Computational Complexity of Experiment Design in Civil Engineering

Olga Kosheleva[1], Yan Wang[2], and Vladik Kreinovich[1]
[1]University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
olgak@utep.edu, vladik@utep.edu
[2]Woodruff School of Mechanical Engineering
Georgia Institute of Technology
813 Ferst Drive
Atlanta, GA 30332-0405, USA
yan.wang@me.gatech.edu

1. Need to Measure Mechanical Characteristics of Engineering Structures

- Reliability and safety of a structure is a very important issue in civil engineering.
- We need to make sure that a bridge will withstand a typical load and/or a typical wind thrust.
- We need to make sure that a building will withstand an earthquake typical for the given area.
- To simulate the effect of all these loads and disruptions, we need to know the mechanical properties of the corresponding construction.
- For the long-standing constructions, mechanical properties change with time.
- The actual values of the corresponding mechanical characteristics need to be determined from measurements.

2. Linearization Is Usually Possible

- The mechanical characteristics describe how the displacement depend on the forces.
- In most cases, the displacements are relatively small.
- So we can safely ignore quadratic and higher order terms and assume that the dependence is linear.
- Such a dependence is known as *Hooke's law*.
- It is well known that linear equations are easier to solve and to analyze.
- So the fact that we can limit ourselves to linear equations is, from the practical viewpoint, very beneficial.

3. Need for Experiment Design

- Measurements are often not easy of the existing large-scale engineering structures, be it bridges or buildings.
- Each such measurement is costly and time-consuming.
- It is therefore necessary to carefully design the corresponding measurements, so as not to overspend on these measurements.
- After we have already performed several measurements, the first task is to check whether the existing measurements have been sufficient.
- At first glance, it may seem that since all the equations are linear, checking whether additional measurements are possible is easy.
- Indeed, there are many efficient algorithms for solving systems of linear equations.
- If we take into account measurement uncertainty, then even in the linear case, we may get an NP-hard problem.
- However, in the ideal case when all the measurements are accurate, it may seem that the problems should be feasible.

4. In Reality, the Experiment Design Problem Is Complicated

- The problem is that it is not possible to place sensors at all the points on the bridge.
- When we only measure *some* of the quantities – even if we measure accurately – many computational problems become NP-hard.
- In this talk, we show that the experiment design problem also becomes NP-hard.
- The fact that the problem is NP-hard means that:
  - if – as most computer scientists believe – NP $\neq$ P,
  - no feasible algorithm is possible that would always check whether a given set of measurement results is sufficient.

5. Practical Consequences of This Result

- Theoretically, there exists the most economical way to perform the corresponding safety analysis.
- However, in practice, finding such a way is not feasible.
- Thus, when performing measurement, overspending is inevitable.
- This may be one of the reasons why it is often cheaper to demolish a building and rebuild it from scratch rather than repair it.

6. Towards Formulating the Problem in Precise Terms

- In general, the dependence on forces f_α at different locations α on different displacement ε_β is non-linear.
- In this talk, we consider the case when displacements are small.
- In this case, we can ignore terms which are quadratic or higher order in terms of ε_β.
- So, we can assume that the dependence of each force component f_α on all the components ε_β of displacements at different locations β is linear.
- Taking into account that in the absence of forces, there is no displacement, we conclude that, for some coefficients $K_{\alpha,\beta}$,

$$f_\alpha = \sum_\beta K_{\alpha,\beta} \cdot \varepsilon_\beta.$$

- These coefficients $K_{\alpha,\beta}$ describe the mechanical properties of the body.
- It is therefore desirable to experimentally determine these coefficients.

7. Ideal Case and Real Case

- In the ideal case, we measure displacements ε_β and forces f_α at all possible locations.
- Each such measurement results in an equation which is linear in terms of the unknowns $K_{\alpha,\beta}$.
- Thus, after performing sufficiently many measurements, we get an easy-to-solve system of linear equations that enables us to find $K_{\alpha,\beta}$.
- In reality, we only measure displacements and forces at *some* locations – i.e., we know only some values f_α and ε_β.
- In this case, since both $K_{\alpha,\beta}$ and some values ε_β are unknown, the corresponding system of equations becomes quadratic.
- After sufficiently many measurements, we may still uniquely determine $K_{\alpha,\beta}$, but the reconstruction is more complex.

