

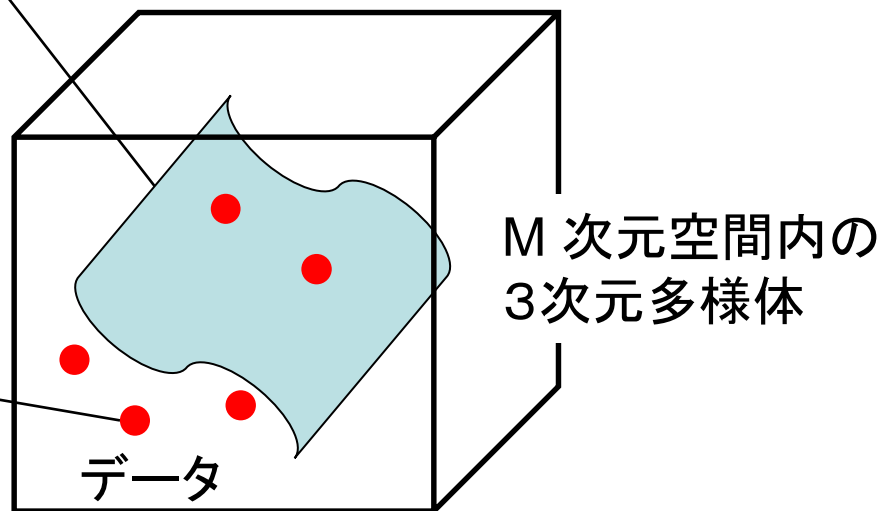
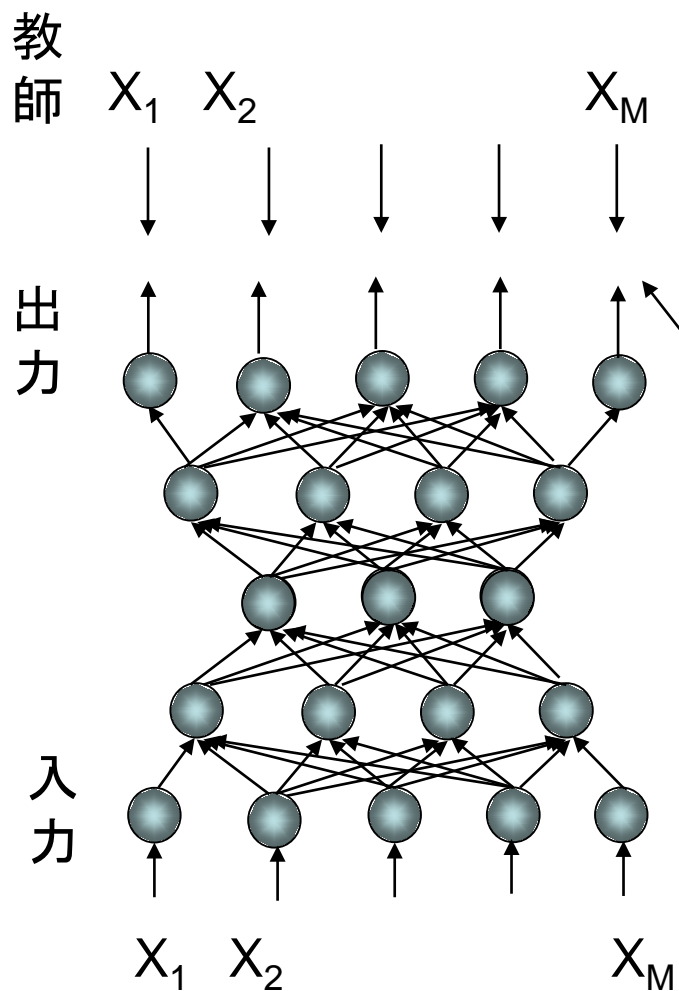
学習理論の練習 3

砂時計型ニューラルネット(オートエンコーダー)

入力と出力に同じデータを用意して学習

入力よりも少ない個数の中間ユニットを設定することで、入力を作る低次元多様体の表現が中間ユニットに作られる・・・と期待される。

(Auto-Encoder には別の型もあります)



