カーネル多変量解析

第3.5章
カーネル独立成分分析

新納浩幸

独立成分分析(1)

d 個の信号発生源

$$s_1(t), s_2(t), \cdots, s_d(t)$$

d 本のマイクへの入力

$$x_1(t), x_2(t), \cdots, x_d(t)$$

$$x_i(t) = \sum_{j=1}^{d} A_{ij} s_j(t)$$

$A = (A_{ij})$ は未知

$$\mathbf{x}(t) = (x_1(t), x_2(t), \cdots, x_d(t))^T$$ から

$$\mathbf{s}(t) = (s_1(t), s_2(t), \cdots, s_d(t))^T$$ を復元

独立成分分析(2)

仮定　$s_i(t)$ は互いに独立

注意)　上記の条件だけでは、$s_i(t)$ と $s_j(t)$ を入れ替えてもOK

$s_i(t)$ をスカラー倍してもOK

$\mathbf{s}(t)$ の推定のために以下を考える

$$\mathbf{y}(t) = W\mathbf{x}(t)$$

無相関からの推定(1)

$$X = (\mathbf{x}(1), \mathbf{x}(2), \cdots, \mathbf{x}(n))^T$$

特異値分解

$$X = U\Lambda V^T \quad \text{s.t.} \quad U^T U = V^T V = V V^T = I_d$$

$$\Rightarrow \quad W^T = \frac{V}{\sqrt{n}}$$

$$W X^T = X W^T$$

$$= \frac{1}{\sqrt{n}} U\Lambda V^T V = \frac{1}{\sqrt{n}} X V = \frac{1}{\sqrt{n}} U\Lambda$$

無相関からの推定(2)

得られた成分の分散を1にスケーリング　(白色化)

$$\Rightarrow \quad \mathbf{y}(t)$$

注意) 得られた成分は無相関、しかし無相関でも
独立とは限らない、逆は成立

カーネルを利用した独立性の規準(1)

信号自身の相関　$Cor[y_i(t), y_j(t)]$

上記の相関が 0 でも独立とは言えない

信号の関数どうしの相関

$$\rho = Cor[f(y_i(t)), g(y_j(t))]$$

どんな関数 f と g でもρが 0

$s_i(t)$ は互いに独立

カーネルを利用した独立性の規準(2)

$$\rho = Cor[f(y_i(t)), g(y_j(t))]$$

どんな関数 f と g でもρが 0

⇒ 任意の関数 f と g に対してρの最大値が小さい

⇒ f と g をカーネル関数の線形和に限定

$$f(x) = \sum_j \alpha_j k(x_j, x)$$

⇒ ρを最大にするf と g はカーネル正準相関分析から求まる

カーネル独立成分分析

カーネル独立成分分析の問題

・最終の W を求めるのは計算が大変

・「すべての関数」を「カーネルの線形和」に限定したことで、独立性の保証が崩れる?

ほぼ大丈夫

カーネル関数の線形和でかける f と g について

$$\max_{ij} \rho = 0$$

は独立性と等価

R メモ

R の独立成分分析のパッケージは fastICA

```
> library("fastICA")
> data(iris)
> a <- fastICA(iris[,-5],2)
> plot(a$S)
```