Regularization Theory for Convex and some Non-Convex Regularizers

Clemens Kirisits & Otmar Scherzer

Computational Science Center, University Vienna, Austria
&
Radon Institute of Computational and Applied Mathematics (RICAM), Linz, Austria

Outline

Regularization and Splines

Problem setting:

- $y = y(x)$ is a smooth function on $0 \leq x \leq 1$
- Noisy samples $\tilde{y}_i$ of $y(x_i)$ at the points of a uniform grid $\Delta = \{0 = x_0 < x_1 < \cdots < x_n = 1\}$, $h = x_{i+1} - x_i$

$$\boxed{|\tilde{y}_i - y(x_i)| \leq \delta}$$

Boundary data are known exactly: $\tilde{y}_0 = y(0)$ and $\tilde{y}_n = y(1)$

Regularization and Splines

Problem setting:

- $y = y(x)$ is a smooth function on $0 \leq x \leq 1$
- Noisy samples $\tilde{y}_i$ of $y(x_i)$ at the points of a uniform grid $\Delta = \{0 = x_0 < x_1 < \cdots < x_n = 1\}$, $h = x_{i+1} - x_i$

$$\boxed{|\tilde{y}_i - y(x_i)| \leq \delta}$$

Boundary data are known exactly: $\tilde{y}_0 = y(0)$ and $\tilde{y}_n = y(1)$

Goal: Find a smooth approximation, $u'(x)$, of $y'(x)$

M. Hanke and O. Scherzer
Inverse problems light: numerical differentiation
Amer. Math. Monthly 108.6. 2001

M. Hanke and O. Scherzer
Error analysis of an equation error method for the identification of the diffusion coefficient in a quasi-linear parabolic differential equation
SIAM J. Appl. Math. 59.3. 1999

Strategy I

Minimize

$$\boxed{\left\| u'' \right\| = \left\| u'' \right\|_{L^2}}$$

among all smooth functions u satisfying $u(0) = y(0)$, $u(1) = y(1)$, and

$$\boxed{\frac{1}{n-1} \sum_{i=1}^{n-1} \left(\tilde{y}_i - u(x_i) \right)^2 \leq \delta^2}$$

Take the derivative, u_*', of the minimizing element u_* as an approximation of y'

Strategy II

Minimize

$$\boxed{\mathcal{T}_\alpha(u, \tilde{y}) \equiv \frac{1}{n-1}\sum_{i=1}^{n-1}\left(\tilde{y}_i - u(x_i)\right)^2 + \alpha \left\| u'' \right\|^2}$$

among all smooth functions u satisfying $u(0) = y(0)$, $u(1) = y(1)$, where α is such that the minimizing element u_α satisfies

$$\boxed{\frac{1}{n-1}\sum_{i=1}^{n-1}\left(\tilde{y}_i - u_\alpha(x_i)\right)^2 = \delta^2}$$

The derivative u_α' is the approximation of y'

Strategy I and II are equivalent

Error estimates for strategy II: exact data y

Let $y'' \in L^2(0,1)$. Then

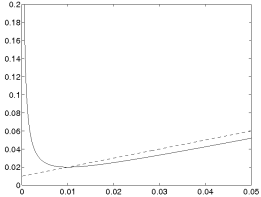
$$\boxed{\|u_*' - y'\| \leq \sqrt{8}\Big(h\,\|y''\| + \sqrt{\delta}\,\|y''\|^{1/2}\Big)}$$

Numerical differentiation versus Tikhonov regularization

Let $y \in C^2[0,1]$ (smoothness (source) condition), then

$$\left| \frac{\tilde{y}_{i+1} - \tilde{y}_i}{h} - y'(x) \right| \leq \mathcal{O}(h + \delta/h), \qquad x_i \leq x \leq x_{i+1}$$

The right hand side attains a minimal value of $\mathcal{O}(\sqrt{\delta})$ for $h \sim \sqrt{\delta}$



