Proximal Operators for Nonnegative Inverse Problems

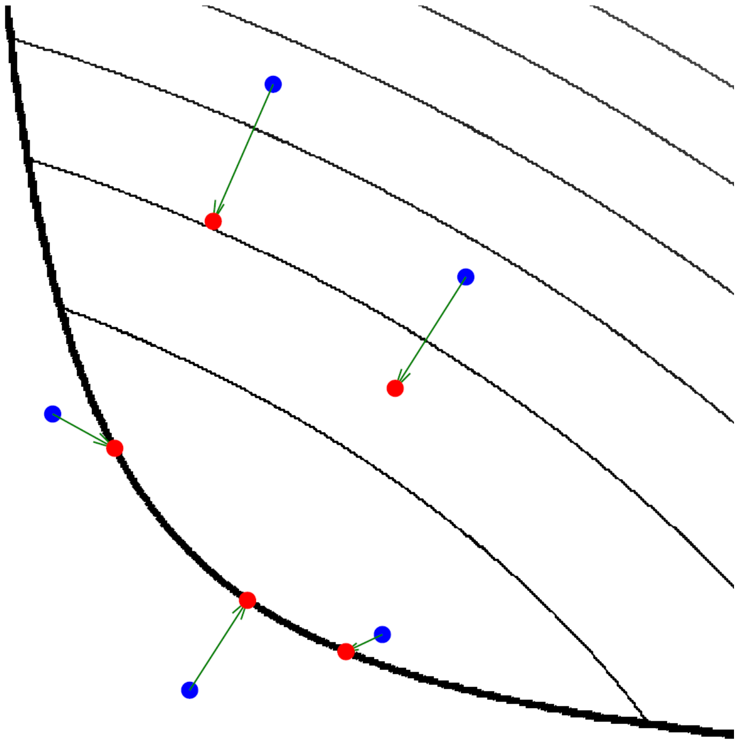
Alejandro Noguerón

Supervised by

Dr. Pol del Aguila Pla

Prof. Michael Unser

Content



Parikh, Boyd, *Proximal Algorithms,* Foundations and Trends in Optimization, vol 1 (2013)

Imaging as an Inverse Problem

An imaging problem takes the form:

$$\mathbf{y} = \mathrm{H}\,\mathbf{x} + \mathbf{w}$$

Where the challenge is to recover **x** from the measurements **y.**

Approaches:

- Minimizing noise
- Introducing prior knowledge

$$\arg\min_{\mathbf{x}\in\mathbb{R}} ||\mathbf{H}\mathbf{x} - \mathbf{y}||_2^2 + \mathcal{R}(\mathbf{L}\mathbf{x}) + \delta_{\mathbb{R}_+^{\mathrm{N}}}(\mathbf{x})$$

Imaging as an Inverse Problem

An imaging problem takes the form:

$$\mathbf{y} = \mathrm{H}\,\mathbf{x} + \mathbf{w}$$

Where the challenge is to recover **x** from the measurements **y.**

Approaches:

- Minimizing noise
- Introducing prior knowledge

$$\arg\min_{\mathbf{x}\in\mathbb{R}} ||\mathbf{H}\mathbf{x} - \mathbf{y}||_2^2 + \mathcal{R}(\mathbf{L}\mathbf{x}) + \delta_{\mathbb{R}_+^{\mathrm{N}}}(\mathbf{x})$$

smooth, convex

Imaging as an Inverse Problem

An imaging problem takes the form:

$$\mathbf{y} = \mathrm{H}\,\mathbf{x} + \mathbf{w}$$

Where the challenge is to recover **x** from the measurements **y.**

Approaches:

- Minimizing noise
- Introducing prior knowledge

$$\arg\min_{\mathbf{x}\in\mathbb{R}} \|\mathbf{H}\mathbf{x} - \mathbf{y}\|_2^2 + \mathcal{R}(\mathbf{L}\mathbf{x}) + \delta_{\mathbb{R}_+^{\mathrm{N}}}(\mathbf{x})$$

Imaging as an Inverse Problem

An imaging problem takes the form:

$$\mathbf{y} = \mathrm{H}\,\mathbf{x} + \mathbf{w}$$

Where the challenge is to recover **x** from the measurements **y.**

Approaches:

- Minimizing noise
- Introducing prior knowledge

$$\arg\min_{\mathbf{x}\in\mathbb{R}} \|\mathbf{H}\mathbf{x} - \mathbf{y}\|_2^2 + \mathcal{R}(\mathbf{L}\mathbf{x}) + \delta_{\mathbb{R}_+^{\mathrm{N}}}(\mathbf{x})$$

non-smooth, convex

Imaging as an Inverse Problem

An imaging problem takes the form:

$$\mathbf{y} = \mathrm{H}\,\mathbf{x} + \mathbf{w}$$

Where the challenge is to recover **x** from the measurements **y.**

Approaches:

- Minimizing noise
- Introducing prior knowledge

$$\arg\min_{\mathbf{x}\in\mathbb{R}} \|\mathbf{H}\mathbf{x} - \mathbf{y}\|_2^2 + \mathcal{R}(\mathbf{L}\mathbf{x}) + \delta_{\mathbb{R}_+^{\mathrm{N}}}(\mathbf{x})$$

Imaging as an Inverse Problem

An imaging problem takes the form:

$$\mathbf{y} = \mathrm{H}\,\mathbf{x} + \mathbf{w}$$

Where the challenge is to recover **x** from the measurements **y.**

Approaches:

- Minimizing noise
- Introducing prior knowledge

$$\arg\min_{\mathbf{x}\in\mathbb{R}} \|\mathbf{H}\mathbf{x} - \mathbf{y}\|_2^2 + \mathcal{R}(\mathbf{L}\mathbf{x}) + \delta_{\mathbb{R}_+^{\mathrm{N}}}(\mathbf{x})$$

non-smooth, convex

Imaging as an Inverse Problem

An imaging problem takes the form:

$$\mathbf{y} = \mathrm{H}\,\mathbf{x} + \mathbf{w}$$

Where the challenge is to recover **x** from the measurements **y.**

Approaches:

- Minimizing noise
- Introducing prior knowledge

$$\arg\min_{\mathbf{x}\in\mathbb{R}} ||\mathbf{H}\mathbf{x} - \mathbf{y}||_2^2 + \mathcal{R}(\mathbf{L}\mathbf{x}) + \delta_{\mathbb{R}_+^{\mathrm{N}}}(\mathbf{x})$$

Image Regularization

- Enforcing nonnegativity

$$\delta_{\mathbb{R}_+^N}(\mathbf{x}) = \begin{cases} 0 & \text{if } \mathbf{x} \in \mathbb{R}_+^N \\ +\infty & \text{if } \mathbf{x} \notin \mathbb{R}_+^N \end{cases}$$

- Promoting smoothness – TV

$$\mathcal{R}(\mathbf{x}) = \sum_{n=1}^{N} \|\mathbf{D}_1\mathbf{x}\|_n + \|\mathbf{D}_2\mathbf{x}\|_n$$

$$\mathcal{R}(\mathbf{x}) = \sum_{n=1}^{N} \sqrt{[\mathbf{D_1}\mathbf{x}]_n^2 + [\mathbf{D_2}\mathbf{x}]_n^2}$$

Image Regularization

- Promoting smoothness – Hessian-Schatten norm

$$\mathcal{R}(\mathbf{x}) = \sum_{n=1}^{N} \left\| \begin{bmatrix} [D_{11}\mathbf{x}]_n & [D_{12}\mathbf{x}]_n \\ [D_{21}\mathbf{x}]_n & [D_{22}\mathbf{x}]_n \end{bmatrix} \right\|_{S_p}$$

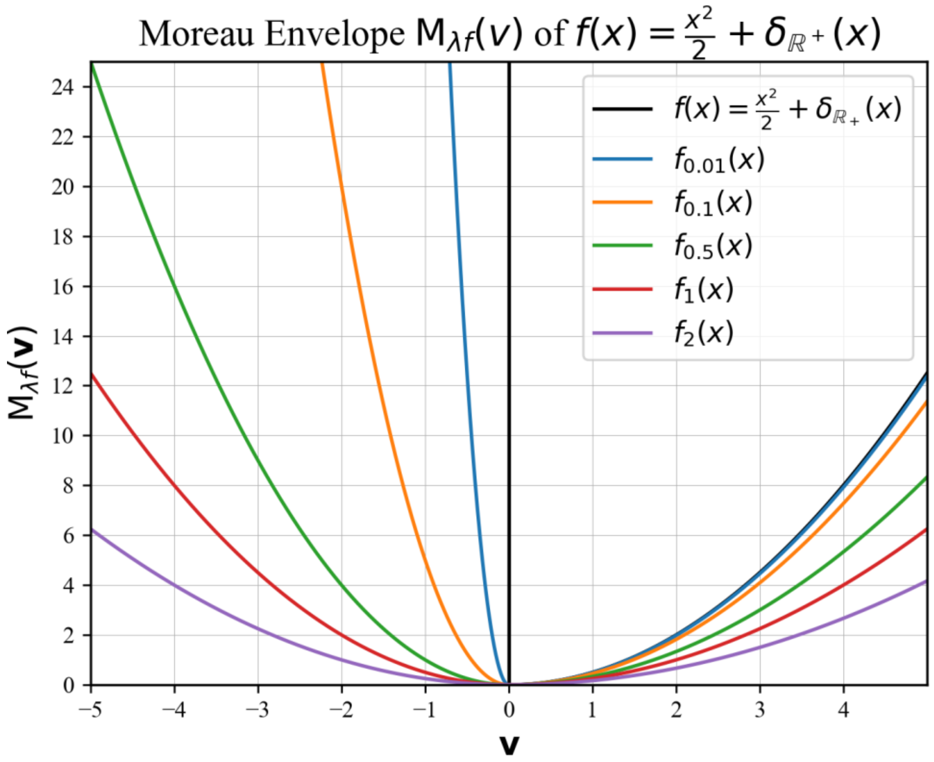
- Promoting sparsity – Group Sparsity

$$\|\mathbf{x}\|_{p,q} = \left(\sum_{i=1}^{S} \|\mathbf{x}_i\|_p^q \right)^{\frac{1}{q}}$$

The Proximal Operator

$$\boxed{\text{prox}_{\lambda f}(\mathbf{v}) = \arg\min_{\mathbf{x}}(f(\mathbf{x}) + \frac{1}{2\lambda}||\mathbf{x} - \mathbf{v}||_2^2)}$$

$$\text{M}_{\lambda f}(\mathbf{v}) = \inf_{\mathbf{x}}(f(\mathbf{x}) + \frac{1}{2\lambda}||\mathbf{x} - \mathbf{v}||_2^2)$$



Common Regularizers and their Proximals

- Non-negativity constraint

$$\Pi_C(\mathbf{v}) = \arg\min_{\mathbf{x}\in\mathcal{C}} \|\mathbf{x}-\mathbf{v}\|_2$$

- ℓ_p Norm

$$\text{prox}_{\lambda\ell_p}(\mathbf{v}) = \mathbf{v} - \lambda\Pi_{\mathcal{B}_p}(\frac{\mathbf{v}}{\lambda})$$

