Exact recovery of partially sparse vectors

Christoph Brauer
Dirk A. Lorenz
Institute of Analysis and Algebra, TU Braunschweig
38092 Braunschweig, Germany
Email: {ch.brauer, d.lorenz}@tu-braunschweig.de

Andreas M. Tillmann
Chair of Operations Research & Visual Computing Institute
RWTH Aachen University
52072 Aachen, Germany
Email: tillmann@or.rwth-aachen.de

***Abstract*—We investigate the recovery of partially sparse signals with regularization of the signal part for which no sparsity prior is known, to overcome the complete loss of recoverability by previous approaches without such modifications. For certain mixed ℓ^1-ℓ^2-norms and Luxemburg norms, we present optimality conditions for recovery and some numerical experiments.**

I. Introduction

We consider the problem of reconstructing a vector $\boldsymbol{x}$ from underdetermined measurements $\boldsymbol{Ax} = \boldsymbol{b}$ when we only know that a part of $\boldsymbol{x}$ is sparse. To fix notation, let $\boldsymbol{x} \in \mathbb{R}^n$ and $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ and assume that $\boldsymbol{x} = [\boldsymbol{x}_1^\top\, \boldsymbol{x}_2^\top]^\top$ with $\boldsymbol{x}_1 \in \mathbb{R}^s$ and $\boldsymbol{x}_2 \in \mathbb{R}^{n-s}$, where $\boldsymbol{x}_1$ is supposed to be sparse and $\boldsymbol{x}_2$ is expected to be dense. Problems of this type may appear in the context of compressed sensing or sparse recovery when the sparsity assumption fails for a part of the variable $\boldsymbol{x}$. Indeed, the problem of partial sparse recovery appeared previously in the literature, see, e.g., [1], [4], [5], [6]. There, the focus is on the problem of minimizing the ℓ^1-norm of $\boldsymbol{x}_1$ under the constraint $\boldsymbol{A}_1\boldsymbol{x}_1 + \boldsymbol{A}_2\boldsymbol{x}_2 = \boldsymbol{b}$, where the partition $\boldsymbol{A} = [\boldsymbol{A}_1\, \boldsymbol{A}_2]$ complies with the partition of $\boldsymbol{x}$ into a sparse and a dense part. The sparsity assumption on $\boldsymbol{x}_1$ is naturally incorporated in terms of the ℓ^1-norm penalty, whereas the objective function includes no further regularization of the dense part $\boldsymbol{x}_2$. Studies of this problem were motivated by the situation when one has partial knowledge of the support of the sought solution.

Note that, whenever $\boldsymbol{A}_2$ has full row rank, this approach cannot recover a nonzero $\boldsymbol{x}_1$, since $\boldsymbol{A}_2\boldsymbol{x}_2 = \boldsymbol{b}$ then always has a solution and thus, choosing $\boldsymbol{x}_1 = \boldsymbol{0}$ is optimal. To overcome this restriction, we propose to include a regularization term for $\boldsymbol{x}_2$ in the objective. Specifically, we consider related approaches

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \; f(\boldsymbol{x}_1, \boldsymbol{x}_2) \quad \text{s.t. } \boldsymbol{A}_1\boldsymbol{x}_1 + \boldsymbol{A}_2\boldsymbol{x}_2 = \boldsymbol{b}, \tag{1}$$

where the function f depends on both the sparse *and* the dense part of $\boldsymbol{x}$ and is either a weighted ℓ^1-ℓ^2-norm or a Luxemburg norm. Both are introduced in the following section before we investigate recovery conditions for the respective versions of (1).