Qualitative behavior of the two error bounds, $h + \delta/h$ (numerical differentiation) and $h + \sqrt{\delta}$ (Tikhonov regularization) versus h for fixed δ

Comments

u^* solving Strategy II is a natural cubic spline, i.e.,

- a function that is twice continuously differentiable on $[0,1]$ with
- $u_*''(0) = u_*''(1) = 0$, and coincides on each subinterval $[x_{i-1}, x_i]$ of Δ with some cubic polynomial

u_* is uniquely determined by connecting the jumps of u_*''' at the interior nodes $x = x_i$ with the values $u_*(x_i)$ through

$$u_*'''(x_i+) - u_*'''(x_i-) = \frac{1}{\alpha(n-1)} \left(\tilde{y}_i - u_*(x_i) \right), \qquad i = 1, \dots, n-1$$

Tikhonov regularization $\Leftrightarrow$ Natural spline approximation

Ill-Posed Problems

Operator equation

$$\boxed{Lu = y}$$

Setting:

- Available data y^δ of y are noisy
- Ill-posed:

$$\boxed{y^\delta \to y \not\Rightarrow u^\delta \to u^\dagger}$$

- L is an operator between infinite dimensional spaces (before discretization)

Exampels:

- L...Radon transform, electrical impedance tomography,...
- Numerical differentiation : $L = T_\Delta \approx I$: Trace on the sampling points

General Variational Methods: Setting

- H_1 and H_2 are Hilbert spaces
- $L : H_1 \to H_2$ linear and bounded
- $L_h : H_1 \to H_2$ linear and bounded ($\approx L$)
- $\rho : H_2 \times H_2 \to \mathbb{R}_+$ similarity functional
- $\Psi : H_1 \to \mathbb{R}_+$ an energy functional
- δ: estimate for the amount of noise

Three Kind of Variational Methods

1. *Residual method ($\tau \geq 1$):*

$$\boxed{u_\alpha^\delta = \text{argmin}\Psi(u) \to \min \qquad \text{subject to } \rho(Lu, y^\delta) \leq \tau\delta}$$

2. *Tikhonov regularization with discrepancy principle ($\tau \geq 1$):*

$$\boxed{u_\alpha^\delta := \text{argmin}\left\{\rho^2(Lu, y^\delta) + \alpha\Psi(u)\right\} ,}$$

where $\alpha > 0$ is chosen according to *Morozov's discrepancy principle*, i.e., the minimizer u_α^δ of the Tikhonov functional satisfies

$$\rho(Lu_\alpha^\delta, y^\delta) = \tau\delta$$

3. *Tikhonov regularization with a–priori parameter choice:* $\alpha = \alpha(\delta)$

Relation between Methods

E.g. Ψ convex and $\rho^2(a, b) = \|a - b\|^2$

Residual Method $\equiv$ Tikhonov with discrepancy principle

Note, this was exactly the situation in the spline example!

Analysis of variational regularization

L might have a null-space.
The Ψ-Minimal Solution is denoted by $u^\dagger$ and satisfies:

$$\boxed{\Psi(u^\dagger) = \inf\{\Psi(u) : Lu = y\}}$$

Unique: for instance if Ψ is strictly convex

Regularization Method

A method is called a regularization method if the following holds:

- Stability for fixed α: $y^\delta \to y \Rightarrow u_\alpha^\delta \to u_\alpha$
- Convergence: There exists a parameter choice $\alpha = \alpha(\delta) > 0$ such that $y^\delta \to y \Rightarrow u_{\alpha(\delta)}^\delta \to u^\dagger$

Regularization Method

A method is called a regularization method if the following holds:

- Stability for fixed α: $y^\delta \to y \Rightarrow u_\alpha^\delta \to u_\alpha$
- Convergence: There exists a parameter choice $\alpha = \alpha(\delta) > 0$ such that $y^\delta \to y \Rightarrow u_{\alpha(\delta)}^\delta \to u^\dagger$