- ℓ_1 Norm

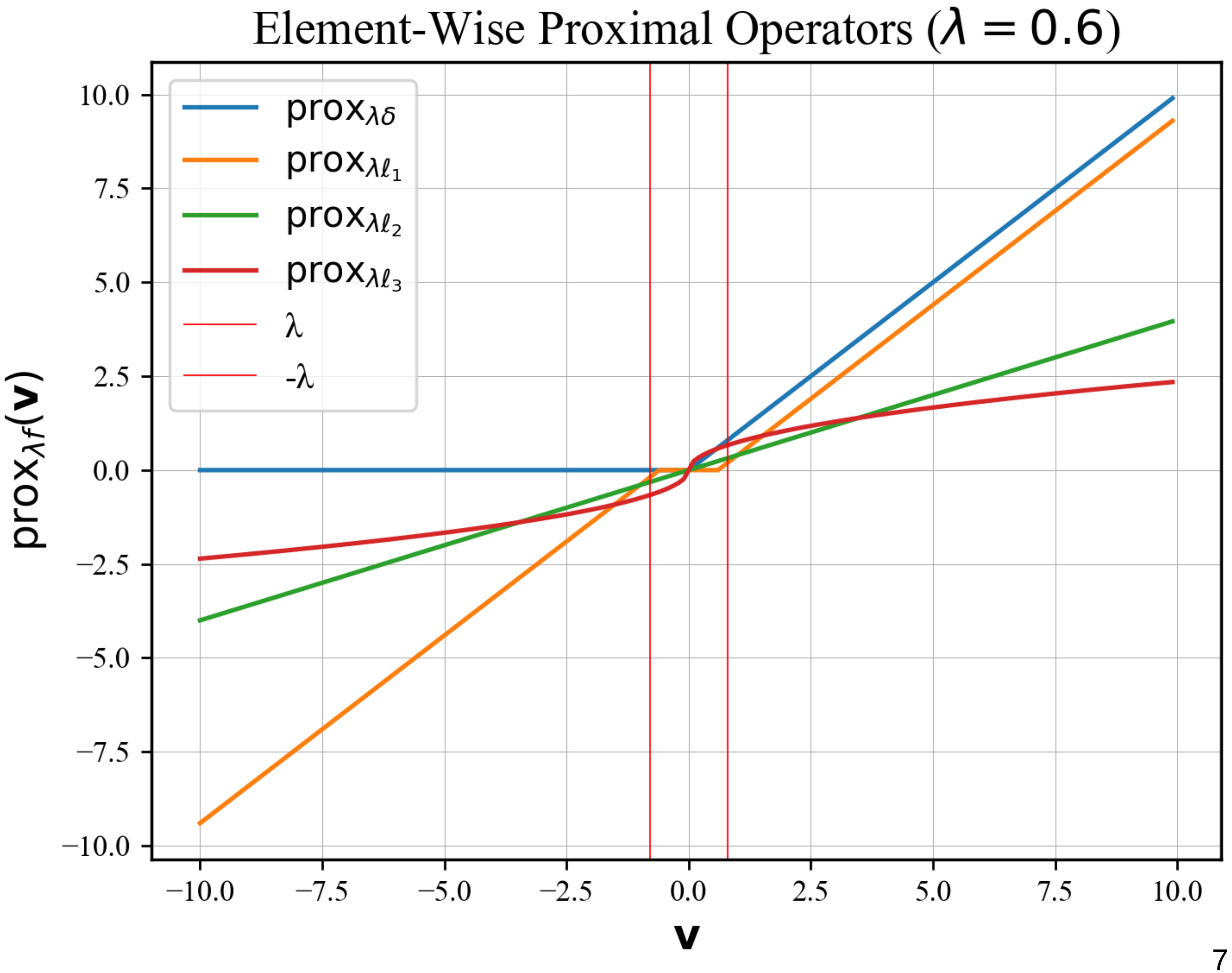
$$\text{sgn}(\mathbf{x})\max(|\mathbf{x}| - \lambda, 0)$$

- ℓ_2 Norm

$$\left(1 - \frac{\lambda}{\min(\|\mathbf{v}\|_2, \lambda)}\right)\mathbf{v}$$

Common Regularizers and their Proximals

- Non-negativity constraint
- ℓ_p Norm
- ℓ_1 Norm
- ℓ_2 Norm

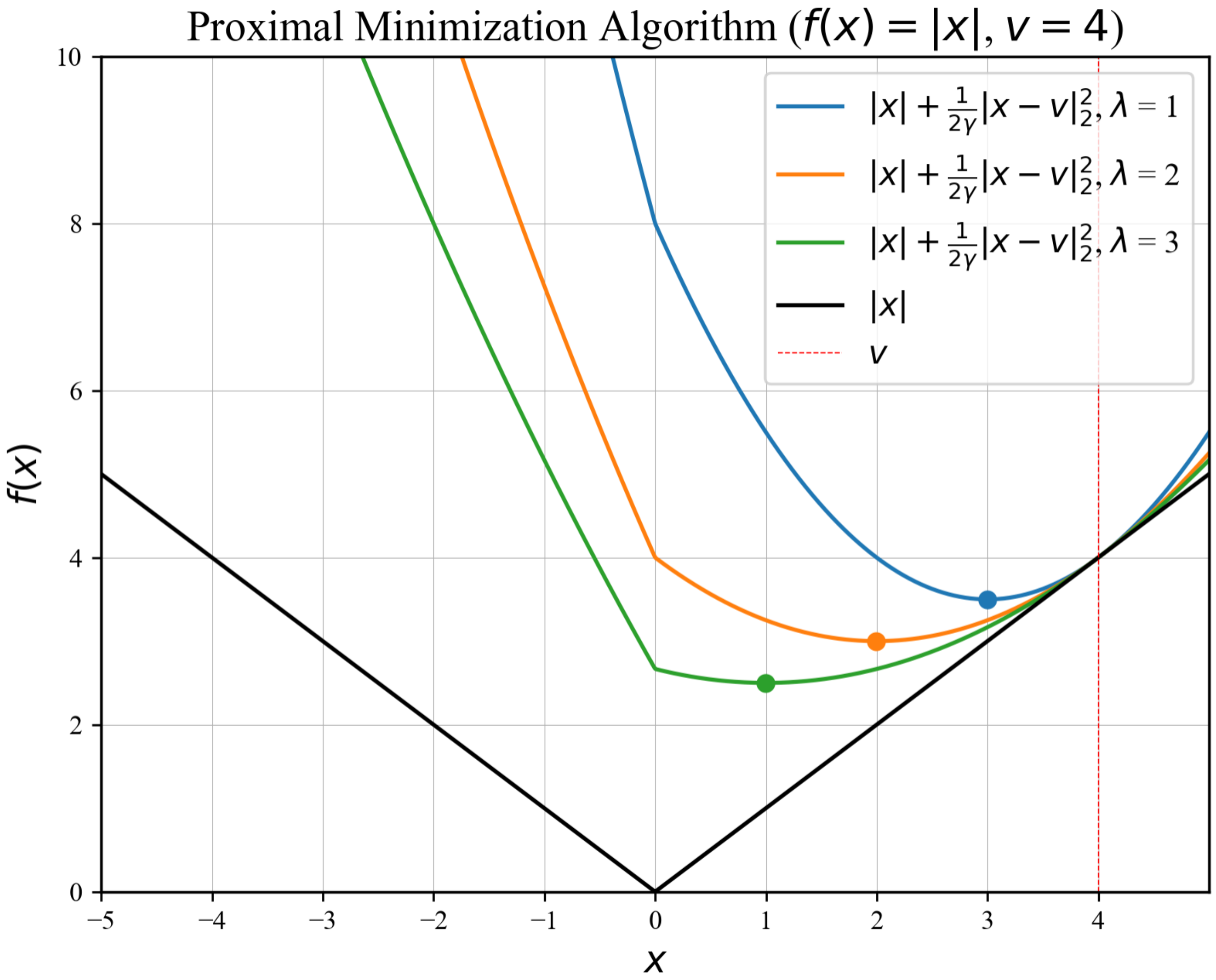


Proximal Algorithms

- Proximal Minimization

$$\text{minimize } f(\mathbf{x})$$

$$\mathbf{x}^{k+1} := \mathrm{prox}_{\lambda f}(\mathbf{x}^k)$$



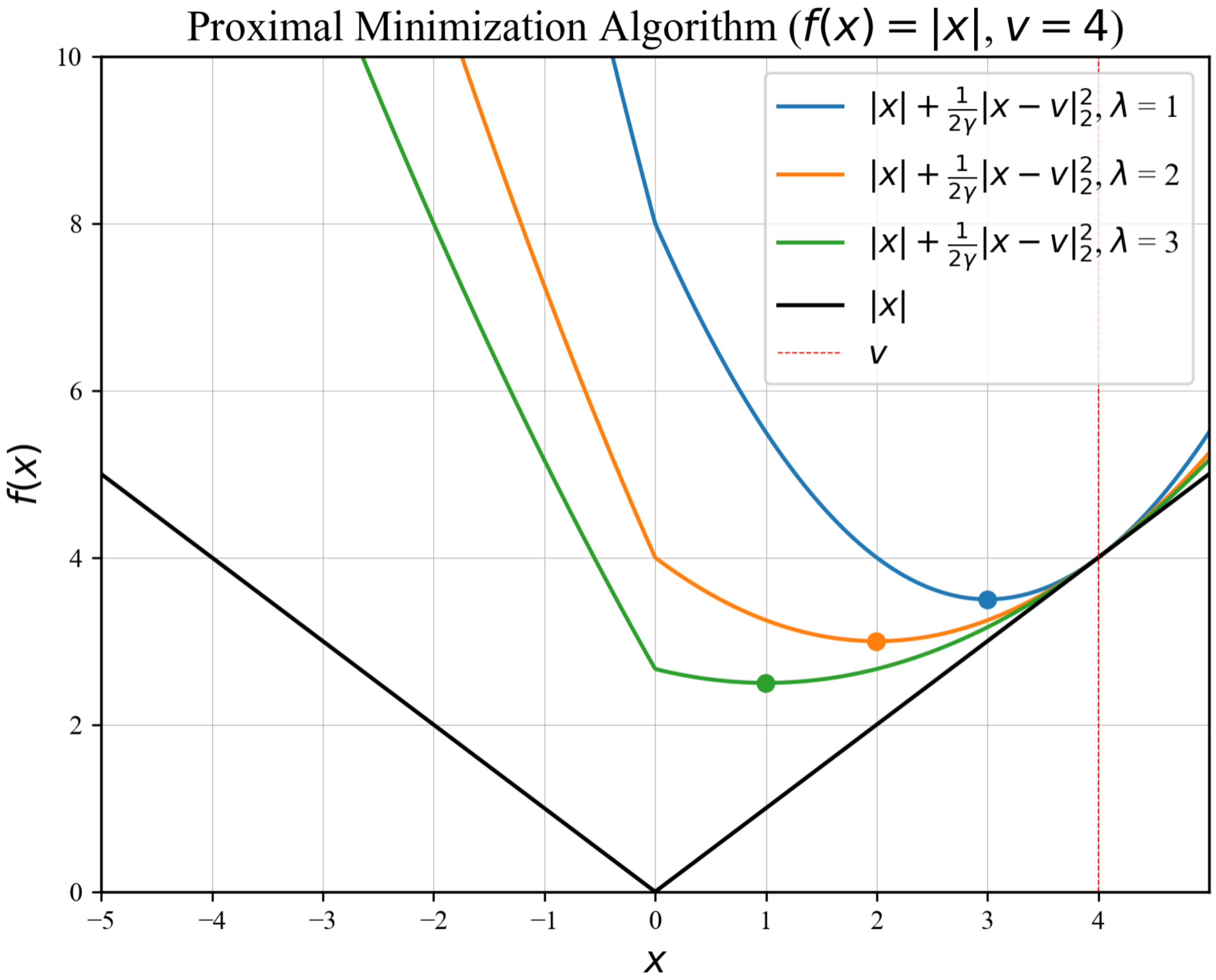
Proximal Algorithms

- Proximal Minimization

$$\text{minimize}\ f(\mathbf{x})$$

non-smooth, convex

$$\mathbf{x}^{k+1} := \mathrm{prox}_{\lambda f}(\mathbf{x}^k)$$



Proximal Algorithms

- Proximal Gradient Descent

$$\text{minimize } f(\mathbf{x}) + g(\mathbf{x})$$

$$\mathbf{x}^{k+1} := \text{prox}_{\lambda g}(\mathbf{x}^k - \lambda^k \nabla f(\mathbf{x}^k))$$

