MATH 437 Homework 7 (20 points)

1. Consider the system

$$x_1^2 - x_2^2 = 0,$$
$$x_1^3 + x_2^3 - 16 = 0.$$

   (a) (2 points) Write down Newton's method for this system.

   (b) (2 points) Compute 7 iterations of Newton's method for this system, starting with the initial condition $x_1 = x_2 = 1$. Report the iterations in a table with columns n, $x_1^{(n)}$, $x_2^{(n)}$, where $x_i^{(n)}$ is the value of x_i at the nth iteration and $n = 0, \dots, 7$.

   (a) *Hint.* Define a function $F : \mathbb{R}^2 \to \mathbb{R}^2$ where $F(x)$ is related to the given system. Then, compute the Jacobian matrix

$$\nabla F(x)_{i,j} := \partial_j F_i(x).$$

   Newton's method is then

$$x^{(n+1)} = x^{(n)} - \nabla F(x^{(n)})^{-1} F(x^{(n)}).$$

   Note: the original system in the homework posted by the professor cannot be solved with Newton's method, because the Jacobian of F becomes singular at the root for that system. Thus, I modified the problem to be solveable. □

   (b) *Hint.* See `problem_1b.py`. □

2. (4 points) Consider the minimization problem

$$\min_{x_1, x_2 \in \mathbb{R}} \left\{ (x_1 + x_2 - 4)^2 + (x_1^2 + x_2^2 - 8)^2 \right\}.$$

Write the gradient descent method for this problem and compute 2 iterations starting with initial guess $x_1 = x_2 = 1$.

Hint. Define a function $g : \mathbb{R}^2 \to \mathbb{R}$ where $g(x)$ is related to the minimization problem. Then, compute its gradient

$$\nabla g(x)_j := \partial_j g(x).$$

The gradient descent method involves computing an optimal step size $0 < \alpha_n \leq 1$ such that

$$x^{(n+1)} := x^{(n)} - \alpha_n \frac{g(x^{(n)})}{\|g(x^{(n)})\|_{\ell^2}}$$

so that $g(x^{(n+1)}) \leq g(x^{(n)})$. Explain how to find α_n by minimizing a related quadratic polynomial. For the implementation, see `problem_2.py`. □

3. (4 points) Find a singular-value decomposition $A = U \Sigma V^T$ for the matrix

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

Hint. Recall that U is a $m \times m$ orthogonal matrix, Σ is a $m \times n$ matrix with the singular values on the main diagonal, and V is a $n \times n$ orthogonal matrix. Find the singular values and the columns of V by first finding the eigenvalues and eigenvectors of $A^T A$. Use the columns of V to find the first n columns of U, and then extend to an orthonormal basis of $\mathbb{R}^m$ to find the remaining columns of U. $\square$

4. Let $M \geq N$. Given N vectors $a_1, \ldots, a_N \in \mathbb{R}^M$, we form the $M \times N$ matrix $A := [a_1 \cdots a_N]$. Let $\sigma_1 \geq \cdots \geq \sigma_N \geq 0$ denote the singular values of A, and let $u_1, \ldots, u_N \in \mathbb{R}^M$ be the columns of U in a singular-value decomposition $A = U\Sigma V^T$, where u_j corresponds to σ_j and σ_j is the ith diagonal entry of Σ. It is known that u_1 is a minimizer for

$$\min_{\substack{u\in\mathbb{R}^M \\ \|u\|_2=1}} \sum_{j=1}^{N} \|a_j - \langle a_j, u\rangle_2 u\|_2^2,$$

where

$$\langle v, w\rangle_2 := \sum_{i=1}^{M} v_i w_i$$

and

$$\|w\|_2^2 := \sum_{i=1}^{M} w_i^2.$$

We investigate this via a computational example.

(a) (2 points) Let $M = 80$ and $N = 20$. Construct a $M \times N$ matrix A with random entries. Perform a singular-value decomposition of A to find the unit eigenvector u_1 corresponding to the largest singular value.

(b) (1 point) Compute the projection error

$$\mathrm{Err} := \sum_{j=1}^{N} \|a_j - \langle a_j, u_1\rangle_2 u_1\|_2^2$$

and the relative projection error

$$E := \frac{\mathrm{Err}}{\sum_{j=1}^{N} \|a_j\|_2^2}.$$

(c) (1 point) Construct a random unit vector $u \in \mathbb{R}^M$ and compute the projection

$$\sum_{j=1}^{N} \|a_j - \langle a_j, u\rangle_2 u\|_2^2.$$

Hint. See `problem_4.py`. $\square$

5. Given $f, g : \mathbb{R}^2 \to \mathbb{R}$, consider the system

$$\begin{aligned} x &= f(x, y), \\ y &= g(x, y), \end{aligned}$$

and the non-standard fixed-point iteration

$$\begin{aligned} x_{n+1} &= f(x_n, y_n), \\ y_{n+1} &= g(x_{n+1}, y_n). \end{aligned}$$

(a) (2 points) Determine the convergence criteria for this method.

(b) (2 points) Let $f(x, y) = 2y$ and $g(x, y) = (x - y)/2$. Find the matrices T_1 and T_2 such that the non-standard fixed point iteration in matrix form is

$$\begin{bmatrix} x_{n+1} \\ y_{n+1} \end{bmatrix} = T_1 \begin{bmatrix} x_n \\ y_n \end{bmatrix},$$

and the standard fixed-point iteration in matrix form is

$$\begin{aligned} x_{n+1} &= f(x_n, y_n), \\ y_{n+1} &= g(x_n, y_n). \end{aligned} \implies \begin{bmatrix} x_{n+1} \\ y_{n+1} \end{bmatrix} = T_2 \begin{bmatrix} x_n \\ y_n \end{bmatrix}.$$

Find the spectral radius of the matrices T_1 and T_2. Based on this, which method, if any, converges?

(a) *Hint.* Let x_{n+1}, y_{n+1} be obtained from the method above, and let x, y be the solution to

$$\begin{aligned} x &= f(x, y), \\ y &= g(x, y). \end{aligned}$$

Assume that f and g are smooth functions. To determine a convergence criteria, we look at the error in each component

$$\begin{aligned} x_{n+1} - x &= f(x_n, y_n) - f(x, y), \\ y_{n+1} - y &= g(x_{n+1}, y_n) - g(x, y). \end{aligned}$$

By using Taylor's theorem on each equation, obtain an equation in matrix-vector form:

$$\begin{bmatrix} x_{n+1} - x \\ y_{n+1} - y \end{bmatrix} = T_1(\xi_n, \eta_n) \begin{bmatrix} x_n - x \\ y_n - y \end{bmatrix},$$

where $T_1(\xi_n, \eta_n)$ is a 2×2 matrix whose entries are composed of partial derivatives of f and g at some points ξ_n and η_n that come from Taylor's theorem. The method converges if and only if the spectral radius of this matrix is strictly less than 1. □

(b) *Hint.* You already computed the matrix T_1 in the previous step. Just substitute for f and g and find its largest eigenvalue in magnitude. The matrix T_2 is just the Jacobian of the function $F(x, y) = (f(x, y), g(x, y))$. The class notes discuss this in more detail. □