TUKR のアルゴリズム

2022 年 5 月 17 日

TUKR のアルゴリズム実装で扱う変数を表 1 に示す.

表 1　TUKR における変数記号表

表記	説明
I	ドメイン 1 のデータ数
J	ドメイン 2 のデータ数
D	データの次元数
$\underline{\mathbf{X}}$	データの集合
	太字下付きは 3 階のテンソルの意味
	$\underline{\mathbf{X}} = (\mathbf{x_{ij}}) \in R^{I \times J \times D}$
	$\mathbf{x_{ij}} = (\mathbf{x_{ij}^1}, ..., \mathbf{x_{ij}^d})^{\mathrm{T}}$
$\underline{\mathbf{Y}}$	観測データの推定値の集合
	$\underline{\mathbf{Y}} = (\mathbf{y_{ijd}}) \in R^{I \times J \times D}$
$\mathbf{Z}$	潜在変数の集合
$\mathbf{U}$	ドメイン 1 の潜在変数の集合
	$\mathbf{U} = (\mathbf{u_{il_1}}) \in R^{I \times L_1}$
$\mathbf{V}$	ドメイン 2 の潜在変数の集合
	$\mathbf{V} = (\mathbf{v_{jl_2}}) \in R^{J \times L_2}$
L_1	ドメイン 1 の潜在変数の次元数
L_2	ドメイン 2 の潜在変数の次元数
T	総学習回数
η_1, η_2	各ドメインの学習率
σ_1, σ_2	各ドメインの平滑化カーネルのカーネル幅

1　TUKR のシミュレーションコードの作成手順

1.1　人口データの作成

鞍型の人工データ $\underline{\mathbf{X}}$ を作成する.

1.2　アルゴリズム部の作成

1.2.1　初期化

潜在変数 $\mathbf{z}$ を乱数によって初期化し、学習をスタートする。その際、配列のサイズに注意する.

1.2.2　目的関数が最小となるように勾配法で潜在変数の更新を学習回数 T 買い繰り返す.

- 写像の推定

$$f(u_i, v_j) = \frac{\sum_i \sum_j k(u, u_i)k(v, v_j)\mathbf{x_{ij}}}{\sum_{i'} \sum_{j'} k(u, u_{i'})k(v, v_{j'})} \tag{1}$$

$$k(\mathbf{u}, \mathbf{u}_i) = exp\left(-\frac{1}{2\sigma^2}\|\mathbf{u} - \mathbf{u}_i\|^2\right) \tag{2}$$

$$k(\mathbf{v}, \mathbf{v}_j) = exp\left(-\frac{1}{2\sigma^2}\|\mathbf{v} - \mathbf{v}_j\|^2\right) \tag{3}$$

- 潜在変数の推定
  誤差関数

$$E = \frac{1}{IJ}\sum_i \sum_j \|\mathbf{x_{ij}} - \mathbf{f}(\mathbf{u_i}, \mathbf{v_j})\|^2 + \lambda R(\mathbf{u_i}) + \lambda R(\mathbf{v_j}) \tag{4}$$

- 事前分布 $R(\mathbf{u_i})$ は潜在変数に対する正則化項である。(λ は正則化項の強さを決めるハイパラメータ).
  ガウス事前分布

$$\sum_i \|\mathbf{u}_i\|^2 = \sum_i \sum_l u_{il}^2 \tag{5}$$

正方形一様分布

$$\sum_i \|\mathbf{u}_i\|_p^p = \sum_i \sum_l u_{il}^p \tag{6}$$

p は十分大きい数（たとえば $p = 10$）
Early stopping

$$\frac{1}{N}\sum_i \|\mathbf{u}_i\|^2 \tag{7}$$

一定値（たとえば 1）に達したら停止する
勾配法による潜在変数の更新

$$\mathbf{u_i}^{new} = \mathbf{u_i}^{new} - \eta_1 \frac{\partial E}{\partial \mathbf{u_i}^{old}} \tag{8}$$

$$\mathbf{v_j}^{new} = \mathbf{v_j}^{new} - \eta_2 \frac{\partial E}{\partial \mathbf{v_j}^{old}} \tag{9}$$

1.2.3 近傍半径

TUKR では (基本的に) 近傍半径 σ は固定し，スケジューリングは行わない．