Problem Set 2: Classification
5. Written problems
[30 points; all points assigned to one random problem]

5.1 Derivation of LDA

Show that the log-odds decision function a(x) for LDA

$$a = \ln \frac{p(\mathbf{x}|\mathcal{C}_1)p(\mathcal{C}_1)}{p(\mathbf{x}|\mathcal{C}_2)p(\mathcal{C}_2)}$$

is linear in x, that is, we can express $a(x) = \theta^T x$ for some θ. Show all your steps.

5.3. Maximum likelihood for Logistic Regression

Showing all steps, derive the LR cost function using maximum likelihood. Assume that the probability of y given x is described by

$$\begin{aligned} P(y = 1 \mid x; \theta) &= h_\theta(x) \\ P(y = 0 \mid x; \theta) &= 1 - h_\theta(x) \end{aligned}$$

5.2 LR Classification with Label Noise

Suppose you are building a logistic regression classifier for images of dogs, represented by a feature vector x, into one of two categories $y \in \{0,1\}$, where 0 is “terrier” and 1 is “husky.” You decide to use the logistic regression model $p(y = 1|x) = h_\theta(x) = \theta^T x$. You collected an image dataset $\boldsymbol{D} = \{x^{(i)}, t^{(i)}\}$, however, you were very tired and made some mistakes in assigning labels $t^{(i)}$. You estimate that you were correct in about τ fraction of all cases.

(a) Write down the equation for the posterior probability $p(t = 1|x)$ of the label being 1 for some point x, in terms of the probability of the true class, $p(y = 1|x)$.

(b) Derive the modified cost function in terms of $\theta, x^{(i)}, t^{(i)}$ and τ.