

第 8 章

回帰

8.1 目的

数値特徴を入力して数値を出力する関数をデータから学習する回帰について学びます。線形回帰では、正則化の効果について学びます。回帰木では性能を高める方法について学びます。

8.2 Weka

実践演習 8-1

<http://www.liquidasset.com/winedata.html> に掲載されているアッシェンフェルターのワイン方程式を導く元となったデータ（Variables 覧の 1980 年までのデータ）から arff 形式のファイルを作成し、Weka の LinerRegression を用いて回帰式を導け。

実践演習 8-2

例題 8.1 に従い、Weka の LinerRegression を用いて cpu データで回帰を行え。その際、正則化係数を変化させ、学習された線形関数の重みを観察せよ。

実践演習 8-3

Weka の LinerRegression を用いて cpu.with.vendor データで回帰を行い、学習結果を解釈せよ。

実践演習 8-4

例題 8.2 に従い、Weka の REPTree を用いて cpu データで回帰を行え。その際、木の大きさを変えて相関係数の変化を観察せよ。

実践演習 8-5

例題 8.3 に従い、Weka の M5P を用いて cpu データで回帰を行え。その際、木の大きさを変えて相関係数の変化を観察せよ。

8.3 sklearn

実践演習 8-6

sklearn の `LinerRegression` を用いて boston データで回帰を行え。その際、正則化係数を変化させ、学習された線形関数の重みを観察せよ。

1. boston データの読み込み（パターン行列 X 、ターゲット y ）
2. `LinearRegression` の学習
3. 学習結果オブジェクトの `coef_` の値を表示して係数を確認
4. `sklearn.model_selection.cross_val_score` を使って交差確認を行い、決定係数を表示
5. L2 正則化の `Ridge` で同様の手順
6. L1 正則化の `Lasso` で同様の手順

実践演習 8-7

sklearn の `DecisionTreeRegressor` を用いて boston データで回帰木を作成せよ。その際、木の大きさを変えて、誤差の変化を観察せよ。