【線形回帰モデル】

教師データがある場合に使用できる回帰モデル。回帰モデルとは入力から出力を予想数モデルで、出力が連続値となる。説明変数ｘ、目的変数ｙとした場合、ｙ＝ax+bとなる比例関係の式を使用する。ここで、a,bの2つパラメータがある。説明変数ｘが1つであるとは限らずｘがｍ個あった場合、パラメータはｍ+1個用意する必要がある。また、目的変数ｙも1つである場合だけでなく一般的にはｎ個存在する場合を考える。数式的な表現ではｎ行（ｍ+1）列のパラメータ行列と入力ｘのベクトルとの線形結合（内積）からｙの予測値を表現する。パラメータｗの行列の決定方法は目的変数の教師データとモデル出力の平均二乗誤差を最小にするように決定される。すなわち、ｗを変数としたときの平均二乗誤差の最小値問題に帰着する。

・ハンズオン

実応用上ではscikit-learnなどのpythonライブラリを用いることで、簡単にモデル実装が可能。Pandasとnumpyによってデータを扱う。今回はscikit-learnのbostonデータ（ボストン市の住宅価格データセット）に対し、線形回帰モデルを適用した。モデルを適用する前にはデータセットから、説明変数と目的変数がどういう意味合いのデータのか、データに外れ値や欠陥値がないかの確認を行う必要がある。Pandasによってカラム名を指定しデータフレームを作成すると、よりデータを理解しやすくなる。データフレームからどの変数を説明変数とするのかを選定し抽出する。目的変数も同様に抽出する。線形回帰モデルは

from sklearn.linear\_model import LinearRegression

でモジュールをインポートできる。.fitで学習（モデルパラメータ計算）.predictで予測ができる。予測値が妥当な値となっているかを確認するべきである。

【非線形回帰モデル】

非線形回帰モデルは教師データがある場合に使用できる回帰モデル。線形回帰では説明変数は単なる1次式ｘだったが、ここが任意の関数φ（ｘ）に置き換わると非線形回帰モデルとなる。このパラメータに線形結合される説明変数の関数（基底関数）は2次以上の多項式関数であったり、三角関数であったり様々なパターンが考えられる。特に、基底関数にはガウス関数が用いられることが多い。ここで注目すべきは線形回帰モデルと同様、非線形回帰モデルのパラメータｗは定数で線形なままであるということである。

パラメータｗの決定方法は線形回帰モデルと同様、目的変数の教師データとモデル出力の平均二乗誤差が最小となるようなｗを決定する。

学習データに対して小さな誤差が得られない（未学習）あるいは、学習データに対して過剰にフィッティングしている（過学習）場合には注意が必要である。いずれも未知のデータに対して誤差が最小になる予測ができていない。（汎化性能が確保できていない）

過学習に対しては、①学習データ数を増やす、②不要な既定関数を削除➂正則化法を利用

の3つの方法がある。正則化法とはモデルの複雑さ大きくなるほど、評価値が大きくなるような罰則項を導入して評価する方法である。罰則項がＬ２ノルムの場合Ridge推定量、Ｌ１ノルムの場合Lasso推定量という。

【ロジスティックス回帰モデル】

分類問題で使用するモデルの一種であり、目的変数が例えば0，1の場合に適用できる。

ロジスティックス回帰は識別的アプローチで分類を行う。識別的アプローチとはある確率（説明）変数に対して、その分類になる確率を直接モデル化する方法のことである。予測値は実数全体をとるが、出力は0～1である必要があるため、シグモイド関数を使い0～1の範囲に収める計算処理が行われる。すなわち、線形回帰モデルにシグモイド関数を適用したモデルがロジスティクス回帰モデルである。

ここで、重みパラメータｗをどう決定するかについて説明する。ベルヌーイ分布を用いて、データから分布を推定する方法がある。また、データを固定しパラメータを変化させ、最もフィットする関数を予想するのが最尤推定であり、尤度関数を最大化するｗ（パラメータ）を求める。

計算上の問題として確率は0～1しかとらないため、それを複数回かけるととても小さな値になり桁落ちする可能性がある。したがって計算上の工夫として、対数関数で評価する。

シグモイド関数が含まれているので解析的に解を出すことが難しいため、勾配降下法でｗを更新し、負の対数尤度の最小値を求める方法がよく用いられる。

微分項の計算は連鎖律を利用する

データすべてに対して、一括で勾配降下法を計算すると、計算コストが膨大になるので1つづつランダムに選んでパラメータを更新する確率的勾配降下法（ミニバッチ法の一種）

【主成分分析】

次元を削減する手法でありデータの散らばりが最大になるように（多くの散らばりをなるべく残せるように）次元を決定し、データを圧縮する。

【K-近傍法】

教師なし学習のモデルの一種でクラスタリング、分類タスクを行う機械学習方法である。最近傍のデータをｋ個とってきてそれらが最も多く所属するクラスに識別する。一つの着目する説明変数を中心に円を広げそれにｋ個属するかどうかを調べる。

【ｋ-means】

教師なし学習のモデルの一種でクラスタリングタスクを行う機械学習方法である。各クラスタの中心の初期値を設定し、各データ点と設定した初期値との距離を計算して最も距離が近いクラスタを決定する。そののちクラスタ内の全データ平均、重心ベクトルをとり初期値を更新する。以後同じ手順を繰り返す。初期値があまりにも近いとうまくクラスタリングできない場合があるので、初期値の設定に注意が必要である。