Convergence Analysis of (Statistical) Inverse Problems under Conditional Stability Estimates

Frank Werner

joint work with Bernd Hofmann

Institut für Mathematische Stochastik

Universität Göttingen

Model

In this talk we consider **nonlinear** inverse problems

$$F(f) = g$$

with $F : D(F) \subset \mathcal{X} \to \mathcal{Y}$, $\mathcal{X}, \mathcal{Y}$ Hilbert spaces in the data model

$$g^{\mathrm{obs}} = g^{\dagger} + \sigma Z + \delta \xi$$

with

- exact data $g^{\dagger} = F\left(f^{\dagger}\right)$
- stochastic noise given by a (centered) **Gaussian white noise** process $Z : \mathcal{Y} \to \mathbf{L}^2(\Omega, \mathcal{A}, \mathbb{P})$
- deterministic noise $\xi \in \mathcal{Y}$, $\|\xi\|_{\mathcal{Y}} \leq 1$
- and corresponding noise levels $\sigma, \delta \geq 0$.

Model (cont')

$$g^{\mathrm{obs}} = g^{\dagger} + \sigma Z + \delta\xi$$

As $Z \notin \mathcal{Y}$, whole model has to be understood in the weak sense:
For each $g \in \mathcal{Y}$ we can access the corresponding *coefficient*

$$\left\langle g^{\mathrm{obs}}, g \right\rangle = \left\langle g^{\dagger}, g \right\rangle + \sigma \left\langle Z, g \right\rangle + \delta \left\langle \xi, g \right\rangle$$

under additive

deterministic noise $\delta \left\langle \xi, g \right\rangle \leq \delta \left\| g \right\|_{\mathcal{Y}}$ and

stochastic noise $\sigma \left\langle Z, g \right\rangle \sim \mathcal{N}\left(0, \sigma^2 \left\| g \right\|_{\mathcal{Y}}^2\right)$.

Note: For $g_1, g_2 \in \mathcal{Y}$ it holds

$$\mathrm{Cov}\left[\left\langle Z, g_1 \right\rangle, \left\langle Z, g_2 \right\rangle\right] = \left\langle g_1, g_2 \right\rangle.$$

Examples

$$g^{\mathrm{obs}} = g^{\dagger} + \sigma Z + \delta \xi$$

allows for

- completely deterministic models: $\sigma = 0$
- continuous Gaussian white noise models: $\delta = 0$
- mixtures of both, e.h. discretized Gaussian white noise models (with ξ being the normalized discretization error)

N. Bissantz, T. Hohage, A. Munk, and F. Ruymgaart.
Convergence rates of general regularization methods for statistical inverse problems and applications.
SIAM Journal on Numerical Analysis 45(6): 2610-2636, 2007

Hilbert scales

We assume that there is a Hilbert scale $\{\mathcal{X}_\nu\}_{\nu\in\mathbb{R}}$ such that

- $\mathcal{X}_\nu := D(L^\nu)$ with a densely defined linear self-adjoint $L : D(L) \subset \mathcal{X} \to \mathcal{X}$
- $\mathcal{X}_0 := \mathcal{X}$
- $\|f\|_\nu := \|L^\nu f\|_{\mathcal{X}}$
- Interpolation for $-a < t \leq s$:

$$\|f\|_t \leq \|f\|_{-a}^{\frac{s-t}{s+a}} \|f\|_s^{\frac{t+a}{s+a}}$$

Regularization

We consider variational regularization of the form

$$\hat{f}_\alpha \in \operatorname*{argmin}_{f \in D(F)} \left[\frac{1}{2} \|F(f)\|_{\mathcal{Y}}^2 - \left\langle F(f), g^{\mathrm{obs}} \right\rangle + \alpha \|f\|_s^2 \right]$$

$$\hat{=} \operatorname*{argmin}_{f \in D(F)} \left[\frac{1}{2} \left\| F(f) - g^{\mathrm{obs}} \right\|_{\mathcal{Y}}^2 + \alpha \|f\|_s^2 \right]$$

Well-known properties of minimizers (if $D(F)$ is closed and convex, and F is weak-to-weak sequentially continuous):

- existence (with probability 1 in the stochastic case)
- stability if $s \geq 0$ (w.r.t. ξ clear, w.r.t. Z more complicated)
- convergence in the deterministic case (if α is chosen appropriately)

Here we will focus on **rates of convergence**!

Assumptions (on F, $f^\dagger$ and the method)

Conditional stability estimate (CSE)

There exists a concave **index function** φ, a set Q and a constant $R > 0$ such that

$$\left\| f - f^\dagger \right\|_{-a} \leq R\varphi\left(\left\| F(f) - g^\dagger \right\|_{\mathcal{Y}} \right)$$

for all $f \in Q$.

