

4-2012

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

Martine Ceberio

The University of Texas at El Paso, mceberio@utep.edu

Olga Kosheleva

The University of Texas at El Paso, olgak@utep.edu

Vladik Kreinovich

The University of Texas at El Paso, vladik@utep.edu

Follow this and additional works at: https://scholarworks.utep.edu/cs_techrep



Part of the [Computer Sciences Commons](#), and the [Mathematics Commons](#)

Comments:

Technical Report: UTEP-CS-12-16

To appear in *Proceedings of the Fifth International Workshop on Constraint Programming and Decision Making CoProD'12*, Novosibirsk, Russia, September 23, 2012.

Recommended Citation

Ceberio, Martine; Kosheleva, Olga; and Kreinovich, Vladik, "Simplicity Is Worse Than Theft: A Constraint-Based Explanation of a Seemingly Counter-Intuitive Russian Saying" (2012). *Departmental Technical Reports (CS)*. 696.

https://scholarworks.utep.edu/cs_techrep/696

This Article is brought to you for free and open access by the Computer Science at ScholarWorks@UTEP. It has been accepted for inclusion in Departmental Technical Reports (CS) by an authorized administrator of ScholarWorks@UTEP. For more information, please contact lweber@utep.edu.

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,olgak,vladik}@utep.edu

Abstract. In many practical situations, simplified models, models that enable us to gauge the quality of different decisions reasonably well, lead to far-from-optimal situations when used in searching for an optimal decision. There is even an appropriate Russian saying: simplicity is worse than theft. In this paper, we provide a mathematical explanation of this phenomenon.

In science, simplicity is good. The world around us is very complex. One of the main objectives of science is to simplify it – and since has indeed greatly succeeded in doing it. For example, when Isaac Newton discovered his dynamical equations, it allowed him to explain the complex pictures of celestial bodies motion in terms of simple laws, laws that enable us to predict their positions for hundreds of years ahead.

From this viewpoint, a simplicity of a description is desirable. Yes, to achieve this simplicity, we sometimes ignore minor factors – but without such simplifying assumptions, it is difficult to make predictions, and predictions made based on these simplifying assumptions is usually reasonably good. For example, in his celestial mechanics studies, Newton ignored the fact that the planets and the Sun have finite size, and treated them as points with mass. To some extent, this simplifying assumption was justified: the gravitational field of a rotationally symmetric body is indeed the same as the field generated by the point with the same mass. However, real bodies are not perfectly symmetric, and there is a small discrepancy between the actual field and Newton’s approximate values.

In real life, simplified models – that seem to be working well for prediction – are sometimes disastrous when we move to decision making. One of the main purposes of science is to explain the world, to be able to predict what is happening in the world. Once this understanding is reached, once we have acquired the knowledge about the world, we use this knowledge to come up with actions that would make the world a better place.

For example, once the scientists get a better understanding of how cracks propagate through materials, it helps engineers design more stable constructions. Once the scientists learn about the life cycle of viruses, it helps develop medicines that prevent and cure the 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. Numerous examples of such an inefficiency can be found in the Soviet experiment with the global planning of economy; see, e.g., [11] and references therein. In spite of using latest techniques of mathematical economics, including the ideas of the Nobelist Wassily Leontieff [5] who started his research as a leading USSR economist, the results were sometimes disastrous.

For example, during the Soviet times, buckwheat – which many Russian like to eat – was often difficult to buy. A convincing after-the-fact explanation is based on the fact that if we describe the economy in too many details, the corresponding optimization problem becomes too complex to solve. To make it solvable, the problem has been simplified, usually by grouping several similar quantities together. For example, many different types of grains were grouped together into a single grain rubric. The corresponding part of the optimization task became as follows: optimize the overall grain production under the given costs. The problem is that for the same expenses, we can get slightly more wheat than buckwheat. As a result, when we optimize the overall grain production, buckwheat is replaced by wheat – and thus, the buckwheat production shrank.

