Predicting the Integer Decomposition Property via Machine Learning

Brian Davis

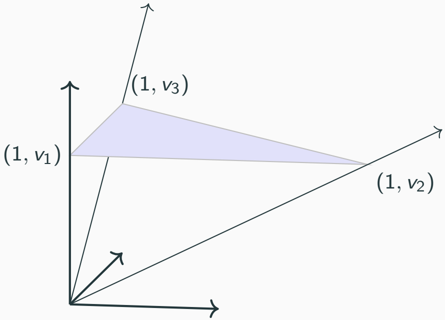
2 Aug 2018

University of Kentucky

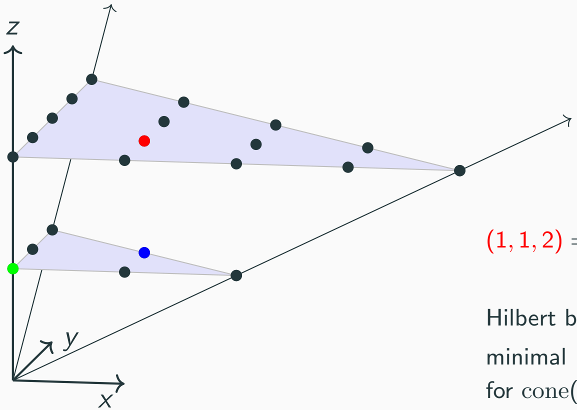
Hilbert Basics

Cone over a lattice simplex

$$\boldsymbol{v} = \{v_1, \ldots, v_{d+1}\} \subset \mathbb{Z}^d \qquad \mathrm{cone}(\boldsymbol{v}) = \mathbb{R}_{\geq 0}\big\langle (1, v_1), \ldots, (1, v_{d+1}) \big\rangle \subset \mathbb{R}^{d+1}$$



Hilbert basis = additively minimal lattice points



$(1,1,2) = (0,0,1) + (1,1,1)$

Hilbert basis $\mathrm{HB}(\boldsymbol{v})$:

minimal additive generating set for $\mathrm{cone}(\boldsymbol{v}) \cap \mathbb{Z}^{d+1}$

The Integer Decomposition Property (IDP)

We say that the simplex with vertices $\{v_1, \ldots, v_{d+1}\}$ has the **Integer Decomposition Property (IDP)** if for all elements z of the Hilbert basis $\mathrm{HB}(\boldsymbol{v})$,

$$\text{height}(z) := z_0$$

is equal to 1.

WANTED: Large, diverse set of examples of IDP simplices

WANTED: Large, diverse set of examples of IDP simplices

PROPOSED SOLUTION:

Construct very large test set and use Normaliz to reject non-IDP examples

WANTED: Large, diverse set of examples of IDP simplices

PROPOSED SOLUTION:

Construct very large test set and use Normaliz to reject non-IDP examples

PROBLEM:

Computing Hilbert basis over very large test sets is expensive (computationally)

WANTED: Large, diverse set of examples of IDP simplices

PROPOSED SOLUTION:

Construct very large test set and use Normaliz to reject non-IDP examples

PROBLEM:

Computing Hilbert basis over very large test sets is expensive (computationally)

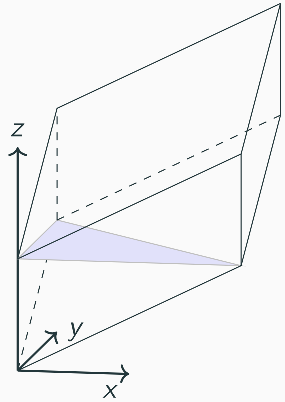
PROPOSED SOLUTION:

Consider $\mathbb{IDP}$ to be a 0/1 function and approximate it with an easily evaluated function $\widehat{\mathbb{IDP}}$.

Apply $\widehat{\mathbb{IDP}}$ to very large test set to get likely IDP candidates, then validate candidates with Normaliz.

