Regularization and Inverse Problems

Caroline Sieger

Host Institution: Universität Bremen

Home Institution: Clemson University

August 5, 2009

Outline

1. Preliminaries
   - The Inverse Problem
   - The Moore-Penrose Generalized Inverse
   - Eigensystems and Singular Systems
2. Regularization
   - Classical Tikhonov Regularization
   - Tikhonov Regularization with Sparsity Constraints

Preliminaries

Notation:

$T : \mathcal{X} \to \mathcal{Y}$ is a bounded linear operator
$K : \mathcal{X} \to \mathcal{Y}$ is a compact linear operator
$\mathcal{X}$ and $\mathcal{Y}$ are Hilbert spaces

Definition

A problem is **ill-posed** if one or more of the following holds:

- a solution does not exist

Preliminaries

Notation:

$T : \mathcal{X} \to \mathcal{Y}$ is a bounded linear operator

$K : \mathcal{X} \to \mathcal{Y}$ is a compact linear operator

$\mathcal{X}$ and $\mathcal{Y}$ are Hilbert spaces

Definition

A problem is **ill-posed** if one or more of the following holds:

- a solution does not exist
- the solution is not unique

Zentrum für Technomathematik

Preliminaries

Notation:

$T : \mathcal{X} \to \mathcal{Y}$ is a bounded linear operator

$K : \mathcal{X} \to \mathcal{Y}$ is a compact linear operator

$\mathcal{X}$ and $\mathcal{Y}$ are Hilbert spaces

Definition

A problem is **ill-posed** if one or more of the following holds:

- a solution does not exist
- the solution is not unique
- the solultion does not depend continuously on the data

The Inverse Problem

Find $x \in \mathcal{X}$ such that $Tx = y$.

The Inverse Problem

Find $x \in \mathcal{X}$ such that $Tx = y$.

There may be some problems with this, i.e.

- T^{-1} may not exist
- $NS(T) \neq \{0\}$ ($\implies$ non-unique solutions)

Thus, we seek an approximate solution.

The Inverse Problem

Find $x \in \mathcal{X}$ such that $Tx = y$.

There may be some problems with this, i.e.

- T^{-1} may not exist
- $NS(T) \neq \{0\}$ ($\implies$ non-unique solutions)

Thus, we seek an approximate solution.

Types of Solutions

- $x \in \mathcal{X}$ is a **least-squares solution** to $Tx = y$ if $||Tx - y|| = \inf\{||Tz - y|| \mid z \in \mathcal{X}\}$

Zentrum für Technomathematik

The Inverse Problem

Find $x \in \mathcal{X}$ such that $Tx = y$.

There may be some problems with this, i.e.

- T^{-1} may not exist
- $NS(T) \neq \{0\}$ ($\implies$ non-unique solutions)

Thus, we seek an approximate solution.

Types of Solutions

- $x \in \mathcal{X}$ is a **least-squares solution** to $Tx = y$ if
  $||Tx - y|| = \inf\{||Tz - y|| \mid z \in \mathcal{X}\}$
- $x \in \mathcal{X}$ is a **best-approximate solution** to $Tx = y$ if x is a least-squares solution and
  $||x|| = \inf\{||z|| \mid z \text{ is a least-squares solution of } Tx = y\}$

Zentrum für Technomathematik

The Moore-Penrose Generalized Inverse

Definition: Moore-Penrose Generalized Inverse

$T^\dagger$ is the unique linear extension of $\tilde{T}^{-1}$ to $dom(T^\dagger)$ with $NS(T^\dagger) = range(T)^\perp$ where $\tilde{T} := T\big|_{NS(T)^\perp} : NS(T)^\perp \to range(T)$.

Zentrum für Technomathematik

The Moore-Penrose Generalized Inverse

Definition: Moore-Penrose Generalized Inverse

$T^\dagger$ is the unique linear extension of $\tilde{T}^{-1}$ to $dom(T^\dagger)$ with $NS(T^\dagger) = range(T)^\perp$ where $\tilde{T} := T\big|_{NS(T)^\perp} : NS(T)^\perp \to range(T)$.

