

ERRATUM: A Continuous Exact l0 penalty (CEL0) for least squares regularized problem

Emmanuel Soubies, Laure Blanc-Féraud, Gilles Aubert

► **To cite this version:**

Emmanuel Soubies, Laure Blanc-Féraud, Gilles Aubert. ERRATUM: A Continuous Exact l0 penalty (CEL0) for least squares regularized problem. *SIAM Journal on Imaging Sciences*, Society for Industrial and Applied Mathematics, 2016, 9 (1), pp.490-494. 10.1137/15M1038384 . hal-01267679

HAL Id: hal-01267679

<https://hal.inria.fr/hal-01267679>

Submitted on 4 Feb 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ERRATUM: A Continuous Exact ℓ_0 penalty (CEL0) for least squares regularized problem.

Emmanuel Soubies[†] Laure Blanc-Féraud[†] Gilles Aubert[‡]

February 4, 2016

Abstract

Lemma 4.4 in [E. Soubies, L. Blanc-Féraud and G. Aubert, *SIAM J. Imaging Sci.*, 8 (2015), pp. 1607-1639] is wrong for local minimizers of the CEL0 functional. The argument used to conclude the proof of this lemma is not sufficient in the case of local minimizers. In this note, we supply a revision of this Lemma where new results are established for local minimizers. Theorem 4.8 in that paper remains unchanged but the proof has to be rewritten according to the new version of the lemma. Finally, some remarks of this paper are also rewritten using the corrected lemma.

Key words inverse problems, ℓ_0 regularization, sparse modelling, underdetermined linear systems, global minimizers, local minimizers, minimizers equivalence, Continuous Exact ℓ_0 penalty, nonconvex nonsmooth penalty

Let $\mathcal{C} \subset \mathbb{R}^N$ be a closed subset of \mathbb{R}^N and $F : \mathbb{R}^N \rightarrow \mathbb{R}$. Assume that F is constant on \mathcal{C} and that $\text{int}(\mathcal{C}) = \emptyset$ (where int stands for interior). Then, the fact that \mathcal{C} contains a local (not global) minimizer \hat{x} of F does *not* imply that any $\bar{x} \in \mathcal{C}$ is a local minimizer of F , as it was argued in the proof of [2, Lemma 4.4] which is thus wrong for local (not global) minimizers¹. In such a case, only points $\bar{x} \in \mathcal{C} \cap \text{int}(\mathcal{V})$, where $\mathcal{V} \ni \hat{x}$ is the largest neighborhood such that $\forall x \in \mathcal{V} F(\hat{x}) \leq F(x)$, are local minimizers of F .

1 Corrections to the paper

To correct this problem, we propose in this note a new version of [2, Lemma 4.4] stated and proved in Section 2. Then, [2, Theorem 4.8] remains unchanged but its proof has to be rewritten which is done in Section 3. Finally, some remarks of the original paper are also updated. In the following, we use the notation of [2].

[†]Université de Nice Sophia Antipolis, CNRS, Laboratoire I3S UMR 7271, 06903 Sophia Antipolis, France (soubies@i3s.unice.fr, blancf@i3s.unice.fr)

[‡]Université de Nice Sophia Antipolis, CNRS, Laboratoire J.A. Dieudonné UMR 7351, 06100 Nice, France (gilles.aubert@unice.fr)

¹The claim remains true for global minimizers.

2 Correction of Lemma 4.4

First of all, in the proof of [2, Lemma 4.4], only the conclusion (for local minimizers) is wrong. The fact that

$$\forall i \in \sigma^+(\hat{x}), \forall t \in \left[0, \frac{\sqrt{2\lambda}}{\|a_i\|}\right], G_{\text{CELO}}(\hat{x}^{(i)} - s_i e_i t) = C \quad (2.1)$$

(where C is a constant of \mathbb{R}^+), is true and will be used to prove the new version of [2, Lemma 4.4] which reads as follows:

Lemma 2.1 (Lemma 4.4 of [2]). *For $d \in \mathbb{R}^M$ and $\lambda > 0$, let G_{CELO} have a minimum at $\hat{x} \in \mathbb{R}^N$. Define $\hat{\sigma}^+ := \sigma^+(\hat{x})$ and let $s_i = \text{sign}(\langle a_i, A\hat{x}^{(i)} - d \rangle)$. Then,*

$$(i) \quad \forall i \in \hat{\sigma}^+, \exists \mathcal{T}_i \subseteq \left[0, \frac{\sqrt{2\lambda}}{\|a_i\|}\right], \text{ a non-degenerate interval of } \mathbb{R}, \text{ s.t. } \hat{x}_i \in \mathcal{T}_i \text{ and } \forall t \in \mathcal{T}_i, \\ \bar{x} = \hat{x}^{(i)} - s_i e_i t \text{ is a minimizer of } G_{\text{CELO}}. \quad (2.2)$$

Furthermore $\forall i \in \hat{\sigma}^+, t \mapsto G_{\text{CELO}}(\hat{x}^{(i)} - s_i e_i t)$ is constant on $\left[0, \frac{\sqrt{2\lambda}}{\|a_i\|}\right]$.

(ii) if \hat{x} is global, then $\forall i \in \hat{\sigma}^+, \mathcal{T}_i = \left[0, \frac{\sqrt{2\lambda}}{\|a_i\|}\right]$ and \bar{x} is global.

Proof. (i) Since \hat{x} is a minimizer of G_{CELO} , there exists $\varepsilon > 0$ such that,

$$\forall x \in \mathcal{B}_2(\hat{x}, \varepsilon), G_{\text{CELO}}(\hat{x}) \leq G_{\text{CELO}}(x). \quad (2.3)$$

where $\mathcal{B}_2(\hat{x}, \varepsilon)$ denotes the open ℓ_2 -ball with center \hat{x} and radius ε . Therefore, from (2.1) and (2.3), we get the result in (2.2). Indeed, $\forall i \in \sigma^+(\hat{x})$ let

$$\mathcal{T}_i = \left\{ t \in \left[0, \frac{\sqrt{2\lambda}}{\|a_i\|}\right], x = \hat{x}^{(i)} - s_i e_i t \in \mathcal{B}_2(\hat{x}, \varepsilon) \right\}, \quad (2.4)$$

Clearly, since $\varepsilon > 0$, \mathcal{T}_i is a non-degenerate interval of \mathbb{R} . Then,

$$\forall t \in \mathcal{T}_i, \exists \varepsilon' \in (0, \varepsilon), \text{ s.t. } \mathcal{B}_2(\bar{x}, \varepsilon') \subset \mathcal{B}_2(\hat{x}, \varepsilon), \quad (2.5)$$

where $\bar{x} = \hat{x}^{(i)} - s_i e_i t$, and we get

$$\forall x \in \mathcal{B}_2(\bar{x}, \varepsilon'), G_{\text{CELO}}(\bar{x}) \stackrel{(2.1)}{=} G_{\text{CELO}}(\hat{x}) \stackrel{(2.3) \& (2.5)}{\leq} G_{\text{CELO}}(x), \quad (2.6)$$

which proves (2.2).

(ii) Using the fact that \hat{x} is a global minimizer of G_{CELO} , (2.1) completes the proof. \square

The comment after [2, Remark 4.2] about the strict minimizers of G_{CELO} is still valid and its justification has to be rewritten according to the results in Lemma 2.1:

Lemma 2.1 ensures that all strict minimizers of G_{CELO} — i.e. $\hat{x} \in \mathbb{R}^N$ such that there exists a neighborhood $\mathcal{V} \subset \mathbb{R}^N$ containing \hat{x} for which $\forall y \in \mathcal{V} \setminus \{\hat{x}\}, G_{\text{CELO}}(\hat{x}) < G_{\text{CELO}}(y)$ — verify $\sigma^+(\hat{x}) = \emptyset$. Indeed, suppose that this claim is not verified by a strict minimizer $\hat{x} = \hat{x}^{(i)} - s_i e_i \hat{t}$ of G_{CELO} , then Lemma 2.1 states that $\forall i \in \sigma^+(\hat{x}) \exists \mathcal{T}_i \subseteq [0, \sqrt{2\lambda}/\|a_i\|]$, a non-degenerate interval of \mathbb{R} containing \hat{t} , s.t. $\forall t \in \mathcal{T}_i \setminus \{\hat{t}\}, \bar{x} = \hat{x}^{(i)} - s_i e_i t$ is also a minimizer of G_{CELO} and contradicts the fact that \hat{x} is strict.

