独立成分分析ゼミ

3　Infomax, KL-divergence, 自然勾配法

発表日 : 2007/10/17

発表者 : 阿部 竜之介

エントロピー

エントロピー

■確率変数の曖昧さの尺度

定義

$$H(X_1, X_2) = E_{P_{X_1X_2}}[-\log P_{X_1X_2}(X_1, X_2)]$$
$$= -\int P_{X_1X_2}(x_1, x_2)\log P_{X_1X_2}(x_1, x_2)dx_1dx_2$$

※確率変数 X の確率密度を $P_X(x)$ とする
　結合分布の密度関数: $P_{X_1X_2}(x_1, x_2)$

条件付きエントロピー

条件付きエントロピー

- X_1を知った時に残るX_2の曖昧さ
- 条件付きエントロピーが小さいということは従属性を表す

⇒ 独立な変数が欲しい場合は大きければいい

定義

$$H(X_2 \mid X_1) = \int H(X_2 \mid x_1) P_{X_1}(x_1) dx_1$$

$$= -\int P_{X_1}(x_1) P_{X_2|X_1}(x_2 \mid x_1) \log P_{X_2|X_1}(x_2 \mid x_1) dx_2 dx_1$$

相互情報量

<u>相互情報量</u>

- 独立性を測る量
- X_1 に関する曖昧さから X_2 を知ったあとに残る X_1 の曖昧さを引いたもの

定義　$I(X_1;X_2)=H(X_1)-H(X_1\,|\,X_2)$

KL-ダイバージェンス

■相互情報量を密度関数で書き直すと、

$$I(X_1;X_2)=E_{P_{X_1X_2}}\left[\log\frac{P_{x_1}(X_1,X_2)}{P_{X_1}(X_1)P_{X_2}(X_2)}\right]$$

　これは結合分布と各周辺分布の積の間の距離KL-ダイバージェンスと呼ばれる量で測ったものになっている

自然勾配法

yの各成分が独立になるようなWを求めるには

■各成分間の相互情報量を計算し、それを最小化するようなWを求める

■自然勾配に基づく最急降下法を用いる

→ 自然勾配法

最急降下方向

■εを小さな数として、WをW+εUと変化させたときに一番 $\psi(W+\varepsilon U)-\psi(W)$ を小さくする方向

$\psi(W)$:スカラー関数

自然勾配法

- 自然勾配（計量がある場合の本来のグラジエント）

$$U = -\frac{\partial \psi(W)}{\partial W} W^T W$$

- この関係を使い、

$$\Delta W = -\frac{\partial I(Y)}{\partial W} W^T W$$

$$= -\{W^{-T} - E_{P_X}[\varphi(y) x^T]\} W^T W$$

$$= -(I - E_{P_Y}[\varphi(Y) Y^T]) W$$

としてWを変化させていけば良い

Infomax

Infomax

観測値Xをある通信路Gに通してYを得たとし、

$$Y = G(X) + N$$

と書くことにする。Nは通信の途中に入ってくる雑音。

- このYが持っているXの情報量を最大化するように通信路を設計してやればよいという考え方
- 雑音を考えずに最急降下法によってこれを求めるアルゴリズムを考えると、自然勾配に基づく最急降下法と同じになる