II. Mixed norms

One way to penalize also the dense part $\boldsymbol{x}_2$ is to use a weighted ℓ^1-ℓ^2-norm

$$\|\boldsymbol{x}\|_{M,\alpha} := \|\boldsymbol{x}_1\|_1 + \alpha\|\boldsymbol{x}_2\|_2, \tag{2}$$

where $\alpha > 0$ is a tuning parameter. This is a special case of the group lasso [3] where each entry in $\boldsymbol{x}_1$ is in a single group and the full vector $\boldsymbol{x}_2$ forms another group. Note that replacing $\|\cdot\|_2$ by its squared version, as one might consider for the benefit of increased smoothness, does not yield a norm due to lacking homogeneity. As a consequence, the recovery problem is not homogeneous, i.e. if we would recover $\boldsymbol{x}$ from some given $\boldsymbol{b}$, we would *not* recover $\lambda\boldsymbol{x}$ from $\lambda\boldsymbol{b}$ for any $\lambda \neq 1$. By contrast, like (2), the related Luxemburg norm

$$\begin{aligned}\|\boldsymbol{x}\|_{L,\beta} &:= \inf\left\{\lambda > 0 : \left\|\tfrac{\boldsymbol{x}_1}{\lambda}\right\|_1 + \beta\left\|\tfrac{\boldsymbol{x}_2}{\lambda}\right\|_2^2 \leq 1\right\} \\ &= \tfrac{\|\boldsymbol{x}_1\|_1}{2} + \sqrt{\tfrac{\|\boldsymbol{x}_1\|_1^2}{4} + \beta\|\boldsymbol{x}_2\|_2^2}\end{aligned} \tag{3}$$

is indeed a norm in the mathematical sense, see [2]. Prior to a valid comparison of (2) and (3) we have to discuss how the parameter β should be chosen in relation to α. To that end, we first observe that both norms are equal in case $\boldsymbol{x}_1 = \boldsymbol{0}$ and $\beta = \alpha^2$. Based on that idea, we can further deduce that $\beta = \alpha^2$ also minimizes the ratio of the largest and smallest values of $\|\cdot\|_{L,\beta}$ on the unit sphere of $\|\cdot\|_{M,\alpha}$. Moreover, this choice leads to the desirable property that the unit spheres of both considered norms intersect specific coordinate axes at the same points, as illustrated in Fig. 1. Therefore, we use the setting $\beta = \alpha^2$ in the following section in order to compare the respective recovery conditions.

III. Recovery conditions

Similarly to the pure ℓ^1 case, recovery conditions for the mixed and the Luxemburg norm are based on optimality conditions for the respective minimization problems. We have the following result:

Theorem 1: A point $\boldsymbol{x}^*$ with $\boldsymbol{x}_2^* \neq \boldsymbol{0}$ is a solution of

$$\min_{\boldsymbol{Ax}=\boldsymbol{Ax}^*} \|\boldsymbol{x}\|_{M,\alpha}$$

if and only if there exists $\boldsymbol{w}^* \in \mathbb{R}^m$ such that

$$(\boldsymbol{A}_1)^\top\boldsymbol{w}^* \in \partial\|\boldsymbol{x}_1^*\|_1 \tag{4}$$

$$\text{and} \quad (\boldsymbol{A}_2)^\top\boldsymbol{w}^* = \tfrac{\alpha}{\|\boldsymbol{x}_2^*\|_2} \cdot \boldsymbol{x}_2^*. \tag{5}$$

For the Luxemburg norm, we have a similar result:

Theorem 2: A point $\boldsymbol{x}^* \neq \boldsymbol{0}$ is a solution of

$$\min_{\boldsymbol{Ax}=\boldsymbol{Ax}^*} \|\boldsymbol{x}\|_{L,\alpha^2}$$

if and only if there exists $\boldsymbol{w}^* \in \mathbb{R}^m$ such that

$$(\boldsymbol{A}_1)^\top\boldsymbol{w}^* \in \left(\frac{1}{2} + \frac{\|\boldsymbol{x}_1^*\|_1}{4\sqrt{\frac{1}{4}\|\boldsymbol{x}_1^*\|_1^2 + \alpha^2\|\boldsymbol{x}_2^*\|_2^2}}\right)\partial\|\boldsymbol{x}_1^*\|_1 \tag{6}$$

$$\text{and} \quad (\boldsymbol{A}_2)^\top\boldsymbol{w}^* = \left(\frac{\alpha^2}{\sqrt{\frac{1}{4}\|\boldsymbol{x}_1^*\|_1^2 + \alpha^2\|\boldsymbol{x}_2^*\|_2^2}}\right) \cdot \boldsymbol{x}_2^*. \tag{7}$$

