# CSCI 145 Problem 3: One Quadratic, Three Gradients

August 27, 2026

## Submission Instructions

Please upload your work to Gradescope by **11:59pm on DATE**.

- Turn in a single PDF of *handwritten* solutions. Typed (including LaTeX) solutions will not be accepted.
- Problems are grouped into three-week **cycles**. Once per cycle, you will meet with the grader and present the solution to *one* problem from that cycle, selected at random when you arrive — so bring your work and be ready to explain any of them.

## Problem 3: One Quadratic, Three Gradients — 2 points

Consider the general quadratic function

$$q(\mathbf{x}) = \mathbf{x}^\top \mathbf{B}\mathbf{x} + \mathbf{a}^\top \mathbf{x} + c,$$

where $\mathbf{x}, \mathbf{a} \in \mathbb{R}^d$, $\mathbf{B} \in \mathbb{R}^{d\times d}$, and $c \in \mathbb{R}$. Use the directional definition of the gradient to derive a single formula for $\nabla_{\mathbf{x}} q$. Your argument must work even when $\mathbf{B}$ is not symmetric and should explain why only the symmetric part of $\mathbf{B}$ can affect the scalar $\mathbf{x}^\top \mathbf{B}\mathbf{x}$.

Use your formula to analyze the least-squares loss $L(\mathbf{x}) = \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2$ for $\mathbf{A} \in \mathbb{R}^{n\times d}$ and $\mathbf{b} \in \mathbb{R}^n$. Derive its gradient, characterize its stationary points geometrically in terms of the residual $\mathbf{A}\mathbf{x} - \mathbf{b}$ and the columns of $\mathbf{A}$, and decide whether a stationary point can be a strict local maximum. Justify the last decision by examining the loss along an arbitrary line $\mathbf{x} + t\mathbf{v}$, rather than by quoting a theorem about convex functions.

Your general formula should also recover the two special cases $\nabla_{\mathbf{x}}(\mathbf{a}^\top \mathbf{x}) = \mathbf{a}$ and $\nabla_{\mathbf{x}}(\mathbf{x}^\top \mathbf{B}\mathbf{x}) = (\mathbf{B} + \mathbf{B}^\top)\mathbf{x}$.

**Results you may use.** For every direction $\mathbf{v}$, the gradient is defined by $q(\mathbf{x} + \epsilon\mathbf{v}) = q(\mathbf{x}) + \epsilon\langle \nabla_{\mathbf{x}} q, \mathbf{v}\rangle + o(\epsilon)$. A scalar equals its transpose, and $\|\mathbf{z}\|_2^2 = \mathbf{z}^\top \mathbf{z}$.