Chapter 1
Theory

Variational Image Processing
Summer School on Inverse Problems 2015

Michael Moeller
Computer Vision
Department of Computer Science
TU München

Why do we see this unstable behavior?

Representation of images:

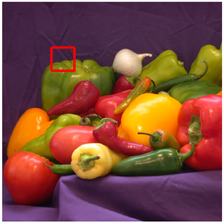
- **Discrete**: Pixel values stored in a matrix $\mathbb{R}^{n\times m\times c}$
  - Representation for computer screen
  - Important for algorithms

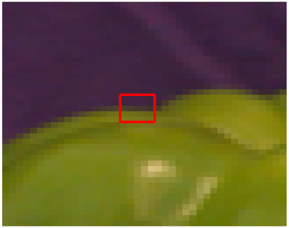
Why do we see this unstable behavior?

Representation of images:

- **Discrete**: Pixel values stored in a matrix $\mathbb{R}^{n \times m \times c}$
  - Representation for computer screen
  - Important for algorithms
- **Continuous**: Function $f : \Omega \subset \mathbb{R}^2 \to \mathbb{R}^c$
  - Preferable to understand certain behaviors

What happened on a discrete level?





36	35	35	36	36
36	38	38	36	39
65	65	62	58	55
112	113	112	109	104
122	121	119	119	115

0.0029	0.0133	0.0219	0.0133	0.0029
0.0133	0.0596	0.0983	0.0596	0.0133
0.0219	0.0983	0.1621	0.0983	0.0219
0.0133	0.0596	0.0983	0.0596	0.0133
0.0029	0.0133	0.0219	0.0133	0.0029

At each blurry pixel is formed as a weighted average over the sharp pixels. The weights for the averaging are given in the convolution kernel.

What happened on a discrete level?

Sharp image: $u \in \mathbb{R}^{n \times m}$ (grayscale)

Blur kernel: $b \in \mathbb{R}^{r \times r}$, typically $r << \min\{n, m\}$, for example

$$b = \begin{pmatrix} 0.0030 & 0.0133 & 0.0219 & 0.0133 & 0.0030 \\ 0.0133 & 0.0596 & 0.0983 & 0.0596 & 0.0133 \\ 0.0219 & 0.0983 & 0.1621 & 0.0983 & 0.0219 \\ 0.0133 & 0.0596 & 0.0983 & 0.0596 & 0.0133 \\ 0.0030 & 0.0133 & 0.0219 & 0.0133 & 0.0030 \end{pmatrix}$$

Blurry image:

$$f_{i,j} = \sum_{h=1}^{r} \sum_{l=1}^{r} b_{h,r} \; u_{i+h-\frac{r+1}{2}, j+l-\frac{r+1}{2}}$$

Linear equations! Can be written as

$$\vec{f} = B\vec{u}$$

Understanding the continuous case

Continuous model for a blurred image:

$$f(x) = \int_{\Omega} b(x, y)u(y)\, dy$$

with a kernel $b \in L^2(\Omega \times \Omega)$ (typically being a function of $x - y$).

Understanding the continuous case

Continuous model for a blurred image:

$$f(x) = \int_\Omega b(x,y)u(y)\,dy$$

with a kernel $b \in L^2(\Omega \times \Omega)$ (typically being a function of $x - y$).

Linear map

$$\begin{aligned} B : L^2(\Omega) &\to L^2(\Omega) \\ u &\mapsto \int_\Omega b(\cdot,y)u(y)\,dy \end{aligned}$$

seems to be difficult to ”invert”.

Well-posedness

Well-posedness

- A solution exists
- The solution is unique
- The solution depends continuously on the data

Linear inverse problems

Consider a problem of recovering u from $f = Au$ for a linear operator $A : X \to Y$.

Linear inverse problems

Consider a problem of recovering u from $f = Au$ for a linear operator $A : X \to Y$.

There could not be a solution!

