Regularization of Inverse Problems

Otmar Scherzer

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

Inverse Problems

desire to calculate or estimate

causal factors

from a set of observations

Inverse Problems are often Ill-Posed

Operator equation:

$$\boxed{Lu = y}$$

Setting:

- Available data y^δ of y are noisy
- Focus: L is a linear operator
- Ill-posed: Let $u^\dagger$ be a solution:

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

A. N. Tikhonov
On the stability of inverse problems
Doklady Akademii Nauk SSSR 39. 1943

Examples of Ill–Posed Problems: $L=$

1. Identity operator: Measurements of noisy data. Applications: A.e.
2. X-Ray transform: Measurements of averages over lines. Application: Computerized Tomography (CT)
3. Radon transform: Measurements of averages over hyperplanes. Application: Cryo-EM
4. Spherical Radon transform: Measurements of averages over spheres. Application: Photoacoustic Imaging
5. Circular Radon transform: Measurements of averages over circles. Applications: Ground Penetrating Radar and Photoacoustics

An Application

GPR: Location of avalanche victims

Project with Wintertechnik AG and Alps

Various Philosophies

- Continuous approach: $L : H_1 \to H_2$. H_i infinite dimensional spaces
- Semi-continuous approach: $L : H \to \mathbb{R}^n$. H infinite dimensional space, finitely many measurements
- Discrete Setting: $L : \mathbb{R}^m \to \mathbb{R}^n$. Large scale inverse problems
- Bayesian approach: $L : \mathbb{R}^m \to \mathbb{R}^n$. Stochastic inverse problems

A. N. Tikhonov
Solution of incorrectly formulated problems and the regularization methods
Soviet Mathematics. Doklady 4. 1963

H. Engl, M. Hanke, and A. Neubauer
Regularization of inverse problems
Kluwer Academic Publishers Group, 1996

M. Unser, J. Fageot, and J. P. Ward
Splines are universal solutions of linear inverse problems with generalized TV regularization
SIAM Review 59.4. 2017

C. R. Vogel
Computational Methods for Inverse Problems
SIAM, 2002

M. Hanke and P. C. Hansen
Regularization methods for large-scale problems
Surveys on Mathematics for Industry 3.4. 1994

J. Kaipio and E. Somersalo
Statistical and Computational Inverse Problems
Springer Verlag, 2005

Deterministic Setting

From Continuous approach: $L : H_1 \to H_2$. H_i infinite dimensional spaces $\Rightarrow$

- Semi-Continuous approach: $P \circ L$ with $P[f] = (f(x_i))_{i=1,\dots,n}$
- Discrete approach: $P \circ L \circ Q$ with $Q[(c_i)_{i=1,\dots,m}](x) = \sum_{i=1}^{m} c_i \phi(x)$. (ϕ_i) family of test functions.

A. Neubauer and O. Scherzer
Finite-dimensional approximation of Tikhonov regularized solutions of nonlinear ill-posed problems
Numer. Funct. Anal. Optim. 11.1-2. 1990

C. Pöschl, E. Resmerita, and O. Scherzer
Discretization of variational regularization in Banach spaces
Inverse Probl. 26.10. 2010

Various Methods to Solve

- Backprojection formulas
- Iterative Methods for linear and nonlinear inverse problems
- Flow methods: Showalter's methods and Inverse scale space methods
- Variational methods: Tikhonov type regularization.
- ...

J. Radon
Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten
Berichte über die Verhandlungen der Königlich-Sächsischen Gesellschaft der Wissenschaften zu Leipzig, Mathematisch-Physische Klasse 69. 1917

M. Hanke and P. C. Hansen
Regularization methods for large-scale problems
Surveys on Mathematics for Industry 3.4. 1994

B. Kaltenbacher, A. Neubauer, and O. Scherzer
Iterative regularization methods for nonlinear ill-posed problems
Walter de Gruyter, 2008

O. Scherzer and C. W. Groetsch (2001). "Inverse Scale Space Theory for Inverse Problems". In: *Scale-Space and Morphology in Computer Vision*. Ed. by M. Kerckhove. Vol. 2106. Lecture Notes in Computer Science. Vancouver, Canada: Springer, pp. 317–325. ISBN: 978-3-540-42317-1. DOI: 10.1007/3-540-47778-0_29. URL: http://dx.doi.org/10.1007/3-540-47778-0_29

H. Engl, M. Hanke, and A. Neubauer
Regularization of inverse problems
Kluwer Academic Publishers Group, 1996

O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen
Variational methods in imaging
Springer, 2009

Outline

Numerical Differentiation as an Inverse Problem

- $y = y(x)$ is a smooth function on $0 \leq x \leq 1$
- Given: Noisy samples y_i^δ of $y(x_i)$ on a uniform grid

$$\Delta = \{0 = x_0 < x_1 < \cdots < x_n = 1\}, h = x_{i+1} - x_i$$

satisfying

$$\boxed{\left| y_i^\delta - y(x_i) \right| \leq \delta}$$

Boundary data are known exactly: $y_0^\delta = y(0)$ and $y_n^\delta = y(1)$

- Goal: Find a smooth approximation u' of y'

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: Constrained Minimization

Approach: Continuous to discrete

1. $\|u''\|_{L^2}^2 = \int_0^1 (u'')^2 \, dx \to \min$ among smooth functions u satisfying
   - $u(0) = y(0)$, $u(1) = y(1)$,
   - Constraint:

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

2. Minimizer u_*: $u_*' \approx y'$

Strategy II: Tikhonov Regularization

1. Let $\alpha > 0$. Minimization among smooth functions u satisfying $u(0) = y(0)$, $u(1) = y(1)$, of

$$u_\alpha^\delta = \text{argmin}\Phi[u], \quad \Phi[u] = \frac{1}{n-1} \sum_{i=1}^{n-1} \left(y_i^\delta - u(x_i)\right)^2 + \alpha \left\| u'' \right\|_{L^2}^2$$

2. ${u_\alpha^\delta}' \approx y'$

Strategy II: Tikhonov Regularization + Discrepancy Principle

Theorem

If α is selected according to the discrepancy principle

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

Then Strategy I and II are equivalent: $u_\alpha^\delta = u_*$

V. A. Morozov
Methods for Solving Incorrectly Posed Problems
Springer, 1984

Analysis of Constrained Optimization

Let

1. $y'' \in L^2(0,1)$ (assumption on the data to be reconstructed) and
2. u_* be the minimizer of Strategy II

Then

$$\left\|u_*' - y'\right\|_{L^2} \leq \sqrt{8}\Big(\underbrace{h \left\|y''\right\|_{L^2}}_{\text{approx. error}} + \underbrace{\sqrt{\delta \left\|y''\right\|_{L^2}}}_{\text{noise influence}} \Big)$$

I. J. Schoenberg
Spline interpolation and the higher derivatives
Proceedings of the National Academy of Sciences of the USA 51.1. 1964

M. Unser
Splines: a perfect fit for signal and image processing
IEEE Signal Processing Magazine 16.6. 1999

Textbook Example: Numerical Differentiation

Let $y \in C^2[0,1]$, then

$$\left|\frac{y^\delta_{i+1} - y^\delta_i}{h} - y'(x)\right| \leq \mathcal{O}(\underbrace{h}_{\text{approx. error}} + \underbrace{\delta/h}_{\text{noise error}}), \quad x_i \leq x \leq x_{i+1}$$

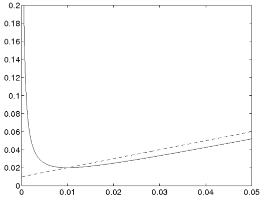

Figure: $h \to h + \delta/h$ (numerical differentiation) and $h \to h + \sqrt{\delta}$ (Tikhonov regularization) for fixed δ

$y'' \in L^2$: $h \to h + \delta/h$ is minimal for $h \sim \sqrt{\delta}$. Optimal rates for Strategy I, II and numerical differentiation is $\mathcal{O}(\sqrt{\delta})$

The rate $\mathcal{O}(\sqrt{\delta})$ does not hold if $y'' \notin L^2(0,1)$

Properties of u^*

Theorem

- *a solution u^* of Strategy I exists*
- *u^* is a natural cubic spline, i.e.,*
  - *a function that is twice continuously differentiable over $[0, 1]$ with*
  - *$u_*''(0) = u_*''(1) = 0$, and coincides on each subinterval $[x_{i-1}, x_i]$ of Δ with some cubic polynomial*

Generalizations of the ideas to non-quadratic regularization and general inverse problems in Adcock and A. C. Hansen 2015; Unser, Fageot, and Ward 2017

B. Adcock and A. C. Hansen
Generalized sampling and the stable and accurate reconstruction of piecewise analytic functions from their Fourier coefficients
Math. Comp. 84.291. 2015

M. Unser, J. Fageot, and J. P. Ward
Splines are universal solutions of linear inverse problems with generalized TV regularization
SIAM Review 59.4. 2017

General Variational Methods: Setting

General:

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

Numerical differentiation:

- $H_1 = W_0^2(0,1) = \{w : w, w' \in L^2(0,1)\}$ and $H_2 = \mathbb{R}^{n-1}$
- $L : W_0^2(0,1) \to \mathbb{R}^{n-1}$, $u \mapsto (u(x_i))_{1 \leq i \leq n-1}$
- $\rho(\xi, \nu) = \frac{1}{n-1} \sum_{i=1}^{n-1} (\xi_i - \nu_i)^2$
- $\mathcal{R}[u] = \int_0^1 (u'')^2 \, dx$
- $\frac{1}{n-1} \sum_{i=1}^{n-1} (y_i - y_i^\delta)^2 \leq \delta^2$.

Three Kind of Variational Methods ($\tau \geq 1$)

1. Residual method:

$$u_\alpha^\delta = \text{argmin}\mathcal{R}(u) \to \min \quad \text{subject to } \rho(Lu, y^\delta) \leq \tau\delta$$

Three Kind of Variational Methods ($\tau \geq 1$)

1. Residual method:

$$u_\alpha^\delta = \text{argmin}\mathcal{R}(u) \to \min \quad \text{subject to } \rho(Lu, y^\delta) \leq \tau\delta$$

2. Tikhonov regularization with discrepancy principle:

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

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

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

Three Kind of Variational Methods ($\tau \geq 1$)

1. Residual method:

$$u_\alpha^\delta = \text{argmin} \mathcal{R}(u) \to \min \quad \text{subject to } \rho(Lu, y^\delta) \leq \tau\delta$$

2. Tikhonov regularization with discrepancy principle:

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

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. $\mathcal{R}$ convex and $\rho^2(a,b) = \|a-b\|^2$

$$\boxed{\text{Residual Method} \equiv \text{Tikhonov with discrepancy principle}}$$

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

V. K. Ivanov, V. V. Vasin, and V. P. Tanana
Theory of linear ill-posed problems and its applications
VSP, 2002

$\mathcal{R}$-Minimal Solution

If L has a null-space, we concentrate on a particular solution. The $\mathcal{R}$-Minimal Solution is denoted by $u^\dagger$ and satisfies:

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

Uniqueness of $\mathcal{R}$-minimal solution: For instance if $\mathcal{R}$ is strictly convex

Regularization Method

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

- Stability for fixed α: $y^\delta \to_{H_2} y \Rightarrow u_\alpha^\delta \to_{H_1} u_\alpha$
- Convergence: There exists a parameter choice $\alpha = \alpha(\delta) > 0$ such that $y^\delta \to_{H_2} y \Rightarrow u_{\alpha(\delta)}^\delta \to_{H_1} u^\dagger$

Regularization Method

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

- Stability for fixed α: $y^\delta \to_{H_2} y \Rightarrow u_\alpha^\delta \to_{H_1} u_\alpha$
- Convergence: There exists a parameter choice $\alpha = \alpha(\delta) > 0$ such that $y^\delta \to_{H_2} y \Rightarrow u_{\alpha(\delta)}^\delta \to_{H_1} u^\dagger$

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

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

where

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

Importance of Topologies

It is important to specify the topology of the convergence. Typically Sobolev or Besov spaces.

Example

Differentiation is well-posed from $W_0^1(0,1)$ into $L^2(0,1)$, but not from $L^2(0,1)$ into itself. Take

$$x \to f_n(x) := \frac{1}{n}\sin(2\pi nx)$$

Then

$$x \to f_n'(x) := 2\pi\cos(2\pi nx)$$

Note

$$\|f_n\|^2_{W_0^1(0,1)} = \frac{1}{2}\frac{1}{n^2} + \pi \to \pi \sim \left\|f_n'\right\|^2_{L^2(0,1)} = \pi \text{ but } \|f_n\|^2_{L^2(0,1)} = \frac{1}{2}\frac{1}{n^2} \to 0$$

Quadratic Regularization in Hilbert Spaces

$$u_\alpha^\delta = \operatorname{argmin} \left\{ \|Lu - y^\delta\|_{H_2}^2 + \alpha \|u - u_0\|_{H_1}^2 \right\}$$

Results:

- Stability ($\alpha > 0$): $y^\delta \to_{H_2} y \Rightarrow u_\alpha^\delta \to_{H_1} u_\alpha$
- Convergence: Choose

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

  If $\delta \to 0$, then $u_\alpha^\delta \to u^\dagger$
  Note that $u^\dagger$ is the $\mathcal{R}(\cdot) = \|\cdot - u_0\|^2$ minimal solution

Convergence Rates (The Simplest Case)

Assumptions:

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

Result:

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

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

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

- If $L \in \mathbb{R}^{m \times n}$, then $L^* = L^T \in \mathbb{R}^{n \times m}$
- If $L =$ Radon transform, the L^* is backprojection operator

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

H. Engl, M. Hanke, and A. Neubauer
Regularization of inverse problems
Kluwer Academic Publishers Group, 1996

Convergence Rates for the Spline Example

Recall $Lu = u(0.5)$ (just one sampling point) and $\Delta = \{0, 0.5, 1\}$.
Adjoint operator of $L : W_0^2(0,1) \to \mathbb{R}$, $L^* : \mathbb{R} \to W_0^2(0,1)$.

Let z be the solution of

$$z^{(IV)}(x) = \delta_{0.5}(x)$$

satisfying $z(0) = z(1) = z''(0) = z''(1) = 0$ and $z(0.5) = 1$ and C^2-smoothness, i.e. it is a fundamental solution.

Then z is a natural cubic spline! [1]

[1] Note that a cubic spline is infinitely often differentiable between sampling point and the third derivative jumps. Thus fourth derivative is a δ-distribution at the sampling points

Adjoint for the Spline Example

Let $v \in \mathbb{R}$

$$\langle Lu, v\rangle_{\mathbb{R}} = Luv = v \int_0^1 u(x)\delta_{0.5}(x)\,dx = v \int_0^1 u(x) z^{(IV)}(x)\,dx$$
$$= \int_0^1 u''(x)\left(v z''(x)\right)\,dx = \langle u, vz\rangle_{W_0^2(0,1)}$$

Thus $L^* v(x) = v z(x)$.
A convergence rate $\mathcal{O}(\sqrt{\delta})$ holds if the solution is a natural cubic spline and $u^{\dagger\prime\prime} \in L^2(0,1)$ (integration by parts)

Classical Convergence Rates - Spectral Decomposition

First, let $L \in \mathbb{R}^{n \times m}$ be a **matrix:**

$$L = \Psi^T \Lambda \Phi \text{ with } \Phi \in \mathbb{R}^{m \times m}, \Psi \in \mathbb{R}^{n \times n} \text{ orthogonal}$$

and Λ diagonal with rank $\leq \min\{m, n\}$.
Then

$$L^* L = L^T L = \Phi^T \Lambda \Psi \Psi^T \Lambda D \Phi = \Phi^T \Lambda^2 \Phi$$

which rewrites to

$$L^* L u = \sum_{n=1}^{\min\{m,n\}} \lambda_n^2 \langle u, \phi_n \rangle \phi_n = \int_0^\infty \lambda^2 \underbrace{\langle u, \phi_n \rangle \delta_{\lambda_n} dx}_{= de(\lambda) u}$$

Classical Convergence Rates (Generalized)

Spectral Theory:

- L^*L is a bounded, positive definitive, self-adjoint operator
- $L^*Lu = \int_0^\infty \lambda^2 de(\lambda)u$, where $e(\lambda)$ denotes the spectral measure of L^*L
- If L is compact, then

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

where (λ_n^2, ϕ_n) are the spectral values of L

Classical Convergence Rates

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

Result:

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

Note, that for $\nu = 1/2$

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

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

Non-Quadratic Regularization

$$\frac{1}{2}\left\|Lu - y^\delta\right\|^2 + \alpha\mathcal{R}[u] \to \min$$

Examples:

- Total Variation regularization:

$$TV[u] = \sup\left\{\int_\Omega u\nabla\cdot\phi\,dx : \phi \in C_0^\infty(\Omega;\mathbb{R}^m), \|\phi\|_{L^\infty} \le 1\right\}$$

  the total variation semi-norm.

