WESTFÄLISCHE
WILHELMS-UNIVERSITÄT
MÜNSTER

The Wonderful World of Bregman Distances

Skiseminar, Zafernahütte

wissen.leben
WWU Münster

Martin Benning
martin.benning@wwu.de

21.02.11

› Outline

› Fundamentals

› Fundamentals

- Many problems in applied math can be modelled as inverse problems like

Inverse Problem with Exact Data

$$K\tilde{u} = g \tag{1}$$

with a (linear) operator K, and given *exact data* g and the unknown *exact solution* $\tilde{u}$

› Fundamentals

- Many problems in applied math can be modelled as inverse problems like

Inverse Problem with Exact Data

$$K\tilde{u} = g \tag{1}$$

with a (linear) operator K, and given *exact data* g and the unknown *exact solution* $\tilde{u}$, or - since the exact data is usually not available - like

Inverse Problem with Noisy Data

$$Ku = f \tag{2}$$

with given *noisy data* f and unknown u

- Since the operator K is usually not (continuously) invertible and since g is usually not available, it is common to look for approximate solutions $\hat{u}$ via the variational framework

Variational Minimization Scheme

$$\hat{u} \in \arg\min_{u \in \mathrm{dom}(J)} \left\{ E(u) \right\} = \arg\min_{u \in \mathrm{dom}(J)} \left\{ H_f(Ku) + \alpha J(u) \right\} \tag{3}$$

- Since the operator K is usually not (continuously) invertible and since g is usually not available, it is common to look for approximate solutions $\hat{u}$ via the variational framework

Variational Minimization Scheme

$$\hat{u} \in \arg\min_{u \in \mathrm{dom}(J)} \{E(u)\} = \arg\min_{u \in \mathrm{dom}(J)} \left\{H_f(Ku) + \alpha J(u)\right\} \tag{3}$$

- The term $H_f(Ku)$ is called *fidelity term*, while $J(u)$ is called *regularization term* with *regularization parameter* $\alpha > 0$

- Since the operator K is usually not (continuously) invertible and since g is usually not available, it is common to look for approximate solutions $\hat{u}$ via the variational framework

Variational Minimization Scheme

$$\hat{u} \in \arg\min_{u \in \mathrm{dom}(J)} \{E(u)\} = \arg\min_{u \in \mathrm{dom}(J)} \left\{H_f(Ku) + \alpha J(u)\right\} \tag{3}$$

- The term $H_f(Ku)$ is called *fidelity term*, while $J(u)$ is called *regularization term* with *regularization parameter* $\alpha > 0$
- The fidelity term measures the deviation between $K\hat{u}$ and f, usually considering the distribution of the noise in the data

- Since the operator K is usually not (continuously) invertible and since g is usually not available, it is common to look for approximate solutions $\hat{u}$ via the variational framework

Variational Minimization Scheme

$$\hat{u} \in \arg\min_{u \in \mathrm{dom}(J)} \{E(u)\} = \arg\min_{u \in \mathrm{dom}(J)} \left\{H_f(Ku) + \alpha J(u)\right\} \tag{3}$$

- The term $H_f(Ku)$ is called *fidelity term*, while $J(u)$ is called *regularization term* with *regularization parameter* $\alpha > 0$
- The fidelity term measures the deviation between $K\hat{u}$ and f, usually considering the distribution of the noise in the data
- The regularization term incorporates a-priori knowledge on the desired solution $\hat{u}$

› Fidelities and Regularization Energies: Examples

- Typical fidelities are L^p-norms, i.e. $H_f(Ku) = \|Ku - f\|_{L^p(\Sigma)}$, usually to the power of p (i.e. $H_f(Ku) = \frac{1}{p}\|Ku - f\|^p_{L^p(\Sigma)}$), for $p \geq 1$ (in particular $p = 2$)

› Fidelities and Regularization Energies: Examples

- Typical fidelities are L^p-norms, i.e. $H_f(Ku) = \|Ku - f\|_{L^p(\Sigma)}$, usually to the power of p (i.e. $H_f(Ku) = \frac{1}{p}\|Ku - f\|^p_{L^p(\Sigma)}$), for $p \geq 1$ (in particular $p = 2$)
- Standard regularization energies, to mention a few:
    - Squared L^2-norm, with operator: $J(u) = \frac{1}{2}\|Bu\|^2_{L^2(\Omega)}$ (e.g. with $B = I$, $B = \nabla$ etc.)

› Fidelities and Regularization Energies: Examples

- Typical fidelities are L^p-norms, i.e. $H_f(Ku) = \|Ku - f\|_{L^p(\Sigma)}$, usually to the power of p (i.e. $H_f(Ku) = \frac{1}{p}\|Ku - f\|^p_{L^p(\Sigma)}$), for $p \geq 1$ (in particular $p = 2$)
- Standard regularization energies, to mention a few:
  - Squared L^2-norm, with operator: $J(u) = \frac{1}{2}\|Bu\|^2_{L^2(\Omega)}$ (e.g. with $B = I$, $B = \nabla$ etc.)
  - ℓ^1-norm: $J(u) = \|u\|_{\ell^1}$

› Fidelities and Regularization Energies: Examples

- Typical fidelities are L^p-norms, i.e. $H_f(Ku) = \|Ku - f\|_{L^p(\Sigma)}$, usually to the power of p (i.e. $H_f(Ku) = \frac{1}{p}\|Ku - f\|^p_{L^p(\Sigma)}$), for $p \geq 1$ (in particular $p = 2$)
- Standard regularization energies, to mention a few:
  - Squared L^2-norm, with operator: $J(u) = \frac{1}{2}\|Bu\|^2_{L^2(\Omega)}$ (e.g. with $B = I$, $B = \nabla$ etc.)
  - ℓ^1-norm: $J(u) = \|u\|_{\ell^1}$
  - Total Variation: $J(u) = \mathrm{TV}(u) = \sup_{\substack{q \in C_0^\infty(\Omega;\mathbb{R}^n) \\ \|q\|_{L^\infty(\Omega)} \leq 1}} \int_\Omega u \, \mathrm{div} q \, dx \approx \int_\Omega \|\nabla u\|_{\ell^1} \, dx$

› Fidelities and Regularization Energies: Examples

- Typical fidelities are L^p-norms, i.e. $H_f(Ku) = \|Ku - f\|_{L^p(\Sigma)}$, usually to the power of p (i.e. $H_f(Ku) = \frac{1}{p}\|Ku - f\|^p_{L^p(\Sigma)}$), for $p \geq 1$ (in particular $p = 2$)
- Standard regularization energies, to mention a few:
  - Squared L^2-norm, with operator: $J(u) = \frac{1}{2}\|Bu\|^2_{L^2(\Omega)}$ (e.g. with $B = I$, $B = \nabla$ etc.)
  - ℓ^1-norm: $J(u) = \|u\|_{\ell^1}$
  - Total Variation: $J(u) = \text{TV}(u) = \sup_{\substack{q \in C_0^\infty(\Omega;\mathbb{R}^n) \\ \|q\|_{L^\infty(\Omega)} \leq 1}} \int_\Omega u \,\text{div} q \, dx \approx \int_\Omega \|\nabla u\|_{\ell^1} \, dx$
  - And many many more...

› Fidelities and Regularization Energies: Examples

- Typical fidelities are L^p-norms, i.e. $H_f(Ku) = \|Ku - f\|_{L^p(\Sigma)}$, usually to the power of p (i.e. $H_f(Ku) = \frac{1}{p}\|Ku - f\|^p_{L^p(\Sigma)}$), for $p \geq 1$ (in particular $p = 2$)
- Standard regularization energies, to mention a few:
  - Squared L^2-norm, with operator: $J(u) = \frac{1}{2}\|Bu\|^2_{L^2(\Omega)}$ (e.g. with $B = I$, $B = \nabla$ etc.)
  - ℓ^1-norm: $J(u) = \|u\|_{\ell^1}$
  - Total Variation: $J(u) = \text{TV}(u) = \sup_{\substack{q \in C_0^\infty(\Omega;\mathbb{R}^n) \\ \|q\|_{L^\infty(\Omega)} \leq 1}} \int_\Omega u \,\text{div} q \, dx \approx \int_\Omega \|\nabla u\|_{\ell^1} \, dx$
  - And many many more...
- For us of particular interest:

ROF (Rudin-Osher-Fatemi)

$$\hat{u} \in \arg\min_{u \in \text{BV}(\Omega)} \left\{ \frac{1}{2}\|Ku - f\|^2_{L^2(\Sigma)} + \alpha \text{TV}(u) \right\}$$

› Fidelities and Regularization Energies: Examples

- Typical fidelities are L^p-norms, i.e. $H_f(Ku) = \|Ku - f\|_{L^p(\Sigma)}$, usually to the power of p (i.e. $H_f(Ku) = \frac{1}{p}\|Ku - f\|^p_{L^p(\Sigma)}$), for $p \geq 1$ (in particular $p = 2$)
- Standard regularization energies, to mention a few:
  - Squared L^2-norm, with operator: $J(u) = \frac{1}{2}\|Bu\|^2_{L^2(\Omega)}$ (e.g. with $B = I$, $B = \nabla$ etc.)
  - ℓ^1-norm: $J(u) = \|u\|_{\ell^1}$
  - Total Variation: $J(u) = \text{TV}(u) = \sup_{\substack{q \in C_0^\infty(\Omega;\mathbb{R}^n) \\ \|q\|_{L^\infty(\Omega)} \leq 1}} \int_\Omega u \operatorname{div} q \, dx \approx \int_\Omega \|\nabla u\|_{\ell^1} \, dx$
  - And many many more...
- For us of particular interest:

Compressed Sensing

$$\hat{u} \in \arg\min_{u \in \ell^1} \left\{ \frac{1}{2}\|Ku - f\|^2_{\ell^2} + \alpha\|u\|_{\ell^1} \right\}$$

- Another way to obtain approximate solutions $\hat{u}$ is via constrained minimization

Constrained Minimization

$$\hat{u} \in \arg\min_{u \in \mathrm{dom}(J)} \{J(u)\} \text{ subject to } Ku = f \tag{4}$$

- Another way to obtain approximate solutions $\hat{u}$ is via constrained minimization

Constrained Minimization

$$\hat{u} \in \arg\min_{u \in \mathrm{dom}(J)} \{J(u)\} \text{ subject to } Ku = f \tag{4}$$

- The goal of this talk is to show that under appropriate conditions on the data f and on the operator K, and with the right tool - namely the Bregman distance - we are able to guarantee $\hat{u} = \tilde{u}$

› Bregman Distance

- The Bregman distance of a functional J is defined as

Bregman Distance

$$D_J^p(u, v) := J(u) - J(v) - \langle p, u - v \rangle \, , p \in \partial J(v) \, , \tag{5}$$

with $\partial J(v)$ denoting the subdifferential at position v, i.e.