One difference between these theorems becomes apparent in the case $s = 1$: Then, the mixed norm is equal to the ℓ^1-norm and hence, exact recovery depends on the sign of $\boldsymbol{x}$ only, but the same is not true for the Luxemburg norm. Numerical experiments with $s = n-1$ confirm this: For some random $\boldsymbol{A}_1 \in \mathbb{R}^{60\times 200}$, $\boldsymbol{A}_2 \in \mathbb{R}^{60}$ being the vector of all ones, some sparse $\boldsymbol{x}_1^* \in \mathbb{R}^{200}$ (with 15 nozero entries) and some α, the mixed norm $\|\cdot\|_{M,\alpha}$ recovers $\boldsymbol{x}_2 = 0$ regardless of the value of $\boldsymbol{x}_2^*$, while the Luxemburg norm $\|\cdot\|_{L,\alpha^2}$ does recover the whole $\boldsymbol{x}$ exactly for a broad range of $\boldsymbol{x}_2 \neq 0$, cf. Figures 4, 5, and 6.

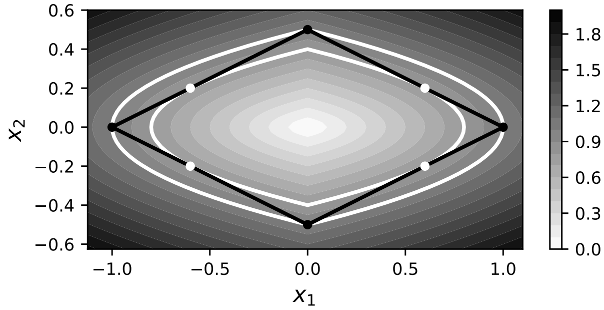

Fig. 1. A comparison of the unit spheres of $\|\cdot\|_{M,\alpha}$ (black sphere) and $\|\cdot\|_{L,\alpha^2}$ (outer white sphere) for $\alpha = 2$ shows that both have identical intersection points with the coordinate axes (black dots). If we use the parameter $\beta = \alpha^2$ in the Luxemburg norm, then the radius of the largest $\|\cdot\|_{L,\beta}$ sphere intersecting the $\|\cdot\|_{M,\alpha}$ unit sphere is always 1 and the radius of the smallest such sphere (inner white sphere) is always $4/5$. The ratio $5/4$ is minimal among all possible choices for β. Both unit spheres intersect at all locations where either $\|\boldsymbol{x}_1\|_1 = 0$ or $\|\boldsymbol{x}_2\|_2 = 0$. The unit sphere of $\|\cdot\|_{M,\alpha}$ and the $4/5$-sphere of $\|\cdot\|_{L,\alpha^2}$ intersect at all points where $\|\boldsymbol{x}_2\|_2 = 2/(5\alpha)$ (white dots).

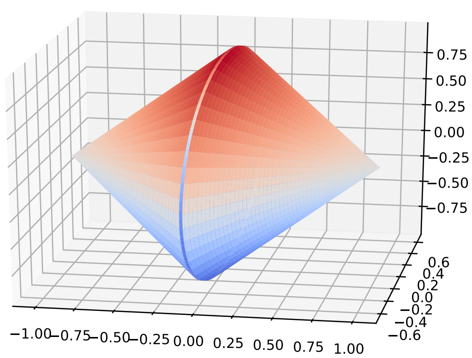

Fig. 2. The unit ball of the mixed norm $\|\cdot\|_{M,1}$ in three dimensions where $\boldsymbol{x}_1 = x_1$ and $\boldsymbol{x}_2 = [x_2\ x_3]$

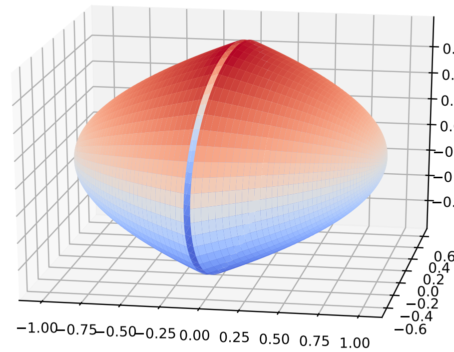

Fig. 3. The unit ball of the Luxemburg norm $\|\cdot\|_{L,1}$ in three dimensions where $\boldsymbol{x}_1 = x_1$ and $\boldsymbol{x}_2 = [x_2\ x_3]$

