CS427: Mathematics for Data Science **Group 1**

Project 1

Instructor: Dr. Bharath B.N

TAs: Chandan N., Jayanth S., Shruti M., Sawan S. M

Editor: Kushagra Khatwani, Rahul Ankamreddi

Contributions:

Rahul Ankamreddi:

- Problem 1.
- Problem 2.
- Latex Part

Kushagra Khatwani :

- Problem 3
- Problem 4
- Problem 5
- Github Repository

Dadi Swamy Vinay :

- No contribution.

Problem 1

Given optimization problem:

$$\begin{aligned} \min_{s} \quad & ||s||_0 \\ \text{s.t.} \quad & ||y - Cs||_2^2 \end{aligned}$$

To prove that above optimization problem is not convex it will be sufficient enough for us to show that L_0 norm is not convex which will eventually show that objective function is not convex hence the problem is not a convex optimization problem.

Firstly let us see what L_0 norm is:
It is the number of non-zero entries in a vector.

Now back to our proof, consider two vectors $x_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and $x_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\lambda \in [0, 1]$.
If a function is convex then it should satisfy the following inequality

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2) \tag{0.1}$$

For our function, $f(x) = ||x||_0$; So, assume that L_0 norm is convex. Then, it must satisfies (0.1).

$$||\lambda x_1 + (1-\lambda)x_2||_0 \leq \lambda||x_1||_0 + (1-\lambda)||x_2||_0 \tag{0.2}$$

$$||\lambda \begin{bmatrix} 0 \\ 1 \end{bmatrix} + (1-\lambda) \begin{bmatrix} 1 \\ 0 \end{bmatrix} ||_0 \leq \lambda|| \begin{bmatrix} 0 \\ 1 \end{bmatrix} ||_0 + (1-\lambda)|| \begin{bmatrix} 1 \\ 0 \end{bmatrix} ||_0 \tag{0.3}$$

Let $\lambda = \frac{1}{2}$ then substituting back in (0.3) we get,

$$||0.5 \begin{bmatrix} 0 \\ 1 \end{bmatrix} + 0.5 \begin{bmatrix} 1 \\ 0 \end{bmatrix} ||_0 \leq 0.5|| \begin{bmatrix} 0 \\ 1 \end{bmatrix} ||_0 + 0.5|| \begin{bmatrix} 1 \\ 0 \end{bmatrix} ||_0$$
$$2 \leq 1$$

Above result is a contradiction as $2 \geq 1$ hence our initial assumption that L_0 norm is convex is wrong.
As our objective function is not convex so the optimization problem given below is not a convex optimization.

$$\begin{aligned} \min_s \quad & ||s||_0 \\ \text{s.t.} \quad & ||y - Cs||_2^2 \end{aligned}$$

Hence Proved

Problem 2

Given the optimization problem:

$$\begin{aligned} \min_s \quad & ||s||_1 \\ \text{s.t.} \quad & ||y - Cs||_2^2 \end{aligned}$$

We have to show that following optimization problem is convex relaxation of non-convex optimization given in Problem 1.
Again reiterating the fact that if a function is convex then it must satisfy (0.1).
Using property of norms that are:

1. For x_1 and $x_2 \in V$ where V is a vector space then $||x_1 + x_2||_1 \leq ||x_1|| + ||x_2||_1$.
2. Also for $\theta \in \mathbb{R}, \theta||x_1||_1 = ||\theta x_1||_1$

Then we can apply the same for convex property for $\lambda \in [0, 1]$ and show that:

$$||\lambda x_1 + (1-\lambda)x_2||_1 \leq \lambda||x_1||_1 + (1-\lambda)||x_2||_1$$

So, from the above equation L_1 norm is a convex function and also the constraint for the given problem is affine so the optimization problem is a convex relaxation of the problem in previous part.

$$\begin{aligned} \min_s \quad & ||s||_1 \\ \text{s.t.} \quad & ||y - Cs||_2^2 \end{aligned}$$

Hence, L_1 norm is convex relaxation of problem with L_0 norm.

