2023年5月26日　統計数理研究所　オープンハウス

群の作用による系列データからの深層表現学習

福水　健次　　数理・推論研究系　教授

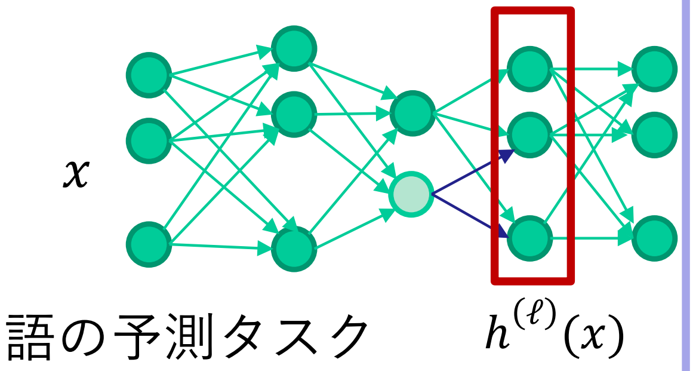
Unsupervised Learning of Equivariant Structure from Sequences
Neurips 2022 with T. Miyato (Preferred Networks/University of Tübingen) and M. Koyama (Preferred Networks)

■ 背景：表現学習，群の作用

● 表現学習

個別タスク専用のニューラルネットを学習させるのでなく，大量のデータを用いて，多くのタスクに有用な表現/特徴ベクトルを学習．（画像，自然言語処理で成功している方法論）

- NNの表現学習
  大量のデータを用いて識別/予測タスクなどを学習（Pretraining）
  - e.g. Large Language Model (GPT4)では，欠損単語の予測タスク



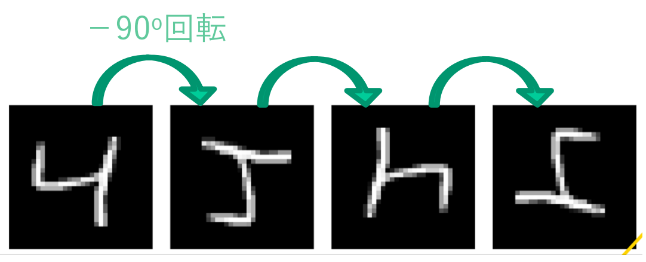
途中の層/モジュールに「意味のある特徴」が現れるはず．
→ 情報の「表現」として用いる

目的
- 他のタスクでもよい性能が得られる．
- 情報の分解

● 対称性の利用　– 不変性, 同変性

「対称性」の数学的定式化：群の作用による不変性
群：「回転」や「折り返し」を一般化したもの
e.g. 可逆行列（掛け算），並べ替え

・群の作用 $G\times\mathcal{X}\to\mathcal{X}$　$(g,x)\mapsto g\cdot x$　（e.g. 行列をベクトルに作用）
群の演算を保つ： $e\cdot x = x$ (e:単位元)，　$(gh)\cdot x = g\cdot(h\cdot x)$



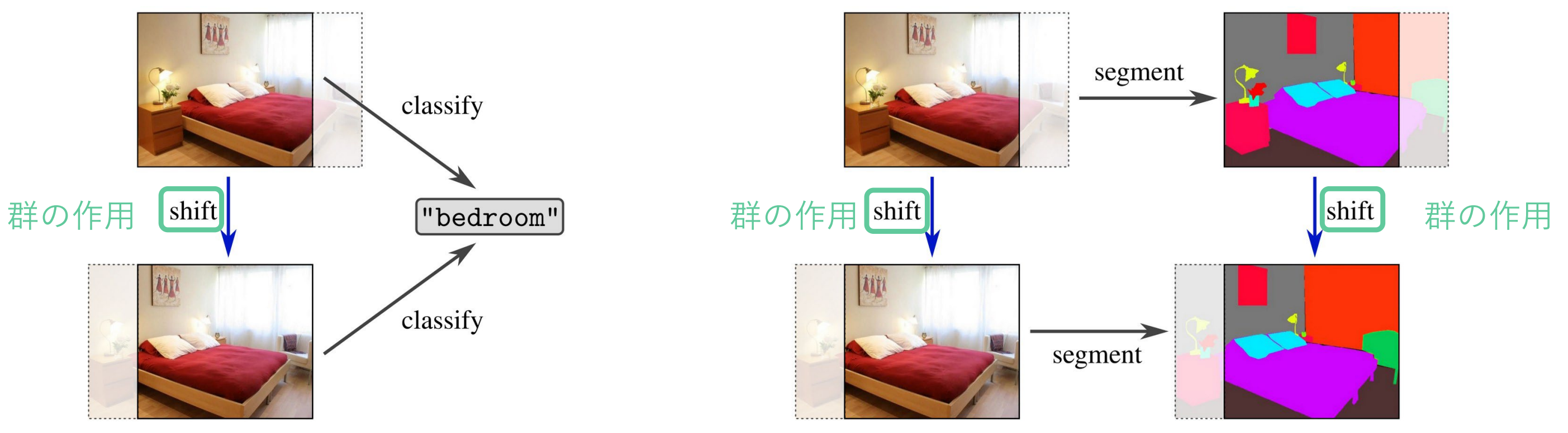
$a = R(-90^o)$
$G = \{1, a, a^2, a^3\}$ の画像への作用

