555 HW1: Due February 4, 2021

1 Theoretical Problems

1. Consider the function $f(x, y) = e^{x^2+2y^2}$.
   (a) Compute the Hessian matrix of $f(x, y)$.
   (b) Is the Hessian matrix symmetric positive definite for $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$?
   (c) Is the Hessian matrix symmetric positive definite for all $\begin{pmatrix} x \\ y \end{pmatrix}$?
2. Consider the function $f(x, y, z) = yz + e^{xyz}$. At the point $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$, find the direction along which the function decreases most rapidly.
3. Consider $f(x, y) = \frac{1}{6}x^2 - \frac{1}{3}xy + \frac{1}{6}y^2 + 1$.
   (a) Find the Hessian matrix H of $f(x, y)$.
   (b) Find the eigenvalues $\lambda_1, \lambda_2 \in \mathbb{R}$ and eigenvectors $v_1, v_2 \in \mathbb{R}^2$ of H.
4. Consider $w, b \in \mathbb{R}$ and the function that
$$f(w, b) = e^{x_1 w + b},$$
   for $x_1 \in \mathbb{R}$.
   (a) Consider the Hessian matrix of f defined by
$$H(w, b) = \nabla^2 f(w, b) = \begin{pmatrix} \frac{\partial^2 f}{\partial w^2} & \frac{\partial^2 f}{\partial w \partial b} \\ \frac{\partial^2 f}{\partial b \partial w} & \frac{\partial^2 f}{\partial b^2} \end{pmatrix}.$$
   Verify that
$$H(w, b) = f(w, b)\boldsymbol{x}\boldsymbol{x}^T,$$
   where
$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ 1 \end{pmatrix}.$$
   (b) Prove that
$$v^T H(w, b) v \geq 0,$$
   for any $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \in \mathbb{R}^2$ and $(w, b) \in \mathbb{R}^2$.

2 Computational Problems

1. Use GD method to solve the following problem.

   (a) Consider the following objective function:

   $$f(x_1, x_2) = (x_1 - 1)^2 + \alpha(x_2 - 1)^2, \tag{1}$$

   for a parameter $\alpha > 1$.

   (b) Implement gradient descent to solve this problem with initial guess $(0, 0)$.

   (c) Test the algorithm with different stepsizes and different values of α (say for $\alpha = 1, 10, 100, 1000$).

   (d) What conclusions can you draw about gradient descent?

2. Given

   $$A_0 = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix} \in R(A_0)$$

   (a) Prove that A_0 is positive semi-definite and find all the eigenvalues and eigenvectors of A_0.

   (b) Prove that $(A_0 + \epsilon I)x = b$ is uniquely solvable for $\epsilon > 0$ and also solvable (and unique up to a constant vector) for $\epsilon = 0$.

   (c) Let $A_\epsilon = A_0 + \epsilon I$ and use gradient descent method to solve

   $$x = \underset{y \in \mathbb{R}^3}{\arg\min} \left(\frac{1}{2}(A_\epsilon y, y) - (b, y) \right)$$

   with the initial guess $x^0 = b$ and learning rate $\eta = \frac{1}{3}$, record the minimal number of iteration m satisftying the stopping criterion that $\|Ax^m - b\| \leq 10^{-8}$:

ε	# of iter $= m$
1.	
10^{-1}	
10^{-2}	
10^{-3}	
10^{-4}	
0. [singular case]	

   (d) Preconditioned gradient descent: let

   $$P = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

   and $\underset{\sim}{A} = P^T A_\epsilon P, \underset{\sim}{b} = P^T b$. Use gradient descent method to solve

   $$\underset{\sim}{x} = \underset{\underset{\sim}{y} \in \mathbb{R}^4}{\arg\min} \left(\frac{1}{2}(\underset{\sim}{A}\underset{\sim}{y}, \underset{\sim}{y}) - (\underset{\sim}{b}, \underset{\sim}{y}) \right)$$

   with the initial guess $\underset{\sim}{x}^0 = \underset{\sim}{b}$ and learning rate $\eta = \frac{1}{3}$, record the minimal number of iteration m satisftying the stopping criterion that $\|\underset{\sim}{A}\underset{\sim}{x}^m - \underset{\sim}{b}\| \leq 10^{-8}$:

ε	# of iter $= m$
1.	
10^{-1}	
10^{-2}	
10^{-3}	
10^{-4}	