Quadratically fast IRLS for sparse signal recovery

Quadratically fast IRLS for sparse signal recovery

Chiara Ravazzi and Enrico Magli
Dipartimento di elettronica e delle telecomunicazioni, Politecnico di Torino, Italia.

***Abstract*—We present a new class of iterative algorithms for sparse recovery problems that combine iterative support detection and estimation. More precisely, these methods use a two state Gaussian scale mixture as a proxy for the signal model and can be interpreted both as iteratively reweighted least squares (IRLS) and Expectation/Maximization (EM) algorithms for the constrained maximization of the log-likelihood function. Under certain conditions, these methods are proved to converge to a sparse solution and to be quadratically fast in a neighborhood of that sparse solution, outperforming classical IRLS for ℓ_τ-minimization. Numerical experiments validate the theoretical derivations and show that these new reconstruction schemes outperform classical IRLS for ℓ_τ-minimization with $\tau \in (0, 1]$ in terms of rate of convergence and sparsity-undersampling tradeoff.**

I. Sparse recovery via IRLS for ℓ_τ-minimization

The theory of compressed sensing has demonstrated that a k-sparse signal $x^\star \in \mathbb{R}^n$ (*i.e.*, it has at most k nonzero entries) can be recovered from a smaller number $m \ll n$ of linear measurements $y = Ax^\star \in \mathbb{R}^m$ than traditional sampling theory believed necessary [1]. The literature describes a large number of algorithms to recover a sparse signal from an under-determined linear system. An attractive solution is provided by the constrained ℓ_τ-minimization with $\tau \in (0, 1]$ (see [2]) that consists in selecting the element which is compatible with the observations which has minimal ℓ_τ-norm with $\tau \in (0, 1]$:

$$\min_{x\in\mathbb{R}^n} \|x\|_{\ell_\tau} \qquad \text{s.t. } y = Ax. \tag{1}$$

The optimization problem in (1) can be solved by an iteratively rewighted least squares method (IRLS, [3]). More precisely, given an initial guess $x^{(0)}$, at each iteration this class of algorithms requires to solve a constrained weighted least-squares problem:

$$x^{(t+1)} = \underset{y=Ax}{\arg\min} \sum_{i=1}^{n} w_i^{(t)} x_i^2$$

with $w_i^{(t+1)} = ((\epsilon^{(t)})^2 + (x_i^{(t)})^2)^{\tau/2-1}$ and a suitable non-increasing sequence $\epsilon^{(t)}$. Sufficient conditions ensure the convergence of these methods to a sparse solution globally linearly fast when $\tau = 1$ and locally superlinearly fast with rate $2-\tau$ for $\tau \in (0, 1)$.

Although classical IRLS algorithms appear very appealing for their simplicity and for their theoretical guarantees, the superlinear convergence is valid only in a neighborhood of the desired solution. Numerical experiments show that exact recovery is achieved when τ is not too small (*i.e.* $\tau > 1/2$) and tends to be trapped in local minima when $\tau < 1/2$ [4]. The design of heuristic techniques to avoid local minima is still an open issue.

II. GSM-IRLS

In the proposed methods, which we call GSM-IRLS, the elements of the signal are modeled as a two state gaussian mixture (GSM, [5]):

$$x_i^\star = z_i \sqrt{\alpha} u_i + (1 - z_i)\sqrt{\beta} u_i$$

where u_i are identically and independently distributed (i.i.d.) zero mean Gaussians and z_i are i.i.d. Bernoulli variables with probability mass function $\mathbb{P}(z_i = 1) = 1 - p$, $p = k/n$, $\alpha \approx 0$, and $\beta \gg 0$. In the last years, several authors devoted their attention to sparse signal recovery using prior information on the support. In our case the considered model is only used as a proxy for sparse signals [6]. The recovery is performed via the minimization of the negative log-likelihood function subject to the constraint $y = Ax$. One possible procedure is provided by the constrained Expectation Maximization based IRLS (EM-IRLS), whose updates are presented in Algorithm 1.