・不変性，同変性：近年の深層学習で有効に利用されている．

同変 (equivariance)
Def. 集合$\mathcal{X}$, $\mathcal{Y}$ に G が作用すると仮定する．
写像 $\varphi:\mathcal{X}\to\mathcal{Y}$ が同変であるとは，
$\varphi(g\circ x) = g\circ\varphi(x)$ $(\forall g\in G, x\in\mathcal{X})$.

$$\begin{tikzcd}
\mathcal{X} \arrow[r, "g\,\circ"] \arrow[d, "\varphi"'] \arrow[dr, phantom, "\circlearrowright"] & \mathcal{X} \arrow[d, "\varphi"] \\
\mathcal{Y} \arrow[r, "g\,\circ"'] & \mathcal{Y}
\end{tikzcd}$$

不変な物体識別　　同変な画像領域分割


classify
群の作用
shift
"bedroom"
classify
segment
群の作用
shift
shift
群の作用
segment

・ Architecture: モデル構造により同変性実現
Convolutional Neural Network (CNN):
シフトに関する同変性　[Fukushima 1980; Zhang 1988; LeCun et al 1989.]
G-convolution NN: **一般の群作用に関する同変性**.
[Cohen & Welling 2016, 2017; Cohen et al 2018.]
E(n)-Equivariant graph NN. [Satorras et al 2021]

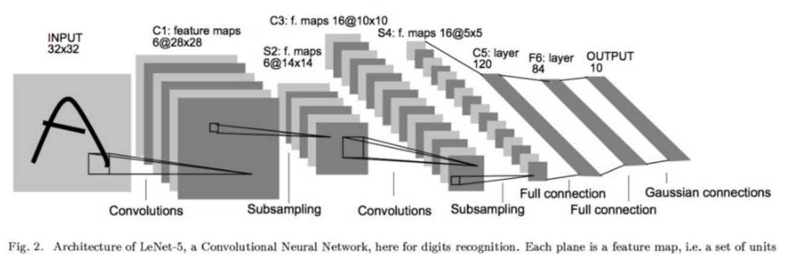
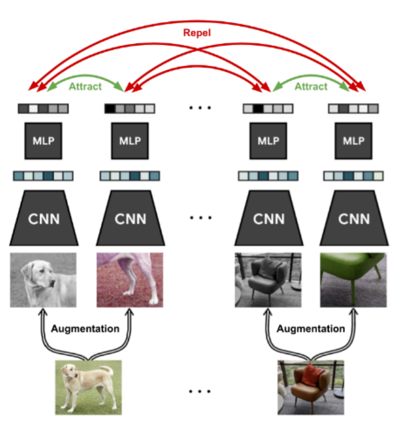

INPUT 32x32
C1: feature maps 6@28x28
S2: f. maps 6@14x14
C3: f. maps 16@10x10
S4: f. maps 16@5x5
C5: layer 120
F6: layer 84
OUTPUT 10
Convolutions
Subsampling
Convolutions
Subsampling
Full connection
Full connection
Gaussian connections

Fig. 2. Architecture of LeNet-5, a Convolutional Neural Network, here for digits recognition. Each plane is a feature map, i.e. a set of units whose weights are constrained to be identical.

