Re-Weighted ℓ_1 Algorithms within the Lagrange Duality Framework

Matías Valdés, Marcelo Fiori
Universidad de la República, Uruguay

Problem of interest

"Sparsest" solution of

$$\Phi x = b, \quad \Phi \in \mathbb{R}^{m\times n}$$

Equivalently

$$x^* \in \underset{\Phi x = b}{\operatorname{argmin}} \; \|x\|_0 \qquad (P_0)$$

Non-Convex and NP-hard

Alternative: Weighted ℓ_1

$$x^w \in \underset{\Phi x = b}{\operatorname{argmin}} \sum_{i=1}^{n} w_i |x_i| \qquad (P_1 W)$$

How to choose w?

Candès, Wakin, Boyd Re-Weighted ℓ_1

$$w^k \;\; / \;\; w_i^{k+1} = \frac{1}{|x_i^k| + \epsilon_k}, \; k \geq 0,$$

$$x^k \in \underset{\Phi x = b}{\operatorname{argmin}} \sum_{i=1}^{n} w_i^k |x_i|$$

We propose using Lagrange Dual

Methodology with Oracle

x^* solution of (P_0) $\Rightarrow$ ideal problem

$$\underset{\substack{\Phi x = b \\ |x_i| \leq |x_i^*|, \; \forall i}}{\operatorname{argmin}} \; 0 \qquad (P)$$

w as Lagrange multiplier

$$d(w) = \min_{\Phi x = b} \sum_{i=1}^{n} w_i |x_i| - \sum_{i=1}^{n} w_i |x_i^*|$$

Or as Dual solution

$$w^* \in \underset{w \geq 0}{\operatorname{argmax}} \; d(w) \qquad (D)$$

RWℓ_1 Subgradient Algorithm

Projected subgradient for Dual solution

$$w^{k+1} = \max\{0, w^k + \alpha_k g^k\}$$

Subgradient

$$g^k \; / \; g_i^k = g_i(x^k) := |x_i^k| - |x_i^*|$$

$$x^k \in \underset{\Phi x = b}{\operatorname{argmin}} \sum_{i=1}^{n} w_i^k |x_i|$$

Algorithm 1 RWℓ_1 subgradient (with oracle)

Require: $\Phi \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $w^0 \geq 0$, RWIter ≥ 0

```
k = 0
while k ≤ RWIter do
    x^k ∈ argmin_{Φx = b} Σ_{i=1}^n w_i^k |x_i|

    g_i^k = g_i(x^k) = |x_i^k| − |x_i^*|  {subgradient}

    w_i^{k+1} = max(0, w_i^k + α_k g_i^k)  {update}

    k = k + 1
end while
return x^{k−1}
```

Without Oracle

Replace x^* by x^k ("amplified")

$$\underset{\substack{\Phi x = b \\ |x_i| \leq (1+\epsilon_k)\,|x_i^k|, \; \forall i}}{\operatorname{argmin}} \; 0 \qquad (P^k)$$

Dual function

$$d^k(w) = \min_{\Phi x = b} \sum_{i=1}^{n} w_i |x_i| - \sum_{i=1}^{n} w_i \,(1+\epsilon_k)\, |x_i^k|$$

Update w^k applying subgradient to d^k

$$g_i^k = |x_i^k| - (1+\epsilon_k)\,|x_i^k|$$

With Noise: RW-LASSO

$$\underset{\frac{1}{2}\|\Phi x - b\|_2^2 \leq \frac{\eta^2}{2}}{\operatorname{argmin}} \; \|x\|_0 \qquad (P_0^\eta)$$

Convex alternative

$$\underset{\frac{1}{2}\|\Phi x - b\|_2^2 \leq \frac{\eta^2}{2}}{\operatorname{argmin}} \sum_{i=1}^{n} w_i |x_i| \qquad (P_1^\eta W)$$

