}|\nabla v|^\alpha + \frac{1}{\alpha'}|y^*|^{\alpha'} - \nabla v \cdot y^*\right) dx + \\ + \; & \frac{C_F^{\alpha'}}{\alpha'(1-\lambda)^{\alpha'}} \|\operatorname{div} y^* - \ell\|_{\alpha'}^{\alpha'} + \\ + \; & \left(\frac{1}{\lambda^{\alpha'}} - 1\right) \frac{1}{\alpha'} \|y^*\|^{\alpha'} + \int_\Omega (y^* \cdot \nabla v - \ell v) dx. \end{aligned}$$

Conclusion:

(a) The majorant is fully computable.

(b) if $\|\operatorname{div} y^* + \ell\|_{\alpha'}$ is small then λ can be set small, three last terms are small and the main part of the error majorant is $D(\nabla v, y^*)$,

(c) in this case, $D(\nabla v, y^*)$ is a good error indicator for mesh refinement.

Decomposition Theorem (generalization of Helmgholtz decomposition) is also expressed in terms of D.
Simplified version for $\Lambda v = \nabla v$:

$$Y^*_\Lambda(\Omega) := \{y^* \in Y^*(\Omega) \mid \exists\, v \in V : D(\nabla v, y^*) = 0\},$$

and

$$Y^*_f(\Omega) := \{y^* \in Y^*(\Omega) \mid \forall\, v \in V_0(\Omega) : \int_\Omega (y^* \cdot \nabla v - fv)\, dx = 0\}.$$

Theorem (*St. Petersburg Math. J.*, 2000)

The sets $Y^*_f(\Omega)$ and $Y^*_\Lambda(\Omega)$ are closed subsets of $Y^*(\Omega)$. The intersection of these sets consists of the single element – solution of Problem A. For any function $y^* \in Y^*(\Omega)$, there exists a unique decomposition

$$y^* = y^*_\Lambda + y^*_f\,, \qquad y^*_f \in Y^*_f(\Omega) \text{ and } y^*_\Lambda \in Y^*_\Lambda(\Omega).$$

The theory is extendable to a much wider class of problems

$$\inf_{w \in V} \left\{ G(\Lambda w) + F(w) \right\} \qquad (B)$$

$G : Y \to \mathbb{R}_+$: is defined as before,
$F : V \to \mathbb{R}_+$: convex, continuous on V.

In particular, this class includes variational inequalities.

Example. The obstacle problem, $\Omega \subset \mathbb{R}^2$

$$K = K_{\phi\psi} := \{v \in V_0 \mid \phi \leq v \leq \psi \text{ a.e. in}\Omega\}.$$

Exact solution of the problem

$$\int_\Omega A\nabla u \cdot \nabla(w-u)dx \geq \int_\Omega f(w-u)dx \quad \forall w \in K_{\phi\psi}$$

generates three sets:

$$\begin{aligned}
\Omega^u_{\oplus} &:= \{x \in \Omega \mid \quad u(x) = \psi(x)\} \text{ upper coincidence set,} \\
\Omega^u_{\ominus} &:= \{x \in \Omega \mid \quad u(x) = \phi(x)\} \text{ lower coincidence set} \\
\Omega^u_0 &:= \{x \in \Omega \mid \phi(x) < u(x) < \psi(x)\}\ .
\end{aligned}$$

Open set Ω^u_0 is the complementary set, where a solution satisfies the differential equation.

Theorem

For any $v \in K$ and $y \in H(\Omega, \mathrm{div}\,)$,

$$\|\nabla(u-v)\|_A \leq \|A\nabla v - y\|_{A^{-1}} + C_\Omega \| \mid f + \mathrm{div} y \mid^{\pm}_{v} \| := M_{obs}(v, y).$$

$$\mid f + \mathrm{div} y \mid^{\pm}_{v} := \begin{cases} (f + \mathrm{div} y)_{\ominus} & \text{at a.e. points of } \Omega^v_{\oplus}, \\ f + \mathrm{div} y & \text{at a.e. points of } \Omega^v_0, \\ (f + \mathrm{div} y)_{\oplus} & \text{at a.e. points of } \Omega^v_{\ominus}, \end{cases}$$

$M_{obs}(v, y)$ vanishes if and only if $v = u$ and y coincides with the exact flux.

$M_{obs}(v, y)$ is based on the coincidence set generated by known v!

Thank you for attention