## *Note: accuracy of these solutions cannot be guaranteed – feel free to comment or fix any errors you may see. Also, it is recommended to use the desktop version of Word due to the equations (from the web version, File -> Info).*

## Question 1]]#

This question was recycled from the course textbook (parts a and c), and its solution is reproduced below as a result.

Text, letter

Description automatically generated

Text, letter

Description automatically generated

## Question 2

### Part a

*Note that there are multiple correct answers to this part. For instance, the answer below uses greater than or equal to constraints, can be easily changed to a traditional less than or equal to case. Similarly, we can set with an appropriate change to the KKT system, as described in the lectures.*

Write down the conditions:

where

The KKT system is hence

Take appropriate derivatives:

The KKT conditions hence include (noting that )

But notice that . This means that is positive when . This isn’t possible, and hence we get .

Now, for the other condition. Find :

Now, when , . When . Then take :

… which is strictly positive in the desired range (-2 to 0). Hence, there’s a value of which is between -2 and 2, as we know such a value will satisfy .

### Part b

First write down the “trivial” conditions:

* (note: cannot be the case that μ is 0)

As with the first part, then set up the KKT system:

Take appropriate derivatives:

The KKT conditions hence include and . The first expression follows directly by multiplying both sides of the previous equation by 2 – we get and showing that is not possible.

Now, to solve the main problem. Notice that

We have that (where **1** is a vector of size *n*: **1** = ), which simply means that

Notice that , and also that . Using the previous observation, we can say that

This can be used to find μ:

and hence find *c*:

### Part c

*Problem 1*

It’s more of a Lagrangian problem than a KKT problem, but still.

*Problem 2*

Setting up the KKT system,

Taking appropriate derivatives gets us the conditions

Now the case for why this is a clear fail. Linear independence cannot be guaranteed because , which will end up as a 0 for an point *x*! Hence the KKT condition cannot be applied – main issue is because of *h*, nothing else.

## Question 3

### Part a

Split into two functions:

Then, take the second derivative for the positive case:

For non-negative *x*, is positive, and so is for . Hence we get convexity by the second derivative test. The same can be shown for the negative case (note: the reason splitting was done is because is not differentiable at ).

Now consider the case where . Then, if -> it’s not convex. The case also forbids showing concavity.

The last section of this part can be done by individually considering the two cases: we want to show that -*f* is convex across these two intervals. Also note that

which is positive for as long as is not 0.

### Part b

The first section can be shown “directly from the definition”. Given that *f* is convex, the below is true:

But,

The below can easily be shown because it must be the case that , and substituting this into the above equation will get the desired result.

Now, for the converse. We’ll use the given hint. First show that is quasi-convex:

Now, notice that . Note that as with part (a), the two cases for positive and negative *x* should be separated, and the proof is similarly symmetric.

The last part is showing the non-convexity of *f*. This has already been done in part (a).

### Part c

To show: is convex for positive *x* and *y*. For that, we need to take the Hessian:

Take out the common factor:

Now, the goal is to show that is positive semidefinite (call this *M*). Find the characteristic polynomial:

That is, either or , which must be nonnegative for non-negative *x* and *y*. Either way, this shows that the Hessian is a positive semidefinite matrix, and hence convex.

## Question 4

### Part 1

To show the steepest descent part, it suffices to find *d*, which is bounded by . Notice that and hence .

Now, we find the optimal step size. Consider that the goal is to find α such that

is minimised. This is

Then,

This is the desired result – and it can be shown that we have a local minima by taking the second derivative, which is positive.

### Part 2

Get the Hessian:

It can be quickly shown that the eigenvalues are 2 and 200, with the condition number being as a result.

### Part 3

The scaling matrix **D**k is

The descent direction is then

Then, much like part 1, the goal is to find α such that

is minimised. This is

This is the desired result – and it can be shown that we have a local minima by taking the second derivative, which is positive.