

概要

- データ生成モデルの学習 (i.e., 密度推定) に特徴量同士の関係性についての事前知識を活用したい
- 特徴量の部分的／あいまいな関係性しかわからない場合、モデルの設計にはそのまま利用しにくい ☹
(例) 知識グラフなどの形で特徴量の関係性が与えられても、どのようなモデルとして反映すればよいかはわからない
- 本研究では、そのような特徴量の関係性を**正則化**のために利用する．そのために独立性尺度を用いる

直接的・詳細な関係性

- $\omega = \text{const. if } V > V_{\text{th}}$
- $P \propto V^3$
- $\omega f(\varepsilon) \propto P$
- $\varepsilon = g(P, \omega, T)$ ……

→ モデリングにそのまま利用できるが
高コスト／未知のため用意しにくい



データ

特徴量①	V	風速
特徴量②	T	気温
特徴量③	P	発電量
特徴量④	ω	回転数
特徴量⑤	ε	翼歪
	:	

部分的／あいまいな関係性

- ① ←(関係あり)→ ③
- ④ ←(関係あり)→ ⑤
- ② ←(関係なし)→ ①
- ⑤ ←(未知)→ ② ……

→ 低コストで用意できるが
既存手法ではそのまま利用しにくい



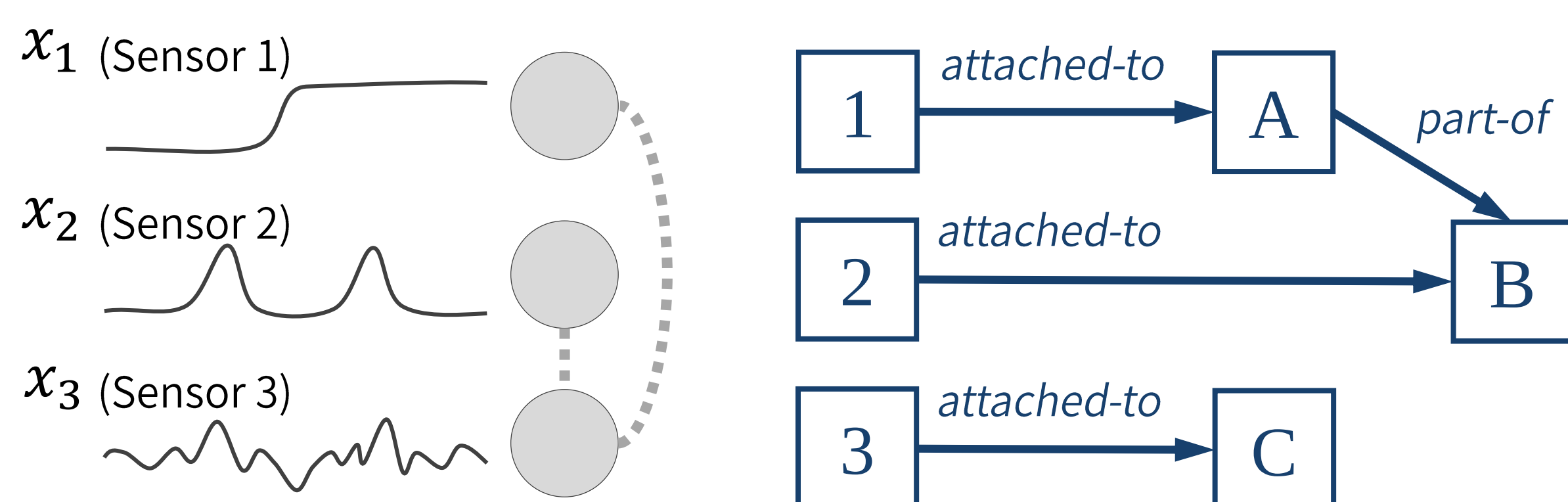
背景

問題設定

- 特徴量 $x_1, \dots, x_D$ からなる D 次元データをもつ
- 特徴量同士の関係性について、次の情報をもつ
 - ① x_i と x_j は独立
 - ② x_i と x_j は (x_i と x_k よりも) 非独立
- このような情報を生成モデルの正則化に利用したい

