Minimizers of Sparsity Regularized Huber Loss Function

Deniz Akkaya & Mustafa Ç. Pınar

Journal of Optimization Theory and Applications

ISSN 0022-3239
Volume 187
Number 1

J Optim Theory Appl (2020) 187:205-233
DOI 10.1007/s10957-020-01745-3

Springer

Journal of Optimization Theory and Applications (2020) 187:205–233
https://doi.org/10.1007/s10957-020-01745-3

Minimizers of Sparsity Regularized Huber Loss Function

Deniz Akkaya[1] · **Mustafa Ç. Pınar**[1]

Received: 27 January 2020 / Accepted: 2 September 2020 / Published online: 19 September 2020

Abstract
We investigate the structure of the local and global minimizers of the Huber loss function regularized with a sparsity inducing L0 norm term. We characterize local minimizers and establish conditions that are necessary and sufficient for a local minimizer to be strict. A necessary condition is established for global minimizers, as well as non-emptiness of the set of global minimizers. The sparsity of minimizers is also studied by giving bounds on a regularization parameter controlling sparsity. Results are illustrated in numerical examples.

Keywords Sparse solution of linear systems · Regularization · local minimizer · Global minimizer · Huber loss function · L0-norm

Mathematics Subject Classification 15A29 · 62J05 · 90C26 · 90C46

1 Introduction

The search for an approximate and regularized solution (i.e., a solution with some desirable properties such as sparsity) to a possibly inconsistent system of linear equations is a ubiquitous problem in applied mathematics. It aims at finding a solution vector x, that minimizes both objectives ($\|Ax - b\|, \|x\|$) with respect to a norm or another appropriate error measure. Here, we use the Huber loss function and the ℓ_0-norm, respectively.

The Huber loss function has been used in several engineering applications since the 1970s (for a recent account, see, e.g., [1]) while the search for sparse solutions of

Communicated by Panos M. Pardalos.

✉ Mustafa Ç. Pınar
mustafap@bilkent.edu.tr

Deniz Akkaya
deniz.akkaya@bilkent.edu.tr

1 Bilkent University, Ankara, Turkey

linear systems has been a popular topic of the last decade. The purpose of the present paper is to initiate an investigation of a combination of the two.

In the world of statistical data analysis, a common assumption is the normality of errors in the measurements. The normality assumption leads to methods yielding closed form solutions and thus is quite convenient. However, many real-world applications in engineering present the modeler with data deviating from the normality assumption. Robust statistics or robust methods in engineering aim at alleviating the effects of departure from normality by being largely immune to its negative ones. One of the proposals for robust statistical procedures was put forward in the 1970s and 1980s by Huber [2]. The so-called Huber loss function (a.k.a. Huber's M-estimator) coincides with the quadratic error measure up to a range beyond which a linear error measure is adopted. Huber established that the resulting estimator corresponds to a maximum likelihood estimate for a perturbed normal law. The Huber loss function has numerous applications in statistics and engineering as documented among others in the recent monograph by Zoubir et al. [1]. The present paper studies a case where sparsity is incorporated directly (instead of an approximation, e.g., using an ℓ_1 term, of the sparsity inducing ℓ_0-norm term) into an estimation effort based on the Huber loss function. Sparsity in estimation of linear models has become popular due to the presence of high-dimensional measurement and prediction spaces. The popularity of sparse estimation is largely due to the success of the least absolute shrinkage and selection operator (Lasso) by Hastie and Tibshirani [3], which had a vast influence on engineering and sciences. This development was further enhanced by the advent of techniques such as compressed sensing (e.g., [4–14]), where the predictors or the signal vector were assumed to be sparse (i.e., having a few nonzero or large components). Indeed, in many applications the measured signal can have a sparse representation with respect to a suitable basis (e.g., a wavelet basis) [8,15].

In all the aforementioned studies, the error criterion is the least squares criterion (for applications in engineering see, e.g., the references in [16]), and the sparsity of the solution is usually controlled using the ℓ_1 norm as a proxy for the number of nonzero components of a vector since the latter leads to non-convex, non-differentiable optimization problems. The pertinent research question is then to find necessary and sufficient conditions under which the solutions obtained with the ℓ_1 approximation correspond to the sought-after sparse vector (i.e., the one that would result from the use of the true measure counting the nonzero elements). There are few papers that study the structure of optimal solutions to the ℓ_0-regularized problem (namely, the $\ell_2 - \ell_0$ problem), e.g., [16–18]. In [17], necessary conditions for optimality are developed along with algorithms for sparsity constrained optimization problems with a continuously differentiable objective function. The reference [16] studies the structure of local and global solutions to the least-squares problem regularized with the ℓ_0 term. In [18], a stationarity-based optimality condition is given in the appendix. Both references [16,17] contain extensive lists of relevant research articles from the last decade on the subject of sparse optimization. It is also noteworthy that in recent (yet unpublished as of this writing) research Chancelier and De Lara [19–21] investigate conjugacy and duality for optimization problems involving ℓ_0 terms. Here, we follow the footsteps of [16] as we characterize local minimizers, strict local minimizers and we investigate properties of global minimizers. A more recent paper following the line of investigation in [16,17] is the reference [22] which gives a review of necessary conditions for the $\ell_2 - \ell_0$ problem as well as their applications in numerical algorithms.

However, the analysis of the present paper is significantly more complicated than that of the least-squares case treated in [16] due to the piecewise quadratic nature of the Huber loss function. While the minimizers of the problem Huber loss-ℓ_0 has not been studied previously, to the best of our knowledge, the connection of Huber loss to sparsity was also investigated in a recent line of work by Selesnick and others in a series of papers, see, e.g., [23–26]. In these papers, the Huber loss function and its generalization are used as the basis of a minimax-concave penalty in the context of regularized least squares for sparse recovery in engineering applications. Additionally, in [27] a unified model for robust regularized Extreme Learning Machine regression using iteratively reweighted least squares which employs the Huber loss function for robustness among other criteria is given as well as a comprehensive study on the robust loss function and regularization term for robust ELM regression. The reference [28] gives an overview of robust nonlinear reconstruction strategies for sparse signals based on replacing the commonly used L2 norm by M-estimators as data fidelity functions. A recent reference [29] studies the relationship between the class of M-estimators and the probability density functions of residuals in regression analysis using maximum likelihood estimation and entropy maximization. An interesting line of future work would be to investigate the impact of ℓ_0 norm on this relationship.

The contributions of the present paper are the following:

- We show how to find local minimizers in Sect. 4 after some preliminaries in Sects. 2 and 3.
- We develop necessary and sufficient conditions for strict local minimizers in Sect. 5. The conditions given are verifiable numerically.
- We give a necessary condition for global minimizers, which is useful in setting meaningful values of the scalarization parameter in Sect. 6.
- We prove the non-emptiness property of the set of global minimizers and discuss the choice of a the regularization parameter for controlling sparsity of global minimizers.

We also relate our results to the optimality criteria (support optimality, L-stationarity and partial coordinate-wise optimality) of [17]. It is the hope of the authors that the results presented here will be useful for further studies on development of algorithms.

2 Preliminaries

Let $A \in \mathbb{R}^{M \times N}$ for $M < N$, where the positive integers M and N are fixed, with $\mathrm{rank}(A) = M$. Given a data vector $d \in \mathbb{R}^M$ and $\beta > 0$, we consider an objective function $\mathcal{F}_d : \mathbb{R}^N \to \mathbb{R}$ of the form

$$\mathcal{F}_d(u) = \Psi(u) + \beta \|u\|_0 , \tag{1}$$

where $\Psi(u)$ stands for a variant of the Huber-γ function which is

$$\Psi(u) = \sum_{i=1}^{M} \psi(c_i(u)), \text{ where } \psi(c_i(u)) = \begin{cases} \dfrac{c_i(u)^2}{2\gamma}, & \text{for } |c_i(u)| < \gamma, \\ |c_i(u)| - \frac{1}{2}\gamma, & \text{for } |c_i(u)| \geq \gamma, \end{cases} \tag{2}$$

and $c_i(u) = \langle Au - d, e_i \rangle$. Finally, $\|u\|_0 = |\sigma(u)|$, where $\sigma(u)$ is the support of u. The Huber-γ function is C^1, with Lipschitz continuous derivative (and gradient) and convex. For the Lipschitz continuity property of the gradient, a derivation of the Lipschitz constant is given in "Appendix" for completeness.[1] The C^1 property is trivial, and convexity comes from the fact that c_i's are affine functions, and non-negative weighted sums imply that $\Psi(u)$ is convex. Unfortunately, $\Psi(u)$ is not a coercive function. Let $u \in \ker(A)$ then for any $\delta \in \mathbb{R}$, we have $\delta u \in \ker(A)$. Then, $\Psi(\delta u) = \Psi(u) = \sum_{i=1}^{M} \psi(d_i)$, this shows that there exists a direction where the function does not go to infinity.

If we define $\phi : \mathbb{R} \to \{0, 1\}$ as

$$\phi(t) = \begin{cases} 0, & \text{if } t = 0, \\ 1, & \text{if } t \neq 0, \end{cases} \tag{3}$$

then we can rewrite $\|u\|_0 = \sum_{i=1}^{N} \phi(u[i]) = \sum_{i \in \sigma(u)} \phi(u[i])$. Then, $\mathcal{F}_d$ is equivalently:

$$\mathcal{F}_d(u) = \Psi(u) + \beta \sum_{i \in \sigma(u)} \phi(u[i]). \tag{4}$$

We are looking for all (local and global) minimizers $\hat{u}$ of an objective $\mathcal{F}_d$ of the form (1):

$$\hat{u} \in \mathbb{R}^N \text{ such that } \mathcal{F}_d(\hat{u}) = \min_{u \in \mathcal{O}} \mathcal{F}_d(u). \tag{5}$$

Instead of the minimization of $\mathcal{F}_d$ in (1) one can also study its constrained variants:

$$\begin{cases} \text{given } \varepsilon \geq 0, & \text{minimize } \|u\|_0 \text{ subject to } \Psi(u) \leq \varepsilon, \\ \text{given } K \in \mathbb{I}_M, & \text{minimize } \Psi(u) \text{ subject to } \|u\|_0 \leq K, \end{cases} \tag{6}$$

where $\mathbb{I}_M$ stands for the *totally* and *strictly ordered* index set, $\mathbb{I}_M := (\{1, \ldots, M\}, <)$, although, due to non-convexity, one cannot in general speak of equivalence between these problems. Here, we focus exclusively on the minimizers of $\mathcal{F}_d$ in the spirit of [16].

Given $u \in \mathbb{R}^N$ and $\rho > 0$, the *open ball* at u of radius ρ with respect to the ℓ_p-norm for $1 \leq p \leq \infty$ is defined as $B_p(u, \rho) := \{v \in \mathbb{R}^N : \|v - u\|_p < \rho\}$. The notation $\|\cdot\|$

[1] While everyone dealing with the Huber function uses the constant, we were not able to find a derivation, so it is provided.

is used to mean the standard ℓ_2 norm when not explicitly indicated. The ith column of a matrix is denoted by a_i. Without loss of generality we assume that $a_i \neq 0$ for any $i \in \mathbb{I}_N$. For any $\omega \subseteq \mathbb{I}_N$, we use the following notation for submatrices and subvectors

$$A_\omega := (a_{\omega[1]}, \ldots, a_{\omega[|\omega|]}) \in \mathbb{R}^{M \times |\omega|},$$
$$u_\omega := (u[\omega[1]], \ldots, u[\omega[|\omega|]]) \in \mathbb{R}^{|\omega|}.$$

The *zero padding operator* $Z_\omega : \mathbb{R}^{|\omega|} \to \mathbb{R}^N$ is used for inversion:

$$u = Z_\omega(u_\omega), \quad u[i] = \begin{cases} 0, & \text{if } i \notin \omega, \\ u_\omega[k], & \text{for the unique } k \text{ such that } \omega[k] = i. \end{cases}$$

Remark 2.1 For any $u \in \mathbb{R}^N$ and $\omega \subseteq \mathbb{I}_N$ with $\omega \supseteq \sigma(u)$, we have $Au = A_\omega u_\omega$.

We conclude this section with a numerical example to motivate the research effort of the present paper.