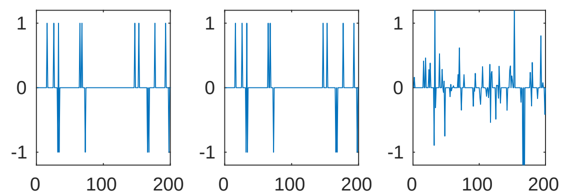

Fig. 4. Reconstruction from $\boldsymbol{A}_1\boldsymbol{x}_1 + \boldsymbol{A}_2 t = b$ with gaussian $\boldsymbol{A}_1 \in \mathbb{R}^{60\times 200}$, $\boldsymbol{A}_2 \in \mathbb{R}^{60}$ vector of all ones, some sparse $\boldsymbol{x}_1^*$ (left plot), and $t = 2$. Middle: recovered $\boldsymbol{x}_1$ by the Luxemburg norm $\|\cdot\|_{L,\alpha^2}$, right: recovered $\boldsymbol{x}_1$ by the mixed norm $\|\cdot\|_{M,\alpha}$ with $\alpha = 4$.

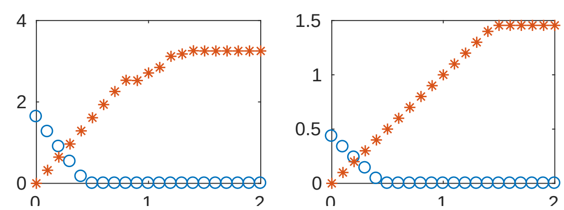

Fig. 5. Recovery error in the setting of Figure 4: Left: error $\|\boldsymbol{x}_1^* - \boldsymbol{x}_1\|_2$ recovered with the Luxemburg norm (blue circles) and the mixed norm (red stars) as a function of t^*. Right: error $|t^* - t|$ as a function of t^* (same color coding). We conclude that the Luxemburg norms recovers $\boldsymbol{x}_1^*$ and t^* exactly for t^* large enough, the mixed norm for $t^* = 0$ only.

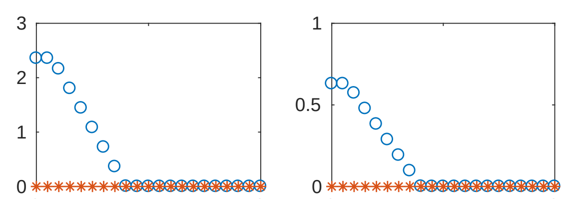

Fig. 6. Same as Figure 5, but with $\alpha = 3$: Left: error $\|\boldsymbol{x}_1^* - \boldsymbol{x}_1\|_2$ recovered with the Luxemburg norm (blue circles) and the mixed norm (red stars) as a function of t^*. Right: error $|t^* - t|$ as a function of t^* (same color coding). We conclude that the Luxemburg recovers $\boldsymbol{x}_1^*$ and t^* exactly for t^* large enough; the mixed norm for all t^*

References

[1] A. S. Bandeira and K. Scheinberg and L. N. Vicente, *On Partial Sparse Recovery*, arXiv:1304.2809 [cs.IT], 2013.

[2] Simon, Barry. *Convexity: An analytic viewpoint*. Vol. 187. Cambridge University Press, 2011.

[3] M. Yuan and Y. Lin, *Model selection and estimation in regression with grouped variables ,Journal of the Royal Statistical Society, Series B*, 68(1), 49–67, 2007.

[4] C. Herzet and C. Soussen and J. Idier and R. Gribonval, *Exact Recovery Conditions for Sparse Representations With Partial Support Information*, IEEE Trans. Inf. Theory 59(11), pp. 7509–7524, 2013.

[5] F. Ahmad and M. G. Amin, *Partially sparse reconstruction of behind-the-wall scenes*, Proc. SPIE 8365, 2012.

[6] H. Gao and H. Zhang and Z. Li and K. Tu, *Optimality analysis on partial l_1-minimization recovery*, J. Glob. Optim. 70(1), pp. 159–170, 2018.