Thought Base

I Thoughts

1. We have looked again at Arnold's and Landau's physics books. We find them in a sense way too informal. What we are looking for is a very mathematical treatment, where physics merely comes as a layer of commentary on top, linking that to the applications.

2. We see a hint to this in Landau's book, where the earth circling around the sun can be modeled using Newton's particle dynamics, while in other cases, it cannot. This is clear, but one notices internal biases that feel the contrary, that feel that particle dynamics 'is' about small granular matter on earth. This is not! how these models should be seen! The model is a good model of what it is a good model for, no more, no less.

3. We would like to keep a clear separation of mathematics and 'interpretation'. We now understand the necessity of a longer treatment (and 'seemingly') more abstract when going with physics from the route of 'mathematical physics'. Every non-mathematical concept has to be subdued by some mathematical concept.

4. The above explains the universal importance of the cartesian product. Per example, for rigid bodies, $SO(3) \times R^3$ is a product of two spaces. In bundle adjustment [@c1], all the features, and the camera configuration $SO(3) \times R^3$ and the camera calibration are amalgamated using a cartesian product into the total space.

5. We can always form the cartesian product of sets. For applications, one must look for possible extract structure, per example, for topologies, and look at the properties of the resulting topology?

6. The covariance matrix is simply the multivariate generalization of (co)variance ([@c2]).

> This density over the variable x is characterized by two sets of parameters: The mean μ and the covariance Σ. The mean μ is a vector that possesses the same dimensionality as the state x. The covariance is a quadratic matrix that is symmetric and positivesemidefinite. Its dimension is the dimensionality of the state x squared. Thus, the number of elements in the covariance matrix depends quadratically on the number of elements in the state vector.

7. After reading [@c3] and [@c4 (p.43)], we have a better way to explain a statistical inference view of a NN classifying an image with two classes:

We would like to find the model for P(dog|pixels). First we fix a class of models by the structure of the neural network. Then identifying the exact model happen during training. The model, that is, our sought for P(dog|pixels), is the one with coordinates set to the values of all the weights. But, we have not found the right features, what is the perfect thing to replace 'pixels' with? this happens as we choose the weights. Now we would like that we find the features such that our probability distribution is very sharp. We must find the space of features where dogs are near dogs. And that 'location' is assigned a very high probability of 'dog'. When an image is taken as input, it is found ‘where it is’, and if it is in a spot with lots of dogs around it, it will be highly probably a dog. The reason we call it a model is that this is a probabilistic model of the probability that when this picture was taken, there was a dog present in front of the camera, without extra knowledge. So in that sense, we are finding the right distribution, but also the right features.

We need to know more about ‘sampling theory’

> A statistical model is a class of mathematical model, which embodies a set of assumptions concerning the generation of some sample data, and similar data from a larger population. A statistical model represents, often in considerably idealized form, the data-generating process.

But according to [@c4]:

> A model is an argument. Models are collections of various premises which we assign to an observable proposition (or just “observable”). That is, modelling reverses the probability equation: the proposition of interest or conclusion, i.e. the observable Y, is specified first after which premises X thought probative of the observable are sought or discovered. The ultimate goal is to discover just those premises X which cause or which determine Y. Absent these—and there may be many causes of Y—it is hoped to find X which give Y probabilities close to 0 or 1, given X in its various states.

Bibliography

Briggs, William. 2016. ***Uncertainty: The Soul of Modeling, Probability & Statistics***. Springer.

Mellor, David Hugh. 2004. ***Probability: A Philosophical Introduction***. Routledge.

Thrun, Sebastian, Wolfram Burgard, and Dieter Fox. 2005. ***Probabilistic Robotics***. MIT press.

Triggs, Bill, Philip F McLauchlan, Richard I Hartley, and Andrew W Fitzgibbon. 1999. “Bundle Adjustment—a Modern Synthesis.” In ***International Workshop on Vision Algorithms***, 298–372. Springer.