65.

命題 49 の証明

任意の $g = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{i,j} e_{X,i}(x) e_{Y,j}(y) \in H_X \otimes H_Y$，$(x, y) \in E$ について，その評価値が有界である。実際，

$$|g(x,y)| \le \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} |a_{i,j}| \cdot |e_{X,i}(x)| \cdot |e_{Y,j}(y)| \le \sum_{i=1}^{\infty} |e_{X,i}(x)| \cdot \left(\sum_{j=1}^{\infty} e_{Y,j}^2(y)\right)^{1/2} \left(\sum_{j=1}^{\infty} a_{i,j}^2(y)\right)^{1/2} \tag{5.21}$$

となる($\sum_{j=1}^{\infty}$ に関して Cauchy-Schwarz の不等式 (2.5) を用いた)。ここで，$k_Y(y, \cdot) = \sum_j h_j(y) e_{Y,j}(\cdot)$ とおくと，$\langle e_{Y,i}(\cdot), k_Y(y, \cdot)\rangle = e_{Y,i}(y)$ より，$h_i(y) = e_{Y,i}(y)$ となり，$k_Y(y, \cdot) = \sum_{j=1}^{\infty} e_{Y,j}(y) e_{Y,j}(\cdot)$ が成立する。したがって，

$$\sum_{j=1}^{\infty} e_{Y,j}^2(y) = k_Y(y, y) \tag{5.22}$$

および

$$\sum_{i=1}^{\infty} |e_{X,i}(x)| \cdot \left(\sum_{j=1}^{\infty} a_{i,j}^2\right)^{1/2} \le \left(\sum_{i=1}^{\infty} e_{X,i}^2(x)\right)^{1/2} \left(\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{i,j}^2\right)^{1/2} = \sqrt{k_X(x,x)}\|g\| \tag{5.23}$$

が成立する($\sum_{i=1}^{\infty}$ に関して Cauchy-Schwarz の不等式 (2.5) を用いた)。(5.21)–(5.23) は，$|g(x,y)| \le \sqrt{k_X(x,x)}\sqrt{k_Y(y,y)}\|g\|$ を意味する。したがって，$H_X \otimes H_Y$ は RKHS である。

a

そして，$k(x, x', y, y') := k_X(x, x') k_Y(y, y')$ は，$k_X(x, \cdot) \in H_X$，$k_Y(y, \cdot) \in H_Y$ より，$k(x, \cdot, y, \star) := k_X(x, \cdot) k_Y(y, \cdot) \in H_X \otimes H_Y$ が成立する。

b

さらに，

$$\begin{aligned} g(x,y) &= \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{i,j} e_{X,i}(x) e_{Y,j}(y) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{i,j} \langle e_{X,i}(\cdot), k_X(x, \cdot)\rangle_{H_X} \langle e_{Y,i}(\star), k_Y(y, \star)\rangle_{H_Y} \\ &= \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{i,j} \langle e_{X,i}(\cdot) e_{Y,j}(\star), k(x, \cdot, y, \star)\rangle_H = \left\langle \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{i,j} e_{X,i}(\cdot) e_{Y,j}(\cdot), k(x, \cdot, y, \star)\right\rangle_H \\ &= \langle g(\cdot, \star), k(x, \cdot, y, \star)\rangle \end{aligned}$$

より，k は $H_X \otimes H_Y$ の再生核である。 □

c

66. $\mathbb{E}[k_X(X,X)]$ と $\mathbb{E}[k_Y(Y,Y)]$ は有限であると仮定する。このとき、
$k_X(x,\cdot)k_Y(y,\cdot) \in H_X \otimes H_Y$ を XY の同時分布で平均をとっても $\mathbb{E}_{XY}[k_X(X,\cdot)k_Y(Y,\cdot)]$ について

$$\mathbb{E}_{XY}[\|k_X(X,\cdot)k_Y(Y,\cdot)\|_{H_X\otimes H_Y}] = \mathbb{E}_{XY}[\|k_X(X,\cdot)\|_{H_X}\|k_Y(Y,\cdot)\|_{H_Y}]$$
$$= \mathbb{E}_{XY}\left[\sqrt{k_X(X,X)k_Y(Y,Y)}\right]$$
$$\leq \sqrt{\mathbb{E}_X[k_X(X,X)]\,\mathbb{E}_Y[k_Y(Y,Y)]}$$

