



絶対誤差

「データとモデルの差」を測るなら、次のような絶対値を用いた絶対誤差を使ってもよさそうだと考えられる

$$J^{abs}(\boldsymbol{w}) = \sum_{n=1}^N |y_n - f(\boldsymbol{x}_n)|$$

しかし、絶対値は微分不可能な部分を生むため、二乗誤差の方がよく使われる



ref: 線形代数の半歩先
p127~128
ref: なっとく! 機械学
習 p66~67

誤差関数を比較する観点

目的によっては、二乗誤差以外の誤差関数を選ぶ必要もある

ref: 線形代数の半歩先
p128~129

最小値からのズレを許しやすいか

原点付近の様子を見ると、二次関数は最小値からずれても関数の値が急激には増加しない

絶対値関数の方が、少しでもズレが生じると、関数の値が急に増える

つまり、二乗誤差より絶対誤差の方が、最小値からのズレを許しづらくなる

ズレの大きさを無視するか

二次関数は、大きくズれると二乗の大きさに効く

絶対値関数は、原点以外は直線（比例）であるため、大きなズレはどれも同じような扱いをする



平均を用いた誤差関数

実際には、「合計」ではなく「平均」を求める、**平均絶対誤差 (MAE: mean absolute error)** や **平均二乗誤差 (MSE: mean square error)** がよく使われる

ref: なっとく！機械学習 p67~68

合計は総数に影響される

たとえば、次の 2 つのデータセットを使って、誤差またはモデルを比較しようとしているとする

- 10 個のデータ点が含まれるデータセット
- 100 万個のデータ点が含まれるデータセット

誤差関数がデータ点ごとに 1 つの数字を「合計」したものだとなれば、合計する数字の量が多い 100 万個のデータ点を含むデータセットの方が、誤差（合計値）がはるかに大きくなってしまう

これらのデータセットを正しく比較したい場合は、「平均値」を求める必要がある

さらに平方根を使う

よく使われている誤差関数として、**二乗平均平方根誤差 (RMSE: root mean square error)** と呼ばれるものもある

RMSE は、**MSE** の平方根として定義されるもので、次のような目的で使われる

- 問題の単位を合わせる目的
- モデルが予測を行うときにどれくらい誤差が生じるのかをよく理解する目的

たとえば、住宅の価格を予測しようとしていて、価格の正解値と予測値の単位がドルだとする

二乗平均の単位は二乗ドルになってしまうため、平方根をとることで、正しい単位が得られる

これにより、住宅 1 軒あたりのモデルの大まかな誤差をドル単位でより正確に知ることができる