1　ロジスティック回帰

1.1　分類問題

ある入力（数値）からクラスに分類する問題。

1.2　ロジスティック線形回帰モデル

分類問題を解くための教師あり機械学習モデル線形回帰モデルの出力をシグモイド関数に入力し、$y = 1$ となる確率を出力する。

$$P(Y = 1|\boldsymbol{x}) = \sigma(w_0 + \sum_{j_1}^{m} w_j x_j)$$

1.3　シグモイド関数

入力は実数、出力は 0 1 の単調増加関数。

$$\sigma(x) = \frac{1}{1 + \exp(-ax)}$$

DNN での活性化関数としてもよく用いられる。

1.4　最尤推定

ロジスティック回帰モデルでの最尤推定ではベルヌーイ分布を仮定する。

$$P(y) = p^y(1 - p)^y$$

1.5　尤度関数

データがある分布に従っていると仮定したとき、そのデータを生成したであろう分布の尤もらしさを表す関数。

$$
\begin{aligned}
P(y_1, y_2, \cdots, y_n | w_0, w_1, \cdots, w_n) &= \prod_{i=1}^{n} p_i^{y_i} (1 - p_i)^{1-y_i} \\
&= \prod_{i=1}^{n} \sigma(\boldsymbol{w}^T \boldsymbol{x}_i)^{y_i} (1 - \sigma(\boldsymbol{w}^T \boldsymbol{x}_i))^{1-y_i} \\
&= L(\boldsymbol{w})
\end{aligned}
$$

推定の際は総和で計算できるよう対数をとり、マイナスをかけた対数尤度関数 $-\log L(\boldsymbol{w})$ を最小化する。

1.6　勾配降下法（SG）

反復学習によりパラメータを逐次的に更新することで、対数尤度関数の最小化を行う。

パラメータが更新されなくなった場合、勾配が 0 になっているので、探索範囲内では最適な解が求められたことになる（ローカルミニマム）。

1.7　確率的勾配降下法（SGD）

データを一つずつランダムに選んでパラメータを更新する。

1.8　分類の評価方法

混同行列を使用して評価指標を計算する。

1.8.1　再現率（Recall）

検出に抜け漏れを少なくしたい場合に採用する。

$$\frac{TP}{TP+FN}$$

1.8.2　適合率（Precision）

見逃しが多くても正確に検出したい場合に採用する。

$$\frac{TP}{TP+FP}$$

1.8.3　F 値

Recall と Precision の調和平均