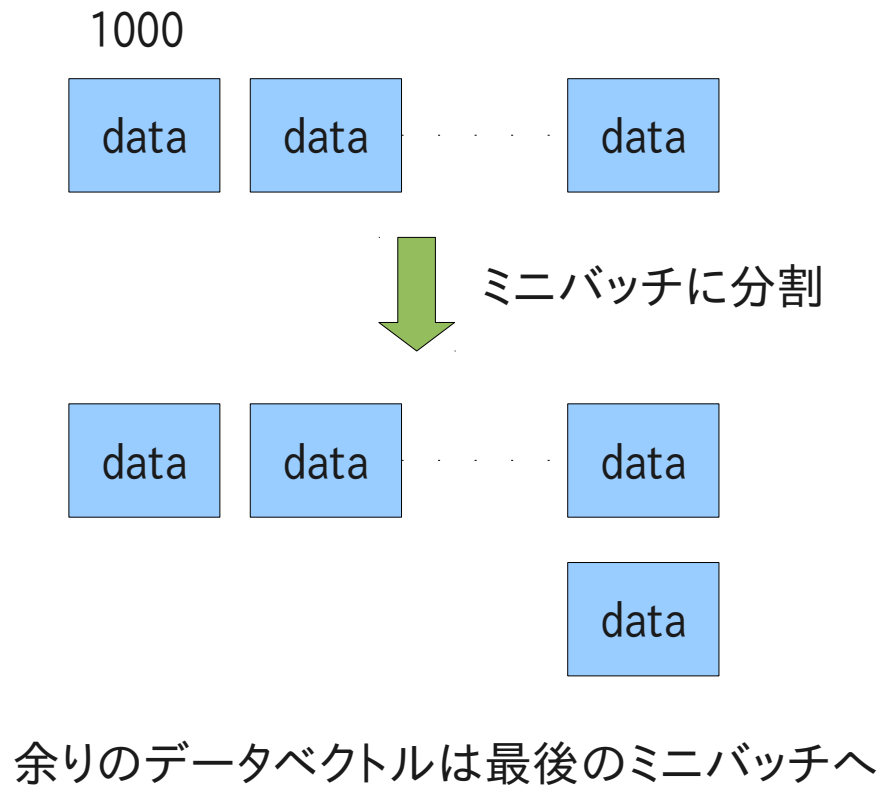
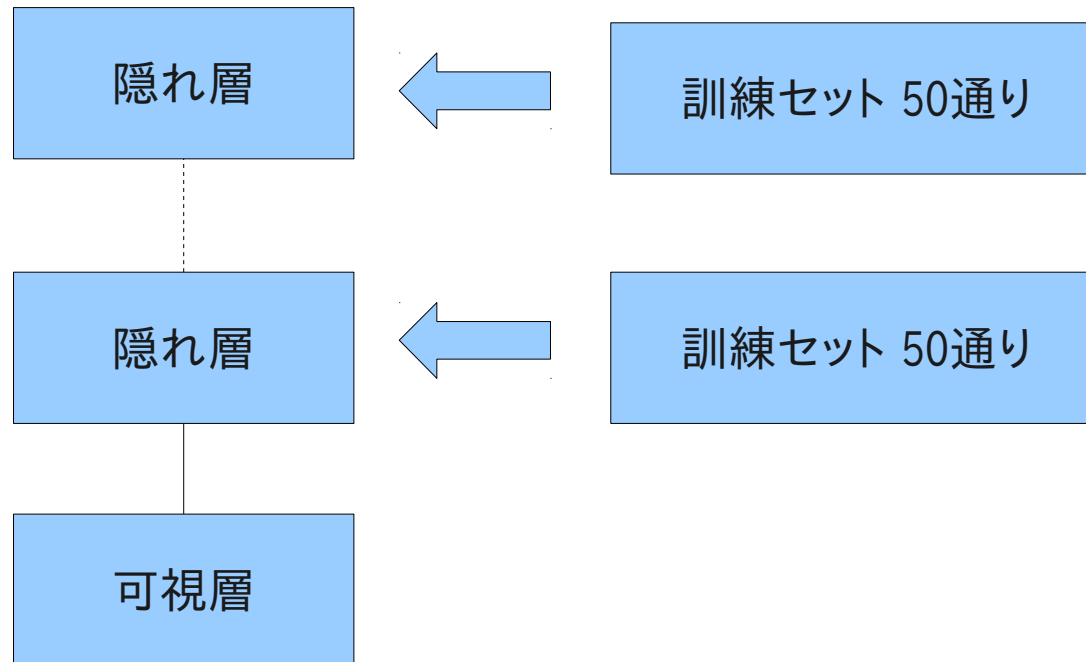


Supporting Online Material for Reducing the Dimensionality of Data with Neural Networks

Details of the pretraining(1)



Details of the pretraining(2)



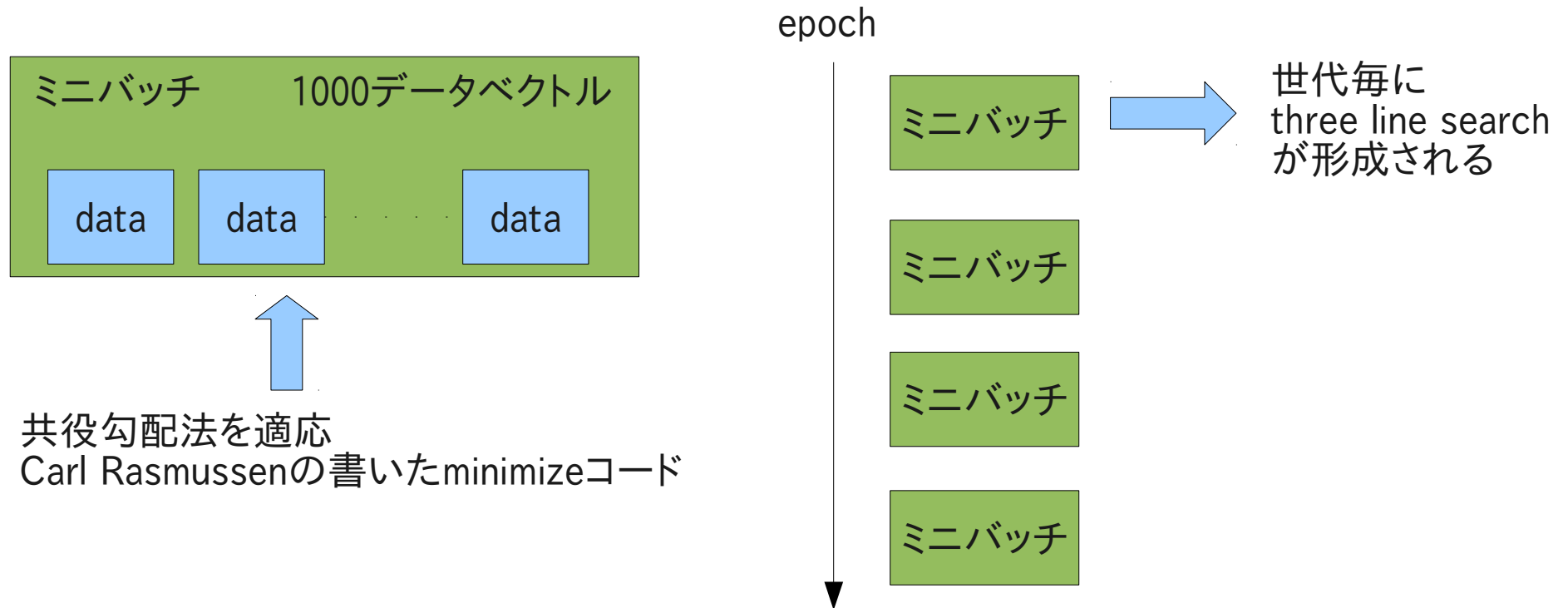
各隠れ層：訓練セット50通りについて事前学習

学習率：0.1

$\text{重み}(t+1) = \text{重み}(t) + 0.9 \times \text{重み}(t) - 0.00002 \times \text{重み}(t)$

