Global Optimality for Non-linear Constrained Restoration Problems via Invexity

Science and Technology Facilities Council

Samuel Pinilla, Jeyan T.

Introduction

The class of functions that any stationary point is a global minimizer is defined as follows.

Definition (Invexity). *Let $f:\mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz; then f is invex if there exists a function $\eta:\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ such that*

$$f(\boldsymbol{x}) - f(\boldsymbol{y}) \geq \zeta^T \eta(\boldsymbol{x},\boldsymbol{y}),$$

$\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n, \forall \zeta \in \partial f(\boldsymbol{y})$.

Hierarchy of optimizable functions

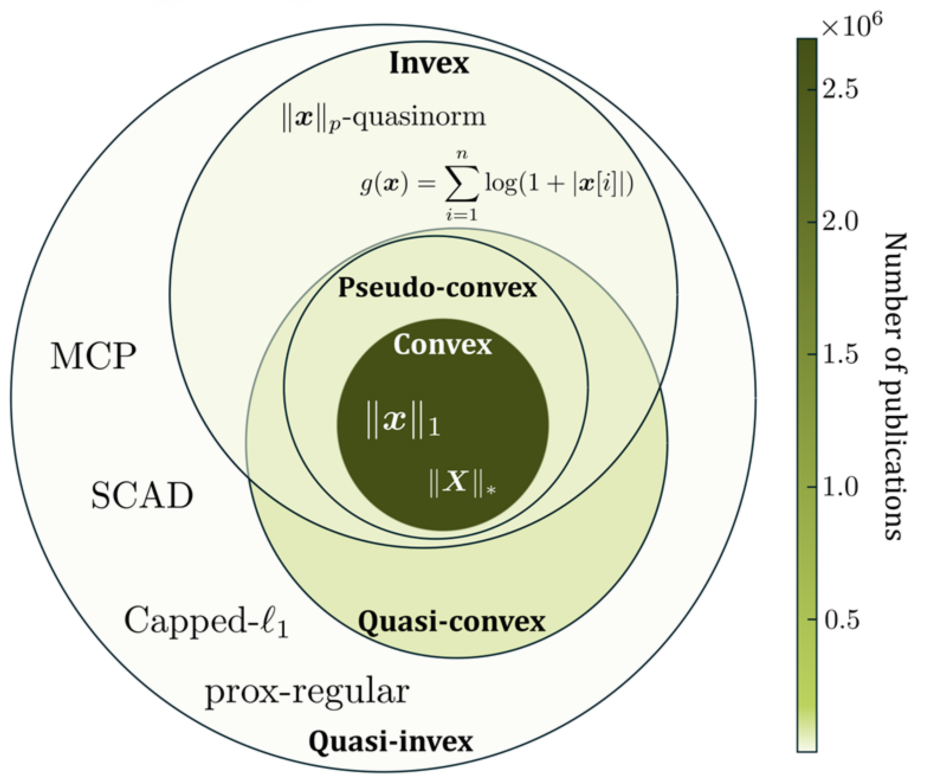


Fig 1. Our contribution is identifying invex and quasi-invex functions relevant for imaging applications.

Background

A reconstruction task is the solution of:

$$\underset{\boldsymbol{x}\in\mathbb{R}^n}{\text{minimize}} \;\; g(\boldsymbol{x}) \text{ subject to } f(\boldsymbol{x}) \leq \epsilon$$

where $\epsilon > 0$, $f(\boldsymbol{x})$ is a mapping constructed as a reconstruction error (fidelity term), and $g(\boldsymbol{x})$ is a regularizer.

Limitations

- Global guarantees are not available for non-convex mappings.
- We extend the capability of ADMM and accelerated proximal gradient methods to handle.

Contact Information

Email: samuel.pinilla@stfc.ac.uk

References

- Yu Wang, Wotao Yin, and Jinshan Zeng, "Global convergence of ADMM in nonconvex nonsmooth optimization," Journal of Scientific Computing.
- Pinilla, S., Mu, T., Bourne, N., Thiyagalingam, J. (2022). Improved imaging by invex regularizers with global optima guarantees. Advances in Neural Information Processing Systems, 35, 10780-10794.