Algorithm 1 EM-IRLS

Input: Measurements $y \in \mathbb{R}^n$, data matrix $A \in \mathbb{R}^n$

1: Initialization: $\alpha^{(0)} = \alpha_0,\ \beta^{(0)} = \beta_0,\ \pi^{(0)} \in [0, 1]^n$
2: **for** $t = 0, 1, \ldots, StopIter$ **do**
3: Weights update: $w_i^{(t+1)} = \pi_i^{(t)}/\alpha^{(t)} + (1 - \pi_i^{(t)})/\beta^{(t)}$
4: Constrained weighted least squares:
$$x^{(t+1)} = \underset{x\in\mathbb{R}^n : y=Ax}{\arg\min} \sum_{i=1}^{n} w_i (x_i^2 + |\epsilon^{(t)}|^2)$$
5: Posterior beliefs of the signal coefficients:
$$\pi_i^{(t+1)} = \frac{f(x_i^{(t+1)}, \alpha^{(t)}, 1-p)}{f(x_i^{(t+1)}, \alpha^{(t)}, 1-p) + f(x_i^{(t+1)}, \beta^{(t)}, p)}$$
$$f(s, \sigma, q) = \exp\left(-\frac{s^2}{2\sigma} - \frac{\log(\sigma)}{2} + \log(q)\right)$$
6: Parameters update:
$$\alpha^{(t+1)} = \frac{\sum_{i=1}^n \pi_i^{(t+1)} |x_i^{(t+1)}|^2 + |\epsilon^{(t)}|^2}{\sum_{i=1}^n \pi_i^{(t+1)}}$$
$$\beta^{(t+1)} = \frac{\sum_{i=1}^n (1 - \pi_i^{(t+1)}) |x_i^{(t+1)}|^2 + |\epsilon^{(t)}|^2}{\sum_{i=1}^n (1 - \pi_i^{(t+1)})}$$
7: **end for**

Moreover, we consider two other versions of IRLS, which we call ML-IRLS and K-EM-IRLS Algorithm, which differ from EM-IRLS as the beliefs are discrete variables or obtained by thresholding and taking into account that we are seeking a K-sparse signal, respectively. Besides the design of the algorithms, we prove that, under suitable conditions, the sequence of provided estimations converges to a fixed point of the map that rules their dynamics. Moreover, we derive conditions for exact recovery that are verifiable a posteriori. Finally, the algorithms turn out to be quadratically fast in a neighborhood of a sparse solution (see Fig.1). Figures 2-4 compare the performance of GSM-IRLS with classical IRLS methods, Basis Pursuit (BP, [7]), Iterative support detection (Threshold-ISD, [8]) and Orthogonal Matching Pursuit (OMP, [9]), in terms of the empirical recovery success rate, averaged over 50 experiments, as a function of the sparsity level and number of measurements. The recovery is considered successfully when the reconstruction error is below 10^{-4}. Finally, we show that GSM-IRLS converge even in presence of noise (see Fig. 5) and are robust against noise (see Fig. 6), in that small errors on the measurements produce small perturbation in the reconstruction. Form more details the reader can refer to [6].

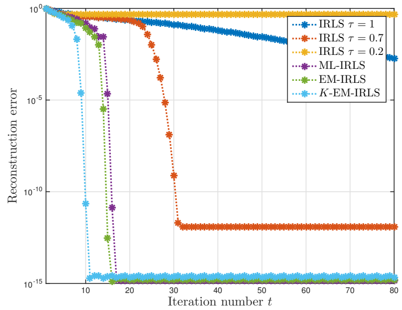

Fig. 1. A typical evolution of the approximation error $E(t) = \|x^{(t+1)} - x^\star\|/\|x^\star\|$ for classical IRLS algorithms (with $\tau = 1, 0.7, 0.2$) and IRLS based on ML and EM. The nonzero components of the signal $x^\star$ are drawn from a uniform distribution $\mathsf{U}([-10, 10])$.

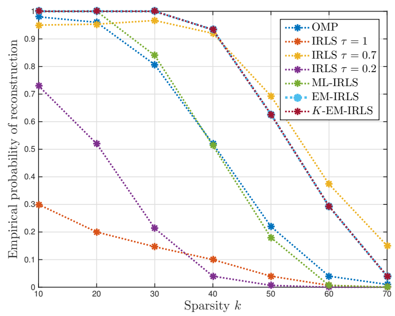

Fig. 2. Empirical probability of successful recovery as a function of the sparsity value k with $n = 512$ and $m = 160$. The nonzero components of the signal $x^\star$ are drawn from a uniform distribution $\mathsf{U}([-10, 10])$.