LeCun 1998

・Data augmentation： データ拡張に利用
データを変換により拡張して学習する．
自己教師あり学習のキー技術．変換したものは近い表現
SimCLR　[Chen et al 2020].
Contrastive predictive coding [van der Ord et al 2018]　　etc.

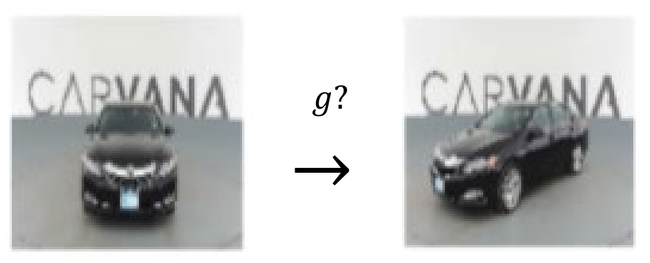

Repel
Attract
Attract
MLP
MLP
MLP
MLP
CNN
CNN
CNN
CNN
Augmentation
Augmentation

SimCLR

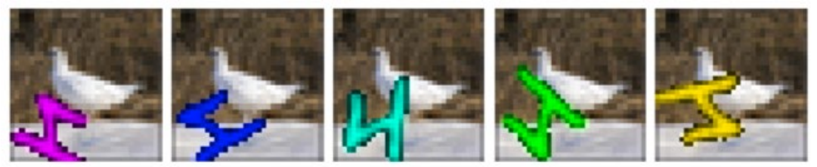
既存のアプローチでは群 G はknown

● 本研究の目的

・潜在的な群作用を用いた表現学習
「動き」の中に有効な特徴が潜在している



・群作用により生成される系列データからの表現学習
・群の作用は潜在的，陽には知らない
・定常的な系列データを仮定
・多数の系列データから表現学習
・群論（表現論）によるdisentanglement

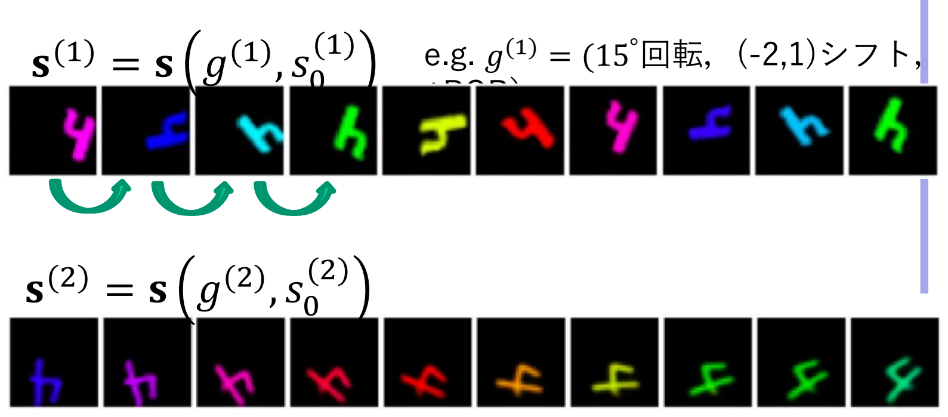


■ 提案法： Meta-Sequential Prediction

● 問題設定

データの生成モデル：群作用による定常な系列データ
G: 群，$\mathcal{S}$ に作用.
系列： $\mathbf{s} = (s_0, s_1, s_2, \ldots, s_T)$.　$s_t = g^t\circ s_0$,　$\mathbf{s} = \mathbf{s}(g, s_0)$
$g\in G$ は系列ごとに異なってよいが系列内では一定.

多数の系列データからの教師無し学習
G も $g\in G$ も作用も知らない


$\mathbf{s}^{(1)} = \mathbf{s}\left(g^{(1)}, s_0^{(1)}\right)$
e.g. $g^{(1)}$ = (15°回転，(-2,1)シフト，
$\mathbf{s}^{(2)} = \mathbf{s}\left(g^{(2)}, s_0^{(2)}\right)$