Material and Methods

Proposed family of functions

Definition Let $h:\mathbb{R}^n \to \mathbb{R}$ such that $h(\boldsymbol{x}) = \sum_{i=1}^n s(|\boldsymbol{x}[i]|)$, where $s:[0,\infty) \to [0,\infty)$ and $s'(w) > 0$ for $w \in (0,\infty)$. If s with $s(0)=0$ such that $s(w)/w^2$ is non-increasing on $(0,\infty)$, then $h(\boldsymbol{x})$ is said to be an *admissible function*.

Properties of proposed family of functions

Theorem 1. Let $f,g:\mathbb{R}^n \to \mathbb{R}$ be two admissible functions as in Definition , such that $f(\boldsymbol{x}) = \sum_{i=1}^n s_f(|\boldsymbol{x}[i]|)$, and $g(\boldsymbol{x}) = \sum_{i=1}^n s_g(|\boldsymbol{x}[i]|)$. Then the following holds:

- $f(\boldsymbol{x})$, and $g(\boldsymbol{x})$ are invex;
- $h(\boldsymbol{x}) = \alpha f(\boldsymbol{x}) + \beta g(\boldsymbol{x})$ is an admissible function (therefore invex) for every $\alpha, \beta \geq 0$;
- $h(\boldsymbol{x}) = \sum_{i=1}^n (s_f \circ s_g)(|\boldsymbol{x}[i]|)$ is admissible function.
- $h(\boldsymbol{x}) = \sum_{i=1}^n \min(s_f(|\boldsymbol{x}[i]|), s_g|\boldsymbol{x}[i]|)$ is admissible function.
- $h(\boldsymbol{x}) = \sum_{i=1}^n \max(s_f(|\boldsymbol{x}[i]|), s_g|\boldsymbol{x}[i]|)$ is admissible function.

Examples of Invex functions

$$\text{(Cauchy)} \quad f(\boldsymbol{x}) = \sum_{i=1}^{m} \log\left(1 + \frac{\boldsymbol{x}^2[i]}{\delta^2}\right) \tag{4}$$

$$\text{(Geman-McClure)} \quad f(\boldsymbol{x}) = \sum_{i=1}^{m} \frac{2\boldsymbol{x}^2[i]}{\boldsymbol{x}^2[i] + 4\delta^2} \tag{5}$$

$$\text{(Welsh)} \quad f(\boldsymbol{x}) = \sum_{i=1}^{n} 1 - \exp\left(\frac{-\boldsymbol{x}^2[i]}{2\delta^2}\right) \tag{6}$$

