Solving the l_1 regularized least square problem via a one-layer neural network

Majid Mohammadi, Wout Hofman, Yaohua Tan and S. Hamid Mousavi

Abstract—The l_1 regularized least square problem, or the lasso, is a non-smooth convex minimization which is widely-used in diverse fields. However, solving such a minimization is not straightforward since it is not differentiable. In this paper, an equivalent smooth minimization with box constraints is obtained, and it is proved to be equivalent to the lasso problem. Accordingly, an efficient recurrent neural network is developed which guarantees to globally converge to the solution of the lasso. Further, it is investigated that the property "the dual of dual is primal" holds for the l_1 regularized least square problem. The experiments on image and signal recovery illustrate the reasonable performance of the proposed neural network.

Index Terms—sparse, l_1 regularization, smooth, neural network.

I. INTRODUCTION

Finding the solution to the least square problem with the l_1 regularization is of utmost importance due to its myriad applications including but not limited to pattern recognition [30], [31], feature selection [33], [35], image processing [8], [34] and bioinformatics [14], [25]. The minimization to obtain such a solution is

$$\min_{x} \frac{1}{2} ||Ax - b||_{2}^{2} + \lambda ||x||_{1}$$
 (1)

where $x \in R^l$ is the coefficient vector, $A \in R^{n \times l}$ is a basis, $b \in R^{n \times 1}$ is a regression vector, λ is the non-negative regularization parameter, and $\|.\|_p$ is the p-norm.

The challenge of finding the optimal solution of minimization (1) is its non-smoothness, thereby impeding us to leverage fast optimization methods. Therefore, developing efficient algorithms to find the optimal solution of this minimization has drawn a lot of attentions in the recent decade.

One approach to solving the problem is to use two auxiliary variables into which x is split [10]. The resulting problem, though smooth, has the double dimension with respect to the minimization (1) so that obtaining its solution becomes more time- and memory-consuming. Another approach is to utilize the dual of the minimization (1), which is a smooth boxconstrained problem [16], [36]. The main drawback is the

calculation of the primal solution x from the dual solution, which needs the computation of $(A^TA)^{-1}$. In a general sense, however, A^TA is not necessarily non-singular, and is not thus invertible. Even if it is invertible, it would be prolonged to compute $(A^TA)^{-1}$ for large-scale problems.

Yet another strategy to solve the problem (1) is based on the subgradient [5]. Subgradient-based methods are also too slow in comparison with the gradient-based techniques which makes them inappropriate in the case of large-scale problems.

Neural networks have been long considered for solving optimization problems. There are diverse neural solutions to various types of optimizing problems; linear [29] and nonlinear [13], smooth [32] and non-smooth [27], and so on. Recently, a novel neural network is proposed to solve the absolute value equation (APV) [19]. The neural network is guaranteed to converge to the exact solution of the APV. Identical to the lasso, the APV is non-smooth and its challenge is the existence of the absolute value operator in its equation. However, it is a much simpler problem rather than the minimization (1).

In this article, we firstly utilize the dual problem of the minimization (1) and show that the property "the dual of dual is primal" is correct for the minimization (1). Such a property holds for all convex linear programming but does not generally hold for all convex nonlinear problems. We further derive a smooth problem with box constraints which is equivalent to the minimization (1). Moreover, a new recurrent neural network for the derived smooth problem is proposed, and its global convergence is guaranteed. On account of the simplicity, we initially suppose that A^TA is invertible, but neither the consequent smooth problem nor the proposed neural network would be predicated on the invertibility of A^TA .

This article is organized as follows. The smooth equivalent problem is derived in Section II, and an efficient neural network is proposed in Section III. The convergence and stability of the proposed neural solution is discussed in Section IV, and it follows by an extensive experiments over multiple realworld problem in Section V. Finally, the paper is concluded in Section VI.

II. SMOOTH EQUIVALENT PROBLEM

In this section, we firstly introduce the dual of the minimization (1) and prove that the dual of dual is primal. Then, a smooth equivalent problem to minimization (1) is obtained.

A. The dual of dual is primal

Since the minimization (1) is not smooth, its dual problem cannot be derived immediately. However, the dual problem can

M. Mohammadi and Yaohua Tan are with the Faculty of Technology, Policy and Management, Delft University of Technology, Netherlands. E-mail: m.mohammadi@tudelft.nl and y.tan@tudelft.nl.

Wout Hofman is a senior researcher at TNO, Netherlands. E-mail: wout.hofman@tno.nl.