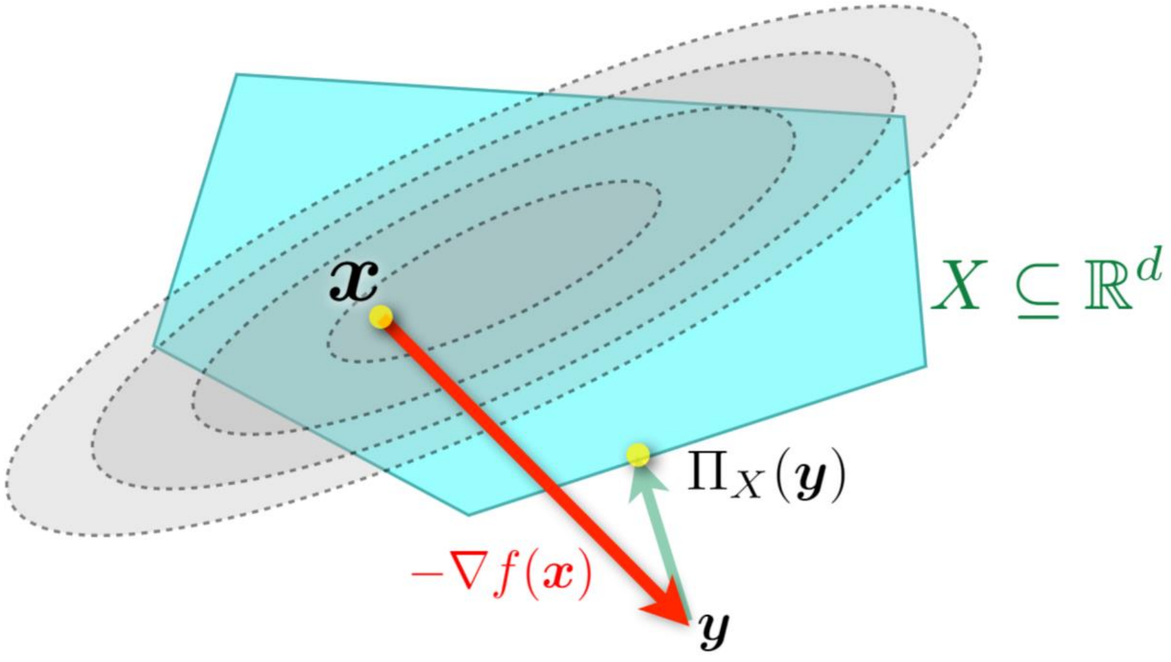


Figure reused from CS-439 course notes

Proximal Algorithms

- Proximal Gradient Descent

smooth, convex,
differentiable

$$\text{minimize}\ f(\mathbf{x}) + g(\mathbf{x})$$

$$\mathbf{x}^{k+1} := \text{prox}_{\lambda g}(\mathbf{x}^k - \lambda^k \nabla f(\mathbf{x}^k))$$

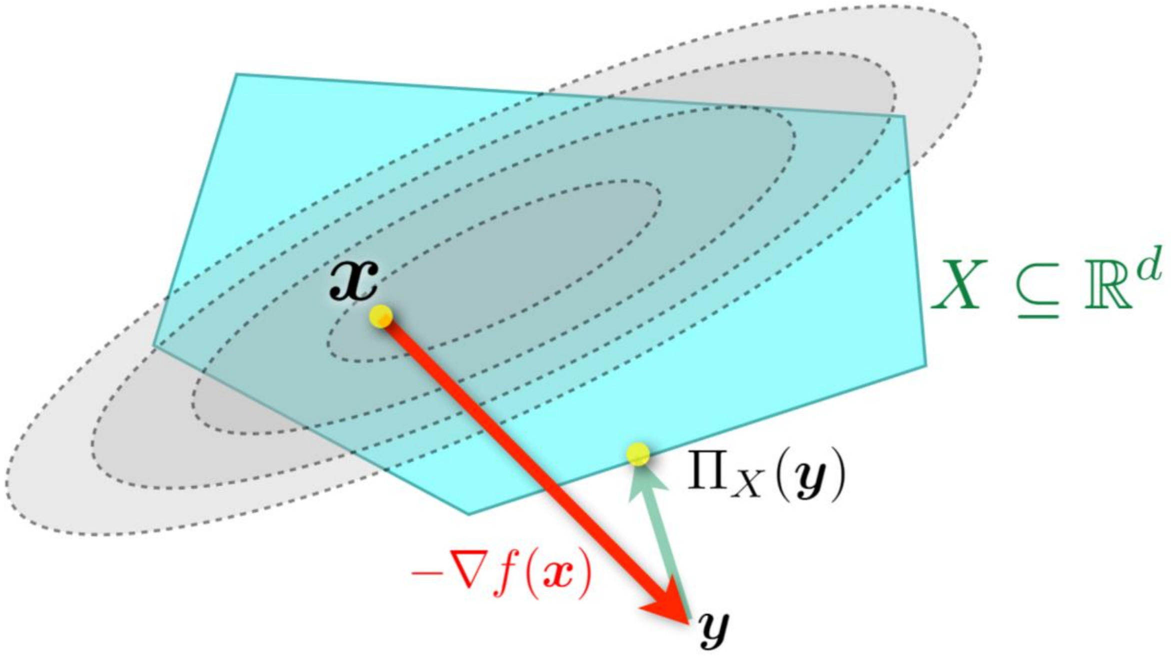


Figure reused from CS-439 course notes

Proximal Algorithms

- Proximal Gradient Descent

smooth, convex, differentiable

non-smooth, convex

$$\text{minimize}\ f(\mathbf{x}) + g(\mathbf{x})$$

$$\mathbf{x}^{k+1} := \text{prox}_{\lambda g}(\mathbf{x}^k - \lambda^k \nabla f(\mathbf{x}^k))$$

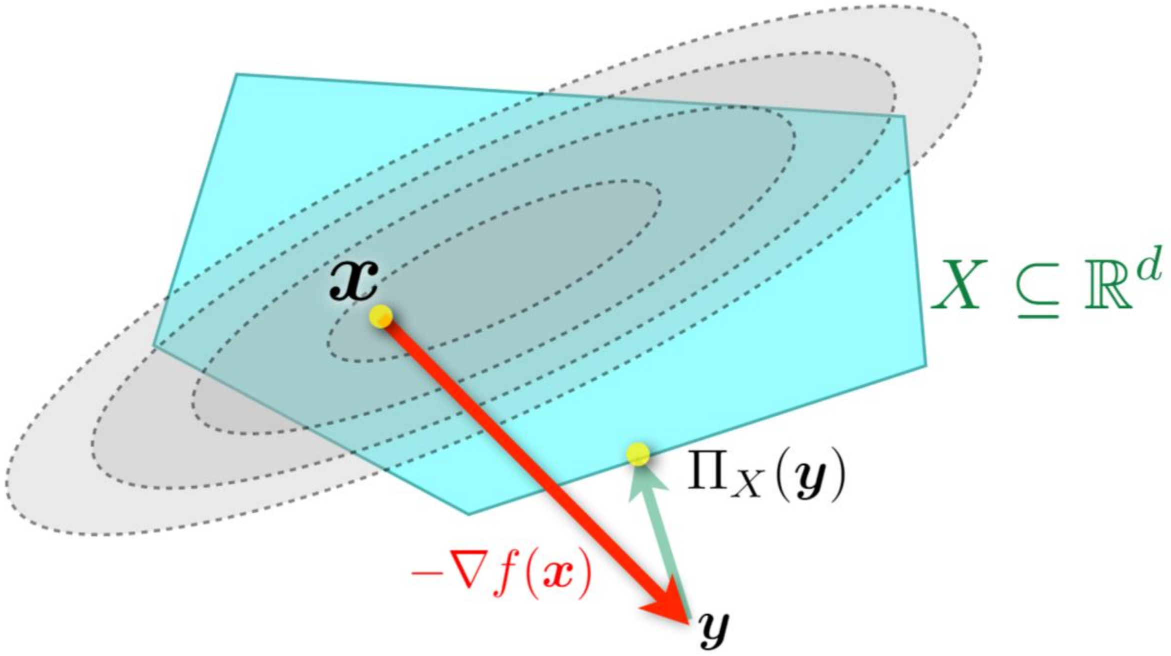


Figure reused from CS-439 course notes

Proximal Algorithms

- Alternating Direction Method of Multipliers (ADMM)

$$\text{minimize } f(\mathbf{x}) + \sum_{n=1}^{N} g_n(\mathbf{x}) \rightarrow$$

$$\begin{aligned} &\text{minimize } f(\mathbf{x}) + \sum_{n=1}^{N} g_n(\mathbf{z}) \\ &\text{subject to } \mathrm{A}\,\mathbf{x} + \mathrm{B}\,\mathbf{z} = \mathbf{c} \end{aligned}$$

$$\begin{aligned} \mathbf{x}^{k+1} &:= \mathrm{prox}_{\lambda f}(\mathbf{z}^{k} - \mathbf{u}^{k}) \\ \mathbf{z}_n^{k+1} &:= \mathrm{prox}_{\lambda g_n}(\mathbf{x}^{k+1} - \mathbf{u}^{k}) \\ \mathbf{u}^{k+1} &:= \mathbf{u}^{k+1} + \mathrm{A}\,\mathbf{x}^{k+1} + \mathrm{B}\,\mathbf{z}^{k+1} \end{aligned}$$

$$\mathbf{z}_n, \ \mathbf{u} \ \in \mathbb{R}^N$$

Proximal Algorithms

- Alternating Direction Method of Multipliers (ADMM)

non-smooth, convex

$$\text{minimize } f(\mathbf{x}) + \sum_{n=1}^{N} g_n(\mathbf{x}) \rightarrow$$

$$\begin{aligned} &\text{minimize } f(\mathbf{x}) + \sum_{n=1}^{N} g_n(\mathbf{z}) \\ &\text{subject to } \mathrm{A}\,\mathbf{x} + \mathrm{B}\,\mathbf{z} = \mathbf{c} \end{aligned}$$

