@CVPR 2023 (Highlight)

THU-AM-127

JHU vision lab

On the Convergence of IRLS and Its Variants in Outlier-Robust Estimation

Liangzu Peng* **Christian Kümmerle°** **René Vidal***

*: Johns Hopkins → UPenn

°: University of North Carolina at Charlotte

IRLS + GNC for Outlier-Robust Estimation

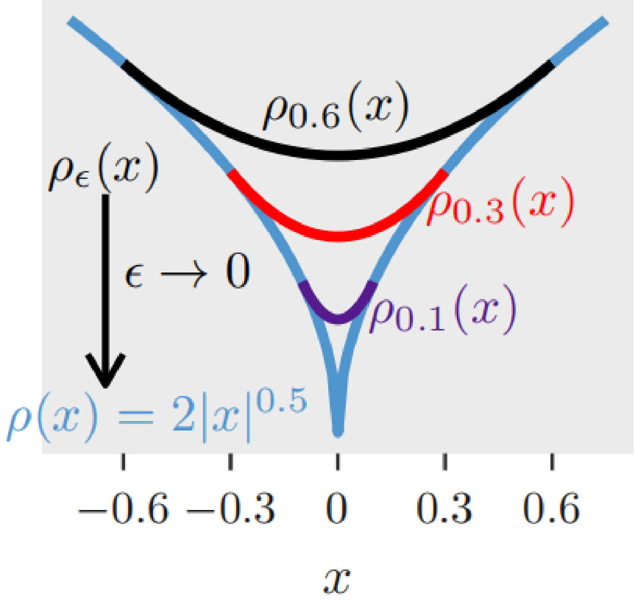
- Outlier-Robust Estimation:

$$\min_{v \in \mathcal{C}} \sum_{i=1}^{m} \rho(\, r(v, d_i)\,)$$

  - $r(v, d_i)$: residual ρ: outlier-robust loss

- Graduated Non-Convexity (GNC):
  - smoothing ρ by some ρ_ϵ

- Contribution:
  - we prove IRLS + GNC converges

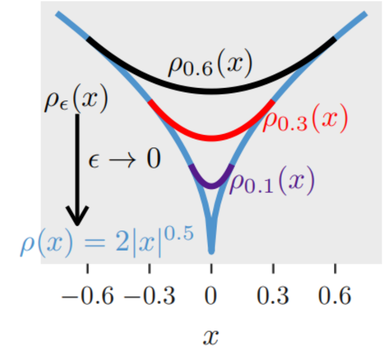


IRLS + GNC for Outlier-Robust Estimation

- Outlier-Robust Estimation:

$$\min_{v \in \mathcal{C}} \sum_{i=1}^{m} \rho(\, r(v, d_i)\,)$$

  - $r(v, d_i)$: residual ρ: outlier-robust loss



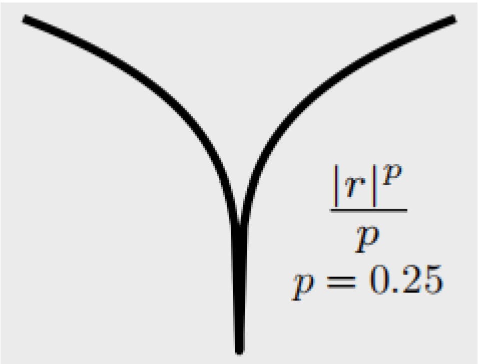
- Graduated Non-Convexity (GNC):
  - smoothing ρ by some ρ_ϵ

- Iteratively Reweighted Least-Squares (IRLS)

IRLS + GNC for Outlier-Robust Estimation

- Outlier-Robust Estimation:

$$\min_{v \in \mathcal{C}} \sum_{i=1}^{m} \rho(\, r(v, d_i) \,)$$



  - $r(v, d_i)$: residual

ρ: outlier-robust loss

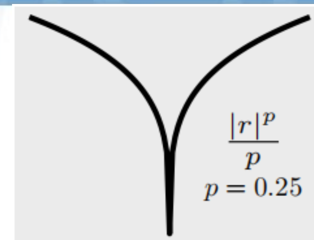
- Linear Regression:
  - $d_i = (x_i, y_i)$
  - $r(v, d_i) = |x_i^\top v - y_i|$
- Point Cloud Registration:
  - $d_i = (x_i, y_i)$, $v = (R, t)$
  - $r(v, d_i) = ||y_i - R x_i - t||_2$
- Essential Matrix Estimation:
  - $d_i = (x_i, y_i)$, $v = E$
  - $r(v, d_i) = |x_i^\top E y_i|$