S. H. Mousavi is with the Department of Medical Physics and Acoustics and Cluster of Excellence Hearing4all, Carl von Ossietzky University of Oldenburg, Germany. E-mail: hamid.mousavi@uni-oldenburg.de.

be obtained by a change of variable and using a theorem in [4] as (see [28] for the whole procedure):

$$\min_{z} \frac{1}{2} z^{T} (A^{T} A)^{-1} z + z^{T} (A^{T} A)^{-1} A^{T} b$$

$$s.t. -\lambda \le z_i \le \lambda i = 1, 2, ..., l. (2)$$

Further, the relation between the primal and dual variables is

$$x = (A^T A)^{-1} (A^T b - z). (3)$$

The minimization (2) is smooth and convex and can be solved efficiently by convex minimization methods. Afterward, the solution z is replaced in Eq. (3) to obtain the primal solution x.

Now, the following theorem indicates that the dual of the minimization (2) is the minimization (1).

Theorem 1: Let u and v be the Lagrangian multipliers for the minimization (2). Then

- (i) the solution of the primal problem (1) is the difference between the Lagrangian multipliers u, v in the dual problem (2), i.e. x = u v;
- (ii) The property "the dual of dual is primal" holds for the minimization (1).

Proof: (i) Let u and v be the Lagrangian multipliers for the problem (2). According to the K.K.T conditions, we have

$$(A^{T}A)^{-1}z - (A^{T}A)^{-1}A^{T}b + u - v = 0.$$
 (4)

Replacing $z = A^T b - A^T A x$ from the Eq. (3) in the above equation, we have x = u - v.

(ii) To get the dual, the Lagrangian of the dual problem is firstly written as

$$J = \frac{1}{2}z^{T}(A^{T}A)^{-1}z - z^{T}(A^{T}A)^{-1}A^{T}b$$
$$+u^{T}(z-\lambda) + v^{T}(-\lambda - z)$$

Since $z = A^T b - (A^T A)(u - v)$ based on Eq. (4), the Lagrangian function can be rewritten as

$$J = \frac{1}{2}(u - v)^T A^T A(u - v) + \frac{1}{2}(u - v)^T A^T b$$
$$- \frac{1}{2}b^T A(u - v) + u^T (A^T b - (A^T A)(u - v) - \lambda)$$
$$+ v^T (-\lambda - A^T b + A^T A(u - v))$$
$$\Rightarrow J = -\frac{1}{2}(u - v)^T A^T A(u - v) + (u - v)^T A^T b - \lambda (u - v).$$

Since $x=u-v, u, v\geq 0$ and uv=0 according to K.K.T conditions of the dual problem (2), we have $\|x\|_1=u+v$. Hence, the dual problem based on the above equation is

$$\max_{x} - \|Ax - b\|_{2}^{2} - \lambda \|x\|_{1} \tag{5}$$

which is equivalent to the minimization (1) and the proof is complete.

B. Smooth problem

The smooth equivalent problem to the minimization (1) can be easily obtained by replacing z in the dual problem. Thanks to Eq. (3), we have

$$z = A^T b - A^T A x. (6)$$

Replacing (6) into the minimization (2) and doing some calculus, we obtain

$$\min_{x} x^{T} A^{T} A x$$

$$s.t. -\lambda \le A^T A x - A^T b \le \lambda. \tag{7}$$

The problem (7) is convex with box constraints. As it is smooth, the optimization methods for constrained problems can be applied to find its solution. In contrast to the dual problem, there is no need to compute the inverse matrix $(A^TA)^{-1}$ for solving this problem. Further, it directly obtains the solution that is desired, e.g. x.

In further sections, an efficient recurrent neural network is proposed for solving the minimization (7) and its convergence is meticulously examined.

III. ONE-LAYER NEURAL NETWORK

The minimization (7) is convex, the Karush-Kuhn-Tucker (K.K.T.) conditions are thus sufficient for optimality [1]. Using the K.K.T conditions, the problem (7) is turned into an equation whose solution is the same as the minimization.

Theorem 2: x is the optimal solution of the minimization (7) if and only if the following equality holds

$$P_{\Omega}(A^T A x - A^T b - x) = A^T A x - A^T b \tag{8}$$

where $P_{\Omega}(.)$ is a piecewise function defined as

$$(P_{\Omega}(w))_{i} = \begin{cases} \lambda & w_{i} > \lambda \\ w_{i} & |w_{i}| \leq \lambda \\ -\lambda & w_{i} < -\lambda \end{cases}$$

Proof: The equation (8) can be easily obtained by writing the K.K.T conditions for the smooth problem (7) (see [2] for details).

