Simplicity Is Worse Than Theft: A Constraint-Based Explanation of a Seemingly Counter-Intuitive Russian Saying

Martine Ceberio, Olga Kosheleva, and Vladik Kreinovich

University of Texas at El Paso
El Paso, TX 79968, USA
mceberio@utep.edu, olgak@utep.edu,
vladik@utep.edu

1. In Science, Simplicity Is Good

- The world around us is very complex.
- One of the main objectives of science is to simplify it.
- Science has indeed greatly succeeded in doing it.
- Example: Newton's equations explain the complex motions of celestial bodies motion by simple laws.
- From this viewpoint, simplicity of the description is desirable.
- To achieve this simplicity, we sometimes ignore minor factors.
- Example: Newton treated planets as points, while they have finite size.
- As a result, there is a small discrepancy between Newton's theory and observations.

2. In Practice, Simplified Models Are Not Always Good in Decision Making

- One of the main purposes of science is to gain knowledge.
- Once this knowledge is gained, we use it to improve the world; examples:
  - knowing how cracks propagate helps design more stable constructions.
  - knowing the life cycle of viruses helps cure diseases caused by these viruses.
- What happens sometimes is that the simplified models,
  - models which have led to very accurate *predictions*,
  - are not as efficient when we use them in *decision making*.

3. Simplified Approximate Models Leads to Bad Decisions: Examples

- Numerous examples can be found in the Soviet experiment with the global planning of economy.
- *Good ideas:* Nobelist Wassily Leontieff started his research as a leading USSR economist.
- However, the *results* were sometimes *not* so *good*.
- *Example:* buckwheat – which many Russian like to eat – was often difficult to buy.
- *Explanation:* to solve a complex optimization problem, we need to simplify the problem.
- *How to simplify:* similar quantities (e.g., all grains) are grouped together.

4. Examples (cont-d)

- *Reminder:* all grains are grouped together.
- *Problem:* we get slightly less buckwheat per area than wheat.
- So, to optimize grain production, we replace all buckwheat with wheat.
- *Example:* optimizing transportation.
- When trucks are stuck in traffic or under-loaded, we decrease *tonne-kilometers*.
- *At first glance:* maximizing tonne-kilometers is a good objective.
- *"Optimal" plan:* fully-loaded trucks circling Moscow :(
- *General saying:* Simplicity is worse than theft.

5. Question

- There is an anecdotal evidence of situations in which:
  - the use of simplified models in optimization
  - leads to absurd solutions.
- How frequent are such situations? Are they typical or rare?
- To answer this question, let us analyze this question from the mathematical viewpoint.

6. Reformulating the Question in Precise Terms

- In a general decision making problem:
  - we have a finite amount of resources, and
  - we need to distribute them between n possible tasks, so as to maximize the resulting outcomes.
- Examples:
  - a farmer allocates money to different crops, to maximize profits;
  - a city allocates police to different districts, to minimize crime.
- For simplicity, assume that all resources are of one type.
- We must distribute x_0 resources between n tasks, i.e., find $x_1, \ldots, x_n \geq 0$ such that $\sum\limits_{i=1}^{n} x_i = x_0$.

7. Reformulating the Question in Precise Terms (cont-d)

- We must distribute x_0 resources between n tasks, i.e., find $x_1, \ldots, x_n \geq 0$ such that $\sum\limits_{i=1}^{n} x_i = x_0$.
- In many practical problems, the amount of resources is reasonably small.
- So, we can safely linearize the objective function:

$$f(x_1, \ldots, x_n) \approx c_0 + \sum_{i=1}^{n} c_i \cdot x_i.$$

- So, the problem is:
  - maximize $c_0 + \sum\limits_{i=1}^{n} c_i \cdot x_i$
  - under the constraint $\sum\limits_{i=1}^{n} x_i = x_0$.

