SPARSE SIGNAL RECOVERY VIA CONVEX OPTIMIZATION

Michaela Novosadová
Doctoral Degree Programme (2.), FEEC BUT
E-mail: xnovos11@stud.feec.vutbr.cz

Supervised by: Pavel Rajmic
E-mail: rajmic@feec.vutbr.cz

Abstract: We propose recovering 1D piecewice linear signal using a sparsity-based method consisting of two steps. The first step is signal segmentation via optimization algorithms solving sparsity based model. Second step consists of applying an ordinary mean square method on each detected segment of the signal. We show results of our experiments on two types of the signal.

Keywords: convex optimization, Douglas-Rachford algorithm, Forward-Backward algorithm, signal recovery, signal segmentation

1 INTRODUCTION

We explore signal recovering by use combination of convex optimization methods for segmenting of noisy piecewise linear signal and signal denoising by use of mean square method at each detected segment of the signal. The number of segments is considerably lower than the number of signal samples, which suggests using sparse signal processing techniques [1].

2 PROBLEM FORMULATION

The overparameterization model should be introduced first. We suppose 1D piecewise linear signal $\mathbf{f}$. Each element of the signal $\mathbf{f}$ can be described by two parameterization coefficients: constant offset (parameter a) and constant slope (parameter b). So the i-th element of the signal is defined as $f_i = a_i + b_i i$, alternatively

$$\mathbf{f} = \mathbf{a} + \mathbf{D}\mathbf{b} \quad \text{i.e.} \quad \mathbf{f} = [\mathbf{I}\,\mathbf{D}] \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = \mathbf{A}\mathbf{x}, \tag{1}$$

where $\mathbf{a}$, $\mathbf{b} \in \mathbb{R}^N$, $\mathbf{I} = \mathbf{I}_N$ is the identity matrix, $\mathbf{D} = \operatorname{diag}(1,2,\ldots,N)$ is the diagonal matrix with the values $1,2,...,N$ on its main diagonal. Note that vectors $\mathbf{a}$ and $\mathbf{b}$ are piecewise constant in each segment so they are sparse under the difference operation. Due to this assumption we can formulate the recovery problem using total variation (TV):

$$\hat{\mathbf{a}}, \hat{\mathbf{b}} = \underset{\mathbf{a},\mathbf{b}}{\arg\min}\ \frac{1}{2} \left\| \mathbf{y} - [\mathbf{I}\,\mathbf{D}] \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} \right\|_2^2 + \tau_{\mathbf{a}} \operatorname{TV}(\mathbf{a}) + \tau_{\mathbf{b}} \operatorname{TV}(\mathbf{b}), \tag{2}$$

where $\mathbf{y}$ is observed signal which is corrupted by uncorrelated Gaussian noise $\mathbf{e}$ with zero mean and positive variance $\mathbf{y} = \mathbf{f} + \mathbf{e}$, and $\operatorname{TV}(\cdot)$ is the total variation functional defined as

$$\operatorname{TV}(\mathbf{z}) = \|\nabla \mathbf{z}\|_1 = \sum_{i=1}^{N-1} |z_{i+1} - z_i|, \tag{3}$$

and finally $\hat{\mathbf{a}}, \hat{\mathbf{b}}$ are achieved optimizers. The nonzero values in $\nabla\hat{\mathbf{a}}$ and $\nabla\hat{\mathbf{b}}$ indicate possible segment borders, the nonzero values should be on shared positions. Positive constants $\tau_{\mathbf{a}}, \tau_{\mathbf{b}}$ are regularization weights, with their values depending on the properties of the signal and noise level. The regularization weights should be carefully tuned.

3 METHODOLOGY

The solution of signal denoising is divided into two parts. First, the breakpoints are detected describing each segment of the signal. This is done by proximal splitting methods. Second is the denoising of each segment of the signal by least squared method.

3.1 SIGNAL SEGMENTATION

Signal segmentation is achieved by finding the breakpoints. To identify breakpoints, it is necessary to solve the optimalization problem (2). For solving this problem proximal splitting algorithms can be used. Proximal splitting algorithm is an iterative way to minimize a sum of convex functions by repetitive evaluation of their gradients or proximal operators. The basic optimalization problem is

$$\arg\min_{\mathbf{x}} \; f_1(\mathbf{x}) + f_2(\mathbf{x}). \tag{4}$$

For solving this optimalization problem, two appropriate algorithms are available, namely the Forward-Backward and Douglas-Rachford algorithm [2]. Is proven that proximal algorithms converge to the optimal value which is minimum in case that (4) is convex. The convergence is in practice influenced by the character of functions and the choice of parameters of algorithm.