Based on this theorem, an efficient neural network is proposed with the dynamical equation being given by

$$\frac{dx}{dt} = P_{\Omega}(A^T A x - A^T b - x) + A^T b - A^T A x. \tag{9}$$

The dynamic system can be readily recognized as a recurrent neural network with one-layer structure. In the next section, the convergence of such neural network is discussed.

IV. CONVERGENCE ANALYSIS

In this section, we study the proposed neural network (9) and investigate its properties. We first show that there exists a unique solution for the system (9), and then examine its convergence and stability.

Definition 1: A continuous-time neural network is globally convergent if the trajectory of its dynamical system converges to an equilibrium point for any arbitrary initialization.

Lemma 1 ([12]): For any closed convex set Ω ,

(i)
$$(v - P_{\Omega}(v))^T (P_{\Omega}(v) - x) \ge 0, \quad v \in \mathbb{R}^l, \quad x \in \Omega;$$

(ii)
$$||P_{\Omega}(u) - P_{\Omega}(v)|| \le ||u - v|, \quad u, v \in \mathbb{R}^l$$
 (10)

where $\|.\|$ is the l_2 norm.

Lemma 2: Given any arbitrary initialization, there is a unique continuous solution x(t) for the dynamic system (9). Further, the equilibrium of this system solves the minimization (1).

Proof: The right-hand side of the equation (9) is Lipschitz continuous, thanks to Lemma 1. Therefore, there is a unique continuous solution x(t) for any given initial point. Moreover, the equilibrium of the system (9) solves the minimization (1) according to Theorem 2.

Theorem 3: The proposed neural network (9) with the initial point $x(t_0) \in \mathbb{R}^n$ is stable in the sense of Lyapunov and globally converges to the solution of the minimization (7).

Proof: Assume that x(t) is the trajectory obtained from the system (9) with initial point x_0 . Setting $v = A^TAx - A^Tb - x$ and $u = A^TAx^* - A^Tb$ in the first equation of Lemma 1, where x^* is the equilibrium of the system (9), we have

$$(P_{\Omega}(A^{T}Ax - A^{T}b - x) - A^{T}Ax^{*} + A^{T}b)^{T}$$
$$(A^{T}Ax - A^{T}b - x - P_{\Omega}(A^{T}Ax - A^{T}b - x)) \ge 0(11)$$

On the other hand, $(v - A^TAx + A^Tb)^Tx^* \ge 0$ for each $v \in \Omega$ since x^* is the equilibrium of the dynamical system (9). Summing this inequality with Eq. (11) and doing some calculus, it follows

$$(x - x^*)^T (I + A^T A) \left(P_{\Omega} (A^T A x - A^T b - x) - A^T A x + A^T b \right)$$

$$\leq -(x - x^*)^T (A^T A) (x - x^*)$$

$$- \| P_{\Omega} (A^T A x - A^T b - x) - A^T A x + A^b \|^2.$$
(12)

Now, consider the Lyapunov function

$$V(x) = \frac{1}{2} ||K(x - x^*)||$$

where $K^TK = (I + A^TA)$. Then,

$$\frac{d}{dt}V(x) = \left(\frac{dV}{du}\right)^T \frac{du}{dt}$$

$$= (x - x^*)(I + A^T A) \left(P_{\Omega}(A^T A x - A^T b - x) - A^T A x + A^T b\right)$$

$$\leq -(x - x^*)^T A^T A(x - x^*)$$

$$-\|P_{\Omega}(A^{T}Ax - A^{T}b - x) - A^{T}Ax + A^{T}b\|^{2}$$
(13)

where the last inequality is derived using Eq. (12). Based on Eq. (13), $dv/dt \leq 0$ and the dynamical system (9) is stable in the sense of Lyapunov. Further, it is easy to verify that dv/dt = 0 if and only if dx/dt = 0. Besides, $V(x) > \beta || u - u^* ||^2 / 2$ where β is the smallest eigenvalue of $I + A^T A$. Thus, the proposed neural network is globally convergent to to the solution of the lasso, and that completes the proof.

V. EXPERIMENTAL RESULTS

In the section, the experimental results are presented for three different problems. First, a randomly-generated sparse signal is recovered by solving the l_1 regularized least square problem. It follows by two experiments over the image restoration and aCGH signal recovery.

