# Homework 1

Optimization, 18-660/18-460
**Out:** Jan 20, 2026
**Due:** Feb. 5, 2026 by 11:59 pm EST

## Instructions

- **Collaboration Policy:** See syllabus for collaboration policy. You need to fill out the declaration page at the end of this PDF and append it to your homework submission. ***Failure to submit the declaration page will lead to a panelty on assignment credit.***
- **Submission Instructions:** Assignments should be submitted as pdfs on Gradescope. Each problem should be on a separate page. When submitting to Gradscope, make sure to label each question using the interface provided by Gradescope. If you do not do this, your assignment WILL NOT be graded!
- **Programming:** Programming assignments should also be submitted to Gradescope. If you are not planning to use MATLAB or Python for your submission, please clear your choice with the instructors before submitting.
- **Late Policy:** You may use up to 3 late days on a single assignment. The maximum total number of late days allowed across all assignments is 5.

## Problem 1: Convex Functions [20 pts]

1. **[20 pts]** (Examples of convex functions) Show whether each of the following functions is $\mu$-strongly convex/concave for some $\mu > 0$, strictly convex/concave, convex/concave, or neither convex nor concave. You must identify the most precise category, i.e. when a function is both convex and strictly convex, you should identify the function as strictly convex.

   (a) $f(x) = \frac{1}{x^2}$ over domain $(0, \infty)$.

   (b) $f(x) = \min\{x, 1 - x\}$ over domain $\mathbb{R}$.

   (c) A McCulloch-Pitts neuron with ReLU activation: $f(x) = (\theta^T x)_+$ over domain $\mathbb{R}^n$. Here $\theta \in \mathbb{R}^n$ is a constant vector and notation $(y)_+$ means $\max(y, 0)$. These "neurons" are a fundamental component of neural networks.

   (d) A McCulloch-Pitts neuron with sigmoid activation: $f(x) = \text{sigm}(\theta^T x)$ over domain $\mathbb{R}^n$. Here $\theta \in \mathbb{R}^n$ is a constant vector and $\text{sigm}(x) = \frac{1}{1+e^{-x}}$ is the sigmoid function.

## Problem 2: Convex Sets [30 pts]

1. **[18 pts]** (Properties of convex sets) Let $S \subseteq \mathbb{R}^n$ and $T \subseteq \mathbb{R}^n$ denote convex sets. For each of the following, prove the set is convex or provide a counter example.

   (a) Intersection of sets: $S \cap T$

   (b) Union of sets: $S \cup T$

   (c) Set difference: $S \setminus T = \{x | x \in S, x \notin T\}$

2. **[12 pts]** (Examples) For each of the following sets $S$, prove that they are either convex or not.

   (a) $S = \{x \in \mathbb{R}^n : ||x||_2 \leq 1\}$.

   (b) Let $A, B \subseteq \mathbb{R}^n$ where $A$ is convex. $S = \{x | x + B \subseteq A\}$, where $x + B$ means $\{x + y : y \in B\}$.

   (c) Rank-constrained matrices: $S = \{X \in \mathbb{S}^n \mid \text{rank}(X) = k\}$ for some $k \in \mathbb{Z}$ such that $0 < k < n$. (Here recall $\mathbb{S}^n$ is the set of $n$-by-$n$ symmetric matrices)

   (d) $S = \{x \in \mathbb{R}^n \mid \text{The number of non-zero entries in } x \text{ is no more than } k\}$ for some $k \in \mathbb{Z}$ such that $0 < k < n$.

## Problem 3: Strong Convexity, Smoothness [30 pts]

Let $f$ be a convex, twice continuously differentiable function over $\mathbb{R}^n$ .