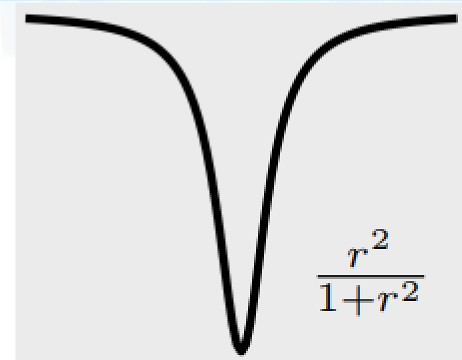
- ℓ_p:
  - $\rho(r) = \frac{|r|^p}{p}$

IRLS + GNC for Outlier-Robust Estimation

- Outlier-Robust Estimation:

$$\min_{v \in \mathcal{C}} \sum_{i=1}^{m} \rho(\, r(v, d_i) \,)$$





  - $r(v, d_i)$: residual

ρ: outlier-robust loss

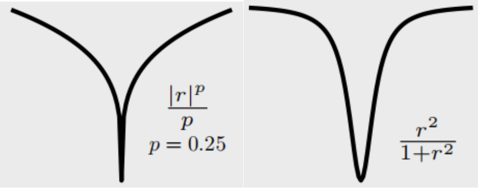
- Linear regression:
  - $d_i = (x_i, y_i)$
  - $r(v, d_i) = |x_i^\top v - y_i|$
- Point Cloud Registration:
  - $d_i = (x_i, y_i)$, $v = (R, t)$
  - $r(v, d_i) = ||y_i - Rx_i - t||_2$
- Essential Matrix Estimation:
  - $d_i = (x_i, y_i)$, $v = E$
  - $r(v, d_i) = |x_i^\top E y_i|$

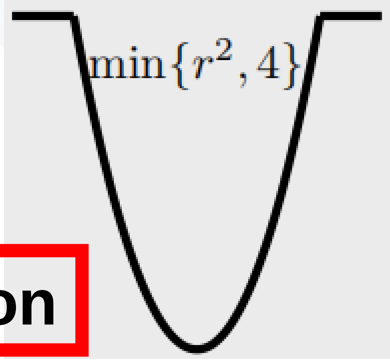
- ℓ_p:
  - $\rho(r) = \frac{|r|^p}{p}$
- Geman-McClure (GM):
  - $\rho(r) = \frac{r^2}{1+r^2}$

IRLS + GNC for Outlier-Robust Estimation

- Outlier-Robust Estimation:

$$\min_{v \in \mathcal{C}} \sum_{i=1}^{m} \rho(\, r(v, d_i)\,)$$





Challenge: Non-Smooth Non-Convex Optimization

- $r(v, d_i)$: residual
- ρ: outlier-robust loss

- Linear regression:
  - $d_i = (x_i, y_i)$
  - $r(v, d_i) = |x_i^\top v - y_i|$
- Point Cloud Registration:
  - $d_i = (x_i, y_i)$, $v = (R, t)$
  - $r(v, d_i) = ||y_i - Rx_i - t||_2$
- Essential Matrix Estimation:
  - $d_i = (x_i, y_i)$, $v = E$
  - $r(v, d_i) = |x_i^\top E y_i|$

- ℓ_p:
  - $\rho(r) = \frac{|r|^p}{p}$
- Geman-McClure (GM):
  - $\rho(r) = \frac{r^2}{1+r^2}$
- TLS:
  - $\rho(r) = \min\{r^2, c^2\}$

IRLS

$$\min_{v \in \mathcal{C}} \sum_{i=1}^{m} \rho\big(r(v, d_i)\big)$$

- IRLS:
  - Initialize v^0
  - Alternate:
    - Weight Update:

$$w_i^{t+1} \leftarrow \frac{\rho'\big(r_i(v^t)\big)}{r(v^t, d_i)}$$

    - Variable Update:

$$v^{t+1} \in \min_{v \in \mathcal{C}} \sum_{i=1}^{m} w_i^{t+1} \cdot r(v^t, d_i)^2$$

- Weight Update
  - Typically done in closed form
- Variable Update:
  - Assume it can be done efficiently
- Not well-defined if
  - ρ is not differentiable

Challenge: Proof of Correctness & Convergence (Rates)!