$$\partial J(v) = \{ p \in \mathcal{X}^* \mid J(u) - J(v) - \langle p, u - v \rangle_{\mathcal{X}} \geq 0 \, , \forall u \in \mathcal{X} \} \tag{6}$$

› Examples

- Subdifferential Example:

› Examples

- Subdifferential Example: Let us consider the absolute value function, i.e. $J(v) : \mathbb{R} \to \mathbb{R}_{\geq 0}$ with $J(v) := |v|$

› Examples

- Subdifferential Example: Let us consider the absolute value function, i.e. $J(v) : \mathbb{R} \to \mathbb{R}_{\geq 0}$ with $J(v) := |v|$
- In that case the subdifferential is given as

Subdifferential $\partial|u|$

$$\partial|v| = \text{sign}(v) = \begin{cases} \{1\} & \text{for } v > 0 \\ [-1, 1] & \text{for } v = 0 \\ \{-1\} & \text{for } v < 0 \end{cases}$$

› Examples

- Subdifferential Example: Let us consider the absolute value function, i.e. $J(v) : \mathbb{R} \to \mathbb{R}_{\geq 0}$ with $J(v) := |v|$
- In that case the subdifferential is given as

Subdifferential $\partial|u|$

$$\partial|v| = \text{sign}(v) = \begin{cases} \{1\} & \text{for } v > 0 \\ [-1, 1] & \text{for } v = 0 \\ \{-1\} & \text{for } v < 0 \end{cases}$$

- This can easily be verified by case differentiation:
  - $v = 0$: $|u| \geq pu$ is obviously verified for any $u \in \mathbb{R}$ iff $p \in [-1, 1]$

› Examples

- Subdifferential Example: Let us consider the absolute value function, i.e. $J(v) : \mathbb{R} \to \mathbb{R}_{\geq 0}$ with $J(v) := |v|$
- In that case the subdifferential is given as

Subdifferential $\partial|u|$

$$\partial|v| = \text{sign}(v) = \begin{cases} \{1\} & \text{for } v > 0 \\ [-1, 1] & \text{for } v = 0 \\ \{-1\} & \text{for } v < 0 \end{cases}$$

- This can easily be verified by case differentiation:
  - $v = 0$: $|u| \geq pu$ is obviously verified for any $u \in \mathbb{R}$ iff $p \in [-1, 1]$
  - $v > 0$: $|u| \geq pu + (1 - p)v$ is fulfilled for every $u \in \mathbb{R}$ iff $p = 1$

› Examples

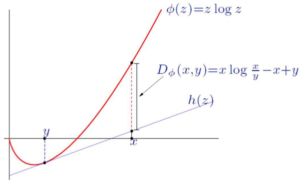
- Subdifferential Example: Let us consider the absolute value function, i.e. $J(v) : \mathbb{R} \to \mathbb{R}_{\geq 0}$ with $J(v) := |v|$
- In that case the subdifferential is given as

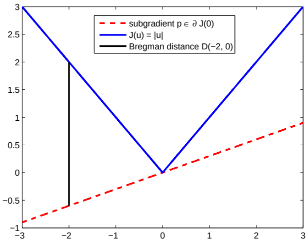
Subdifferential $\partial|u|$

$$\partial|v| = \text{sign}(v) = \begin{cases} \{1\} & \text{for } v > 0 \\ [-1, 1] & \text{for } v = 0 \\ \{-1\} & \text{for } v < 0 \end{cases}$$

- This can easily be verified by case differentiation:
  - $v = 0$: $|u| \geq pu$ is obviously verified for any $u \in \mathbb{R}$ iff $p \in [-1, 1]$
  - $v > 0$: $|u| \geq pu + (1 - p)v$ is fulfilled for every $u \in \mathbb{R}$ iff $p = 1$
  - $v < 0$: $|u| \geq pu - (1 + p)v$ is fulfilled for each $u \in \mathbb{R}$ iff $p = -1$

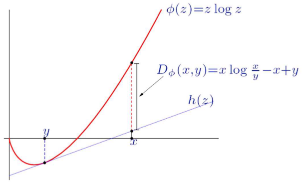
› Examples

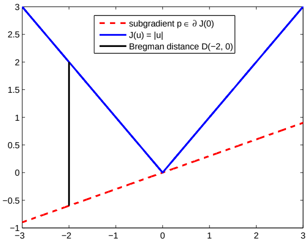




- The Bregman distance measures the difference between J and a tangent of J at v in u

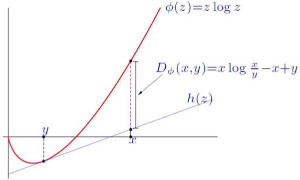
› Examples

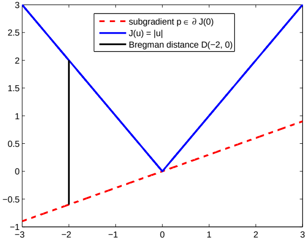




- The Bregman distance measures the difference between J and a tangent of J at v in u
- For convex J the Bregman distance is always non-negative; for strictly convex J we even have $D_J(u,v) = 0$ iff $u = v$

› Examples





- The Bregman distance measures the difference between J and a tangent of J at v in u
- For convex J the Bregman distance is always non-negative; for strictly convex J we even have $D_J(u, v) = 0$ iff $u = v$
- The Bregman distance is no metric; it is usually not symmetric and does not satisfy a triangular inequality

› Bregman Distance as a Regularizer

› Bregman Distance as a Regularizer

- Replacing the regularization term $J(u)$ with the Bregman distance $D_J^{p_k}(u, u_k)$, for $p_k \in \partial J(u_k)$, implies the following iterative procedure

Bregman Iteration

$$u_k = \arg\min_{u \in \mathrm{dom}(J)} \left\{ H_f(Ku) + \alpha D_J^{p_{k-1}}(u, u_{k-1}) \right\} \tag{7}$$

› Bregman Distance as a Regularizer

- Replacing the regularization term $J(u)$ with the Bregman distance $D_J^{p_k}(u, u_k)$, for $p_k \in \partial J(u_k)$, implies the following iterative procedure

Bregman Iteration

$$u_k = \arg\min_{u \in \mathrm{dom}(J)} \left\{ H_f(Ku) + \alpha D_J^{p_{k-1}}(u, u_{k-1}) \right\} \tag{7}$$

- Considering the optimality condition $K^* H_f'(Ku_k) + \alpha(p_k - p_{k-1}) = 0$ as a backward-Euler-discretization with stepsize α, Bregman iteration can be seen as the discrete counterpart of

Inverse Scale Space Flow

$$\partial_t p = -K^* H_f'(Ku) \tag{8}$$

- In the case of $H_f(Ku) = \frac{1}{2}\|Ku - f\|^2_{L^2(\Sigma)}$ one can prove that solving (7) is equivalent to solving the constrained minimization problem (4)

- In the case of $H_f(Ku) = \frac{1}{2}\|Ku - f\|^2_{L^2(\Sigma)}$ one can prove that solving (7) is equivalent to solving the constrained minimization problem (4)
- For the sake of simplicity we want to focus on $H_f(Ku) = \frac{1}{2}\|Ku - f\|^2_{L^2(\Sigma)}$; the variational scheme (3) and the inverse scale space formulation (8) therefore modify to

Bregman Tools for the Remainder of this Talk

$$\hat{u} \in \arg\min_{u \in \mathrm{dom}(J)} \left\{ \frac{1}{2}\|Ku - f\|^2_{L^2(\Sigma)} + \alpha J(u) \right\} \tag{9}$$

$$\partial_t p = K^* \left(f - Ku \right) \tag{10}$$

› Exact Recovery

- When are we able to recover signals exactly?

› Exact Recovery

- When are we able to recover signals exactly?
- In order to answer this question, we want to introduce the term *Eigenfunction*

› Exact Recovery

- When are we able to recover signals exactly?
- In order to answer this question, we want to introduce the term *Eigenfunction*

Groundstates and Eigenfunctions

Let $J : \mathrm{dom}(J) \subseteq L^2(\Omega) \to \mathbb{R} \cup \{\infty\}$ be a convex functional and $K : L^2(\Omega) \to L^2(\Sigma)$ a linear operator. Then, an Eigenfunction $\hat{u}$ with Eigenvalue λ satisfies

$$\lambda K^* K \hat{u} \in \partial J(\hat{u}) \,. \tag{11}$$

- Eigenfunctions of the two-homogeneous functional $J(u) = \frac{1}{2}\|\nabla u\|^2_{L^2(\Omega)}$:

- Eigenfunctions of the two-homogeneous functional $J(u) = \frac{1}{2}\|\nabla u\|^2_{L^2(\Omega)}$:

Eigenfunction of $\frac{1}{2}\|\nabla u\|^2_{L^2(\Omega)}$

$$\lambda K^* K \hat{u} \in \partial J(\hat{u}) = \{-\Delta \hat{u}\} \tag{12}$$

For $K = I$ we have the classical Eigenfunction-problem of the Laplace operator, i.e. $-\lambda \hat{u} = \Delta \hat{u}$

- Eigenfunctions of the two-homogeneous functional $J(u) = \frac{1}{2}\|\nabla u\|^2_{L^2(\Omega)}$:

Eigenfunction of $\frac{1}{2}\|\nabla u\|^2_{L^2(\Omega)}$

$$\lambda K^* K \hat{u} \in \partial J(\hat{u}) = \{-\Delta \hat{u}\} \tag{12}$$

For $K = I$ we have the classical Eigenfunction-problem of the Laplace operator, i.e. $-\lambda \hat{u} = \Delta \hat{u}$

- Eigenfunctions of $J(u) = \mathrm{TV}(u)$ are more complicated to discover;

- Eigenfunctions of the two-homogeneous functional $J(u) = \frac{1}{2}\|\nabla u\|^2_{L^2(\Omega)}$:

Eigenfunction of $\frac{1}{2}\|\nabla u\|^2_{L^2(\Omega)}$

$$\lambda K^* K\hat{u} \in \partial J(\hat{u}) = \{-\Delta \hat{u}\} \tag{12}$$

For $K = I$ we have the classical Eigenfunction-problem of the Laplace operator, i.e. $-\lambda\hat{u} = \Delta\hat{u}$

- Eigenfunctions of $J(u) = \mathrm{TV}(u)$ are more complicated to discover; let us therefore consider the subdifferential $\partial\mathrm{TV}$ first

TV-Subdifferential

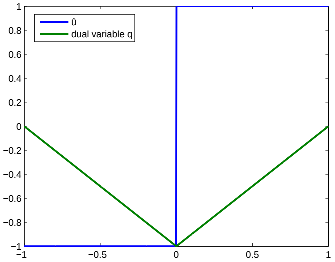
$$\partial\mathrm{TV}(u) = \{\mathrm{div} q \mid q \in C_0(\Omega), \|q\|_\infty = 1, \langle u, \mathrm{div} q\rangle = \mathrm{TV}(u)\} \tag{13}$$

