

Why Polynomial Formulas in Soft Computing, Decision Making, etc.?

Olga Kosheleva, *Member, IEEE*, and Martine Ceberio

Abstract—We show that in many application areas including soft constraints reasonable requirements of scale-invariance lead to polynomial formulas for combining degrees (of certainty, of preference, etc.)

I. PARTIAL ORDERS NATURALLY APPEAR IN MANY APPLICATION AREAS

One of the main objectives of science and engineering is to help people select decisions which are the most beneficial to them. To make these decisions,

- we must know people’s preferences,
- we must have the information about different events – possible consequences of different decisions, and
- since information is never absolutely accurate and precise, we must also have information about the degree of certainty.

All these types of information naturally lead to partial orders:

- For preferences, $a < b$ means that b is preferable to a . This relation is used in *decision theory*; see, e.g., [5].
- For events, $a < b$ means that a can influence b . This causality relation is used in *space-time physics*.
- For uncertain statements, $a < b$ means that a is less certain than b . This relation is used in logics describing uncertainty such as fuzzy logic (see, e.g., [2], [7]) and in soft constraints.

II. NUMERICAL CHARACTERISTICS RELATED TO PARTIAL ORDERS

While an order may be a natural way of describing a relation, orders are difficult to process, since most data processing algorithms process numbers. Because of this, in all three application areas, numerical characteristics have appeared that describe the corresponding orders:

- in decision making, *utility* describes preferences:
 $a < b$ if and only if $u(a) < u(b)$;
- in space-time physics, *metric* (and time coordinates) describes causality relation;
- in logic and soft constraints, numbers from the interval $[0, 1]$ are used to describe degrees of certainty; see, e.g., [2], [7].

Olga Kosheleva and Martine Ceberio are with the University of Texas at El Paso, 500 W. University, El Paso, TX 79968, USA (emails: olgak@utep.edu, mceberio@utep.edu).

III. NEED TO COMBINE NUMERICAL CHARACTERISTICS, AND THE EMERGENCE OF POLYNOMIAL AGGREGATION FORMULAS

- In decision making, we need to combine utilities

$$u_1, \dots, u_n$$

of different participants. Nobelist Josh Nash showed that reasonable conditions lead to

$$u = u_1 \cdot \dots \cdot u_n;$$

see, e.g., [5], [6].

- In space-time geometry, we need to combine coordinates x_i and x'_i of two events into a “metric”, i.e., into the formula for the proper time s between the two events’ reasonable conditions lead to polynomial metrics such as Minkowski metric in which

$$s^2 = c^2 \cdot (x_0 - x'_0)^2 -$$

$$[(x_1 - x'_1)^2 + (x_2 - x'_2)^2 + (x_3 - x'_3)^2]$$

and of a more general (“Riemann”) metric in which the proper time Δs between the two nearby events whose coordinates differ by Δx_i is determined by the formula

$$(\Delta s)^2 = \sum_{i,j} g_{ij}(x) \cdot \Delta x^i \cdot \Delta x^j$$

for some function $g_{ij}(x)$.

- In fuzzy logic and soft constraints, we must combine degrees of certainty d_i in A_i into a degree $d_{\&}$ for $A_1 \& A_2$ or a degree d_{\vee} for $A_1 \vee A_2$; reasonable conditions lead to polynomial functions like

$$d_{\&} = d_1 \cdot d_2$$

and

$$d_{\vee} = d_1 + d_2 - d_1 \cdot d_2.$$

IV. IN MATHEMATICAL TERMS, POLYNOMIAL FORMULAS ARE TENSOR-RELATED

A general polynomial dependence

$$y = f(x_1, \dots, x_n) = f_0 + \sum_{i=1}^n f_i \cdot x_i +$$

$$\sum_{i=1}^n \sum_{j=1}^n f_{ij} \cdot x_i \cdot x_j +$$

$$\sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n f_{ijk} \cdot x_i \cdot x_j \cdot x_k + \dots$$

looks like the first few terms of the general Taylor series. The difference between the Taylor series and a polynomial dependence is that in general, a Taylor series contains infinitely many terms, while in a polynomial dependence, there are only finitely many different terms.

In other words, to describe a general polynomial dependence, we need a finite collection of finite groups of coefficients $f_0, \{f_i\}, \{f_{ij}\}, \{f_{ijk}\}, \dots$

Here, f_0 is a single (scalar) value. The coefficients f_i naturally form a *vector*. The set of all these coefficients is also a vector in the following formal sense.