3.1.1 Used proximal splitting algorithms

We assign: $f_1(\mathbf{x}) = \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2$, $f_2(\mathbf{x}) = f_2(\mathbf{a}, \mathbf{b}) = \tau_{\mathbf{a}}\,\mathrm{TV}(\mathbf{a}) + \tau_{\mathbf{b}}\,\mathrm{TV}(\mathbf{b})$.

Forward-Backward (FB) splitting algorithm solves problem (4) where $f_1(\mathbf{x})$ is convex and differentiable with a β-Lipschitz continuous gradient $\nabla f_1(\mathbf{x}) = \mathbf{A}^\top(\mathbf{A}\mathbf{x} - \mathbf{y})$ [3]. Proximal operator of $f_2(\mathbf{x})$ is $\mathrm{prox}_{f_2}(\mathbf{x}) = \mathrm{prox}_{f_2}(\mathbf{a}, \mathbf{b}) = \begin{bmatrix} \mathrm{prox}_{\tau_{\mathbf{a}}\mathrm{TV}(\cdot)}(\mathbf{a}) \\ \mathrm{prox}_{\tau_{\mathbf{b}}\mathrm{TV}(\cdot)}(\mathbf{b}) \end{bmatrix}$. The proximal operator of $\mathrm{TV}(\cdot)$ for 1D signals can be computed fast using the Condat's algorithm [4]. The algorithm consists of a forward (gradient) step using function f_1 and a backward (proximal) step using function f_2.

Douglas-Rachford (DR) algorithm does not require function f_1 having β-Lipschitz continuous gradient. This algorithm consists of two proximal steps using functions f_1 and f_2. Instead of the gradient step, proximal operator of f_1 is used, which is $\mathrm{prox}_{f_1}(\mathbf{x}) = (\mathbf{I} + \tau\mathbf{A}^\top\mathbf{A})^{-1}(\mathbf{x} + \tau\mathbf{A}^\top\mathbf{y})$ [3].

3.1.2 Detection of breakpoints

From the solution $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$, the breakpoints can be established. Nonzero values in $\nabla\hat{\mathbf{a}}$ and $\nabla\hat{\mathbf{b}}$ indicate possible segment borders. In theory, the nonzero values should share the same position, but in practice they are not and there are found more segments than there should be. This happens because it is complicated to set regularization parameters to achieve piecewise constant solution in $\mathbf{a}$ and $\mathbf{b}$ so that in $\nabla\hat{\mathbf{a}}$ and $\nabla\hat{\mathbf{b}}$ are found more nonzero values. Because of this we used thresholding with threshold λ_a for $\nabla\hat{\mathbf{a}}$ and threshold λ_b for $\nabla\hat{\mathbf{b}}$ to find appropriated segment borders. After thresholding is created vector of breakpoints as:

$$\mathbf{bp} = [0, sort(\mathbf{bp}_a \cup \mathbf{bp}_b), N], \tag{5}$$

where $\mathbf{bp}_a$ and $\mathbf{bp}_b$ are positions of nonzero values in $\nabla\hat{\mathbf{a}}$ satisfying condition $|\hat{a}_i| > \lambda_a$ and in $\nabla\hat{\mathbf{b}}$ satisfying condition $|\hat{b}_i| > \lambda_b$, respectively. N is length of signal $\mathbf{y}$.

3.2 Signal denoising

In each detected segment we perform simple denoising of the signal by least squared method finding optimal parameters:

$$\beta_k = (\mathbf{X}_k^T \mathbf{X}_k)^{-1} \mathbf{X}_k^T \mathbf{y}(bp_k : bp_{k+1}), \tag{6}$$

where $\beta_k = \begin{bmatrix} \dot{a}_k & \dot{b}_k \end{bmatrix}^T$

is vector with parameter $\dot{a}$ and $\dot{b}$ (offset and slope) belonging to the k-th segment of signal $\mathbf{y}$, and $\mathbf{X}_k = \begin{bmatrix} 1 & 1 & \dots & 1 & 1 \\ bp_k & bp_k + 1 & \dots & bp_{k+1} - 1 & bp_{k+1} \end{bmatrix}^T$.