- In order to show that a particular function $\hat{u}$ is an Eigenfunction of TV with Eigenvalue λ, we need to show that there exists a q with $\operatorname{div} q = \lambda K^* K \hat{u}$ that satisfies all the subgradient properties of (13)

- In order to show that a particular function $\hat{u}$ is an Eigenfunction of TV with Eigenvalue λ, we need to show that there exists a q with $\text{div}q = \lambda K^* K \hat{u}$ that satisfies all the subgradient properties of (13)
- Example:

- In order to show that a particular function $\hat{u}$ is an Eigenfunction of TV with Eigenvalue λ, we need to show that there exists a q with $\text{div}q = \lambda K^* K\hat{u}$ that satisfies all the subgradient properties of (13)
- Example: We are going to show that for $K = I$ the function $\hat{u}(x) = \text{sign}(x)$ is an Eigenfunction of TV with Eigenvalue $\lambda = 1/L$ for $x \in [-L, L]$:

- In order to show that a particular function $\hat{u}$ is an Eigenfunction of TV with Eigenvalue λ, we need to show that there exists a q with div$q = \lambda K^* K \hat{u}$ that satisfies all the subgradient properties of (13)
- Example: We are going to show that for $K = I$ the function $\hat{u}(x) = \text{sign}(x)$ is an Eigenfunction of TV with Eigenvalue $\lambda = 1/L$ for $x \in [-L, L]$:

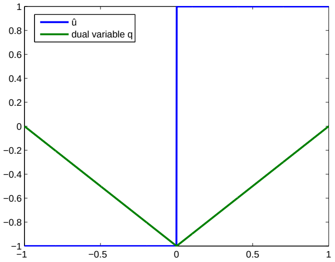


Eigenfunction-Example for $J(u) = \text{TV}(u)$

The function $q(x) = (|x| - L)/L$ satisfies

- $q(L) = q(-L) = 0$

- In order to show that a particular function $\hat{u}$ is an Eigenfunction of TV with Eigenvalue λ, we need to show that there exists a q with $\text{div}q = \lambda K^* K \hat{u}$ that satisfies all the subgradient properties of (13)
- Example: We are going to show that for $K = I$ the function $\hat{u}(x) = \text{sign}(x)$ is an Eigenfunction of TV with Eigenvalue $\lambda = 1/L$ for $x \in [-L, L]$:

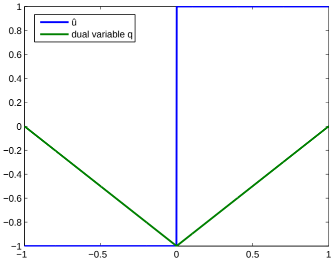


Eigenfunction-Example for $J(u) = \text{TV}(u)$

The function $q(x) = (|x| - L)/L$ satisfies

- $q(L) = q(-L) = 0$
- $q'(x) = \lambda \text{sign}(x)$ (in a weak sense)

- In order to show that a particular function $\hat{u}$ is an Eigenfunction of TV with Eigenvalue λ, we need to show that there exists a q with div$q = \lambda K^* K \hat{u}$ that satisfies all the subgradient properties of (13)
- Example: We are going to show that for $K = I$ the function $\hat{u}(x) = \text{sign}(x)$ is an Eigenfunction of TV with Eigenvalue $\lambda = 1/L$ for $x \in [-L, L]$:

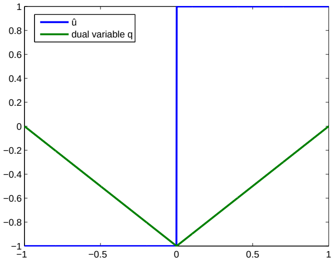


Eigenfunction-Example for $J(u) = \text{TV}(u)$

The function $q(x) = (|x| - L)/L$ satisfies

- $q(L) = q(-L) = 0$
- $q'(x) = \lambda \text{sign}(x)$ (in a weak sense)
- $\|q\|_\infty = 1$

- In order to show that a particular function $\hat{u}$ is an Eigenfunction of TV with Eigenvalue λ, we need to show that there exists a q with $\text{div}q = \lambda K^* K \hat{u}$ that satisfies all the subgradient properties of (13)
- Example: We are going to show that for $K = I$ the function $\hat{u}(x) = \text{sign}(x)$ is an Eigenfunction of TV with Eigenvalue $\lambda = 1/L$ for $x \in [-L, L]$:



Eigenfunction-Example for $J(u) = \text{TV}(u)$

The function $q(x) = (|x| - L)/L$ satisfies

- $q(L) = q(-L) = 0$
- $q'(x) = \lambda \text{sign}(x)$ (in a weak sense)
- $\|q\|_\infty = 1$
- $\langle \hat{u}, q' \rangle = \text{TV}(\hat{u}) = 2$

- As a last example we want to consider $J(u) = \|u\|_{\ell^1}$; for an element $q \in \partial\|u\|_{\ell^1}$ we have $q_j = \text{sign}(u_j), \ \forall j \in \{1, \dots, n\}$

- As a last example we want to consider $J(u) = \|u\|_{\ell^1}$; for an element $q \in \partial\|u\|_{\ell^1}$ we have $q_j = \operatorname{sign}(u_j), \ \forall j \in \{1, \dots, n\}$
- Example:

- As a last example we want to consider $J(u) = \|u\|_{\ell^1}$; for an element $q \in \partial\|u\|_{\ell^1}$ we have $q_j = \mathrm{sign}(u_j)$, $\forall j \in \{1, \dots, n\}$
- Example: We are going to show that $\hat{u} = \delta_i/\gamma = 1/\gamma \begin{cases} 1 & \text{if } j = i \\ 0 & \text{else} \end{cases}$ is an Eigenfunction of ℓ^1 with Eigenvalue $\lambda = 1/\gamma$ if the columns of the $m \times n$-matrix K are normed with respect to the 2-norm

- As a last example we want to consider $J(u) = \|u\|_{\ell^1}$; for an element $q \in \partial\|u\|_{\ell^1}$ we have $q_j = \text{sign}(u_j), \ \forall j \in \{1, \dots, n\}$
- Example: We are going to show that $\hat{u} = \delta_i/\gamma = 1/\gamma \begin{cases} 1 & \text{if } j = i \\ 0 & \text{else} \end{cases}$ is an Eigenfunction of ℓ^1 with Eigenvalue $\lambda = 1/\gamma$ if the columns of the $m \times n$-matrix K are normed with respect to the 2-norm

Eigenfunction-Example for $J(u) = \|u\|_{\ell^1}$

- Considering $K\hat{u}$ yields the i-th column of K multiplied with λ, i.e. $\lambda K_{m \times i}$

- As a last example we want to consider $J(u) = \|u\|_{\ell^1}$; for an element $q \in \partial\|u\|_{\ell^1}$ we have $q_j = \text{sign}(u_j),\ \forall j \in \{1, \dots, n\}$
- Example: We are going to show that $\hat{u} = \delta_i/\gamma = 1/\gamma \begin{cases} 1 & \text{if } j = i \\ 0 & \text{else} \end{cases}$ is an Eigenfunction of ℓ^1 with Eigenvalue $\lambda = 1/\gamma$ if the columns of the $m \times n$-matrix K are normed with respect to the 2-norm

Eigenfunction-Example for $J(u) = \|u\|_{\ell^1}$

- Considering $K\hat{u}$ yields the i-th column of K multiplied with λ, i.e. $\lambda K_{m\times i}$
- Computing $\hat{q}_j = \lambda(K^T K_{m\times i})_j$ is equivalent to $\hat{q}_j = \lambda\langle K^T_{j\times m}, K_{m\times i}\rangle$;

- As a last example we want to consider $J(u) = \|u\|_{\ell^1}$; for an element $q \in \partial\|u\|_{\ell^1}$ we have $q_j = \text{sign}(u_j)$, $\forall j \in \{1, \ldots, n\}$
- Example: We are going to show that $\hat{u} = \delta_i/\gamma = 1/\gamma \begin{cases} 1 & \text{if } j = i \\ 0 & \text{else} \end{cases}$ is an Eigenfunction of ℓ^1 with Eigenvalue $\lambda = 1/\gamma$ if the columns of the $m \times n$-matrix K are normed with respect to the 2-norm

Eigenfunction-Example for $J(u) = \|u\|_{\ell^1}$

- Considering $K\hat{u}$ yields the i-th column of K multiplied with λ, i.e. $\lambda K_{m\times i}$
- Computing $\hat{q}_j = \lambda(K^T K_{m\times i})_j$ is equivalent to $\hat{q}_j = \lambda\langle K^T_{j\times m}, K_{m\times i}\rangle$; hence, for normalized columns of K we have $\hat{q}_i = 1$, while $|\hat{q}_j| \leq 1$, $\forall j \neq i$

- As a last example we want to consider $J(u) = \|u\|_{\ell^1}$; for an element $q \in \partial\|u\|_{\ell^1}$ we have $q_j = \operatorname{sign}(u_j)$, $\forall j \in \{1, \dots, n\}$
- **Example:** We are going to show that $\hat{u} = \delta_i/\gamma = 1/\gamma \begin{cases} 1 & \text{if } j = i \\ 0 & \text{else} \end{cases}$ is an Eigenfunction of ℓ^1 with Eigenvalue $\lambda = 1/\gamma$ if the columns of the $m \times n$-matrix K are normed with respect to the 2-norm

Eigenfunction-Example for $J(u) = \|u\|_{\ell^1}$

- Considering $K\hat{u}$ yields the i-th column of K multiplied with λ, i.e. $\lambda K_{m\times i}$
- Computing $\hat{q}_j = \lambda(K^T K_{m\times i})_j$ is equivalent to $\hat{q}_j = \lambda\langle K^T_{j\times m}, K_{m\times i}\rangle$; hence, for normalized columns of K we have $\hat{q}_i = 1$, while $|\hat{q}_j| \leq 1$, $\forall j \neq i \Rightarrow \hat{q} \in \partial\|\hat{u}\|_{\ell^1}$

- What are Eigenfunctions good for?

- What are Eigenfunctions good for?
  - We can recover them exactly up to a constant factor for an appropriately chosen regularization parameter

- What are Eigenfunctions good for?
  - We can recover them exactly up to a constant factor for an appropriately chosen regularization parameter
  - With the use of *Inverse Scale Space*-methods (ISS) we can recover them exactly, even in the presence of noise!

