ChE 597 Computational Optimization

Homework 1 Solutions

January 21st 11:59 pm

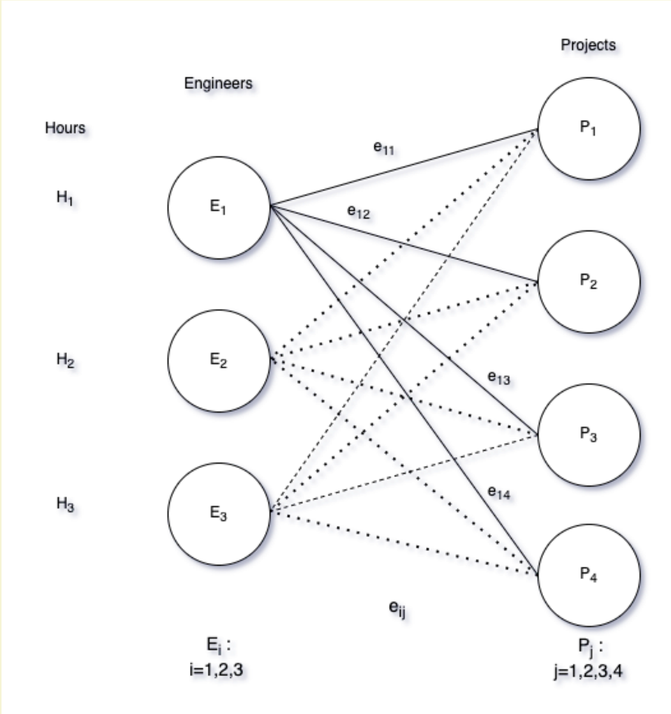
1. A small engineering consulting firm has 3 senior designers available to work on the firm's 4 current projects over the next 2 weeks. Each designer has 80 hours to split among the projects, and the following table shows the manager's scoring ($0 =$ nil to $100 =$ perfect) of the capability of each designer to contribute to each project, along with his estimate of the hours that each project will require.

	Project			
Designer	**1**	**2**	**3**	**4**
1	90	80	10	50
2	60	70	50	65
3	70	40	80	85
Required	70	50	85	35

(a) Formulate an allocation LP to choose an optimal work assignment.

Solution: We model the problem as: Let the design engineers be set E with $E_i\ : i \in [1,2,3]$ and the projects be P with $P_j\ : j \in [1,2,3,4]$

We can look at the problem as allocation of number of hours H_{ij} with each design engineer E_i $\forall i$ that are being put on to the projects P_j $\forall j$, given the i^{th} engineer E works on j^{th} project with given score e_{ij} (below figure can help to illustrate the allocation).



Let the maximum hours available with each engineer be H_{max} and the required number of hours for each project P_j be R_j $\forall j$. Thus, then the mathematical formulation to get optimal assignment in terms of score can be made as:

$$Maximize \sum_{i \in E} \sum_{j \in P} H_{ij} e_{ij}$$

Subject to:

$$\sum_{j \in P} H_{ij} \leq H_{max} \ \ \forall i \in E$$

$$\sum_{i \in E} H_{ij} \geq R_j \ \ \forall j \in P$$

$$H_{ij} \geq 0 \ \ \forall i, j$$

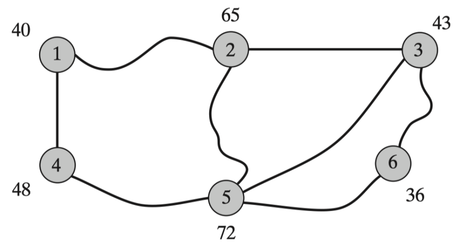
(b) Solve your model using pyomo.

Solution: We find hours worked by Engineers as:
E_1 =E_2 =E_3: 80 Hours
With the individual distribution across projects as:

	P_1	P_2	P_3	P_4
E_1	70	10	0	0
E_2	0	40	5	35
E_3	0	0	80	0

The link to the Jupyter notebook is:
`https://github.com/li-group/ChE-597-Computational-Optimization/blob/main/HW%201/HW1%20Q1.ipynb`

2. The following map shows the 6 intersections at which automatic traffic monitoring devices might be installed. A station at any particular node can monitor all the road links meeting that intersection. Numbers next to nodes reflect the monthly cost (in thousands of dollars) of operating a station at that location. This problem is known as the set covering problem in combinatorial optimization.



