MACHINE LEARNING THE DIMENSION OF A POLYTOPE

TOM COATES, JOHANNES HOFSCHEIER, AND ALEXANDER M. KASPRZYK

Abstract. We use machine learning to predict the dimension of a lattice polytope directly from its Ehrhart series. This is highly effective, achieving almost 100% accuracy. We also use machine learning to recover the volume of a lattice polytope from its Ehrhart series, and to recover the dimension, volume, and quasi-period of a rational polytope from its Ehrhart series. In each case we achieve very high accuracy, and we propose mathematical explanations for why this should be so.

1. Introduction

Let $P \subset \mathbb{Z}^d \otimes_{\mathbb{Z}} \mathbb{Q}$ be a convex lattice polytope of dimension d, that is, let P be the convex hull of finitely many points in $\mathbb{Z}^d$ whose $\mathbb{Q}$-affine combinations generate $\mathbb{Q}^d$. A fundamental invariant of P is the number of lattice points that it contains, $|P \cap \mathbb{Z}^d|$. More generally, let $L_P(m) \coloneqq |mP \cap \mathbb{Z}^d|$ count the lattice points in the mth dilation mP of P, where $m \in \mathbb{Z}_{\geq 0}$. Then L_P is given by a polynomial of degree d called the *Ehrhart polynomial* [15]. The corresponding generating series, called the *Ehrhart series* and denoted by Ehr_P, can be expressed as a rational function with numerator a polynomial of degree at most d [29]:

$$\mathrm{Ehr}_P(t) \coloneqq \sum_{m \geq 0} L_P(m) t^m = \frac{\delta_0 + \delta_1 t + \cdots + \delta_d t^d}{(1-t)^{d+1}}, \qquad \delta_i \in \mathbb{Z}.$$

The coefficients $(\delta_0, \delta_1, \ldots, \delta_d)$ of this numerator, called the *δ-vector* or *h^*-vector* of P, have combinatorial meaning [14]:

(i) $\delta_0 = 1$;
(ii) $\delta_1 = |P \cap \mathbb{Z}^d| - d - 1$;
(iii) $\delta_d = |P^\circ \cap \mathbb{Z}^d|$, where $P^\circ = P \setminus \partial P$ is the strict interior of P;
(iv) $\delta_0 + \cdots + \delta_d = \mathrm{Vol}(P)$, where $\mathrm{Vol}(P) = d!\,\mathrm{vol}(P)$ is the lattice-normalised volume of P.

The polynomial L_P can be expressed in terms of the δ-vector:

$$L_P(m) = \sum_{i=0}^{d} \delta_i \binom{d+m-i}{d}.$$

From this we can see that the leading coefficient of L_P is $\mathrm{vol}(P)$, the Euclidean volume of P.

Given $d+1$ terms of the Ehrhart series, one can recover the δ-vector and hence the invariants $|P \cap \mathbb{Z}^d|$, $|P^\circ \cap \mathbb{Z}^d|$, and $\mathrm{Vol}(P)$. This, however, assumes knowledge of the dimension d.

Question 1. *Given a lattice polytope P, can machine learning recover the dimension d of P from sufficiently many terms of the Ehrhart series* Ehr_P*?*

There has been recent success using machine learning (ML) to predict invariants such as $\mathrm{Vol}(P)$ directly from the vertices of P [4], and to predict numerical invariants from a geometric analogue of the Ehrhart series called the Hilbert series [5]. As we will see in §§2.1–2.3, ML is also extremely effective at answering Question 1. In §2.4 we propose a possible explanation for this.

2020 *Mathematics Subject Classification.* 52B20, 68T05 (Primary); 14M25 (Secondary).
Key words and phrases. Machine learning, convex polytope, Ehrhart series, quasi-period.

1.1. **The quasi-period of a rational convex polytope.** Now let P be a convex polytope with *rational* vertices, and let $k \in \mathbb{Z}_{>0}$ be the smallest positive dilation of P such that kP is a lattice polytope. One can define the Ehrhart series of P exactly as before:

$$L_P(m) := |mP \cap \mathbb{Z}^d|.$$

In general L_P will no longer be a polynomial, but it is a *quasi-polynomial* of degree d and period k [15,27]. That is, there exist polynomials $f_0, \ldots, f_{k-1}$, each of degree d, such that:

$$L_P(qk + r) = f_r(q), \qquad \text{whenever } 0 \leq r < k.$$

The leading coefficient of each f_r is $\mathrm{vol}(kP)$, the Euclidean volume of the kth dilation of P.

It is sometimes possible to express L_P using a smaller number of polynomials. Let ρ be the minimum number of polynomials needed to express L_P as a quasi-polynomial; ρ is called the *quasi-period* of P, and is a divisor of k. Each of these ρ polynomials is of degree d with leading coefficient $\mathrm{vol}(\rho P)$. When $\rho < k$ we say that P exhibits *quasi-period collapse* [7,8,16,23,26]. The Ehrhart series of P can be expressed as a rational function:

$$\mathrm{Ehr}_P(t) = \frac{\delta_0 + \delta_1 t + \cdots + \delta_{\rho(d+1)-1} t^{\rho(d+1)-1}}{(1-t^\rho)^{d+1}}, \qquad \delta_i \in \mathbb{Z}.$$

As in the case of lattice polytopes, the δ-vector carries combinatorial information about P [6, 7,20,21,25]. Given $\rho(d+1)$ terms of the Ehrhart series, one can recover the δ-vector. This, however, requires knowing both the dimension d and the quasi-period ρ of P.

Question 2. *Given a rational polytope P, can machine learning recover the dimension d and quasi-period ρ of P from sufficiently many terms of the Ehrhart series* Ehr_P*?*

ML again performs very well here – see §3. We propose a mathematical explanation for this, in terms of forward differences and affine hyperplanes, in §3.4.

