

# 非線形次元推定と生成品質評価に基づく VAE ボトルネック次元の検討

## Investigating VAE Bottleneck Dimensions via Nonlinear Intrinsic Dimension Estimation and Generation Quality Assessment

小端 千佳<sup>1)</sup> 代 美月<sup>1)</sup> 神野 健哉<sup>1)</sup>

Obata Chika Mizuki Dai Kenya Jin'no

### 概要

本研究では、多様体仮説に基づき、データ内在次元の推定手法と深層生成モデルのボトルネック次元設定の関連性を実験的に検証した。ベンチマークデータを用いて内在次元推定の精度と安定性を比較し、非線形データには TwoNN および MLE が高精度かつ高安定性を示すことを確認した。次に、MNIST データセットに対し潜在次元を 1~100 で変化させて VAE を学習し、CKA で生成画像を評価することで最適なボトルネック次元を算出した。この結果は「最適な潜在次元は訓練データに内在している」という仮説を示している。

### 1 はじめに

深層学習モデルで高次元データを低次元空間へ写像する際、潜在空間の次元（ボトルネック次元）の選定は経験則に頼らざるを得ない。一方、多様体仮説は次元削減に理論的根拠を与える [1][2]。本稿では、多様体仮説に基づき、「適切な潜在次元は訓練データに内在することの妥当性を実験的に検証する。

### 2 研究目的

本研究の目的は、以下の二点である。

- ベンチマークデータを用いて内在次元推定し、高精度かつ高安定な次元推定手法を特定する。
- MNIST データセットを対象に、潜在次元を 1 から 100 までスweepして学習した Variational Autoencoder (VAE) [3] モデル群を、生成画像品質を CKA で評価し、最適なボトルネック次元を決定する。さらに、第 1 の結果で特定した内在次元推定結果とを検証する。

### 3 多様体仮説の概要

多様体仮説とは、高次元空間に分布するデータが低次元の多様体上に存在するという仮説である。多様体次元を把握できれば、データの本質的な次元でパラメータ空間を適切にモデル化でき、過学習抑制や計算コスト低減

に寄与する。従来の Isomap, LLE, t-SNE, UMAP などは事前に多様体次元を指定する必要があり、最適次元の選定が課題である。

本研究では、まず「内在次元を高精度かつ高安定に推定する手法」を検討し、次にその推定結果を VAE のボトルネック次元設定に応用することで、多様体仮説に基づくモデル設計の有効性を検証する。

### 4 次元推定手法の比較

次元推定実験では、スイスロール、S 字カーブ、およびハイパー平面の三種類のベンチマークデータ (図 1) を用いた。各データについて、真の内在次元を  $d$  とし、ランダムに抽出したデータの 80% を対象に次元推定を 10 回繰り返した。比較対象とした手法は、PCA, TwoNN[4], MLE[5], ESS[6], DANCo[7] の 5 つである (表 1)。各試行における推定結果の中央値  $\hat{d}$  と四分位範囲 (IQR) を算出し、以下の評価基準を定義した。

- $Error = |\hat{d} - d| \leq d \times 0.05$
- $IQR \leq 1$

これら両条件を満たす手法を、「高精度かつ高安定性を有する次元推定手法」と位置付ける。

表 2 に各手法の推定結果を示す。非線形構造を持つスイスロールおよび S 字カーブでは、TwoNN と MLE がいずれも Error および IQR の基準を満たし、最も高い精度と安定性を示した。一方、線形構造を持つハイパー平面では PCA が真の次元を完全に復元し、IQR も 0 となった。ESS は大きな誤差や変動を示し、DANCo は非線形データで推定次元を過大評価した。

以上の結果から、非線形データ一般に対しては TwoNN または MLE が最も有効、線形データに対しては PCA が

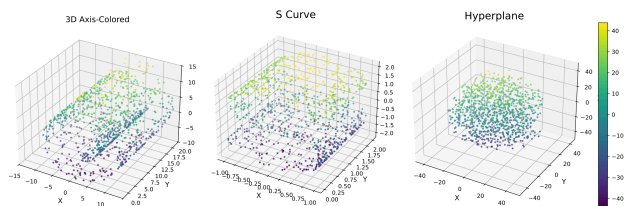


図 1: 使用したベンチマークデータ (左: スイスロール, 中央: S カーブ, 右: ハイパー平面)

1) 東京都市大学, Tokyo City University

表1: 本実験で使用した次元推定手法の概要

手法名	カテゴリ	概要	特徴
PCA	線形	分散最大化	計算高速/大域構造把握
TwoNN	局所距離比	近傍距離比	パラメータ不要/非線形対応
MLE	最大尤度	尤度最大化	統計モデル/高精度
ESS	単体スキュー	Simplexのゆがみ測定	局所構造評価/ノイズ耐性
DANCo	角度+ノルム	角度&ノルム分布統計	非線形多様体全域評価