$$\begin{aligned} \mathbf{x}^{k+1} &:= \operatorname{prox}_{\lambda f}(\mathbf{z}^k - \mathbf{u}^k) \\ \mathbf{z}_n^{k+1} &:= \operatorname{prox}_{\lambda g_n}(\mathbf{x}^{k+1} - \mathbf{u}^k) \\ \mathbf{u}^{k+1} &:= \mathbf{u}^{k+1} + \mathrm{A}\,\mathbf{x}^{k+1} + \mathrm{B}\,\mathbf{z}^{k+1} \end{aligned}$$

$$\mathbf{z}_n,\ \mathbf{u} \in \mathbb{R}^N$$

Proximal Algorithms

- Alternating Direction Method of Multipliers (ADMM)

non-smooth, convex

non-smooth, convex

$$\text{minimize } f(\mathbf{x}) + \sum_{n=1}^{N} g_n(\mathbf{x}) \rightarrow$$

$$\begin{aligned} &\text{minimize } f(\mathbf{x}) + \sum_{n=1}^{N} g_n(\mathbf{z}) \\ &\text{subject to } \mathrm{A}\,\mathbf{x} + \mathrm{B}\,\mathbf{z} = \mathbf{c} \end{aligned}$$

$$\begin{aligned} \mathbf{x}^{k+1} &:= \mathrm{prox}_{\lambda f}(\mathbf{z}^k - \mathbf{u}^k) \\ \mathbf{z}_n^{k+1} &:= \mathrm{prox}_{\lambda g_n}(\mathbf{x}^{k+1} - \mathbf{u}^k) \\ \mathbf{u}^{k+1} &:= \mathbf{u}^{k+1} + \mathrm{A}\,\mathbf{x}^{k+1} + \mathrm{B}\,\mathbf{z}^{k+1} \end{aligned}$$

$$\mathbf{z}_n, \;\; \mathbf{u} \;\in \mathbb{R}^N$$

Proximal Algorithms

- Alternating Direction Method of Multipliers (ADMM)

$$\text{minimize } f(\mathbf{x}) + \sum_{n=1}^{N} g_n(\mathbf{x}) \rightarrow$$

$$\begin{aligned} &\text{minimize } f(\mathbf{x}) + \sum_{n=1}^{N} g_n(\mathbf{z}) \\ &\text{subject to } \mathrm{A}\,\mathbf{x} + \mathrm{B}\,\mathbf{z} = \mathbf{c} \end{aligned}$$

$$\begin{aligned} \mathbf{x}^{k+1} &:= \mathrm{prox}_{\lambda f}(\mathbf{z}^k - \mathbf{u}^k) \\ \mathbf{z}_n^{k+1} &:= \mathrm{prox}_{\lambda g_n}(\mathbf{x}^{k+1} - \mathbf{u}^k) \\ \mathbf{u}^{k+1} &:= \mathbf{u}^{k+1} + \mathrm{A}\,\mathbf{x}^{k+1} + \mathrm{B}\,\mathbf{z}^{k+1} \end{aligned}$$

$$\mathbf{z}_n, \;\; \mathbf{u} \;\in \mathbb{R}^N$$

Proximal Algorithms

- Alternating Direction Method of Multipliers (ADMM)

$$\text{minimize } f(\mathbf{x}) + \sum_{n=1}^{N} g_n(\mathbf{x}) \rightarrow$$

$$\text{minimize } f(\mathbf{x}) + \sum_{n=1}^{N} g_n(\mathbf{z})$$

$$\text{subject to } \mathrm{A}\,\mathbf{x} + \mathrm{B}\,\mathbf{z} = \mathbf{c}$$

$$\begin{aligned}
\mathbf{x}^{k+1} &:= \mathrm{prox}_{\lambda f}(\mathbf{z}^{k} - \mathbf{u}^{k}) \\
\mathbf{z}_n^{k+1} &:= \mathrm{prox}_{\lambda g_n}(\mathbf{x}^{k+1} - \mathbf{u}^{k}) \\
\mathbf{u}^{k+1} &:= \mathbf{u}^{k+1} + \mathrm{A}\,\mathbf{x}^{k+1} + \mathrm{B}\,\mathbf{z}^{k+1}
\end{aligned}$$

$$\mathbf{z}_n,\ \mathbf{u} \in \mathbb{R}^N$$

$$\mathrm{prox}_{\lambda g}, \mathrm{prox}_{\lambda f} \rightarrow \text{optimization problems}$$

Scope of the Project

- Go back to proximal gradient descent algorithm
- Find regularizers that satisfy:

$$\mathrm{prox}_{\mathcal{R}+\delta_{\mathbb{R}^N_+}}(\mathbf{v}) = \mathrm{prox}_{\delta_{\mathbb{R}^N_+}}(\mathrm{prox}_{\mathcal{R}}(\mathbf{v}))?$$

$$\mathrm{prox}_{\mathbb{R}+\delta_{\mathbb{R}^N_+}}(\mathbf{v}) = \mathrm{prox}_{\mathcal{R}}(\mathrm{prox}_{\delta_{\mathbb{R}^N_+}}(\mathbf{v}))?$$

$$\text{where} \quad \mathbf{v} = \mathbf{x} - \lambda \nabla f(\mathbf{x})$$

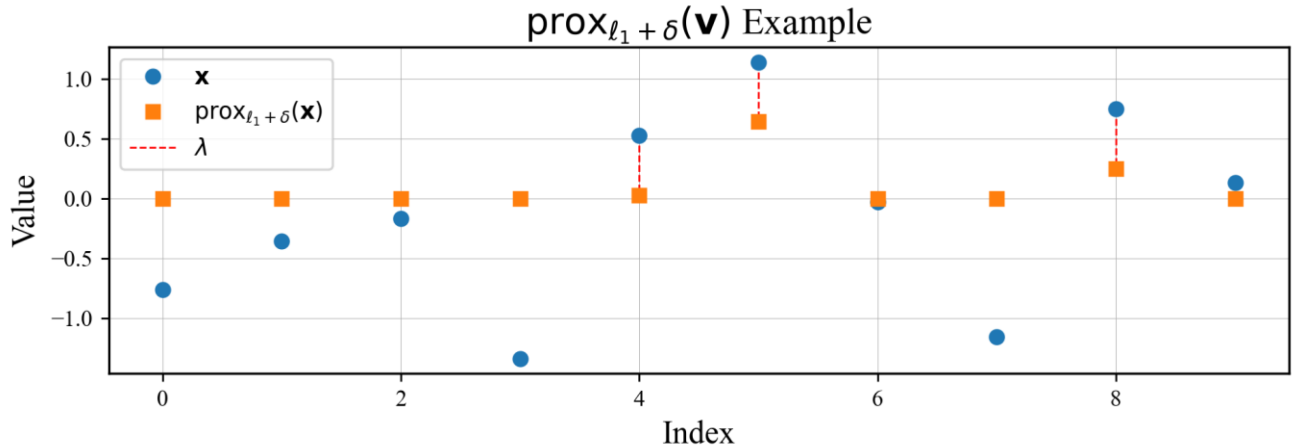
Setup

- Python 3.9
  - NumPy, SciPy, Matplotlib
- CVXPy
  - Modelling language for convex optimization
  - MOSEK solver
- Methodology
  - Generate random vectors
    - Uniform and Gaussian Distributions
  - Solve 3 cases we are considering
  - Check equality

CVXPy Example

```python
# Define size of vector, lambda, and initialize vector to apply prox to
n = 10; lambda_ = 0.5
v = np.random.normal(size = (n))
print(f'Value of v: {np.round(v, 3)}')

# Define CVXPy variable of the same size as v, constrained to the nonnegative domain
x_nonneg = cp.Variable(n, nonneg = True)

# Define cost function: |x|_1 + |x-v|^2_2 * (1/2 lambda_)
obj = cp.norm(x_nonneg, 1) + cp.norm(x_nonneg - v, 2)**2/(2*lambda_)

# Solve problem
prox_nonneg_prob = cp.Problem(cp.Minimize(obj))
prox_nonneg_prob.solve()
print(f'The optimal value of x: {np.round(x_nonneg.value, 3)}')
```

```
Value of v: [-0.765 -0.358 -0.167 -1.341  0.529  1.144 -0.03  -1.159  0.753  0.133]
The optimal value of x: [0.    0.    0.    0.    0.029 0.644 0.    0.    0.253 0.   ]
```



- CVXPy functions:
  - Norms,
  - Eigenvalues,
  - Sum of squares, etc.

