４－２ 凸計画問題を用いたカーネル多変量解析

05T4007T 江口 晃

サポートベクトル回帰（1）

- サポートベクトル回帰

  -回帰問題の場合にも二乗誤差を区分線形関数の誤差にすることによってロバストでスパースな関数を得ることができる。

- ε -不感応関数

$$
r_{\varepsilon}(z)=\begin{cases} z-\varepsilon & (\varepsilon \leqq z\text{のとき}) \\ 0 & (-\varepsilon \leqq z < \varepsilon\text{のとき}) \\ -z-\varepsilon & (z < -\varepsilon\text{のとき}) \end{cases}
$$

※目的値 y、関数の出力 f(x)、z=y-f(x)

サポートベクトル回帰（2）

- 損失関数

$$\min_{\xi,\alpha} \sum_{i=1}^{n} \xi_i + \lambda \alpha^T K \alpha$$

- 制約式

$$\xi_i \geq -(y^{(i)} - f(x^{(i)})) - \varepsilon$$

$$\xi_i \geq 0$$

$$\xi_i \geq y_i - f(x^{(i)}) - \varepsilon \qquad\qquad i = 1, \ldots, n$$

双対問題の導出（1）

サポートベクトル回帰の双対問題は、3つの制約式に対応するβ i、γ i+、γ i-という3種類のラグランジュ乗数を使う。

式変形すると、β iには依存しない最大化問題となる。

- 最大化問題

$$L_{dual}(\gamma^+,\gamma^-) = -\frac{1}{2\lambda}\sum_{i,j}(\gamma_i^+ - \gamma_i^-)(\gamma_j^+ - \gamma_j^-)K_{ij}$$

$$-\sum_{i=1}^{n}\gamma_i^-(y_i+\varepsilon) - \sum_{i=1}^{n}\gamma_i^+(-y_i+\varepsilon)$$

双対問題の導出（2）

- 制約式

$$0 \leq \gamma_i^+ + \gamma_i^- \leq 1$$

- サポートベクトル回帰で求めたい関数

$$f(x) = \frac{1}{\lambda} \sum_{x^{(i)} \in SV} (\gamma_i^+ - \gamma_i^-) k(x^{(i)}, x)$$

サポートベクトル回帰のスパース性

- サポートベクトル回帰のスパース性

γ_i^+とγ_i^-のどちらか一方は必ず0であり、サポートベクトルは関数からのずれがε以上であるようなサンプルである。

損失関数の一般化

これまでは、損失関数が区分線形関数の形で表現されており、そこから凸二次計画問題に帰着させることができた。

一般化は、関数そのものを使う代わりに、その２乗を損失として使う。

$$\min_{\xi,\alpha} \sum_{i=1}^{n} \xi_i^2 + \frac{\lambda}{2} \alpha^T K \alpha$$

損失関数も最適化する

- 区分線形関数の位置パラメータ

区分線形関数(式4.4)を平行移動して、パラメータ $\rho>0$を持つ関数を誤差に取る。

$$\gamma_{\rho}(yf(x)) = \max\{0, \rho - yf(x)\}$$

最適化関数に－ρに比例した罰金項を加えて、ρ が小さくなりすぎないようにする。

ν -サポートベクトルマシン（1）

- 最小化問題

ξ i≧0, $\xi_i \geq \rho - y_i \sum_{j=1}^{n} \alpha_j K_{ij}$, ρ ≧0

という制約のもとで、

$$\min_{\xi,\alpha,\rho} \frac{1}{n} \sum_{i=1}^{n} \xi_i + \frac{1}{2} \alpha^T K \alpha - \upsilon \rho$$

という最小化問題となる。ただしν は正の定数で正則化パラメータの役割を果たす。

区分線形の損失関数の傾きが変化する点のパラメータを最適化する仕組みをνトリックと呼ぶ。

ν -サポートベクトルマシン（2）

- 双対問題

-最大化問題

$$L_{dual}(\gamma) = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} y^{(i)} y^{(j)} \gamma_i \gamma_j K_{ij}$$

-制約条件

$$0 \leq \gamma_i \leq \frac{1}{n} \quad , \qquad \upsilon \leq \sum_{i=1}^{n} \gamma_i$$

-識別関数

$$f(x) = \sum_{x^{(i)} \in SV} \gamma_i y^{(i)} k(x^{(i)}, x)$$

ν -サポートベクトルマシンの特徴

-サポートベクトルの全体の割合(|SV|/n)の下限値が ν になっている。

-識別の誤り率(Err/n)の上限値が ν である。

-ゆるやかな正則条件のもとで、下限|SV|/nも上限 Err/nも、サンプル数が増えるにつれて漸近的に ν に収束する。

サポートベクトルの数や誤り率をあらかじめ制限したい場合には有効な性質である。

ν -サポートベクトル回帰

- 最適化関数

$$\min_{\xi,\alpha,\varepsilon} \sum_{i=1}^{n} \xi_i + \lambda\alpha^T K\alpha + \upsilon\varepsilon$$

- 制約式

$$\xi_i \geq -(y^{(i)} - f(x^{(i)})) - \varepsilon \qquad \xi_i \geq 0 \qquad \xi_i \geq y_i - f(x^{(i)}) - \varepsilon$$

罰金項としてεに比例した値を最適化関数に加えておく。