Question 1:

1. For each of the following functions $f(x_1, x_2)$, find all critical points (i.e, all x_1, x_2 such that $\nabla f(x_1, x_2) = \mathbf{0}$).

(a) $f(x_1, x_2) = (4x_1^2 - x_2)^2$
(b) $f(x_1, x_2) = 2x_2^3 - 6x_2^2 + 3x_1^2 x_2$
(c) $f(x_1, x_2) = (x_1 - 2x_2)^4 + 64x_1 x_2$
(d) $f(x_1, x_2) = x_1^2 + 4x_1 x_2 + x_2^2 + x_1 - x_2$

Answer :

Question 2:

2. Find the gradient of the following functions, where the space $\mathbb{R}$ and $\mathbb{R}^{n\times n}$ are equipped with the standard inner product.
(a) $f(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2 + \lambda\|\boldsymbol{x}\|_2^2$, where $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, and $(\lambda > 0)$ are given.
(b) $f(\boldsymbol{X}) = \boldsymbol{b}^T\boldsymbol{X}\boldsymbol{c}$, where $\boldsymbol{X} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{b}, \boldsymbol{c} \in \mathbb{R}^n$
(c) $f(\boldsymbol{X}) = \boldsymbol{b}\boldsymbol{X}^T\boldsymbol{X}c$, where $\boldsymbol{X} \in \boldsymbol{R}^{n\times n}$ and $\boldsymbol{b}, \boldsymbol{c} \in \mathbb{R}^n$

Answer :

Question 3:

3. Let $\{\boldsymbol{x}_i, y_i\}_{i=1}^N$ be given with $\boldsymbol{x}_i \in \mathbb{R}$ and $y_i \in \mathbb{R}$. Assume $N < n$. Consider the ridge regression

$$\text{minimize}_{\boldsymbol{a}\in\mathbb{R}^N} \sum_{i=1}^{N} (\langle \boldsymbol{\alpha}, \boldsymbol{x}_i\rangle - y_i)^2 + \lambda\|\boldsymbol{a}\|_2^2,$$

where $\lambda \in \mathbb{R}$ is a regularization parameter, and we set the bias $b = 0$ for simplicity.

(a) Prove that the solution must be in the form of $\boldsymbol{a} = \sum_{i=1}^N c_i\boldsymbol{x}_i$ for some $\boldsymbol{c} = [c_1, c_2, \cdots, c_N]^T \in \mathbb{R}^N$.
(*hint: similar to the proof of the representer theorem.*)
(b) Re-express the minimization in terms of $c \in \mathbb{R}^N$, which has fewer unknowns than the original formulation as $N < n$.

Answer :

Question 4:

4. Let $f(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} + 2\boldsymbol{b}^T\boldsymbol{x} + c$, where $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is a symmetric positive semidefinite matrix, $b \in \mathbb{R}^n$, and $c \in \mathbb{R}$.
(a) Prove that $\boldsymbol{x}$ is a global minimizer of f if and only if $\boldsymbol{A}\boldsymbol{x} = -\boldsymbol{b}$.
(b) Prove that f is bounded below over $\mathbb{R}^n$ if and only if $\boldsymbol{b} \in \{\boldsymbol{A}\boldsymbol{y} : \boldsymbol{y} \in \mathbb{R}^n\}$.

Answer :