Why Superellipsoids: A Probability-Based Explanation

Pedro Barragan and Vladik Kreinovich
Department of Computer Science
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
pabarraganolague@miners.utep.edu
vladik@utep.edu

Abstract

In many practical situations, it turns out that the set of possible values of the deviation vector is (approximately) a super-ellipsoid. In this paper, we provide a theoretical explanation for this empirical fact – an explanation based on the natural notion of scale-invariance.

1 Formulation of the Problem

Need to describe uncertainty domains. The intent of the mass production of a gadget is to produce gadgets with identical values $(x_1, \ldots, x_n)$ of the desired characteristics x_i. In reality, of course, different gadgets end up having slightly different values $\widetilde{x}_i$ of these characteristics: $\Delta x_i \stackrel{\text{def}}{=} \widetilde{x}_i - x_i \neq 0$. For each of these characteristics x_i, we usually have a tolerance bound Δ_i for which $|\Delta x_i| \leq \Delta_i$, so that possible values of Δx_i form an interval $[-\Delta_i, \Delta_i]$. Thus, possible values of the deviation vector $\Delta x = (\Delta x_1, \ldots, \Delta x_n)$ are located in the box $[-\Delta_1, \Delta_1] \times \ldots \times [-\Delta_n, \Delta_n]$.

In practice, not all vectors Δx from this box are possible. It is therefore desirable to describe the set of all possible deviation vectors Δx. This set is known as the *uncertainty domain.*

Shall not we also determine probabilities? At first glance, it seems that we should be interested not only in finding out which deviation vectors Δx are possible and which are not, but also in how frequent different possible vectors are. In other words, we should be interested not only in the uncertainty domain, but also in the probability distribution on this domain.

In reality, however, it is not possible to find these probabilities. Indeed, the manufacturing process may slightly change (and often does change). After each such change, the tolerance intervals and the resulting uncertainty domain remain largely unchanged, but the probabilities change (often drastically).

Empirical shapes of uncertainty domains. Empirical analysis has shows that in many practical cases, the uncertainty domain can be well approximated by a *super-ellipsoid* $\sum_{i=1}^{n}\left(\frac{|\Delta x_i|}{\sigma_i}\right)^p \le C$ for some values σ_i, p, and C, and the accuracy of this approximation is higher than for other approximation families with the same number of parameters; see, e.g., [4, 5].

Historical comment. Super-ellipsoids were first successfully used to describe uncertainty domain in [3]. Super-ellipsoids are also actively used in image processing, to describe different components of an image; see, e.g., [2, 6, 7, 10].

What we do in this paper. In this paper, we provide a theoretical explanation for this empirical phenomenon.

2 Our Idea

Let us apply probabilistic approach. In reality, there *is* some probability distribution $\rho_i(\Delta x_i)$ for each of the random variables Δx_i. Since we have no reason to assume that positive values of Δx_i are more probable than negative values or vice versa, it makes sense to assume that they are equally probable, i.e., that each distribution $\rho_i(\Delta x_i)$ is symmetric: $\rho_i(\Delta x_i) = \rho_i(|\Delta x_i|)$.

Similarly, since we have no reasons to believe that different deviations are statistically dependent, it makes sense to assume that the corresponding random variables are independent. In this case, the overall probability density function (pdf) has the form $\rho(\Delta x) = \prod_{i=1}^{n} \rho_i(|\Delta x_i|)$.

Usually, we consider a deviation vector possible if its probability exceed a certain threshold t. Thus, the desired set has the form $S_t \stackrel{\text{def}}{=} \{\Delta x : \rho(\Delta x) \ge t\}$.

Scaling and scale-invariance: an informal description. Numerical values of the deviations Δx_i depend on the choice of a measuring unit; if we replace the original unit by a unit which is λ times smaller, then for the exact same physical situation, we get the new numerical values $\Delta x'_i = \lambda \cdot \Delta x_i$.

Since the physics remains the same, it makes sense to require that the uncertainty domains do not change under such a re-scaling.