となり、この左辺は仮定より有限の値をとる。したがって、
$f \in H_X \otimes H_Y$ に対して、

$$\mathbb{E}_{XY}[f(X,Y)] = \mathbb{E}_{XY}[\langle f, k_X(X,\cdot)k_Y(Y,\cdot)\rangle]$$
$$\leq \|f\|_{H_X\otimes H_Y}\,\mathbb{E}_{XY}[\|k_X(X,\cdot)k_Y(Y,\cdot)\|_{H_X\otimes H_Y}]$$

が成立し、Riesz の表現定理より、

$$\mathbb{E}_{XY}[f(X,Y)] = \langle f, m_{XY}\rangle$$

を満たす $m_{XY} \in H_X \otimes H_Y$ が存在する。

このように平均を定義する。

67. 任意の $g \in H_Y$ について、線形汎関数

$$T_g : H_X \ni f \mapsto \langle fg, m_{XY} - m_X m_Y \rangle_{H_X \otimes H_Y} \in \mathbb{R}$$

は、$\langle fg, m_{XY} - m_X m_Y \rangle_{H_X \otimes H_Y}$

$$\leq \|f\|_{H_X} \|g\|_{H_Y} \|m_{XY} - m_X m_Y\|_{H_X \otimes H_Y}$$

より有界であり、Rieszの表現定理より、

$T_g f = \langle f, h_g \rangle_{H_X}$ なる $h_g \in H_X$ が存在する。すなわち、

$$\langle fg, m_{XY} - m_X m_Y \rangle_{H_X \otimes H_Y} = \langle f, \Sigma_{XY} g \rangle_{H_X}$$

なる $\Sigma_{XY} : H_Y \ni g \mapsto h_g \in H_X$ が存在する。

また、

$$\|\Sigma_{XY} g\|_{H_X} = \|h_g\|_{H_X} = \|T_g\| \leq \|g\|_{H_Y} \|m_{XY} - m_X m_Y\|_{H_X \otimes H_Y}$$

より、Σ_{XY} は有界である。 図

68. $\mathcal{F} := \{ f \in H \mid \|f\|_H \leq 1 \}$

$MMD := \sup_{f \in \mathcal{F}} \{ E_P[f(x)] - E_Q[f(x)] \}$ とする。

このとき、

$$MMD^2 = \sup_{f \in \mathcal{F}} \{ \langle m_P, f \rangle - \langle m_Q, f \rangle \}^2$$

$$= \sup_{f \in \mathcal{F}} \{ \langle m_P - m_Q, f \rangle \}^2 = \| m_P - m_Q \|_H^2$$

負のケースは？

が成り立つ。($MMD = \| m_P - m_Q \|_H$ は言える？？)

さらに、X' は X と、(Y' は Y と) 同一の分布に従う独立な確率変数とする。

このとき、

$$MMD^2 = \langle m_P, m_P \rangle^2 + \langle m_Q, m_Q \rangle^2 - 2 \langle m_P, m_Q \rangle$$

$$= \langle E_X[k(X, \cdot)], E_{X'}[k(X', \cdot)] \rangle + \langle E_Y[k(Y, \cdot)], E_{Y'}[k(Y', \cdot)] \rangle$$

$$- 2 \langle E_X[k(X, \cdot)], E_Y[k(Y, \cdot)] \rangle$$

$$= E_{XX'}[k(X, X')] + E_{YY'}[k(Y, Y')] - 2 E_{XY}[k(X, Y)]$$

が成り立つ。

図

69.

$$\frac{1}{m(m-1)}\sum_{i=1}^{m}\sum_{j\neq i}k(x_i,x_j)+\frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}k(y_i,y_j)-\frac{2}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}k(x_i,y_j) \qquad (5.4)$$

