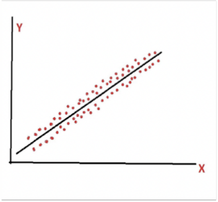
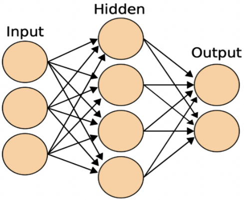
Seminar Series Kickoff

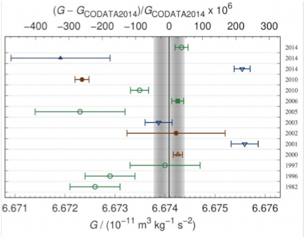
Various Remarks, “Data & Modeling”

Conor Rowan

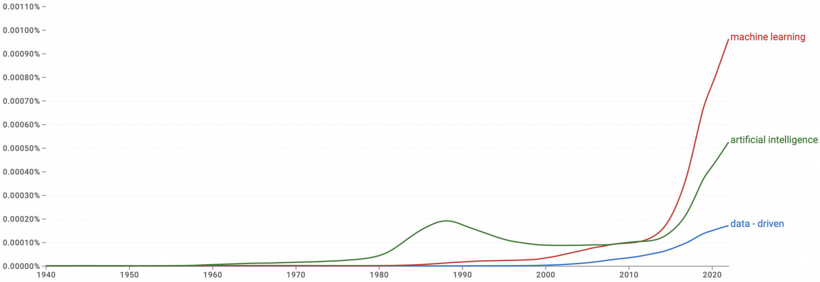
University of Colorado Boulder, Aerospace Engineering





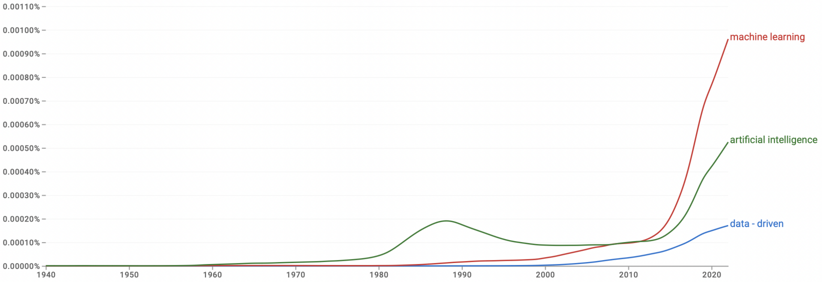


The Data Revolution



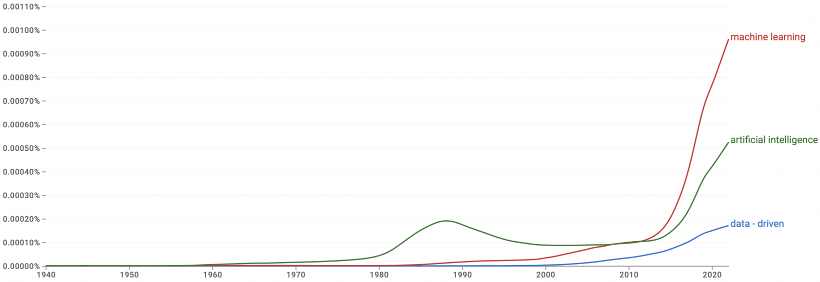
- "Big data" has emerged as a new paradigm in science

The Data Revolution



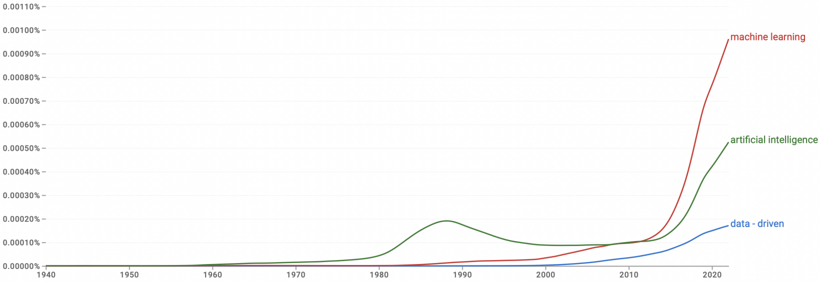
- "Big data" has emerged as a new paradigm in science
- Lots of success stories: AlphaFold, AlphaGo, DeepBlue, ChatGPT, etc.

The Data Revolution



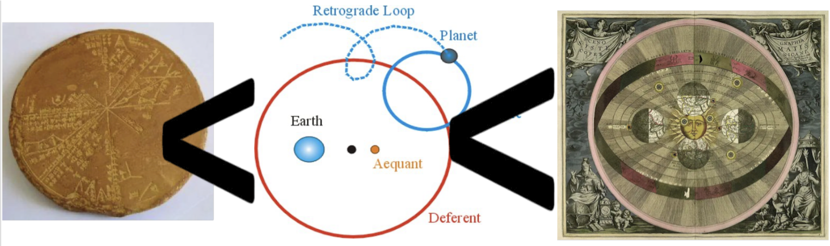
- "Big data" has emerged as a new paradigm in science
- Lots of success stories: AlphaFold, AlphaGo, DeepBlue, ChatGPT, etc.
- Data-driven models are often contrasted with more traditional rule- and physics-based methods–**does this make sense?**

The Data Revolution



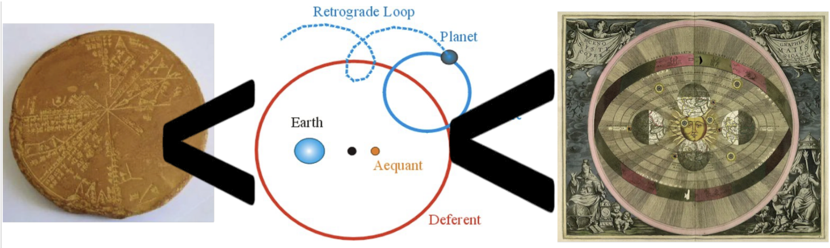
- "Big data" has emerged as a new paradigm in science
- Lots of success stories: AlphaFold, AlphaGo, DeepBlue, ChatGPT, etc.
- Data-driven models are often contrasted with more traditional rule- and physics-based methods–**does this make sense?**
- Bertrand Russell: "the point of philosophy is to start with something so simple as to seem not worth stating, and to end with something so paradoxical that no one will believe it."