Graduated Non-Convexity

$$\min_{v \in \mathcal{C}} \sum_{i=1}^{m} \rho\big(r(v, d_i)\big)$$

- IRLS + GNC:
  - Initialize v^0 and ϵ^0
  - Alternate:
    - Weight Update:

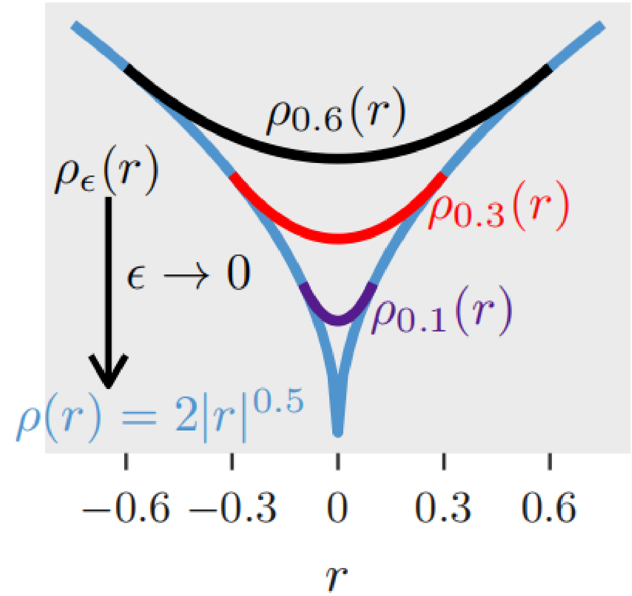
$$w_i^{t+1} \leftarrow \boxed{\frac{\rho'_{\epsilon^t}\big(r(v^t, d_i)\big)}{r(v^t, d_i)}}$$

    - Variable Update:

$$v^{t+1} \in \min_{v \in \mathcal{C}} \sum_{i=1}^{m} w_i^{t+1} \cdot r(v^t, d_i)^2$$

    - Smoothing Parameter Update: decrease ϵ^t to obtain ϵ^{t+1}

- GNC
  - Construct a smoothing approximation ρ_ϵ of ρ ($\epsilon > 0$)
  - Minimize ρ_ϵ instead of ρ, decrease ϵ, and repeat



Challenge: Proof of Convergence (Rates)!

Main Results

$$\min_{v \in \mathcal{C}} \sum_{i=1}^{m} \rho\big(r(v, d_i)\big)$$

- IRLS + GNC:
  - Initialize v^0 and ϵ^0
  - Alternate:
    - Weight Update:

$$w_i^{t+1} \leftarrow \boxed{\frac{\rho'_{\epsilon^t}\big(r(v^t, d_i)\big)}{r(v^t, d_i)}}$$

    - Variable Update:

$$v^{t+1} \in \min_{v \in \mathcal{C}} \sum_{i=1}^{m} w_i^{t+1} \cdot r(v^t, d_i)^2$$

    - Smoothing Parameter Update: decrease ϵ^t to obtain ϵ^{t+1}

- Theorem (Informal):

  Under mild assumptions,
  - IRLS + GNC converges eventually.
  - IRLS converges at a sublinear rate [1].
  - For linear regression with ℓ_p-loss, IRLS + GNC converges at a linear rate if $p = 1$, or a superlinear rate if $0 < p < 1$

[1] L. Peng and R. Vidal. Block Coordinate Descent on Smooth Manifolds. *arXiv:2305.14744 [math.OC]* 2023.

More Information

Please Come to Our Poster Session

THU-AM-127

Vision Lab @ JHU
http://www.vision.jhu.edu

Innovation in Data Engineering and Science (IDEAS) @ UPenn
https://research.seas.upenn.edu/initiatives/data-science/

Thank You!