(a) Formulate the problem of providing full coverage at minimum total cost as a set covering integer program.

Solution: Let the set of nodes be N, indexed as $N_i = i$ in : $i = 1, 2, ..., n$.

The objective of monitoring all road links implies covering one end point of every edge of the graph thus yielding a vertex cover problem. Additionally, selecting nodes usually comes with certain costs thus making it an optimization problem.

If we represent the undirected graph $G = (N, E)$ with the vertex/node weight $c : N$. Introducing a decision variable x_n for selecting each node, such that $x_n = 1$ if selected, $x_n = 0$ otherwise. The ILP program formulating this minimum node cover is written as:

$$min \sum_{n \in N} c_n x_n$$

subject to

$$x_n + x_v \geq 1 \ \ \forall e = (n, v) \in E$$

$$x_n \in \{0, 1\} \ \ \forall N$$

The first constraint helps to ensure that each edge is being supervised at least by some station, and the second one by our definition signifies the decision to choose x_n or not (this can also be defined in the domain of the variable).

The edges can be defined with the help of a matrix A with elements a_{nv}, also called the adjacency matrix:

$$a_{nv} = \begin{cases} 1 & \text{if node n is connected to node v } (n \neq v) \\ 0 & \text{otherwise} \end{cases}$$

(Note: Adjacency matrix is symmetric)
Using these matrix elements, the constraint can be written as:

$$x_n + x_v \geq a_{nv} \ \ (n, v) \in E$$

(b) Use pyomo to solve part (a)

Solution: Solving the ILP yields minimum total cost: of 155.0 with nodes 1,3 and 5.

The link to the Jupyter notebook is:
`https://github.com/li-group/ChE-597-Computational-Optimization/blob/main/HW%201/HW1%20Q2b.ipynb`

(c) Revise your formulation of part (a) to obtain an ILP minimizing the number of uncovered road links while using at most 2 stations.

Solution: The objective of minimizing the number of uncovered road links is equivalent to maximizing the number of road links that can be covered. Let the decision variable representing the selection of edge $(n,v) \in E$ be Y_{nv}. Then, the objective can be written as:

$$\text{Max} \sum_{(n,v)\in E} \frac{Y_{nv}}{2}$$

(Note: We divide by 2 since $Y_{nv} = Y_{vn}$, as because selection of at least one node of an edge would result in the edge being covered)

Logically we would have the following equivalence:

$$Y_{nv} \Leftrightarrow Y_{vn}$$

Further, the edge selection and hence the objective would be implicitly governed by the node selection decision. Representing the node selection decision with $x_n \; \forall \, n \in N$, the logic for edge selection is written as:

$$Y_{nv} \Leftrightarrow (x_n \vee x_v)$$

The logic facilitates the selection of the edge whenever either node of an edge is chosen. The above logic can be broken down to following linear constraints as:

$$Y_{nv} \leq X_n + X_v$$
$$Y_{nv} \geq X_n$$
$$Y_{nv} \geq X_v$$

subject to:

$$\sum_{n\in N} x_n \leq 2$$

$$x_n \in \{0,1\} \;\; \forall N$$

The knowledge of node selection would yield the set of uncovered edges and thus their count.

(d) Use pyomo to solve part (c)

Solution: Model solution yields selection of nodes 3 and 5, which results in two uncovered edges (2,1) and (1,4).

The link to the Jupyter notebook is:
`https://github.com/li-group/ChE-597-Computational-Optimization/blob/`

```
main/HW%201/HW1%20Q2d.ipynb
```

3. Prove the following results covered in class. Hint: Use the definition of convex functions, the first and second order characterizations.