Smoothing properties

Furthermore $f^\dagger \in \mathcal{X}_u$ is the unique solution to $F\left(f^\dagger\right) = g^\dagger$ and the indices satisfy

- $a \geq 0$ (smoothing property of F)
- $-a < s < u$ (smoothing of the regularization, undersmoothing case)
- $u \leq 2s + a$ (smoothness of $f^\dagger$, saturation)

Examples

Conditional stability estimates hold true whenever

- F' is ill-posed of degree a (i.e. $\|h\|_{-a} \leq \bar{K}\,\|F'(f^\dagger)\,h\|_{\mathcal{Y}}$) and a tangential-cone-type condition holds true ($\|F'(f^\dagger)(f-f^\dagger)\|_{\mathcal{Y}} \leq \tilde{K}\,\varphi(\|F(f)-F(f^\dagger)\|_{\mathcal{Y}}$)
- F' is ill-posed of degree a and a Hölder continuity condition holds true ($\|F'(f)-F'(f^\dagger)\|_{\mathcal{L}(\mathcal{X},\mathcal{Y})} \leq \check{K}\,\|f-f^\dagger\|^{\eta}$)

$\rightsquigarrow$ by interpolation this leads to $\varphi(t) = t$

But they can also be validated independently of nonlinearity restrictions:

B. Hofmann and M. Yamamoto.
On the interplay of source conditions and variational inequalities for nonlinear ill-posed problems.
Appl. Anal., 89:1705–1727, 2010.

T. Hohage and F. Weidling.
Verification of a variational source condition for acoustic inverse medium scattering problems.
Inverse Problems 31:075006, 2015.

Notation

- For an index function h denote by

$$h^*(y) := \sup_{x \geq 0} [xy - h(x)]$$

  the **Fenchel** conjugate of h.
- Introduce

$$\psi_{u,s,a}(t) := \left(\varphi\left(\sqrt{t}\right)\right)^{\frac{2(u-s)}{a+u}}, \qquad t \geq 0$$

  and

$$\varphi_{\mathrm{app}}(\alpha) := (-\psi_{u,s,a})^* \left(-\frac{1}{\alpha}\right), \qquad \alpha \geq 0$$

Deterministic case ($\sigma = 0$)

Error estimates

Whenever $\hat{f}_\alpha \in Q$ (validity set of CSE), then it holds

$$\left\| \hat{f}_\alpha - f^\dagger \right\|_s^2 \leq \frac{\delta^2}{\alpha} + C\varphi_{\mathrm{app}}\left(8C\alpha\right)$$

with a constant $C = C\left(R, \left\|f^\dagger\right\|_u, u, s, a\right)$.

The function φ_{app} (approximation error) is

- non-negative
- monotonically increasing
- satisfies $\varphi_{\mathrm{app}}\left(C\alpha\right) \leq \max\left\{1, C^{\frac{u-s}{a+s}}\right\}\varphi_{\mathrm{app}}\left(\alpha\right)$ for all $\alpha, C > 0$.
- $\varphi_{\mathrm{app}}\left(\alpha\right) \to 0$ as $\alpha \to 0$

Deterministic case ($\sigma = 0$)

Convergence rates

Assume additionally that $\psi_{u,s,a}$ is concave, and choose $\alpha = \alpha_*$ such that

$$-\frac{1}{\alpha^*} \in \partial \left(-\psi_{u,s,a}\right) \left(\delta^2\right).$$

Then we obtain the convergence rate

$$\left\| \hat{f}_{\alpha_*} - f^\dagger \right\|_s = \mathcal{O}\left(\sqrt{\psi_{u,s,a}\left(\delta^2\right)}\right) = \mathcal{O}\left(\left(\varphi\left(\delta\right)\right)^{\frac{u-s}{a+u}}\right) \qquad \text{as} \qquad \delta \to 0.$$

- If F is linear and φ a power function (Hölder case), then this rate is order optimal.
- A posteriori choice (Lepskiĭ) is also possible.

Smoothing properties of F

In the stochastic case ($\sigma > 0$) we assume additionally:

Interpolation in $\mathcal{Y}$

- There is a Gelfand triple $(\mathcal{V}, \mathcal{Y}, \mathcal{V}')$ s. th. $\iota : \mathcal{V} \hookrightarrow \mathcal{Y}$ is Hilbert-Schmidt.
- F satisfies the interpolation inequality

$$\left\| F(f) - g^\dagger \right\|_{\mathcal{V}} \leq C \left\| F(f) - g^\dagger \right\|_{\mathcal{Y}}^{\theta} \left\| f - f^\dagger \right\|_{s}^{1-\theta}$$

for all $f \in Q$ with some constant $C > 0$ and $\theta \in (0,1)$.

Note: ι being Hilbert-Schmidt implies

$$\mathbb{E}\left[\|Z\|_{\mathcal{V}'}^2\right] = \operatorname{trace}\left(\iota^* \operatorname{Cov}[Z]\, \iota\right) < \infty,$$

i.e. $\|Z\|_{\mathcal{V}'} < \infty$ a.s.