特徴量の関係性はどのように得られるか？

- side information (の類似度) から
- 知識グラフ (のノードの類似度等) から



提案手法

- θ : 観測モデル (decoder) $p_\theta(x|z)$ のパラメタ
- 周辺分布 $p_\theta(x) = p_\theta(x|z)p(z)$ からのサンプリングが可能であるとする

独立性による正則化項 (x_i と x_j は独立)

$$\text{minimize } \gamma_1 \sum_{\ell} \text{HSIC}(p_\theta(x_{i_\ell}), p_\theta(x_{j_\ell}))$$

非独立性による正則化項 (x_i と x_j は非独立)

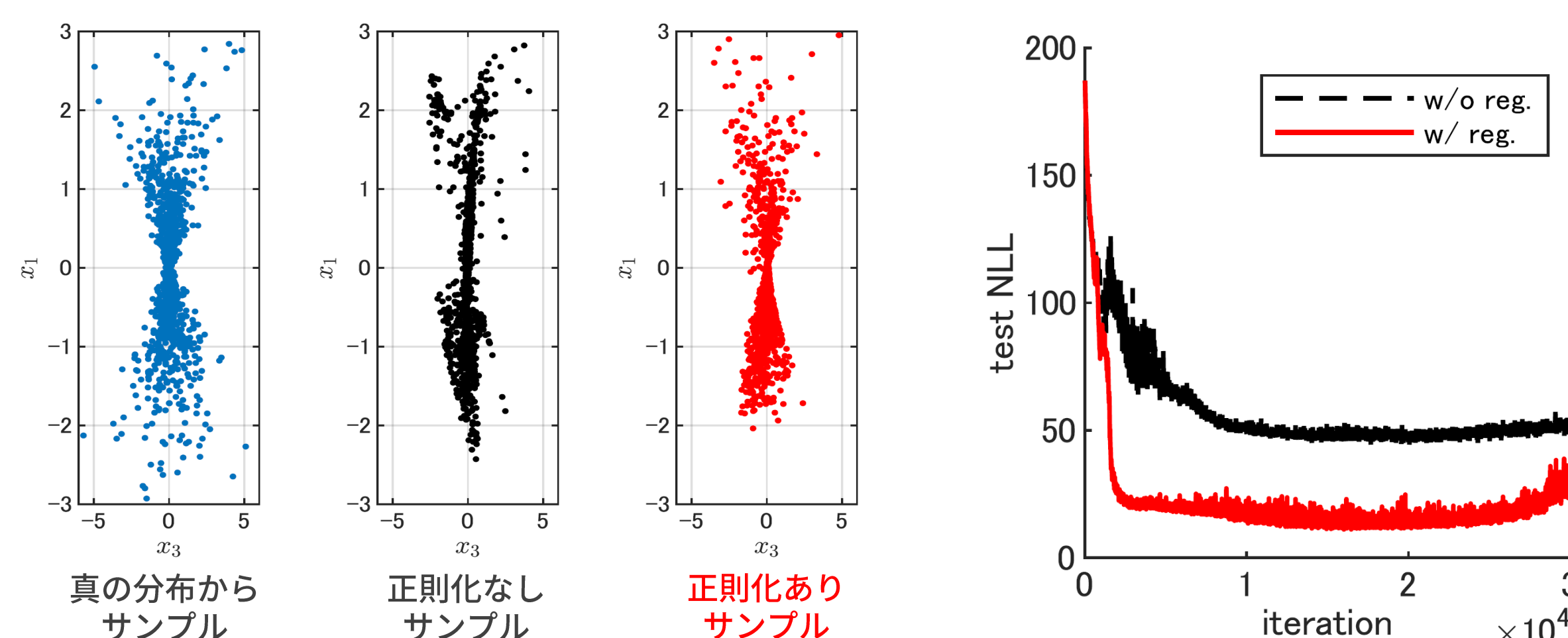
$$\text{min. } \gamma_2 \sum_{\ell} \max(0, \text{HSIC}(x_{i_\ell}, x_{k_\ell}) - \text{HSIC}(x_{i_\ell}, x_{j_\ell}) + \varepsilon)$$

- $p_\theta(x)$ からのサンプルをもとに計算できる
- 非独立性の強さを正しく指定する必要がない
- shift-inv. kernelならRFFで計算を高速化できる

