TAM 598 Lecture 1:

Introduction to Predictive Modeling

Announcements:

- course website
- join course slack
- next week: online only

b Machine Learning – making models out of data

examples: regression is a statistical / ML approach to model relationship between variables

goal: understand how one or more independent variables (predictors, features) influence a dependent variable (outcome, response)

eg)

features		response
characteristics of house	$\longrightarrow$	price
	mathematical model	

⊛ CAUSALITY

	feature		response	
Model 1 :	character of house	⟶	price	causal model
Model 2 :	price	⟶	character	statistical correlation

A causal model attempts to capture the mechanisms that govern a phenomenon

scientific ML → aim to make our models causal

J. Pearl → formalized framework for causality to distinguish between correlation & causation

"Structural causal models" (SCMs) have two components:

- set of variable:

  endogenous - variables that our model is trying to explain

  exogenous - variables that are needed

- a set of functions that give value to the variables

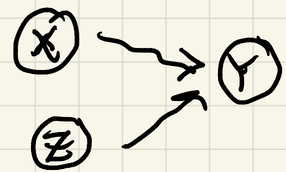
From the two we can construct a directed, acyclic graph (DAG) that represents the SCM and establishes causality

variables → nodes
causal correlations → directed edges

Example: asthma model

treatment X
lung function Y } goal: causal relationship
local air pollution Z

$$Y = f(X, Z)$$



(*) predictive modeling – models that accurately predict outcomes

Not enough to have an excellent causal model, as the model may require inputs that are unknown or random. This introduces uncertainty into predictions examples of uncertain inputs:

- initial values of model parameter
- initial conditions of an ODE
- boundary conditions of a PDE
- value of an experimental measurement that hasn't been done
- mathematical form of the model itself

We categorize uncertainty into two types:

- ALEATORIC - true, inherent randomness (cannot be reduced)
- EPISTEMIC - uncertainty due to a lack of knowledge. Can be reduced by getting more data, improve measurements

probabilistic modeling can deal w/ both.

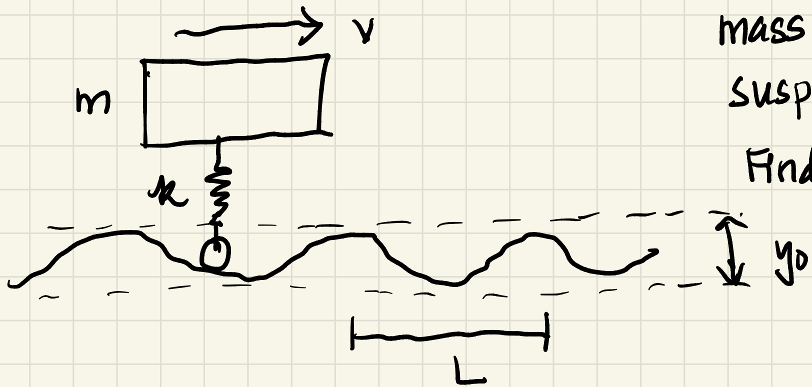
⊛ <u>scientific</u> vs traditional ML

- want causal models
- need to deal w/ uncertainty
- we have underlying physics ("inductive bias", physics-informed)

UNCERTAINTY PROPAGATION PROBLEM

given a causal model that relates input variables to output variables, and given uncertainty in the inputs (some or all), how can we predict the output and quantify the uncertainty in the prediction?

Example: trailer on a rough road



mass m, speed v

suspension spring constant k

Find: vibration amplitude X

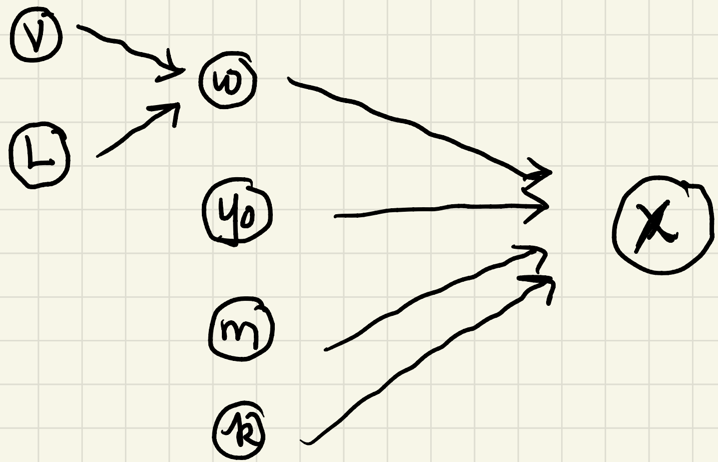
solution: $$X = \left| \frac{k y_0}{k - m\omega^2} \right|$$

$$\boxed{\omega = 2\pi v / L}$$

$$\int p(\omega)\, d\omega = 1$$

$$\int p(x)\, dx = 1$$

Draw the causal graph (DAG):



solving uncertainty propagation problem :

- requires that we estimate / quantify probability statements for all uncertain variables (not easy!)

