13장 부록

정리 13-1.2. 증명

입력공간 $\mathcal{X}$ 위에서 정의된 이변량함수 $k(x, x')$가 PDS (positive definite symmetric) 조건을 만족한다고 가정하자. 이때 다음과 같이 함수를 정의한다.

- $\phi(x) : \mathcal{X} \to \mathbb{R}$를 $\phi(x)(\cdot) = k(\cdot, x)$로 정의한다.
- 벡터공간 $\mathbb{H}_0$는 $k(\cdot, x_i)$의 유한 선형결합 형태의 함수들로 구성된다:

$$\mathbb{H}_0 = \left\{ \sum_{i=1}^{m} \alpha_i k(\cdot, x_i) : m \in \mathbb{N},\ \alpha_i \in \mathbb{R},\ x_i \in \mathcal{X} \right\} \tag{1}$$

이제 $\mathbb{H}_0$ 위에 내적 $\langle \cdot, \cdot \rangle$를 다음과 같이 정의한다.
$f(\cdot) = \sum_{i=1}^{m} \alpha_i k(\cdot, x_i)$, $g(\cdot) = \sum_{j=1}^{m'} \beta_j k(\cdot, x'_j)$에 대해

$$\langle f, g \rangle := \sum_{i=1}^{m} \sum_{j=1}^{m'} \alpha_i \beta_j k(x_i, x'_j) = \sum_{i=1}^{m} \alpha_i g(x_i) = \sum_{j=1}^{m'} \beta_j f(x'_j) \tag{2}$$

이 내적은 다음과 같은 재생 속성(reproducing property)을 갖는다:

$$f(x) = \langle \phi(x), f \rangle = \langle k(\cdot, x), f \rangle, \qquad \langle k(\cdot, x), k(\cdot, x') \rangle = k(x, x') \tag{3}$$

따라서, 내적의 정의로부터 $\langle \phi(x), \phi(x') \rangle = k(x, x')$가 성립한다.
또한 이 내적은 다음 성질을 만족한다:

- **대칭성(symmetric)**: $\langle f, g \rangle = \langle g, f \rangle$
- **선형성(bilinear)**: $\langle af + bg, h \rangle = a\langle f, h \rangle + b\langle g, h \rangle$

PDS 조건에 따라 다음이 성립한다:

$$\langle f, f \rangle = \sum_{i=1}^{m} \sum_{j=1}^{m} \alpha_i \alpha_j k(x_i, x_j) \geq 0 \tag{4}$$

또한, 임의의 $f_1, \ldots, f_m \in \mathbb{H}_0$와 $c_1, \ldots, c_m \in \mathbb{R}$에 대해

$$\sum_{i,j=1}^{m} c_i c_j \langle f_i, f_j \rangle = \left\langle \sum_{i=1}^{m} c_i f_i, \sum_{j=1}^{m} c_j f_j \right\rangle \geq 0 \tag{5}$$

이므로, $\langle \cdot, \cdot \rangle$는 $\mathbb{H}_0$에서 정의된 양의정부호 대칭(PDS) 내적이다.

이제 코시-슈바르츠 부등식을 이용하면 다음이 성립한다:

$$f(x)^2 = \langle f, \phi(x) \rangle^2 \leq \langle f, f \rangle \cdot \langle \phi(x), \phi(x) \rangle = \langle f, f \rangle \cdot k(x, x) \tag{6}$$

따라서 $\langle\cdot,\cdot\rangle$는 $\mathbb{H}_0$ 위에서 내적공간의 정의를 만족한다.

이제 완비성(completeness)을 보장하기 위해, 고정된 입력 $x \in \mathcal{X}$에 대해, $\mathbb{H}_0$에 속하는 함수열 f_n이 코시 수열(Cauchy sequence)라고 하자. 즉 $f_n(\cdot), f_m(\cdot)$에 대해

$$f_n(x) - f_m(x) = \langle f_n - f_m, k(x,\cdot)\rangle \tag{7}$$

이때 코시-슈바르츠 부등식을 적용하면

$$(f_n(x) - f_m(x))^2 \le \|f_n - f_m\|^2 \cdot \langle k(x,\cdot), k(x,\cdot)\rangle = \|f_n - f_m\|^2 \cdot k(x,x) \tag{8}$$

