Lecture 2: Tikhonov-Regularization

Bastian von Harrach

harrach@math.uni-stuttgart.de

Chair of Optimization and Inverse Problems, University of Stuttgart, Germany

Advanced Instructional School on
Theoretical and Numerical Aspects of Inverse Problems
TIFR Centre For Applicable Mathematics
Bangalore, India, June 16–28, 2014.

Setting

Let

- X and Y be Hilbert spaces
- $A \in \mathcal{L}(X, Y)$, i.e., $A: \; X \to Y$ is linear and continuous
- A be injective, i.e., there exists left inverse

$$A^{-1}: \; \mathcal{R}(A) \subseteq Y \to X.$$

- A^{-1} linear but possibly discontinuous (unbounded)

Setting

Consider

- true solution $\hat{x} \in X$,
- exact measurements $\hat{y} = A\hat{x} \in Y$,
- noisy measurements $y^\delta \in Y$, $\left\|y^\delta - \hat{y}\right\| \leq \delta$.

Given exact measurements $\hat{y}$, we could calculate $\hat{x} = A^{-1}\hat{y}$.
But how can we approximate $\hat{x}$ from noisy measurements y^δ?

Problems:

- $A^{-1}y^\delta$ may not be well-defined (for $y^\delta \notin \mathcal{R}(A)$)
- A^{-1} discontinuous $\rightsquigarrow A^{-1}y^\delta \not\to A^{-1}\hat{y} = \hat{x}$ for $\delta \to 0$.
- A^{-1} unbounded. Possibly, $\left\|A^{-1}y^\delta\right\|_X \to \infty$

Goal

Can we approximate $\hat{x} = A^{-1}\hat{y}$ from noisy measurements $y^\delta \approx \hat{y}$?

Goal: Find reconstruction function $R(y^\delta, \delta)$ so that

$$R(y^\delta, \delta) \to A^{-1}\hat{y} = \hat{x} \quad \text{for } \delta \to 0.$$

Note: $R(y^\delta, \delta) := A^{-1}y^\delta$ does not work!

In this lecture: $R(y^\delta, \delta) := (A^*A + \delta I)^{-1}A^*y^\delta$ works, i.e.,

$$(A^*A + \delta I)^{-1}A^*y^\delta \to A^{-1}\hat{y} = \hat{x}.$$

Motivation

Can we approximate $\hat{x} = A^{-1}\hat{y}$ from noisy measurements $y^\delta \approx \hat{y}$?

- Standard approach ($\rightsquigarrow x^\delta = A^{-1}y^\delta$)

  Minimize data-fit (residuum) $\left\| Ax - y^\delta \right\|_Y$!

- Tikhonov regularization:

  Minimize $\left\| Ax - y^\delta \right\|_Y^2 + \alpha \left\| x \right\|_X^2$!

  with regularization parameter $\alpha > 0$

  α small $\rightsquigarrow$ solution will fit measurements well,
  α large $\rightsquigarrow$ solution will be regular (small norm).

Motivation

Tikhonov regularization:

$$\text{Minimize} \quad \left\| Ax - y^\delta \right\|_Y^2 + \alpha \left\| x \right\|_X^2 \;!$$

Equivalent formulation:

$$\text{Minimize} \quad \left\| \begin{pmatrix} A \\ \sqrt{\alpha} I \end{pmatrix} x - \begin{pmatrix} y^\delta \\ 0 \end{pmatrix} \right\|_{X \times Y}^2 = \left\| \begin{pmatrix} Ax - y^\delta \\ \sqrt{\alpha} x \end{pmatrix} \right\|_{X \times Y}^2 \;!$$

Formal use of normal equations yields

$$\begin{pmatrix} A^* & \sqrt{\alpha} I \end{pmatrix} \begin{pmatrix} A \\ \sqrt{\alpha} I \end{pmatrix} x = \begin{pmatrix} A^* & \sqrt{\alpha} I \end{pmatrix} \begin{pmatrix} y^\delta \\ 0 \end{pmatrix},$$

and thus

$$(A^* A + \alpha I) x = A^* y^\delta.$$

Invertibility of $A^*A + \alpha I$

Theorem 2.1. Let $A \in \mathcal{L}(X, Y)$. For each $\alpha > 0$, the operators

$$A^*A + \alpha I \in \mathcal{L}(X) \quad \text{and} \quad AA^* + \alpha I \in \mathcal{L}(Y)$$

are continuously invertible and they fulfill

$$\left\|(A^*A + \alpha I)^{-1}\right\|_{\mathcal{L}(X)} \leq \frac{1}{\alpha}, \quad \left\|(AA^* + \alpha I)^{-1}\right\|_{\mathcal{L}(Y)} \leq \frac{1}{\alpha}.$$

Minimizer of Tikhonov Functional

Theorem 2.2. Let $A \in \mathcal{L}(X, Y)$. $x_\alpha^\delta := (A^*A + \alpha I)^{-1} A^* y^\delta$ is the unique minimizer of the Tikhonov functional

$$J_\alpha(x) := \left\| Ax - y^\delta \right\|_Y^2 + \alpha \left\| x \right\|_X^2$$

An auxiliary result

Theorem 2.3. Let $A \in \mathcal{L}(X, Y)$ be injective.

(a) $\mathcal{R}(A^*A)$ is dense in X

(b) If a sequence $(x_k)_{k\in\mathbb{N}} \subset X$ fulfills

$$A^*Ax_k \to A^*Ax \quad \text{and} \quad \|x_k\|_X \leq \|x\|_X$$

then $x_k \to x$.

Convergence of Tikhonov regularization

Theorem 2.4. Let

- $A \in \mathcal{L}(X, Y)$ be injective (with a possibly unbounded inverse),
- $A\hat{x} = \hat{y}$
- $(y^\delta)_{\delta>0} \subseteq Y$ be noisy measurements with $\left\| y^\delta - \hat{y} \right\|_Y \leq \delta$.

If we choose the regularization parameter so that

$$\alpha(\delta) \to 0 \quad \text{and} \quad \frac{\delta^2}{\alpha(\delta)} \to 0,$$

then

$$(A^*A + \alpha I)^{-1} A^* y^\delta \to \hat{x} \quad \text{for } \delta \to 0.$$

Image deblurring

$\hat{x}$

$\hat{y} = A\hat{x}$

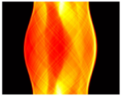
y^δ

$A^{-1}y^\delta$

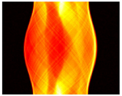
$(A^*A + \delta I)^{-1}A^*y^\delta$

Computerized tomography

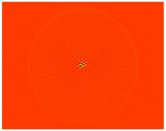
$\hat{x}$



$\hat{y} = A\hat{x}$



y^δ



$A^{-1}y^\delta$

$(A^*A + \delta I)^{-1}A^*y^\delta$

Conclusions and remarks

Conclusions

- For ill-posed inverse problems, the best data-fit solutions generally do not converge against the true solution.
- The regularized solutions do converge against the true solution.

More sophisticated parameter choice rule

- Discrepancy principle: Choose α such that $\left\| Ax_\alpha^\delta - y^\delta \right\| \approx \delta$

Strategies for non-linear inverse problems $F(x) = y$:

- First linearize, then regularize.
- First regularize, then linearize.

A-priori information

- Regularization can be used to incorporate a-priori knowledge (promote sparsity or sharp edges, include stochastic priors, etc.)