Supplementary material to Multi-choice solutions for feature and parameter importance

Daniel Fryer[1(✉)][0000-0001-6032-0522], David Lowing[2][0000-0002-9456-790X], Inga Strümke[3][0000-0003-1820-6544], and Hien Nguyen[1][0000-0002-9958-432X]

[1] Department of Engineering, School of Computing, Engineering and Mathematical Sciences, La Trobe University, Bundoora, Australia
{dfryer, h.nguyen5}@ltu.edu.au
[2] Industrial Engineering Research Department, Centrale Supélec, Université Paris-Saclay
david.lowing@centralesupelec.fr
[3] Department of Computer Science, The Norwegian University of Science and Technology, Norway
inga.strumke@ntnu.no

1 Axiomatic characterisations

In this section, we establish the values of MC-games on an axiomatic basis, detailing the characterisations of the DP, PZ, and LT values. Some of the axioms introduced here are straightforward adaptations of principles from TU-games to MC-games. The analytical and numerical methods serve as a validation for the axiomatic approach. In this perspective, the axioms prompt us to look for practical counterexamples, whereas the examples encourage us to refine the axioms. By including the axioms here, we hope to encourage further research into the optimal choice of value for polynomial regression framework and any similar frameworks introduced in future.

The initial axiom embodies the straightforward concept that the entire value of the game should be distributed among the players. It's important to note that the principle of Efficiency is relevant only when $\boldsymbol{m}$ is established. There are axioms akin to Efficiency that don't presume this condition. For example, the work of [1] examines the Cohesive efficiency axiom, which dictates that the aggregate of individual payoffs should match the highest collective value the players can attain by forming into groups.

1 (MC-Efficiency) *For each* $(\mathcal{M}, v) \in \mathcal{G}$*,*

$$\sum_{i \in N} \sum_{j \leq m_i} \varphi_{ij}(\mathcal{M}, v) = v(\boldsymbol{m}).$$

When dealing with a problem, it can be useful to divide it into several smaller problems. By doing so, it is desirable that the sum of the outcomes of the smaller problems is equal to the outcome of the original problem.

2 (MC-Additivity) *For each* $(\mathcal{M}, v), (\mathcal{M}, w) \in \mathcal{G}$*,*

$$\boldsymbol{\varphi}(\mathcal{M}, v + w) = \boldsymbol{\varphi}(\mathcal{M}, v) + \boldsymbol{\varphi}(\mathcal{M}, w).$$

The next axiom requires that if the maximal participation level of each player reduces to a certain level j, then the payoff of each player for their j-th participation level should remain unchanged.

3 (Independence of higher levels) *For each* $(\mathcal{M}, v) \in \mathcal{G}$*,*

$$\forall j \leq m_\top, \quad \varphi_{ij}(\mathcal{M}, v) = \varphi_{ij}((\min\{j, m_k\})_{k \in N}, v).$$

If a player's participation at a certain level produces nothing, it seems reasonable to penalise them accordingly. [7] introduce an axiom indicating that non-productive participation levels should not receive anything from the value.

4 (Non-productive level) *For each* $(\mathcal{M}, v) \in \mathcal{G}$*, if there is a* $i \in N$ *and* $j \leq m_i$ *that verifies*

$$\forall \boldsymbol{s} \in \mathcal{M}, \quad v(\boldsymbol{s}_{-i}, j-1) = v(\boldsymbol{s}_{-i}, j),$$

then $\varphi_{ij}(\mathcal{M}, v) = 0$*.*

A player's participation level is said to be inessential if that player stops being productive past that level. [4] introduce an axiom indicating that inessential participation levels should not receive anything from the value.[4] The following is a weaker axiom than Non-productive level.

5 (Inessential level) *For each* $(\mathcal{M}, v) \in \mathcal{G}$*, if there is a* $i \in N$ *and a* $j \leq m_i$ *that verifies*

$$\forall \boldsymbol{s} \in \mathcal{M}, \forall l \geq j, \quad v(\boldsymbol{s}_{-i}, l-1) = v(\boldsymbol{s}_{-i}, l),$$

then $\varphi_{ij}(\mathcal{M}, v) = 0$*.*

A necessary level represents the level of participation of a player under which the worth of any profile is null. [6] introduce an axiom stating that two necessary levels should receive the same payoffs.[5]

6 (Necessary level) *For each* $(\mathcal{M}, v) \in \mathcal{G}$*, if there are two* $i, i' \in N$ *and two* $j \leq m_i$*,* $j' \leq m_{i'}$*, verifying*

$$\forall \boldsymbol{s} \in \mathcal{M},\ s_i < j \text{ or } \boldsymbol{s}_{i'} < j', \quad v(\boldsymbol{s}) = 0,$$

then $\varphi_{ij}(\mathcal{M}, v) = \varphi_{i'j'}(\mathcal{M}, v)$*.*

If a player has the same performances at two distinct participation levels, then it seems reasonable that these two levels obtain the same payoff.