In many practical situations, it makes sense to replace the original quantities x_i with their linear combinations $x'_{i'} = \sum_{i=1}^n a_{i'i} \cdot x_i$. In this case, the original quantities can be reconstructed as $x_i = \sum_{i'=1}^n b_{ii'} \cdot x'_{i'}$, where the matrix $\{b_{ii'}\}$ is the inverse matrix to $\{a_{i'i}\}$. For example, this replacement makes perfect sense when the quantities x_i are coordinates: if we choose a different coordinate system, we get the new values $x'_{i'}$ which are linearly dependent on the original ones. If we substitute the expression for x_i in terms of $x'_{i'}$ into the above formula, we thus get an expression of the desired quantity y in terms of the new coordinates:

$$y = f'(x'_1, \dots, x'_n) = f'_0 + \sum_{i'=1}^n f'_{i'} \cdot x'_{i'} + \sum_{i'=1}^n \sum_{j'=1}^n f'_{i'j'} \cdot x'_{i'} \cdot x'_{j'} + \sum_{i'=1}^n \sum_{j'=1}^n \sum_{k'=1}^n f'_{i'j'k'} \cdot x'_{i'} \cdot x'_{j'} \cdot x'_{k'} + \dots,$$

where

$$\begin{aligned} f'_0 &= f_0; \\ f'_{i'} &= \sum_{i=1}^n f_i \cdot b_{ii'}, \\ f'_{i'j'} &= \sum_{i=1}^n \sum_{j=1}^n f_{ij} \cdot b_{ii'} \cdot b_{jj'}, \\ f'_{i'j'k'} &= \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n f_{ijk} \cdot b_{ii'} \cdot b_{jj'} \cdot b_{kk'}, \end{aligned}$$

etc. In mathematical, there are general terms for combinations that transform this way under quantity replacement; see, e.g., [1], [3], [4]:

- combinations $\{f_i\}$ that, under quantity replacement, transform as above, are called *vectors*, or, alternatively, *tensors of order 1*;
- combinations $\{f_{ij}\}$ that, under quantity replacement, transform as above, are called *tensors of order 2*;
- combinations $\{f_{ijk}\}$ that, under quantity replacement, transform as above, are called *tensors of order 3*;
- etc.

In these terms, we can say that to describe a general polynomial dependence, we need a finite collection of tensors $f_0, \{f_i\}, \{f_{ij}\}, \{f_{ijk}\}, \dots$, of different orders.

V. TOWARDS A GENERAL JUSTIFICATION OF POLYNOMIAL FORMULAS

The fact that similar polynomials appear in different application areas indicates that there is a common reason behind them. In this paper, we provide such a general justification.

We want to find a finite-parametric class F of analytical functions $f(x_1, \dots, x_n)$ approximating the actual complex aggregation. It is reasonable to require that this class F be invariant with respect to addition and multiplication by a constant, i.e., that it is a (finite-dimensional) linear space of functions.

The invariance with respect to multiplication by a constant corresponds to the fact that the aggregated quantity is usually defined only modulo the choice of a measuring unit. If we replace the original measuring unit by a one which is λ times smaller, then all the numerical values get multiplied by this factor λ :

$$f(x_1, \dots, x_n)$$

is replaced with

$$\lambda \cdot f(x_1, \dots, x_n).$$

Similarly, in all three areas, the numerical values x_i are defined modulo the choice of a measuring unit. If we replace the original measuring unit by a one which is λ times smaller, then all the numerical values get multiplied by this factor λ : x_i is replaced with $\lambda \cdot x_i$. It is therefore reasonable to also require that the finite-dimensional linear space F be invariant with respect to such re-scalings, i.e., if $f(x_1, \dots, x_n) \in F$, then for every $\lambda > 0$, the function

$$f_\lambda(x_1, \dots, x_n) \stackrel{\text{def}}{=} f(\lambda \cdot x_1, \dots, \lambda \cdot x_n)$$

also belongs to the family F .

Under this requirement, we prove that all elements of F are polynomials.