- What are Eigenfunctions good for?
  - We can recover them exactly up to a constant factor for an appropriately chosen regularization parameter
  - With the use of *Inverse Scale Space*-methods (ISS) we can recover them exactly, even in the presence of noise! (but only if the regularization functional is one-homogeneous and under appropriate conditions on the noise)

- What are Eigenfunctions good for?
  - We can recover them exactly up to a constant factor for an appropriately chosen regularization parameter
  - With the use of *Inverse Scale Space*-methods (ISS) we can recover them exactly, even in the presence of noise! (but only if the regularization functional is one-homogeneous and under appropriate conditions on the noise)

Clean Data

Let $J : \mathrm{dom}(J) \subseteq L^2(\Omega) \to \mathbb{R} \cup \{\infty\}$ be a convex and one-homogeneous functional and let $K : L^2(\Omega) \to L^2(\Sigma)$ be a linear operator. Furthermore, let $\hat{u}$ be an Eigenfunction of J with corresponding Eigenvalue λ. Then, if the data f is given via $f = \gamma K \hat{u}$ for a positive constant γ, the solution of (9) is $u = c\hat{u}$ with

$$c = \gamma - \alpha\lambda\,, \tag{14}$$

if $\gamma > \alpha\lambda$ is satisfied.

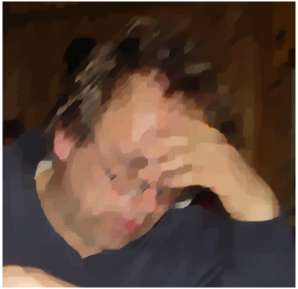
- How to prove it?

- How to prove it? Just write it down!

- How to prove it? Just write it down!

Proof

$$
\begin{aligned}
u &= \arg\min_u \left\{ \frac{1}{2}\|Ku - \gamma K\hat{u}\|^2 + \alpha J(u) \right\} \\
&= \arg\min_u \left\{ \frac{1}{2}\|Ku - cK\hat{u}\|^2 + \alpha J(u) - \alpha J(c\hat{u}) \right. \\
&\qquad \left. - \frac{\gamma - c}{\lambda}\langle \lambda K^* K\hat{u}, u - c\hat{u}\rangle \right\}
\end{aligned}
$$



- How to prove it? Just write it down!

Proof

$$u = \arg\min_u \left\{ \frac{1}{2} \|Ku - \gamma K\hat{u}\|^2 + \alpha J(u) \right\}$$
$$= \arg\min_u \left\{ \frac{1}{2} \|Ku - cK\hat{u}\|^2 + \alpha J(u) - \alpha J(c\hat{u}) - \frac{\gamma - c}{\lambda} \langle \lambda K^* K\hat{u}, u - c\hat{u} \rangle \right\}$$

- For $c = \gamma - \alpha\lambda > 0$ we have

- How to prove it? Just write it down!

Proof

$$
\begin{aligned}
u &= \arg\min_u \left\{ \frac{1}{2} \|Ku - \gamma K\hat{u}\|^2 + \alpha J(u) \right\} \\
&= \arg\min_u \left\{ \frac{1}{2} \|Ku - cK\hat{u}\|^2 + \alpha J(u) - \alpha J(c\hat{u}) \right. \\
&\qquad \left. - \frac{\gamma - c}{\lambda} \langle \lambda K^* K\hat{u}, u - c\hat{u} \rangle \right\}
\end{aligned}
$$

- For $c = \gamma - \alpha\lambda > 0$ we have

Crucial Point

$$
u = \arg\min_u \left\{ \frac{1}{2} \|Ku - cK\hat{u}\|^2 + \alpha D_J^q(u, c\hat{u}) \right\} \Rightarrow u = c\hat{u}
$$

- How to prove it? Just write it down!

Proof

$$u = \arg\min_u \left\{ \frac{1}{2}\|Ku - \gamma K\hat{u}\|^2 + \alpha J(u) \right\}$$
$$= \arg\min_u \left\{ \frac{1}{2}\|Ku - cK\hat{u}\|^2 + \alpha J(u) - \alpha J(c\hat{u}) - \frac{\gamma - c}{\lambda}\langle \lambda K^* K\hat{u}, u - c\hat{u}\rangle \right\}$$

- For $c = \gamma - \alpha\lambda > 0$ we have (with $q = \lambda K^* K\hat{u} \in \partial J(\hat{u}) = \partial J(c\hat{u})$)

Crucial Point

$$u = \arg\min_u \left\{ \frac{1}{2}\|Ku - cK\hat{u}\|^2 + \alpha D_J^q(u, c\hat{u}) \right\} \Rightarrow u = c\hat{u}$$

- Under appropriate assumptions on the noise we can also recover multiples of Eigenfunctions in the presence of noise

- Under appropriate assumptions on the noise we can also recover multiples of Eigenfunctions in the presence of noise

Noisy Data

Let the same assumptions hold as in the previous theorem. Furthermore, the data f is assumed to be corrupted by noise n, i.e. $f = \gamma K\hat{u} + n$ for a positive constant γ, such that there exist positive constants μ and η with

$$\mu K^* K\hat{u} + \eta K^* n \in \partial J(\hat{u}) \,. \tag{15}$$

Then, the solution of (9) is given via $u = c\hat{u}$ for

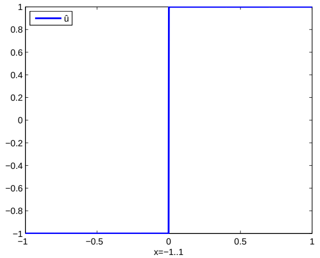
$$c = \gamma - \alpha\lambda + \frac{\lambda - \mu}{\eta} \,, \tag{16}$$

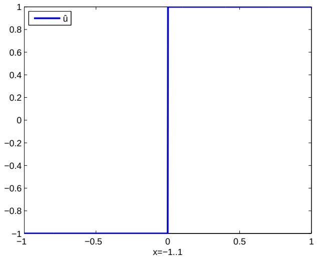
if γ satisfies the SNR-condition $\gamma > \mu/\eta$ and if $1/\eta \leq \alpha < \gamma/\lambda + 1/\eta$ holds.

› Examples

› Examples

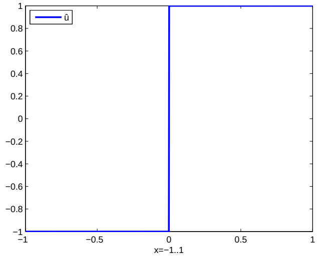
- We again consider $\hat{u}(x) = \text{sign}(x)$ for $x \in [-1, 1]$

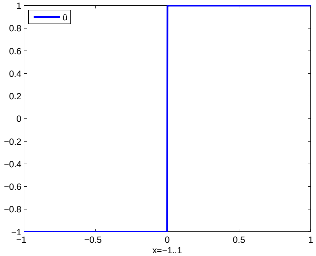




› Examples

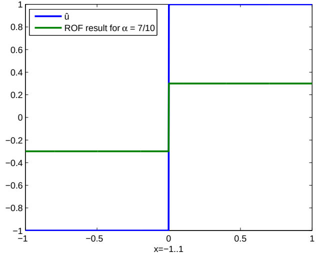
- We again consider $\hat{u}(x) = \text{sign}(x)$ for $x \in [-1, 1]$
- The following plots show the ROF-computations for $\alpha = 7/10, 3/10, 1/20$

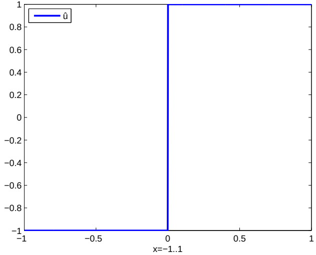




› Examples

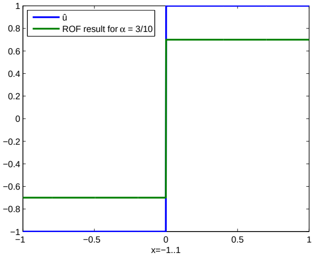
- We again consider $\hat{u}(x) = \text{sign}(x)$ for $x \in [-1, 1]$
- The following plots show the ROF-computations for $\alpha = 7/10, 3/10, 1/20$

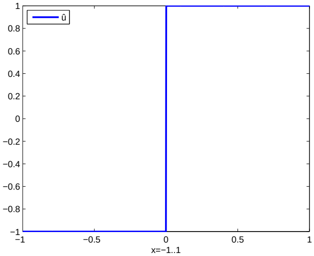




› Examples

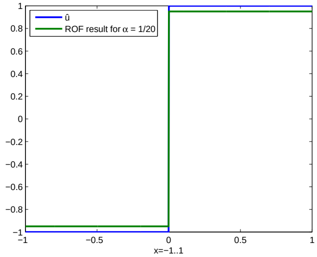
- We again consider $\hat{u}(x) = \text{sign}(x)$ for $x \in [-1, 1]$
- The following plots show the ROF-computations for $\alpha = 7/10, 3/10, 1/20$

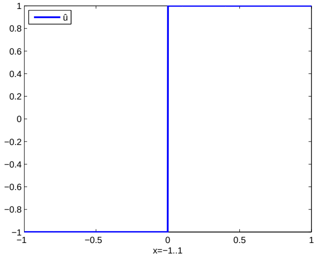




› Examples

- We again consider $\hat{u}(x) = \text{sign}(x)$ for $x \in [-1, 1]$
- The following plots show the ROF-computations for $\alpha = 7/10, 3/10, 1/20$

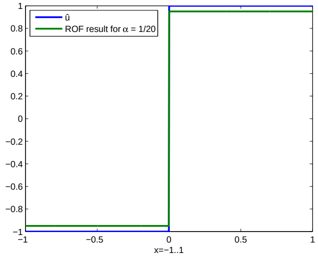


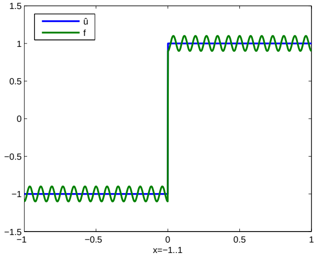


› Examples

- We again consider $\hat{u}(x) = \text{sign}(x)$ for $x \in [-1, 1]$
- The following plots show the ROF-computations for $\alpha = 7/10, 3/10, 1/20$

- Now we add noise $n(x) = -\cos(26\pi x)/10$

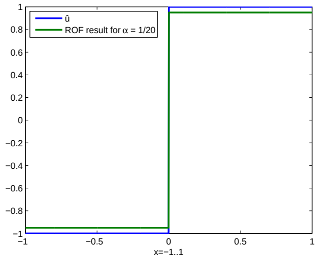


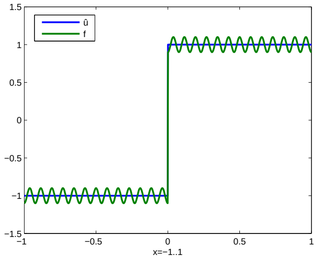


› Examples

- We again consider $\hat{u}(x) = \text{sign}(x)$ for $x \in [-1, 1]$
- The following plots show the ROF-computations for $\alpha = 7/10, 3/10, 1/20$

- Now we add noise $n(x) = -\cos(26\pi x)/10$
- Again, we show the ROF-computations for $\alpha = 7/10, 3/10, 1/20$