[4] This axiom and the Non-productive level level axiom are both equivalent to the Null player axiom [9] when $\boldsymbol{m} = \mathbf{1}$.

[5] This axiom is equivalent to the Necessary player axiom [2] when $\boldsymbol{m} = \mathbf{1}$.

7 (Intra symmetry) *For each* $(\mathcal{M}, v) \in \mathcal{G}$*, if there is a* $i \in N$ *and* $j, j' \leq m_i$ *verifying,* $\forall \boldsymbol{s} \in \mathcal{M}$*,*

$$v(\boldsymbol{s}_{-i}, j) - v(\boldsymbol{s}_{-i}, j-1) = v(\boldsymbol{s}_{-i}, j') - v(\boldsymbol{s}_{-i}, j'-1),$$

then $\varphi_{ij}(\mathcal{M}, v) = \varphi_{ij'}(\mathcal{M}, v)$*.*

[7] propose a straightforward extension of Anonymity from TU-games to MC-games. Denote by $\overline{\mathcal{G}}$ the sub-class of MC-games in which all player have the same number of participation levels. This is just for convenience, as fictive participation levels could be introduced to equalize the number of levels for each player.[6]

8 (MC-Anonymity) *For each* $(\mathcal{M}, v) \in \overline{\mathcal{G}}$*, each* $t \in \mathcal{M}$ *and each order* $\pi \in P(N)$*, we define* πt *as* $\pi t_{\pi(i)} = t_i$ *for each* $i \in N$*, and* πv *as* $\pi v(\pi t) = v(t)$*. Then, it holds that*

$$\varphi_{ij}(\mathcal{M}, v) = \varphi_{\pi(i)j}(m, \pi v).$$

[5] propose an axiom which guarantees that two players with the same performance at a given participation level should receive the same payoff for that level.[7]

9 (MC-Symmetry) *For each* $(\mathcal{M}, v) \in \mathcal{G}$*, if there are two* $i, i' \in N$ *and a* $j \leq m_i$*,* $j \leq m_{i'}$*, verifying,* $\forall \boldsymbol{s} \in \mathcal{M}$*,*

$$v(\boldsymbol{s}_{-i}, j) - v(\boldsymbol{s}_{-i}, j-1) = v(\boldsymbol{s}_{-i'}, j) - v(\boldsymbol{s}_{-i'}, j-1),$$

then $\varphi_{ij}(\mathcal{M}, v) = \varphi_{i'j}(\mathcal{M}, v)$*.*

Lastly, we present characterisations of the DP, PZ and LT values from the game theory literature. Since the definitions required for the proofs are tedious, we do not reproduce them here. Instead, for each theorem we reference a paper to which the reader can refer for the proof. [4] provide the following characterisation of the DP value. The proof of in [4] proceeds by first showing that DP satisfies the four stated axioms, and then showing that an arbitrary value satisfying the four axioms will coincide with DP on each game in a certain basis of the space of MC-games. The result then follows from the MC-Additivity axiom.

Theorem 1 ([4]). *A value* $\boldsymbol{\varphi}$ *on* $\mathcal{G}$ *satisfies MC-Efficiency, MC-Additivity, Necessary level and Inessential level if and only if* $\boldsymbol{\varphi} = \mathrm{DP}$*.*

[7] provide the following axiomatic characterisation of the PZ value. Similar to [4], the proof of in [7] proceeds by first showing that PZ satisfies the five axioms, and then showing that an arbitrary value satisfying the five axioms will coincide with PZ on an appropriate basis.

[6] To permute the labels of two players it is necessary that these players have the same number of participation levels. Therefore, this axiom only holds on the class of MC-games in which all the players have the same maximal activity level. Clearly, this axiom is equivalent to the Anonymity axiom [8] when $\boldsymbol{m} = \boldsymbol{1}$.

[7] This axiom is equivalent to the Symmetry axiom [9] when $\boldsymbol{m} = \boldsymbol{1}$.