1.2. **Some motivating examples.**

Example 1.1. *The 2-dimensional lattice polytope $P := \mathrm{conv}\{(-1,-1),(-1,2),(2,-1)\}$ has volume $\mathrm{Vol}(P) = 9$, $|P \cap \mathbb{Z}^2| = 10$, and $|P^\circ \cap \mathbb{Z}^2| = 1$. The Ehrhart polynomial of P is:*

$$L_P(m) = \frac{9}{2}m^2 + \frac{9}{2}m + 1,$$

and the Ehrhart series of P is generated by:

$$\mathrm{Ehr}_P(t) = \frac{1 + 7t + t^2}{(1-t)^3} = 1 + 10t + 28t^2 + 55t^3 + 91t^4 + \cdots.$$

Example 1.2. *The smallest dilation of the triangle $P := \mathrm{conv}\{(5,-1),(-1,-1),(-1,1/2)\}$ giving a lattice triangle is $2P$. Hence L_P can be written as a quasi-polynomial of degree 2 and period 2:*

$$f_0(q) = 18q^2 + 9q + 1, \qquad f_1(q) = 18q^2 + 27q + 10.$$

This is not, however, the minimum possible period. As in Example 1.1,

$$L_P(m) = \frac{9}{2}m^2 + \frac{9}{2}m + 1$$

and thus P has quasi-period $\rho = 1$.

This striking example of quasi-period collapse is developed further in Example 4.1.

1.3. **Code and data availability.** The datasets used in this work were generated with V2.25-4 of Magma [9]. We performed our ML analysis with scikit-learn [28], a standard machine learning library for Python, using scikit-learn v0.24.1 and Python v3.8.8. All data, along with the code used to generate it and to perform the subsequent analysis, is available from Zenodo [10,11] under a permissive open source license (MIT for the code, CC0 for the data).

2. Question 1: Dimension

In this section we investigate whether machine learning can predict the dimension d of a lattice polytope P from sufficiently many terms of the Ehrhart series Ehr_P. We calculate terms $L_P(m)$ of the Ehrhart series, for $0 \leq m \leq 1100$, and encode these in a *logarithmic Ehrhart vector*:

$$(\log y_0, \log y_1, \ldots, \log y_{1100}), \qquad \text{where } y_m \coloneqq L_P(m).$$

We find that standard ML techniques are extremely effective, recovering the dimension of P with almost 100% accuracy from the logarithmic Ehrhart vector. We then ask whether ML can recover $\mathrm{Vol}(P)$ from this Ehrhart data; again, this is achieved with near 100% success.

2.1. **Data generation.** A dataset [11] containing 2918 distinct entries, with $2 \leq d \leq 8$, was generated using Algorithm 2.1 below. The distribution of this data is summarised in Table 1.

Algorithm 2.1.

Input: A positive integer d.
Output: A vector

$$\big(\log y_0, \log y_1, \ldots, \log y_{1100}, d, \mathrm{Vol}(P)\big)$$

for a d-dimensional lattice polytope P, where $y_m \coloneqq L_P(m)$.

(i) Choose $d + k$ lattice points $\{v_1, \ldots, v_{d+k}\}$ uniformly at random in a box $[-5, 5]^d$, where k is chosen uniformly at random in $\{1, \ldots, 5\}$.
(ii) Set $P \coloneqq \mathrm{conv}\{v_1, \ldots, v_{d+k}\}$. If $\dim(P) \neq d$ return to step (i).
(iii) Calculate the coefficients $y_m \coloneqq L_P(m)$ of the Ehrhart series of P, for $0 \leq m \leq 1100$.
(iv) Return the vector $\big(\log y_0, \log y_1, \ldots, \log y_{1100}, d, \mathrm{Vol}(P)\big)$.

We deduplicated on the vector $\big(\log y_0, \log y_1, \ldots, \log y_{1100}, d, \mathrm{Vol}(P)\big)$ to get a dataset with distinct entries. In particular, two polytopes that are equivalent under the group of affine-linear transformations $\mathrm{GL}_d(\mathbb{Z}) \ltimes \mathbb{Z}^d$ give rise to the same point in the dataset.

2.2. **Machine learning the dimension.** We reduced the dimensionality of the dataset by projecting onto the first two principal components of the logarithmic Ehrhart vector. As one would expect from Figure 1, a linear support vector machine (SVM) classifier trained on these features predicted the dimension of P with 100% accuracy. Here we used a scikit-learn pipeline consisting of a `StandardScaler` followed by an `SVC` classifier with linear kernel and regularization hyperparameter $C = 0.1$, using 50% of the data for training the classifier and tuning the hyperparameter, and holding out the remaining 50% of the data for model validation.

In §2.4 below we give a mathematical explanation for the structure observed in Figure 1, and hence for why ML is so effective at predicting the dimension of P. Note that the discussion in §3.4 suggests that one should also be able to extract the dimension using ML on the *Ehrhart vector*

$$(y_0, y_1, \ldots, y_{1100})$$

rather than the logarithmic Ehrhart vector

$$(\log y_0, \log y_1, \ldots, \log y_{1100}).$$

Dimension	2	3	4	5	6	7	8
Total	431	787	812	399	181	195	113

Table 1. The distribution of the dimensions appearing in the dataset for Question 1.

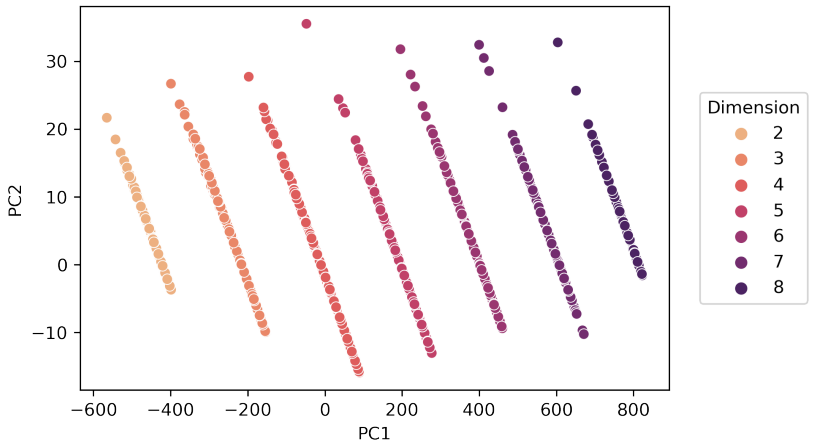

Figure 1. Projection onto the first two principal components of the logarithmic Ehrhart vector coloured by the dimension of P.