平均 0.0、標準偏差0.1の正規分布から初期化

Details of the fine-tuning



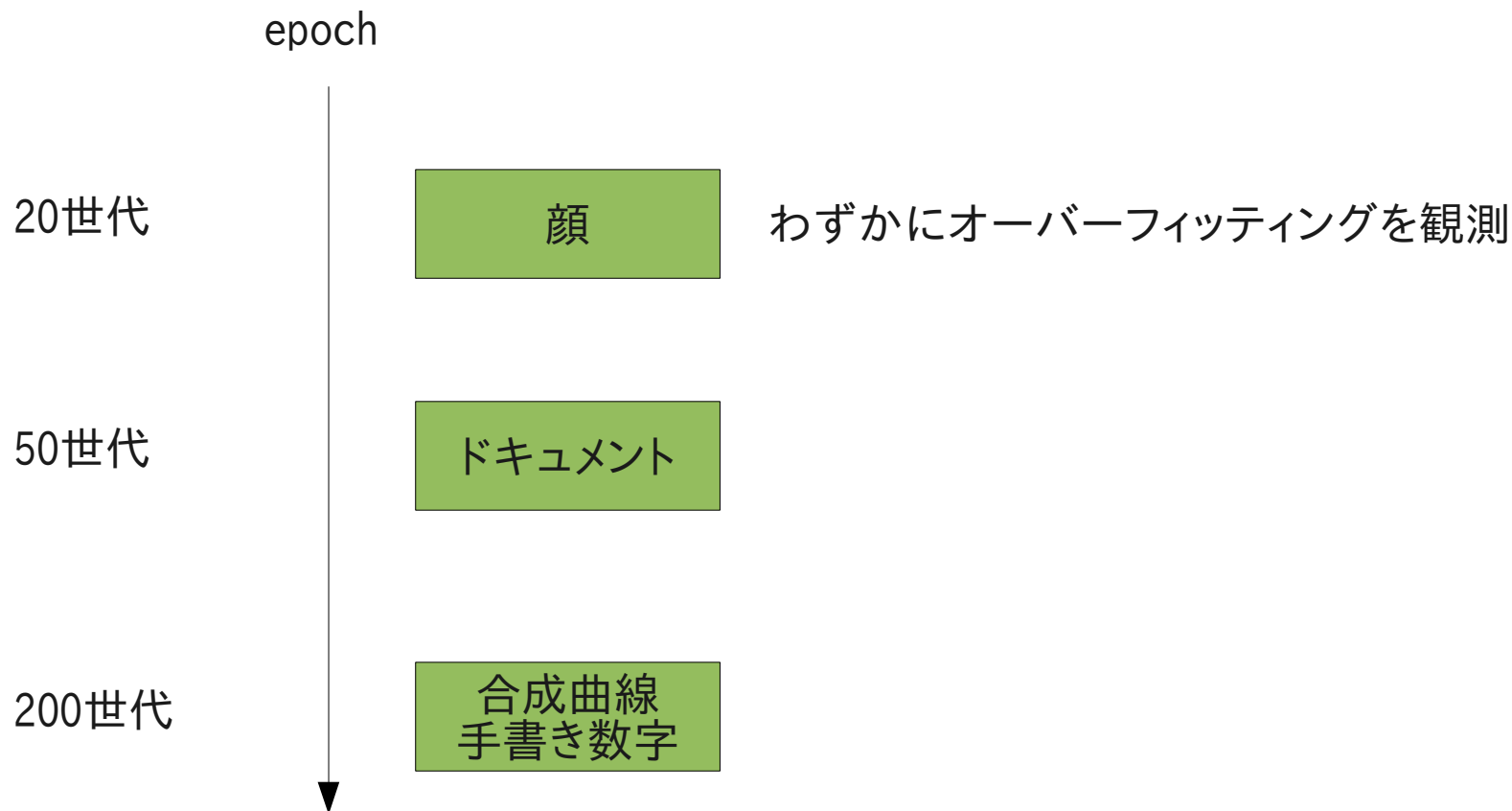
各隠れ層は訓練セット50通りについて事前学習

学習率 0.1

重み(t+1)=重み(t)+0.9×重み(t)-0.00002×重み(t)

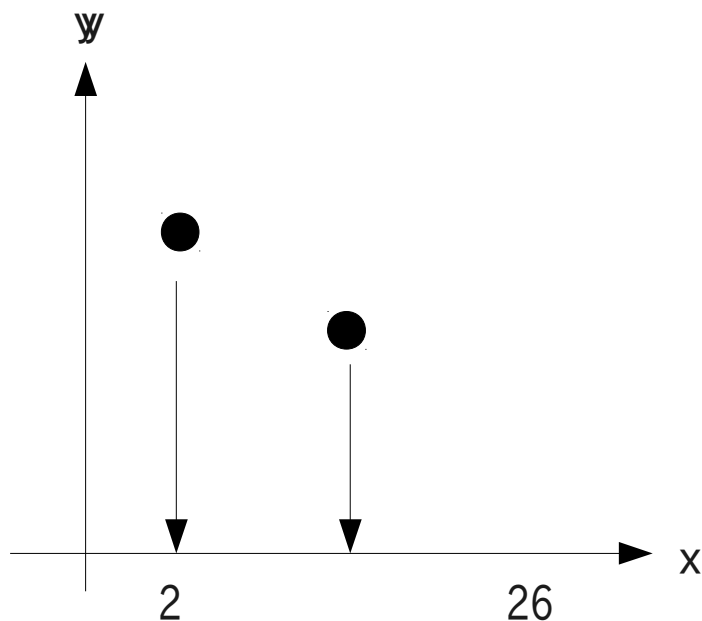
平均 0.0,標準偏差0.1の正規分布から初期化

Details of the fine-tuning

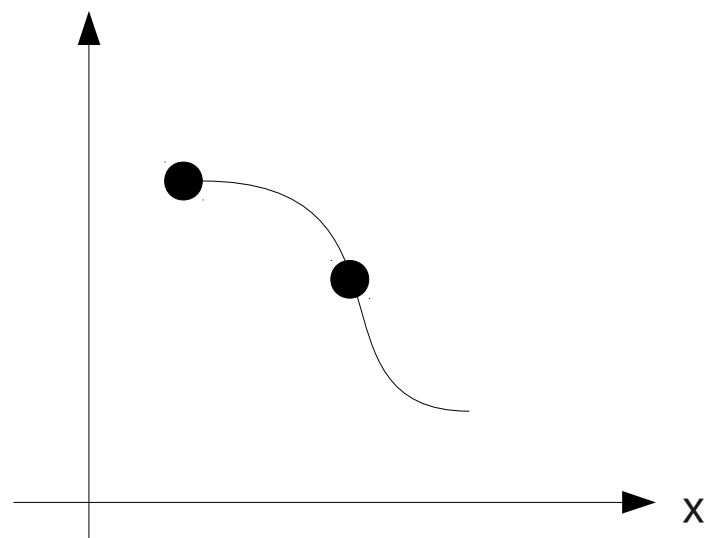


オーバーフィッティングは訓練データが終わりに近いことを意味する
reconstructionはまだ改善中で検証セットよりも悪い
学習率、モーメントム、weight-decay等、様々なパラメータについて実験