- ℓ^p regularization: $\mathcal{R}[u] = \sum_i w_i \left|\langle u, \phi_i\rangle\right|^p, \quad 1 \le p \le 2$

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

Functional Analysis, Basics I

- Let (u_n) be a sequence in a Hilbert space H, then $u_n \rightharpoonup_H u$ iff
$$\langle u_n, \phi \rangle_H \to \langle u, \phi \rangle_H \quad \forall \phi \in H$$
- The set
$$\left\{ u : u \in L^1(\Omega) \text{ and } TV[u] < \infty \right\}$$
with the norm
$$\|u\|_{BV} := \|u\|_{L^1(\Omega)} + TV[u]$$
is a Banach space and is called Space of Functions of Bounded Variation
- A sequence in $BV \cap L^2(\Omega)$ is weak* convergent, $u_n \rightharpoonup_* u$, iff
$$\langle u_n, \phi \rangle_{L^2(\Omega)} \to \langle u, \phi \rangle_{L^2(\Omega)} quad \forall \phi \in L^2(\Omega) \text{ and } TV[u_n] \to TV[u]$$
- If $u \in C^1(\Omega)$, then $TV[u] = \int_\Omega |\nabla u| \, dx$

L. Ambrosio, N. Fusco, and D. Pallara
Functions of bounded variation and free discontinuity problems
Oxford University Press, 2000

Functional Analysis, Basics II

Let H be a Hilbert space

- $\mathcal{R} : H \to \mathbb{R} \cup \{+\infty\}$ is called proper if $\mathcal{R} \neq \infty$
- $\mathcal{R}$ is weakly lower semi-continuous if for $u_n \rightharpoonup_H u$

$$\mathcal{R}[u] \leq \liminf \mathcal{R}[u_n]$$

R. T. Rockafellar
Convex Analysis
Princeton University Press, 1970

Non-Quadratic Regularization

Assumptions:

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

Results:

- *Stability:* $y^\delta \to_{H_2} y \Rightarrow u_\alpha^\delta \rightharpoonup_{H_2} u_\alpha$ and $\mathcal{R}[u_\alpha^\delta] \to \mathcal{R}[u_\alpha]$
- *Convergence:* $y^\delta \to_{H_2} y$ and $\alpha = \alpha(\delta)$ such that $\delta^2/\alpha \to 0$, then

$$\boxed{u_\alpha^\delta \rightharpoonup_{H_2} u^\dagger \text{ and } \mathcal{R}[u_\alpha^\delta] \to \mathcal{R}[u^\dagger]}$$

Asplund property: For quadratic regularization in H-spaces weak convergence and convergence of the norm gives strong convergence

Some Convex Analysis: The Subgradient

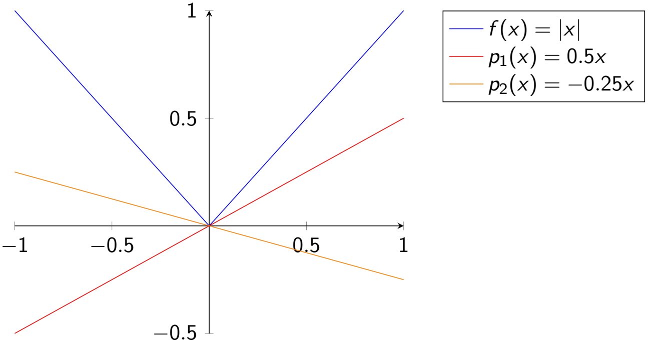

Illustration of the function $f : (-1,1) \to \mathbb{R}$, $f(x) = |x|$, and the graphs of two of its subgradients $p_1, p_2 \in \partial f(0) = \{p \in \mathbb{R}^* \mid p(x) = cx,\ c \in [-1,1]\}$

Some Convex Analysis: The Bregman Distance

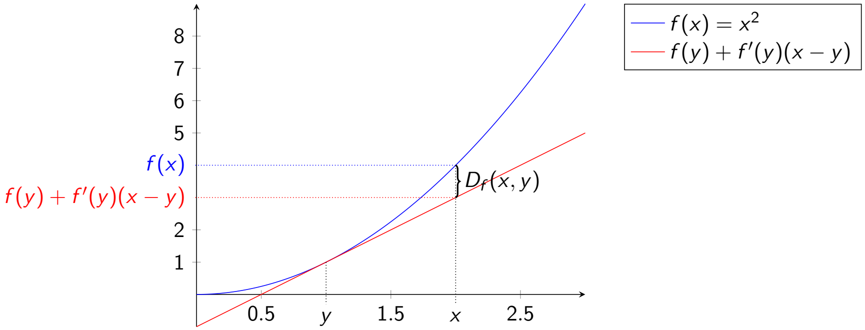

Illustration of the Bregman distance
$D_{f'(y)}(x,y) = f(x) - f(y) - f'(y)(x-y)$ for the function $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$, between the points $x = 2$ and $y = 1$

