Question 1 — 10 pts

In a statistical learning problem, we are interested in predicting the response variable. Given the type of response variables, specify the type of learning problems.

Quantitative 定量 — Regression problems

Qualitative 定質 — Classification problems

qualitative → classification

quantitative → regression

定量

Question 2 — 10 pts

We are given a set of data to predict whether an individual will default on his or her credit card payment, on the basis of annual income and monthly credit card balance. And we decided to apply Linear discriminant analysis (LDA) to build the model. Select all that apply.

- [x] The method assumes that we model each class density as multivariate Gaussian.
- [x] In total, we have 2 sets of parameters to estimate, the mean of each class and the covariance matrix.
- [ ] The model assumes that each class has its own covariance matrix. There are no constraints applied to the covariance matrices.

- Linear discriminant analysis (LDA) arises in the special case when we assume that the classes have a common covariance matrix.
- In addition to the mean of each class, the common covariance matrix, LDA also needs the estimate of the prior probability of each class.

Question 3 — 10 pts

Quadratic discriminant analysis (QDA) assumes that the distribution of each class is Gaussian. The decision boundary is linear.

- ○ True
- ◉ False

Question 4 — 10 pts

Using logistic regression, the prediction is in the range between 0 and 1, and we can interpret them as probabilities.

- ◉ True
- ○ False

Question 5 — 10 pts

Suppose we want to apply Newton's method to solving the root of a function f, i.e., solving f(x)=0. Which of the following updating rules is correct?

(1) $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$

(2) $x_{n+1} = x_n + \frac{f(x_n)}{f'(x_n)}$

(3) $x_{n+1} = x_n - \frac{f'(x_n)}{f''(x_n)}$

(4) $x_{n+1} = x_n + \frac{f'(x_n)}{f''(x_n)}$

- ◉ (1)
- ○ (2)
- ○ (3)
- ○ (4)

Question 6 — 10 pts

LDA and logistic regression explicitly try to separate the data into different classes as much as possible by constructing linear decision boundaries.

- ○ True
- ◉ False

Logistic Regression can also generate linear decision boundaries under certain circumstances, but its essence and purpose is to estimate the probability that data points belong to a certain category, and does not emphasize separating data points as much as possible.

Question 7 — 10 pts

Let $f(x) = \beta_0 + \beta x = 0$ be the hyperplane. The signed distance of any point x to the hyperplane can be described as which of the following. Select all that apply.

(1) $\frac{1}{\|\beta\|}(\beta^T x + \beta_0)$

(2) $\frac{f(x)}{\|f'(x)\|}$

(3) $(\beta^T x + \beta_0)$

(4) $\frac{1}{\|f'(x)\|}$

- [x] (1)
- [x] (2)
- [ ] (3)
- [ ] (4)

Question 8 — 10 pts

There are some problems with the perceptron algorithm. Select all the apply.

- [x] When the data are separable, there are many solutions, and which one is found depends on the starting values.
- [x] The "finite" number of steps can be very large. The smaller the gap, the longer the time to find it.
- [x] When the data are not separable, the algorithm will not converge, and cycles develop. The cycles can be long and therefore hard to detect.

Question 9 — 10 pts

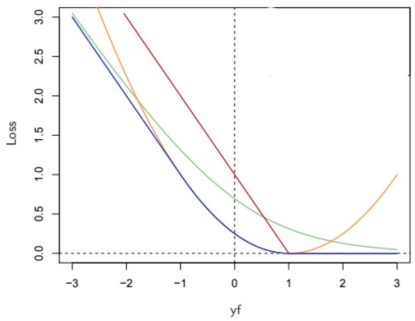
The optimal separating hyperplanes works only for the case when two classes are linearly separable.

- (●) True
- (✗) False

The optimal separating hyperplanes works only for the case when two classes are linearly separable.

Question 10 — 10 pts

Which curve in the figure below is the hinge loss? All are shown as a function of yf rather than f, in which f is the classifier function, because of the symmetry between the y = +1 and y = −1 case.



- ○ The blue one.
- ○ The orange one.
- ○ The green one.
- (●) The red one.