Optimization and Newton's Method Questions & Answers

1. What is the primary advantage of Newton's method over gradient descent?

Answer: b. Quadratic convergence rate

Explanation:

- Newton's method converges quadratically near the optimal solution, making it much faster than gradient descent.
- Gradient descent has a lower per-iteration cost, making it more suitable for large-scale problems.

2. True/False: Newton's method requires the Hessian matrix to be positive definite for minimization problems.

Answer: True

Explanation:

- If the Hessian is indefinite or singular, Newton's method may diverge or converge to a saddle point.

3. Why might Newton's method fail to converge for non-convex functions?

Answer:

- If the function is non-convex, the Hessian matrix may be indefinite.
- This can lead to steps moving away from the optimal point.
- If the Hessian is singular, Newton's method cannot proceed.

4. The steepest descent direction for minimizing f(x) is:

Answer: b. -gradf(x)

Explanation:

- The steepest descent direction is given by the negative gradient -grad f(x), which points in the direction of the fastest decrease.

5. True/False: The steepest descent method converges linearly for quadratic functions.

Answer: True

Explanation:

- For quadratic functions, the steepest descent method exhibits linear convergence.

6. Calculate the steepest descent direction for $f(x,y) = x^2 + 3y^2$ at (1, -1).

Solution:

- Gradient: grad f(x, y) = (2x, 6y)
- Evaluating at (1, -1): (2, -6)
- Steepest descent direction: (-2, 6)

7. The directional derivative of f(x,y) = xy at (2,3) in direction u = (1/sqrt2, 1/sqrt2) is:

Answer: a. 5/sqrt2

Solution:

- Gradient: grad f(x, y) = (y, x)
- Evaluating at (2,3): (3,2)
- Dot product with u: 5/sqrt2

8. True/False: The maximum directional derivative of a function equals the magnitude of its gradient.

Answer: True

Explanation:

- The maximum directional derivative occurs when the direction u aligns with the gradient grad f.

9. BFGS and DFP are examples of:

Answer: b. Hessian approximation methods

Explanation:

- BFGS and DFP are quasi-Newton methods.

10. Why are quasi-Newton methods preferred over Newton's method for large-scale problems?

Answer:

- No need to compute the full Hessian matrix, which is costly.
- Lower memory requirements.
- Superlinear convergence.
- More stable when the Hessian is singular or indefinite.

11. Perform one iteration of Newton's method for $f(x) = x^4 - 3x^2 + 2$ starting at $x_0 = 1$.

Solution:

- Compute derivatives:

$f(x) = x^4 - 3x^2 + 2$

$f'(x) = 4x^3 - 6x$

- Evaluate at $x_0 = 1$:

  $f(1) = 0$

- Newton's method update:

  $x_1 = 1 - (0/-2) = 1$

Observation: Newton's method converged in one step.