- If P and Q are orthogonal projectors onto $NS(T)$ and $\overline{range(T)}$

1.) $TT^\dagger T = T$ 2.) $T^\dagger T T^\dagger = T^\dagger$

3.) $T^\dagger T = I - P$ 4.) $TT^\dagger = Q\big|_{dom(T^\dagger)}$

Zentrum für Technomathematik

The Moore-Penrose Generalized Inverse

Definition: Moore-Penrose Generalized Inverse

$T^\dagger$ is the unique linear extension of $\tilde{T}^{-1}$ to $dom(T^\dagger)$ with $NS(T^\dagger) = range(T)^\perp$ where $\tilde{T} := T\big|_{NS(T)^\perp} : NS(T)^\perp \to range(T)$.

- If P and Q are orthogonal projectors onto $NS(T)$ and $\overline{range(T)}$

$$1.)\ TT^\dagger T = T \qquad 2.)\ T^\dagger T T^\dagger = T^\dagger$$

$$3.)\ T^\dagger T = I - P \qquad 4.)\ TT^\dagger = Q\big|_{dom(T^\dagger)}$$

- $T^\dagger$ is bounded if $range(T)$ is closed.

Zentrum für Technomathematik

The Moore-Penrose Generalized Inverse

Definition: Moore-Penrose Generalized Inverse

$T^\dagger$ is the unique linear extension of $\tilde{T}^{-1}$ to $dom(T^\dagger)$ with $NS(T^\dagger) = range(T)^\perp$ where $\tilde{T} := T\big|_{NS(T)^\perp} : NS(T)^\perp \to range(T)$.

- If P and Q are orthogonal projectors onto $NS(T)$ and $\overline{range(T)}$

1.) $TT^\dagger T = T$ 2.) $T^\dagger TT^\dagger = T^\dagger$

3.) $T^\dagger T = I - P$ 4.) $TT^\dagger = Q\big|_{dom(T^\dagger)}$

- $T^\dagger$ is bounded if $range(T)$ is closed.
- If $y \in dom(T^\dagger)$ then $\exists$ a unique best-approximate solution, $x^\dagger := T^\dagger y$

The Moore-Penrose Generalized Inverse

Definition: Moore-Penrose Generalized Inverse

$T^\dagger$ is the unique linear extension of $\tilde{T}^{-1}$ to $dom(T^\dagger)$ with $NS(T^\dagger) = range(T)^\perp$ where $\tilde{T} := T\big|_{NS(T)^\perp} : NS(T)^\perp \to range(T)$.

- If P and Q are orthogonal projectors onto $NS(T)$ and $\overline{range(T)}$

$$1.)\ \ TT^\dagger T = T \qquad 2.)\ \ T^\dagger TT^\dagger = T^\dagger$$
$$3.)\ \ T^\dagger T = I - P \qquad 4.)\ \ TT^\dagger = Q\big|_{dom(T^\dagger)}$$

- $T^\dagger$ is bounded if $range(T)$ is closed.
- If $y \in dom(T^\dagger)$ then $\exists$ a unique best-approximate solution, $x^\dagger := T^\dagger y$

Definition: Gaussian Normal Equation

For $y \in dom(T^\dagger)$, $x \in \mathcal{X}$ is a least-squares solution $\Leftrightarrow$ $T^*Tx = T^*y$.

Eigensystems and Singular Systems

Definition

A selfadjoint K has the **eigensystem** $(\lambda_n; v_n)$ where the λ_n are non-zero eigenvalues and the v_n are corresponding eigenvectors.

We may diagonalize K by $Kx = \sum_{n=1}^{\infty} \lambda_n \langle x, v_n \rangle v_n$ for all $x \in \mathcal{X}$.

Eigensystems and Singular Systems

Definition

A selfadjoint K has the **eigensystem** $(\lambda_n; v_n)$ where the λ_n are non-zero eigenvalues and the v_n are corresponding eigenvectors.

We may diagonalize K by $Kx = \sum_{n=1}^{\infty} \lambda_n \langle x, v_n \rangle v_n$ for all $x \in \mathcal{X}$.

Definition

A non-selfadjoint K has the **singular system** $(\sigma_n; v_n, u_n)$ where

- K^* is the adjoint of K
- $\{\sigma_n^2\}_{n \in \mathbb{N}}$ are the non-zero eigenvalues of K^*K (and KK^*)

Zentrum für Technomathematik

Eigensystems and Singular Systems

Definition

A selfadjoint K has the **eigensystem** $(\lambda_n; v_n)$ where the λ_n are non-zero eigenvalues and the v_n are corresponding eigenvectors.