Linear inverse problems

Consider a problem of recovering u from $f = Au$ for a linear operator $A : X \to Y$.

There could not be a solution!

Definition

We call u a *least-squares solution* of $Au = f$ if

$$\|Au - f\| = \inf\{\|Av - f\| \mid v \in X\}$$

Linear inverse problems

Consider a problem of recovering u from $f = Au$ for a linear operator $A : X \to Y$.

There could not be a solution!

Definition

We call u a *least-squares solution* of $Au = f$ if

$$\|Au - f\| = \inf\{\|Av - f\| \mid v \in X\}$$

The least-squares solution might not be unique.

Linear inverse problems

Consider a problem of recovering u from $f = Au$ for a linear operator $A : X \to Y$.

There could not be a solution!

Definition

We call u a *least-squares solution* of $Au = f$ if

$$\|Au - f\| = \inf\{\|Av - f\| \mid v \in X\}$$

The least-squares solution might not be unique.

Definition

We call u a *minimal-norm solution* of $Au = f$ if

$$\|u\| = \inf\{\|v\| \mid v \text{ is least-squares solution of } Au = f\}$$

Linear inverse problems

How can we recover minimal-norm solutions?

Moore-Penrose inverse

One can define a linear operator $A^\dagger$, such that for any $f \in \mathcal{R}(A) + \mathcal{R}(A)^\perp$, the equation $Ax = f$ has a unique minimal-norm solution given by

$$x^\dagger := A^\dagger f.$$

Linear inverse problems

How can we recover minimal-norm solutions?

Moore-Penrose inverse

One can define a linear operator $A^\dagger$, such that for any $f \in \mathcal{R}(A) + \mathcal{R}(A)^\perp$, the equation $Ax = f$ has a unique minimal-norm solution given by

$$x^\dagger := A^\dagger f.$$

Remember the SVD

$$Au = \sum_n \sigma_n \langle u_n, u \rangle v_n$$

Then

$$A^\dagger f = \sum_n \frac{1}{\sigma_n} \langle v_n, f \rangle u_n$$

Linear inverse problems

How can we recover minimal-norm solutions?

Moore-Penrose inverse

One can define a linear operator $A^\dagger$, such that for any $f \in \mathcal{R}(A) + \mathcal{R}(A)^\perp$, the equation $Ax = f$ has a unique minimal-norm solution given by

$$x^\dagger := A^\dagger f.$$

Remember the SVD

$$Au = \sum_n \sigma_n \langle u_n, u \rangle v_n$$

Then

$$A^\dagger f = \sum_n \frac{1}{\sigma_n} \langle v_n, f \rangle u_n$$

Back to deblurring

In finite dimensions: For $\|f - f^\delta\| \leq \delta$ do we have

$$f^\delta \approx f \quad \overset{?}{\Rightarrow} \quad B^\dagger \vec{f} \approx B^\dagger \vec{f^\delta} \quad ?$$

Back to deblurring

In finite dimensions: For $\|f - f^\delta\| \leq \delta$ do we have

$$f^\delta \approx f \qquad \overset{?}{\Rightarrow} \qquad B^\dagger \vec{f} \approx B^\dagger \vec{f^\delta} \ \ ?$$

Possible answers:

- **Yes** – in the sense that $\lim_{\delta \to 0} B^\dagger \vec{f^\delta} = B^\dagger \vec{f}$ (continuity).
- **No** – since B could be very ill-conditioned!

Observation in practice: Finer resolution $\rightarrow$ worse condition of B. Does the continuous case reveal a problem?

Understanding the continuous case

Definition: Compact linear operator

A linear operator $A : X \to Y$ is said to be *compact* if for every bounded sequence $\{x_n\} \subset X$, $\{Ax_n\}$ has a convergent subsequence.