It is an efficient regularization method if there exists a parameter choice $\alpha = \alpha(\delta)$ such that

$$\boxed{D(u_{\alpha(\delta)}^\delta, u^\dagger) \le f(\delta)\,,}$$

where

- D is an appropriate distance measure
- f rate ($f \to 0$ for $\delta \to 0$)

Quadratic regularization in Hilbert spaces

$$\boxed{u_\alpha^\delta = \operatorname{argmin}\left\{\|Lu - y^\delta\|^2 + \alpha\|u - u_0\|^2\right\}}$$

Results:

- *Stability* ($\alpha > 0$)*:* $y^\delta \to y \Rightarrow u_\alpha^\delta \to u_\alpha$
- *Convergence:* Choose

$$\alpha = \alpha(\delta) \text{ such that } \delta^2/\alpha \to 0$$

If $\delta \to 0$, then $u_\alpha^\delta \to u^\dagger$, which solves $Lu^\dagger = y$

Note $u^\dagger$ is the $\Psi(\cdot) = \|u - u_0\|^2$ minimal solution

Convergence rates

Assumptions:

- *Source Condition:* $\boxed{u^\dagger - u_0 \in L^*\eta}$
- $\alpha = \alpha(\delta) \sim \delta$

Result:

$$\boxed{\left\| u_\alpha^\delta - u^\dagger \right\|^2 = \mathcal{O}(\delta) \text{ and } \left\| L u_\alpha^\delta - y \right\| = \mathcal{O}(\delta)}$$

Here L^* is the adjoint of L, i.e.,

$$\langle Lu, y \rangle = \langle u, L^* y \rangle$$

C. W. Groetsch.
The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind.
Pitman, Boston, 1984.

Spectral Theory

- L^*L is a bounded, positive definitive, self-adjoint operator
- $L^*Lu = \int_0^\infty \lambda de(\lambda)u$, where $e(\lambda)$ denotes the spectral measure of L^*L
- If L is compact, then

$$\boxed{L^*Lu = \sum_{n=0}^{\infty} \lambda_n^2 \langle u, u_n \rangle u_n \,,}$$

where (λ_n, u_n, v_n) are the spectral values of L (SVD)

Classical Convergence Rates

- *Source Condition:* $\boxed{u^\dagger - u_0 \in (L^*L)^\nu \eta, \nu \in (0,1]}$
- $\alpha = \alpha(\delta) \sim \delta^{\frac{2}{2\nu+1}}$

Result:

$$\boxed{\left\| u_\alpha^\delta - u^\dagger \right\| = \mathcal{O}(\delta^{\frac{2\nu}{2\nu+1}}) \text{ and } \left\| Lu_\alpha^\delta - y \right\| = \mathcal{O}(\delta)}$$

Note, that when $\nu = 1/2$, then

$$\mathcal{R}((L^*L)^{1/2}) = \mathcal{R}(L^*)$$

C. W. Groetsch.
The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind.
Pitman, Boston, 1984.

Numerical differentiation

- $L^* = T_h^*$: Adjoint of a restriction operator: Prolongation operator
- $L \approx I : H^2 \to L^2$ (spaces do not exactly match) $\to L^* : L^2 \to H^2$. Source condition requires $u^\dagger - u_0 (= 0) \in H^2$ (in contrast to C^2)

Convex regularization

$$\boxed{\frac{1}{2}\left\|Lu - y^\delta\right\|^2 + \alpha R(u) \to \min}$$

Examples:

- Total Variation regularization: $R(u) = \int |\nabla u|$ the total variation semi-norm
- ℓ^p regularization: $R(u) = \sum_i w_i \left|\langle u, \phi_i \rangle\right|^p, \quad 1 \le p \le 2$

  ϕ_i is an orthonormal basis of a Hilbert space with inner product $\langle \cdot, \cdot \rangle$, w_i are appropriate weights - we take $w_i \equiv 1$

Non-Quadratic Regularization

Assumptions:

- L is a bounded operator between Hilbert spaces H_1 and H_2 with closed and convex domain $\mathcal{D}(F)$
- R is weakly lower semi-continuous

