

卒研ゼミ 「深層学習」

5.11 ~ 5.11.2(第4段落)

2019年11月25日?

岡崎健人

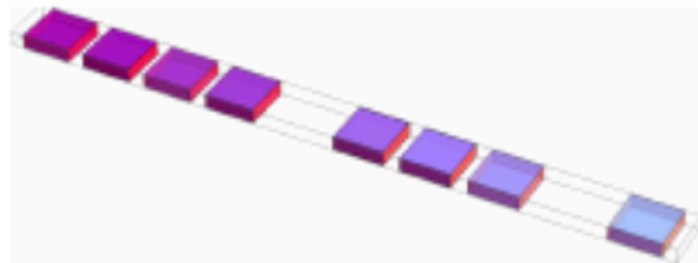
5.11 Challenges Motivating Deep Learning

この節では次の点について説明する

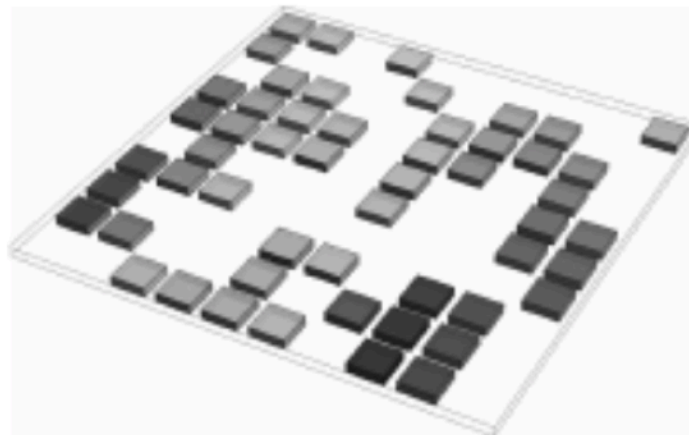
- 高次元のデータを扱う際に、新しい事例に対して汎化することが指数関数的に難しくなること
- 高次元空間の複雑な関数を学習するのに、これまでの機械学習のメカニズムが不十分であること
- 高次元空間のせいで、計算コストが大きくなること

5.11.1 The Curse of Dimensionality (次元の呪い)

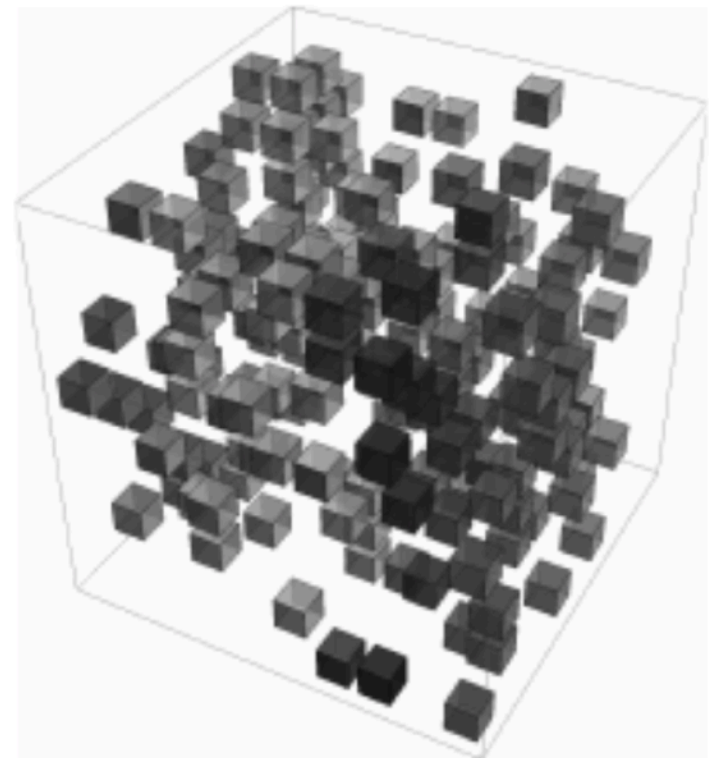
データの次元が大きくなると、その識別可能な領域が指数関数的に多くなる



10領域



100領域



1000領域

d 次元の空間をそれぞれの軸にそって v 個に分割するならば、
領域数は v^d になる。訓練事例は $O(v^d)$ ほど必要。

5.11.1 The Curse of Dimensionality (次元の呪い)

データの次元が大きくなると、その識別可能な領域は指数関数的に多くなる

- 次元が大きくなると、(領域の数) $>$ (訓練事例の数)
- すると訓練集合によっては、ある領域の事例が存在しない場合がでてくる \rightarrow どうする？
- \rightarrow その領域に対する出力は、周りの出力と近似的に同じになると仮定する

5.11.2 Local Constancy and Smoothness Regularization

うまく汎化させるために、アルゴリズムに「事前の信念」を与える必要がある

- 事前分布の「関数形」として与える
- ある方向にバイアスがかかっているアルゴリズムを採用して表現することもある
- よく使われる事前分布：**平滑化事前分布** (smoothness prior)
または**局所一様事前分布** (local constancy prior)
- どちらも小さい領域で大きく変化しないことを「事前の信念」として与えている

5.11.2 Local Constancy and Smoothness Regularization

「小さい領域ではあまり変化しないでほしい」

→ 平滑化事前分布 *or* 局所一様事前分布

- 単純なアルゴリズムでよく用いられる
- 人工知能レベルのタスクには手が届かない
- ここではなぜ平滑化事前分布のみでは高度なタスクに不十分かを説明する

5.11.2 Local Constancy and Smoothness Regularization

「小さい領域ではあまり変化しないほしい」

$$f^*(x) \approx f^*(x + \epsilon)$$

- となる関数 f^* を学習させるよう設計
- 点 x の周辺値を平均したり「内挿」したりすることで、近傍の点 $x + \epsilon$ での値を生成
- 担当はここまで。今井くんよろしく👍