演習問題

@amoO_O

2.61 （基本）K 近傍密度モデルでは，全空間上での積分が発散する変則分布になることを示せ．

データ点を x_i とする．

$$R = \max(\|x_i\|)$$

とおく．三角不等式より，

$$\|x - x_i\| \leq \|x\| + \|x_i\| \leq \|x\| + R$$

が成り立つ．

(2.246)より，K 近傍密度モデル p(x)は，x を中心とする小球の体積 V に反比例する．V はちょうど K 個のデータ点が入るまで拡大した球だから，次の不等式が成り立つ．

$$\frac{1}{p(x)} \propto V \leq \frac{S_D}{D}(\max\|x - x_i\|)^D \leq \frac{S_D}{D}(\|x\| + R)^D$$

$$\mathrm{p(x)} \geq \frac{const}{V} \geq \frac{const}{(\|x\| + R)^D}$$

この不等式を使って，全空間上で積分をすると，

$$\int_{\mathbf{R}^{\mathrm{D}}} p(x)dx \geq \int_{\mathbf{R}^{\mathrm{D}}} \frac{const}{(\|x\| + R)^D} dx$$

$$\geq const\, S_D \int_0^\infty \frac{r^{D-1}}{(r + R)^D} dr$$

右辺の積分を s=r+R と変数変換すると，

$$\int_{\mathrm{R}}^\infty \frac{(s - R)^{D-1}}{s^D} ds = \int_R^\infty \left(\frac{1}{s} - (D-1)\frac{R}{s^2} + \cdots\right) ds$$

$$= [\ln(s)]_R^\infty + const = \infty$$

以上より，K 近傍密度モデルは変則分布であることが示された．