8. What We Prove

- We prove it is NP-hard to check, after the measurement,
  - whether additional measurements are needed,
  - or whether we already have enough information to determine the value of the desired quantity.

9. Definitions and the Main Result

- Let K be a natural number. This number will be called *the number of experiments.*
- By a *problem of checking whether additional measurements are needed*, we mean the following problem.
  - We know that for every k from 1 to K, we have $f_\alpha^{(k)} = \sum\limits_\beta K_{\alpha,\beta} \cdot \varepsilon_\beta^{(k)}$ for some values $f_\alpha^{(k)}$ and $\varepsilon_\beta^{(k)}$.
  - For each k, we know some of the values $f_\alpha^{(k)}$ and $\varepsilon_\beta^{(k)}$.
  - We need to check if for given α_0 and β_0, the above equations uniquely determine the value K_{α_0,β_0}.

Proposition. *The problem of checking whether additional measurements are needed is NP-hard.*

10. Main Idea: Reduction to Subset Sum

- By definition, NP-hard means that all the problems from a certain class NP can be reduced to this problem.
- It is known that the following *subset sum* problem is NP-hard:
  - given $m+1$ natural numbers $s_1, \ldots, s_M, S$,
  - check whether it is possible to find the values $x_i \in \{0,1\}$ for which
$$\sum_{i=1}^{M} s_i \cdot x_i = S;$$
  - in other words, check whether it is possible to find a subset of the values $s_1, \ldots, s_M$ whose sum is equal to the given value S.
- The fact that the subset sum problem is NP-hard means that every problem from the class NP can be reduced to this problem.
- So, if we reduce the subset problem to our problem, that would mean, by transitivity of reduction, that our problem is indeed NP-hard.

11. Corresponding Physical Quantities

- Let $s_1, \ldots, s_M, S$, be the values that describe an instance of the subset sum problem.
- Let us reduce it to the following instance of our problem.
- Let us denote $m \stackrel{\text{def}}{=} M + 1$.
- In this instance, we have $2m+1$ variables $\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_m, \varepsilon_{m+1}, \ldots, \varepsilon_{2m}$.
- We also have $m + 1$ different values f_α, $\alpha = 0, 1, \ldots, m$.

12. First Series of Experiments

- In the first series of experiments $k = 1, \ldots, m$, for each $i = 1, \ldots, m$, we have

$$\varepsilon_i^{(i)} = 1, \varepsilon_{m+i}^{(i)} = -1, \text{ and} \varepsilon_j^{(i)} = 0 \text{ for all } j \neq i.$$

- The only value of f_α that we measure in each of these experiments is the value $f_0^{(i)} = 0$.
- The corresponding equation takes the form

$$0 = f_0^{(i)} = \sum_\beta K_{0,\beta} \cdot \varepsilon_\beta^{(i)} = K_{0,i} - K_{0,m+i}.$$

- So, we can conclude that $K_{0,m+i} = K_{0,i}$.

13. Second Series of Experiments

- In the second series of experiments $k = m+1, \ldots, m+i, \ldots, 2m$, where $i = 1, \ldots, m$, for each $k = m + i$:
  - we measure the values $\varepsilon_j^{(m+i)} = 0$ for all $j \neq k$, and
  - we measure the values $f_0^{(m+i)} = f_i^{(m+i)} = 1$.
- From the corresponding equations, we conclude that

$$1 = K_{0,m+i} \cdot \varepsilon_{m+i}^{(m+i)} \text{ and} 1 = K_{i,m+i} \cdot \varepsilon_{m+i}^{(m+i)}.$$

- We do not know the value $\varepsilon_{m+i}^{(m+i)}$, but we can find it from the first equation and substitute into the second equation.
- As a result, we conclude that $K_{0,m+i} = K_{i,m+i}$.
- Combining this equality with the equality derived from the first experiment, we conclude that $K_{0,i} = K_{i,m+i}$.