Understanding the continuous case

Definition: Compact linear operator

A linear operator $A : X \to Y$ is said to be *compact* if for every bounded sequence $\{x_n\} \subset X$, $\{Ax_n\}$ has a convergent subsequence.

Theorem: Ill-posedness of compact linear operators

Let the linear operator linear operator $A : X \to Y$ be compact, and let the dimension of its range, $\mathcal{R}(A) \subset Y$, be infinite.

Then $A^\dagger$ is not continuous.

Compact linear operators

What kind of operators are compact?

Compact linear operators

What kind of operators are compact?

Theorem: Operators with Hilbert-Schmidt kernel are compact

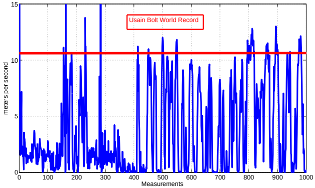
Let

$$Au(x) = \int_\Omega k(x,y)u(y)\ dy$$

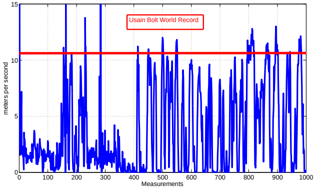
with kernel $k \in L^2(\Omega \times \Omega)$. Then $A \in \mathcal{L}(L^2(\Omega), L^2(\Omega))$ is compact.

A kernel $k \in L^2(\Omega \times \Omega)$ is called a *Hilbert-Schmidt kernel* from $\Omega \times \Omega \to \mathbb{R}$.

Differentiation: Finding $u(x)$ for given $\int_0^x u(y)dy$

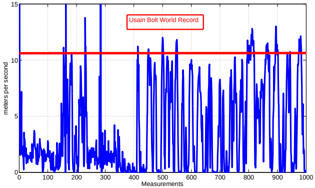


Differentiation: Finding $u(x)$ for given $\int_0^x u(y)dy$



Deconvolution: Finding $u(x)$ for given $\int_\Omega b(x-y)u(y)dy$

Differentiation: Finding $u(x)$ for given $\int_0^x u(y)dy$



Deconvolution: Finding $u(x)$ for given $\int_\Omega b(x-y)u(y)dy$

Ill-posed inverse problems!

Variational methods

Let f^δ be such that $\|f^\delta - f\|_2^2 \leq \delta$, $f = A\hat{u}$ for some (compact) linear operator A, (and $\hat{u}$ being a minimal norm solution).

Variational methods

Let f^δ be such that $\|f^\delta - f\|_2^2 \leq \delta$, $f = A\hat{u}$ for some (compact) linear operator A, (and $\hat{u}$ being a minimal norm solution).

We know A, f^δ and δ. Task: Find a good approximation of $\hat{u}$!

Variational methods

Let f^δ be such that $\|f^\delta - f\|_2^2 \leq \delta$, $f = A\hat{u}$ for some (compact) linear operator A, (and $\hat{u}$ being a minimal norm solution).

We know A, f^δ and δ. Task: Find a good approximation of $\hat{u}$!

Variational regularization methods

Determine

$$u(\alpha) = \arg\min_u \frac{1}{2}\|Au - f^\delta\|_2^2 + \alpha J(u)$$

for a suitable *regularization functional* J and a *regularization parameter* α based on δ and f^δ.

Existence and uniqueness of solutions

Before we can tackle the question if

$$u(\alpha) = \arg \min_{u \in L^2(\Omega)} \frac{1}{2} \|Au - f^\delta\|_2^2 + \alpha J(u)$$

is a good approximation of $\hat{u}$ with $f = A\hat{u}$, we have to consider:

1. Does $u(\alpha)$ even exist?
2. If yes, is $u(\alpha)$ unique?

Existence and uniqueness of solutions

Before we can tackle the question if

$$u(\alpha) = \arg \min_{u \in L^2(\Omega)} \frac{1}{2}\|Au - f^\delta\|_2^2 + \alpha J(u)$$

is a good approximation of $\hat{u}$ with $f = A\hat{u}$, we have to consider:

1. Does $u(\alpha)$ even exist?
2. If yes, is $u(\alpha)$ unique?

General setting we will consider:

$$u(\alpha) \in \arg \min_{u \in X} E_\alpha(u)$$

for a Banach space X.