A similar example related to transportation is described in [10]. One of the main symptoms of an inefficient use of trucks is that sometimes, instead of delivering goods, trucks spend too much time waiting to be loaded, or getting stuck in traffic. Another symptom is when a truck is under-loaded, so that a small load is inefficiently transported by an unnecessarily large truck. In view of these symptoms, a natural way to gauge the efficiency of a transportation company is to measure the amount of tonne-kilometers that it produced during a given time period. If this amount is close to the product of the overall truck capacity and the distance which can be covered during this time period, the company is more efficient; if this amount is much smaller, there is probably room for improvement. In view of this criterion, when the first efficient reasonably large-scale optimization algorithms appeared in the last 1960s, scientists decided to use these algorithms to find the optimal transportation schedule in and around Moscow – by optimizing the total number of tonne-kilometers. The program did find a mathematically optimal solution, but this solution was absurd: load full weight on all the trucks in the morning and let them go round and round the Circular Highway around Moscow :-)

In all these anecdotal examples, a simplified model – which works reasonably well in estimating the relative quality of the existing situations – leads to absurd solutions when used for optimization. Such situations were so frequent that there is a colorful Russian saying appropriate for this phenomenon: simplicity is worse than theft.

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.

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. For example, a farmer has a finite amount of money, and we need to allocate them to different possible crops so as to maximize the income. Police has a finite amount of officers, and we need to allocate them to different potential trouble spots so as to minimize the resulting amount of crime, etc. In some practical problems, we have limitations on several different types of resources, but for simplicity, we will assume that all resources are of one type.

Let x_0 be the total amount of resources, let n be the total number of possible tasks, and let x_1, \dots, x_n be the amounts allocated to different tasks; then, $x_i \geq 0$ and $x_1 + \dots + x_n = x_0$. Let $f(x_1, \dots, x_n)$ be the outcome corresponding to the allocation $x = (x_1, \dots, x_n)$. In many practical problems, the amount of resources is reasonably small. In such cases, we can expand the dependence $f(x_1, \dots, x_n)$ in the Taylor series and keep only linear terms in this expansion. In this case, the objective function takes a linear form $f(x_1, \dots, x_n) = c_0 + \sum_{i=1}^n c_i \cdot x_i$. The question is then to find the values $x_1, \dots, x_n \geq 0$ for which the sum $c_0 + \sum_{i=1}^n c_i \cdot x_i$ is the largest possible under the constraint that $\sum_{i=1}^n x_i = x_0$.

What does simplification means in this formulation. For this problem, simplification – in the sense of the above anecdotal examples – means that we replace individual variables by their sum. This can be always done if for two variables x_i and x_j , the coefficients c_i and c_j are equal. In this case, the sum of the corresponding terms in the objective function takes the form $c_i \cdot x_i + c_j \cdot x_j = c_i \cdot (x_i + x_j)$, so everything depends only on the sum $x_i + x_j$ (and does not depend on the individual values of these variables).

Since this replacement can be done exactly when the coefficients c_i and c_j are equal, it makes sense to perform a similar replacement when the coefficients c_i and c_j are close to each other. In this case, we replace both coefficients c_i and c_j , e.g., by their average. Similarly, if we have several variables with similar coefficients c_i , we replace all these coefficients by the average value.

Not all the variables have similar coefficients. Let us assume that for all other variables x_k , we have already selected some values, so only the variables with similar coefficients remain. In this case, the objective problem reduces to optimizing the sum $\sum c_i \cdot x_i$ over remaining variables, and the constraint take the form $\sum x_i = X_0$, where X_0 is equal to x_0 minus the sum of already allocated resources. If we now rename the remaining variables as x_1, \dots, x_m , we arrive at the following situation:

- the original problem is to maximize the sum $f(x_1, \dots, x_m) = \sum_{i=1}^m c_i \cdot x_i$ under the constraint $\sum_{i=1}^m x_i = X_0$;

- for simplicity, we replace this original problem by the following one: maximize the sum $s(x_1, \dots, x_m) = \sum_{i=1}^m c \cdot x_i$ under the constraint $\sum_{i=1}^m x_i = X_0$.

The simplified description provides, in general, a reasonable estimate for the objective function. Let us first show that the question is to estimate the value of the objective function corresponding to given allocations $x = (x_1, \dots, x_m)$, then the estimation provided by the simplified expression is reasonably accurate.

