پروژه اول درس یادگیری ماشین

پریا پاسه‌ورز

شماره دانشجویی: 810101393

سوال سوم

در یک مسئله احتمالاتی رگرسیون خطی، رابطه میان داده‌های مسئله به شرح زیر است:

$$y_i = wx_i$$

$$\boldsymbol{\varepsilon}_i = N(0,1)$$

که در آن w پارامتر مدل و ε_i یک نویز گاوسی است.

با فرض i.i.d بودن داده‌ها، تابع log likelihood را تشکیل داده و نشان دهید بیشینه کردن این تابع معادل با کمینه کردن تابع زیر است:

$$arg\ max(w)\ \mathbf{lo}g\ P(\boldsymbol{D}|\boldsymbol{w}) = arg\ min(w) \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} (\boldsymbol{y_i} - \boldsymbol{wx_i})^2$$

می‌دانیم اگر $\mathrm{y}^{(1)}, \mathrm{y}^{(2)}, \ldots, \mathrm{y}^{(\mathrm{n})}$ مشاهدات ما باشند و بدانیم که i.i.d هستند، یک توزیع نرمال با میانگین $\mathrm{wx}^{(\mathrm{i})}$ و واریانس σ^2 خواهد داشت.

همچنین تعریف تابع likelihood به شرح زیر است:

$$L(\theta) = \prod_{i=1}^{n} f_{y_i}\left(y^{(i)}\right)$$

که در اینجا چون نرمال گاوسی است، پس:

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{\left(y^{(i)} - \mathrm{w}x^{(i)}\right)^2}{2\sigma^2}}$$

حال اگر از این عبارت، لگارتیم بگیریم، خواهیم داشت:

$$\log L\,(\theta) = \sum_{i=1}^{n} \log \frac{1}{\sqrt{2\pi\sigma}} - \frac{1}{2\sigma^2} \sum_{i=1}^{n} \left(y^{(i)} - \mathrm{w}x^{(i)}\right)^2$$

برای بیشینه کردن این عبارت، باید عبارت $\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y^{(i)} - \mathrm{w}x^{(i)}\right)^2$ را کمینه کنیم، چون عبارت $\sum_{i=1}^{n}\log\frac{1}{\sqrt{2\pi\sigma}}$ ثابت است. پس:

$$arg\ max(w)\ \mathbf{lo}g\ P(\boldsymbol{D}|\boldsymbol{w}) = arg\ min(w)\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}}(\boldsymbol{y_i} - \boldsymbol{wx_i})^{\mathbf{2}}$$