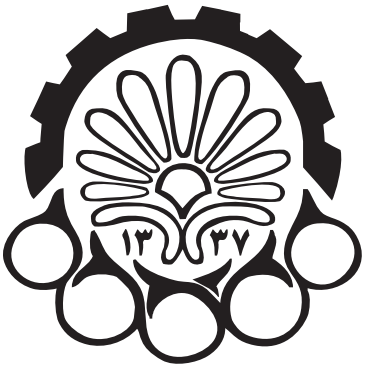


دانشگاه صنعتی امیرکبیر

(پلی‌تکنیک تهران)

دانشکده ریاضی و علوم کامپیوتر

عنوان

یادگیری ماشین

استاد درس

دکتر فاطمه شاکری

تمرین اول

رگرسیون خطی و گرادیان کاهشی

مهر ۱۴۰۳

۱ مقدمه

در این تمرین، هدف شما پیاده‌سازی رگرسیون خطی با استفاده از گرادیان کاهشی است. اما علاوه بر آن، باید از منظم‌سازی L2 (همچنین به‌عنوان رگرسیون Ridge شناخته می‌شود) نیز استفاده کنید تا مدل از بیش‌برازش جلوگیری کند.

رگرسیون خطی مدلی ساده برای پیش‌بینی متغیر وابسته y با استفاده از متغیر مستقل X است. در این مدل فرض می‌شود که رابطه خطی بین متغیرها برقرار است، اما برای جلوگیری از پیچیدگی بیش‌ازحد مدل، منظم‌سازی به این فرمول اضافه می‌شود.

۲ فرمول‌بندی مسئله

رگرسیون خطی در حالت ساده دوبعدی به شکل زیر تعریف می‌شود:

$$h_\theta(x) = \theta_0 + \theta_1 x$$

که در آن:

- $h_\theta(x)$ پیش‌بینی مدل است.
- θ_0 و θ_1 پارامترهای مدل (یا وزن‌ها) هستند.
- x متغیر مستقل (ویژگی) است.

در این تمرین، ما همچنین از L2 Norm Regularization استفاده می‌کنیم. تابع هزینه همراه با منظم‌سازی به شکل زیر تغییر می‌کند:

$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m}\left(h_\theta(x_i) - y_i\right)^2 + \lambda\sum_{j=1}^{n}\theta_j^2$$

که در آن:

- λ مقدار منظم‌سازی است. هرچه این مقدار بیشتر باشد، منظم‌سازی قوی‌تر عمل می‌کند و مدل ساده‌تر می‌شود.

۳ الگوریتم گرادیان کاهشی با منظم‌سازی $L2$

الگوریتم گرادیان کاهشی به شکل زیر تغییر می‌کند تا منظم‌سازی L2 را لحاظ کند:

$$\theta_j := \theta_j - \alpha\left(\frac{1}{m}\sum_{i=1}^{m}\left(h_\theta(x_i) - y_i\right)x_i + \frac{\lambda}{m}\theta_j\right)$$

۴ مراحل تمرین

- **گام ۱:** یک کلاس پایتون با نام LinearRegressor پیاده‌سازی کنید که شامل منظم‌سازی L2 باشد. متدهای زیر باید در این کلاس وجود داشته باشند:
  - train: این متد الگوریتم گرادیان کاهشی را با منظم‌سازی L2 پیاده‌سازی می‌کند.
  - plot: این متد داده‌های ورودی و خط رگرسیون به‌دست‌آمده را رسم می‌کند.
  - mse: این متد مقدار MSE را محاسبه و برمی‌گرداند.
  - R Squared: این متد مقدار R Squared را محاسبه و برمی‌گرداند.
- **گام ۲:** داده‌های ورودی را تولید کنید. ۲۰۰ نمونه برای هر یک از دو مجموعه داده تصادفی X_1 و X_2 را با استفاده از توزیع تصادفی تولید کنید که به‌ترتیب رابطه‌های زیر را دنبال می‌کنند:

$$y_1 = 4 + 3X_1 + \text{نویز}$$

$$y_2 = 10 - 2X_2 + \text{نویز}$$

برای گام‌های ادامه، مجموعه داده را به دو بخش آموزش (۷۰ درصد) و آزمایش (۳۰ درصد) با حفظ توزیع (stratify) تقسیم کنید. از مجموعه‌ی آموزش برای تمرین مدل استفاده کرده و سپس بر روی مجموعه‌ی آزمایش مدل خود را بسنجید. موارد خواسته شده مانند دو خطای MSE و R Squared برای هر دو بخش بدست آورده و تحلیل کنید.

