EECS 127/227AT Optimization Models in Engineering UC Berkeley Spring 2023

Homework 6

Self grades are due at 11 PM on March 9, 2023.

1. **Midsemester Survey**

   Please complete this mid-semester survey at the following link: link. You will get a code at the end of the survey; write it in as the solution for this problem.

2. Convex or Concave

Determine whether the following functions are convex, strictly convex, concave, strictly concave, both or neither.

(a) $f(x) = \mathrm{e}^x - 1$ on $\mathbb{R}$.

Solution: $f(x) = \mathrm{e}^x - 1$ on $\mathbb{R}$.
This is strictly convex since $\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x) = \mathrm{e}^x > 0$ for all $x \in \mathbb{R}$.

(b) $f(x_1, x_2) = x_1 x_2$ on $\mathbb{R}^2_{++}$ (i.e. when $x_1 > 0$ and $x_2 > 0$).

Solution: $f(x_1, x_2) = x_1 x_2$ on $\mathbb{R}^2_{++}$.
This is neither convex nor concave. The Hessian of f is

$$\nabla^2 f(x) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{1}$$

which has eigenvalues ± 1 which implies the Hessian is neither positive semidefinite nor negative semidefinite.

(c) The log-likelihood of a set of points $\{x_1, \ldots, x_n\}$ that are normally distributed with mean μ and finite variance $\sigma > 0$ is given by:

$$f(\mu, \sigma) = n \log\left(\frac{1}{\sqrt{2\pi}\sigma}\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2 \tag{2}$$

i. Show that if we view the log likelihood for fixed σ as a function of the mean, i.e

$$g(\mu) = n \log\left(\frac{1}{\sqrt{2\pi}\sigma}\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2 \tag{3}$$

then g is strictly concave (equivalently, we say f is strictly concave in μ).

ii. **(OPTIONAL)** Show that if we view the log likelihood for fixed μ as a function of the inverse of the variance, i.e

$$h(z) = n \log\left(\frac{\sqrt{z}}{\sqrt{2\pi}}\right) - \frac{z}{2}\sum_{i=1}^{n}(x_i - \mu)^2 \tag{4}$$

then h is strictly concave (equivalently, we say f is strictly concave in $z = \frac{1}{\sigma^2}$). Note that we have used the dummy variable z to denote $\frac{1}{\sigma^2}$.

iii. **(OPTIONAL)** Show that f is not jointly concave in $\mu, \frac{1}{\sigma^2}$. *HINT: We say a function $w(x, y)$ with $x \in \mathcal{R}^m$ and $y \in \mathcal{R}^n$ is jointly convex if*

$$w\left(\lambda(x_1, y_1) + (1-\lambda)(x_2, y_2)\right) \leq \lambda w((x_1, y_1)) + (1-\lambda) w((x_2, y_2)). \tag{5}$$

This is the same as letting $z = (x, y)$ and saying f is convex in z. We can define joint concavity in a similar fashion by reversing the inequalities.

Solution: For $g(\mu)$ we have,

$$\nabla g(\mu) = \sum_{i=1}^{n} \frac{x_i - \mu}{\sigma^2} \tag{6}$$

$$\nabla^2 g(\mu) = -\frac{n}{\sigma^2} < 0. \tag{7}$$

Since σ is finite, g is strictly concave (equivalently f is strictly concave in μ).
For $h(z)$ we have,

$$\nabla h(z) = \frac{n}{2z} - \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{2} \tag{8}$$

$$\nabla^2 h(z) = -\frac{n}{2z^2} < 0. \tag{9}$$

Since z^2 is finite ($\sigma > 0$), h is strictly concave (equivalently f is strictly concave in σ^2). For $f(\mu, \frac{1}{\sigma^2})$, we find the second order partial derivatives and stack them in the Hessian. We have,

$$\nabla^2 f(\mu, \frac{1}{\sigma^2}) = \begin{bmatrix} -\frac{n}{\sigma^2} & \sum_{i=1}^{n}(x_i - \mu) \\ \sum_{i=1}^{n}(x_i - \mu) & -\frac{n\sigma^4}{2} \end{bmatrix}. \tag{10}$$

The determinant of the Hessian is given by,

$$\det\left(\nabla^2 f\right) = \frac{n^2\sigma^2}{2} - (\sum_{i=1}^{n}(x_i - \mu))^2. \tag{11}$$

and the trace of the Hessian is given by,

$$\operatorname{tr}\left(\nabla^2 f\right) = -\frac{n}{\sigma^2} - \frac{n\sigma^4}{2} < 0 \tag{12}$$

