Laboratory Assignment 4
Identifying and Characterising Critical Points

CSC372-M72: Optimisation

2022-23

1 Objectives.

- To identify critical points.
- To characterise critical points, applying the necessary and sufficient conditions for optimal solutions of unconstrained objective functions.

2 Background

To identify a critical point and classify it, we need to follow the steps below.

1. Compute $\nabla f(\mathbf{x})$. You may find the Appendix B useful for this step.
2. Solving $\nabla f(\mathbf{x}) = 0$ and reporting the critical point.
3. Compute $\nabla^2 f(\mathbf{x})$.
4. Solving $\det[\nabla^2 f(x) - \lambda I] = 0$. You may find the Appendix C and D useful for this step.
5. Final judgement on the nature of the critical point.

3 Tasks

You are given a function $f(\mathbf{x}) = 5x_1^2 + 3x_2 - 5x_1x_2 + 5$ where $\mathbf{x} = (x_1, x_2)^\top \in \mathbb{R}^2$ is a two-dimensional real vector. Identify the critical points, and characterise them.

Answer: They need to find the gradient vector $\nabla f(\mathbf{x})$. Then solve the equations $\nabla f(\mathbf{x}) = \mathbf{0}$: this will give them the critical point(s).

$$\nabla f(\mathbf{x}) = \begin{bmatrix} 10x_1 - 5x_2 \\ 3 - 5x_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$\Rightarrow x_1 = \frac{3}{5} = 0.6, \text{ and } x_2 = \frac{6}{5} = 1.2.$$

So, there is one critical point: $\mathbf{x}^* = (0.6, 1.2)^\top$.

They need to find the Hessian matrix $\nabla^2 f(\mathbf{x})$. Then if they solve the equations $\det[\nabla^2 f(x) - \lambda I] = 0$ to find λ. If the λs are all positive at a point of tangency, then the point is a minimum. If all are negative, then the point is a maximum. If there is a mix between positive and negative λs, then it is a saddle point.

The Hessian matrix is: $\nabla^2 f(x) = \begin{bmatrix} 10 & -5 \\ -5 & 0 \end{bmatrix}$. Solving $\det[\nabla^2 f(x) - \lambda I] = 0$ should yield $\lambda_1 \approx 12.07$ and $\lambda_2 \approx -2.07$. Since, both of these eigenvalues have different signs, we can deduce that the critical point is a **saddle point**.

A Necessary and Sufficient Conditions

For an unconstrained problem in n-dimension, $\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x})$:

First Order Necessary Condition (FONC). If $\mathbf{x}^*$ is a strong local minimiser (or maximiser) of $f(\mathbf{x})$, and $f(\mathbf{x})$ is continuously differentiable in an open neighbourhood of $\mathbf{x}^*$, then $\nabla f(\mathbf{x}) = \left(\frac{\delta f(\mathbf{x})}{\delta x_1}, \ldots, \frac{\delta f(\mathbf{x})}{\delta x_n}\right)^\top = (0, \ldots, 0)^\top = \mathbf{0}$.

Second Order Necessary Condition (SONC). For checking curvature, we need to consider the characteristics of Hessian matrix:

Convex. $\nabla^2 f(\mathbf{x})$ is positive definite, i.e. local minimum.

Concave. $\nabla^2 f(\mathbf{x})$ is negative definite, i.e. local maximum.

Saddle. $\nabla^2 f(\mathbf{x})$ has both positive and negative **eigenvalues**.

B Rules of Derivates

Here are some of the basic rules of derivatives.

Notation. $f'(x) = \frac{df(x)}{dx}$.

Constant rule. If $f(x) = c$ where c is a constant, then $f'(x) = 0$.

Constant multiplier rule. If $f(x) = cg(x)$, then $f'(x) = cg'(x)$.

Power rule. If $f(x) = x^c$, then $f'(x) = cx^{c-1}$.

Sum and difference rule. If $h(x) = f(x) \pm g(x)$, then $h'(x) = f'(x) \pm g'(x)$.

Multi-variate functions. If the independent variable is a n-dimensional vector $\mathbf{x} = (x_1, \ldots, x_n)^\top$, then the gradients form a vector:

$$\nabla f(\mathbf{x}) = \left(\frac{df(\mathbf{x})}{dx_1}, \ldots, \frac{df(x)}{dx_n}\right)^\top,$$

and the second derivatives construct the Hessian matrix:

$$\nabla^2 f(\mathbf{x}) = \begin{bmatrix} \frac{d^2 f(x)}{dx_1^2} & \cdots & \frac{d^2 f(x)}{dx_1 dx_n} \\ \vdots & \ddots & \vdots \\ \frac{d^2 f(x)}{dx_1 dx_n} & \cdots & \frac{d^2 f(x)}{dx_n^2} \end{bmatrix}.$$

C Eigenvalues.

To find the eigenvalues λs of a matrix A, we solve the following system of equations:

$$\det(A - \lambda I) = 0.$$

D Solving an Equation.

Given an equation $a\lambda^2 + b\lambda + c = 0$, we can find λ as follows:

$$\lambda = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$