



بنام خدا  
دانشگاه تهران  
پردیس دانشکده‌های فنی  
دانشکده برق و کامپیوتر



درس سیستم‌های هوشمند

تمرین شماره 1

آبان ۱۴۰۱

## فهرست سوالات

- سوال 1 – رگرسیون لجیستیک باینری ..... 3
- الف- محاسبه‌ی گرادیان (تحلیلی) ..... 3
- ب- محاسبه‌ی گرادیان نزولی ..... 3
- ج- محاسبه‌ی دقت ..... 3
- د- گرادیان نزولی تصادفی ..... 3
- سوال 2 – بهینه سازی در توابع غیر محدب ..... 4
- الف- روش نیوتن (تحلیلی) ..... 4
- ب) روش نیوتن (شبیه سازی) ..... 4
- ج) روش فراابتکاری ..... 5
- سوال 3 – ماشین بردار پشتیبان ..... 6
- نکات تحویل: ..... 7

## سوال ۱ - رگرسیون لجیستیک باینری

دادگان [MNIST](#) را در نظر بگیرید که می‌خواهیم یک طبقه‌بندی باینری برای آن داشته باشیم؛ بطوریکه تشخیص دهد یک رقم، عدد ۲ است یا ۷. به این منظور، برای ارقام ۷ در دادگان فرض می‌کنیم  $Y = 1$  و برای ارقام ۲ در دادگان فرض می‌کنیم  $Y = -1$  باشد. با در نظر گرفتن یک مجموعه داده‌ی طبقه‌بندی با دو کلاس،  $\{(x_i, y_i)\}_{i=1}^n$ ، می‌توان از تابع زیر به‌عنوان تابع هزینه استفاده کرد:

$$J(w, b) = \frac{1}{n} \sum_{i=1}^n \log(1 + \exp(-y_i(b + x_i^T w)))$$

با در نظر گرفتن اینکه  $\mu_i(w, b) = \frac{1}{1 + \exp(-y_i(b + x_i^T w))}$

### الف- محاسبه‌ی گرادیان (تحلیلی)

بصورت دستی گرادیان‌های  $\nabla_w J(w, b)$  و  $\nabla_b J(w, b)$  را محاسبه کنید. پاسخ خود را بر حسب  $\mu_i(w, b)$  بنویسید.

### ب- محاسبه‌ی گرادیان نزولی

روش گرادیان نزولی را با نظر گرفتن مقادیر اولیه‌ی صفر پیاده‌سازی کنید. مقادیر متفاوت برای اندازه‌ی گام را امتحان کرده و برای هرکدام نمودار تغییرات  $J(w, b)$  را بر حسب تعداد تکرار الگوریتم رسم کنید. اینکار را بر روی هردو مجموعه‌ی آموزش و آزمون انجام دهید.

### ج- محاسبه‌ی دقت

برای هردو مجموعه‌ی آموزش و آزمون داده، با توجه به فرمول  $\text{sign}(b + x_i^T w)$  درصد درستی الگوریتم خود را محاسبه کنید.

### د- گرادیان نزولی تصادفی

بخش ب و ج را با پیاده‌سازی گرادیان نزولی تصادفی و در نظر گرفتن اندازه‌ی بسته‌ی برابر با ۱، تکرار کنید. سپس همین کارها را با اندازه‌ی دسته‌ی ۱۰۰ نیز تکرار کنید.

\*نکته ۱: شما در این قسمت از یک بردار وزن  $w$  و یک عدد بایاس  $b$ ، که در تابع هزینه به کار برده شده‌اند، استفاده می‌کنید.

\*نکته ۲: از مجموعه داده‌ی آزمون برای به‌روز رسانی پارامترها استفاده نکنید.

## سوال ۲ - بهینه سازی در توابع غیر محدب

تابع زیر را در نظر بگیرید:

$$f(x) = 2x_1^2 + 2x_2^2 - 17x_2 \cos(0.2\pi x_1) - x_1 x_2$$

### الف- روش نیوتن (تحلیلی)

با شروع از نقطه  $(0, 0)$ ، جهت گرادینان نزولی<sup>۱</sup> را پیدا کنید. سپس با استفاده از روش نیوتن، مقدار جدید نقطه شروع را در یکبار به‌روزرسانی به صورت تحلیلی به دست آورید.