Results:

- *Stability:* $y^\delta \to y \Rightarrow u_\alpha^\delta \rightharpoonup u_\alpha$ and $R(u_\alpha^\delta) \to R(u_\alpha)$
- *Convergence:* $y^\delta \to y$ and $\alpha = \alpha(\delta)$ such that $\delta^2/\alpha \to 0$, then

$$\boxed{u_\alpha^\delta \rightharpoonup u^\dagger \text{ and } R(u_\alpha^\delta) \to R(u^\dagger)}$$

Note, for quadratic regularization in H-spaces weak convergence and convergence of the norm gives strong convergence

Convergence Rates, R convex

Assumptions:

- *Source Condition:* There exists η such that $\boxed{\xi = F^*\eta \in \partial R(u^\dagger)}$
- $\alpha \sim \delta$

Result: $\boxed{D_\xi(u_\alpha^\delta, u^\dagger) = \mathcal{O}(\delta) \text{ and } \left\| L u_\alpha^\delta - y \right\| = \mathcal{O}(\delta)}$

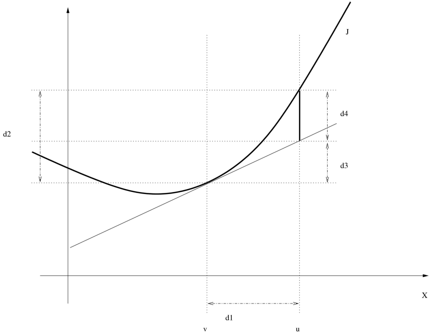
Comments:

1. $\partial R(v)$ is the subgradient of R at v, i.e., all elements ψ that satisfy $D_\psi(u,v) := R(u) - R(v) - \langle \psi, u - v \rangle \geq 0$ for all u
2. If $R(u) = \frac{1}{2}\|u - u_0\|^2 \Rightarrow \partial R(u^\dagger)u = u - u^\dagger)$
3. $D_\xi(u_\alpha^\delta, u^\dagger)$ is the Bregman distance

M. Burger and S. Osher
Convergence rates of convex variational regularization
Inverse Problems 20.5. 2004

B. Hofmann, B. Kaltenbacher, C. Pöschl, and O. Scherzer
A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators
Inverse Probl. 23.3. 2007

Bregman Distance



1. In general not a distance measure: It may be *non*-symmetric and may vanish for different elements
2. If $R(\cdot) = \frac{1}{2}\|u - u_0\|^2$, then $D_\xi(u, v) = \frac{1}{2}\|u - v\|^2$. Thus generalizes the H-space results

Compressed Sensing

Let ϕ_i be an orthonormal basis of a Hilbert space H_1. $L : H_1 \to H_2$

Constrained optimization problem:

$$\boxed{R(u) = \sum_i |\langle u, \phi_i \rangle| \to \min \qquad \text{such that } Lu = y}$$

Goal: Recover *sparse solutions*: $\operatorname{supp}(u) := \{i : \langle u, \phi_i \rangle \neq 0\}$ is finite

Compressed Sensing

Let ϕ_i be an orthonormal basis of a Hilbert space H_1. $L : H_1 \to H_2$

Constrained optimization problem:

$$\boxed{R(u) = \sum_i |\langle u, \phi_i \rangle| \to \min \quad \text{such that } Lu = y}$$

Goal: Recover *sparse solutions*: $\operatorname{supp}(u) := \{i : \langle u, \phi_i \rangle \neq 0\}$ is finite

Comments:

1. Infinite dimensional setting
2. For noisy data: Residual method

$$\boxed{R(u) \to \min \quad \text{subject to } \left\| Lu - y^\delta \right\| \le \tau\delta}$$

E. J. Candès, J. K. Romberg, and T. Tao
Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information
IEEE Transactions on Information Theory 52.2. 2006