8. What Simplification Means in This Formulation

- Simplification means that we replace variables x_i with close values c_i with their sum.
- Let us assume that for all other variables x_k, we have already selected some values.
- Then, the problem is distributing the remaining resources X_0 to remaining tasks $x_1, \ldots, x_m$.
- The original problem is to maximize the sum $f(x_1, \ldots, x_m) = \sum\limits_{i=1}^{m} c_i \cdot x_i$ under the constraint $\sum\limits_{i=1}^{m} x_i = X_0$.
- The simplified problem is to maximize $s(x_1, \ldots, x_m) = \sum\limits_{i=1}^{m} c \cdot x_i$ under the constraint $\sum\limits_{i=1}^{m} x_i = X_0$.

9. The Simplified Description Provides a Reasonable Estimate for the Objective Function

- The approximation error $a \stackrel{\text{def}}{=} f(x_1, \ldots, x_m) - s(x_1, \ldots, x_m)$ is $a = \sum\limits_{i=1}^m \Delta c_i \cdot x_i$, where $\Delta c_i \stackrel{\text{def}}{=} c_i - c$.
- Let's assume that Δc_i are i.i.d., w/mean 0 and st. dev. σ.
- Thus, a has mean 0 and st. dev. $\sigma[a] = \sigma \cdot \sqrt{\sum\limits_{i=1}^m x_i^2}$.
- When resources are $\approx$ equally distributed $x_i \approx \dfrac{X_0}{m}$, we get $\sigma[a] = X_0 \cdot \dfrac{\sigma}{\sqrt{m}}$ and $s(x_1, \ldots, x_m) = c \cdot \sum\limits_{i=1}^m x_i = c \cdot X_0$.
- Thus, the relative inaccuracy of approximating f by s is $\dfrac{\sigma[a]}{s} = \dfrac{\sigma}{c \cdot \sqrt{m}}$; it is small when m is large.

10. Optimization: Case of the Original Objective Function

- The original problem is to maximize the sum $f(x_1, \ldots, x_m) = \sum\limits_{i=1}^{m} c_i \cdot x_i$ under the constraint $\sum\limits_{i=1}^{m} x_i = X_0$.
- From the mathematical viewpoint, this optimization problem is easy to solve:
  - to get the largest gain $\sum\limits_{i=1}^{m} c_i \cdot x_i$,
  - we should allocate all the resources X_0 to the task with the largest gain c_i per unit resource.
- In this case, the resulting gain is equal to $X_0 \cdot \max\limits_{i=1,\ldots,m} c_i$.

11. Optimization: Case of the Simplified Objective Function

- The simplified problem is to maximize $s(x_1, \ldots, x_m) = \sum\limits_{i=1}^{m} c \cdot x_i$ under the constraint $\sum\limits_{i=1}^{m} x_i = X_0$.
- For the simplified objective function, its value is the same no matter how we distribute the resources.
- In this case, the resulting gain is equal to $X_0 \cdot c$.
- Reminder: for the original objective function, the gain is $X_0 \cdot \max\limits_{i=1,\ldots,m} c_i$.
- For random variables, the largest value $\max c_i$ is often much larger than the average c.
- Moreover, the larger the sample size m, the more probable it is that the max is much larger than the average.

12. For Optimization, the Simplified Objective Function Can Lead to Drastic Non-Optimality

- Reminder: for the original objective function, the gain is $X_0 \cdot \max\limits_{i=1,\ldots,m} c_i$.
- Reminder: for the simplified objective function, the gain is $X_0 \cdot c$, where c is the average of c_i.
- In many application areas, especially in economics and finance, we encounter power-law distributions

$$\rho(x) \sim x^{-\alpha}.$$

- These distributions have heavy tails, with a high probability of c_i exceeding the average.
- Thus, the simplified model can indeed lead to very non-optimal solutions.

13. Acknowledgments

This work was supported in part:

- by the National Science Foundation grants HRD-0734825 and DUE-0926721, and
- by Grant 1 T36 GM078000-01 from the National Institutes of Health.