Baré Alexandre s172388
Wallon Robin s171708

LIÈGE université
Sciences Appliquées

Project : Compressed Sensing

MATH0461 - Introduction to Numerical Optimization

Université de Liege
Faculté des Sciences Appliquées
Année 2020-2021

1 Modelling

1.1 Formulate the problem using the l_0 "norm", which counts the number of non-zero entries in a given input vector. Show that the resulting problem is non-convex.

We have that $r = \Psi * x$ and that $m = \Phi * r$. Thus, $m = \Phi * \Psi * x$.

The optimisation problem should promote sparse solutions x, we have thus

$$\begin{aligned} &\min ||x||_{l0} \\ &\text{s.t. } m = \Phi * \Psi * x \end{aligned}$$

$$\Phi \in \mathbb{R}^{M\times N}, m \in \mathbb{R}^M, \Psi \in \mathbb{R}^{N\times N}, x \in \mathbb{R}^N$$

This problem can be written equivalently as

$$\begin{aligned} &\min |\{x_j \neq 0 | j = 1, ..., N\}| \\ &\text{s.t } m = \Phi * \Psi * x \end{aligned}$$

$$\Phi \in \mathbb{R}^{M\times N}, m \in \mathbb{R}^M, \Psi \in \mathbb{R}^{N\times N}, x \in \mathbb{R}^N$$

where $|.|$ denotes the size of the underlying set

Let $g(z) = ||z||_{l0}$. We can find a counter-example that proves that the problem is non-convex because the objective function $g(z)$ is non-convex. For instance, if we take $N = 2$, $z = (1, 0)$ and $w = (0, 1)$ and for a $\lambda = \frac{1}{2}$,

$$g(\lambda z + (1 - \lambda)w) \nleq \lambda g(z) + (1 - \lambda)g(w)$$

Indeed, we have

$$g(\lambda z + (1 - \lambda)w) = ||\frac{1}{2}(1, 0) + (1 - \frac{1}{2})(0, 1)||_{l0} = ||(\frac{1}{2}, \frac{1}{2})||_{l0} = 2$$

And

$$\lambda g(z) + (1 - \lambda)g(w) = \frac{1}{2}||(1, 0)||_{l0} + (1 - \frac{1}{2})||(0, 1)||_{l0} = \frac{1}{2} * 1 + \frac{1}{2} * 1 = 1$$

As expected, $2 \nleq 1$ which means that the problem is non-convex.

1.2 Formulate the problem using the l_1 norm and show that it can be expressed as a linear program.

$$\begin{aligned} &\min ||x||_{l_1} \\ &\text{s.t. } m = \Phi * \Psi * x \end{aligned}$$

$$\Phi \in \mathbb{R}^{M\times N}, m \in \mathbb{R}^M, \Psi \in \mathbb{R}^{N\times N}, x \in \mathbb{R}^N$$

The problem is equivalent to minimizing the sum of the absolute values of the residuals :

$$\begin{aligned} &\min \textstyle\sum_{j=1}^{N} |x_j| \\ &\text{s.t } m = \Phi * \Psi * x \end{aligned}$$

$$\Phi \in \mathbb{R}^{M\times N}, m \in \mathbb{R}^M, \Psi \in \mathbb{R}^{N\times N}, x \in \mathbb{R}^N$$

where $|.|$ denotes the absolute value

In epigraph form, we obtain

$$\begin{aligned} \min & \textstyle\sum_{j=1}^N t_j \\ \text{s.t. } & x_j \leq t_j \\ & -t_j \leq x_j \\ & m = \Phi * \Psi * x \end{aligned}$$

$$t \in \mathbb{R}_+^N, \Phi \in \mathbb{R}^{M\times N}, m \in \mathbb{R}^M, \Psi \in \mathbb{R}^{N\times N}, x \in \mathbb{R}^N$$

In this form, we have a linear program as the objective function and the constraints are all linear.