Example 2.1 Let $p(x) = 8x^{13} - 2x^{11} - 4x^7 + 5x^6 + 3x^3 - x^2 + 1$ be a polynomial. We have t defined as ten test points randomly chosen between $[-1, 1]$. Then, let $A \in \mathbb{R}^{10\times 20}$ be the Vandermonde matrix generated from t and $d = p(t)$. Assuming a large noise $\eta = \left[\, 100\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0 \,\right]^T$, we perturb the data as $\tilde{d} = d + \eta$. We are looking for an approximation of p with a small number of terms using $\gamma = 0.1$ and $\beta = 0.6$. From Fig. 1, we observe that the Huber loss approximates the polynomial better than the least squares criterion under large amount of noise, especially in the interval [0.3, 0.6]. Beyond this interval, the Huber approximation and the polynomial are almost indistinguishable.

3 Minimizers of (HR$_\omega$)

Now, let $\omega \subseteq \mathbb{I}_N$, problem (HR$_\omega$) reads as

$$\min_{u \in \mathbb{R}^N} \Psi(u) \qquad \text{subject to } u[i] = 0,\ \forall i \in \omega^c. \tag{HR$_\omega$}$$

Let K_ω denote the subspace $K_\omega := \{v \in \mathbb{R}^N : v[i] = 0, \forall i \in \omega^c\}$. Then, one can rewrite the problem (HR$_\omega$) using this subspace:

$$\min_{u \in K_\omega} \Psi(u). \tag{HR$_\omega$}$$

This problem is equivalent to (ZPHR$_\omega$) with the help of the zero-padding operator:

$$\min_{v \in \mathbb{R}^{|\omega|}} \sum_{i=1}^{M} \psi(\langle A_\omega v - d, e_i \rangle), \qquad |\omega| \geq 1. \tag{ZPHR$_\omega$}$$

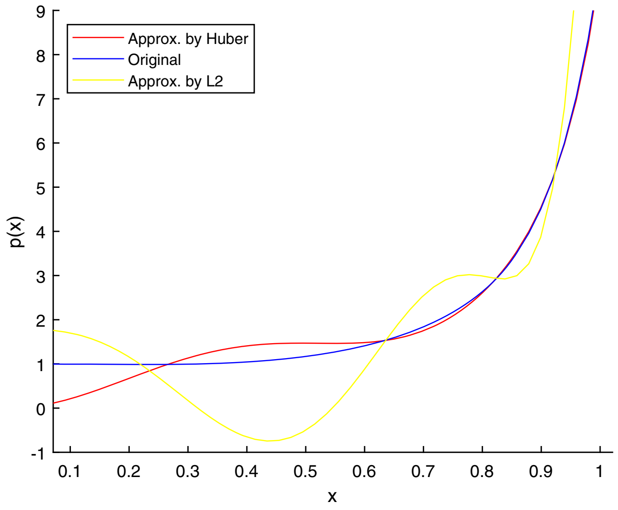

Fig. 1 Comparison of L2 and Huber for approximation of a 13-degree polynomial under large perturbations

Proposition 3.1 *The optimality condition of* ($ZPHR_\omega$) *is* $A_\omega^T clip(A_\omega v - d) = 0 \in \mathbb{R}^{|\omega|}$ *where the clip function is defined as*

$$clip(A_\omega v - d)[i] = \begin{cases} \dfrac{\langle A_\omega v - d, e_i \rangle}{\gamma}, & if\ |\langle A_\omega v - d, e_i\rangle| < \gamma, \\ sgn(\langle A_\omega v - d, e_i\rangle), & if\ |\langle A_\omega v - d, e_i\rangle| \geq \gamma. \end{cases} \tag{7}$$

Proof If we let $\Psi_\omega(v) = \sum_{i=1}^{M} \psi(\langle A_\omega v - d, e_i\rangle)$, we have to check $\nabla \Psi_\omega(v)$. For all $j \in \mathbb{I}_{|\omega|}$:

$$\begin{aligned} \frac{\partial \Psi_\omega(v)}{\partial v_j} = \sum_{i=1}^{M} \frac{\partial \psi(\langle A_\omega v - d, e_i\rangle)}{\partial v_j} &= \sum_{i=1}^{M} \text{clip}(\langle A_\omega v - d, e_i\rangle) \frac{\partial \langle v, A_\omega^T e_i\rangle}{\partial v_j} \\ &= \sum_{i=1}^{M} \text{clip}(\langle A_\omega v - d, e_i\rangle)(A_\omega^T)_{ji} \\ &= \sum_{i=1}^{M} \text{clip}(\langle A_\omega v - d, e_i\rangle)(A_\omega)_{ij} \\ &= \langle a_{\omega j}, \text{clip}(A_\omega v - d)\rangle. \end{aligned}$$

This shows that $A_\omega^T \text{clip}(A_\omega v - d) = 0_{|\omega|}$ is the optimality condition. □

3.1 Existence of Global Minimizers of (HR_ω)

While the problem (HR_ω) has been extensively studied since the 1980s (e.g., [30,31]), an existence result for the minimizer is difficult to find in the literature. For the sake of completeness, it is provided below. We start by defining a "sign vector" s_γ and its associated diagonal matrix W_s to rewrite (HR_ω) as a quadratic optimization problem. Let $s_\gamma(v) = [s_{\gamma 1}(v), \ldots, s_{\gamma m}(v)]$ with

$$s_{\gamma i}(v) = \begin{cases} -1, & \text{if } c_i^\omega(v) \le -\gamma, \\ 0, & \text{if } \left|c_i^\omega(v)\right| < \gamma, \\ 1, & \text{if } c_i^\omega(v) \ge \gamma, \end{cases}$$

where $c_i^\omega(v) = \langle A_\omega v - d, e_i \rangle$. Also let $W_s = \mathrm{diag}(w_1, \ldots, w_m)$ where $w_i = 1 - s_i^2$. Then, our problem (ZPHR_ω) is rewritten as

$$\min \Psi_\omega(v) = \frac{1}{2\gamma}(c^\omega)^T W_s c^\omega + s_\gamma^T \left[c^\omega - \frac{1}{2}\gamma s_\gamma \right] \text{ subject to } |\omega| \ge 1.$$

The dual problem turns out to be

$$\min_\nu \left\{ \nu^T d + \frac{\gamma}{2} \nu^T \nu \right\} \text{ subject to } \|\nu\|_\infty \le 1 \text{ and } A_\omega^T \nu = 0. \tag{$-\mathcal{D}$}$$

Since $A_\omega : \mathbb{R}^{|\omega|} \to \mathbb{R}^M$ is a continuous linear operator, $\ker(A_\omega)$ is a closed subspace of $\mathbb{R}^{|\omega|}$. Thus, the dual problem has a non-empty (zero vector) compact feasible region with a continuous objective function. Therefore, the dual problem always admits a solution. Now, computing the dual of $(-\mathcal{D})$ one immediately obtains that the dual of the $(\mathcal{D})$ is (HR). Since $(\mathcal{D})$ has a non-empty bounded feasibility set, this shows that (HR) has an optimal solution, independently from the choice of $\omega \subseteq \mathbb{I}_N$. Therefore, $A_\omega^T \mathrm{clip}(A_\omega v - d) = 0$ always admits a solution.

Proposition 3.2

$$\left[\hat{u} \in \mathbb{R}^N \textit{ solves } (HR_\omega) \right] \Leftrightarrow \left[\hat{u}_\omega \in \mathbb{R}^{|\omega|} \textit{ satisfies } A_\omega^T \mathit{clip}(A_\omega \hat{u}_\omega - d) = 0 \textit{ and } \hat{u} = Z_\omega(\hat{u}_\omega) \right], \tag{8}$$

where Z_ω is the zero padding operator.

Proof Let $|\omega| = 0$, then the result follows immediately. If $|\omega| \ge 1$, then one can use that $A\hat{u} = A_\omega \hat{u}_\omega$ for $\sigma(\hat{u}) = \omega$. Since Z_ω acts as a bijection between $\mathbb{R}^{|\omega|}$ and $\mathbb{R}^N$, we have the equivalence. □

4 Local Minimizers

We begin with a useful technical lemma which is a counterpart to Lemma 2.1 of [16].

Lemma 4.1 *Let $d \in \mathbb{R}^M$, $\beta > 0$, and $\hat{u} \in \mathbb{R}^N \setminus \{0\}$ be arbitrary. For $\hat{\sigma} := \sigma(\hat{u})$, set*

$$\rho := \min\left\{ \min_{i \in \hat{\sigma}} \left|\hat{u}[i]\right|, \frac{\beta}{\|A\|_{1,1} + 1} \right\}.$$

Then, $\rho > 0$, and we have

$$v \in B_\infty(0, \rho) \Rightarrow \sum_{i \in \mathbb{I}_N} \phi(\hat{u}[i] + v[i]) = \sum_{i \in \hat{\sigma}} \phi(\hat{u}[i]) + \sum_{i \in \hat{\sigma}^c} \phi(v[i]). \tag{i}$$

$$v \in B_\infty(0, \rho) \cap (\mathbb{R}^N \setminus K_{\hat{\sigma}}) \Rightarrow \mathcal{F}_d(\hat{u} + v) \geq \mathcal{F}_d(\hat{u}), \tag{ii}$$

where the inequality is strict whenever $\hat{\sigma}^c \neq \emptyset$.

Proof For (i), $v \in B_\infty(0, \rho)$ implies that $\max_{i \in \mathbb{I}_N}\{|v[i]|\} \leq \min_{i \in \mathbb{I}_N}\{\left|\hat{u}[i]\right|\}$, then $\operatorname{sgn}(\hat{u}[i] + v[i]) = \operatorname{sgn}(\hat{u}[i])$ for $i \in \hat{\sigma}$. Then, we have

$$\sum_{i \in \mathbb{I}_N} \phi(\hat{u}[i] + v[i]) = \sum_{i \in \hat{\sigma}} \phi(\hat{u}[i] + v[i]) + \sum_{i \in \hat{\sigma}^c} \phi(v[i]) = \sum_{i \in \hat{\sigma}} \phi(\hat{u}[i]) + \sum_{i \in \hat{\sigma}^c} \phi(v[i]),$$

as expected.

For (ii), we have to define a four set partition of $\mathbb{I}_M$:

$$\begin{aligned}
H_1 &:= \{i \in \mathbb{I}_M : \left|c_i(\hat{u} + v)\right| < \gamma\} \cap \{i \in \mathbb{I}_M : \left|c_i(\hat{u})\right| < \gamma\},\\
H_2 &:= \{i \in \mathbb{I}_M : \left|c_i(\hat{u} + v)\right| < \gamma\} \cap \{i \in \mathbb{I}_M : \left|c_i(\hat{u})\right| \geq \gamma\},\\
H_3 &:= \{i \in \mathbb{I}_M : \left|c_i(\hat{u} + v)\right| \geq \gamma\} \cap \{i \in \mathbb{I}_M : \left|c_i(\hat{u})\right| < \gamma\},\\
H_4 &:= \{i \in \mathbb{I}_M : \left|c_i(\hat{u} + v)\right| \geq \gamma\} \cap \{i \in \mathbb{I}_M : \left|c_i(\hat{u})\right| \geq \gamma\}.
\end{aligned}$$