VI. MAIN RESULT

Definition 1. Let n be an arbitrary integer. We say that a finite-dimensional linear space F of analytical functions of n variables is scale-invariant if for every $f \in F$ and for every $\lambda > 0$, the function

$$f_\lambda(x_1, \dots, x_n) \stackrel{\text{def}}{=} f(\lambda \cdot x_1, \dots, \lambda \cdot x_n)$$

also belongs to the family F .

Theorem 1. For every scale-invariant finite-dimensional linear space F of analytical functions, every element $f \in F$ is a polynomial.

VII. PROOF OF THEOREM 1

Let F be a scale-invariant finite-dimensional linear space of analytical functions, and let $f(x_1, \dots, x_n)$ be a function from this family F .

By definition, an analytical function $f(x_1, \dots, x_n)$ is an infinite series consisting of monomials $m(x_1, \dots, x_n)$ of the type

$$a_{i_1 \dots i_n} \cdot x_1^{i_1} \cdot \dots \cdot x_n^{i_n}.$$

For each such term, by its *total order*, we will understand the sum $i_1 + \dots + i_n$. The meaning of this total order is simple: if we multiply each input of this monomial by λ , then the value of the monomial is multiplied by λ^k :

$$\begin{aligned} m(\lambda \cdot x_1, \dots, \lambda \cdot x_n) &= \\ a_{i_1 \dots i_n} \cdot (\lambda \cdot x_1)^{i_1} \cdot \dots \cdot (\lambda \cdot x_n)^{i_n} &= \\ \lambda^{i_1 + \dots + i_n} \cdot a_{i_1 \dots i_n} \cdot x_1^{i_1} \cdot \dots \cdot x_n^{i_n} &= \\ \lambda^k \cdot m(x_1, \dots, x_n). \end{aligned}$$

For each order k , there are finitely many possible combinations of integers i_1, \dots, i_n for which $i_1 + \dots + i_n = k$, so there are finitely many possible monomials of this order. Let $P_k(x_1, \dots, x_n)$ denote the sum of all the monomials of order k from the series describing the function $f(x_1, \dots, x_n)$. Then, we have

$$f(x_1, \dots, x_n) = P_0 + P_1(x_1, \dots, x_n) + P_2(x_1, x_2, \dots, x_n) + \dots$$

Some of these terms may be zeros – if the original expansion has no monomials of the corresponding order. Let k_0 be the first index for which the term $P_{k_0}(x_1, \dots, x_n)$ is not identically 0. Then,

$$f(x_1, \dots, x_n) = P_{k_0}(x_1, \dots, x_n) + P_{k_0+1}(x_1, x_2, \dots, x_n) + \dots$$

Since the family F is scale-invariant, it also contains the function

$$f_\lambda(x_1, \dots, x_n) = f(\lambda \cdot x_1, \dots, \lambda \cdot x_n).$$

At this re-scaling, each term P_k is multiplied by λ^k ; thus, we get

$$f_\lambda(x_1, \dots, x_n) = \lambda^{k_0} \cdot P_{k_0}(x_1, \dots, x_n) + \lambda^{k_0+1} \cdot P_{k_0+1}(x_1, x_2, \dots, x_n) + \dots$$

Since F is a linear space, it also contains a function

$$\lambda^{-k_0} \cdot f_\lambda(x_1, \dots, x_n) = P_{k_0}(x_1, \dots, x_n) + \lambda \cdot P_{k_0+1}(x_1, x_2, \dots, x_n) + \dots$$

Since F is finite-dimensional, it is closed under turning to a limit. In the limit $\lambda \rightarrow 0$, we conclude that the term $P_{k_0}(x_1, \dots, x_n)$ also belongs to the family F .

Since F is a linear space, this means that the difference

$$f(x_1, \dots, x_n) - P_{k_0}(x_1, \dots, x_n) =$$

$$P_{k_0+1}(x_1, x_2, \dots, x_n) + P_{k_0+2}(x_1, x_2, \dots, x_n) + \dots$$

also belongs to F . If we denote, by k_1 , the first index $k_1 > k_0$ for which the term $P_{k_1}(x_1, \dots, x_n)$ is not identically 0, then we can similarly conclude that this term $P_{k_1}(x_1, \dots, x_n)$ also belongs to the family F , etc.

We can therefore conclude that for every index k for which term $P_k(x_1, \dots, x_n)$ is not identically 0, this term $P_k(x_1, \dots, x_n)$ also belongs to the family F .

