HOMEWORK # 4 SOLUTIONS

Ahmad F. Taha

October 2, 2015

Homework Instructions:

1. Type your solutions in the LaTeX homework template file. Otherwise, you can use any other typesetting tool or you can provide handwritten solutions, assuming everything is clear.
2. **Due date:** Wednesday, September 30th, @ 5:00pm on Blackboard, **AND** drop off a copy of your solutions (slip it under the office door if I'm away).
3. **Collaboration policy:** you can collaborate with your classmates, under the assumption that everyone is required to write their own solutions. If you choose to collaborate with anyone, list their name(s).
4. You don't show your work ⇒ You don't get credit.
5. Solutions that are unclear won't be graded.
6. Before you start with this homework assignment, make sure that you have grasped the content of Module 04.

Problem 1 — Convexity Property

A function $f : \mathbb{R}^n \to R$ is continuously differentiable. Also, assume that $f(x)$ **is concave on a convex set** $\mathcal{X}$. Given the aforementioned properties of $f(x)$, prove that for all $x_1, x_2 \in \mathcal{X}$, $f(x)$ satisfies this property:

$$f(x_2) \leq f(x_1) + Df(x_1)(x_2 - x_1).$$

Hint: *Back to basics—what is the basic definition of a derivative?*

Solutions:

By definition, a function $f(x)$ is concave if $g(x) = -f(x)$ is convex. By the definition of a convex function, we know that $\forall\, \beta \in [0,1]$, $g(x) = -f(x)$ satisfies:

$$g(\beta x_2 + (1-\beta)x_1) \leq \beta g(x_2) + (1-\beta) g(x_1).$$

The above inequality can be written as:

$$\frac{g(x_1 + \beta(x_2 - x_1)) - g(x_1)}{\beta} \leq g(x_2) - g(x_1).$$

Applying the basic definition of a derivative by setting $\beta \to 0$, we obtain:

$$Dg(x_1)(x_2 - x_1) \leq g(x_2) - g(x_1), \text{ or}$$

$$g(x_2) \geq g(x_1) + Dg(x_1)(x_2 - x_1).$$

Since $g(x) = -f(x)$, we obtain:

$$f(x_2) \leq f(x_1) + Df(x_1)(x_2 - x_1).$$

Problem 2 — Convexity of a Disc

Show that the set Ω given by $\Omega = \{y \in \mathbb{R}^2; \|y\|^2 \leq 4\}$ is convex, where $\|y\|^2 = y^\top y$.

Hint: Show that if $z = \beta x + (1-\beta)y$, then $\|z\|^2 \leq 4$. You might find the submultiplicative matrix-vector property to be useful too.

Solutions:

The set Ω is convex if $x, y \in \Omega$, then $z = \beta x + (1-\beta)y$ is a point on the line joining x and y should also be in Ω, where $0 \leq \beta \leq 1$. Hence, the problem reduces to showing that $z = \beta x + (1-\beta)y \in \Omega$.

$$\begin{aligned}
\|z\|^2 &= z^\top z \\
\|z\|^2 &= (\beta x + (1-\beta)y)^\top(\beta x + (1-\beta)y) \\
&= \beta^2\|x\|^2 + 2\beta(1-\beta)x^\top y + (1-\beta)^2\|y\|^2 \\
&\leq \beta^2\|x\|^2 + 2\beta(1-\beta)\|x\|\|y\| + (1-\beta)^2\|y\|^2 \\
&\leq 4\beta^2 + 8\beta(1-\beta) + 4(1-\beta)^2 \\
&= 4\beta^2 + 8\beta - 8\beta^2 + 4 + 4\beta^2 - 8\beta \\
&= 4
\end{aligned}$$

Hence $\|z\|^2 \leq 4$, which proves that z is in Ω and that Ω is a convex set.