These partitions help rewrite $\mathcal{F}_d(\hat{u} + v)$, where $v \in B_\infty(0, \rho) \cap (\mathbb{R}^N \setminus K_{\hat{\sigma}})$:

$$\begin{aligned}
\mathcal{F}_d(\hat{u} + v) &= \Psi(\hat{u} + v) + \beta \sum_{i \in \mathbb{I}_N} \phi(\hat{u}[i] + v[i])\\
&= \sum_{i \in H_1 \cup H_2} \frac{c_i(\hat{u} + v)^2}{2\gamma} + \sum_{i \in H_3 \cup H_4} \left(\left|c_i(\hat{u} + v)\right| - \frac{\gamma}{2} \right) + \beta \sum_{i \in \mathbb{I}_N} \phi(\hat{u}[i] + v[i])\\
&\geq \sum_{i \in H_1} \frac{c_i(\hat{u} + v)^2}{2\gamma} + \sum_{i \in H_2 \cup H_4} \left(\left|c_i(\hat{u} + v)\right| - \frac{\gamma}{2} \right)\\
&\quad + \sum_{i \in H_3} \frac{c_i(\hat{u})^2}{2\gamma} + \beta \sum_{i \in \mathbb{I}_N} \phi(\hat{u}[i] + v[i])\\
&\geq \sum_{i \in H_1} \frac{c_i(\hat{u} + v)^2}{2\gamma} + \sum_{i \in H_2 \cup H_4} \left(\left|c_i(\hat{u})\right| - \frac{\gamma}{2} \right)\\
&\quad + \sum_{i \in H_3} \frac{c_i(\hat{u})^2}{2\gamma} - \sum_{i \in H_2 \cup H_4} \left|c_i(\hat{u} + v) - c_i(\hat{u})\right|
\end{aligned}$$

$$
\begin{aligned}
&+ \beta \sum_{i\in\mathbb{I}_N} \phi(\hat{u}[i] + v[i]) \\
&\geq \sum_{i\in H_1\cup H_3} \frac{c_i(\hat{u})^2}{2\gamma} + \sum_{i\in H_2\cup H_4} \left(\left|c_i(\hat{u})\right| - \frac{\gamma}{2} \right) \\
&+ \sum_{i\in H_1} \frac{c_i(\hat{u})\langle Av, e_i\rangle}{\gamma} - \sum_{i\in H_2\cup H_4} \left|\langle v, A^T e_i\rangle\right| \\
&+ \beta \sum_{i\in\mathbb{I}_N} \phi(\hat{u}[i] + v[i]) \\
&\geq \Psi(\hat{u}) - \sum_{i\in H_1\cup H_2\cup H_4} \left|\langle v, A^T e_i\rangle\right| + \beta \sum_{i\in\mathbb{I}_N} \phi(\hat{u}[i] + v[i]) \\
&\geq \Psi(\hat{u}) - \sum_{i\in\mathbb{I}_M} \|v\|_\infty \left\|A^T e_i\right\|_1 + \beta \sum_{i\in\hat{\sigma}} \phi(\hat{u}[i]) + \beta \sum_{i\in\hat{\sigma}^c} \phi(v[i]) \\
&= \mathcal{F}_d(\hat{u}) - \|v\|_\infty \sum_{i\in\mathbb{I}_M} \left\|A^T e_i\right\|_1 + \beta \|v_{\hat{\sigma}^c}\|_0 \\
&= \mathcal{F}_d(\hat{u}) - \|v\|_\infty \|A\|_{1,1} + \beta \|v_{\hat{\sigma}^c}\|_0 \geq \mathcal{F}_d(\hat{u}).
\end{aligned}
$$

For $v = 0$, inequality is trivial. Otherwise, $\|v_{\hat{\sigma}^c}\|_0 \geq 1$ and the choice of the radius provides the inequality. □

Lemma 4.2 *For any $d \in \mathbb{R}^M$ and for all $\beta > 0$, $\mathcal{F}_d$ has a strict (local) minimum at $\hat{u} = 0 \in \mathbb{R}^N$.*

Proof Using the fact that

$$
\mathcal{F}_d(0) = \sum_{i\in H_0} \frac{d[i]^2}{2\gamma} + \sum_{i\in {H_0}^c} |d[i]| - \frac{\gamma}{2} \geq 0 \text{ where } H_0 := \{i \in \mathbb{I}_M : |d[i]| < \gamma\}
$$

we have

$$
\begin{aligned}
\mathcal{F}_d(v) &= \Psi(v) + \beta \|v\|_0 \\
&\geq \Psi(0) - \|v\|_\infty \|A\|_{1,1} + \beta \|v\|_0 \\
&\geq \Psi(0) - \|v\|_\infty \|A\|_{1,1} + \beta.
\end{aligned}
$$

Then, $v \in B_\infty\left(0, \dfrac{\beta}{\|A\|_{1,1} + 1}\right)$ implies $\mathcal{F}_d(v) > \mathcal{F}_d(0)$. □

Proposition 4.1 *Let $d \in \mathbb{R}^M$. Given $\omega \subseteq \mathbb{I}_N$, let $\hat{u}$ solve problem (HR_ω). Then, for any $\beta > 0$, the objective $\mathcal{F}_d$ reaches a (local) minimum at $\hat{u}$ and $\sigma(\hat{u}) \subseteq \omega$.*

Proof Let $\hat{u}$ solve (HR_ω), then $\hat{\sigma} = \sigma(\hat{u}) \subseteq \omega$ comes from the constraint of (HR_ω). Let $\hat{u} \neq 0$, then $1 \leq \left|\hat{\sigma}\right| \leq |\omega|$. The inclusion implies that $K_{\hat{\sigma}} \subseteq K_\omega$ and this shows: $v \in K_{\hat{\sigma}} \Rightarrow \Psi(\hat{u} + v) \geq \Psi(\hat{u})$. Also, $v \in K_{\hat{\sigma}}$ implies $v[i] = 0$ for all $i \in \hat{\sigma}^c$. Then, for all $v \in B_\infty(0, \rho) \cap K_{\hat{\sigma}}$

$$\begin{aligned}\mathcal{F}_d(\hat{u}+v) &= \Psi(\hat{u}+v) + \beta \sum_{i \in \mathbb{I}_N} \phi(\hat{u}[i]+v[i]) \\ &= \Psi(\hat{u}+v) + \beta \sum_{i \in \hat{\sigma}} \phi(\hat{u}[i]) \\ &\geq \Psi(\hat{u}) + \beta \sum_{i \in \hat{\sigma}} \phi(\hat{u}[i]) = \mathcal{F}_d(\hat{u}).\end{aligned}$$

Then, we have $\mathcal{F}_d(\hat{u}+v) \geq \mathcal{F}_d(\hat{u})$ for all $v \in B_\infty(0,\rho)$. For $\hat{u}=0$, it was proven before. □

Lemma 4.3 *For $d \in \mathbb{R}^M$ and $\beta > 0$, let $\mathcal{F}_d$ have a (local) minimum at $\hat{u}$. Then, $\hat{u}$ solves $(HR_{\hat{\sigma}})$ for $\hat{\sigma} = \sigma(\hat{u})$.*

Remark 4.1 Solving (HR_ω) for some $\omega \subseteq \mathbb{I}_N$ is equivalent to finding a (local) minimizer of $\mathcal{F}_d$.

Corollary 4.1 *For $d \in \mathbb{R}^M$ and $\beta > 0$, let $\hat{u}$ be a (local) minimizer of $\mathcal{F}_d$. Set $\hat{\sigma} = \sigma(\hat{u})$. Then,*

$$\hat{u} = Z_{\hat{\sigma}}(\hat{u}_{\hat{\sigma}}), \text{ where } \hat{u}_{\hat{\sigma}} \text{ satisfies } A_{\hat{\sigma}}^T \mathit{clip}(A_{\hat{\sigma}}\hat{u}_{\hat{\sigma}} - d) = 0. \tag{9}$$

Remark 4.2 Given $d \in \mathbb{R}^M$, for any $\omega \subseteq \mathbb{I}_N$, $\mathcal{F}_d$ has a (local) minimizer $\hat{u}$ defined by (9) and obeying $\sigma(\hat{u}) \subseteq \omega$.

In closing this section, we note that local minimizers coincide with the support optimal solutions of [17].

5 (Local) Strict Minimizers

Now, we shall concentrate on strict local minimizers.

Definition 5.1 Given a matrix $A \in \mathbb{R}^{M \times N}$, for any $r \in \mathbb{I}_M$, we define Ω_r as the subset of all r-length supports that correspond to a full column rank $M \times r$ sub matrix of A,

$$\Omega_r = \Big\{ \omega \subseteq \mathbb{I}_N : |\omega| = r = \operatorname{rank}(A_\omega) \Big\}. \tag{10}$$

If $\Omega_0 = \emptyset$, then we define

$$\Omega = \bigcup_{r=0}^{M-1} \Omega_r \text{ and } \Omega_{\max} = \Omega \cup \Omega_M. \tag{11}$$

This definition leads to: $\operatorname{rank}(A) = r \geq 1 \Leftrightarrow \Omega_r \neq \emptyset$ and $\Omega_t = \emptyset, \forall t \geq r+1$.

Proposition 5.1 *Given $d \in \mathbb{R}^M$ and $\beta > 0$, let $\hat{u}$ be a (local) minimizer of $\mathcal{F}_d$. Define $\hat{\sigma} = \sigma(\hat{u})$. $\operatorname{rank}(A_{\hat{\sigma}}) = \left|\hat{\sigma}\right|$ if and only if $\hat{\sigma} \in \Omega_{\max}$.*

Theorem 5.1 *(Strict Minimizers I) Given $d \in \mathbb{R}^M$ and $\beta > 0$, let $\hat{u}$ be a (local) minimizer of $\mathcal{F}_d$. Define $\hat{\sigma} = \sigma(\hat{u})$. If the (local) minimum that $\mathcal{F}_d$ has at $\hat{u}$ is strict, then* $\operatorname{rank}(A_{\hat{\sigma}}) = \left|\hat{\sigma}\right|$.

Proof Let $\hat{u} \neq 0$. Suppose $\dim\ker(A_{\hat{\sigma}}) \geq 1$. Then, $v \in B_\infty(0, \rho) \cap K_{\hat{\sigma}}$ and $v_{\hat{\sigma}} \in \ker(A_{\hat{\sigma}})$ implies that

$$\begin{aligned}\mathcal{F}_d(\hat{u} + v) &= \Psi(\hat{u} + v) + \beta \left\|\hat{u} + v\right\|_0 \\ &= \sum_{i \in \mathbb{I}_M} \psi(\langle A\hat{u} + Av - d, e_i\rangle) + \beta \left\|\hat{u} + v\right\|_0 \\ &= \sum_{i \in \mathbb{I}_M} \psi(\langle A_{\hat{\sigma}}\hat{u}_{\hat{\sigma}} + A_{\hat{\sigma}}v_{\hat{\sigma}} - d, e_i\rangle) + \beta \left\|\hat{u} + v\right\|_0 \\ &= \sum_{i \in \mathbb{I}_M} \psi(\langle A\hat{u} - d, e_i\rangle) + \beta \left\|\hat{u} + v\right\|_0 = \mathcal{F}_d(\hat{u})\end{aligned}$$

Then, $\hat{u}$ cannot be a strict minimizer, which is a contradiction. If $\hat{u} = 0$, then $\hat{\sigma} = \emptyset$; hence $A_{\hat{\sigma}} \in \mathbb{R}^{M\times 0}$ and $\operatorname{rank}(A_{\hat{\sigma}}) = \left|\hat{\sigma}\right| = 0$. □

5.1 Equivalent Formulation of (*HR*)

For the purposes of this section, we restate the Huber loss problem (HR) with fixed matrix $A \in \mathbb{R}^{M\times N}$ and vector $d \in \mathbb{R}^M$ with scaling parameter γ as follows

$$\min_{u \in \mathbb{R}^N} \Psi(u). \tag{HR}$$

Throughout this subsection, we will assume $M \geq N$ and A to be full rank, i.e., we are working with an overdetermined system. Next, we introduce a different formulation for Huber regression. We define the following collection $\Upsilon := \{(\upsilon_\gamma, \upsilon_+, \upsilon_-) \in \{0,1\}^M \times \{0,1\}^M \times \{0,1\}^M : \upsilon_\gamma + \upsilon_+ + \upsilon_- = \mathbb{1}_M\}$. Using this collection, we define the following sets in $\mathbb{R}^N$ for all $\upsilon \in \Upsilon$:

$$F_\upsilon := \left\{ u \in \mathbb{R}^N : \begin{array}{ll} |c_i(u)| < \gamma, & \text{if } \upsilon_\gamma(i) = 1 \\ c_i(u) \geq \gamma, & \text{if } \upsilon_+(i) = 1 \\ c_i(u) \leq -\gamma, & \text{if } \upsilon_-(i) = 1 \end{array} \right\}.$$

These sets are convex and pairwise disjoint. Also, they satisfy $\bigcup_{\upsilon \in \Upsilon} F_\upsilon = \mathbb{R}^N$. If we denote the indicator function as

$$\tilde{I}_{F_\upsilon}(u) = \begin{cases} 0, & \text{if } u \in F_\upsilon, \\ +\infty, & \text{if } u \notin F_\upsilon, \end{cases}$$

we define the Domain Considering Huber Regression (DCHR) as follows

$$\min_{u\in\mathbb{R}^N}\min_{\upsilon\in\Upsilon}\Big[\Psi(u)+\tilde{I}_{F_\upsilon}(u)\Big]. \tag{DCHR}$$