To be more precise, the pdf threshold t may change, but the family of such sets should remain unchanged: $\{S'_t\}_t = \{S_t\}_t$, where S'_t corresponds to the re-scaled pdf $\rho'(\Delta x) = \text{const} \cdot \rho(\lambda \cdot \Delta)$.

We will prove that under this scale-invariance, the corresponding sets S_t are exactly super-ellipsoids. Thus, we will get the desired explanation.

3 Definitions and the Main Result

Definition. *Let $n > 1$, and let $\rho(y) = (\rho_1(y_1), \ldots, \rho_n(y_n))$ be a tuple of positive symmetric ($\rho_i(-y_i) = \rho_i(y_i)$) smooth functions of one variable.*

- *For every $t > 0$, let us denote the set* $\left\{(y_1, \ldots, y_n) : \prod\limits_{i=1}^{n} \rho_i(y_i) \geq t\right\}$ *by $S_t(\rho)$.*
- *We say that a tuple $\rho(y)$ is* bounded *if the set $S_t(\rho)$ is bounded for every t.*
- *For every $\lambda > 0$, by a* λ-re-scaling *of the tuple $\rho(x)$, we mean a tuple $\rho_\lambda(y)$, for which $\rho_{\lambda,i}(y_i) \stackrel{\text{def}}{=} \dfrac{1}{\lambda} \cdot \rho_i(\lambda \cdot y_i)$.*
- *We say that a tuple $\rho(y)$ is* scale-invariant *if for every $\lambda > 0$, re-scaling does not change the family S_t: $\{S_t(\rho)\}_t = \{S_t(\rho_\lambda)\}_t$.*

Main Result. *If the tuple $\rho(y)$ is bounded and scale-invariant, then each set $S_t(\rho)$ is a super-ellipsoid.*

Comments.

- Vice versa, it is easy to prove that each super-ellipsoid can be represented as a set S_t for the bounded and scale-invariant distributions of the type $\rho_i(y_i) = \text{const} \cdot \exp\left(-\dfrac{|y_i|^p}{\sigma_i^p}\right)$. Such probability distributions indeed occur as probability distributions of measuring errors corresponding to some measuring instruments; see, e.g., [9].
- Processing super-ellipsoids is similar to processing ellipsoids; see, e.g., [8].

Proof. For convenience, let us consider logarithms $\psi_i(y_i) \stackrel{\text{def}}{=} -\log(\rho_i(y_i))$. Once we take the negative logarithm of both sides of the inequality $\prod\limits_{i=1}^{n} \rho_i(y_i) \geq t$ that describes the set $S_t(\rho)$, we get an equivalent description $\sum\limits_{i=1}^{n} \psi_i(y_i) \leq c$, where we denoted $c \stackrel{\text{def}}{=} -\log(t)$. In these terms, scale-invariance means that the corresponding family of sets is the same for all c.

In terms of the new functions $\psi_i(y_i)$, scaling means

$$\psi_{\lambda,i}(y_i) = -\ln(\rho_{\lambda,i}(y_i)) = -\log\left(\frac{1}{\lambda} \cdot \rho_i(\lambda \cdot y_i)\right) =$$

$$\log(\lambda) - \log(\rho_i(\lambda \cdot y_i)) = \psi_i(\lambda \cdot y_i) + \log(\lambda), \tag{1}$$

i.e., has the form $\psi_{\lambda,i}(y_i) = \psi_i(\lambda \cdot y_i) + \log(\lambda)$.

In these terns, the fact that scaling does not change the family of sets S_t means that if two tuples $(y_1, \ldots, y_n)$ and $(z_1, \ldots, z_n)$ always belong or not belong to the same sets – i.e., have the same value of the corresponding sum