Existence of solutions

Fundamental theorem of optimization

Let $E : (\mathcal{X}, \tau) \to \mathbb{R} \cup \{\infty\}$ be a functional on a topological space $\mathcal{X}$ with topology τ such that the following two conditions are met:

Existence of solutions

Fundamental theorem of optimization

Let $E : (\mathcal{X}, \tau) \to \mathbb{R} \cup \{\infty\}$ be a functional on a topological space $\mathcal{X}$ with topology τ such that the following two conditions are met:

- **Lower semi-continuity**: For $u_k \to u$ in the topology τ it holds that

$$E(u) \leq \liminf_k E(u_k)$$

Existence of solutions

Fundamental theorem of optimization

Let $E : (\mathcal{X}, \tau) \to \mathbb{R} \cup \{\infty\}$ be a functional on a topological space $\mathcal{X}$ with topology τ such that the following two conditions are met:

- **Lower semi-continuity**: For $u_k \to u$ in the topology τ it holds that

$$E(u) \leq \liminf_k E(u_k)$$

- **Precompactness of the sub-level-sets**: There exists an $\beta \in \mathbb{R}$ such that

$$S_\beta := \{u \in \mathcal{X} \mid E(u) \leq \beta\}$$

is not empty and precompact ($\overline{S_\beta}$ compact) in τ.

Existence of solutions

Fundamental theorem of optimization

Let $E : (\mathcal{X}, \tau) \to \mathbb{R} \cup \{\infty\}$ be a functional on a topological space $\mathcal{X}$ with topology τ such that the following two conditions are met:

- **Lower semi-continuity**: For $u_k \to u$ in the topology τ it holds that

$$E(u) \leq \liminf_k E(u_k)$$

- **Precompactness of the sub-level-sets**: There exists an $\beta \in \mathbb{R}$ such that

$$S_\beta := \{u \in \mathcal{X} \mid E(u) \leq \beta\}$$

  is not empty and precompact ($\overline{S_\beta}$ compact) in τ.

Then there exists a minimizer $\hat{u} \in \mathcal{X}$, i.e.

$$E(\hat{u}) = \inf_{u \in \mathcal{X}} E(u).$$

Uniqueness of solutions

Convex and strictly convex functions

Let $E : M \to \mathbb{R}$, $M \subset X$ being a convex subset of a Banach space.

- E is called convex if

$$E(\beta x_1 + (1-\beta)x_2) \leq \beta E(x_1) + (1-\beta)E(x_2)$$

  holds for all $x_1, x_2 \in M$ and $\beta \in [0,1]$.
- E is called strictly convex if the strict inequality holds for all $x_1, x_2 \in M$, $x_1 \neq x_2$ and $\beta \in]0,1[$.

Uniqueness of solutions

Convex and strictly convex functions

Let $E : M \to \mathbb{R}$, $M \subset X$ being a convex subset of a Banach space.

- E is called convex if

$$E(\beta x_1 + (1-\beta)x_2) \leq \beta E(x_1) + (1-\beta)E(x_2)$$

  holds for all $x_1, x_2 \in M$ and $\beta \in [0,1]$.
- E is called strictly convex if the strict inequality holds for all $x_1, x_2 \in M$, $x_1 \neq x_2$ and $\beta \in]0,1[$.

Uniqueness of minimizers

- Any local minimum of a convex function E is a global minimum.
- If E is strictly convex, the minimum is unique.

Variational methods

Back to our original question:

- Noisy data f^δ (given)
- Clean data f (unknown)
- Noise level $\delta = \|f - f^\delta\|_2$ (known)
- Data generation model: $\hat{u} = A^\dagger f$; linear op. A known.