Monomials of different total order are linearly independent. Thus, if there were infinitely many non-zero terms P_k in the expansion of the function $f(x_1, \dots, x_n)$, we would have infinitely many linearly independent function in the family F – which contradicts to our assumption that the family F is a finite-dimensional linear space.

So, in the expansion of the function $f(x_1, \dots, x_n)$, there are only finitely many non-zero terms. Hence, the function $f(x_1, \dots, x_n)$ is a sum of finitely many monomials – i.e., a polynomial.

The statement is proven.

VIII. TOWARDS AN ALTERNATIVE JUSTIFICATION BASED ON OPTIMALITY

Alternatively, we would like to select the *optimal* finite-dimensional family of analytical functions F . To make this selection, we must figure out what “optimal” means.

In general, what is an optimality criterion? It is when we can decide, for every two families F and F' ,

- whether F is better than F' (denoted $F' \prec F$)
- or F' is better than F ($F \prec F'$)
- or F' is of the same quality as F (denoted $F \equiv F'$).

An example of an optimality criterion is a numerical criterion: $F \prec F' \Leftrightarrow J(F) < J(F')$ for some functional $J(F)$.

However, it is possible to consider optimality criteria which are more general than numerical ones. For example, if we have two different families F and F' for which $J(F) = J(F')$, e.g., for average approximation accuracy $J(F)$, then we can still choose between F and F' based on some other criteria J' (e.g., computational simplicity).

The resulting criterion is non-numerical:

$$F \prec F' \Leftrightarrow$$

$$J(F) < J(F') \vee (J(F) = J(F') \ \& \ J'(F) < J'(F')).$$

It is therefore reasonable to consider a general case when optimality is described as a general (pre)-ordering relation \preceq – i.e., a relation which is transitive but for which $F \preceq F'$ and $F' \preceq F$ does not necessarily imply that $F = F'$.

In our case, a natural requirement on this optimality criterion \preceq is that which operation is better should not depend on the choice of measuring unit:

$$F \prec F' \Leftrightarrow F_\lambda \prec F'_\lambda,$$

where $F_\lambda \stackrel{\text{def}}{=} \{f_\lambda : f \in F\}$.

Thus, we arrive at the following definitions.

IX. OPTIMIZATION APPROACH: DEFINITIONS AND THE MAIN RESULT

Definition 2. We consider the set A of all finite-dimensional spaces of analytical functions.

- By an optimality criterion, we mean a pre-ordering (i.e., a transitive, reflexive relation) \preceq on the set A .
- An optimality criterion \preceq on the class of all finite-dimensional is called scale-invariant if for all F, F' , and $\lambda \neq 0$, we have $F \preceq F'$ implies $F_\lambda \preceq F'_\lambda$.
- An optimality criterion \preceq is called final if there exists one and only one space F that is preferable to all the others, i.e., for which $F' \preceq F$ for all $F' \neq F$.

Comment. What is the motivation for the finality?

- If no space is optimal relative to some criterion, then this criterion is useless.
- If the criterion selects several spaces F as equally good, then we can also optimize something else.

If the criterion is not final, e.g., if F and F' have the same average approximation accuracy, then we can select, among them, the one which is easier to compute. Thus, such criteria can be adjusted. So, for the final criterion, the optimal space is unique.

Theorem 2. Let a family F_{opt} be optimal w.r.t. some scale-invariant and final optimality criterion. Then, all elements of the family F_{opt} are polynomials.

X. OPTIMIZATION APPROACH: PROOF

Optimality of the family F_{opt} means that $F \preceq F_{\text{opt}}$ for all $F \in A$. In particular, $F_{\lambda^{-1}} \preceq F_{\text{opt}}$ for all $F \in A$.

Due to scale-invariance of \preceq , we have $F \preceq (F_{\text{opt}})_\lambda$ for all $F \in A$. Thus, $(F_{\text{opt}})_\lambda$ is optimal.

Since there is only one optimal space, we have

$$(F_{\text{opt}})_\lambda = F_{\text{opt}}.$$

Thus, the space F_{opt} is scale-invariant. We already know (see Theorem 1) that in this case, all $f \in F_{\text{opt}}$ are polynomials. The statement is proven.

XI. WHAT IF $f(x_1, \dots, x_n)$ IS ONLY SMOOTH?

What if we do not require that a function $f(x_1, \dots, x_n)$ be analytical, only that it be smooth? In this case, we can still prove a similar result – but we will need a strong invariance condition.

