SOLUTION 11 – MATH-250 Advanced Numerical Analysis I

(⋆) **Problem 4.** Consider points $(t_i, y_i)_{i=1,\ldots,m}$ and the weights $w_i > 0, i = 1, \ldots, m$. The goal of this exercise is to find the parameters x_1 and x_2 in $g(t) = x_1 + x_2 t$ such that

$$\sum_{i=1}^{m} w_i (g(t_i) - y_i)^2$$

is minimized.

(a) Show that the minimization problem is solved by

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (A^\top W A)^{-1} A^\top W \mathbf{y} \tag{1}$$

for a certain choice of A, W, and $\mathbf{y}$.

(b) Use (1) to complete the function `weighted_linear_least_squares(t, y, w)` which implements the weighted least squares.

(c) Consider the weight function

$$w(t) = (1 + \exp(-10t))^{-1}. \tag{2}$$

Let the weights $w_i = w(t_i), i = 1, \ldots, m$. Use these weights to fit the data given by the arrays `t` and `y` in the Jupyter notebook.

(d) Consider the function $f(t) = \cos(t)$ and the points $t_1 = -1$, $t_2 = 0$, and $t_3 = 1$. Compute the coefficients x_1 and x_2 which minimize

$$\sum_{i=1}^{m} w(t_i)(g(t_i) - f(t_i))^2$$

for the unweighted case $w(t) = 1$ and for the weight function (2). Comment on the result.

Solution.

(a) We rewrite the minimization target () as

$$f(\mathbf{x}) = \|W^{1/2}(A\mathbf{x} - \mathbf{y})\|_2^2,$$

where

$$A = \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_m \end{pmatrix} \in \mathbb{R}^{m \times 2}, \ \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} \in \mathbb{R}^m, \ \text{and } W = \begin{pmatrix} w_1 & & & \\ & w_2 & & \\ & & \ddots & \\ & & & w_m \end{pmatrix} \in \mathbb{R}^{m \times m}$$

Then we compute the gradient of f at $\mathbf{x}$ by perturbing $\mathbf{x}$:

$$\begin{aligned} f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) &= \|W^{1/2}(A(\mathbf{x} + \mathbf{h}) - \mathbf{y})\|_2^2 - \|W^{1/2}(A\mathbf{x} - \mathbf{y})\|_2^2 \\ &= (A(\mathbf{x} + \mathbf{h}) - \mathbf{y})^\top W (A(\mathbf{x} + \mathbf{h}) - \mathbf{y}) - (A\mathbf{x} - \mathbf{y})^\top W (A\mathbf{x} - \mathbf{y}) \\ &= 2\mathbf{h}^\top A^\top W A\mathbf{x} - 2\mathbf{h}^\top A^\top W \mathbf{y} + O(\|\mathbf{h}\|_2^2). \end{aligned}$$

Hence,

$$\nabla f(\mathbf{x}) = 2(A^\top W A\mathbf{x} - A^\top W\mathbf{y})$$

The gradient is zero if and only if the normal equations are satisfied:

$$A^\top W A\mathbf{x} - A^\top W\mathbf{y} = 0$$

Under the assumption that A has rank 2, we then get the unique solution

$$\mathbf{x} = (A^\top W A)^{-1} A^\top W\mathbf{y}.$$

(b – d) Available in the Jupyter notebook `homework11-sol.ipynb` on Moodle.