が、MMD^2 の不偏推定量であることを示す。

まず、68.より、

$$MMD^2 = \mathbb{E}_{XX'}[k(X,X')] + \mathbb{E}_{YY'}[k(Y,Y')] - 2\mathbb{E}_{XY}[k(X,Y)]$$

が成り立つ。

$$\mathbb{E}\left[\frac{1}{m(m-1)}\sum_{i=1}^{m}\sum_{j\neq i}k(X_i,X_j)\right]$$

$$= \frac{1}{m}\sum_{i=1}^{m}\mathbb{E}_{X_i}\left[\frac{1}{m-1}\sum_{j\neq i}\mathbb{E}_{X_j}[k(X_i,X_j)]\right] \quad (\because \text{期待値の線形性})$$

$$= \mathbb{E}_{XX'}[k(X,X')]$$

であり、第2項も同様に、

$$\mathbb{E}\left[\frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}k(Y_i,Y_j)\right] = \mathbb{E}_{YY'}[k(Y,Y')]$$ であり、また、

$$\mathbb{E}\left[\frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}k(X_i,Y_j)\right] = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\mathbb{E}_{X_iY_j}[k(X_i,Y_j)]$$

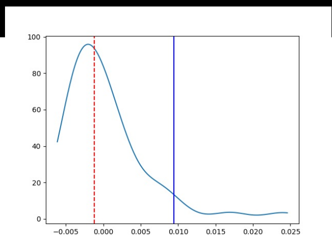
$$= \mathbb{E}_{XY}[k(X,Y)]$$

以上より、

$$\mathbb{E}[(5.4)] = MMD^2$$ が成り立つ。 ∎

70.

```
"""問題70"""
from pathlib import Path

import matplotlib.pyplot as plt
import numpy as np
from scipy.stats import gaussian_kde

sigma = 1
def k(x: np.ndarray, y: np.ndarray) -> float:
    """ガウシアンカーネル"""
    return np.exp(-(x - y)**2 / (2 * sigma**2))

# データの生成
m = 150
n = 100
rs = np.random.RandomState(42)
xx = rs.randn(m)
yy = rs.randn(n)

# 帰無分布の計算
T: list[float] = []
for _ in range(100):
    index1 = rs.choice(m, size=int(m/2), replace=False)
    index2 = rs.choice(n, size=int(n/2), replace=False)
    x_perm = np.concatenate((xx[index1], yy[index2]))
    y_perm = np.concatenate((xx[np.setdiff1d(np.arange(m), index1)], yy[np.setdiff1d(np.arange(n),
index2)]))
    m_x_perm = x_perm.shape[0]
    n_y_perm = y_perm.shape[0]
    # (5.4)の計算
    term1 = sum(k(x_perm[i], x_perm[j]) for i in range(m_x_perm) for j in range(m_x_perm) if i != j) /
(m_x_perm * (m_x_perm - 1))
    term2 = sum(k(y_perm[i], y_perm[j]) for i in range(n_y_perm) for j in range(n_y_perm) if i != j) /
(n_y_perm * (n_y_perm - 1))
    term3 = sum(k(x_perm[i], y_perm[j]) for i in range(m_x_perm) for j in range(n_y_perm)) * 2 /
(m_x_perm * n_y_perm)
    T.append(term1 + term2 - term3)
v = np.quantile(T, 0.95)

# 統計量の計算
term1 = sum(k(xx[i], xx[j]) for i in range(m) for j in range(m) if i != j) / (m * (m - 1))
term2 = sum(k(yy[i], yy[j]) for i in range(n) for j in range(n) if i != j) / (n * (n - 1))
term3 = sum(k(xx[i], yy[j]) for i in range(m) for j in range(n)) * 2 / (m * n)
u = term1 + term2 - term3

# グラフの図示
x = np.linspace(min(*T, u, v), max(*T, u, v), 200)
density = gaussian_kde(T)
plt.plot(x, density(x))
plt.axvline(x=u, c="r", linestyle="--")
plt.axvline(x=v, c="b")

filename = Path("src/exercise/Chap5/out/problem70/result.png")
filename.parent.mkdir(parents=True, exist_ok=True)
plt.savefig(filename)
```



