Part B: The distinguished role of smoothness in variational regularization for the solution of nonlinear inverse problems in Banach spaces

BERND HOFMANN

TU Chemnitz
Faculty of Mathematics
D-09107 Chemnitz, GERMANY

Thematic Program "Parameter Identification in Mathematical Models"

IMPA, Rio de Janeiro, October 2017

Research supported by Deutsche Forschungsgemeinschaft (DFG-grant HO 1454/10-1)

Email: hofmannb@mathematik.tu-chemnitz.de

Internet: www.tu-chemnitz.de/mathematik/ip/index.php?en=1

Outline

Outline

Outline

Outline

Outline

Outline

1. Introduction
2. Learning from Hilbert space situations
3. Rates based on variational source conditions
4. When do variational inequalities occur?
5. No common source conditions but variational inequalities in ℓ^1-regularization when the sparsity assumption fails
6. References

Outline

We recall:

Mathematical description of nonlinear inverse problems

Let X, Y be infinite dimensional Banach spaces with norms $\|\cdot\|_X, \|\cdot\|_Y$, dual spaces X^*, Y^* and dual pairings $\langle \cdot, \cdot \rangle_{X^* \times X}$. Moreover we denote by τ_X, τ_Y topologies in X, Y which are weaker than the norm topology.

$F : \mathcal{D}(F) \subseteq X \longrightarrow Y$ **forward operator** with domain $\mathcal{D}(F)$.

We must treat the operator equation

$$\boxed{F(x) = y \qquad (x \in \mathcal{D}(F) \subseteq X,\ y \in Y) \qquad (**)}$$

with solution $x^\dagger \in \mathcal{D}(F)$ and exact right-hand side $y = F(x^\dagger)$, which is in most cases **ill-posed** and **nonlinear**.

Tikhonov-type regularization

For the stable approximate solution of $(**)$ we consider with convex and stabilizing functional $\mathcal{R} : \mathcal{D}(\mathcal{R}) \subseteq X :\to \mathbb{R}$ and for noisy data y^δ assuming a deterministic noise model

$$\|y - y^\delta\|_Y \le \delta$$

variational regularization (Tikhonov-type regularization)

$$\boxed{T_\alpha^\delta(x) := \frac{1}{p}\,\|F(x) - y^\delta\|_Y^p + \alpha\,\mathcal{R}(x) \to \min,}$$

subject to $x \in \mathcal{D} := \mathcal{D}(F) \cap \mathcal{D}(\mathcal{R})$, with exponents $1 \le p < \infty$, regularization parameters $\alpha > 0$ and minimizers $x_\alpha^\delta \in \mathcal{D}(F)$.

Assumption 1

- X, Y are Banach spaces and $\mathcal{D}(F)$ is a convex subset of X.
- F is weak-to-weak τ_X-τ_Y-sequentially continuous and $\mathcal{D}(F)$ is τ_X-weakly closed, hence F weak-to-weak closed.
- $\mathcal{R}$ is **convex** and τ_X-weakly lower semi-continuous.
- $\mathcal{D} = \mathcal{D}(F) \cap \mathcal{D}(\mathcal{R}) \neq \emptyset$.
- $\mathcal{R}$ is **stabilizing**, which means that for every $c \geq 0$ the sublevel sets

  $$\mathcal{M}^{\mathcal{R}}(c) := \{x \in \mathcal{D}(F) : \mathcal{R}(x) \leq c\} \ ,$$

  are τ_X-weakly sequentially pre-compact in the sense that every sequence $\{x_k\}$ in $\mathcal{M}^{\mathcal{R}}(c)$ has a subsequence, which is τ_X-convergent in X to some element from X.

Stabilizing functionals and coercivity

a): For a **reflexive Banach space X** choose **weak convergence** $\rightharpoonup$ as τ_X-convergence.
If $\sup_{x \in \mathcal{M}^{\mathcal{R}}(c)} \|x\|_X < \infty$ for all $c \geq 0$, then $\mathcal{R}$ is stabilizing since the closed unit ball in X is weakly sequentially pre-compact.

b): For a **non-reflexive Banach space X** $=$ **Z*** **with predual separable Banach space Z** choose **weak*** **convergence** $\rightharpoonup^*$ as τ_X-convergence.
If $\sup_{x \in \mathcal{M}^{\mathcal{R}}(c)} \|x\|_X < \infty$ for all $c \geq 0$, then $\mathcal{R}$ is stabilizing since the closed unit ball in X is weak* sequentially pre-compact (sequential Banach-Alaoglu theorem).

An element $x^\dagger \in \mathcal{D}$ is called an $\mathcal{R}$**-minimizing solution** to $(**)$ if

$$\mathcal{R}(x^\dagger) = \min \{\mathcal{R}(x) : x \in \mathcal{D},\ F(x) = y\}\,.$$

$\mathcal{R}$-minimizing solutions always exist under Assumption 1 and attainability, i.e. if, for given $y \in Y$, $(**)$ has a solution $x \in \mathcal{D}$.

Results on **existence, stability and convergence** of $\mathcal{R}$-minimizing solutions $x^\dagger$ and **regularized solutions** x_α^δ for arbitrary $\alpha > 0$ can be found in

$\triangleright$ H./KALTENBACHER/P./SCHERZER 2007, $\triangleright$ PÖSCHL 2008.

We introduce a general non-negative **error measure** $E(x, x^\dagger)$ applied to any approximate solution x for evaluating its quality.

The standard case is the **norm error**

$$\boxed{E(x, x^\dagger) := \|x - x^\dagger\|_X \,,}$$

but in (reflexive) Banach space regularization we often exploit

$$\boxed{E(x, x^\dagger) := B_{\xi^\dagger}^{\mathcal{R}}(x, x^\dagger)\,,}$$

the **Bregman distance** (cf. ▷ BURGER/OSHER 2004) at $x^\dagger \in \mathcal{D}(\mathcal{R}) \subseteq X$ and $\xi^\dagger \in \partial\mathcal{R}(x^\dagger) \subseteq X^*$ for $\mathcal{R}$ with subdifferential $\partial\mathcal{R}$ defined as

$$B_{\xi^\dagger}^{\mathcal{R}}(x, x^\dagger) := \mathcal{R}(x) - \mathcal{R}(x^\dagger) - \langle \xi^\dagger, x - x^\dagger \rangle_{X^* \times X}.$$

$\mathcal{D}_B(\mathcal{R}) := \{x \in \mathcal{D}(\mathcal{R}) : \partial\mathcal{R}(x) \neq \emptyset\}$ is called Bregman domain.

Partially we also need:

Assumption 2

Let F, $\mathcal{R}$, X, Y and $\mathcal{D}$ satisfy Assumption 1.

