

گروه پژوهشی ایک



دانشگاه صنعتی خواجه نصیرالدین طوسی
دانشکده مهندسی برق - گروه مهندسی کنترل

یادگیری ماشین آزمون میان ترم

نام و نام خانوادگی	علیرضا امیری
شماره دانشجویی	۴۰۲۰۲۴۱۴
تاریخ	خرداد ماه ۱۴۰۴

فهرست مطالب

۱	پاسخ سوال هماهنگ شده اول
۲	پاسخ سوال هماهنگ شده دوم
۳	پاسخ سوال ۱
۳	پاسخ سوال ۲
۳	پاسخ سوال ۳
۳	پاسخ سوال ۴

۱ پاسخ سوال هماهنگ شده اول

برای کلاس بندی دو دسته داده که هر یک دارای توزیع نرمال گوسی با میانگین μ ، کوواریانس σ و ویژگی های x هستند، از روابط احتمالی این توزیع ها و تعریف توزیع نرمال به شرح زیر استفاده می کنیم.

$$p(\mathbf{x} | C_i) = \mathcal{N}(\mathbf{x}; \boldsymbol{\mu}_i, \Sigma_i) = \frac{1}{(2\pi)^{d/2} |\Sigma_i|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_i)^\top \Sigma_i^{-1}(\mathbf{x} - \boldsymbol{\mu}_i)\right), \quad i = 1, 2. \quad (۱)$$

حال، برای دسته بندی این دو کلاس، طبق رابطه ی زیر با استفاده از لگاریتم و، likelihood با فرض $P(C_1) = P(C_2)$ خواهیم داشت:

$$\delta(\mathbf{x}) = \ln p(\mathbf{x} | C_1) - \ln p(\mathbf{x} | C_2) \underset{C_2}{\underset{C_1}{>}} 0. \quad (۲)$$

با جایگذاری فرم گوسی داده ها در این رابطه، رابطه ی تصمیم گیری به فرم کوادراتیک به صورت زیر به دست می آید:

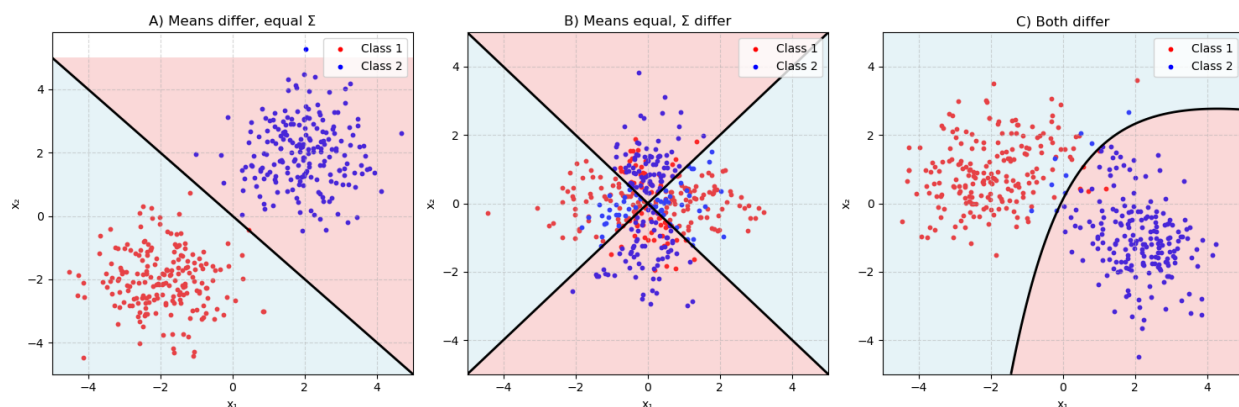
$$\begin{aligned} \delta(\mathbf{x}) &= \left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_1)^\top \Sigma_1^{-1}(\mathbf{x} - \boldsymbol{\mu}_1) - \frac{1}{2} \ln |\Sigma_1|\right] \\ &\quad - \left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_2)^\top \Sigma_2^{-1}(\mathbf{x} - \boldsymbol{\mu}_2) - \frac{1}{2} \ln |\Sigma_2|\right] \\ &= -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_1)^\top \Sigma_1^{-1}(\mathbf{x} - \boldsymbol{\mu}_1) + \frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_2)^\top \Sigma_2^{-1}(\mathbf{x} - \boldsymbol{\mu}_2) - \frac{1}{2} \ln \frac{|\Sigma_1|}{|\Sigma_2|}. \end{aligned} \quad (۳)$$