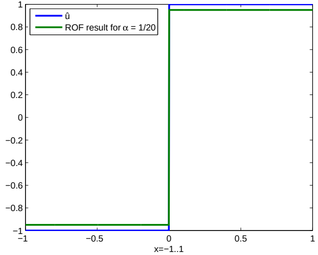


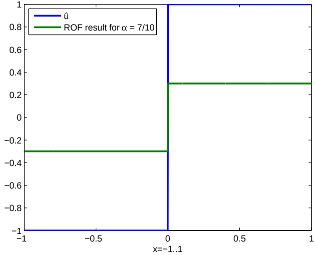


› Examples

- We again consider $\hat{u}(x) = \text{sign}(x)$ for $x \in [-1, 1]$
- The following plots show the ROF-computations for $\alpha = 7/10, 3/10, 1/20$

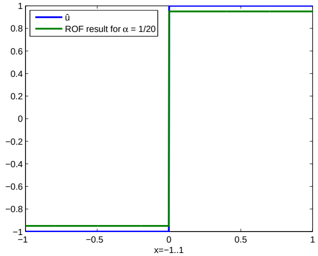
- Now we add noise $n(x) = -\cos(26\pi x)/10$
- Again, we show the ROF-computations for $\alpha = 7/10, 3/10, 1/20$

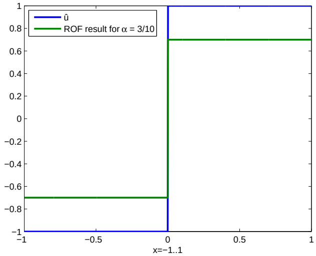




› Examples

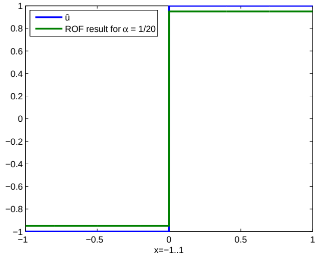
- We again consider $\hat{u}(x) = \text{sign}(x)$ for $x \in [-1, 1]$
- The following plots show the ROF-computations for $\alpha = 7/10, 3/10, 1/20$
- Now we add noise $n(x) = -\cos(26\pi x)/10$
- Again, we show the ROF-computations for $\alpha = 7/10, 3/10, 1/20$

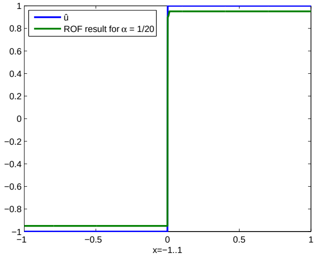




› Examples

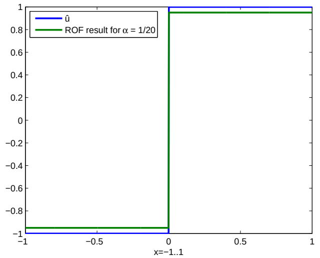
- We again consider $\hat{u}(x) = \text{sign}(x)$ for $x \in [-1, 1]$
- The following plots show the ROF-computations for $\alpha = 7/10, 3/10, 1/20$
- Now we add noise $n(x) = -\cos(26\pi x)/10$
- Again, we show the ROF-computations for $\alpha = 7/10, 3/10, 1/20$



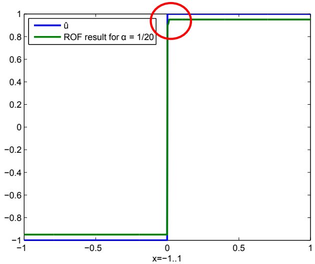


› Examples

- We again consider $\hat{u}(x) = \text{sign}(x)$ for $x \in [-1, 1]$
- The following plots show the ROF-computations for $\alpha = 7/10, 3/10, 1/20$



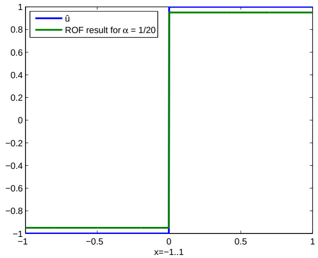
- Now we add noise $n(x) = -\cos(26\pi x)/10$
- Again, we show the ROF-computations for $\alpha = 7/10, 3/10, 1/20$

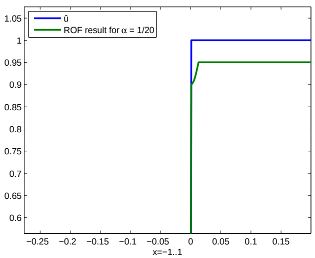


› Examples

- We again consider $\hat{u}(x) = \text{sign}(x)$ for $x \in [-1, 1]$
- The following plots show the ROF-computations for $\alpha = 7/10, 3/10, 1/20$

- Now we add noise $n(x) = -\cos(26\pi x)/10$
- Again, we show the ROF-computations for $\alpha = 7/10, 3/10, 1/20$





- In the noise-free case we have $c = \gamma - \alpha\lambda = 1 - \alpha$ as the loss of contrast predicted by (14), which is exactly what we observe

- In the noise-free case we have $c = \gamma - \alpha\lambda = 1 - \alpha$ as the loss of contrast predicted by (14), which is exactly what we observe
- In the noisy case, according to (16), the loss of contrast can be computed via $c = 1 - \alpha + (1 - \mu)/\eta$

- In the noise-free case we have $c = \gamma - \alpha\lambda = 1 - \alpha$ as the loss of contrast predicted by (14), which is exactly what we observe
- In the noisy case, according to (16), the loss of contrast can be computed via $c = 1 - \alpha + (1 - \mu)/\eta$
- Condition (15) reads as $\mu \operatorname{sign}(x) - \eta \cos(26\pi x)/10 \in \partial TV(\operatorname{sign}(x))$; we therefore consider q with

$$q(x) = \mu(1 - |x|) + (\eta \sin(26\pi x))/(260\pi)$$

- In the noise-free case we have $c = \gamma - \alpha\lambda = 1 - \alpha$ as the loss of contrast predicted by (14), which is exactly what we observe
- In the noisy case, according to (16), the loss of contrast can be computed via $c = 1 - \alpha + (1 - \mu)/\eta$
- Condition (15) reads as $\mu \text{sign}(x) - \eta \cos(26\pi x)/10 \in \partial TV(\text{sign}(x))$; we therefore consider q with

$$q(x) = \mu(1 - |x|) + (\eta \sin(26\pi x))/(260\pi)$$

Properties of q

- $|q(0)| = \mu$

- In the noise-free case we have $c = \gamma - \alpha\lambda = 1 - \alpha$ as the loss of contrast predicted by (14), which is exactly what we observe
- In the noisy case, according to (16), the loss of contrast can be computed via $c = 1 - \alpha + (1 - \mu)/\eta$
- Condition (15) reads as $\mu \text{sign}(x) - \eta \cos(26\pi x)/10 \in \partial TV(\text{sign}(x))$; we therefore consider q with

$$q(x) = \mu(1 - |x|) + (\eta \sin(26\pi x))/(260\pi)$$

Properties of q

- $|q(0)| = \mu \Rightarrow \mu = 1$

- In the noise-free case we have $c = \gamma - \alpha\lambda = 1 - \alpha$ as the loss of contrast predicted by (14), which is exactly what we observe
- In the noisy case, according to (16), the loss of contrast can be computed via $c = 1 - \alpha + (1 - \mu)/\eta$
- Condition (15) reads as $\mu \text{sign}(x) - \eta \cos(26\pi x)/10 \in \partial \text{TV}(\text{sign}(x))$; we therefore consider q with

$$q(x) = \mu(1 - |x|) + (\eta \sin(26\pi x))/(260\pi)$$

Properties of q

- $|q(0)| = \mu \Rightarrow \mu = 1 \Rightarrow c = 1 - \alpha$

- In the noise-free case we have $c = \gamma - \alpha\lambda = 1 - \alpha$ as the loss of contrast predicted by (14), which is exactly what we observe
- In the noisy case, according to (16), the loss of contrast can be computed via $c = 1 - \alpha + (1 - \mu)/\eta$
- Condition (15) reads as $\mu \text{sign}(x) - \eta \cos(26\pi x)/10 \in \partial TV(\text{sign}(x))$; we therefore consider q with

$$q(x) = \mu(1 - |x|) + (\eta \sin(26\pi x))/(260\pi)$$

Properties of q

- $|q(0)| = \mu \Rightarrow \mu = 1 \Rightarrow c = 1 - \alpha$
- $q(-1) = q(1) = 0$

- In the noise-free case we have $c = \gamma - \alpha\lambda = 1 - \alpha$ as the loss of contrast predicted by (14), which is exactly what we observe
- In the noisy case, according to (16), the loss of contrast can be computed via $c = 1 - \alpha + (1 - \mu)/\eta$
- Condition (15) reads as $\mu \text{sign}(x) - \eta \cos(26\pi x)/10 \in \partial \text{TV}(\text{sign}(x))$; we therefore consider q with

$$q(x) = \mu(1 - |x|) + (\eta \sin(26\pi x))/(260\pi)$$

Properties of q

- $|q(0)| = \mu \Rightarrow \mu = 1 \Rightarrow c = 1 - \alpha$
- $q(-1) = q(1) = 0$
- $\langle q', \hat{u} \rangle = \text{TV}(\hat{u}) = 2$

- In the noise-free case we have $c = \gamma - \alpha\lambda = 1 - \alpha$ as the loss of contrast predicted by (14), which is exactly what we observe
- In the noisy case, according to (16), the loss of contrast can be computed via $c = 1 - \alpha + (1 - \mu)/\eta$
- Condition (15) reads as $\mu \operatorname{sign}(x) - \eta \cos(26\pi x)/10 \in \partial \mathrm{TV}(\operatorname{sign}(x))$; we therefore consider q with

$$q(x) = \mu(1 - |x|) + (\eta \sin(26\pi x))/(260\pi)$$

Properties of q

- $|q(0)| = \mu \Rightarrow \mu = 1 \Rightarrow c = 1 - \alpha$
- $q(-1) = q(1) = 0$
- $\langle q', \hat{u} \rangle = \mathrm{TV}(\hat{u}) = 2$
- What about $\|q\|_\infty \leq 1$?

- In the noise-free case we have $c = \gamma - \alpha\lambda = 1 - \alpha$ as the loss of contrast predicted by (14), which is exactly what we observe
- In the noisy case, according to (16), the loss of contrast can be computed via $c = 1 - \alpha + (1 - \mu)/\eta$
- Condition (15) reads as $\mu \operatorname{sign}(x) - \eta \cos(26\pi x)/10 \in \partial \mathrm{TV}(\operatorname{sign}(x))$; we therefore consider q with

$$q(x) = \mu(1 - |x|) + (\eta \sin(26\pi x))/(260\pi)$$

Properties of q

- $|q(0)| = \mu \Rightarrow \mu = 1 \Rightarrow c = 1 - \alpha$
- $q(-1) = q(1) = 0$
- $\langle q', \hat{u} \rangle = \mathrm{TV}(\hat{u}) = 2$
- What about $\|q\|_\infty \leq 1$?

