Math 390.4   1/29/18 Lec 1

abstraction/
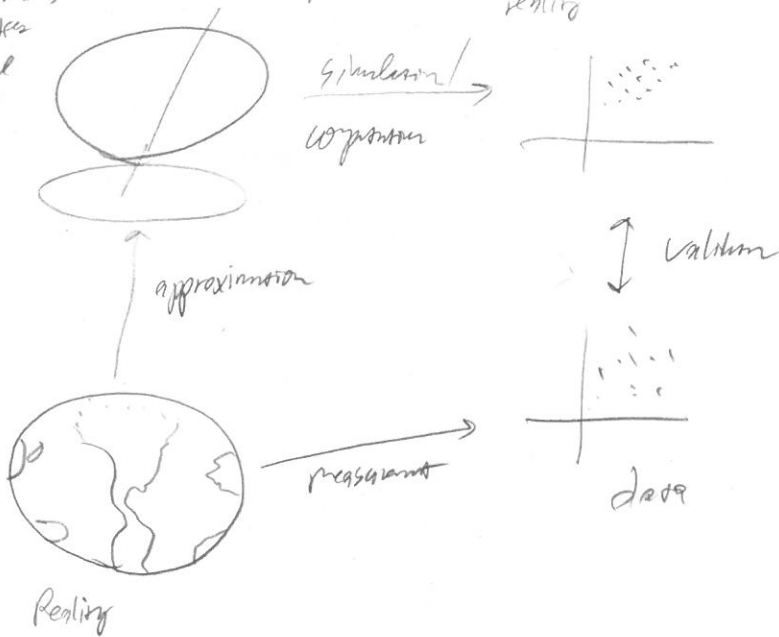
"Models are approximations to reality / absolute truth / system / phenomenon.

— model airplane ——→ airplane ————————————→ allows engineers to evaluate the real product
— map ——————→ locations of streets in a city ——→ allows you to find your way
— wind tunnel ——————→ fast moving air ————————→ allows for testing vehicles traveling at high speeds
— early to bed, early to rise, makes a man healthy wealthy & wise ——→ success of human beings

George Box   "All Models are wrong but some are useful"
                                          ↑   ↑
Models are useful for two general goals: because they're   they serve some
                                             not   useful function
① Prediction → what will happen in the   reality   even though they're
                real world under set   not reality
                of conditions
② Explanation →
what makes
this model
tick?

approximation

simulation/
comparison ——————→

validation

Reality

measurement ——————→

data

"Data" is the result of system or phenomenon which is "measured".

"Validation" is the comparison of the "data" to the "prediction" from the model.
If data and predictions are "close", the model is "good". If not,
we can rebuild/improve the model. The process is usually iterative.

Consider the model

"Early to bed, early to rise makes a man healthy wealthy & wise."

What is this saying?

$$\begin{bmatrix} health \\ wealth \\ wisdom \end{bmatrix} = f\left(\begin{array}{c} bedtime \\ waketime \end{array} \right)$$

outputs   inputs

~every model allows~ for input(s) and output(s)
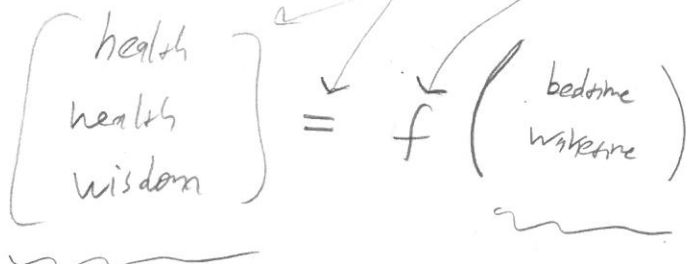
a means of combining the inputs to produce the outputs i.e. a functional relationship

What goes on inside of f?

Well, the earlier the bedtime $\Rightarrow$ all 3 outputs increase, the earlier the waketime $\Rightarrow$ all 3 outputs increase

This aphorism is extremely imprecise / ambiguous!

Why?

— bedtime ill-defined
— waketime ''
— health ''
— wealth ''
— wisdom ''
— f ''

Why function? It wouldn't make sense to get two different versions of health, or would it? (we'll see soon enough)

How to make this precise?   How to measure?