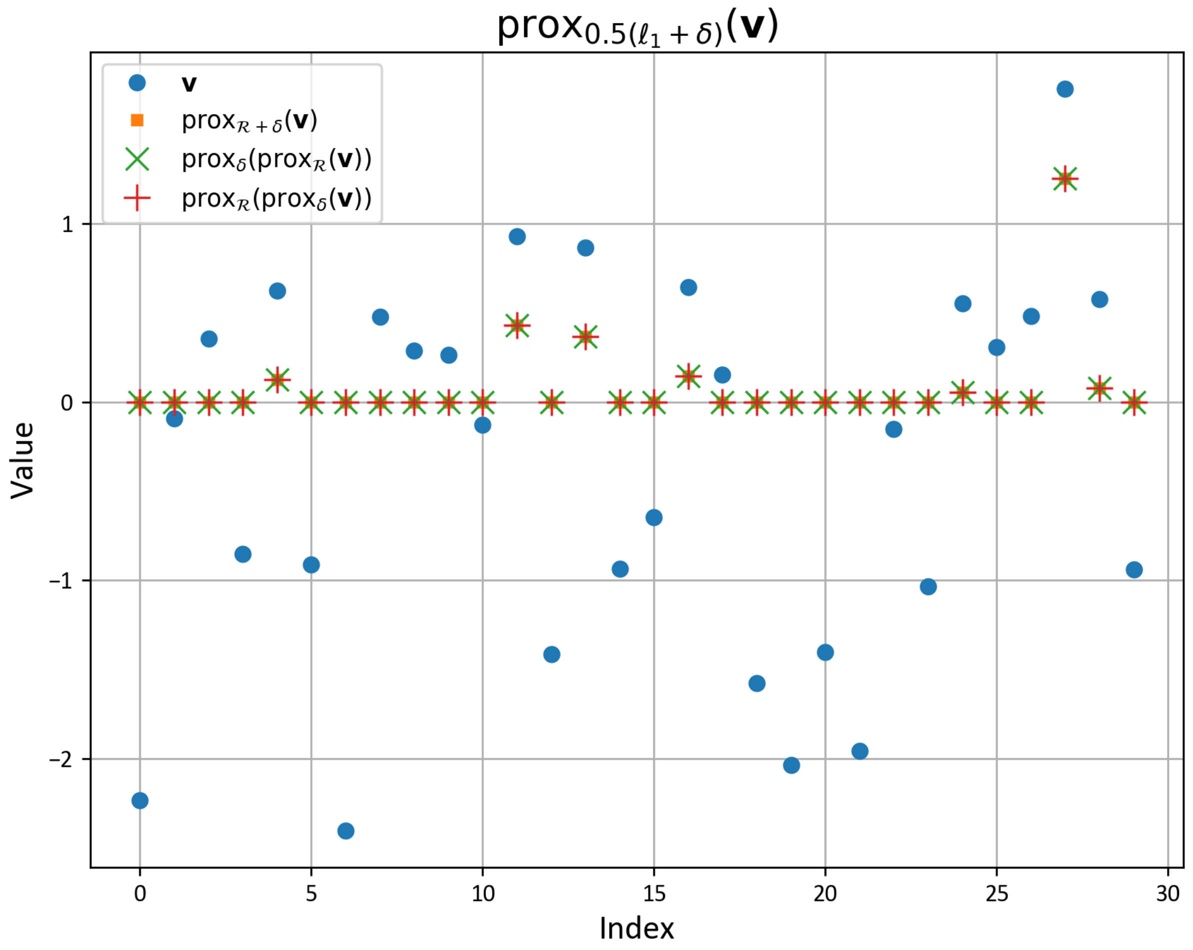
Results Summary

Dimension	$\mathbf{L}$	R	$\text{prox}_\delta(\text{prox}_\mathcal{R}(\mathbf{x}))$	$\text{prox}_\mathcal{R}(\text{prox}_\delta(\mathbf{x}))$	Parameters
1, 2	I	$\|\cdot\|_1$	$\checkmark(10^{-5})$	$\checkmark(10^{-7})$	$\mathcal{N}(0,1)$, $\mathbf{x}\in\mathbb{R}^{100}$
1, 2	I	$\|\cdot\|_2$	$\times(10^{-3})$	$\checkmark(10^{-6})$	$\mathcal{N}(0,1)$, $\mathbf{x}\in\mathbb{R}^{100}$
1, 2	I	$\|\cdot\|_p$	$\times(10^{-3})$	$\checkmark(10^{-7})$	$\mathcal{N}(0,1)$, $\mathbf{x}\in\mathbb{R}^{100}$
1, 2	I	$\|\cdot\|_p^p$	$\checkmark(10^{-6})$	$\checkmark(10^{-6})$	$\mathcal{N}(0,1)$, $\mathbf{x}\in\mathbb{R}^{100}$
2	I	$\|\cdot\|_{S_p}$	$\times\ (10^{-2})$	? (10^{-3})	$\mathcal{N}(0,100)$, $\mathbf{x}\in\mathbb{R}^{2\times 2}$
1	∇_1	$\|\cdot\|_1$	$\checkmark(10^{-6})$	$\times\ (10^{-1})$	$\mathcal{N}(0,1)$, $\mathbf{x}\in\mathbb{R}^{50}$
2	∇_2	$\|\cdot\|_{1,1}$	$\checkmark(10^{-7})$	$\times\ (10^{-1})$	$\mathcal{N}(0,1)$, $\mathbf{x}\in\mathbb{R}^{40\times 40}$
1, 2	I	$\|\cdot\|_{1,1}$	$\checkmark(10^{-7})$	$\checkmark(10^{-7})$	$\mathcal{N}(0,1)$, $\mathbf{x}\in\mathbb{R}^{40\times 40}$
1, 2	I	$\|\cdot\|_{p,q}$	$\times\ (10^{-2})$	$\checkmark(10^{-6})$	$\mathcal{N}(0,1)$, $\mathbf{x}\in\mathbb{R}^{40\times 40}$
1, 2	I	$\|\cdot\|_{p,q}^{p,q}$	$\times\ (10^{-4})$	$\checkmark(10^{-7})$	$\mathcal{N}(0,1)$, $\mathbf{x}\in\mathbb{R}^{40\times 40}$
1, 2	I	$\|\cdot\|_{1,q}$	$\checkmark(10^{-7})$	$\checkmark(10^{-7})$	$\mathcal{N}(0,1)$, $\mathbf{x}\in\mathbb{R}^{40\times 40}$
1, 2	I	$\|\cdot\|_{p,1}^{p,q}$	$\checkmark(10^{-7})$	$\checkmark(10^{-7})$	$\mathcal{N}(0,1)$, $\mathbf{x}\in\mathbb{R}^{40\times 40}$
2	HS	$\|\cdot\|_{S_p,q}$	$\times\ (10^{-1})$	$\times\ (10^{-1})$	$\mathcal{N}(0,10)$, $\mathbf{x}\in\mathbb{R}^{15\times 15}$

Results: Validation

- ℓ_1 norm, $\lambda = 0.5$

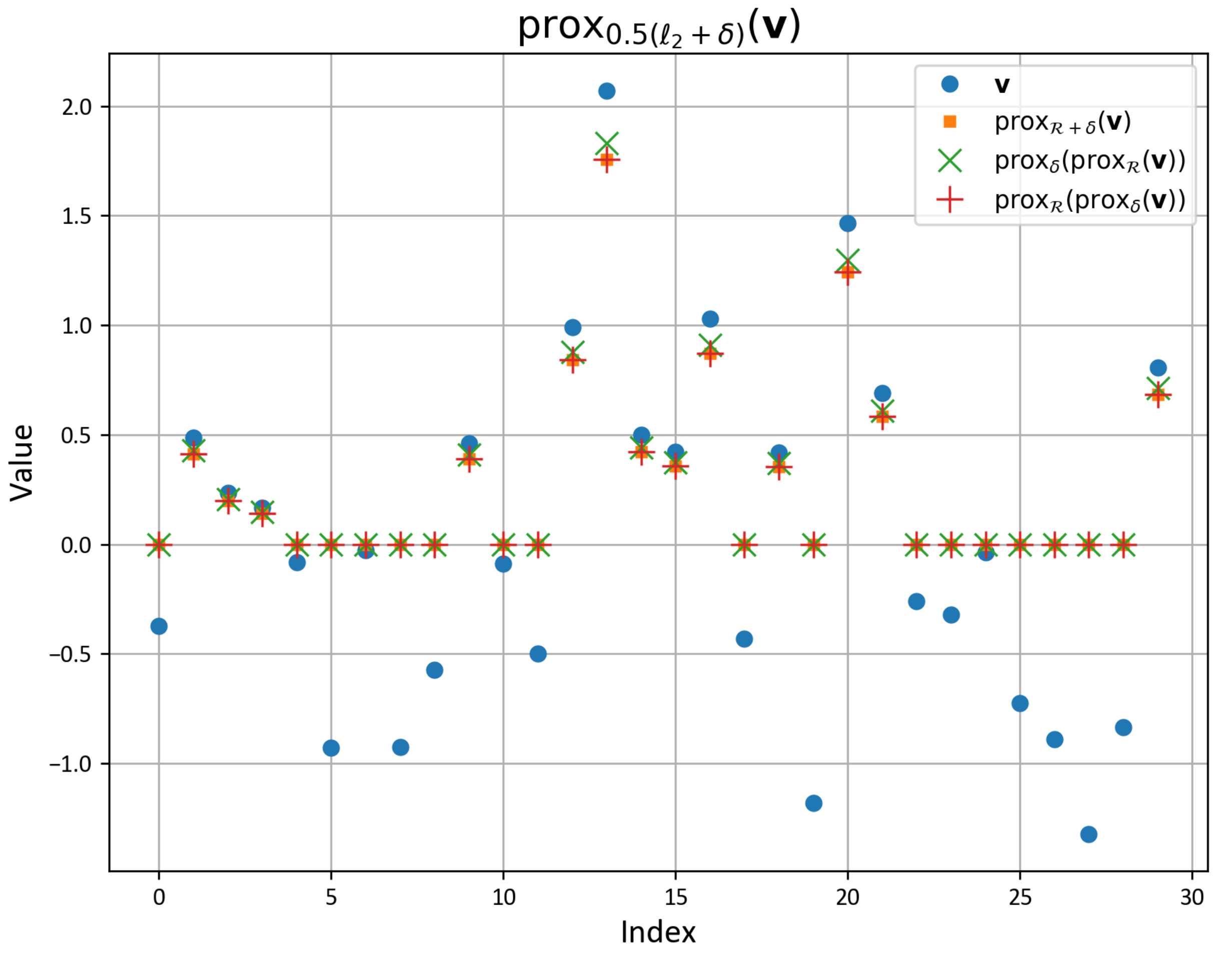
```
L1 Norm:
prox_nonneg(prox_reg(v)) SEEMS equal to prox(reg + nonneg)
Max absolute error: 0.000e+00
Average absolute error: 0.000e+00

prox_reg(prox_nonneg(v)) SEEMS equal to prox(reg + nonneg)
Max absolute error: 0.000e+00
Average absolute error: 0.000e+00
```



Results: Validation

- ℓ_2 norm, $\lambda = 0.5$

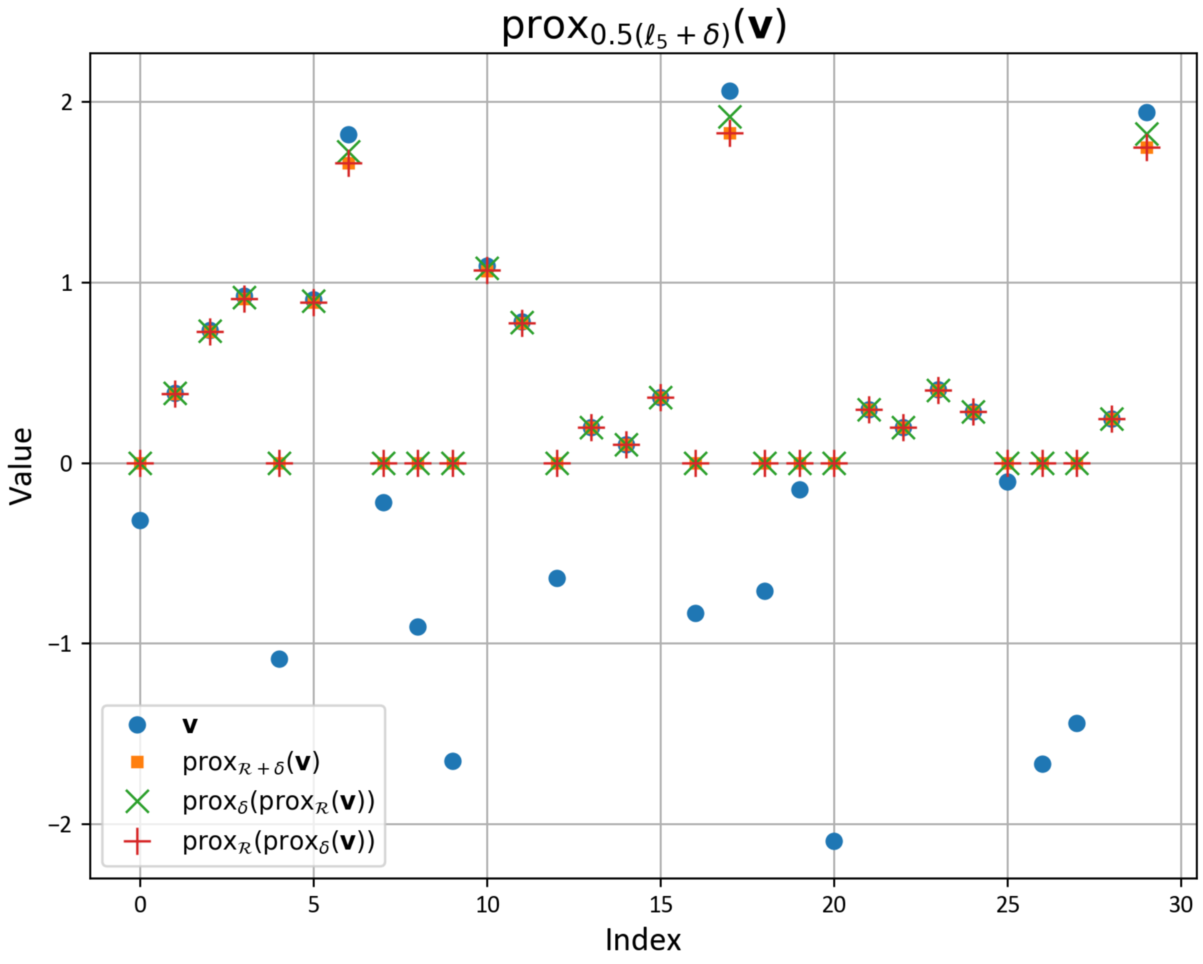
```
L2 Norm:
prox_nonneg(prox_reg(v)) IS NOT equal to prox(reg + nonneg)
Max absolute error: 1.891e-01
Average absolute error: 1.586e-02

prox_reg(prox_nonneg(v)) SEEMS equal to prox(reg + nonneg)
Max absolute error: 0.000e+00
Average absolute error: 0.000e+00
```



