

# DensityRatioEstimation

大下 範晃

February 2019

## 1 イントロダクション

$\mathcal{X} \subset \mathbf{R}^d$  はデータドメインとする

$$\{x_i^{\text{nu}}\}_{i=1}^{n_{\text{nu}}} \overset{\text{i.i.d.}}{\sim} p_{\text{nu}}^*(x) \quad \text{and} \quad \{x_j^{\text{de}}\}_{j=1}^{n_{\text{de}}} \overset{\text{i.i.d.}}{\sim} p_{\text{de}}^*(x) \quad (1)$$

ドメイン  $\mathcal{X}$  に対して  $p^*(x)$  は完全に正と仮定する.  
つまり,  $p^*(x) > 0, x \in \mathcal{X}$   
ゴールは以下に示す密度比を推定すること

$$r(x)^* = \frac{p_{\text{nu}}^*(x)}{p_{\text{de}}^*(x)} \quad (2)$$

### 1.1 基本的枠組み

次のサンプルを仮定する.

$$\{x_k\}_{k=1}^n \overset{\text{i.i.d.}}{\sim} p^*(x) \quad (3)$$

密度推定のゴールは  $\{x_k\}_{k=1}^n$  からの真の密度  $p(x)^*$  の推定量  $\hat{p}(x)$  を取得すること.

簡易的な方法

$$\hat{r} = \frac{\hat{p}_{\text{nu}}(x)}{\hat{p}_{\text{de}}(x)} \quad (4)$$

として求める.

### 1.2 密度推定の統一化

分母密度  $p_{\text{de}}^*(x)$  が一様だった時、密度比  $r^*(x)$  は分子密度  $p_{\text{nu}}^*(x)$  に比例する.  
これを数式で書けば

$$r^*(x) \propto p_{\text{nu}}^*(x) \quad (5)$$

である.

したがって、分母密度が一様になるように入力変数  $x$  を他のドメインに変換すると、標準密度推定法によって密度比を推定することができる。

入力ドメインの次元数が有限領域  $[a, b]$  で  $d = 1$  の時  $-\infty < a < b < \infty$  では  $p_{de}^*$  の累積分布関数  $p_{de}^*(x)$  を使用して上記のアイデアを実装できる。

$$p_{de}^* := \int_a^x p_{nu}^*(x) dx \quad (6)$$

## References

- [1] DensityRatioEstimation in Machine Learning Masashi Sugiyama Taiji Suzuki Takafumi Kanamori