CSE 250B Winter 2019 - Quiz 3

Feb 8, 2019 - 20 Minutes **Name:** **PID:**

---

(1) This is a closed book, closed notes exam. Switch off your cell phone and do not communicate with anyone other than an exam proctor.
(2) Start writing when instructed. Stop writing when your time is up.
(3) Remember that your work is graded on the quality of your writing and explanation as well as the validity of the mathematics.

Alice has collected a dataset of dependent and independent variables $\{(x^{(1)}, y^{(1)}), \ldots, (x^{(n)}, y^{(n)})\}$; she does linear regression on it and obtains the weight vector w_{Alice}. Bob also collects the same dataset, but while recording the independent variables (the $x^{(i)}$'s) he uses different units. Specifically, Bob's dataset is $\{(z^{(1)}, y^{(1)}), \ldots, (z^{(n)}, y^{(n)})\}$, where for each i, $z^{(i)} = cx^{(i)}$, where $c > 0$ is a scalar. Bob does linear regression on this dataset and obtains a weight vector w_{Bob}.

(1) (5 points) Do Alice and Bob have the same training loss? Is $w_{Alice} = w_{Bob}$? In either case, justify your answer.

If X is Alice's data matrix, and Z is Bob's data matrix, then $Z = cX$. $w_{Bob} = (Z^\top Z)^{-1} Z^\top y = (c^2 X^\top X)^{-1} c X^\top y = w_{Alice}/c$; thus when $c \neq 1$, $w_{Alice} \neq w_{Bob}$.

The training loss of Bob is: $\|Z w_{Bob} - y\|^2 = \|cX \cdot w_{Alice}/c - y\|^2 = \|X w_{Alice} - y\|^2$, which by definition is equal to the training loss of Alice.

(2) (5 Points) Suppose now that Bob records each feature in a different unit; that is, for each coordinate j, $z_j^{(i)} = c_j x_j^{(i)}$ for all i, where $c_j > 0$ is a scalar, and the c_j's are not all equal. Do Alice and Bob still have the same training loss and is $w_{Alice} = w_{Bob}$? If yes, justify your answer. If no, provide an example of a dataset where this is not the case.

The answer will still be the same. Let C be a diagonal matrix whose i-th diagonal entry is c_i. Observe that Bob's data matrix $Z = X \cdot C$. Moreover, for any w, $ZC^{-1}w = Xw$. Hence, the optimal solution to the training loss for Bob $w_{Bob} = C^{-1} w_{Alice}$, and the training loss for Bob is: $\|Z w_{Bob} - y\|^2 = \|X w_{Alice} - y\|^2$ which is equal to Alice's training loss.