   (a) Univariate functions:

   - Exponential function: e^{ax} is convex for any a over $\mathbb{R}$
   - Power function: x^a is convex for $a \geq 1$ or $a \leq 0$ over $\mathbb{R}_+$ (nonnegative reals)
   - Power function: x^a is concave for $0 \leq a \leq 1$ over $\mathbb{R}_+$
   - Logarithmic function: $\log x$ is concave over $\mathbb{R}_{++}$

   Solution: For the above mentioned univariate functions defined over **convex domains**, we shall use the second-order conditions for the evaluations:
   a.1. $f(x) = e^{ax} \implies f^{"}(x) = a^2 e^{ax}$ Since, $f^{"}(x) \geq 0$ for dom f, thus f(x) is (strictly) convex.

   a.2. $f(x) = x^a \implies f^{"}(x) = a(a-1)x^{a-2}$
   Since, $x \in \mathbb{R}_+$, $\implies x^{a-2} \geq 0$
   and given *dom a* s.t. $a(a-1) \geq 0$
   $\therefore f^{"}(x) \geq 0$, which makes it convex.

   a.3. As seen in previous part, for $f(x) = x^a$, we get $f^{"}(x) = a(a-1)x^{a-2}$
   for $0 \leq a \leq 1$, $a(a-1) \leq 0$
   $\therefore f^{"}(x) \leq 0$, which makes it concave.

   a.4. $f(x) = log(x) \implies f^{"}(x) = \frac{-1}{x^2}$
   Now since, $\frac{-1}{x^2} < 0$ over the given domain, we get $f^{"}(x) < 0$, which makes it concave.

   (b) Affine function: $a^T x + b$ is both convex and concave

   Solution: Given the function f, we have for $x_1, x_2 \in \mathbb{R}^n$ and $\lambda \in [0,1]$:
   Convexity:
   $$f((1-\lambda)x_1 + \lambda x_2) \leq (1-\lambda)f(x_1) + \lambda f(x_2)$$
   Concavity:
   $$f((1-\lambda)x_1 + \lambda x_2) \geq (1-\lambda)f(x_1) + \lambda f(x_2)$$
   Therefore, it follows that for function to be both:
   $$f((1-\lambda)x_1 + \lambda x_2) = (1-\lambda)f(x_1) + \lambda f(x_2)$$
   Since, the affine function is linear, making it both convex and concave.

   (c) Quadratic function: $\frac{1}{2}x^T Q x + b^T x + c$ is convex provided that $Q \succeq 0$ (positive semidefinite)

   Solution: We enforce the second-order conditions to evaluate with the help of the Hessian.
   Given $f(x) = \left(\frac{1}{2}x^T Q x + b^T x + c\right)$, then $\nabla^2 f(x) = Q$.
   Given that $Q \succeq 0$ (positive semidefinite), this implies $\nabla^2 f(x) \succeq 0$, making it convex.

(d) Least squares loss: $\|y - Ax\|_2^2$ is always convex (since $A^T A$ is always positive semidefinite)

Solution: $f(x) = \|y - Ax\|_2^2$ can be written as $f(x) = (y - Ax)^T (y - Ax)$
With this we can evaluate its derivatives as: $\nabla f(x) = 2A^T(Ax - y)$
Further, $\nabla^2 f(x) = 2A^T A$ Now since, $A^T A$ is always positive semidefinite, $\implies$ $\nabla^2 f(x) \succeq 0$ making it convex.

(e) **Norm:** $\|x\|$ is convex for any norm; e.g., ℓ_p norms,

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p} \quad \text{for } p \geq 1, \quad \|x\|_\infty = \max_{i=1,\ldots,n} |x_i|$$

Solution: Let $f : \mathbb{R}^n \to \mathbb{R}$ be defined as $f(x) = \|x\|_p$, where $x \in \mathbb{R}^n$ and $p \geq 1$.
The ℓ_p norm of a vector x in $\mathbb{R}^n$ is defined as:

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}}$$

To prove that $f(x) = \|x\|_p$ is convex, we'll show that for any $x_1, x_2 \in \mathbb{R}^n$ and $\lambda \in [0,1]$, the inequality

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$$

holds.
Consider $x_1, x_2 \in \mathbb{R}^n$ and $\lambda \in [0,1]$. Then, by the definition of the ℓ_p norm:

$$f(\lambda x_1 + (1-\lambda)x_2) = \|\lambda x_1 + (1-\lambda)x_2\|_p = \left(\sum_{i=1}^n |\lambda x_{1i} + (1-\lambda)x_{2i}|^p\right)^{\frac{1}{p}}$$

Using Minkowski inequality i.e.

