
Question1 (6 points)

The negative of a convex function is known to be a concave function . The negative of a quasiconvex function is known to be a quasiconcave function . For each of the following functions determine whether it is convex, concave, quasiconvex, or quasiconcave.

(a) (1 point) The function $f \colon \mathbb{R} \to \mathbb{R}$ given by

$$f(x) = e^x - 1.$$

(b) (1 point) The function $f \colon \mathbb{R}^2_{++} \to \mathbb{R}$ given by

$$f(\mathbf{x}) = x_1 x_2.$$

(c) (1 point) The function $f \colon \mathbb{R}^2_{++} \to \mathbb{R}$ given by

$$f(\mathbf{x}) = \frac{1}{x_1 x_2}.$$

(d) (1 point) The function $f \colon \mathbb{R}^2_{++} \to \mathbb{R}$ given by

$$f(\mathbf{x}) = \frac{x_1}{x_2}.$$

(e) (1 point) The function $f \colon \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$ given by

$$f(\mathbf{x}) = \frac{x_1^2}{x_2}.$$

(f) (1 point) The function $f \colon \mathbb{R}^2_{++} \to \mathbb{R}$ given by

$$f(\mathbf{x}) = x_1^{\alpha} x_2^{1-\alpha} \qquad \text{where} \quad 0 \leq \alpha \leq 1.$$

Question2 (7 points)

Let $f \colon \mathbb{R}^n \to \mathbb{R}$. Prove the following theorems.

(a) (1 point) If f is continuously differentiable in an open neighbourhood of a local minimiser $\mathbf{x}^*$, then the gradient $\nabla f(\mathbf{x}^*) = \mathbf{0}$

(b) (2 points) If the Hessian $\mathbf{H}$ of f is continuous in an open neighbourhood of a local minimiser $\mathbf{x}^*$, then the gradient $\nabla f(\mathbf{x}^*) = \mathbf{0}$ and the Hessian $\mathbf{H}$ at $\mathbf{x} = \mathbf{x}^*$ is positive semi-definite .

(c) (2 points) If f is differentiable and convex, then $\mathbf{x}^*$ is a global minimum if and only if

$$\nabla f(\mathbf{x}^*) = \mathbf{0}$$

(d) (2 points) If f is twice differentiable, then f is convex if and only if the Hessian of f is *positive semi-definite* for all $\mathbf{x} \in \mathbb{R}$.

Question3 (2 points)

Let $f_0, f_1, \ldots, f_n \colon \mathbb{R} \to \mathbb{R}$ be continuous functions. Consider the problem of approximating f_0 as a linear combination of $f_1, \ldots f_n$. For $\boldsymbol{\alpha} \in \mathbf{R}^n$, we say that

$$f = \alpha_1 f_1 + \cdots \alpha_n f_n$$

approximates f_0 with tolerance $\epsilon > 0$ over the interval $[0, T]$ if

$$|f(t) - f_0(t)| \leq \epsilon \qquad \text{for} \qquad 0 \leq t \leq T$$

For a fixed tolerance $\epsilon > 0$, we define the approximation width as the largest T such that f approximates f_0 over the interval $[0, T]$, that is

$$W(\boldsymbol{\alpha}) = \sup \{T \mid |f(t) - f_0(t)| \leq \epsilon \quad \text{for} \quad 0 \leq t \leq T\}$$

Show that W is quasiconcave.

Question4 (2 points)
Choose a suitable numerical method to find the global minimum of the function

$$f(x) = \frac{\sin\left(\dfrac{1}{x}\right)}{(x - 2\pi)^2 + \pi}$$

correct to 3 decimal places. Explain your choice.

Question5 (8 points)
Suppose f is unimodal and differentiable. Consider the following algorithm which explicitly uses f' when we know the minimum point $x \in [a_0, b_0]$. The essential idea is to approximate the function f on the interval $[a_{k-1}, b_{k-1}]$ with a cubic polynomial of the form

$$P(x) = \alpha_{k-1}(x - a_{k-1})^3 + \beta_{k-1}(x - a_{k-1})^2 + \gamma_{k-1}(x - a_{k-1}) + \rho_{k-1}$$

which has the same value and derivative as f at the endpoints a_{k-1} and b_{k-1}. Then using the minimum point c_{k-1} of the cubic polynomial to tell how to squeeze the interval

$$[a_{k-1}, b_{k-1}] \qquad \text{to} \qquad [a_k, b_k]$$

(a) (1 point) Discuss why if $f'(c_{k-1}) > 0$, then we shall set $a_k = a_{k-1}$ and $b_k = c_{k-1}$, else we shall set $a_k = c_{k-1}$ and $b_k = b_{k-1}$.

