The University of Alabama

College of Engineering · Computer Science

Assignment 4

Submission Instructions

- Submit your solutions as a **separate PDF file**.
- Clearly label each problem number and part (e.g., Problem 1a, Problem 2c).
- For Python problems, include screenshots of your code and output.
- Ensure your work is legible and well-organized. Show all steps for full credit.

Problem 1: Derivatives and Critical Points (15 Points)

Consider the function $f(x) = 2x^3 - 9x^2 + 12x - 4$.

(a) Compute $f'(x)$ using the power rule.

(b) Find all critical points of f (i.e., values of x where $f'(x) = 0$).

(c) For each critical point, use the second derivative test to determine whether it is a local minimum, local maximum, or neither.

(d) Evaluate $f'(0)$. Is the function increasing or decreasing immediately to the right of $x = 0$? Justify your answer.

Solution:

■

Problem 2: Gradient Ascent by Hand (10 Points)

Consider maximizing $f(x) = -(x-3)^2 + 9$.

(a) Starting from $x_0 = 0$ with learning rate $\alpha = 0.5$, perform **two** manual gradient ascent steps. That is, compute x_1 and x_2 using the update rule

$$x_{k+1} = x_k + \alpha \cdot f'(x_k).$$

Show all intermediate computations.

(b) What is the analytical maximum of $f(x)$, and where does it occur?

(c) Explain why gradient ascent is guaranteed to find the **global** maximum for this particular function (Hint: think about the shape of the function).

Solution:

■

Problem 3: Partial Derivatives and the Gradient (20 Points)

Let $f(x_1, x_2) = 3x_1^2 x_2 - x_2^3 + 2x_1$.

(a) Compute the partial derivatives $\dfrac{\partial f}{\partial x_1}$ and $\dfrac{\partial f}{\partial x_2}$.

(b) Evaluate the gradient ∇f at the point $(1, 2)$.

(c) In which direction does f increase most steeply at the point $(1, 2)$? Express your answer as a unit vector.

(d) The gradient of a scalar-valued function $f : \mathbb{R}^n \to \mathbb{R}$ is a vector. What is the dimension of ∇f in general? Explain what each component represents.

Solution:

■

Problem 4: Chain Rule (15 Points)

Consider the composite function $h(x) = \ln(x^2 + 1)$, which can be written as $h(x) = g(f(x))$ where $f(x) = x^2 + 1$ and $g(u) = \ln(u)$.

(a) Compute $h'(x)$ directly by treating it as a single function (using standard derivative rules).

(b) Now apply the **chain rule**: compute $f'(x)$ and $g'(u)$ separately, then form $h'(x) = g'(f(x)) \cdot f'(x)$. Verify your answer matches part (a).

(c) Approximate $h'(2)$ using the **central difference** formula with $\Delta = 0.001$:

$$h'(x) \approx \frac{h(x+\Delta) - h(x-\Delta)}{2\Delta}.$$

Compare this approximation to the exact value from part (a).

Solution:

■

Problem 5: Jacobian Matrix (10 Points)

Consider the vector-valued function

$$\mathbf{f}(x,y) = \begin{pmatrix} x^2 y \\ x + e^y \end{pmatrix}.$$

(a) Compute the Jacobian matrix

$$J = \begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{pmatrix}.$$

(b) Evaluate J at the point $(x,y) = (1,0)$.

(c) What is the dimension of the Jacobian for a function $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$? In the special case $m = 1$, what does the Jacobian reduce to?

Solution:

■

Problem 6: Loss Functions and the Gradient Descent Update (15 Points)

Suppose we have a single training example $(x, y) = (2, 5)$ and a linear model $y' = w_1 x + w_0$, initialized with $w_1 = 1$ and $w_0 = 0$.

(a) Compute the model's prediction y' and the squared loss $\ell = (y - y')^2$ at the current parameter values.

(b) Using the formulas derived in lecture,

$$\frac{d\ell}{dw_1} = 2(y' - y)\,x, \qquad \frac{d\ell}{dw_0} = 2(y' - y),$$

compute the gradients at the current parameters.

(c) Perform one gradient descent update with learning rate $\eta = 0.1$:

$$w_1 \leftarrow w_1 - \eta\frac{d\ell}{dw_1}, \qquad w_0 \leftarrow w_0 - \eta\frac{d\ell}{dw_0}.$$

Report the updated values w_1 and w_0.

(d) Compute the new squared loss after the update. Did it decrease compared to part (a)?

Solution:

■