Variational methods

Back to our original question:

- Noisy data f^δ (given)
- Clean data f (unknown)
- Noise level $\delta = \|f - f^\delta\|_2$ (known)
- Data generation model: $\hat{u} = A^\dagger f$; linear op. A known.

Continuous dependence on the data?

Does

$$u(\alpha) = \arg\min_u \frac{1}{2}\|Au - f^\delta\|_2^2 + \alpha J(u)$$

for a suitable *regularization functional* J and a *regularization parameter* α based on δ and f^δ converge to $\hat{u}$ for $\delta \to 0$?

Tikhonov regularization

Tikhonov regularization

$$u(\alpha) = \arg\min_u \frac{1}{2}\|Au - f^\delta\|_2^2 + \frac{\alpha}{2}\|u\|_2^2$$

Tikhonov regularization

Tikhonov regularization

$$u(\alpha) = \arg\min_u \frac{1}{2}\|Au - f^\delta\|_2^2 + \frac{\alpha}{2}\|u\|_2^2$$

Optimality condition

$$0 = A^*(Au(\alpha) - f^\delta) + \alpha u(\alpha)$$
$$\Rightarrow u(\alpha) = (A^*A + \alpha I)^{-1} A^* f^\delta$$

Tikhonov regularization

Tikhonov regularization

$$u(\alpha) = \arg\min_u \frac{1}{2}\|Au - f^\delta\|_2^2 + \frac{\alpha}{2}\|u\|_2^2$$

Optimality condition

$$0 = A^*(Au(\alpha) - f^\delta) + \alpha u(\alpha)$$
$$\Rightarrow u(\alpha) = (A^*A + \alpha I)^{-1}A^*f^\delta$$

Denote $R_\alpha = (A^*A + \alpha I)^{-1}A^*$. Then

$$\begin{aligned}\|u(\alpha) - \hat{u}\|_2 &= \|R_\alpha f^\delta - A^\dagger f\|_2 \\ &= \|R_\alpha f - A^\dagger f + R_\alpha(f^\delta - f)\|_2 \\ &\leq \underbrace{\|R_\alpha f - A^\dagger f\|_2}_{\text{Approximation error}} + \underbrace{\|R_\alpha(f^\delta - f)\|_2}_{\text{Data error}}\end{aligned}$$

Tikhonov regularization

$$\|u(\alpha) - \hat{u}\|_2 \leq \underbrace{\|R_\alpha f - A^\dagger f\|_2}_{\text{Approximation error}} + \underbrace{\|R_\alpha (f^\delta - f)\|_2}_{\text{Data error}}$$

$$\leq \|R_\alpha f - A^\dagger f\|_2 + \delta \|R_\alpha\|_2$$

What happens for $\delta \to 0$? Do we get $u(\alpha) \to \hat{u}$?

Tikhonov regularization

$$\|u(\alpha) - \hat{u}\|_2 \leq \underbrace{\|R_\alpha f - A^\dagger f\|_2}_{\text{Approximation error}} + \underbrace{\|R_\alpha(f^\delta - f)\|_2}_{\text{Data error}}$$

$$\leq \|R_\alpha f - A^\dagger f\|_2 + \delta \|R_\alpha\|_2$$

What happens for $\delta \to 0$? Do we get $u(\alpha) \to \hat{u}$?

- We need to choose a rule $\alpha = \alpha(\delta)$.
  (Also called *a-priori parameter choice* rule)

Tikhonov regularization

$$\|u(\alpha) - \hat{u}\|_2 \leq \underbrace{\|R_\alpha f - A^\dagger f\|_2}_{\text{Approximation error}} + \underbrace{\|R_\alpha(f^\delta - f)\|_2}_{\text{Data error}}$$

$$\leq \|R_\alpha f - A^\dagger f\|_2 + \delta \|R_\alpha\|_2$$

What happens for $\delta \to 0$? Do we get $u(\alpha) \to \hat{u}$?

