Computational Complexity of Experiment Design in Civil Engineering

Olga Kosheleva

Department of Teacher Education, University of Texas at El Paso, USA. E-mail: olgak@utep.edu

Yan Wang

School of Mechanical Engineering, Georgia Institute of Technology, USA. E-mail: yan.wang@me.gatech.edu

Vladik Kreinovich

Department of Computer Science, University of Texas at El Paso, USA. E-mail: vladik@utep.edu

To guarantee reliability and safety of engineering structures, we need to regularly measure their mechanical properties. Such measurements are often expensive and time-consuming. It is therefore necessary to carefully plan the corresponding measurement experiments, to minimize the corresponding expenses.
It is known that, in general, experiment design is NP-hard. However, the previous proofs dealt either with nonlinear systems, or with situations with low measurement accuracy. In civil engineering, however, most systems are well-described by linear systems, and measurements are reasonably accurate. In this paper, we show that experiment design is NP-hard even for civil engineering problems. We show that even checking whether the results of the previous measurements are sufficient to determine the value of the desired mechanical quantity – or additional measurement are needed – even this problem is, in general, NP-hard. So, crudely speaking, no feasible algorithm can always answer this question – and thus, overspending on measurements is inevitable.

Keywords: Civil engineering, NP-hard, experiment design, linear elasticity.

1. Formulation of the Practical Problem

1.1. *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, whose mechanical properties have changed with time, the actual values of the corresponding mechanical characteristics need to be determined from measurements.

1.2. *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.

1.3. *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.

1.4. *Experiment design: first task*

After we have already performed several measurements, the first task is to check whether the existing measurements have been sufficient, or whether additional measurements are needed.

1.5. *At first glance, the first task may sound simple*

At first glance, it may seem that since all the equations are linear, checking whether additional measurements are possible is not that computationally difficult: indeed, while solving systems of quadratic and higher order equations is known to be computationally intractable (NP-hard) Kreinovich et al. (1997); Papadimitriou (1993), there are many efficient algorithms for the linear case.

Proceedings of the 29th European Safety and Reliability Conference.
Edited by Michael Beer and Enrico Zio

ISBN: 981-973-0000-00-0 :: doi: 10.3850/981-973-0000-00-0_tr19-24

Yes, if we take into account measurement uncertainty, then even in the linear case, we may get an NP-hard problem Kreinovich et al. (1997). However, in the ideal case when all the measurements are accurate, it may seem that the problems should be feasible.

1.6. *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; see, e.g., Wang et al. (2019).

In this paper, we show that the experiment design problem also becomes NP-hard.

1.7. *How complicated is it?*

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 to reconstruct the value of the desired quantity.

Comment. Readers interested in more technical description of NP-hardness are welcome to see, e.g., Kreinovich et al. (1997); Papadimitriou (1993).

1.8. *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.

Comment. This may be one of the reasons why it is often cheaper to demolish a building and rebuild it from scratch rather than repair and try to salvage what can be saved.

2. Towards Formulating the Problem in Precise Terms

2.1. *Linear elasticity*

In general, the dependence on forces f_α at different locations α on different displacement ε_β is non-linear. In this paper, we consider the case when displacements are small. In this case, we can ignore terms quadratic or higher order in terms of ε_β and 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

$$f_\alpha = \sum_\beta K_{\alpha,\beta} \cdot \varepsilon_\beta \tag{1}$$

for some coefficients $K_{\alpha,\beta}$. These coefficients $K_{\alpha,\beta}$ describe the mechanical properties of the body.

It is therefore desirable to experimentally determine these coefficients.

2.2. *Ideal 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 (1) 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 the values $K_{\alpha,\beta}$.

2.3. *In practice, we only measure some values*

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.

2.4. *What we prove*

We prove that even the problem of checking, after the measurement:

- whether additional measurements are needed,
- or whether we already have enough information to determine the value of the desired quantity,

this problem is already NP-hard.

3. Definitions and the Main Result

From the computational viewpoint, the above problem can be formulated as follows.

Definition 1. *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*

$$f_\alpha^{(k)} = \sum_\beta K_{\alpha,\beta} \cdot \varepsilon_\beta^{(k)} \tag{2}$$

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.*

Comment. Our proof is similar to the proof of a similar problem presented in Wang et al. (2019). The difference is that:

- in Wang et al. (2019), we considered the problem of checking whether a given value of a mechanical characteristic is consistent with the measurement results;
- in this paper, we consider a different problem: checking
  - whether, based on the given measurement results, we can uniquely determine the value of the desired mechanical characteristic,
  - or whether additional measurements are needed to find the value of this mechanical characteristic.

4. Proof

4.1. *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; see, e.g., Kreinovich et al. (1997); Papadimitriou (1993). It is known that the following *subset sum* problem is NP-hard:

- given $m+1$ natural numbers
$$s_1, \ldots, s_M, S, \tag{3}$$
- check whether it is possible to find the values $x_i \in \{0,1\}$ for which
$$\sum_{i=1}^{M} s_i \cdot x_i = S \tag{4}$$
(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 every problem from the class NP can be reduced to our problem as well – i.e., that our problem is indeed NP-hard.

This is exactly how we will construct this proof – by showing that the subset sum problem can be reduced to our problem.

4.2. *Corresponding physical quantities*

Let (3) 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
$$\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_m, \varepsilon_{m+1}, \ldots, \varepsilon_{2m}. \tag{5}$$
We also have $m+1$ different values f_α, $\alpha = 0, 1, \ldots, m$.

