機械学習勉強メモ --- ロジスティック回帰編

ロジスティック回帰とは

- 実装が簡単で，線形分離可能なクラスでは非常に性能が良い分類モデル
- 産業界において最も広く使われている分類アルゴリズの1つ
- ADLINEとの違いは，以下の2点である．
    - 活性化関数$\phi(z)$がシグモイド関数になっている
    - コスト関数が誤差平方和から対数尤度になっている
- zはADLINEと変わらず総入力である．
- シグモイド関数を以下に示す．

$$\phi(z) = \frac{1}{1+e^{-z}}$$

重みの学習方法

- ロジスティック回帰では，コスト関数に対数尤度を用いて勾配降下法により最適な重みを見つける．
- 対数尤度(コスト関数)を以下に示す．

$$J(w) = \sum_{i=1}^{n}[-y^{(i)}log(\phi(z^{(i)})) - (1-y^{(i)})log(1-\phi(z^{(i)}))]$$

正則化

- 機械学習では，過学習や学習不足という問題がよく発生する．
- 過学習とは，訓練データでは上手く機能するモデルが未知のデータでは上手く汎化しないという問題である．
- 学習不足とは学習が十分でなくモデルが上手く複雑さに対応していないことである．
- これらの対処法にモデルの複雑さを調整する正規化というものがある．
- 最も一般的な正則化にL2正則化がある．以下に式を示す．

$$\frac{\lambda}{2}||w||^2 = \frac{\lambda}{2}\sum_{j=1}^{m} w_j^2$$

    - λは正規化パラメータである．

- 正則化を適用するにはコスト関数に対して，正則化の項を追加する．

$$J(w) = C\sum_{i=1}^{n}[-y^{(i)}log(\phi(z^{(i)})) - (1-y^{(i)})log(1-\phi(z^{(i)}))] + \frac{1}{2}||w||^2$$

- C=$\frac{1}{\lambda}$である．