Computational Models and Empirical Evaluation

Kyle A. Joyce

Department of Political Science
University of California, Davis

August 13, 2014

de Marchi (CUP, 2005)

Goals

- Develop a new modeling framework broadly speaking.
- Introduce computational modeling.

Introduction

- What are the assumptions/parameters of the model?

Introduction

- What are the assumptions/parameters of the model?
- Do the values chosen for the parameters come from qualitative or empirical research, or are they chosen arbitrarily (i.e., for convenience)?

Introduction

- What are the assumptions/parameters of the model?
- Do the values chosen for the parameters come from qualitative or empirical research, or are they chosen arbitrarily (i.e., for convenience)?
- Do the assumptions follow from a consideration of the problem itself, or are they unrelated to the main logic of the model?

Introduction

- Is there any assurance that the results of the model are immune to small perturbations of the parameters?

Introduction

- Is there any assurance that the results of the model are immune to small perturbations of the parameters?
- Is there an equivalence class where the model yields the same results for a neighborhood around the chosen parameters?

Introduction

- Is there any assurance that the results of the model are immune to small perturbations of the parameters?
- Is there an equivalence class where the model yields the same results for a neighborhood around the chosen parameters?
- Or, is the model brittle?

Introduction

- Do the results of the model map directly to a dependent variable, or is the plan to make analogies from the model to the empirical referent?

Introduction

- Do the results of the model map directly to a dependent variable, or is the plan to make analogies from the model to the empirical referent?
- Although toy models (e.g., IPD) have their place in developing intuition, they are difficult to falsify, and even more difficult to build on in a cumulative fashion.

Introduction

- Are the results of the model verified by out-of-sample tests?

Introduction

- Are the results of the model verified by out-of-sample tests?
- The only appropriate out-of-sample tests are: 1) a large-N statistical approach that tests the model directly, 2) a logical implication derived deductively from the model (e.g., one that is novel and uniquely connected to the model).

Introduction

- Curse of dimensionality: is the parameter space of the model too large to span with the available data?

Introduction

- Curse of dimensionality: is the parameter space of the model too large to span with the available data?
- To reduce the size of the parameter space, should you use a domain-specific encoding, provide a feature space, or something else?

Problems: Statistical Models

- A researcher, given time, can produce almost any result that is desired.
- Journals only report positive results and only the final model.
- Any guessing or cheating is never seen in print.

Problems: Statistical Models

- A researcher, given time, can produce almost any result that is desired.
- Journals only report positive results and only the final model.
- Any guessing or cheating is never seen in print.
- Think of every choice made by an empirical modeler as a parameter; results are conditional upon the set of parameter values chosen.
- These parameter spaces are large and thus one cannot have much faith in the final report of in-sample performance.

Problems: Game-Theoretic Models

- Game theorists often cite logical consistency as one benefit of formalizing a model.
- The standard argument is that any pattern of evidence can be matched with some conclusion of a logically <u>inconsistent</u> theory.

Problems: Game-Theoretic Models

- Game theorists often cite logical consistency as one benefit of formalizing a model.
- The standard argument is that any pattern of evidence can be matched with some conclusion of a logically <u>inconsistent</u> theory.
- But one may also achieve any outcome one desires with a <u>consistent</u> theory; all it takes is the right combination of assumptions, solution concepts, etc.

Problems: Game-Theoretic Models

- Game theorists often cite logical consistency as one benefit of formalizing a model.
- The standard argument is that any pattern of evidence can be matched with some conclusion of a logically inconsistent theory.
- But one may also achieve any outcome one desires with a consistent theory; all it takes is the right combination of assumptions, solution concepts, etc.
- Additionally, the class of consistent games that provide any given result is infinite, which means that only some of the models correspond with the real world process.

Problems: Game-Theoretic Models

- When technical assumptions drive model results, one has to question how comfortable one is with the idea that an assumption that is unrelated to the phenomenon in question ends up accounting for the model's brittle results.

Problems: Game-Theoretic Models

- When technical assumptions drive model results, one has to question how comfortable one is with the idea that an assumption that is unrelated to the phenomenon in question ends up accounting for the model's brittle results.
- The fact that one can prove something to be true is not in and of itself useful since there are an infinite number of models that prove any given result is true.

Problems: Game-Theoretic Models

- For any given empirical referent, 10 different researchers might construct 10 different games that all yield the "right" answer.

Problems: Game-Theoretic Models

- For any given empirical referent, 10 different researchers might construct 10 different games that all yield the "right" answer.
- How does one know which models are "right"?

Problems: Game-Theoretic Models

- For any given empirical referent, 10 different researchers might construct 10 different games that all yield the “right” answer.
- How does one know which models are “right”?
- Absent empirical validation, comparing game-theoretic models forces one to argue about assumptions.

Problems: Computational Models

- You can write a computational model (infinitely many) that will produce any given result; with the right components, parameter settings, etc.

Problems: Computational Models

- You can write a computational model (infinitely many) that will produce any given result; with the right components, parameter settings, etc.
- Researchers present results (as in game-theoretic models) that serve as an existence proof rather than anything dispositive.

Overfitting

- Searching through the space of possible models (an infinite set) until one finds a model that "works" for the existing sample.