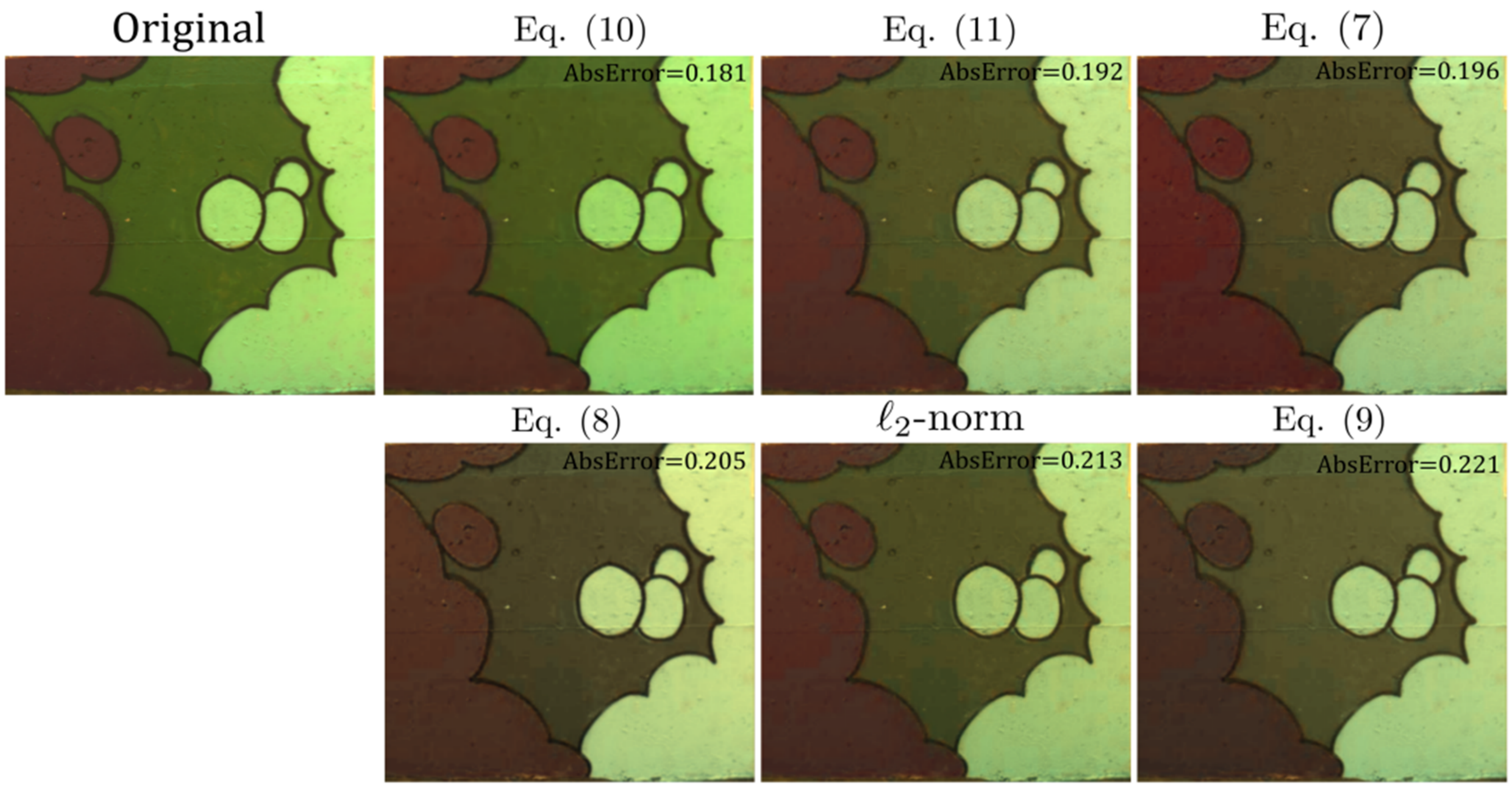
$$\text{(Adaptive robust)} \quad f(\boldsymbol{x}) = \sum_{i=1}^{n} \frac{|\alpha-2|}{\alpha}\left(\left(\frac{(\boldsymbol{x}[i]/c)^2}{|\alpha-2|}+1\right)^{\alpha/2} - 1\right) \tag{7}$$

$$f(\boldsymbol{x}) = \sum_{i=1}^{m} \log\left(1+\boldsymbol{x}^2[i]\right) - \frac{\boldsymbol{x}^2[i]}{2\boldsymbol{x}^2[i]+2} \tag{8}$$

Applications of Invex functions

Invex	Application
equation (3)	Compressive sensing
Invex $g_{TV}(\boldsymbol{x})$	Tomography Total variation filtering
equation (4)	Robust Learning
equation (5)	Neural Radiance Fields
equations (6),(7)	Adaptive Filtering
equation (8)	Supervised Learning

Experiments



Experiment 2 (Loss functions used to train MST++)

Metrics	equation (4)	equation (5)	equation (6)	equation (7)	equation (8)	ℓ_2-norm
AbsError	0.1975	0.2056	0.2241	0.1830	0.1900	0.2145
SqrtError	1.3752	1.3982	1.4465	1.4219	1.3530	1.3315
RMSE	2.6847	2.9390	3.2465	2.1315	2.4710	2.2887
LogRMSE	0.2683	0.2851	0.2943	0.2534	0.2606	0.2765

Experiment 3 (Total variation filtering using ADMM)

Metrics	TV-ℓ_p	Invex TV-LL_p	Quasi-invex TV-LL_p	ℓ_1-norm
SSIM	0.6137	0.6320	0.6227	0.6050
MS-SSIM	0.9192	0.9235	0.9149	0.9106
ADMM-residual	2.3×10^{-3}	1.8×10^{-3}	2.0×10^{-3}	2.8×10^{-3}

Conclusion

- We identified invex/quasi-invex functions to support real-world signal processing problems - signal restoration.
- We provided the proof for the invex behaviours of these functions and global optimality with their convergence rate.
- Numerical results show significant benefits of using the proposed family of invex/quasi-invex functions from theoretical and empirical aspects.

Acknowledgment

This work is partially supported by the EPSRC grant, Blueprinting for AI for Science at Exascale (BASE-II, EP/X019918/1), and by STFC Facilities Fund.