This is indeed the case, although here it is important not to reduce the dimensionality of the data too much[1]. A linear SVM classifier trained on the full Ehrhart vector predicts the dimension of P with 98.7% accuracy, with a pipeline exactly as above except that $C = 50000$. A linear SVM classifier trained on the first 30 principal components of the Ehrhart vector (the same pipeline, but with $C = 20$) gives 93.7% accuracy, but projecting to the first two components reduces accuracy to 43.3%.

2.3. **Machine learning the volume.** To learn the normalised volume of a lattice polytope from its logarithmic Ehrhart vector, we used a scikit-learn pipeline consisting of a `StandardScaler` followed by an `SVR` regressor with linear kernel and regularization hyperparameter $C = 340$. We restricted attention to the roughly 75% of the data with volume less than 10 000, thereby removing outliers. We used 50% of that data for training and hyperparameter tuning, selecting the training set using a shuffle stratified by volume; this corrects for the fact that the dataset contains a high proportion of polytopes with small volume. The regression had a coefficient of determination (R^2) of 0.432, and gave a strong hint (see Figure 2) that we should repeat the analysis replacing the logarithmic Ehrhart vector with the Ehrhart vector.

Using the same pipeline (but with $C = 1000$) and the Ehrhart vector gives a regression with $R^2 = 1.000$; see Figure 2. This regressor performs well over the full dataset, with volumes ranging up to approximately 4.5 million: over the full dataset we still find $R^2 = 1.000$. The fact that Support Vector Machine methods are so successful in recovering the volume of P from the Ehrhart vector is consistent with the discussion in §3.4.

2.4. **Crude asymptotics for $\log y_k$.** Since

$$L_P(m) = \mathrm{vol}(P)m^d + \text{lower order terms in } m$$

[1]This is consistent with the discussion in §3.4, which suggests that we should try to detect whether the Ehrhart vector lies in a union of linear subspaces that have fairly high codimension.

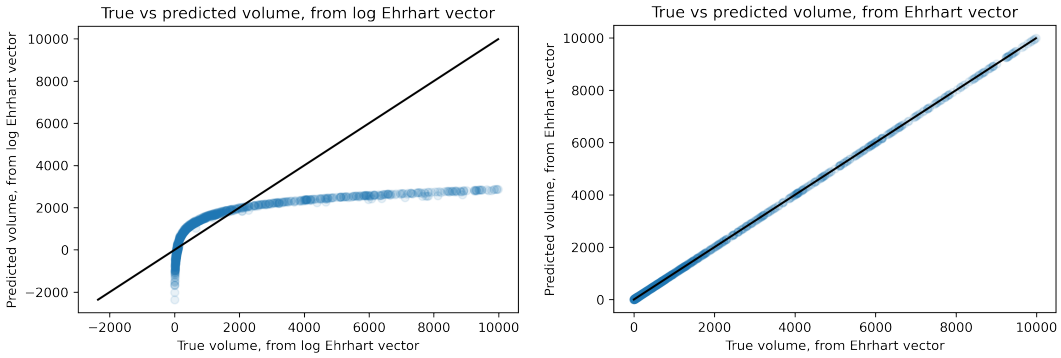

Figure 2. Plotting the true versus the predicted normalised volume.

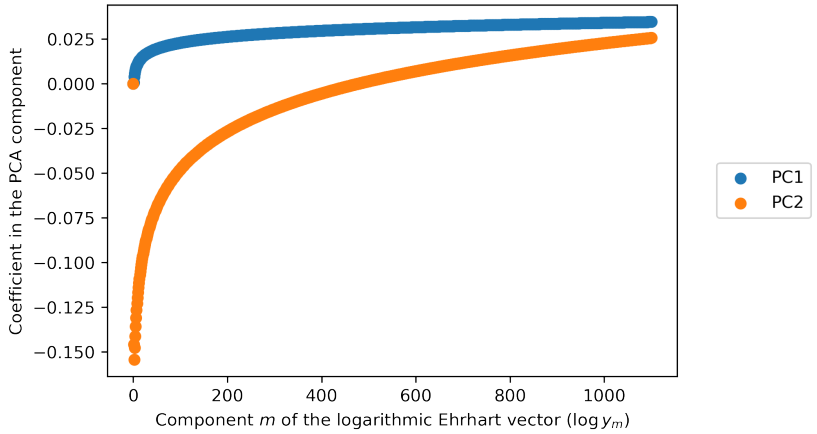

Figure 3. Contribution of $\log y_m$ to the first two principal components of the logarithmic Ehrhart vector, as m varies.

we have that

$$\log y_m \sim d \log m + \log \mathrm{vol}(P).$$

For $m \gg 0$, therefore, we see that the different components $\log y_m$ of the logarithmic Ehrhart vector depend approximately affine-linearly on each other as P varies. It seems intuitively plausible that the first two PCA components of the logarithmic Ehrhart vector should depend non-trivially on $\log y_m$ for $m \gg 0$, and in fact this is the case – see Figure 3. Thus the first two PCA components of the logarithmic Ehrhart vector should vary approximately affine-linearly as P varies, with constant slope and with a translation that depends only on the dimension of P.

3. Question 2: Quasi-period

Here we investigate whether ML can predict the quasi-period ρ of a d-dimensional rational polytope P from sufficiently many terms of the Ehrhart vector of P. Once again, standard ML techniques based on Support Vector Machines are highly effective, achieving classification accuracies of up to 95.3%. We propose a potential mathematical explanation for this in §3.4.

3.1. Data generation.

In this section we consider a dataset [10] containing 84 000 distinct entries. This was obtained by using Algorithm 3.1 below to generate a larger dataset, followed by random downsampling to a subset with 2000 datapoints for each pair (d, ρ) with $d \in \{2, 3, 4\}$ and $\rho \in \{2, 3, \ldots, 15\}$.

Algorithm 3.1.

Input: A positive integer d.
Output: A vector

$$\big(\log y_0, \log y_1, \ldots, \log y_{1100}, d, \rho\big)$$

for a d-dimensional rational polytope P with quasi-period ρ, where $y_m := L_P(m)$.