Results: ℓ_p Norm

- ℓ_5 norm, $\lambda = 0.5$

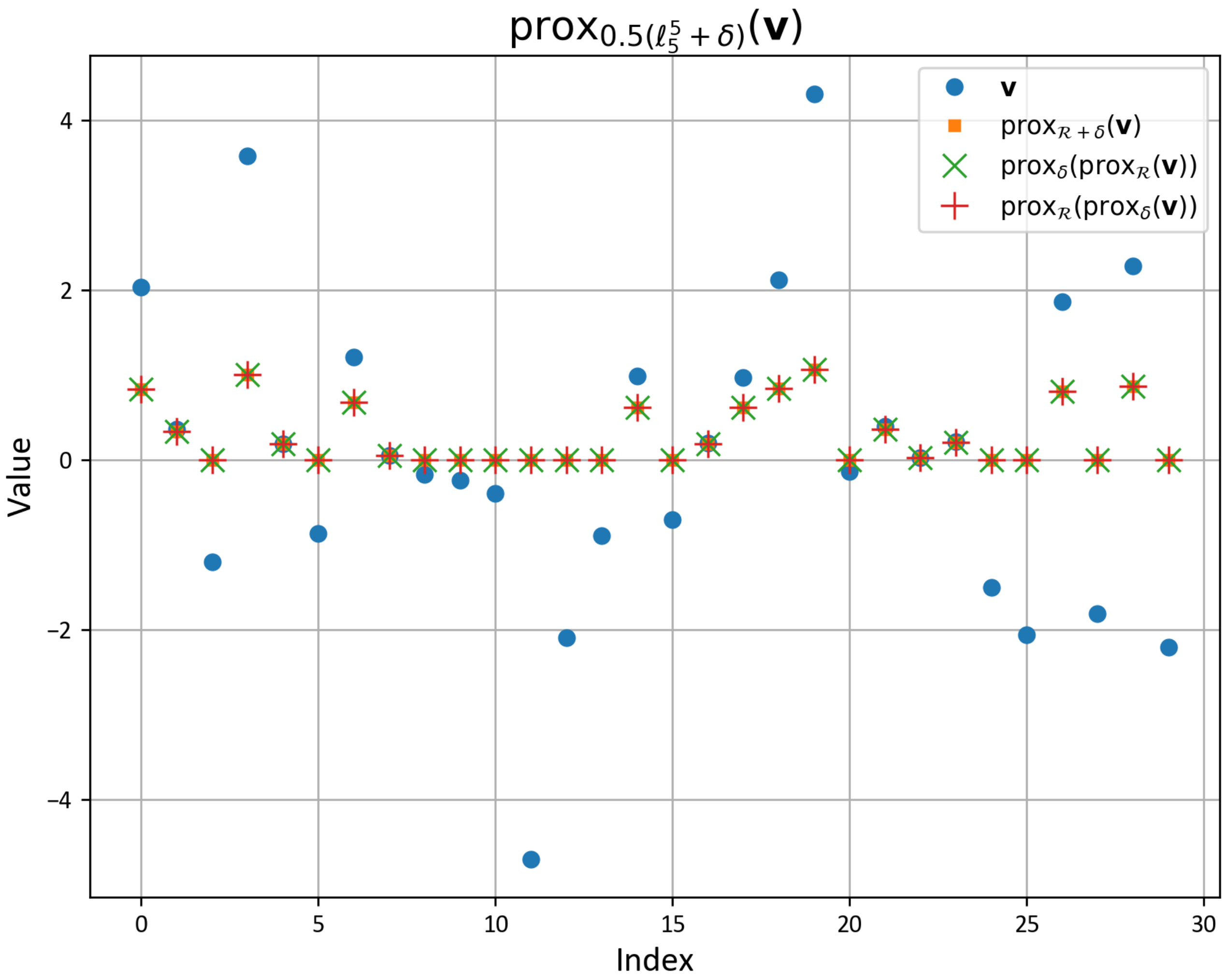
```
||p norm where p = 5
prox_nonneg(prox_reg(v)) IS NOT equal to prox(reg + nonneg)
Max absolute error: 3.467e-01
Average absolute error: 9.999e-03

prox_reg(prox_nonneg(v)) SEEMS equal to prox(reg + nonneg)
Max absolute error: 5.454e-04
Average absolute error: 3.652e-06
```



Results: ℓ_p^p Norm

- ℓ_5^5 norm, $\lambda = 0.5$

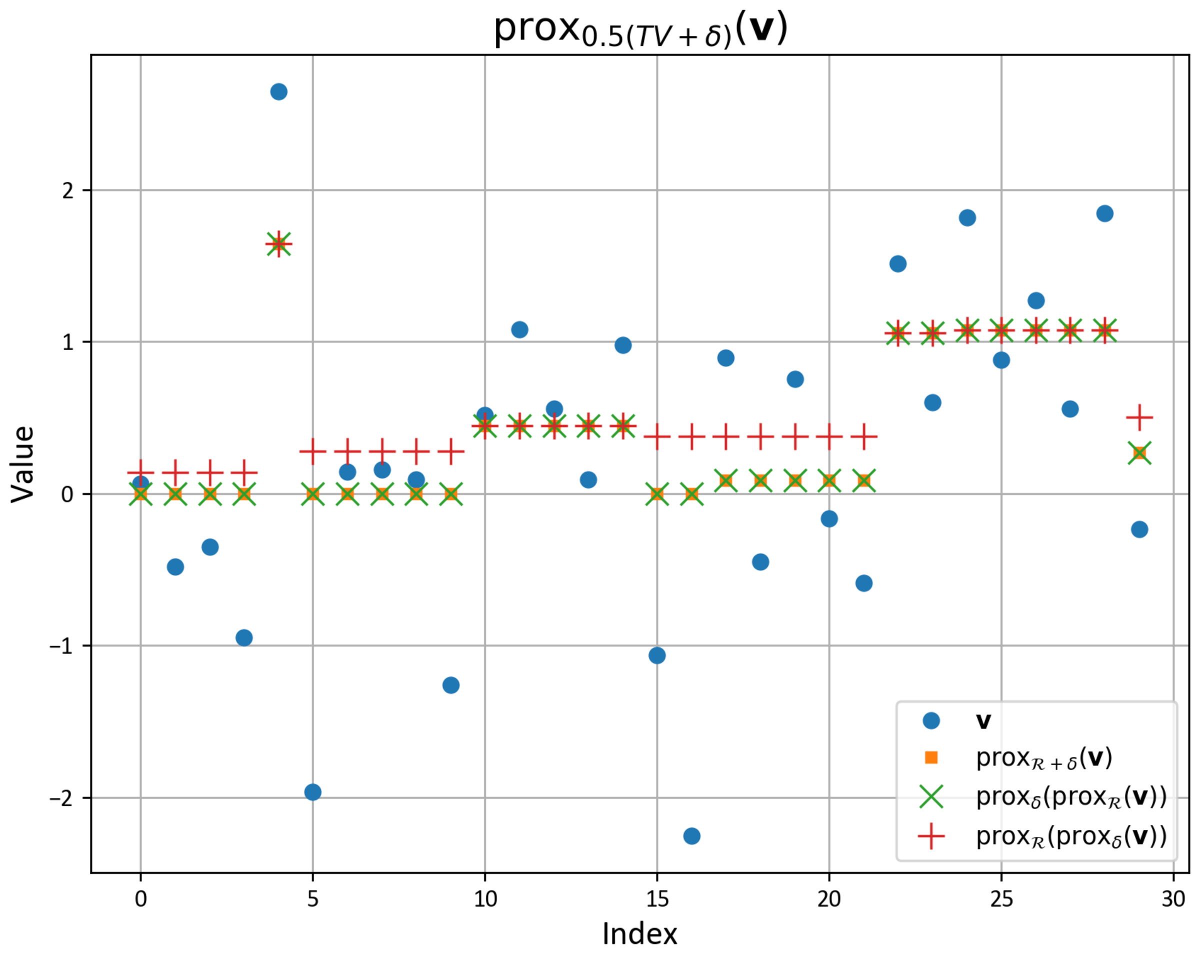
```
||p^p norm where p = 5
prox_nonneg(prox_reg(v)) SEEMS equal to prox(reg + nonneg)
Max absolute error: 1.808e-04
Average absolute error: 4.213e-06

prox_reg(prox_nonneg(v)) SEEMS equal to prox(reg + nonneg)
Max absolute error: 1.084e-04
Average absolute error: 2.542e-06
```



