

سوال پنجم

در این سوال قرار است انواع رگرسیون را پیاده سازی کنید. داده های این سوال در فایل های data.mat و data.npz قرار دارند. این داده ها براساس رابطه ی زیر تولید شده اند :

$$y = 4 x_2^2 x_1 + 2 x_2^2 + 3 x_1 + 1$$

بنابراین نمونه ها به صورت $[x_1, x_2]$ و خروجی متناظرشان y است.

در فایل data.mat و data.npz شش آرایه ی یک بعدی، $y, x_1, x_2, y_{test}, x_{1,test}, x_{2,test}$ قرار دارند. یک مثال از دسترسی به آرایه ی y فایل در زبان پایتون، چنین است :

```
#python
a = np.load('data.npz')
print(a['y'])
```

آرایه ی y_{test} ، خروجی صحیح روی داده های $(x_{1,test}, x_{2,test})$ است و شما باید خروجی کد خود را با این آرایه مقایسه کنید.

تابع هزینه ی ما تابع SSE (Sum of squared errors) خواهد بود.

رگرسیون خطی را با استفاده از Gradient decent و Stochastic gradient decent پیاده سازی کرده و نتایج را روی داده های تست با هم مقایسه کنید.

امتیازی: برای داده های تست، نتیجه ی کد خودتان و مقادیر صحیح خروجی که در آرایه ی y_{test} است را بر

حسب $x_{1,test}$ و $x_{2,test}$ به صورت نمودار سه بعدی نمایش دهید.

مقدار تابع خطا روی داده های آموزش و داده های تست را گزارش نمایید.