Using the properties of the previously defined regions, we have $\Psi(u) = \min_{\upsilon\in\Upsilon}\Big[\Psi(u)+\tilde{I}_{F_\upsilon}(u)\Big]$, and this shows that (DCHR) and (HR) are equivalent. Since we can change the order of minimization, we can define an updated version of (DCHR)

$$\min_{\upsilon\in\Upsilon}\min_{u\in\mathbb{R}^N} W_\upsilon(u), \tag{DCHR2}$$

where $W_\upsilon(u) = \Psi(u)+\tilde{I}_{F_\upsilon}(u)$ is a convex function of u, hence $\min_{u\in\mathbb{R}^N} W_\upsilon(u)$ is a convex optimization problem. Now, suppose we have the following quadratic expression

$$Q_\upsilon(u) = \frac{1}{2\gamma}c(u)^T D_\gamma c(u) + (\upsilon_+-\upsilon_-)^T\Big[c(u)-\frac{\gamma}{2}(\upsilon_+-\upsilon_-)\Big],$$

where $D_\gamma = \operatorname{diag}(\upsilon_\gamma)$. Using this quadratic expression, we obtain another equivalent optimization problem to (HR)

$$\min_{\upsilon\in\Upsilon}\min_{u\in\mathbb{R}^N} Q_\upsilon(u)+\tilde{I}_{F_\upsilon}(u) = \min_{\upsilon\in\Upsilon}\min_{u\in F_\upsilon} Q_\upsilon(u). \tag{QHR}$$

Lemma 5.1 (Unique Solution of (QHR), [32]) *Let $\gamma > 0$, $A \in \mathbb{R}^{M\times N}$ and $d \in \mathbb{R}^M$. If $\hat{u}$ minimizes (HR), then there is a unique $\hat{\upsilon} \in \Upsilon$ which minimizes (QHR) such that $\hat{u} \in F_{\hat{\upsilon}}$. Furthermore, $D_\gamma Au = D_\gamma A\hat{u}$ for all u solving $\min_{u\in F_{\hat{\upsilon}}} Q_\upsilon(u)$.*

The following simple observation is used in the proof of the theorem following it.

Lemma 5.2 *Consider the linear programming problem $\min_{u\in P} c^T u$ with $c \neq 0$, then interior points cannot be optimal.*

Proof Assume that an interior point $\hat{u}$ is an optimal solution to $\min_{u\in P} c^T u$. Then, there is some $\varepsilon > 0$ such that $c^T(\hat{u}-\varepsilon c) = c^T\hat{u} - \varepsilon\,\|c\|_2^2 < c^T\hat{u}$ contradicting the optimality of $\hat{u}$. If the polyhedron contains equality constraints, $Eu = d$, say, then let $\hat{c} = \mathbf{P}_{N(E)}c$ denote the projection of c onto the null space of E with the projection matrix $\mathbf{P}_{N(E)}$. We still have $c^T(\hat{u}-\varepsilon\hat{c}) = c^T\hat{u} - \varepsilon\,\|\mathbf{P}_{N(E)}c\|_2^2 < c^T\hat{u}$ by the properties of the projection matrix. □

Theorem 5.2 *Let $\gamma > 0$, $A \in \mathbb{R}^{M\times N}$ and $d \in \mathbb{R}^M$. Let $\mathbb{B}_\gamma = \{u \in \mathbb{R}^N : \forall i \in \mathbb{I}_M$ st. $|c_i(u)| \neq \gamma\}$.[2] If $\hat{u}$ minimizes (HR) such that $\hat{u} \in \mathbb{B}_\gamma$, then $\mathbb{1}_M^T\hat{\upsilon}_\gamma \geq N$, where $\hat{\upsilon}$ solves (QHR).*

[2] This set appears in our subsequent results and is shown to be dense in $\mathbb{R}^N$ in "Appendix."

Proof Assume that $\mathbb{1}_M^T \hat{\upsilon}_\gamma < N$. Since $\hat{\upsilon}$ minimizes (QHR), we can solve

$$\min_{u \in F_{\hat{\upsilon}}} Q_{\hat{\upsilon}}(u) = \min_{u \in F_{\hat{\upsilon}}} \frac{1}{2\gamma} c(u)^T D_\gamma c(u) + (\hat{\upsilon}_+ - \hat{\upsilon}_-)^T \left[c(u) - \frac{\gamma}{2}(\hat{\upsilon}_+ - \hat{\upsilon}_-) \right],$$

instead of (HR). Since $D_\gamma Au = D_\gamma A\hat{u}$ for any solution of this problem, we have

$$\begin{aligned} \min_{u \in F_{\hat{\upsilon}}} Q_{\hat{\upsilon}}(u) = \frac{1}{2\gamma} c(\hat{u})^T D_\gamma c(\hat{u}) - \frac{\gamma}{2} \left\| \hat{\upsilon}_+ - \hat{\upsilon}_- \right\|_2^2 \\ - (\hat{\upsilon}_+ - \hat{\upsilon}_-)^T d + \min_{\substack{u \in F_{\hat{\upsilon}} \\ D_\gamma A(u - \hat{u}) = 0}} (\hat{\upsilon}_+ - \hat{\upsilon}_-)^T Au. \end{aligned}$$

Hence, one can solve the following linear programming problem to identify minimizers of (HR)

$$\begin{aligned} \text{minimize } & (\hat{\upsilon}_+ - \hat{\upsilon}_-)^T Au \\ \text{subject to } & D_\gamma A(u - \hat{u}) = 0, \\ & D_+ Au \geq \gamma \hat{\upsilon}_+, \\ & D_- Au \leq -\gamma \hat{\upsilon}_-, \end{aligned} \tag{HR-LP}$$

where D_+ and D_- are $\operatorname{diag}(\hat{\upsilon}_+)$ and $\operatorname{diag}(\hat{\upsilon}_-)$, respectively. Since $\mathbb{1}_M^T \hat{\upsilon}_\gamma < N$, we have

$$\dim(\ker(D_\gamma A)) \geq 1,$$

; therefore, there exists points different from $\hat{u}$ in feasible region of (HR-LP). But since $\hat{u} \in \mathbb{B}_\gamma$ is an interior point, by previous lemma there is some extreme point u^* solving (HR-LP) with

$$(\hat{\upsilon}_+ - \hat{\upsilon}_-)^T Au^* < (\hat{\upsilon}_+ - \hat{\upsilon}_-)^T A\hat{u},$$

and $\Psi(u^*) < \Psi(\hat{u})$, which contradicts the fact that $\hat{u}$ is a minimizer of (HR). Hence, we have $\mathbb{1}_M^T \hat{\upsilon}_\gamma \geq N$. □

Example 5.1 Suppose we have some (local) minimizer $\hat{u}$ of $(\mathcal{F}_d)$ with $\gamma = 0.1$. Then, $\hat{u}$ solves ($\text{HR}_{\hat{\sigma}}$) with

$$A_{\hat{\sigma}} = \begin{bmatrix} 17 & 20 & 9 & 14 & 6 \\ 19 & 20 & 19 & 16 & 1 \\ 3 & 4 & 16 & 15 & 2 \\ 19 & 20 & 20 & 8 & 17 \\ 13 & 20 & 14 & 14 & 14 \\ 2 & 10 & 1 & 4 & 7 \\ 6 & 17 & 17 & 15 & 20 \\ 11 & 3 & 19 & 1 & 1 \end{bmatrix}, \quad d = \begin{bmatrix} 9 \\ 8 \\ 16 \\ 16 \\ 4 \\ 10 \\ 9 \\ 13 \end{bmatrix}, \text{ and } \hat{u}_{\hat{\sigma}} = \begin{bmatrix} 2.0536 \\ -2.5221 \\ -0.2379 \\ 1.4282 \\ 1.1116 \end{bmatrix}.$$

Now, we can examine $c(\hat{u}) = [-0.0073\ 0.0192\ \ -0.0877\ \ -1.8577\ 4.4822\ \ -17.8572\ 0.0578\ 0.0432]$, hence $\hat{u} \in \mathbb{B}_\gamma$. Also, these values return a $\hat{\upsilon}$ such that

$$\upsilon_\gamma = \left[1\ 1\ 1\ 0\ 0\ 0\ 1\ 1\right]^T,$$

and $\mathbb{1}_M^T \upsilon_\gamma = 5 \geq 5 = \left|\hat{\sigma}\right|$ as expected. If we try this with $\gamma = 1$ and $\gamma = 10$, we obtain

$$\begin{aligned} \hat{u}_{\hat{\sigma}} &= \left[1.9554\ -2.3685\ -0.1777\ 1.2945\ 1.0848\right]^T, \\ c(\hat{u}) &= \left[-0.0958\ 0.2030\ -0.8637\ -0.9730\ 4.8732\ -17.1799\ 0.5616\ 0.4071\right]^T, \\ \mathbb{1}_M^T \upsilon_\gamma &= 6 \geq 5 = \left|\hat{\sigma}\right|, \end{aligned}$$

and

$$\begin{aligned} \hat{u}_{\hat{\sigma}} &= \left[-0.1914\ 0.1431\ 0.7172\ -0.0476\ -0.0205\right]^T, \\ c(\hat{u}) &= \left[-3.7240\ 4.0729\ -5.2803\ -3.1570\ 5.4637\ -8.5678\ 3.3550\ -1.1160\right]^T, \\ \mathbb{1}_M^T \upsilon_\gamma &= 8 \geq 5 = \left|\hat{\sigma}\right|, \text{ respectively.} \end{aligned}$$

Example 5.2 One should also observe that there are instances with $\hat{u} \in \mathbb{B}_\gamma^c$. Let $\hat{u}$ be a local minimizer of $(\mathcal{F}_d)$ with $\gamma = 0.1$. Then, $\hat{u}$ solves $(\mathrm{HR}_{\hat{\sigma}})$ with

$$A_{\hat{\sigma}} = \begin{bmatrix} -0.0742 & -0.5227 & 0.5227 & 0.5227 \\ 0.9961 & -0.0273 & 0.0273 & 0.0273 \\ -0.0273 & 0.8074 & 0.1926 & 0.1926 \\ 0.0273 & 0.1926 & 0.8074 & -0.1926 \\ 0.0273 & 0.1926 & -0.1926 & 0.8074 \end{bmatrix},$$

$$d = \begin{bmatrix} 3.69156058 \\ 10 \\ 10 \\ 10 \\ 10 \end{bmatrix}, \text{ and } \hat{u}_{\hat{\sigma}} = \begin{bmatrix} 9.81743962 \\ 8.70987760 \\ 6.02898938 \\ 11.29012240 \end{bmatrix}.$$

From this information, we investigate $c(\hat{u})$

$$c(\hat{u}) = \left[0.08003211\ 0.01418370\ 0.1\ -5.36113302\ -0.1\right]^T.$$

This shows that $\hat{u}$ is a (local) minimizer of $(\mathcal{F}_d)$ with $\hat{u} \in \mathbb{B}_\gamma^c$.

5.2 Sufficient Conditions for Strict Minimality

Lemma 5.1 and Theorem 5.2 have a very important role in the proof of the following theorem.