Note that the trace is the sum of the eigenvalues, and the determinant is the product of the eigenvalues. Since the trace is always negative, if the determinant is negative it must imply that one eigenvalue is positive and another is negative; that is, we have f is neither convex nor concave. It is easy to see that $\det\left(\nabla^2 f\right)$ can sometimes be negative – for example, if we choose σ^2 to be close to zero and μ away from x_i, the second negative term dominates and make $\det\left(\nabla^2 f\right) \leq 0$. **Aside:** Note however, in the maximum likelihood estimates, the Hessian is negative semi-definite implying that locally the function is concave. More concretely, at

$$\widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n}, \quad \widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \widehat{\mu})^2 \tag{13}$$

we have $\nabla^2 f(\widehat{\mu}, 1/\widehat{\sigma}^2) \preceq 0$

(d) $f(x) = \log(1 + \mathrm{e}^x)$. Note that this implies that $g(x) = -f(x) = \log\left(\frac{1}{1+\mathrm{e}^x}\right)$ is concave. Compare this to $h(x) = \frac{1}{1+\mathrm{e}^x}$, is $h(x)$ convex or concave?

Solution: We will do this by verifying the second order sufficient conditions for convexity. We have the derivatives of f can be computed using the chain rule as follows:

$$f'(x) = \frac{\partial f}{\partial x}(x) = \frac{\mathrm{e}^x}{1 + \mathrm{e}^x}$$

$$f''(x) = \frac{\partial^2 f}{\partial x^2}(x) = \frac{\mathrm{e}^x}{1 + \mathrm{e}^x} + \frac{-\mathrm{e}^x}{(1 + \mathrm{e}^x)^2}\mathrm{e}^x = \frac{\mathrm{e}^x}{(1 + \mathrm{e}^x)^2} > 0.$$

Since we have $f''(x) > 0$ for all x, we conclude that the function f is strictly convex.
Now consider $h(x) = \frac{1}{1+\mathrm{e}^x}$. We use the second order condition for convexity, and calculate

$$\nabla h(x) = \frac{-\mathrm{e}^x}{(1 + \mathrm{e}^x)^2}; \quad \nabla^2 h(x) = \frac{(\mathrm{e}^x - 1)\mathrm{e}^x}{(\mathrm{e}^x + 1)^3}.$$

The second derivative is positive for $x > 0$, and negative for $x < 0$, hence the function is neither convex nor concave.

3. Further characterizations of convexity

Show that $\sigma_1 : \mathbb{R}^{m\times n} \to \mathbb{R}_+$, the function that maps a matrix to its largest singular value, is a convex function, with domain $\mathbb{R}^{m\times n}$.

HINT: You may express $\sigma_1(A)$ using the ℓ^2 operator norm of A:

$$\sigma_1(A) = \max_{\vec{x}\in\mathbb{R}^n : \|\vec{x}\|_2 = 1} \|A\vec{x}\|_2 ,$$

This question proves that this norm is convex, so you may not use the fact that norms are convex.

Solution: We have

$$\sigma_1(A) = \max_{\vec{x}\in\mathbb{R}^n : \|\vec{x}\|_2 = 1} \|A\vec{x}\|_2 ,$$

which is the characterization of the largest singular value of a matrix as its induced ℓ^2 norm. Note that this expresses the function

$$\sigma_1 : \mathbb{R}^{m\times n} \to \mathbb{R}_+,$$

as the supremum of a family of function, one for each $x \in \mathbb{R}^n$ with $\|\vec{x}\|_2 = 1$, which we may temporarily call ψ_x for convenience, given by

$$\psi_x(A) \coloneqq \|A\vec{x}\|_2 \, \forall A \in \mathbb{R}^{m\times n}.$$

If we could prove that each $\psi_x : \mathbb{R}^{m\times n} \to \mathbb{R}_+$ is convex then we could be done, because the supremum of a family of convex functions is convex.

To show that $\psi_x(\cdot)$ is convex, consider two matrices $A, B \in \mathbb{R}^{m\times n}$ and their linear combination $\theta A + (1-\theta)B$, where $\theta \in [0, 1]$.

$$\begin{aligned}\psi_x(\theta A + (1-\theta)B) &= \|(\theta A + (1-\theta)B)\vec{x}\|_2 = \|\theta A\vec{x} + (1-\theta)B\vec{x}\|_2 \\ &\le \theta \|A\vec{x}\|_2 + (1-\theta)\|B\vec{x}\|_2 = \theta\psi_x(A) + (1-\theta)\psi_x(B),\end{aligned}$$

where the inequality comes from the triangle inequality on the ℓ^2 norm. Hence, $\psi_x(\cdot)$ is convex.