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p$$

or

$$\left(\sum_{k=1}^n |x_k + y_k|^p\right)^{1/p} \leq \left(\sum_{k=1}^n |x_k|^p\right)^{1/p} + \left(\sum_{k=1}^n |y_k|^p\right)^{1/p}$$

We can write,

$$\left(\sum_{i=1}^n |\lambda x_{1i} + (1-\lambda)x_{2i}|^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^n |\lambda x_{1i}|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^n |(1-\lambda)x_{2i}|^p\right)^{\frac{1}{p}}$$
$$= \lambda \|x_1\|_p + (1-\lambda)\|x_2\|_p$$
$$= \lambda f(x_1) + (1-\lambda) f(x_2)$$

Thus, $f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$. This inequality holds for any $x_1, x_2 \in \mathbb{R}^n$ and $\lambda \in [0,1]$.

Therefore, $f(x) = \|x\|_p$ is a convex function.

Refer: Note on Minkowski Inequality

(f) **Support function:** for any set C (convex or not), its support function

$$I_C^*(x) = \max_{y \in C} x^T y$$

is convex.

Solution: For this part and one that follows, we first establish If f_1 and f_2 are convex functions, then their pointwise maximum f defined by $f(x) = \max\{f_1(x), f_2(x)\}$, with $\text{dom } f = \text{dom } f_1 \cap \text{dom } f_2$, is also convex.
This can be verified as:

$$\begin{aligned} f(\theta x + (1-\theta)y) &\leq \max\{f_1(\theta x + (1-\theta)y), f_2(\theta x + (1-\theta)y)\} \\ &\leq \max\{\theta f_1(x) + (1-\theta)f_1(y), \theta f_2(x) + (1-\theta)f_2(y)\} \\ &\leq \theta \max\{f_1(x), f_2(x)\} + (1-\theta)\max\{f_1(y), f_2(y)\} \\ &= \theta f(x) + (1-\theta)f(y) \end{aligned}$$

which establishes convexity of f.
This result can easily be extended to state $f(x) = \max\{f_1(x), \ldots, f_m(x)\}$ is also convex if $f_1, \ldots, f_m$ are convex functions.

Now for the given support function $I_C^*(x) = \max_{y \in C} x^T y$, with $C \neq \emptyset$
For each $y \in C$, $x^T y$ is a linear function of x. Therefore, $I_C^*(x)$ is the pointwise maximum of a family of linear functions, hence convex.

Refer: Boyd, S., & Vandenberghe, L. (2004). Convex optimization. Cambridge university press. (Chapter 3)

(g) **Max function:** $f(x) = \max\{x_1, \ldots, x_n\}$ is convex.

Solution: As verified in part (f), if f_1 and f_2 are convex functions, then their pointwise maximum f defined by $f(x) = \max\{f_1(x), f_2(x)\}$ being convex, with the result being true in general for $f_1, f_2, .., f_m$; upto any m number of functions.

Here, we are given $f(x)$ as max of n linear functions: $x_i's$, It is convex since it is the pointwise maximum of affine (convex) functions.

4. Prove the following properties of convex sets.

   (a) **Intersection**: the intersection of convex sets is convex.

   Solution: Proof: Suppose A and B are convex sets. We aim to show that $A \cap B$ is also convex.
   Take $x, y \in A \cap B$ (i.e., x, y belong to both A and B). Now, consider $\lambda \in [0, 1]$.
   Since $x, y \in A$ and A is convex, $\lambda x + (1 - \lambda)y$ must also belong to A because A is convex.
   Similarly, since $x, y \in B$ and B is convex, $\lambda x + (1 - \lambda)y$ must also belong to B because B is convex.
   Therefore, $\lambda x + (1 - \lambda)y$ belongs to both A and B, implying that $\lambda x + (1 - \lambda)y$ belongs to $A \cap B$.
   Hence, by definition, $A \cap B$ is convex.

   This result can be extended to the intersection of an infinite number of sets: if S_α is convex for every $\alpha \in A$, then $\bigcap_{\alpha \in A} S_\alpha$ is convex.

   Refer: Boyd, S., & Vandenberghe, L. (2004). Convex optimization. Cambridge university press. (Chapter 2)

   (b) **Scaling and translation**: if C is convex, then

   $$aC + b = \{ax + b : x \in C\}$$

   is convex for any a, b.

