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

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

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

$$r(x)^* = \frac{p(x)_{nu}^*}{p(x)_{de}^*} \tag{2}$$

1.1 基本的枠組み

次のサンプルを仮定する．

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

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 $p(\hat{x})$ を取得すること．

簡易的な方法

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

として求める．

References

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