1.3 Formulate the problem using the l2 norm and show that it can be expressed as a second- order cone program.

The problem can be expressed as

$$\begin{aligned} \min & ||x||_{l2} \\ \text{s.t. } & m = \Phi * \Psi * x \end{aligned}$$

$$\Phi \in \mathbb{R}^{M\times N}, m \in \mathbb{R}^M, \Psi \in \mathbb{R}^{N\times N}, x \in \mathbb{R}^N$$

It is equivalent to

$$\begin{aligned} \min & \sqrt{\textstyle\sum_{j=1}^N x_j^2} \\ \text{s.t. } & m = \Phi * \Psi * x \end{aligned}$$

$$\Phi \in \mathbb{R}^{M\times N}, m \in \mathbb{R}^M, \Psi \in \mathbb{R}^{N\times N}, x \in \mathbb{R}^N$$

In epigraph form, we obtain

$$\begin{aligned} \min & \, t \\ \text{s.t. } & \sqrt{\textstyle\sum_{j=1}^N x_j^2} \leq t \\ & m = \Phi * \Psi * x \end{aligned}$$

$$\Phi \in \mathbb{R}^{M\times N}, m \in \mathbb{R}^M, \Psi \in \mathbb{R}^{N\times N}, x \in \mathbb{R}^N$$

We note that the inequality constraint represents a quadratic cone. We can thus emphasize it in the problem as

$$\begin{aligned} \min & \, t \\ \text{s.t. } & (t, x) \in \mathbf{K}_2 \\ & m = \Phi * \Psi * x \end{aligned}$$

$$\Phi \in \mathbb{R}^{M\times N}, m \in \mathbb{R}^M, \Psi \in \mathbb{R}^{N\times N}, x \in \mathbb{R}^N$$

We thus have a conic program where the objective function is linear, the first constraint defines a quadratic cone and the second constraint defines affine hyperplanes.

1.4 Provide a closed-form solution to the l2-norm problem. *Hint : use optimality conditions for the primal and dual problems.*

The primal problem is

$$\begin{aligned} &\min t \\ &\text{s.t. } (t,x) \in \mathbf{K}_2 \\ &\quad m = \Phi * \Psi * x \end{aligned}$$

$$\Phi \in \mathbb{R}^{M\times N}, m \in \mathbb{R}^M, \Psi \in \mathbb{R}^{N\times N}, x \in \mathbb{R}^N$$

It can be rewritten as

$$\begin{aligned} &\min c^T w \\ &\text{s.t } (Aw - b) \in \mathbf{K}_2 \\ &\quad m = \Phi * \Psi * x \end{aligned}$$

with

$$\begin{aligned} &w = (t, x_1, ..., x_N)^T \in \mathbb{R}^{N+1} \\ &c = (1, 0, ..., 0)^T \in \mathbb{R}^{N+1} \\ &A = \mathbf{I} \in \mathbb{R}^{(N+1)\times(N+1)} \\ &b = (0, ..., 0)^T \in \mathbb{R}^{N+1} \end{aligned}$$

$$\Phi \in \mathbb{R}^{M\times N}, m \in \mathbb{R}^M, \Psi \in \mathbb{R}^{N\times N}, x \in \mathbb{R}^N$$

The dual problem is

$$\begin{aligned} &\max p^T b \\ &\text{s.t } p^T A = c^T \\ &\quad p \in \mathbf{K}_{2*} \\ &\quad m = \Phi * \Psi * x \end{aligned}$$

$$\Phi \in \mathbb{R}^{M\times N}, m \in \mathbb{R}^M, \Psi \in \mathbb{R}^{N\times N}, x \in \mathbb{R}^N$$

where $p = (p_0, ..., p_N)^T \in \mathbb{R}^{N+1}$
Or equivalently,

$$\begin{aligned} &\max 0 \\ &\text{s.t } p_0 = 1 \\ &\quad p_j = 0 \, for \, j = 1, ..., N \\ &\quad p \in \mathbf{K}_{2*} \\ &\quad m = \Phi * \Psi * x \end{aligned}$$

$$\Phi \in \mathbb{R}^{M\times N}, m \in \mathbb{R}^M, \Psi \in \mathbb{R}^{N\times N}, x \in \mathbb{R}^N$$