Theorem 2 ([7]). *A value $\boldsymbol{\varphi}$ on $\overline{\mathcal{G}}$ satisfies MC-Efficiency, MC-Additivity, MC-Anonymity, Non-productive level and Intra symmetry if and only if $\boldsymbol{\varphi} = \mathrm{PZ}$.*

Finally, we provide the following characterisation of the LT value. The proof is direct from Theorem 1 and Corollary 1 in [5]. To understand this, see that Corollary 1 simply states that the value characterised in Theorem 1 is equal to (**??**). Theorem 1 then gives the desired characterisation, though we note that the names given to some of the axioms are different to the names we have used here.

Theorem 3. *A value $\boldsymbol{\varphi}$ on $\mathcal{G}$ satisfies MC-Efficiency, MC-Additivity, MC-Symmetry, Independence of higher levels and Inessential level if and only if $\boldsymbol{\varphi} = \mathrm{LT}$. The proof is direct from Theorem 1 and Corollary 1 in [5].*

2 Polynomial regression

The following two examples are an instructive comparison between TU-games and MC-games for polynomial regression. They are included here aid in understanding the difference between these two frameworks and the reason for the gain in computational efficiency in the MC-game framework. Example 1 covers a TU-game, and Example 2 covers an MC-game.

Example 1 (TU-game framework). Consider a two-dimensional first-degree polynomial with first-order interaction term, over a vector $\boldsymbol{x}^\top = (x_1, x_2) \in \mathbb{R}^2$, with coefficients $\boldsymbol{\beta} \in \mathbb{R}^4$,

$$p_{\boldsymbol{x}} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2. \tag{1}$$

Removing terms from (1) generates 8 sub-polynomials, which we represent as a function of binary vectors (see Example 1). Note, we do not consider the constant β_0 to be a term.
Given a sample of *iid* observations $\boldsymbol{Z}^\top = (Y, \boldsymbol{X}^\top)$, with $\boldsymbol{X}^\top = (X_1, X_2)$, and a random error term ε (having finite mean and variance), define the full multivariate polynomial regression model,

$$Y = p_{\boldsymbol{X}}(111) + \varepsilon,$$

as well as its submodels (e.g., $Y = p_{\boldsymbol{X}}(001) + \varepsilon = \beta_0 + \beta_3 X_1 X_2 + \varepsilon$). Fitting each submodel with ordinary least squares (OLS), we can compute all goodness-of-fit measurements corresponding to the 8 submodels (i.e., $R^2_{\boldsymbol{Z}}(111), R^2_{\boldsymbol{Z}}(101)$, etc., where $R^2_{\boldsymbol{Z}}(\boldsymbol{s})$ denotes the coefficient of multiple correlation for the submodel with sub-polynomial $p_{\boldsymbol{X}}(\boldsymbol{s})$).

To directly apply the Shapley framework of [3], the interaction term in (1) is treated as a third variable $X_3 = X_1 X_2$. The decomposition results in goodness-of-fit allocations $\mathrm{Sh}_1(\boldsymbol{Z}), \mathrm{Sh}_2(\boldsymbol{Z}), \mathrm{Sh}_3(\boldsymbol{Z})$ for X_1, X_2 and X_1X_2 respectively, where

$$R^2_{\boldsymbol{Z}}(111) = \mathrm{Sh}_1(\boldsymbol{Z}) + \mathrm{Sh}_2(\boldsymbol{Z}) + \mathrm{Sh}_3(\boldsymbol{Z}). \tag{2}$$

The type of allocation (2) may be desirable. Indeed, it includes a summary of the goodness-of-fit contribution of X_1X_2 to all submodels in Figure 1.

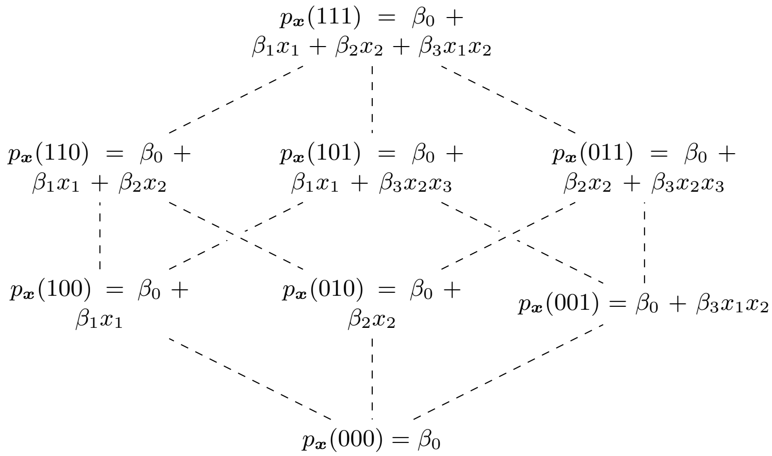

Fig. 1: Lattice diagram of the 8 sub-polynomials of (1), ordered by inclusion.