Sparsity Regularization

Unconstrained Optimization

$$\boxed{\left\| Lu - y^\delta \right\|^2 + \alpha R(u) \to \min}$$

General theory for sparsity regularization:

- *Stability:* $y^\delta \to y \Rightarrow u_\alpha^\delta \rightharpoonup u_\alpha$ and $\left\| u_\alpha^\delta \right\|_{\ell^1} \to \left\| u_\alpha \right\|_{\ell^1}$
- *Convergence:* $y^\delta \to y \Rightarrow u_\alpha^\delta \rightharpoonup u_\alpha^\delta$ and $\left\| u_\alpha^\delta \right\|_{\ell^1} \to \left\| u^\dagger \right\|_{\ell^1}$ if $\delta^2/\alpha \to 0$.

If α is chosen according to the discrepancy principle, then Sparsity Regularization $\equiv$ Compressed Sensing

Convergence Rates: Sparsity Regularization

Assumptions:

- *Source Condition:* There exists η such that

$$\boxed{\xi = L^* \eta \in \partial R(u^\dagger)}.$$

  Formally this means that $\xi_i = \mathrm{sgn}(u_i^\dagger)$ and $u^\dagger$ is sparse (means in the domain of ∂R)

- $\alpha \sim \delta$

Result:

$$\boxed{D_\xi(u_\alpha^\delta, u^\dagger) = \mathcal{O}(\delta) \text{ and } \left\| L u_\alpha^\delta - y \right\| = \mathcal{O}(\delta)}$$

Comment: Rate is *optimal* for a choice $\alpha \sim \delta$

Convergence Rates: Compressed Sensing

Assumption: Source condition

$$\boxed{\xi = L^* \eta \in \partial R(u^\dagger)}$$

Then

$$\boxed{D_\xi(u^\delta, u^\dagger) \leq 2 \|\eta\| \, \delta}$$

for every

$$u^\delta \in \operatorname{argmin} \left\{ R(u) : \left\| Lu - y^\delta \right\| \leq \delta \right\}$$

Remark: Candes et al have rate δ with respect to the *finite dimensional Euclidean norm* and not w.r.t. the Bregman distance

M. Grasmair, M. Haltmeier, and O. Scherzer
Necessary and sufficient conditions for linear convergence of l^1-regularization
Comm. Pure Appl. Math. 64.2. 2011

$0 < p < 1$: Nonconvex sparsity regularization

$$\boxed{\left\| Lu - y^\delta \right\|^2 + \sum |\langle u, \phi_i \rangle|^p \to \min}$$

is stable, convergent, and well–posed in the Hilbert-space norm

- Zarzer: $\mathcal{O}(\sqrt{\delta})$
- Grasmair + IP $\Rightarrow \mathcal{O}(\delta)$

C. A. Zarzer
On Tikhonov regularization with non-convex sparsity constraints
Inverse Problems 25. 2009

M. Grasmair
Non-convex sparse regularisation
J. Math. Anal. Appl. 365.1. 2010

Image registration

- Given: Image $I_1, I_2 : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$
- Find $u : \Omega \to \mathbb{R}^2$ satisfying

$$\boxed{L(u) := I_2 \circ u = I_1}$$

u should be a diffeomorphism (no twists)

Calculus of variations: Notions of convexity

$$
\begin{aligned}
f : \mathbb{R}^N \times \mathbb{R}^n \times \mathbb{R}^{N\times n} &\to \mathbb{R}\,, \\
(x,u,v) &\to f(x,u,v)
\end{aligned}
$$

f **convex** $\Rightarrow$ $\boxed{\textit{polyconvex}}$ $\Rightarrow$ **quasi-convex** $\Rightarrow$ **rank-one convex**

Up to quasi-convex:

$$u \to \int_{\mathbb{R}^n} f(x,u,\nabla u)\,dx \text{ is weakly lower semicontinuous on } W^{1,p}$$