### ب) روش نیوتن (شبیه سازی)

با استفاده از روش نیوتن و با شروع از نقاط  $(3 و 1)$ ، نقطه کمینه این تابع را با استفاده از شبیه سازی کامپیوتری به دست آورید. همچنین به ازای تمامی نقاط  $0 < x_2 < 10, -5 < x_1 < 5$  با  $stepsize = 0.5$ ، مقدار کمینه را به دست آورید و آنها را بر اساس فاصله‌شان تا  $-36.4$ ، به سه دسته نزدیک<sup>۲</sup>، دور<sup>۳</sup> و دورتر<sup>۴</sup> تقسیم بنده کنید و نمودار این نقاط را همانند شکل ۱-۲ رسم کنید و نتایج خود را به صورت مختصر توضیح دهید. (نمودار ۱-۲ برای رفع ابهام است و لزومی ندارد که نمودار به دست آمده توسط شما شبیه به این نمودار باشد).

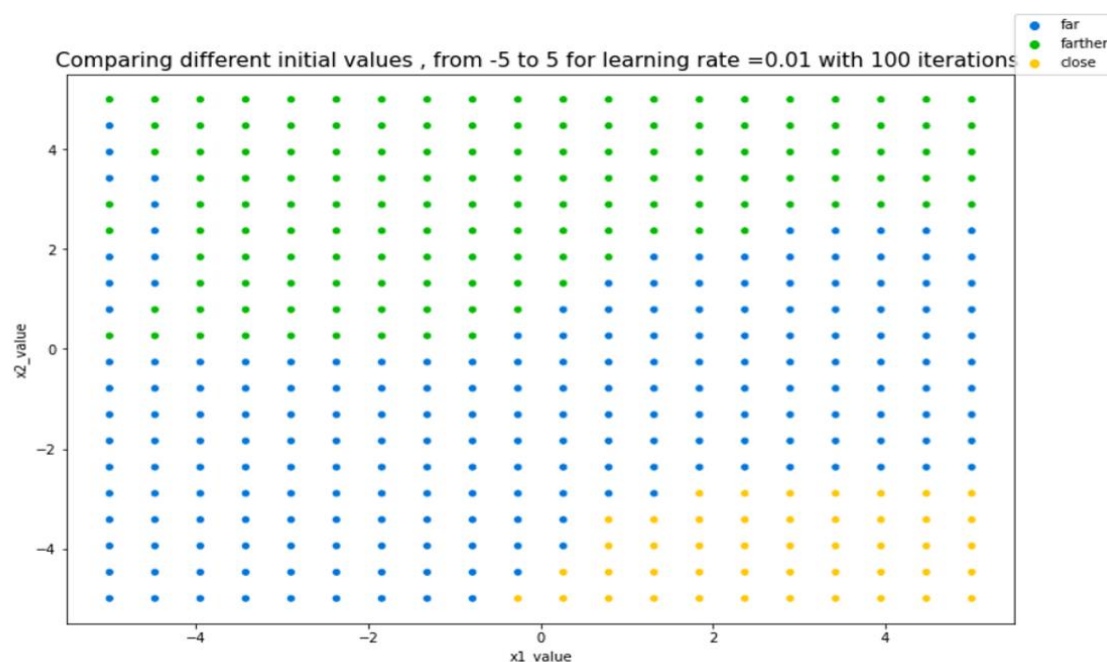
---

<sup>1</sup> Gradient Descent

<sup>2</sup> Close

<sup>3</sup> Far

<sup>4</sup> Farther



شکل ۲-۱: یک نمونه از نمودار خواسته شده برای بخش ب سوال ۲

### ج) روش فراابتکاری

با استفاده از یکی از روشهای فراابتکاری (تبرید شبیه سازی<sup>۵</sup> یا ژنتیک<sup>۶</sup>)، مقدار کمینه محلی تابع را با شرط (۱-۲) به دست آورید. در صورت استفاده از هر کدام از روشها، پارامترها را با استفاده از شرایط مسئله در نظر بگیرید و علت انتخاب خود را بیان کنید.

$$(x_1^0, x_2^0) \in [-15, 15]; \quad (1-2)$$

نکات:

1. در این تابع، مقدار بهینه تابع در حدود ۳۶.۴- است.
2. در قسمت ج، امکان به دام افتادن در کمینه محلی وجود دارد. لذا با تغییر پارامترهای مسئله، همگرایی یا عدم همگرایی برنامه خود را بررسی کنید.
3. در قسمت ج، میزان فاصله از مقدار بهینه به صورت حدودی قابل قبول است.