An intermediate step

The fundamental parallelepiped



The fundamental parallelepiped Π is the set

$$\left\{ \sum_{i=1}^{d+1} \gamma_i (1, v_i) \;:\; 0 \leq \gamma_i < 1 \right\}$$

FACT:
$\mathrm{HB}(\boldsymbol{v})$ is the union of $\{(1, v_1), \ldots, (1, v_{d+1})\}$ and the additively minimal elements of

$$\Pi \cap \mathbb{Z}^{d+1}$$

Partition the fundamental parallelepiped

Map $z \in \Pi$ to $\{0, \ldots, n-1\}^{d+1}$

by

$$z \mapsto \left(\lfloor n\, \gamma_1 \rfloor, \ldots, \lfloor n\, \gamma_{d+1} \rfloor \right)$$

Partition the fundamental parallelepiped

Map $z \in \Pi$ to $\{0, \ldots, n-1\}^{d+1}$

by

$$z \mapsto \left(\lfloor n\, \gamma_1 \rfloor, \ldots, \lfloor n\, \gamma_{d+1} \rfloor \right)$$

EX: $n = d = 2$

$$z = \frac{3}{4}(1, v_1) + \frac{2}{4}(1, v_2) + \frac{3}{4}(1, v_3)$$

$$z \mapsto \left(\left\lfloor 2 \cdot \frac{3}{4} \right\rfloor, \left\lfloor 2 \cdot \frac{2}{4} \right\rfloor, \left\lfloor 2 \cdot \frac{3}{4} \right\rfloor \right)$$
$$= (1, 1, 1) \in \{0, 1\}^3$$

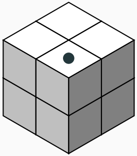
Partition the fundamental parallelepiped

Map $z \in \Pi$ to $\{0, \ldots, n-1\}^{d+1}$

by

$$z \mapsto \left(\lfloor n\, \gamma_1 \rfloor, \ldots, \lfloor n\, \gamma_{d+1} \rfloor \right)$$

Record the box containing z



EX: $n = d = 2$

$$z = \frac{3}{4}(1, v_1) + \frac{2}{4}(1, v_2) + \frac{3}{4}(1, v_3)$$

$$z \mapsto \left(\left\lfloor 2 \cdot \frac{3}{4} \right\rfloor, \left\lfloor 2 \cdot \frac{2}{4} \right\rfloor, \left\lfloor 2 \cdot \frac{3}{4} \right\rfloor \right)$$
$$= (1, 1, 1) \in \{0, 1\}^3$$

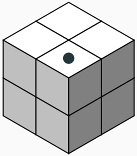
Partition the fundamental parallelepiped

Map $z \in \Pi$ to $\{0, \ldots, n-1\}^{d+1}$

by

$$z \mapsto \left(\lfloor n\,\gamma_1 \rfloor, \ldots, \lfloor n\,\gamma_{d+1} \rfloor \right)$$

Record the box containing z



$\downarrow$

EX: $n = d = 2$

$$z = \frac{3}{4}(1, v_1) + \frac{2}{4}(1, v_2) + \frac{3}{4}(1, v_3)$$

$$z \mapsto \left(\left\lfloor 2 \cdot \frac{3}{4} \right\rfloor, \left\lfloor 2 \cdot \frac{2}{4} \right\rfloor, \left\lfloor 2 \cdot \frac{3}{4} \right\rfloor \right)$$
$$= (1, 1, 1) \in \{0, 1\}^3$$

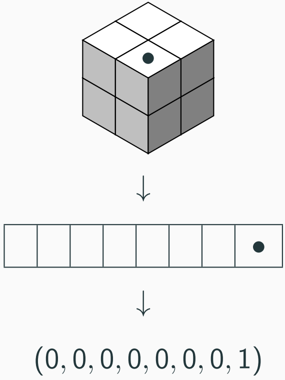
Partition the fundamental parallelepiped

Map $z \in \Pi$ to $\{0, \ldots, n-1\}^{d+1}$

by

$$z \mapsto \left(\lfloor n\,\gamma_1 \rfloor, \ldots, \lfloor n\,\gamma_{d+1} \rfloor \right)$$

Record the box containing z



EX: $n = d = 2$

$$z = \frac{3}{4}(1, v_1) + \frac{2}{4}(1, v_2) + \frac{3}{4}(1, v_3)$$

$$z \mapsto \left(\left\lfloor 2 \cdot \frac{3}{4} \right\rfloor, \left\lfloor 2 \cdot \frac{2}{4} \right\rfloor, \left\lfloor 2 \cdot \frac{3}{4} \right\rfloor \right)$$
$$= (1, 1, 1) \in \{0, 1\}^3$$