Then, the comment before [2, Theorem 4.5] concerning the nonstrict minimizers should be read only for nonstrict *global* minimizers of G_{CELO} .

3 Correction of the proof of Theorem 4.8

The claim of Theorem 4.8 is still valid but its proof has to be rewritten. In order to rewrite this proof, we need first to prove the following lemma:

Lemma 3.1. *Let $d \in \mathbb{R}^M$, $\lambda > 0$, and G_{CELO} have a minimum at $\hat{x} \in \mathbb{R}^N$. Then,*

$$\forall i \in \sigma^+(\hat{x}), \forall j \in \sigma(\hat{x}) \setminus \{i\}, \langle a_i, a_j \rangle = 0. \quad (3.1)$$

Proof. Let $\hat{\sigma} = \sigma(\hat{x})$, $\hat{\sigma}^- = \sigma^-(\hat{x})$ and $\hat{\sigma}^+ = \sigma^+(\hat{x})$. Let $i \in \hat{\sigma}^+$, \mathcal{T}_i be defined by Lemma 2.1, $\hat{x} = \hat{x}^{(i)} - s_i e_i \hat{t}$ and $t \in \mathcal{T}_i \setminus \{\hat{t}\}$ then, from (2.2), $\bar{x} = \hat{x}^{(i)} - s_i e_i t$ is a minimizer of G_{CELO} . Thus it verifies the conditions of [2, Lemma 4.1] (critical point). By construction,

$$\forall j \in \mathbb{I}_N \setminus \{i\}, \hat{x}_j = \bar{x}_j. \quad (3.2)$$

Then, the conditions of [2, Lemma 4.1] lead to

$$\forall j \in \hat{\sigma} \setminus \{i\}, \begin{cases} |\langle a_j, A\bar{x}^{(j)} - d \rangle| = \sqrt{2\lambda} \|a_j\| & \text{if } j \in \hat{\sigma}^-, \\ \bar{x}_j = -\frac{\langle a_j, A\bar{x}^{(j)} - d \rangle}{\|a_j\|^2} & \text{if } j \in \hat{\sigma} \setminus \hat{\sigma}^-, \end{cases} \quad (3.3)$$

From (3.2), $\forall j \in \mathbb{I}_N \setminus \{i\}$, $\bar{x}^{(j)} = \hat{x}^{(j)} - \hat{x}_i e_i + \bar{x}_i e_i$ and (3.3) can be rewritten as follows,

$$\forall j \in \hat{\sigma} \setminus \{i\}, \begin{cases} |\langle a_j, A\hat{x}^{(j)} - d \rangle + \langle a_j, a_i \rangle (\bar{x}_i - \hat{x}_i)| = \sqrt{2\lambda} \|a_j\| & \text{if } j \in \hat{\sigma}^-, \\ \bar{x}_j = -\frac{\langle a_j, A\hat{x}^{(j)} - d \rangle}{\|a_j\|^2} + \frac{\langle a_j, a_i \rangle}{\|a_j\|^2} (\hat{x}_i - \bar{x}_i) & \text{if } j \in \hat{\sigma} \setminus \hat{\sigma}^-. \end{cases} \quad (3.4)$$

