4033/5033: Assignment 2

Your Name

Due: Sep 30 (by 11:59pm)

Previously, we studied a weighted least square (WLS) learning technique

$$J(\beta) = \sum\nolimits_{i=1}^{n} w_i \cdot (x_i^T \beta - y_i)^2 \qquad (1)$$

where $w_i \in \mathbb{R}$ is the weight for instance $x_i \in \mathbb{R}^p$.

Please design some probabilistic assumptions and prove that, under those assumptions, the MLE estimation of your distribution parameter is the same as the solution to the above WLS problem.

You need to

– Clearly list ALL the assumptions you made

– Clearly explain the meaning of every notation, especially if it is not commonly used in the lectures

– Clearly explain the dimension of every matrix or vector e.g., $X \in \mathbb{R}^{n \times p}$

– Clearly elaborate the arguments to derive from MLE to WLS

– Clearly point out which part of your derived results corresponds to the weight. (For example, in lecture we show the $\frac{\sigma^2}{t^2}$ in MAP estimate corresponds to the regularization coefficient λ in ridge regression.)

<u>Submission Instruction</u>

Please submit a single pdf file 'hw2.pdf' to Canvas.

You can directly type the answers in Latex and compile them into pdf, or first hand-write the answers and scan them into pdf. In the latter case, please make sure your hand-writing is clear. If the grader does not understand the writing, you will lose points at that part.

Parts of the answers will be graded based on hard criterion e.g., lose X points if an assumption is missing. The other parts will be graded based on subjective evaluation e.g., how much partial credit can be given to an argument that does not lead to the correct result or contains flaws. The subjective evaluation will not be open for negotiation.