Data-driven, as opposed to what?



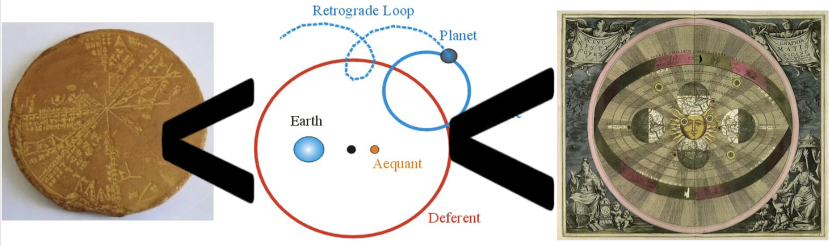
- We off-load the hard work of constructing theories to past generations

Data-driven, as opposed to what?



- We off-load the hard work of constructing theories to past generations
- Obviously data contributes to the development of basic theory

Data-driven, as opposed to what?



- We off-load the hard work of constructing theories to past generations
- Obviously data contributes to the development of basic theory
- Kuhn on theory construction: *"In the absence of a paradigm . . . all of the facts that could possibly pertain to the development of a given science are likely to seem equally relevant. As a result, early fact gathering is a far more nearly random activity than the one that subsequent scientific development makes familiar."* [1]

Theories & data

Egyptian myth

"When bodies were placed in sarcophagi the flesh disappeared, leaving only the bones: so the body was thought to have been absorbed, or eaten, by the stone . . . this is an interesting case of a generalized observation serving as evidence for a magical or mythical view." [2]

Egyptian myth

"When bodies were placed in sarcophagi the flesh disappeared, leaving only the bones: so the body was thought to have been absorbed, or eaten, by the stone . . . this is an interesting case of a generalized observation serving as evidence for a magical or mythical view." [2]

On astrology

"It was, therefore, a matter of plain observation that there were correspondences or harmonies between happenings in the Heavens . . . and happenings on the Earth and in man himself." [3]

Theories & data

Egyptian myth

"When bodies were placed in sarcophagi the flesh disappeared, leaving only the bones: so the body was thought to have been absorbed, or eaten, by the stone . . . this is an interesting case of a generalized observation serving as evidence for a magical or mythical view." [2]

On astrology

"It was, therefore, a matter of plain observation that there were correspondences or harmonies between happenings in the Heavens . . . and happenings on the Earth and in man himself." [3]

Medieval challenges to Greek science

"[The] attack on science took the form of 'occasionalism'–the doctrine that whatever happens is a singular occasion, governed not by any laws of nature but directly by the will of God." [4]

Assume we have a (good) theory

- For the sake of argument, assume that the theory/model is "correct" (e.g. gravity is an inverse square law, continuum mechanics is the right way to model matter, etc.)

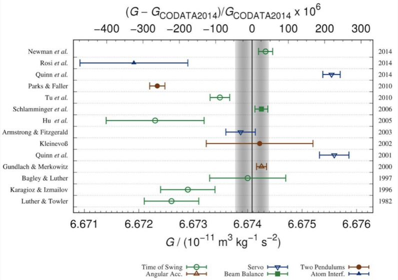


Figure: Measuring the gravitational constant G in $F = GMm/r^2$

Assume we have a (good) theory

- For the sake of argument, assume that the theory/model is "correct" (e.g. gravity is an inverse square law, continuum mechanics is the right way to model matter, etc.)
- *Models still have empirical parameters which need to be estimated from data*

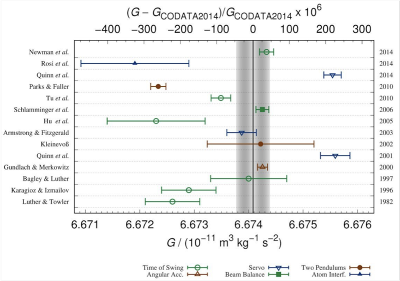


Figure: Measuring the gravitational constant G in $F = GMm/r^2$

Assume we have a (good) theory

- For the sake of argument, assume that the theory/model is "correct" (e.g. gravity is an inverse square law, continuum mechanics is the right way to model matter, etc.)
- *Models still have empirical parameters which need to be estimated from data*
- **So what is the difference between physics-based and data-driven modeling?**

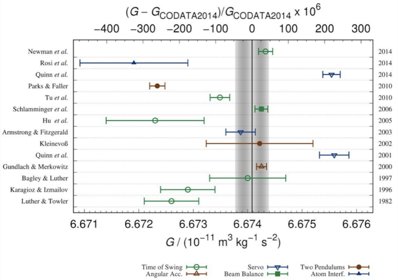


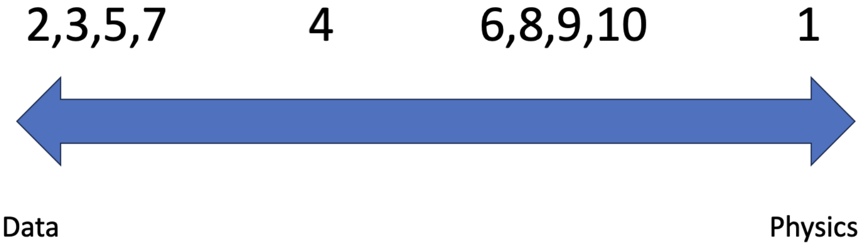
Figure: Measuring the gravitational constant G in $F = GMm/r^2$