Bregman Distance

1. We consider Bregman distance for functionals
2. If $\mathcal{R}[u] = \frac{1}{2}\|u - u_0\|^2 \Rightarrow \partial\mathcal{R}[u^\dagger] = u - u^\dagger$
3. and $D_\xi(u, v) = \frac{1}{2}\|u - v\|^2$.
4. In general not a distance measure: It may be *non*-symmetric and may vanish for non-equal elements
5. Bregman distance can be a weak measure and difficult to interpret

Convergence Rates, $\mathcal{R}$ convex

Assumptions:

- *Source Condition:* There exists η such that

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

- $\alpha \sim \delta$

Result:

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

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

Compressed Sensing

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

$$\mathcal{R}[u] = \sum_i |\langle u, \phi_i \rangle| \to \min \qquad \text{such that } Lu = y$$

Goal is to recover *sparse solutions*:

$$\text{supp}(u) := \{i : \langle u, \phi_i \rangle \neq 0\} \text{ is finite}$$

Compressed Sensing

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

Constrained optimization problem:

$$\mathcal{R}[u] = \sum_i |\langle u, \phi_i \rangle| \to \min \quad \text{such that } Lu = y$$

Goal is to recover *sparse solutions*:

$$\text{supp}(u) := \{i : \langle u, \phi_i \rangle \neq 0\} \text{ is finite}$$

For noisy data: Residual method

$$\mathcal{R}[u] \to \min \quad \text{subject to } \left\| Lu - y^\delta \right\| \leq \tau\delta$$

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

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

General theory for sparsity regularization:

- *Stability:* $y^\delta \to_{H_2} y \Rightarrow u_\alpha^\delta \rightharpoonup_{H_1} u_\alpha$ and $\left\| u_\alpha^\delta \right\|_{\ell^1} \to \left\| u_\alpha \right\|_{\ell^1}$
- *Convergence:* $y^\delta \to_{H_2} y \Rightarrow u_\alpha^\delta \rightharpoonup_{H_1} u^\dagger$ 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

$$H_2 \ni \xi = L^*\eta \in \partial\mathcal{R}[u^\dagger] = \partial\left(\sum_i \left|\langle u^\dagger, \phi_i\rangle\right|\right) = \sum_i \underbrace{\operatorname{sgn}(\langle u^\dagger, \phi_i\rangle)}_{=:\xi_i}\phi_i$$

$\Rightarrow u^\dagger$ is sparse (means in the domain of $\partial\mathcal{R}$)

- $\alpha \sim \delta$

Result:

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

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

O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen
Variational methods in imaging
Springer, 2009

Analogous Convergence Rates: Compressed Sensing

Assumption: Source condition

$$\boxed{\xi = L^*\eta \in \partial\mathcal{R}[u^\dagger]}$$

Then

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

for every

$$u_* \in \text{argmin}\left\{\mathcal{R}[u] : \left\|Lu - y^\delta\right\| \leq \delta\right\}$$

Note: u_* is the constraint solution

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

M. Grasmair, M. Haltmeier, and O. Scherzer
The residual method for regularizing ill-posed problems
Appl. Math. Comput. 218.6. 2011

What can we deduce from the Bregman Distance?

Because we assume $(\phi_i)_{i\in\mathbb{N}}$ to be an orthonormal basis, the Bregman distance simplifies to

$$\begin{aligned} D_\xi(u,u^\dagger) &= \mathcal{R}[u] - \mathcal{R}[u^\dagger] - \langle \xi, u - u^\dagger \rangle \\ &= \mathcal{R}[u] - \langle \xi, u \rangle \\ &= \sum_i \Big(|\langle u, \phi_i \rangle| - \underbrace{\langle \xi, \phi_i \rangle}_{=\xi_i} \underbrace{\langle u, \phi_i \rangle}_{=u_i} \Big) \end{aligned}$$

Note, by the definition of the subgradient $|\langle \xi, \phi_i \rangle| \leq 1$

FWF

Rates with respect to the norm: On the infinite set!

Recall source condition $\xi = L^*\eta \in \partial\mathcal{R}[u^\dagger]$
Define

$$\Gamma(\eta) := \{i : |\xi_i| = 1\} \text{ (which is finite – solution is sparse)}$$

and the number (take into account that the coefficients of ζ are in ℓ^2)

$$m_\eta := \max\{|\xi_i| : i \notin \Gamma(\eta)\} < 1$$

Then

$$D_\xi(u_*, u^\dagger) = \sum_i |u_{*,i}| - \xi_i u_{*,i} \geq (1 - m_\eta) \sum_{i \notin \Gamma(\eta)} |u_{*,i}|$$

Consequently, since $\|\cdot\|_{\ell^1} \geq \|\cdot\|_{\ell^2}$, we get

$$\left\| \pi_{\mathbb{N}\setminus\Gamma(\eta)}(u_*) - \underbrace{\pi_{\mathbb{N}\setminus\Gamma(\eta)}(u^\dagger)}_{=0} \right\|_{H_1} \leq C D_\xi(u_*, u^\dagger) \leq C\delta$$

Rates with respect to the Norm: On the small Set

Additional Assumption: *Restricted injectivity:*

The mapping $L_{\Gamma(\eta)}$ is injective

Thus on $\Gamma(\eta)$ the problem is well–posed on the small set and consequently

$$\left\| \pi_{\Gamma(\eta)}(u_*) - \pi_{\Gamma(\eta)}(u^\dagger) \right\|_{H_1} \leq C\delta$$

Together with previous slide:

$$\left\| u_* - u^\dagger \right\|_{H_1} \leq C\delta$$

M. Grasmair
Linear convergence rates for Tikhonov regularization with positively homogeneous functionals
Inverse Probl. 27.7. June 2011

