Computational Intelligence For Engineering Solutions: Invariance-Based Approach (a brief overview of the Fall 2022 class CS 5354/CS 4365)

Vladik Kreinovich
Department of Computer Science
University of Texas at El Paso, El Paso, Texas 79968, USA
vladik@utep.edu
https://www.cs.utep.edu/vladik/cs5354.22

1. Pre-Requisites

- For graduate students:
  - no special pre-requisites;
  - graduate level standing is sufficient.
- For undergraduate students:
  - ideally, Statistics and Linear Algebra, but this is not required,
  - we will recall the needed material anyway.

2. Introduction: formulation of the problems

- The main purpose of computers is to process real-life data, so that we will able:
  - to understand the current state of the system and
  - to predict its future behavior.
- In some situations – e.g., in basic mechanics – we have fundamental from-first-principles laws that enable us to make the corresponding predictions.
- However, in many other situations, especially in engineering, we only have approximate empirical formulas.
- For example, it is not possible to predict, based on the first principles:
  - how pavement will deteriorate with time, or
  - how people will change their opinions about goods.

3. Introduction: formulation of the problems (cont-d)

- In such situations, we face the following problems:
- *Why these formulas?*
- Users are usually reluctant to use purely empirical formulas.
- Reason: there is no guarantee that theses formulas will work in a new situations.
- It is therefore desirable to come up with theoretical explanations for these formulas.
- *Maybe these formulas are not the best.*
- Within these theoretical explanations, are the current formulas most adequate?
- And if they are not the best, what are the better formulas?

4. Introduction: formulation of the problems (cont-d)

- *What next?*
- Empirical formulas are usually approximate.
- If we want a more accurate description, we need more complex, more detailed formulas.
- Of course, the ultimate test is comparing with the observations and measurement results.
- In view of the theoretical explanations, what are good candidates for such more complex formulas?
- Similarly, in many engineering applications, there are semi-empirical methods for solving the corresponding problems.

5. Introduction: formulation of the problems (cont-d)

- In such cases, similar problems appear.
- *Why these methods?*
- Users are usually reluctant to use purely empirical methods, since there is no guarantee that these methods will work in a new situations.
- It is therefore desirable to come up with theoretical explanations for these methods.
- *Maybe these methods are not the best.*
- Within these theoretical explanations, are the current methods the most adequate?
- And if not, what are the better methods?

6. Introduction: formulation of the problems (cont-d)

- *What next?*
- Empirical methods are usually imperfect.
- If we want better results, we need more complex, more detailed methods.
- Of course, the ultimate test is testing these methods on the real data.
- In view of the theoretical explanations, what are good candidates for such more complex methods?

7. First topic: how to solve these problems – the idea of invariance

- A natural way to make predictions is to look what happened in similar situations in the past.
- And what does "similar" mean? It means that:
  - some important features of the current situations
  - are the same (or at least almost the same) as the same important features of the past situation.
- In other words:
  - there may have been some changes between the two situations,
  - but the important features did not change.

8. First topic: how to solve these problems – the idea of invariance (cont-d)

- In mathematical terms:
  - changes are called *transformations*,
  - if a feature does not change under a transformation, we say that this feature is *invariant* under this transformation, and
  - transformations under which some features are invariant are known as *symmetries*.
- Not surprisingly, symmetries and invariances are, at present, one of most effective tools in theoretical physics (and in other disciplines).
- In line with this reasoning, in this class, we will study the invariance-based approach to solving the above problems.

9. A simple example

- How do we know that:
  - if we drop a pen,
  - then it will start falling down with the acceleration of 9.81 m/sec^2?
- Because we – and others:
  - repeated this experiment in many different locations, at many moments of time,
  - and always observed the same result.
- In this case, the current situation can be obtained from the previous one by shifts in space and time, maybe by rotation.
- So the observed phenomenon is invariant with respect to all these transformations.

10. Second topic: what transformations we will study in this class

- Some transformations studied in physics are very complicated.
- Example: changing particles to corresponding anti-particles.
- In this class, we will focus on the simplest and most natural transformations.
- These transformation come from the fact that:
  - while we want to process the *actual* values of different quantities like acceleration,
  - in practice, we deal with *numerical* values.
- For most physical quantities, numerical values depend on the choice of the measuring unit.

11. Second topic: what transformations we will study in this class (cont-d)

- For example:
  - if instead of meters, we use centimeters, then
  - the same acceleration of 9.81 m/sec^2 becomes described by a different numerical value 981 cm/sec^2.
- For some quantities like temperature or time, there is no fixed starting point.
- So numerical values also depend on what starting point we choose.
- For example:
  - the difference between Fahrenheit and Celsius scales for measuring temperature is that
  - these two scales use different measuring units and different starting points.