For each segment k we get parameter $\dot{a}_k$ and parameter $\dot{b}_k$. With obtained parameters we can reconstructed the signal $\mathbf{y}$, and observed the denoised signal $\hat{\mathbf{y}}$ according to:

$$\hat{\mathbf{y}} = \dot{\mathbf{a}} + \mathbf{D}\dot{\mathbf{b}}, \tag{7}$$

where $\dot{\mathbf{a}}$ and $\dot{\mathbf{b}}$ contain parameter $\dot{a}_k$ resp. parameter $\dot{b}_k$ for each element of k-th segment of $\mathbf{y}$.

4 EXPERIMENTAL RESULTS

Two experiments of signal denoising were performed. First experiment was performed with sawtooth signal and the second with the randomly generated signal.

4.1 Results for Sawtooth signal

A periodic sawtooth signal of length $N = 150$ with line slope equal to 1 was generated. Vector $\mathbf{b} \in \mathbb{R}^N$ is a vector of ones. Vector $\mathbf{a} \in \mathbb{R}^N$ is, according to the assumption, piecewise constant, and for each subsequent segment, it takes value lowered by $L = 25$, which is the selected period length. Computing $[\mathbf{I}\,\mathbf{D}][\mathbf{a}\,\mathbf{b}]^\top$ and adding Gaussian IID noise then synthesizes the noisy sawtooth.

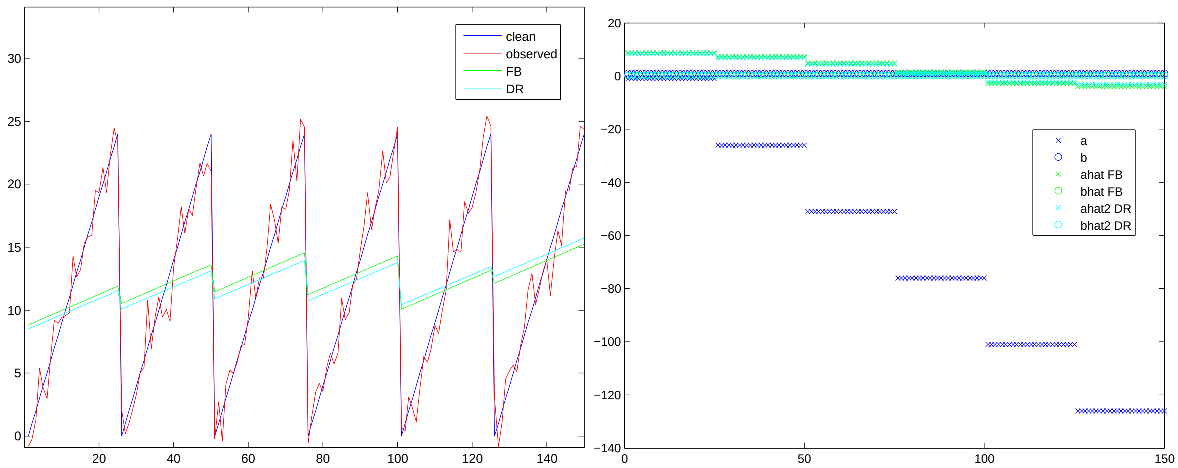

Figure 1: Approximating sawtooth signal after first step with $\tau_\mathbf{a} = 51.8$ and $\tau_\mathbf{b} = 12540$. We have $\mathrm{TV}(\mathbf{b}) = 0$, $\mathrm{TV}_{FB}(\hat{\mathbf{b}}) = 0$, $\mathrm{TV}_{CR}(\hat{\mathbf{b}}) = 0$, $\mathrm{TV}(\mathbf{a}) = 55.9$, $\mathrm{TV}_{FB}(\hat{\mathbf{a}}) = 6.29$, $\mathrm{TV}_{DR}(\hat{\mathbf{a}}) = 5{,}65$. Signal to noise ratio (SNR) of the observed signal is $SNR = 19.68dB$, recovered signal after first step has $SNR_{FB} = 6.83dB$, $SNR_{DR} = 6.77dB$.

Regularization parameters $\tau_\mathbf{a}, \tau_\mathbf{b}$ and threshold parameters λ_a and λ_b were carefully tuned to obtain stepwise parameter vector $\mathbf{a}$ producing signal that has breakpoints at the same positions as the clean one does. The resulting signal of first step and the parameters found are depicted in Fig. 1. Recovered sawtooth signal $\hat{\mathbf{y}}$ and its parameters are depicted in Fig. 2.