71. 帰無仮説 $X=Y$ の下で，

$h_1 \equiv 0$ （恒等的に 0 である関数）を示す。

$$h_1(Z_1) = \mathbb{E}_{Z_2}[h(Z_1, Z_2)]$$
$$= \mathbb{E}[k(x_i, X_j)] + \mathbb{E}[k(y_i, Y_j)] - \mathbb{E}[k(x_i, Y_j)] - \mathbb{E}[k(x_j, Y_i)]$$
$$= 0.$$

$\boxtimes$

また，$\tilde{h}_2 \equiv h$ を示す。

帰無仮説の下で，

$$\Theta := \mathbb{E}\, h(Z_1, \dots, Z_m) = 0 \quad \text{であり，}$$
$$\tilde{h}_c(Z_1, \dots, Z_c) := h_c(Z_1, \dots, Z_c) - \Theta$$

であるので，

$$\tilde{h}_2(Z_1, Z_2) = h_2(Z_1, Z_2) \quad \text{である}$$

$\boxtimes$

72. $X \sim N(\mu_X, \sigma_X^2)$, $Y \sim N(\mu_Y, \sigma_Y^2)$ とする。

(i) X と Y が独立 $\Rightarrow \rho_{XY} = 0$ を示す。

X と Y が独立であるとき, $E[XY] = E[X]E[Y]$ が成り立つため,

$\mathrm{Cov}[X,Y] = E[XY] - E[X]E[Y] = 0$ である。

よって $\rho_{XY} = \dfrac{\mathrm{Cov}[X,Y]}{\sigma_X \sigma_Y} = 0$ である。

(ii) $\rho_{XY} = 0 \Rightarrow$ X と Y が独立を示す。

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{XY}^2}} \exp\left\{-\frac{1}{2(1-\rho_{XY}^2)}\left(\frac{x^2}{\sigma_X^2} + \frac{y^2}{\sigma_Y^2} - \frac{2\rho xy}{\sigma_X\sigma_Y}\right)\right\}$$

であるが, $\rho_{XY} = 0$ のとき,

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y}\exp\left\{-\frac{1}{2}\left(\frac{x^2}{\sigma_X^2} + \frac{y^2}{\sigma_Y^2}\right)\right\}$$

$$= f_X(x) f_Y(y)$$

であるため, X と Y は独立である

以上 (i), (ii) より示せた。

また, $\theta \sim U(0, 2\pi)$, $X = \cos\theta$, $Y = \sin\theta$ とすると明らかに,

$X \not\perp\!\!\!\perp Y$ である。一方,

$$\mathrm{Cov}[X,Y] = E[XY] - \overset{0}{\overset{\shortparallel}{E[X]}}\,\overset{0}{\overset{\shortparallel}{E[Y]}} = E[\cos\theta\sin\theta] = \frac{1}{2}E[\sin 2\theta] = 0$$

であり, $\rho_{XY} = \dfrac{\mathrm{Cov}[X,Y]}{\sigma_X\sigma_Y} = 0$ である。

73. $\|\cdot\|_{X\otimes Y}, \langle\cdot,\cdot\rangle_{X\otimes Y}$ を $\|\cdot\|, \langle\cdot,\cdot\rangle$ と略記する。

$$
\begin{aligned}
\|m_{XY}\|^2 &= \langle \mathbb{E}_{XY}[k_X(X,\cdot)k_Y(Y,\cdot)], \mathbb{E}_{X'Y'}[k_X(X',\cdot)k_Y(Y',\cdot)]\rangle \\
&= \mathbb{E}_{XY}\mathbb{E}_{X'Y'}[\langle k_X(X,\cdot)k_Y(Y,\cdot), k_X(X',\cdot)k_Y(Y',\cdot)\rangle] \\
&= \mathbb{E}_{XYX'Y'}[k_X(X,X')k_Y(Y,Y')]
\end{aligned}
$$
,

