MLE

TI

$N(\mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$

Consider the following very simple model for stock pricing. The price at the end of each day is the price of the previous day multiplied by a fixed, but unknown, rate of return, α, with some noise, w. For a two-day period, we can observe the following sequence

$y_2 = \alpha y_1 + w_1$

$y_1 = \alpha y_0 + w_0$

where the noises w_0, w_1 are iid with the distribution $N(0, \sigma^2)$, $y_0 \sim N(0, \lambda)$ is independent of the noise sequence. σ^2 and λ are known, while α is unknown.

$y_0 \sim N(0, \lambda)$

$y_1 \sim N(\alpha y_0, \sigma^2) \leftarrow$

$y_2 \sim N(\alpha y_1, \sigma^2)$

$\therefore p(y_0, y_1, y_2 \mid \alpha)$

$= P(y_0|\alpha) P(y_1|\alpha) P(y_2|\alpha)$

T1. Find the MLE of the rate of return, α, given the observed price at the end of each day y_2, y_1, y_0. In other words, compute for the value of α that maximizes $p(y_2, y_1, y_0|\alpha)$

Hint: This is a Markov process, e.g. y_2 is independent of y_0 given y_1. In general, a process is Markov if $p(y_n|y_{n-1}, y_{n-2}, ...) = p(y_n|y_{n-1})$. In other words, the present is independent of the past $(y_{n-2}, y_{n-3}, ...)$, conditioned on the immediate past y_{n-1}. You may also find the steps of the proof for logistic regression we did in class useful.

$\therefore \log P(y_0, y_1, y_2|\alpha) = \log P(y_0|\alpha) + \log P(y_1|\alpha) + \log(y_2|\alpha)$

$\frac{d}{d\alpha} P(y_0, y_1, y_2|\alpha) = 0$

$\therefore \frac{d}{d\alpha}\left(\frac{y_0^2}{\lambda} + \frac{(y_1-\alpha y_0)^2}{\sigma^2} + \frac{(y_2-\alpha y_1)^2}{\sigma^2}\right) = 0$

$\downarrow$

$\alpha = \frac{y_0 y_1 + y_1 y_2}{y_1^2 + y_2^2}$

OT1. Consider the general case, where

$y_{n+1} = \alpha y_n + w_n, n = 0, 1, 2, ...$

Find the MLE given the observed price $y_{N+1}, y_N, ..., y_0$

OT1

$\rightarrow$ similiar to T1 ;

$P(y_0, y_1, y_2, \dots, y_{N+1} \mid \alpha) = P(y_0|\alpha) \cdots P(y_{N+1}|\alpha)$

$\frac{d}{d\alpha} = 0$; $0 = \frac{d}{d\alpha}\left(\frac{y_0^2}{\lambda} + \frac{(y_1-\alpha y_0)^2}{\sigma^2} + \frac{(y_2-\alpha y_1)^2}{\sigma^2} + \dots + \frac{(y_{N+1}-\alpha y_N)^2}{\sigma^2}\right)$

$\hookrightarrow \alpha = \frac{y_0 y_1 + y_1 y_2 + \dots + y_N y_{N+1}}{y_1^2 + y_2^2 + \dots + y_{N+1}^2}$

Simple Bayes Classifier

A student in Pattern Recognition course had finally built the ultimate classifier for cat emotions. He used one input features: the amount of food the cat ate that day, x (Being a good student he already normalized x to standard Normal). He proposed the following likelihood probabilities for class 1 (happy cat) and 2 (sad cat)

$P(x|w_1) = N(4, 2)$

$P(x|w_2) = N(0, 2)$

Figure 1: The sad cat and the happy cat used in training

T2. Plot the posteriors values of the two classes on the same axis. Using the likelihood ratio test, what is the decision boundary for this classifier? Assume equal prior probabilities.

T3. What happen to the decision boundary if the cat is happy with a prior of 0.75?

OT2

$P(x_0|w_1) = P(x_0|w_2)$

$\frac{1}{\sqrt{2\pi}\sigma} e^{\frac{(x_0-\mu_1)^2}{\sigma^2}} = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{(x_0-\mu_2)^2}{\sigma^2}}$

$\therefore (x_0-\mu_1)^2 = (x_0-\mu_2)^2$

$x_0^2 - 2x_0\mu_1 + \mu_1^2 = x_0^2 - 2x_0\mu_2 + \mu_2^2$

$x_0 = \frac{\mu_1+\mu_2}{2}$

$\leftarrow$ **OT2.** For the ordinary case of $P(x|w_1) = N(\mu_1, \sigma^2)$, $P(x|w_2) = N(\mu_2, \sigma^2)$, $p(w_1) = p(w_2) = 0.5$, prove that the decision boundary is at $x = \frac{\mu_1+\mu_2}{2}$

OT3. If the student changed his model to

$P(x|w_1) = N(4, 2) \rightarrow \frac{1}{\sqrt{2\pi}(2)} \exp(-(x-4)^2/2^2)$

$P(x|w_2) = N(0, 4) \rightarrow \frac{1}{\sqrt{2\pi}(4)} \exp(-(x-0)^2/4^2)$

$\Rightarrow e^{-\frac{(x-4)^2}{4}} = \frac{1}{2} e^{-\frac{x^2}{16}}$

Plot the posteriors values of the two classes on the same axis. What is the decision boundary for this classifier? Assume equal prior probabilities.

$-\frac{(x-4)^2}{4} = \log\left(\frac{1}{2}\right) - \frac{x^2}{16}$

$-(x-4)^2$