(b) (1 point) Show that

$$c_{k-1} = a_{k-1} + \frac{-\beta_{k-1} + \sqrt{\beta_{k-1}^2 - 3\alpha_{k-1}\gamma_{k-1}}}{3\alpha_{k-1}}$$

(c) (1 point) Find formulas for α_{k-1}, β_{k-1}, γ_{k-1} and ρ_{k-1} in terms of a_{k-1} and b_{k-1}.

(d) (1 point) Use this method to find the minimum of

$$f(x) = e^x + 2x + \frac{x^2}{2}$$

on the interval $[-2.4, -1.6]$ correct to 5 decimal places.

(e) (2 points) Compare this method with Newton's method in terms of assumptions about f and convergence rate.

(f) (2 points) Compare this method with Golden Section Search in terms assumptions about f and convergence rate.

Question6 (2 points)

Consider the following function

$$f(x, y) = e^x \left(4x^2 + 2y^2 + 4xy + 2y + 1\right)$$

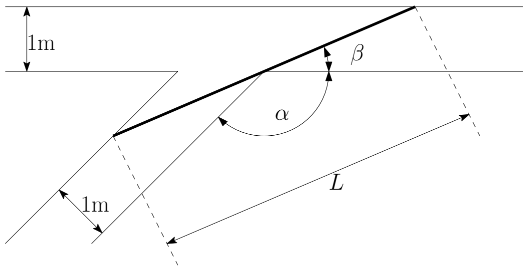
Produce an informative Matlab plot of $g(y) = \min_x f(x, y)$, which gives the minimum of

$$f(x, y)$$

as y changes.

Question7 (3 points)

The length, L, of the longest ladder that can pass around the corner of two corridors depends on the angle α shown in the figure below.



Produce a Matlab plot of L versus α ranging from 45° to 135° by first solving a minimisation problem using numerical methods that we have discussed so far.

Question8 (0 points)

A popular global optimization algorithm for difficult functions, especially if there are many local minima, is called the method of simulated annealing . It involves no derivatives or an initial guess that needs to be sufficiently close to the minimum. Suppose $f\colon \mathbb{R}^n \to \mathbb{R}$ has a global minimum at $\mathbf{x}^*$ and the k-iteration $\mathbf{x}_k$ has been computed. It iterates by the following scheme:

1. Generating a number of random points $\mathbf{u}_1$, $\mathbf{u}_2$, …, $\mathbf{u}_m$ in a large neighborhood of $\mathbf{x}_k$.
2. Computing $f(\mathbf{u}_1)$, $f(\mathbf{u}_2)$, …, $f(\mathbf{u}_m)$.
3. Finding the index j such that $f(\mathbf{u}_j) = \min\{f(\mathbf{u}_1), f(\mathbf{u}_2), \ldots, f(\mathbf{u}_m)\}$.
4. Assigning $\mathbf{x}_{k+1} = \mathbf{u}_j$ if $f(\mathbf{x}_k) > f(\mathbf{u}_j)$. Otherwise assigning a probability

$$p_i = \frac{\exp\Big(\alpha\big(f(\mathbf{x}_k) - f(\mathbf{u}_i)\big)\Big)}{\sum_{\ell=1}^{m} \exp\Big(\alpha\big(f(\mathbf{x}_k) - f(\mathbf{u}_\ell)\big)\Big)} \qquad \text{where } \alpha \text{ is a positive real,}$$

to $\mathbf{u}_i$ for each $i = 1, \ldots, m$. Then making a random choice among $\mathbf{u}_1$, $\mathbf{u}_2$, …, $\mathbf{u}_m$ according to the probabilities p_i and assigning this randomly chosen $\mathbf{u}_\ell$ to be $\mathbf{x}_{k+1}$.

With some minor modifications, this can be used for function $Q\colon \mathcal{X} \to \mathbb{R}$, where $\mathcal{X}$ is any set. For example, in the *traveling salesman problem*, $\mathcal{X}$ is the set of all permutations of a set of integers. Consider the *Euclidean traveling salesman problem* (ETSP): Given a set of points in $\mathbb{R}^2$ representing positions of cities on a map, we wish to visit each city exactly once while minimizing the total distance traveled.

(a) (1 point (bonus)) Implement simulated annealing to solve ETSP with different α.

(b) (1 point (bonus)) Propose another optimization method to solve ETSP. Analyze how its efficiency would compare to that of simulated annealing.