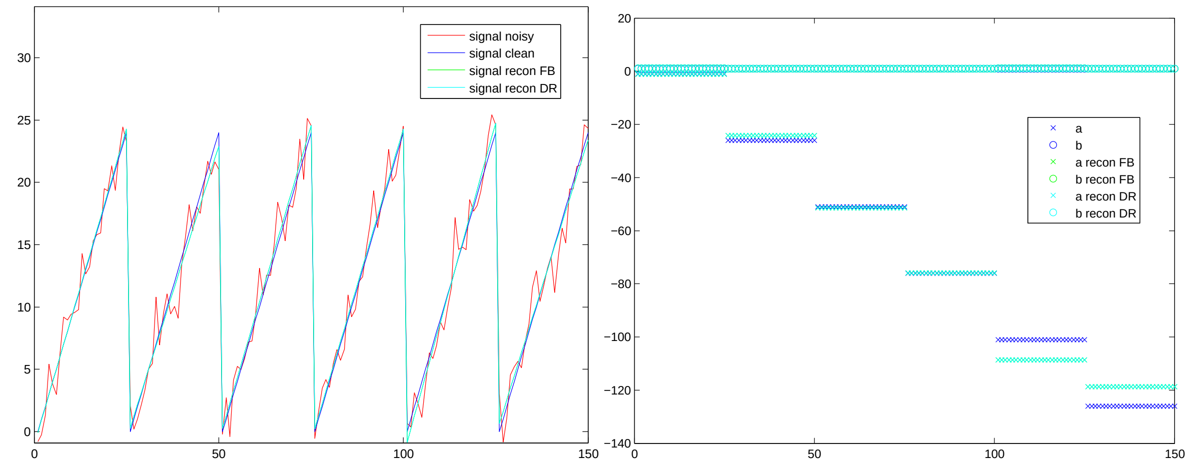

Figure 2: Approximating sawtooth signal after second step with $\lambda_{\mathbf{a}} = 0.075$ and $\lambda_{\mathbf{b}} = 0.075$. SNR is $SNR = 19.68dB$, of the recovered signal after second step $SNR_{FB} = 30.77dB$, $SNR_{DR} = 30.77dB$.

4.2 RESULTS FOR RANDOMLY GENERATED SIGNAL

Second, we performed a similar experiment on a randomly generated signal of length $N = 150$ with five linear segments. Vectors $\mathbf{a} \in \mathbb{R}^N$ and $\mathbf{b} \in \mathbb{R}^N$ are, according to the assumption, piecewise constant. Computing $[\mathbf{I}\,\mathbf{D}][\mathbf{a}\,\mathbf{b}]^\top$ and adding Gaussian IID noise then synthesizes the noisy signal. The resulting signal of breakpoints detection and the parameters found are depicted in Fig. 3. Recovered random signal $\hat{\mathbf{y}}$ and its parameters are depicted in Fig. 4.

