

توضیح کوتاه کد:

در ابتدا سه نوع تابع خطایی که فرمول ریاضی آنها در سؤال نوشته شده است، پیاده سازی شده و سپس برای هر کدام از این توابع مسئله ای در کتابخانه `cvxpy` تعریف شده است و `solver` آنها صدا زده می شود و پاسخ بهینه آنها بدست می آید.

خروجی برنامه به صورت زیر است:

Part a Results:

Case: X_std, y_std

norm-2 of theta-theta_gen = 0.03794704271621642

Case: X_std, y_outliers

norm-2 of theta-theta_gen = 6.667494364700117

Case: X_outliers, y_outliers

norm-2 of theta-theta_gen = 4.456298954200446

Case: X_outliers, y_std

norm-2 of theta-theta_gen = 2.955629794968814

Part b results:

Case: X_std, y_std

norm-2 of theta-theta_gen = 0.04879219084010686

Case: X_std, y_outliers

norm-2 of theta-theta_gen = 0.05668813624629632

Case: X_outliers, y_outliers

norm-2 of theta-theta_gen = 1.0837618574397203

Case: X_outliers, y_std

norm-2 of theta-theta_gen = 2.7887408108696627

Part c results:

Case: X_std, y_std

norm-2 of theta-theta_gen = 0.04879219084010669

Case: X_std, y_outliers

norm-2 of theta-theta_gen = 0.056688136246311305

Case: X_outliers, y_outliers

norm-2 of theta-theta_gen = 0.0271113525187579

Case: X_outliers, y_std

norm-2 of theta-theta_gen = 0.016333160443689876

پاسخ بخش (d):

خطا از نوع **quadratic** چون نسبت به داده‌های پرت حساس هست پس خطایش برای این شکل از داده‌ها زیاد خواهد بود که نتایج هم نشان می‌دهند چنین شده‌است. در مورد خطای **normilized** به این دلیل که **normalization** داریم برای خطای هر سمپل پس انتظار داریم که خطا به مراتب از دو خطای دیگر کمتر شود که همین اتفاق نیز افتاده‌است. به طور کلی خطای از نوع **norm** اول از خطای از نوع **norm** دوم کمتر است اما زمانی که مدل تخمینی به خوبی بر داده فیت شود، این خطای **norm** دوم خواهد بود که کمتر خواهد شد به دلیل وجود توان ۲ که عدد کوچکتر از ۱ با این توان ۲ کمتر نیز خواهد شد که نتایج بالا این درستی این فرضیه را نشان می‌دهد.