We may diagonalize K by $Kx = \sum_{n=1}^{\infty} \lambda_n \langle x, v_n \rangle v_n$ for all $x \in \mathcal{X}$.

Definition

A non-selfadjoint K has the **singular system** $(\sigma_n; v_n, u_n)$ where

- K^* is the adjoint of K
- $\{\sigma_n^2\}_{n \in \mathbb{N}}$ are the non-zero eigenvalues of K^*K (and KK^*)
- $\{v_n\}_{n \in \mathbb{N}}$ are eigenvectors of K^*K
- $\{u_n\}_{n \in \mathbb{N}}$ are eigenvectors of KK^* defined by $u_n := \frac{Kv_n}{||Kv_n||}$

Eigensystems and Singular Systems

Properties of a singular system:

- $Kv_n = \sigma_n u_n$
- $K^* u_n = \sigma_n v_n$
- $Kx = \sum_{n=1}^{\infty} \sigma_n \langle x, v_n \rangle u_n$
- $Ky = \sum_{n=1}^{\infty} \sigma_n \langle y, u_n \rangle v_n$

Zentrum für Technomathematik

Eigensystems and Singular Systems

Properties of a singular system:

- $Kv_n = \sigma_n u_n$
- $K^* u_n = \sigma_n v_n$
- $Kx = \sum_{n=1}^{\infty} \sigma_n \langle x, v_n \rangle u_n$
- $Ky = \sum_{n=1}^{\infty} \sigma_n \langle y, u_n \rangle v_n$

Further, iff K has finite dimensional range

$\implies$ K has finitely many singular values

$\implies$ infinite series involving singular values degenerate to finite sums.

Eigensystems and Singular Systems

Properties of a singular system:

- $Kv_n = \sigma_n u_n$
- $K^* u_n = \sigma_n v_n$
- $Kx = \sum_{n=1}^{\infty} \sigma_n \langle x, v_n \rangle u_n$
- $Ky = \sum_{n=1}^{\infty} \sigma_n \langle y, u_n \rangle v_n$

Further, iff K has finite dimensional range

$\implies$ K has finitely many singular values

$\implies$ infinite series involving singular values degenerate to finite sums.

Theorem (Engl,Hanke,Neubauer)

For compact linear operator K with singular system $(\sigma_n; v_n, u_n)$ and $y \in \mathcal{Y}$ we have:

1. $y \in dom(K^\dagger) \Leftrightarrow \sum_{n=1}^{\infty} \frac{|\langle y, u_n \rangle|^2}{\sigma_n^2} < \infty$ *(Picard Criterion for existence of a best-approximate solution.)*
2. *For $y \in dom(K^\dagger)$, $K^\dagger y = \sum_{n=1}^{\infty} \frac{\langle y, u_n \rangle}{\sigma_n} v_n$*

Regularization

- Recall: we seek to approximate $x^\dagger := T^\dagger y$

Regularization

- Recall: we seek to approximate $x^\dagger := T^\dagger y$
- Normally, we only have an approximation of y, i.e. y^δ such that $||y^\delta - y|| \leq \delta$
- in the ill-posed case, $T^\dagger y^\delta$ is not a good approximation of $x^\dagger$ because of the unboundedness of $T^\dagger$.

Zentrum für Technomathematik

Regularization

- Recall: we seek to approximate $x^\dagger := T^\dagger y$
- Normally, we only have an approximation of y, i.e. y^δ such that $||y^\delta - y|| \leq \delta$
- in the ill-posed case, $T^\dagger y^\delta$ is not a good approximation of $x^\dagger$ because of the unboundedness of $T^\dagger$.
- We seek an approximation x_α^δ of $x^\dagger$ such that
  1. x_α^δ depends continuously on the noisy data y^δ (this allows stable computation of x_α^δ)
  2. the noise level $\delta \to 0$ and for appropriate α, $x_\alpha^\delta \to x^\delta$

Regularization

Definition

Let $\alpha_0 \in (0, \infty]$ then $\forall \alpha \in (0, \alpha_0]$ let $R_\alpha : \mathcal{Y} \to \mathcal{X}$ be a continuous operator.