学習はかんたん

深層学習のプログラムがそのまま使えます。

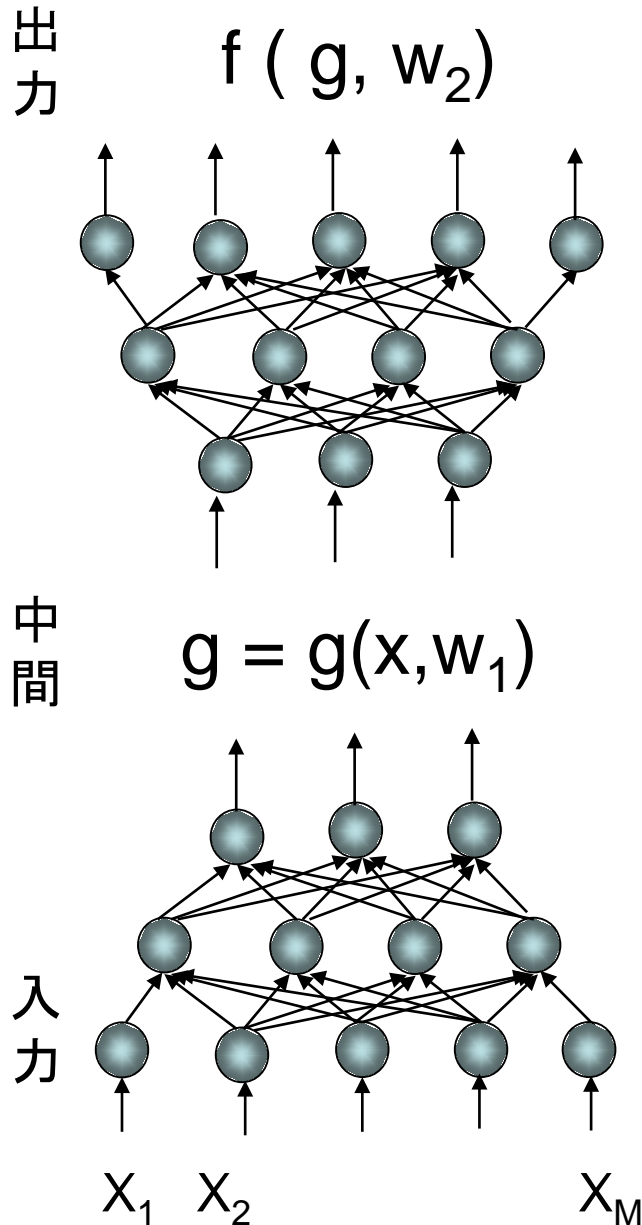
$$w = (w_1, w_2)$$

$$E(w) = \sum \| x_i - f(g(x_i, w_1), w_2) \|^2$$

最急降下法 (Lasso など)