How the curves were generated



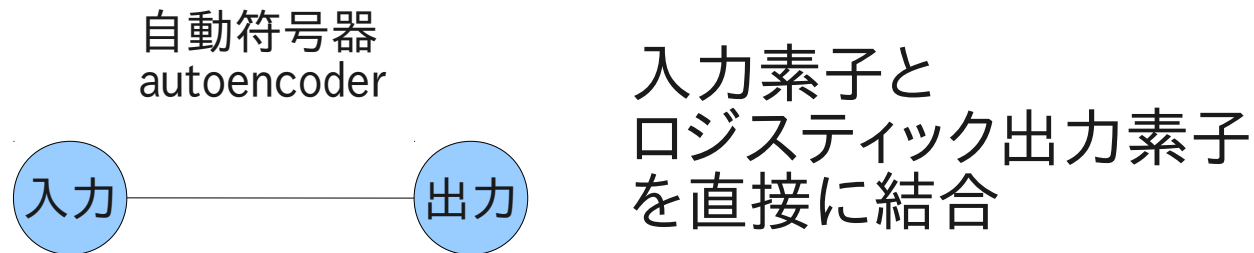
X座標より少なくとも2つの
良い各点のx座標を制約



28x28画素の画像生成に3次スプライン

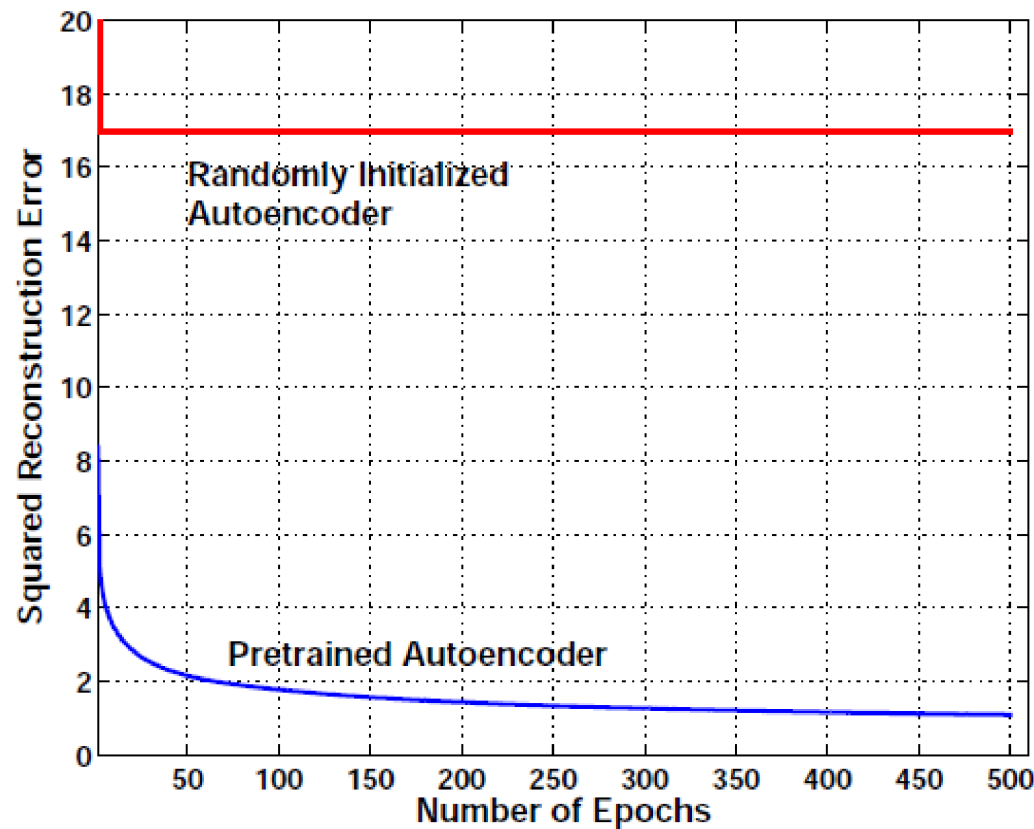
Fitting logistic PCA

ロジスティック主成分分析をフィットさせるために
自動符号器 (autoencoder) を使う



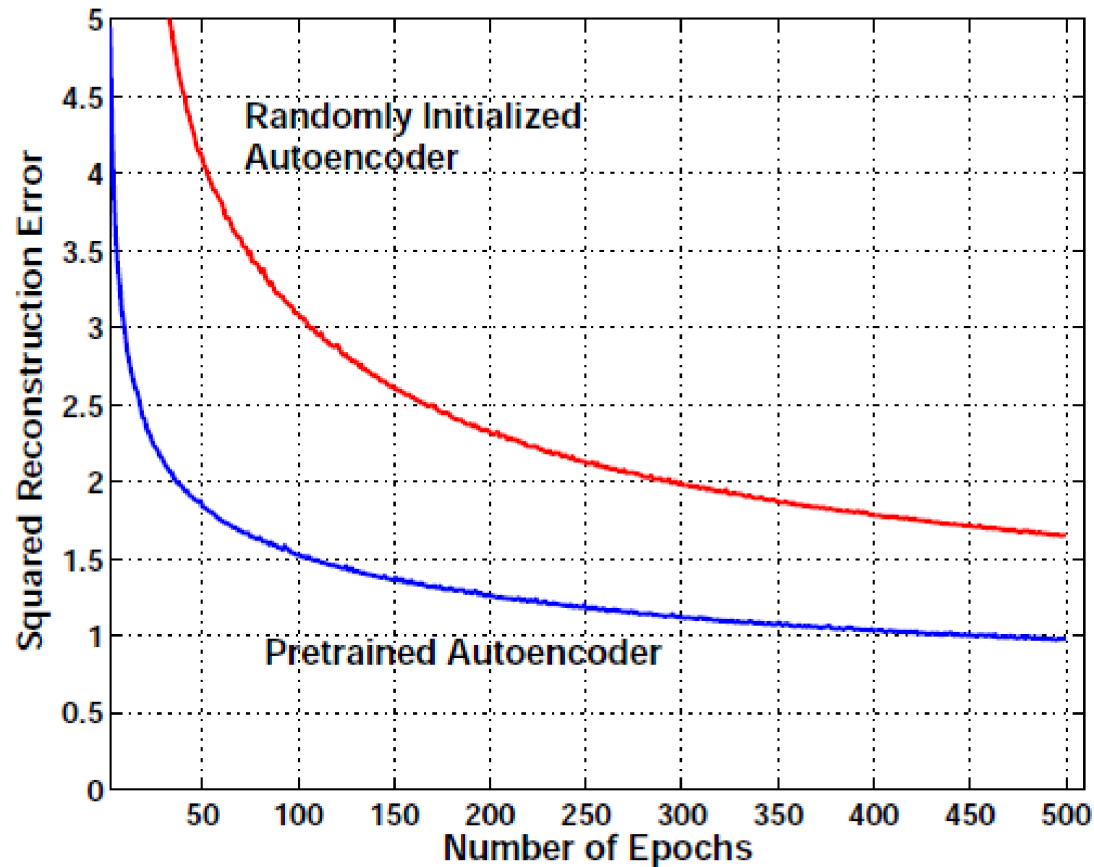
共役勾配法を使いクロスエントロピー誤差を最小化

How pretraining affects fine-tuning in deep and shallow autoencoder(1)



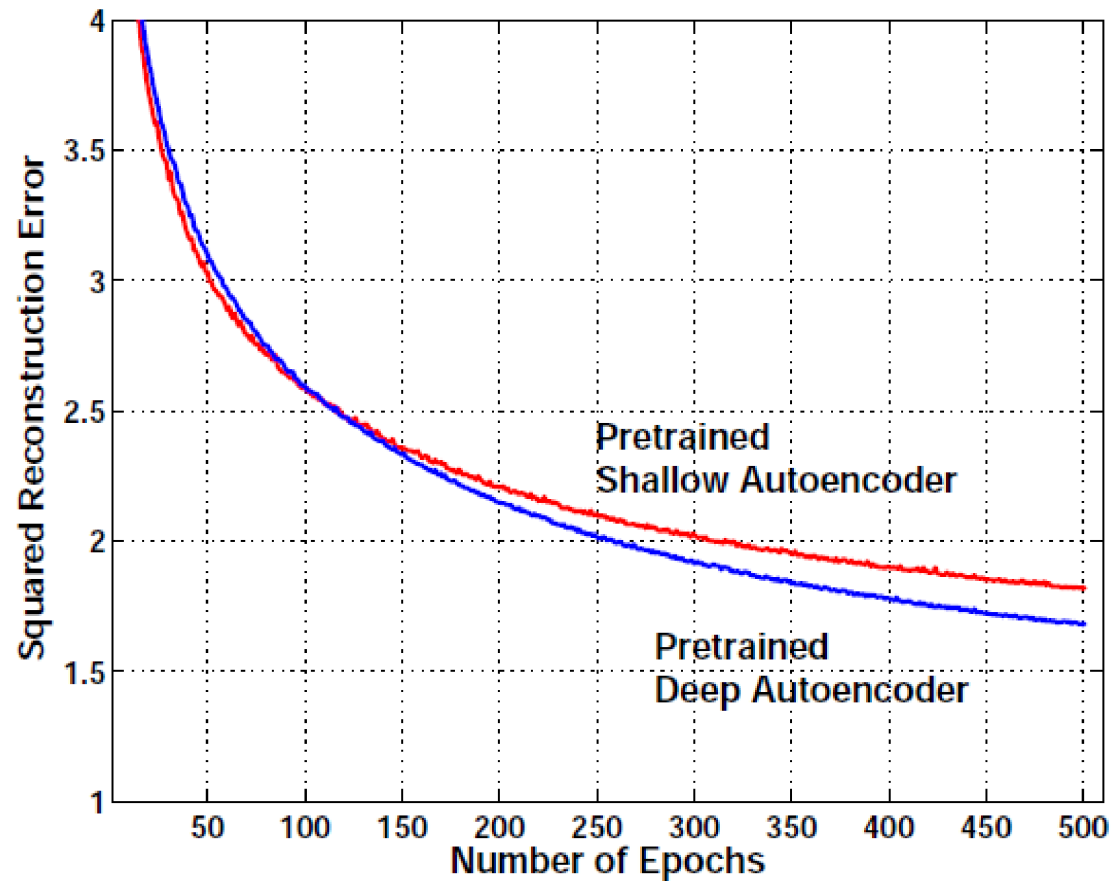
赤：事前訓練無し、ランダムに初期化
青：事前訓練有り

How pretraining affects fine-tuning in deep and shallow autoencoder(2)



