Math 390.4
First Theoretical Lecture

Joseph Peltroche

1/29/2018

Understanding and Applying Models

Definition 0.1. Model: A model is an abstraction/approximation to reality/absolute truth/the system/a phenomena.

Examples of models include the a model airplane which would be an abstraction or an approximation to an airplane, a city street map modelling a city streets, and a wind tunneling modelling airflow.

"All models are wrong but some are useful." -George Box

In the George Box quote, the word "wrong" is used to show the model is not an exact every way to the reality it is mimicking. Models are useful by providing **1.** predictions and **2.** explanations.

Definition 0.2. <u>Validation:</u> A comparison of the measured data to the predictions.

If the comparisons are "close", then the model is "good". If not, we can rebuild the model, iterate, and get closer.

Considering the following old-saying:

"Eary to bed
early to rise
makes a man healthy and wise"

Mathematically speaking, this quote can be represented in the following manner:

$$\begin{bmatrix} health \\ wealth \\ wisdom \end{bmatrix} = f \begin{pmatrix} waketime \\ bedtime \end{pmatrix}$$

where the outputs are on the left-hand side of the equation and the inputs on the right-hand side. Yet, measuring this crude model becomes overcomplicated when considering how to numerically represent health and wisdom which leads us to a requirement of a mathematical model: **mathematical models need numbers and numerical measurements**. They must have numerical inputs and outputs which are related by an equal sign.

Some of the most famous examples are in physics such as the classical Newtownian relation $F = ma = f(m, a)$ and Einstein's relaitivistic equation $E = mc^2 = f(m, c)$. These mathematical models all make the assumption that the universe is explicable with mathematics. For our purposes, we will assume this to be true.

If we represent reality by y, such that it is exact, then we can state an arbitrary relationship such that it and its inputs are true.

$$y = t(z_1, z_2, \cdots, z_n)$$

In this case, y is exact and known as the output, response, outcome, endpoint or dependent variable. The variables $z_1, z_2, \cdots, z_n$ are the true causal input information, and t is the true relationship between the causal input and output.

(Example) Consider a real-world example of credit-worthiness.The reality is either a person can creditworthy or not creditworthy, which can be represented by a 1 or a 0, respectifully.

$$y \in \{0, 1\} = Y$$

The true causal inputs can then later be defined as

- z_1 : The variable indicating whether the person in consideration has enough money at the time the loan is due; $z_1 \in \{0, 1\}$
- z_2: The variable responsible for an unforesable emergency; $z_2 \in \{0, 1\}$
- z_3: The variable respresenting criminal intent

and the true relation can be represented as

$$y = z_1(1 - z_2)(1 - z_3)$$

Note: The true relationship above is completly made up for the pedagogical purposes. The main focus here is the conceptual groundwork of the causal inputs.

The biggest problem with this set up is the nature of the true causal inputs. They are true, yes, by the premise that y is exact and by the assumption that the universe is explicable with mathematics. Yet, the inputs are *unobserable* and *unaccessible.* The features are nice to consider, but we cannot go any further in extracting data to be able to make any predictions.

Please see the following lecture for a continuation of this discussion