Problem 3 — Minimizing a Function

Given a multivariable function $f(x)$, many optimization solvers use the following algorithm to solve $\min_x f(x)$:

1. Choose an initial guess, $x^{(0)}$
2. Choose an initial real, symmetric positive definite matrix $H^{(0)}$
3. Compute $d^{(k)} = -H^{(k)}\nabla_x f(x^{(k)})$
4. Find $\beta^{(k)} = \arg\min_\beta f(x^{(k)} + \beta^{(k)}d^{(k)}),\ \beta \geq 0$
5. Compute $x^{(k+1)} = x^{(k)} + \beta^{(k)}d^{(k)}$

For this problem, we assume that the given function is a typical quadratic function ($x \in \mathbb{R}^n$), as follows:

$$f(x) = \frac{1}{2}x^\top Qx - x^\top b + c, \quad Q = Q^\top \succ 0.$$

Answer the following questions:

1. Find $f(x^{(k)} + \beta^{(k)}d^{(k)})$ for the given quadratic function.
2. Obtain $\nabla_x f(x^{(k)})$ for $f(x)$.
3. Using the chain rule, and given that $\beta^{(k)} = \arg\min_\beta f(x^{(k)} + \beta^{(k)}d^{(k)})$, find a closed form solution for $\beta^{(k)}$ in terms of the given matrices ($H^{(k)}, \nabla f(x^{(k)}), d^{(k)}, Q$).
4. Since it is required that $\beta^{(k)} \geq 0$, provide a sufficient condition related to $H^{(k)}$ that guarantees the aforementioned condition on $\beta^{(k)}$.

Solutions:

1. $f(x^{(k)} + \beta^{(k)}d^{(k)}) = \frac{1}{2}(x^{(k)} + \beta^{(k)}d^{(k)})^\top Q(x^{(k)} + \beta^{(k)}d^{(k)}) - (x^{(k)} + \beta^{(k)}d^{(k)})^\top b + c$
2. $\nabla_x f(x^{(k)}) = Qx^{(k)} - b$
3. Using the chain rule, and since we're minimizing with respect to β, we obtain:

$$\frac{d}{d\beta}f(x^{(k)} + \beta d^{(k)}) = \left(x^{(k)} + \beta d^{(k)}\right)^\top Qd^{(k)} - (d^{(k)})^\top b = 0$$

$$\Rightarrow \left((x^{(k)})^\top Q - b^\top\right) d^{(k)} = -\beta(d^{(k)})^\top Qd^{(k)}.$$

   Note that $d^{(k)} = -H^{(k)}\nabla_x f(x^{(k)})$, and since $Q = Q^\top \succ 0$, we obtain:

$$\boxed{\beta_k^* = -\frac{\nabla f(x^{(k)})^\top d^{(k)}}{(d^{(k)})^\top Qd^{(k)}} = \frac{\nabla f(x^{(k)})^\top H^{(k)}\nabla f(x^{(k)})}{(d^{(k)})^\top Qd^{(k)}}.}$$

4. Clearly, the condition is $H^{(k)} = (H^{(k)})^\top \succ 0$, since $Q = Q^\top \succ 0$.

Problem 4 — KKT Conditions, 1

Using the KKT conditions discussed in class, obtain all the candidate strict local minima for the following nonlinear optimization problem:

$$\begin{aligned} \max \quad & -x_1^2 - 2x_2^2 \\ \text{subject to} \quad & x_1 + x_2 \geq 3 \\ & x_2 - x_1^2 \geq 1 \end{aligned}$$

There are many cases to consider. Make sure that you don't miss any.

After solving the problem analytically, code the problem on `NEOS` solver (http://www.neos-server.org/neos/solvers/index.html), using any solver of your choice and any modeling language (GAMS, AMPL, ...).