- There exists an $\mathcal{R}$-minimizing solution $x^\dagger$ which is an element of the Bregman domain $\mathcal{D}_B(\mathcal{R})$.
- There is a bounded linear operator $F'(x^\dagger) : X \to Y$ such that we have for the one-sided directional derivative at $x^\dagger$ and for every $x \in \mathcal{D}$ the equality

$$\lim_{t \to 0+} \frac{1}{t}\Big(F(x^\dagger + t(x - x^\dagger)) - F(x^\dagger)\Big) = F'(x^\dagger)(x - x^\dagger)\,.$$

The operator $F'(x^\dagger)$ has Gâteaux derivative like properties, and there is an adjoint operator $F'(x^\dagger)^* : Y^* \to X^*$

Example: Standard situation in Hilbert spaces

X, Y **Hilbert spaces**,

$\mathcal{R}(x) := \|x - \bar{x}\|_X^2$, $\quad x^\dagger$ is called $\bar{x}$-minimum norm solution

$$T_\alpha^\delta(x) := \frac{1}{2}\,\|F(x) - y^\delta\|_Y^2 + \alpha\,\|x - \bar{x}\|_X^2$$

$\mathcal{D}(\mathcal{R}) = \mathcal{D}_B(\mathcal{R}) = X$, since $\partial\mathcal{R}(x)$ is singleton

$$\xi^\dagger := \mathcal{R}'(x^\dagger) = 2(x^\dagger - \bar{x})$$

$$B_{\xi^\dagger}^{\mathcal{R}}(x, x^\dagger) = \|x - x^\dagger\|_X^2.$$

Example: Regularization with differential operators

X, Y **Hilbert spaces**, $\quad p := 2$

$\mathcal{R}(x) := \|Sx\|_X^2$ with unbounded s.a. operator $S : \mathcal{D}(S) \subset X \to X$

$$T_\alpha^\delta(x) := \frac{1}{2}\|F(x) - y^\delta\|_Y^2 + \alpha\,\|Sx\|_X^2$$

$\mathcal{D}(\mathcal{R}) = \widetilde{X}$ Hilbert space with stronger norm $\|x\|_{\widetilde{X}} := \|Sx\|_X$

$$\xi^\dagger := \mathcal{R}'(x^\dagger) = 2S^2x^\dagger$$

$$B_{\xi^\dagger}^{\mathcal{R}}(x, x^\dagger) = \|S(x - x^\dagger)\|_X^2 \quad \text{with} \quad \mathcal{D}_B(\mathcal{R}) = \mathcal{D}(S^2)$$

Example: Power-type penalties in Banach spaces

$$X, Y \ \textbf{Banach spaces}, \qquad \mathcal{R}(x) := \frac{1}{q}\,\|x\|_X^q,$$

$$T_\alpha^\delta(x) := \frac{1}{p}\,\|F(x) - y^\delta\|_Y^p + \frac{\alpha}{q}\,\|x\|_X^q \qquad (p, q \geq 1)$$

$\mathcal{D}(\mathcal{R}) = \mathcal{D}_B(\mathcal{R}) = X$, since $\mathcal{R}(x)$ is differentiable with

$\xi^\dagger := \mathcal{R}'(x^\dagger) = J_q(x^\dagger)$ with $J_q : X \to X^*$ duality mapping

$$B_{\xi^\dagger}^{\mathcal{R}}(x, x^\dagger) = \frac{1}{q}\,\|x\|_X^q - \frac{1}{q}\,\|x^\dagger\|_X^q - \langle J_q(x^\dagger), x - x^\dagger \rangle_{X^* \times X}.$$

Factors influencing the error and link conditions

We search for **convergence rates**

$$E(x_\alpha^\delta, x^\dagger) = \mathcal{O}(\varphi(\delta)) \quad \text{as} \quad \delta \to 0$$

with **index functions** φ.

We call a function $\varphi : (0, \infty) \to (0, \infty)$ index function if it is continuous and strictly increasing with $\lim_{t \to +0} \varphi(t) = 0$.

Rate results require

- Appropriate choices of the regularization parameter **a priori** as $\alpha = \alpha(\delta)$ and **a posteriori** as $\alpha = \alpha(\delta, y^\delta)$.
- The appropriate interplay of all model components.

The attainability of convergence rates will depend on the interplay of the following four relevant ingredients, as these are:

(i) the **smoothness of the solution** $x^\dagger$,

(ii) the **nonlinearity structure** of the forward operator F,

(iii) properties of the **penalty** $\mathcal{R}$,

(iv) and the character of the **error measure** $E(x, x^\dagger)$.

Link conditions are necessary for combining the four factors.

In Hilbert spaces solution smoothness can be expressed by variable Hilbert scales and **general source conditions** (see $\triangleright$ PEREVERZYEV, MATHÉ, HEGLAND).

Outline

Learning from Hilbert space situations

For Hilbert spaces X and Y we consider now the classical version of **Tikhonov regularization.** We omit for simplicity norm indices in this section.

$$\boxed{T_\alpha^\delta(x) := \|F(x)-y^\delta\|^2+\alpha\,\|x-\overline{x}\|^2 \to \text{min}, \;\; \text{s.t. } x \in \mathcal{D}(F)}$$

with regularization parameters $\alpha>0$ and

minimizers (**regularized solutions**) $x_\alpha^\delta \in \mathcal{D}(F)$.

$\overline{x} \in X$ plays the role of a **reference element** (initial guess).

First preliminary assumption

- X, Y are Hilbert spaces and $\mathcal{D}(F)$ is a convex subset of X.
- F is weakly sequentially closed.

Under this assumption there exist **$\overline{\mathbf{x}}$-minimum-norm solutions** $x^\dagger \in \mathcal{D}(F)$ of $(**)$ with $F(x^\dagger) = y$ and

$$\|x^\dagger - \overline{x}\| = \min\{\|x - \overline{x}\| : \; F(x) = y, \; x \in \mathcal{D}(F)\}$$

for arbitrarily chosen reference elements $\overline{x} \in X$.

Moreover, there exist **regularized solutions** x_α^δ for all $\alpha > 0$ and arbitrary data elements $y^\delta \in Y$.

Second preliminary assumption

- For all $\overline{x}$-minimum-norm solutions $x^\dagger$ there exists a bounded linear operator $F'(x^\dagger) : X \to Y$ such that

$$\lim_{t \to +0} \frac{F(x^\dagger + t(x - x^\dagger)) - F(x^\dagger)}{t} = F'(x^\dagger)(x - x^\dagger)$$

  holds for all $x \in \mathcal{D}(F)$.

- For all $\overline{x}$-minimum-norm solutions $x^\dagger$ there are a constant $K > 0$ and a radius $r > 0$ of the ball $\overline{B}_r(x^\dagger) := \{z \in X : \|z - x^\dagger\| \le r\}$ such that

$$\|F(x) - F(x^\dagger) - F'(x^\dagger)(x - x^\dagger)\| \le K \, \|x - x^\dagger\|^2 \qquad (Lip)$$

  holds for all $x \in \mathcal{D}(F) \cap \overline{B}_r(x^\dagger)$.

Seminal conv. rate result by ▷ ENGL/KUNISCH/NEUBAUER 1989

Under both preliminary assumptions and for an a priori parameter choice $\alpha = \alpha(\delta) \sim \delta$ we have a convergence rate

$$\|x_\alpha^\delta - x^\dagger\| = \mathcal{O}(\sqrt{\delta}) \qquad \text{as} \qquad \delta \to 0$$

if the benchmark source condition

$$x^\dagger = \bar{x} + \frac{1}{2} F'(x^\dagger)^* v \qquad\qquad (BSC)$$

for an $\bar{x}$-minimum-norm solution $x^\dagger$ and the smallness condition

$$K \|v\| < 1 \qquad\qquad (SMC)$$

for the source element $v \in Y$ are satisfied.

Whenever for the choice of $\alpha > 0$ the limit conditions

$$\alpha \to 0 \qquad \text{and} \qquad \frac{\delta^2}{\alpha} \to 0 \qquad \text{as} \qquad \delta \to 0$$

hold, then x_α^δ converges in the sense of subsequences to $\overline{x}$-minimum-norm solutions of $(**)$.

Consequently, if multiple $\overline{x}$-minimum-norm solutions exist, then only one of them can satisfy the benchmark source condition together with the smallness condition.

If the benchmark source condition (*BSC*) or at least (*SMC*) fail, then **more qualified nonlinearity conditions** are required for **convergence rates** under **low order source conditions**:

1. Hölder source conditions with small exponents $0 < \nu < \frac{1}{2}$:

$$x^\dagger = \overline{x} + (F'(x^\dagger)^* F'(x^\dagger))^\nu \, w, \quad w \in X,$$

for $\alpha \sim \delta^{\frac{2}{2\nu+1}}$ yielding: $\quad \|x_\alpha^\delta - x^\dagger\| = \mathcal{O}\left(\delta^{\frac{2\nu}{2\nu+1}}\right) \quad$ as $\quad \delta \to 0$.

2. Logarithmic source conditions:

$$x^\dagger = \overline{x} + f(F'(x^\dagger)^* F'(x^\dagger)) \, w, \quad w \in X, \quad f(t) := (-\log t)^{-\mu}, \;\; \mu > 0,$$

for $\alpha \sim \delta$ yielding: $\quad \|x_\alpha^\delta - x^\dagger\| = \mathcal{O}\left((-\log \delta)^{-\mu}\right) \quad$ as $\quad \delta \to 0$.

Nonlinearity conditions

Most powerful is the **tangential cone condition**

$$\|F(x) - F(x^\dagger) - F'(x^\dagger)(x - x^\dagger)\| \le C\,\|F(x) - F(x^\dagger)\| \qquad (TCC)$$

for some constant $0 < C < \infty$ and all $x \in \overline{B}_r(x^\dagger) \cap \mathcal{D}(F)$, often with focus on $0 < C < 1$ in iterative regularization methods. But the verification is still missing or cannot be proven for large relevant classes of nonlinear inverse problems.

The same can be said for weaker conditions of the form

$$\|F(x)-F(x^\dagger)-F'(x^\dagger)(x-x^\dagger)\| \le C\,\varphi(\|F(x)-F(x^\dagger)\|), \qquad (Phi)$$

where φ is a concave *index function* $\varphi : (0,\infty) \to (0,\infty)$.

Convergence rates under (*BSC*) and (*Phi*) ▷ H./MATHÉ 2012

Provided that (*Lip*) is replaced by (*Phi*), then we have under the benchmark source condition (*BSC*) a convergence rate

$$\|x_\alpha^\delta - x^\dagger\| = \mathcal{O}(\sqrt{\varphi(\delta)}) \qquad \text{as} \qquad \delta \to 0$$

if the regularization parameter $\alpha > 0$ is selected a priori as $\alpha(\delta) = \frac{\delta^2}{\varphi(\delta)}$ or a posteriori by using the sequential discrepancy principle.

By using the **method of approximate source conditions**:

Conv. rates under approximate SC and (*Phi*) ▷ Boţ/H. 2010

Provided that (*Lip*) is replaced by (*Phi*), then we have with the auxiliary function $\Psi(R) = d_{x^\dagger}(R)^2/R$ and for the distance function

$$d_{x^\dagger}(R) = \min\{\|x^\dagger - \overline{x} - \frac{1}{2}F'(x^\dagger)^* w\| : \, w \in Y, \|w\| \le R\} \to 0$$

as $R \to \infty$ a convergence rate

$$\|x_\alpha^\delta - x^\dagger\| = \mathcal{O}\left(d_{x^\dagger}(\Psi^{-1}(\varphi(\delta))\right) \qquad \text{as} \qquad \delta \to 0$$

if the regularization parameter $\alpha > 0$ is selected appropriately.

Technical assumption here: $F'(x^\dagger)$ is **injective** linear operator.

To obtain Hölder rates 1. and logarithmic rates 2.
there are two more options $\triangleright$ KALTENBACHER JIIP 2008:

Left-side rotation:

$$F'(x) = R(x,x^\dagger)F'(x^\dagger), \quad \|R(x,x^\dagger)-I\|_{Y\to Y} \le C_R\,\|x-x^\dagger\|^\kappa \quad (L)$$

for $0<\kappa\le 1,\ 0<C_R<\infty$, and all $x\in \overline{B}_r(x^\dagger)\subseteq \mathcal{D}(F)$

Right-side rotation:

$$F'(x) = F'(x^\dagger)R(x,x^\dagger), \quad \|R(x,x^\dagger)-I\|_{Y\to Y} \le C_R\,\|x-x^\dagger\|^\kappa \quad (R)$$

for $0<\kappa\le 1,\ 0<C_R<\infty$, and all $x\in \overline{B}_r(x^\dagger)\subseteq \mathcal{D}(F)$

For (L) the mean value theorem in integral form yields

$$\|F(x) - F(x^\dagger) - F'(x^\dagger)(x - x^\dagger)\|$$

$$= \| \int_0^1 [F'(x^\dagger + t(x - x^\dagger)) - F'(x^\dagger)](x - x^\dagger) dt\|$$

$$\leq \| \int_0^1 [R(x^\dagger + t(x - x^\dagger), x^\dagger) - I]\, F'(x^\dagger)(x - x^\dagger)\, dt\|$$

$$\leq C_R \left(\int_0^1 t^\kappa \, dt \right) \|F'(x^\dagger)(x - x^\dagger)\| \, \|x - x^\dagger\|^\kappa$$

and hence

$$\|F(x) - F(x^\dagger) - F'(x^\dagger)(x - x^\dagger)\| \leq \frac{C_R}{1+\kappa} \|F'(x^\dagger)(x - x^\dagger)\| \, \|x - x^\dagger\|^\kappa.$$

$$\|F(x)-F(x^\dagger)-F'(x^\dagger)(x-x^\dagger)\| \le \frac{C_R}{1+\kappa}\,\|F'(x^\dagger)(x-x^\dagger)\|\,\|x-x^\dagger\|^\kappa$$

implies on the one hand that

$$\|F(x)-F(x^\dagger)-F'(x^\dagger)(x-x^\dagger)\| \le \widetilde{C}\,\|F'(x^\dagger)(x-x^\dagger)\| \quad (Prime)$$

holds for some constant $0<\widetilde{C}<\infty$ and all $x \in \overline{B}_r(x^\dagger)$.