Since \hat{x} is a critical point of G_{CELO} we get, $\forall j \in \hat{\sigma} \setminus \{i\}$

$$\begin{aligned} & \begin{cases} |s_j \sqrt{2\lambda} \|a_j\| + \langle a_j, a_i \rangle (\bar{x}_i - \hat{x}_i)| = \sqrt{2\lambda} \|a_j\| & \text{if } j \in \hat{\sigma}^-, \\ \bar{x}_j = \hat{x}_j + \frac{\langle a_j, a_i \rangle}{\|a_j\|^2} (\hat{x}_i - \bar{x}_i) & \text{if } j \in \hat{\sigma} \setminus \hat{\sigma}^-, \end{cases} \\ \xrightarrow{\hat{x}_i \neq \bar{x}_i \text{ \& } (3.2)} & \begin{cases} \langle a_j, a_i \rangle = 0 & \text{if } j \in \hat{\sigma}^-, \\ \langle a_j, a_i \rangle = 0 & \text{if } j \in \hat{\sigma} \setminus \hat{\sigma}^-, \end{cases} \end{aligned} \quad (3.5)$$

which, together with the fact that $\hat{\sigma}^- \subseteq \hat{\sigma}$, complete the proof. \square

We are now able to rewrite the proof of [2, Theorem 4.8]. Let us recall this theorem:

Theorem 3.2 (Theorem 4.8 of [2]). *Let $d \in \mathbb{R}^M$, $\lambda > 0$, and G_{CELO} have a local minimum (not global) at $\hat{x} \in \mathbb{R}^N$. Then \hat{x}^0 is a local minimizer (not global) of G_{ℓ_0} and $G_{\text{CELO}}(\hat{x}) = G_{\text{CELO}}(\hat{x}^0) = G_{\ell_0}(\hat{x}^0)$.*

Proof. Let $\hat{\sigma} = \sigma(\hat{x})$, $\hat{\sigma}^- = \sigma^-(\hat{x})$ and $\hat{\sigma}^0 = \sigma(\hat{x}^0)$. From Lemma 3.1, since \hat{x} is a minimizer of G_{CELO} , we have (3.1). Moreover, \hat{x} is a critical point of G_{CELO} and from [2, Lemma 4.1] we get

$$\begin{aligned} \forall i \in \hat{\sigma}^0 = \hat{\sigma} \setminus \hat{\sigma}^-, \hat{x}_i = \hat{x}_i^0 &= -\frac{\langle a_i, A\hat{x}^{(i)} - d \rangle}{\|a_i\|^2} \\ \iff \hat{x}_i^0 &= -\frac{\langle a_i, A_{\hat{\sigma}^0}(\hat{x}_{\hat{\sigma}^0}^0)^{(i)} + \sum_{j \in \hat{\sigma}^-} a_j \hat{x}_j - d \rangle}{\|a_i\|^2} \\ \stackrel{(3.1)}{\iff} \hat{x}_i^0 &= -\frac{\langle a_i, A_{\hat{\sigma}^0}(\hat{x}_{\hat{\sigma}^0}^0)^{(i)} - d \rangle}{\|a_i\|^2} - \frac{1}{\|a_i\|^2} \sum_{j \in \hat{\sigma}^-} \underbrace{\langle a_i, a_j \rangle}_{=0} \hat{x}_j \\ \iff \langle a_i, A_{\hat{\sigma}^0} \hat{x}_{\hat{\sigma}^0}^0 - d \rangle &= 0 \end{aligned}$$

Finally we have $(A_{\hat{\sigma}^0})^T A_{\hat{\sigma}^0} \hat{x}_{\hat{\sigma}^0}^0 = (A_{\hat{\sigma}^0})^T d$ which, with [1, Corollary 2.5], ensure that \hat{x}^0 is a local minimizer of G_{ℓ_0} . Then,

