Advanced Machine Learning Subsidary Notes

Lecture 4: Symmetry

Adam Prügel-Bennett

January 29, 2024

1 Keywords

- Inductive Bias, Symmetry, Invariance, Group theory

2 Main Points

2.1 Inductive Bias

- Learning machines are not universal approximators, but rather have a propensity to learn a specific set of rules
- This is known as their *inductive bias*
- We want to choose the learning machine so the inductive bias reflects the properties of the problem
- In almost all problems we use some smoothness property to select slowly varying functions

2.2 Invariances

- For many problems there are symmetries that you want your learning machine to respect
- For example in image classification we consider the class of an object to be invariant to its position in an image
- Learning machines whose outputs are invariant to translations have a much smaller set of functions to learn and therefore are much less likely to overfit
- CNNs are equivariant in that if we translate the input the the output will be translated in the same way
- Building classifies using layers of CNN builds in a lot of translational invariance
- Ideally we want a classifier system to also be scale and rotation invariant, although we don't know how to this well
- Sometimes we use pyramid systems where we rescale the image and run the same operations on all scales
- However, in my datasets the foreground object is rescaled to fill most of the image with smaller object treated as the background so scale invariance is not necessarily required
- Often we try to learn invariance by augmenting the training set by artificially transforming the image (e.g. by reflecting, cropping, shearing and rescaling an image)
- Data augmentation usually gives a very significant improvement in performance
- We can sometimes achieve invariance by suitably normalising our data. Again when done right this can give a significant improvement in performance

2.3 Group Theory

- The mathematical language of symmetry is group theory
- You should know group theory exists, but I'm not going to examine you on details of group theory
- This considers a set of objects (we can think of these as transformations)
- Together with a binary operation we can think of as composition
- That is if $a, b \in \mathcal{G}$ are two transformations the $a \cdot b$ can be interpreted as applying transformation b followed by transformation a
- For the set, G, and binary operation, "$\cdot$" to form a group they must satisfy four conditions
  1. Closure: if $a, b \in \mathcal{G}$ the $a \cdot b \in \mathcal{G}$
  2. Associativity: for all $a, b, c \in \mathcal{G}$, we have $(a \cdot b) \cdot c = a \cdot (b \cdot c)$
  3. Existence of an identity element where $e \cdot a = a$
  4. Every element, $a \in \mathcal{G}$ has an inverse such that $a^{-1}a = e$
- (Group theory doesn't really care if the elements of the group are transformations, just that the axioms are satisfied)
- An example is the cyclic group C_2
  - We can think of this as tossing a coin with two actions
    - flip the coin: f
    - leave the coin alone: e
  - These form a group as
    1. Closure: $f \cdot e = f$, $f \cdot f = e$, $e, \cdot e = e$ and $e \cdot f = f$ (if you don't believe this get out a coin and try. The first identity is don't do anything then flip the coin is the same as just flipping the coin)
    2. Associatity; this follows automatically from the operation being a composition (we will see this in innerProduct lecture where we look at matrix multiplication). You can also check rather boringly, e,g, $(e \cdot f) \cdot f = f \cdot f = e$ while $e \cdot (f \cdot f) = e \cdot e = e$ thus $(e \cdot f) \cdot f = e \cdot (f \cdot f)$. You can try this for all triples
    3. e is clearly the identity
    4. e^{-1}=e and $f^{-1} = f$
  - The cyclic group C_3 corresponds to the rotations of an equilateral triangle maps the triangle to itself. While C_n are the rotations that keep an n sided polygon with identical sides onto itself
  - The *four group* is a second group of size 4 (besides C_4) that can be thought of as describing the flipping around a vertical and horizontal axis a rectangle, or rotating it by 180 degress around its centre of mass. We can describe the group in terms of a *multiplication table* where the first row and first column corresponds to the elements a_i and the i^{th} row and j^{th} column corresponds to the multiplication $a_j \cdot a_i$

	e	v	h	r
e	e	v	h	r
v	v	e	r	h
h	h	r	e	v
r	r	h	v	e

    - This is an "Abelian group" meaning that for any elements a, b we have $a \cdot b = b \cdot a$

- The group of permutations of n objects forms a group of size $n!$ called the symmetric group S_n
    - This is a group as the composition of a permutation is a permutation
    - There is an identity (doing nothing)
    - For every element there is an inverse permutation (which is a permutation)
    - For $n > 2$ (where $S_2 = C_2$) the group is non-Abelian meaning that there are permutations p_i and p_j where $p_1 \neq p_2$
    - Despite this composing permutations is associative
    - Permutations have a lot of structure and S_n is very well studied
- There exists groups with an infinite number of elements. The set of integers with the operation of addition form a group with identity, 0, and the inverse of n is $-n$
    - You can think of these describing an infinite set of discrete shifts
    - You first started studying this group when you first learn to count
- The set of rational numbers under addition also forms a group
- The reals under addition form a group describing any one dimensional translation
- The set of reals excluding 0 forms a group under multiplication with an identity 1 and where the inverse of x is $1/x$
    - These can be thought of as describing scale transformations
- Groups can be interpreted in different ways with their elements having different names (e.g.)one, two, three, $\cdots$ or un, deux, trois, $\cdots$). The permutation of two elements has the same structure (is isomorphic to) flipping a coin. What counts is the structure of the multiplication time (up to re-ordering the elements)

2.3.1 Lie Groups

- When we consider continuous transformation in high dimensional space then we find a set of very interesting groups known as Lie groups
- These describe things like general translational invariance, rotational invariance, relativistic invariance
- The set on transformations can be represented as matrices, with the composition operator being matrix multiplication
- Different groups correspond to different sets of matrices.
- A well known group is the set of $n \times n$ orthogonal matrices $O(n)$
- By definition of orthogonal matrices $\mathbf{M}^{\mathsf{T}}\mathbf{M} = \mathbf{M}\,\mathbf{M}^{\mathsf{T}} = \mathbf{I}$
- These correspond to rotations and reflections
- Note that, in general, $\det(\mathbf{M}^{\mathsf{T}}) = \det(\mathbf{M})$ and $\det(\mathbf{A}\,\mathbf{B}) = \det(\mathbf{A}) \times \det \mathbf{B}$
- As a consequence $\det(\mathbf{M}^{\mathsf{T}}\mathbf{M}) = \det(\mathbf{I}) = 1$ so that $\det(\mathbf{M})^2 = 1$ or $\det(\mathbf{M}) = \pm 1$
- If $\det(\mathbf{M}) = -1$ then the transformation involves a reflection
- The special orthogonal group, $SO(n)$, corresponds to a subgroup of matrices in $O(2)$ with $\det(\mathbf{M}) = 1$. These correspond to rotations only
- The group $SO(2)$ corresponds to 2-d rotations where the elements of the group are matrices

$$\begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$$