- **گام ۳:** داده‌های دو مجموعه را در یک مجموعه ادغام کرده و یک مدل رگرسیون خطی با منظم‌سازی L2 بر روی آن آموزش دهید. سپس وزن‌ها و MSE و R Squared نهایی را چاپ کنید و نمودار داده‌ها و خط رگرسیون را رسم کنید.
- **گام ۴:** دو مدل جداگانه برای هر مجموعه داده (X_1 و X_2) آموزش داده و نمودارها و خطاهای گفته شده را مانند گام سوم برای آن‌ها را رسم کنید. مقدار λ را تغییر دهید تا اثر آن بر مدل را مشاهده کنید.
- **گام ۵:** نتایج به‌دست‌آمده از مدل‌های جداگانه و مدل ادغام‌شده را با یکدیگر مقایسه کرده و نمودارهای MSE و R Squared آن‌ها را در یک نمودار واحد رسم کنید.
- **گام ۶:** آیا امکان دارید در این مجموعه داده‌ی ادغام شده، شاهد Overfitting باشیم؟ چرا؟

۵ آزمایش با پارامترهای مختلف

نرخ یادگیری (Learning Rate)

نرخ یادگیری (α) یکی از پارامترهای کلیدی در الگوریتم گرادیان کاهشی است. اگر این مقدار بیش از حد کوچک باشد، الگوریتم خیلی کند به جواب می‌رسد و اگر این مقدار خیلی بزرگ باشد، ممکن است الگوریتم نوسانات

زیادی داشته باشد یا حتی به نتیجه‌ای نرسد. در این تمرین باید با نرخ‌های یادگیری مختلف آزمایش کنید تا تاثیر آن‌ها بر دقت مدل و همگرایی را مشاهده کنید.

وزن منظم‌سازی (Regularization Weight)

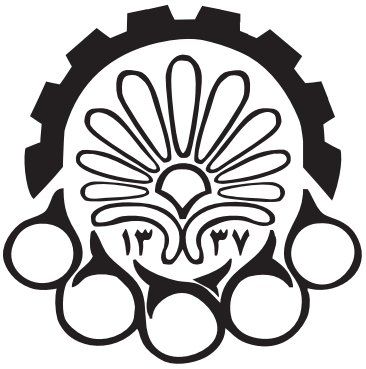
پارامتر λ که وزن منظم‌سازی L2 است نیز تاثیر بسزایی در عملکرد مدل دارد. اگر مقدار این پارامتر بیش از حد کوچک باشد، مدل ممکن است به‌طور کامل از داده‌ها بیش‌برازش کند (Overfitting). از طرف دیگر، اگر این مقدار خیلی بزرگ باشد، مدل بیش از حد ساده می‌شود و نتایج ضعیفی ارائه می‌دهد.

۶ گام‌های اضافی

- **گام ۸:** با نرخ‌های یادگیری متفاوت مانند 0.001، 0.01، 0.1 و 1 الگوریتم خود را اجرا کرده و نمودار MSE و R Squared مربوط به هر حالت را رسم کنید. سپس مشاهده کنید که آیا نرخ یادگیری بالا (مثل 1) باعث نوسان و خراب شدن نتیجه می‌شود یا خیر.
- **گام ۹:** تاثیر وزن منظم‌سازی λ را بررسی کنید. مقادیر مختلفی برای λ مانند 0.01، 0.1، 1 و 10 را امتحان کنید. مشاهده کنید که مقدار خیلی کوچک (مثل 0.01) باعث بیش‌برازش می‌شود و مقدار خیلی بزرگ (مثل 10) باعث می‌شود مدل بیش از حد ساده شود و دقت کاهش یابد.
- **گام ۱۰:** نمودارها و مقادیر به‌دست‌آمده را تحلیل کرده و توضیح دهید که چرا برخی از پارامترها باعث بهبود عملکرد مدل می‌شوند و چرا برخی دیگر باعث خراب شدن نتیجه می‌شوند.

۷ نکات مهم

- نرخ یادگیری α باید به‌دقت تنظیم شود تا الگوریتم به‌خوبی همگرا شود.
- وزن منظم‌سازی λ نیز باید تنظیم شود تا مدل به‌طور مناسب از داده‌ها یادگیری کند و از پیچیدگی یا سادگی بیش‌ازحد جلوگیری کند.



Amirkabir University of Technology

(Tehran Polytechnic)

Department Of Mathematics and Computer Science

Course

Machine Learning

Instructor

Dr. Fatemeh Shakeri

Exercise One

Linear Regression and Gradient Descent

Sep 2024