Minsky and Papert's Group Invariance Theorem

Notes for Geoffrey Hinton's *Neural Networks for Machine Learning*
Lecture 2e

Jonathan Merriman

April 13, 2017

Here we will consider a simplified example of the pattern recognition problem presented in the Lecture 2e slides, which is adapted from Minsky and Papert's XOR example in the second chapter of *Perceptrons* [2]. This simplified example considers images with only four pixels, denoted x_1 through x_4. Each x has a value 1 if the pixel is on and 0 if it is off. The image is completely defined by the tuple (x_1, x_2, x_3, x_4).

Group Transformations

Define a transformation T that shifts all of the pixels one step to the right:

$$T(a, b, c, d) = (d, a, b, c)$$

The transformation T can be repeated to form new transformations:

$$T(T(a, b, c, d)) = (T \cdot T)(a, b, c, d) = T^2(a, b, c, d) = (c, d, a, b)$$

Repeated application of T yields the original pattern:

$$T^4(a, b, c, d) = T^0(a, b, c, d) = (a, b, c, d)$$

This transformation has an inverse:

$$T^{-1}(a, b, c, d) = T^3(a, b, c, d) = (b, c, d, a)$$

These properties characterize a mathematical structure known as a *group*. A group consists of a set G augmented with a binary operation $\cdot$ that satisfies the following axioms:

1. The set is closed under the operation $\cdot$:
   $\forall a, b \in G \ . \ a \cdot b \in G.$

2. The operation $\cdot$ is associative:
   $\forall a, b, c \in G \ . \ a \cdot (b \cdot c) = (a \cdot b) \cdot c.$

3. The set has an identity element e:
   $\forall a \in G \ . \ \exists e \ . \ a \cdot e = e \cdot a = a.$

4. Each element in the set has an inverse:
   $\forall a \in G \ . \ \exists a^{-1} \ . \ a \cdot a^{-1} = e.$

The set of transformations $\{T^0, T^1, T^2, T^3\}$ forms a group when paired with function composition.

Group Invariance Theorem

A Perceptron defined on the space of four-pixel images assigns a weight to each pixel and outputs a 1 if the sum of each weighted pixel is above a threshold. We will use weights $\alpha, \beta, \gamma, \delta$ and threshold θ to define the decision function:

$$\Psi(X) = \begin{cases} 1 & \text{if } \alpha x_1 + \beta x_2 + \gamma x_3 + \delta x_4 > 0 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

Minsky and Papert's *Group Invariance Theorem* is concerned with recognizing patterns that are invariant under a group transformation. For a pattern to be invariant under the group of transformations generated by T, $\Psi(X) = \Psi(T^n(X))$ for $n \in 0, 1, 2, 3$. A solution that correctly classifies a given pattern must therefore satisfy four inequalities corresponding to the four distinct transformations generated by T:

$$\begin{aligned} \alpha x_1 + \beta x_2 + \gamma x_3 + \delta x_4 > \theta \\ \alpha x_4 + \beta x_1 + \gamma x_2 + \delta x_3 > \theta \\ \alpha x_3 + \beta x_4 + \gamma x_1 + \delta x_2 > \theta \\ \alpha x_2 + \beta x_3 + \gamma x_4 + \delta x_1 > \theta \end{aligned} \tag{2}$$

These can be reduced to a single inequality:

$$\begin{aligned} \alpha(x_1 + x_2 + x_3 + x_4) + \beta(x_1 + x_2 + x_3 + x_4) \\ +\gamma(x_1 + x_2 + x_3 + x_4) + \delta(x_1 + x_2 + x_3 + x_4) \quad > 4\theta \end{aligned}$$

Letting $\sigma = \frac{1}{4}(\alpha + \beta + \gamma + \delta)$ this can further simplified to:

$$\sigma(x_1 + x_2 + x_3 + x_4) > \theta \tag{3}$$

This is where we get into trouble! Let's say we want to distinguish between two patterns that have exactly two pixels inked. Specifically, let's call the images that can be generated from the transformation T on the prototype $(1,0,1,0)$ examples of Pattern A and those from the prototype $(1,1,0,0)$ Pattern B. There are $\binom{4}{2} = 6$ possible ways to ink exactly two pixels; two of which are instances of Pattern A and four are of Pattern B.

Equation 3 says that if a pattern is invariant under a group transformation, then we must be able to replace the individual weights with a shared weight σ. It's obvious that this would make the sums of the input from both Pattern A and Pattern B equal. Moreover, any images with more than $\lceil \frac{\theta}{\sigma} \rceil$ pixels inked would unavoidably evaluate to the same result.

Refer to §2.3 of *Perceptrons* for the formal definition of the Group Invariance Theorem. The proof proceeds in a similar fashion to the example above but obviously is generalized to any finite group.

References

[1] Hinton, Geoffrey, et al. "Lecture 2: What perceptrons can't do." *Neural Networks for Machine Learning* Coursera.

[2] Minsky, Marvin and Seymour Papert. *Perceptrons - an introduction to computational geometry*. MIT Press, 1987.