On the other hand, by using the triangle inequality we even derive the tangential cone condition (*TCC*) in the case of sufficiently small $r>0$, which is also a consequence of (*L*).

There is a complete **deficit in low order convergence rates** if

(a) the benchmark source condition fails, which means that $x^\dagger$ is **too non-smooth**.

and moreover

(b) (L) and (R) fail and there is no concave index function φ such that (*Phi*) holds, which means that the structure of nonlinearity of F is **too poor**.

The autoconvolution equation in the real space $L^2(0,1)$

In this paragraph, we consider the autoconvolution operator F on the space $X = Y = L^2(0,1)$ of quadratically integrable real functions over the unit interval $[0,1]$. Then $(**)$ attains the form

$$[F(x)](s) := \int_0^s x(s-t)x(t)dt = y(s), \quad 0 \le s \le 1, \qquad (**)$$

with $F: L^2(0,1) \to L^2(0,1)$ and $\mathcal{D}(F) = L^2(0,1)$. This operator equation of quadratic type occurs in physics of spectra, in optics and in stochastics, often as part of a more complex task.

We recall for nonlinear operator equations $(**)$ the local well-posedness and ill-posedness concept with some verbal reformulations:

Local ill-posedness and well-posedness

We call a nonlinear operator equation $(**)$ locally well-posed at a solution point $x^\dagger \in \mathcal{D}(F)$ if there is a closed ball $\overline{B}_r(x^\dagger) := \{x \in X :\ \|x - x^\dagger\| \le r\}$ around $x^\dagger$ with radius $r > 0$ such that, for every sequence $\{x_n\}_{n=1}^\infty \subset \overline{B}_r(x^\dagger) \cap \mathcal{D}(F)$, the limit condition $\lim\limits_{n\to\infty} \|F(x_n) - F(x^\dagger)\| = 0$ implies that $\lim\limits_{n\to\infty} \|x_n - x^\dagger\| = 0$. Otherwise the equation is called locally ill-posed at $x^\dagger \in \mathcal{D}(F)$, which means that, for arbitrarily small radii $r > 0$, there exist sequences $\{x_n\}_{n=1}^\infty \subset \overline{B}_r(x^\dagger) \cap \mathcal{D}(F)$ such that $\lim\limits_{n\to\infty} \|F(x_n) - F(x^\dagger)\| = 0$, but $\lim\limits_{n\to\infty} \|x_n - x^\dagger\| = 0$ fails.

Local ill-posedness everywhere for autoconvolution

The simple example of a sequence belonging to $\overline{B}_r(x^\dagger)$,

$$x_n(t) = \begin{cases} x^\dagger(t) & \text{if} \quad 0 \le t \le 1 - \frac{1}{n} \\ x^\dagger(t) + r\sqrt{n} & \text{if} \quad 1 - \frac{1}{n} < t \le 1 \end{cases} \qquad (n = 2, 3, ...),$$

with $\|x_n - x^\dagger\| = r$, but

$$\|F(x_n) - F(x^\dagger)\| \le 2r \int_0^{1/n} |x^\dagger(t)| dt \le \frac{2r}{\sqrt{n}} \|x^\dagger\| \to 0 \text{ as } n \to \infty,$$

shows that the equation $(**)$ is locally ill-posed at every point $x^\dagger \in L^2(0, 1)$.

This ill-posedness occurs although the nonlinear autoconvolution operator F is **not compact**, but has a **compact Fréchet derivative**

$$\left[F'(x)h\right](s) = 2\int_0^s x(s-t)h(t)dt, \quad 0 \le s \le 1, \quad h \in L^2(0,1).$$

Based on Titchmarsh's theorem it can be shown that $F'(x^\dagger)$ is just an **injective** operator if

$$\sup\{\,\bar{t} \in [0,1] : \; x^\dagger(t) = 0 \;\text{ a.e. on }\; [0,\bar{t}]\} \;=\; 0. \quad (Inj)$$

If a solution $x^\dagger$ to $(**)$ satisfies the condition (Inj), then $x^\dagger$ and $-x^\dagger$ are the two solutions of this equation.

Moreover, F is **weakly sequentially closed** and $F'(x)$ is Lipschitz continuous and satisfies the condition

$$\|F(x) - F(x^\dagger) - F'(x^\dagger)(x - x^\dagger)\| = \|F(x - x^\dagger)\|^2 \le \|x - x^\dagger\|^2$$

for all $x, x^\dagger \in L^2(0,1)$, hence (Lip) with $K = 1$ and for all $r > 0$.

Proposition

For the autoconvolution operator F under consideration here, mapping in $L^2(0,1)$, and any element $x^\dagger \in L^2(0,1)$ there is no index function η in combination with a radius $r > 0$ such that

$$\|F(x) - F(x^\dagger)\| \le \hat{C}\,\eta(\|F'(x^\dagger)(x - x^\dagger)\|) \quad (Eta)$$

for some constant $0 < \hat{C} < \infty$ and all $x \in \overline{B}_r(x^\dagger)$.

Proof: To construct a contradiction it is enough to find a sequence $\{x_n\}_{n=1}^{\infty} \subset \overline{B}_r(x^\dagger)$ such that $\|F'(x^\dagger)(x_n - x^\dagger)\| \to 0$ as $n \to \infty$, but $\lim\limits_{n\to\infty} \|F(x_n) - F(x^\dagger)\| > 0$. We consider the sequence of functions $x_n = x^\dagger + \Delta_n \in \overline{B}_r(x^\dagger)$ with $\Delta_n(t) = \sqrt{2} r \sin(\pi n t)$ and $\|\Delta_n\| = r > 0$. Taking into account the weak convergence $x_n - x^\dagger \rightharpoonup 0$ in $L^2(0,1)$ we have $\|F'(x^\dagger)(x_n - x^\dagger)\| \to 0$ and for any index function η also $\eta(\|F'(x^\dagger)(x_n - x^\dagger)\|) \to 0$ as $n \to \infty$, because $F'(x^\dagger)$ is a compact operator. However, F is not compact and $\lim\limits_{n\to\infty} \|F(x_n) - F(x^\dagger)\| = \lim\limits_{n\to\infty} \|(2x^\dagger + \Delta_n) * \Delta_n\| = \lim\limits_{n\to\infty} \|\Delta_n * \Delta_n\| = \frac{r^2}{\sqrt{6}} > 0$. This proves the proposition.

Note that we have used in this context the limit $\lim\limits_{n\to\infty} \|x^\dagger * \Delta_n\| = 0$, which is again a consequence of the compactness of linear convolution operators.

