12-2016

A Simple Geometric Explanation of Occam's Razor

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

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

Comments:
Technical Report: UTEP-CS-16-99

A Simple Geometric Explanation of Occam's Razor

Olga Kosheleva and Vladik Kreinovich
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
olgak@utep.edu, vladik@utep.edu

Abstract

Occam's razor states that out of possible explanations, plans, and designs, we should select the simplest one. It turns out that in many practical situations, the simplest explanation indeed turns out to be the correct one, the simplest plan is often the most successful, etc. But why this happens is not very clear. In this paper, we provide a simple geometric explanation of Occam's razor.

1 Formulation of the Problem

Occam razor: reminder. The famous principle – attributed to a medieval philosopher William of Ockham – is that we should look for the simplest possible explanation to different phenomena, for the simplest possible plans and designs, etc.

Metaphorically, if we have an explanation or a design that is too complicated, we should try to "cut off" unnecessary parts – just like a razor can cut off unnecessary parts of the beard. This is the reason why this principle is known as Occam's razor.

Occam razor is a very efficient idea. Interestingly, the Occam razor principle not only leads to a simple explanations and designs, but it also helps to find an explanation or a design in the first place; see, e.g., [1, 2, 3].

With the advent of computers, is this principle still important? In the past, before computers were invented and all the computations had to be performed largely by hand, having a simple easy-to-calculate model was a necessity. However, nowadays, with powerful computers readily available, it should not matter whether a model is slightly more complicated than necessary – computations are fast anyway.

Yes, this principle is still important. It should not matter, but, surprisingly, it does matter: in many practical situations, when we find a way to sim-

plify a model, not only we decrease the computation time, but we also increase the accuracy and efficiency of the model predictions.

Specifically, in many cases, the simplest explanation later turns out to be more adequate than more complex ones, the simplest plan turns out to be more efficient than more complex ones, etc.

But why? But why are simpler models more adequate? There seems to be a law of nature according to which nature prefers simple models.

What we do in this paper. In this paper, we provide a simple geometric explanation for Occam'a razor.

2 Explanation

Geometric reformulation of the problem. All the above problems – explanation, planning, and design – can be interpreted in geometric terms, as the desire to go from point A to point B. Here, point A is the starting point, in which we have a phenomenon that we try to explain or a task that we need to perform, and point B is the endpoint, when the task is explained and the task has been performed.

Usual features of the corresponding geometric transition. The transition from point A (the original state) to point B (the desired state) is rarely done in one step. Usually, there are several consequent steps.

For example, we first explain some part of the phenomena, then another part, etc. Similarly, a plan usually consists of several steps.

In general, to get from point A to point B, we use several intermediate steps: first, we go from point A to the first intermediate point A_1, then we go from A_1 to the second intermediate point A_2, etc., until after several steps, we go from the last intermediate state A_s to the desired point B.

Reformulating Occam's razor in geometric terms. In the first approximation, the complexity of a plan can be characterized by the number of steps in this plan – the larger number of steps, the more complicated the corresponding plan.

In these geometric terms, Occam's razor means simply that the number s of intermediate steps in the transition from point A to point B should be the smallest possible.

What happens in reality? To understand the efficiency of Occam's razor, let us recall what happens to plans in real life, when these plans are being implemented.

What happens is that no plan is implemented exactly, there is always some inevitable deviation from the original plan – sometimes a minor deviation, sometimes a more significant deviation. This is life.

How does this affect the result of implementing a plan. On each stage of the plan, whether we go from $\mathrm{A}_0 \stackrel{\text{def}}{=} \mathrm{A}$ to A_1, or from A_1 to A_2, ..., or from

A_s to $\mathrm{A}_{s+1} \stackrel{\text{def}}{=} \mathrm{B}$, we add a little inaccuracy to the result. We add such an inaccuracy on each of the $s+1$ stages, so the resulting inaccuracy I is equal to the sum of $s+1$ inaccuracies I_i corresponding to each step: $I = I_1 + \ldots + I_{s+1}$.

Inaccuracies are not easy to predict: if they were predictable, we would have taken them into account. Thus, inaccuracies can be viewed as random variables. In our general case, there is no reason to believe that some stages can be implemented more accurately than others. Therefore, when simulating these uncertainties, it makes sense to assume that they have the same inaccuracy, i.e., that the corresponding random variables have the same standard deviation σ.

Similarly, in the general case, there is no reason to believe that there is positive or negative correlation between inaccuracies corresponding to different stages. Therefore, it makes sense to assume that all $s+1$ random variables are independent from each other.

It is known that the variance of the sum of independent variables is equal to the sum of the variances of these variables. In the sum $I = I_1 + \ldots + I_{s+1}$, the variance of each of the terms I_i is equal to σ^2. Thus, the variance of the sum I is equal to $V_I = (s+1) \cdot \sigma^2$, and so, the standard deviation σ_I of the sum I is equal to $\sigma_I = \sqrt{V_I} = \sqrt{s+1} \cdot \sigma$. Now, we are ready for the desired explanation.

Resulting explanation. The variable I describes the deviation of the actual result of implementing this plan from the desired point B. The larger the number s of intermediate stages in a plan, the larger the standard deviation of I, and thus, on average, the further we are from the desired destination point B.

Thus, in general, the fewer stages the plan has, the closer will this plan bring us to the desired state. This is exactly what we observe in the practice, that Occam's razor is an efficient planning tool. Therefore, the above simple geometric model indeed explains why Occam's razor is empirically efficient.

Comments.

- In the above text, we talked about imperfection of plans, but we could as well talk about imperfection of design or imperfections of an explanation. Thus, the above argument also explains why Occam's razor is efficient in design and explanations.

- Occam's razor provides a possible explanation for a somewhat mysterious Biblical commandment not to pronounce God's name in vain. Indeed, from the religious viewpoint, each stage of a plan, design, or explanation needs God's help – and thus, often requires an explicit mention of God's name in the plead for this help. From this viewpoint, this commandment simply means that we should minimize the number of stages as much as possible – and this is exactly Occam's razor.

Acknowledgments

This work was supported by the National Science Foundation grants HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and DUE-0926721, and by an award "UTEP and Prudential Actuarial Science Academy and Pipeline Initiative" from Prudential Foundation.

References

[1] B. H. Friesen and V. Kreinovich, "Ockham's razor in interval identification", *Reliable Computing*, 1995, Vol. 1, No. 3, pp. 225–238.

[2] O. Kosheleva and V. Kreinovich, "Occam's Razor Explains Matthew Effect", *Journal of Innovative Technology and Education*, 2015, Vol. 2, No. 1, pp. 47–50.

[3] M. Li and P. M. B. Vitányi, *An Introduction to Kolmogorov Complexity*, Springer-Verlag, Berlin, Heidelberg, New York, 2008.