ヒルベルト空間における Maximum Mean Discrepancy の一致推定量の漸近正規性

千葉大・融合理工学府　牧草 夏実
千葉大・理学研究院　内藤 貫太

ヒルベルト空間 $\mathcal{H}$ 上の分布 Q を既知の分布とし，P を未知の分布とする．正定値カーネルを k とし，k に対応する再生核ヒルベルト空間を $H(k)$ とする．k を characteristic kernel とし，$\mathbb{E}_{X\sim P}[\sqrt{k(X,X)}] < \infty$ かつ $\mathbb{E}_{X\sim Q}[\sqrt{k(X,X)}] < \infty$ を仮定することで，$\mathcal{H}$ 上の 2 つの確率分布 P と Q の違いは，Maximum Mean Discrepancy (MMD) と呼ばれる距離

$$\mathrm{MMD}(P,Q) = ||\mu(P) - \mu(Q)||_{H(k)}$$

によって測ることができる．ここで，$\mu(P) = \mathbb{E}_{X\sim P}[k(\cdot, X)]$，$\mu(Q) = \mathbb{E}_{X\sim Q}[k(\cdot, X)]$ はそれぞれ確率分布 P，Q のカーネル平均埋め込みと呼ばれるものである．$D^2 = \mathrm{MMD}(P,Q)^2$ のナイーブな推定量は，$\mathcal{H}$ におけるデータ $X_1, \ldots, X_n \overset{i.i.d.}{\sim} P$ を用いて，

$$\widehat{D}^2 = \left|\left|\frac{1}{n}\sum_{i=1}^{n} k(\cdot, X_i) - \mu(Q)\right|\right|_{H(k)}^2 = \frac{1}{n^2}\sum_{i,j=1}^{n} k(X_i, X_j) - \frac{2}{n}\sum_{i=1}^{n}\mu(Q)(X_i) + ||\mu(Q)||_{H(k)}^2$$

によって構成できる．この $\widehat{D}^2$ は $P \neq Q$ のもとでは漸近分布は正規分布になり，$P = Q$ のもとでは退化 V 統計量であるので，漸近分布は重み付きカイ二乗分布の無限和になる．この $P = Q$ のもとでの重みを陽に求めることは非常に困難となっていた．

このヒルベルト空間での分布の距離 MMD に対し，L_2 空間での仮説の分布とデータを出した分布の L_2 距離の推定量として，Cramér-von Mises 統計量が知られている．Cramér-von Mises 統計量も同様に，帰無仮説のもとでは漸近分布が重み付きカイ二乗分布の無限和であったが，Cramér-von Mises 統計量については，漸近正規性をもつ一致推定量への ‘修正方法’ がすでに議論されている．

この D^2 のナイーブな推定量 $\widehat{D}^2$ に対して， Cramér-von Mises 統計量の場合と同様に推定量のクロスタームに重みをかけた，

$$\widehat{D}_\gamma^2 = \frac{1}{n^2}\sum_{i,j=1}^{n} k(X_i, X_j) - \frac{2}{n}\sum_{i=1}^{n} K_{i,n}(\gamma)\mu(Q)(X_i) + ||\mu(Q)||_{H(k)}^2$$

によって D^2 の一致推定量を構築する．ただし，$\{K_{i,n}(\gamma)\}_{i=1}^n$ は適当な条件を満たす重みである．

本発表では，この $\widehat{D}_\gamma^2$ が漸近正規性をもつ D^2 の一致推定量であることについて報告を行う．また，このような修正方法は，未知パラメータを含む正規分布族からの MMD についても同様に適用可能であることを示す．