تکلیف چهارم درس شناسایی الگو

وحید ملکی
شماره دانشجویی: ۴۰۳۱۳۰۰۴

۷ نوامبر ۲۰۲۵

۱ سوال ۱۶

تابع چگالی احتمال زیر را در نظر بگیرید:

$$p_\theta(x) = 2\theta x e^{-\theta x^2}$$

که در آن، پارامتر θ مقداری حقیقی و مثبت دارد. فرض کنید N نمونه‌ی مستقل x_i از این توزیع در دسترس باشد. هدف آن است که تخمین پارامتر θ با استفاده از روش بیشینه‌ی درست‌نمایی (Maximum Likelihood Estimation) به‌دست آورده شود.

۱.۱ جواب سوال ۱۶

برای به‌دست آوردن برآورد بیشینه درست‌نمایی (MLE) پارامتر θ، مراحل زیر را طی می‌کنیم:

۱.۱.۱ ۰۱ تشکیل تابع درست‌نمایی (Likelihood Function)

تابع درست‌نمایی $L(\theta)$، برابر با حاصل‌ضرب چگالی احتمال برای تمام N نمونه‌ی مستقل (i.i.d.) است:

$$L(\theta) = p(x_1, \dots, x_N|\theta) = \prod_{i=1}^{N} p_\theta(x_i)$$

با جایگذاری تابع چگالی داده شده:

$$L(\theta) = \prod_{i=1}^{N} \left(2\theta x_i e^{-\theta x_i^2}\right)$$

می‌توانیم این حاصل‌ضرب را ساده‌تر کنیم:

$$L(\theta) = (2\theta)^N \left(\prod_{i=1}^{N} x_i\right) \left(\prod_{i=1}^{N} e^{-\theta x_i^2}\right)$$

$$L(\theta) = (2\theta)^N \left(\prod_{i=1}^{N} x_i\right) \exp\left(-\theta \sum_{i=1}^{N} x_i^2\right)$$

۲.۱.۰۱ ۲. تشکیل تابع لگاریتم-درست‌نمایی (Log-Likelihood Function)

برای سادگی محاسبات، به جای بیشینه‌سازی $L(\theta)$، لگاریتم آن یعنی $\mathcal{L}(\theta) = \ln(L(\theta))$ را بیشینه می‌کنیم. چون لگاریتم یک تابع یکنواخت صعودی است، نقطه‌ی بیشینه‌ی هر دو تابع یکسان خواهد بود.

$$\mathcal{L}(\theta) = \ln\left[(2\theta)^N\left(\prod_{i=1}^{N} x_i\right)\exp\left(-\theta\sum_{i=1}^{N} x_i^2\right)\right]$$

با استفاده از خواص لگاریتم ($\ln(abc) = \ln(a) + \ln(b) + \ln(c)$ و $\ln(e^x) = x$):

$$\mathcal{L}(\theta) = \ln((2\theta)^N) + \ln\left(\prod_{i=1}^{N} x_i\right) + \ln\left(\exp\left(-\theta\sum_{i=1}^{N} x_i^2\right)\right)$$

$$\mathcal{L}(\theta) = N\ln(2\theta) + \sum_{i=1}^{N}\ln(x_i) - \theta\sum_{i=1}^{N} x_i^2$$

باز هم ساده‌تر می‌کنیم ($\ln(2\theta) = \ln(2) + \ln(\theta)$):

$$\mathcal{L}(\theta) = N\ln(2) + N\ln(\theta) + \sum_{i=1}^{N}\ln(x_i) - \theta\sum_{i=1}^{N} x_i^2$$

۳.۱.۰۱ ۳. مشتق‌گیری و یافتن بیشینه

برای یافتن θ که $\mathcal{L}(\theta)$ را بیشینه می‌کند، از $\mathcal{L}(\theta)$ نسبت به θ مشتق می‌گیریم و آن را برابر صفر قرار می‌دهیم.

$$\frac{\partial\mathcal{L}}{\partial\theta} = \frac{\partial}{\partial\theta}\left[N\ln(2) + N\ln(\theta) + \sum_{i=1}^{N}\ln(x_i) - \theta\sum_{i=1}^{N} x_i^2\right]$$

مشتق $N\ln(2)$ (ثابت) برابر صفر است. مشتق $N\ln(\theta)$ برابر $\frac{N}{\theta}$ است. مشتق $\sum\ln(x_i)$ (ثابت نسبت به θ) برابر صفر است. مشتق $-\theta\sum x_i^2$ برابر $-\sum x_i^2$ است.

بنابراین، مشتق کلی برابر است با:

$$\frac{\partial\mathcal{L}}{\partial\theta} = \frac{N}{\theta} - \sum_{i=1}^{N} x_i^2$$

۴.۱.۰۱ ۴. حل معادله برای $\hat{\theta}_{MLE}$

اکنون مشتق را برابر صفر قرار می‌دهیم و معادله را برای θ (که آن را $\hat{\theta}$ می‌نامیم) حل می‌کنیم:

$$\frac{N}{\hat{\theta}} - \sum_{i=1}^{N} x_i^2 = 0$$

$$\frac{N}{\hat{\theta}} = \sum_{i=1}^{N} x_i^2$$

و در نهایت، برآورد بیشینه درست‌نمایی برای θ به‌دست می‌آید:

$$\hat{\theta}_{MLE} = \frac{N}{\sum_{i=1}^{N} x_i^2}$$

این نتیجه نشان می‌دهد که برآورد MLE برای θ، معکوس میانگینِ مجذورات نمونه‌ها ($1/\left(\frac{1}{N}\sum x_i^2\right)$) است.