How to choose μ ?

- $\alpha = 7/10$

- In the noise-free case we have $c = \gamma - \alpha\lambda = 1 - \alpha$ as the loss of contrast predicted by (14), which is exactly what we observe
- In the noisy case, according to (16), the loss of contrast can be computed via $c = 1 - \alpha + (1 - \mu)/\eta$
- Condition (15) reads as $\mu \operatorname{sign}(x) - \eta \cos(26\pi x)/10 \in \partial \mathrm{TV}(\operatorname{sign}(x))$; we therefore consider q with

$$q(x) = \mu(1 - |x|) + (\eta \sin(26\pi x))/(260\pi)$$

Properties of q

- $|q(0)| = \mu \Rightarrow \mu = 1 \Rightarrow c = 1 - \alpha$
- $q(-1) = q(1) = 0$
- $\langle q', \hat{u} \rangle = \mathrm{TV}(\hat{u}) = 2$
- What about $\|q\|_\infty \leq 1$?

How to choose μ ?

- $\alpha = 7/10 \Rightarrow \eta \geq 10/7$

- In the noise-free case we have $c = \gamma - \alpha\lambda = 1 - \alpha$ as the loss of contrast predicted by (14), which is exactly what we observe
- In the noisy case, according to (16), the loss of contrast can be computed via $c = 1 - \alpha + (1 - \mu)/\eta$
- Condition (15) reads as $\mu \text{sign}(x) - \eta \cos(26\pi x)/10 \in \partial \text{TV}(\text{sign}(x))$; we therefore consider q with

$$q(x) = \mu(1 - |x|) + (\eta \sin(26\pi x))/(260\pi)$$

Properties of q

- $|q(0)| = \mu \Rightarrow \mu = 1 \Rightarrow c = 1 - \alpha$
- $q(-1) = q(1) = 0$
- $\langle q', \hat{u} \rangle = \text{TV}(\hat{u}) = 2$
- What about $\|q\|_\infty \leq 1$?

How to choose μ ?

- $\alpha = 7/10 \Rightarrow \eta \geq 10/7 \Rightarrow \|q\|_\infty \leq 1$

- In the noise-free case we have $c = \gamma - \alpha\lambda = 1 - \alpha$ as the loss of contrast predicted by (14), which is exactly what we observe
- In the noisy case, according to (16), the loss of contrast can be computed via $c = 1 - \alpha + (1-\mu)/\eta$
- Condition (15) reads as $\mu \operatorname{sign}(x) - \eta\cos(26\pi x)/10 \in \partial \mathrm{TV}(\operatorname{sign}(x))$; we therefore consider q with

$$q(x) = \mu(1 - |x|) + (\eta \sin(26\pi x))/(260\pi)$$

Properties of q

- $|q(0)| = \mu \Rightarrow \mu = 1 \Rightarrow c = 1 - \alpha$
- $q(-1) = q(1) = 0$
- $\langle q', \hat{u} \rangle = \mathrm{TV}(\hat{u}) = 2$
- What about $\|q\|_\infty \leq 1$?

How to choose μ ?

- $\alpha = 7/10 \Rightarrow \eta \geq 10/7 \Rightarrow \|q\|_\infty \leq 1$
- $\alpha = 3/10 \Rightarrow \eta \geq 10/3 \Rightarrow \|q\|_\infty \leq 1$

- In the noise-free case we have $c = \gamma - \alpha\lambda = 1 - \alpha$ as the loss of contrast predicted by (14), which is exactly what we observe
- In the noisy case, according to (16), the loss of contrast can be computed via $c = 1 - \alpha + (1 - \mu)/\eta$
- Condition (15) reads as $\mu \text{sign}(x) - \eta \cos(26\pi x)/10 \in \partial \text{TV}(\text{sign}(x))$; we therefore consider q with

$$q(x) = \mu(1 - |x|) + (\eta \sin(26\pi x))/(260\pi)$$

Properties of q

- $|q(0)| = \mu \Rightarrow \mu = 1 \Rightarrow c = 1 - \alpha$
- $q(-1) = q(1) = 0$
- $\langle q', \hat{u} \rangle = \text{TV}(\hat{u}) = 2$
- What about $\|q\|_\infty \leq 1$?

How to choose μ ?

- $\alpha = 7/10 \Rightarrow \eta \geq 10/7 \Rightarrow \|q\|_\infty \leq 1$
- $\alpha = 3/10 \Rightarrow \eta \geq 10/3 \Rightarrow \|q\|_\infty \leq 1$
- $\alpha = 1/20 \Rightarrow \eta \geq 20$

- In the noise-free case we have $c = \gamma - \alpha\lambda = 1 - \alpha$ as the loss of contrast predicted by (14), which is exactly what we observe
- In the noisy case, according to (16), the loss of contrast can be computed via $c = 1 - \alpha + (1 - \mu)/\eta$
- Condition (15) reads as $\mu \operatorname{sign}(x) - \eta \cos(26\pi x)/10 \in \partial \mathrm{TV}(\operatorname{sign}(x))$; we therefore consider q with

$$q(x) = \mu(1 - |x|) + (\eta \sin(26\pi x))/(260\pi)$$

Properties of q

- $|q(0)| = \mu \Rightarrow \mu = 1 \Rightarrow c = 1 - \alpha$
- $q(-1) = q(1) = 0$
- $\langle q', \hat{u}\rangle = \mathrm{TV}(\hat{u}) = 2$
- What about $\|q\|_\infty \leq 1$?

How to choose μ ?

- $\alpha = 7/10 \Rightarrow \eta \geq 10/7 \Rightarrow \|q\|_\infty \leq 1$
- $\alpha = 3/10 \Rightarrow \eta \geq 10/3 \Rightarrow \|q\|_\infty \leq 1$
- $\alpha = 1/20 \Rightarrow \eta \geq 20 \Rightarrow \|q\|_\infty > 1$

- In the noise-free case we have $c = \gamma - \alpha\lambda = 1 - \alpha$ as the loss of contrast predicted by (14), which is exactly what we observe
- In the noisy case, according to (16), the loss of contrast can be computed via $c = 1 - \alpha + (1 - \mu)/\eta$
- Condition (15) reads as $\mu \text{sign}(x) - \eta \cos(26\pi x)/10 \in \partial \text{TV}(\text{sign}(x))$; we therefore consider q with

$$q(x) = \mu(1 - |x|) + (\eta \sin(26\pi x))/(260\pi)$$

Properties of q

- $|q(0)| = \mu \Rightarrow \mu = 1 \Rightarrow c = 1 - \alpha$
- $q(-1) = q(1) = 0$
- $\langle q', \hat{u} \rangle = \text{TV}(\hat{u}) = 2$
- What about $\|q\|_\infty \leq 1$?

How to choose μ ?

- $\alpha = 7/10 \Rightarrow \eta \geq 10/7 \Rightarrow \|q\|_\infty \leq 1$
- $\alpha = 3/10 \Rightarrow \eta \geq 10/3 \Rightarrow \|q\|_\infty \leq 1$
- $\alpha = 1/20 \Rightarrow \eta \geq 20 \Rightarrow \|q\|_\infty > 1$
- Condition (15) is violated!

- For one-homogeneous functionals the ISS allows us to exactly recover Eigenfunctions in finite time

- For one-homogeneous functionals the ISS allows us to exactly recover Eigenfunctions in finite time

Clean Data

Let $J : \mathrm{dom}(J) \subseteq L^2(\Omega) \to \mathbb{R} \cup \{\infty\}$ be a convex and one-homogeneous functional and let $K : L^2(\Omega) \to L^2(\Sigma)$ be a linear operator. Furthermore, let $\hat{u}$ be an Eigenfunction with corresponding Eigenvalue λ. Then, if the data f is given via $f = \gamma K\hat{u}$ for a positive constant γ, the solution of the Inverse Scale Space Flow (10) at time $t > t_* = \lambda/\gamma$ is $u(t) = \hat{u}$.

- For one-homogeneous functionals the ISS allows us to exactly recover Eigenfunctions in finite time

Clean Data

Let $J : \mathrm{dom}(J) \subseteq L^2(\Omega) \to \mathbb{R} \cup \{\infty\}$ be a convex and one-homogeneous functional and let $K : L^2(\Omega) \to L^2(\Sigma)$ be a linear operator. Furthermore, let $\hat{u}$ be an Eigenfunction with corresponding Eigenvalue λ. Then, if the data f is given via $f = \gamma K \hat{u}$ for a positive constant γ, the solution of the Inverse Scale Space Flow (10) at time $t > t_* = \lambda / \gamma$ is $u(t) = \hat{u}$.

How to prove?

- For one-homogeneous functionals the ISS allows us to exactly recover Eigenfunctions in finite time

Clean Data

Let $J : \mathrm{dom}(J) \subseteq L^2(\Omega) \to \mathbb{R} \cup \{\infty\}$ be a convex and one-homogeneous functional and let $K : L^2(\Omega) \to L^2(\Sigma)$ be a linear operator. Furthermore, let $\hat{u}$ be an Eigenfunction with corresponding Eigenvalue λ. Then, if the data f is given via $f = \gamma K \hat{u}$ for a positive constant γ, the solution of the Inverse Scale Space Flow (10) at time $t > t_* = \lambda/\gamma$ is $u(t) = \hat{u}$.

How to prove? Yes, just write it down again!

- For one-homogeneous functionals the ISS allows us to exactly recover Eigenfunctions in finite time

Clean Data

Let $J : \mathrm{dom}(J) \subseteq L^2(\Omega) \to \mathbb{R} \cup \{\infty\}$ be a convex and one-homogeneous functional and let $K : L^2(\Omega) \to L^2(\Sigma)$ be a linear operator. Furthermore, let $\hat{u}$ be an Eigenfunction with corresponding Eigenvalue λ. Then, if the data f is given via $f = \gamma K \hat{u}$ for a positive constant γ, the solution of the Inverse Scale Space Flow (10) at time $t > t_* = \lambda / \gamma$ is $u(t) = \hat{u}$.

How to prove? Yes, just write it down again!

Proof

- Yves Meyer $\Rightarrow u \equiv 0$ for $t < t_*$

- For one-homogeneous functionals the ISS allows us to exactly recover Eigenfunctions in finite time

Clean Data

Let $J : \mathrm{dom}(J) \subseteq L^2(\Omega) \to \mathbb{R} \cup \{\infty\}$ be a convex and one-homogeneous functional and let $K : L^2(\Omega) \to L^2(\Sigma)$ be a linear operator. Furthermore, let $\hat{u}$ be an Eigenfunction with corresponding Eigenvalue λ. Then, if the data f is given via $f = \gamma K \hat{u}$ for a positive constant γ, the solution of the Inverse Scale Space Flow (10) at time $t > t_* = \lambda / \gamma$ is $u(t) = \hat{u}$.

