MTH377/577 CONVEX OPTIMIZATION
Winter Semester 2022
Indraprastha Institute of Information Technology Delhi
MIDSEM EXAM (Time: 1 hour 15 minutes, Total Points: 30)

Please Note:

1. The stipulated time to workout the exam is 1 hour 15 minutes. After that, scan and upload. Submission received after the deadline will incur a penalty.

2. First attempt submitting via Google Classroom. If you encounter technical problems in submitting via Google Classroom, send it to shreyat@iiitd.ac.in directly via email. If you have sent it via email well within time, then there is no need to upload it on Google Classroom.

3. While you can consult all the material at hand, discussing the problems with any person is a violation of academic integrity.

Q1. (a) (2 points). Are the canonical basis vectors $e_1, \ldots, e_n$ in $\mathbb{R}^n$ affinely independent ? If yes, prove it. If not, argue why not.

A1. (a) Yes, they are affinely independent. To verify this, we only need to show that the set $e_2 - e_1, \ldots, e_n - e_1$ is linearly independent in $\mathbb{R}^n$. Suppose some linear combination of these vectors is the zero vector i.e. $\lambda_2(e_2 - e_1) + \ldots + \lambda_n(e_n - e_1) = 0$ where $\lambda_2, \ldots, \lambda_n$ are real numbers. This implies that

$$-(\lambda_2 + \ldots + \lambda_n)e_1 + \lambda_2 e_2 + \ldots + \lambda_n e_n = 0$$

Since $e_1, \ldots, e_n$ are linearly independent vectors, it must be that $\lambda_2 = \ldots = \lambda_n = 0$. Hence $e_2 - e_1, \ldots, e_n - e_1$ is linearly independent set of vectors and so $e_1, \ldots, e_n$ are affinely independent.

Grading: The particular vector that one subtracts is not important. I have subtracted e_1 but it could be any e_i. Allocate 1 point if you see that a student understands the definition of affine independence. This is the second sentence above. Allocate 1 point to the rest of the answer which is about applying this definition to answer the question.

(b) (2 point). Is the line passing through the point $(3/2, 1)$ and slope -2 a hyperplane ? If yes, identify this hyperplane with its (normal vector, scalar) and indicate the positive and negative halfspaces associated with it. If not, argue why not.

A1. (b) Yes, a line in the plane is a hyperplane in $\mathbb{R}^2$. Its equation is $2x + y = 4$. So its normal vector is $(2, 1)$ and its scalar is 4.

(c) (2 points). Consider the function $f : \mathbb{R}_{++} \mapsto \mathbb{R}$ defined by

$$f(x) = 2e^{(2x+5)\log(2x+5)} + \frac{5}{x}$$

If f convex or concave ? Why ?

A1. (c) Let us build up f from some elementary functions.
1. From our basic library, recognize that $h(x) = e^{x\log(x)}$ is convex on $\mathbb{R}_{++}$, $g(x) = 2x + 5$ is affine on $\mathbb{R}_{++}$, and $f_2(x) = 1/x$ is convex on $\mathbb{R}_{++}$.
2. $f_1(x) = h(g(x)) = e^{(2x+5)\log(2x+5)}$ is convex as a convex transformation of an affine function.
3. $f = 2f_1 + 5f_2$ is convex as a nonnegative weighted sum of two convex functions f_1 and f_2.

Grading: 1 point for the first step, and 0.5 points each for the second and the third step.

(d) (2 points). Is the following optimization problem convex ? Argue why or why not.

$$\begin{aligned} \max_{x_1,x_2} \quad & 2\log(x_1 - 2) + 3\log(x_2 - 3) \\ \text{subject to} \quad & 2x_1 + 3x_2 \leq 25 \end{aligned}$$

A1. (d) The feasible set of the problem is given by the set

$$F = \{(x_1, x_2) : x_1 > 2, x_2 > 3, 2x_1 + 3x_2 \leq 25\}$$

F is a solution set of linear inequalities and so a polyhedron. Therefore, F is a convex set.

For the objective function, 1. $(x_1, x_2) \mapsto x_1 - 2$ and $(x_1, x_2) \mapsto x_2 - 3$ are affine functions.
2. $\log(x_1 - 2)$ is a concave function as a concave transformation of an affine function. For a similar reason, $\log(x_2 - 3)$ is concave.
3. So the objective function $2\log(x_1-2)+3\log(x_2-3)$ is a concave function as a nonnegative weighted sum of two concave functions.

Hence, the problem is convex.

Grading: 1 point for arguing that the feasible set is convex; 1 point for showing the objective function is concave.

(e) (2 points). Draw and precisely write as a set, the 1-sublevel set of e^x and the 0-superlevel set of $\log(x)$, where x is a real valued variable.

A1. (e) The 1-sublevel set of e^x is the nonpositive real line $(-\infty, 0]$; the 0-superlevel set of $\log(x)$ is the interval $[1, \infty)$.

Grading: 1 point for each part.

Q2. (5 points). Consider the function $f : \mathbb{R}^2 \mapsto \mathbb{R}$ defined by $f(x_1, x_2) = \log(e^{x_1} + e^{x_2})$. Is f convex or concave ? Either way, prove it.