$$\begin{aligned} G_{\text{CELO}}(\hat{x}) &= \frac{1}{2} \|A_{(\hat{\sigma}^-)^c} \hat{x}_{(\hat{\sigma}^-)^c} + A_{\hat{\sigma}^-} \hat{x}_{\hat{\sigma}^-} - d\|^2 + \sum_{i \in \mathbb{I}_N} \phi(\|a_i\|, \lambda; \hat{x}_i) \\ &= \frac{1}{2} \|A_{(\hat{\sigma}^-)^c} \hat{x}_{(\hat{\sigma}^-)^c} - d\|^2 + \langle A_{\hat{\sigma}^-} \hat{x}_{\hat{\sigma}^-}, A_{(\hat{\sigma}^-)^c} \hat{x}_{(\hat{\sigma}^-)^c} - d \rangle + \frac{1}{2} \|A_{\hat{\sigma}^-} \hat{x}_{\hat{\sigma}^-}\|^2 + \sum_{i \in \mathbb{I}_N} \phi(\|a_i\|, \lambda; \hat{x}_i) \\ &= G_{\text{CELO}}(\hat{x}^0) + \sum_{i \in \hat{\sigma}^-} \left(\phi(\|a_i\|, \lambda; \hat{x}_i) + \hat{x}_i \langle a_i, A_{(\hat{\sigma}^-)^c} \hat{x}_{(\hat{\sigma}^-)^c} - d \rangle + \frac{\hat{x}_i}{2} \langle a_i, A_{\hat{\sigma}^-} \hat{x}_{\hat{\sigma}^-} \rangle \right) \\ &\stackrel{(3.1)}{=} G_{\text{CELO}}(\hat{x}^0) + \sum_{i \in \hat{\sigma}^-} \left(\phi(\|a_i\|, \lambda; \hat{x}_i) + \hat{x}_i \langle a_i, A \hat{x}^{(i)} - d \rangle + \frac{\|a_i\|^2}{2} \hat{x}_i^2 \right), \end{aligned}$$

and, since \hat{x} is a critical point of G_{CELO} , we have

$$G_{\text{CELO}}(\hat{x}) = G_{\text{CELO}}(\hat{x}^0) + \sum_{i \in \hat{\sigma}^-} \left(\phi(\|a_i\|, \lambda; \hat{x}_i) - \text{sign}(\hat{x}_i) \sqrt{2\lambda} \|a_i\| \hat{x}_i + \frac{\|a_i\|^2}{2} \hat{x}_i^2 \right) = G_{\text{CELO}}(\hat{x}^0). \quad (3.6)$$

Here we used the fact that from the definition of ϕ in [2, eq. (2.9)] one has

$$\forall i \in \hat{\sigma}^-, \phi(\|a_i\|, \lambda; \hat{x}_i) - \text{sign}(\hat{x}_i) \sqrt{2\lambda} \|a_i\| \hat{x}_i + \frac{\|a_i\|^2}{2} \hat{x}_i^2 = 0. \quad (3.7)$$

Finally, $G_{\text{CELO}}(\hat{x}^0) = G_{\ell_0}(\hat{x}^0)$ comes from the same arguments as in the proof of [2, Theorem 4.5] and completes the proof. \square

In [2, Remark 4.4], the claim is based on the argument that when \hat{x} is a local (not global) minimizer of G_{CELO} , \hat{x}^0 is also a local (not global) minimizer of G_{CELO} (and thus a critical point). However we have seen that this argument is wrong. From this erratum we see that \hat{x}^0 is a local (not global) minimizer of G_{ℓ_0} (Theorem 3.2) but it may not be a critical point of G_{CELO} . Therefore, an additional assumption is required in [2, Remark 4.4] that is “ \hat{x}^0 is a critical point of G_{CELO} ” which can be reduced to

$$\forall i \notin \sigma(\hat{x}^0), |\langle a_i, A \hat{x}^0 - d \rangle| \leq \sqrt{2\lambda} \|a_i\|, \quad (3.8)$$

since other conditions of [2, Lemma 4.1] are verified, using Lemma 3.1, for such an \hat{x}^0 .