(i) Choose $r \in \{2, 3, \ldots, 15\}$ uniformly at random.
(ii) Choose $d + k$ lattice points $\{v_1, \ldots, v_{d+k}\}$ uniformly at random in a box $[-5r, 5r]^d$, where k is chosen uniformly at random in $\{1, \ldots, 5\}$.
(iii) Set $P := \mathrm{conv}\{v_1, \ldots, v_{d+k}\}$. If $\dim(P) \neq d$ return to step (ii).
(iv) Choose a lattice point $v \in P \cap \mathbb{Z}^d$ uniformly at random and replace P with the translation $P - v$. (We perform this step to ensure that the resulting rational polytope always contains a lattice point; this avoids complications when taking log in step (viii).)
(v) Replace P with the dilation P/r.
(vi) Calculate the coefficients $y_m := L_P(m)$ of the Ehrhart series of P, for $0 \leq m \leq 1100$.
(vii) Calculate the quasi-period ρ.
(viii) Return the vector $\big(\log y_0, \log y_1, \ldots, \log y_{1100}, d, \rho\big)$.

As before, we deduplicated the dataset on the vector $\big(\log y_0, \log y_1, \ldots, \log y_{1100}, d, \rho\big)$.

3.2. **Recovering the dimension and volume.** Figure 4 shows the first two principal components of the logarithmic Ehrhart vector. As in §2.2, this falls into widely-separated linear clusters according to the value of $\dim(P)$, and so the dimension of the rational polytope P can be recovered with high accuracy from its logarithmic Ehrhart vector. Furthermore, as in §2.3, applying a linear `SVR` regressor (with $C = 1000$) to the Ehrhart vector predicts the volume of a rational polytope P with high accuracy ($R^2 = 1.000$).

3.3. **Machine learning the quasi-period.** To learn the quasi-period of a rational polytope from its Ehrhart vector, we used a scikit-learn pipeline consisting of a `StandardScaler` followed by a `LinearSVC` classifier. We fixed a dimension $d \in \{2, 3, 4\}$, moved to PCA coordinates, and used 50% of the data ($N = 14000$) for training the classifier and hyperparameter tuning, holding out the remaining 50% of the data for model validation. Results are summarised on the left-hand side of Table 2, with learning curves in the left-hand column of

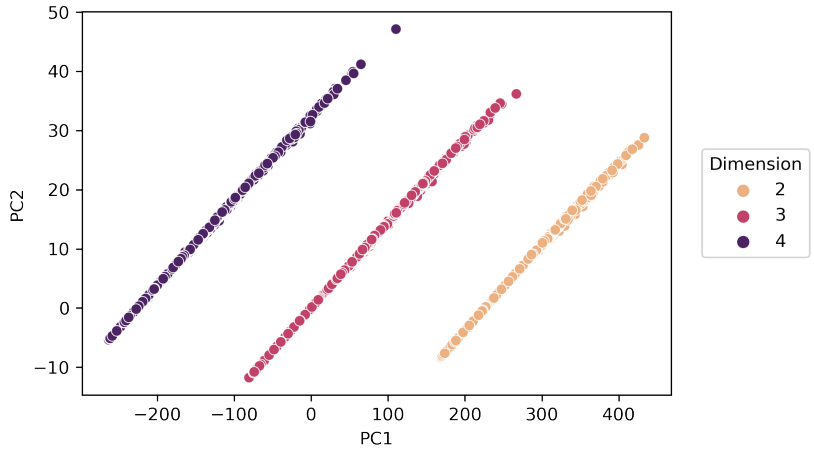

Figure 4. Projection onto the first two principal components of the logarithmic Ehrhart vector data coloured by the dimension of P.

Ehrhart vector			log Ehrhart vector		
Dimension	C	Accuracy	Dimension	C	Accuracy
2	0.01	80.6%	2	1	79.7%
3	0.001	94.0%	3	1	85.2%
4	0.001	95.3%	4	1	83.2%

Table 2. The regularization hyperparameter C and accuracy for a `LinearSVC` classifier predicting the quasi-period from the Ehrhart vector and logarithmic Ehrhart vector of rational polytopes P.

Figure 5 and confusion matrices in the left-hand column of Figure 6. The confusion matrices hint at some structure in the misclassified data.

One could also use the same pipeline but applied to the logarithmic Ehrhart vector rather than the Ehrhart vector. Results are summarised on the right-hand side of Table 2, with learning curves in the right-hand column of Figure 5 and confusion matrices in the right-hand column of Figure 6. Using the logarithmic Ehrhart data resulted in a less accurate classifier, but the learning curves suggest that this might be improved by adding more training data. Again there are hints of structure in the misclassified data.

3.4. **Forward differences.** Let y denote the sequence $(y_m)_{m=0}^{\infty}$. Recall the forward difference operator Δ defined on the space of sequences:

$$\Delta y = (y_{m+1} - y_m)_{m=0}^{\infty}.$$

A sequence y depends polynomially on m, that is

$$y_m = a_0 + a_1 m + \cdots + a_d m^d \qquad \text{for some } a_0, \ldots, a_d \in \mathbb{R},$$

if and only if y lies in the kernel of Δ^{d+1}. Furthermore, in this case, $\Delta^d y$ is the constant sequence with value $d! a_d$.

Thus a sequence $(y_m)_{m=0}^{\infty}$ is quasi-polynomial of degree d and period k, in the sense of §1.1, if and only if it lies in the kernel of Δ_k^{d+1} where Δ_k is the k-step forward difference operator,

$$\Delta_k \left((y_m)_{m=0}^{\infty} \right) = (y_{m+k} - y_m)_{m=0}^{\infty},$$

and $(y_m)_{m=0}^{\infty}$ does not lie in the kernel of Δ_k^d. Furthermore, in this case, we can determine the leading coefficients of the polynomials $f_0, \ldots, f_{k-1}$ by examining the values of the k-periodic sequence $\Delta_k^d y$. When y arises as the Ehrhart series of a rational polytope P, all of these constant terms equal the volume of kP, and so the value of the constant sequence $\Delta_k^d y$ determines the normalised volume $\mathrm{Vol}(P)$.