The function $\mathbb{HB}$

Partition Π into $(d+1)^{d+1}$ boxes indexed by $\alpha \in \{0, \ldots, d\}^{d+1}$

The function $\mathbb{HB}$

Partition Π into $(d+1)^{d+1}$ boxes indexed by $\alpha \in \{0, \ldots, d\}^{d+1}$

$$(v_1, \ldots, v_{d+1}) \xrightarrow{\mathbb{HB}} \{0,1\}^{{d+1}^{d+1}}$$

The function $\mathbb{HB}$

Partition Π into $(d+1)^{d+1}$ boxes indexed by $\alpha \in \{0, \ldots, d\}^{d+1}$

$$(v_1, \ldots, v_{d+1}) \xrightarrow{\mathbb{HB}} \{0,1\}^{{d+1}^{d+1}}$$

$$\mathbb{HB}(\boldsymbol{v})_\alpha = \begin{cases} 1 & \text{if there exists a Hilbert basis element in box } \alpha \\ 0 & \text{otherwise} \end{cases}$$

Why bother?

FACT:

If z in $\Pi \cap \mathbb{Z}^{d+1}$ lies in box with indices $\alpha = (i_1, \ldots, i_{d+1})$, then

$$\text{height}(z) = \left\lceil \frac{i_1 + \cdots + i_{d+1}}{d+1} \right\rceil$$

Why bother?

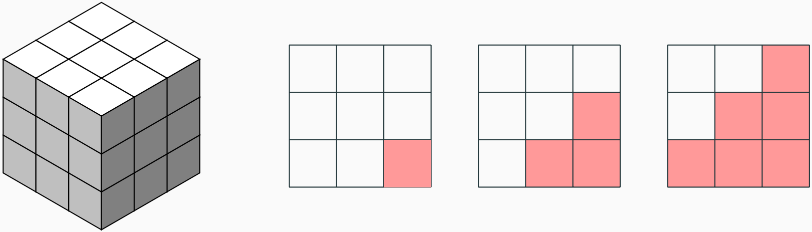
FACT:

If z in $\Pi \cap \mathbb{Z}^{d+1}$ lies in box with indices $\alpha = (i_1, \ldots, i_{d+1})$, then

$$\text{height}(z) = \left\lceil \frac{i_1 + \cdots + i_{d+1}}{d+1} \right\rceil$$

CONSEQUENCE:

We can detect IDP by looking at support of $\mathbb{HB}(\boldsymbol{v})$.



Defining $\widehat{\mathbb{IDP}}$ via $\widehat{\mathbb{HB}}$

$\mathbb{IDP}$ is the the composite function:

$$\boldsymbol{v} \xrightarrow{\mathbb{HB}} \{0,1\}^{d+1^{d+1}} \xrightarrow{\mathsf{supp}} \{0,1\}$$

Defining $\widehat{\mathbb{IDP}}$ via $\widehat{\mathbb{HB}}$

$\mathbb{IDP}$ is the the composite function:

$$\boldsymbol{v} \xrightarrow{\mathbb{HB}} \{0,1\}^{d+1^{d+1}} \xrightarrow{\text{supp}} \{0,1\}$$

Let $\widehat{\mathbb{HB}}$ be any function from $\mathbf{v} \rightarrow (0,1)^{d+1^{d+1}}$

Defining $\widehat{\mathbb{IDP}}$ via $\widehat{\mathbb{HB}}$

$\mathbb{IDP}$ is the the composite function:

$$\boldsymbol{v} \xrightarrow{\mathbb{HB}} \{0,1\}^{d+1^{d+1}} \xrightarrow{\text{supp}} \{0,1\}$$

Let $\widehat{\mathbb{HB}}$ be any function from $\mathbf{v} \to (0,1)^{d+1^{d+1}}$

Pick $0 < \eta < 1$

$$\text{cutoff}(x) = \begin{cases} 1 & \text{if } x \geq \eta \\ 0 & \text{if } x < \eta \end{cases}$$