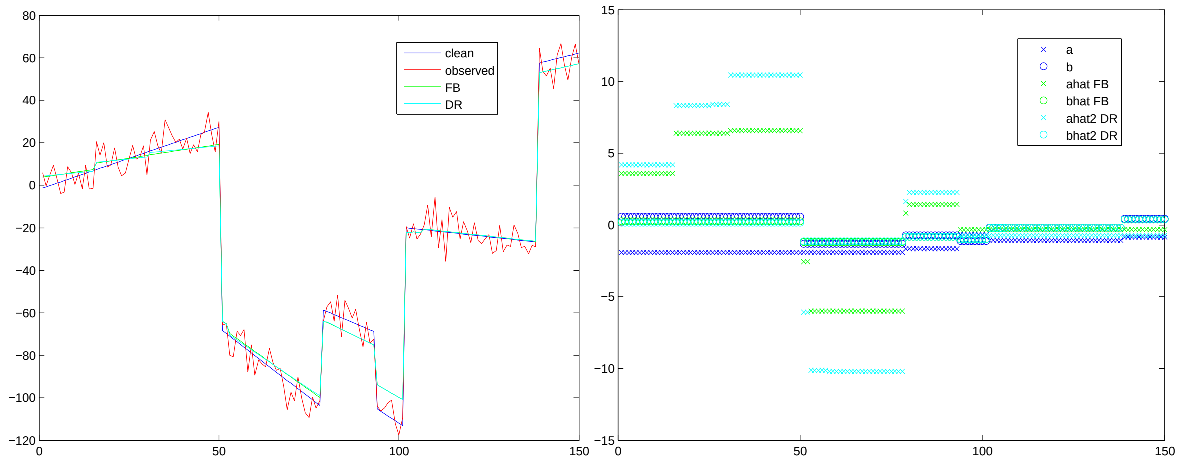

Figure 3: Approximating randomly generated signal after first step $\tau_{\mathbf{a}} = 43.27$ and $\tau_{\mathbf{b}} = 3850$. We have $\mathrm{TV}(\mathbf{b}) = 2.3$, $\mathrm{TV}_{FB}(\hat{\mathbf{b}}) = 1.8$, $\mathrm{TV}_{CR}(\hat{\mathbf{b}}) = 1.66$, $\mathrm{TV}(\mathbf{a}) = 1.1$, $\mathrm{TV}_{FB}(\hat{\mathbf{a}}) = 12.4$, $\mathrm{TV}_{DR}(\hat{\mathbf{a}}) = 21.45$. SNR of the observed signal is $SNR = 18.91dB$, recovered signal after first step has $SNR_{FB} = 21.86dB$, $SNR_{DR} = 21.8dB$.

5 CONCLUSION

The experiments present the fact that the suggested approach gives relatively good results. Detected positions of breakpoints in $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ are not the same although it is supposed that they should be on the same positions. Such a problem, however, can be solve by enforcing joint breakpoints using a group-sparse model, which can be provided by usage of a ℓ_{21} mixed norm in (2) instead of a used ℓ_1 norm. This will be the topic of further research.

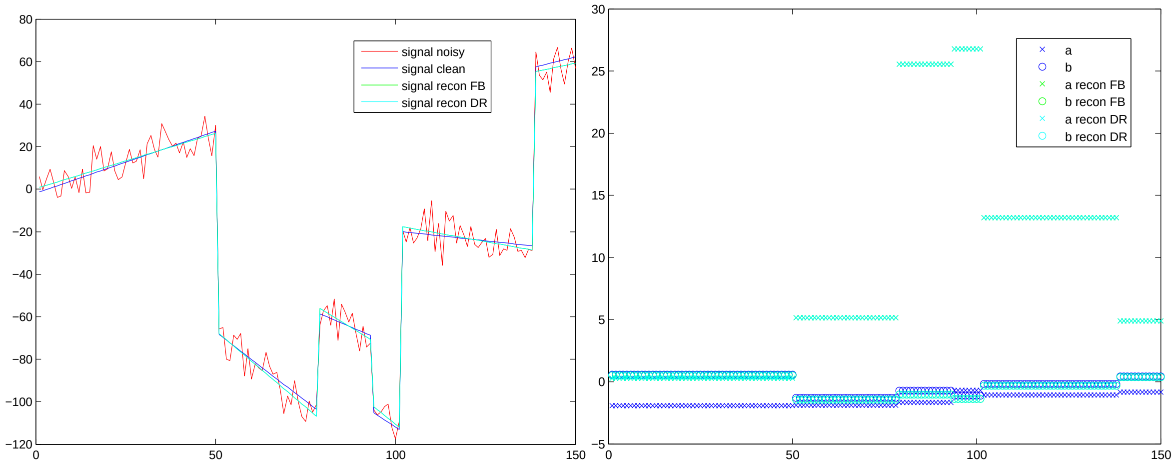

Figure 4: Approximating randomly generated signal after second step $\lambda_\mathbf{a} = 8$ and $\lambda_\mathbf{b} = 0.15$. SNR of the observed signal is $SNR = 18.91dB$, of the recovered signal after second step $SNR_{FB} = 31.03dB$, $SNR_{DR} = 31.03dB$.

REFERENCES

[1] Giryes, R.; Elad, M.; Bruckstein, A.M.: Sparsity Based Methods for Overparameterized Variational Problems, SIAM journal on imaging sciences, vol. 8, no. 3, pp. 2133 - 2159, 2015.

[2] Combettes, P.L.; Pesquet, J.C.: Proximal splitting methods in signal processing, Fixed-Point Algorithms for Inverse Problems in Science and Engineering, Springer, 2011, DOI: 10.1007/978-1-4419-9569-8_10

[3] Condat, L.: A Generic Proximal Algorithm for Convex Optimization—Application to Total Variation Minimization, Signal Processing Letters, IEEE, vol. 21, no. 8, pp. 985-989, 2014, DOI: 10.1109/LSP.2014.2322123

[4] Condat, L.: A Direct Algorithm for 1-D Total Variation Denoising, Signal Processing Letters, IEEE, vol. 20, no.11, pp. 1054-1057, 2013, DOI: 10.1109/LSP.2013.2278339