If $N = 1$ or $n = 1$, then all convexity definitions are equivalent

Polyconvex functions

Let $N, n \in \mathbb{N}$ and $N \wedge n = \min(N, n)$.
For $A \in \mathbb{R}^{N \times n}$ and $1 \leq s \leq N \wedge n$

$$\mathrm{adj}_s(A) \text{ consists of all } s \times s \text{ minors of } A$$

Properties:

$$\mathrm{adj}_1(A) = A\,, \quad \mathrm{adj}_s(A) \in \mathbb{R}^{\sigma(s)}, \sigma(s) = \binom{N}{s}\binom{n}{s}\,, \quad \tau(N,n) = \sum_{s=1}^{N \wedge n} \sigma(s)$$

$f : \mathbb{R}^{N \times n} \to \mathbb{R} \cup \{+\infty\}$ is polyconvex if $f = F \circ T$, where
$F : \mathbb{R}^{\tau(N,n)} \to \mathbb{R} \cup \{+\infty\}$ is convex and

$$T : \mathbb{R}^{N \times n} \to \mathbb{R}^{\tau(N,n)}\,, \quad A \to (A, \mathrm{adj}_2(A), \ldots, \mathrm{adj}_m(A))$$

Reduced map of T considered: $T_2(A) = (\mathrm{adj}_2(A), \ldots, \mathrm{adj}_m(A))$

Generalized Bregman distances

Let W be a family of real-valued functions defined on U

- The *W-subdifferential* of a functional $\mathcal{R}$ is defined by

$$\boxed{\partial_W \mathcal{R}(u) = \{w \in W : \mathcal{R}(v) \geq \mathcal{R}(u) + w(v) - w(u)\,, \forall v \in U\}}$$

- For $w \in \partial_W \mathcal{R}(u)$ the *W-Bregman distance* is defined by

$$\boxed{D_w^W(v;u) = \mathcal{R}(v) - \mathcal{R}(u) - w(v) + w(u)}$$

M. Grasmair
Generalized Bregman distances and convergence rates for non-convex regularization methods
Inverse Probl. 26.11. Oct. 2010

I. Singer
Abstract convex analysis
John Wiley & Sons Inc., 1997

Bregman distances of polyconvex integrands

Let $p \in [1, \infty)$ and $U = W^{1,p}(\Omega, \mathbb{R}^N)$.

$$T_2(\nabla u) \in \prod_{s=2}^{N \wedge n} L^{\frac{p}{s}}(\Omega, \mathbb{R}^{\sigma(s)}) =: S_2.$$

We define

$$\begin{aligned} W_{\text{poly}} := \{ w : U \to \mathbb{R} : \exists (u^*, v^*) \in U^* \times S_2^* \text{ s.t.} \\ w(u) = \langle u^*, u \rangle_{U^*, U} + \langle v^*, T_2(\nabla u) \rangle_{S_2^*, S_2} \} \end{aligned}$$

Remark:

- $W_{\text{poly}} = (U \times S_2)^*$. However, functionals are non-linear
- W_{poly}-Bregman distance:

$$\begin{aligned} D_w^{\text{poly}}(u; \bar{u}) &= \mathcal{R}(u) - \mathcal{R}(\bar{u}) - w(u) + w(\bar{u}) \\ &= \mathcal{R}(u) - \mathcal{R}(\bar{u}) - \langle u^*, u - \bar{u} \rangle_{U^*, U} \\ &\quad - \langle v^*, T_2(\nabla u) + T_2(\nabla \bar{u}) \rangle_{S_2^*, S_2} \end{aligned}$$

Polyconvex subgradient

- $\Omega \subset \mathbb{R}^n$ and $U = W^{1,p}(\Omega, \mathbb{R}^N)$
- $F : \Omega \times \mathbb{R}^N \times \mathbb{R}^{\tau(N,n)} \to \mathbb{R} \cup \{+\infty\}$ is Carathéodory
- For $x \in \Omega$, the map $(u, A) \mapsto F(x, u, A)$ is convex and differentiable
- $\mathcal{R}(u) = \int_\Omega F(x, u(x), T(\nabla u(x)))\, dx$

