

Problem Sheet 1

Problems in Part A will be discussed in class. Problems in Part B come with solutions and should be tried at home, they will be discussed in class if time permits.

Part A

(1.1) Find examples of

- (a) A function $f \in C^2(\mathbb{R})$ with a strict minimizer x such that $f''(x) = 0$ (that is, the second derivative is not positive definite).
- (b) A function $f: \mathbb{R} \rightarrow \mathbb{R}$ with a strict minimizer x^* that is not an isolated local minimizer. **Hint:** Consider a rapidly oscillating function that has minima that are arbitrary close together, but not equal.

(1.2) For this problem you might want to recall some linear algebra.

- (a) Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be a symmetric matrix, $\mathbf{b} \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Show that the quadratic function

$$f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^\top \mathbf{A} \mathbf{x} + \mathbf{b}^\top \mathbf{x} + c \quad (1)$$

with symmetric \mathbf{A} is convex if and only if \mathbf{A} is positive semidefinite.

- (b) Now let $\mathbf{A} \in \mathbb{R}^{m \times n}$ be an arbitrary matrix. Show that the function

$$f(\mathbf{x}) = \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2$$

is convex (the 2-norm is defined as $\|\mathbf{x}\|_2 = \sqrt{\mathbf{x}^\top \mathbf{x}}$). Moreover, if $m \geq n$ and the matrix \mathbf{A} has rank m , then it is strictly convex and the unique minimizer is

$$\mathbf{x}^* = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{b}.$$

(1.3) A set $S \subseteq \mathbb{R}^n$ is called *convex*, if for any $\mathbf{x}, \mathbf{y} \in S$ and $\lambda \in [0, 1]$,

$$\lambda \mathbf{x} + (1 - \lambda) \mathbf{y} \in S.$$

In words, for any two points in S , the line segment joining them is also in S . Which of the following sets are convex?

- (a) $S = \{\mathbf{x} \in \mathbb{R}^3 : \|\mathbf{x}\|_2 = 1\}$ (the unit sphere);
- (b) $S = \{\mathbf{x} \in \mathbb{R}^2 : 1 \leq x_1 - x_2 < 2\}$;
- (c) $S = \{\mathbf{x} \in \mathbb{R}^n : |x_1| + \cdots + |x_d| \leq 1\}$;
- (d) $S = \mathcal{S}_+^n \subset \mathbb{R}^{n \times n}$, the set of symmetric, positive semidefinite matrices.

(1.4) For this problem we generalize the notion of convexity to function not necessarily defined on all of \mathbb{R}^n . Denote by $\text{dom} f$ the *domain* of f , i.e., the set of \mathbf{x} on which $f(\mathbf{x})$ attains a finite value. A function f is called *convex*, if $\text{dom} f$ is a convex set and for all $\mathbf{x}, \mathbf{y} \in \text{dom} f$ and $\lambda \in [0, 1]$,

$$f(\lambda \mathbf{x} + (1 - \lambda)\mathbf{y}) \leq \lambda f(\mathbf{x}) + (1 - \lambda)f(\mathbf{y}).$$

Which of the following functions are convex?

- (a) $f(x) = \log(x)$ on \mathbb{R}_{++} (the positive real numbers);
- (b) $f(x) = x^4$ on \mathbb{R} ;
- (c) $f(\mathbf{x}) = x_1 x_2$ on \mathbb{R}_{++}^2 ;
- (d) $f(\mathbf{x}) = x_1/x_2$ on \mathbb{R}_{++}^2 ;
- (e) $f(x) = e^x - 1$ on \mathbb{R} ;
- (f) $f(\mathbf{x}) = \max_i x_i$ on \mathbb{R}^n .

Part B

(1.5) In engineering applications¹ one sometimes encounters a problem of the form

$$\text{minimize} \quad \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_\infty, \quad (2)$$

with $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$ and $\|\mathbf{x}\|_\infty = \max_{1 \leq i \leq n} |x_i|$ is the ∞ -norm.

- (a) Draw the “unit circle” $\{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\|_\infty \leq 1\}$.
- (b) Formulate a linear programming problem \mathcal{P} with decision variables (\mathbf{x}, t) , such that if (\mathbf{x}^*, t^*) is the unique minimizer of \mathcal{P} , then \mathbf{x}^* is the unique minimizer of (2).

Even though (2) is not a linear programming problem (the objective is not linear), it is *equivalent* to one, in the sense that a minimizer can be read off the solution of a linear programming problem.

(1.6) Using Python or another computing system, compute and plot the sequence of points \mathbf{x}_k , starting with $\mathbf{x}_0 = (0, 0)^\top$, for the gradient descent algorithm for the problem

$$\text{minimize} \quad \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2$$

with data

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \\ -1 & 0 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 10 \\ -1 \\ 0 \end{pmatrix}.$$

¹For example in the synthesis of linear time-invariant dynamical systems.

(1.7) Consider the **Rosenbrock function** in \mathbb{R}^2 ,

$$f(\mathbf{x}) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2.$$

Compute the gradient ∇f and the Hessian $\nabla^2 f$. Show that $\mathbf{x}^* = (1, 1)^\top$ is the only local minimizer of this function, and that the Hessian at this point is positive definite.

Using Python or another computing system, draw a contour plot of the Rosenbrock function.