

Pythonではじめる 機械学習

4章 データの表現と特徴量エンジニアリング
3.3 相互作用と多項式

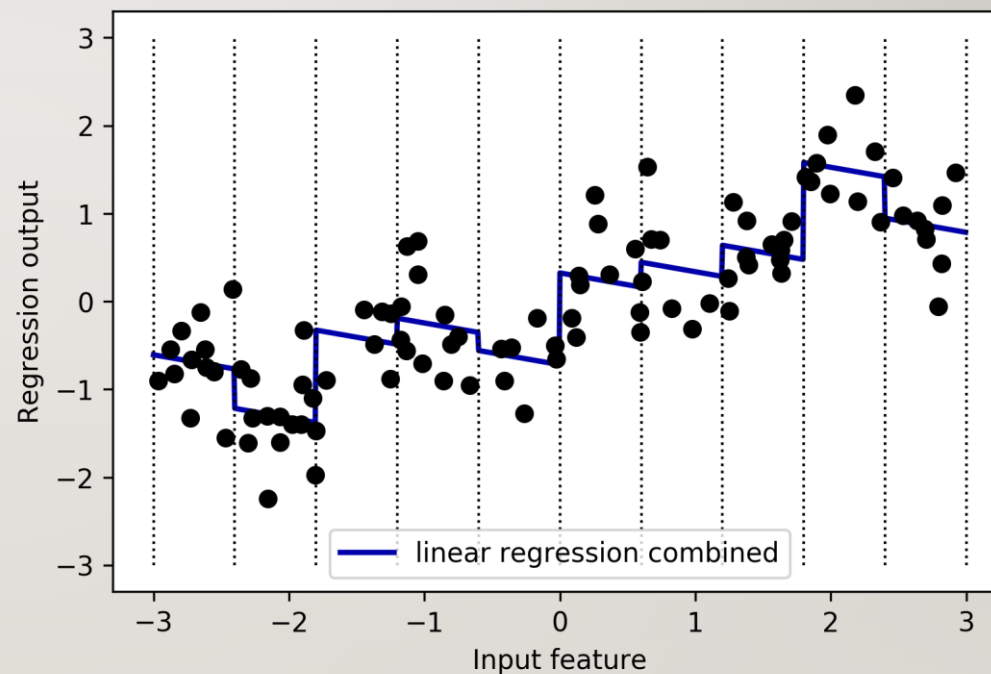
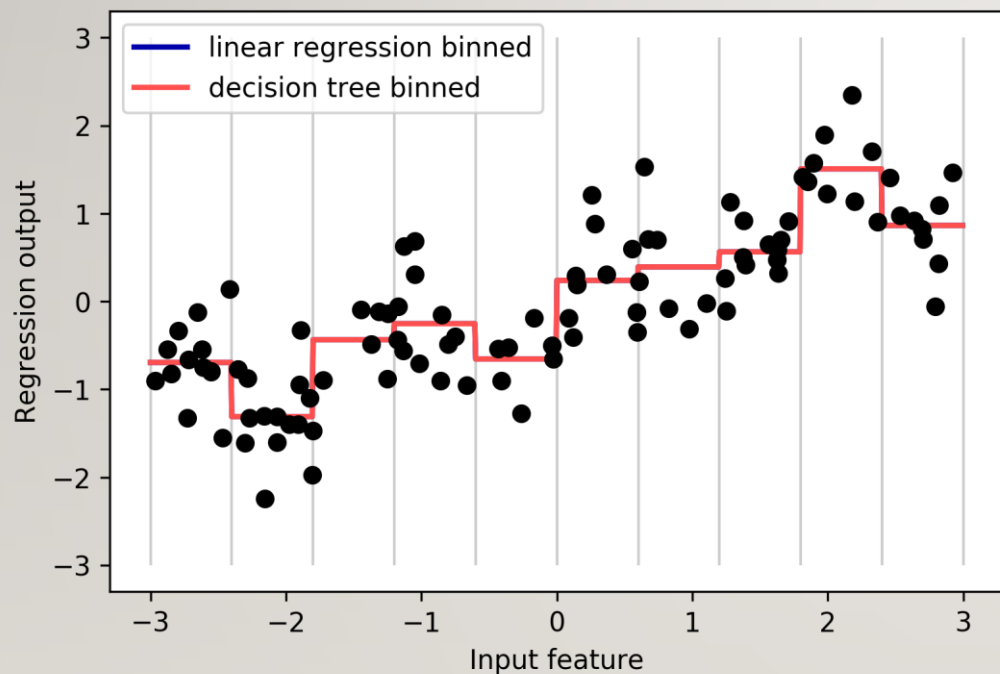
4.3 交互作用と多項式

特徴量表現をより豊かにする方法

- 交互作用特徴量と多項式特徴量を加える(特に線形モデルに有効)