Outlining and classifying some techniques in modeling

- 1) Numerical simulations, 2) regression, 3) classification, 4) equation discovery, 5) generative models, 6) reduced-order models, 7) surrogate models, 8) physics-informed machine learning, 9) data assimilation, 10) constitutive modeling

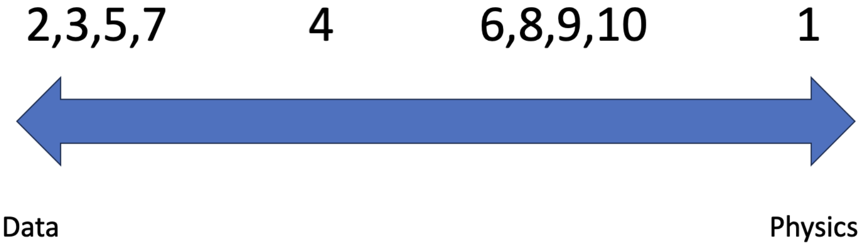
Outlining and classifying some techniques in modeling

- 1) Numerical simulations, 2) regression, 3) classification, 4) equation discovery, 5) generative models, 6) reduced-order models, 7) surrogate models, 8) physics-informed machine learning, 9) data assimilation, 10) constitutive modeling



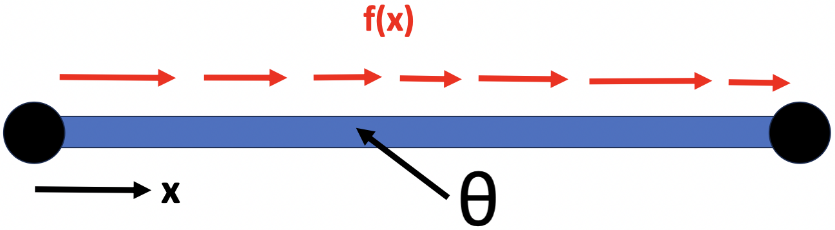
Outlining and classifying some techniques in modeling

- 1) Numerical simulations, 2) regression, 3) classification, 4) equation discovery, 5) generative models, 6) reduced-order models, 7) surrogate models, 8) physics-informed machine learning, 9) data assimilation, 10) constitutive modeling



- **Can we make this intuition more precise?**

Canonical mechanics problem



$$\theta \frac{\partial^2 u}{\partial x^2} = f(x), \quad u(0) = u(1) = 0$$

$$u(x) = \int^x \int^y \frac{f(z)}{\theta} dzdy + cx := \mathcal{G}(f(x); \theta)(x)$$

- How do we estimate the material parameter θ if it is unknown?

Constitutive modeling

- Assume that we have data $\hat{u}(x)$ for a given force $f(x)$ at discrete points $x_1, x_2, \ldots, x_N$:

$$\mathcal{L}(\theta) = \frac{1}{2}\sum_{i=1}^{N}\Big(\mathcal{G}(f(x);\theta)(x_i) - \hat{u}(x_i)\Big)^2 \implies \theta = \underset{\theta}{\text{argmin}}\ \mathcal{L}(\theta)$$

Constitutive modeling

- Assume that we have data $\hat{u}(x)$ for a given force $f(x)$ at discrete points $x_1, x_2, \dots, x_N$:

$$\mathcal{L}(\theta) = \frac{1}{2}\sum_{i=1}^{N}\Big(\mathcal{G}(f(x);\theta)(x_i) - \hat{u}(x_i)\Big)^2 \implies \theta = \underset{\theta}{\operatorname{argmin}}\ \mathcal{L}(\theta)$$

- Model (operator) $\mathcal{G}$ is parameterized in terms of θ, takes the input $f(x)$ to the output $u(x)$

Constitutive modeling

- Assume that we have data $\hat{u}(x)$ for a given force $f(x)$ at discrete points $x_1, x_2, \ldots, x_N$:

$$\mathcal{L}(\theta) = \frac{1}{2}\sum_{i=1}^{N}\Big(\mathcal{G}(f(x);\theta)(x_i) - \hat{u}(x_i)\Big)^2 \implies \theta = \underset{\theta}{\operatorname{argmin}}\ \mathcal{L}(\theta)$$

- Model (operator) $\mathcal{G}$ is parameterized in terms of θ, takes the input $f(x)$ to the output $u(x)$
- The model is "trained" by minimizing discrepancy with data

Constitutive modeling

- Assume that we have data $\hat{u}(x)$ for a given force $f(x)$ at discrete points $x_1, x_2, \ldots, x_N$:

$$\mathcal{L}(\theta) = \frac{1}{2}\sum_{i=1}^{N}\Big(\mathcal{G}(f(x);\theta)(x_i) - \hat{u}(x_i)\Big)^2 \implies \theta = \underset{\theta}{\operatorname{argmin}}\ \mathcal{L}(\theta)$$

- Model (operator) $\mathcal{G}$ is parameterized in terms of θ, takes the input $f(x)$ to the output $u(x)$
- The model is "trained" by minimizing discrepancy with data
- Something like this has gone on behind the scenes of every physics-based model

Constitutive modeling

- Assume that we have data $\hat{u}(x)$ for a given force $f(x)$ at discrete points $x_1, x_2, \ldots, x_N$:

$$\mathcal{L}(\theta) = \frac{1}{2}\sum_{i=1}^{N}\Big(\mathcal{G}(f(x);\theta)(x_i) - \hat{u}(x_i)\Big)^2 \implies \theta = \underset{\theta}{\operatorname{argmin}}\ \mathcal{L}(\theta)$$