Shallow 784-532-6自動符号器
赤:事前訓練無し、ランダムに初期化
青:事前訓練有り

How pretraining affects fine-tuning in deep and shallow autoencoder(3)

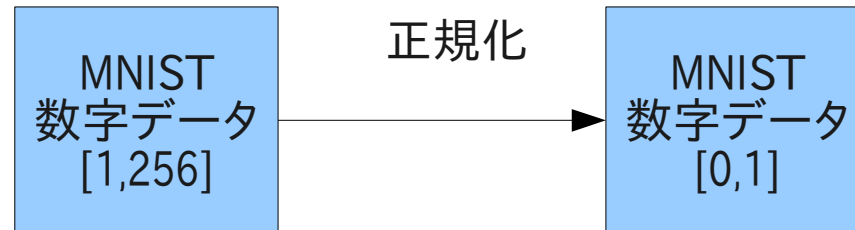


事前訓練有り、自動符号器

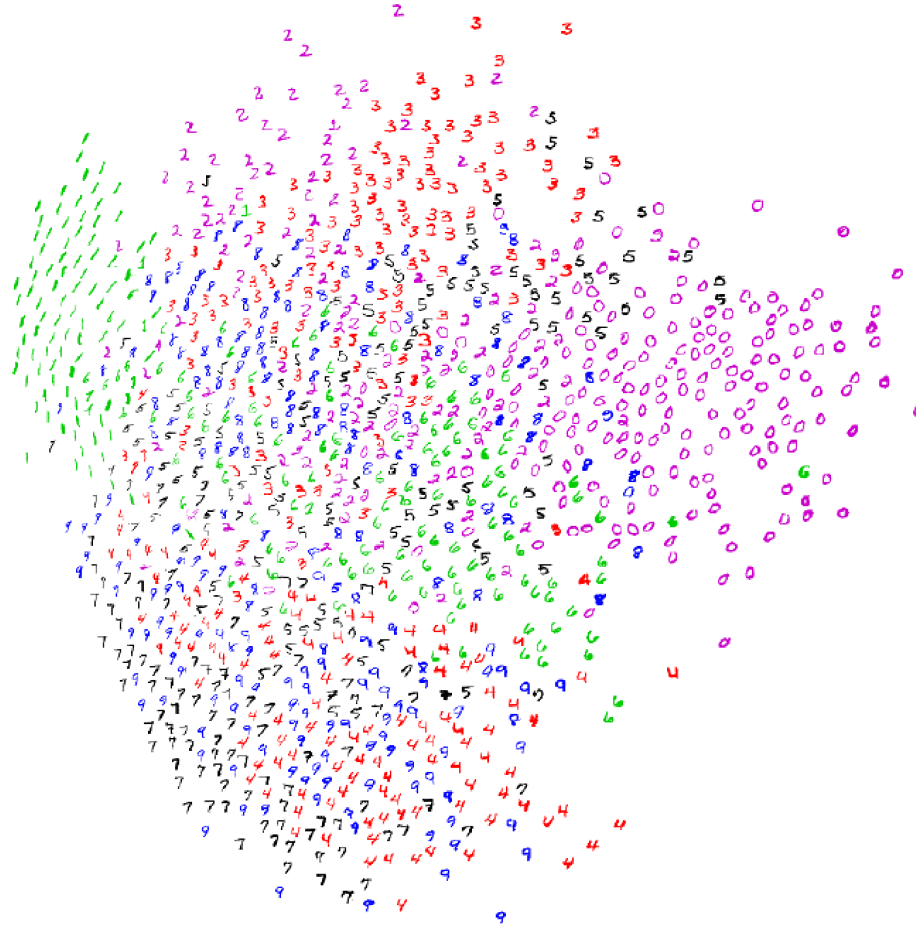
赤: shallow 784-108-6

青: deep 784-100-50-25-6

Details of finding codes for the MNIST digit (1)

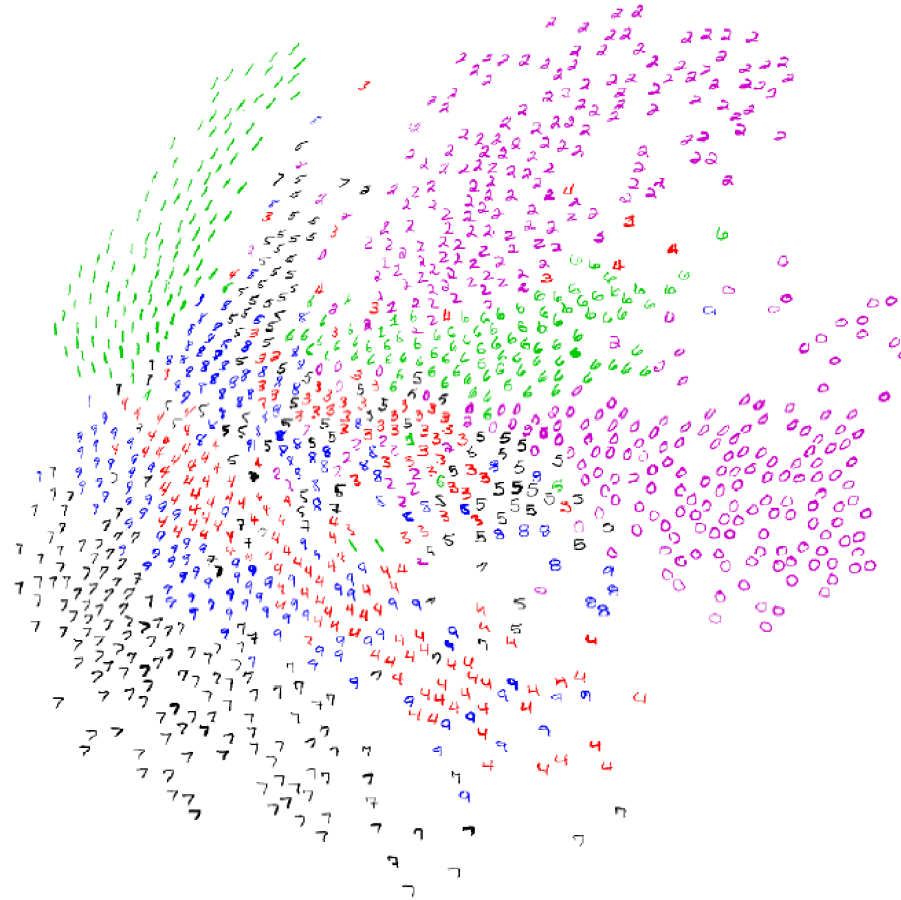


Details of finding codes for the MNIST digit (2)