A2. f is twice differentiable, so use the second order characterization of differentiable convex/concave functions to check convexity/concavity.

$$Df(x_1, x_2) = \begin{bmatrix} \frac{e^{x_1}}{e^{x_1} + e^{x_2}} & \frac{e^{x_2}}{e^{x_1} + e^{x_2}} \end{bmatrix}$$

$$D^2 f(x_1, x_2) = \begin{bmatrix} \frac{e^{x_1+x_2}}{(e^{x_1}+e^{x_2})^2} & -\frac{e^{x_1+x_2}}{(e^{x_1}+e^{x_2})^2} \\ -\frac{e^{x_1+x_2}}{(e^{x_1}+e^{x_2})^2} & \frac{e^{x_1+x_2}}{(e^{x_1}+e^{x_2})^2} \end{bmatrix}$$

The eigenvalues of the Hessian matrix $D^2 f(x_1, x_2)$ are 0 and $\frac{2e^{x_1+x_2}}{(e^{x_1}+e^{x_2})^2}$. Both of these are nonnegative no matter what the point (x_1, x_2) is. So f is convex.

Grading: 1 point for computing the derivative, 1 points for computing the Hessian, 2 points for computing the eigenvalues and 1 point for the convexity inference.

Q3. (5 points). Let A be an $m \times n$ matrix. Is the set K defined below convex ? Why or why not ?

$$K = \{y \in \mathbb{R}^m : \exists x \in \mathbb{R}^n \text{ such that } ||x|| \leq 1 \text{ and } y = Ax\}$$

A3. (i) The unit norm ball in $\mathbb{R}^n$ defined as $B(0, 1) = \{x \in \mathbb{R}^n : ||x|| \leq 1\}$ is a convex set.
(ii) The mapping $x \mapsto Ax$ is an affine (in fact linear) map from $\mathbb{R}^n$ to $\mathbb{R}^m$.
(iii) K is the image of a convex set (specified in (i)) under an affine map (specified in (ii)) and so convex.

Grading: 2 points for (i), 1 point for (ii), and 2 points for (iii).

Q4. (5 points). Let f_1 and f_2 be concave functions from $\mathbb{R}^n$ to $\mathbb{R}$. Let $f : \mathbb{R}^n \mapsto \mathbb{R}$ be defined as the pointwise minimum $f(x) = \min(f_1(x), f_2(x))$. Is f convex or concave ? Either way, prove it.

A4. In the first step, prove that $hypo(f) = hypo(f_1) \cap hypo(f_2)$. This is shown as follows.

$$
\begin{aligned}
(x,t) \in hypo(f) &\text{ iff } x \in dom(f) \text{ and } t \leq f(x) \\
&\text{ iff } x \in dom(f) \text{ and } (t \leq f_1(x) \text{ and } t \leq f_2(x)) \\
&\text{ iff } (x \in dom(f_1) \text{ and } t \leq f_1(x)) \text{ and } (x \in dom(f_2) \text{ and } t \leq f_2(x)) \\
&\text{ iff } (x,t) \in hypo(f_1) \text{ and } (x,t) \in hypo(f_2) \\
&\text{ iff } (x,t) \in hypo(f_1) \cap hypo(f_2)
\end{aligned}
$$

In the second step, since f_1 and f_2 are concave, $hypo(f_1)$ and $hypo(f_2)$ are convex sets. Since convexity is preserved under set intersection, $hypo(f)$ is a convex set. But this immediately implies, by the hypograph characterization of concavity, that f is concave.

Grading: 2 points for the first step, 3 points for the second step.

Q5. (5 points). Consider the function $f(x,y) = x^3 + y^2 - 4xy - 3x$. Find all the local minima, maxima or saddle points of f.

A5. Since f is twice differentiable, compute the Jacobian and the Hessian matrix for f

$$
\begin{aligned}
Df(x,y) &= \begin{bmatrix} 3x^2 - 4y - 32y - 4x \end{bmatrix} \\
D^2f(x,y) &= \begin{bmatrix} 6x & -4 \\ -4 & 2 \end{bmatrix}
\end{aligned}
$$

The first-order necessary conditions for optimality $Df(x,y) = 0$ give the pair of equations:

$$3x^2 - 4y - 3 = 0, \quad 2y - 4x = 0$$

which can be solved to get the critical points $(3,6)$ and $(-1/3, -2/3)$.

$$D^2f(3,6) = \begin{bmatrix} 18 & -4 \\ -4 & 2 \end{bmatrix} \qquad D^2f(-1/3,-2/3) = \begin{bmatrix} -2 & -4 \\ -4 & 2 \end{bmatrix}$$

$D^2f(3,6)$ has both eigenvalues $\frac{20+/-\sqrt{192}}{2}$ strictly positive and $D^2f(-1/3,-2/3)$ has one eigenvalue positive ($\sqrt{20}$) and one eigenvalue negative ($-\sqrt{20}$). By second-order sufficiency conditions for optimality, the critical point $(3,6)$ is a local minimum while the critical point $(-1/3,-2/3)$ is a saddle point.

Grading: 2 points for finding the critical points using first order conditions, 3 points for their identification as a local minimum and saddle point using second order conditions.