— define bedtime as 24 hr time (avg?)   (b)
—   ''   ''   ''   ''   (avg?)   (w)
— define health as longevity   (h)
— define wealth as net worth at age 60   (h)
— define wisdom as... good luck! (create a test and measure a score!)   (s)
— provide functional form of f

$\Rightarrow$ Mathematical model $\subset$ models

where inputs & outputs are numerical quantities related by an equals sign
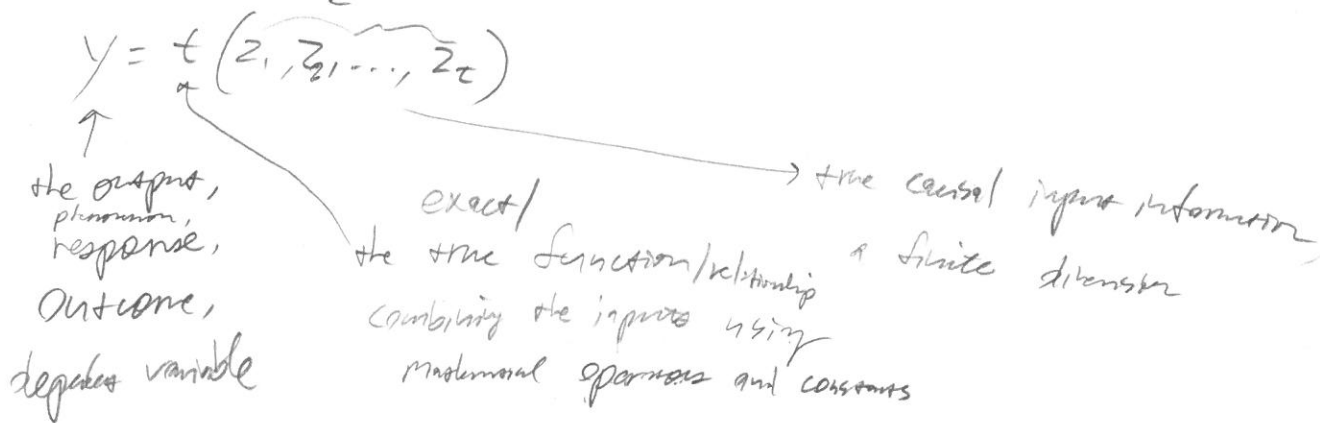
models

mathematical

Mathematical models are ideas, not concrete such as the model airplane.

We've been building mathematical models since ≈ 2000 BC. We are pretty good at them:

$$F = ma, \quad E = mc^2$$

↑ ↑↑   ↑   ↑↑

output inputs

Physics, Chemistry, Economics, ... all disciplines use these.

This rests on an assumption: that the world is explicable mathematically. We will agree for the purposes of this course to ~~believe~~ this.

In which case, all phenomena can be described via:

$$y = \underset{a}{t}\left(\overset{\vec{z}}{\overbrace{z_1, z_2, \ldots, z_{\tau}}}\right)$$

↑

the output,
phenomenon,
response,
outcome,
dependent variable

exact/
the true function/relationship
combining the inputs using
mathematical operators and constants

→ true causal input information
a finite dimension

For the purposes of this class we will assume this is one-dimensional

Note that we assume y exactly equals $t(\vec{z})$ which means the phenomenon is perfectly explicable

there are those that ~~debate~~ determinism, but it doesn't matter for the purposes of this class

e.g. the canonical example is a model which explains "creditworthiness" for a credit application for bob. Here

$$y \in \{ \text{creditworthy}, \text{uncreditworthy} \}$$

If we build a mathematical model, y must be a #.

$$\Rightarrow y \in \{0, 1\} \underset{\text{creditworthy}}{=} \text{Creditworthy (pay back loan)}$$
$$= y \quad \text{(the output space)}$$

The inputs, $z_1, z_2, \ldots z_t$ are characteristics about a person.

What are the the characteristics that are causal? I don't know, but e.g.

z_1: bob has sufficient funds to pay back loan at the time it is due $\in \{0, 1\}$

z_2: bob has unforeseen emergencies occur during this time $\in \{0, 1\}$

z_3: bob has criminal intentions $\in \{0, 1\}$

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

If you had z_1, z_2, z_3 and t you can explain y exactly? YES.

Do you have z_1, z_2, z_3? No... they are inputs assessed in the future and the future hasn't happened yet or they are impossible to assess. Criminal intention? How do you measure that??

Further, in general, we don't know t either. In the mock situation here, we do. (Secondary problem)

Primary problem is you don't observe the truly causal input information.

What's the next best thing? Get information/inputs that best approximate the z's.

x_1: Salary. How to measure; historical avg. salary?

x_2: previous loan repayments. How to measure: did they ever miss a payment?

x_3: historical criminal record. How to measure: previous crime type?