Dual Problem:

Now lets form the lagrangian for the given optimization problem.
Before making the lagrangian lets write the optimization problem in a different form:

$$\begin{aligned} \min_s \quad & ||s||_1 \\ \text{s.t.} \quad & y - Cs = 0 \end{aligned}$$

Now,

$$\mathcal{L}(s, V) = ||s||_1 + V^T(y - Cs) \tag{0.4}$$

Here, $y - Cs$ is a collection of equality constraints and V is a collection of lagrangian equality multipliers μ.

Dual Problem:

$$g(V) = \inf_s(||s||_1 + V^T(y - Cs)) \tag{0.5}$$

$$g(V) = \inf_s(||s||_1 + (C^T V)^T + V^T y) \tag{0.6}$$

Lets now veiw our dual $g(V)$ in terms of two cases which are:

1. If $C^T V > 1$ (component wise each element should follow this inequality)
   Here if i make $s \to \infty$ then $-(C^T V)^T \to -\infty$ so,
   $g(V) \to -\infty$
2. if $C^T V \leq 1$ (component wise each element should follow this inequality)
   We know, $||C^T V||_\infty = \max\{\text{all components}\}$ Above follows by the definition of infinity norm.

   For two vectors x and y:
   $$x^T y \leq ||x||_1 ||y||_\infty \tag{0.7}$$
   as,
   $$\begin{aligned} x_1 y_1 \leq& x_1 ||y||_\infty \\ x_2 y_2 \leq& x_2 ||y||_\infty \\ &\vdots \\ x_n y_n \leq& x_n ||y||_\infty \end{aligned}$$
   Adding all the above inequalities results in (0.7). Using (0.7) we get:
   $$(C^T V)^T s \leq ||s||_1 ||C^T V||_\infty \tag{0.8}$$
   $$(C^T V)^T s \leq ||s||_1 \quad (C^T V \leq 1 \quad \text{from our assumption}) \tag{0.9}$$
   so, $||s||_1 - (C^T V)^T s > 0$ thus $g(V) = V^T y$ by substitution in (0.6).

Now the dual problem can be defined as follows:

$$g(v) = \begin{cases} -\infty & , C^T V > 1 \\ V^T y & , C^T V \leq 1 \end{cases}$$

Now lets solve the dual problem obtained above, keeping in mind that the dual $g(V)$i is always a concave function we can frame the optimization problem as follows:

$$\begin{aligned} \max_V \quad & y^T V \\ \text{s.t.} \quad & ||C^T V||_\infty \leq 1 \end{aligned}$$

The above optimization problem is a convex as objective function is concave and the constraints are convex too.

Lets first state slater's condition,

Slater's condition:

There exists an $x \in relintD$ such that,

$$f_i(x) < 0, i = 1, 2, \ldots, m, Ax = b$$

Such a point is known as strictly feasible as their is strong inequality and Slater's condition states that if this holds then the problem has strong duality i.e duality gap is zero.

Duality Gap:

Lets assume that our duality gap is zero i.e strong duality exists. Then we can find a solution to the primal using the dual. If the solution is **unique** for the primal problem then we can be sure that strong duality exists.
Sticking to our assumption let V^* be the solution to the dual and s^* be the solution to the primal.
Then,

$$
\begin{aligned}
y^T V^* &= ||s^*||_1 \quad \text{(Strong Duality)} && (0.10)\\
(Cs^*)^T V^* &= ||s^*||_1 && (0.11)\\
(s^*)^T (C^T V^*) &= ||s^*||_1 && (0.12)\\
\sum_{i=1}^{n} (C^T V^*)[i] s^*[i] &= ||s^*||_1 && (0.13)
\end{aligned}
$$