Regularization

Definition

Let $\alpha_0 \in (0, \infty]$ then $\forall \alpha \in (0, \alpha_0]$ let $R_\alpha : \mathcal{Y} \to \mathcal{X}$ be a continuous operator. The family $\{R_\alpha\}$ is a **regularization** iff $\forall y \in dom(T^\dagger)$ $\exists$ a parameter choice rule $\alpha = \alpha(\delta, y^\delta)$ such that

$$\lim_{\delta \to 0} \sup \left\{ ||R_{\alpha(\delta, y^\delta)} y^\delta - T^\dagger y|| \mid y^\delta \in \mathcal{Y},\ ||y^\delta - y|| \leq \delta \right\} = 0 \qquad (1)$$

Regularization

Definition

Let $\alpha_0 \in (0, \infty]$ then $\forall \alpha \in (0, \alpha_0]$ let $R_\alpha : \mathcal{Y} \to \mathcal{X}$ be a continuous operator. The family $\{R_\alpha\}$ is a **regularization** iff $\forall y \in dom(T^\dagger)$ $\exists$ a parameter choice rule $\alpha = \alpha(\delta, y^\delta)$ such that

$$\lim_{\delta \to 0} \sup \left\{ ||R_{\alpha(\delta, y^\delta)} y^\delta - T^\dagger y|| \mid y^\delta \in \mathcal{Y},\ ||y^\delta - y|| \leq \delta \right\} = 0 \qquad (1)$$

and $\alpha : \mathbb{R}^+ \times \mathcal{Y} \to (0, \alpha_0)$ such that

$$\lim_{\delta \to 0} \sup \left\{ \alpha(\delta, y^\delta) \mid y^\delta \in \mathcal{Y},\ ||y^\delta - y|| \leq \delta \right\} = 0 \qquad (2)$$

Zentrum für Technomathematik

Regularization

Definition

Let $\alpha_0 \in (0, \infty]$ then $\forall \alpha \in (0, \alpha_0]$ let $R_\alpha : \mathcal{Y} \to \mathcal{X}$ be a continuous operator. The family $\{R_\alpha\}$ is a **regularization** iff $\forall y \in dom(T^\dagger)$ $\exists$ a parameter choice rule $\alpha = \alpha(\delta, y^\delta)$ such that

$$\lim_{\delta \to 0} \sup \left\{ ||R_{\alpha(\delta,y^\delta)} y^\delta - T^\dagger y|| \mid y^\delta \in \mathcal{Y},\ ||y^\delta - y|| \leq \delta \right\} = 0 \tag{1}$$

and $\alpha : \mathbb{R}^+ \times \mathcal{Y} \to (0, \alpha_0)$ such that

$$\lim_{\delta \to 0} \sup \left\{ \alpha(\delta, y^\delta) \mid y^\delta \in \mathcal{Y},\ ||y^\delta - y|| \leq \delta \right\} = 0 \tag{2}$$

For a specific $y \in dom(T^\dagger)$, a pair (R_α, α) is a **regularization method** if (1) and (2) hold.

Zentrum für Technomathematik

Regularization

- If the parameter choice rule does not depend on y^δ, we say it is **a-priori** and we write $\alpha = \alpha(\delta)$.
- Otherwise, it is an **a-posteriori** parameter choice rule.

Regularization

- If the parameter choice rule does not depend on y^δ, we say it is **a-priori** and we write $\alpha = \alpha(\delta)$.
- Otherwise, it is an **a-posteriori** parameter choice rule.

Proposition (Engl,Hanke,Neubauer)

Further, (R_α, α) is convergent (for linear R_α) iff $\lim_{\delta \to 0} \alpha(\delta) = 0$ and $\lim_{\delta \to 0} \delta ||R_{\alpha(\delta)}|| = 0$.

Zentrum für Technomathematik

Regularization

- If the parameter choice rule does not depend on y^δ, we say it is **a-priori** and we write $\alpha = \alpha(\delta)$.
- Otherwise, it is an **a-posteriori** parameter choice rule.

Proposition (Engl,Hanke,Neubauer)

Further, (R_α, α) is convergent (for linear R_α) iff $\lim\limits_{\delta \to 0} \alpha(\delta) = 0$ and $\lim\limits_{\delta \to 0} \delta ||R_{\alpha(\delta)}|| = 0$.

$$\mathcal{X}_{\mu,\rho} := \{x \in \mathcal{X} \mid x = (T^*T)^\mu \omega,\ ||\omega|| \leq \rho\},\ \mu > 0$$

Proposition (Engl,Hanke,Neubauer)

If *range*(T) is non-closed, a regularization algorithm cannot converge to zero faster than $\delta^{\frac{2\mu}{2\mu+1}} \rho^{\frac{1}{2\mu+1}}$ as $\delta \to 0$ for $x^\dagger \in \mathcal{X}_{\mu,\rho}$.

