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 $f\colon \mathbb{R} \to \mathbb{R}$ 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 $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ by a symmetric matrix, $\boldsymbol{b} \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Show that the quadratic function

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

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

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

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

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

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

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

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

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) $S = \{\boldsymbol{x} \in \mathbb{R}^3 : \|\boldsymbol{x}\|_2 = 1\}$ (the unit sphere);

(b) $S = \{\boldsymbol{x} \in \mathbb{R}^2 : 1 \leq x_1 - x_2 < 2\}$;

(c) $S = \{\boldsymbol{x} \in \mathbb{R}^n : |x_1| + \cdots + |x_n| \leq 1\}$;

(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 $\mathrm{dom} f$ the *domain* of f, i.e., the set of $\boldsymbol{x}$ on which $f(\boldsymbol{x})$ attains a finite value. A function f is called *convex*, if $\mathrm{dom} f$ is a convex set and for all $\boldsymbol{x}, \boldsymbol{y} \in \mathrm{dom} f$ and $\lambda \in [0, 1]$,

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

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) $f(\boldsymbol{x}) = x_1 x_2$ on $\mathbb{R}^2_{++}$;

(d) $f(\boldsymbol{x}) = x_1/x_2$ on $\mathbb{R}^2_{++}$;

(e) $f(x) = e^x - 1$ on $\mathbb{R}$;

(f) $f(\boldsymbol{x}) = \max_i x_i$ on $\mathbb{R}^n$.

Part B

(1.5) In engineering applications[1] one sometimes encounters a problem of the form

$$\text{minimize} \quad \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_\infty, \tag{2}$$

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

(a) Draw the "unit circle" $\{\boldsymbol{x} \in \mathbb{R}^2 : \|\boldsymbol{x}\|_\infty \leq 1\}$.

(b) Formulate a linear programming problem $\mathcal{P}$ with decision variables $(\boldsymbol{x}, t)$, such that if $(\boldsymbol{x}^*, t^*)$ is the unique minimizer of $\mathcal{P}$, then $\boldsymbol{x}^*$ 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 $\boldsymbol{x}_k$, starting with $\boldsymbol{x}_0 = (0,0)^\top$, for the gradient descent algorithm for the problem

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

with data

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

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

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

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

Compute the gradient ∇f and the Hessian $\nabla^2 f$. Show that $\boldsymbol{x}^* = (1, 1)^\top$ 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.