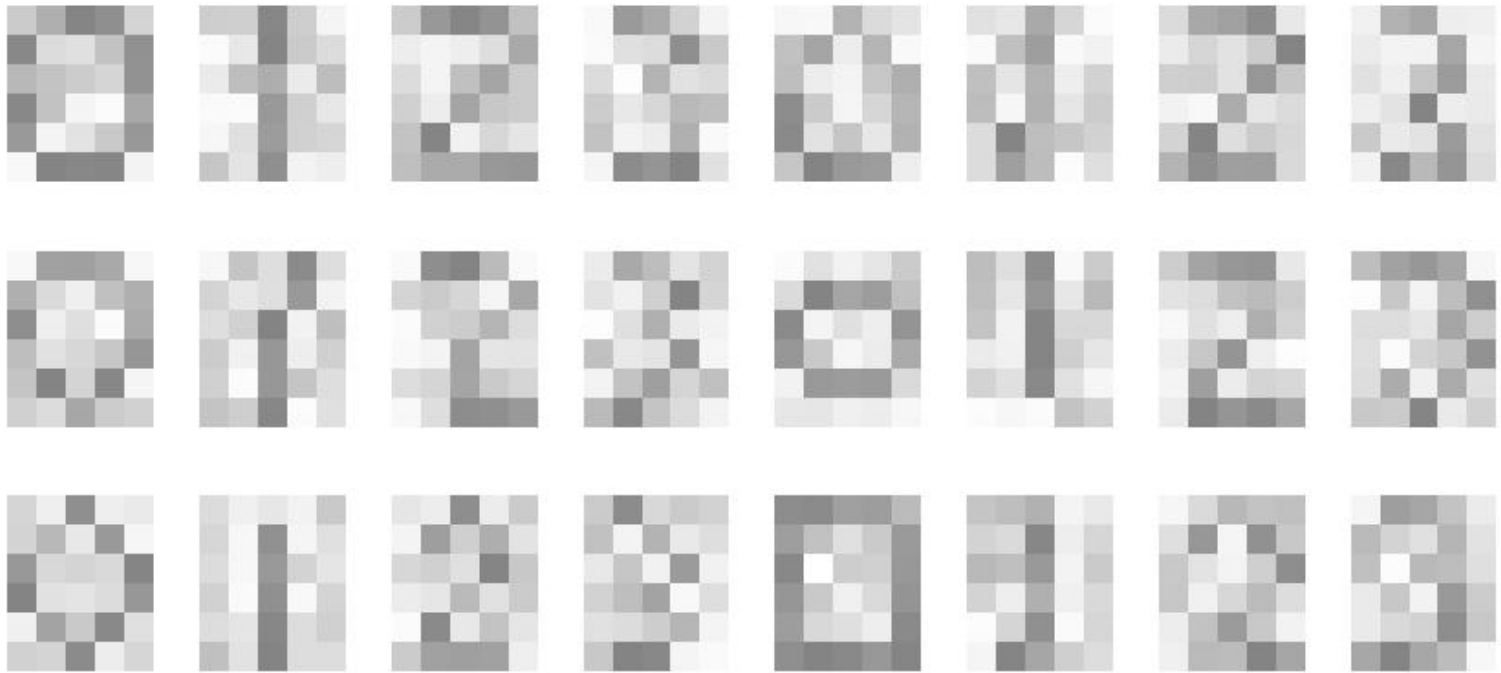
$$dw/dt = -\nabla E(w)$$

※ 誤差は0にはなりません。それは何ですか。誤差が大きい x_i は何ですか。



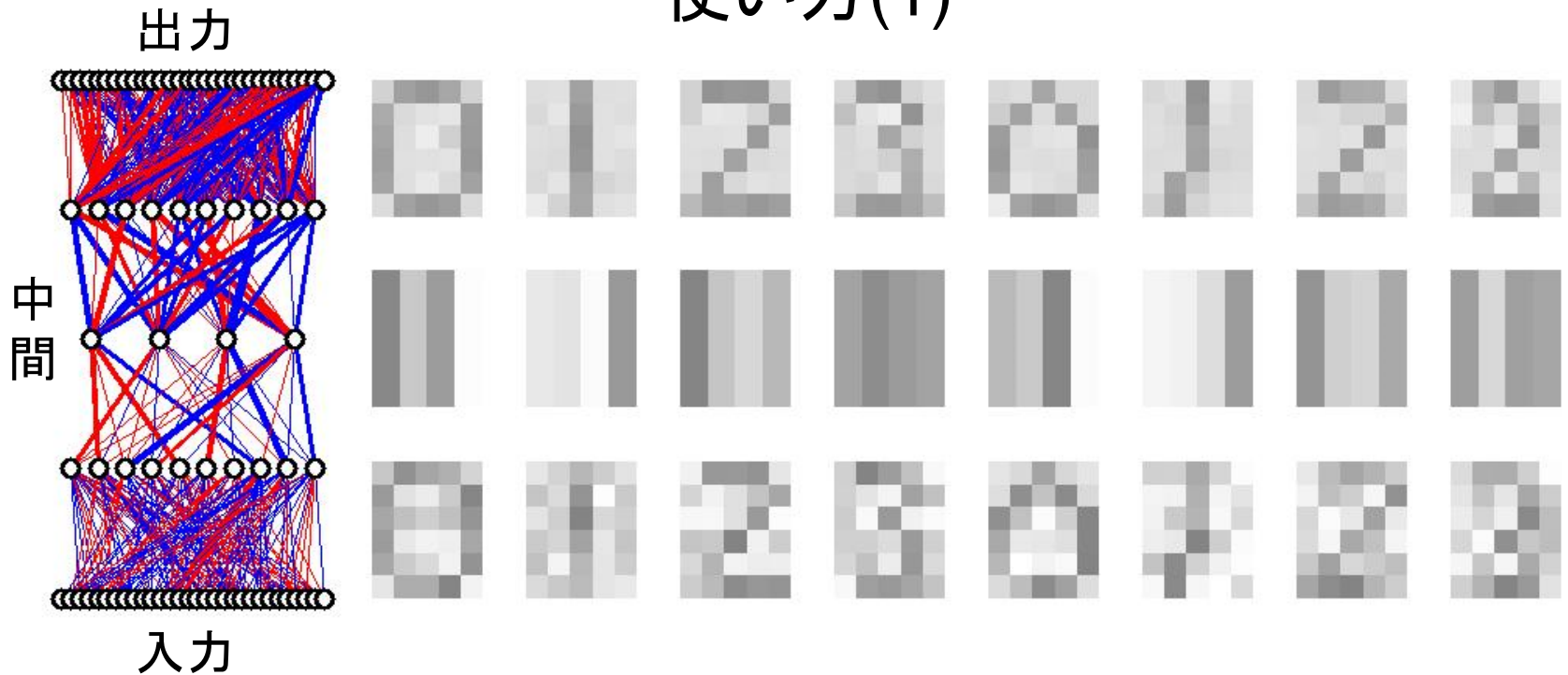
データの例

5 × 6 画像(30次元入力ベクトル)が1000個ある。