Regularization

$\Delta(\delta, \mathcal{M}, R_\alpha) := \sup \left\{ ||R_\alpha y^\delta - x|| \mid x \in \mathcal{M}, y^\delta \in \mathcal{Y}, ||Tx - y^\delta|| \leq \delta \right\}$ for some $\mathcal{M} \subset \mathcal{X}$

Regularization

$\Delta(\delta, \mathcal{M}, R_\alpha) := \sup \left\{ ||R_\alpha y^\delta - x|| \mid x \in \mathcal{M}, y^\delta \in \mathcal{Y}, ||Tx - y^\delta|| \leq \delta \right\}$ for some $\mathcal{M} \subset \mathcal{X}$

Definition

We say (R_α, α) is **optimal** in $\mathcal{X}_{\mu,\rho}$ if $\Delta(\delta, \mathcal{X}_{\mu,\rho}, R_\alpha) = \delta^{\frac{2\mu}{2\mu+1}} \rho^{\frac{1}{2\mu+1}}$ holds $\forall\, \delta > 0$.

Zentrum für Technomathematik

Regularization

$\Delta(\delta, \mathcal{M}, R_\alpha) := \sup \left\{ ||R_\alpha y^\delta - x|| \mid x \in \mathcal{M}, y^\delta \in \mathcal{Y}, ||Tx - y^\delta|| \leq \delta \right\}$ for some $\mathcal{M} \subset \mathcal{X}$

Definition

We say (R_α, α) is **optimal** in $\mathcal{X}_{\mu,\rho}$ if $\Delta(\delta, \mathcal{X}_{\mu,\rho}, R_\alpha) = \delta^{\frac{2\mu}{2\mu+1}} \rho^{\frac{1}{2\mu+1}}$ holds $\forall\, \delta > 0$.

(R_α, α) is of **optimal order** in $\mathcal{X}_{\mu,\rho}$ if $\exists\, C \geq 1$ such that

$\Delta(\delta, \mathcal{X}_{\mu,\rho}, R_\alpha) \leq C\delta^{\frac{2\mu}{2\mu+1}} \rho^{\frac{1}{2\mu+1}} \quad \forall\, \delta > 0$

Regularization

$\Delta(\delta, \mathcal{M}, R_\alpha) := \sup\left\{||R_\alpha y^\delta - x|| \mid x \in \mathcal{M}, y^\delta \in \mathcal{Y}, ||Tx - y^\delta|| \leq \delta\right\}$ for some $\mathcal{M} \subset \mathcal{X}$

Definition

We say (R_α, α) is **optimal** in $\mathcal{X}_{\mu,\rho}$ if $\Delta(\delta, \mathcal{X}_{\mu,\rho}, R_\alpha) = \delta^{\frac{2\mu}{2\mu+1}} \rho^{\frac{1}{2\mu+1}}$ holds $\forall\ \delta > 0$.
(R_α, α) is of **optimal order** in $\mathcal{X}_{\mu,\rho}$ if $\exists\ C \geq 1$ such that
$\Delta(\delta, \mathcal{X}_{\mu,\rho}, R_\alpha) \leq C\delta^{\frac{2\mu}{2\mu+1}} \rho^{\frac{1}{2\mu+1}} \quad \forall\ \delta > 0$

Theorem (Engl,Hanke,Neubauer)

Let $\tau > \tau_0 \geq 1$, then if (R_α, α_τ) is of optimal order in $\mathcal{X}_{\mu,\rho}$ for some $\mu > 0$ and $\forall \rho > 0$ then all $(R_\alpha, \alpha_{\tilde{\tau}})$ with $\tilde{\tau} > \tau_0$ are convergent for $y \in range(T)$ and of optimal order $\forall \mathcal{X}_{\nu,\rho}$ with $0 < \nu \leq \mu$ and $\rho > 0$.

Classical Tikhonov Regularization

The Tikhonov Functional

$\Phi(x) := ||Tx - y^\delta||^2 + \alpha||x||^2$

Theorem (Engl,Hanke,Neubauer)

$x_\alpha^\delta := (T^*T + \alpha I)^{-1}T^*y^\delta$ *is the unique minimizer of* $\Phi(x)$.