Examples

$$\left\| F(f) - g^\dagger \right\|_{\mathcal{V}} \leq C \left\| F(f) - g^\dagger \right\|_{\mathcal{Y}}^{\theta} \left\| f - f^\dagger \right\|_{s}^{1-\theta}$$

This condition holds true whenever

- $(\mathcal{V}, \mathcal{Y}, \mathcal{V}')$ is part of a Hilbert scale $\{\mathcal{Y}_\mu\}_{\mu \in \mathbb{R}}$ (i.e. $\mathcal{V} = \mathcal{Y}_t$, $\mathcal{V}' = \mathcal{Y}_{-t}$, $\mathcal{Y} = \mathcal{Y}_0$) and
- $F : \mathcal{X}_s \to \mathcal{Y}_r$ is Lipschitz continuous for some $r > t$

Proof:

$$\begin{aligned}\left\| F(f) - g^\dagger \right\|_{\mathcal{V}} &\leq \left\| F(f) - g^\dagger \right\|_{\mathcal{Y}}^{\theta} \left\| F(f) - g^\dagger \right\|_{\mathcal{Y}_r}^{1-\theta} \\ &\leq L^{1-\theta} \left\| F(f) - g^\dagger \right\|_{\mathcal{Y}}^{\theta} \left\| f - f^\dagger \right\|_{s}^{1-\theta}\end{aligned}$$

Classical example: Sobolev spaces $\mathcal{V} = H^s(\Omega)$, $\mathcal{Y} = \mathbf{L}^2(\Omega)$, $s > d/2$.

Stochastic case ($\sigma > 0$)

Error estimates

Whenever $\hat{f}_\alpha \in Q$ (validity set of CSE), then it holds (surely) that

$$\left\| \hat{f}_\alpha - f^\dagger \right\|_s^2 \leq C \left[\sigma^2 \left\| Z \right\|_{\mathcal{V}'}^2 \alpha^{\theta-2} + \frac{\delta^2}{\alpha} + \varphi_{\mathrm{app}} \left(8 C \alpha \right) \right]$$

with a constant $C > 0$.

Note: $\left\| Z \right\|_{\mathcal{V}'}$ is an a.s. finite random variable. If the unit ball in $\mathcal{V}$ has a countable dense subset w.r.t. $\left\| \cdot \right\|_{\mathcal{Y}}$ it can be proven that

- $\left\| Z \right\|_{\mathcal{V}'}$ is sub-Gaussian and hence
- $\mathbb{P}\left[\left| \left\| Z \right\|_{\mathcal{V}'} - \mathbf{E}\left[\left\| Z \right\|_{\mathcal{V}'} \right] \right| \geq t \right] \leq 2 \exp\left(-\frac{2t^2}{\pi^2 \left\| \iota \right\|^2} \right)$ for all $t \geq 0$.

Stochastic case ($\sigma > 0$)

Convergence rates

Let

$$\Sigma(\alpha) = \sqrt{\alpha}\sqrt{\varphi_{\mathrm{app}}(\alpha)} \qquad \text{and} \qquad \widetilde{\Sigma}(\alpha) = \alpha^{1-\frac{\theta}{2}}\sqrt{\varphi_{\mathrm{app}}(\alpha)}, \quad \alpha > 0$$

and choose α such that

$$\alpha \sim \left(\Sigma^{-1}(\delta) + \widetilde{\Sigma}^{-1}(\sigma)\right) \qquad \text{as} \qquad \max\{\delta, \sigma\} \to 0.$$

Then we obtain the a.s. convergence rate

$$\left\|\hat{f}_\alpha - f^\dagger\right\|_s = \mathcal{O}\left(\sqrt{\varphi_{\mathrm{app}}\left(\Sigma^{-1}(\delta) + \widetilde{\Sigma}^{-1}(\sigma)\right)}\right)$$

as $\max\{\delta, \sigma\} \to 0$.

Conclusion and outlook

- Statistical Inverse Problems differ from deterministic ones ...
  - ... by the fact, that the data is not an element of the space $\mathcal{Y}$.
  - ... and hence regularization has to be treated differently.
- Convergence analysis under conditional stability estimates ...
  - ... avoids nonlinearity assumptions on the operator.
  - ... can be carried out both in the deterministic and stochastic case.
- Future research should address ...
  - ... validity of conditional stability estimates ($\rightsquigarrow$ Q and φ).
  - ... what can be done to ensure $\hat{f}_\alpha \in Q$ under weaker conditions.

F. Werner and B. Hofmann.
Convergence Analysis of (Statistical) Inverse Problems under Conditional Stability Estimates.
arXiv preprint 1905.09765, 2019.

Thank you for your attention!