OPTIMIZATION. HOMEWORK 1

OSCAR DALMAU

(1) Let $f_1(x_1, x_2) = x_1^2 - x_2^2$, $f_2(x_1, x_2) = 2x_1x_2$. Represent the level sets associated with $f_1(x_1, x_2) = 12$ and $f_2(x_1, x_2) = 16$ on the same figure using Python. Indicate on the figure, the points $\boldsymbol{x} = [x_1, x_2]^T$ for which $f(\boldsymbol{x}) = [f_1(x_1, x_2), f_2(x_1, x_2)]^T = [12, 16]^T$.

(2) Consider the function $f(\boldsymbol{x}) = (\boldsymbol{a}^T\boldsymbol{x})(\boldsymbol{b}^T\boldsymbol{x})$, where $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{x}$ are n-dimensional vectors.
   - Compute the gradient $\nabla f(\boldsymbol{x})$ and the Hessian $\nabla^2 f(\boldsymbol{x})$.

(3) Compute the gradient of

$$f(\theta) \stackrel{def}{=} \frac{1}{2}\sum_{i=1}^{n} [g(\boldsymbol{x}_i) - g(\mathbf{A}\boldsymbol{x}_i + \boldsymbol{b})]^2$$

with respect to θ, where $\theta = [a_{11}, a_{12}, a_{21}, a_{22}, b_1, b_2]^T$, $\boldsymbol{x}_i \in \mathbb{R}^2$, $\mathbf{A} \in \mathbb{R}^{2\times 2}$, $\boldsymbol{b} \in \mathbb{R}^2$ are defined as follows

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \\ b &= [b_1, b_2]^T \end{aligned}$$

and $g : \mathbb{R}^2 \to \mathbb{R} \in \mathcal{C}^1$.

(4) Let $f(r, \theta)$ be $\mathbb{R}^2 \to \mathbb{R}$ with $r = \sqrt{x^2 + y^2}$ and $\theta = \arctan \frac{y}{x}$. Compute $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$

(5) The directional derivative $\frac{\partial f}{\partial v}(x_0, y_0, z_0)$ of a differentiable function f are $\frac{3}{\sqrt{2}}$, $\frac{1}{\sqrt{2}}$ and $-\frac{1}{\sqrt{2}}$ in the directions of vectors $[0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}]^T$, $[\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}]^T$ and $[\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0]^T$. Compute $\nabla f(x_0, y_0, z_0)$.

(6) Show that the level curves of the function $f(x, y) = x^2 + y^2$ are orthogonal to the level curves of $g(x, y) = \frac{y}{x}$ for all (x, y).

(7) Let f, g, h be differentiable functions, with $f : \mathbb{R}^n \to \mathbb{R}^3$, $g : \mathbb{R}^n \to \mathbb{R}^3$ and $h : \mathbb{R}^n \to \mathbb{R}$

$$h(\boldsymbol{x}) = f(\boldsymbol{x})^T g(\boldsymbol{x})$$

show that

$$Dh(\boldsymbol{x}) = f(\boldsymbol{x})^T Dg(\boldsymbol{x}) + g(\boldsymbol{x})^T Df(\boldsymbol{x})$$

(8) Consider the **induced matrix norm**

$$\|\mathbf{A}\|_p = \max_{\boldsymbol{x}\neq\mathbf{0}} \frac{\|\mathbf{A}\boldsymbol{x}\|_p}{\|\boldsymbol{x}\|_p}$$

where $\|\cdot\|_p$ is the ℓ_p norm, ie

$$\|\boldsymbol{x}\|_p = (|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{1/p}$$

Show that

$$\|\mathbf{AB}\|_p \leq \|\mathbf{A}\|_p \|\mathbf{B}\|_p$$