Corollary

For the autoconvolution operator F from $(**)$ mapping in $L^2(0,1)$ a condition (*Prime*) and consequently a nonlinearity condition (*L*) cannot hold. Moreover also the tangential cone condition (*TCC*) cannot hold with a small constant $0 < C < 1$.

Proof: From (*Prime*) would have by the triangle inequality

$$\|F(x)-F(x^\dagger)\| \le \|F(x)-F(x^\dagger)-F'(x^\dagger)(x-x^\dagger)\|+\|F'(x^\dagger)(x-x^\dagger)\|$$

$\le (\widetilde{C}+1)\,\|F'(x^\dagger)(x-x^\dagger)\|$ and hence (*Eta*) with $\eta(t)=t$, which contradicts the above proposition. Moreover, (*TCC*) would yield

$$\|F(x)-F(x^\dagger)\| \le \|F(x)-F(x^\dagger)-F'(x^\dagger)(x-x^\dagger)\|+\|F'(x^\dagger)(x-x^\dagger)\|$$

$\le C\,\|F(x)-F(x^\dagger)\| + \|F'(x^\dagger)(x-x^\dagger)\|$, and in particular with $0 < C < 1$

$$\|F(x)-F(x^\dagger)\| \le \frac{1}{1-C}\,\|F'(x^\dagger)(x-x^\dagger)\|,$$

which contradicts again the proposition.

The above proposition, however, says nothing about the validity of the tangential cone condition (TCC) with constants $C \geq 1$.

Conjecure

For the autoconvolution operator F from $(**)$ mapping in $L^2(0,1)$ and $x^\dagger \neq 0$ there is no concave index function φ in combination with a radius $r > 0$ such that

$$\|F'(x^\dagger)(x - x^\dagger)\| \leq \widetilde{C}\,\varphi(\|F(x) - F(x^\dagger)\|) \quad (Dif)$$

holds for some constant $0 < \widetilde{C} < \infty$ and all $x \in \overline{B}_r(x^\dagger)$.

If the conjecture is true, then for the autoconvolution operator also (Phi) and in particular (TCC) cannot hold for $x^\dagger \neq 0$.

Critical inspection of (TCC): What does it mean for $x^\dagger \equiv 1$?

$$H(s) := \int_0^s h(t)dt \quad \text{for} \quad h \in L^2(0,1)$$

$$\|h*h\| \le C\,\|2H+h*h\| \quad \text{for all} \quad h \in L^2(0,1):\ \|h\| \le r \quad (TCC)$$

Is there always a sequence $\{h_n\}$ with primitives $\{H_n\}$ such that

$$\lim_{n\to\infty} \frac{\|2H_n + h_n * h_n\|}{\|h_n * h_n\|} = 0\,?$$

One more chance: nonlinearity condition (R)? But:

Proposition

For the autoconvolution operator F from $(**)$ mapping in $L^2(0,1)$ a nonlinearity condition (R) cannot hold.

A proof was given by Steven Bürger 2015 in his PhD thesis (two full pages A4).

To our best knowledge convergence rates for this autoconvolution equation $(**)$ have been established only if the benchmark source condition

$$x^\dagger(t) = \overline{x}(t) + \int_t^1 x^\dagger(s-t)\, v(s) ds, \quad {\scriptstyle 0 \le t \le 1,\ v \in L^2(0,1)}, \qquad (BSC)$$

is satisfied under the smallness condition

$$\|v\| < 1\,. \qquad (SMC)$$

Proposition

Apart from the trivial case $\overline{x} = x^\dagger$, for the autoconvolution operator F from $(**)$ mapping in $L^2(0,1)$, the conditions (BSC) and (SMC) can only hold if the reference element $\overline{x} \in L^2(0,1)$ is chosen such that

$$\frac{\|x^\dagger - \overline{x}\|}{\|x^\dagger\|} < 1 \qquad\qquad (Rel)$$

and $x^\dagger - \overline{x}$ is a continuous function on $[0,1]$ with $\overline{x}(1) = x^\dagger(1)$. Hence, for the appropriate choice of $\overline{x}$ the value $x^\dagger(1)$ must be known. Furthermore, for the choice $\overline{x} = 0$ there is no $x^\dagger \neq 0$ which satisfies both conditions.

Remark: For $\overline{x} = 0$ the solutions $x^\dagger$ and $-x^\dagger$ have the same distance to the reference element and if $x^\dagger$ satisfies both conditions so also does $-x^\dagger$. This is a contradiction.

Proof: For $\overline{x} = x^\dagger$, both conditions are always satisfied with $v = 0$. By using the norm-conserving linear transformation $v \mapsto \widetilde{v}$ in $L^2(0,1)$ defined as $\widetilde{v}(t) := v(1-t),\ 0 \le t \le 1$, we can rewrite (BSC) as

$$x^\dagger(1-t) - \overline{x}(1-t) = \int_0^t \widetilde{v}(t-s)\, x^\dagger(s) ds, \quad 0 \le t \le 1,$$

or short in convolution form as $x^\dagger - \overline{x} = \widetilde{v} * x^\dagger$. Therefore, the transformation $x(t) \mapsto \overline{x}(t) + \int_t^1 x(s-t) v(s) ds$ in $L^2(0,1)$ is a contractive, affine linear mapping and, for fixed $\|v\| < 1$, by Banach's fixed point theorem there is a uniquely determined solution $x^\dagger \in L^2(0,1)$ satisfying (BSC).

For $\overline{x} = 0$ we have $x^\dagger = 0$ as solution to that equation for all such source elements v.

Now we can estimate $\|x^\dagger - \overline{x}\| \le \|x^\dagger\|\|v\| < \|x^\dagger\|$, for all nonzero solutions $x^\dagger$, which yields the necessary condition (*Rel*). Moreover, $x^\dagger - \overline{x}$ is a continuous function as the result of the convolution of the two functions $\widetilde{v}$ and $x^\dagger$ from $L^2(0,1)$, and thus we have $\overline{x}(1) = x^\dagger(1)$ as another necessary condition.

This proves the proposition.

Outline

Rates based on variational source conditions

Now we are back to Banach spaces X and Y from the introduction. For expressing solution smoothness we use **variational source conditions** (variational inequalities) in a form developed independently by FLEMMING and GRASMAIR 2010-11

Assumption 3 (variational source condition - VSC)

We assume to have a constant $0 < \beta \le 1$*, and a* **concave** *index function* φ *such that*

$$\beta\, E(x, x^\dagger) \le \mathcal{R}(x) - \mathcal{R}(x^\dagger) + \varphi(\|F(x) - F(x^\dagger)\|_Y) \quad \text{for all } \ x \in \mathcal{M}.$$

The set $\mathcal{M}$ of the validity of (VSC) must be large enough such that it contains $x^\dagger$ and all regularized solutions x_α^δ under consideration for $0 < \delta \le \delta_{max}$. This is for example the case if $\mathcal{M} = \mathcal{M}^{\mathcal{R}}(\mathcal{R}(x^\dagger) + c)$ for some $c > 0$.