Thus, $p^* = (1, 0, ..., 0)^T \in \mathbb{R}^{N+1}$
We have the optimal primal-dual pair (w^*, p^*) if and only if $c^T w^* = (p^*)^T b$, if and only if $p^{*T}(Aw^* - b) = 0$. Therefore,

$$c^T w^* = 0$$

$$p^{*T} \mathbf{I} w^* = 0$$

Thus, $t^* = 0$ and $w^* = (0, x_1^*, ..., x_N^*) \in \mathbb{R}^{N+1}$

The closed-form solution of the l2-norm results in the initial configuration where $x^* = (\Phi * \Psi)^{-1} * m$. We will thus reobtain the initial signal $r^* = \Psi * x^* = \Psi * (\Phi * \Psi)^{-1} * m = \Phi^{-1} * m$. The decompression with the L2-norm was of no use...

1.5 Formulate at least two robust variants of the l1-norm problem, whereby the reconstructed signal may not exactly match the measurements, up to some prespecified tolerance .

By specifying the tolerance as $\epsilon_1 \in \mathbb{R}_+^M$, we can do

$$\begin{aligned} &\min ||x||_{l_1} \\ &\text{s.t. } |\Phi * \Psi * x - m| <= \epsilon_1 \end{aligned}$$

$$\Phi \in \mathbb{R}^{M \times N}, m \in \mathbb{R}^M, \Psi \in \mathbb{R}^{N \times N}, x \in \mathbb{R}^N$$

In epigraph form, we obtain

$$\begin{aligned} &\min \textstyle\sum_{j=1}^{N} t_j \\ &\text{s.t. } x_j \leq t_j \\ &\quad -t_j \leq x_j \\ &\quad \Phi * \Psi * x - m <= \epsilon_1 \\ &\quad -\epsilon_1 <= \Phi * \Psi * x - m \\ &\quad j = 1, ...N \end{aligned}$$

$$t \in \mathbb{R}_+^N, \Phi \in \mathbb{R}^{M \times N}, m \in \mathbb{R}^M, \Psi \in \mathbb{R}^{N \times N}, x \in \mathbb{R}^N$$

This method results in a linear program as all constraints and the objective function are linear.

With here the tolerance $\epsilon_2 \in \mathbb{R}_+$, a second approach consists in using the L2-norm.

$$\begin{aligned} &\min ||x||_{l_1} \\ &\text{s.t. } ||\Phi * \Psi * x - m||_{l2} <= \epsilon_2 \end{aligned}$$

$$\Phi \in \mathbb{R}^{M \times N}, m \in \mathbb{R}^M, \Psi \in \mathbb{R}^{N \times N}, x \in \mathbb{R}^N$$

In epigraph form, we obtain

$$\begin{aligned} &\min \textstyle\sum_{j=1}^{N} t_j \\ &\text{s.t. } x_j \leq t_j \\ &\quad -t_j \leq x_j \\ &\quad (\epsilon_2, \Phi * \Psi * x - m) \in \mathbf{K}_2 \\ &\quad j = 1, ...N \end{aligned}$$

$$t \in \mathbb{R}_+^N, \Phi \in \mathbb{R}^{M \times N}, m \in \mathbb{R}^M, \Psi \in \mathbb{R}^{N \times N}, x \in \mathbb{R}^N$$

This method results in a conic program as there is a second-order cone as last constraint while the objective function and the other constraints are linear.

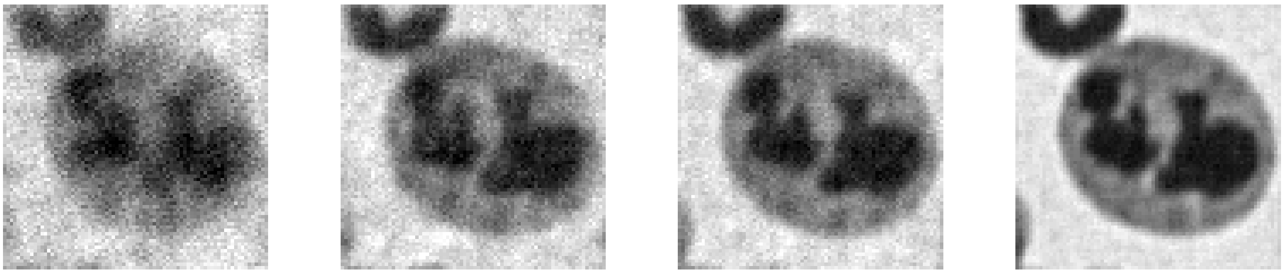
2 Numerical Experiments

2.1 Code up the formulations proposed earlier in Julia JuMP.

See `l1norm.jl`, `l2norm.jl`, `l1normRobust.jl` and `l1normRobust2.jl`

2.2 Solve the l1 and l2-norm problems numerically for the set of uncorrupted measurements. Discuss the performance of each method.

