Exercise: Linear Models

Due: Optional

Problem 1 (Gradient for quadratic functions)

In this problem we develop gradients for quadratic functions.

(1) We first consider the *two-dimensional case*. Let

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \qquad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2.$$

and

$$f(\mathbf{x}) = \mathbf{x}^\top A \mathbf{x} = a_{1,1}x_1^2 + (a_{1,2} + a_{2,1})x_1 x_2 + a_{2,2}x_2^2.$$

Prove that

$$\nabla f(\mathbf{x}) = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} a_{1,1} & a_{2,1} \\ a_{1,2} & a_{2,2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

(2) We now turn to more *general cases*. Let (this is a challenging question:))

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,d} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,d} \\ \vdots & \vdots & \ldots & \vdots \\ a_{d,1} & a_{d,2} & \cdots & a_{d,d} \end{pmatrix} \quad \text{and} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_d. \end{pmatrix} \tag{1}$$

Define

$$f(\mathbf{x}) = \mathbf{x}^\top A \mathbf{x} = \sum_{i,j=1}^{d} a_{i,j} x_i x_j.$$

Prove that

$$\nabla\left(\mathbf{x}^\top A \mathbf{x}\right) = A\mathbf{x} + A^\top \mathbf{x}.$$