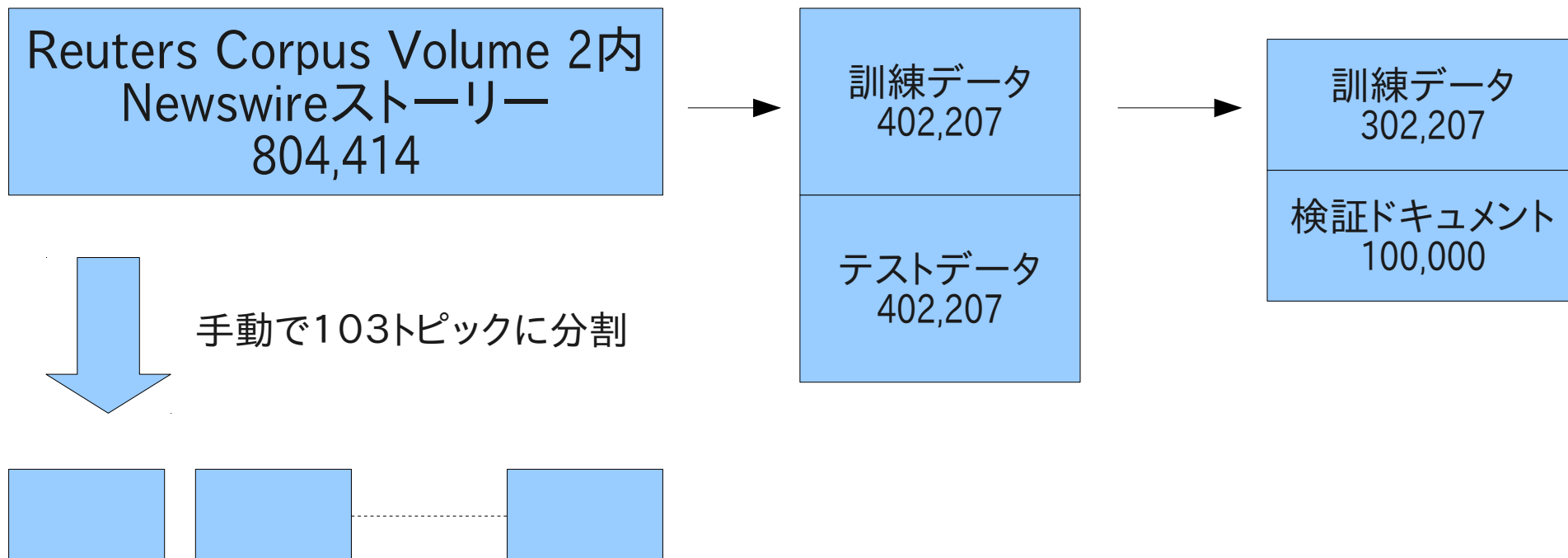
- 2次元コードを可視化
- バイナリ素子をともなう自動符号器
- PCAで生成された2次元コード

Details of finding codes for the MNIST digit (3)



- 784-1000-500-250-2自動符号器
- 全60,000訓練画像による訓練
- 数字の5000画像(500毎クラス)

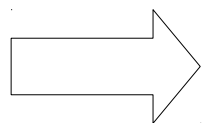
Details of finding codes for the Reuters documents



隠れ素子層

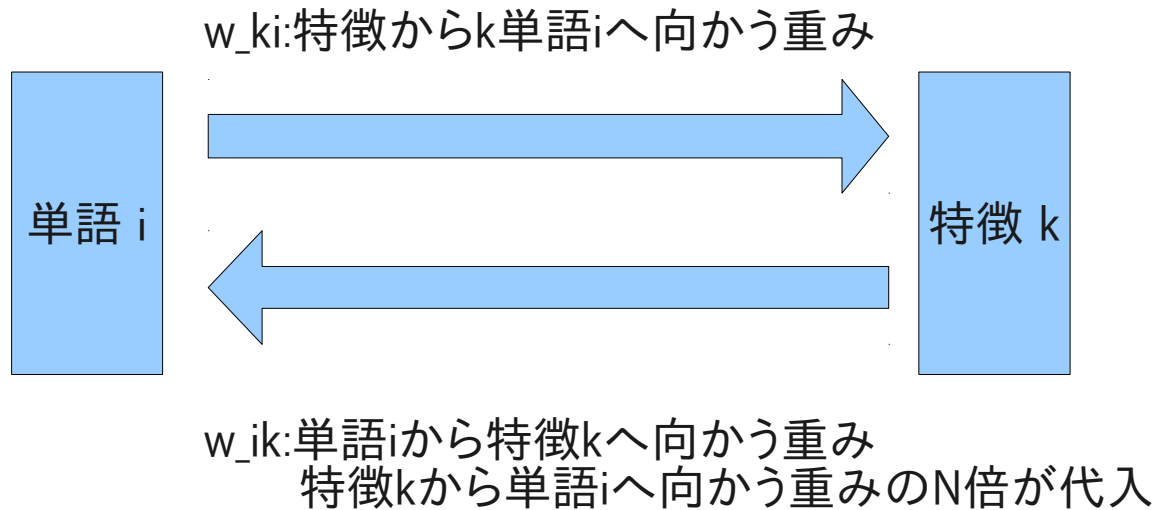
隠れ素子層

可視素子層



confabulated activityは
softmaxを使うことで計算される

Details of finding codes for the Reuters document



単語 i と特徴付け活動 X_i
によるモデル化された確率

$$\hat{p}_i = \frac{\exp(x_i)}{\sum_j \exp(x_j)}$$

単語のカウント数:

$$c_i$$

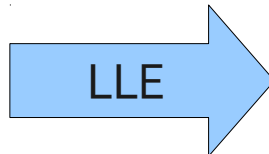
Latent Semantic Analysis: $\log(1 + c_i)$

A comparison with Local Linear Embedding(1)

Local Linear Embedding: S.T.Roweis, L.K.Saul, Science 290,2323(2000)

復元処理での比較が困難

他のノンパラメトリックな次元を減少させる手続きのように
テスト画像を復元する単純な方法を提供しない



新しい問い合わせドキュメントについて
高次元空間の最近傍の発見に関与する

ドキュメント修復は低次元コードが
どう良いか評価する方法を与える

A comparison with Local Linear Embedding(2)

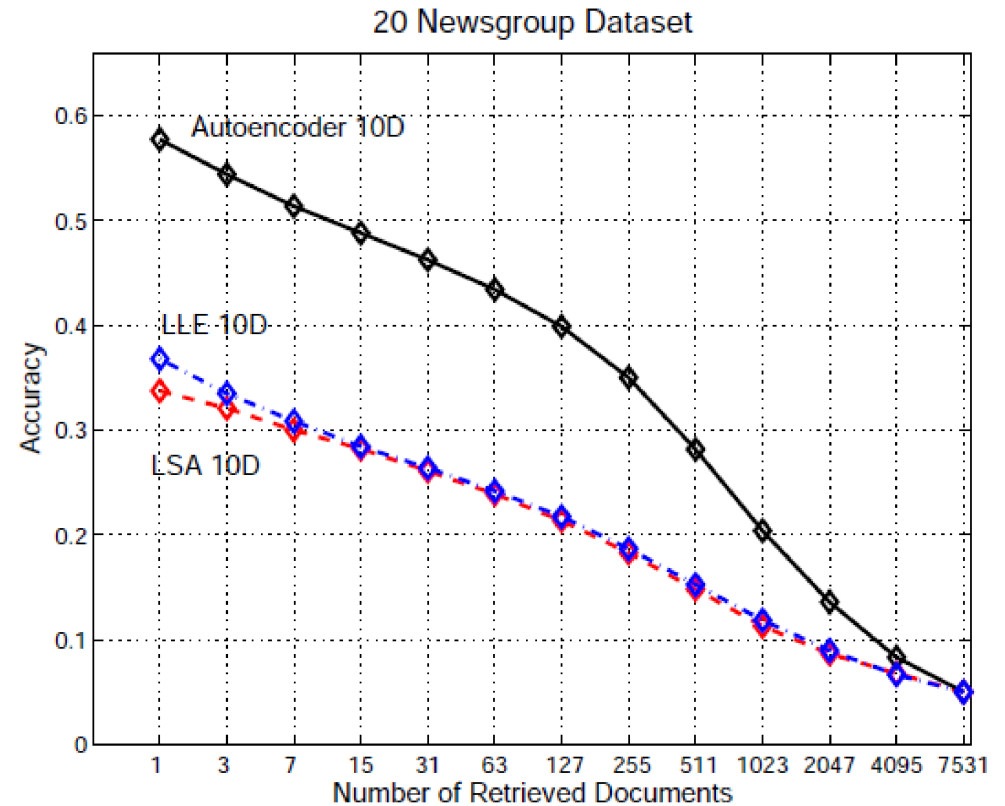
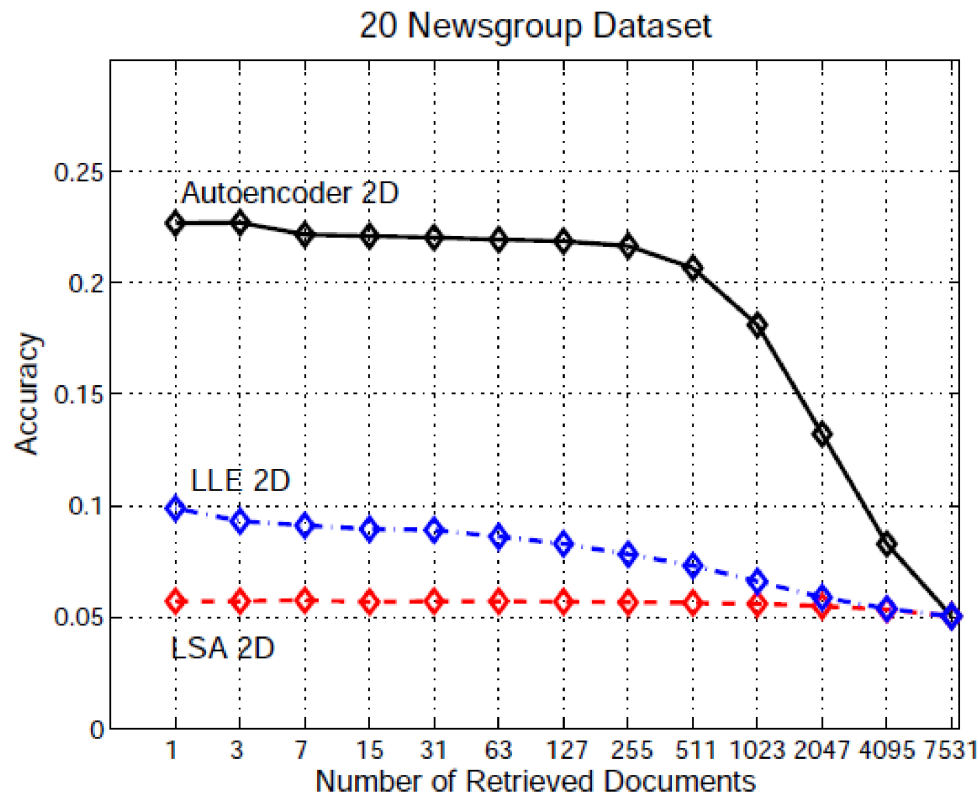
- LLEフットィング
 - 訓練ケース数の2次方程式をとる
 - 20 newsgroup コーパスを使う
 - _ 訓練ドキュメント 11,314
 - _ テストドキュメント 7,531
 - Reuters corpusに関する訓練データ内の最頻出単語2000
 - 最近傍を求めるために、counter vectorの角度の余弦を使う