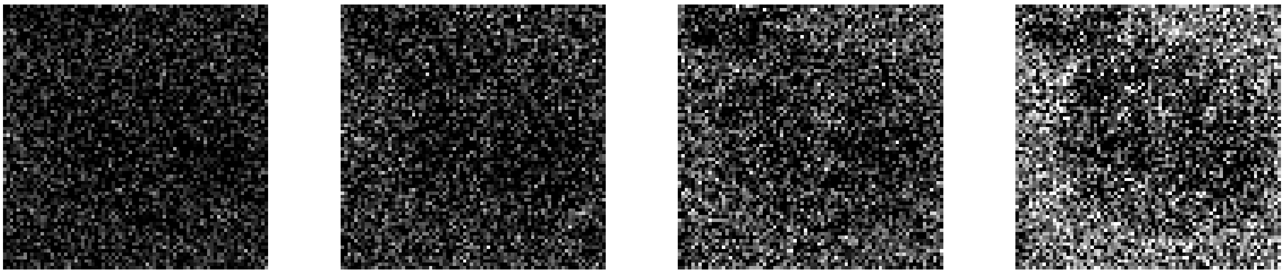
The L1-norm problem can be cast as a linear problem. So, it is solved quite fast and the reconstituted image is not bad. We can still see that all pixels are not exactly corresponding to those of the `GrayscaleCellImage`. But the correspondence becomes better when more measurements are made.



(a) 608 (b) 1014 (c) 1521 (d) 3042

FIGURE 1 – uncorrupted measurements.

The L2-norm however does not yield good results. It is quite fast to compute though the problem is cast as a conic program. Unfortunately, the reconstituted image is unexploitable. In fact, this method does not seem to promote sparse solutions as opposed to the L1-norm.



(a) 608 (b) 1014 (c) 1521 (d) 3042

FIGURE 2 – L2 norm with uncorrupted measurements.

2.3 Provide an interpretation of the dual variables associated with equality constraints in the l1-norm formulation.

The strong duality property is

$$c^T x^* = (p^*)^T b$$

Let there be a change in the right-hand side b.

$$b_i' = b + \Delta b_i e_i$$

with $i = 1, ..., N$ and e_i being the vector whose $i^t h$ entry is 1 and all others are 0.

Updating b to b' will lead to a new primal solution x_2^*. The strong duality property becomes

$$c^T x_2^* = (p^*)^T b_i' = (p*)^T b + \Delta b_i p_i^* = c^T x^* + \Delta b_i p_i^*$$

For $i = 1, ..., N$, the difference between the 2 primal objective values is

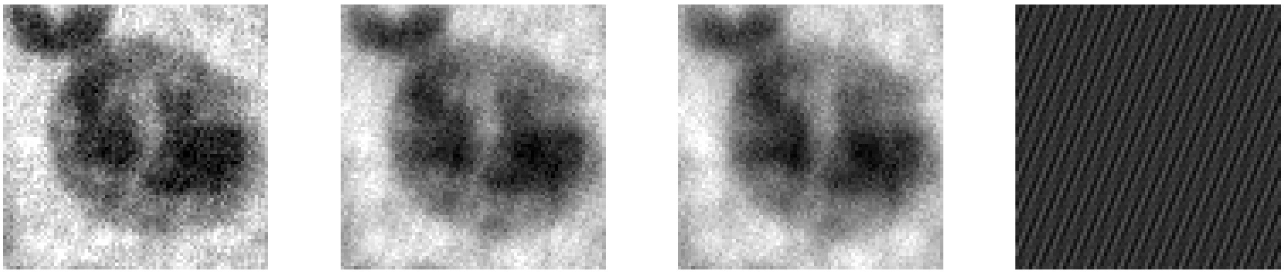
$$\begin{aligned} \Delta c_i &= c^T x_2^* - c^T x^* && (1) \\ &= (c^T x^* + \Delta b_i p_i^*) - c^T x^* && (2) \\ &= \Delta b_i p_i^* && (3) \end{aligned}$$