Results: 1-dimensional TV

- $\|Qx\|_1$, $\lambda = 0.3$

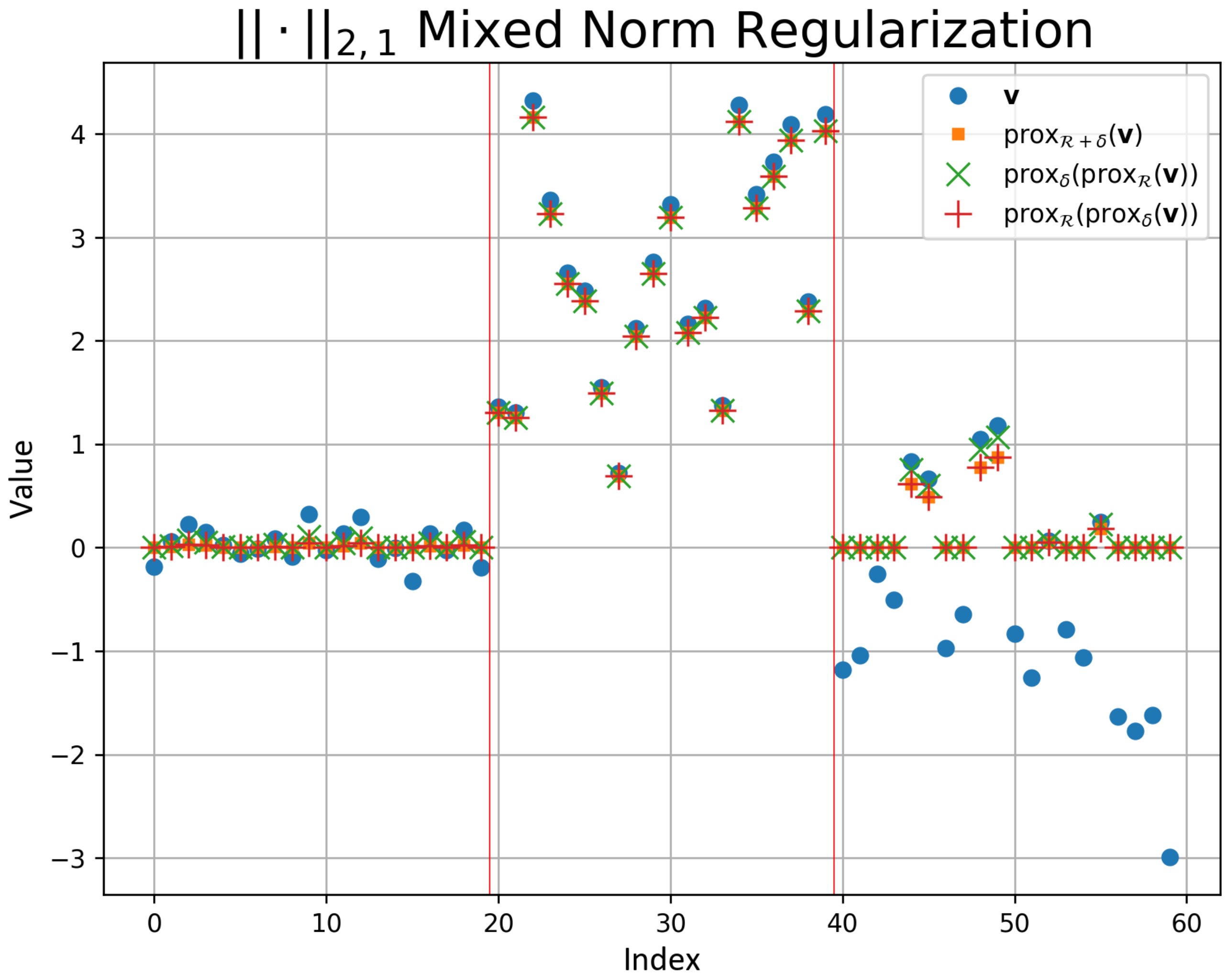
```
||Qx||1 Norm, where Q is the finite differences:
prox_nonneg(prox_reg(v)) SEEMS equal to prox(reg + nonneg)
Max absolute error: 1.650e-06
Average absolute error: 1.319e-07

prox_reg(prox_nonneg(v)) IS NOT equal to prox(reg + nonneg)
Max absolute error: 5.000e-01
Average absolute error: 1.786e-01
```



Results: 1-dimensional Group Sparsity

- $\|\cdot\|_{2,1}$, $\lambda = 0.5$

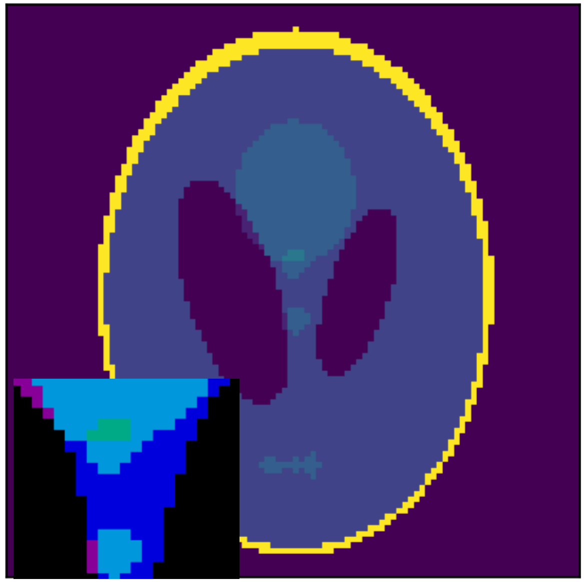
```
||p norm where p = 2, q = 1
prox_nonneg(prox_reg(v)) IS NOT equal to prox(reg + nonneg)
Max absolute error: 3.689e-01
Average absolute error: 3.957e-02

prox_reg(prox_nonneg(v)) SEEMS equal to prox(reg + nonneg)
Max absolute error: 1.667e-04
Average absolute error: 4.217e-06
```

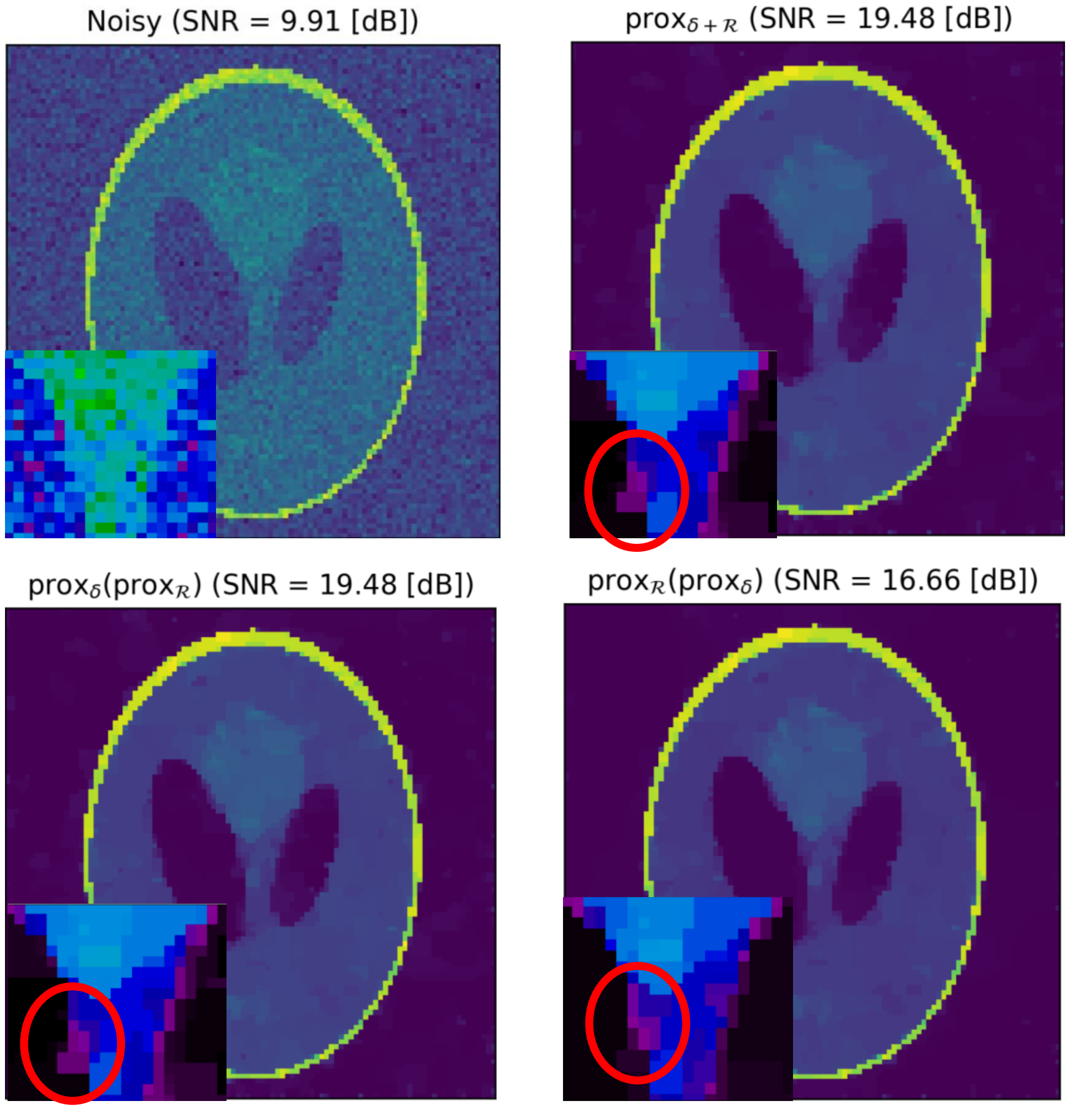


Nonisotropic TV

```
TV with 1,1 mixed norm regularizer:
prox_nonneg(prox_reg(v)) SEEMS equal to prox(reg + nonneg).
Max absolute error: 2.403e-04
Average absolute error: 3.037e-07

prox_reg(prox_nonneg(v)) IS NOT equal to prox(reg + nonneg)
Max absolute error: 7.712e-01
Average absolute error: 1.661e-01
```