K. Bredies and D. Lorenz
Linear convergence of iterative soft-thresholding
Journal of Fourier Analysis and Applications
14.5-6. 2008

Restricted Isometry Property (RIP)

Candès, Romberg, and Tao 2006: Key ingredient in proving linear convergence rates for the finite dimensional ℓ^1-residual method:
The *s-restricted isometry constant* ϑ_s of L is defined as the smallest number $\vartheta \geq 0$ that satisfies

$$(1-\vartheta)\|u\|^2 \leq \|Lu\|^2 \leq (1+\vartheta)\|u\|^2$$

for all s-sparse $u \in X$. The (s,s')-*restricted orthogonality constant* $\vartheta_{s,s'}$ of L is defined as the smallest number $\vartheta \geq 0$ such that

$$\left|\langle Lu, Lu'\rangle\right| \leq \vartheta \|u\| \left\|u'\right\|$$

for all s-sparse u and s'-sparse u' with $\operatorname{supp}(u) \cap \operatorname{supp}(u') = \emptyset$.
The mapping L satisfies the s-restricted isometry property, if $\vartheta_s + \vartheta_{s,s} + \vartheta_{s,2s} < 1$

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

Linear Convergence of Candes & Rhomberg & Tao

Assumptions:

1. L satisfies the s-restricted isometry property
2. $u^\dagger$ is s-sparse

Result:

$$\left\| u_* - u^\dagger \right\|_{H_1} \leq c_s \delta$$

However: These condition imply the source condition and the restricted injectivity

$0 < p < 1$: Nonconvex sparsity regularization

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

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

- Zarzer 2009: $\mathcal{O}(\sqrt{\delta})$
- Grasmair 2010b: $\Rightarrow \mathcal{O}(\delta)$

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

An Application: Wintertechnik AG and Alps

Ground Penetrating Radar: Location of avalanche victims

GPR: L ist the spherical mean operator

Assumption: GPR which focused radar wave

Figure: Simulations with noise free synthetic data: Left: Data. Middle: Reconstruction by Kirchhoff migration. Right: Reconstruction with sparsity regularization

M. Grasmair, M. Haltmeier, and O. Scherzer
Sparsity in Inverse Geophysical Problems
Handbook of Geomathematics
ed. by W. Freeden, M. Z. Nashed, and T. Sonar
Springer Berlin Heidelberg, 2015

GPR: Simulations with noisy data

Figure: **Noisy data.** Left: Data. Middle: Reconstruction by Kirchhoff migration. Right: Reconstruction with sparsity regularization

Reconstruction with real data

Figure: **Reconstruction from real data.** Left: Data. Middle: Reconstruction by Kirchhoff migration. Right: Reconstruction with sparsity regularization

TV-Regularization

Let Ω, Σ two two open sets. TV minimization consists in calculating

$$u_\alpha^\delta := \text{argmin}_{u \in L^2(\Omega)} \left\{ \frac{1}{2} \left\| Lu - y^\delta \right\|_{L^2(\Sigma)}^2 + \alpha TV[u] \right\}$$

L. I. Rudin, S. Osher, and E. Fatemi
Nonlinear total variation based noise removal algorithms
Physica D. Nonlinear Phenomena 60.1–4. 1992

TV-Regularization

- Assumption: L is a bounded operator between $L^2(\Omega)$ and $L^2(\Sigma)$
- Fact: TV is weakly lower semi-continuous on $L^2(\Omega)$

Results:

- *Stability:* $y^\delta \to_{L^2(\Sigma)} y \Rightarrow u_\alpha^\delta \rightharpoonup_{L^2(\Omega)} u_\alpha$ and $TV[u_\alpha^\delta] \to TV[u_\alpha]$
- *Convergence:* $y^\delta \to_{L^2(\Sigma)} y$ and $\alpha = \alpha(\delta)$ such that $\delta^2/\alpha \to 0$, then

$$\boxed{u_\alpha^\delta \rightharpoonup_{L^2(\Omega)} u^\dagger \text{ and } TV[u_\alpha^\delta] \to TV[u^\dagger]}$$

TV-Regularization: Source Condition

$u^\dagger$ satisfies the *source condition* if there exist $\xi \in L^2(\Omega)$ and $\eta \in L^2(\Sigma)$ such that

$$\boxed{\xi = L^*\eta \in \partial TV[u^\dagger]}$$

Then for $\alpha \sim \delta$

$$\boxed{TV[u_\alpha^\delta] - TV[u^\dagger] - \langle \xi, u_\alpha^\delta - u^\dagger \rangle_{L^2(\Omega)} = D_\xi TV(u_\alpha^\delta, u^\dagger) = \mathcal{O}(\delta)}$$

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

O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen
Variational methods in imaging
Springer, 2009

Source Condition for the Circular Radon Transform

Notation: $\Omega := B(0,1) \subseteq \mathbb{R}^2$ open, $\varepsilon \in (0,1)$. We consider the Circular Radon transform

$$\mathbb{S}_{circ}[u] := t \int_{\mathbb{S}^1} u(z + tw) d\mathcal{H}^1(w)$$

for functions from

$$L^2(B(0,1-\varepsilon)) := \left\{ u \in L^2(\mathbb{R}^2) : \operatorname{supp}(u) \subseteq \overline{B(0,1-\varepsilon)} \right\}$$

- is well-defined
- bounded from $L^2(B(0,1-\varepsilon))$ into $L^2(\mathbb{S}^1 \times (0,1))$
- and $\|\mathbb{S}_{circ}\| \leq 2\pi$