Proof

Zentrum für Technomathematik

Classical Tikhonov Regularization

The Tikhonov Functional

$\Phi(x) := ||Tx - y^\delta||^2 + \alpha||x||^2$

Theorem (Engl,Hanke,Neubauer)

$x_\alpha^\delta := (T^*T + \alpha I)^{-1}T^*y^\delta$ *is the unique minimizer of* $\Phi(x)$.

Proof

Theorem (Engl,Hanke,Neubauer)

For $x_\alpha^\delta := (T^*T + \alpha I)^{-1}T^*y^\delta$, $y \in range(T)$, $||y - y^\delta|| \leq \delta$ *if* $\alpha = \alpha(\delta)$ *such that* $\lim\limits_{\delta \to 0} \alpha(\delta) = 0$ *and* $\lim\limits_{\delta \to 0} \frac{\delta^2}{\alpha(\delta)} = 0$ *then* $\lim\limits_{\delta \to 0} x_{\alpha(\delta)}^\delta = T^\dagger y$.

Zentrum für Technomathematik

Classical Tikhonov Regularization

Proposition (Engl,Hanke,Neubauer)

As long as $\mu \leq 1$, Tikhonov regularization with the a-priori choice rule $\alpha \sim \left(\frac{\delta}{\rho}\right)^{\frac{2}{2\mu+1}}$ is of optimal order in $\mathcal{X}_{\mu,\rho}$, the best possible convergence rate for $\mu = 1$ is:

$$\alpha \sim \left(\frac{\delta}{\rho}\right)^{\frac{2}{3}} \quad \text{and} \quad ||x_\alpha^\delta - x^\dagger|| = O\left(\delta^{\frac{2}{3}}\right)$$

Tikhonov Regularization with Sparsity Constraints

- In some applications, we require a sparse solution
- So use an ℓ^p norm of the coefficients of x wrt an orthonormal basis $\{\varphi_i\}_{i\in\mathbb{N}}$ with $1 \leq p \leq 2$

Zentrum für Technomathematik

Tikhonov Regularization with Sparsity Constraints

- In some applications, we require a sparse solution
- So use an ℓ^p norm of the coefficients of x wrt an orthonormal basis $\{\varphi_i\}_{i\in\mathbb{N}}$ with $1 \le p \le 2$
- Decreasing p from 2 to 1, we increase the penalty on "small" coefficients *and* decrease the penalty on "large" coefficients,
- i.e. we encourage representations with large, but fewer coefficients wrt the orthonormal basis $\{\varphi_i\}_{i\in\mathbb{N}}$

Zentrum für Technomathematik

Tikhonov Regularization with Sparsity Constraints

- In some applications, we require a sparse solution
- So use an ℓ^p norm of the coefficients of x wrt an orthonormal basis $\{\varphi_i\}_{i\in\mathbb{N}}$ with $1 \leq p \leq 2$
- Decreasing p from 2 to 1, we increase the penalty on "small" coefficients *and* decrease the penalty on "large" coefficients,
- i.e. we encourage representations with large, but fewer coefficients wrt the orthonormal basis $\{\varphi_i\}_{i\in\mathbb{N}}$
- Generally, we restrict ourselves to $p \geq 1$ since the functional is no longer convex for $p < 1$.

Tikhonov Regularization with Sparsity Constraints

- In some applications, we require a sparse solution
- So use an ℓ^p norm of the coefficients of x wrt an orthonormal basis $\{\varphi_i\}_{i\in\mathbb{N}}$ with $1 \leq p \leq 2$
- Decreasing p from 2 to 1, we increase the penalty on "small" coefficients *and* decrease the penalty on "large" coefficients,
- i.e. we encourage representations with large, but fewer coefficients wrt the orthonormal basis $\{\varphi_i\}_{i\in\mathbb{N}}$
- Generally, we restrict ourselves to $p \geq 1$ since the functional is no longer convex for $p < 1$.

The Tikhonov Functional with ℓ^1 Penalty

$\Psi(x) := ||Tx - y^\delta||^2 + \alpha \sum_i |\langle \varphi_i, x\rangle|$

Denote x_α^δ to be the minimizer of $\Psi(x)$.