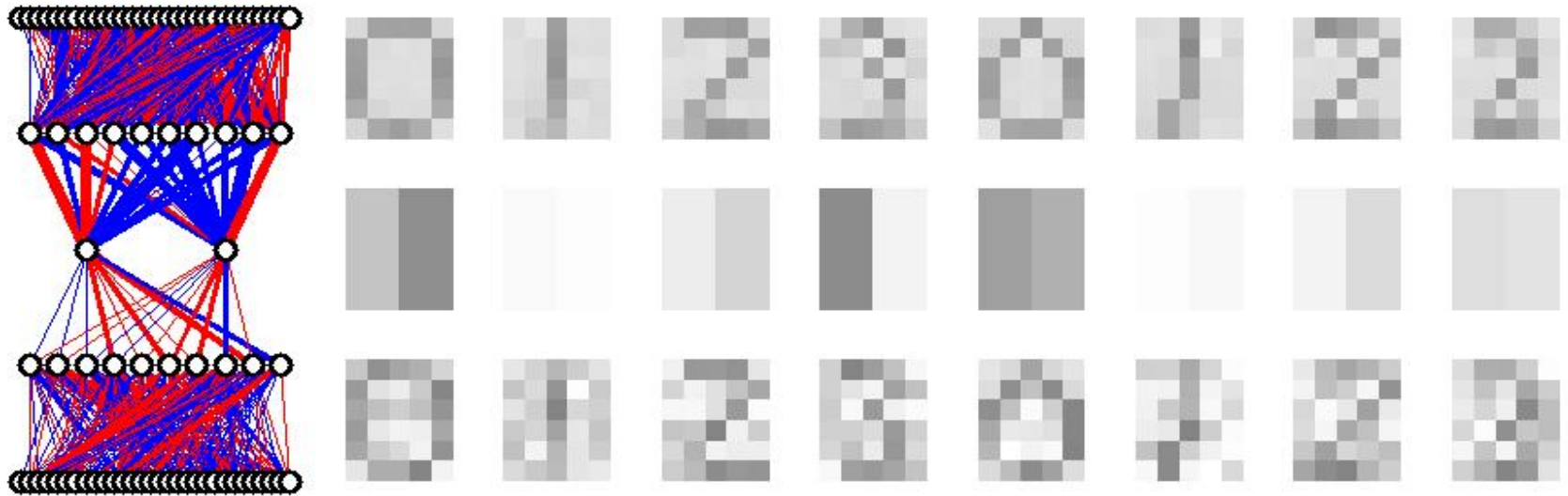
注意: 人間は画像を見て何かわかるが、一般の高次元空間の入力はよくわからないことが普通である。たとえば「千人の人が30個の商品をいくつ買ったかを表すデータ」は、これと同じ形式であるが、見ただけでは購買者の特徴がわからないので解析が必要になる。

使い方(1)