Oracle Primal

$$\underset{\substack{\frac{1}{2}\|\Phi x - b\|_2^2 \leq \frac{\eta^2}{2} \\ |x_i| \leq |x_i^*|, \; \forall i}}{\operatorname{argmin}} \; 0 \qquad (P^\eta)$$

Relax (all) constraints

$$d(w,\lambda) = \left(\min_{x \in \mathbb{R}^n} \frac{\lambda}{2}\|\Phi x - b\|_2^2 + \sum_{i=1}^{n} w_i |x_i| \right) + C\,(w,\lambda)$$

Same methodology but solving Weighted-LASSO

$$x^k \in \underset{x \in \mathbb{R}^n}{\operatorname{argmin}} \frac{\lambda^k}{2}\|\Phi x - b\|_2^2 + \sum_{i=1}^{n} w_i^k |x_i|$$

Experimental Results

Noise-free Setting

$m = 100$, $n = 256$, Gaussian $\Phi x = b$

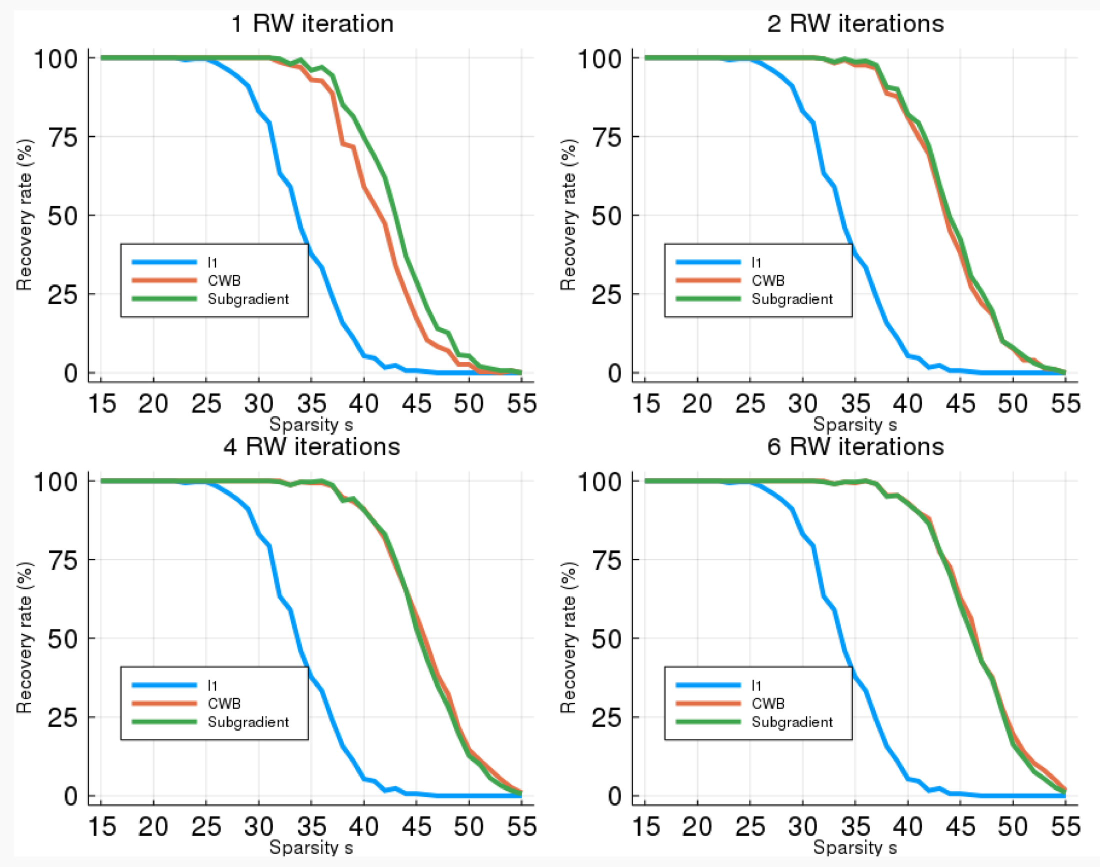

Figure 1: Recovery rate RWℓ_1 (300 problems $\forall$ s)