$\sum\limits_{i=1}^n \psi_i(y_i) = \sum\limits_{i=1}^n \psi_i(z_i)$, then the re-scaled functions should also have the same value of the sum, i.e., $\sum\limits_{i=1}^n \psi_{\lambda,i}(y_i) = \sum\limits_{i=1}^n \psi_{\lambda,i}(z_i)$. Substituting the above expression for $\psi_{\lambda,i}(y_i)$ into this formula, we get

$$\sum_{i=1}^n (\psi_i(\lambda \cdot y_i) + \log(\lambda)) = \sum_{i=1}^n (\psi_i(\lambda \cdot z_i) + \log(\lambda), \tag{2}$$

i.e.,

$$n \cdot \lambda + \sum_{i=1}^n \psi_i(\lambda \cdot y_i) = n \cdot \lambda + \sum_{i=1}^n \psi_i(\lambda \cdot z_i). \tag{3}$$

Subtracting $n \cdot \lambda$ from both sides of this equality, we get

$$\sum_{i=1}^n \psi_i(\lambda \cdot y_i) = \sum_{i=1}^n \psi_i(\lambda \cdot z_i). \tag{4}$$

Thus, we have the following property:

- if $\sum\limits_{i=1}^n \psi_i(y_i) = \sum\limits_{i=1}^n \psi_i(z_i)$,
- then $\sum\limits_{i=1}^n \psi_i(\lambda \cdot y_i) = \sum\limits_{i=1}^n \psi_i(\lambda \cdot z_i)$.

In particular, this property holds if we perform very small changes to only two of the values y_i, i.e., if for some $a \neq b$, we replace y_a with $z_a = y_a + \delta_a$ and y_b with $z_b = y_b + \delta_b$; for $i \neq a, b$, we take $z_i = y_i$.

In this case,

$$\psi_a(z_a) = \psi_a(y_a + \delta_a) = \psi_a(y_a) + \psi'_a(y_a) \cdot \delta_a + o(\delta), \tag{5}$$

where $\psi'_a(y_a)$, as usual, denotes the derivative of the function $\psi_a(y_a)$. Similarly, we have

$$\psi_b(z_b) = \psi_b(y_b + \delta_b) = \psi_b(y_b) + \psi'_b(y_b) \cdot \delta_b + o(\delta). \tag{6}$$

Thus,

$$\sum_{i=1}^n \psi_i(z_i) = \sum_{i=1}^n \psi_i(y_i) + \psi'_a(y_a) \cdot \delta_a + \psi'_b(y_b) \cdot \delta_a + o(\delta), \tag{7}$$

and the original equality $\sum\limits_{i=1}^n \psi_i(y_i) = \sum\limits_{i=1}^n \psi_i(z_i)$ takes the form

$$\psi'_a(y_a) \cdot \delta_a + \psi'_b(y_b) \cdot \delta_b + o(\delta) = 0. \tag{8}$$

Similarly, we have

$$\psi_a(\lambda \cdot z_a) = \psi_a(\lambda \cdot (y_a + \delta_a)) = \psi_a(\lambda \cdot y_a + \lambda \cdot \delta_a) =$$

$$\psi_a(\lambda \cdot y_a) + \lambda \cdot \psi'_a(\lambda \cdot y_a) \cdot \delta_a + o(\delta) \tag{9}$$

and

$$\psi_b(\lambda \cdot z_b) = \psi_b(\lambda \cdot (y_b + \delta_b)) = \psi_b(\lambda \cdot y_b + \lambda \cdot \delta_b) =$$
$$\psi_b(\lambda \cdot y_b) + \lambda \cdot \psi'_b(\lambda \cdot y_b) \cdot \delta_b + o(\delta). \tag{10}$$

Thus, the equality $\sum\limits_{i=1}^n \psi_i(\lambda \cdot y_i) = \sum\limits_{i=1}^n \psi_i(\lambda \cdot z_i)$ takes the form

$$\lambda \cdot \psi'_a(\lambda \cdot y_a) \cdot \delta_a + \lambda \cdot \psi'_b(\lambda \cdot y_b) \cdot \delta_b + o(\delta) = 0, \tag{11}$$

i.e., equivalently,

$$\psi'_a(\lambda \cdot y_a) \cdot \delta_a + \psi'_b(\lambda \cdot y_b) \cdot \delta_b + o(\delta) = 0. \tag{12}$$

So, the scale-invariance condition takes the following form:

- if $\psi'_a(y_a) \cdot \delta_a + \psi'_b(y_b) \cdot \delta_b + o(\delta) = 0$,
- then $\psi'_a(\lambda \cdot y_a) \cdot \delta_a + \psi'_b(\lambda \cdot y_b) \cdot \delta_b + o(\delta) = 0$.