回転，平行移動，色の回転

初期画像 s_0 と群の元 $g\in G$ を決めると，ひとつの系列 $\mathbf{s}$ が生成される

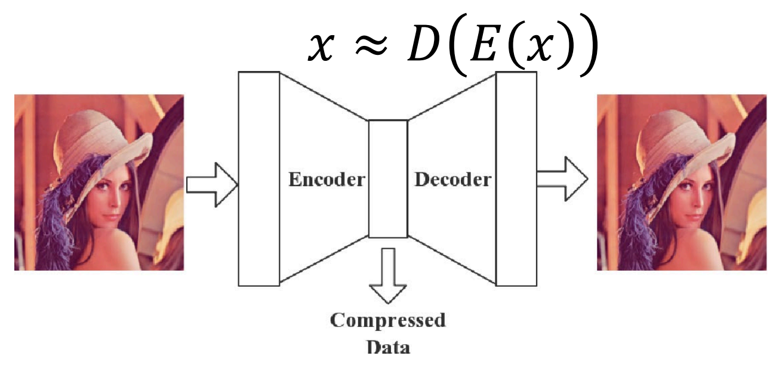
● 基礎モデル：Autoencoder

データを圧縮するモデル

$$\min_{w,\theta}\sum\nolimits_x \left\|x - D_w\left(E_\theta(x)\right)\right\|^2$$

$z = E_\theta(x)$: encoder (NN)
$x' = D_w(z)$: decoder (NN)

$x\approx D\left(E(x)\right)$


Encoder
Decoder
Compressed Data

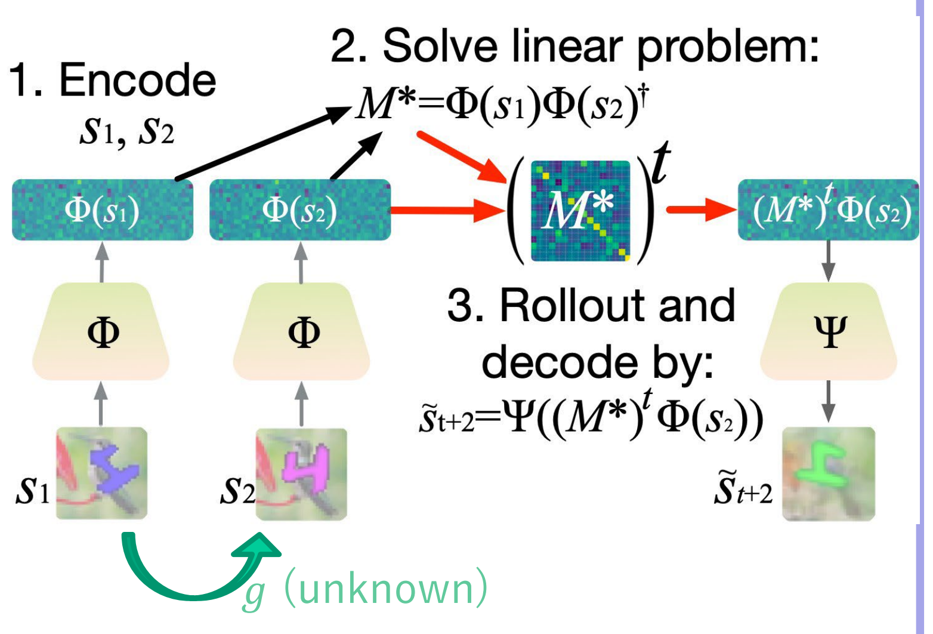
Latent representation（潜在表現）

● AEによる線形な遷移則の学習

$\mathbf{s} = (s_t)_{t=0}^T$.　$s_{t+1} = g\circ s_t$
Autoencoder (AE):
Encoder $\Phi: \mathcal{X}\to\mathbb{R}^{a\times m}$，Decoder $\Psi: \mathbb{R}^{a\times m}\to\mathcal{X}$

潜在表現の遷移則は線形：
行列 $M_{\mathbf{s}}$　系列依存.
潜在空間は行列：
表現能力を落とさずに $M_{\mathbf{s}}$の次元を小さくできる.
$M_{\mathbf{s}}, \Phi, \Psi$ の学習