Namely, for any fixed parameter choice
$\alpha_* = \alpha_*(\delta)$ or $\alpha_* = \alpha_*(y^\delta, \delta)$ satisfying

$$\alpha_* \to 0 \quad \text{and} \quad \frac{\delta^p}{\alpha_*} \to 0 \quad \text{as} \quad \delta \to 0 \qquad (+)$$

we have convergence for both

$$\mathcal{R}(x_{\alpha_*}^\delta) \to \mathcal{R}(x^\dagger) \quad \text{and} \quad \|F(x_{\alpha_*}^\delta) - F(x^\dagger)\|_Y \to 0 \quad \text{as} \quad \delta \to 0. \quad (++)$$

Moreover if $\delta_n \to 0$ then the regularized solutions $x_{\alpha_*}^{\delta_n}$ converge (in the sense of subsequences) with respect to the (weaker) topology τ_X of X to $\mathcal{R}$-minimizing solutions $x^\dagger$.

On a posteriori choices $\alpha_* = \alpha_*(y^\delta, \delta)$: **discrepancy principles**

A **strong discrepancy principle** was used in the literature for Banach spaces and (VI) ($\triangleright$ RAMLAU ET AL.): For two constants $1 < \tau_1 < \tau_2 < \infty$ the regularization parameter α_* has to satisfy the condition

$$\tau_1 \delta \le \|F(x^\delta_{\alpha_*}) - y^\delta\|_Y \le \tau_2 \delta.$$

Duality gaps may destroy its applicability. To avoid this we suggest to use of the **sequential discrepancy principle (SDP)** for which the variational inequality (VSC) is also strong enough to ensure convergence rates.

Here we restrict the selection of the regularization parameter to a discrete exponential grid. Precisely, we select $0 < q < 1$, choose a parameter $\alpha_0 > 0$ large enough and consider the set

$$\Delta_q := \left\{ \alpha_j : \quad \alpha_j := q^j \alpha_0, \quad j = 1, 2, \ldots \right\}.$$

Definition

For prescribed $\tau > 1$ we say that the regularization parameter $\alpha_* \in \Delta_q$ is chosen according to the sequential discrepancy principle (SDP) if

$$\|F(x_{\alpha_*}^\delta) - y^\delta\|_Y \le \tau\delta < \|F(x_{\alpha_*/q}^\delta) - y^\delta\|_Y.$$

In ▷ ANZENGRUBER, H., MATHÉ 2013 we have proven:

Proposition

For $\alpha_* > 0$ from (SDP) we have

$$\alpha_* \to 0 \quad \text{and} \quad \frac{\delta^p}{\alpha_*} \to 0 \quad \text{as} \quad \delta \to 0 \qquad (+)$$

whenever the **exact penalization veto** is satisfied.

Definition

We say that the exact penalization veto is satisfied if, with the exception of singular cases, for arbitrary $\alpha > 0$ an $\mathcal{R}$-minimizing solution $x^\dagger$ cannot be a minimizer of

$$T_\alpha^0(x) := \frac{1}{p}\,\|F(x) - y\|_Y^p + \alpha\,\mathcal{R}(x) \to \min\,.$$

The veto is often failed in the case $p = 1$.

Convergence rates for a natural a priori parameter choice and for the sequential discrepancy principle

Theorem ▷ HOFMANN/MATHÉ 2012

Suppose that $x^\dagger$ obeys (VSC) for some concave index function φ and some set $\mathcal{M}$.

(i) For $p > 1$ let $\alpha_* = \alpha_*(\delta) > 0$ be selected according to the a priori parameter choice $\alpha_* := \frac{\delta^p}{\varphi(\delta)}$.

(ii) For prescribed $\tau > 1$ let $\alpha_* = \alpha_*(\delta, y^\delta) > 0$ be chosen according to the sequential discrepancy principle (SDP).

Provided that $x^\delta_{\alpha_*} \in \mathcal{M}$ for all $0 < \delta \le \delta_{max}$ and some $\delta_{max} > 0$ we have for both parameter choices (i) and (ii) the convergence rates

$$E(x^\delta_{\alpha_*}, x^\dagger) = \mathcal{O}(\varphi(\delta)), \quad \|F(x^\delta_{\alpha_*}) - F(x^\dagger)\|_Y = \mathcal{O}(\delta), \quad \text{and}$$

$$|\mathcal{R}(x^\delta_{\alpha_*}) - \mathcal{R}(x^\dagger)| = \mathcal{O}(\varphi(\delta)) \qquad \text{as } \delta \to 0.$$

Outline

When do variational inequalities occur?

I. The benchmark case

Here we assume that $x^\dagger \in \mathcal{D}_B(\mathcal{R})$ and the subdifferential $\xi^\dagger \in X^*$ fulfills the **benchmark source condition**

$$\xi^\dagger = F'(x^\dagger)^* v \in \partial\mathcal{R}(x^\dagger), \qquad \text{for some } v \in Y^*. \qquad (\$)$$

Such information allows us to bound for all $x \in X$

$$\begin{aligned} &\langle \xi^\dagger, x^\dagger - x\rangle_{X^*\times X} \\ = &\langle (F'(x^\dagger))^* v, x^\dagger - x\rangle_{X^*\times X} = \langle v, F'(x^\dagger)(x^\dagger - x)\rangle_{Y^*\times Y} \\ &\le \|v\|_{Y^*}\|F'(x^\dagger)(x - x^\dagger)\|_Y. \end{aligned}$$

After adding the term $\mathcal{R}(x) - \mathcal{R}(x^\dagger)$ on both sides this yields that

$$B^{\mathcal{R}}_{\xi^\dagger}(x, x^\dagger) \le \mathcal{R}(x) - \mathcal{R}(x^\dagger) + \|v\|_{Y^*}\|F'(x^\dagger)(x - x^\dagger)\|_Y, \; x \in \mathcal{M} := \mathcal{D}(\mathcal{R}).$$

The special case of Hilbert space regularization

X Hilbert space, $\quad \mathcal{R}(x) = \|x - \bar{x}\|_X^2, \quad B_{\xi^\dagger}^{\mathcal{R}}(x, x^\dagger) = \|x - x^\dagger\|_X^2.$

This implies that

$$\|x - x^\dagger\|_X^2 \le \|x - \bar{x}\|_X^2 - \|x^\dagger - \bar{x}\|_X^2 + \|v\|_{Y^*} \|F'(x^\dagger)(x - x^\dagger)\|_Y, \; x \in X,$$

and for a bounded linear operator $F := A : X \to Y$ we have (VSC)

with $\mathcal{M} = X,\; E(x, x^\dagger) = \|x - x^\dagger\|_X^2,\; \beta = 1$ and $\varphi(t) = \|v\|_{Y^*}\, t$.

In this Hilbert space setting for linear ill-posed problems solution smoothness can always be expressed by variational inequalities (VSC) with general index functions φ.