O. Scherzer, M. Grasmair, H. Grossauer,
M. Haltmeier, and F. Lenzen
Variational methods in imaging
Springer, 2009

Finer Properties of the Circular Radon Transform

- There exists a constant $C_\varepsilon > 0$, such that

$$C_\varepsilon^{-1} \|\mathbb{S}_{circ} u\|_2 \leq \|i^*(u)\|_{1/2,2} \leq C_\varepsilon \|\mathbb{S}_{circ} u\|_2 \ , \quad u \in L^2(B(0,1-\varepsilon))$$

  where i^* is the adjoint of the embedding
  $i : W^{1/2,2}(B(0,1)) \to L^2(B(0,1))$

- For every $\varepsilon \in (0,1)$ we have

$$\boxed{W^{1/2,2}(B(0,1-\varepsilon)) = \mathcal{R}(\mathbb{S}_{circ}^*) \cap L^2(B(0,1-\varepsilon))}$$

O. Scherzer, M. Grasmair, H. Grossauer,
M. Haltmeier, and F. Lenzen
Variational methods in imaging
Springer, 2009

Wellposedness of TV-minimization for $\mathbb{S}_{circ}$

Minimization of the TV-functional with $L = \mathbb{S}_{circ}$ is

- well-posed, stable, and convergent
- Let $\varepsilon \in (0, 1)$ and $u^\dagger$ the TV-minimizing solution. Moreover, if the Source Condition

$$\xi \in \partial TV[u^\dagger] \cap W^{1/2,2}(B(0, 1 - \varepsilon))$$

is satisfied, then

$$\boxed{TV[u_\alpha^\delta] - TV[u^\dagger] - \langle \xi, u_\alpha^\delta - u^\dagger \rangle = \mathcal{O}(\delta)}$$

Functions that satisfy the Source Condition

- Let $\rho \in C_0^\infty(\mathbb{R}^2)$ be an adequate mollifier and ρ_μ the scaled function of ρ. Moreover, let $x_0 = (0.2, 0)$, $a = 0.1$, and $\mu = 0.3$. Then

$$u^\dagger := 1_{B(x_0, a+\mu)} * \rho_\mu$$

  satisfies the source condition
- Let $u^\dagger := 1_F$ be the indicator function of a bounded subset of $\mathbb{R}^2$ with smooth boundary

Convergence of Level-Sets

$\Omega \subset \mathbb{R}^2$!

$$\frac{1}{2}\left\|Lu - y^\delta\right\|_{L^2(\Sigma)}^2 + \alpha TV[u] \to \min$$

for

$$u \in L^2(\Omega) \cong \left\{u \in L^2(\mathbb{R}^2) : \text{supp}(u) \subset \overline{\Omega}\right\}$$

A. Chambolle, V. Duval, G. Peyré, and C. Poon
Geometric properties of solutions to the total variation denoising problem
Inverse Problems 33.1. 2017

J. A. Iglesias, G. Mercier, and O. Scherzer
A note on convergence of solutions of total variation regularized linear inverse problems
Inverse Probl. 35.5. 2018

Convergence of Level-Sets

t-super level-set of u_α^δ:

$$U_\alpha^\delta(t) := \left\{ x \in \Omega : u_\alpha^\delta(x) \geq t \right\} \quad \text{for } t \geq 0$$
$$U_\alpha^\delta(t) := \left\{ x \in \Omega : u_\alpha^\delta(x) \leq t \right\} \quad \text{for } t < 0$$

Theorem

Assume that source condition holds! Let $\delta_n, \alpha_n \to 0^+$ such that $\frac{\delta_n}{\alpha_n} \leq \sqrt{\pi}/2$. Then, up to a subsequence and for almost all $t \in \mathbb{R}$, denoting $U_n := U_{\alpha_n}^{\delta_n}$,

$$\lim_{n\to\infty} |U_n(t) \Delta U^\dagger(t)| = 0, \quad \textit{and} \quad \lim_{n\to\infty} \partial U_n(t) = \partial U^\dagger(t).$$

A. Chambolle, V. Duval, G. Peyré, and C. Poon
Geometric properties of solutions to the total variation denoising problem
Inverse Problems 33.1. 2017

J. A. Iglesias, G. Mercier, and O. Scherzer
A note on convergence of solutions of total variation regularized linear inverse problems
Inverse Probl. 35.5. 2018

A Deblurring Result

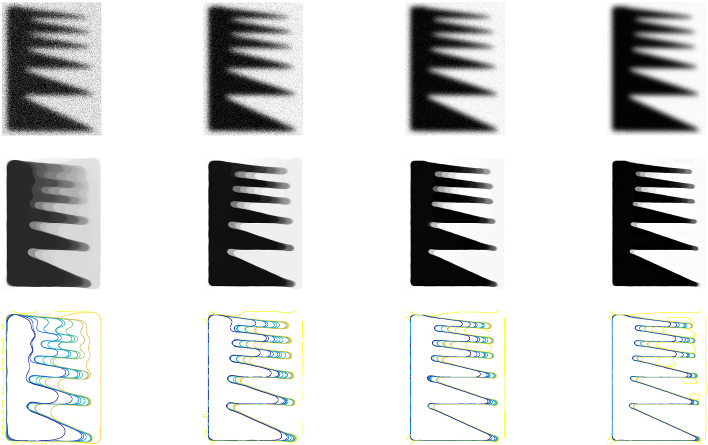

Figure: Deblurring of a characteristic function by total variation regularization with Dirichlet boundary conditions. First row: Input image blurred with a known kernel and with additive noise. Second row: numerical deconvolution results. Third row: some level lines of the results.