4. Convex and strictly convex functions

(a) Recall that a function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be strictly convex if it satisfies Jensen's inequality with strict inequality, i.e., $\forall \vec{x} \neq \vec{y} \in \mathbb{R}^n$ and $\forall t \in (0,1)$, we have

$$f(t\vec{x} + (1-t)\vec{y}) < tf(\vec{x}) + (1-t)f(\vec{y})$$

Show that for a strictly convex function $f : \mathbb{R}^n \to \mathbb{R}$, the problem

$$\min_{\vec{x} \in \mathbb{R}^n} f(\vec{x}) \tag{14}$$

has at most one solution.

HINT: Try to argue by contradiction assuming that there are two solutions $\vec{x}_1, \vec{x}_2$ which achieve the minimum value. Argue that using these two points you can find another point in $\mathbb{R}^n$ with strictly smaller function value.

Solution: Assume that $f^* = \min_{\vec{x} \in \mathbb{R}^n} f(\vec{x})$ has at least two different optimal solutions $\vec{x}_1, \vec{x}_2 \in \mathbb{R}^n$. Hence $f^* = f(\vec{x}_1) = f(\vec{x}_2)$. Consider $\vec{z} = (\vec{x}_1 + \vec{x}_2)/2$.

$$\begin{aligned} f(\vec{z}) &= f((\vec{x}_1 + \vec{x}_2)/2) \\ &< (f(\vec{x}_1) + f(\vec{x}_2))/2 \\ &= f^* \end{aligned}$$

where the inequality follows from strict convexity. Hence we've shown that $\vec{z}$ has functional value strictly smaller than f^* and hence f^* was not the optimal value giving us a contradiction.

(b) Prove that for all convex optimization problems $\min_{\vec{x} \in \mathcal{X}} f(\vec{x})$, where f is a convex function and $\mathcal{X}$ is a convex set, all local minima are global minima. You may not assume that f is differentiable.

HINT: Start with assuming $\vec{x}^\star$ is a local minimum that is not global, and $\vec{\tilde{x}}$ is a global minimum. Use the definition of the convexity of a function to prove by contradiction.

Solution: To arrive at a contradiction, suppose $\vec{x}^\star$ is a local minimum that is not global. Let $\vec{\tilde{x}}$ be a global minimum. Thus we have $f(\vec{\tilde{x}}) < f(\vec{x}^\star)$. Then by convexity $\lambda\vec{x}^\star + (1-\lambda)\vec{\tilde{x}} \in \mathcal{X}$ and hence

$$\begin{aligned} f(\lambda\vec{x}^\star + (1-\lambda)\vec{\tilde{x}}) &\leq \lambda f(\vec{x}^\star) + (1-\lambda) f(\vec{\tilde{x}}) \\ &< \lambda f(\vec{x}^\star) + (1-\lambda) f(\vec{x}^\star) \\ &= f(\vec{x}^\star) \end{aligned}$$

Thus, for all $\lambda \in [0,1)$, we have $f(\lambda\vec{x}^\star + (1-\lambda)\vec{\tilde{x}}) < f(\vec{x}^\star)$. Then we can make $\lambda\vec{x}^\star + (1-\lambda)\vec{\tilde{x}}$ arbitrarily close to $\vec{x}^\star$, contradicting that $\vec{x}^\star$ is a local minimum.

5. Direction of Steepest Ascent

For a differentiable function $f : \mathbb{R}^n \mapsto \mathbb{R}$ we want to show that the gradient $\nabla f(\vec{x})$ is the direction of steepest ascent at the point $\vec{x}$.

(a) Let us define the rate of change of the function $f(\vec{x})$ at the point $\vec{x}$ along an arbitrary unit vector $\vec{u}$ as:

$$D_{\vec{u}} f(\vec{x}) = \lim_{h \to 0} \frac{f(\vec{x} + h\vec{u}) - f(\vec{x})}{h}. \tag{15}$$

We call this the directional derivative. Show that the directional derivative can be equivalently expressed as $D_{\vec{u}} f(\vec{x}) = \vec{u}^\top [\nabla f(\vec{x})]$.

HINT: Use Taylor approximation of the function around the point $\vec{x}$ and evaluate it at the point $\vec{x} + h\vec{u}$.