Solutions:

First, rewrite the optimization problem in standard form:

$$\begin{aligned} \min \quad & x_1^2 + 2x_2^2 \\ \text{subject to} \quad & -x_1 - x_2 + 3 \leq 0 \\ & -x_2 + x_1^2 + 1 \leq 0 \end{aligned}$$

Then, construct the Lagrangian:

$$L(x_1, x_2, \mu_1, \mu_2) = x_1^2 + 2x_2^2 + \mu_1(-x_1 - x_2 + 3) + \mu_2(-x_2 + x_1^2 + 1).$$

The KKT conditions are:

1. $\nabla_{x_1} L(x_1, x_2, \mu_1, \mu_2) = 2x_1 - \mu_1 + 2\mu_2 x_1 = 0$
2. $\nabla_{x_2} L(x_1, x_2, \mu_1, \mu_2) = 4x_2 - \mu_1 - \mu_2 = 0$
3. $\mu_1(x_1 + x_2 - 3) = 0$
4. $\mu_2(x_2 - x_1^2 - 1) = 0$
5. $\mu_1, \mu_2 \geq 0$
6. $-x_1 - x_2 + 3 \leq 0$
7. $-x_2 + x_1^2 + 1 \leq 0$

There are few cases to consider:

Case 1— $\mu_1 = \mu_2 = 0 \Rightarrow x_1 = x_2 = 0$. However, condition 6 would be violated. Thus, this point doesn't satisfy the KKT condition and is not a candidate for a minimizer.

Case 2— $\mu_1 \geq 0, \mu_2 = 0$. Given this assumption, and solving

$$x_2 - x_1^2 - 1 = 0, 2x_1 - \mu_1 = 0, 4x_2 - \mu_1 = 0,$$

we obtain a unique solution: $x_1 = 2, x_2 = 1$. However, this solution violates condition 7. Thus, this point doesn't satisfy the KKT condition and is not a candidate for a minimizer.

Case 3— Similar to Case 2, we choose $\mu_2 \geq 0, \mu_1 = 0$. The solution obtained is $x_1 = 0$ and $x_2 = 1$, which violates condition 6. Thus, this point doesn't satisfy the KKT condition and is not a candidate for a minimizer.

Case 4— We now consider the case when $\mu_1, \mu_2 > 0$. This case implies that

$$\begin{aligned} -x_1 - x_2 + 3 &= 0 \\ -x_2 + x_1^2 + 1 &= 0, \end{aligned}$$

which implies that $(3 - x_1) - x_1^2 - 1 = 0$ or $x_1^2 + x_1 - 2 = 0$. This quadratic polynomial has two solutions: $x_1^{(1)} = -2, x_2^{(2)} = 1$. By substitution, the two solutions generate $x_2^{(1)} = 5$ and $x_2^{(2)} = 2$.

The first candidate point, $x_1^{(1)} = -2, x_2^{(2)} = 5$ implies that $-4 - \mu_1 - 4\mu_2$ (from Condition 1), which is impossible for two positive variables μ_1 and μ_2.

The second candidate point, $x_1^{(2)} = 1, x_2^{(2)} = 2$ implies that $\mu_1 = 6, \mu_2 = 2$ (from Conditions 1 and 2)—satisfying all the KKT conditions. Therefore, $x_1^* = 1, x_2^* = 2$.

Problem 5 — KKT Conditions, 2

Using the KKT conditions discussed in class, obtain all the candidate strict local minima for the following nonlinear optimization problem:

$$\begin{aligned} \min \quad & x_1 + x_2^2 \\ \text{subject to} \quad & x_1 - x_2 = 5 \\ & x_1^2 + 9x_2^2 \leq 36 \end{aligned}$$

There are many cases to consider. Make sure that you don't miss any.

After solving the problem analytically, code the problem on `NEOS` solver, using any solver of your choice.

Solutions:

Solution approach is similar to Problem 4.

Problem 6 — Convexity Range

For the following function, find the set of values for β such that the function is convex.

$$f(x,y,z) = x^2 + y^2 + 5z^2 - 2xz + 2\beta xy + 4yz$$

Solutions:

After representing $f(x,y,z)$ in a quadratic, symmetric form, find the principal minors. To guarantee the convexity of $f(x,y,z)$, all the principal minors should be non-negative.

$f(x,y,z)$ can be written as:

$$\begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 1 & \beta & -1 \\ \beta & 1 & 2 \\ -1 & 2 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x & y & z \end{bmatrix} Q \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Evaluating the principal minors of Q, we obtain the following conditions:

$$1 - \beta^2 \geq 0$$

$$-5\beta^2 - 4\beta \geq 0.$$

Combining the above conditions on β, the range of β that guarantees convexity of f is:

$$\boxed{-\frac{4}{5} \leq \beta \leq 0.}$$

Problem 7 — CVX Programming

The objective of this problem is to get you started with CVX—the convex optimization solver on MATLAB. Do the following:

1. Watch this CVX introductory video: https://www.youtube.com/watch?v=N2b_B4TNfUM
2. Download and install CVX on your machine: http://cvxr.com/cvx/download/
3. Read the first few pages of the CVX User's Guide: http://web.cvxr.com/cvx/doc/
4. Solve Problems 4 and 5 using CVX.

Solutions:

```
% Problem 4
cvx_begin
variable x1
variable x2
minimize( x1^2+2*x2^2)
subject to
x1+x2-3>=0
x2-x1^2-1 >= 0
cvx_end
x1
x2
Status: Solved
Optimal value (cvx_optval): +9


x1 =

1.0000


x2 =

2.0000
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

clc
clear all
% Problem 5
cvx_begin
variable x1
variable x2
minimize( x1+x2^2)
subject to
x1-x2-5==0
x1^2+9*x2^2-36 <= 0
cvx_end
x1
x2
Status: Solved
Optimal value (cvx_optval): +4.75


x1 =
```

```
4.5000


x2 =

-0.5000
```

Problem 8 — Solving LMIs using CVX

Using CVX, solve the following LMI for P:

$$\begin{aligned} A^\top P + PA &< 0 \\ B^\top P + PB &< 0 \\ P = P^\top &> 0.1I, \text{ where:} \\ A &= \begin{bmatrix} -3 & 1 \\ 0 & -1 \end{bmatrix} \\ B &= \begin{bmatrix} -2 & 0 \\ 1 & -1 \end{bmatrix} \end{aligned}$$

What happens if you try to solve the same LMI when $B = \begin{bmatrix} -2 & 3 \\ 1 & -1 \end{bmatrix}$?

Justify the results.

Solutions:

For the first part of the problem, this code can be used to generate the answer. Note that the objective cannot be empty. Hence, a constant in the objective function (which is a convex cost function) suffices.

```
clc
clear all
% Problem 8-a
A=[-3 1;0 -1];
B=[-2 0; 1 -1];
n=length(A);
cvx_begin

variable P(n,n) symmetric

minimize(1)
subject to
A'*P + P*A < 0
B'*P + P*B < 0
P > 0.1*eye(n)
cvx_end
% Solution:
Status: Solved
Optimal value (cvx_optval): +1


P =

4.1413    4.8195
4.8195    8.5548
```

When $B = \begin{bmatrix} -2 & 3 \\ 1 & -1 \end{bmatrix}$, the solution won't converge, since B has a positive eigenvalue. Recall the sufficiency condition on the existence of a solution to the Lyapunov equation...

```
clc
clear all
% Problem 8-b
A=[-3 1;0 -1];
B=[-2 3; 1 -1];
```

```
cvx_begin

variable P(2,2) symmetric

minimize(norm(P))
subject to
A'*P + P*A < 0
B'*P + P*B < 0
P > 0.1*eye(2)
cvx_end
P
eig(B)
Status: Infeasible
Optimal value (cvx_optval): +Inf


P =

NaN   NaN
NaN   NaN


ans =

-3.3028
0.3028 % a +ve evalue
```