- Model (operator) $\mathcal{G}$ is parameterized in terms of θ, takes the input $f(x)$ to the output $u(x)$
- The model is "trained" by minimizing discrepancy with data
- Something like this has gone on behind the scenes of every physics-based model
- (Experimentalists can correct me if this is the wrong way of thinking about things)

Same problem, different approach

- We have the same data (and boundary conditions) but no knowledge of a governing equation
- Want to fit an input-output relationship–we are going to need more parameters!

$$u(x) = \mathcal{N}(f(x); \theta_1, \theta_2, \dots)(x)$$

$$\mathcal{L}(\underline{\theta}) = \frac{1}{2}\sum_{i=1}^{N}\Big(\mathcal{N}(f(x); \underline{\theta})(x_i) - \hat{u}(x_i)\Big)^2 \implies \underline{\theta} = \underset{\underline{\theta}}{\text{argmin}}\ \mathcal{L}(\underline{\theta})$$

Same problem, different approach

- We have the same data (and boundary conditions) but no knowledge of a governing equation
- Want to fit an input-output relationship–we are going to need more parameters!

$$u(x) = \mathcal{N}(f(x); \theta_1, \theta_2, \dots)(x)$$

$$\mathcal{L}(\underline{\theta}) = \frac{1}{2}\sum_{i=1}^{N}\Big(\mathcal{N}(f(x); \underline{\theta})(x_i) - \hat{u}(x_i)\Big)^2 \implies \underline{\theta} = \underset{\underline{\theta}}{\operatorname{argmin}}\ \mathcal{L}(\underline{\theta})$$

- This is just a funky regression problem–$f(x)$ is the "independent variable" and $u(x)$ is the "dependent variable"

Same problem, different approach

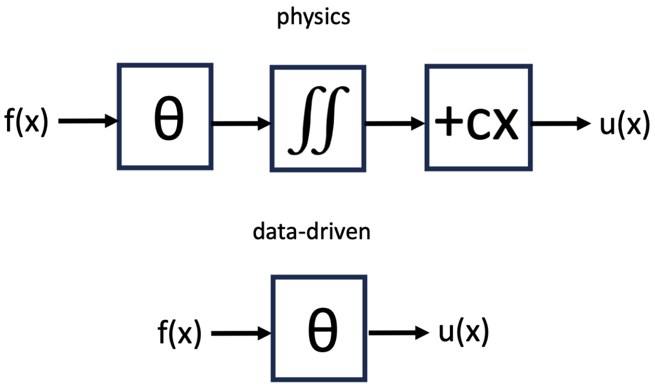
- We have the same data (and boundary conditions) but no knowledge of a governing equation
- Want to fit an input-output relationship–we are going to need more parameters!

$$u(x) = \mathcal{N}(f(x); \theta_1, \theta_2, \dots)(x)$$

$$\mathcal{L}(\underline{\theta}) = \frac{1}{2}\sum_{i=1}^{N}\Big(\mathcal{N}(f(x); \underline{\theta})(x_i) - \hat{u}(x_i)\Big)^2 \implies \underline{\theta} = \underset{\underline{\theta}}{\text{argmin}}\ \mathcal{L}(\underline{\theta})$$

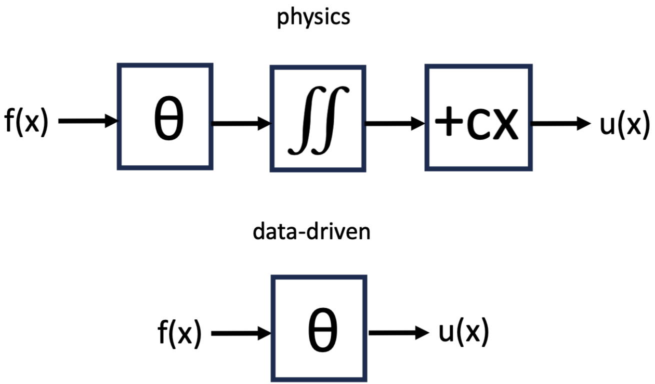
- This is just a funky regression problem–$f(x)$ is the "independent variable" and $u(x)$ is the "dependent variable"
- Choosing the structure of $\mathcal{N}$ is like choosing a regression model (linear, quadratic, neural network, etc.)

What's the difference?



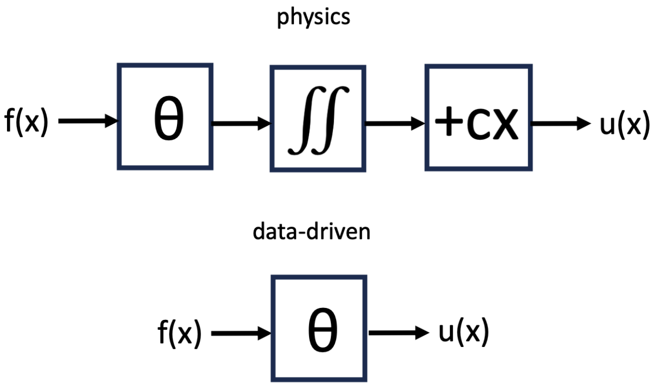
- In both cases we are using data to "train" a parameterized operator which maps inputs to outputs

What's the difference?



- In both cases we are using data to "train" a parameterized operator which maps inputs to outputs
- Physics model has fewer parameters (?)

What's the difference?



