Preliminaries, Optimality Conditions and Gradient Descent.

1. Let $G \in \mathbb{R}^{n\times n}$ be a symmetric matrix, $c \in \mathbb{R}^n$, and define
$$f(x) := \frac{1}{2}x^T Gx + c^T x$$
Using Taylor's theorem, argue that $\nabla f(x) = Gx + c$.

2. Let $A \in \mathbb{R}^{m\times n}$. Define $\|A\|_2 := \max\limits_{\|x\|_2=1} \|Ax\|_2$.
   (a) Show that
   $$\|A\|_2 = \sqrt{\lambda_{\max}(A^T A)}.$$
   (b) Show that
   $$\|AB\|_2 \le \|A\|_2\|B\|_2$$
   for any $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{n\times k}$.

3. Compute the gradient and Hessian of the Rosenbrock function:
$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2.$$
Show that $x^* = \begin{bmatrix}1 & 1\end{bmatrix}^T$ is its only minimizer and the Hessian is positive definite at x^*.

4. For the following functions, find all the stationary points and determine their nature, if possible.
   (a) $f(x) = x_1^4 + x_2^4 - 4x_1x_2 + 2$;
   (b) $f(x) = x_1^2 - x_1x_2 + x_2^2 + 9x_1 - 6x_2 + 10$;
   (c) $f(x) = x_1x_2 - x_1^3 - x_2^2$;
   (d) $f(x) = 3x_1x_2 - x_1^2x_2 - x_1x_2^2$.

5. Consider the function $f(x) = (x_1 + x_2^2)^2$, the point $x^* = \begin{bmatrix}1 & 0\end{bmatrix}^T$ and the direction $d^* = \begin{bmatrix}-1 & 1\end{bmatrix}^T$.
Show that d^* is a descent direction of f at x^*, and find all stepsizes that satisfy the *exact line search* criterion (Slide 18 of Lecture 2) at x^* along d^*.

6. Consider the function $f : \mathbb{R}^3 \to \mathbb{R}$ defined by
$$f(x) = (x_1)^2 + (x_2)^2 + (x_3)^2 + \big(x_1 + x_2 - 1\big)^2$$
Consider an iterate of the following form:
$$x^{k+1} = x^k + \alpha_k d^k,$$
where $\alpha_k > 0$ is a stepsize and d^k is a search direction.
   (a) Let $d^k = -D\nabla f(x^k)$, where D is the following **positive definite matrix**:
   $$D = \begin{bmatrix}2 & 1 & 0\\ 1 & 2 & 0\\ 0 & 0 & 1\end{bmatrix},$$
   and let α_k be chosen to satisfy the Armijo rule. Suppose that x^k is nonstationary for all k.
      i. Find all eigenvalues of D.

ii. Show that $\{x^k\}$ is bounded.

(b) Let $d^k = -0.1\nabla f(x^k)$. Show that if α_k is set to be 3.321 for all k, then any accumulation point of $\{x^k\}$ is stationary.

7. Consider the function $f : \mathbb{R}^3 \to \mathbb{R}$ defined by

$$f(x) = (x_1)^2 + (x_2)^2 + (x_3)^2 + \ln(1 + (x_1 + x_2 - x_3 - 7)^4) + \ln(1 + (3x_1 + x_3 + 2)^4)$$

(a) Let $h : \mathbb{R}^2 \to \mathbb{R}$ be defined by $h(y) = \ln(1 + (y_1)^4) + \ln(1 + (y_2)^4)$. Show that

$$\|\nabla^2 h(y)\|_2 \leq 6 \quad \text{for all } y \in \mathbb{R}^2.$$

(b) Suppose that the steepest descent with constant stepsize $\alpha = 0.0305$ is applied to minimize f. Argue that any accumulation point of the sequence generated is a stationary point of f.

8. Consider the function $f : \mathbb{R}^3 \to \mathbb{R}$ defined by

$$f(x) = \ln(1 + (2x_1 + x_2 + x_3 + 2022)^2) + \frac{1}{16}(x_1 - x_2 + 2x_3)^2 + \frac{1}{2}\|x\|_2^2.$$

Argue that any accumulation point of the sequence generated by steepest descent with constant stepsize $\alpha = \frac{1}{\pi^2}$ is a stationary point of f.

9. Consider the function $f : \mathbb{R}^3 \to \mathbb{R}$ defined by

$$f(x) = \ln(1 + (x_1 + x_2 - x_3)^2) + \ln(1 + \exp(x_1 - 3x_2 + x_3 + 2)) + \frac{1}{2}\|x\|_2^2.$$

Argue that any accumulation point of the sequence generated by steepest descent with constant stepsize $\alpha = \frac{\pi}{40}$ is a stationary point of f.

10. Consider the function

$$f(x) = \sum_{i=1}^{m} \ln(1 + e^{-a_i^T x}) + \frac{1}{2}\|x\|_2^2,$$

where $a_i \in \mathbb{R}^n$ for each $i = 1, \ldots, m$. Let $A \in \mathbb{R}^{m\times n}$ with the ith row being a_i^T and suppose that $\|A\|_2 = 1$. Argue that any accumulation point of the sequence generated by steepest descent with constant stepsize $\alpha = 1.5$ is a stationary point of f.

11. Consider the function $f(x) = e^{-x}$. Consider an iterate of the following form

$$x_{k+1} = x_k + \alpha_k d_k,$$

where $d_k = -f'(x_k)$ and α_k is obtained via Armijo line search by backtracking with $\bar{\alpha}_k \equiv 1$ and $\sigma = 0.1$. Start with $x_0 = 0$.

(a) Show that $e^{-y} \leq 1 - 0.1y$ whenever $y \in [0, 1]$.

(b) Show that $\alpha_0 = 1$ and $x_1 = 1$.

(c) Show that, for all $k \geq 0$, it holds that $x_{k+1} > 0$ and $\alpha_k = 1$.

12. Let $Q \succ 0$ and $b \in \mathbb{R}^n$. Define

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

(a) Suppose that $\bar{x}$ is not a stationary point of f and suppose the steepest descent method with exact line search is applied to minimizing f starting from $\bar{x}$.

Show that the stepsize that satisfies the *exact line search* criterion at $\bar{x}$ along $-\nabla f(\bar{x})$ is given by

$$\frac{\|\nabla f(\bar{x})\|^2}{[\nabla f(\bar{x})]^T Q \nabla f(\bar{x})}.$$

(b) Suppose that x^* is the unique minimizer of f (we shall discuss solution existence in Lecture 10) and let v be any eigenvector of Q.

i. Let $x^0 = x^* + v$. Show that x^0 is not a stationary point of f.

ii. Show that the steepest descent method with exact line search initialized at $x^0 = x^* + v$ gives $x^1 = x^*$.

Remarks on the level of difficulties:

(I) Q1 and Q2 are extensions of lecture discussions. They are not exam-type problems.

(II) Q3 thru Q10 are of basic level.

(III) Q11 and Q12 are of advanced level.