12. Third topic: corresponding invariances

- In many situations:
  - there is no fixed measuring unit,
  - the choice of a measuring unit is simply a matter of convention.
- In this case, it is reasonable to assume that:
  - the desired empirical formula has the same form
  - no matter what measuring unit we select.
- Of course, each formula relates the values of several quantities; so:
  - if we change the measuring unit for one of the quantities,
  - we need to appropriately change the measuring unit for other quantities.
- As an example, let us consider the formula $d = v \cdot t$ that describes:
  - the traveled distance d
  - as a function of velocity v and time t.

13. Third topic: corresponding invariances (cont-d)

- This formula remains true whether we measure distance in kilometers or in miles; however:
  - for this formula to remain true when we switch from miles to kilometers,
  - we also need to change the units for measuring velocity from miles per hour to kilometers per hour.
- Similarly, in some formulas:
  - if we change the starting point for measuring one of the quantities,
  - we may need to appropriately re-scale other quantities.

14. Fourth topic: which dependencies are invariant with respect to these transformations

- This is what we will study first:
- Which functions are invariant with respect to changing the measuring unit and changing the starting point.
- We will show that, in effect, the only invariant dependencies are:
  - the power law $y = A \cdot x^b$,
  - the exponential dependence $y = A \cdot \exp(b \cdot x)$,
  - the logarithmic dependence $y = A \cdot \log(x) + b$, and
  - the linear dependence $y = A \cdot x + b$.
- We will also describe engineering situations where such dependencies appear, including:
  - how people change their ratings of different goods,
  - inverse distance weighting in geosciences, and
  - the use of so-called soft-max in deep learning.

15. Fourth topic: which dependencies are invariant with respect to these transformations (cont-d)

- In all the above cases, x and y are two different – but related – quantities.
- We will also consider a special case when x and y are related values of the same quantity.
- In this case, the corresponding invariance will:
  - explain the appearance of Rectified Linear Unit (ReLU)
  $$y = \max(0, x);$$
  - this is the main activation function in deep neural networks.

16. Fifth topic: optimization and how it is related to invariance

- Several of the above-described problems are:
  - about looking for the *best* formula or the best method,
  - i.e., about *optimization*.
- We consider situations when we have a symmetry – i.e., a transformation with respect to which important features are invariant.
- It is reasonable to assume that:
  - the relative quality of different alternatives (formulas or algorithms) should not change
  - if we simply change the measuring unit and/or the starting point of the corresponding quantity.

17. Fifth topic: optimization and how it is related to invariance (cont-d)

- We will show that in this case, the optimal alternative should also be invariant.
- This explains the effectiveness of invariant empirical formulas in engineering.
- In particular, it explains the effectiveness of ReLU units and of softmax in deep learning.

18. Sixth topic: what next – possible ideas

- As we have mentioned, often, simple invariant transformations provide only an approximation to the actual dependence.
- In some cases, it is not a very accurate approximation, it is desirable to have more accurate formulas.
- In other cases, it is a reasonably accurate approximation, but still a more accurate formula is desirable.
- In both cases, it is desirable to come up with more accurate formulas.
- To come up with such formulas, let us recall what simplifying assumptions we made in the above description of invariant dependencies

$$y = f(x).$$

- We can think of weakening these assumptions and thus, coming up with more accurate formulas.

19. Sixth topic: what next – possible ideas (cont-d)

- We made the following assumptions:
  - that the quantity y directly depends on the quantity x;
  - that there is only one way how x affects y;
  - that we are looking for a universal formula $y = f(x)$ describing all possible situations; and
  - that the only invariances are invariances with respect to the change of the measuring unit and/or starting point.
- In practice, all these assumptions may be violated.
- This leads to four directions in which more complex formulas can appear.

20. Seventh topic: what if the dependence of y on x is indirect

- In many practical situations, the dependence on the desired quantity y on the known quantity x is indirect:
  - the quantity y depends on some auxiliary quantity z, and
  - this auxiliary quantity z, in its turn, depends on x.
- We may have two or more such auxiliary quantities.
- In this case, the dependence of y on x is described:
  - not by one of the above four formulas describing the direct dependence,
  - but by the composition of such formulas.
- For example, if $y = f(z)$ and $z = g(x)$, then $y = f(g(x))$.
- We will study such composition functions.
- We will describe engineering examples where such compositions are indeed experimentally observed.

21. Eighth topic: what if there are several ways how x affects y

- In some situations, there are two or more different ways how x can influence y.
- We have empirical formulas describing each of these ways.
- They usually correspond to specific cases when one of these ways is dominant and other can be ignored.
- To describe a general situation:
  - in which all these ways are important,
  - we need to combine the known formulas.
- We will show how invariance requirements can help select an appropriate combination.
- We will provide engineering examples where the resulting combined formulas work well.

22. Ninth topic: what if in different situations, the dependence of y on x is different?

- In such situations, we cannot use a single function.
- We need to come up with a *family* of functions.
- A natural – and probably the simplest – way to get a family of functions is to take linear combinations of a few selected functions $e_i(x)$:

$$f(x) = c_1 \cdot e_1(x) + \ldots + c_n \cdot e_n(x).$$

- In the case when we were looking for a single function, it was reasonable to assume that this function is invariant.
- Similarly, in this case, it is also reasonable to assume that the *family* of functions is invariant.
- We will study such invariant families.
- Examples will include a biomedical application: how to describe the growth of a tumor.