Theorem 5.3 *(Strict Minimizers II) Given $d \in \mathbb{R}^M$ and $\beta > 0$, let $\hat{u}$ be a (local) minimizer of $\mathcal{F}_d$. Define $\hat{\sigma} = \sigma(\hat{u})$. If $\operatorname{rank}(A_{\hat{\sigma}}) = |\hat{\sigma}|$ and $\hat{u} \in \mathbb{B}_\gamma$, then the (local) minimum that $\mathcal{F}_d$ has at $\hat{u}$ is strict.*

Proof $|\hat{\sigma}| = 0$ implies $\hat{u} = 0$ and this is a strict minimizer, proven before. Let $|\hat{\sigma}| \geq 1$. Given that $\hat{u} \in \mathbb{B}_\gamma$, there exists some positive number $\hat{\xi} := \min_{i \in \mathbb{I}_M} \left\{ \min\{|c_i(\hat{u}) - \gamma|, |c_i(\hat{u}) + \gamma|\}\right\}$, and one can define a positive radius $\rho^* := \min\left\{ \min_{i \in \hat{\sigma}} |\hat{u}[i]|, \frac{\beta}{\|A\|_{1,1} + 1}, \frac{\hat{\xi}}{\|A\|_\infty} \right\}$, where $\|A\|_\infty = \max_{i \in \mathbb{I}_M} \sum_{j \in \mathbb{I}_N} |A_{ij}|$. Also define three index sets

$$H_\gamma := \{j \in \mathbb{I}_M : |c_j(\hat{u})| < \gamma\}, \quad H_- := \{j \in \mathbb{I}_M : c_j(\hat{u}) < -\gamma\}, \quad H_+ := \{j \in \mathbb{I}_M : c_j(\hat{u}) > \gamma\}.$$

If $v \in B_\infty(0, \rho^*) \setminus K_{\hat{\sigma}}$, we have $\mathcal{F}_d(\hat{u}) < \mathcal{F}_d(\hat{u} + v)$ as shown in Lemma 4.1 since $\hat{\sigma}^c \neq \emptyset$. Now, suppose $v \in B_\infty(0, \rho^*) \cap K_{\hat{\sigma}}$, then we have $\beta \|\hat{u} + v\|_0 = \beta \|\hat{u}\|_0$ for all positive β. Then, $\mathcal{F}_d(\hat{u} + v) \geq \mathcal{F}_d(\hat{u})$ implies that $\Psi(\hat{u} + v) \geq \Psi(\hat{u})$. Thus, we can say that $A_{\hat{\sigma}}^T \operatorname{clip}(A_{\hat{\sigma}} \hat{u}_{\hat{\sigma}} - d) = 0_{|\hat{\sigma}|}$. Now, assume that $A_{\hat{\sigma}}^T \operatorname{clip}(A_{\hat{\sigma}} \hat{u}_{\hat{\sigma}} + A_{\hat{\sigma}} v_{\hat{\sigma}} - d) = 0_{|\hat{\sigma}|}$. Then, for all $i \in \hat{\sigma}$

$$\begin{aligned}
0 &= \sum_{j \in H_\gamma} (A_{\hat{\sigma}}^T)_{ij} (A_{\hat{\sigma}} \hat{u}_{\hat{\sigma}} + A_{\hat{\sigma}} v_{\hat{\sigma}} - d)[j] + \sum_{j \in H_+} \gamma (A_{\hat{\sigma}}^T)_{ij} - \sum_{j \in H_-} \gamma (A_{\hat{\sigma}}^T)_{ij} \\
&\quad - \sum_{j \in H_\gamma} (A_{\hat{\sigma}}^T)_{ij} (A_{\hat{\sigma}} \hat{u}_{\hat{\sigma}} - d)[j] - \sum_{j \in H_+} \gamma (A_{\hat{\sigma}}^T)_{ij} + \sum_{j \in H_-} \gamma (A_{\hat{\sigma}}^T)_{ij} \\
&= \sum_{j \in H_\gamma} (A_{\hat{\sigma}}^T)_{ij} \left[\sum_{k \in \sigma} (A_{\hat{\sigma}})_{jk} v_{\hat{\sigma}}[k] \right].
\end{aligned}$$

By previous results, we know that $|H_\gamma| \geq |\hat{\sigma}|$, then $[A_{\hat{\sigma}}^T]_{H_\gamma} [A_{\hat{\sigma}}]_{H_\gamma}$ is an invertible matrix with dimension $|\sigma|$. Hence, $[A_{\hat{\sigma}}^T]_{H_\gamma} [A_{\hat{\sigma}}]_{H_\gamma} v_{\hat{\sigma}} = 0$ implies that $v_{\hat{\sigma}} = 0$, so $v = 0$. □

Example 5.3 Using this result and the data in Example 5.1 with $\gamma = 0.1$ and $\beta = 5$, $\hat{u}$ is a (local) strict minimizer in the ball $B_\infty(\hat{u}, \rho^*)$, where $\rho^* = \min\left\{0.2379, \frac{5}{475}, \frac{0.0123}{20}\right\} = 6.15 \times 10^{-4}$.

Corollary 5.1 *Let $d \in \mathbb{R}^M$. Given an arbitrary $\omega \in \Omega_{\max}$, let $\hat{u}$ solve (HR_ω) with $\hat{u} \in \mathbb{B}_\gamma$ and $|\omega| \leq |H_\gamma|$. Then,* *(i)* $\hat{u}$ *reads as*

$$\hat{u} = Z_\omega(\hat{u}_\omega), \quad \textit{where} \quad \hat{u}_\omega = ([A_\omega^T]_{H_\gamma}[A_\omega]_{H_\gamma})^{-1} A_\omega^T \zeta(\hat{u}_\omega)$$
$$\textit{with} \quad \zeta(\hat{u}_\omega) = \begin{cases} \gamma, & j \in H_-, \\ -\gamma, & j \in H_+, \\ d_j, & j \in H_\gamma, \end{cases}$$

and obeys $\hat{\sigma} = \sigma(\hat{u}) \subseteq \omega$ *and* $\hat{\sigma} \in \Omega_{\max}$*;*

(ii) for any $\beta > 0$*,* $\hat{u}$ *is a strict (local) minimizer of* $\mathcal{F}_d$*;*

(iii) $\hat{u}$ *solves* $(HR_{\hat{\sigma}})$.

Proof (i) Comes from the fact that $[A_\omega^T]_{H_\gamma}[A_\omega]_{H_\gamma}$ is invertible under the given conditions. Then, Proposition 4.1 implies part (ii) and Lemma 4.3 implies part (iii). □

Remark 5.1 One can easily compute a strict (local) minimizer $\hat{u}$ of $(\mathcal{F}_d)$ without knowing the value of the regularization parameter β. However, Corollary 5.1 is useful when the unique solution υ is known for the problem (QHR$_\omega$). In other words, if one knows for which $\upsilon \in \Upsilon$ a solution u^* of (HR$_\omega$) satisfies $u^* \in F_\upsilon$ then one can use Corollary 5.1 to find u^* exactly.

6 Global Minimizers

In this section, we first show that one can prove a necessary condition for a global minimizer. Then, non-emptiness of the set of global minimizers and their sparsity will be studied.

6.1 A Necessary Condition and Boundaries of β

The following result gives a necessary condition for global minimizers. It is also useful for finding meaningful values for the parameter β as we shall discuss below.

Theorem 6.1 *For* $d \in \mathbb{R}^M$ *and* $\beta > 0$*, let* $\mathcal{F}_d$ *have a global minimum at* $\hat{u}$*. Then,*

$$i \in \sigma(\hat{u}) \Rightarrow \left|\hat{u}[i]\right| \geq \frac{\beta}{\sqrt{M}\,\|a_i\|}. \tag{12}$$

Proof Let $\hat{u}$ be a global minimizer of $\mathcal{F}_d$. For $\hat{u} = 0$, there is nothing to prove. Let $\left|\sigma(\hat{u})\right| \geq 1$, for all $i \in \mathbb{I}_N$ define $g_i(u) : \mathbb{I}_N \times \mathbb{R}^N \to \mathbb{R}^N$ as

$$g_i(u) = (u[1], \ldots, u[i-1], 0, u[i+1], \ldots, u[N]).$$

Then, again for all $i \in \mathbb{I}_N$ we have

$$\mathcal{F}_d(\hat{u}) = \mathcal{F}_d(g_i(\hat{u}) + e_i \hat{u}[i]) = \Psi(g_i(\hat{u}) + e_i \hat{u}) + \beta \sum_{j \in \mathbb{I}_N} \Big(\phi(g_i(\hat{u})[j]) + \phi(\hat{u}[i]) \Big).$$

Then, we have $\mathcal{F}_d(\hat{u}) \le \mathcal{F}_d(g_i(\hat{u}))$ and $\Psi(g_i(\hat{u})) - \Psi(\hat{u}) \ge \beta$. Using these, we obtain

$$\begin{aligned} \left|\hat{u}[i]\right|^2 &= \frac{\|a_i\|_2^2 \left|\hat{u}[i]\right|^2}{\|a_i\|_2^2} \ge \frac{\|a_i\|_2^2 \left|\hat{u}[i]\right|^2 \|\mathrm{clip}(Ag_i(u) - d)\|_2^2}{M \|a_i\|_2^2} \\ &\ge \frac{\left|\hat{u}[i]\mathrm{clip}(Ag_i(\hat{u}) - d)^T a_i\right|^2}{M \|a_i\|_2^2} = \frac{\left|\hat{u}[i]\mathrm{clip}(Ag_i(\hat{u}) - d)^T Ae_i\right|^2}{M \|a_i\|_2^2}. \end{aligned}$$

Taking square roots will lead to

$$\begin{aligned} \left|\hat{u}[i]\right| &\ge \frac{\left|\hat{u}[i]\mathrm{clip}(Ag_i(\hat{u}) - d)^T Ae_i\right|}{\sqrt{M} \|a_i\|_2} \ge \frac{-\hat{u}[i]\mathrm{clip}(Ag_i(\hat{u}) - d)^T Ae_i}{\sqrt{M} \|a_i\|_2} \\ &= \frac{-\hat{u}[i]\nabla\Psi(g_i(\hat{u}))^T e_i}{\sqrt{M} \|a_i\|_2} = \frac{\nabla\Psi(g_i(\hat{u}))^T (g_i(\hat{u}) - \hat{u})}{\sqrt{M} \|a_i\|_2} \\ &\ge \frac{\Psi(g_i(\hat{u})) - \Psi(\hat{u})}{\sqrt{M} \|a_i\|_2} \ge \frac{\beta}{\sqrt{M} \|a_i\|_2}. \end{aligned}$$

□

The lower bound provided above for the support is independent of d. This property is also true for local minimizers of $\mathcal{F}_d$ satisfying $\mathcal{F}_d(\hat{u}) \le \mathcal{F}_d(\hat{u} + \rho e_i), \quad \forall \rho \in \mathbb{R}, \quad \forall i \in \mathbb{I}_N$.

Example 6.1 Consider the following numerical example with $M = 8$ and $N = 17$ and the matrix A:

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 8 & 1 & 1 & 8 & 9 & 6 & 8 & 0 & 6 & 8 & 6 & 0 \\ -1 & 0 & -5 & 3 & -1 & 9 & 3 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 9 & 0 \\ 0.5 & -17 & 15 & -7 & 5 & -7 & 7 & 5 & -7 & 1 & -7 & 1 & 1 & 7 & 11 & 7 & 11 \\ -1 & 2 & -1 & 2 & -1 & 2 & 8 & -1 & 2 & 5 & 3 & 50 & 5 & 3 & 50 & 3 & 0 \\ 4 & 5 & 8 & 1 & 2 & 7 & -1 & 2 & 9 & 1 & 19 & 11 & 31 & 29 & 11 & 20 & 10 \\ 5 & 7 & 0 & 1 & 3 & 10 & 6 & 7 & -1 & 4 & -1 & 4 & 14 & -1 & 24 & -1 & 26 \\ 6 & 1 & 10 & -2 & 1 & 0 & 1 & 2 & 3 & 38 & 15 & 0 & 5 & 8 & 23 & 8 & 28 \\ 9 & 0 & 7 & 5 & 9 & 1 & 0 & 0 & 3 & 1 & 20 & 4 & 0 & 2 & 21 & 2 & 8 \end{bmatrix}$$

and

$$d = [-1, 2, 4, 7, -3, -10, 20, 3]^T.$$

Choosing $\gamma = 0.1$ and $\beta = 3$, we have a global minimizer u^* with the only nonzero entries 4.944, −1.664 , −4.608, 1.049, −2.217 at components 1, 2, 5, 7, 8. Then, it is easy to verify that the necessary condition (12) in Proposition 6.1 is fulfilled. The values $\frac{\beta}{\sqrt{M}\,\|a_i\|}$ for the entries 1, 2, 5, 7, 8 are 0.0835, 0.0553, 0.0956, 0.0836, 0.1157 which are all smaller than the respective entries of $\hat{u}$ in absolute value.

Remark 6.1 Let $d \in \mathbb{R}^M$ and $\beta > \Psi(0)$. Then, $(\mathcal{F}_d)$ has a strict global minimum at $\hat{u} = 0$, since for any nonzero u we have $\|u\|_0 \geq 1$ and $\mathcal{F}_d(0) = \Psi(0) < \beta \leq \Psi(u) + \beta \|u\|_0$. This shows that $(\mathcal{F}_d)$ has nonzero global minima only for finite values of β. Using these two bounds, one can use the interval $\Big(0, \Psi(0)\Big]$ for β and adjust γ as desired.

6.2 Non-triviality of the Global Minimizers

In this section, we shall prove that the set of global minimizers is non-empty. We start by recalling some useful definitions below.