- In both cases we are using data to "train" a parameterized operator which maps inputs to outputs
- Physics model has fewer parameters (?)
- **Why do we expect the physics model to generalize better?**

A complicated story with a simple missing piece

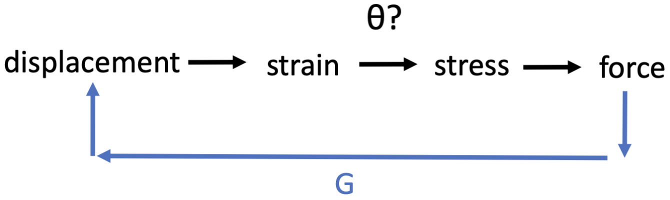
$$\frac{f(x)}{\theta} = \frac{\partial^2 u}{\partial x^2} \implies u(x) = \mathcal{G}(f(x), \theta)(x)$$

- "Applied forces create curvature in the displacement field at a rate controlled by an empirical parameter θ"
- This knowledge is encoded in the operator $\mathcal{G}$, only a small and simple piece of the "story" needs to be filled in
- Small amount of data required

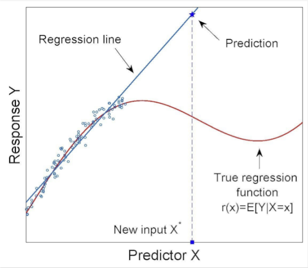
A complicated story with a simple missing piece

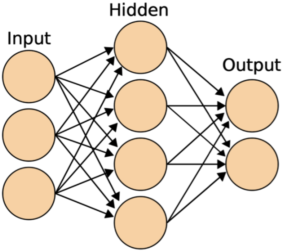
$$\frac{f(x)}{\theta} = \frac{\partial^2 u}{\partial x^2} \implies u(x) = \mathcal{G}(f(x), \theta)(x)$$

- "Applied forces create curvature in the displacement field at a rate controlled by an empirical parameter θ"
- This knowledge is encoded in the operator $\mathcal{G}$, only a small and simple piece of the "story" needs to be filled in
- Small amount of data required



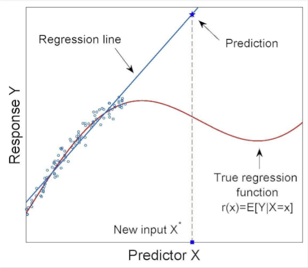
The data-driven model may not learn the story

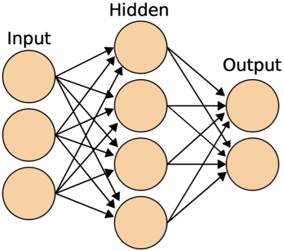




- For the data-driven model to truly generalize, it needs to learn that forces cause curvature in the displacement

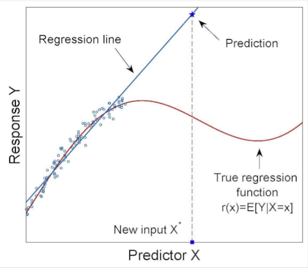
The data-driven model may not learn the story

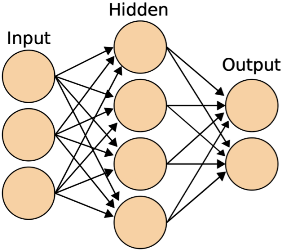




- For the data-driven model to truly generalize, it needs to learn that forces cause curvature in the displacement
- There are lots of ways to fit the relationship between $f_i(x)$ and $u_i(x)$ for $i = 1, 2, \ldots, N < \infty$ without understanding this

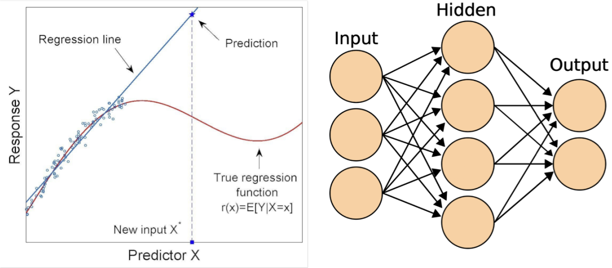
The data-driven model may not learn the story





- For the data-driven model to truly generalize, it needs to learn that forces cause curvature in the displacement
- There are lots of ways to fit the relationship between $f_i(x)$ and $u_i(x)$ for $i = 1, 2, \ldots, N < \infty$ without understanding this
- The data-driven operator $\mathcal{N}$ is "unstructured" in the sense that it doesn't encode any principles/laws

The data-driven model may not learn the story



- For the data-driven model to truly generalize, it needs to learn that forces cause curvature in the displacement
- There are lots of ways to fit the relationship between $f_i(x)$ and $u_i(x)$ for $i = 1, 2, \ldots, N < \infty$ without understanding this
- The data-driven operator $\mathcal{N}$ is "unstructured" in the sense that it doesn't encode any principles/laws
- Needs a lot of data to even have a chance of being useful

My argument

- There is no sharp distinction between data-driven and physics-based models

My argument

- There is no sharp distinction between data-driven and physics-based models
- **Any distinction between the two relies not on whether data is used, or how data is used, rather on how much structure the operator taking inputs to outputs has**

My argument

- There is no sharp distinction between data-driven and physics-based models
- **Any distinction between the two relies not on whether data is used, or how data is used, rather on how much structure the operator taking inputs to outputs has**
- The amount of structure in the operator lies on a spectrum (it is not binary)

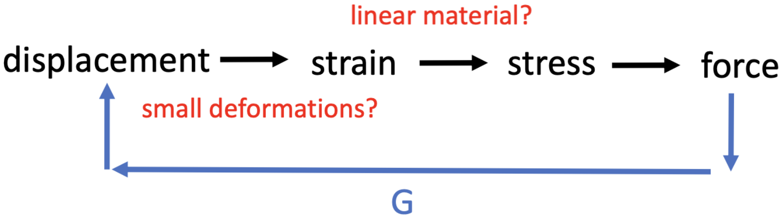
My argument

- There is no sharp distinction between data-driven and physics-based models
- **Any distinction between the two relies not on whether data is used, or how data is used, rather on how much structure the operator taking inputs to outputs has**
- The amount of structure in the operator lies on a spectrum (it is not binary)
- This should encourage machine learning skeptics to rethink their argument of "it's just curve-fitting"