This discussion suggests that an SVM classifier with linear kernel should be able to learn the quasi-period and volume with high accuracy from the Ehrhart vector of a rational polytope, at least if we consider only polytopes of a fixed dimension d. Having quasi-period k amounts to the Ehrhart vector lying in (a relatively open subset of) a certain subspace $\ker \Delta_k^{d+1}$; these subspaces, being linear objects, should be easily separable using hyperplanes. Similarly, having fixed normalised volume amounts to lying in a given affine subspace; such affine subspaces should be easily separable using affine hyperplanes.

From this point of view it is interesting that an SVM classifier with linear kernel also learns the quasi-period with reasonably high accuracy from the logarithmic Ehrhart vector. Passing from the Ehrhart vector to the logarithmic Ehrhart vector replaces the linear subspaces $\ker \Delta_k^{d+1}$ by non-linear submanifolds. But our experiments above suggest that

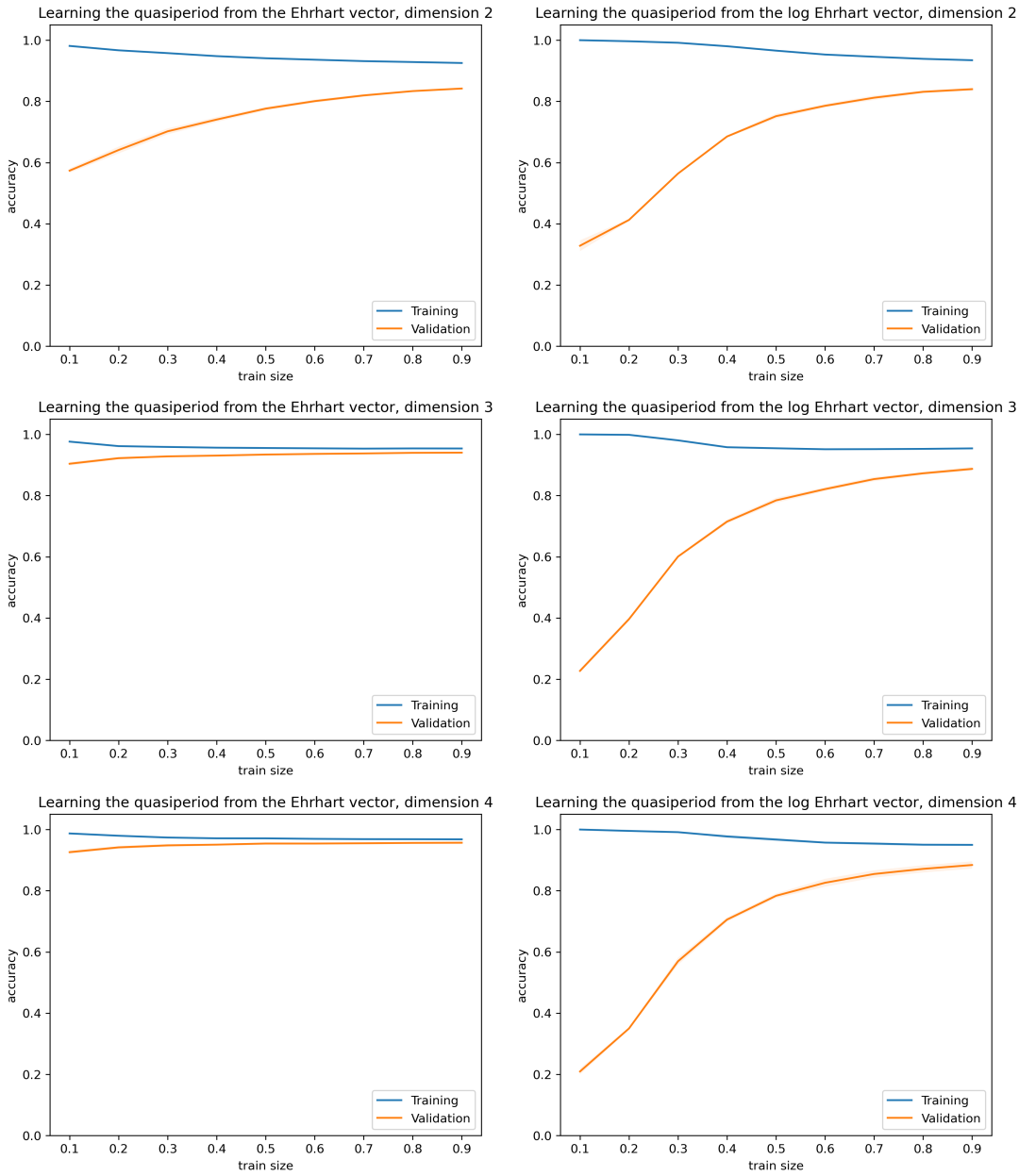

Figure 5. Quasi-period learning curves for a linear SVM classifier, for both the Ehrhart vector (left-hand column) and the logarithmic Ehrhart vector (right-hand column).

these non-linear submanifolds must nonetheless be close to being separable by appropriate collections of affine hyperplanes.

4. A remark on the Gorenstein index

In this section we discuss a geometric question where machine learning techniques failed. This involves a more subtle combinatorial invariant called the Gorenstein index, and was the question that motivated the rest of the work in this paper. We then suggest why in retrospect we should not have expected to be able to answer this question using machine learning (or at all).