A. Signal recovery

We generated a sparse signal $x_0 \in R^{4096}$ with 160 spikes; each spike has amplitude $\{-1,1\}$. This signal is plotted at the top of Fig. 1. A Gaussian noise with the standard variation $\sigma=0.1$ is added to x_0 to generate the observation y. Then, the measurement matrix $A \in R^{1024 \times 4096}$ is generated whose entries are i.i.d. according to the standard normal distribution. The rows of this matrix are then orthogonalized as done in [6]. The regularization parameter λ is also set as $\lambda=0.01\|A^Tb\|_{\infty}$ as any value greater than $\|A^Tb\|_{\infty}$ for λ leads to the solution zero for the minimization (1) [11], [17]. Given A, y and λ , it is the time estimate x_0 by solving the minimization (1).

The result of the recovery is shown in Fig. 1. The top plot in this figure corresponds to the original noise-free signal which is randomly generated. The bottom plot is the recovered signals by the proposed method, and the middle one is the minimal L_2 norm of the system Ax = b. This figure confirms that the recovered signal is faithful in spite of having a few measurements.

B. Image restoration

The proposed neural network is applied to the image restoration problem and its result is compared with the proximal gradient. To do so, an MRI image is chosen which is plotted in Fig. 2(a). Further, a random Gaussian noise with $\sigma=0.05$ is added to the image and the resulting image is depicted in Fig. 2(b). Given the image, the restoration is done by the minimization (1) solved by the proposed method and the proximal gradient [23]. The recovered images are shown in Fig. 2 (c) and (d) for the proposed method and the proximal gradient, respectively. The mean square of the neural network is 2.6×10^{-3} while that of the proximal gradient is 4.08×10^{-4} .

Fig. 1: Sparse signal recovery by the proposed method. Top: the randomly generated signal $x_0 \in R^{4096}$ with 160 spikes. Middle: the minimal norm solution obtained by A^Tx_0 . Bottom: the recovered signal by the proposed method.

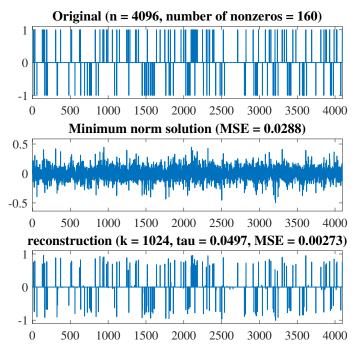
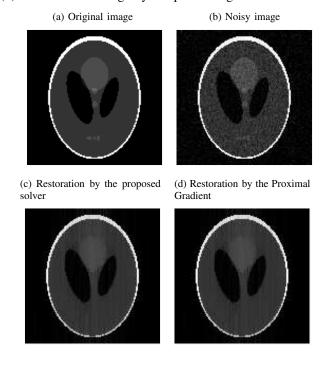


Fig. 2: The restoration image experiment across an MRI image. (a) the original image, (b) the corrupted image with a noise distributed according to the normal distribution with $\sigma=0.05$, (c) the recovered image by the proposed method, (d) the recovered image by the proximal gradient.



C. CGH array data recovery

The array comparative genome hybridization (aCGH or CGH array) is a powerful technique to discover the genome-wide DNA copy number variations [24]. However, the experiential aCGH data are highly corrupted by various noises thereby disabling us to find the change-points from the raw data [9].

One underlying assumption in aCGH data is that the contiguous chromosome has the identical copy number unless an alteration has happened. Based on this critical assumption, myriad methods have utilized the total variation regularization, whether they process individual samples separately [15], [18], [20] or process multisample data simultaneously [3], [21], [22], [37], [38].

We apply the proposed neural network to the CGH array from two breast cancer datasets. The Pollack et al. dataset [26] consists of 6691 human mapped genes for 44 primary breast tumors, and the Chin et al. dataset [7] has 2149 clones from 141 primary breast tumors.

Multiple recovered profiles from the aforementioned datasets are plotted in Figure 3. In this figure, the profiles from the methods TVSp [38] and PLA [37] are also shown. The red dots indicates the raw data and the blue lines are the recovered data by methods. Figure 3 (a) and (b) correspond to two samples from Pollack et al. [26] and Chin et al. [7] datasets, respectively. It is plain to grasp that the proposed method has successfully recovered smooth data from the noisy observations. The recovered profiles by the proposed solver are much smoother than PLA and are competitive with TVSp.

