Why LINEX (Linear Exponential) Loss Functions?

Emilio Ramirez and Vladik Kreinovich
Department of Computer Science, University of Texas at El Paso
El Paso, TX 79968, USA, ejramirez5@miners.utep.edu, vladik@utep.edu

Empirical fact. In design and control, we want to select the values of the parameters that minimize the losses. Traditionally, the dependence of loss on the parameter is described by a quadratic function. However, for this function, positive and negative deviations from the optimal value lead to the exact same increase in loss, while in practice, the effects are often different: e.g., for a refrigerator, a decrease in temperature is rarely harmful, while an increase can spoil the food.

It turns out (see, e.g., [1, 2, 3]) that in many practical situations, in the next approximation, the best "asymmetric" loss function is a linear combination (called LINEX) of a linear and exponential functions: $L(x) = c \cdot \exp(a \cdot x) - b \cdot x + a_0$. How can we explain this empirical fact?

Our explanation. Quadratic functions are linear combinations of smooth functions 1, x, and x^2. The class of such functions does not change if we change the starting point for measuring x, i.e., if we replace x with $x + x_0$ for some x_0. Let us thus look for smooth families of the type $c_1 \cdot f_1(x) + c_2 \cdot f_2(x) + c_3 \cdot f_3(x)$ which are similarly "shift-invariant". Shift-invariance means, in particular, that all shifted functions $f_i(x + x_0)$ belong to the same family, i.e., that

$$f_i(x + x_0) = \sum_j c_{ij}(x_0) \cdot f_j(x) \tag{1}$$

for some value $c_{ij}(x_0)$.

In particular, for each i and x_0, by selecting three different values x_k, we get a linear system $f_i(x_k + x_0) = \sum\limits_j c_{ij}(x_0) \cdot f_j(x_k)$ for three unknowns $c_{ij}(x_0)$, $j = 1, 2, 3$. By Cramer's rule, the values $c_{ij}(x_0)$ are rational functions of values $f_i(x_k + x_0)$ and $f_j(x_k)$ that smoothly depend on x_0. Thus, the functions $c_{ij}(x_0)$ are also differentiable. So, we can differentiate the equality (1) with respect to x_0. After the differentiation, we can take $x_0 = 0$ and get $f_i'(x) = \sum a_{ij} \cdot f_j(x)$, where $a_{ij} \stackrel{\text{def}}{=} c_{ij}'(0)$. Solutions to such linear differential equations with constant coefficients are well known: they are linear combinations of terms $x^k \cdot \exp(\lambda \cdot x)$, where λ is an eigenvalue of the matrix c_{ij}, and $k = 0, 1, 2, \ldots$; $k > 0$ corresponds to the case of equal eigenvalues.

Quadratic functions correspond to the case when all three eigenvalues are equal to 0. The simplest modification is when one of the eigenvalues becomes different from 0, while the other two remain 0s. This corresponds to the loss functions $c_1 \cdot \exp(\lambda \cdot x) + c_2 \cdot x + c_2$, i.e., exactly to LINEX. Thus, the practical efficiency of LINEX loss functions can be naturally explained.

References

[1] Yen-Chang Chang and Wen-Liang Hung, "LINEX loss functions with applications to determining the optimum process parameters", *Quality & Quantity*, 2007, Vol. 41, pp. 291–301.

[2] H. R. Varian, "A Bayesian approach to real estate assessment", In: S. E. Fienberg and A. Zellner (eds.), *Studies in Bayesian Econometrics and Statistics in Honor of Leonard L. Savage*, Noth Holland, Amsterdam, 1975, 1975, pp. 195–208.

[3] A. Zellner, "Bayesian estimation and prediction using asymmetric loss functions", *Journal of the American Statistical Association*, 1986, Vol. 81, pp. 446–451.