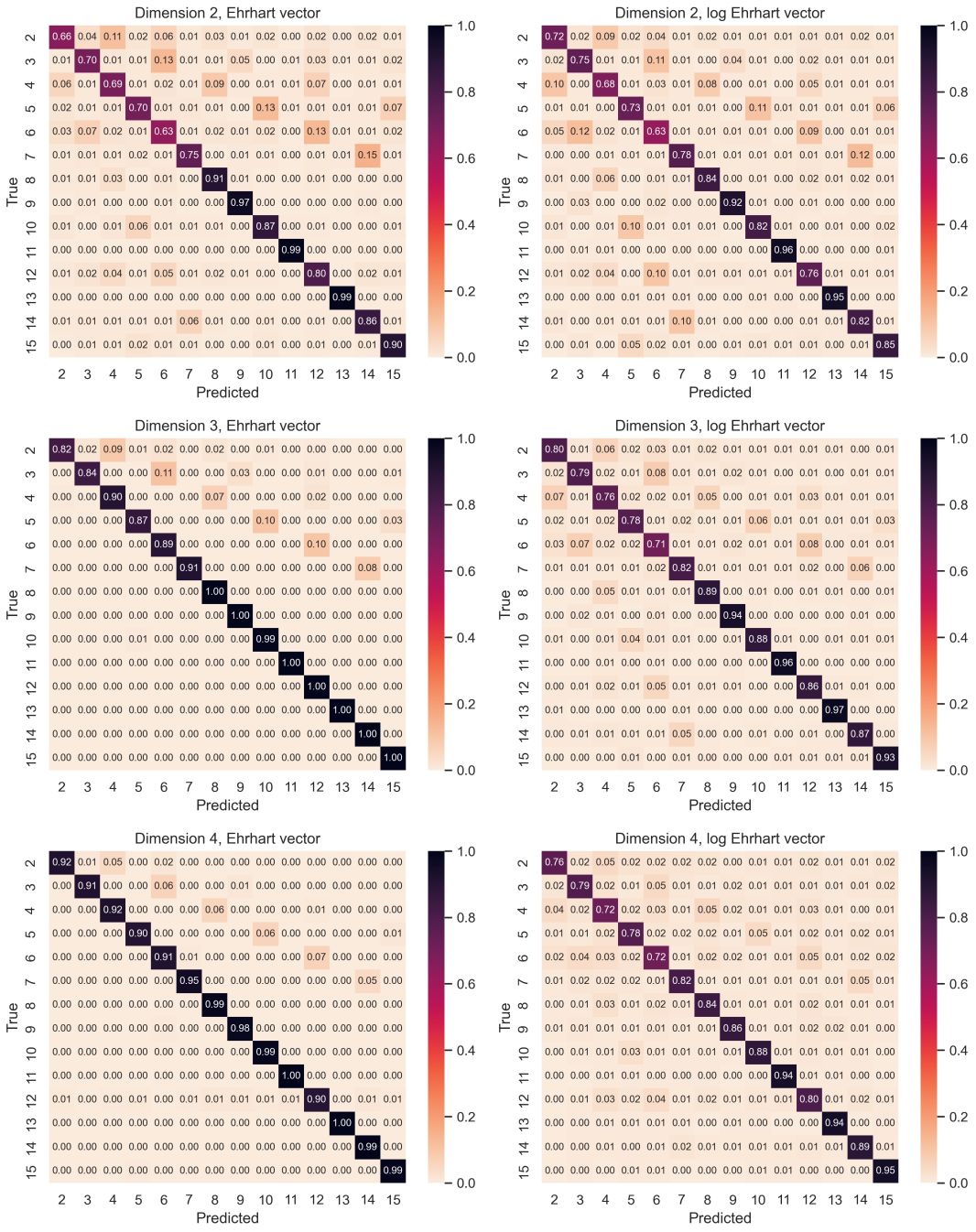

Figure 6. Row-normalised confusion matrices for a linear SVM classifier for the quasi-period, for both the Ehrhart vector (left-hand column) and the logarithmic Ehrhart vector (right-hand column).

4.1. **The Gorenstein index.** Fix a rank d lattice $N \cong \mathbb{Z}^d$, write $N_{\mathbb{Q}} := N \otimes_{\mathbb{Z}} \mathbb{Q}$, and let $P \subset N_{\mathbb{Q}}$ be a lattice polytope. The *polar polyhedron* of P is given by

$$P^* := \{u \in M \otimes_{\mathbb{Z}} \mathbb{Q} \mid u(v) \geq -1 \text{ for all } v \in P\}$$

where $M := \mathrm{Hom}(N, \mathbb{Z}) \cong \mathbb{Z}^d$ is the lattice dual to N. The polar polyhedron P^* is a convex polytope with rational vertices if and only if the origin lies in the strict interior of P. In this case, $(P^*)^* = P$. The smallest positive integer k_P such that $k_P P^*$ is a lattice polytope is called the *Gorenstein index* of P.

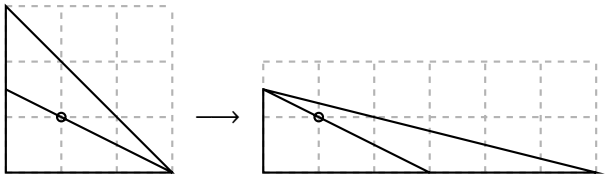

Figure 7. The mutation from $P_{(1,1,1)}$ to $P_{(1,1,2)}$ gives a scissors congruence between the polar polytopes.

The Gorenstein index arises naturally in the context of *Fano toric varieties*, and we will restrict our discussion to this setting. Let $P \subset N_{\mathbb{Q}}$ be a lattice polytope such that the vertices of P are primitive lattice vectors and that the origin lies in the strict interior of P; such polytopes are called *Fano* [22]. The spanning fan Σ_P of P – that is, the complete fan whose cones are generated by the faces of P – gives rise to a Fano toric variety X_P [12, 17]. This construction gives a one-to-one correspondence between $\mathrm{GL}_d(\mathbb{Z})$-equivalence classes of Fano polytopes and isomorphism classes of Fano toric varieties. Let $P \subset N_{\mathbb{Q}}$ be a Fano polytope that corresponds to a Fano toric variety $X := X_P$. The polar polytope $P^* \subset M_{\mathbb{Q}}$ then corresponds to a divisor on X called the *anticanonical divisor*, which is denoted by $-K_X$. In general $-K_X$ is an ample $\mathbb{Q}$-Cartier divisor, and X is $\mathbb{Q}$-Gorenstein. The Gorenstein index k_P of P is equal to the smallest positive multiple k of the anticanonical divisor such that $-kK_X$ is Cartier.