– Let C be a non-empty set in $\mathbb{R}^N$. Then, the *asymptotic cone* of the set C, denoted by C_∞, is the set of vectors $t \in \mathbb{R}^N$ that are limits in direction of the sequences $\{x_k\} \subset C$, namely

$$C_\infty = \left\{ t \in \mathbb{R}^N : \exists y_k \to +\infty, \exists x_k \in C \text{ with } \lim_{k\to\infty} \frac{x_k}{y_k} = t \right\}.$$

The following definitions are for $f : \mathbb{R}^N \to \mathbb{R}$,

– The *level set* of f for λ is defined as $\operatorname{lev}(f, \lambda) := \{v \in \mathbb{R}^N : f(v) \leq \lambda\}$ for $\lambda > \inf f$.
– The *epigraph* of f is defined as $\operatorname{epi} f := \{(v, \lambda) \in \mathbb{R}^N \times \mathbb{R} : f(v) \leq \lambda\}$.
– For any proper function $f : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$, there exists a unique function $f_\infty : \mathbb{R}^N \to \mathbb{R} \cup \{\infty\}$ associated with f, called the *asymptotic function*, such that $\operatorname{epi} f_\infty = (\operatorname{epi} f)_\infty$.

Definition 6.1 Let $f : \mathbb{R}^N \to \mathbb{R} \cup \{\infty\}$ be lower semi-continuous and proper. Then, f is said to be asymptotically level stable if for each $\rho > 0$, each bounded sequence $\{\lambda_k\} \in \mathbb{R}$, and each sequence $\{v_k\} \in \mathbb{R}^N$ satisfying

$$v_k \in \operatorname{lev}(f, \lambda_k), \quad \|v_k\| \to \infty, \quad v_k \|v_k\|^{-1} \to \overline{v} \in \ker((f)_\infty),$$

where $(f)_\infty$ denotes the asymptotic (or recession) function of f, there exists k_0 such that

$$v_k - \rho \overline{v} \in \operatorname{lev}(f, \lambda_k) \quad \forall k \geq k_0.$$

Remark 6.2 $\mathcal{F}_d$ defined in (1) is a non-negative function which is not constantly $+\infty$, which shows it is a proper function. We show that all level sets of our function are closed to establish lower-semi continuity. Let $\alpha \in \mathbb{R}$ be arbitrary, and let $x \in \mathbb{R}^N \setminus \{0\}$

be such that $\mathcal{F}_d(v) > \alpha$. Then, $\mathcal{F}_d = \alpha + \delta$ for some $\delta > 0$. Let ρ be a positive radius defined in Lemma 4.1 and $\rho^* = \min\left\{\rho, \dfrac{\delta}{\|A\|_{1,1,}}\right\}$, then we have

$$\mathcal{F}_d(x+v) \geq \mathcal{F}_d(v) - \|x\|_\infty \|A\|_{1,1} + \beta \|x_{\sigma^c}\|_0,$$

for $x \in B_\infty(0, \rho^*)$ and $\sigma = \sigma(v)$. If $x \in B_\infty(0, \rho^*) \setminus K_\sigma$, we have

$$\mathcal{F}_d(x+v) \geq \mathcal{F}_d(v) - \|x\|_\infty \|A\|_{1,1} + \beta \|x_{\sigma^c}\|_0 \geq \mathcal{F}_d(v) > \alpha.$$

Otherwise, $x \in B_\infty(0, \rho^*) \cap K_\sigma$ and

$$\mathcal{F}_d(x+v) \geq \mathcal{F}_d(v) - \|x\|_\infty \|A\|_{1,1} = \alpha + \delta - \|x\|_\infty \|A\|_{1,1} > \alpha.$$

Therefore, we have shown that $\mathcal{F}_d$ is a lower-semi continuous function.

Proposition 6.1 *Let $\mathcal{F}_d : \mathbb{R}^N \to \mathbb{R}$ be of the form (1). Then,* $\ker((\mathcal{F}_d)_\infty) = \ker(A)$, *and $\mathcal{F}_d$ is asymptotically level stable.*

Proof Since $\mathcal{F}_d$ is a proper function, the asymptotic function may be written explicitly

$$\begin{aligned}(\mathcal{F}_d)_\infty(v) &= \liminf_{\substack{v' \to v \\ x \to \infty}} \frac{\mathcal{F}_d(xv')}{x} = \liminf_{\substack{v' \to v \\ x \to \infty}} \frac{\sum_{i \in \mathbb{I}_M} \psi(c_i(xv')) + \beta \left\|xv'\right\|_0}{x} \\ &= \liminf_{\substack{v' \to v \\ x \to \infty}} \frac{\sum_{i \in \mathbb{I}_M} \psi(c_i(xv')) + \beta \left\|v'\right\|_0}{x}.\end{aligned}$$

Then, for $v \notin \ker(A)$ there is some k such that $\sum_{j \in \mathbb{I}_N} A_{kj} v_j \neq 0$. Since $\ker(A)$ is closed then there is some open ball around v which does not belong to the kernel. Hence,

$$\begin{aligned}(\mathcal{F}_d)_\infty(v) &= \liminf_{x \to \infty} \frac{\sum_{i \in \mathbb{I}_M} \psi(c_i(xv)) + \beta \|v\|_0}{x} \geq \liminf_{x \to \infty} \frac{\psi(c_k(xv)) + \beta \|v\|_0}{x} \\ &= \liminf_{x \to \infty} \frac{\left|x(\sum_{j \in \mathbb{I}_N} A_{kj} v_j) - d_k\right| - \frac{\gamma}{2} + \beta \|v\|_0}{x} = \left|\sum_{j \in \mathbb{I}_N} A_{kj} v_j\right| > 0.\end{aligned}$$

Now, let $v \in \ker(A)$. We want to show that $(\mathcal{F}_d)_\infty(v) = 0$. Then, it suffices to show there exists a direction such that $\dfrac{\mathcal{F}_d(xv')}{x}$ is zero, since it is a non-negative expression. Now, let v'_n be a sequence in $\ker(A)$ converging to v. Then, we have

$$\lim_{\substack{n \to \infty \\ x \to \infty}} \frac{\mathcal{F}_d(xv'_n)}{x} = \lim_{\substack{n \to \infty \\ x \to \infty}} \frac{\sum_{i \in \mathbb{I}_M} \psi(c_i(xv'_n)) + \beta \|v\|_0}{x} = \lim_{x \to \infty} \frac{\sum_{i \in \mathbb{I}_M} \psi(d_i) + \beta \|v\|_0}{x} = 0.$$

Combining these two results we obtain that $\ker((\mathcal{F}_d)_\infty) = \ker(A)$ where $\ker((\mathcal{F}_d)_\infty) = \{v \in \mathbb{R}^N : (\mathcal{F}_d)_\infty(v) = 0\}$. Now, let $\{\lambda_k\}_{k\in\mathbb{N}}$ be a bounded sequence of real numbers, $\rho > 0$ and $\{v_k\}_{k\in\mathbb{N}}$ be an arbitrary sequence in $\mathbb{R}^N$. Then, we compare $\|v_k - \rho\overline{v}\|_0$ and $\|v_k\|_0$. Let $i \in \sigma(\overline{v})$, then

$$\lim_{k\to\infty} \frac{v_k}{\|v_k\|} \neq 0 \Rightarrow |v_k| > 0 \text{ for } k \geq k_i,$$

for some fixed constant $k_i \in \mathbb{N}$. Otherwise, let $i \in \sigma^c(\overline{v})$, then $\|v_k - \rho\overline{v}\|_0 \leq \|v_k\|_0$. We define $k_0 = \max_{i\in\sigma(\overline{v})} k_i$, then we obtain

$$\begin{aligned}\mathcal{F}_d(v_k - \rho\overline{v}) &= \Psi(v_k - \rho\overline{v}) + \beta\,\|v_k - \rho\overline{v}\|_0 \\ &\leq \Psi(v_k) + \beta\,\|v_k\|_0 = \mathcal{F}_d(v_k) \leq \lambda_k,\end{aligned}$$

since $\overline{v} \in \ker(A)$. This shows that $v_k - \rho\overline{v} \in \operatorname{lev}(\mathcal{F}_d, \lambda_k)$ for $k \geq k_0$, $\mathcal{F}_d$ is asymptotically level stable. □

Lemma 6.1 *(Non-triviality of the Optimal Set, [33]) Let* $\mathcal{F}_d : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ *be asymptotically level stable with* $\inf \mathcal{F}_d > -\infty$. *Then, the optimal set* $\hat{U}$,

$$\hat{U} = \{\hat{u} \in \mathbb{R}^N : \mathcal{F}_d(\hat{u}) = \min_{u\in\mathbb{R}^N} \mathcal{F}_d(u)\},$$

is non-empty.

Theorem 6.2 *Let* $d \in \mathbb{R}^M$, $\beta > 0$ *and* $\gamma > 0$. *Then, the set of global minimizers of* $\mathcal{F}_d$, $\hat{U}$, *is non-empty.*

Proof By Proposition 6.1 and the definition of $\mathcal{F}_d$, we have $\mathcal{F}_d$ asymptotically level stable, and $\inf \mathcal{F}_d \geq 0$. □

Previously we had proven in Sect. 3 that (HR) itself has an optimal solution. As a result of the above theorem, we know that $(\mathcal{F}_d)$ always admits a minimizer. Below, we provide additional results for global minimizers in $\mathbb{B}_\gamma$, which we denote by $\mathcal{UB}$.

Theorem 6.3 *Let* $d \in \mathbb{R}^M$, $\beta > 0$ *and* $\gamma > 0$. *Then, every* $\hat{u} \in \mathcal{UB}$ *is a strict (local) minimizer of* $(\mathcal{F}_d)$, *i.e.,*

$$\sigma(\hat{u}) \in \Omega_{\max};$$

hence $\|\hat{u}\|_0 \leq M$.

Proof Let $\hat{u} \in \hat{U} \cap \mathbb{B}_\gamma$ and define $\sigma(\hat{u}) = \hat{\sigma}$. If $\hat{u} = 0$, this is done in Lemma 4.2. Suppose that the global minimizer $\hat{u} \neq 0$ of $(\mathcal{F}_d)$ is non-strict. Then, Theorem 5.3 fails to hold; i.e.,

$$\dim(\ker(A_{\hat{\sigma}})) \geq 1.$$

Choose $v_{\hat{\sigma}} \in \ker(A_{\hat{\sigma}}) \setminus \{0\}$ and set $v = Z_{\hat{\sigma}}(v_{\hat{\sigma}})$. Select an i satisfying $v[i] \neq 0$. Define u^* by $u^* := \hat{u} - \hat{u}[i]\dfrac{v}{v[i]}$. Then, we have $u^*[i] = 0$ while $\hat{u}[i] \neq 0$ and this implies $u^* \subsetneqq \hat{u}$ and $|\sigma^*| \leq \left|\hat{\sigma}\right| - 1$ where $\sigma^* = \sigma(u^*)$. By the choice of v, we have $A\hat{u} = A_{\hat{\sigma}}\hat{u} = A_{\hat{\sigma}}u^* = A_{\sigma^*}u^* = Au^*$, then

$$\mathcal{F}_d(u^*) = \Psi(u^*) + \beta \left\|u^*\right\|_0 \leq \Psi(\hat{u}) + \beta \left\|\hat{u}\right\|_0 - \beta = \mathcal{F}_d(\hat{u}) \Rightarrow \mathcal{F}_d(u^*) < \mathcal{F}_d(\hat{u}).$$

This contradicts the fact that $\hat{u}$ is a global minimizer, hence $\operatorname{rank}(A_{\hat{\sigma}}) = \left|\hat{\sigma}\right|$. Therefore, $\hat{u}$ is a strict minimizer, $\hat{\sigma} \in \Omega_{\max}$ and $\left\|\hat{u}\right\|_0 \leq M$. □

6.3 *K*-Sparse Global Minimizers for $K \leq M - 1$

Since A has full rank one can find invertible M-dimensional square submatrices. A consequence of this fact is that for small β values one may end up with multiple global minimizers as one can express d as a linear combination of the columns in the invertible $M \times M$ submatrix and obtain $\mathcal{F}_d(\hat{u}) = \beta M$ for different $\hat{u}$. Hence, it is interesting to examine $M-1$-dimensional (or, lower dimensional) submatrices. For any $K \in \mathbb{I}_{M-1}$, let

$$\overline{\Omega}_K := \bigcup_{r=0}^{K} \Omega_r,$$

where Ω_r was set up in Definition 5.1.

