Statistical Learning for Engineers (EN.530.641)
Homework 1

Jin Seob Kim, Ph.D.
Senior Lecturer, ME Dept., LCSR, JHU

Out: 09/09/2022
Due: 09/16/2022 by midnight EST

1 (a) In the class, we have discussed the properties that define a vector space. Let $\mathbf{x} \in \mathbb{R}^n$. Prove that

$$0\,\mathbf{x} = \mathbf{0}$$

only by using those properties. Do not use component form calculations. Here 0 denotes scalar zero, while $\mathbf{0}$ means the zero vector in $\mathbb{R}^n$.

(b) Using part (a), prove that $-\mathbf{x} = (-1)\mathbf{x}$. Note that through this, one can define the difference between two vectors $\mathbf{x}$, $\mathbf{y}$ in a vector space, i.e., $\mathbf{x} - \mathbf{y}$.

(c) Prove that $\mathbb{R}^{m\times n}$ (the set of all real $m \times n$ matrices) with matrix addition, $+$, and scalar multiplication, $\cdot$, forms a vector space.

2 (a) Let's consider a system of equations

$$\begin{pmatrix} 1 & 1 & 2 \\ 1 & -1 & 2 \\ 2 & 0 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}.$$

Does this system of equations have a solution? How do you know?

(b) Suppose that, for $\mathbb{R}^2$, we have a basis set $\{\mathbf{v}_1, \mathbf{v}_2\}$ that is not orthonormal. How do you obtain the orthonormal basis set? Can you generalize the process into $\mathbb{R}^n$?

3 In the class, we have learned FONC as $\mathbf{v}^T \nabla f(\mathbf{x}^*) \geq 0$ for a local minimizer $\mathbf{x}^*$ ($\mathbf{v}$ is a feasible direction). If $\mathbf{x}^*$ is located inside the domain, then FONC becomes

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

Prove that this is indeed true.

4 Given a function $f : \mathbb{R}^2 \to \mathbb{R}$ which is explicitly written as $f(x_1, x_2) = x_1^2 + 2x_1x_2 + x_2^3$,

(a) Compute the gradient ∇f.

(b) Compute the Hessian $D^2 f$.

(c) Expand $f(x_1, x_2)$ by using the Taylor series about the point $(1, 2)$ up to the second order.

5 Let $f : \mathbb{R}^2 \to \mathbb{R}$ have the following form

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

where $\mathbf{x} = [x_1, x_2]^T$.

(a) Draw a contour plot using python. (Hint: look for `contourf` in numpy).

(b) Can you find any minimum point? (do it analytically, you are to do it numerically in the next problem). Is it a local or a global minimizer?

6 Now you are asked to write your python code for finding a minimum of the function given in Problem **5**. Choose any initial point that is not close to the minimum. Then implement the following algorithms:

(a) Gradient descent algorithm, where α_k is constant. You will have to play with this parameter.

(b) Newton's algorithm.

As the outcomes:

- Report $\min f$ and $\mathbf{x}^*$ (the corresponding minimum point). And compare it with the results in Problem **5** (b).
- Generate the plot of function values vs. step numbers.
- Generate plots that looks like Fig. 1.

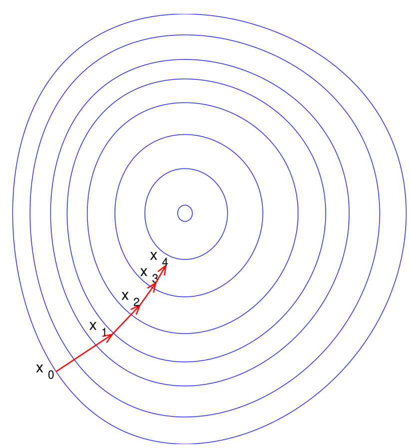

Figure 1: Problem **6**: An exemplary figure that shows the steps and movements during the gradient-based optimization process, with contour plot of the function. This is from wikipedia.

Submission Guideline

- Submit the analytic parts of the homework (Problem 1 to 5) in pdf format, including plots, to "HW1_analytical" on Gradescope.
  - No more than two (2) homework problems may be on the same page. In other words, for each problem your answers should be on a separate set of pages (e.g., for Problem 1, your answer is on page 1–3, and for Problem 2, your answer is on page 4-5, and so on).

  - When submitting, you should assign the pages to each problem on Gradescope. You can scan your answers or use an app (e.g., Adobe Scan) to generate a pdf file. *Show your work.*

- Submit to "HW1_computational" on Gradescope a single zip file that contains all the codes for each computational problem (if any, codes for analytic parts as well). Name your single zip file as "YourName_HW1.zip". For example, "JinSeobKim_HW1.zip" for a single zip file. Submission will be done through the Gradescope.
- Just in case you have related separate files, please make sure to include *all the necessary files*. If TAs try to run your function and it does not run, then your submission will have a significant points deduction.
- Make as much comments as possible so that the TAs can easily read your codes.