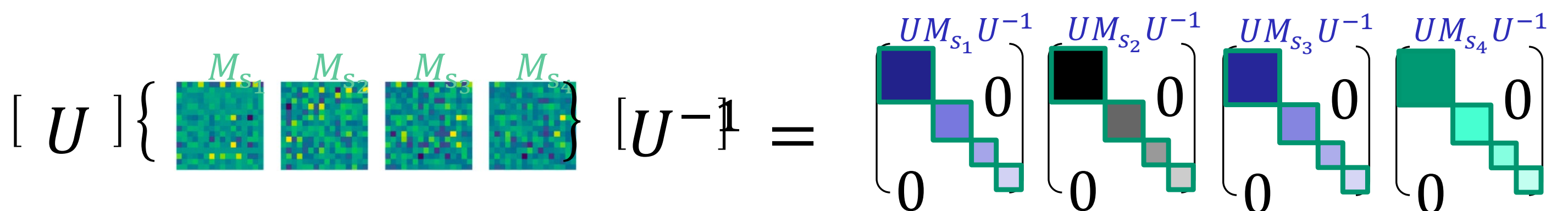
$$M_{\mathbf{s}}\Phi(s_t)\approx\Phi(s_{t+1})\quad[\text{同変性}]$$
$$\Psi\left(M_{\mathbf{s}}^{\ell}\Phi(s_t)\right)\approx s_{t+\ell}\quad[\text{Prediction}]$$


1. Encode
s_1, s_2
2. Solve linear problem:
$M^* = \Phi(s_1)\Phi(s_2)^\dagger$
$\Phi(s_1)$
$\Phi(s_2)$
$\left(M^*\right)^t$
$(M^*)^t\Phi(s_2)$
Φ
Φ
Ψ
3. Rollout and decode by:
$\tilde{s}_{t+2} = \Psi((M^*)^t\Phi(s_2))$
s_1
s_2
$\tilde{s}_{t+2}$
g (unknown)

MSPモデル

● 群の表現論によるdisentanglement

群の表現論（表現の既約分解）→ $\{M_{\mathbf{s}}\}_{\mathbf{s}}$ の同時ブロック対角化が可能
系列に依存しない基底変換により，同時にブロック対角化可能
各ブロックは基本的な線形変換 → 分解された表現.

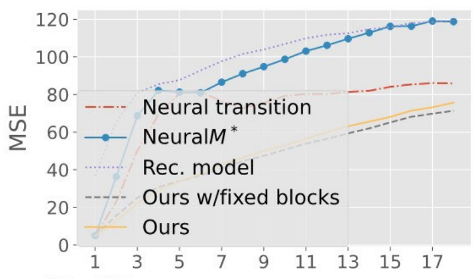

$[\,U\,]$
M_{s_1}
M_{s_2}
M_{s_3}
M_{s_4}
$[U^{-1}]$
=
$UM_{s_1}U^{-1}$
$UM_{s_2}U^{-1}$
$UM_{s_3}U^{-1}$
$UM_{s_4}U^{-1}$
0
0