1. **[6 pts]** Show that $f$ is convex if and only if $(\nabla f(x)-\nabla f(y))^T(x-y) \geq 0$. You can use this fact later if needed. (Hint: use the first order condition for both $(x,y)$ and $(y,x)$, and you may use the fundamental theorem of calculus. )

2. **[12 pts, optional for 18-460 students]** Show that the following statements are equivalent: (Hint: to prove these 4 conditions are equivalent, one possible way is to prove (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv), (iv) $\Rightarrow$ (ii), (iii) $\Rightarrow$ (i). You may use Cauchy–Schwarz and the second-order Taylor expansion with residual: $f(y) = f(x) + \nabla f(x)^T(y-x) + \frac{1}{2}(y-x)^T\nabla^2 f(\xi)(y-x)$ for some $\xi = x + t(y-x)$ with $t \in (0,1)$. )

    i. $\|\nabla f(x) - \nabla f(y)\|_2 \leq L\|x-y\|_2$ for all $x, y$.

    ii. $(\nabla f(x) - \nabla f(y))^T(x-y) \leq L||x-y||_2^2$ for all $x, y$.

    iii. $\nabla^2 f(x) \preceq LI$ for all $x$, where $I$ denotes the $n \times n$ identity matrix

    iv. $f(y) \leq f(x) + \nabla f(x)^T(y-x) + \frac{L}{2}||y-x||_2^2$ for all $x, y$

3. **[12 pts, optional for 18-460 students]** Show that the following statements are equivalent: (Hint: use the fact that $\mu$-strongly convex means $f(x) - \frac{\mu}{2}\|x\|^2$ is convex and apply conditions for convexity. )

    i. $f$ is strongly convex with constant $\mu$

    ii. $(\nabla f(x) - \nabla f(y))^T(x-y) \geq \mu||x-y||_2^2$ for all $x, y$.

    iii. $\nabla^2 f(x) \succeq \mu I$ for all $x$, where $I$ denotes the $n \times n$ identity matrix

    iv. $f(y) \geq f(x) + \nabla f(x)^T(y-x) + \frac{\mu}{2}||y-x||_2^2$ for all $x, y$

## Problem 4: Programming with CVX [20 pts]

CVX is a tool that was created to make solving optimization problems easy. It is not particularly fast (compared to other, more specialized solvers), but its versatility makes it a great educational tool. In this exercise, you will set up CVX and solve some simple optimization problems.

Download the CVX variant of your choosing:

- Matlab - http://cvxr.com/cvx/
- Python - http://www.cvxpy.org/en/latest/

Read over the documentation, and familiarize yourself with CVX's operation. You should try out some basic problems (e.g., least squares) to make sure you understand how CVX works and that it matches analytical solutions to the problem.

We will use CVX to solve a simple resource allocation problem given as follows,

$$
\begin{aligned}
&\min_{x_1,\ldots,x_n} \sum_{i=1}^{n} f_i(x_i) \\
\text{s.t.} \quad & x_i \geq 0, \forall i = 1, 2, \ldots, n \\
& x_1 + x_2 + \ldots x_n \leq D.
\end{aligned}
$$

1. **[10 pts]** Set $n = 10$ and $D = 10$, and set $f_i(x_i) = -ix_i$. Solve the problem using CVX and plot the values of $x_i$ as a function of $i$.

2. **[10 pts]** In the previous problem, you will find all resources will concentrate on one single $x_i$. To promote fairness, one way is to incorporate a log function, i.e. change the optimization problem to

$$
\begin{aligned}
&\min_{x_1,\ldots,x_n} \sum_{i=1}^{n} f_i(x_i) - \tau \sum_{i=1}^{n} \log x_i \\
\text{s.t.} \quad & x_i \geq 0, \forall i = 1, 2, \ldots, n \\
& x_1 + x_2 + \ldots x_n \leq D,
\end{aligned}
$$

where $n$, $D$, and $f_i$ are the same as the previous problem. Try at least three different positive values of $\tau$, solve the problem and plot $x_i$ again. Describe what you see.

## Declaration of External Resources

Please fill out the following information with handwriting/typing and append this page to your homework submission. ***Failure to submit this declaration page will lead to a panelty on assignment credit.*** You can use extra sheets of paper to answer the following questions if you need more space.

1. Have you collaborated with other students? If yes, name the collaborators and briefly describe the nature of collaboration.

Yes ☐ No ☐

2. Have you used GenAI in helping solve the problem? If yes, provide a brief description of how you used GenAI. You can also optionally append your chat history to your homework submission.

Yes ☐ No ☐

3. Cite all other external resources you used to finish the assignment.