Defining $\widehat{\mathbb{IDP}}$ via $\widehat{\mathbb{HB}}$

$\mathbb{IDP}$ is the the composite function:

$$\boldsymbol{v} \xrightarrow{\mathbb{HB}} \{0,1\}^{d+1^{d+1}} \xrightarrow{\mathsf{supp}} \{0,1\}$$

Let $\widehat{\mathbb{HB}}$ be any function from $\mathbf{v} \to (0,1)^{d+1^{d+1}}$

Pick $0 < \eta < 1$

$$\mathrm{cutoff}(x) = \begin{cases} 1 & \text{if } x \geq \eta \\ 0 & \text{if } x < \eta \end{cases}$$

$\widehat{\mathbb{IDP}}$ is the the composite function:

$$\boldsymbol{v} \xrightarrow{\widehat{\mathbb{HB}}} (0,1)^{d+1^{d+1}} \xrightarrow{\mathrm{cutoff}} \{0,1\}^{d+1^{d+1}} \xrightarrow{\mathsf{supp}} \{0,1\}$$

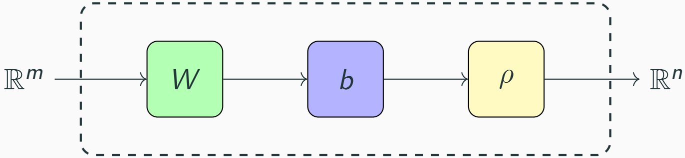
A general method for creating piece-wise linear approximations

The building blocks

matrix: $W \in \mathbb{R}^{n \times m}$ (weights)

vector: $b \in \mathbb{R}^{n}$ (biases)

function: $\rho(z) = \max(0, z)$ coordinatewise

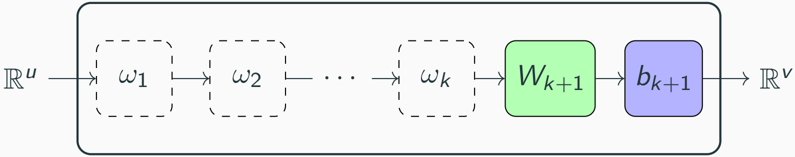


$$\omega(x) = \rho(Wx + b)$$

An initial approximation of $f : \mathbb{R}^u \longrightarrow \mathbb{R}^v$

Pick positive integers k and $\ell_1, \ldots, \ell_k$,

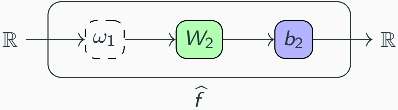
Set weights W_i and biases b_i randomly for $\omega_i : \mathbb{R}^{\ell_i} \longrightarrow \mathbb{R}^{\ell_{i+1}}$



approximation $\widehat{f}$

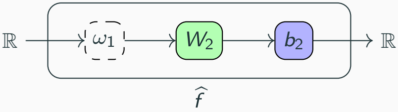
Note: $\widehat{f}$ is piece-wise linear

EXAMPLE: $f(x) = \log(x)$ on interval $[1, 3]$



$$W_1 = [0.75, -0.5]^T \quad b_1 = [-0.75, 1] \quad W_2 = [1, 1] \quad b_2 = [-0.5]$$

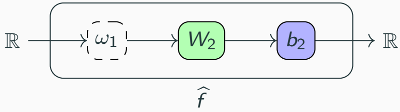
EXAMPLE: $f(x) = \log(x)$ on interval $[1, 3]$



$$W_1 = [0.75, -0.5]^T \quad b_1 = [-0.75, 1] \quad W_2 = [1, 1] \quad b_2 = [-0.5]$$

$\widehat{f}(x) = [1, 1]\omega_1(x) + [-0.5]$

EXAMPLE: $f(x) = \log(x)$ on interval $[1, 3]$

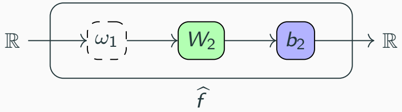


$$W_1 = [0.75, -0.5]^T \quad b_1 = [-0.75, 1] \quad W_2 = [1, 1] \quad b_2 = [-0.5]$$

$$\widehat{f}(x) = [1, 1]\omega_1(x) + [-0.5]$$

$$= [1, 1]\rho\left(\begin{bmatrix} 0.75 \\ -0.5 \end{bmatrix}[x] + \begin{bmatrix} -0.75 \\ 1 \end{bmatrix}\right) + [-0.5]$$