Image Registration: Model Problems

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

$$\boxed{L[u] := I_2 \circ u = I_1}$$

u should be a diffeomorphism (no twists)

Calculus of Variations: Notions of Convexity

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

Hierarchy:

f **convex** $\Rightarrow$ *polyconvex* $\Rightarrow$ **quasi-convex** $\Rightarrow$ **rank-one convex**

Up to quasi-convexity:

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

$H_1 := W^{1,p}(\Omega, \mathbb{R}^n)$ with $1 \leq p \leq \infty$

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

Polyconvex functionals are used in elasticity theory

C. B. Morrey
Multiple Integrals in the Calculus of Variations
Springer Verlag, 1966

B. Dacorogna
Direct Methods in the Calculus of Variations
Springer Verlag, 1989

Polyconvex Functions

For $A \in \mathbb{R}^{m \times n}$ and $1 \leq s \leq m \wedge n$

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

$f : \mathbb{R}^{m \times n} \to \mathbb{R} \cup \{+\infty\}$ is polyconvex if

$$f = F \circ T,$$

where $F : \mathbb{R}^{\tau(m,n)} \to \mathbb{R} \cup \{+\infty\}$ is convex and

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

Typical example:

$$f(A) = (\det[A])^2$$

J. M. Ball
Convexity conditions and existence theorems in nonlinear elasticity
Archive for Rational Mechanics and Analysis 63. 1977

Polyconvex Regularization

Assumptions:

- $\mathcal{R}[u] := \int_\Omega F \circ T[u](x)\,dx$.
- L is a non-linear continous operator between $W^{1,p}(\Omega, \mathbb{R}^n)$ and H_2 (sometimes needs to be a Banach space) with closed and convex domain of definition $\mathcal{D}(L)$

Results:

- *Stability:* $y^\delta \to_{H_2} y \Rightarrow u_\alpha^\delta \rightharpoonup_{W^{1,p}} u_\alpha$ and $\mathcal{R}[u_\alpha^\delta] \to \mathcal{R}[u_\alpha]$
- *Convergence:* $y^\delta \to_{H_2} y$ and $\alpha = \alpha(\delta)$ such that $\delta^2/\alpha \to 0$, then

$$\boxed{u_\alpha^\delta \rightharpoonup_{W^{1,p}} u^\dagger \text{ and } \mathcal{R}[u_\alpha^\delta] \to \mathcal{R}[u^\dagger]}$$

Generalized Bregman Distances

Let W be a family of functionals on $H_1 = W^{1,p}(\Omega, \mathbb{R}^n)$

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

$$\boxed{\partial_W \mathcal{R}[u] = \{w \in W : \mathcal{R}[v] \geq \mathcal{R}[u] + w[v] - w[u]\,, \forall v \in H_1\}}$$

- 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 $H_1 = W^{1,p}(\Omega, \mathbb{R}^n)$.

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

We define

$$\begin{aligned} W_{\text{poly}} := \{ w : H_1 \to \mathbb{R} : \exists (u^*, v^*) \in H_1^* \times S_2^* \text{ s.t.} \\ w[u] = \langle u^*, u \rangle_{H_1^*, H_1} + \langle v^*, T(\nabla u) \rangle_{S_2^*, S_2} \} \end{aligned}$$

Remark:

- $W_{\text{poly}} = (H_1 \times S_2)^*$. However, functionals w are non-linear
- W_{poly}-Bregman distance:

$$\begin{aligned} D_w^{\text{poly}}(u, \bar{u}) &= \mathcal{R}[u] - \mathcal{R}(\bar{u}) - w[u] + w(\bar{u}) \\ &= \mathcal{R}[u] - \mathcal{R}(\bar{u}) - \langle u^*, u - \bar{u} \rangle_{H_1^*, H_1} \\ &\quad - \langle v^*, T(\nabla u) - T(\nabla \bar{u}) \rangle_{S_2^*, S_2} \end{aligned}$$

Polyconvex Subgradient

- $\Omega \subset \mathbb{R}^m$ and $H_1 = W^{1,p}(\Omega, \mathbb{R}^n)$
- For $x \in \Omega$, the map $(u, A) \mapsto F(x, u, A)$ is convex and differentiable
- $\mathcal{R}[u] = \int_\Omega F(x, u(x), T(\nabla u(x)))dx$

Definition

If $\mathcal{R}[\bar{v}] \in \mathbb{R}$ and the function $x \mapsto F'_{u,A}(x, \bar{v}(x), T(\nabla \bar{v}(x)))$ lies in

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

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

Rates result

Let $H_1 = W^{1,p}(\Omega, \mathbb{R}^n)$ and consider regularization by

$$u \to \left\| L[u] - y^\delta \right\|^2 + \alpha \mathcal{R}[u]$$

Assumptions:

- $\mathcal{R}$ has a W_{poly}-subgradient w at $u^\dagger$
- Let $\alpha(\delta) \sim \delta$ and $\exists \beta_1 \in [0,1), \beta_2$ such that in a neighborhood

$$\boxed{w[u^\dagger] - w[u] \leq \beta_1 D_w^{\mathrm{poly}}(u, u^\dagger) + \beta_2 \| L[u] - y \|}$$

Results:

$$D_w^{\mathrm{poly}}(u_\alpha^\delta, u^\dagger) = \mathcal{O}(\delta) \quad \text{and} \quad \left\| L[u] - y^\delta \right\| = \mathcal{O}(\delta)$$

Note, that for polyconvex regularization one requires a stronger condition than for convex regularization.

C. Kirisits and O. Scherzer
Convergence rates for regularization functionals with polyconvex integrands
Inverse Probl. 33.8, Aug. 2017

Thank you for your attention