「ベイズ推論による機械学習」輪読会

第 1 章: 機械学習とベイズ学習

正好 奏斗 (@cosnomi)

この章で扱うこと

- そもそも機械学習って何？
  - 具体的にはどんな種類・タスクがあるの？
- 確率の基本
  - コインとか赤玉白玉とか
- ベイズ推論
  - ベイズの基本で重要な考え方
  - 事前分布・事後分布って何？

機械学習とは

- データの特徴を抽出する
- その特徴に基づいて未知の現象に対する予測を行う
- 特徴
  - 例えば「比例関係」「一次関数」など
  - 比例定数や切片を既知のデータから推測
    - このような値をパラメータという
  - 未知のデータ(x)が与えられたら y を予測できる
  - 実際はもっと複雑な関数

ここからは具体的な機械学習について見ていきます

- １つ１つを詳しく理解する必要は無いと思います
- 全体を俯瞰するのがこの章の目的

線形回帰とは

- **回帰(regression)**とは M 次元の入力$\boldsymbol{x} = (x_1, ..., x_M) \in \mathbb{R}^M$ から$y \in \mathbb{R}$を求めること
- $M = 1$なら$y = wx + b$とか $(y, w, x, b \in \mathbb{R})$

線形回帰の式

- 一般化すると、

$$y = w_1x_1 + w_2x_2 + \cdots + w_Mx_M$$

- という式の$\boldsymbol{w} = \left(w_1, w_2, \ldots, w_M\right) \in \mathbb{R}^M$を既知のデータから求めたい
- ベクトルで書くと、

$$y = \boldsymbol{w}^T\boldsymbol{x}$$

- 内積を転置で表していることに注意
  - $\boldsymbol{a} \cdot \boldsymbol{b} = \boldsymbol{a}^T\boldsymbol{b}$

- しかし既知のデータもすべて綺麗にこの式に従うわけではない
- 各サンプル(既知のデータのこと)のn番目について次のような式を考えられる

$$y_n = \boldsymbol{w}^T \boldsymbol{x}_n + \epsilon_n$$

- ϵがn番目のサンプルの誤差あるいはノイズを表している
  - 多くの場合、ϵは正規分布に従う(が、その話は後で)
- 切片は？2 次以上は考えられないの？
  - $\boldsymbol{x} = (1, x_1, {x_1}^2, x_2, {x_2}^2, \ldots)$などとしてやれば良い
  - これはグラフだと曲線になるけど「線形回帰」という

回帰から分類へ

- 分類も機械学習においてよく出てくるタスク
  - $y \in \{0, 1\}$のように離散的な値を取る
  - 例えば 0 が陰性、1 が陽性みたいな
  - 多クラス(0: 陰性, 1: 軽症, 2: 重症みたいな)は後で
- 線形回帰では連続値yを予測した
  - 離散的な値はどう予測する？

連続値を確信度とみなす

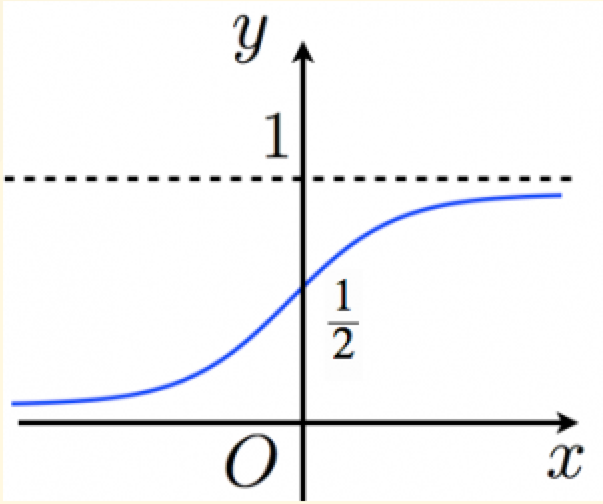
- 回帰で求めた値(μ_nとしますう)は実数全体を取りうるのでややこしい
- $f : \mathbb{R} \to (0, 1)$みたいな関数によって$\mu_n$を$(0, 1)$に移せれば簡単
- μを「クラス 1 に分類される確率」として考えられる
  - $f(\mu) \geq 0.5$なら$y_n = 1$、そうでないなら$y_n = 0$とすればよい
- どんな関数が良いだろう…

Sigmoid function

- 有名な関数: シグモイド関数(**sigmoid function**)

$$\mathrm{Sig}(a) = \frac{1}{1+e^{-a}}$$

- $\sigma(a)$とか表記されることもある



ここまでを数式でまとめる

- 今やりたいのは
  - 入力$\boldsymbol{x}_n = (x_1, x_2, \ldots x_M)$が与えられて
  - 出力$y_n \in \{0, 1\}$を求めたい
- 線形回帰

$$\mu_n = \boldsymbol{w}^T \boldsymbol{x_n}$$

- 画像
- $\mathrm{Sig}(\mu_n)$をとって、0.5 以上なら 1、そうでないなら 0