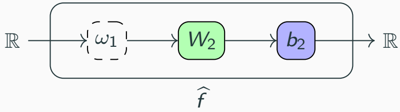
EXAMPLE: $f(x) = \log(x)$ on interval $[1, 3]$



$$W_1 = [0.75, -0.5]^T \quad b_1 = [-0.75, 1] \quad W_2 = [1, 1] \quad b_2 = [-0.5]$$

$$\begin{aligned}
\widehat{f}(x) &= [1, 1]\omega_1(x) + [-0.5] \\
&= [1, 1]\rho\left(\begin{bmatrix} 0.75 \\ -0.5 \end{bmatrix}[x] + \begin{bmatrix} -0.75 \\ 1 \end{bmatrix}\right) + [-0.5] \\
&= 1 \cdot \rho(0.75x - 0.75) \\
&\qquad + 1 \cdot \rho(-0.5x + 1) - 0.5
\end{aligned}$$

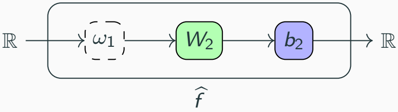
EXAMPLE: $f(x) = \log(x)$ on interval $[1, 3]$



$$W_1 = [0.75, -0.5]^T \quad b_1 = [-0.75, 1] \quad W_2 = [1, 1] \quad b_2 = [-0.5]$$

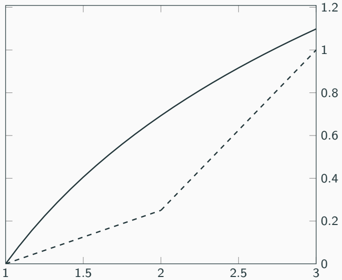
$$\begin{aligned}
\widehat{f}(x) &= [1, 1]\omega_1(x) + [-0.5] \\
&= [1, 1]\rho\left(\begin{bmatrix} 0.75 \\ -0.5 \end{bmatrix}[x] + \begin{bmatrix} -0.75 \\ 1 \end{bmatrix}\right) + [-0.5] \\
&= 1 \cdot \rho(0.75x - 0.75) \\
&\quad + 1 \cdot \rho(-0.5x + 1) - 0.5 \\
&= \begin{cases} 0.25x - 0.25 & 1 \leq x \leq 2 \\ 0.75x - 1.25 & 2 < x \leq 3 \end{cases}
\end{aligned}$$

EXAMPLE: $f(x) = \log(x)$ on interval $[1, 3]$



$$W_1 = [0.75, -0.5]^T \quad b_1 = [-0.75, 1] \quad W_2 = [1, 1] \quad b_2 = [-0.5]$$

$$\begin{aligned}
\widehat{f}(x) &= [1,1]\omega_1(x) + [-0.5] \\
&= [1,1]\rho\left(\begin{bmatrix}0.75\\-0.5\end{bmatrix}[x] + \begin{bmatrix}-0.75\\1\end{bmatrix}\right) + [-0.5] \\
&= 1\cdot\rho(0.75x - 0.75) \\
&\quad + 1\cdot\rho(-0.5x + 1) - 0.5 \\
&= \begin{cases} 0.25x - 0.25 & 1 \le x \le 2 \\ 0.75x - 1.25 & 2 < x \le 3 \end{cases}
\end{aligned}$$



f and $\widehat{f}$ (dashed)

Refining the approximation $\widehat{f}$

Define a **loss function** $L(x)$ which measures the accuracy of $\widehat{f}$.

Refining the approximation $\widehat{f}$

Define a **loss function** $L(x)$ which measures the accuracy of $\widehat{f}$.

Consider L as a function of the parameters

$$p = (W_1, b_1, \ldots, W_{k+1}, b_{k+1})$$

and compute the gradient ∇L.

Refining the approximation $\widehat{f}$

Define a **loss function** $L(x)$ which measures the accuracy of $\widehat{f}$.