Indeed, due to many different factors, the actual values c_i differ from the average c . There are so many different reasons for this deviation, that it makes sense to assume that the deviations $\Delta c_i \stackrel{\text{def}}{=} c_i - c$ are independent identically distributed random variables, with mean 0 and some standard deviation σ . In this case, the approximation error $a \stackrel{\text{def}}{=} f(x_1, \dots, x_m) - s(x_1, \dots, x_m)$ takes the form $a = \sum_{i=1}^m \Delta c_i \cdot x_i$. Since all Δc_i are independent, with mean 0 and standard deviation σ , their linear combination a has mean 0 and standard deviation $\sigma[a] = \sigma \cdot \sqrt{\sum_{i=1}^m x_i^2}$. In particular, when the resources are approximately equally distributed between different tasks, i.e., $x_i \approx \frac{X_0}{m}$, this standard deviation is equal to $\sigma[a] = X_0 \cdot \frac{\sigma}{\sqrt{m}}$. The actual value of the objective function is approximately equal to $s(x_1, \dots, x_m) = c \cdot \sum_{i=1}^m x_i = c \cdot X_0$. Thus, the relative accuracy of approximating f by s can be described as the ratio $\frac{\sigma[a]}{s} = \frac{\sigma}{c \cdot \sqrt{m}}$. When m is large, this ratio is small, meaning that this simplification indeed leads to a very accurate estimation.

For optimization, the simplified objective function can lead to drastic non-optimality. From the mathematical viewpoint, the above optimization problem is easy to solve: to get the largest gain $\sum_{i=1}^m c_i \cdot x_i$, we should allocate all the resources X_0 to the task that leads to the largest amount of gain per unit resource, i.e., to the task with the largest possible value c_i . In this case, the resulting gain is equal to $X_0 \cdot \max_{i=1, \dots, m} c_i$.

On the other hand, for the simplified objective function, its value is the same no matter how we distribute the resources, and is equal to $X_0 \cdot c$. In this simplified problem, it does not matter how we allocate the resources between the tasks, so we can as well allocate them equally. In this case, the resulting gain is indeed equal to $X_0 \cdot c$.

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 we will observe values which are much larger than the average. This is especially true for power-law distributions which are frequent in economics and

finance; see, e.g., [1–4, 6–9, 12–14]. These distributions have heavy tails, with a high probability of c_i exceeding the mean. Thus, the simplified model can indeed lead to very non-optimal solutions.

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.

References

1. Chakrabarti, B. K., Chakraborti, A., Chatterjee, A.: *Econophysics and Sociophysics: Trends and Perspectives*, Wiley-VCH, Berlin (2006)
2. Chatterjee, A., Yarlagadda, S., Chakrabarti, B. K.: *Econophysics of Wealth Distributions*, Springer-Verlag Italia, Milan (2005)
3. Farmer, J. D., Lux, T. (eds.), *Applications of statistical physics in economics and finance*, a special issue of the *Journal of Economic Dynamics and Control*, **32**(1) (2008) 1–320
4. Gabaix, X., Parameswaran, G., Vasiliki, P., Stanley, H. E.: Understanding the cubic and half-cubic laws of financial fluctuations, *Physica A*, **324** (2003) 1–5
5. Leontieff, W.: *Input-Output Economics*, Oxford University Press, New York (1986)
6. Mantegna, R. N., Stanley, H. E.: *An Introduction to Econophysics: Correlations and Complexity in Finance*, Cambridge University Press, Cambridge, Massachusetts, (1999)
7. McCauley, J.: *Dynamics of Markets, Econophysics and Finance*, Cambridge University Press, Cambridge, Massachusetts (2004)
8. Rachev, S. T., Mittnik, S.: *Stable Paretian Models in Finance*, Wiley Publishers, New York (2000)
9. Roehner, B.: *Patterns of Speculation - A Study in Observational Econophysics*, Cambridge University Press, Cambridge, Massachusetts (2002)
10. Romanovsky, J. V.: *Lectures on Mathematical Economics*, St. Petersburg University, Russia (1972)
11. Shmelev, N. Popov, V.: *The Turning Point*, Doubleday, New York (1989)
12. Stanley, H. E., Amaral, L. A. N., Gopikrishnan, P., Plerou, V.: Scale invariance and universality of economic fluctuations, *Physica A*, **283** (2000) 31–41.
13. Stoyanov, S. V., Racheva-Iotova, B., Rachev, S. T., Fabozzi, F. J.: Stochastic models for risk estimation in volatile markets: a survey, *Annals of Operations Research*, **176** (2010) 293–309
14. Vasiliki, P., Stanley, H. E.: Stock return distributions: tests of scaling and universality from three distinct stock markets, *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, **77**(3) (2008) Publ. 037101