Proposition 6.2 *Let $d \in \mathbb{R}^M$. For any $K \in \mathbb{I}_{M-1}$, there exists $\beta_K \geq 0$, such that if $\beta > \beta_K$, then each global minimizer of $\hat{u}$ of $\mathcal{F}_d$ satisfies $\left\|\hat{u}\right\|_0 \leq K$. Furthermore, for all K-sparse vectors $\hat{u} \in \mathcal{UB}$, $\sigma(\hat{u}) \in \overline{\Omega}_K$.*

Proof Given $K \in \mathbb{I}_{M-1}$, set

$$U_{K+1} := \bigcup_{\omega \subset \mathbb{I}_N} \{\overline{u} : \overline{u} \text{ solves } (\mathrm{HR}_\omega) \text{ and } \|\overline{u}\|_0 \geq K+1\}.$$

Let $U_{K+1} \neq \emptyset$. Then, for any $\beta > 0$, $\mathcal{F}_d$ has a (local) minimum at each $\overline{u} \in U_{K+1}$. Therefore,

$$\overline{u} \text{ is a (local) minimizer of } \mathcal{F}_d \text{ and } \|u\|_0 \geq K+1 \Leftrightarrow \overline{u} \in U_{K+1}.$$

Then, for any $\beta > 0$ we have $\mathcal{F}_d(\overline{u}) \geq \beta(K+1)$ for all $\overline{u} \in U_{K+1}$. Let $\tilde{u}$ solve (HR_ω) for some $\omega \in \Omega_K$. Then, $\|\tilde{u}\|_0 \leq K$. Set β and β_K according to $\beta > \beta_K = \Psi(\tilde{u})$. For such a β, we have

$$\mathcal{F}_d(\tilde{u}) = \Psi(\tilde{u}) + \beta \|\tilde{u}\|_0 < \beta + \beta K \leq \mathcal{F}_d(\overline{u}),$$

for all $\overline{u} \in U_{K+1}$. Then, for any global minimizer $\hat{u} \in \hat{U}$ we have

$$\mathcal{F}_d(\hat{u}) \leq \mathcal{F}_d(\tilde{u}) < \mathcal{F}_d(\overline{u}),$$

for $\overline{u} \in U_{K+1}$, therefore $\left\|\hat{u}\right\|_0 \leq K$. Now, suppose $U_{K+1} = \emptyset$. Then, $\overline{u}$ solving (HR_ω) for $\omega \subset \mathbb{I}_N$ with $|\omega| \geq K+1$, we have $\|\overline{u}\|_0 \leq K$. Let $\hat{u}$ be a global minimizer of $\mathcal{F}_d$, then $\left\|\hat{u}\right\|_0 \leq K$. If one has $\hat{u} \in \mathcal{UB}$ with $\|u\|_0 \leq K$, we have shown that $\hat{u}$ is a strict minimizer and $\sigma(\hat{u}) \in \overline{\Omega}_K$. □

We have found a meaningful region for β satisfying $\beta \in \Big(0, \Psi(0)\Big]$ in previous sections. It is easy to show that for any $K \in \mathbb{I}_{M-1}$, we have $\beta_K \leq \Psi(0)$, for β_K provided in the proof of Proposition 6.2. Therefore, the region can be made finer by specifying

$$\beta \in \Big(\beta_K, \Psi(0)\Big],$$

to investigate K-sparse global minimizers. As stated in the proof, a solution to (HR_ω) for $\omega \in \Omega_K$ is sufficient for a lower bound β_K. But one can relax this lower bound by taking

$$\beta_K^* := \min_{\omega \in \Omega_K} \{\Psi(\tilde{u}) : \tilde{u} \text{ solves } (\mathrm{HR}_\omega)\},$$

and restate the region as

$$\beta \in \Big(\beta_K^*, \Psi(0)\Big].$$

Example 6.2 Let $d \in \mathbb{R}^5$ be fixed and defined as $d = \begin{bmatrix}1 & 9 & 9 & 4 & 8\end{bmatrix}^T$. We investigate the choice of regularization parameter β along with γ. We can pick any regularization parameter in the region

$$\beta \in (0, \Psi(0)].$$

Suppose we have a data matrix $A \in \mathbb{R}^{5\times 10}$ as below

$$A = \begin{bmatrix} 1 & 6 & 7 & 7 & 8 & 4 & 4 & 2 & 4 & 7 \\ 4 & 0 & 7 & 0 & 6 & 3 & 4 & 6 & 9 & 2 \\ 9 & 8 & 3 & 2 & 3 & 7 & 6 & 6 & 3 & 5 \\ 7 & 9 & 6 & 0 & 9 & 7 & 7 & 1 & 5 & 6 \\ 9 & 6 & 1 & 0 & 0 & 1 & 7 & 1 & 2 & 8 \end{bmatrix},$$

and we look for k-sparse global minimizers. We have the plots below showing the relevant regions for γ and β for K-sparse global minimizers of $\mathcal{F}_d$ (Figs. 2, 3, 4, 5).

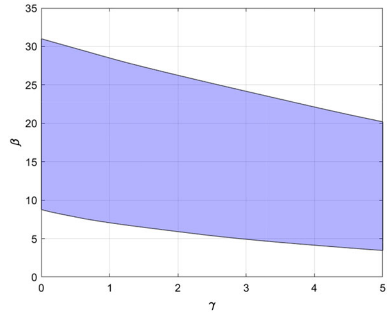

Fig. 2 $K = 1$

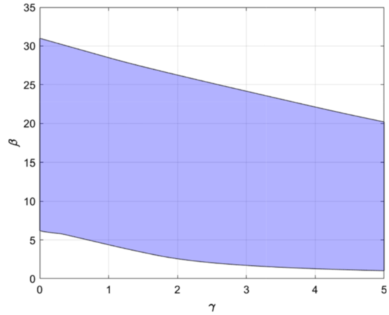

Fig. 3 $K = 2$

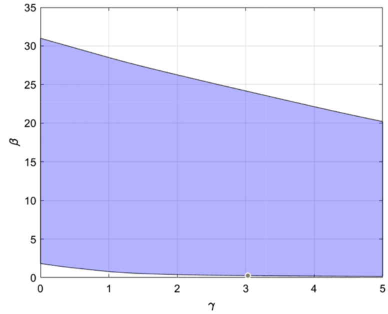

Fig. 4 $K = 3$

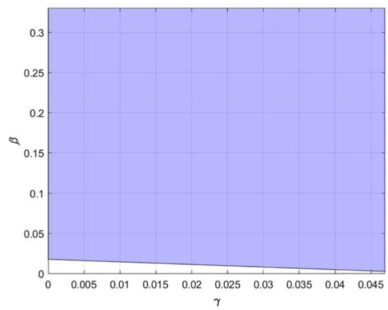

Fig. 5 $K = 4$

Proposition 6.3 *Let A_ω be a $M \times |\omega|$ dimensional submatrix of A, with* $\mathrm{rank}(A_\omega) < \omega \le M$*. If u solves HR_ω, then there exists a subset ω^* of ω such that $\mathcal{F}_d(u) \ge \mathcal{F}_d(\hat{u})$ for $\hat{u}$ solving HR_{ω^*}.*

Proof Since A_ω is not full rank, there is some submatrix of it A_{ω^*} which is $M \times |\omega^*|$ dimensional and $\mathrm{rank}(A_{\omega^*}) = \omega^* < \omega \le M$. Without loss of generality, we assume that first $|\omega^*|$ columns of A_ω are linearly independent and A_{ω^*} consist of first $|\omega^*|$ columns. Then, for any u in the feasible region of HR_ω we can write

$$\begin{aligned} A_\omega u &= a_{\omega[1]}u_1 + \cdots + a_{\omega[|\omega^*|]}u_{|\omega^*|} + \cdots + a_{\omega[|\omega|]}u_{|\omega|} \\ &= a_{\omega^*[1]}u_1^* + \cdots + a_{\omega^*[|\omega^*|]}u_{|\omega^*|}^* = A_{\omega^*}u^*. \end{aligned}$$

Second equality comes from the fact that each column $a_{|\omega^*|+1}, \ldots, a_{|\omega|}$ can be written as a linear combination of the remaining ones. Therefore, if u solves HR_ω we have

$$\mathcal{F}_d(u) = \Psi(u) + \beta \, \|u\|_0 = \Psi(u^*) + \beta \, \|u\|_0 \, .$$

In this case, we always have $\|u\|_0 \ge \|u^*\|_0$, because if u had some zero entries we could shrink A_ω accordingly and apply the steps above. Therefore, we have $\mathcal{F}_d(u) \ge \mathcal{F}_d(u^*) \ge \mathcal{F}_d(\hat{u})$. □

Remark 6.3 As stated in the beginning of this subsection, when searching for a global minimum it suffices to check submatrices with at most M columns. Then, one can solve an overdetermined Huber regression problem using (QHR) or (QHR2) (see Remark 8.2) in each of those subproblems and then compare them after adding penalty generated by the ℓ_0 norm. During the search, it suffices to check submatrices with full rank by Proposition 6.3. Using these results, we provide in Appendix an enumeration-based procedure for searching the global minimizers in small dimensional examples.

7 Connection to Other Optimality Criteria

Beck and Hallak [17] introduced the concept of L-stationary points for studying optimality conditions in sparsity constrained problems. Now, we relate our results to the concept of L-stationarity (see also [18,34] for similar results).

Lemma 7.1 *(Equivalent Form of L-stationary, [17]) Let $L > 0$. A vector $u \in \mathbb{R}^N$ is called an L-stationary point if and only if*

$$|\nabla_k \Psi(u)| \begin{cases} = 0, & k \in \sigma(u), \\ \le \sqrt{2\beta L}, & k \notin \sigma(u). \end{cases}$$

Proposition 7.1 *Let $\hat{u}$ be a strict local minimum of $\mathcal{F}_d$. Then, $\hat{u}$ is an L-stationary point for*

$$L \ge \max_{k \notin \hat{\sigma}} \frac{(\nabla_k \Psi(\hat{u}))^2}{2\beta}.$$

Proof Let $\hat{u}$ be a strict local minimum of $\mathcal{F}_d$. Then, by Proposition 4.3 we know that $\hat{u}$ solves ($\mathrm{HR}_{\hat{\sigma}}$). Then, KKT conditions imply

$$\begin{aligned}\left|\nabla_k \Psi(\hat{u})\right| &= 0, \quad k \in \hat{\sigma},\\ \left|\nabla_k \Psi(\hat{u})\right| &= \lambda_k, \quad k \notin \hat{\sigma},\end{aligned}$$

where λ_k are Lagrange multipliers. Then, $\hat{u}$ is an L-stationary point for

$$\max_{k \notin \hat{\sigma}} |\nabla_k \Psi(u)| \leq \sqrt{2\beta L}.$$

□

Another optimality concept developed in [17] (since the definition of this optimality criterion is quite involved we refer the reader to Section 4.4: Definition 4.13 of [17]) is the "Partial Coordinate-wise Optimality" (PCW-optimal) criterion which is stronger than the support optimality and L-stationarity (provided $L > L_f$ where L_f is the Lipschitz constant for the gradient mapping) criteria. A strict local minimizer needs not be a PCW-optimal point as the following example shows.

Example 7.1 Let us define $\gamma = 0.1$,

$$A = \begin{bmatrix} 1 & 4 & 9 \\ 6 & 2 & 5 \end{bmatrix} \text{ and } d = \begin{bmatrix} 1 \\ 9 \end{bmatrix}.$$

Suppose we have $K = \{2\} \subseteq \mathbb{I}_3$. Then, $\hat{u} = [\,0\ 0.263\ 0\,]^T$ is a support optimal point of $\mathcal{F}_d$. This is a strict minimizer since A_K has only one column and the residuals are 0.05 and -8.475. This minimizer leads to $\Psi(\hat{u}) = 8.4375$ and $\mathcal{F}_d(\hat{u}) = 8.435 + \beta$ for some $\beta > 0$. Then, $1 \in \mathbb{I}_3 \setminus K$ has the smallest partial derivative between $\{1, 3\}$. Then, $K_{\text{swap}} = \{1\}$ and $\bar{u} = [\,1.497\ 0\ 0\,]^T$ is a support optimal point of $\mathcal{F}_d$ and this point has the objective $\mathcal{F}_d(\bar{u}) = 0.449 + \beta$ which is smaller than $\mathcal{F}_d(\hat{u})$. This example shows that a strict local minimum of $\mathcal{F}_d$ is not partial-CW in general.