Consider L as a function of the parameters

$$p = (W_1, b_1, \ldots, W_{k+1}, b_{k+1})$$

and compute the gradient ∇L.

Update the parameters $p \mapsto p'$ by

$$p' = p - \epsilon \nabla L$$

Refining the approximation $\widehat{f}$

Define a **loss function** $L(x)$ which measures the accuracy of $\widehat{f}$.

Consider L as a function of the parameters

$$p = (W_1, b_1, \ldots, W_{k+1}, b_{k+1})$$

and compute the gradient ∇L.

Update the parameters $p \mapsto p'$ by

$$p' = p - \epsilon \nabla L$$

For sufficiently small ϵ, we expect that $L_{p'}(x) < L_p(x)$

Our new approximation is $\widehat{f}$ with the updated parameters p'

$f(x) = \log(x)$ continued

Let $x = 1.5$ and use Euclidean distance loss function

$$L(x) = \left\| \log(x) - \widehat{f}(x) \right\|$$

$$\nabla L = \left\langle \frac{\partial L}{\partial W_1}, \frac{\partial L}{\partial b_1}, \frac{\partial L}{\partial W_2}, \frac{\partial L}{\partial b_2} \right\rangle_{x=1.5}$$

$$= \langle -1.5, -1.5, -1, -1, -0.375, -0.25, -1 \rangle .$$

$f(x) = \log(x)$ **continued**

Let $x = 1.5$ and use Euclidean distance loss function

$$L(x) = \left\| \log(x) - \widehat{f}(x) \right\|$$

$$\nabla L = \left\langle \frac{\partial L}{\partial W_1}, \frac{\partial L}{\partial b_1}, \frac{\partial L}{\partial W_2}, \frac{\partial L}{\partial b_2} \right\rangle_{x=1.5}$$

$$= \langle -1.5, -1.5, -1, -1, -0.375, -0.25, -1 \rangle .$$

For $\epsilon = 0.02$, the update $p' = p - \epsilon \nabla L$ is given by

$$W_1 = [0.78, -0.47]^T \quad b_1 = [-0.73, 1.02]$$

$$W_2 = [1.0075, 1.0075] \quad b_2 = [-0.48]$$

$f(x) = \log(x)$ continued

Let $x = 1.5$ and use Euclidean distance loss function

$$L(x) = \left\| \log(x) - \widehat{f}(x) \right\|$$

$$\nabla L = \left\langle \frac{\partial L}{\partial W_1}, \frac{\partial L}{\partial b_1}, \frac{\partial L}{\partial W_2}, \frac{\partial L}{\partial b_2} \right\rangle_{x=1.5}$$

$$= \langle -1.5, -1.5, -1, -1, -0.375, -0.25, -1 \rangle .$$

For $\epsilon = 0.02$, the update $p' = p - \epsilon \nabla L$ is given by

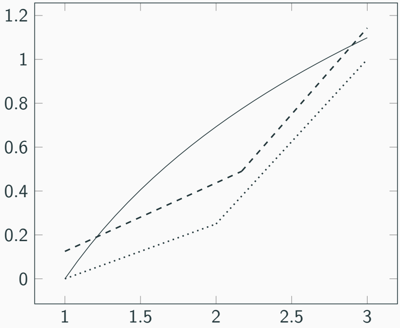
$$W_1 = [0.78, -0.47]^T \quad b_1 = [-0.73, 1.02]$$

$$W_2 = [1.0075, 1.0075] \quad b_2 = [-0.48]$$

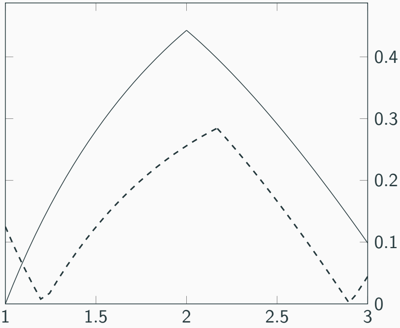
The updated approximation is

$$\widehat{f}(x) = \begin{cases} 0.312x - 0.187 & 1 \leq x \leq 2.17 \\ 0.786x - 1.215 & 2.17 < x \leq 3 \end{cases}$$