We thus finally have that $p_i^* = \frac{\Delta c_i}{\Delta b_i}$

In our case, the primal objective represents the sparsification of x (by the minimisation of the L1-norm of x). Thus, the dual variable of the equality constraint can be interpreted the following way : it is the marginal improvement of the sparsification of x as the input measurements become less and less noisy. So, it quantifies the quality of the reconstruction of the image over the variation of the input measurements.

2.4 Solve the l1-norm formulation and its robust variants numerically for the set of noisy measurements. Test your methods for different values of ϵ and discuss their performance.

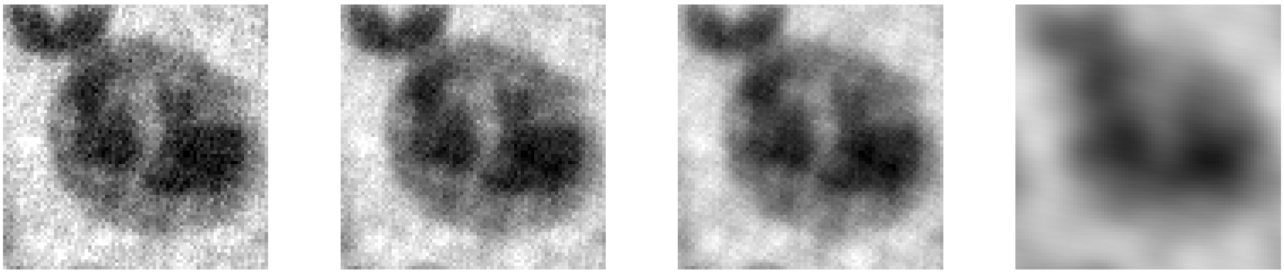
Our first approach yields good results and the computation time is quite acceptable. It consists in directly assuming that the measurement matrix m might not be equal to the measured reconstructed signal $\Phi * \Psi * x$ up to some prespecified tolerance. This tends to uniformize the noisy pixels over the whole image. This method is quite sensitive to the values of the tolerance ϵ_1. For very low ϵ_1, we are left with a very noisy image. For too high ϵ_1, we are left with a uniformization of the image pixels, i.e. a black image. In-between (typically in the range $[0.0001, 0.01]$), we obtain exploitable results. The values of the tolerance are here independent of the size of the measurement matrix.



(a) tolerance = 0.0001 (b) tolerance = 0.0075 (c) tolerance = 0.01 (d) tolerance =0.1

Figure 3 – L1 norm Robust Variant 1 with 1014 noisy measurements

Our second approach yields the best results and the fastest computation time. It consists in assuming that the difference between the measurement matrix m and the measured reconstructed signal $\Phi * \Psi * x$ should be bounded in L2-norm by the prespecified tolerance. It is even more sensitive to the values of the tolerance ϵ_2. For very low ϵ_2, we are left with a very noisy image. For too high ϵ_1, we are left with an image of randomly colored pixels that is completely unexploitable. In-between, we obtain exploitable results. The values of the tolerance should here be adapted to the size of the measurement matrix.



(a) tolerance = 0.01 (b) tolerance = 0.1 (c) tolerance = 0.2 (d) tolerance = 0.4

Figure 4 – L1 norm Robust Variant 2 with 1014 noisy measurements.

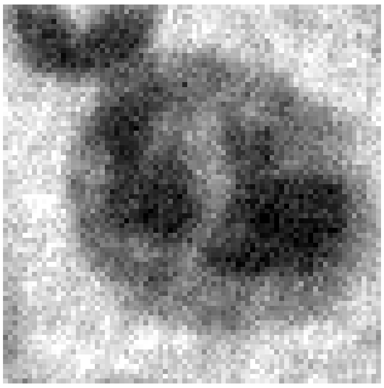


Figure 5 – L1 norm with 1014 noisy measurements