$$
\begin{aligned}
&\langle m_{XY}, m_X m_Y\rangle \\
&= \langle \mathbb{E}_{XY}[k_X(X,\cdot)k_Y(Y,\cdot)], \mathbb{E}_{X'}[k_X(X',\cdot)]\mathbb{E}_{Y'}[k_Y(Y',\cdot)]\rangle \\
&= \mathbb{E}_{XY}\left[\mathbb{E}_{X'}[\langle k_X(X,\cdot)k_Y(Y,\cdot), k_X(X',\cdot)\mathbb{E}_{Y'}[k_Y(Y',\cdot)]\rangle]\right] \\
&= \mathbb{E}_{XY}\left[\mathbb{E}_{X'}[k_X(X,X')]\mathbb{E}_{Y'}[\langle k_Y(Y,\cdot), k_Y(Y',\cdot)\rangle]\right] \\
&= \mathbb{E}_{XY}\left[\mathbb{E}_{X'}[k_X(X,X')]\mathbb{E}_{Y'}[k_Y(Y,Y')]\right]
\end{aligned}
$$
,

$$
\begin{aligned}
&\|m_X m_Y\|^2 \\
&= \langle \mathbb{E}_X[k_X(X,\cdot)]\mathbb{E}_Y[k_Y(Y,\cdot)], \mathbb{E}_{X'}[k_X(X',\cdot)]\mathbb{E}_{Y'}[k_Y(Y',\cdot)]\rangle \\
&= \mathbb{E}_X\mathbb{E}_{X'}[k_X(X,X')]\mathbb{E}_Y\mathbb{E}_{Y'}[k_Y(Y,Y')]
\end{aligned}
$$

であるので、

$$
\|m_{XY} - m_X m_Y\|^2 = \|m_{XY}\|^2 - 2\langle m_{XY}, m_X m_Y\rangle + \|m_X m_Y\|^2
$$

= (示したい右辺：(5.9)式)　よって示せた。

D4. traceで計算できることの証明

$$\mathrm{tr}(K_X H K_Y H) = \sum_i (K_X H K_Y H)_{i,i}$$

$$= \sum_i \sum_j (K_X H)_{i,j} (K_Y H)_{j,i}$$

$$= \sum_i \sum_j \Big\{ \sum_k k_X(x_i, x_j) k_Y(y_i, y_j) - \frac{1}{N} k_X(x_i, x_j) \sum_k k_Y(y_i, y_k)$$

$$- \frac{1}{N} k_Y(y_i, y_j) \sum_k k_X(x_i, x_k)$$

$$+ \frac{1}{N^2} \sum_k k_X(x_i, x_k)$$

$$= \sum_i \sum_j k_X(x_i, x_j) k_Y(y_i, y_j) - \frac{2}{N} \sum_i \Big\{ \sum_j k_X(x_i, x_j) \sum_k k_Y(y_i, y_k) \Big\}$$