有効であると結論付けた。本研究ではTwoNNとMLEを高性能手法として採用する。

## 5 ボトルネック次元推定に基づくVAE評価

MNISTデータに対し、潜在次元を1~100に変化させてVAEを学習した。VAEは全結合層と活性化関数から成るエンコーダ/デコーダ構造で、バッチサイズ100、学習率 $1 \times 10^{-3}$ 、エポック50とした。学習後、CKA[8]で入力画像と再構成画像の一致度を測定し、最大となる次元を最適ボトルネック次元とした。その値をTwoNNおよびMLEの推定値と比較する。

実験の結果(図2)から、CKAが最大となる潜在次元数は13であった。この値はTwoNNによる推定値13.34およびMLE推定値11.97に近く、TwoNN推定値が最も良く一致した。以上から、MNISTデータセットでのVAEの最適ボトルネック次元は、訓練データから推定可能であることが示唆された。

## 6 結論と今後の展望

本研究では、多様体仮説に基づき、ベンチマークデータに対する次元推定手法の比較と、推定結果に基づくVAEボトルネック次元の生成品質評価を行った。その結果、TwoNNおよびMLEが非線形データの内在次元推定に高い精度と安定性を示し、さらにMNISTにおいてはTwoNN推定値に基づく13次元がVAE生成品質の最適点と一致した。これにより、「深層生成モデルの適切

表2: 各手法の推定結果

Dataset	Estimator	Median	IQR	Error
Swiss Roll	PCA	3.00	0.000	1.000
	TwoNN	1.92	0.104	0.079
	MLE	1.93	0.071	0.073
S Curve	PCA	3.00	0.000	1.000
	TwoNN	1.90	0.128	0.105
	MLE	1.94	0.019	0.058
Hyperplane	PCA	10.00	0.000	0.000
	TwoNN	9.82	0.290	0.178
	ESS	10.25	0.230	0.249

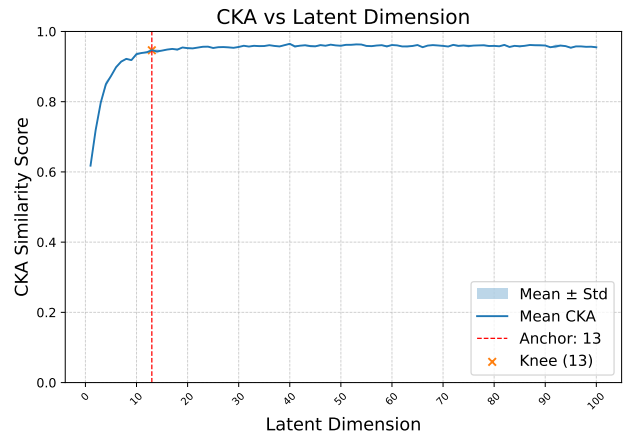


図2: 潜在次元とCKA類似度の関係

な潜在次元は訓練データに内在している」という仮説の実証的根拠を得たと言える。

今後は、多様体次元の推移を捉えた動的な次元設定手法の開発を検討する。また、画像以外の時系列データや高解像度自然画像など多様なデータセットを用い、手法の汎用性と実用性をさらに評価する予定である。

### 謝辞

本研究は、科研費JP23K11266, JP23K28077, JP24-K15115, 東北大学電気通信研究所共同プロジェクト研究【R06/B14】「深層学習における表現学習に関する研究」の助成によるものです。

### 参考文献

- 小端 千佳, 神野 健哉, “高次元データの内在次元に基づく深層学習モデルの次元最適化”, 電子情報通信学会 非線形問題研究会, NLP2024-115, MMS2024-74, 2025.
- 小端 千佳, 神野 健哉, “深層学習における多様体仮説の検証と次元削減手法の評価”, 2025年度人工知能学会 全国大会 (JSAI2025), 1Win4-02, 2025
- D. P. Kingma and M. Welling, “Auto-Encoding Variational Bayes,” in Proc. 2nd International Conference on Learning Representations (ICLR), 2014.
- E. Facco, M. d’Errico, A. Rodriguez, and A. Laio, “Estimating the intrinsic dimension of datasets by a minimal neighborhood information,” Scientific Reports, vol. 7, Art. no. 12140, 2017.
- P. Levina and P. J. Bickel, “Maximum Likelihood Estimation of Intrinsic Dimension,” Advances in Neural Information Processing Systems, vol. 17, pp. 777–784, 2005.
- K. Johnsson, “Structures in high-dimensional data: Intrinsic dimension and cluster analysis,” Ph.D. thesis, Lund University, Sweden, 2016.
- C. Ceruti, A. Staiano, N. Vanello, P. Migliorelli, F. Ambrogi, and A. Vinciarelli, “DANCo: Dimensionality from Angle and Norm Concentration,” Machine Learning, vol. 93, no. 2, pp. 169–191, 2014.
- S. Kornblith, M. Norouzi, H. Lee, and G. Hinton, “Similarity of Neural Network Representations Revisited,” in Proc. 36th Int. Conf. Machine Learning (ICML), vol. 97, pp. 4365–4374, 2019.