Under the correspondence just discussed, the Ehrhart series Ehr_{P^*} of the polar polytope P^* coincides with the Hilbert series $\mathrm{Hilb}_X(-K_X)$. The Hilbert series is an important numerical invariant of X, and it makes sense to ask whether the Gorenstein index of X is determined by the Hilbert series. Put differently:

Question 3. *Given a Fano polytope P, can ML recover the Gorenstein index k_P of P from sufficiently many terms of the Ehrhart series Ehr_{P^*} of the polar polytope P^*?*

There are good reasons, as we discuss below, to expect the answer to Question 3 to be 'no'. But part of the power of ML in mathematics is that it can detect or suggest structure that was not known or expected previously (see e.g. [13]). That did not happen on this occasion: applying the techniques discussed in §2 and §3 did not allow us to predict the Gorenstein index of P from the Ehrhart series of the polar polytope Ehr_{P^*}.

4.2. **Should we have expected this?** The Hilbert series is preserved under an important class of deformations called *qG-deformations* [24]. But the process of *mutation* [2] can transform a Fano polytope P to a Fano polytope Q with $\mathrm{Ehr}_{P^*} = \mathrm{Ehr}_{Q^*}$. Mutation gives rise to a qG-deformation from X_P to X_Q [1], but need not preserve the Gorenstein index: k_P need not be equal to k_Q. Thus the Gorenstein index is not invariant under qG-deformation. It might have been unrealistic to expect that a qG-deformation invariant quantity (the Hilbert series) could determine an invariant (the Gorenstein index) which can vary under qG-deformation.

Quasiperiod collapse. Although the phenomenon of quasi-period collapse remains largely mysterious from a combinatorial view-point, in the context of toric geometry one possible explanation arises from mutation and qG-deformation [23]. The following example revisits Examples 1.1 and 1.2 from this point of view, and illustrates why Question 3 cannot have a meaningful positive answer.

Example 4.1. *Let $P_{(a,b,c)} \subset N_{\mathbb{Q}}$ denote the 2-dimensional Fano polytope associated with weighted projective space $\mathbb{P}(a^2, b^2, c^2)$, where a, b, c are pairwise coprime positive integers. Then $P_{(1,1,1)} :=$*

$\mathrm{conv}\{(1,0),(0,1),(-1,-1)\}$ *is the Fano polygon associated with* $\mathbb{P}^2$*, with polar polygon* $P^*_{(1,1,1)} \subset M_{\mathbb{Q}}$ *the lattice triangle appearing in Example 1.1.*

The graph of mutations of $P_{(1,1,1)}$ *has been completely described [3,19]. Up to* $\mathrm{GL}_2(\mathbb{Z})$*-equivalence, there is exactly one mutation from* $P_{(1,1,1)}$*: this gives* $P_{(1,1,2)}$ *corresponding to* $\mathbb{P}(1,1,4)$*. The polar polygon* $P^*_{(1,1,2)}$ *is the rational triangle in Example 1.2. As Figure 7 illustrates, a mutation between polytopes gives rise to a scissors congruence [18] between polar polytopes. Mutation therefore preserves the Ehrhart series of the polar polytope, and this explains why we have quasi-period collapse in this example.*

We can mutate $P_{(1,1,2)}$ *in two ways that are distinct up to the action of* $\mathrm{GL}_2(\mathbb{Z})$*: one returns us to* $P_{(1,1,1)}$*, whilst the other gives* $P_{(1,2,5)}$*. Continuing to mutate, we obtain an infinite graph of triangles* $P_{(a,b,c)}$*, where the* (a,b,c) *are the* Markov triples*, that is, the positive integral solutions to the Markov equation:*

$$3xyz = x^2 + y^2 + z^2.$$

The Gorenstein index of $P_{(a,b,c)}$ *is* abc*. In particular, the Gorenstein index can be made arbitrarily large whilst the Ehrhart series, and hence quasi-period, of* $P^*_{(a,b,c)}$ *is fixed. See [23] for details.*

5. Conclusion

We have seen that Support Vector Machine methods are very effective at extracting the dimension and volume of a lattice or rational polytope P, and the quasi-period of a rational polytope P, from the initial terms of its Ehrhart series. We have also seen that ML methods are unable to reliably determine the Gorenstein index of a Fano polytope P from the Ehrhart series of its polar polytope P^*. The discussions in §2.4, §3.4, and §4.2 suggest that these results are as expected: that ML is detecting known and understood structure in the dimension, volume, and quasiperiod cases, and that there is probably no structure to detect in the Gorenstein index case. But there is a more useful higher-level conclusion to draw here too: when applying ML methods to questions in pure mathematics, one needs to think carefully about methods and results. Questions 2 and 3 are superficially similar, yet one is amenable to ML and the other is not. Furthermore, applying standard ML recipes in a naive way would have led to false negative results. For example, since the Ehrhart series grows so fast, it would have been typical to suppress the growth rate by taking logarithms, and also to pass to principal components. Taking logarithms is a good idea for some of our questions but not for others; this reflects the mathematical realities underneath the data, and not just whether the vector-components y_m involved grow rapidly with m or not. Passing to principal components is certainly a useful tool, but naive feature extraction would have retained only the first principal component, which is responsible for more than 99.999% of the variation (in both the logarithmic Ehrhart vector and the Ehrhart vector). This would have left us unable to detect the positive answers to Questions 1 and 2, in the former case because projection to one dimension amalgamates clusters (see Figure 1) and in the latter case because we need to detect whether the Ehrhart vector lies in a certain high-codimension linear subspace and that structure is destroyed by projection to a low-dimensional space.

Acknowledgments. TC is supported by ERC Consolidator Grant 682603 and EPSRC Programme Grant EP/N03189X/1. JH is supported by a Nottingham Research Fellowship. AK is supported by EPSRC Fellowship EP/N022513/1.

References

1. Mohammad Akhtar, Tom Coates, Alessio Corti, Liana Heuberger, Alexander M. Kasprzyk, Alessandro Oneto, Andrea Petracci, Thomas Prince, and Ketil Tveiten, *Mirror symmetry and the classification of orbifold del Pezzo surfaces*, Proc. Amer. Math. Soc. **144** (2016), no. 2, 513–527.