How to prove? Yes, just write it down again!

Proof

- Yves Meyer $\Rightarrow u \equiv 0$ for $t < t_* \Rightarrow \partial_t p = K^* f = \gamma K^* K \hat{u}$

- For one-homogeneous functionals the ISS allows us to exactly recover Eigenfunctions in finite time

Clean Data

Let $J : \mathrm{dom}(J) \subseteq L^2(\Omega) \to \mathbb{R} \cup \{\infty\}$ be a convex and one-homogeneous functional and let $K : L^2(\Omega) \to L^2(\Sigma)$ be a linear operator. Furthermore, let $\hat{u}$ be an Eigenfunction with corresponding Eigenvalue λ. Then, if the data f is given via $f = \gamma K \hat{u}$ for a positive constant γ, the solution of the Inverse Scale Space Flow (10) at time $t > t_* = \lambda / \gamma$ is $u(t) = \hat{u}$.

How to prove? Yes, just write it down again!

Proof

- Yves Meyer $\Rightarrow u \equiv 0$ for $t < t_* \Rightarrow \partial_t p = K^* f = \gamma K^* K \hat{u}$
- $p(t) = t \gamma K^* K \hat{u}$

- For one-homogeneous functionals the ISS allows us to exactly recover Eigenfunctions in finite time

Clean Data

Let $J : \mathrm{dom}(J) \subseteq L^2(\Omega) \to \mathbb{R} \cup \{\infty\}$ be a convex and one-homogeneous functional and let $K : L^2(\Omega) \to L^2(\Sigma)$ be a linear operator. Furthermore, let $\hat{u}$ be an Eigenfunction with corresponding Eigenvalue λ. Then, if the data f is given via $f = \gamma K \hat{u}$ for a positive constant γ, the solution of the Inverse Scale Space Flow (10) at time $t > t_* = \lambda/\gamma$ is $u(t) = \hat{u}$.

How to prove? Yes, just write it down again!

Proof

- Yves Meyer $\Rightarrow u \equiv 0$ for $t < t_* \Rightarrow \partial_t p = K^* f = \gamma K^* K \hat{u}$
- $p(t) = t\gamma K^* K \hat{u} \Rightarrow t_* = \lambda/\gamma$

- For one-homogeneous functionals the ISS allows us to exactly recover Eigenfunctions in finite time

Clean Data

Let $J : \mathrm{dom}(J) \subseteq L^2(\Omega) \rightarrow \mathbb{R} \cup \{\infty\}$ be a convex and one-homogeneous functional and let $K : L^2(\Omega) \rightarrow L^2(\Sigma)$ be a linear operator. Furthermore, let $\hat{u}$ be an Eigenfunction with corresponding Eigenvalue λ. Then, if the data f is given via $f = \gamma K\hat{u}$ for a positive constant γ, the solution of the Inverse Scale Space Flow (10) at time $t > t_* = \lambda/\gamma$ is $u(t) = \hat{u}$.

How to prove? Yes, just write it down again!

Proof

- Yves Meyer $\Rightarrow u \equiv 0$ for $t < t_* \Rightarrow \partial_t p = K^* f = \gamma K^* K\hat{u}$
- $p(t) = t\gamma K^* K\hat{u} \Rightarrow t_* = \lambda/\gamma$ (continuous extension)

- Even in the presence of noise exact recovery can be possible

- Even in the presence of noise exact recovery can be possible

Noisy Data

Let the same assumptions hold as in the previous theorem. Furthermore the data f is assumed to be corrupted by noise n, i.e. $f = \gamma K\hat{u} + n$ for a positive constant γ, such that there exist positive constants μ and η that satisfy (15). Then, the solution of the Inverse Scale Space Flow (10) for time $t_* = (\lambda\eta)/(\gamma\eta + \lambda - \mu) \leq t < t_{**}$ is given via $u(t) = c\hat{u}$ for

$$c = \gamma + \frac{\lambda - \mu}{\eta} . \tag{17}$$

- Even in the presence of noise exact recovery can be possible

Noisy Data

Let the same assumptions hold as in the previous theorem. Furthermore the data f is assumed to be corrupted by noise n, i.e. $f = \gamma K\hat{u} + n$ for a positive constant γ, such that there exist positive constants μ and η that satisfy (15). Then, the solution of the Inverse Scale Space Flow (10) for time $t_* = (\lambda\eta)/(\gamma\eta + \lambda - \mu) \leq t < t_{**}$ is given via $u(t) = c\hat{u}$ for

$$c = \gamma + \frac{\lambda - \mu}{\eta} \,. \tag{17}$$

- Remarkable:

- Even in the presence of noise exact recovery can be possible

Noisy Data

Let the same assumptions hold as in the previous theorem. Furthermore the data f is assumed to be corrupted by noise n, i.e. $f = \gamma K\hat{u} + n$ for a positive constant γ, such that there exist positive constants μ and η that satisfy (15). Then, the solution of the Inverse Scale Space Flow (10) for time $t_* = (\lambda\eta)/(\gamma\eta + \lambda - \mu) \leq t < t_{**}$ is given via $u(t) = c\hat{u}$ for

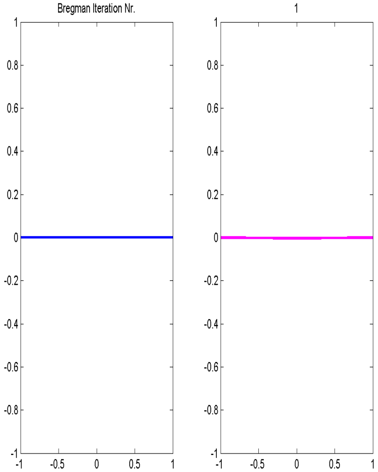
$$c = \gamma + \frac{\lambda - \mu}{\eta} \,. \tag{17}$$

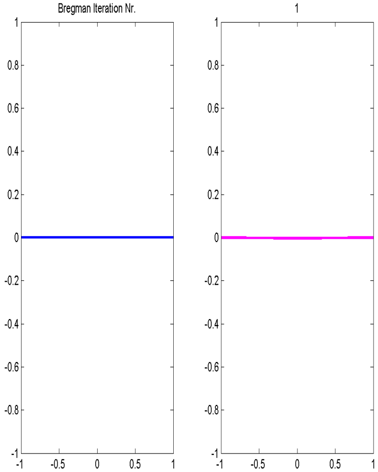
- Remarkable: assume $\gamma = 1$; for $\lambda = \mu$ we have $u(t) = \hat{u}$, no matter what value η takes (as long as η and μ satisfy (15))

Essen ist fertig!

› Examples

› Examples


Bregman Iteration Nr.
1
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-0.5
0.5


Bregman Iteration Nr.
1
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-0.5
0.5

› Adaptive Inverse Scale Space Method for Compressed Sensing

› Adaptive Inverse Scale Space Method for Compressed Sensing

- The ISS formulation can also be used to develop a fast and stable compressed sensing algorithm

CS Setup 1

$$u \in \arg\min_{u \in \ell^1} \left\{ \frac{1}{2} \|Ku - f\|_{\ell^2}^2 + \alpha \|u\|_{\ell^1} \right\} \tag{18}$$

› Adaptive Inverse Scale Space Method for Compressed Sensing

- The ISS formulation can also be used to develop a fast and stable compressed sensing algorithm

CS Setup 2

$$u_{k+1} \in \arg\min_{u \in \ell^1} \left\{ \frac{1}{2} \| Ku - f \|_{\ell^2}^2 + \alpha D_{\|\cdot\|_{\ell^1}}^{p_k}(u, u_k) \right\} \tag{18}$$

› Adaptive Inverse Scale Space Method for Compressed Sensing

- The ISS formulation can also be used to develop a fast and stable compressed sensing algorithm

CS Setup 2

$$u_{k+1} \in \arg\min_{u \in \ell^1} \left\{ \frac{1}{2} \|Ku - f\|_{\ell^2}^2 + \alpha D_{\|\cdot\|_{\ell^1}}^{p_k}(u, u_k) \right\} \tag{18}$$

- If we consider the related Inverse Scale Space Flow, i.e. $\partial_t p_i = (K^T(f - Ku))_i$, $(p \in \ell^\infty)$ we can also think of CS setup 2 as

CS Setup 2

$$\min \|u\|_{\ell^1} \text{ subject to } Ku = f \tag{19}$$

- We know $p \in \partial \|u\|_{\ell^1} \Leftrightarrow p_i = \text{sign}(u_i)$ and hence, for $u \not\equiv 0$ we have $\|p\|_{\ell^\infty} = 1$
- Due to Yves Meyer we know that for the ISS we have $u(t) \equiv 0$ for $t \in [0, t_1[$;

- We know $p \in \partial \|u\|_{\ell^1} \Leftrightarrow p_i = \text{sign}(u_i)$ and hence, for $u \not\equiv 0$ we have $\|p\|_{\ell^\infty} = 1$
- Due to Yves Meyer we know that for the ISS we have $u(t) \equiv 0$ for $t \in [0, t_1[$; hence, we can easily compute t_1 and $p(t_1)$ via

First non-trivial p at t_1

- Meyer $\Rightarrow u \equiv 0$ in $[0, t_1[$

- We know $p \in \partial \|u\|_{\ell^1} \Leftrightarrow p_i = \text{sign}(u_i)$ and hence, for $u \not\equiv 0$ we have $\|p\|_{\ell^\infty} = 1$
- Due to Yves Meyer we know that for the ISS we have $u(t) \equiv 0$ for $t \in [0, t_1[$; hence, we can easily compute t_1 and $p(t_1)$ via

First non-trivial p at t_1

- Meyer $\Rightarrow u \equiv 0$ in $[0, t_1[\Rightarrow \partial_t p_i = \left(K^T f\right)_i$ in $[0, t_1[$

- We know $p \in \partial \|u\|_{\ell^1} \Leftrightarrow p_i = \text{sign}(u_i)$ and hence, for $u \neq 0$ we have $\|p\|_{\ell^\infty} = 1$
- Due to Yves Meyer we know that for the ISS we have $u(t) \equiv 0$ for $t \in [0, t_1[$; hence, we can easily compute t_1 and $p(t_1)$ via

First non-trivial p at t_1

- Meyer $\Rightarrow u \equiv 0$ in $[0, t_1[\Rightarrow \partial_t p_i = \left(K^T f\right)_i$ in $[0, t_1[$
- Continuous extension $\Rightarrow p_i(t_1) = t_1 / \left(K^T f\right)_i$