The first condition is satisfied if we choose

$$-\frac{\delta_b}{\delta_a} = \frac{\psi'_a(y_a)}{\psi'_b(y_b)} + o(\delta). \tag{13}$$

The satisfaction of the second condition then means that

$$-\frac{\delta_b}{\delta_a} = \frac{\psi'_a(\lambda \cdot y_a)}{\psi'_b(\lambda \cdot y_b)} + o(\delta), \tag{14}$$

i.e., that

$$\frac{\psi'_a(\lambda \cdot y_a)}{\psi'_b(\lambda \cdot y_b)} = \frac{\psi'_a(y_a)}{\psi'_b(y_b)} + o(\delta). \tag{15}$$

Since this is true for all δ, then we can take $\delta \to 0$ and conclude that

$$\frac{\psi'_a(\lambda \cdot y_a)}{\psi'_b(\lambda \cdot y_b)} = \frac{\psi'_a(y_a)}{\psi'_b(y_b)}. \tag{16}$$

This equality is equivalent to

$$\frac{\psi'_a(\lambda \cdot y_a)}{\psi'_a(y_a)} = \frac{\psi'_b(\lambda \cdot y_b)}{\psi'_b(y_b)}. \tag{17}$$

The left-hand side of this equality does not depend on y_b; thus, the right-hand side does not depend on y_b either. Hence, this ratio depends only on λ. Let us denote this common ratio by $r(\lambda)$. Then, for each a, we have

$$\frac{\psi'_a(\lambda \cdot y_a)}{\psi'_a(y_a)} = r(\lambda), \tag{18}$$

i.e., equivalently,

$$\psi_a'(\lambda \cdot y_a) = r(\lambda) \cdot \psi_a'(y_a). \tag{19}$$

The derivative of a smooth function is always measurable, and thus, the function $r(\lambda)$ is also measurable, as a ratio of two measurable functions.

Now, let us take arbitrary values $\lambda_1 > 0$ and $\lambda_2 > 0$. Then, we can re-scale first by λ_1, then by λ_1, or we can right away re-scale by $\lambda = \lambda_1 \cdot \lambda_2$. In the first case, the above formula has the form

$$\psi_a'(\lambda_2 \cdot y_a) = r(\lambda_2) \cdot \psi_a'(y_a) \tag{20}$$

and then

$$\psi'(\lambda_1 \cdot (\lambda_2 \cdot y_a)) = r(\lambda_1) \cdot \psi'(\lambda_2 \cdot y_a) = r(\lambda_1) \cdot r(\lambda_2) \cdot \psi_a'(y_a), \tag{21}$$

i.e.,

$$\psi'(\lambda_1 \cdot \lambda_2 \cdot y_a) = r(\lambda_1) \cdot r(\lambda_2) \cdot \psi_a'(y_a). \tag{22}$$

In the second case, we get

$$\psi'(\lambda_1 \cdot \lambda_2 \cdot y_a) = r(\lambda_1 \lambda_2) \cdot \psi_a'(y_a). \tag{23}$$

Since the left-hand sides of the two equalities (22) and (23) coincide, their right-hand sides must coincide as well, i.e., we must have $r(\lambda_1 \cdot \lambda_2) = r(\lambda_1) \cdot r(\lambda_2)$.

It is known (see, e.g., [1]) that all measurable functions satisfying this property have the form $r(\lambda) = \lambda^\beta$ for some real number β. Now, from the condition

$$\psi_a'(\lambda \cdot y_a) = r(\lambda) \cdot \psi_a'(y_a) = \lambda^\beta \cdot \psi_a'(y_a), \tag{24}$$

for $\lambda = z$ and $y_a = 1$, we conclude that $\psi_a'(z) = \psi_a'(1) \cdot z^\beta$, i.e., that $\psi_a'(y_a) = c_a \cdot y_a^\beta$ for some constant c_a.

