



## تمرین اول

### توضیحات

این تمرین شامل ۵ مسئله است. در مسئله ۱، مباحثی از جمله امید ریاضی، واریانس و توضیح احتمال مرور می شوند. مسئله ۲ به رگرسیون خطی می پردازد. مسئله ۳ تخمین پارامتر با استفاده از روش تخمین بیشینه درست نمایی<sup>۱</sup> را مورد بررسی قرار می دهد. مسئله ۴ رگرسیون را با روش نیوتن ترکیب می کند و مسئله ۵ که امتیازی است، برای این طراحی شده که مروری بر جبر خطی داشته باشید.

### بارمبندی

- مسئله ۱: ۲۰ امتیاز
  - بخش ۱: ۵ امتیاز
  - بخش ۲: ۵ امتیاز
  - بخش ۳: ۱۰ امتیاز
- مسئله ۲: ۳۰ امتیاز
- مسئله ۳: ۲۰ امتیاز
- مسئله ۴: ۳۰ امتیاز
- مسئله ۵: ۱۰ امتیاز اضافه

### مسئله ۱. (مرور آمار و احتمال)

فرض کنید  $X$  و  $Y$  دو متغیر تصادفی باشند؛ به طوری که داشته باشیم:

$$\text{Range}(X) = \text{Range}(Y) = \{1, 2, \dots, n\}$$

. ثابت کنید گزاره های زیر برقرار هستند.

$$\mathbb{E}_Y[\mathbb{E}_X[X|Y]] = \mathbb{E}[X] \quad ۱.$$

$$\text{Var}(X) = \mathbb{E}_Y[\text{Var}[X|Y]] + \text{Var}[\mathbb{E}_Y[X|Y]] \quad ۲.$$

۳. فرض کنید نمادگذاری ما به این شکل باشد:  $\mathbb{P}[X = i] = \mathbb{P}[X_i]$  و  $\mathbb{P}[Y = i] = \mathbb{P}[Y_i]$ . نشان دهید گزاره زیر برقرار است:

$$\sum_{i=1}^n \mathbb{P}[X_i] \log(\mathbb{P}[X_i]) + \sum_{i=1}^n \mathbb{P}[Y_i] \log(\mathbb{P}[Y_i]) \leq \sum_{i=1}^n \sum_{j=1}^n \mathbb{P}[X = i, Y = j] \log(\mathbb{P}[X = i, Y = j])$$

## مسئله ۲. (رگرسیون خطی و چندجمله‌ای)

فرض کنید مدلی با ورودی  $X$  و خروجی  $Y$  داریم. به دو طریق مدل را آموزش می‌دهیم؛ در حالت اول  $Y$  را به شکل ترکیب خطی و در دیگری به شکل چند جمله‌ای درجه ۳ بر حسب  $X$  در نظر می‌گیریم.

۱. اگر معیار کمینه‌سازی، کمینه مربعات باشد و  $Y$  بر حسب  $X$  خطی به همراه نویز، یعنی داشته باشیم

$$Y = \beta_1 X + \beta_0 + \epsilon$$

میزان تابع هزینه روی داده‌های آموزشی را در این دو مدل مقایسه کنید.

۲. فرض کنید که  $Y$  الزاماً یک تابع خطی بر حسب  $X$  نباشد. در اینصورت نشان دهید که  $Y$  می‌تواند تابعی از  $X$  باشد که بعد از آموزش مدل با چندجمله‌ای درجه یک و سه، میزان هزینه روی داده‌های تست ناچیز باشد.

## مسئله ۳. (تخمین پارامتر)

تیم تدریس می‌خواهند بدانند که چند درصد از دانشجویان به درس یادگیری ماشین علاقه دارند. این کمیت را با  $\mu$  نشان می‌دهیم. برای تخمین  $\mu$ ، یک نظرسنجی ناشناس ایجاد شده که حاوی این سوال است: «دوره مقدماتی یادگیری ماشین را دوست دارید؟ بله یا خیر؟» هر دانشجو فقط یک بار می‌تواند به این سوال پاسخ دهد و ما فرض می‌کنیم که پاسخ‌ها i.i.d هستند.

۱. مقدار  $\mu$  را با روش MLE تخمین بزنید.

۲. فرض کنید  $\hat{\mu}$  یک کران بالا برای  $\mu$  باشد. برای برقراری عبارت زیر باید از چند دانشجو نظرسنجی شود؟

$$\mathbb{P}(|\hat{\mu} - \mu| > 0.1) < 0.05.$$

## مسئله ۴. (رگرسیون و بهینه‌سازی)

در این سوال قصد داریم با روش نیوتون، دو مسئله بهینه‌سازی را حل کنیم.

۱. در مسئله اول می خواهیم پارامترهای یک مدل رگرسیون خطی با ورودی‌های  $x^i \in \mathbb{R}^n$  و خروجی‌های  $y^i \in \mathbb{R}$  را بر حسب تابع هزینه کمینه مربعات بهینه کنیم. به عبارتی می‌خواهیم  $\theta \in \mathbb{R}^b$  را پیدا کنیم؛ به طوری که

$$J(\theta) = \sum_{i=1}^m (\theta^T x^i - y^i)^2$$

کمینه شود. مسئله بهینه‌سازی برای متغیر  $\theta$  را با روش نیوتون حل کنید.

۲. می‌خواهیم مسئله‌ی دسته‌بندی با ورودی‌های  $x^i \in \mathbb{R}^n$  و  $y^i \in \{0, 1\}$  را که تابع هزینه آن به شکل زیر تعریف می‌شود، مورد بررسی قرار دهیم.

$$J(\theta) = \frac{1}{m} \sum_{i=1}^m y_i \log h_\theta(x^i) + (1 - y_i) \log(1 - h_\theta(x^i))$$

که در آن  $h_\theta(x)$  به صورت زیر تعریف می‌شود:

$$h_\theta(x) = \frac{1}{e^{-(\theta^T x + \theta_0)} + 1}.$$

روند بهینه‌سازی تابع  $J(\theta)$  را با استفاده از روش نیوتون توضیح دهید.

(توجه: برای آشنایی با روش نیوتون، می‌توانید به این اسلاید رجوع کنید: لینک).

## مسئله‌ی ۵. (جبر خطی)

فرض کنید  $A$  و  $B$  دو ماتریس متقارن  $n \times n$  باشند و داشته باشیم:  $A, B \preceq I/2$ . ثابت کنید اگر

$$A(I - A) \preceq B(I - B)$$

برقرار باشد، آنگاه داریم:  $A \preceq B$ .

(یادآوری:  $C \preceq D$  اگر و تنها اگر  $D - C$  یک ماتریس مثبت نیمه‌معین باشد).