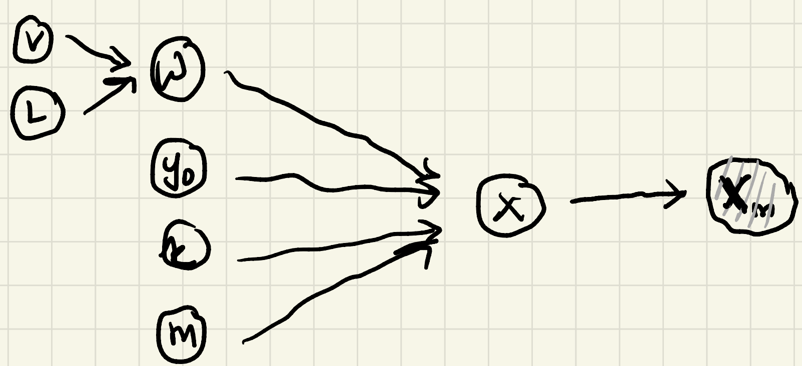
- then propagate the uncertainty through the causal model to characterize the uncertainty in the output

one approach : sampling via Monte Carlo (MCMC)

THE MODEL CALIBRATION PROBLEM :

- the inverse of the uncertainty prop. problem, "the inverse problem"
- you observe a quantity that the model predicts, and you want to characterize how the observation changes the state of knowledge about model parameters

Example: trailer, measure $X = X_m$

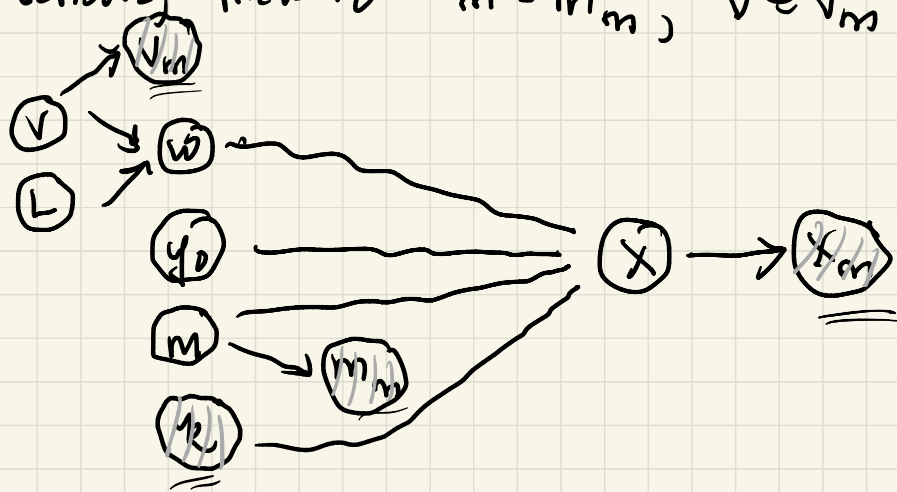


calibration problem: given x_m, find v, L, k, m, y_0

This problem is ill-defined

Need to improve our state of knowledge

- call manufacturer, $k = k$ w/ no uncertainty
- use sensors, measure $m = m_m$, $V = V_m$



Now, we can find L, y_0

Formally: how to solve an inverse problem

1. Specify all structural equations represented by the causal graph by introducing unknown parameters as required.

   → here we have all structural equations except the ones that connect the sensors to the variables (called the measurement model or the likelihood)

2. Quantify our prior knowledge about unknown parameters by assigning them probability densities.

   → Resulting model is called a probabilistic graph, or Bayesian network, or hierarchical Bayesian model

Formally: how to solve an inverse problem

③ Use Bayes rule to condition our prior knowledge of observations. The updated knowledge is called our posterior knowledge.

→ This is Bayesian inference

④ The posterior is rarely analytically available. Practically, we sample from the posterior using Markov chain Monte Carlo (MCMC) or variational inference.