   Solution: Proof: Given C is a convex set. We aim to show that for any a, b, the set $aC + b = \{ax + b : x \in C\}$ is convex.
   Let $x, y \in aC + b$, i.e., $x = ax_1 + b$ and $y = ax_2 + b$ for some $x_1, x_2 \in C$ since x and y are elements of $aC + b$.
   For any $\lambda \in [0, 1]$, we consider:

   $$\lambda x + (1 - \lambda)y = \lambda(ax_1 + b) + (1 - \lambda)(ax_2 + b) = a(\lambda x_1 + (1 - \lambda)x_2) + b$$

   Since C is convex, $\lambda x_1 + (1 - \lambda)x_2 \in C$. Hence, $a(\lambda x_1 + (1 - \lambda)x_2) + b$ is of the form $aX + b$ where $X \in C$, implying that $\lambda x + (1 - \lambda)y$ belongs to $aC + b$.
   Therefore, $aC + b$ is convex given that C is a convex set.

   (c) **Affine images and preimages**: if $f(x) = Ax + b$ and C is convex then

   $$f(C) = \{f(x) : x \in C\}$$

   is convex, and if D is convex then

   $$f^{-1}(D) = \{x : f(x) \in D\}$$

   is convex.

   Solution: Proof: For Affine Image:
   Let $f(x) = Ax + b$ and suppose C is a convex set. We aim to show that $f(C) = \{f(x) : x \in C\}$ is convex.

Take $y_1, y_2 \in f(C)$ such that $y_1 = f(x_1)$ and $y_2 = f(x_2)$ for some $x_1, x_2 \in C$.
For any $\lambda \in [0,1]$, consider:

$$\begin{aligned}\lambda y_1 + (1-\lambda)y_2 &= \lambda f(x_1) + (1-\lambda) f(x_2)\\ &= \lambda(Ax_1+b) + (1-\lambda)(Ax_2+b)\\ &= A(\lambda x_1 + (1-\lambda)x_2) + b\end{aligned}$$

Since C is convex, $\lambda x_1 + (1-\lambda)x_2 \in C$. Hence, $A(\lambda x_1 + (1-\lambda)x_2) + b$ is of the form $f(\lambda x_1 + (1-\lambda)x_2)$, implying that $\lambda y_1 + (1-\lambda)y_2$ belongs to $f(C)$.
Therefore, $f(C)$ is convex given that C is a convex set and $f(x) = Ax + b$.

Proof: For preimage:
Let $f(x) = Ax + b$ and suppose D is a convex set. We aim to show that the inverse image of D, denoted as $f^{-1}(D) = \{x : f(x) \in D\}$, is convex.
Take $x_1, x_2 \in f^{-1}(D)$ such that $f(x_1), f(x_2) \in D$.
For any $\lambda \in [0,1]$, consider:

$$\lambda f(x_1) + (1-\lambda) f(x_2) \in D$$

Given that $f(x) = Ax + b$, we have:

$$\lambda f(x_1) + (1-\lambda) f(x_2) = \lambda(Ax_1+b) + (1-\lambda)(Ax_2+b) = A(\lambda x_1 + (1-\lambda)x_2) + b$$

Therefore, $\lambda x_1 + (1-\lambda)x_2$ belongs to $f^{-1}(D)$.
Hence, $f^{-1}(D)$ is convex given that D is a convex set and $f(x) = Ax + b$.

(d) **Perspective images and preimages:** the perspective function is $\mathbf{P} : \mathbb{R}^n \times \mathbb{R}_{++} \to \mathbb{R}^n$ (where $\mathbb{R}_{++}$ denotes positive reals),

$$\mathbf{P}(x,z) = \frac{x}{z}$$

for $z > 0$. If $C \subseteq \mathrm{dom}(\mathbf{P})$ is convex then so is $\mathbf{P}(C)$, and if D is convex then so is $\mathbf{P}^{-1}(D)$

Solution: **Proof:** Perspective Image
Consider any $y_1, y_2 \in P(C)$ where $y_1 = \frac{x_1}{z_1}$ and $y_2 = \frac{x_2}{z_2}$. To prove for any $\alpha \in [0,1]$, $y_3 = \alpha y_1 + (1-\alpha)y_2 \in f(S)$, we seek $(x_3, z_3) \in S$ such that $y_3 = \frac{x_3}{z_3}$.
Given (x_1, z_1) and (x_2, z_2) as two points in the convex set S, their convex combinations also lie in S. Assume $x_3 = \theta x_1 + (1-\theta)x_2$ and $z_3 = \theta z_1 + (1-\theta)z_2$.
The goal is to find $\theta \in [0,1]$ satisfying:

$$\alpha\frac{x_1}{z_1} + (1-\alpha)\frac{x_2}{z_2} = \frac{\theta x_1 + (1-\theta)x_2}{\theta z_1 + (1-\theta)z_2} \qquad (1)$$

It can be verified that (1) holds when:

$$\theta = \frac{\alpha z_2}{(1-\alpha)z_1 + \alpha z_2}$$

Note that since $\alpha \in [0,1]$, the θ given by the equation above lies in $[0,1]$.

