Assignment for Lecture 4

GRADIENT DESCENT, REGULARIZATION, AND MECHANISM OF MOMENTUM

Lecture Date: 3/25/2026

"C" denotes for "computational" problems, language suggestion: Python/Julia

please include codes and results with analyses for computational problems

please write in pdf format and submit to `bjcai@fudan.edu.cn` *before the lecture of 4/1/2026*

1. For the function $f(x) = x^{2n}$ with n an integer greater than 2, work out the expression for the number of the effective steps to terminate at the condition $|x^{(i+1)} - x^{(i)}| \lesssim \sigma$ using the gradient descent algorithm.

2. Write down the expression for $\vec{H}\vec{g}$ using the five-point algorithm.

3. The condition number is defined as $\kappa(f) = x'f(x)/f(x)$ for the 1-dimensional function $f(x)$. Write down the condition number for the function $f(x) = e^x$, $\sin x$, $\cos x$, $\tan x$, $\arcsin x$, $\arccos x$, $\arctan x$, $\cosh x$, $\sinh x$ and $\tanh x$.

4. [C] Use the gradient descent with/without the exact-line search mechanism and the gradient descent with Polyak/Nesterov scheme to find the minimum of the objective function

$$f(\mathbf{x}) = 1 + \sum_{i=2}^{10} \left[100\left(x_i - x_{i-1}^2\right)^2 + (1 - x_{i-1})^2 \right], \tag{1-1}$$

with initial points set as $x_i^{(0)} = i/10$, $i = 1, 2, \cdots, 10$.