機械学習 課題2

高林秀
三宅研究室 博士前期課程１年
V-CampusID：23vr008n

July 23, 2023

Abstract

本稿は本年度必修授業の機械学習の第２回レポートの答案用紙である。
本稿は、第５回授業～第１１回授業までの範囲を対象とし、各回で課された課題に対する解答を記載する。
答案の問題番号は各章のタイトルに記載している。

1 第５回授業： 5/16 問題２

1.1 問題文

次のモデルは、どのようなデータを学習する際に役に立ちそうでしょうか？具体的な応用先のタスク・データを想定してみてください。パラメータが三つある事の意味・役割も解説してください。またこのモデルの学習は、勾配降下法で十分でしょうか？このモデルの学習の流れを説明しなさい

$$\hat{y}(x) = a_0 + a_1 \sin x + a_2 \cos x$$

1.2 解答

このモデルは、$\sin, \cos$ といった周期関数が含まれているため、周期性をもつデータの学習に適していると考えられる。

具体的な応用可能なタスクとして、「時刻毎の気温の変化」や「月間の株価の変動」など周期性や季節性をもつデータの学習が想定される。

また各パラメータ a_0, a_1, a_2 の役割は、モデルの式から次のように読み取ることができる。

- a_0：バイアス項. モデルの予測に対して追加される定数。データの偏りによるモデルの予測値を補正する為に使用される。モデルの複雑さを抑制し過学習を予防する役割も持つ。
- a_1： 周期関数 $\sin x$ に対する係数。$\sin x$ の振幅を調整し、モデルに対する $\sin x$ の影響度を制御する。データの周期性の大きさを調整する役割を持つ。

- a_2：周期関数 $\cos x$ に対する係数。$\cos x$ の振幅を調整し、モデルに対する $\cos x$ の影響度を制御する。データの周期性の大きさを調整する役割を持つ。ただし、$\sin x$ の周期とは異なる位相で周期性を調整する。

上記 3 つのパラメータが独立して調整されることで、このモデルはバイアスの調整、$\sin x$ と $\cos x$ の周期性と影響度の調整を独立して学習することができ、よりデータに適合したモデルとすることができる。

また、このモデルの学習は、勾配降下法で十分であると考えられ、その際の学習の流れは次のようになる。

1. パラメータ a_0, a_1, a_2 を適当な値で初期化する。
2. データ全体に対して、モデルの予測値 $\hat{y}$ と実際の値 y の誤差を計算する。
   - ここでは、誤差として二乗誤差:MSE などを用いる。
3. 誤差関数の勾配を計算し、勾配降下法を用いてパラメータを更新する。
   - 各パラメータを、その勾配に学習率をかけた値だけ引き算して更新する。
4. 一定の回数（エポック数）ないしは、モデルの予測値 $\hat{y}$ と実際の値 y の誤差が十分小さくなるまで、2. と 3. を繰り返す。

この流れを前提に、このモデルの誤差関数のグラフを描画してみると、次のようになる。パラメータ $a_0 = 1, a_1 = 2$ とし入力 $x = \frac{\pi}{4}$ をした場合の,a_2 と真の値 y との誤差:Error をプロットしたものである。

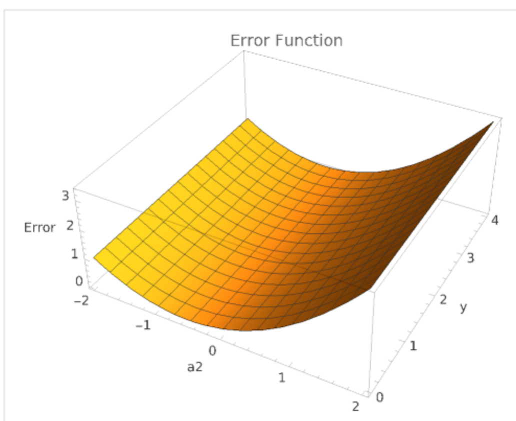

Figure 1: 誤差関数のグラフ

このように、誤差関数は a_2 に対して凸関数となっているため、勾配降下法によって最適解を求めることができると考えられる。

2　第 6 回授業：問題１

2.1　問題文

今回説明した手法を使って、重回帰の練習をしましょう。まず自分で何らかのデータセットを用意します（web から回帰向きのデータを検索）。そして、そのデータの前処理、重回帰モデルの学習、モデルの改善、ベストなモデルの選択、テスト性能の評価までを一通り行い、得られた結果について考察を加えてください。LASSO だけではなく、Ridge も試してみましょう。モデルの学習結果から、データに関する何らかの仮説が立てられたり洞察が得られるとベストです

2.2　解答

本課題に対する解答は、以下の GoogleColab に記載している。ここでは、本課題に対して解答したソースコードのリンクのみを記載する。

- ファイル (code.ipynb)： `https://colab.research.google.com/drive/1Zx4wYZWsR5Ge_Phy0yRFJbOT9IO5puMr?usp=sharing`
- 使用データセット (SSDSE-C-2023.csv)： `https://drive.google.com/file/d/1-LLPqbwy09z1qwkAQnVIXE2dTMomZiyc/view?usp=drive_link`