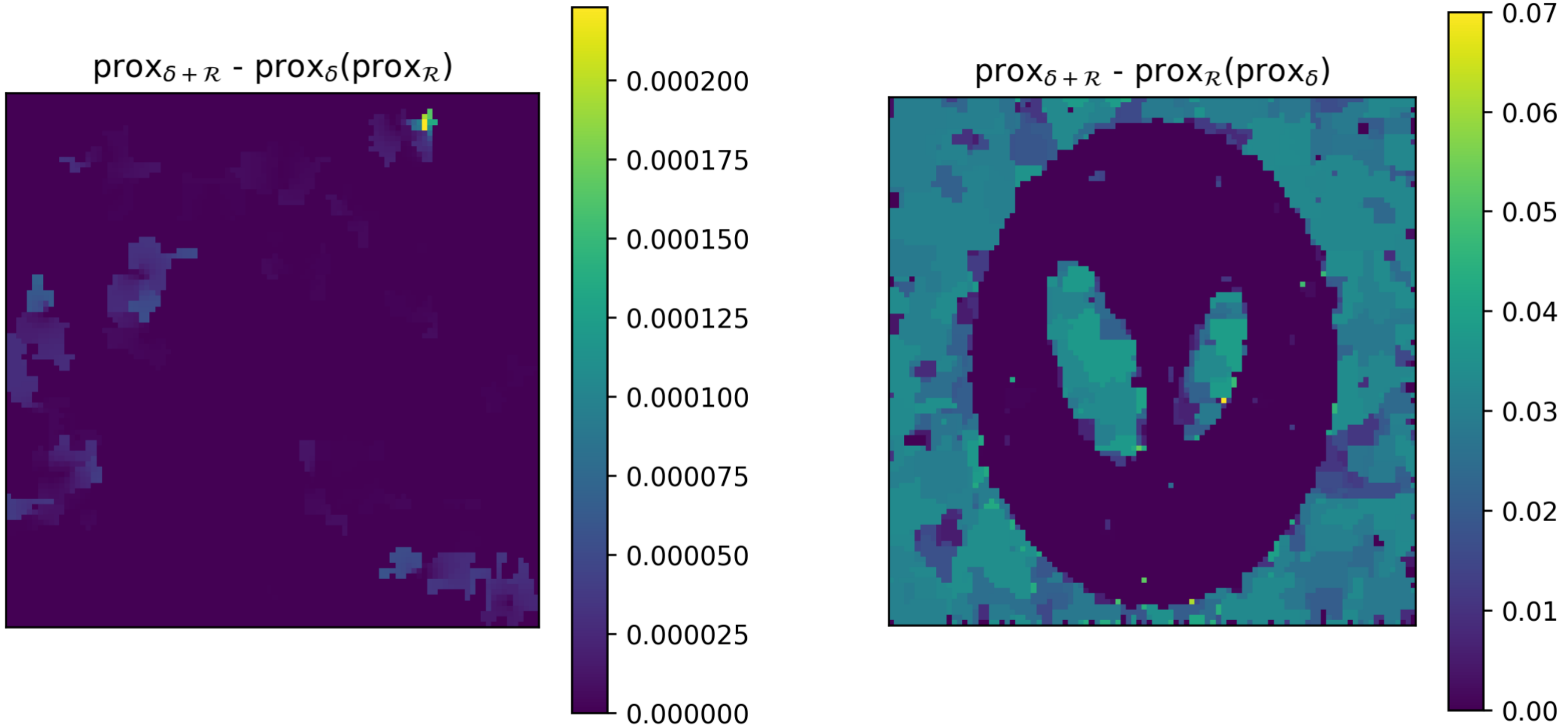
Original





Nonisotropic TV Error Maps

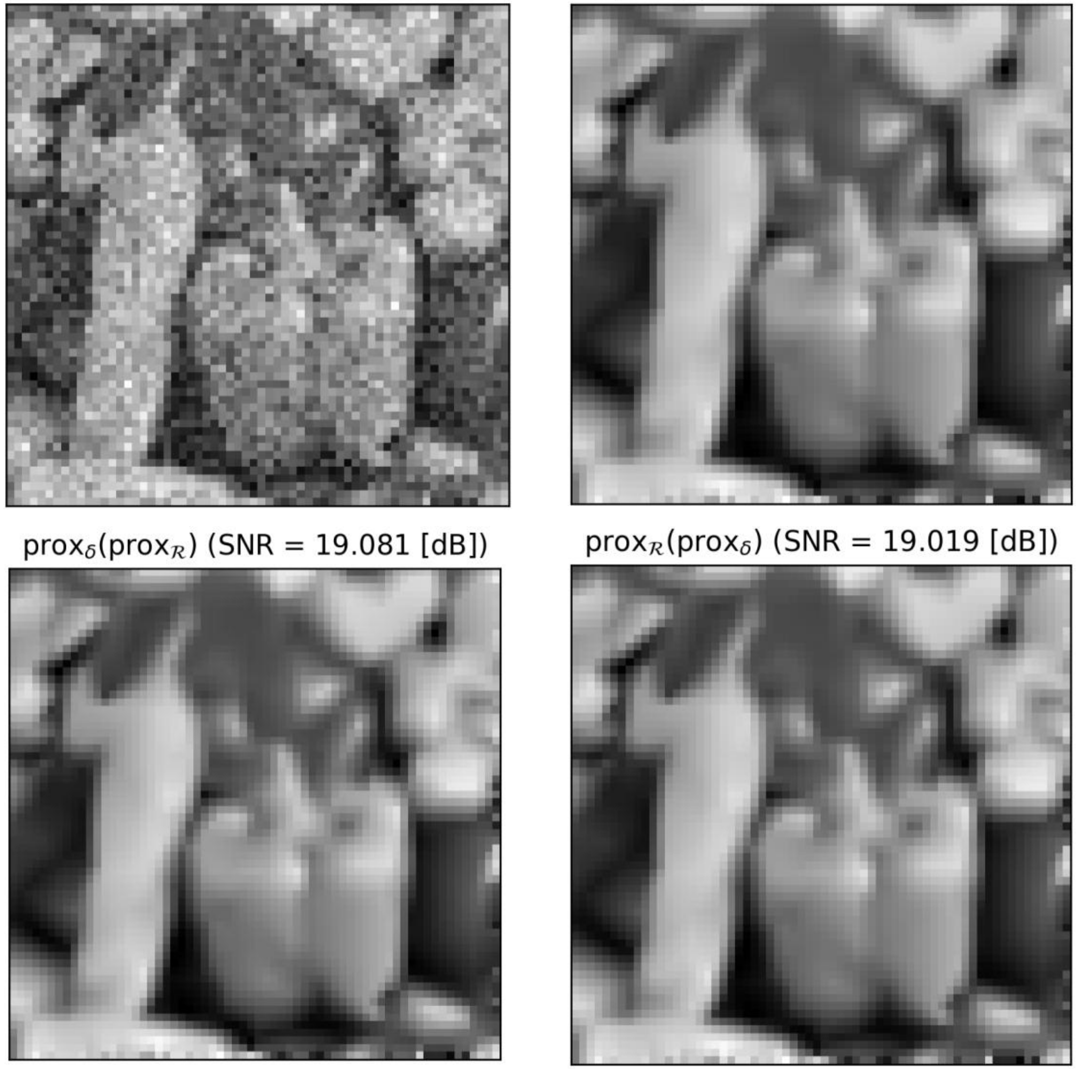
$$\boxed{\mathrm{prox}_{\mathrm{TV}+\delta_{\mathbb{R}^N_+}}(\mathbf{v}) = \mathrm{prox}_{\delta_{\mathbb{R}^N_+}}(\mathrm{prox}_{TV}(\mathbf{v}))!!}$$



Hessian-Schatten Norm

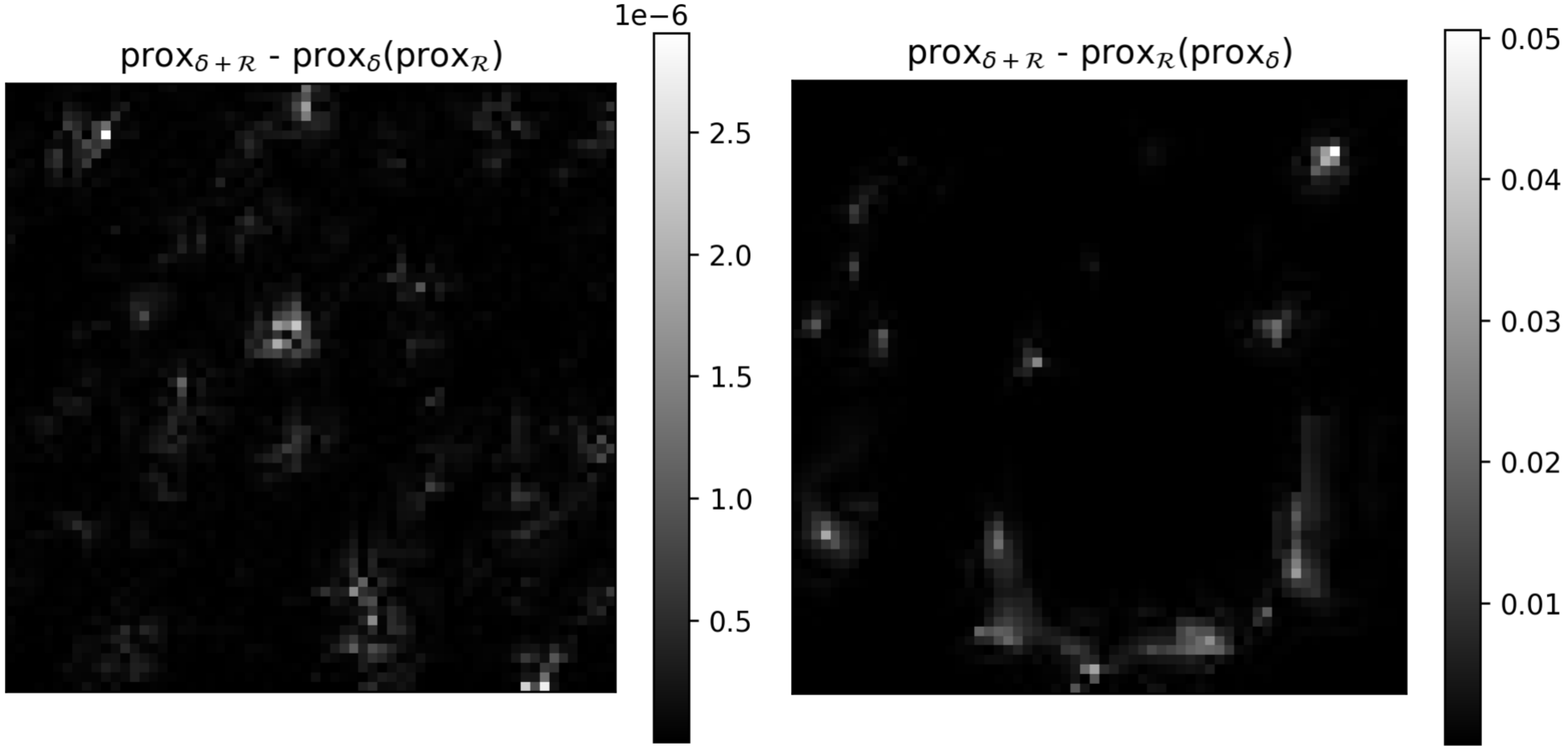
```
Hessian-Schatten norm:
prox_nonneg(prox_reg(v)) IS NOT equal to prox(reg + nonneg)
Max absolute error: 3.079e+00
Average absolute error: 2.511e-01

prox_reg(prox_nonneg(v)) IS NOT equal to prox(reg + nonneg)
Max absolute error: 4.921e+00
Average absolute error: 7.353e-01
```

Original



Hessian-Schatten Norm Error Maps

$$\boxed{\mathrm{prox}_{\mathrm{HS}+\delta_{\mathbb{R}^N}}(\mathbf{x}) \approx \mathrm{prox}_{\delta_{\mathbb{R}^N}}(\mathrm{prox}_{\mathrm{HS}}(\mathbf{x}))!}$$



Conclusions

- Solve a regularized imaging problem
- With proximal gradient descent

$$\arg\min_{x\in\mathbb{R}^N} \|\mathrm{H}\,\mathbf{x}-\mathbf{y}\|_2^2+\mathcal{R}(\mathbf{Lx})+\delta_{\mathbb{R}_+^N}$$

$$\mathrm{prox}_{\mathcal{R}+\delta_{\mathbb{R}_+^N}}(\mathbf{v})=\mathrm{prox}_{\delta_{\mathbb{R}_+^N}}(\mathrm{prox}_{\mathcal{R}}(\mathbf{v}))?$$

$$\mathrm{prox}_{\mathbb{R}+\delta_{\mathbb{R}_+^N}}(\mathbf{v})=\mathrm{prox}_{\mathcal{R}}(\mathrm{prox}_{\delta_{\mathbb{R}_+^N}}(\mathbf{v}))?$$

Dimension	$\mathbf{L}$	R	Reduced Splitting
1, 2	I	$\|\cdot\|_p$	$\checkmark(10^{-7})$
1, 2	I	$\|\cdot\|_p^p$	$\checkmark(10^{-6})$
2	I	$\|\cdot\|_{S_p}$	? (10^{-3})
1	∇_1	$\|\cdot\|_1$	$\checkmark(10^{-6})$
1	∇_1	$\|\cdot\|_p$	$\times$ (10^{-2})
2	∇_2	$\|\cdot\|_{1,1}$	$\checkmark(10^{-7})$
2	∇_2	$\|\cdot\|_{1,1}$	$\times$ (10^{-7})
1, 2	I	$\|\cdot\|_{p,p}^{(p),(q)}$	$\checkmark(10^{-7})$
2	HS	$\|\cdot\|_{S_{1,q}}$	? (10^{-1})

Further Work

- Prove in closed-form the properties studied here
- Test in optimization libraries, such as GlobalBioIm
- Quantify performance improvement