My argument

- There is no sharp distinction between data-driven and physics-based models
- **Any distinction between the two relies not on whether data is used, or how data is used, rather on how much structure the operator taking inputs to outputs has**
- The amount of structure in the operator lies on a spectrum (it is not binary)
- This should encourage machine learning skeptics to rethink their argument of "it's just curve-fitting"
- Maybe this tempers optimism as well? [5]

Another point on generalization



- When an operator is structured, it is built up from a number of steps which have clear assumptions
- Knowing the assumptions helps assess/predict generalization properties of the model

Thought experiment 1

- Suppose you are an executive at an international company, and you need to communicate to employees in another country that their weekly pay will be late by 1 day, but it will be double when it arrives

Thought experiment 1

- Suppose you are an executive at an international company, and you need to communicate to employees in another country that their weekly pay will be late by 1 day, but it will be double when it arrives
- The message is dictated to a secretary, written down, sent in the mail, translated to the local language, then read aloud by a manager

Thought experiment 1

- Suppose you are an executive at an international company, and you need to communicate to employees in another country that their weekly pay will be late by 1 day, but it will be double when it arrives
- The message is dictated to a secretary, written down, sent in the mail, translated to the local language, then read aloud by a manager
- You trust the secretary, the mail system, the translator, but are not so sure about the manager (apparently she is prone to exaggeration and "has it out" for the company)

Thought experiment 1

- Suppose you are an executive at an international company, and you need to communicate to employees in another country that their weekly pay will be late by 1 day, but it will be double when it arrives
- The message is dictated to a secretary, written down, sent in the mail, translated to the local language, then read aloud by a manager
- You trust the secretary, the mail system, the translator, but are not so sure about the manager (apparently she is prone to exaggeration and "has it out" for the company)
- You learn that upon receiving the message, the employees were furious and threatened to quit

Thought experiment 1

- Suppose you are an executive at an international company, and you need to communicate to employees in another country that their weekly pay will be late by 1 day, but it will be double when it arrives
- The message is dictated to a secretary, written down, sent in the mail, translated to the local language, then read aloud by a manager
- You trust the secretary, the mail system, the translator, but are not so sure about the manager (apparently she is prone to exaggeration and "has it out" for the company)
- You learn that upon receiving the message, the employees were furious and threatened to quit
- *What happened? What would you expect to happen in a similar situation in the future?*

Thought experiment 2

- Suppose you are an executive at an international company, and you need to communicate to employees in another country that their weekly pay will be late by 1 day, but it will be double when it arrives

Thought experiment 2

- Suppose you are an executive at an international company, and you need to communicate to employees in another country that their weekly pay will be late by 1 day, but it will be double when it arrives
- You dictate the message to a secretary who you have never met before, and you have no knowledge of the chain of events which lead to the employees receiving the message

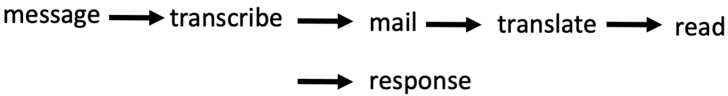
Thought experiment 2

- Suppose you are an executive at an international company, and you need to communicate to employees in another country that their weekly pay will be late by 1 day, but it will be double when it arrives
- You dictate the message to a secretary who you have never met before, and you have no knowledge of the chain of events which lead to the employees receiving the message
- You learn that upon receiving the message, the employees were furious and threatened to quit

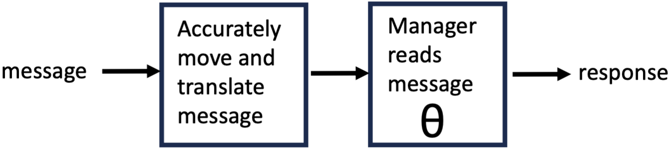
Thought experiment 2

- Suppose you are an executive at an international company, and you need to communicate to employees in another country that their weekly pay will be late by 1 day, but it will be double when it arrives
- You dictate the message to a secretary who you have never met before, and you have no knowledge of the chain of events which lead to the employees receiving the message
- You learn that upon receiving the message, the employees were furious and threatened to quit
- *What happened? What would you expect to happen in a similar situation in the future?*

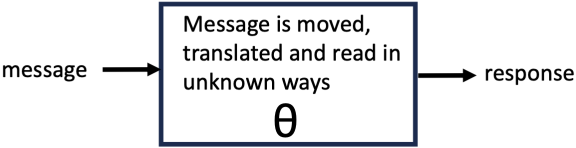
Visualizing the two cases



Case 1



Case 2



Interpretation

- This is like the difference between physics-based and data-driven models–knowing what is going on inside the operator taking inputs to outputs allows for interpretation and insight into generalization

Interpretation

- This is like the difference between physics-based and data-driven models–knowing what is going on inside the operator taking inputs to outputs allows for interpretation and insight into generalization
- This is what I mean by “structure”

Interpretation

- This is like the difference between physics-based and data-driven models–knowing what is going on inside the operator taking inputs to outputs allows for interpretation and insight into generalization
- This is what I mean by "structure"
- With Case 1, you have a sense of what to expect if you send another message to the same place, but not what to expect if you send a message to a different country–the domain of applicability is well-established and the model is likely robust within that domain