Also in Banach spaces we obtain for bounded linear operators such variational inequalities (VSC) with $\beta = 1$,
$E(x, x^\dagger) = B^{\mathcal{R}}_{\xi^\dagger}(x, x^\dagger)$ and $\varphi(t) = \|v\|_{Y^*} t,\ t > 0$ on $\mathcal{M} = X$.

If the mapping F is nonlinear then we may use certain **structure of nonlinearity** to bound
$\|F'(x^\dagger)(x - x^\dagger)\|_Y$ in terms of $\|F(x^\dagger) - F(x)\|_Y$.

Provided that

$$\|F'(x^\dagger)(x - x^\dagger)\|_Y \le \sigma(\|F(x) - F(x^\dagger)\|_Y), \quad x \in \mathcal{M}, \qquad (\&)$$

holds for some concave index function σ on some set $\mathcal{M} \subset \mathcal{D}(F)$, then we derive (VSC) on $\mathcal{M}$ with
$\beta = 1,\ E(x, x^\dagger) = B^{\mathcal{R}}_{\xi^\dagger}(x, x^\dagger)$ and $\varphi(t) = \|v\|_{Y^*}\sigma(t),\ t > 0$.

An **alternative structural condition** is given in the form

$$\|F(x)-F(x^\dagger)-F'(x^\dagger)(x-x^\dagger)\|_Y \le \eta B^{\mathcal{R}}_{\xi^\dagger}(x,x^\dagger), \quad x \in \mathcal{M}, \quad (\&\&)$$

again for some set $\mathcal{M} \subset \mathcal{D}(F)$ (cf. ▷ RESMERITA, SCHERZER). This allows us to bound

$$\|F'(x^\dagger)(x-x^\dagger)\|_Y \le \eta B^{\mathcal{R}}_{\xi^\dagger}(x,x^\dagger) + \|F(x)-F(x^\dagger)\|_Y,\ x \in \mathcal{M}$$

and further as (VSC) under the **smallness condition**

$$\eta\|v\|_{Y^*} < 1 \qquad (\$\$)$$

with $0 < \beta = 1-\eta\|v\|_{Y^*} \le 1$, $E(x,x^\dagger) = B^{\mathcal{R}}_{\xi^\dagger}(x,x^\dagger)$ and $\varphi(t) = \|v\|_{Y^*}t,\ t > 0$ on $\mathcal{M}$.

II. Violation of the benchmark

If the source condition ($) is violated then we may use the **method of approximate source conditions** to derive variational using the **distance function**

$$d_{\xi^\dagger}(R) \;:= \inf\{\|\xi^\dagger - \xi\|_{X^*} : \, \xi = F'(x^\dagger)^* v, \; v \in Y^*, \; \|v\|_{Y^*} \le R\},$$

which is nonincreasing, continuous and concace for all $R > 0$ and should obey the limit condition

$$d_{\xi^\dagger}(R) \to 0 \quad \text{as} \quad R \to \infty.$$

As mentioned in BOŢ/HOFMANN 2010 this is the case when $F'(x^\dagger)^{**} : X^{**} \to Y^{**}$ is injective.
Additionally this approach presumes q-coercivity

$$B_{\xi^\dagger}^{\mathcal{R}}(x, x^\dagger) \ge c_q \, \|x - x^\dagger\|_X^q \qquad \text{for all} \quad x \in \mathcal{M}, \quad q \ge 2, \; c_q > 0.$$

Such assumption is for example fulfilled if $\mathcal{R}(x) := \|x\|_X^q$ and X is a q-convex Banach space.

Then, for $R > 0$ one can find $v_R \in Y^*$ and $u_R \in X^*$ such that

$$\xi^\dagger = \left(F'(x^\dagger)\right)^* v_R + u_R \qquad \text{with} \qquad \|v_R\|_{Y^*} = R, \;\; \|u_R\|_{X^*} \le d_{\xi^\dagger}(R)\,,$$

and we can estimate for all $R > 0$ and $x \in \mathcal{M}$ as

$$\begin{aligned} -\langle \xi^\dagger, x - x^\dagger \rangle_{X^* \times X} &= -\langle \left(F'(x^\dagger)\right)^* v_R + u_R, x - x^\dagger \rangle_{X^* \times X} \\ &= -\langle v_R, F'(x^\dagger)(x - x^\dagger) \rangle_{Y^* \times Y} + \langle u_R, x^\dagger - x \rangle_{X^* \times X} \\ &\le R\,\|F'(x^\dagger)(x - x^\dagger)\|_Y + d_{\xi^\dagger}(R)\,\|x - x^\dagger\|_X\,. \end{aligned}$$

Adding again the difference $\mathcal{R}(x) - \mathcal{R}(x^\dagger)$ gives for $x \in \mathcal{M}$

$$B_{\xi^\dagger}^{\mathcal{R}}(x, x^\dagger) \le \mathcal{R}(x) - \mathcal{R}(x^\dagger) + R\,\|F'(x^\dagger)(x - x^\dagger)\|_Y + d_{\xi^\dagger}(R)\,\|x - x^\dagger\|_X.$$

Using Young's inequality and the q-coercivity, for the linear case $F'(x^\dagger) = A$, we equilibrate the second and the third term, depending of R and $d_{\xi^\dagger}(R)$, respectively, by means of the auxiliary continuous and strictly decreasing function

$$\Phi(R) := \frac{\left(d_{\xi^\dagger}(R)\right)^{q^*}}{R}, \quad R > 0, \quad \frac{1}{q} + \frac{1}{q^*} = 1\,.$$

By setting $R := \Phi^{-1}\left(\|A(x - x^\dagger)\|_Y\right)$ and introducing the index function

$$\zeta(t) := \left[d_{\xi^\dagger}(\Phi^{-1}(t))\right]^{q^*} \quad (t > 0)$$

we get again a variational inequality (VSC):

$$\beta B_{\xi^\dagger}^{\mathcal{R}}(x, x^\dagger) \le \mathcal{R}(x) - \mathcal{R}(x^\dagger) + \hat{K}\,\zeta(\|A(x - x^\dagger)\|_Y), \quad x \in \mathcal{M}.$$

Outline

No common source conditions but variational inequalities in ℓ^1-regularization when the sparsity assumption fails

▷ BURGER/FLEMMING/H. 2012/2013 and ▷ BOŢ/H. 2013

Under a sparsity expectation we consider for $X = \ell^1 = (c_0)^*$ with the weak*-topology as τ_X in ℓ^1 and $F : \mathcal{D}(F) \subseteq \ell^1 \to Y$ ℓ^1-regularized solutions x_α^δ as minimizers of

$$T_\alpha^\delta(x) := \frac{1}{p} \|F(x) - y^\delta\|_Y^p + \alpha \|x\|_{\ell^1} \to \min .$$

We are searching for convergence rates with respect to the ℓ^1-norm minimizing solution $x^\dagger$.