### سوال ۳ - ماشین بردار پشتیبان

#### الف) تحلیلی

۱ - فرمولاسیون زیر را برای طبقه بند ماشین بردار پشتیبان<sup>۷</sup> در نظر بگیرید:

$$\text{Min} \frac{1}{2} \|w^2\| + C \sum_{i=1}^n \xi_i \xi_i = \max(0, 1 - y_i(W^T x_i + b))$$

اگر برای یک نمونه مقدار  $\xi_i$  برابر صفر شود، در مورد آن چه نتیجه ای میتوان گرفت؟

۲ - فرض کنید سه بردار پشتیبان به صورت  $x_3 = (2,2), x_2 = (2,1), x_1 = (1,0)$  داریم. دو بردار اول نمونه کلاس مثبت و بردار سوم نمونه کلاس منفی است. اگر  $X$  و  $Y$  مختصات اول و دوم در این فضا را نشان دهند (یعنی هر بردار بصورت  $(X, Y)$  بیان شده است) مرز تصمیم بهینه کدام مورد است. (لطفاً به طور کامل استدلال خود را بیان کنید)

- 1)  $y = x + \frac{1}{2}$
- 2)  $y = -x + \frac{1}{2}$
- 3)  $y = x - \frac{1}{2}$
- 4)  $y = -x - \frac{1}{2}$

#### ب) پیاده سازی

در این بخش قصد داریم با استفاده از ماشین بردار پشتیبان، طبقه بندی را برای مجموعه داده گان iris انجام دهیم. ابتدا با استفاده از قطعه کد زیر داده گان را لود کنید:

```
from sklearn import datasets
iris = datasets.load_iris()
data = iris.data[:, :2] #data
label = iris.target #label
```

\*دقت کنید که در قطعه کد بالا فقط از دو ستون اول که مربوط به دو ویژگی طول کاسبرگ و عرض کاسبرگ است، استفاده کردیم.

حال با پیاده سازی ماشین بردار پشتیبان و با استفاده از روش یک در قیاس با بقیه<sup>۸</sup> ناحیه طبقه های مختلف را مشخص و رسم کنید.

علاوه بر رسم نواحی مختلف، دقت داده آموزش، ماتریس آشفتگی<sup>۹</sup> و ماتریس اطمینان<sup>۱۰</sup> را نیز در گزارش خود بیاورید و روابط میان آنها را تفسیر کنید.

\*راهنمایی : میتوانید از کتابخانه sklearn برای این منظور استفاده کنید. به خاطر داشته باشید که ماشین بردار پشتیبان مورد نظر ما خطی است.

### نکات تحویل:

- مهلت تحویل این تمرین 21 آبان میباشد.
- انجام این تمرین به صورت یک نفره است.
- برای انجام این تمرین تنها مجاز به استفاده از زبان برنامه نویسی پایتون هستید.
- در صورت وجود تقلب نمره تمامی افراد شرکت کننده در آن ۰ لحاظ میشود.
- لطفا پاسخ تمرین خود را (به همراه کد/گزارش سوال کامپیوتری) به صورت زیر در صفحه درس آپلود نمایید:

HW [HW number] \_ [Last name] \_ [Student number].zip

- در صورت وجود هر گونه ابهام یا مشکل میتوانید از طریق ایمیل با مسئولان حل تمرین در تماس باشید:

مسئول تمرین سوال ۱: شیوا شاکری ( [shiva.shakeri@ut.ac.ir](mailto:shiva.shakeri@ut.ac.ir) )

مسئول تمرین سوال ۲: روزبه نهاوندی ( [roozbeh.nahavandi@ut.ac.ir](mailto:roozbeh.nahavandi@ut.ac.ir) )

مسئول تمرین سوال ۳: محمدرضا تیموریان فرد ( [mr.teymoorian@ut.ac.ir](mailto:mr.teymoorian@ut.ac.ir) )

One vs Rest<sup>8</sup>

Confusion Matrix<sup>9</sup>

Confidence Matrix<sup>10</sup>