Thus, we've proved that for any two points $y_1, y_2 \in f(S)$ and any $\alpha \in [0,1]$, their convex combination y_3 can be represented as $\frac{x_3}{z_3}$ where $(x_3, z_3) \in C$. This implies $y_3 \in f(C)$.
By definition, $f(C)$ is convex.

Proof: Preimage
First note that $P^{-1}(C)$ can be explicitly written as:

$$P^{-1}(C) = \left\{ (x,z) : \frac{x}{z} \in C, t > 0 \right\}$$

For any $(x_1, z_1), (x_2, z_2) \in P^{-1}(C)$, and $\alpha \in [0,1]$, we aim to show that:

$$\alpha \frac{x_1}{z_1} + (1-\alpha)\frac{x_2}{z_2} \in P^{-1}(C)$$

In other words, we want to show that:

$$y = \frac{\alpha x_1 + (1-\alpha)x_2}{\alpha z_1 + (1-\alpha)z_2} \in C$$

Assume y can be represented as $y = \theta \frac{x_1}{z_1} + (1-\theta)\frac{x_2}{z_2}$, where $\theta \in [0,1]$. Then the goal is to find $\theta \in [0,1]$ such that the following equality holds:

$$\begin{aligned} &\alpha \left(\frac{1}{\alpha z_1 + (1-\alpha)z_2} \right) x_1 + (1-\alpha) \left(\frac{1}{\alpha z_1 + (1-\alpha)z_2} \right) x_2 \\ &= \theta \left(\frac{x_1}{z_1} \right) + (1-\theta) \left(\frac{x_2}{z_2} \right) \end{aligned}$$

Matching the coefficients of x_1 and x_2 in the equation above gives:

$$\theta = \frac{\alpha z_1}{(1-\alpha)z_2 + \alpha z_1}$$

Since $\alpha \in [0,1]$, $\theta \in [0,1]$. Hence, y is a convex combination of $\frac{x_1}{z_1}$ and $\frac{x_2}{z_2}$, which are in C.
Due to the assumption that C is convex, it follows that $y \in C$, and thus $P^{-1}(C)$ is a convex set.

(e) **Linear-fractional images and preimages:** the perspective map composed with an affine function,

$$f(x) = \frac{Ax+b}{c^T x + d}$$

is called a **linear-fractional function**, defined on $c^T x + d > 0$. If $C \subseteq \text{dom}(f)$ is convex then so is $f(C)$, and if D is convex then so is $f^{-1}(D)$

Solution: Proof: Image
The intuition behind the proof is that $f(x)$ is a perspective transform of an affine function. Define:

$$g(x) = \begin{pmatrix} A \\ c^T \end{pmatrix} x + \begin{pmatrix} b \\ d \end{pmatrix}$$

Now sine $g(x)$ is convex, the image of C under this mapping is convex i.e. $g(C)$ is convex. Now, upon taking perspective we get $f(C) = \text{perspective}(g(C))$.
Since perspective preserves convexity $\Rightarrow f(C)$ is convex.

Proof: Pre-Image
Since, $f(x) = \textbf{perspective}(g(x))$ (where g is convex).

Now, since we know the inverse image of a convex set under the perspective function is also convex, which directly leads us to conclude $f^{-1}(D)$ is also convex.

5. Prove the following properties of convex functions

(a) **Nonnegative linear combination:** $f_1, \ldots, f_m$ convex implies $a_1 f_1 + \ldots + a_m f_m$ convex for any $a_1, \ldots, a_m \geq 0$

Solution: **Proof:** Let $f_1(x)$ and $f_2(x)$ be two convex functions defined on a convex set $C \subset \mathbb{R}^n$. We aim to show that $f(x) = f_1(x) + f_2(x)$ is also convex.