Example 2 (MC-game framework). As an alternative to the approach in Example 1, suppose we wish to treat only the variables X_1 and X_2 as features – regarding the $X_1 X_2$ term as a consequence of including both features[8]. This reduces the 8 node lattice diagram in Figure 1 to the 4 node diagram in Figure 2.

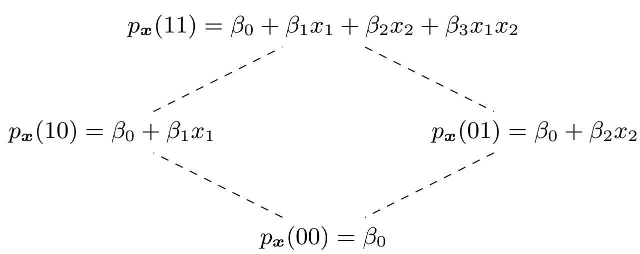

Fig. 2: Lattice diagram of the reduced set of sub-polynomials of (1), where the interaction $x_1 x_2$ is viewed as a consequence of including both x_1 and x_2.

In general, we require a framework that can account for higher degree polynomial terms and higher order interactions. Consider the two-dimensional second-degree

8

Note 1. The Owen value framework (see Section 2.3 of [3]) allows for disjoint grouping of regressor variables, but no partition is useful here, since $X_1 X_2$ would need to be included in either the group with X_1, or the group with X_2 (but not both).

polynomial with first-order interaction term,

$$\mathcal{B}_{\boldsymbol{x}} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \beta_4 x_1^2 + \beta_5 x_2^2. \tag{3}$$

In the TU-game framework of Example 1, there are $2^5 = 32$ sub-polynomials of (3). The MC-game framework, introduced formally in **??**, reduces this to $3 \times 3 = 9$ sub-polynomials, represented as a function $\mathcal{B}_{\boldsymbol{x}}(\boldsymbol{s})$ of $\boldsymbol{s} \in \{0,1,2\}^2$, with

$$\begin{aligned}
\mathcal{B}_{\boldsymbol{x}}(0,0) &= \beta_0 \\
\mathcal{B}_{\boldsymbol{x}}(0,1) &= \beta_0 + \beta_2 x_2 \\
\mathcal{B}_{\boldsymbol{x}}(0,2) &= \beta_0 + \beta_2 x_2 + \beta_5 x_2^2 \\
\mathcal{B}_{\boldsymbol{x}}(1,1) &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 \\
&\vdots \\
\mathcal{B}_{\boldsymbol{x}}(2,2) &= \mathcal{B}_{\boldsymbol{x}}.
\end{aligned}$$

The reduction in the number of sub-polynomials of (3) (from 32 to 9), occurs because the sub-polynomials are now a function of the participation levels of the two features x_1, x_2 (rather than the inclusion of individual terms seen in Example 1). The key difference is that goodness-of-fit will be attributed to levels of feature participation, rather than to individual terms.

References

1. Béal, S., Casajus, A., Rémila, E., Solal, P.: Cohesive efficiency in tu-games: axiomatizations of variants of the shapley value, egalitarian values and their convex combinations. Annals of Operations Research **302**(1), 23–47 (2021)
2. Béal, S., Navarro, F.: Necessary versus equal players in axiomatic studies. Operations Research Letters **48**(3), 385–391 (2020)
3. Huettner, F., Sunder, M.: Axiomatic arguments for decomposing goodness of fit according to Shapley and Owen values. Electronic Journal of Statistics **6**(none), 1239 – 1250 (2012). `https://doi.org/10.1214/12-EJS710`, `https://doi.org/10.1214/12-EJS710`
4. Klijn, F., Slikker, M., Zarzuelo, J.: Characterizations of a multi-choice value. International Journal of Game Theory **28**(4), 521–532 (1999)
5. Lowing, D., Techer, K.: Marginalism, egalitarianism and efficiency in multi-choice games. Social Choice and Welfare pp. 1–47 (2022)
6. van den Nouweland, C.G.A.: Games and graphs in economic situations. tilburg university (1993)
7. Peters, H., Zank, H.: The egalitarian solution for multichoice games. Annals of Operations Research **137**(1), 399–409 (2005)
8. Shapley, L.S.: A value for n-person games. Contributions to the Theory of Games **2**(28), 307–317 (1953)
9. Shubik, M.: Incentives, decentralized control, the assignment of joint costs and internal pricing. Management Science **8**(3), 325–343 (1962)