# Optimization

## Assignment 1

Yonatan Simson, 015764921

Amit Kristal, 025602624

### Q1

Inserting the midpoint rule (2):

Into with eq. (1)

To parameterize the line :

Where and

Where is derived from

### Q2

The relationship between and is linear:

And is calculated using the diagram in figure (1) and the equation (\*) from the previous question in the included matlab script.

### Q3

The derivative defined in the question is the same as matlabs [Gx, Gy] = gradient(X) function.

Calculating matrices Dx, Dy is feasible with sparse matrices and is done using the matlab function [Dx, Dy] = CreateDerivativeOperators(X\_rows, X\_cols)

and for the 3D case

[Dx, Dy, Dz] = CreateDerivativeOperators3D(rows, cols, dim);

### Q4

In the following figures we will display the images X1, X2, X3 and the magnitude of their gradients.





****

### Q5

The number of unknown parameters is 5x5=25. The number of observations is 8. The rank of A is 8.

## Q6

Let's take the equation

(\*)

From the KKT conditions we know that the Lagrangian for this is:

This is the dual form. If we set then eq. (\*) is equivalent to

(3)

Likewise by taking the following

(\*\*)

Using KKT conditions the Lagrangian is:

We can set and and then by minimizing the Lagrangian we get the equivalent of eq. (3)

### Q7

TBC

### Q8

A simple closed form solution for eq. (3) can be developed by expressing the equation as

This equation can be minimized by the close form equation:

### Q9

We will use the equivalent form of:

,

So instead we will minimize:

Where we will define and

Which will give us the same expression as in class:

### Q10

The algorithm given in class went as follows:

Input: function of f, initial point

Output: the local minimizer of f

Init:

For k = 1,… until convergence

End

The version modified for our problem would be:

Init:

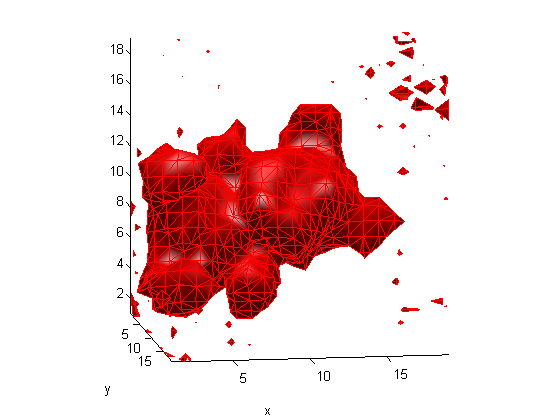
For k = 1,… until convergence

End

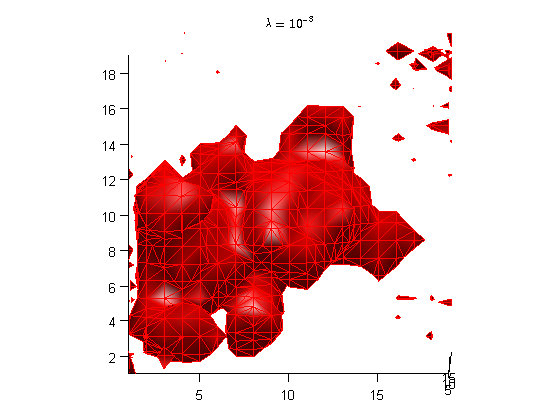
### Q11

When using only CG on the small problem we get a teddy bear that also includes noise:

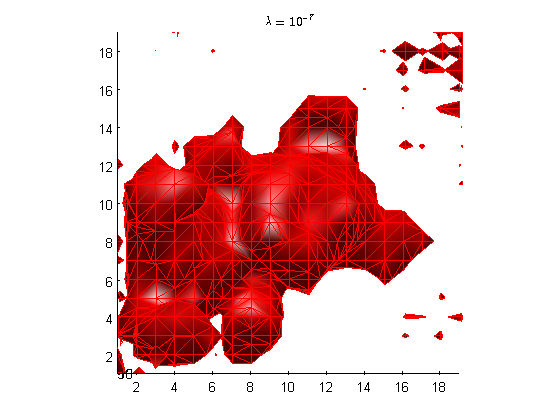
For :



For :



For :



As seen in the figure different values of lambda are not capable of removing the extra noise. For this we need better regularization as will be shown in the next exercises.

### Q12

L1 regularization is convex but not strictly convex. The reason for this is the square shape of the unit ball. If both x and y are on the boundary the line going from x to y may touch the boundary in more places than just x and y. We won't have this problem in a L2 ball.

### Q13

From the figures below we can see that f1 favors the L1 norm over L2 norm and f2 favors the L2 norm over the L1 norm:



### Q14

a)

Prove that

Proof:

For a regularization matrix and unknown vector

The of can be written as:

Let's begin with a derivative of

Due to chain rule and the derivative of

So it follows that

Note:

QED

b)

Prove that and find the entries of :

Proof:

From the previous section we know that:

Where

In order to accommodate for we can define

QED

c)

First order optimality condition for

Is

Begin by opening up the eq. 6

We get the first order optimality condition:

Or equivalently as

Where

,

From here we can derive an iterative reweighted least squares solution and by fixing W as constant at each step

To summarize

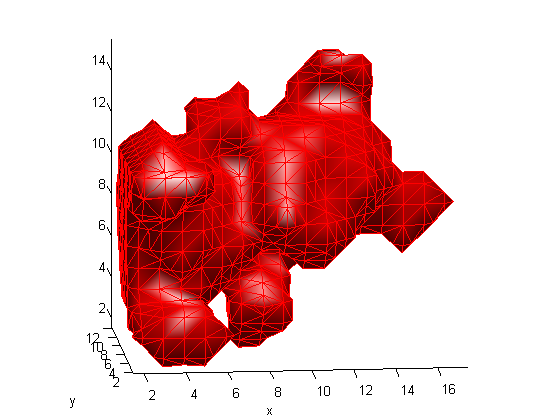
Initialization:

For k = 1,2,…

end

### Q15

The result for the small bag was a teddy bear:



As we can see the L1 norm got rid of noise we could not get rid of using L2 norm regularization.

### Q16

The result for the large bag was a gun:

