HOMEWORK # 4
Hafez Bazrafshan
October 1, 2015

To facilitate checking, questions are color-coded blue and pertinent answers follow in regular font. The printed version however, will be in black and white.

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?*

Response. First assume $f : \mathbb{R} \to \mathbb{R}$ is continously differentiable and concave (which implies its domain $\mathcal{X}$ is convex). Since f is concave, for any $t \in [0, 1]$ and $x_1, x_2 \in \mathcal{X}$ the following holds:

$$f(x_1 + t(x_2 - x_1)) \geq tf(x_2) + (1 - t)f(x_1). \tag{1}$$

We can easily rearrange (1) to obtain the following

$$\frac{f(x_1 + t(x_2 - x_1))}{t} - \frac{1 - t}{t} f(x_1) \geq f(x_2). \tag{2}$$

which yields:

$$\frac{f(x_1 + t(x_2 - x_1)) - f(x_1)}{t} + f(x_1) \geq f(x_2). \tag{3}$$

Taking the limit of both sides of (3) when $t \to 0$ yields the desired result:

$$f'(x_1)(x_2 - x_1) + f(x_1) \geq f(x_2) \quad \forall x_1, x_2 \in \mathcal{X}. \tag{4}$$

Now we prove the general case for $f : \mathbb{R}^n \to \mathbb{R}$, f concave and differentiable. For a given $x_1, x_2 \in \mathcal{X}$ define $g(t) = f(x_1 + t(x_2 - x_1))$ for $t \in [0, 1]$. Notice that $g(1) = f(x_1 + x_2 - x_1) = f(x_2)$, $g(0) = f(x_1)$ and $g'(0) = Df(x_1)(x_2 - x_1)$. Also notice that $g(t)$ is concave since for any $t_1, t_2 \in [0, 1]$ and $\theta \in [0, 1]$ we have the following:

$$\theta t_1 + (1 - \theta)t_2 \geq 0 \tag{5}$$
$$\theta t_1 + (1 - \theta)t_2 \leq \theta + (1 - \theta)1 = 1 \tag{6}$$
$$\Rightarrow \theta t_1 + (1 - \theta)t_2 \in [0, 1]. \tag{7}$$

$$\begin{aligned} g(\theta t_1 + (1 - \theta)t_2) &= f[x_1 + (\theta t_1 + (1 - \theta)t_2)(x_2 - x_1)] \\ &= f[\theta x_1 + \theta t_1(x_2 - x_1) + (1 - \theta)x_1 + (1 - \theta)t_2(x_2 - x_1)] \\ &\geq \theta f(x_1 + t_1(x_2 - x_1)) + (1 - \theta)f(x_1 + t_2(x_2 - x_1)) \quad \text{(due to concavity of } f) \\ &\geq \theta g(t_1) + (1 - \theta)g(t_2) \text{(by definition of } g(t)). \end{aligned} \tag{8}$$

Since $g(t)$ is concave on $t \in [0, 1]$ we can use the result in (4) for $t_1 = 0$ and $t_2 = 1$:

$$g(1) \leq g'(0)(1 - 0) + g(0) \Rightarrow f(x_2) \leq Df(x_1)(x_2 - x_1) + f(x_1). \tag{9}$$

This proof is almost identical to the proof of first-order conditions by Convex Optimization book by Boyd, except that I prove why g is concave.

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.

Response. Assume $x, y \in \Omega$ so we know $||x||^2 \leq 4$ and $||y||^2 \leq 4$. Also note that this implies $||x|| \leq 2$ and $||y|| \leq 2$. Now assume $\theta \in [0,1]$ and let's assess what the value of $||\theta x + (1-\theta)y||^2$ will be.

$$
\begin{aligned}
||\theta x + (1-\theta)y||^2 &= [\theta x + (1-\theta)y]^T[\theta x + (1-\theta)y] \\
&= \theta^2||x||^2 + 2\theta(1-\theta)x^Ty + (1-\theta)^2||y||^2 \qquad (10) \\
&\leq \theta^2 \times 4 + 2\theta(1-\theta)x^Ty + (1-\theta)^2 \times 4. \qquad (11)
\end{aligned}
$$

Due to Cauchy-Shwartz inequality, we know that $x^Ty \leq ||x||||y|| \leq 2 \times 2 = 4$. Hence we can extend (11) to the following:

$$
\begin{aligned}
||\theta x + (1-\theta)y||^2 &\leq \theta^2 \times 4 + 2\theta(1-\theta) \times 4 + (1-\theta)^2 \times 4 \\
&= 4(\theta^2 + 2\theta(1-\theta) + (1-\theta)^2) = 4(\theta + 1 - \theta)^2 = 4 \qquad (12)
\end{aligned}
$$

Hence $||\theta x + (1-\theta)y||^2 \leq 4$ which means $\theta x + (1-\theta)y \in \Omega$.
I did not use the hint since I didn't know what sub-multiplicative property was.

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 $\beta^{(k)}$.

Response.

1.

$$f(x^{(k)} + \beta^{(k)} d^{(k)}) = \frac{1}{2}[x^{(k)} + \beta^{(k)} d^{(k)}]^T Q [x^{(k)} + \beta^{(k)} d^{(k)}] - [x^{(k)} + \beta^{(k)} d^{(k)}]^T b + c. \tag{13}$$

2.

$$\nabla_x f(x^{(k)}) = Q x^{(k)} - b. \tag{14}$$

3. We find $\beta^{(k)}$ by setting the derivative of (13) to zero, which results in the following equation:

$$(x^{(k)})^T Q d^{(k)} + \beta^{(k)} (d^{(k)})^T Q d^{(k)} - (d^{(k)})^T b = 0 \tag{15}$$

In (15), we replace $d^{(k)}$ with its equivalent $d^{(k)} = -H^{(k)} \nabla_x f(x^{(k)})$, which will yield the following:

$$\beta^{(k)} [H^{(k)} \nabla_x f(x^{(k)})]^T Q [H^{(k)} \nabla_x f(x^{(k)})] = (x^{(k)})^T Q [H^{(k)} \nabla_x f(x^{(k)})] - b^T [H^{(k)} \nabla_x f(x^{(k)})] \tag{16}$$

$$\rightarrow \beta^{(k)} = \frac{[(x^{(k)})^T Q - b^T] H^{(k)} \nabla_x f(x^{(k)})}{[H^{(k)} \nabla_x f(x^{(k)})]^T Q [H^{(k)} \nabla_x f(x^{(k)})]}. \tag{17}$$

Notice in (17) that $(x^{(k)})^T Q - b^T = \nabla_x f(x^{(k)})^T$ according to (14). Finally we find the best $\beta^{(k)}$ as:

$$\beta^{(k)} = \frac{\nabla_x f(x^{(k)})^T H^{(k)} \nabla_x f(x^{(k)})}{[H^{(k)} \nabla_x f(x^{(k)})]^T Q [H^{(k)} \nabla_x f(x^{(k)})]}. \tag{18}$$

4. Since $Q \succ 0$, the denominator of (18) is always positive. Therefore, in order to have $\beta^{(k)} \geq 0$, it is sufficient to have $H^{(K)} \succeq 0$.

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, ...). **Show your code and outputs.**

Response. We can write the problem as a minimization problem and convert it to a standard form (for using Lagrange multipliers):

$$\begin{aligned} \min \quad & x_1^2 + 2x_2^2 \\ \text{subject to} \quad & -x_1 - x_2 + 3 \leq 0 \leftarrow \mu_1 \\ & x_1^2 - x_2 + 1 \leq 0 \leftarrow \mu_2 \end{aligned}$$

The Lagrangian is given as:

$$\mathcal{L}(x, \mu_1, \mu_2) = x_1^2 + 2x_2^2 + \mu_1(-x_1 - x_2 + 3) + \mu_2(x_1^2 - x_2 + 1) \tag{19}$$

The KKT conditions:

1.

$$\nabla_x \mathcal{L}(x, \mu_1, \mu_2) = \begin{bmatrix} 2x_1 - \mu_1 + 2\mu_2 x_1 \\ 4x_2 - \mu_1 - \mu_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{20}$$

2.

$$\mu_1, \mu_2 \geq 0. \tag{21}$$

3.

$$\mu_1(-x_1 - x_2 + 3) = 0 \tag{22}$$
$$\mu_2(x_1^2 - x_2 + 1) = 0. \tag{23}$$

4.

$$-x_1 - x_2 + 3 \leq 0 \tag{24}$$
$$x_1^2 - x_2 + 1 \leq 0. \tag{25}$$

5. $\nabla_x^2 \mathcal{L}$ is always positive semidefinite since $\mu_2 \geq 0$.

$$\nabla_x^2 \mathcal{L} = \begin{bmatrix} 2 + 2\mu_2 & 0 \\ 0 & 4 \end{bmatrix} \succeq 0. \tag{26}$$

We solve the KKT conditions by conditioning μ_1, μ_2 based on (22) and (23), and solving the remaining equations as follows:

1. Case $\mu_1 = \mu_2 = 0 \to x_1 = 0, x_2 = 0$. This is not acceptable since (24) and (25) are violated.
2. Case $\mu_1 = 0 \to \mu_2 = 4 \to x_1 = 0, x_2 = 1$. This is not acceptable since (24) is violated.
3. Case $\mu_2 = 0 \to \mu_1 = 4 \to x_1 = 2, x_2 = 1$. This is not acceptable since (25) is violated.
4. Case $\mu_1, \mu_2 \neq 0$. In this case, due to (22) and (23) we have:

$$-x_1 - x_2 + 3 = 0 \tag{27}$$
$$x_1^2 - x_2 + 1 = 0 \tag{28}$$

   The system above has two sets of answers:

   (a) $x_1 = -2, x_2 = 5 \to \mu_1 = 28, \mu_2 = -8$. This is not acceptable since it contradicts condition (21).

   (b) $x_1 = 1, x_2 = 2 \to \mu_1 = 6, \mu_2 = 2$. This is an acceptable answer.

To summarize, we found that $x_1 = 1, x_2 = 2$ is the minimizer (or the maximizer for the original problem). Since the problem is convex (concave for the original), this minimizer is the global minimum of the problem. The code for this problem is pasted below (using AMPL with Filter in NEOS):

```
# model file:
var x1;
var x2;

minimize c:
    x1^2+2*x2^2;

subject to A:
    x1+x2>=3;

subject to B:
    x2-x1^2>=1;

# command file:
solve;
display x1, x2;
```

```
# NEOS email:
Using 64 bit binary
File exists
You are using the solver filter.
Executing AMPL.
processing data.
processing commands.
Executing on neos-5.neos-server.org

2 variables, all nonlinear
2 constraints; 4 nonzeros
        1 nonlinear constraint
        1 linear constraint
        2 inequality constraints
1 nonlinear objective; 2 nonzeros.

filterSQP (20020316): Optimal solution found, objective = 9
4 iterations (0 for feasibility)
Evals: obj = 5, constr = 6, grad = 6, Hes = 5
x1 = 1
x2 = 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. **Show your code and outputs.**

Response.

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

The Lagrangian is as follows:

$$\mathcal{L}(x, \lambda, \mu) = x_1 + x_2^2 + \lambda(x_1 - x_2 - 5) + \mu(x_1^2 + 9x_2^2 - 36) \tag{29}$$

The KKT conditions are given:

1.

$$\nabla_x \mathcal{L}(x, \lambda, \mu) = \begin{bmatrix} 1 + \lambda + 2\mu x_1 \\ 2x_2 - \lambda + 18\mu x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{30}$$

2.

$$\mu \geq 0 \tag{31}$$

3.
$$\mu(x_1^2 + 9x_2^2 - 36) = 0 \tag{32}$$

4.
$$x_1^2 + 9x_2^2 - 36 \leq 0 \tag{33}$$

5.
$$x_1 - x_2 - 5 = 0 \tag{34}$$

6. $\nabla_x^2 \mathcal{L}(x, \lambda, \mu) \succeq 0$ always holds since $\mu \geq 0$.
$$\nabla_x^2 \mathcal{L}(x, \lambda, \mu) = \begin{bmatrix} 2\mu & 0 \\ 0 & 2 + 18\mu \end{bmatrix} \succeq 0. \tag{35}$$

We solve this problem by conditioning μ using (32):

1. Case $\mu = 0 \rightarrow \lambda = -1, x_1 = \frac{9}{2}, x_2 = -\frac{1}{2}$. This is acceptable since it does not violate any other KKT.
2. Case $\mu \neq 0$. Thus we have the following system:
$$x_1^2 + 9x_2^2 = 36 \tag{36}$$
$$x_1 = x_2 + 5 \tag{37}$$

   This system has two sets of solutions, both of which are not acceptable:

   (a) $x_1 = \frac{45}{10} + \frac{\sqrt{135}}{10}, x_2 = \frac{-5}{10} + \frac{\sqrt{135}}{10} \rightarrow \mu = \frac{-1-2x_2}{2x_1+18x_2} < 0.$

   (b) $x_1 = \frac{45}{10} - \frac{\sqrt{135}}{10}, x_2 = \frac{-5}{10} - \frac{\sqrt{135}}{10} \rightarrow \mu = \frac{-1-2x_2}{2x_1+18x_2} < 0.$

The code for this problem is pasted below:

```
# model file:
var x1;
var x2;

minimize c:
    x1+x2^2;

subject to A:
    x1-x2-5=0;

subject to B:
    x1^2+9*x2^2-36<=0;

# command file:
solve;
display x1, x2;

# NEOS email:
Using 64 bit binary
File exists
You are using the solver filter.
Executing AMPL.
processing data.
processing commands.
```

```
Executing on neos-5.neos-server.org

2 variables, all nonlinear
2 constraints; 4 nonzeros
        1 nonlinear constraint
        1 linear constraint
        1 equality constraint
        1 inequality constraint
1 nonlinear objective; 2 nonzeros.

filterSQP (20020316): Optimal solution found, objective = 4.75
1 iterations (0 for feasibility)
Evals: obj = 2, constr = 3, grad = 3, Hes = 2
x1 = 4.5
x2 = -0.5
```

Problem 6 — Convexity Range

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

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

Response. We can organize $f(x,y,z)$ as $f(x,y,z) = [x,y,z]P \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ where

$$P = \begin{bmatrix} 1 & \beta & -1 \\ \beta & 1 & 2 \\ -1 & 2 & 5 \end{bmatrix}. \tag{38}$$

For $f(x,y,z)$ to be convex, P needs to be postivie semidefinite. Necessary and sufficient condition for positive semidefinite symmetic matrices is that all of its principal minors be non-negative. Matrix P is a 3×3 matrix and has 7 principal minors, 5 of which are independent of β and (straightforward to check) positive. Only the determinant, and one of the principal minors of order 2 are dependent on β:

$$\Delta_3 = det(P) = \begin{vmatrix} 1 & 2 \\ 2 & 5 \end{vmatrix} - \beta \begin{vmatrix} \beta & 2 \\ -1 & 5 \end{vmatrix} + (-1) \begin{vmatrix} \beta & 1 \\ -1 & 2 \end{vmatrix} \geq 0 \rightarrow -\frac{4}{5} \leq \beta \leq 0 \tag{39}$$

$$\Delta_2 = 1 - \beta^2 \geq 0 \rightarrow -1 \leq \beta \leq 1. \tag{40}$$

Thus, for $-\frac{4}{5} \leq \beta \leq 0$ the function is convex.

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. **Show your code and outputs.**

Response. Code for problem 4:

```
cvx_begin
variables x1 x2

maximize(-x1-2*x2^2)
subject to:

x1+x2>=3;
x2-x1^2>=1;
cvx_end

% -------------------------------------------------------------
% Status: Solved
% Optimal value (cvx_optval): -9
%
%
% x1 =
%
%       1.0000
%
% x2 =
%
%       2.0000
```

Code for problem 5:

```
cvx_begin
variables x1 x2

minimize(x1+x2^2)
subject to:

x1-x2-5==0
x1+9*x2^2-36<=0;
cvx_end

%
% -------------------------------------------------------------
% 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:

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

Show your code and outputs.

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

Justify the results.

Response.

For $B = \begin{bmatrix} -2 & 0 \\ 1 & -1 \end{bmatrix}$:

```
A=[-3  1;  0  -1];
B=[-2  0;  1  -1];
cvx_begin sdp
variable P(2,2) symmetric;

minimize(0)
subject to:

A.' *P +P*A < 0;
B.'*P+P*B < 0;
P >= 0.1*eye(2);

cvx_end


% ------------------------------------------------------------
% Status: Solved
% Optimal value (cvx_optval): +0


% eig(A.' *P +P*A )
%
% ans =
%
%    -17.8518
%    -10.2769

% eig(B.'*P+P*B)
%
% ans =
%
%    -14.6261
%     -7.6468
```

```
% eig(P)
%
% ans =
%
%       2.6109
%       6.6896
```

For $B = \begin{bmatrix} -2 & 3 \\ 1 & -1 \end{bmatrix}$:

```
A=[-3 1; 0 -1];
B=[-2 3; 1 -1];
cvx_begin sdp
variable P(2,2) symmetric;

minimize(0)
subject to:

A.' *P +P*A < 0;
B.'*P+P*B < 0;
P >= 0.1*eye(2);

cvx_end


% ------------------------------------------------------------
% Status: Infeasible
% Optimal value (cvx_optval): +Inf
```

According to Lyapenov's theorem, real parts of the eigenvalues of matrices A, B are negative if and ONLY if there exists a symmetric positive definite matrix P such that:

$$A^\top P + PA < 0 \tag{41}$$
$$B^\top P + PB < 0. \tag{42}$$

The B in the first problem has negative eigenvalues and hence there exists a positive definite matrix P. The B in the second problem has a positive eigenvalue and hence the problem will become infeasible.

```
eig([-2 0; 1 -1])
ans =

    -1
    -2

eig([-2 3; 1 -1])

ans =

   -3.3028
    0.3028
```