4.3. *First series of experiments: description*

In the first series of experiments $k = 1, \ldots, m$, for each $i = 1, \ldots, m$, we have
$$\varepsilon_i^{(i)} = 1, \varepsilon_{m+i}^{(i)} = -1, \text{ and}$$
$$\varepsilon_j^{(i)} = 0 \text{ for all } j \neq i. \tag{6}$$
The only value of f_α that we measure in each of these experiments is the value $f_0^{(i)} = 0$.

4.4. *What we can conclude from the results of the first series of measurements*

From the corresponding equation
$$0 = f_0^{(i)} = \sum_\beta K_{0,\beta} \cdot \varepsilon_\beta^{(i)} = K_{0,i} - K_{0,m+i}, \tag{7}$$
we conclude that
$$K_{0,m+i} = K_{0,i}. \tag{8}$$

4.5. *Second series of experiments: description*

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
$$f_0^{(m+i)} = f_i^{(m+i)} = 1. \tag{9}$$

4.6. *What we can conclude from the results of the second series of measurements*

From the corresponding equations, we conclude that
$$1 = K_{0,m+i} \cdot \varepsilon_{m+i}^{(m+i)} \tag{10}$$

and

$$1 = K_{i,m+i} \cdot \varepsilon_{m+i}^{(m+i)}. \tag{11}$$

We do not know the value $\varepsilon_{m+i}^{(m+i)}$, but we can find it from the equation (10) and substitute into the the equation (11). As a result, we conclude that

$$K_{0,m+i} = K_{i,m+i}. \tag{12}$$

Combining equalities (8) and (12), we conclude that

$$K_{0,i} = K_{i,m+i}. \tag{13}$$

4.7. *Third series of experiments: description*

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 $f_i^{(2m+i)} = 1$.

4.8. *What we can conclude from the results of the third series of measurements*

The corresponding equation implies that

$$K_{i,i} = 1. \tag{14}$$

4.9. *Fourth series of experiments: description*

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
$$f_0^{(3m+i)} = f_i^{(3m+i)} = 0. \tag{15}$$

4.10. *What we can conclude from the results of the fourth series of measurements*

In this case, the corresponding equations lead to

$$K_{0,i} \cdot \varepsilon_i^{(3m+i)} - K_{0,m+i} = 0 \tag{16}$$

and

$$K_{i,i} \cdot \varepsilon_i^{(3m+i)} - K_{i,m+i} = 0. \tag{17}$$

Since, due to (14), we have $K_{i,i} = 1$, the equation (17) simply means that

$$\varepsilon_i^{(3m+i)} = K_{i,m+i}. \tag{18}$$

Due to formula (13), this implies that

$$\varepsilon_i^{(3m+i)} = K_{0,i}. \tag{19}$$

Substituting this expression for $\varepsilon_i^{(3m+i)}$ into the equation (16) and taking into account that, due to (8), we have $K_{0,m+i} = K_{0,i}$, we conclude that

$$K_{0,i}^2 - K_{0,i} = 0. \tag{20}$$

From

$$K_{0,i} \cdot (K_{0,i} - 1) = 0, \tag{21}$$

we conclude that either $K_{0,i} = 0$ or $K_{0,i} = 1$. Thus, for each i from 1 to m, we have

$$K_{0,i} \in \{0, 1\}. \tag{22}$$

4.11. *Fifth series of experiments: description*

The fifth, final series of experiments consists of only one experiment $k = 4m + 1$. In this experiment, we measure the values,

$$\varepsilon_1^{(4m+1)} = s_1, \ldots, \varepsilon_M^{(4m+1)} = s_M, \tag{23}$$

$$\varepsilon_m^{(4m+1)} = -S, \text{ and } \varepsilon_{m+i}^{(4m+1)} = 0 \tag{24}$$

for all $i = 1, \ldots, m$.

We also measure

$$f_0^{(4m+1)} = 0. \tag{25}$$

4.12. *What we can conclude from the results of the fifth series of measurements*

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 takes the form

$$K_{0,1} \cdot s_1 + \ldots + K_{0,M} \cdot s_M - K_{0,m} \cdot S = 0, \tag{26}$$

i.e., the form

$$K_{0,1} \cdot s_1 + \ldots + K_{0,M} \cdot s_M = K_{0,m} \cdot S \tag{27}$$

for some values $K_{0,i} \in \{0, 1\}$.

The value $K_{0,m}$ is always possible here: for example, in this case, we can have

$$K_{0,1} = \ldots = K_{0,M} = 0. \tag{28}$$

The question is thus whether the value $K_{0,m} = 1$ is possible. For this value, the above formula (27) takes the form

$$K_{0,1} \cdot s_1 + \ldots + K_{0,M} \cdot s_M = S. \tag{29}$$

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 the formula (5), 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.

Acknowledgments

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

References

Kreinovich, V., A. Lakeyev, J. Rohn, and P. Kahl (1997). *Computational complexity and Feasibility of Data Processing and Interval Computations*, Kluwer.

Papadimitriou, C. H. (1993) *Computational Complexity*, Pearson.

Wang, Y., O. Galindo, M. Baca, J. Lasley, and V. Kreinovich (2019). Experimental determination of mechanical properties is, in general, NP-hard – unless we measure everything. *Journal of Uncertain Systems*, to appear.