予備実験

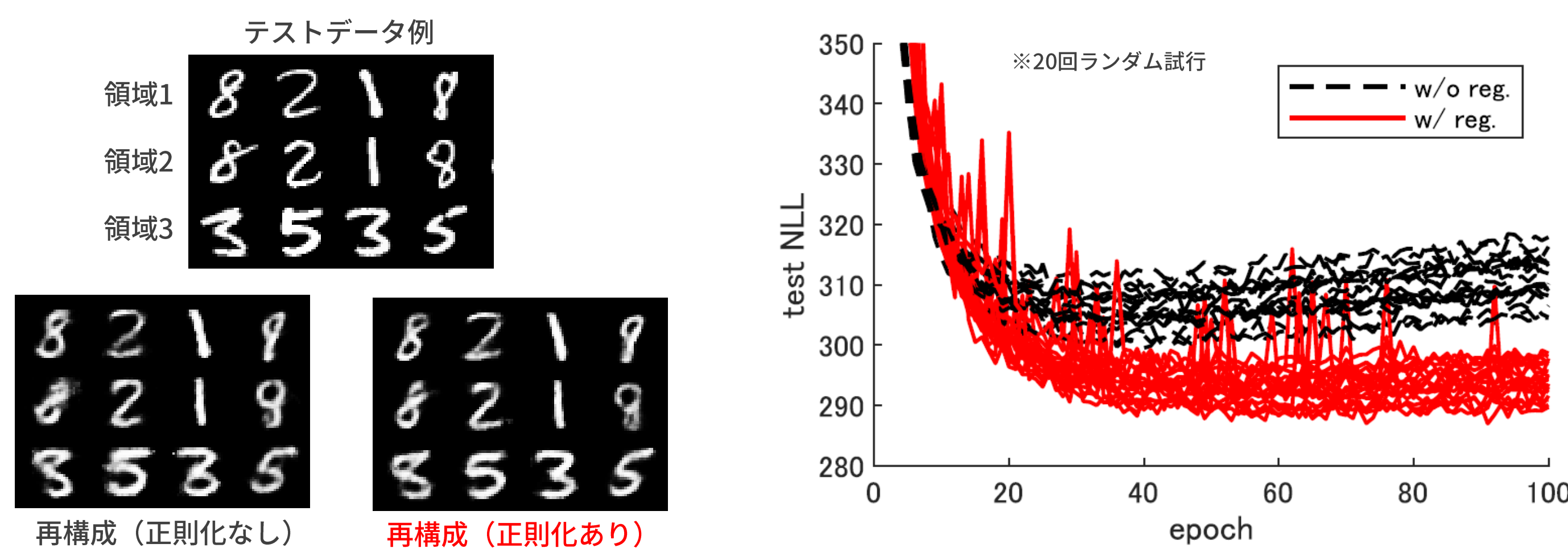
実験 (1) VAE (合成データ)

- 潜在変数 $z_1 \sim \mathcal{N}(0,1)$, $z_2 \sim \mathcal{N}(0,1)$
- 観測 $x_1 = z_1 + e_1$, $x_2 = z_2 + e_2$, $x_3 = z_1 z_2 + e_3$
- 「 x_1 と x_3 は非独立, x_1 と x_2 は独立」を使う



実験 (2) VAE (concatenated MNIST)

- MNISTの3画像を並べて1画像とする (縦に領域1, 2, 3)
- ただし 領域1 と 領域2 は同じラベルから持ってくる
- 「領域1 と 領域2 は非独立」を使う



関連研究 特徴量同士の非独立性に関する情報は feature network と呼ばれ教師あり学習に活用されてきた [Krupka+07,08; Sandler+08; Li&Li08; Li+17; Mollaysa+17]. また、特徴量同士のグラフ構造に基づくスパース正則化が考えられてきた [Huang+11]. さらに、特徴量に関する side information はロバストPCAや協調フィルタリングに利用されてきた [Chiang+16, Porteous+10, Dong+17]. その他に似たような情報を用いる問題設定として、structured prediction などがある. / 独立性尺度をモデル学習の正則化に使う試みはいくつかの文脈でなされてきた. そのひとつは fairness-aware ML で、教師なし学習でのアプローチとしては variational fair autoencoder [Louizous+15] がある. 他には、encoderの出力に任意の独立性を持たせるようにする試みがある [Lopez+18] (一方、本研究ではdecoderの出力に任意の (非) 独立性を持たせようとする).