**Convex Optimizations – Homework 4**

**Q1. Mirror Descent**

* 1. Calculate the Fenchel conjugate of   
       
     We know that where

Using Cauchy Shwarz inequality we get,

Here, is the p-norm and is its dual norm.

Eqn (1) is true for all x, the right-hand side of the equation is quadratic in terms of

To maximize the RHS of the inequality, we can take the derivative of the right part of the equation w.r.t and equate it to 0.

By substituting the value of (2) in (1), we get,

We know that dual norm follows the property,

where is the dual norm of .

Substituting in ,

This shows that

To show the other inequality, let x be any vector with scaled so that

. Then we have for this x,

which shows that

From we can conclude that,

* 1. Calculate the gradient of

We know that where

Differentiating w.r.t ,

Converting the above equation in vector form

Where is a matrix that has terms of as its diagonal elements and the rest of the terms are zero.

and is a vector of the form,

* 1. Calculate the gradient of

From know that

Differentiating w.r.t ,

Converting the above equation in vector form

Where is a matrix that has terms of as its diagonal elements and the rest of the terms are zero.

and is a vector of the form,

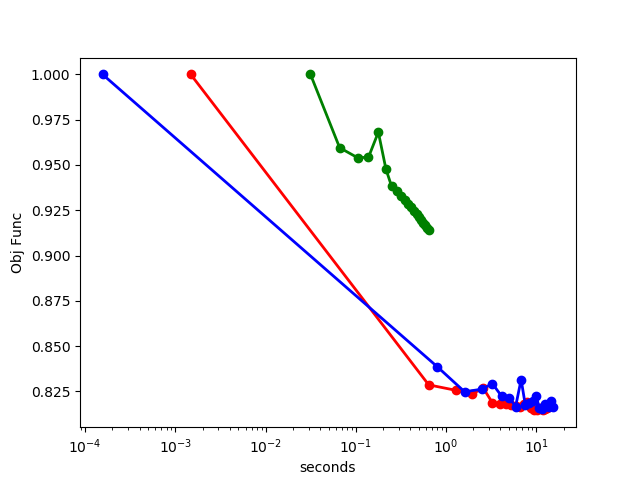
* 1. Write the update rule for mirror descent

The update rule for Mirror Descent can be written as,

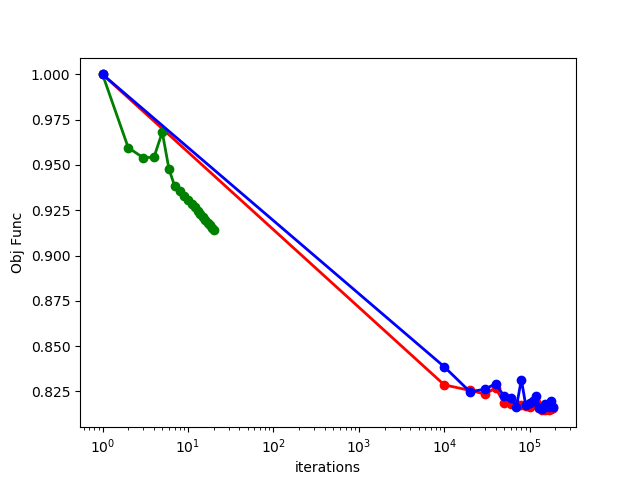
From we get,

From (11) we get,

1. Programming Exercise

Plot 1:

Plot 2:



OUTPUT:

Solution found by stochastic subgradient descent

[[-3.82766340e-01], [-3.81199210e-02], [-1.33802728e-02], [-5.23892968e-01], [ 3.87125284e-04], [ 7.92598865e-01], [-1.63058725e-02], [ 9.88906872e-03], [ 9.18107436e-02], [ 1.54241552e-01], [ 3.11034142e-02], [ 2.14837698e-02], [-4.36862199e-02], [ 1.65145408e-01], [-2.44947948e-02], [ 2.29216069e-02], [-2.31771226e-02], [ 2.58326202e-01], [ 1.33468110e-02], [ 3.95091361e-03], [-3.98102673e-02], [-2.66643596e-01], [ 9.17257424e-01], [ 1.17738321e+00], [ 6.17161348e-01], [-1.51896124e+00], [ 2.42179921e+00], [-3.84489995e+00]]

Objective function 0.8148441242106866

Solution found by stochastic adagrad

[[-2.93993292e-01], [-8.47177834e-02], [-1.56053677e-03], [-4.88253260e-01], [ 2.84568649e-02], [ 7.90104525e-01], [ 6.62801863e-03], [-2.14067903e-03], [ 9.02412127e-02], [ 1.59807984e-01], [-3.86026941e-02], [-2.24010421e-02], [-3.81838505e-02], [ 2.20424331e-01], [ 4.63755831e-02], [-2.89309034e-02], [-5.72662586e-02], [ 2.29611015e-01], [-3.16711860e-02], [ 3.35261481e-03], [-3.68488598e-02], [-2.38980461e-01], [ 8.62666343e-01], [ 1.18104634e+00], [ 6.32419470e-01], [-1.52388920e+00], [ 2.91419386e+00], [-4.42720816e+00]]

Objective function 0.8176521573618533

Solution found by subgradient descent

[[-0.02582542], [-0.01474747], [-0.00618211], [-0.21100792], [ 0.00319675], [ 0.22824581], [-0.00754905], [ 0.00146226], [ 0.14210423], [ 0.12311359], [ 0.01842417], [-0.005359 ], [-0.05565465], [ 0.12481449], [ 0.00288305], [ 0.01901484], [ 0.00618816], [ 0.19496673], [-0.01605185], [ 0.0104227 ], [ 0.05892398], [ 0.05372047], [ 0.14753203], [ 0.20979349], [ 0.08846406], [-0.39405411], [-0.04870572], [-0.1612745 ]]

Objective function 0.9125785012725371