Finally, Remark 4.5 in [2] reads as follows:

Remark 3.1. For a (local) minimizer \hat{x} of G_{CELO} , one can set $\forall i \in \hat{\sigma}^- = \sigma^-(\hat{x})$ either $\hat{x}_i = 0$ or $\hat{x}_i = -s_i \frac{\sqrt{2\lambda}}{\|a_i\|}$ to obtain a (local) minimizer, denoted \tilde{x} , of G_{ℓ_0} . Indeed, let $\{\omega_-, \omega_+\}$ be a partition of $\hat{\sigma}^-$ (i.e. $\omega_- \subseteq \hat{\sigma}^-, \omega_+ \subseteq \hat{\sigma}^-$ such that $\omega_- \cup \omega_+ = \hat{\sigma}^-$ and $\omega_- \cap \omega_+ = \emptyset$) and

$$\forall i \in \mathbb{I}_N, \tilde{x}_i = \begin{cases} \hat{x}_i & \text{if } i \notin (\omega_- \cup \omega_+), \\ 0 & \text{if } i \in \omega_-, \\ -s_i \frac{\sqrt{2\lambda}}{\|a_i\|} & \text{if } i \in \omega_+. \end{cases} \quad (3.9)$$

where $s_i = \text{sign}(\langle a_i, A \hat{x}^{(i)} - d \rangle)$. Since \hat{x} is a critical point of G_{CELO} , [2, Lemma 4.1] lead to

$$\begin{aligned} \forall i \in \sigma(\tilde{x}), \tilde{x}_i &= -\frac{\langle a_i, A \tilde{x}^{(i)} - d \rangle}{\|a_i\|^2} \\ \stackrel{(3.1) \& (3.9)}{\iff} \tilde{x}_i &= -\frac{\langle a_i, A \tilde{x}^{(i)} - d \rangle}{\|a_i\|^2} - \frac{1}{\|a_i\|^2} \sum_{j \in \omega_-} \underbrace{\langle a_i, a_j \rangle}_{=0} \tilde{x}_j - \frac{1}{\|a_i\|^2} \sum_{\substack{j \in \omega_+ \\ j \neq i}} \underbrace{\langle a_i, a_j \rangle}_{=0} \left(\tilde{x}_j + s_j \frac{\sqrt{2\lambda}}{\|a_j\|} \right) \\ \iff & \langle a_i, A_{\sigma(\tilde{x})} \tilde{x}_{\sigma(\tilde{x})} - d \rangle = 0 \end{aligned}$$

which, with [1, Corollary 2.5], ensure that \tilde{x} is a local minimizer of G_{ℓ_0} . There exists $2^{\#\tilde{\sigma}^-}$ of such minimizers. Among them, \hat{x}^0 is the sparsest. Note that this remark can be extended to points \tilde{x} defined by (3.9) with $\{\omega_-, \omega_+\}$ a partition of $\sigma^+(\hat{x})$ such that $\forall (i, j) \in (\omega_+ \setminus \sigma(\hat{x}))^2$, $\langle a_i, a_j \rangle = 0$.

Remark 3.2. As outlined in [2], some local (not global) minimizers of G_{ℓ_0} are not critical points of G_{CELO} and from each local (not global) minimizer of G_{CELO} , we can easily extract a local (not global) minimizer, \hat{x}^0 , of G_{ℓ_0} (Theorem 3.2). However, we are not ensured that this extracted minimizer is a critical point of G_{CELO} . Therefore, when using Theorem 3.2 in practice, it is important to verify if \hat{x}^0 is a critical point of G_{CELO} which is a necessary condition to be global for G_{ℓ_0} . Note that when using the macro algorithm [2, Algorithm 1], the convergence point is ensured to be both a critical point of G_{CELO} and a (local) minimizer of G_{ℓ_0} .

Acknowledgment The authors would like to thank Charles Soussen for having pointed out the technical error in [2, Lemma 4.4].

References

- [1] MILA NIKOLOVA, *Description of the minimizers of least squares regularized with ℓ_0 -norm. Uniqueness of the global minimizer*, SIAM Journal on Imaging Sciences, 6 (2013), pp. 904–937.
- [2] EMMANUEL SOUBIES, LAURE BLANC-FÉRAUD, AND GILLES AUBERT, *A Continuous Exact ℓ_0 Penalty (CEL0) for Least Squares Regularized Problem*, SIAM Journal on Imaging Sciences, 8 (2015), pp. 1607–1639.