Now we know if (0.13) follows then strong duality will exist.
Here,

$$\sum_{i=1}^{n} (C^T V^*)[i] s^*[i] \leq ||s^*||_1$$

as for the dual we had the constraint $(C^T V^*) \leq 1$.
Equality will hold when,

$$(C^T V^*) = sign(s^*)$$

That is,

$$sign(s^*) = \begin{cases} +1 & , s > 0 \\ -1 & , s < 0 \\ 0 & , s = 0 \end{cases}$$

The only thing left to prove now is that s^* will be unique for our primal solution. In general for a primal problem the optimal value p^* and dual optimal value d^* follow,

$$d^* \leq p^* \tag{0.14}$$

From the eqn (0.14), for any feasible s_0^* and a particular solution V^* to dual, we have
Here, $y(V^*)^T \leq ||s_0^*||_1$

$$
\begin{aligned}
||s_0^*||_1 &\leq y(V^*)^T \\
&= (C^T s^*)^T V^* \\
&= s^* (C^T V^*) \\
&= ||s^*||_1 \quad \text{(from(0.13))}
\end{aligned}
$$

This shows that s_0^* is unique and equal to s^* which is obtained from the dual problem.
So, we have strong-duality$\Rightarrow$ unique primal solution.

Hence, strong duality holds that is duality gap = 0.

Github Repository for the codes of the Project:

Click the link to go to the repository: Click Here
NOTE: Go through the README first to get insight about different files used for the project.

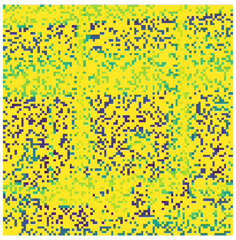
Problem 3 and 4

In problem 3 and 4 we were already given data that includes y,A_inverse,c matrices.We had to use a convex solver and solve the optimization problem obtained above which is as follows:

$$\begin{aligned} \min_{s} \quad & ||s||_1 \\ \text{s.t.} \quad & y - Cs = 0 \end{aligned}$$

While browsing the net we came across two types of solvers Lasso Regularizer which can be used to solve L1 norm optimization problems.Another was CVXPY which was made by Stanford.
We selected CVXPY as it had range of solvers.We tried using **CVXOPT**,**OSQP**,**SCS**.Finally we settled down on **OSQP** as it took the least time and gave comparable results to other solvers.Other solvers took a lot of time to reach similar optimal results.Once we got our s that is sparse representation matrix then we used the inverse cosine transform which was given in **Data for Assignment.py** file to calculate the original image stacked columns.
The results we got our summarised below.

Comparison:



(a) Corrupted Image

(b) Reconstructed

Comparison between corrupted image provided and reconstructed image.

Problem 5

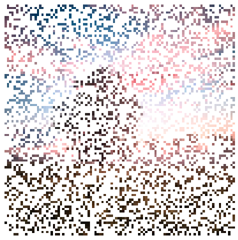
In problem 5 we had to take our own RGB image,preferably 100 $\times$ 100 pixels in order to save time and avoid the risk of program crashing.We tried the program for 400$\times$400 image but it took a long time to compute the data to be used.So,we finally rescaled our image to 100 $\times$ 100 pixels.
Again here we had to use convex optimizer after generating data to obtain final reconstructed image using corrupted image.
Our approach was to separate all three components and then use combine later after applying convex optimization to each component.(explained in github readme)
We tried the code for two corruption values:

1. 70%
2. 40%

We have summarised our results on next page:

Comparison when corruption was 70%:

(a) Original Image



(b) Corrupted Image

(c) Reconstructed Image

Comparison between original,corrupted image and reconstructed image.

Comparison when corruption was 40%:

(a) Original Image

(a) Corrupted Image

(b) Reconstructed Image

Comparison between original,corrupted image and reconstructed image.

We can see from the observations how our convex optimizer works for different values of corruption.For 70% the reconstruction gives us image with some pixels blurred out while for 40% the pixels got sharper and have values closer to actual value.

For understanding how we did the reconstruction and our logic please visit the github repository link mentioned above.