따라서 $f_n(x)$는 수렴하는 실수 수열이며, 이를 통해 점별 수렴하는 극한 함수 $f(x) = \lim_{n\to\infty} f_n(x)$가 존재한다.
이와 같이 $\mathbb{H}_0$에 속한 모든 코시 수열의 극한 함수를 추가하여 얻는 공간을 완비화(completion)된 힐베르트 공간이라 하고, 이를 다음과 같이 정의한다:

$$\mathbb{H} = \overline{\mathrm{span}\{k(x,\cdot) : x \in \mathcal{X}\}} \tag{9}$$

이 공간 $\mathbb{H}$는 함수 $f(x)$의 값을 $k(x,\cdot)$와의 내적으로 재현(reproduce)할 수 있는 구조를 가지며, 이로 인해 **재생 커널 힐베르트 공간(Reproducing Kernel Hilbert Space; RKHS)**이라 불린다.
이상으로, 정리 13-1.2의 증명이 완료된다. ¶

보충설명: 13.4.2절

Triangular Gap 기반 커널 변환 정리 증명[1]

대칭 함수 $D : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ 에 대해, 임의의 기준점 $x_0 \in \mathcal{X}$ 를 고정하고 다음과 같이 대칭 삼각형 간격 (**symmetric triangular gap**) 커널 k_D 를 정의하자:

$$k_D(x,y) := D(x,x_0) + D(y,x_0) - D(x,y) - D(x_0,x_0) \tag{10}$$

그러면 다음이 성립한다:

$$D \text{ is NDS} \iff k_D \text{ is PDS} \tag{11}$$

증명 스케치:

($\Leftarrow$) 방향: k_D가 PDS이면 D는 NDS

$\sum_i c_i = 0$ 인 임의의 $\mathbf{c} \in \mathbb{R}^N$ 에 대해:

$$\sum_{i,j} c_i c_j D(x_i, x_j) = (\sum_j c_j)(\sum_i c_i D(x_i, x_0)) + (\sum_i c_i)(\sum_j c_j D(x_j, x_0)) \\ -(\sum_i c_i)^2 D(x_0, x_0) - \sum_{i,j} c_i c_j k_D(x_i, x_j) = -\sum_{i,j} c_i c_j k_D(x_i, x_j) \leq 0 \tag{12}$$

k_D가 PDS라면 우변은 음이거나 0이므로 D는 NDS가 된다.

($\Rightarrow$) 방향: D가 NDS이면 k_D는 PDS

임의의 $\mathbf{c} = (c_0, c_1, \ldots, c_N)' \in \Re^N$ 에 대해, $c_0 := -\sum_{i=1}^N c_i$ 라 정의하면:

$$0 \geq \sum_{i,j=0}^N c_i c_j D(x_i, x_j) = (\sum_{j=0}^N c_j)(\sum_{i=0}^N c_i D(x_i, x_0)) + (\sum_{i=0}^N c_i)(\sum_{j=0}^N c_j D(x_j, x_0)) \\ -(\sum_{i=0}^N c_i)^2 D(x_0, x_0) - \sum_{i,j=0}^N c_i c_j k_D(x_i, x_j) = -\sum_{i,j=0}^N c_i c_j k_D(x_i, x_j) \tag{13}$$

위 식을 정리하면:

$$\sum_{i,j=0}^N c_i c_j k_D(x_i, x_j) = \sum_{i,j=1}^N c_i c_j k_D(x_i, x_j) + 2c_0 \sum_{i=0}^N c_i k_D(x_i, x_0) + c_0^2 k_D(x_0, x_0) \geq 0 \tag{14}$$

k_D 정의로부터 $k_D(x, x_0) = D(x, x_0) + D(x_0, x_0) - D(x, x_0) - D(x_0, x_0) = 0$ 이므로

$$\sum_{i,j=1}^N c_i c_j k_D(x_i, x_j) \geq 0 \tag{15}$$

가 되어 k_D 가 PDS가 된다.

$$\sum_{i,j=1}^N c_i c_j k_D(x_i, x_j) \geq 0 \tag{16}$$

즉, k_D는 PDS가 된다. ¶

각주

---

1. [Mohri (2012)] M. Mohri, A. Rostamizadeh, and A. Talwalkar Foundation of Machine Learning, MIT Press, 2012. ↩