2. Mohammad Akhtar, Tom Coates, Sergey Galkin, and Alexander M. Kasprzyk, *Minkowski polynomials and mutations*, SIGMA Symmetry Integrability Geom. Methods Appl. **8** (2012), Paper 094, 17.
3. Mohammad E. Akhtar and Alexander M. Kasprzyk, *Mutations of fake weighted projective planes*, Proc. Edinb. Math. Soc. (2) **59** (2016), no. 2, 271–285.
4. Jiakang Bao, Yang-Hui He, Edward Hirst, Johannes Hofscheier, Alexander M. Kasprzyk, and Suvajit Majumder, *Polytopes and Machine Learning*, arXiv:2109.09602 [math.CO], 2021.
5. ———, *Hilbert series, machine learning, and applications to physics*, Phys. Lett. B **827** (2022), Paper No. 136966, 8.
6. Matthias Beck, Benjamin Braun, and Andrés R. Vindas-Meléndez, *Decompositions of Ehrhart h^*-Polynomials for Rational Polytopes*, Discrete Comput. Geom. **68** (2022), no. 1, 50–71.
7. Matthias Beck, Sophia Elia, and Sophie Rehberg, *Rational Ehrhart theory*, arXiv:2110.10204 [math.CO], 2022.
8. Matthias Beck, Steven V. Sam, and Kevin M. Woods, *Maximal periods of (Ehrhart) quasi-polynomials*, J. Combin. Theory Ser. A **115** (2008), no. 3, 517–525.
9. Wieb Bosma, John Cannon, and Catherine Playoust, *The Magma algebra system. I. The user language*, J. Symbolic Comput. **24** (1997), no. 3-4, 235–265, Computational algebra and number theory (London, 1993).
10. Tom Coates, Johannes Hofscheier, and Alexander M. Kasprzyk, *Ehrhart series coefficients and quasi-period for random rational polytopes*, Zenodo, doi:10.5281/zenodo.6614829, 2022.
11. ———, *Ehrhart series coefficients for random lattice polytopes*, Zenodo, doi:10.5281/zenodo.6614821, 2022.
12. David A. Cox, John B. Little, and Henry K. Schenck, *Toric varieties*, Graduate Studies in Mathematics, vol. 124, American Mathematical Society, Providence, RI, 2011.
13. Alex Davies, Petar Veličković, Lars Buesing, Sam Blackwell, Daniel Zheng, Nenad Tomašev, Richard Tanburn, Peter Battaglia, Charles Blundell, András Juhász, Marc Lackenby, Geordie Williamson, Demis Hassabis, and Pushmeet Kohli, *Advancing mathematics by guiding human intuition with AI*, Nature **600** (2021), 70–74.
14. E. Ehrhart, *Sur un problème de géométrie diophantienne linéaire. II. Systèmes diophantiens linéaires*, J. Reine Angew. Math. **227** (1967), 25–49.
15. Eugène Ehrhart, *Sur les polyèdres homothétiques bordés à n dimensions*, C. R. Acad. Sci. Paris **254** (1962), 988–990.
16. Matthew H. J. Fiset and Alexander M. Kasprzyk, *A note on palindromic δ-vectors for certain rational polytopes*, Electron. J. Combin. **15** (2008), no. 1, Note 18, 4.
17. William Fulton, *Introduction to toric varieties*, Annals of Mathematics Studies, vol. 131, Princeton University Press, Princeton, NJ, 1993, The William H. Roever Lectures in Geometry.
18. Christian Haase and Tyrrell B. McAllister, *Quasi-period collapse and $\mathrm{GL}_n(\mathbb{Z})$-scissors congruence in rational polytopes*, Integer points in polyhedra—geometry, number theory, representation theory, algebra, optimization, statistics, Contemp. Math., vol. 452, Amer. Math. Soc., Providence, RI, 2008, pp. 115–122.
19. Paul Hacking and Yuri Prokhorov, *Smoothable del Pezzo surfaces with quotient singularities*, Compos. Math. **146** (2010), no. 1, 169–192.
20. Girtrude Hamm, Johannes Hofscheier, and Alexander M. Kasprzyk, *Half-integral polygons with a fixed number of lattice points*, in preparation.
21. Andrew John Herrmann, *Classification of Ehrhart quasi-polynomials of half-integral polygons*, Master's thesis, San Francisco State University, 2010.
22. Alexander M. Kasprzyk and Benjamin Nill, *Fano polytopes*, Strings, gauge fields, and the geometry behind, World Sci. Publ., Hackensack, NJ, 2013, pp. 349–364.
23. Alexander M. Kasprzyk and Ben Wormleighton, *Quasi-period collapse for duals to Fano polygons: an explanation arising from algebraic geometry*, arXiv:1810.12472 [math.CO], 2018.
24. J. Kollár and N. I. Shepherd-Barron, *Threefolds and deformations of surface singularities*, Invent. Math. **91** (1988), no. 2, 299–338.
25. Tyrrell B. McAllister, *Coefficient functions of the Ehrhart quasi-polynomials of rational polygons*, Proceedings of the 2008 International Conference on Information Theory and Statistical Learning, ITSL 2008, Las Vegas, Nevada, USA, July 14-17, 2008 (Matthias Dehmer, Michael Drmota, and Frank Emmert-Streib, eds.), CSREA Press, 2008.
26. Tyrrell B. McAllister and Kevin M. Woods, *The minimum period of the Ehrhart quasi-polynomial of a rational polytope*, J. Combin. Theory Ser. A **109** (2005), no. 2, 345–352.
27. P. McMullen, *Lattice invariant valuations on rational polytopes*, Arch. Math. (Basel) **31** (1978/79), no. 5, 509–516.
28. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, *Scikit-learn: Machine learning in Python*, Journal of Machine Learning Research **12** (2011), 2825–2830.
29. Richard P. Stanley, *Decompositions of rational convex polytopes*, Ann. Discrete Math. **6** (1980), 333–342.

Department of Mathematics, Imperial College London, 180 Queen's Gate, London, SW7 2AZ, UK
Email address: t.coates@imperial.ac.uk

School of Mathematical Sciences, University of Nottingham, Nottingham, NG7 2RD, UK
Email address: johannes.hofscheier@nottingham.ac.uk

School of Mathematical Sciences, University of Nottingham, Nottingham, NG7 2RD, UK
Email address: a.m.kasprzyk@nottingham.ac.uk