Homework 4 and Homework 5
CSCE 590 - Section 002: Optimization   Instructor - Vignesh Narayanan
Spring 2024
Due Date - April 9

Instructions:

- Submit a single PDF with solutions. Use the following format to name the PDF file - "CSCE790-HW-45-Lastname."
- The solutions should be very clear and should follow all the instructions below.
- If the solutions are not readable, they will not be graded
- If you refer to any resource to get your solutions, add an acknowledgement and list the references (details of the source, e.g., book, website, etc.)
- Include codes to the problems, i.e., link to GitHub or Notebook, and they should be clearly commented
- Do not include colab printed page or .ipynb file in your solution as an additional document. You may add their links in the PDF and the codes should be accessible with the link without needing additional permissions.
- Add good figures; the figures should be generated as pdf, eps, or svg file and added to the solution.
- Figures should have detailed captions
- If you do calculations or derivations, clearly include them and all the steps that led to the final solution

Part A: (Coding)

1. For the objective function $f(x) = 5x_1^2 + x_2^2 + 2x_3^2 + 4x_1x_2 - 14x_1 - 6x_2 + 20$, find the minimizer via:
   (a) Steepest descent (SD) algorithm with Bisection algorithm for stepsize selection
   (b) Newtons'(NR) algorithm

   Plot - (a) cost vs iteration in each case. (b) Cost contours and evolution of the solution sequence on it.

2. Find the quadratic polynomial

$$p(t) = x_0 + x_1 t + x_2 t^2 + \ldots + x_k t^k, \quad \text{for } k \in \mathbb{N}$$

   that best fits the following data in the least squares sense:

t	-2	-1	0	1	2
y	2	-10	0	2	1

   (a) Use the algorithms (a) and (b) from the previous problem to solve this fitting problem.
   (b) Plot the fitted polynomial $(p(t))$ along with the given input (t) - output (y) data.
   (c) Plot error vs k.

3. Find the functional approximation

$$Y = f(X) = x_0 + x_1 \sin(x_2(X)) + x_3 \cos(x_4(X))$$

that best fits the following data points by choosing appropriate x_0, x_1, x_2, x_3, x_4:

$$X = \begin{bmatrix} -1 & -0.9 & -0.8 & -0.7 & -0.6 & -0.5 & -0.4 & -0.3 & -0.2 & -0.1 & 0 & 0.1 & 0.2 & 0.3 & 0.4 & 0.5 & 0.6 & 0.7 & 0.8 & 0.9 & 1 \end{bmatrix}$$

and

$$Y = \begin{bmatrix} -0.96 & -0.577 & -0.073 & 0.377 & 0.641 & 0.66 \\ 0.461 & 0.134 & -0.201 & -0.434 & -0.5 & -0.393 & -0.165 & 0.099 & 0.307 & 0.396 & 0.345 & 0.182 & -0.031 & -0.219 & -0.321 \end{bmatrix}.$$

   (a) State the optimization problem.

   (b) Plot the fitted funciont $(f(X))$ along with the given input (X) - output (Y) data. Use any gradient algorithm to solve the fitting problem.

   (c) Plot: cost vs iteration

4. Let

$$A^T = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 2 \\ 1 & -1 & 0 \\ 1 & 1 & 2 \end{bmatrix}$$

   (a) Use the Gram-Schmidt Process to find $\mathcal{R}(A)$ (i.e., the space spanned by the column vectors of A).

   (b) Graph the column vectors of A and graph the orthonormal vectors that you obtain from the Gram-schmidt procedure.

Part B: (Theory and Derivations)

1. For any given matrix A, the projection $P_{\mathcal{R}(A)}$ onto the range space of A is given by $P_{\mathcal{R}(A)} = AA^\dagger$. For the matrix

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 2 \\ 2 & 3 \end{bmatrix}$$

   (a) Compute the generalized inverse $A^\dagger$.

   (b) Compute $P_{\mathcal{R}(A)}$.

   (c) Let $A \in \mathbb{R}^{m \times n}$ be such that the null space, $\mathcal{N}(A) = \{0\}$. Show that $A^T A$ is invertible.

2. Let $A \in \mathbb{R}^{m \times n}$, $m > n$, $\text{rank}(A) = n$, and $b \in \mathbb{R}^m$. The least squares problem

$$\min \quad \|Ax - b\|^2 \tag{*}$$

can be formulated as the linear algebraic system

$$\begin{bmatrix} I_m & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} r \\ x \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix}, \tag{**}$$

where I_m stands for the $m \times m$ identity matrix and $r \doteq b - Ax \in \mathbb{R}^m$ is the residual.

   (a) Using the normal equations, show that the component x of the solution of (**) solves the least squares problem (*).

3. Consider the iterative process

$$x_{k+1} = \frac{1}{2}(x_k + \frac{a}{x_k}) \tag{1}$$

where $a > 0$. We assume that this process converges.

- To what does it converge?
- What is the order of convergence?

4. The mean value theorem from one-variable calculus states that if a function $f : [a, b] \to \mathbb{R}$ is continuous on the closed interval $[a, b]$ and differentiable in the open interval (a, b), then there is a point $a < c < b$ such that

$$f(b) - f(a) = (b - a)f'(c)$$

Does this theorem remain true for a vector-valued function $f : [a, b] \to \mathbb{R}^2$? Use the example $f : [0, 1] \to \mathbb{R}^2$ defined by

$$f(x) = (f_1(x), f_2(x)) = (x(1 - x), x^2(1 - x))$$

to support your answer.