Definition 3. Let n be an arbitrary integer. We say that a finite-dimensional linear space F of smooth functions of n variables is affine-invariant if for every $f \in F$ and for every linear transformation $T : R^n \rightarrow R^n$, the function

$$f_T(x) \stackrel{\text{def}}{=} f(Tx)$$

also belongs to the family F .

Theorem 3. For every affine-invariant finite-dimensional linear space F of smooth functions, every element $f \in F$ is a polynomial.

XII. PROOF OF THE RESULT ABOUT SMOOTH FUNCTIONS

Let F be an affine-invariant finite-dimensional linear space F of smooth functions. Since the space F is finite-dimensional, it has a finite basis. Let $f_1(x), \dots, f_m(x)$ be the basis of F .

Since the family F is affine invariant, for every $i \leq m$, for every variable x_j and for every $\lambda > 0$, a function

$$f_i(x_1, \dots, x_{j-1}, \lambda \cdot x_j, x_{j+1}, \dots, x_n)$$

also belongs to the family F . Since the functions f_i form a basis of the family F , this means that for some $c_{ik}(\lambda)$, we have

$$f_i(x_1, \dots, x_{j-1}, \lambda \cdot x_j, x_{j+1}, \dots, x_n) = \sum_{k=1}^m c_{ik}(\lambda) \cdot f_k(x_1, \dots, x_{j-1}, x_j, x_{j+1}, \dots, x_n).$$

Differentiating both sides of this equality by λ and taking $\lambda = 1$, we get

$$x_j \cdot \frac{\partial f_i}{\partial x_j} = \sum_{k=1}^m c_{jk} \cdot f_k.$$

For the new variables $X_j \stackrel{\text{def}}{=} \ln(x_j)$, we have a simplified equation $\frac{\partial f_i}{\partial X_j} = \sum_{k=1}^m c_{jk} \cdot f_k$. In terms of X_j , we have a system of linear ODEs with constant coefficients. A general solution to such a system is a linear combination of terms

- $\exp(\alpha \cdot X_j) = x_j^\alpha$ (with possible complex α) and
- $X_j^p \cdot \exp(\alpha \cdot X_j) = x_j^\alpha \cdot \ln^p(x_j)$.

A general linear transformation leads to different terms – except when we have x_j^α for integer $\alpha \geq 0$. Thus, every $f \in F$ is a polynomial in each variable – hence a polynomial in general. The statement is proven.

XIII. CONCLUSION

In many application areas, different reasonable requirements lead to polynomial formulas for combining degrees (of certainty, of preference, etc.). In this paper, we provide a general explanation of why polynomial formulas appear in different situations.

Because of our general explanation, in the cases when we do not yet know how to combine the degrees, it may be a good idea to start with trying polynomial formulas.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation grants HRD-0734825 and DUE-0926721, by Grant 1 T36 GM078000-01 from the National Institutes of Health, by Grant MSM 6198898701 from MŠMT of Czech Republic, and by Grant 5015 “Application of fuzzy logic with operators in the knowledge based systems” from the Science and Technology Centre in Ukraine (STCU), funded by European Union.

The authors are thankful to the anonymous referees for valuable suggestions.

REFERENCES

- [1] D. A. Danielson, *Vectors and Tensors in Engineering and Physics*, Perseus Books, Westview Press, New York, 2003.
- [2] G. J. Klir and B. Yuan, *Fuzzy sets and fuzzy logic: Theory and applications*, Prentice Hall, Upper Saddle River, New Jersey, 1995.
- [3] D. F. Lawden, *Introduction to Tensor Calculus, Relativity and Cosmology*, Dover, New York, 2003.
- [4] L. P. Lebedev and M. J. Cloud, *Tensor Analysis*, World Scientific, Singapore, 2003.
- [5] R. D. Luce and R. Raiffa, *Games and decisions: introduction and critical survey*, New York: Dover, 1989.
- [6] H. T. Nguyen, O. Kosheleva, and V. Kreinovich, "Decision Making Beyond Arrow's Impossibility Theorem", *International Journal of Intelligent Systems*, vol. 24, No. 1, pp. 27–47, 2009.
- [7] H. T. Nguyen and E. A. Walker, *A First Course in Fuzzy Logic*, Boca Raton, Florida: Chapman & Hall/CRC Press, 2006.