Exercises for Week 1

1 Function plots and contour plots in Python

1.1 (in Python)

Consider the function $f(x) = x - \log(x)$ with $x > 0$ and log as the natural logarithm. In Python the function can be implemented as

```
def f(x):
# f(x) = x - ln(x)     x>0
return x - numpy.log(x)
```

- Compute and plot $f(x)$ for $0 < x \leq 2$ with spacing 0.01.
- Compute and plot $f(x)$, $f'(x) = \frac{df}{dx}(x)$, and $f''(x) = \frac{d^2 f}{dx^2}(x)$ in the same figure using `matplotlib.pyplot.subplots`. Add grid lines in the plots. Note that you need calculate f' and f'' by hand or in other tools first.
- Based on the plots, argue that $f(x)$ is convex. You should choose reasonable ranges for the y-axis of the relevant plots.
- Locate the minimizer, x^*, of $f(x)$. What is the corresponding minimum value, $f(x^*)$? Is this minimizer unique? Why?

1.2 (by hand then in Python)

Calculate the gradient and show that the function

$$f(\boldsymbol{x}) = f(x_1, x_2) = 8x_1 + 12x_2 + x_1^2 - 2x_2^2$$

has only one stationary point, and that it is neither a maximum nor a minimum, but a saddle-point.
Then, use Python to draw a contour plot of this function.

2 Gradient and Hessian of functions

2.1 (by hand)

Consider the linear function

$$f(\boldsymbol{x}) = f(x_1, x_2, x_3) = \boldsymbol{g}^T \boldsymbol{x} = \begin{bmatrix} g_1 \\ g_2 \\ g_3 \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = g_1 x_1 + g_2 x_2 + g_3 x_3$$

Compute the gradient, $\nabla f(\boldsymbol{x})$.

In more general case, let $\boldsymbol{x} \in \mathbb{R}^n$ and $\boldsymbol{g} \in \mathbb{R}^n$. Let

$$f(\boldsymbol{x}) = \boldsymbol{g}^T \boldsymbol{x} = \begin{bmatrix} g_1 \\ \vdots \\ g_n \end{bmatrix}^T \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \sum_{i=1}^{n} g_i x_i$$

Compute the gradient $\nabla f(\boldsymbol{x})$.

2.2 (by hand)

Consider the quadratic function

$$f(\boldsymbol{x}) = f(x_1, x_2, x_3) = \boldsymbol{x}^T H \boldsymbol{x} = \frac{1}{2} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}^T \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

1. Expand the expression of $f(\boldsymbol{x})$ as $f(\boldsymbol{x}) = \frac{1}{2}(h_{11}x_1^2 + \cdots)$.
2. Derive an expression for

$$\nabla f(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(x) \\ \frac{\partial f}{\partial x_2}(x) \\ \frac{\partial f}{\partial x_3}(x) \end{bmatrix}$$

3. Derive an expression for

$$\nabla^2 f(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(x) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(x) & \frac{\partial^2 f}{\partial x_1 \partial x_3}(x) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(x) & \frac{\partial^2 f}{\partial x_2 \partial x_2}(x) & \frac{\partial^2 f}{\partial x_2 \partial x_3}(x) \\ \frac{\partial^2 f}{\partial x_3 \partial x_1}(x) & \frac{\partial^2 f}{\partial x_3 \partial x_2}(x) & \frac{\partial^2 f}{\partial x_3 \partial x_3}(x) \end{bmatrix}$$

4. Redo the questions now assuming that H is symmetric, i.e., $H = H^T$.
5. Consider the general quadratic function

$$f(\boldsymbol{x}) = \frac{1}{2} \boldsymbol{x}^T H \boldsymbol{x}$$

with $H \in \mathbb{R}^{n \times n}$ being symmetric and $\boldsymbol{x} \in \mathbb{R}^n$. What is the gradient, $\nabla f(\boldsymbol{x})$, of this function? What is the Hessian, $\nabla^2 f(\boldsymbol{x})$, of this function?

3 Minimizers of univariate and multivariate problems

3.1 (by hand)

Consider the unconstrained optimization problem

$$\min_{\boldsymbol{x} \in \mathbb{R}^2} f(\boldsymbol{x}) = \frac{3}{2}(x_1^2 + x_2^2) + (1 + a)x_1 x_2 - (x_1 + x_2) + b$$

where a and b are real-valued parameters. According to optimality condition, find all values of a and b such that the problem has a unique optimal solution.

3.2 (by hand)

Show that for any unconstrained quadratic problem

$$\min_{\boldsymbol{x}} f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T Q \boldsymbol{x} - \boldsymbol{b}^T \boldsymbol{x}$$

with $Q \succ 0$, $\boldsymbol{x}^*$ is a global minimizer if and only if $\boldsymbol{x}^*$ satisfies the first-order necessary condition. That is, the problem is equivalent to solving $Q\boldsymbol{x} = \boldsymbol{b}$.

3.3 (by hand. Solid proof is not necessary.)

Consider

$$\min_{\boldsymbol{x}} \boldsymbol{b}^T \boldsymbol{x} \quad \text{subject to } \boldsymbol{x} \in \Omega.$$

Suppose that $\boldsymbol{b} \neq \boldsymbol{0}$ and the problem has a global minimizer. Can the minimizer lie in the interior of Ω?

4 Convexity

4.1 (by hand)

Let $A \in \mathbb{R}^{m \times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$. Show that the polyhedron

$$P = \{\boldsymbol{x} \in \mathbb{R}^n | A\boldsymbol{x} \leq \boldsymbol{b}\}$$

is a convex set.

4.2 (in Python)

Consider the functions

$$\begin{aligned}
f_1(x) &= x^2 + x + 1 \\
f_2(x) &= -x^2 + x + 1 \\
f_3(x) &= x^3 - 5x^2 + x + 1 \\
f_4(x) &= x^4 + x^3 - 10x^2 - x + 1
\end{aligned}$$

1. Compute and plot $f_i(x)$, $f_i'(x)$, and $f_i''(x)$ for $i = 1, 2, 3, 4$, using the range $-2 \leq x \leq 2$ for $i = 1, 2$ and $-4 \leq x \leq 4$ for $i = 3, 4$.
2. From the plots, can you recognize which of the functions are convex in the given range?
3. Locate graphically the local minimizers and maximizers of each function.
4. Locate graphically the global minimizers and maximizers of each function.
5. Look at the plots you have generated and state conditions for a local minimum and a local maximum, respectively.

4.3 (by hand)

Consider the following function

$$f(\boldsymbol{x}) = f(x_1, x_2) = 2x_1^2 - 2x_1x_2 + \frac{1}{2}x_2^2 + 3x_1 - x_2.$$

1. Express the function in matrix-vector form.
2. Is the Hessian singular?
3. Is f a convex function?

4.4 (by hand)

Suppose that f is a convex function. Show that the set of global minimizers of f is a convex set.

4.5 (by hand or in Python)

Is the problem

$$\begin{aligned} \max_{\boldsymbol{x}=[x_1,x_2]^T} \quad & x_1 \log x_1 \\ s.t. \quad & x_1^2 + x_2^2 \leq 25 \\ & \boldsymbol{x} \geq \mathbf{0} \end{aligned}$$

a convex problem?