Assignment-1
MA6100: Maths Behind Machine Learning

September 3, 2023

0.1 Computing gradient problems

1. Find the derivative/gradient of the following functions:

   (a) $f : \mathbb{R}^n \to \mathbb{R}$
   $$f(x) = a^\top x$$

   (b) $f : \mathbb{R}^n \to \mathbb{R}^m$
   $$f(x) = Ax$$

   (c) $f : \mathbb{R}^n \to \mathbb{R}$
   $$f(x) = x^\top Ax$$

   (d) $f : \mathbb{R}^n \to \mathbb{R}$
   $$f(x) = \frac{e^{a^\top x}}{1 + e^{a^\top x}}$$

   (e) $f : \mathbb{R}^n \to \mathbb{R}$
   $$f(x) = ||Ax - b||_2^2$$

   (f) $f : \mathbb{R}^n \to \mathbb{R}^n$
   $$f_i(x) = \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}}$$

   (g) $f : \mathbb{R}^n_{>0} \to \mathbb{R}$
   $$f(x) = \sum_{i=1}^{n} x_i \ln \frac{x_i}{\alpha_i}$$

0.2 Basic Optimization problems

1. Let $\mathcal{C} \subseteq \mathbb{R}^n$ be the solution set of the quadratic inequality
   $$\mathcal{C} = \{x \in \mathbb{R}^n | \ x^\top Ax + b^\top x + c \leq 0\}$$
   with $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$ and $c \in \mathbb{R}$

(a) Show that $\mathcal{C}$ is convex if $A \geq 0$[1]

(b) Show that the intersection of $\mathcal{C}$ and the hyperplane defined by $\alpha \cdot x + \beta = 0$ ($\alpha \neq 0$) is convex if $A + \lambda_0 \alpha\alpha^\top \geq 0$ for some $\lambda_0 \in \mathbb{R}$.

2. Show that $x = (1, \frac{1}{2}, -\frac{1}{2})$ is optimal solution for the minimization problem

$$\begin{array}{ll} \min\limits_{x} & \frac{1}{2}x^\top Ax + b^\top x + c \\ \text{s.t.} & -1 \leq x_i \leq 1, \; i = 1, 2, 3 \end{array}$$

where $A = \begin{bmatrix} 13 & 12 & -2 \\ 12 & 17 & 6 \\ -2 & 6 & 12 \end{bmatrix}$, $b = \begin{bmatrix} -22 \\ -14.5 \\ 13 \end{bmatrix}$ and $c = 1$.

3. Consider the problem of minimizing the function $f : \mathbb{R}^n \to R$ which is defined as

$$f(x) = \frac{||Ax - b||_2^2}{c^\top x + d},$$

where $Dom(f) = \{x | c^\top x + d > 0\}$. Assume that $rank(A) = r$ and $b \notin R(A)$. Show that the minimizer x^* of f is given by $x^* = x_1 + x_2$, where $x_1 = (A^\top A)^{-1}A^\top b$, $x_2 = (A^\top A)^{-1}c$ and $t \in \mathbb{R}$.

4. Suppose the convex function f satisfies

$$mI \leq \nabla^2 f(x) \leq MI \;\; (0 < m < M < \infty),$$

I being the identity function. Let ∇x be a descent direction at x, show that the backtracking stopping condition holds for

$$0 < t \leq -\frac{\nabla f(x) \cdot \nabla x}{M||x||_2^2}$$

Use this number to give an upper bound on the number of backtracking iterations.

5. Consider the quadratic objective function on $\mathbb{R}^2$

$$f(x) = \frac{1}{2}x_1^2 + rx_2^2,$$

where $r > 0$. Apply the gradient descent method with exact line search starting at a point $x(0) = (r, 1)$ and find closed form expression for $f(x_k)$.

[1] A being positive semi-definite is denoted as $A \geq 0$.