- We know $p \in \partial \|u\|_{\ell^1} \Leftrightarrow p_i = \mathrm{sign}(u_i)$ and hence, for $u \neq 0$ we have $\|p\|_{\ell^\infty} = 1$
- Due to Yves Meyer we know that for the ISS we have $u(t) \equiv 0$ for $t \in [0, t_1[$; hence, we can easily compute t_1 and $p(t_1)$ via

First non-trivial p at t_1

- Meyer $\Rightarrow u \equiv 0$ in $[0, t_1[\Rightarrow \partial_t p_i = \left(K^T f\right)_i$ in $[0, t_1[$
- Continuous extension $\Rightarrow p_i(t_1) = t_1 / \left(K^T f\right)_i$
- $\|p(t_1)\|_{\ell^\infty} = 1 \Rightarrow t_1 = 1 / \left\|K^T f\right\|_{\ell^\infty}$

- We know $p \in \partial\|u\|_{\ell^1} \Leftrightarrow p_i = \text{sign}(u_i)$ and hence, for $u \not\equiv 0$ we have $\|p\|_{\ell^\infty} = 1$
- Due to Yves Meyer we know that for the ISS we have $u(t) \equiv 0$ for $t \in [0, t_1[$; hence, we can easily compute t_1 and $p(t_1)$ via

First non-trivial p at t_1

- Meyer $\Rightarrow u \equiv 0$ in $[0, t_1[\Rightarrow \partial_t p_i = \left(K^T f\right)_i$ in $[0, t_1[$
- Continuous extension $\Rightarrow p_i(t_1) = t_1 / \left(K^T f\right)_i$
- $\|p(t_1)\|_{\ell^\infty} = 1 \Rightarrow t_1 = 1 / \left\|K^T f\right\|_{\ell^\infty}$
- $p_i(t_1) = \left(K^T f\right)_i / \left\|K^T f\right\|_{\ell^\infty}$

- In order to determine $u(t_1)$ from $p(t_1)$ we want to define the set $I_1 := \{i \mid |p_i(t)| = 1\}$ and denote the projection on I_1 via P_{I_1}; then we can obtain the following result

- In order to determine $u(t_1)$ from $p(t_1)$ we want to define the set $I_1 := \{i \mid |p_i(t)| = 1\}$ and denote the projection on I_1 via P_{I_1}; then we can obtain the following result

Determining $u(t_1)$ from $p(t_1)$

We can determine $u(t_1)$ from

$$u = \arg\min_u \left\{ \|KP_{I_1}u - f\|_{\ell^2}^2 \right\} \tag{20}$$

subject to $P_{I_1^c}u = 0$ and $u_i p_i \geq 0$ for $i \in I_1$, with I_1^c denoting the complement of I_1

- Given $p(t_1)$ and $u(t_1)$, it actually turns out to be easy to proceed to $p(t_2)$

- Given $p(t_1)$ and $u(t_1)$, it actually turns out to be easy to proceed to $p(t_2)$

Update of the Dual Variable

- We know that there exists a time $t_2 > t_1$ such that u remains contstant for $t \in [t_1, t_2[$, i.e. $u(t) = u(t_1)$ in $[t_1, t_2[$

- Given $p(t_1)$ and $u(t_1)$, it actually turns out to be easy to proceed to $p(t_2)$

Update of the Dual Variable

- We know that there exists a time $t_2 > t_1$ such that u remains contstant for $t \in [t_1, t_2[$, i.e. $u(t) = u(t_1)$ in $[t_1, t_2[$
- Hence, we can construct

$$p_i(t) = p_i(t_1) + (t - t_1) \left(K^T (f - Ku(t_1))\right)_i \tag{21}$$

in order to satisfy $\partial_t p_i = \left(K^T (f - Ku(t_1))\right)_i$ and $p_i(t_1) = \left(K^T f\right)_i / \left\|K^T f\right\|_{\ell^\infty}$

- Given $p(t_1)$ and $u(t_1)$, it actually turns out to be easy to proceed to $p(t_2)$

Update of the Dual Variable

- We know that there exists a time $t_2 > t_1$ such that u remains contstant for $t \in [t_1, t_2[$, i.e. $u(t) = u(t_1)$ in $[t_1, t_2[$
- Hence, we can construct

$$p_i(t) = p_i(t_1) + (t - t_1)\left(K^T (f - Ku(t_1))\right)_i \tag{21}$$

in order to satisfy $\partial_t p_i = \left(K^T (f - Ku(t_1))\right)_i$ and $p_i(t_1) = \left(K^T f\right)_i / \left\|K^T f\right\|_{\ell^\infty}$

- Note that we have $P_{I_1} p(t) = P_{I_1} p(t_1)$ and therefore $\|P_{I_1} p(t)\|_{\ell^\infty} = 1$, while for small $t - t_1$ we have $\|P_{I_1^c} p(t)\|_{\ell^\infty} \leq 1$, since $\|P_{I_1^c} p(t_1)\|_{\ell^\infty} < 1$

- Given $p(t_1)$ and $u(t_1)$, it actually turns out to be easy to proceed to $p(t_2)$

Update of the Dual Variable

- We know that there exists a time $t_2 > t_1$ such that u remains contstant for $t \in [t_1, t_2[$, i.e. $u(t) = u(t_1)$ in $[t_1, t_2[$
- Hence, we can construct

$$p_i(t) = p_i(t_1) + (t - t_1)\left(K^T(f - Ku(t_1))\right)_i \tag{21}$$

in order to satisfy $\partial_t p_i = \left(K^T(f - Ku(t_1))\right)_i$ and $p_i(t_1) = \left(K^T f\right)_i / \left\|K^T f\right\|_{\ell^\infty}$

- Note that we have $P_{I_1} p(t) = P_{I_1} p(t_1)$ and therefore $\|P_{I_1} p(t)\|_{\ell^\infty} = 1$, while for small $t - t_1$ we have $\|P_{I_1^c} p(t)\|_{\ell^\infty} \leq 1$, since $\|P_{I_1^c} p(t_1)\|_{\ell^\infty} < 1$
- Idea:

- Given $p(t_1)$ and $u(t_1)$, it actually turns out to be easy to proceed to $p(t_2)$

Update of the Dual Variable

- We know that there exists a time $t_2 > t_1$ such that u remains contstant for $t \in [t_1, t_2[$, i.e. $u(t) = u(t_1)$ in $[t_1, t_2[$
- Hence, we can construct

$$p_i(t) = p_i(t_1) + (t - t_1) \left(K^T \left(f - Ku(t_1)\right)\right)_i \tag{21}$$

in order to satisfy $\partial_t p_i = \left(K^T \left(f - Ku(t_1)\right)\right)_i$ and $p_i(t_1) = \left(K^T f\right)_i / \left\|K^T f\right\|_{\ell^\infty}$

- Note that we have $P_{I_1} p(t) = P_{I_1} p(t_1)$ and therefore $\|P_{I_1} p(t)\|_{\ell^\infty} = 1$, while for small $t - t_1$ we have $\|P_{I_1^c} p(t)\|_{\ell^\infty} \leq 1$, since $\|P_{I_1^c} p(t_1)\|_{\ell^\infty} < 1$
- Idea: find t_2 as the minimal time t such that

$$\left\|P_{I_1^c} \left(p(t_1) + (t - t_1) K^T \left(f - Ku(t_1)\right)\right)\right\|_{\ell^\infty} = 1 \tag{22}$$

- We can summarize the previous considerations to the following algorithm

- We can summarize the previous considerations to the following algorithm

Adaptive Inverse Scale Space for Compressed Sensing

Algorithm 2 Adaptive Inverse Scale Space for Compressed Sensing

1. **Parameters:** K, f, $\delta \geq 0$
2. **Initialization:** $t_1 = 1/\left\|K^T f\right\|_{\ell^\infty}$, $p(t_1) = t_1 K^T f$, $I_1 = \{i \mid |p_i(t_1)| = 1\}$

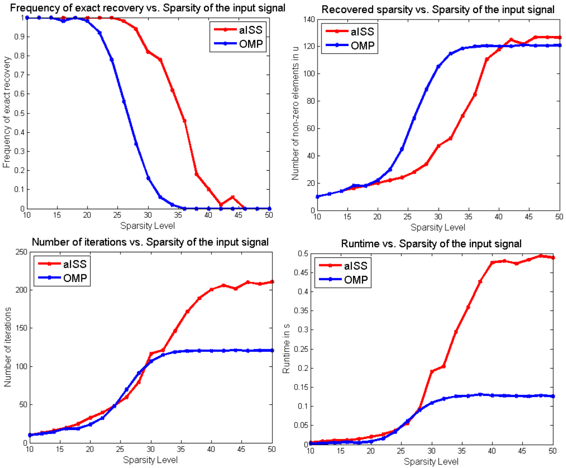
while $\|Ku - f\|_{\ell^2} \leq \delta$ **do**

  Compute $u(t_k)$ from (20) with P_{I_k}
  Obtain t_{k+1} as the minimal time for which (22) holds
  Update the dual variable via (21) with $t = t_{k+1}$
  Compute $I_{k+1} = \{i \mid |p_i(t_{k+1})| = 1\}$

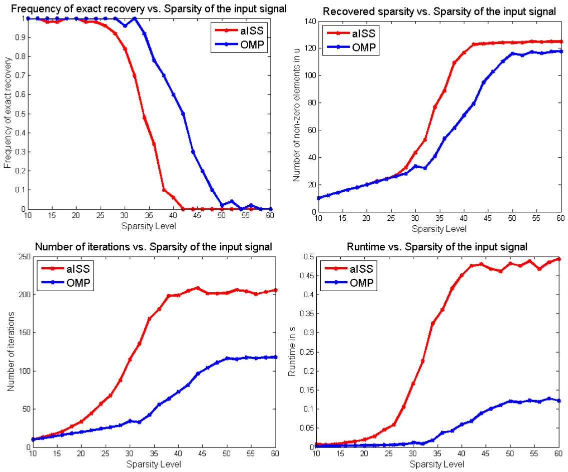
end while
return $u(t_k)$

› Computational Results

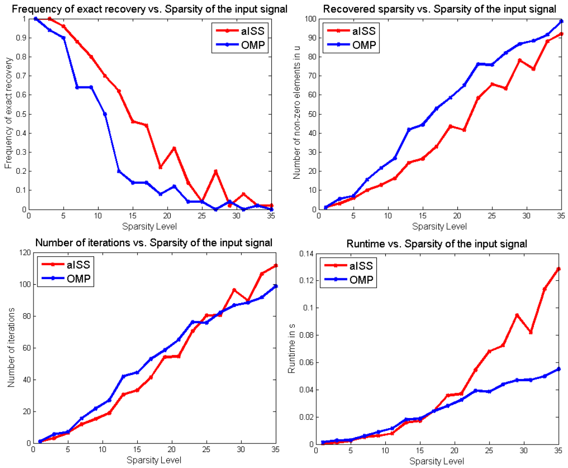
› Computational Results



› Computational Results



› Computational Results



Thank you for attention!