

# Homework 1

APM 523

Arizona State University

## Problem 1

Implement the Golden Section Search algorithm for finding a minimum of

$$f(x) = 10 + x^2 - 10 \cos(2\pi x), \quad x \in [a, b],$$

where  $a = 1.6$  and  $b = 2.4$ . You can use a tolerance  $\epsilon = 10^{-4}$  and maximum iterations 101.0. What is  $x_{\min}$  and what is  $f(x_{\min})$ ?

## Problem 2

Implement the Nelder-Mead simplex algorithm and test it on the even Rosenbrock function

$$f(x) = \sum_{i=1}^{n/2} 100(x_{2i-1}^2 - x_{2i})^2 + (1 - x_{2i-1})^2$$

You can use the following parameters

$X_0 = [0 \quad I_n]$  (initial vertices),  $\epsilon = 10^{-4}$  (stopping),  $M = 1000$  (max iter)

$$\alpha = 1, \quad \beta = \frac{1}{2}, \quad \gamma = 1$$

1. Test the method when  $n = 2$
2. Also, test the method when  $n = 8$

## Problem 3

The contraction of the simplex in the Nelder-Mead algorithm happens when no point improves on  $x^{(n)}$ . This part of the algorithm overwrites the previous vertices by

$$x^{(i)} \leftarrow \frac{1}{2}(x^{(i)} + x^{(n)}), \quad 0 \leq i \leq n$$

Show that if  $f$  is a convex function, the contraction step will not increase the average value of the function over the simplex vertices defined by  $\frac{1}{n+1} \sum_{i=0}^n f(x^{(i)})$ . Show that unless  $f(x^{(0)}) = f(x^{(1)}) = \dots = f(x^{(n)})$ , the average value will in fact decrease. Hint: You may use the definition of convexity:  $f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y)$  for all  $\alpha \in [0, 1]$  and here  $x, y \in \mathbb{R}^n$ .

## Problem 4

Show that the steepest descent method with an “exact” line-search converges in one iteration to  $x^*$  when  $f(x)$  is a convex quadratic function, and  $x_0 - x^*$  is parallel to an eigenvector of  $A \in \mathbb{R}^{n \times n}$

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

Hint: Assume that  $x_0 = x^* + q$  where  $q$  is some eigenvector of  $A$  with eigenvalue  $\lambda$ . Moreover, the value for  $\alpha_k = \min_{\alpha > 0} f(x_k - \alpha \nabla f(x_k))$  in the line-search can be analytically computed.

## Problem 5

Consider a logistic regression problem with two classes. Given a training set  $P$  consisting of datapoint and label pairs  $(z, y)$  where  $z \in \mathbb{R}^n$  and  $y \in \{-1, +1\}$ , we define the objective function  $f$  (or loss) for variables  $x \in \mathbb{R}^n$  (or weight vector) to be

$$f(x) = \sum_{(z,y) \in P} -\ln(\sigma(yx^T z)),$$

where  $\sigma(s) = 1/(1 + \exp(-s))$ . One says that the weight vector  $x$  is a separator for  $P$  if for all  $(z, y) \in P$ ,  $y(x^T z) \geq 0.0$ . A separator is said to be trivial if for all  $(x, y) \in P$ ,  $y(x^T z) = 0.0$ . For example  $x = 0$  is a trivial separator. Depending on the data  $P$ , there may be other trivial separators. Prove whether  $f(x)$  is convex.