VI. CONCLUSION

This paper presented an efficient neural network which is proved to globally converge to the solution of the lasso. To obtain such a neural solution, a smooth box-constrained minimization is derived, and it is demonstrated that it is equivalent to the lasso problem. Extensive experimental results illustrate the reasonable performance of the proposed neural network.

REFERENCES

- Mokhtar S Bazaraa, Hanif D Sherali, and Chitharanjan M Shetty. Nonlinear programming: theory and algorithms. John Wiley & Sons, 2013.
- [2] Dimitri P Bertsekas and John N Tsitsiklis. Parallel and distributed computation: numerical methods, volume 23. Prentice hall Englewood Cliffs. NJ, 1989.
- [3] Kevin Bleakley and Jean-Philippe Vert. The group fused lasso for multiple change-point detection. arXiv preprint arXiv:1106.4199, 2011.
- [4] Stephen Boyd and Lieven Vandenberghe. Convex optimization. Cambridge university press, 2004.
- [5] Stephen Boyd, Lin Xiao, and Almir Mutapcic. Subgradient methods. lecture notes of EE392o, Stanford University, Autumn Quarter, 2004:2004– 2005, 2003.
- [6] Emmanuel Candes and Justin Romberg. 11-magic: A collection of matlab routines for solving the convex optimization programs central to compressive sampling. Avaliable: www. acm. caltech. edu/l1magic, 2006
- [7] Koei Chin, Sandy DeVries, Jane Fridlyand, Paul T Spellman, Ritu Roydasgupta, Wen-Lin Kuo, Anna Lapuk, Richard M Neve, Zuwei Qian, Tom Ryder, et al. Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. *Cancer cell*, 10(6):529–541, 2006.

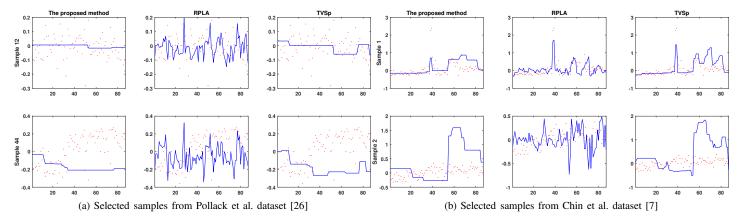


Fig. 3: Recovered profiles by the proposed method, TVSp [38] and PLA [37]. (a) two samples selected from the Pollack et al. dataset [26]; (b) two samples selected from the Chin et al. dataset [7]. Each row is dedicated to each sample, and each column is devoted to each method. The red dots are the raw observations, and the blue lines are the recovered profiles by each method.

- [8] Michael Elad and Michal Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions* on *Image processing*, 15(12):3736–3745, 2006.
- [9] Lars Feuk, Andrew R Carson, and Stephen W Scherer. Structural variation in the human genome. *Nature Reviews Genetics*, 7(2):85–97, 2006.
- [10] Mário AT Figueiredo, Robert D Nowak, and Stephen J Wright. Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems. *IEEE Journal of selected topics in signal* processing, 1(4):586–597, 2007.
- [11] J-J Fuchs. On sparse representations in arbitrary redundant bases. *IEEE transactions on Information theory*, 50(6):1341–1344, 2004.
- [12] Richard Fujimoto. Parallel and distributed simulation. In *Proceedings* of the 2015 Winter Simulation Conference, pages 45–59. IEEE Press, 2015
- [13] Xing-Bao Gao. A novel neural network for nonlinear convex programming. IEEE Transactions on Neural Networks, 15(3):613–621, 2004.
- [14] Xiyi Hang and Fang-Xiang Wu. Sparse representation for classification of tumors using gene expression data. *BioMed Research International*, 2009, 2009.
- [15] Jing Hu, Jian-Bo Gao, Yinhe Cao, Erwin Bottinger, and Weijia Zhang. Exploiting noise in array cgh data to improve detection of dna copy number change. *Nucleic acids research*, 35(5):e35, 2007.
- [16] Jingu Kim and Haesun Park. Fast active-set-type algorithms for 11regularized linear regression. In AISTATS, pages 397–404, 2010.
- [17] S Kim, K Koh, M Lustig, S Boyd, and D Gorinevsky. A method for large-scale 11-regularised least squares problems with applications in signal processing and statistics. *IEEE J. Select. Topics Signal Process*, 2007.
- [18] Yoon-ha Lee, Michael Ronemus, Jude Kendall, B Lakshmi, Anthony Leotta, Dan Levy, Diane Esposito, Vladimir Grubor, Kenny Ye, Michael Wigler, et al. Removing system noise from comparative genomic hybridization data by self-self analysis. arXiv preprint arXiv:1105.0900, 2011
- [19] Amin Mansoori, Mohammad Eshaghnezhad, and Sohrab Effati. An efficient neural network model for solving the absolute value equations. IEEE Transactions on Circuits and Systems II: Express Briefs, 2017.
- [20] Apratim Mitra, George Liu, and Jiuzhou Song. A genome-wide analysis of array-based comparative genomic hybridization (cgh) data to detect intra-species variations and evolutionary relationships. *PloS one*, 4(11):e7978.
- [21] Majid Mohammadi, Ghosheh Abed Hodtani, and Maryam Yassi. A robust correntropy-based method for analyzing multisample acgh data. *Genomics*, 2015.
- [22] Hossein Sharifi Noghabi, Majid Mohammadi, and Yao-Hua Tan. Robust group fused lasso for multisample copy number variation detection under uncertainty. *IET Systems Biology*, 10(6):229–236, 2016.
- [23] Neal Parikh, Stephen Boyd, et al. Proximal algorithms. Foundations and Trends® in Optimization, 1(3):127–239, 2014.