Similar performance   Interpretability of weights

Noisy Setting

$m = 128$, $n = 256$, Gaussian $z = b - \Phi x^*$

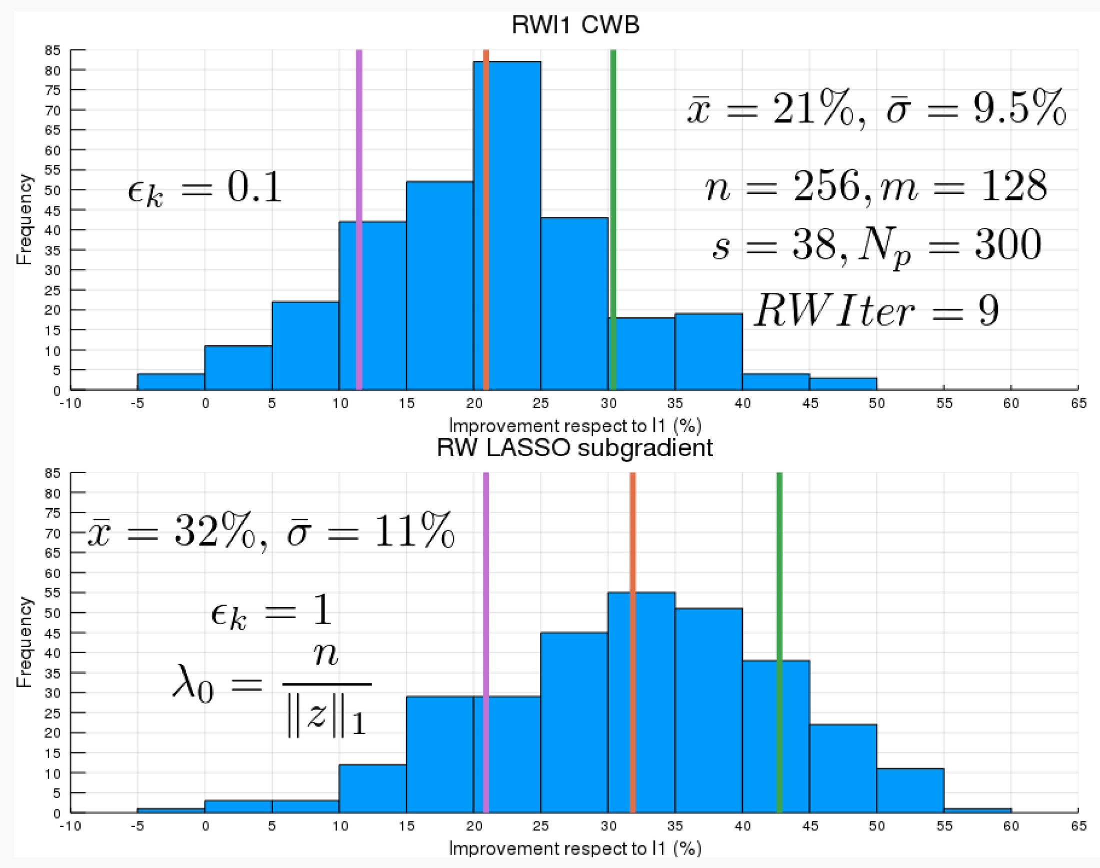

Figure 2: Improvement respect to ℓ_1 (300 tests, $s = 38$)

CWB RWℓ_1: $+21\%$ RW-LASSO subgradient: $+32\%$

References

[1] Candes, E. J., Wakin, M. B., & Boyd, S. P. (2008). Enhancing sparsity by reweighted ℓ1 minimization. Journal of Fourier analysis and applications, 14(5-6), 877-905.