Department of Mathematics & Computer Science — BSc

Optimization Techniques for Machine Learning

Prerequisite Diagnostic Test — **MODEL ANSWERS**

Prepared by: Dr. D Bhanu Prakash **Total Marks: 25** **Time: 30 minutes**

SECTION A Calculus & Algebra Essentials

Q1. Find the derivative of $f(x) = 3x^3 - 5x^2 + 2x - 7$. At which value of x does $f'(x) = 0$? (Solve and identify the critical points.) **[3 marks]**

Solution: Differentiating term by term:

$$f'(x) = 9x^2 - 10x + 2$$

Setting $f'(x) = 0$:

$$9x^2 - 10x + 2 = 0 \quad \Rightarrow \quad x = \frac{10 \pm \sqrt{100 - 72}}{18} = \frac{10 \pm \sqrt{28}}{18} = \frac{10 \pm 2\sqrt{7}}{18} = \frac{5 \pm \sqrt{7}}{9}$$

Critical points: $x_1 = \frac{5 - \sqrt{7}}{9} \approx 0.261$ and $x_2 = \frac{5 + \sqrt{7}}{9} \approx 0.850$

Q2. Compute the **gradient** ∇f of the function $f(x, y) = x^2 + 3xy + y^3$. **[2 marks]**

Solution:

$$\frac{\partial f}{\partial x} = 2x + 3y, \qquad \frac{\partial f}{\partial y} = 3x + 3y^2$$

$$\nabla f = \begin{pmatrix} 2x + 3y \\ 3x + 3y^2 \end{pmatrix}$$

Q3. Solve the system of linear equations. Express your answer as a column vector $\mathbf{x}$: **[3 marks]**

$$\begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 8 \\ 19 \end{pmatrix}$$

Solution: The matrix $A = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}$ has determinant $\det(A) = 2(3) - 1(5) = 1$.

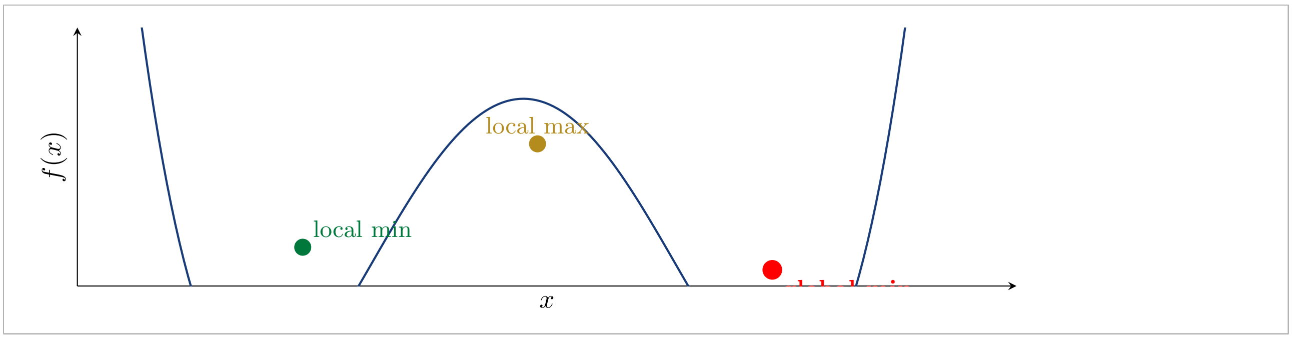
Since $\det(A) \neq 0$, $A^{-1} = \frac{1}{1}\begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix} = \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix}$.

$$\mathbf{x} = A^{-1}\mathbf{b} = \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix} \begin{pmatrix} 8 \\ 19 \end{pmatrix} = \begin{pmatrix} 24 - 19 \\ -40 + 38 \end{pmatrix} = \boxed{\begin{pmatrix} 5 \\ -2 \end{pmatrix}}$$