Solution: Using Taylor's theorem we can express the function $f(\vec{x})$ as

$$f(\vec{x} + h\vec{u}) = f(\vec{x}) + [\nabla f(\vec{x})]^\top [h\vec{u}] + o(h). \tag{16}$$

We rearrange the terms, and dividing both sides by h we get

$$\frac{f(\vec{x} + h\vec{u}) - f(\vec{x})}{h} = [\nabla f(\vec{x})]^\top [\vec{u}] + \frac{o(h)}{h}. \tag{17}$$

Now we take the limit of both sides as $h \to 0$; we get

$$\lim_{h \to 0} \frac{f(\vec{x} + h\vec{u}) - f(\vec{x})}{h} = [\nabla f(\vec{x})]^\top [\vec{u}] + \lim_{h \to 0} \left(\frac{o(h)}{h} \right) \tag{18}$$

$$= [\nabla f(\vec{x})]^\top [\vec{u}]. \tag{19}$$

Note that $\lim_{h \to 0} \frac{o(h)}{h} = 0$ because $o(h)$ decays faster than h as $h \to 0$.

(b) Show that

$$\frac{\nabla f(\vec{x})}{\|\nabla f(\vec{x})\|_2} = \underset{\|\vec{u}\|_2 = 1}{\operatorname{argmax}} \, \vec{u}^\top [\nabla f(\vec{x})]. \tag{20}$$

Solution: Using Cauchy-Schwarz inequality we can write:

$$\vec{u}^\top [\nabla f(\vec{x})] \leq \|\vec{u}\|_2 \|\nabla f(\vec{x})\|_2 \tag{21}$$

$$= \|\nabla f(\vec{x})\|_2 , \tag{22}$$

so the maximum value that the expression can take is $\|\nabla f(\vec{x})\|_2$. Now it remains to show that this value is attained for the choice $\vec{u} = \frac{\nabla f(\vec{x})}{\|\nabla f(\vec{x})\|_2}$.

$$\frac{[\nabla f(\vec{x})]^\top}{\|\nabla f(\vec{x})\|_2} \nabla f(\vec{x}) = \frac{\|\nabla f(\vec{x})\|_2^2}{\|\nabla f(\vec{x})\|_2} \tag{23}$$

$$= \|\nabla f(\vec{x})\|_2 . \tag{24}$$

6. Gradient Descent Algorithm

Given a continuous and differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, the gradient of f at any point $\vec{x}$, $\nabla f(\vec{x})$, is orthogonal to the level curve of f at point $\vec{x}$, and it points in the increasing direction of f (as you showed in the last question). In other words, moving from point $\vec{x}$ in the direction $\nabla f(\vec{x})$ leads to an increase in the value of f, while moving in the direction of $-\nabla f(\vec{x})$ decreases the value of f. This idea gives an iterative algorithm to minimize the function f: the gradient descent algorithm.

(a) Consider $f(x) = \frac{1}{2}(x-2)^2$, and assume that we use the gradient descent algorithm:

$$x_{k+1} = x_k - \eta \nabla f(x_k) \quad \forall k \geq 0, \tag{25}$$

with some random initialization x_0, where $\eta > 0$ is the step size (or the learning rate) of the algorithm. Write $(x_k - 2)$ in terms of $(x_0 - 2)$, and show that x_k converges to 2, which is the unique minimizer of f, when $\eta = 0.2$.

Solution: For the given function, we have $\nabla f(x) = (x-2)$; therefore, the gradient descent algorithm gives

$$x_{k+1} = x_k - \eta(x_k - 2). \tag{26}$$

By subtracting 2 from both sides, we obtain

$$(x_{k+1} - 2) = (1-\eta)(x_k - 2) \implies (x_k - 2) = (1-\eta)^k (x_0 - 2). \tag{27}$$

Given $\eta = 0.2$, we have

$$|x_k - 2| = 0.8^k \, |x_0 - 2| \to 0 \quad \text{as } k \to \infty, \tag{28}$$

which shows that x_k converges to 2.

(b) What is the largest value of η that we can use so that the gradient descent algorithm converges to 2 from all possible initializations in $\mathbb{R}$? What happens if we choose a larger step size?

Solution: From the solution for part (a), we have

$$|x_k - 2| = |1-\eta|^k \, |x_0 - 2| \quad \forall k \in \mathbb{N}. \tag{29}$$

For convergence of the algorithm for every initialization, it is necessary and sufficient to have $|1-\eta| < 1$, which is equivalent to $\eta \in (0, 2)$. If $\eta = 2$, x_k oscillates around 2 while $|x_k - 2|$ remains fixed. If $\eta > 2$, x_k oscillates around 2 while $|x_k|$ grows unboundedly.