$$+ \frac{1}{N^2} \sum_i \sum_k k_X(x_i, x_k) \sum_j \sum_r k_Y(y_j, y_r)$$

$$=: N^2 \widehat{\mathrm{HSIC}}$$

$\blacksquare$

```python
"""問題74"""
import numpy as np


def HSIC(x, y, k_x, k_y):
    n = len(x)
    S = 0
    for i in range(n):
        for j in range(n):
            S = S + k_x(x[i], x[j]) * k_y(y[i], y[j])
    T = 0
    for i in range(n):
        T_1 = 0
        for j in range(n):
            T_1 = T_1 + k_x(x[i], x[j])
        T_2 = 0
        for l in range(n):
            T_2 = T_2 + k_y(y[i], y[l])
        T = T + T_1 * T_2
    U = 0
    for i in range(n):
        for j in range(n):
            U = U + k_x(x[i], x[j])
    V = 0
    for i in range(n):
        for j in range(n):
            V = V + k_y(y[i], y[j])
    return S / n**2 - 2 * T / n**3 + U * V / n**4


def HSIC_trace(x, y, k_x, k_y):
    n = len(x)
    K_x = np.zeros((n, n))
    for i in range(n):
        for j in range(n):
            K_x[i, j] = k_x(x[i], x[j])
    K_y = np.zeros((n, n))
    for i in range(n):
        for j in range(n):
            K_y[i, j] = k_y(y[i], y[j])
    E = np.ones((n, n))
    H = np.identity(n) - E / n
    return np.sum(np.diag(np.diag(K_x.dot(H).dot(K_y).dot(H)))) / n**2

if __name__ == "__main__":
    def k_x(x, y):
        return np. exp(-np.linalg.norm(x - y)**2 / 2)

    k_y = k_x
    k_z = k_x

    n = 100
    rs = np.random.RandomState(42)
    for a in [0, 0.1, 0.2, 0.4, 0.6, 0.8]:
        x = rs.randn(n)
        z = rs.randn(n)
        y = a * x + np.sqrt(1 - a**2) * z
        print(f"{a=}, {HSIC(x, y, k_x, k_y)=}, {HSIC_trace(x, y, k_x, k_y)}")
    """
    a=0, HSIC(x, y, k_x, k_y)=0.0016144347904881728, 0.0016144347904855324
    a=0.1, HSIC(x, y, k_x, k_y)=0.0024845938768538467, 0.002484593876853815
    a=0.2, HSIC(x, y, k_x, k_y)=0.0020244572354856105, 0.0020244572354858425
    a=0.4, HSIC(x, y, k_x, k_y)=0.0032241047523186017, 0.003224104752316417
    a=0.6, HSIC(x, y, k_x, k_y)=0.011470114073025228, 0.011470114073027894
    a=0.8, HSIC(x, y, k_x, k_y)=0.019901742290715563, 0.019901742290714505
    """
```

同一の結果が出力されている。

75.

```python
"""問題75"""
def HSIC_2(x, y, z, k_x, k_y, k_z):
    n = len(x)
    S = 0
    for i in range(n):
        for j in range(n):
            S = S + k_x(x[i], x[j]) * k_y(y[i], y[j]) * k_z(z[i], z[j])
    T = 0
    for i in range(n):
        T_1 = 0
        for j in range(n):
            T_1 = T_1 + k_x(x[i], x[j])
        T_2 = 0
        for l in range(n):
            T_2 = T_2 + k_y(y[i], y[l]) * k_z(z[i], z[j])
        T = T + T_1 * T_2
    U = 0
    for i in range(n):
        for j in range(n):
            U = U + k_x(x[i], x[j])
    V = 0
    for i in range(n):
        for j in range(n):
            V = V + k_y(y[i], y[j]) * k_z(z[i], z[j])
    return S / n**2 - 2 * T / n**3 + U * V / n**4

if __name__ == "__main__":
    import numpy as np

    def k_x(x, y):
        return np. exp(-np.linalg.norm(x - y)**2 / 2)

    k_y = k_x
    k_z = k_x

    n = 100
    rs = np.random.RandomState(42)
    x = rs.randn(n)
    y = 2 * x + rs.randn(n)
    z = -3 * x + rs.randn(n)

    print(HSIC_2(x, y, z, k_x, k_y, k_z))
    # 0.0007463123042284747
```

HSICの値は十分に小さい。

7b.

(1). HSIC_1(y-x, z-xy, k-y, k-z)

(2). HSIC_1(z-x, y-zx, k-z, k-y)

(3). HSIC_1(z-y, x-yz, k-y, k-z)

(4). HSIC_1(x-y, z-xy, k-z, k-y)

(5). HSIC_1(z-y, x-yz, k-z, k-x)

(6). HSIC_1(x-y, z-xy, k-x, k-z)

77.

```
# データの生成
x = np.random.randn(n)
y = np.random.randn(n)
u = HSIC_1(x, y, k_x, k_y)
m = 100
w = []
for i in range(m):
    x = x[np.random.choice(n, n, replace=False)]
    w.append(HSIC_1(x, y, k_x, k_y))
v = np.quantile(w, 0.95)
x = np.linspace(min(min(w), u, v), max(max(w), u, v), 200)

density = kde.gaussian_kde(w)
plt.plot(x, density(x))
plt.axvline(x=v, c="r", linestyle="--")
plt.axvline(x=u, c="b")
```

統計量のHSICを求める.