Verification: $2(5) + 1(-2) = 8$ ✓ $5(5) + 3(-2) = 19$ ✓

SECTION B Conceptual Understanding

Q4. The graph below (sketch your own) shows a function with **two local minima** and **one local maximum**. Mark these on the sketch below, and indicate where the **global minimum** is. Why does finding the global minimum matter in machine learning? **[3 marks]**



Solution: In machine learning, the loss function (e.g., mean squared error) represents model error. We adjust model parameters to *minimise* this loss. The **global minimum** corresponds to the set of parameters giving the lowest possible error — the best model. Local minima may trap optimisation algorithms (like gradient descent) in suboptimal solutions with higher error, which is why distinguishing and reaching the global minimum is a key challenge in ML.

Q5. What does it mean geometrically for a function to be **convex**? Give one example of a convex function and one that is *not* convex. **[2 marks]**

Solution: A function f is **convex** if, geometrically, the line segment joining any two points on its graph lies *on or above* the graph. Equivalently, the function "curves upward" — it looks like a bowl. Formally: $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$ for all $\lambda \in [0, 1]$.

Convex example: $f(x) = x^2$ (a parabola opening upward). Convexity guarantees any local minimum is also the global minimum — very useful in ML optimisation.

Non-convex example: $f(x) = \sin(x)$ (oscillates, has multiple peaks and troughs; a chord between two points on the curve dips below the curve).

SECTION C Application & Reasoning

Q6. Consider $f(x) = (x-3)^2 + 4$. **(a)** What is the minimum value of $f(x)$, and where does it occur? **(b)** What is the maximum value of $f(x)$, and where does it occur? **[5 marks]**

Solution: **(a) Minimum value:** The term $(x-3)^2 \geq 0$ for all real x, with equality when $x = 3$.

Therefore the **minimum value** of $f(x)$ is **4**, occurring at $x = \mathbf{3}$.

(b) Maximum value: Since $(x-3)^2$ grows without bound as $x \to \pm\infty$, the function $f(x)$ has **no finite maximum** — it increases without limit ($f(x) \to +\infty$).

If the domain were restricted (e.g., $x \in [-2, 3]$), the maximum would occur at the endpoint farthest from $x = 3$, i.e., at $x = -2$: $f(-2) = 25 + 4 = 29$.

Q7. Consider the function $f(x) = (x-2)^2$. (a) Without calculus, state the minimum value of $f(x)$ and the value of x where it occurs. (b) Find $f'(x)$ (the derivative). (c) At $x = 5$, is the function increasing or decreasing? How does $f'(x)$ tell you this? **[3 marks]**

Solution: **(a)** Since $(x-2)^2 \geq 0$ always, and equals 0 when $x = 2$: the **minimum value is** 0, occurring at $x = \mathbf{2}$.

(b) $f(x) = (x-2)^2 = x^2 - 4x + 4$, so:

$$f'(x) = 2(x-2) = 2x - 4$$

(c) At $x = 5$: $f'(5) = 2(5-2) = 6 > 0$.

Since $f'(5) > 0$, the function is **increasing** at $x = 5$. A positive derivative means the function's output rises as x increases — we are to the right of the minimum, heading uphill on the parabola.

Q8. **Bonus:** Fill in the blanks using terms from this list: minimum, maximum, zero, positive, negative.

(a) At a local ________ , the derivative of a function equals ________.

(b) If $f'(x) > 0$ at some point, the function is ________ at that point.

(c) To move downhill on a curve, we should move in the ________ direction of the gradient. **[4 (bonus) marks]**

Solution: (a) At a local **minimum** (or maximum), the derivative of a function equals **zero**.

(b) If $f'(x) > 0$ at some point, the function is **increasing** at that point.

(c) To move downhill on a curve, we should move in the **negative** direction of the gradient.

Note on (c): This is the principle behind **gradient descent** — the cornerstone of ML optimisation. We update parameters as $\theta \leftarrow \theta - \eta \nabla f(\theta)$, moving opposite to the gradient.