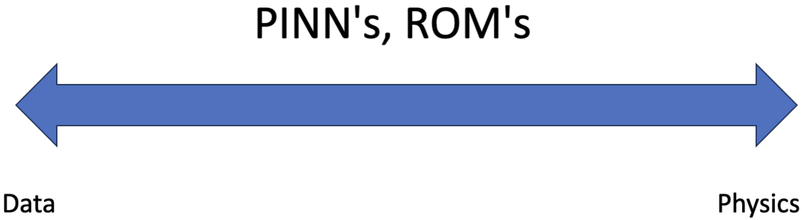
Interpretation

- This is like the difference between physics-based and data-driven models–knowing what is going on inside the operator taking inputs to outputs allows for interpretation and insight into generalization
- This is what I mean by "structure"
- With Case 1, you have a sense of what to expect if you send another message to the same place, but not what to expect if you send a message to a different country–the domain of applicability is well-established and the model is likely robust within that domain
- You would need to send many messages and detect patterns to accomplish this for Case 2

Interpretation

- This is like the difference between physics-based and data-driven models–knowing what is going on inside the operator taking inputs to outputs allows for interpretation and insight into generalization
- This is what I mean by "structure"
- With Case 1, you have a sense of what to expect if you send another message to the same place, but not what to expect if you send a message to a different country–the domain of applicability is well-established and the model is likely robust within that domain
- You would need to send many messages and detect patterns to accomplish this for Case 2
- When we don't know what mediates inputs and outputs, the "black box" is our only hope...

Applying this argument to an example



- Physics-informed neural networks (PINN's) and reduced-order models (ROM's) occupy an intermediate position between data and physics
- Can use the "operator structure" argument to justify this
- We will consider these two examples

PINN's (actually Deep Operator Network)

- Physics-informed Deep Operator Networks [6,7] build physics into a data-driven model in an unusual way
- Using a particular discretization of the solution (neural networks), tune parameters such that data is matched AND governing equations are satisfied:

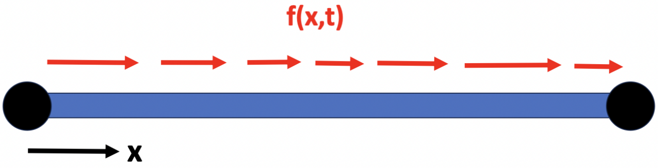
$$u(x) = \sum_i c_i(f(x);\theta)\Psi_i(x;\theta)$$

$$\mathcal{L}^{data}(\theta) = \frac{1}{2}\sum_j \Big(\sum_i c_i(f(x);\theta)\Psi_i(x_j;\theta) - \hat{u}(x_j)\Big)^2$$

$$\mathcal{L}^{physics}(\theta) = \frac{1}{2}\sum_j \Big(\sum_i c_i(f(x);\theta)\frac{\partial^2\Psi_i}{\partial x^2}(x_j;\theta) - f(x_j)\Big)^2$$

$$\theta = \underset{\theta}{\operatorname{argmin}}\ \mathcal{L}^{data}(\theta) + \mathcal{L}^{physics}(\theta)$$

ROM – dynamic bar



$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} - f(x,t), \quad u(x,0) = u(0,t) = u(1,t) = 0$$

$$u(x,t) \approx \sum_{i=1}^{N} a_i(t)\Psi_i(x)$$

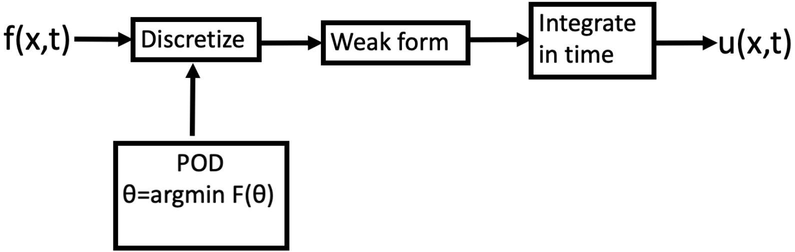
- Cannot write down solution for dynamics of elastic bar, the problem must be solved numerically
- Solution expanded with time varying coefficients $a_i(t)$ scaling given spatial shape functions $\Psi_i(x)$

Proper orthogonal decomposition (POD)

$$\underline{\underline{X}} = \begin{bmatrix} u(x_1, t_1) & u(x_2, t_1) & \dots & u(x_N, t_1) \\ u(x_1, t_2) & u(x_2, t_2) & \dots & u(x_N, t_2) \\ \vdots & \vdots & \dots & \vdots \\ u(x_1, t_T) & u(x_2, t_T) & \dots & u(x_N, t_T) \end{bmatrix}$$

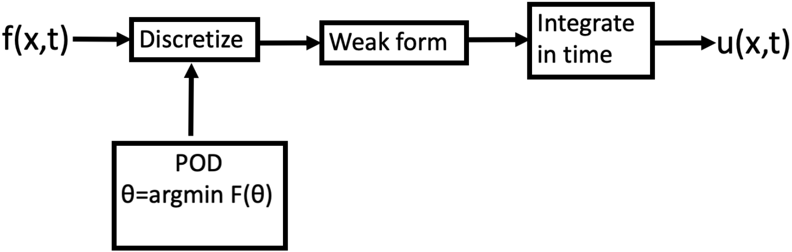
- POD is a common model reduction technique
- Uses data on a dynamical system to find an "optimal" set of spatial shape functions $\Psi_i(x)$ [8]
- Using our same notation, we could say abstractly that $\Psi_i(x) = \Psi_i(x; \theta)$ because they are computed from data