帰無仮説に従うHSICの値を複数得ている.

並び替えることによって、x_i と y_i の相関を無くすことができるため、帰無仮説(X, Y が独立)が得られる。

78. 積分作用素のカーネル k_2（命題53）は、対称なカーネルであり、自己共役であるため、命題27より、固有値 $\{\lambda_i\}_i$ と固有関数 $\{\phi_i\}_i$ が存在する。

79. $k(x,y) = \phi(x-y)$, $\phi(t) = e^{-|t|}$

が特性カーネルであることを示す。

$\phi(t)$の逆フーリエ変換を用いると、

$$\phi(t) = \int_{-\infty}^{\infty} \frac{2}{1+w^2} e^{iwt} dw$$

となるため、$\phi(t) = \int_E e^{iwt} d\eta(w)$ の形で表せる。

また、η の台は $E = \mathbb{R}$ であるため、

命題54から、$k(x,y) = e^{-|x-y|}$ は特性カーネルである。

80、$g(w) := (\varepsilon - \|w\|_2)_+^{\frac{(d+1)}{2}}$, $d=1$ にて。

$$\frac{1}{2\pi}\int_{-\varepsilon}^{\varepsilon} g(w) e^{-iwx}\, dw = \frac{1-\cos(x\varepsilon)}{\pi x^2}$$ を示す。

$$\int_{-\varepsilon}^{\varepsilon} (\varepsilon - \|w\|_2)_+ e^{-iwx}\, dw$$

$$= \varepsilon \int_{-\varepsilon}^{\varepsilon} e^{-iwx}\, dw - \int_{-\varepsilon}^{\varepsilon} \|w\|_2 e^{-iwx}\, dw$$

ここで、

$$\int_{-\varepsilon}^{\varepsilon} e^{-iwx}\, dw = \int_{-\varepsilon}^{\varepsilon} \cos(wx)\, dw - i\int_{-\varepsilon}^{\varepsilon} \sin(wx)\, dw$$

$$= 2\int_0^{\varepsilon} \cos(wx)\, dw = \frac{2}{x}\Big[\sin(wx)\Big]_0^{\varepsilon}$$

$$= \frac{2}{x}\sin(x\varepsilon)$$

$$\int_{-\varepsilon}^{\varepsilon} \|w\|_2 e^{-iwx}\, dw = \int_{-\varepsilon}^{\varepsilon} |w| \cos(wx)\, dw - i\int_{-\varepsilon}^{\varepsilon} |w| \sin(wx)\, dw$$

$$= 2\int_0^{\varepsilon} w\cos(wx)\, dw$$

$$= 2\left(\left[\frac{w}{x}\sin(wx)\right]_0^{\varepsilon} - \int_0^{\varepsilon} \frac{1}{x}\sin(wx)\, dw\right)$$

$$= 2\,\frac{\varepsilon}{x}\sin(x\varepsilon) + \frac{2}{x}\,\frac{1}{x}\Big[\cos wx\Big]_0^{\varepsilon}$$

$$= \frac{2\varepsilon}{x}\sin(x\varepsilon) + \frac{2}{x^2}(\cos(x\varepsilon) - 1)$$

よって、

$$\frac{1}{2\pi}\int_{-\varepsilon}^{\varepsilon} (\varepsilon - \|w\|_2)_+ e^{-iwx}\, dw = \frac{1}{\pi x^2}(1-\cos(x\varepsilon))$$ ∎

81. まず，指数型は，定義より無限次元多項式カーネルである。よって系3より，
E の各コンパクト集合における普遍カーネルである。

また，三角分布に基づく特性カーネルの台は，$(-a, a)$ であって，E がその台以外を含むコンパクト集合であるとき，このカーネルは普遍的でない。

82.