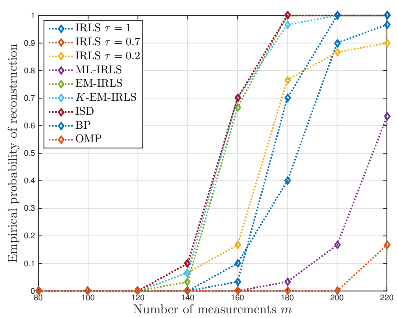

Fig. 3. Empirical probability of successful recovery as a function of the number of measurements $m \in [80, 220]$ with sparse Bernoulli signals $n = 600$ and $k = 40$.

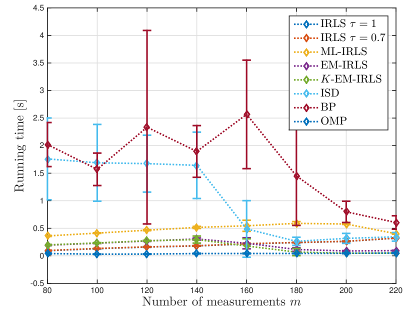

Fig. 4. Average running times (computed over 50 experiments) as a function of the number of measurements $m \in [80, 220]$ with sparse Bernoulli signals $n = 600$ and $k = 40$. The error bar represents the standard deviation of uncertainty.

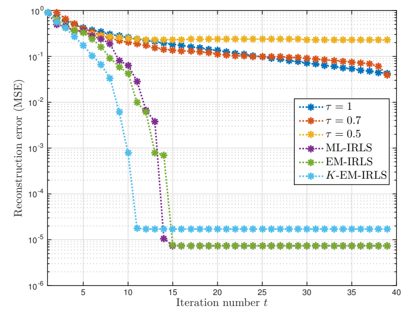

Fig. 5. Noisy scenario: A typical evolution of MSE as a function of the iterations for classical IRLS algorithms (with $\tau = 1, 0.7, 0.2$) and GSM-IRLS. The nonzero components of the signal $x^\star$ are drawn from a uniform distribution $\mathsf{U}([-10, 10])$ and the additive white noise has standard deviation $\sigma = 0.01$.

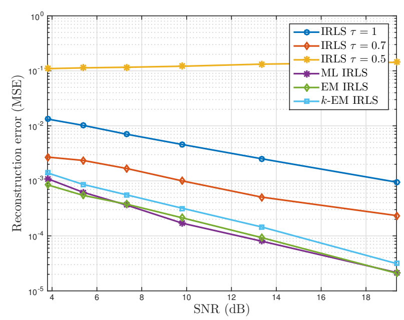

Fig. 6. Robustness: Mean square error after 40 iterations as a function of the SNR for classical IRLS (with $\tau = 1$, $\tau = 0.7$ and $\tau = 0.5$) and the proposed GSM-IRLS with $n = 1500, m = 250, k = 45$. The nonzero components of the signal $x^\star$ are drawn from a uniform distribution $\mathsf{U}([-10, 10])$.

REFERENCES

[1] E. J. Candès, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207 – 1223, 2006.

[2] R. Chartrand, "Exact reconstruction of sparse signals via nonconvex minimization," *IEEE Signal Processing Letters*, vol. 14, no. 10, pp. 707–710, Oct 2007.

[3] I. Daubechies, R. DeVore, M. Fornasier, and C. S. Güntürk, "Iteratively reweighted least squares minimization for sparse recovery," *Comm. Pure Appl. Math.*, vol. 63, no. 1, pp. 1–38, Jan. 2010.

[4] M. Fornasier, *Theoretical Foundations and Numerical Methods for Sparse Recovery*. Radon Series on Computational and Applied Mathematics, 2010.

[5] D. Baron, S. Sarvotham, and R. Baraniuk, "Bayesian compressive sensing via belief propagation," *IEEE Transactions on Signal Processing*, vol. 58, no. 1, pp. 269–280, Jan 2010.

[6] C. Ravazzi and E. Magli, "Gaussian mixtures based IRLS for sparse recovery with quadratic convergence," to appear on *IEEE Transactions on Signal Processing*, pp. 1–16, 2015.

[7] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *Compte Rendus de l'Academie des Sciences. Paris, France, ser. I*, vol. 346, pp. 589 – 592, 2008.

[8] Y. Wang and W. Yin, "Sparse signal reconstruction via iterative support detection," *SIAM J. Img. Sci.*, vol. 3, no. 3, pp. 462–491, aug 2010.

[9] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information theory*, vol. 53, pp. 4655–4666, 2007.