- [24] Daniel Pinkel and Donna G Albertson. Array comparative genomic hybridization and its applications in cancer. *Nature genetics*, 37:S11– S17, 2005.
- [25] Roger Pique-Regi, Jordi Monso-Varona, Antonio Ortega, Robert C Seeger, Timothy J Triche, and Shahab Asgharzadeh. Sparse representation and bayesian detection of genome copy number alterations from microarray data. *Bioinformatics*, 24(3):309–318, 2008.
- [26] Jonathan R Pollack, Therese Sørlie, Charles M Perou, Christian A Rees, Stefanie S Jeffrey, Per E Lonning, Robert Tibshirani, David Botstein, Anne-Lise Børresen-Dale, and Patrick O Brown. Microarray analysis reveals a major direct role of dna copy number alteration in the transcriptional program of human breast tumors. *Proceedings of the National Academy of Sciences*, 99(20):12963–12968, 2002.
- [27] Sitian Qin and Xiaoping Xue. A two-layer recurrent neural network for nonsmooth convex optimization problems. *IEEE transactions on neural* networks and learning systems, 26(6):1149–1160, 2015.
- [28] Mark Schmidt. Two dual problems to 1 1 regularized least squares. 2008.
- [29] Jun Wang. Analysis and design of a recurrent neural network for linear programming. IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications, 40(9):613–618, 1993.
- [30] John Wright, Yi Ma, Julien Mairal, Guillermo Sapiro, Thomas S Huang, and Shuicheng Yan. Sparse representation for computer vision and pattern recognition. *Proceedings of the IEEE*, 98(6):1031–1044, 2010.
- [31] John Wright, Allen Y Yang, Arvind Ganesh, S Shankar Sastry, and Yi Ma. Robust face recognition via sparse representation. *IEEE transactions on pattern analysis and machine intelligence*, 31(2):210–227, 2009.
- [32] Youshen Xia and Jun Wang. A general methodology for designing globally convergent optimization neural networks. *IEEE Transactions* on Neural Networks, 9(6):1331–1343, 1998.
- [33] Allen Y Yang, John Wright, Yi Ma, and S Shankar Sastry. Feature selection in face recognition: A sparse representation perspective. submitted to IEEE Transactions Pattern Analysis and Machine Intelligence, 2007.
- [34] Jianchao Yang, John Wright, Thomas S Huang, and Yi Ma. Image super-resolution via sparse representation. *IEEE transactions on image* processing, 19(11):2861–2873, 2010.
- [35] Meng Yang and Lei Zhang. Gabor feature based sparse representation for face recognition with gabor occlusion dictionary. In European conference on computer vision, pages 448–461. Springer, 2010.
- [36] Pinghua Gong Changshui Zhang. A fast dual projected newton method for 11-regularized least squares. Tsinghua University, Beijing, 2011.
- [37] Xiaowei Zhou, Jiming Liu, Xiang Wan, and Weichuan Yu. Piecewise-constant and low-rank approximation for identification of recurrent copy number variations. *Bioinformatics*, page btu131, 2014.
- [38] Xiaowei Zhou, Can Yang, Xiang Wan, Hongyu Zhao, and Weichuan Yu. Multisample acgh data analysis via total variation and spectral regularization. Computational Biology and Bioinformatics, IEEE/ACM Transactions on, 10(1):230–235, 2013.

