Optimization-Based Data Analysis

Recitation 10

1. Suppose $f, g : \mathbb{R} \to \mathbb{R}$ are convex and twice differentiable. Give conditions so that $f \circ g$ is convex. Does the result still hold under your conditions if we remove differentiability?

   Solution. Let $h = f \circ g$ so that

   $$h'(x) = f'(g(x))g'(x) \quad \text{and} \quad h''(x) = f''(g(x))g'(x)^2 + f'(g(x))g''(x).$$

   This is non-negative if $f'(g(x)) \geq 0$ which occurs if f is increasing.

   Now assume that f, g are just convex. Then

   $$\begin{aligned} h(tx + (1-t)y) &= f(g(tx + (1-t)y)) \\ &\leq f(tg(x) + (1-t)g(y)) \\ &\leq tf(g(x)) + (1-t)f(g(y)), \end{aligned}$$

   for $t \in (0, 1)$ where the first inequality requires the convexity of g and that f is increasing.

2. Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = e^{-x^2/2}$. Is f concave?

   Solution. No. Note that $f'(x) = -xe^{-x^2/2}$ and

   $$f''(x) = -e^{-x^2/2} + x^2 e^{-x^2/2} = -(1 - x^2)e^{-x^2/2}$$

   which changes sign at ± 1.

3. Which of the following functions $f : \mathbb{R}^2 \to \mathbb{R}$ are convex?

   (a) $f(x, y) = x^2 + y^2$

   (b) $f(x, y) = x^2 - y$

   (c) $f(x, y) = x^2 - y^2$

   (d) $f(x, y) = x^2 y^2$

   Solution.

   (a) Yes, $Hf(x, y) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$.

   (b) Yes, sum of convex functions x^2 and $-y$.

   (c) No, $Hf(x, y) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$.

(d) No, $Hf(x,y) = \begin{bmatrix} 2y^2 & 4xy \\ 4xy & 2x^2 \end{bmatrix}$ has determinant $-12x^2y^2$ which is negative when $x, y \neq 0$.

4. Which of the following convex functions are strictly convex?

   (a) $f(\vec{x}) = \|\vec{x}\|_1$ for $\vec{x} \in \mathbb{R}^n$ with $n > 1$

   (b) $f(\vec{x}) = \|\vec{x}\|_2$ for $\vec{x} \in \mathbb{R}^n$ with $n > 1$

   (c) $f(\vec{x}) = \|\vec{x}\|_2^2$ for $\vec{x} \in \mathbb{R}^n$ with $n > 1$

   (d) $f(X) = \|X\|$ for $X \in \mathbb{R}^{n\times n}$ with $n > 1$

   (e) $f(X) = \|X\|_F$ for $X \in \mathbb{R}^{n\times n}$ with $n > 1$

   (f) $f(X) = \|X\|_*$ for $X \in \mathbb{R}^{n\times n}$ with $n > 1$

   Solution.

   (a) No, $\|\vec{e}_1 + \vec{e}_2\|_1 = \|(\vec{e}_1 + \vec{e}_2) + t(\vec{e}_1 - \vec{e}_2)\|_1$ for $t \in [0,1]$. Alternatively, note that the contour lines are straight, or that $f(x) = |x|$ is not strictly convex.

   (b) No, $\|t\vec{x}\|_2 = |t|\|\vec{x}\|_2$ for $t \in [0,1]$ (i.e., consider the line segment between 0 and $\vec{x}$). Alternatively, the function is linear on radial lines.

   (c) Yes, Hessian is $2I$.

   d,e,f) No. Apply previous parts to diagonal matrices.