- We need to choose a rule $\alpha = \alpha(\delta)$.
  (Also called *a-priori parameter choice* rule)
- For the approximation error to go to zero, we need

$$\alpha(\delta) \to 0$$

- The data error increases as $\alpha(\delta) \to 0$. Trade-off!
  How does $\|R_{\alpha(\delta)}\|_2$ increase?

Tikhonov regularization

$$\|u(\alpha) - \hat{u}\|_2 \leq \underbrace{\|R_\alpha f - A^\dagger f\|_2}_{\text{Approximation error}} + \underbrace{\|R_\alpha(f^\delta - f)\|_2}_{\text{Data error}}$$

$$\leq \|R_\alpha f - A^\dagger f\|_2 + \delta \|R_\alpha\|_2$$

What happens for $\delta \to 0$? Do we get $u(\alpha) \to \hat{u}$?

- We need to choose a rule $\alpha = \alpha(\delta)$.
  (Also called *a-priori parameter choice* rule)
- For the approximation error to go to zero, we need

$$\alpha(\delta) \to 0$$

- The data error increases as $\alpha(\delta) \to 0$. Trade-off!
  How does $\|R_{\alpha(\delta)}\|_2$ increase?
  Remember $R_\alpha = (A^* A + \alpha I)^{-1} A^*$

Tikhonov regularization

$$\|u(\alpha) - \hat{u}\|_2 \leq \underbrace{\|R_\alpha f - A^\dagger f\|_2}_{\text{Approximation error}} + \underbrace{\|R_\alpha(f^\delta - f)\|_2}_{\text{Data error}}$$

$$\leq \|R_\alpha f - A^\dagger f\|_2 + \delta \|R_\alpha\|_2$$

What happens for $\delta \to 0$? Do we get $u(\alpha) \to \hat{u}$?

- We need to choose a rule $\alpha = \alpha(\delta)$.
  (Also called *a-priori parameter choice* rule)
- For the approximation error to go to zero, we need

$$\alpha(\delta) \to 0$$

- The data error increases as $\alpha(\delta) \to 0$. Trade-off!
  How does $\|R_{\alpha(\delta)}\|_2$ increase?
  Remember $R_\alpha = (A^*A + \alpha I)^{-1} A^*$
  For singular values σ we have to consider $\frac{\sigma}{\sigma^2 + \alpha(\delta)}$.

Tikhonov regularization

$$\|u(\alpha) - \hat{u}\|_2 \leq \underbrace{\|R_\alpha f - A^\dagger f\|_2}_{\text{Approximation error}} + \underbrace{\|R_\alpha (f^\delta - f)\|_2}_{\text{Data error}}$$
$$\leq \|R_\alpha f - A^\dagger f\|_2 + \delta \|R_\alpha\|_2$$

What happens for $\delta \to 0$? Do we get $u(\alpha) \to \hat{u}$?

- We need to choose a rule $\alpha = \alpha(\delta)$.
  (Also called *a-priori parameter choice* rule)
- For the approximation error to go to zero, we need

$$\alpha(\delta) \to 0$$

- The data error increases as $\alpha(\delta) \to 0$. Trade-off!
  How does $\|R_{\alpha(\delta)}\|_2$ increase?
  Remember $R_\alpha = (A^* A + \alpha I)^{-1} A^*$
  For singular values σ we have to consider $\frac{\sigma}{\sigma^2 + \alpha(\delta)}$.
  Worst case: $\sigma = \sqrt{\alpha(\delta)}$. Therefore

$$\|R_{\alpha(\delta)}\|_2 \leq \frac{1}{2\sqrt{\alpha(\delta)}}$$

Tikhonov regularization

$$\|u(\alpha) - \hat{u}\|_2 \ \leq \|R_\alpha f - A^\dagger f\|_2 + \frac{\delta}{2\sqrt{\alpha(\delta)}}$$