To prove convexity, consider $x, y \in C$ and $\lambda \in [0,1]$.

By the definition of convexity for f_1 and f_2:

$$\begin{aligned}
&\text{For any } x, y \in C \text{ and } \lambda \in [0,1],\\
&f_1(\lambda x + (1-\lambda)y) \leq \lambda f_1(x) + (1-\lambda) f_1(y)\\
&f_2(\lambda x + (1-\lambda)y) \leq \lambda f_2(x) + (1-\lambda) f_2(y)
\end{aligned}$$

Now, consider the function $f(x) = f_1(x) + f_2(x)$:

$$\begin{aligned}
f(\lambda x + (1-\lambda)y) &= f_1(\lambda x + (1-\lambda)y) + f_2(\lambda x + (1-\lambda)y)\\
&\leq \lambda f_1(x) + (1-\lambda) f_1(y) + \lambda f_2(x) + (1-\lambda) f_2(y)\\
&= \lambda (f_1(x) + f_2(x)) + (1-\lambda)(f_1(y) + f_2(y))\\
&= \lambda f(x) + (1-\lambda) f(y)
\end{aligned}$$

This confirms that for any $x, y \in C$ and $\lambda \in [0,1]$, $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$, satisfying the definition of convexity.

Hence, the sum of two convex functions $f_1(x)$ and $f_2(x)$, i.e., $f(x) = f_1(x) + f_2(x)$, is also convex.

Now, since multiplying $f_i(x)$ by some $a_i \geq 0$, would preserve its convexity. Thus, the result of addition of two convex functions can easily be extended for any m.

(b) **Pointwise maximization:** if f_s is convex for any $s \in S$, then $f(x) = \max_{s \in S} f_s(x)$ is convex. Note that the set S here (number of functions f_s) can be infinite

Solution: **Proof:** If f_1 and f_2 are convex functions, then their pointwise maximum f defined by $f(x) = \max\{f_1(x), f_2(x)\}$, with $\mathrm{dom}\, f = \mathrm{dom}\, f_1 \cap \mathrm{dom}\, f_2$, is also convex. This can be verified as:

$$\begin{aligned}
f(\theta x + (1-\theta)y) &\leq \max\{f_1(\theta x + (1-\theta)y), f_2(\theta x + (1-\theta)y)\}\\
&\leq \max\{\theta f_1(x) + (1-\theta) f_1(y), \theta f_2(x) + (1-\theta) f_2(y)\}\\
&\leq \theta \max\{f_1(x), f_2(x)\} + (1-\theta) \max\{f_1(y), f_2(y)\}\\
&= \theta f(x) + (1-\theta) f(y)
\end{aligned}$$

which establishes convexity of f.

This result can easily be extended to state $f(x) = \max\{f_1(x), \ldots, f_m(x)\}$ is also convex if $f_1, \ldots, f_m$ are convex functions, where the number of elements m in S can be infinite.

Note: Here the set S being infinite refers to the cardinality (size) of a set that can be put into a one-to-one correspondence with the natural numbers (1, 2, 3, ...), thus implies countable infinity.

(c) **Partial minimization:** if $g(x,y)$ is convex in x,y, and C is convex, then $f(x) = \min_{y\in C} g(x,y)$ is convex

Solution: **Proof:** We prove this by verifying Jensen's inequality for $x_1, x_2 \in \text{dom } g$. Let $\varepsilon > 0$. Then there are $y_1, y_2 \in C$ such that $g(x_i, y_i) \leq f(x_i)$ for $i = 1, 2$. Now let $\theta \in [0,1]$. We have

$$\begin{aligned} f(\theta x_1 + (1-\theta)x_2) &= min_{y\in C} g(\theta x_1 + (1-\theta)x_2, y) \\ &\leq g(\theta x_1 + (1-\theta)x_2, \theta y_1 + (1-\theta)y_2) \\ &\leq \theta g(x_1, y_1) + (1-\theta)g(x_2, y_2) \\ &\leq \theta f(x_1) + (1-\theta) f(x_2). \end{aligned}$$

Thus by definition, $f(x)$ is convex.

Refer: Boyd, S., & Vandenberghe, L. (2004). Convex optimization. Cambridge university press. (Chapter 3)