

به نام خدا



مبانی هوش محاسباتی

تمرین اول

(linear algebra – linear regression – logistic regression - SGD)

دکتر مائده سادات طاهایی

زمستان 1403

طراحان تمرین: سارا یونسی – جواد فرجی



- در صورت وجود هرگونه ابهام به طراح پیام دهید.
 - انجام تمرین ها تک نفره می باشد.
 - زبان برنامه نویسی پایتون است.
 - موارد ارسال شده به صورت آنلاین تحویل گرفته خواهند شد.
 - کل فایل محتوای ارسالی را داخل فایل زیپ قرار داده و نام آن را شماره دانشجویی خود قرار دهید.
 - تاریخ ریلیز تمرین: 29 اسفند ماه
 - تاریخ تحویل تمرین: 18 فروردین ماه
 - آیدی طراحان در تلگرام: @Javadyes1381 - saraunc
-
-

1. سوالات تئوری

سوال 1) آقای فرجی: کدام یک از تبدیل های زیر، تبدیل خطی هستند؟ چرا؟

$$T(x_1, x_2) = (1+x_1, x_2)$$

$$T(x_1, x_2) = (x_2, x_1)$$

$$T(x_1, x_2) = (x_1^2, x_2)$$

$$T(x_1, x_2) = (\sin x_1, x_2)$$

$$T(x_1, x_2) = (x_1 - x_2, 0)$$

سوال 2) آقای فرجی: ثابت کنید که دو ماتریس زیر، هم ارز سطری نیستند.

$$A = \begin{bmatrix} 2 & 0 & 0 \\ a & -1 & 0 \\ b & c & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 & 2 \\ -2 & 0 & -1 \\ 1 & 3 & 5 \end{bmatrix}$$

سوال 3) آقای فرجی: ماتریس $A_{3 \times 3}$ دارای مقادیر ویژه و بردارهای ویژه زیر است:

مقدار ویژه 1 با بردار ویژه:

$$\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

مقدار ویژه 2 با بردار ویژه:

$$\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}.$$

مقدار ویژه 3 با بردار ویژه:

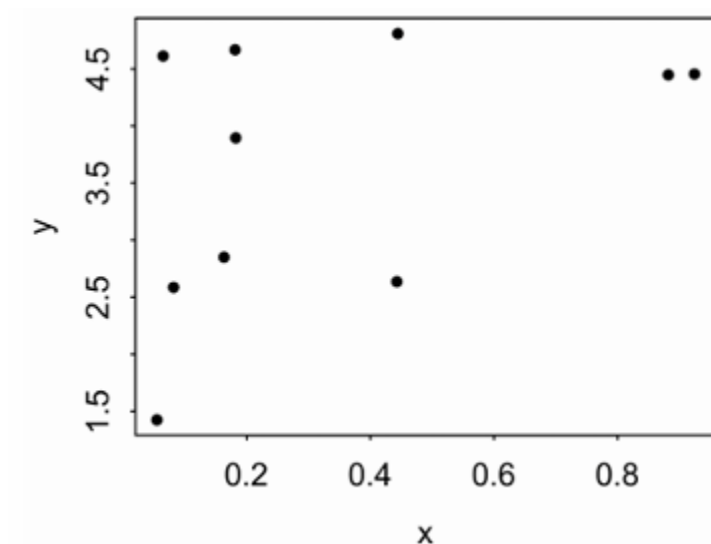
$$\begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix}.$$

حاصل ضرب $A \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}$ را پیدا کنید.

سوال 4) آقای فرجی: روش گرادیان کاهشی (GD ساده) را با روش SGD با $\text{batch-size} = 1$ مقایسه کنید و مزایا و معایب هر کدام را شرح دهید.

سوال 5) آقای فرجی: چرا MSE برای logistic regression مناسب نیست؟ توضیح دهید.

سوال 6) آقای فرجی: یک رگرسیون خطی با دیتاهای زیر fit شده است. کدام نقطه بیشترین تاثیر را بر شیب دارد؟ توضیح دهید.



سوال 7 (خانم یونسی): به سوالات جواب کوتاه بدهید و درستی یا نادرستی جملات خود را با توضیح رد یا قبول کنید.

