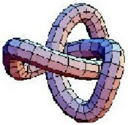
Ergänzungen V zur Vorlesung Inverse Probleme Sommersem. 2017

Nichtlinearitätsbedingungen im Hilbertraum

BERND HOFMANN

TU Chemnitz
Fakultät für Mathematik
D-09107 Chemnitz, GERMANY



Juni 2018

Forschung gefördert durch Deutsche Forschungsgemeinschaft unter HO1454/12-1

Gliederung

Gliederung

1. Einführung
2. Nichtlinearitätsbedingungen
3. Die Selbstfaltungsgleichung im reellen Hilbertraum $L^2(0,1)$

Gliederung

Nonlinear inverse problems

Let X, Y be infinite dimensional separable Hilbert spaces with norms $\|\cdot\|$.

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

We consider an **ill-posed nonlinear** 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 $y = F(x^\dagger)$.

For the stable approximate solution of $(**)$ we consider with noisy data y^δ and a deterministic noise model

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

the classical version of **Tikhonov regularization**

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

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).

Assumption 1

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

Under Assumption 1 there exist **$\bar{\mathbf{x}}$-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 $\overline{x} \in X$.

Under Assumption 1 there exist **regularized solutions** x_α^δ for all $\alpha > 0$ and arbitrary data elements $y^\delta \in Y$.

Assumption 2

- 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 Assumptions 1 and 2 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

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

for an $\overline{x}$-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 $\Psi(R) = d(R)^2/R$ and for the distance function

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

as $R \to \infty$ a convergence rate

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

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

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

Left-side rotation:

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

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:

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

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

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

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**.

Open question: Finding results if x_α^δ is not necessarily a global minimum of the Tikhonov functional T_α^δ, but at least a stationary point with $\nabla T_\alpha^\delta(x_\alpha^\delta) = 0$! Hence, solutions of the Euler equation

$$F'(x_\alpha^\delta)^*(F(x_\alpha^\delta) - y^\delta) + \alpha(x_\alpha^\delta - \overline{x}) = 0$$

are under consideration as regularized solutions.

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

In this section, 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.

For nonlinear operator equations $(**)$ the ill-posedness concept is not as easy as in the linear case, where ill-posedness is equivalent with nonclosed range of the forward operator.

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

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

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

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**

$$\left[F'(x)h\right](s) = 2\int\limits_0^s x(s-t)h(t)dt, \quad 0 \le s \le 1, \quad h \in L^2(0,1).$$

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 1

For the autoconvolution operator F from $(**)$ 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 Proposition 1. 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 Proposition 1.

Proposition 1 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)

Proposition 2

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

Proof: Steven's proof (2 pages A4)

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

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

is satisfied under the smallness condition

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

Proposition 3

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

$$x^\dagger(1-t) - \overline{x}(1-t) = \int\limits_0^t \widetilde{v}(t-s)\, x^\dagger(s)ds, \quad 0 \le t \le 1,$$

or short in convolution form as $x^\dagger - \overline{x} = \widetilde{v} * x^\dagger$. Therefore, the transformation $x(t) \mapsto \overline{x}(t) + \int\limits_t^1 x(s-t)v(s)ds$ 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 $\|x^\dagger - \overline{x}\| \leq \|x^\dagger\| \|v\| < \|x^\dagger\|$, 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.