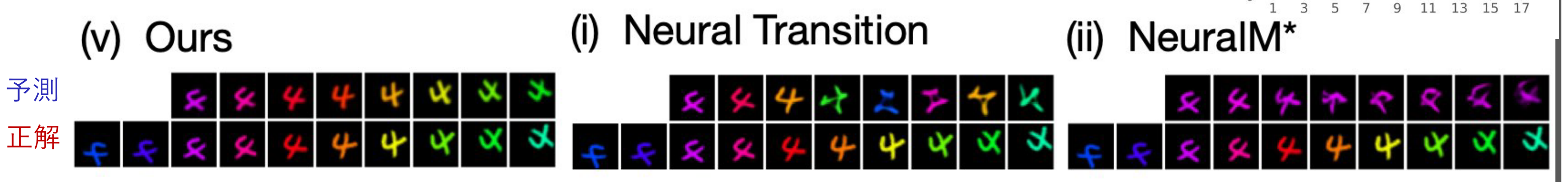
基底変換 U は共通

群の表現論によるDisentanglement

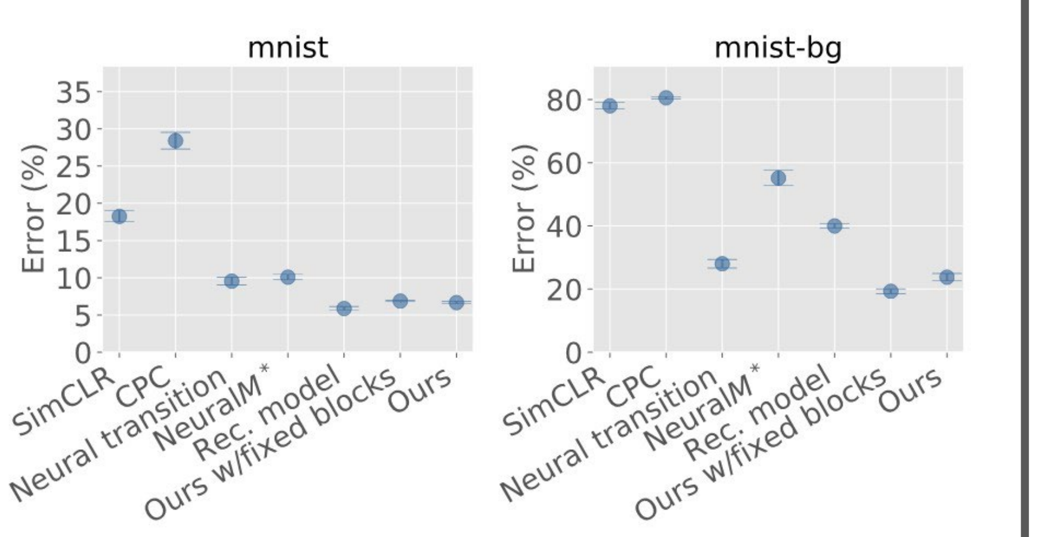
■ 実験

1）予測性能

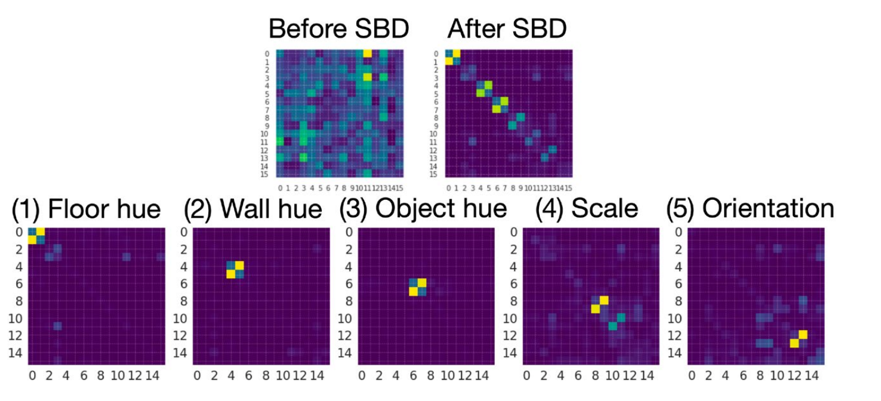



(v) Ours
(i) Neural Transition
(ii) NeuralM*
予測
正解

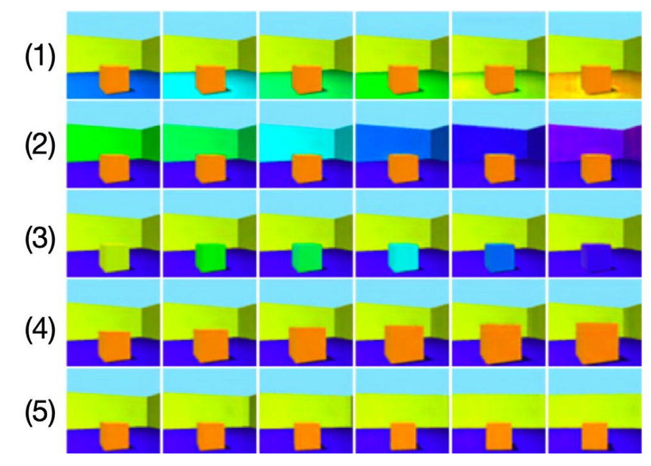
2）獲得された表現による識別性能
潜在変数 $\Phi(s)$ による識別
注）「４」の画像のみで表現学習
→「０」-「９」の１０クラスを線形識別



3）表現論によるDisentanglement



同時ブロック対角化により得られたブロック



各ブロックのみを取り出した M による生成結果

大学共同利用機関法人 情報・システム研究機構
統計数理研究所
The Institute of Statistical Mathematics