23. Tenth topic: transformations more complex than changing the measuring unit and/or the starting point

- As we have mentioned:
  - in many situations, the transformations are
  - more complex than changing the measuring unit and/or the starting point.
- In this course, we will consider the simplest case of such complex transformations:
  - when we transform a single quantity
  - as opposed to, e.g., rotation in a plane, that transforms the numerical values of both coordinates x and y.

24. Tenth topic (cont-d)

- It turns out that this case, the transformations are fractionally linear, i.e., have the form
$$y = \frac{a \cdot x + b}{c \cdot x + d}.$$
- We will analyze functions invariant under such transformations.
- We will show that this explains the efficiency of sigmoid activation function $1/(1 + \exp(-x))$.
- This function is used both in shallow and in deep neural networks.

25. Projects

- The main purpose of all this activity is to help engineering applications.
- Students are therefore required to work on projects.
- Projects should be either related to real-life applications, or have a more general theoretical nature.
- There are two types of projects.
- A *possible* project may consist of reviewing some related paper(s) – and presenting this review to the class.

26. Projects (cont-d)

- If you select to review a paper, please make sure to concentrate on the following questions in your review:
  - what is the general practical problem for whose solution this paper is aimed;
  - what was known before this paper and what were the remaining challenges;
  - what are the results of this paper, and in what sense they are better than what was done before;
  - what are the remaining challenges – if any.
- An *ideal* project should be creative.
- In such a project, students – individually or in groups – will come up with something new.

27. Projects (cont-d)

- How is this possible?
- Invariance-based approach is a new developing topic; so:
  - if there are some interesting empirical formulas that have no convincing theoretical explanations,
  - why not try to explain them by using the techniques we learn in class?
- Or, better yet, why not try to come up with more accurate formulas?

28. Materials used in class

- Pedro Barragan Olague and Vladik Kreinovich, "A Symmetry-Based Explanation for an Empirical Model of Fatigue Damage of Composite Materials", Journal of Uncertain Systems, 2018, Vol. 12, No. 3, pp. 176–179.
https://www.cs.utep.edu/vladik/2017/tr17-34.pdf
- Pedro Barragan Olagues and Vladik Kreinovich, "Why growth of cancerous tumors is Gompertzian: a symmetry-based explanation", Cybernetics and Physics, 2017, Vol. 6, No. 1, pp. 13–18.
https://www.cs.utep.edu/vladik/2016/tr16-106c.pdf
- Laxman Bokati, Aaron Velasco, and Vladik Kreinovich, "Scale-Invariance and Fuzzy Techniques Explain the Empirical Success of Inverse Distance Weighting and of Dual Inverse Distance Weighting in Geosciences", Proceedings of the Annual Conference of the North American Fuzzy Information Processing Society NAFIPS'2020, Redmond, Washington, August 20-22, 2020, pp. 379–390.
https://www.cs.utep.edu/vladik/2020/tr20-24.pdf

29. Materials used in class (cont-d)

- Vladik Kreinovich and Olga Kosheleva, "Optimization under uncertainty explains empirical success of deep learning heuristics", In: Panos Pardalos, Varvara Rasskazova, and Michael N. Vrahatis (eds.), Black Box Optimization, Machine Learning and No-Free Lunch Theorems, Springer, Cham, Switzerland, 2021, pp. 195–220.
https://www.cs.utep.edu/vladik/2019/tr19-49.pdf
- Vladik Kreinovich, Anh H. Ly, Olga Kosheleva, and Songsak Sriboonchitta, "Efficient Parameter-Estimating Algorithms for Symmetry-Motivated Models: Econometrics and Beyond", In: Ly H. Anh, Le Si Dong, Vladik Kreinovich, and Nguyen Ngoc Thach (eds.), Econometrics for Financial Applications, Springer Verlag, Cham, Switzerland, 2018, pp. 134–145.
/vladik/2017/tr17-81.pdf

30. Materials used in class (cont-d)

- Edgar Daniel Rodriguez Velasquez, Vladik Kreinovich, and Olga Kosheleva, "Invariance-Based Approach: General Methods and Pavement Engineering Case Study", International Journal of General Systems, 2021, DOI: 10.1080/03081079.2021.1953005.
https://www.cs.utep.edu/vladik/2020/tr20-122a.pdf

- Julio Urenda, Manuel Hernandez, Natalia Villanueva-Rosales, and Vladik Kreinovich, "How User Ratings Change with Time: Theoretical Explanation of an Empirical Formula", Proceedings of the Annual Conference of the North American Fuzzy Information Processing Society NAFIPS'2020, Redmond, Washington, August 20-22, 2020, pp. 427–432.
https://www.cs.utep.edu/vladik/2019/tr19-100.pdf