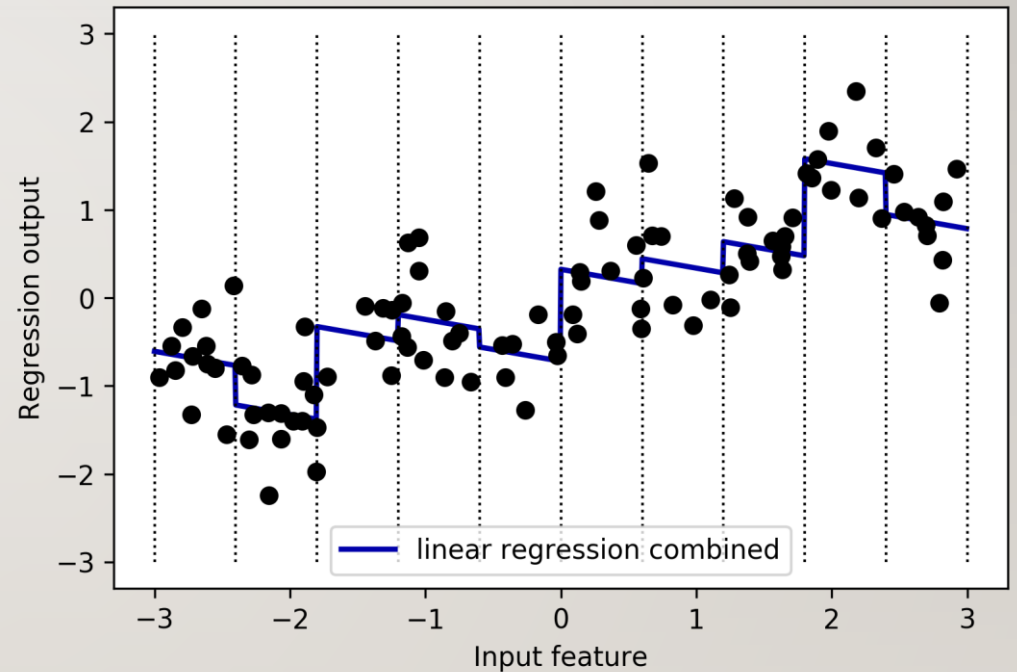
4.3 交互作用と多項式

特徴量を加え直し、ビンに傾きを与えたもの



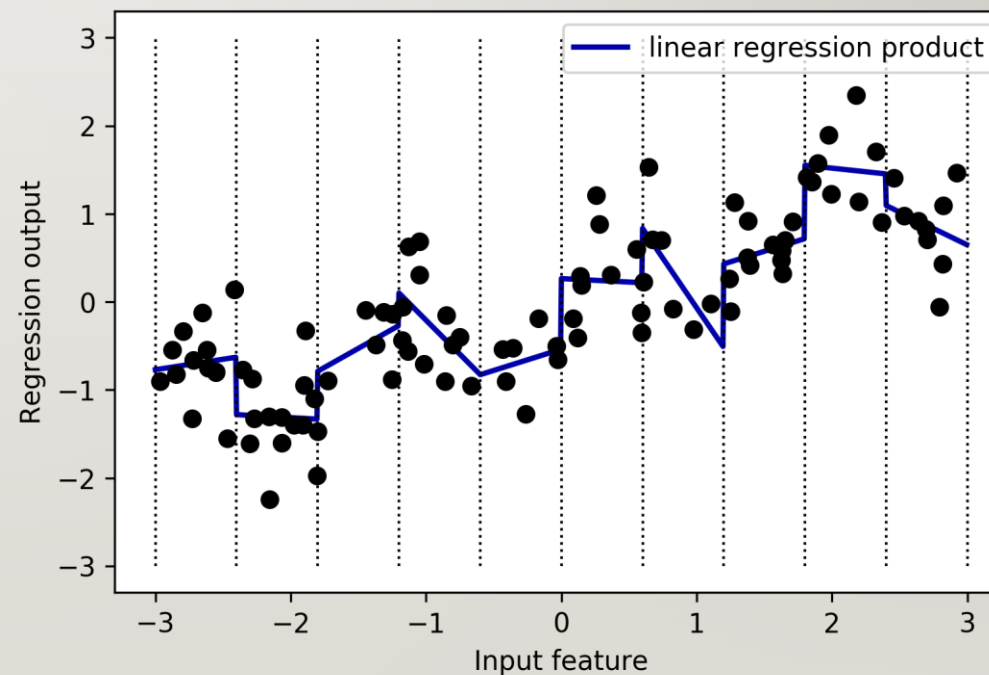
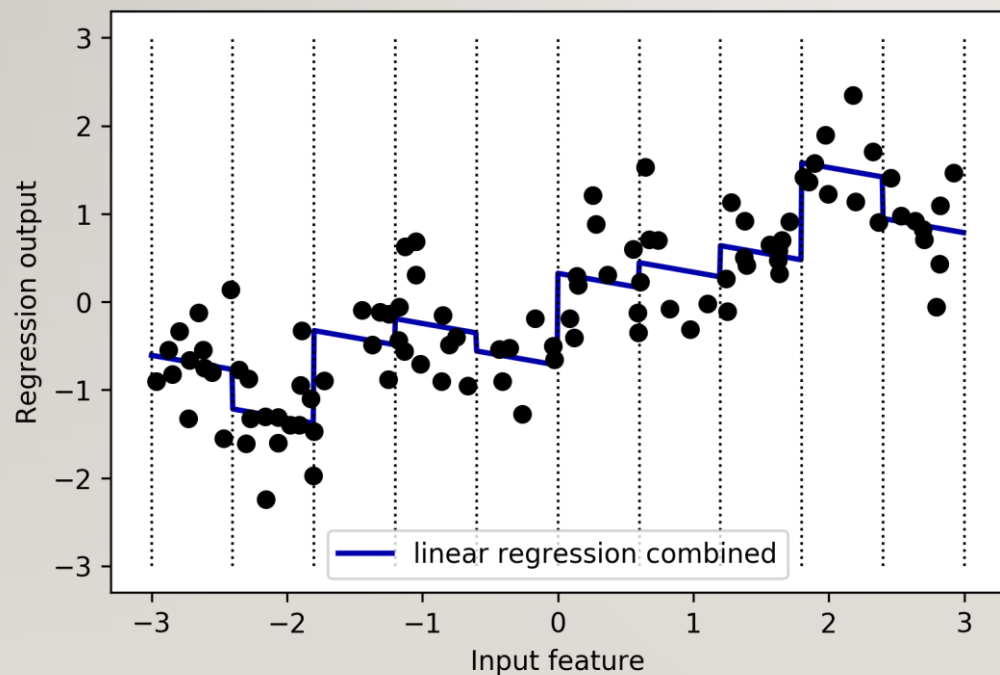
4.3 交互作用と多項式