6. The 20 newsgroups dataset (called 20news-bydate.tar.gz) is available at <http://people.csail.mit.edu/jrennie/20Newsgroups>

A comparison with Local Linear Embedding(3)

- LLEはK近傍の値に依存
 - $K=5, 10, 15, 20, 25, 30$
 - 最適なKを使った結果を報告
- テストフェーズ中
 - 問合せドキュメント q について
 - 訓練セットからの count vector を識別
 - 近傍からcount vectorの復元について
最適な重み W を計算

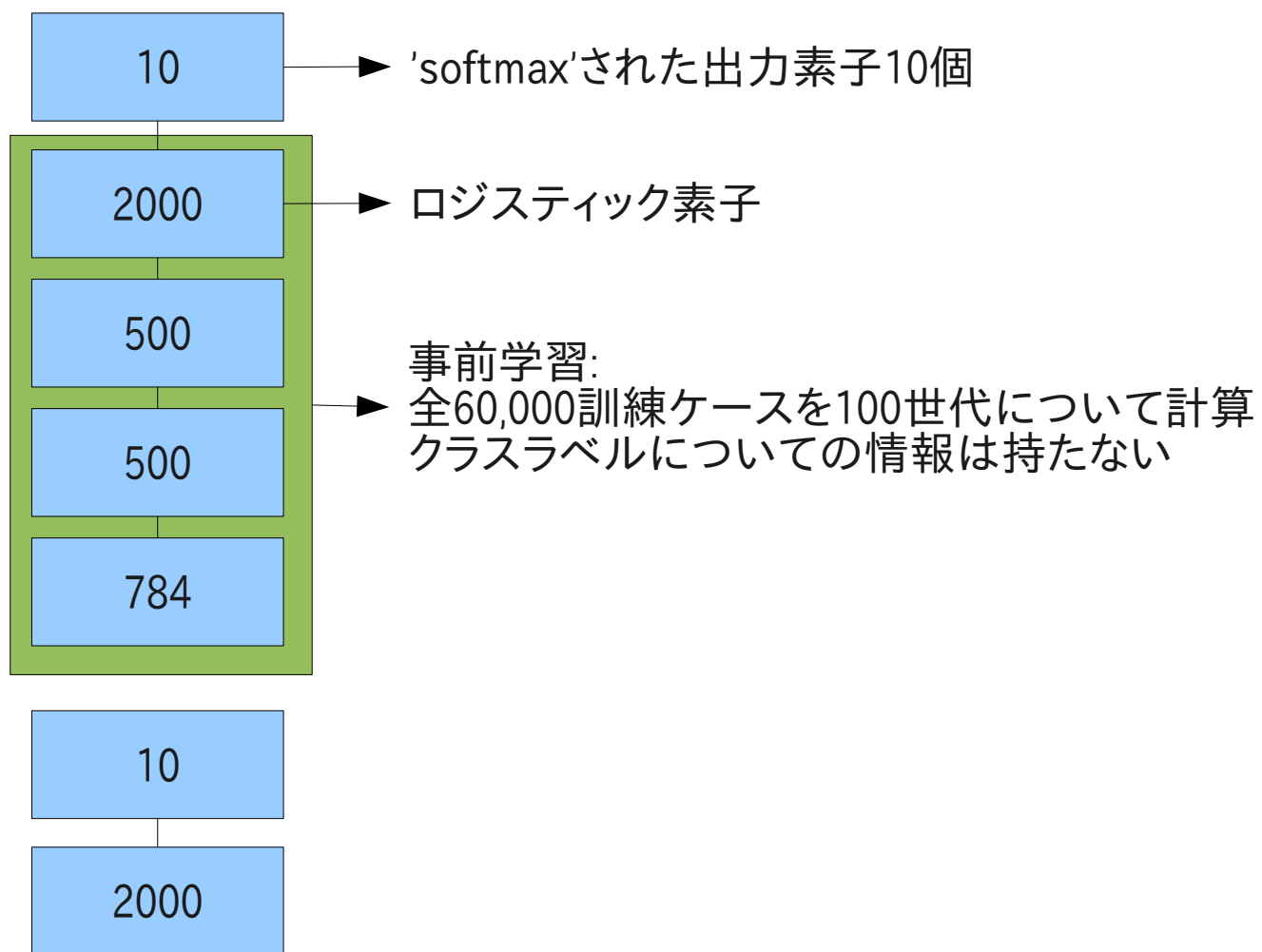
A comparison with Local Linear Embedding(4)