- سوال: در روش SGD، (learning rate) چه تأثیری بر روی همگرایی الگوریتم دارد در صورتی که ثابت نگه داشته شود؟ چگونه می‌توانید این مسئله را مدیریت کنید؟
- سوال: در عملیات جبر خطی، اگر حاصل ضرب داخلی (dot product) دو بردار صفر شود، چه چیزی درباره رابطه این دو بردار می‌توان گفت؟ پاسخ دهید و یک مثال برای آن ارائه دهید.
- در رگرسیون خطی، زمانی که فرض می‌شود خطاها دارای واریانس ناهمسان (heteroscedasticity) هستند، پیش‌بینی‌های مدل همچنان معتبر و قابل اعتماد باقی می‌مانند. این جمله درست است یا نادرست؟ توضیح دهید.
- چرا در رگرسیون لجستیک به جای تابع خطای معمولی (mean squared error)، از Cross-Entropy Loss استفاده می‌شود؟ توضیح دهید چگونه این تابع با ماهیت طبقه‌بندی هماهنگ است.
- کاهش بیش از حد نرخ یادگیری در SGD همیشه همگرایی مدل را تضمین می‌کند. این جمله درست است یا نادرست؟ توضیح دهید.

سوال 8) خانم یونسی: در یک مسئله رگرسیون خطی داریم:

$$y = \underline{w}^T \underline{x}, \quad \underline{x} \in \mathbb{R}^L, \quad y \in \mathbb{R}$$

$$X = [\underline{x}_1, \underline{x}_2, \dots, \underline{x}_N], \quad \underline{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix}, \quad \underline{w} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_L \end{pmatrix}$$

الف) اگر رگرسیون را فقط بر روی ویژگی z انجام دهیم، نشان دهید که:

$$w_j = \frac{X_j \underline{y}}{X_j X_j^T}$$

که X_j سطر z ام ماتریس X است.

ب) فرض کنید ویژگی ها مستقل هستند (یعنی سطرهای ماتریس داده ها مستقل هستند) ثابت کنید که پارامترهای بهینه از آموزش رگرسیون بر روی همه ویژگی ها با پارامترهای بهینه حاصل از آموزش روی هر ویژگی به طور مستقل یکسان است.

پ) فرض کنید $y = w^T x + w_0$ و رگرسیون را فقط بر روی ویژگی z انجام می‌دهیم. w_0 و w_j را به دست آورید.



سوال 9(خانم یونسی): برای بررسی اینکه مهارت‌های واکنشی دانش‌آموزان طی یک سال چگونه بهبود یافته است، هشت دانش‌آموز در ابتدای سال و پایان سال در یک آزمون واکنش شرکت کردند. اینها نمرات آنها هستند:

نام	سارا	امیر	محمد	زهرا	احسان	فرید	فاطمه	تینا	نوید
امتحان اول	56	75	61	61	67	67	72	61	62
امتحان دوم	21	39	34	21	32	32	24	24	29

معادله خط رگرسیون را با توجه به داده‌های داده‌شده پیدا کنید.

سوال 10(خانم یونسی): مقدار مورد انتظار گرادیان در روش SGD چیست؟ با استفاده از فرمول‌ها توضیح دهید و نتیجه نهایی را تشریح کنید.

سوال 11(خانم یونسی): در چه شرایطی روش SGD به مینیمم محلی یا سراسری همگرا می‌شود؟ با استفاده از فرمول آن به این سوال پاسخ دهید.

سوال 12(تیم تدریس‌یاران): فرض کنید y توزیعی بر اساس رابطه زیر است، همچنین فرض می‌کنیم N تا مشاهده را جمع‌آوری کرده ایم $\{y_1, y_2, \dots, y_n\}$

$$f_y(y; a) = a y^{a-1}, y \in (0,1), a > 0$$

بر اساس maximum likelihood اثبات کنید که a از رابطه زیر به دست می‌آید.

$$\hat{a}_{ML} = - \left(\frac{1}{N} \sum_{j=1}^N \ln(y_i) \right)^{-1}$$

2. سوالات عملی

دو نوت بوک برای سوالات عملی در فایل zip ارسال شده قرار گرفته است که باید بررسی و انجام شود.

❖ نوتبوک HWP1 توسط خانم یونسی طراحی شده و نوتبوک HWP2 توسط آقای فرجی.

آنچه تحویل داده میشود:

1. کد اجرایی دو نوتبوک عملی

2. پاسخ های تمرین تئوری در یک فایل PDF

🚩 نکته بسیار مهم: 2 مورد بالا را zip کرده و نام آن را شماره دانشجویی خود گذاشته و

فقط در کوئرا ارسال کنید.