Math-484 Homework #11 (iterative methods)

I will finish the homework before 11am Nov 20. If I spot a mathematical mistake I will let the lecturer know as soon as possible.

I will write clearly and neatly as the grader is not an expert in cryptography. I will sign each paper of my work and indicate if I a 4 credits student.

Do all the exercises (no matter if 3 or 4 credits). You are allowed to use software for computing product of matrices and inverses of matrices.

1: (*Try Newton's Method*)
Compute the first two terms $\mathbf{x}_1, \mathbf{x}_2$ of the Newton's Method sequence $\{\mathbf{x}_k\}$ for minimizing the function

$$f(x_1, x_2) = 2x_1^4 + x_2^2 - 4x_1x_2 + 5x_2$$

with initial point $\mathbf{x}_0 = (0, 0)$.

2: (*Newton's method does not have to converge*)
Show that the function $f(x) : \mathbb{R} \to \mathbb{R}$ by

$$f(x) = |x|^{4/3}$$

has a unique global minimizer at $x^* = 0$ but that, for any nonzero initial point x_0, the Newton's Method sequence $\{x_k\}$ with initial point x_0 for minimizing $f(x)$ diverges.
Hint: Solve for $x_k > 0$ and use symmetry.

3: (*Newton's method works like quadratic approximation*)
(a) Compute quadratic approximation $q(\mathbf{x})$ for

$$f(x_1, x_2) = 8x_1^2 + 8x_2^2 - x_1^4 - x_2^4 - 1$$

at the point $(\frac{1}{2}, \frac{1}{2})$.
(b) Compute the minimum $\mathbf{x}^*$ of the quadratic approximation $q(\mathbf{x})$ at $(\frac{1}{2}, \frac{1}{2})$.

4: (*Try steepest descent method*)
Compute the first two terms $\mathbf{x}_1, \mathbf{x}_2$ of the Steepest Descent sequence $\{\mathbf{x}_k\}$ for minimizing the function

$$f(x_1, x_2) = 2x_1^4 + x_2^2 - 4x_1x_2 + 5x_2$$

with initial point $\mathbf{x}_0 = (0, 0)$.

5: (*Steepest descent method may be slow*)
Consider applying the steepest descent method to the problem

$$\text{minimize } f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T H \mathbf{x},$$

where H is a positive definite matrix. Notice $\mathbf{x} = 0$ is the optimal solution.
a) Recall that steepest descent sequence is defined as $\mathbf{x}_{k+1} = \mathbf{x}_k - t_k \nabla f(\mathbf{x}_k)$ for t_k minimizing

$\varphi(t) = f(\mathbf{x}_k - t_k \nabla f(\mathbf{x}_k))$. Express t_k using $\mathbf{x}_k$ and $\nabla f(\mathbf{x}_k)$.
b) Assume that H is a diagonal matrix (entries not in the diagonal are zeros)

$$H = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

where the entries on the diagonal are $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Let the starting point be

$$\mathbf{x}_0 = \left(\frac{1}{\lambda_1}, 0, \ldots, 0, \frac{1}{\lambda_n} \right)^T .$$

Show that

$$\mathbf{x}_1 = \frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n} \left(\frac{-1}{\lambda_1}, 0, \ldots, 0, \frac{1}{\lambda_n} \right)^T .$$

Give a formula for $\mathbf{x}_k$ (I mean a formula using only k, λ_1, and λ_n but NOT using $\mathbf{x}_{k-1}$ or t_k). Using the formula conclude if the method converges faster when $\lambda_1 = \lambda_n$ or when $\lambda_1 \gg \lambda_n$.