14. Third Series of Experiments

- In the third series of experiments $k = 2m + i$, $i = 1, \ldots, m$, for each i:
  - we measure $\varepsilon_i^{(2m+i)} = 1$, $\varepsilon_j^{(2m+i)} = 0$ for all other j, and
  - we measure $f_i^{(2m+i)} = 1$.
- The corresponding equation implies that $K_{i,i} = 1$.

15. Fourth Series of Experiments

- In the fourth series of experiments $k = 3m + i$, $i = 1, \ldots, m$:
  - we measure the values $\varepsilon_{m+i}^{(3m+i)} = -1$ and $\varepsilon_j^{(3m+i)} = 0$ for all j which are different from i and from $m + i$.
  - We also measure the values $f_0^{(3m+i)} = f_i^{(3m+i)} = 0$.
- The corresponding equations lead to

$$K_{0,i} \cdot \varepsilon_i^{(3m+i)} - K_{0,m+i} = 0 \text{ and } K_{i,i} \cdot \varepsilon_i^{(3m+i)} - K_{i,m+i} = 0.$$

- Since we already know that $K_{i,i} = 1$, the second equation simply means that $\varepsilon_i^{(3m+i)} = K_{i,m+i}$.
- We know that $K_{i,m+i} = K_{0,i}$ so $\varepsilon_i^{(3m+i)} = K_{0,i}$.
- Substituting this expression for $\varepsilon_i^{(3m+i)}$ into the first equation and taking into account that $K_{0,m+i} = K_{0,i}$, we conclude that $K_{0,i}^2 - K_{0,i} = 0$.
- Thus, $K_{0,i} \in \{0, 1\}$.

16. Fifth Series of Experiments

- The fifth, final series of experiments consists of only one experiment $k = 4m + 1$.
- In this experiment, for all $i = 1, \ldots, m$, we measure the values
$$\varepsilon_1^{(4m+1)} = s_1, \ldots, \varepsilon_M^{(4m+1)} = s_M, \varepsilon_m^{(4m+1)} = -S, \text{ and } \varepsilon_{m+i}^{(4m+1)} = 0.$$
- We also measure $f_0^{(4m+1)} = 0$.
- We want to check whether all the measurement results uniquely determine the value $K_{0,m}$.
- We already know that $K_{0,m}$ is equal to either 0 or 1.
- The corresponding equation is $K_{0,1} \cdot s_1 + \ldots + K_{0,M} \cdot s_M - K_{0,m} \cdot S = 0$, i.e.:
$$K_{0,1} \cdot s_1 + \ldots + K_{0,M} \cdot s_M = K_{0,m} \cdot S.$$
- The value $K_{0,m} = 0$ is always possible here: for example, in this case, we can have $K_{0,1} = \ldots = K_{0,M} = 0$.
- The question is thus whether the value $K_{0,m} = 1$ is possible.

17. Fifth Series of Experiments (cont-d)

- For $K_{0,m} = 1$, the above formula takes the form

$$K_{0,1} \cdot s_1 + \ldots + K_{0,M} \cdot s_M = S.$$

- One can easily see that:
  - If the original instance of the subset sum problem has a solution $x_i \in \{0, 1\}$, then the above equality holds for $K_{0,i} = x_i$.
  - Vice versa, if there exist values $K_{0,i}$ that satisfy this formula, then the values $x_i = K_{0,i}$ solve the original subset sum problem.
- So, whether additional measurements are needed depends on whether the corresponding instance of the subset sum problem has a solution.
- Thus, we indeed have a reduction – and hence, our problem is indeed NP-hard.

18. Acknowledgments

This work was supported in part by the US National Science Foundation grant HRD-1242122 (Cyber-ShARE Center of Excellence).