

# カーネル法の基礎 4.2節

## 参考文献

- 金森, 統計的学習理論, 講談社, 2015.

## Table of contents

- カーネル関数

## カーネル関数

(4.2)式の関数 $k(x, x')$ に対して、行列 $K = (k(x_i, x_j))$ は対称非負定値行列になる。  
実際、 $c_1, \dots, c_n \in \mathbb{R}$ に対して、

$$\sum_{i=1}^n \sum_{j=1}^n c_i c_j K_{ij} = \sum_{i=1}^n \sum_{j=1}^n c_i c_j \phi(x_i)^T \phi(x_j) = \left\| \sum_{i=1}^n c_i \phi(x_i) \right\|^2 \geq 0$$

より、示される。

⇨ カーネル関数を以下のように定義する。

## Def 4.1 (カーネル関数)

関数  $k : \mathcal{X}^2 \rightarrow \mathbb{R}$  が次の対称非負定値性を満たすとき、 $k$  を  $\mathcal{X}$  上のカーネル関数 (kernel function) と呼ぶ。

### 対称非負定値性

$\forall x_1, \dots, x_n \in \mathcal{X}, n \geq 1$  に対して、 $n \times n$  行列  $K = (K_{ij})$  を  $K_{ij} = k(x_i, x_j)$  とするとき、 $K$  が対称非負定値行列になる。

つまり、 $K$  は対称行列で、 $\forall c \in \mathbb{R}^n$  に対して  $c^T K c \geq 0$

(i.e.,  $\sum_{i=1}^n \sum_{j=1}^n c_i c_j K_{ij} \geq 0$ )

が成立する。

$K$  をグラム行列 (Gram matrix) という。

- カーネル関数が与えられれば、線形モデル $\mathcal{M}$ を明示せずに推定量 (4.3) が計算できる
  - 学習アルゴリズムによって得られる推定量の統計的性質を調べるときは、カーネル関数に対応する統計モデルを明示して議論する必要がある ( $\rightsquigarrow$  4.3節)

## Th 4.2

$k, k_\ell, \ell = 1, 2, \dots$ を $\mathcal{X}$ 上のカーネル関数とする．このとき、次の性質が成り立つ．

1.  $k(x, x) \geq 0 \ \forall x \in \mathcal{X}$
2.  $a, b \geq 0$ とする． $ak_1 + b, k_1 + k_2, k_1 \cdot k_2$ は $\mathcal{X}$ 上のカーネル関数
3. 各点で極限 $k_\infty = \lim_{\ell \rightarrow \infty} k_\ell$ が存在するとき、 $k_\infty$ は $\mathcal{X}$ 上のカーネル関数

証明は別資料．

以下，カーネル関数の例を紹介する．

#### 例4.1 線形カーネル (linear kernel)

$$k(x, x') = x^T x'$$

- 線形判別を行うカーネル関数
- 学習データの入力ベクトル $x$ の要素に0が多く含まれる場合に有効 (e.g. テキストデータ).

#### 例4.2 多項式カーネル (polynomial kernel)

$$k(x, x') = (x^T x' + 1)^d, \quad d \in \mathbb{N}$$

- 画像データの判別に利用される．
- Th4.2 2.より，カーネル関数であることが示される．

### 例4.3 ガウシアンカーネル (Gaussian kernel)

$$k(x, x') = \exp(-\lambda \|x - x'\|^2), \lambda > 0$$

- データに事前知識がないときに使われる.
- 表現力が高い (普遍カーネル, 4.6).

$$\exp(-\lambda \|x - x'\|^2) = e^{-\lambda \|x\|^2} e^{-\lambda \|x'\|^2} e^{2\lambda x^T x'}$$

前半2項はカーネル関数である. ここで, 関数 $e^{2\lambda x^T x'}$ がカーネル関数であることを示せば, ガウシアンカーネルがカーネル関数であることが示される.

$$e^{2\lambda x^T x'} = \sum_{\ell=0}^{\infty} \frac{(2\lambda)^\ell}{\ell!} (x^T x')^\ell$$

とTaylor展開でき, Th4.2 2.から各項がカーネル関数となり, Th4.2 3.より, 極限もカーネル関数となる.