Benchmark source conditions and approximate source conditions are not applicable.

The situation of ℓ^1-regularization under consideration

Assumption 4

(a) $x^\dagger \in \ell^1$, **but the sparsity assumption fails**, i.e. $x^\dagger \notin \ell^0$;

(b) $F'(x^\dagger)\, e_k \rightharpoonup 0$ for all $k \in \mathbb{N}$;

(c) $e_k = (F'(x^\dagger))^* f_k$ for some $f_k \in Y^*$ and all $k \in \mathbb{N}$.

Theorem

Under the nonlinearity condition

$$\|F'(x^\dagger)(x-x^\dagger)\|_Y \le \sigma(\|F(x)-F(x^\dagger)\|_Y) \qquad (\&)$$

valid for all $x \in \mathcal{M} := \mathcal{M}^{\|\cdot\|_{\ell^1}}(c)$*, some concave index function* σ *and some* $c > \|x^\dagger\|_{\ell^1}$ *we have a variational inequality*

$$\|x-x^\dagger\|_{\ell^1} \le \|x\|_{\ell^1} - \|x\|_{\ell^1} + \varphi(\|F(x)-F(x^\dagger)\|_Y) \quad \textit{for all} \;\; x \in \mathcal{M}$$

with the concave index function

$$\varphi(t) = 2 \inf_{n \in N} \left(\sum_{k=n+1}^{\infty} |x_k^\dagger| + \left(\sum_{k=1}^{n} \|f_k\|_{Y^*} \right) \sigma(t) \right).$$

Example: Hölder rates

Consider a polynomial decay and growth $\sigma(t) \le K_3\, t^\kappa,\ t > 0$,

$$\sum_{k=n+1}^{\infty} |x_k^\dagger| \le K_1\, n^{-\mu}, \qquad \sum_{k=1}^{n} \|f_k\|_{Y^*} \le K_2\, n^{\nu},$$

with exponents $0 < \kappa \le 1$, $\mu, \nu > 0$ and corresponding constants $K_1, K_2, K_3 > 0$. Then by setting $n^{-\mu} \sim n^\nu t^\kappa$ and hence $n \sim t^{\frac{-\kappa}{\nu+\mu}}$ we obtain the Hölder convergence rates

$$\|x_{\alpha_*}^\delta - x^\dagger\|_{\ell^1} = \mathcal{O}\left(\delta^{\frac{\mu\kappa}{\mu+\nu}}\right) \qquad \text{as} \qquad \delta \to 0$$

whenever the regularization parameter $\alpha_* = \alpha(\delta, y^\delta)$ is chosen according to the (SDP). The best possible rate arises from the limit case $\kappa = 1$ expressing the tangential cone condition.

Example: exponentially decaying solution components

In contrast to the last example we consider now an exponential decay of the solution components

$$\sum_{k=n+1}^{\infty} |x_k^\dagger| \le K_1 \, \exp\left(-n^\gamma\right), \qquad \sum_{k=1}^{n} \|f_k\|_{Y^*} \le K_2 \, n^\nu,$$

with exponents $\gamma, \nu > 0$ and corresponding constants $K_1, K_2 > 0$. For simplicity let $\sigma(t) \le K_3 \, t$, only.

By setting $n^\gamma \sim \log(1/t)$ and hence $\exp\left(-n^\gamma\right) \sim t$ the rate

$$\|x_{\alpha_*}^\delta - x^\dagger\|_{\ell^1} = \mathcal{O}\left(\delta \left(\log\left(\frac{1}{\delta}\right)\right)^{\frac{\nu}{\gamma}}\right) \qquad \text{as} \qquad \delta \to 0$$

holds for α_* from (SDP). The factor $\left(\log\left(\frac{1}{\delta}\right)\right)^{\frac{\nu}{\gamma}}$ prevents

$$\|x_{\alpha_*}^\delta - x^\dagger\|_{\ell^1} = \mathcal{O}\left(\delta\right) \qquad \text{as} \qquad \delta \to 0,$$

the rate which occurs for sparse solutions $x^\dagger \in \ell^0$.

Outline

Relevant references:

▷ M. BURGER, S. OSHER: Convergence rates of convex variational regularization. *Inverse Problems* **20** (2004), pp. 1411–1421.

▷ B. HOFMANN, B. KALTENBACHER, C. PÖSCHL, O. SCHERZER: A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators. *Inverse Problems* **23** (2007), pp. 87–1010.

▷ T. HEIN, B. HOFMANN: Approximate source conditions for nonlinear ill-posed problems – chances and limitations. *Inverse Problems* **25** (2009), 035003 (16pp).

▷ O. SCHERZER, M. GRASMAIR, H. GROSSAUER, M. HALTMEIER, F. LENZEN: *Variational methods in imaging*. Springer Series on Applied Mathematical Sciences, Vol. 167, Springer, New York 2009.

▷ M. GRASMAIR: Generalized Bregman distances and convergence rates for non-convex regularization methods. *Inverse Problems* **26** (2010), 115014 (16pp).

▷ B. HOFMANN, M. YAMAMOTO: On the interplay of source conditions and variational inequalities for nonlinear ill-posed problems. *Applicable Analysis* **89** (2010), pp. 1705–1727.

▷ R. I. BOŢ, B. HOFMANN: An extension of the variational inequality approach for obtaining convergence rates in regularization of nonlinear ill-posed problems. *Journal of Integral Equations and Applications* **22** (2010), pp. 369–392.

▷ T. SCHUSTER, B. KALTENBACHER, B. HOFMANN, K.S. KAZIMIERSKI: *Regularization Methods in Banach Spaces*. Verlag Walter de Gruyter, Berlin/Boston 2012.

Further relevant references:

▷ J. FLEMMING: *Generalized Tikhonov Regularization and Modern Convergence Rate Theory in Banach Spaces*. Shaker Verlag, Aachen, 2012.

▷ B. HOFMANN, P. MATHÉ: Parameter choice in Banach space regularization under variational inequalities. *Inverse Problems* **28** (2012), 104006 (17pp).

▷ M. BURGER, J. FLEMMING, B. HOFMANN: Convergence rates in ℓ^1-regularization if the sparsity assumption fails. *Inverse Problems* **29** (2013), 025013 (16pp).

▷ S. W. ANZENGRUBER, B. HOFMANN, P. MATHÉ: Regularization properties of the sequential discrepancy principle for Tikhonov regularization in Banach spaces. *Applicable Analysis* **93** (2014), pp. 1382–1400.

▷ J. FLEMMING, B. HOFMANN AND I. VESELIĆ: On ℓ^1-regularization in light of Nashed's ill-posedness concept. *Comput. Methods Appl. Math* **15** (2015), pp. 279–289.

▷ D. CHEN, B. HOFMANN, J. ZOU: Elastic-net regularization versus ℓ^1-regularization for linear inverse problems with quasi-sparse solutions. *Inverse Problems* **33** (2017), 015004 (17pp).