If $\mathcal{R}(\bar{v}) \in \mathbb{R}$ and the function $x \mapsto \nabla_{u,A} F(x, \bar{v}(x), T(\nabla \bar{v}(x)))$ lies in

$$L^{p^*}(\Omega, \mathbb{R}^N) \times \prod_{s=1}^{N \wedge n} L^{\frac{p}{s}}(\Omega, \mathbb{R}^{\sigma(s)}),$$

then this function is a W_{poly}-subgradient of $\mathcal{R}$ at $\bar{v}$

Example

Let $N = n = 2$, $T(a) = (a, \det a)$ for $a \in \mathbb{R}^{2\times 2}$.

$$F(x, u, a, \det a) = F(\det a) = (\det a)^2$$

Let $p = 4$, then

$$\mathcal{R}(u) = \int_\Omega (\det \nabla u(x))^2 \, dx \in \mathbb{R}$$

and $x \mapsto F'(\det \nabla u(x)) = 2 \det \nabla u(x) \in L^2(\Omega)$.
In particular

$\mathcal{R}$ is W_{poly}-subdifferentiable

Example

Let $p > N = n \geq 2$, $q > 1$, and

$$F(x, u, T(a)) = F(a, \det a) = |a|^p/p + |\det a|^q/q$$

If $\bar{v} \in W^{1,\infty}(\Omega, \mathbb{R}^n)$, then $\mathcal{R}(\bar{v}) = \int_\Omega F(x, \bar{v}, T(\bar{v}))\, dx < +\infty$, then

$$\begin{aligned} x \mapsto & \nabla_A F(\nabla \bar{v}(x), \det \nabla \bar{v}(x)) \\ & = (|\nabla \bar{v}(x)|^{p-2} \nabla \bar{v}(x), |\det \nabla \bar{v}(x)|^{q-2} \det \nabla \bar{v}(x)) \in L^\infty \end{aligned}$$

Thus $\mathcal{R}(\bar{v})$ is lsc and has a W_{poly}-subgradient on $W^{1,\infty}(\Omega, \mathbb{R}^n) \subset U$.

Rates result

Let $U = W^{1,p}(\Omega, \mathbb{R}^N)$.

- $\mathcal{R}$ has a W_{poly}-subgradient w at $u^\dagger$
- $\exists \beta_1 \in [0,1), \beta_2$ such that locally

$$w(u^\dagger) - w(u) \leq \beta_1 D_w^{\mathrm{poly}}(u; u^\dagger) + \beta_2 \|L(u) - v^\dagger\|$$

Then,

1. If $p > 1$ and $\alpha(\delta) \sim \delta^{p-1}$:

$$D_w^{\mathrm{poly}}(u_\alpha^\delta; u^\dagger) = \mathcal{O}(\delta) \quad \text{and} \quad \|L(u_\alpha^\delta) - v^\delta\| = \mathcal{O}(\delta).$$

2. If $p = 1$ and $\alpha(\delta) \sim \delta^\epsilon$ for $\epsilon \in [0,1)$. Then

$$D_w^{\mathrm{poly}}(u_\alpha^\delta; u^\dagger) = \mathcal{O}(\delta^{1-\epsilon}) \quad \text{and} \quad \|L(u_\alpha^\delta) - v^\delta\| = \mathcal{O}(\delta)$$

Applications

Let $p > n$, $q \geq 1$, $U = W^{1,p}(\Omega, \mathbb{R}^n)$, with its weak and $V = L^q(\Omega)$ with its strong topology.
Assume that

$$F(x, u, T(a)) \geq |a|^p \text{ and } \mathcal{R}(u) = \int_\Omega F(x, u(x), T(\nabla u(x)))\, dx$$

Then, minimization of

$$\|L(u) - I_1^\delta\|_{L^q(\Omega)}^q + \alpha \mathcal{R}(u), \quad \alpha > 0,$$

is well-defined and source consdition can be stated
See Kirisits and Scherzer (2017)

Thank you for your attention