



دانشکده‌گان علوم  
دانشکده ریاضی، آمار و علوم کامپیوتر

مهلت تحویل: ۱۸ اردیبهشت

تمرین عملی دوم

مسئله ۱. تابع زیر را در نظر بگیرید.

$$f(x) = x_1^4 + 3x_2^2 + x_1$$

(a) توابعی برای محاسبه گرادیان و ماتریس hessian برای  $f(x)$  بنویسید.

(b) با شروع از نقطه  $x = (5, 4)^T$  تابع  $f(x)$  را به روش گرادیان کاهشی و انتخاب طول گام به روش back-tracking بهینه کنید.

(c) قسمت b را به ازای  $\alpha$  و  $\beta$  های متفاوت انجام دهید و تاثیر این دو پارامتر را بر روی عملکرد مدل تحلیل کنید.

(d) با شروع از نقطه  $x = (5, 4)^T$  تابع  $f(x)$  را با استفاده از روش pure newton بهینه کنید.

(e) این بار از روش نیوتن و انتخاب طول گام با روشی متفاوت، تابع را بهینه کنید.

(f) نتایج قسمت های b و d را مقایسه کنید و مزایا و معایب هر یک از این دو روش را بیان کنید. آیا روشی وجود دارد که بتوانیم از مزایای هر دو روش به طور همزمان استفاده کنیم؟

نکات تکمیلی:

- مقدار تابع را در حین فرایند بهینه سازی برای هر یک از روش ها رسم کنید.
- تعداد iteration هر یک از روش ها را تا رسیدن به جواب بهینه گزارش کنید.
- stopping criterion های متفاوتی وجود دارد، با توجه به مساله مناسب ترین را انتخاب کنید.

مسئله ۲. می‌خواهیم یک چندجمله‌ای درجه دو به داده‌های زیر برازش دهیم.

$$x = [-2, -1/5, -1, 0, 1, 2]$$

$$y = [8/3, 5, 2/0, 0/3, 2/5, 8]$$

فرض کنید فرم چندجمله‌ای به شکل  $f(x, \theta) = \theta_0 + \theta_1 x + \theta_2 x^2$  باشد. می‌خواهیم  $\theta$  بهینه‌ای را پیدا کنیم که تابع loss یعنی میانگین مجموع مربعات خطا (MSE) کمینه شود.

$$L(\theta) = \sum_{i=1}^n \frac{1}{n} (f(x_i, \theta) - y_i)^2$$

( a ) تابع خطا  $L(\theta)$  را بر حسب  $\theta_0$ ،  $\theta_1$  و  $\theta_2$  بنویسید.

( b ) گرادیان و ماتریس hessian را برای  $L(\theta)$  محاسبه کنید.

( c ) با استفاده از gradient descent و نقطه اولیه  $\theta = (0, 0, 0)$  مقدار بهینه  $\theta$  را پیدا کنید.

( d ) از همان نقطه اولیه شروع کرده و این بار با روش نیوتن مقدار بهینه  $\theta$  را بیابید.

( e ) نتایج دو مدل را با هم مقایسه کنید.

( f ) با نقطه شروع متفاوتی هر دو روش را تکرار کنید و تاثیر نقطه شروع را بر روی عملکرد هر دو مدل بررسی کنید.

نکات تکمیلی:

- مقدار تابع خطا را در حین فرایند بهینه‌سازی برای هر یک از روش‌ها رسم کنید.
- تعداد iteration هر یک از روش‌ها را تا رسیدن به جواب بهینه گزارش کنید.
- مقدار بهینه پارامتر  $\theta$  را برای هر دو روش گزارش کنید و نمودار چندجمله‌ای برازش داده شده را در کنار نقاط داده شده در یک نمودار رسم کنید.
- برای انتخاب طول گام در هر دو روش می‌توانید از طول گام ثابت یک استفاده کنید.