Convergence of Tikhonov regularization

For Tikhonov regularization to converge we need

$$\alpha(\delta) \overset{\delta\to 0}{\to} 0 \qquad \text{and} \qquad \frac{\delta}{\sqrt{\alpha(\delta)}} \overset{\delta\to 0}{\to} 0$$

Tikhonov regularization

$$\|u(\alpha) - \hat{u}\|_2 \leq \|R_\alpha f - A^\dagger f\|_2 + \frac{\delta}{2\sqrt{\alpha(\delta)}}$$

Convergence of Tikhonov regularization

For Tikhonov regularization to converge we need

$$\alpha(\delta) \overset{\delta \to 0}{\to} 0 \qquad \text{and} \qquad \frac{\delta}{\sqrt{\alpha(\delta)}} \overset{\delta \to 0}{\to} 0$$

Some remarks about this convergence:

- Without add. assumptions the rate can be arbitrarily slow.

Tikhonov regularization

$$\|u(\alpha) - \hat{u}\|_2 \leq \|R_\alpha f - A^\dagger f\|_2 + \frac{\delta}{2\sqrt{\alpha(\delta)}}$$

Convergence of Tikhonov regularization

For Tikhonov regularization to converge we need

$$\alpha(\delta) \overset{\delta \to 0}{\to} 0 \qquad \text{and} \qquad \frac{\delta}{\sqrt{\alpha(\delta)}} \overset{\delta \to 0}{\to} 0$$

Some remarks about this convergence:

- Without add. assumptions the rate can be arbitrarily slow.
- *Source conditions* allow to derive convergence rates.

Tikhonov regularization

$$\|u(\alpha) - \hat{u}\|_2 \leq \|R_\alpha f - A^\dagger f\|_2 + \frac{\delta}{2\sqrt{\alpha(\delta)}}$$

Convergence of Tikhonov regularization

For Tikhonov regularization to converge we need

$$\alpha(\delta) \stackrel{\delta \to 0}{\to} 0 \qquad \text{and} \qquad \frac{\delta}{\sqrt{\alpha(\delta)}} \stackrel{\delta \to 0}{\to} 0$$

Some remarks about this convergence:

- Without add. assumptions the rate can be arbitrarily slow.
- *Source conditions* allow to derive convergence rates.
- The rate of convergence is always strictly worse than $\mathcal{O}(\delta)$!

Tikhonov regularization

$$\|u(\alpha) - \hat{u}\|_2 \leq \|R_\alpha f - A^\dagger f\|_2 + \frac{\delta}{2\sqrt{\alpha(\delta)}}$$

Convergence of Tikhonov regularization

For Tikhonov regularization to converge we need

$$\alpha(\delta) \overset{\delta \to 0}{\to} 0 \quad \text{and} \quad \frac{\delta}{\sqrt{\alpha(\delta)}} \overset{\delta \to 0}{\to} 0$$

Some remarks about this convergence:

- Without add. assumptions the rate can be arbitrarily slow.
- *Source conditions* allow to derive convergence rates.
- The rate of convergence is always strictly worse than $\mathcal{O}(\delta)$!
- Further readings:
  - H.W. Engl, M. Hanke, G. Neubauer. *Regularization of Inverse Problems.*
  - M. Burger, S. Osher. *Convergence Rates of Convex Variational Regularization.*
  - M. Benning, M. Burger. *Error Estimates for General Fidelities.*

Beyond Tikhonov regularization

Seen: Tikhonov regularization

$$u(\alpha) = \arg\min_u \frac{1}{2}\|Au - f^\delta\|_2^2 + \frac{\alpha}{2}\|u\|_2^2$$

stabilizes the reconstruction. For $\delta \to 0$ and a suitable $\alpha(\delta) \to 0$ one converges to the minimal norm solution.