Integrating, for $\beta \neq -1$, for $y_a > 0$, we get $\psi_a(y_a) = k_a \cdot y_a^p + C_a$ for $p = \beta + 1$, $k_a \stackrel{\text{def}}{=} \dfrac{c_a}{\beta + 1}$, and for some constant C_a. Since each function $\psi_i(y_i)$ is even, we get $\psi_i(y_i) = k_i \cdot |y_i|^p + C_i$.

So, the condition $\sum\limits_{i=1}^n \psi_i(y_i) \le c$ takes the equivalent form

$$\sum_{i=1}^n k_i \cdot |y_i|^p \le c_0 \stackrel{\text{def}}{=} c - \sum_{i=1}^n C_i, \tag{25}$$

i.e., the form of the super-ellipsoid. For this super-ellipsoid to be bounded, we need to have $p > 0$.

To complete the proof, it is sufficient to consider the case when $\beta = -1$. In this case, integration leads to $\psi_i(y_i) = k_i \cdot \ln(|y_i|) + C_i$, so the condition $\sum\limits_{i=1}^n \psi(y_i) \le c$ takes the form $\sum\limits_{i=1}^n k_i \cdot \ln(|y_i|) \le c_0 \stackrel{\text{def}}{=} c - \sum\limits_{i=1}^n C_i$.

Exponentiating both sides, we get an equivalent inequality $\prod\limits_{i=1}^n |y_i|^{k_i} \le \exp(C)$, for which the corresponding set S_t is unbounded.

So, in the bounded cases, we always have a super-ellipsoid. The result is proven.

Acknowledgments

This work was supported in part 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] J. Aczel, *Lectures on Functional Equations and Their Applications*, Dover Publ., New York, 2006.

[2] E. Dura, J. Bell, and D. Lane, "Superellipse fitting for the recovery and classification of mine-like shapes in sidescan sonar images", *IEEE Journal of Oceanic Engineering*, 2008, Vol. 33, No. 4, pp. 434–444.

[3] I. Elishakoff and Y. Bekel, "Application of Lamé's super ellipsoids to model initial imperfection", *ASME Journal of Applied Mechanics*, 2013, Vol. 80, No. 6, paper 061006.

[4] I. Elishakoff and F. Elettro, "Interval, ellipsoid, and super-ellipsoid calculi for experimental and theoretical treatment of uncertainty: which one ought to be preferred?", *International Journal of Solid Structures*, 2014, Vol. 51, No. 7, pp. 1576–1586.

[5] I. Elishakoff, C. M. Fu, C. Jiang, B. Y. Ni, X. Han, and G. S. Chen, "Comparison of uncertainty analyses for crankshaft applications", *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems. Part B: Mechanical Engineering*, Vol. 1, Paper 041002.

[6] D. Köhntopp, B. Lehmann, and A. Birk, "Efficient superellipse fitting based contour extraction for mine-like shape recognition", *Proceedings of the 2nd International Conference on Underwater Acoustincs UA'2014*, 2014.

[7] D. Köhntopp, B. Lehmann, D. Kraus, and A. Birk, "Segmentation and classification using active contours based superellipse fitting on side scan sonar images for marine demining", *Proceedings of the 2015 IEEE International Conference on Robotics and Automation ICRA'2015*, Seattle, Washington, May 26–30, 2015.

[8] V. Kreinovich, A. Neumaier, and G. Xiang, "Towards a combination of interval and ellipsoid uncertainty", *Computational Technologies*, 2008, Vol. 13, No. 6, pp. 5–16.

[9] P. V. Novitskii and I. A. Zograph, *Estimating the Measurement Errors*, Energoatomizdat, Leningrad, 1991 (in Russian).

[10] L. Saroul, O. Bernard, D. Vray, and D. Friboulet, "Prostate segmentation in echographic images: a bariational approach using deformable super-ellipse and Rayleigh distribution", *Proceedings of the 5th IEEE International Symposium on Biomedial Imaging ISBI'2008*, May 2008, pp. 129–132.