Overfitting

- Searching through the space of possible models (an infinite set) until one finds a model that "works" for the existing sample.
- Statistical Models: Out-of-sample testing.
- Statistical Models: Divide data into training (in-sample) and test (out-of-sample) datasets.
- Statistical Models: Assess out-of-sample performance by comparing area under ROC curves.

Overfitting

- Searching through the space of possible models (an infinite set) until one finds a model that "works" for the existing sample.
- Statistical Models: Out-of-sample testing.
- Statistical Models: Divide data into training (in-sample) and test (out-of-sample) datasets.
- Statistical Models: Assess out-of-sample performance by comparing area under ROC curves.
- Game-theoretic and Computational Models: conduct large-N statistical tests of empirical implications or derive logical implications.

Curse of Dimensionality

- Concerns the size of parameter spaces in (statistical, game-theoretic, and computational) models.

Curse of Dimensionality

- Concerns the size of parameter spaces in (statistical, game-theoretic, and computational) models.
- Suppose you have 500 observations, 10 independent variables each with 10 possible values.
- What is the size of the parameter space?

Curse of Dimensionality

- Concerns the size of parameter spaces in (statistical, game-theoretic, and computational) models.
- Suppose you have 500 observations, 10 independent variables each with 10 possible values.
- What is the size of the parameter space?
- $10^{10} = 10,000,000,000$

Curse of Dimensionality

- Concerns the size of parameter spaces in (statistical, game-theoretic, and computational) models.
- Suppose you have 500 observations, 10 independent variables each with 10 possible values.
- What is the size of the parameter space?
- $10^{10} = 10,000,000,000$
- So 500 observations will populate this parameter space very, very sparsely and any results should be suspect.

Curse of Dimensionality

- What do statistical models do?

Curse of Dimensionality

- What do statistical models do?
- OLS, for example, simply imposes the constraint that whenever one finds blank spaces in a parameter space, one continues to draw a line through the voids.

Curse of Dimensionality

- What do statistical models do?
- OLS, for example, simply imposes the constraint that whenever one finds blank spaces in a parameter space, one continues to draw a line through the voids.
- Since almost all statistical models interpolate, how does one choose a particular model?

Curse of Dimensionality

- Debate with BKZ.

Curse of Dimensionality

- Debate with BKZ.
- Number of observations: 23,529

Curse of Dimensionality

- Debate with BKZ.
- Number of observations: 23,529
- Possible parameter combinations in dataset:

Curse of Dimensionality

- Debate with BKZ.
- Number of observations: 23,529
- Possible parameter combinations in dataset: $50,570,904,392 \times 43^7 = 13,746,112,994,258,326,281,944$ (not including the choices made in the modeling process).

Curse of Dimensionality

- Debate with BKZ.
- Number of observations: 23,529
- Possible parameter combinations in dataset: $50,570,904,392 \times 43^7 = 13,746,112,994,258,326,281,944$ (not including the choices made in the modeling process).
- $\approx (0.014^{140}) \times$ the number of grains of sand on the earth.

Curse of Dimensionality

- Debate with BKZ.
- Number of observations: 23,529
- Possible parameter combinations in dataset: $50,570,904,392 \times 43^7 = 13,746,112,994,258,326,281,944$ (not including the choices made in the modeling process).
- $\approx (0.014^{140}) \times$ the number of grains of sand on the earth.
- The parameter space is going to be very, very, very sparse.

Conclusions

- Statistical, game-theoretic, and computational models can overfit a desired end result.

Conclusions

- Statistical, game-theoretic, and computational models can overfit a desired end result.
- Statistical, game-theoretic, and computational models all have parameter spaces.

Conclusions

- Statistical, game-theoretic, and computational models can overfit a desired end result.
- Statistical, game-theoretic, and computational models all have parameter spaces.
- Parameterize all choices that are left to the modeler.

Conclusions

- Statistical, game-theoretic, and computational models can overfit a desired end result.
- Statistical, game-theoretic, and computational models all have parameter spaces.
- Parameterize all choices that are left to the modeler.
- How robust are the results to changes in the assumptions (especially technical assumptions unrelated to the empirical referent)?

Conclusions

- Statistical, game-theoretic, and computational models can overfit a desired end result.
- Statistical, game-theoretic, and computational models all have parameter spaces.
- Parameterize all choices that are left to the modeler.
- How robust are the results to changes in the assumptions (especially technical assumptions unrelated to the empirical referent)?
- Game-theoretic and computational models need to be tested empirically.

Conclusions

- Statistical, game-theoretic, and computational models can overfit a desired end result.
- Statistical, game-theoretic, and computational models all have parameter spaces.
- Parameterize all choices that are left to the modeler.
- How robust are the results to changes in the assumptions (especially technical assumptions unrelated to the empirical referent)?
- Game-theoretic and computational models need to be tested empirically.
- Models should be compared by examining how fruitful their predictions are rather than the accuracy of their assumptions.

Conclusions

- Statistical, game-theoretic, and computational models can overfit a desired end result.
- Statistical, game-theoretic, and computational models all have parameter spaces.
- Parameterize all choices that are left to the modeler.
- How robust are the results to changes in the assumptions (especially technical assumptions unrelated to the empirical referent)?
- Game-theoretic and computational models need to be tested empirically.
- Models should be compared by examining how fruitful their predictions are rather than the accuracy of their assumptions.
- Need out-of-sample work to increase confidence that the model has something to say about the world.

Take-Away Points

- See above.