1.6 情報理論

永田 純平

1.6.1　エントロピー

　乱雑さ（確率変数がどの値をとるかを言い当てにくい状態）を計るための尺度

　離散確率分布Pに対して以下のようなエントロピーH(P)が定義される。

$$H(P) = \sum_{x} -P(X = x) \log P(X = x | Y = y)$$

条件付エントロピー

ある事象が生じたという条件のもとでのエントロピー。

Y=yなる条件のもとでXの条件付きエントロピーをH(X|Y=y)とあらわす。

$$H(X|Y=y)$$
$$= -\sum_{x} P(X=x|Y=y)\log P(X=x|Y=y)$$

1.6.2 カルバックライブラーダイバージェンス

二つの確率分布に対しての、それらの間の異なり具合を表す。

$$D_{KL}(P||Q) = \sum_{x} P(X = x) \log \frac{P(X = x)}{Q(X = x)}$$

※二つの確率変数は同じ事象空間の確率分布である。

※P(x)>0,Q(x)=0のときはKLダイバージェンスは定義することができない。

1.6.3 ジェンセンシャノンダイバージェンス

$$D_{JS}(P||Q) = \frac{1}{2}(D_{KL}(P||R) + D_{KL}(Q||R))$$
$$= \frac{1}{2}(\sum_x P(x)\log\frac{P(x)}{R(x)} + \sum_x Q(x)\log\frac{Q(x)}{R(x)})$$
$$= \frac{1}{2}(\sum_x P(x)\log\frac{P(x)}{\frac{P(x)+Q(x)}{2}} + \sum_x Q(x)\log\frac{Q(x)}{\frac{P(x)+Q(x)}{2}})$$

以上よりJSダイバージェンスは対称性 $D_{JS}(P||Q) = D_{JS}(Q||P)$ を満たす。
※Rは$R(x) = (P(x) + Q(x))/2$で定義される。

1.6.4 自己相互情報量

確率変数の相互依存の尺度

$$PMI(x,y) = \log \frac{P(x,y)}{P(x)P(y)}$$

$$P(x,y) = P(x)P(y) \rightarrow PMI(x,y) = 0$$
$$P(x,y) > P(x)P(y) \rightarrow PMI(x,y) > 0$$
$$P(x,y) < P(x)P(y) \rightarrow PMI(x,y) < 0$$

同時確率P(x,y)から各確率変数での出現影響を差し引くことでより正確な共起を測っている。

1.6.5 相互情報量

$$MI(X,Y) = \sum_{x,y} P(x,y) \log \frac{P(x,y)}{P(x)P(y)}$$

$$MI(X,Y) = D_{KL}(P(X,Y)||P(X)P(Y))$$