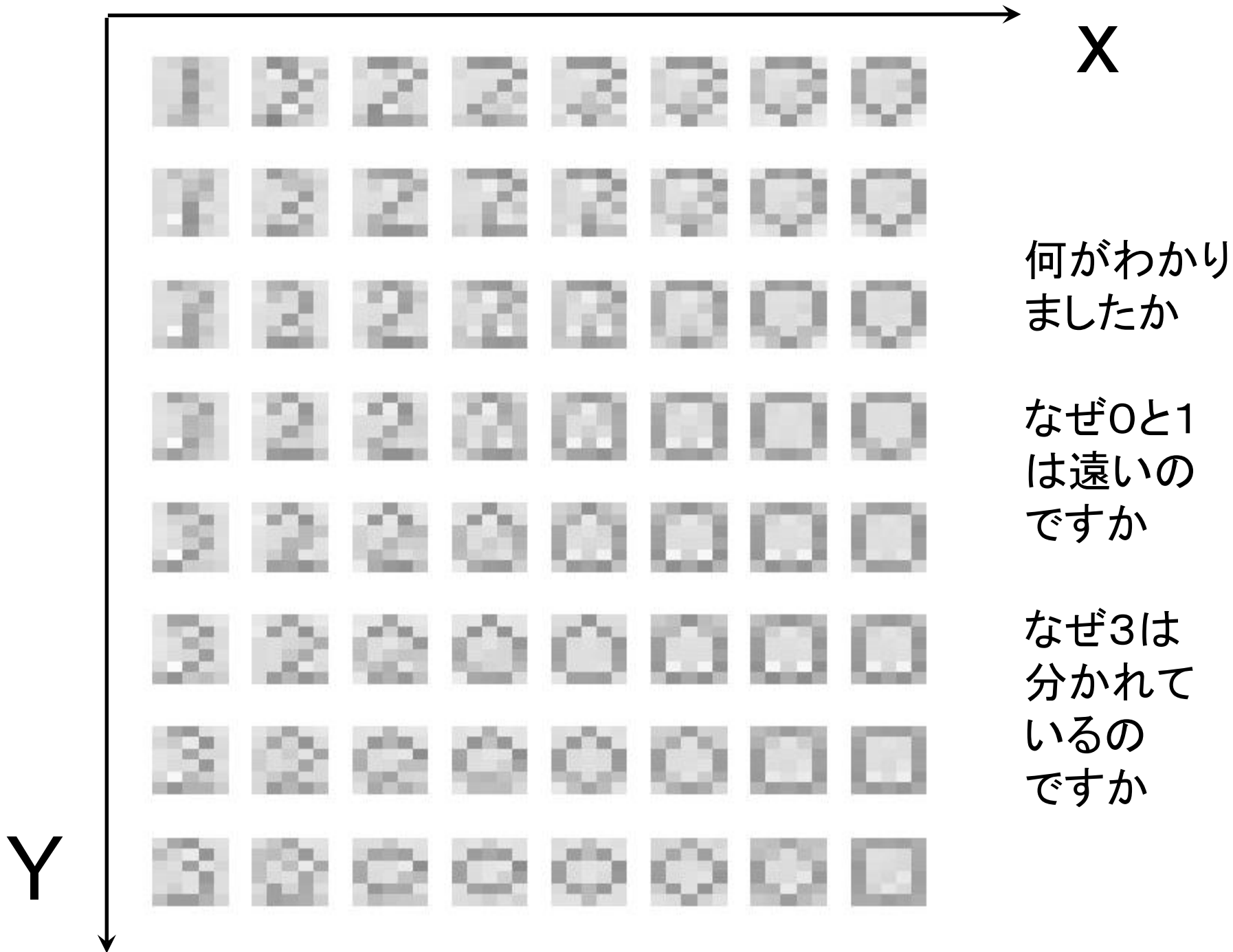


- 30次元のデータを4次元に圧縮することと、圧縮されたデータをだいたい元に戻すことができた。
- 雑音が除去されている。
- 主成分分析の非線形バージョンである。
- 4次元はまだとても大きな空間だ。

使い方(2)



- 30次元のデータを2次元に圧縮した。
- 中間の2次元に $(x,y) \in [0,1]^2$ を入力したときの30次元の出力を見るとデータが作る集合を近似した地図ができる。
- 同じラベルの文字画像でも遠く配置されているものがある。



何がわかり
ましたか

なぜ0と1
は遠いの
ですか

なぜ3は
分かれて
いるの
ですか

課題ですが できなくても気にしない

- できるだけ元画像の情報を残したまま雑音を除去するには中間次元の大きさをどうしたらよいですか。
- 中間次元が3以上のとき、圧縮された情報の意味を理解するためにはどうしたらよいですか。
- 仮にこれが「30商品・千人が買い物」の情報だったとして千人の年齢や趣味などのデータがわかっているとしたらどんな分析をしてみますか。
- これと似た目的に用いられる別の方法も多数ある。比べてみよう。「変分オートエンコーダー」を検索してみよう。
- まったく新しく面白い使い方を考えてみよう。