(c) Now assume that we use the gradient descent algorithm to minimize $f(\vec{x}) = \frac{1}{2}\left\|A\vec{x} - \vec{b}\right\|_2^2$ for some $A \in \mathbb{R}^{m \times n}$ and $\vec{b} \in \mathbb{R}^m$, where A has full column rank. First compute $\nabla f(\vec{x})$. Note that $(A^\top A)^{-1} A^\top \vec{b}$ is the solution to the least-squares problem, and $(\vec{x}_k - (A^\top A)^{-1} A^\top \vec{b})$ is the distance from the solution at time k. Write $(\vec{x}_k - (A^\top A)^{-1} A^\top \vec{b})$ in terms of $(\vec{x}_0 - (A^\top A)^{-1} A^\top b)$.

Solution: We can write $f(\vec{x}) = \frac{1}{2}(\vec{x}^\top A^\top A \vec{x} - \vec{x}^\top A^\top \vec{b} - \vec{b}^\top A \vec{x} + \vec{b}^\top \vec{b})$, so

$$\nabla f(\vec{x}) = A^\top A \vec{x} - A^\top \vec{b}. \tag{30}$$

Then the gradient descent algorithm gives

$$\vec{x}_{k+1} = \vec{x}_k - \eta \left(A^\top A\vec{x}_k - A^\top b\right) = \vec{x}_k - \eta A^\top A \left(\vec{x}_k - (A^\top A)^{-1}A^\top \vec{b}\right). \tag{31}$$

By subtracting $(A^\top A)^{-1}A^\top \vec{b}$ from both sides, we obtain

$$\left(\vec{x}_{k+1} - (A^\top A)^{-1}A^\top \vec{b}\right) = \left(I - \eta A^\top A\right)\left(\vec{x}_k - (A^\top A)^{-1}A^\top \vec{b}\right) \tag{32}$$

and consequently,

$$\left(\vec{x}_k - (A^\top A)^{-1}A^\top \vec{b}\right) = \left(I - \eta A^\top A\right)^k \left(\vec{x}_0 - (A^\top A)^{-1}A^\top \vec{b}\right). \tag{33}$$

(d) Now consider $f(\vec{x}) = \frac{1}{2}\left\|A\vec{x} - \vec{b}\right\|_2^2 + \frac{1}{2}\lambda \|\vec{x}\|_2^2$ for some $A \in \mathbb{R}^{m\times n}$ and $\vec{b} \in \mathbb{R}^m$, where A has full column rank. Suppose we solve this problem via gradient descent with step-size $\eta = \frac{1}{\sigma_1^2+\lambda}$, where σ_1 is the maximum singular value of A. Show the gradient descent converges.

Solution: We can write $f(\vec{x}) = \frac{1}{2}(\vec{x}^\top A^\top A\vec{x} - \vec{x}^\top A^\top \vec{b} - \vec{b}^\top A\vec{x} + \vec{b}^\top \vec{b} + \lambda \vec{x}^\top \vec{x})$, so

$$\nabla f(\vec{x}) = A^\top A\vec{x} - A^\top \vec{b} + \lambda\vec{x} = (A^\top A + \lambda I)\vec{x} - A^\top \vec{b}$$

The least-squares solution to this problem is now $x^\star = (A^\top A + \lambda I)^{-1}A^\top \vec{b}$. Let's consider the distance from $x^\star$ at time k+1

$$\left(\vec{x}_{k+1} - (A^\top A + \lambda I)^{-1}A^\top \vec{b}\right) = \left(I - \eta(A^\top A + \lambda I)\right)\left(\vec{x}_k - (A^\top A + \lambda I)^{-1}A^\top \vec{b}\right) \tag{34}$$

and consequently,

$$\left(\vec{x}_k - (A^\top A + \lambda I)^{-1}A^\top \vec{b}\right) = \left(I - \eta(A^\top A + \lambda I)\right)^k \left(\vec{x}_0 - (A^\top A + \lambda I)^{-1}A^\top \vec{b}\right). \tag{35}$$

For the algorithm to converge, we need the largest eigenvalue of $(I - \eta(A^\top A + \lambda I))$ to be less than 1 in absolute value. We know any eigenvector v_i of $A^\top A$ is also eigenvector of $I - \eta(A^\top A + \lambda I)$ with eigenvalue $1 - \eta(\sigma_i^2 + \lambda)$. With $\eta = \frac{1}{\sigma_1^2+\lambda}$, where σ_1 is the maximum singular value of A given, the following inequality will always be true

$$\forall i,\ |1 - \eta(\sigma_i^2 + \lambda)| < 1. \tag{36}$$

Therefore, the gradient descent eventually converges to the least-squares solution $x^\star$.

7. Homework Process

With whom did you work on this homework? List the names and SIDs of your group members.

NOTE: If you didn't work with anyone, you can put "none" as your answer.