Operator structure with POD



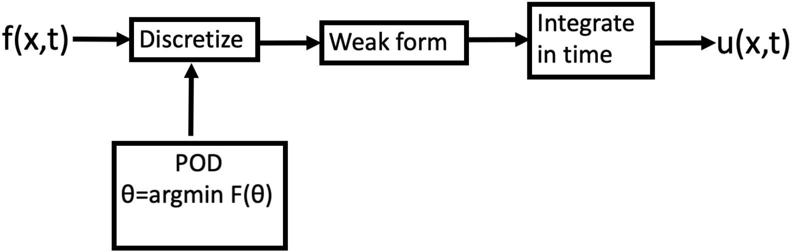
- Data used to compute the spatial shape functions $\Psi_i(x)$

Operator structure with POD



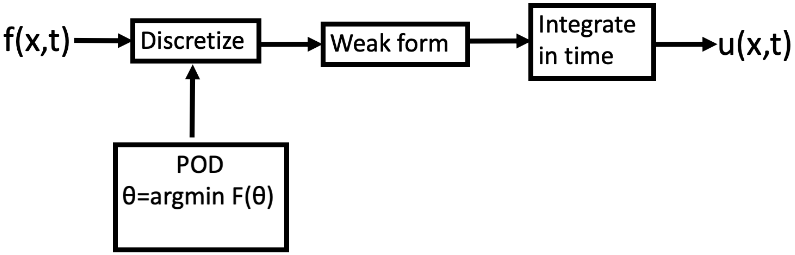
- Data used to compute the spatial shape functions $\Psi_i(x)$
- These inform the “Discretize” step in the operator

Operator structure with POD



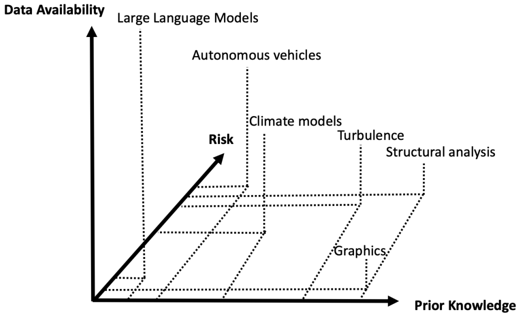
- Data used to compute the spatial shape functions $\Psi_i(x)$
- These inform the "Discretize" step in the operator
- Assumption is that shape functions computed from one forcing will "generalize" to others

Operator structure with POD



- Data used to compute the spatial shape functions $\Psi_i(x)$
- These inform the "Discretize" step in the operator
- Assumption is that shape functions computed from one forcing will "generalize" to others
- Data plays a more complex role in forming the solution operator than determining a single constitutive parameter, but physics still enforced

Final thought: classifying models along three dimensions



- Talked about data and generalization...are there other things to consider?
- Risk refers to the "cost" associated with model error–this depends on the application and determines the importance of generalization
- More prior knowledge of the problem → more "hard-coded" structure in operator → less data needed → more predictable generalization

Conclusion

- Physical principles come from data (in some complex way), and the models which physical principles help construct also rely on data

Conclusion

- Physical principles come from data (in some complex way), and the models which physical principles help construct also rely on data
- **No sharp distinction between physics-based and data-driven models**

Conclusion

- Physical principles come from data (in some complex way), and the models which physical principles help construct also rely on data
- **No sharp distinction between physics-based and data-driven models**
- **Differences are found in how much prior knowledge is baked into the operator mapping inputs to outputs**

Conclusion

- Physical principles come from data (in some complex way), and the models which physical principles help construct also rely on data
- **No sharp distinction between physics-based and data-driven models**
- **Differences are found in how much prior knowledge is baked into the operator mapping inputs to outputs**
- More structured operators hard-code rules and will tend to generalize better and more predictably

Conclusion

- Physical principles come from data (in some complex way), and the models which physical principles help construct also rely on data
- **No sharp distinction between physics-based and data-driven models**
- **Differences are found in how much prior knowledge is baked into the operator mapping inputs to outputs**
- More structured operators hard-code rules and will tend to generalize better and more predictably
- When deploying a model in the "real world," it is necessary to consider risk

Conclusion

- Physical principles come from data (in some complex way), and the models which physical principles help construct also rely on data
- **No sharp distinction between physics-based and data-driven models**
- **Differences are found in how much prior knowledge is baked into the operator mapping inputs to outputs**
- More structured operators hard-code rules and will tend to generalize better and more predictably
- When deploying a model in the "real world," it is necessary to consider risk
- Data-driven models are best-suited for situations with high data availability, minimal prior knowledge, and low risk

Thanks!

Questions? Comments? Concerns?

References

[1] Kuhn, T. S., The structure of Scientific Revolutions, Chicago, Ill: The University of Chicago Press, 2015.
[2] Gregory, R. L., Mind in science: History of explanations in psychology and physics, Penguin, 1993.
[3] Toulmin, S. E., and Goodfield, J., The fabric of the heavens, Chicago: Univ. of Chicago Press, 1999.
[4] Weinberg, S., To explain the world: The discovery of modern science, London: Penguin Books, 2016.
[5] Anderson, C., "The end of theory: The Data Deluge makes the scientific method obsolete," Wired Available: here
[6] Lu, L., Jin, P., Pang, G., Zhang, Z., and Karniadakis, G. E., "Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators," Nature Machine Intelligence, vol. 3, Mar. 2021.

References

[7] Goswami, S., Bora, A., Yu, Y., Karniadakis, G.E., "Physics-Informed Deep Neural Operator Networks," Springer, 2023.
[8] Kutz, N., "Optimal basis elements: the POD expansion," YouTube Available: here

Images, in order of appearance

Neural network, available here
Gravitational constant, available here
Plato's academy, available here
Vienna Circle, available here
Crazy plot, available here
Boring presentation, available here
Babylonian tablet, available here
Ptolemiac system, available here
Copernican system, available here
Generalization error, available here