

パターン認識と機械学習

第6章 カーネル法

橋本龍二

February 10, 2021

ある手法では訓練データ点をテスト時にも使う。データ空間 Ω 上の入力ベクトル \mathbf{x} を特徴空間 H へ移す写像 $\phi(\mathbf{x})$ を考える。今、新しいデータ $\phi(\mathbf{x}')$ が与えられたとき、今までの訓練ベクトル $\phi(\mathbf{x})$ らとの類似度を求めたいと思う。類似度とは、内積

$${}^T\phi(\mathbf{x})\phi(\mathbf{x}')$$

によって与えられる。ところが、 ϕ が複雑な非線形関数の場合、これを計算するのは難しい。そこで、

$$k(\mathbf{x}, \mathbf{x}') = {}^T\phi(\mathbf{x})\phi(\mathbf{x}')$$

という、 \mathbf{x}, \mathbf{x}' のみに依存し、 ϕ の形を考えなくても内積の値が求まる関数 k を使う。 k をカーネル関数という。

1 双対表現

線形回帰モデルで、パラメータが以下のように正則化された二乗和誤差関数を考える。

$$J(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N \{ {}^T\mathbf{w}\phi(\mathbf{x}_n) - t_n \}^2 + \frac{\lambda}{2} {}^T\mathbf{w}\mathbf{w}$$

\mathbf{w} で微分して 0 と置くと、

$$\begin{aligned}\frac{\partial J}{\partial \mathbf{w}} &= \sum_{n=1}^N \{ {}^T \mathbf{w} \phi(\mathbf{x}_n) - t_n \} \phi(\mathbf{x}_n) + \lambda \mathbf{w} \\ \therefore \mathbf{w} &= -\frac{1}{\lambda} \sum_{n=1}^N \{ {}^T \mathbf{w} \phi(\mathbf{x}_n) - t_n \} \phi(\mathbf{x}_n) \\ &= \sum_{n=1}^N a_n \phi(\mathbf{x}_n)\end{aligned}$$

ただし、

$$a_n = -\frac{1}{\lambda} \{ {}^T \mathbf{w} \phi(\mathbf{x}_n) - t_n \}$$

と置いた。したがって、

$$\begin{aligned}\mathbf{w} = \sum_{n=1}^N a_n \phi(\mathbf{x}_n) &= (\phi(\mathbf{x}_1) \quad \dots \quad \phi(\mathbf{x}_n)) \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \\ &= {}^T \Phi \mathbf{a}\end{aligned}$$

ただし、

$$\Phi = \begin{pmatrix} {}^T \phi(\mathbf{x}_1) \\ \vdots \\ {}^T \phi(\mathbf{x}_n) \end{pmatrix}, \quad \mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

と置いた。 $\mathbf{w} = {}^T \Phi \mathbf{a}$ を $J(\mathbf{w})$ に代入して、

$$\begin{aligned}J(\mathbf{a}) &= \frac{1}{2} ({}^T \mathbf{w} {}^T \Phi - {}^T \mathbf{t}) (\Phi \mathbf{w} - \mathbf{t}) + \frac{\lambda}{2} {}^T \mathbf{w} \mathbf{w} \\ &= \frac{1}{2} ({}^T \mathbf{a} \Phi {}^T \Phi - {}^T \mathbf{t}) (\Phi {}^T \Phi \mathbf{a} - \mathbf{t}) + \frac{\lambda}{2} {}^T \mathbf{w} \mathbf{w} \\ &= \frac{1}{2} {}^T \mathbf{a} \Phi {}^T \Phi \Phi {}^T \Phi \mathbf{a} - {}^T \mathbf{a} {}^T \Phi {}^T \Phi \mathbf{t} + \frac{1}{2} {}^T \mathbf{t} \mathbf{t} + \frac{\lambda}{2} {}^T \mathbf{a} \Phi {}^T \Phi \mathbf{a}\end{aligned}$$