定義　　　　　　カーネルトリック

$$R_N(\mathcal{F}) = \mathbb{E}_\sigma\left[\sup_{f\in\mathcal{F}}\left|\frac{1}{N}\sum_{i=1}^{N}\sigma_i f(x_i)\right|\right] = \mathbb{E}_\sigma\left[\sup_{f\in\mathcal{F}}\left|\frac{1}{N}\sum_{i=1}^{N}\sigma_i\langle k(x_i,\cdot), f(\cdot)\rangle_H\right|\right]$$

H-内積を外に出しただけ

$$= \mathbb{E}_\sigma\left[\sup_{f\in\mathcal{F}}\left|\left\langle f, \frac{1}{N}\sum_{i=1}^{N}\sigma_i k(x_i,\cdot)\right\rangle_H\right|\right]$$

$\langle a, b\rangle_H \leq \|a\|_H\|b\|_H$

$$\leq \mathbb{E}_\sigma\left[\sup_{f\in\mathcal{F}}\|f\|_H\sqrt{\left\langle \frac{1}{N}\sum_{i=1}^{N}\sigma_i k(x_i,\cdot), \frac{1}{N}\sum_{i=1}^{N}\sigma_i k(x_i,\cdot)\right\rangle_H}\right]$$

$\|f\|_H \leq 1$

$$\leq \mathbb{E}_\sigma\left[\sqrt{\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sigma_i\sigma_j k(x_i,x_j)}\right]$$

$\mathbb{E}[\sigma_i\sigma_j] = \delta_{i,j}$

$$\leq \sqrt{\mathbb{E}_\sigma\left[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}k(x_i,x_j)\right]} \leq \sqrt{\frac{k_{max}}{N}}$$

k_{max} の定義

83.

$$\mathbb{E}_{X,Y} \sup_{f \in \mathcal{F}} \left| \mathbb{E}_{X'} \left\{ \frac{1}{N} \sum_{i=1}^{N} f(x_i') - \frac{1}{N} \sum_{i=1}^{N} f(x_i) \right\} - \mathbb{E}_{Y'} \left\{ \frac{1}{N} \sum_{j=1}^{N} f(y_j') - \frac{1}{N} \sum_{j=1}^{N} f(y_j) \right\} \right|$$

Jensen の不等式

$$\leq \mathbb{E}_{X,Y,X',Y'} \sup_{f \in \mathcal{F}} \left| \frac{1}{N} \sum_{i=1}^{N} f(x_i') - \frac{1}{N} \sum_{i=1}^{N} f(x_i) - \frac{1}{N} \sum_{i=1}^{N} f(y_i') + \frac{1}{N} \sum_{i=1}^{N} f(y_i) \right|$$

σ_i, σ_i' は独立に等確率に ±1, σ, σ' で平均をとる

$$= \mathbb{E}_{X,Y,X',Y',\sigma,\sigma'} \sup_{f \in \mathcal{F}} \left| \frac{1}{N} \sum_{i=1}^{N} \sigma_i \{ f(x_i') - f(x_i) \} + \frac{1}{N} \sum_{i=1}^{N} \sigma_i' \{ f(y_i') - f(y_i) \} \right|$$

三角不等式

$$\leq \mathbb{E}_{X,X',\sigma} \sup_{f \in \mathcal{F}} \left| \frac{1}{N} \sum_{i=1}^{N} \sigma_i \{ f(x_i') - f(x_i) \} \right| + \mathbb{E}_{Y,Y',\sigma'} \sup_{f \in \mathcal{F}} \left| \frac{1}{N} \sum_{j=1}^{n} \sigma_j' \{ f(y_j') - f(y_j) \} \right|$$

命題60の $R(\mathcal{F},P)$ の定義

$$\leq 2[R(\mathcal{F},P) + R(\mathcal{F},Q)]$$

$$\leq 2[(k_{max}/N)^{1/2} + (k_{max}/N)^{1/2}]$$

命題 60 の結果