

1 線形回帰

1.1 回帰問題

ある入力（離散、連続）から出力（連続）を予測する問題のこと。
（ex. 2次元の入出力データを一次関数で近似）

1.2 線形回帰モデル

回帰問題を解くための機械学習モデルの一つ。

予測タスクに対する教師あり学習でのアプローチで、 m 次元の入力 x からスカラー値の予測値 \hat{y} を出力する。

$$\hat{y} = \mathbf{w}^T \mathbf{x} + w_0 = \sum_{j=1}^m w_j x_j + w_0$$

1.3 モデルのパラメータ

モデルに含まれる推定すべき未知パラメータ。線形回帰では最小二乗法で推定する。

1.4 単回帰モデル

$m = 1$ の時の線形回帰モデル。

$$y = w_0 + w_1 x_1 + \epsilon$$

1.5 重回帰モデル

$m > 1$ の時の線形回帰モデル。

$$y = w_0 + w_1x_1 + w_2x_2 + \epsilon$$

1.6 モデルの検証

モデルの汎化性能を測定するため、データを学習用データと検証用データに分離する。(未知データに対してどれだけ精度よく推定できているかを測りたいため。)

1.7 最小二乗法

学習データの平均二乗誤差 (MSE) を最小とするパラメータを探索する。MSE の勾配が 0 になるようなパラメータを求めればよい。