傾きが下向きで、
すべてのビンで共有されている。
これではあまり役に立たない



4.3 交互作用と多項式

データポイントのビンの場所を示す特徴量と x 軸の場所を示す特徴量の交互作用もしくは積を特徴量として与えたもの



4.3 交互作用と多項式

特徴量の多項式を使用

- preprocessingモジュールの
PolynomialFeaturesに実装
- 多項式特徴量と線形回帰モデルを組み合わせで
多項式回帰モデルになる

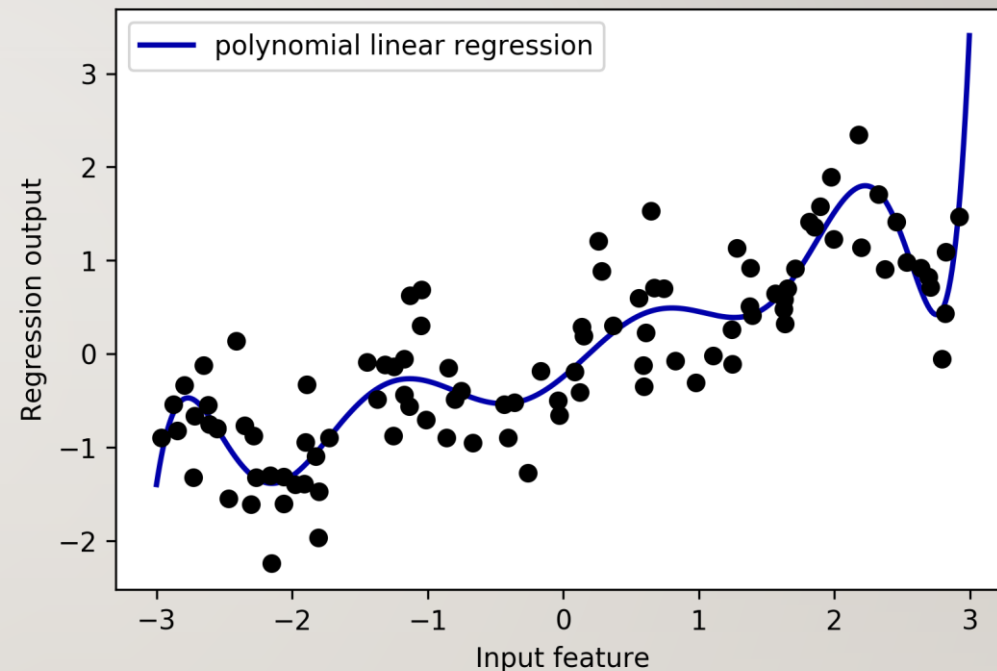
4.3 交互作用と多項式

- 問題点

高次の多項式は

境界近辺やデータが少ない領域では

極端な振る舞いを示す傾向がある



4.3 交互作用と多項式

カーネル法を用いたSVMモデルを
変換していないオリジナルデータに適用

より複雑なモデルを用いると、特徴量に対して明示的な変換を行わなくても、多項式回帰と同じように複雑な予想が可能。