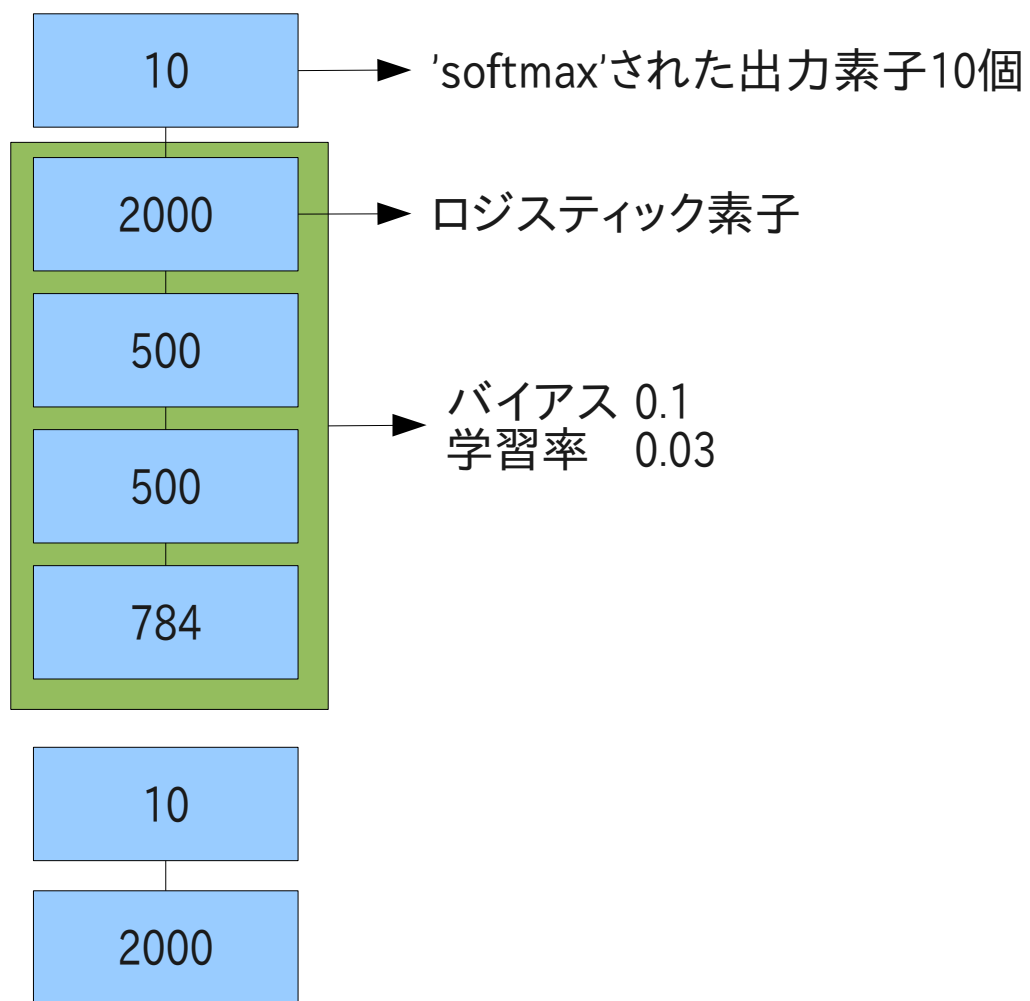


- LLE(K=10)のパフォーマンスはLSAより良い
- LLE(K=25)のパフォーマンスはLSAと似ている

Supporting text

- MNIST数字認識処理の順列不変量版を用いる
- 学習プログラム前
 - 全ての画像はランダムな順列の画素を経験
 - アフィン変換や局所受容野などの配置について事前情報を用いることから学習プログラムを守る
- permutation invariant処理では
 - Support Vector Machine 1.4%
 - BP(784-800-10) 1.6%
 - BP(784-500-500-2000-10) 1.2%





epoch

48世代
テストエラー 1.14%

77世代
テストエラー 1.17%

訓練ケース
50,000

訓練ケース
10,000

fine-tuning

検証セットとして残っている

ミニバッチ

ミニバッチでCG法を適用する
全てのネットワークにfine-tuneを実行

The energy function for real-valued data

モデルに可視ベクトル v を代入した確率

$$p(v) = \sum_{h \in H} p(v, h) = \frac{\sum_h \exp(-E(v, h))}{\sum_{u, g} \exp(-E(u, g))}$$

ガウシアンノイズを持つ線形素子のエネルギー関数

$$E(v, h) = \frac{\sum_{i \in \text{pixels}} (v_i - b_i)^2}{2\sigma_i^2} - \sum_{j \in \text{features}} b_j h_j - \sum_{i, j} \frac{v_i}{\sigma_i} h_j w_{ij}$$

可視素子 i の統計的な更新則

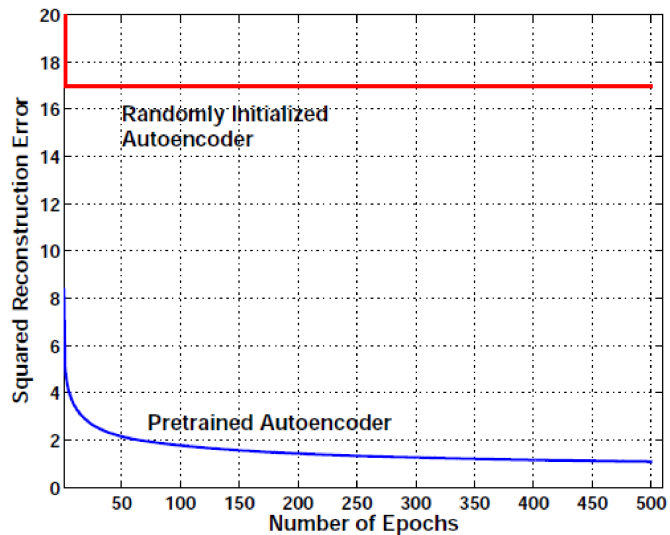
$$b_i + \sigma_i \sum_j h_j w_{ij}$$

The differing goals of RBM and autoencoder

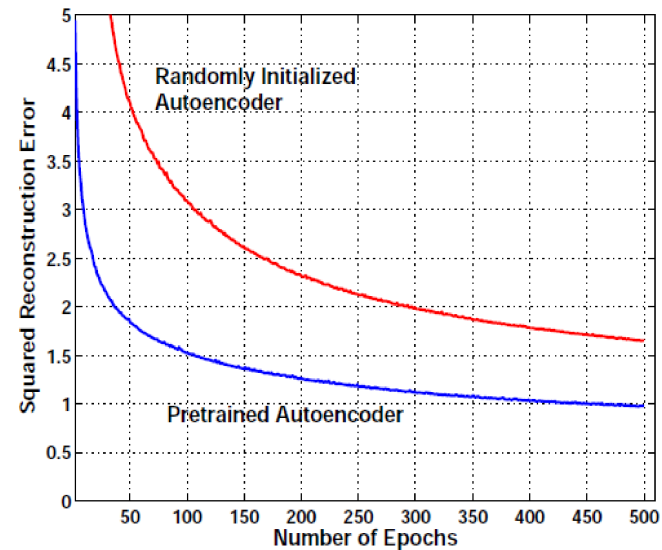
$$p(A) = \sum_B p(B)p(A|B)$$

- autoencoder
 - 訓練画像を復元する
- 事前訓練アルゴリズム
 - 訓練画像の復元が目的ではない
-

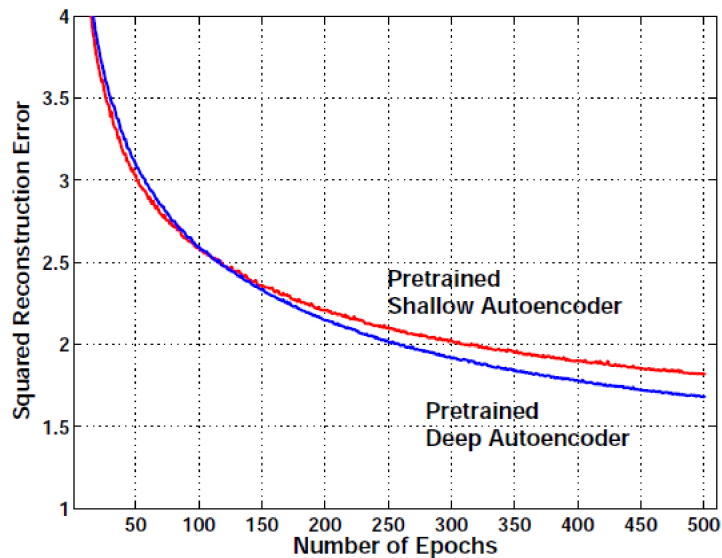
Supporting figures(1)