It is not certain that a PCW-optimal point is a strict local minimizer. However, this can be checked using the conditions developed in Sect. 5. More importantly, especially for algorithm development, one has the ability to check the necessary condition in Theorem 6.1 for a global minimizer after computing a PCW-optimal point using modifications of the algorithms described in [17]. This is left as future work.

8 Conclusions

We investigated the structure of local and global minimizers for the minimization of the Huber loss criterion in the solution of linear systems of equations, coupled with an ℓ_0 term controlling the sparsity of the solution through a regularization parameter β. We characterized local minimizers and gave conditions for local minimizers to be strict. We established non-emptiness of the set of global minimizers as well as a necessary condition for global minimizers. We gave bounds on the choice β to attain a desired level of sparsity. We related our results to existing optimality concepts in the literature. A simple enumeration scheme allowed us to illustrate the results via numerical examples. The development of a full-fledged numerical algorithm incorporating the conditions provided in this paper along with convergence analysis and experimental results, as well as an extension of the results of the present paper to the case $\ell_1 - \ell_0$, are left as future studies. In particular, the problem where the Huber loss is replaced by the ℓ_1 norm imposes a linear structure which opens new possibilities of exploration.

Appendix

Definition 8.1 Let $\Psi : \mathbb{R}^N \to \mathbb{R}$ be a differentiable function over $\mathbb{R}^N$ and its gradient has a *Lipschitz* constant $L_\Psi > 0$:

$$\|\nabla\Psi(x) - \nabla\Psi(y)\|_2 \leq L_\Psi \|x - y\|_2 \text{ for all } x, y \in \mathbb{R}^N.$$

Proposition 8.1 *Ψ has a Lipschitz constant* $\dfrac{\|A\|_2^2}{\gamma}$, *where* $\|A\|_2 = \sup_{\|x\|_2=1} \dfrac{\|Ax\|_2}{\|x\|_2}$.

Proof Let $x, y \in \mathbb{R}^N$,

$$\begin{aligned}
\|\nabla\Psi(x) - \nabla\Psi(y)\|_2 &= \left\|A^T[\operatorname{clip}(Ax - d) - \operatorname{clip}(Ay - d)]\right\|_2 \\
&\leq \|A\|_2 \|\operatorname{clip}(Ax - d) - \operatorname{clip}(Ay - d)\|_2 \\
&= \|A\|_2 \frac{\|\operatorname{clip}(Ax - d) - \operatorname{clip}(Ay - d)\|_2}{\|A(x - y)\|_2} \|A(x - y)\|_2 \\
&\leq \|A\|_2^2 \frac{\|\operatorname{clip}(Ax - d) - \operatorname{clip}(Ay - d)\|_2}{\|(Ax - d) - (Ay - d)\|_2} \|x - y\|_2 \\
&\leq \frac{\|A\|_2^2}{\gamma} \|x - y\|_2 .
\end{aligned}$$

Last inequality comes from continuity of the gradient. □

Remark 8.1 One should use the Frobenius norm for easy computation, which causes the Lipschitz constant to be $\dfrac{\|A\|_F^2}{\gamma}$. This choice is safe since $\|A\|_2 \leq \|A\|_F$.

Proposition 8.2 *$\mathbb{B}_\gamma$ defined in* Theorem 5.2 *is a dense subset of* $\mathbb{R}^N$.

Proof Let $c : \mathbb{R}^N \to \mathbb{R}^M$ be a linear continuous operator defined as $c(x) = Ax - d$. Since A is full rank with $M < N$, c is a surjection. Let $T = c(\mathbb{B}_\gamma)$ where $c(\mathbb{B}_\gamma)$ is the image of set $\mathbb{B}_\gamma$. Hence, $T = \{v \in \mathbb{R}^M : \forall i \in \mathbb{I}_M, |v[i]| \neq \gamma\}$. Let $\mathcal{O}$ be an arbitrary non-empty open set in $\mathbb{R}^M$. Let $\bar{v} \in \mathcal{O}$ and define an index set as follows:

$\bar{v}_\gamma = \{i \in \mathbb{I}_M : |\bar{v}[i]| = \gamma\}$. Now, there is some positive radius $r_{\bar{v}}$ such that $B_\infty(\bar{v}, r_{\bar{v}})$ stays in $\mathcal{O}$. If we define

$$r^* = \min\left\{\min_{i\notin\bar{v}_\gamma}\{|\bar{v}[i]-\gamma|\}, \min_{i\notin\bar{v}_\gamma}\{|\bar{v}[i]+\gamma|\}, r_{\bar{v}}, \gamma\right\},$$

we have $r^* > 0$ and $B_\infty(\bar{v}, r_{\bar{v}})$ stays in $\mathcal{O}$. Then, for any $0 < \delta < r^*$, $v^* = \bar{v} + \delta\mathbb{1}_M$ belongs to $\mathcal{O}$ and T at the same time. Hence, T is a dense set. T is the image of a continuous surjection; therefore, $\mathbb{B}_\gamma$ is a dense set too. □

Remark 8.2 Previous result shows that one can construct a sequence of vectors from $\mathbb{B}_\gamma$ converging to any other vector in $\mathbb{R}^N$. This may be useful for deriving algorithms using second-order methods since the second derivative exists only for vectors in $\mathbb{B}_\gamma$.

For all numerical experiments reported in this paper, we use the following equivalent formulation for (HR)

$$\begin{aligned}
\text{minimize}\ & \frac{1}{2\gamma}\textstyle\sum_{i\in\mathbb{I}_M} p_i^2 + \sum_{i\in\mathbb{I}_M}\left(q_i - \frac{\gamma}{2}\right)\\
\text{subject to}\ & -p - q \le Au - d \le p + q,\\
& 0 \le p \le \gamma\mathbb{1}_M,\\
& 0 \le q.
\end{aligned} \tag{QHR2}$$

Proposition 8.3 *(Equivalent Characterization for (HR), [35]) Any optimal solution to the quadratic program (QHR2) is a minimizer of Ψ, and conversely.*

This alternative eliminates the need to work with piecewise functions and provides an easier computation tool. We used the following algorithm for the numerical examples reported in the paper:

Algorithm 1: GMSRH (Global Minima of Sparse Regularized Huber Regression)

Result: $\hat{u}$ an optimal solution.
Hyperparameters: β, γ;
initialization $A \in \mathbb{R}^{M\times N}$, $d \in \mathbb{R}^M$ and $\hat{U} = \{\mathbf{0}\}$;
for $\omega \subseteq \mathbb{I}_M$ *with* $|\omega| \le N$ **do**
  $r \leftarrow |\omega|$;
  if $\operatorname{rank}(A_\omega) = r$ **then**
    $u_\omega^* \leftarrow \arg\min_{u\in\mathbb{R}^{|\omega|}} \Psi_\omega(u)$;
    $\hat{U} \leftarrow \hat{U} \cup \{Z_\omega(u_\omega^*)\}$;
  end
end
$\hat{u} \leftarrow \arg\min_{u\in\hat{U}} \Psi(u) + \|u\|_0$;

References

1. Zoubir, A., Koivunnen, V., Ollila, E., Muma, M.: Robust Statistics for Signal Processing. Cambridge University Press, Cambridge (2018)
2. Huber, P.J.: Robust Statistics. Wiley, New York (1981)
3. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning, 2nd edn. Springer, New York (2008)
4. El, d'Aspremont A., Ghaoui, L.: Testing the nullspace property using semidefinite programming. Math. Program. **127**(1), 123–144 (2011)
5. Bryan, K., Leise, T.: Making do with less: an introduction to compressed sensing. SIAM Rev. **55**, 547–566 (2013)
6. Candès, E.J., Romberg, J., Tao, T.: Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE Trans. Inf. Theor. **52**(2), 489–509 (2006)
7. Candès, E.J., Tao, T.: Decoding by linear programming. IEEE Trans. Inf. Theory **52**, 4203–4215 (2006)
8. Chen, S., Donoho, D., Saunders, M.: Atomic decomposition by basis pursuit. SIAM J. Sci. Comput. **20**, 33–61 (1998)
9. Chrétien, S.: An alternating ℓ_1 approach to the compressed sensing problem. IEEE Signal Proc. Lett. **17**, 181–184 (2010)
10. Donoho, D.L.: Compressed sensing. IEEE Trans. Inf. Theory **51**, 1289–1306 (2005)
11. Donoho, D.L.: For most large underdetermined systems of linear equations the minimal ℓ_1-norm solution is also the sparsest solution. Comm. Pure Appl. Math. **59**, 797–829 (2006)
12. Elad, M.: Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing. Springer, New York (2010)
13. Foucart, S., Rauhut, H.: A Mathematical Introduction to Compressive Sensing. Springer, New York (2013)
14. Tropp, J.A.: Recovery of short, complex linear combinations via ℓ_1 minimization. IEEE Trans. Inf. Theor. **51**(4), 1568–1570 (2005)
15. Fuchs, J.J.: On sparse representations in arbitrary redundant bases. IEEE Trans. Inf. Theor. **50**(6), 1341–1344 (2004)
16. Nikolova, M.: Description of the minimizers of least squares regularized with ℓ_0-norm. Uniqueness of the global minimizer. SIAM J. Imaging Sci. **6**(2), 904–937 (2013)
17. Beck, A., Hallak, N.: Proximal mapping for symmetric penalty and sparsity. SIAM J. Optim. **28**(1), 496–527 (2018)
18. Tropp, J.A.: Just relax: convex programming methods for identifying sparse signals in noise. IEEE Trans. Inf. Theor. **52**, 1030–1051 (2006)
19. Chancelier, J.-Ph., De Lara, M.: Hidden convexity in the l0 pseudonorm. HAL (2019)
20. Chancelier, J.-Ph., De Lara, M.: Lower bound convex programs for exact sparse optimization. HAL (2019)
21. Chancelier, J.-Ph., De Lara, M.: A suitable conjugacy for the l0 sseudonorm. HAL (2019)
22. Soubies, E., Blanc-Féraud, L., Aubert, G.: New insights on the ℓ_2-ℓ_0 minimization problem. J. Math. Imaging Vis. **62**, 808–824 (2020)
23. Lanza, A., Morigi, S., Selesnick, I.W., Sgallari, F.: Sparsity-inducing non-convex, non-separable regularization for convex image processing. SIAM J. Imaging Sci. **12**(2), 1099–1134 (2019)
24. Selesnick, I.: Sparse regularization via convex analysis. IEEE Trans. Signal Process. **65**(17), 4481–4494 (2017)
25. Wang, S., Chen, X., Dai, W., Selesnick, I.W., Cai, G.: Vector minimax concave penalty for sparse representation. Digital Signal Process. **83**, 165–179 (2018)
26. Wang, J., Zhang, F., Huang, J., Wang, W., Yuan, C.: A non-convex penalty function with integral convolution approximation for compressed sensing. Signal Process. **158**, 116–128 (2019)
27. Chen, K., Lv, Q., Lu, Y., Dou, Y.: Robust regularized extreme learning machine for regression using iteratively reweighted least squares. Neurocomputing **230**, 345–358 (2017)
28. Carrillo, R.E., Ramirez, A.B., Arce, G.R., Barner, K.E., Sadler, B.M.: Robust compressive sensing of sparse signals: a review. EURASIP J. Adv. Signal Process. **2016**, 108 (2016)
29. Grechuk, B., Zabarankin, M.: Regression analysis: likelihood, error and entropy. Math. Program. **174**, 145–166 (2019)
30. Li, W., Swetits, J.J.: The linear ℓ_1 estimator and the Huber M-estimator. SIAM J. Optim. **8**, 457–475 (1998)

31. Madsen, K., Nielsen, H.B.: A finite smoothing algorithm for linear ℓ_1 estimation. SIAM J. Optim. **3**, 223–235 (1993)
32. Madsen, K., Nielsen, H.B., Pınar, M.Ç.: New characterizations of ℓ_1 solutions to overdetermined systems of linear equations. Oper. Res. Lett. **16**, 159–166 (1994)
33. Auslender, A., Teboulle, M.: Asymptotic Cones and Functions in Optimization and Variational Inequalities. Springer, New York (2003)
34. Blumensath, T., Davies, M.E.: Iterative thresholding for sparse approximations. J. Fourier Anal. Appl. **14**, 629–654 (2008)
35. Pınar, M.Ç.: Linear Huber M-estimator under ellipsoidal data uncertainty. BIT **42**(4), 856–866 (2002)