Tikhonov Regularization with Sparsity Constraints

Definition

For $x^\dagger \in \mathcal{X}$ and a regularization R_α, $x^\dagger$ is an **$\mathbf{R}_\alpha$-minimizing solution** if $Tx^\dagger = y$ and $R_\alpha(x^\dagger) = \min\{R_\alpha(x) \mid Tx = y\}$.

Tikhonov Regularization with Sparsity Constraints

Definition

For $x^\dagger \in \mathcal{X}$ and a regularization R_α, $x^\dagger$ is an **$\mathbf{R}_\alpha$-minimizing solution** if $Tx^\dagger = y$ and $R_\alpha(x^\dagger) = \min\{R_\alpha(x) \mid Tx = y\}$.

Recall, $R_\alpha(x) := \alpha \sum_i |\langle \phi_i, x \rangle|$

Tikhonov Regularization with Sparsity Constraints

Definition

For $x^\dagger \in \mathcal{X}$ and a regularization R_α, $x^\dagger$ is an **$\mathbf{R}_\alpha$-minimizing solution** if $Tx^\dagger = y$ and $R_\alpha(x^\dagger) = \min\{R_\alpha(x) \mid Tx = y\}$.

Recall, $R_\alpha(x) := \alpha \sum_i |\langle \phi_i, x \rangle|$

Proposition (Grasmair,Haltmeier,Scherzer)

Assume that T is injective (or finite basis injectivity holds), then $\exists$ a unique minimizer x_α^δ of $\Psi(x)$ and $\exists$ a unique R_α-minimizing solution $x^\dagger$. For $y \in range(T)$ and $||y - y^\delta|| \leq \delta$ if α satisfies $\lim\limits_{\delta \to 0} \alpha(\delta) = 0$ and $\lim\limits_{\delta \to 0} \frac{\delta}{\alpha(\delta)} = 0$, then $\lim\limits_{\delta \to 0} x_{\alpha(\delta)}^\delta = x^\dagger$.

Tikhonov Regularization with Sparsity Constraints

Theorem (Grasmair,Haltmeier,Scherzer)

Assume that $\partial R_\alpha(x^\dagger) \cap range(T^*) \neq \emptyset$, that the finite basis injectivity property holds, and that $Tx = y$ has an R_α-minimizing solution that is sparse wrt $\{\varphi_i\}_{i\in\mathbb{N}}$. Then for parameter choice strategy $\alpha \sim \delta$, we have $||x_\alpha^\delta - x^\dagger|| = O(\delta)$

Tikhonov Regularization with Sparsity Constraints

Theorem (Grasmair,Haltmeier,Scherzer)

Assume that $\partial R_\alpha(x^\dagger) \cap \mathit{range}(T^*) \neq \emptyset$, that the finite basis injectivity property holds, and that $Tx = y$ has an R_α-minimizing solution that is sparse wrt $\{\varphi_i\}_{i\in\mathbb{N}}$. Then for parameter choice strategy $\alpha \sim \delta$, we have $||x_\alpha^\delta - x^\dagger|| = O(\delta)$

In fact, Grasmair,Haltmeier,Scherzer also prove that for $1 \leq p \leq 2$, if we know in advance the solution is sparse in the basis $\{\varphi_i\}_{i\in\mathbb{N}}$ then we obtain $||x_\alpha^\delta - x^\dagger|| = O(\delta^{1/p})$.

Zentrum für Technomathematik

Tikhonov Regularization with Sparsity Constraints

Theorem (Grasmair,Haltmeier,Scherzer)

Assume that $\partial R_\alpha(x^\dagger) \cap range(T^*) \neq \emptyset$, that the finite basis injectivity property holds, and that $Tx = y$ has an R_α-minimizing solution that is sparse wrt $\{\varphi_i\}_{i\in\mathbb{N}}$. Then for parameter choice strategy $\alpha \sim \delta$, we have $||x_\alpha^\delta - x^\dagger|| = O(\delta)$

In fact, Grasmair,Haltmeier,Scherzer also prove that for $1 \leq p \leq 2$, if we know in advance the solution is sparse in the basis $\{\varphi_i\}_{i\in\mathbb{N}}$ then we obtain $||x_\alpha^\delta - x^\dagger|| = O(\delta^{1/p})$.

We have not yet answered how to find this minimizer x_α^δ. However, Jack will discuss this in his presentation.

Zentrum für Technomathematik