Deep 784-400-200-100-50-25-6

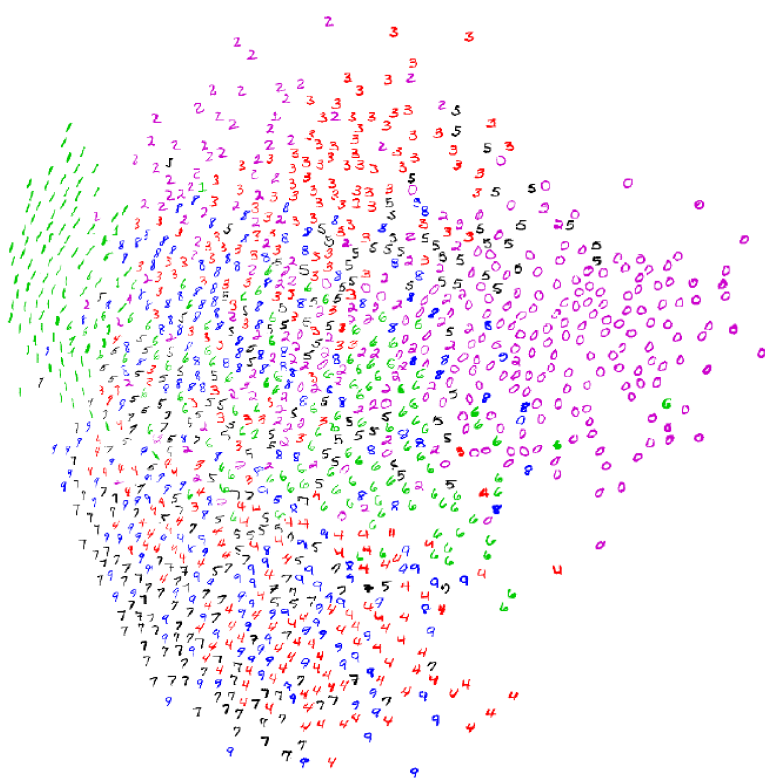


Shallow 784-532-6

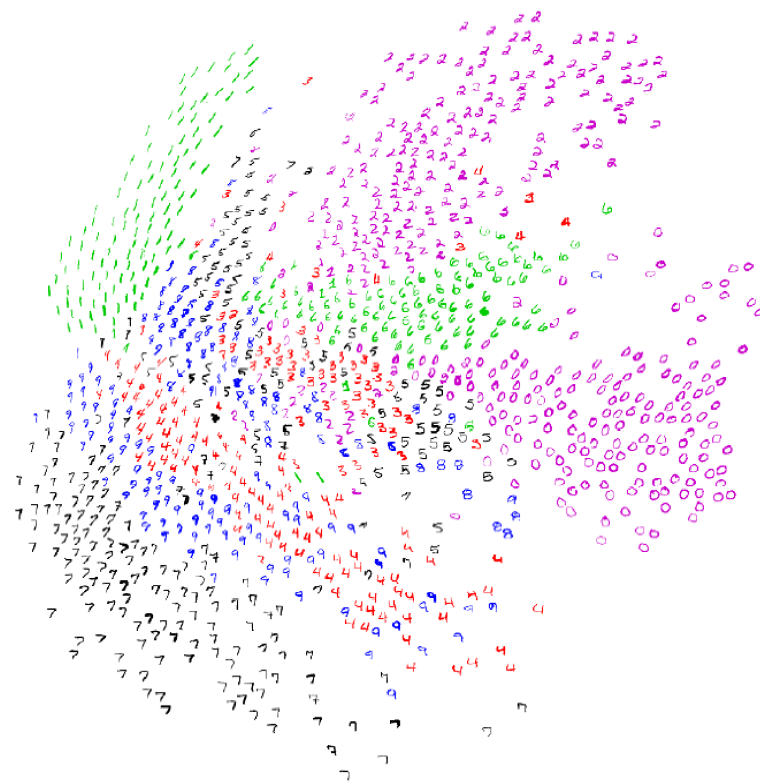


青:784-100-50-25-6 autoencoder
赤:784-108-6 autoencoder

Supporting figures(2)

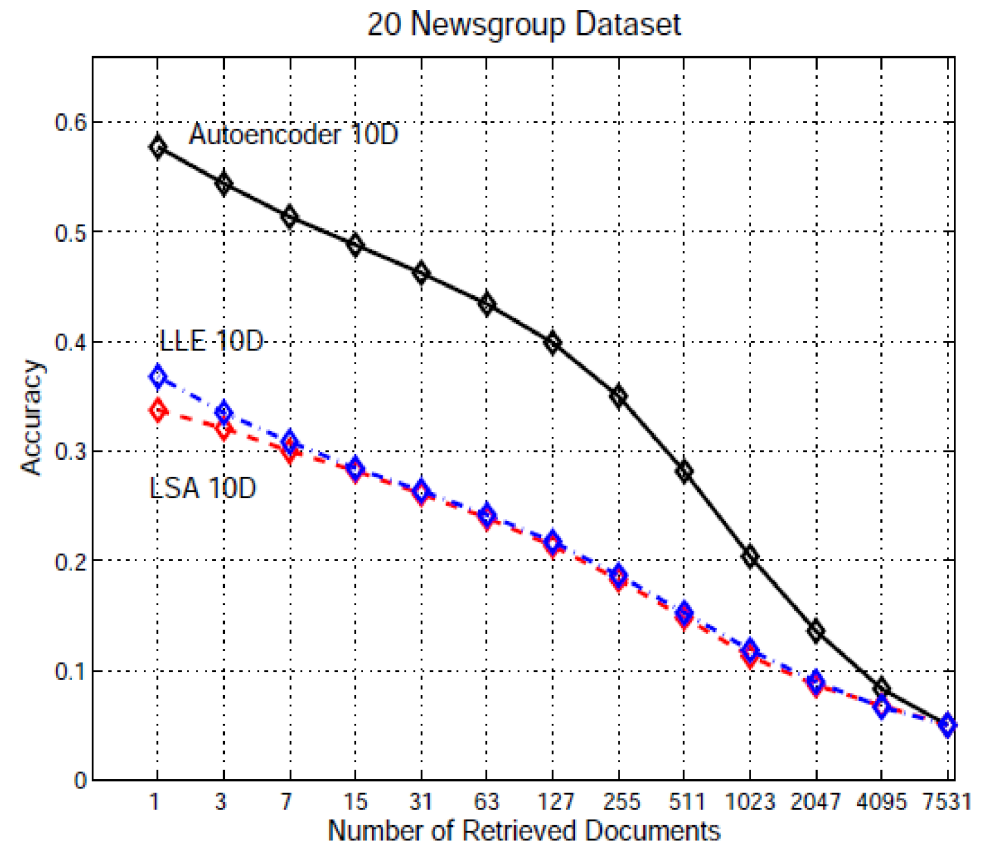
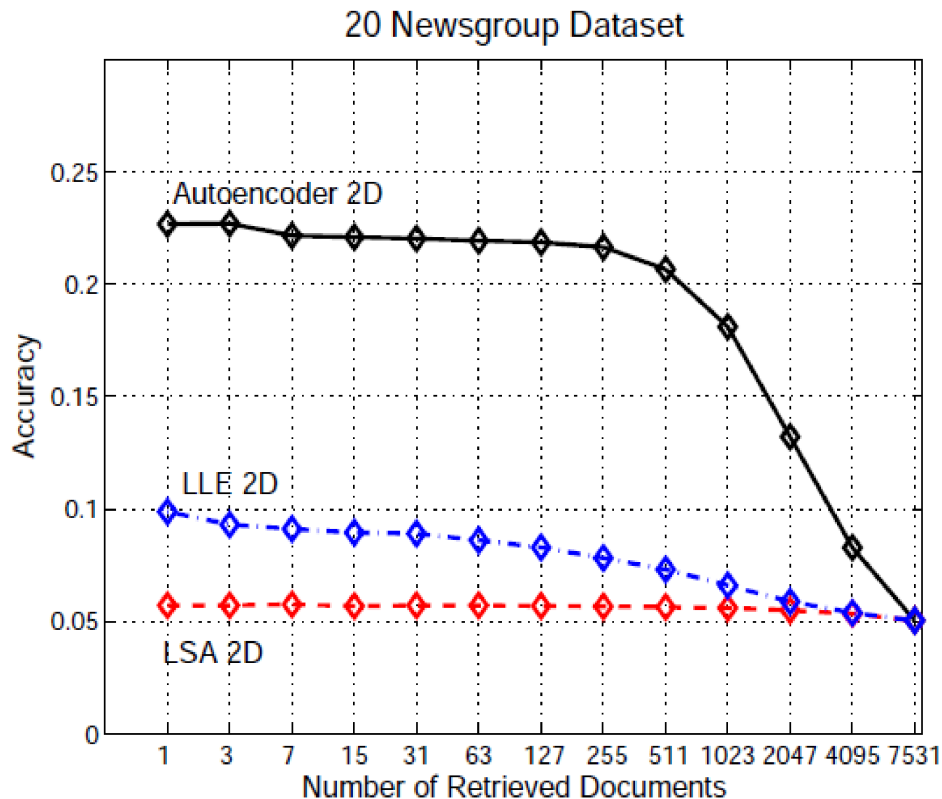


最初の2つの主要な構成をとり生成



784-1000-500-250-2自動符号器により生成

Supporting figures(3)



- テストセットからドキュメントが問い合わせされた際の精度曲線
- 全ての取ることが可能な7,531の問合せを平均化