But is this always desired?

Beyond Tikhonov regularization

Seen: Tikhonov regularization

$$u(\alpha) = \arg\min_u \frac{1}{2}\|Au - f^\delta\|_2^2 + \frac{\alpha}{2}\|u\|_2^2$$

stabilizes the reconstruction. For $\delta \to 0$ and a suitable $\alpha(\delta) \to 0$ one converges to the minimal norm solution.

But is this always desired?

Regularizations can be much more powerful!

$$u(\alpha) = \arg\min_u \frac{1}{2}\|Au - f^\delta\|_2^2 + \alpha J(u)$$

***J* allows to impose a-priori information on the solution!**

What is a good regularization for images?

Task: Distinguish between signal and noise.

Bad! $J(u)$ large!

Good! $J(u)$ small!

Noisy is highly oscillatory!

Regularization $\|u\|_2^2$ does not ”see” the oscillations.

What is a good regularization for images?

Idea: Penalize the gradient of the reconstructed image.

$$J(u) = \frac{1}{2}\|\nabla u\|_2^2$$

What is a good regularization for images?

Idea: Penalize the gradient of the reconstructed image.

$$J(u) = \frac{1}{2}\|\nabla u\|_2^2$$

Problem: Discontinuous functions are not admissible!

$J(u) \approx 170$

$J(u) \approx 1200$

Discrete: The finer the discretization the bigger is the difference between sharp and blurry images.

Functions: Discontinuous functions are not in $W^{1,2}$

What is a good regularization for images?

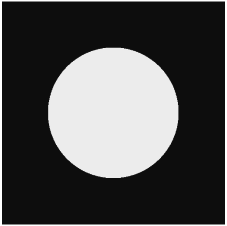
Derivatives should be allowed to be distributions.

$$TV(u) = \sup_{\substack{q \in C_0^\infty(\Omega, \mathbb{R}^2), \\ |q(x)| \leq 1}} \int_\Omega u \, \text{div}(q) \, dx$$

$$\text{For } u \in W^{1,1} : \quad TV(u) = \int_\Omega |\nabla u(x)| \, dx$$

$TV(u) \approx 945$

$\|\nabla u\|^2 \approx 18$



$TV(u) \approx 958$

$\|\nabla u\|^2 \approx 916$

Regularization via Total Variation

Example: TV-Image deblurring

$$u(\alpha) = \arg\min_u \frac{1}{2}\|Au - f^\delta\|_2^2 + \alpha TV(u)$$

Regularization via Total Variation

Example: TV-Image deblurring

$$u(\alpha) = \arg\min_u \frac{1}{2}\|Au - f^\delta\|_2^2 + \alpha TV(u)$$

Regularization via Total Variation

Example: TV-Image deblurring

$$u(\alpha) = \arg\min_u \frac{1}{2}\|Au - f^\delta\|_2^2 + \alpha TV(u)$$

Regularization via Total Variation

Example: TV-Image deblurring

$$u(\alpha) = \arg\min_u \frac{1}{2}\|Au - f^\delta\|_2^2 + \alpha TV(u)$$

Analysis of TV-model (existence, continuous dependence, ...): Important to choose the right topology! Further readings:

- *Image recovery via total variation minimization and related problems.* Chambolle, Lions.
- *A Guide to the TV Zoo.* Burger, Osher.
- *Mathematische Bildverarbeitung.* Bredis, Lorenz. E.g. Satz 6.115.

Generalizations and Extensions

If changing $J(u)$ can improve the results, we might as well consider more general data terms, too!

What are applications beyond image deblurring?

Generalizations and Extensions

If changing $J(u)$ can improve the results, we might as well consider more general data terms, too!

What are applications beyond image deblurring?

General modeling of variational methods of the form

$$u(\alpha) = \arg\min_u H(u, f) + \alpha J(u)$$

To be continued...