$$\mathbf{x}^\top (\Sigma_2^{-1} - \Sigma_1^{-1}) \mathbf{x} + 2(\boldsymbol{\mu}_1^\top \Sigma_1^{-1} - \boldsymbol{\mu}_2^\top \Sigma_2^{-1}) \mathbf{x} + c, c = \boldsymbol{\mu}_2^\top \Sigma_2^{-1} \boldsymbol{\mu}_2 - \boldsymbol{\mu}_1^\top \Sigma_1^{-1} \boldsymbol{\mu}_1 - \ln \frac{|\Sigma_2|}{|\Sigma_1|}. \quad (۴)$$

که با در نظر داشتن متغیر ویژگی های x و عبارت $\mathbf{x} \mathbf{x}^\top = x^2$ فرمت کوادراتیک این معادله مشخص میشود. حال، با تغییر مقادیر میانگین و کوواریانس، می توانیم سطوح مختلفی را به دست بیاوریم. به عنوان مثال، در صورتی که مقادیر کوواریانس دو کلاس با یکدیگر برابر باشد، آنگاه $\Sigma_1 = \Sigma_2 = \Sigma$ و در نتیجه $\Sigma_2^{-1} - \Sigma_1^{-1} = 0$ و در نهایت با حذف ضریب بخش درجه ۲ این معادله، صفحه ی تصمیم به فرمت خطی به صورت زیر به دست می آید:

$$(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^\top \Sigma^{-1} \mathbf{x} - \frac{1}{2}(\boldsymbol{\mu}_1^\top \Sigma^{-1} \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2^\top \Sigma^{-1} \boldsymbol{\mu}_2) = 0, \quad (۵)$$

به عنوان مثال، حالت های مختلف برای تغییرات میانگین و کوواریانس داده های کلاس ها در شکل ۱ رسم شده است.



شکل ۱: وضعیت های مختلف صفحه تصمیم با تغییر میانگین و کوواریانس

۲ پاسخ سوال هماهنگ شده دوم

در صورتی که تابع توزیع احتمال را در اختیار داشته باشیم، اولین گام برای تولید نمونه هایی که دارای توزیعی مطابق با این تابع باشند، محاسبه ی تابع توزیع احتمال تجمعی یا به اختصار CDF است. بنابراین، ابتدا با محاسبه ی انتگرال، این تابع را محاسبه می کنیم:

$$F(x) = \int_{-\infty}^x f(t)dt \quad (6)$$

آنگاه با استفاده از وارون تابع CDF و وارد کردن داده های تصادفی به آن، داده هایی با توزیع مورد نظر بسازیم. این فرایند به طریق زیر انجام خواهد شد:

$$x = F^{-1}(u) \Rightarrow x \sim f(x) \quad (7)$$

برای راحتی در این بخش، می توان داده های تصادفی را با استفاده از دستور $Uniform(0, 1)$ ایجاد کرد. در این حالت، با جایگذاری این داده ها به عنوان ورودی $F^{-1}(u)$ ، می توانیم داده های مورد نظر را به دست بیاوریم. با این حال، این روش تنها در حالتی برقرار است که بتوانیم از CDF به دست آمده وارون بگیریم که این در همه ی شرایط برقرار نیست. علاوه بر این، فرایند تولید اعداد با توزیع یکنواخت نیز در اینجا توضیح داده نشده است که در ادامه به آن خواهیم پرداخت. برای ساخت اعداد تصادفی، الگوریتم های مختلفی مانند $LinearCongruentialGenerator(LCG)$ و $MersenneTwister$ تولید می شوند. در این الگوریتم ها، اعداد رندوم در بازه ی 0 تا 1 از تقسیم عددی بزرگ بر F^{32} یا F^{64} به دست می آیند. در اینجا، عددی که بر این مقدار تقسیم می شود با استفاده از پارامترهای تصادفی در سخت افزار کامپیوتر ایجاد می شود.



۳ پاسخ سوال ۱

۴ پاسخ سوال ۲

۵ پاسخ سوال ۳

۶ پاسخ سوال ۴