3　第 7 回授業：問題２

3.1　問題文

クロスエントロピーの大事な性質。

1. クロスエントロピーは常にゼロ以上であることを示せ。
2. クロスエントロピーがゼロとなるのは、各データ点 x_n に対して、ロジスティック回帰モデルの予測 $\hat{y}(x)$ がどの様な条件を満たすときか。ただし、ラベル y_n は 0,1 の値しかとらないことに注意。

3.2　解答

3.2.1　1. の解答

まず、クロスエントロピーの式を以下で示す。

$$E(th\hat{e}ta) = -\sum_{n=1}^{N} \{y_n \log \hat{y}(x_n) + (1 - y_n) \log(1 - \hat{y}(x_n))\}$$

この式は、各データ点におけるクロスエントロピー誤差の和を示したもので、y_n は n 番目のデータの正解ラベル,$\hat{y}(x_n)$ は n 番目のデータ点に対するモデルの予測確率を表す。

　このとき、$y_n = 0$ の場合、$E(\hat{\theta}) = -\log(1 - \hat{y}(x_n))$ に、$y_n = 1$ の場合、$E(\hat{\theta}) = -\log \hat{y}(x_n)$ となる。

いずれの場合も、$E(\hat{\theta}) = -\log A$ の形をしており、$-\log A$ は A の値が $0 \to 1$ の範囲で変化すると、常に 0 以上の値をとる関数である。

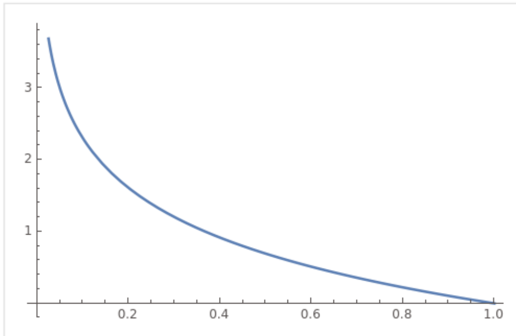

Figure 2: $-\log A$ のグラフ

よって、各項は常に 0 以上であり、全体の和も 0 以上となることから、クロスエントロピーは常に 0 以上であることが示される。

3.2.2　2. の解答

クロスエントロピーがゼロになるということは、各データ点に対するモデルの予測 $\hat{y}(x_n)$ と正解ラベル y_n との誤差が 0 であることを意味する。したがって、$\hat{y}(x_n) = y_n$ であるということになる。

具体的には、

$$\hat{y}(x_n) = y_n = 0 \text{ のとき}$$
$$\hat{y}(x_n) = y_n = 1 \text{ のとき}$$

ということになる。これをクロスエントロピー誤差関数の式に代入すると、$y_n = 0$ の場合は

$$\begin{aligned}
E(\hat{\theta}) &= -\sum_{n=1}^{N} \{y_n \log \hat{y}(x_n) + (1 - y_n) \log(1 - \hat{y}(x_n))\} \\
&= -\sum_{n=1}^{N} \{0 \log \hat{y}(x_n) + (1 - 0) \log(1 - \hat{y}(x_n))\} \\
&= -\sum_{n=1}^{N} \{\log(1 - \hat{y}(x_n))\} \\
&= -\sum_{n=1}^{N} \{\log(1 - 0)\} \\
&= 0
\end{aligned}$$

となり、$y_n = 1$ の場合は

$$
\begin{aligned}
E(\hat{\theta}) &= -\sum_{n=1}^{N} \{y_n \log \hat{y}(x_n) + (1 - y_n) \log(1 - \hat{y}(x_n))\} \\
&= -\sum_{n=1}^{N} \{1 \log \hat{y}(x_n) + (1 - 1) \log(1 - \hat{y}(x_n))\} \\
&= -\sum_{n=1}^{N} \{\log \hat{y}(x_n)\} \\
&= -\sum_{n=1}^{N} \{\log 1\} \\
&= 0
\end{aligned}
$$

となる。

ゆえに、クロスエントロピーが 0 であるには、モデルの予測と正解ラベルが一致することが必要である。

4 第 10 回授業：6/20 問題１

4.1 問題文

多クラス分類と、識別関数の図形的な仕組みについて、以下の問に答えよ。

1. p.18 の絵において、$f_k =$ 一定値 の直線はどの様に図示されるか？傾きは適当で良いが、$k = 0, 1, 2$ に対して（共通の一定値に対して）図示せよ。
2. この図の半直線 $f_0 = f_1$ に沿ったまま無限遠方へ移動したとする。この時３つのクラスに対する確率の値は幾つになるか？
    - 真面目に計算して証明する必要はない。簡単な理由があればいい。

4.2 解答

4.2.1 1. の解答

5 第 11 回授業：6/27 問題１

5.1 問題文

二次計画問題と有効な制約について、以下の問に答えよ。

1. p.51 の問題を解きなさい。
2. p.60 p.63 で得られた双対問題を解いて、① と同じ解が求まることえお確認しなさい。

5.2 解答

5.2.1 1. の解答

5.2.2 2. の解答