$f(x) = \log(x)$ **continued**



$\widehat{f}(x; p)$ (dotted) and
$\widehat{f}(x; p')$ (dashed)



Loss for $\widehat{f}(x; p)$ and
$\widehat{f}(x; p')$ (dashed)

Implementation

Training data

$\mathbb{HB}$ is expensive to compute, so we pre-compute a collection of values which we reuse multiple times to update parameters of $\widehat{\mathbb{HB}}$

Training data

$\mathbb{HB}$ is expensive to compute, so we pre-compute a collection of values which we reuse multiple times to update parameters of $\widehat{\mathbb{HB}}$

We restrict to the case that $d = 4$ and $\boldsymbol{v}$ is of the form:

- $v_i = e_i$ for $1 \leq i \leq d$, and
- $v_{d+1} = (-\lambda_1, -\lambda_2, -\lambda_3, -\lambda_4)$ with $1 \leq \lambda_i \leq 25$

Training data

$\mathbb{HB}$ is expensive to compute, so we pre-compute a collection of values which we reuse multiple times to update parameters of $\widehat{\mathbb{HB}}$

We restrict to the case that $d = 4$ and $\boldsymbol{v}$ is of the form:

- $v_i = e_i$ for $1 \leq i \leq d$, and
- $v_{d+1} = (-\lambda_1, -\lambda_2, -\lambda_3, -\lambda_4)$ with $1 \leq \lambda_i \leq 25$

Using Normaliz, we compute $\mathbb{HB}(\boldsymbol{v})$ for 50,000 such examples drawn uniformly at random.

Hyperparameters

Trainable parameters: $\approx$ 12million

Learning rate: $\epsilon = 10^{-4}$

Number of updates: 200,000 batches of size 25

Loss function

Binary Cross Entropy: For f a $0/1$ function, $\widehat{f}$ real valued,

$$\mathrm{BCE}(x) = (f - 1) \cdot \log\left(1 - \sigma \circ \widehat{f}\right) - f \cdot \log\left(\sigma \circ \widehat{f}\right)$$

where σ is the *sigmoid function* $\sigma(z) = (1 + e^{-z})^{-1}$.

Results

$\lambda = (5, 11, 11, 20)$

$\lambda = (5, 11, 11, 20)$

$$\widehat{f}(\boldsymbol{v}) = (-51.4, -26.9, -62.9, -29.6, -25.2, -30.2, -2.1, \ldots)$$

$\lambda = (5, 11, 11, 20)$

$$\widehat{f}(\boldsymbol{v}) = (-51.4, -26.9, -62.9, -29.6, -25.2, -30.2, -2.1, \dots)$$

$$\widehat{\mathbb{HB}}(\boldsymbol{v}) = (4.4e{-}23, 1.9e{-}12, 4.6e{-}28, 1.3e{-}13, 1.0e{-}11, 7.0e{-}14, 1.0e{-}01, \dots)$$

$\lambda = (5, 11, 11, 20)$

$$\widehat{f}(\boldsymbol{v}) = (-51.4, -26.9, -62.9, -29.6, -25.2, -30.2, -2.1, \ldots)$$

$$\widehat{\mathbb{HB}}(\boldsymbol{v}) = \ (4.4e{-}23, 1.9e{-}12, 4.6e{-}28, 1.3e{-}13, 1.0e{-}11, 7.0e{-}14, 1.0e{-}01, \ldots)$$

$$\text{with } \eta = 0.1, \quad \mathrm{cutoff}\left(\widehat{\mathbb{HB}}(\boldsymbol{v})\right) = (0, 0, 0, 0, 0, 0, 1, \ldots)$$

$\lambda = (5, 11, 11, 20)$

$$\widehat{f}(\boldsymbol{v}) = (-51.4, -26.9, -62.9, -29.6, -25.2, -30.2, -2.1, \ldots)$$

$$\widehat{\mathbb{HB}}(\boldsymbol{v}) = \; (4.4e{-}23, 1.9e{-}12, 4.6e{-}28, 1.3e{-}13, 1.0e{-}11, 7.0e{-}14, 1.0e{-}01, \ldots)$$

with $\eta = 0.1$, $\quad \mathrm{cutoff}\left(\widehat{\mathbb{HB}}(\boldsymbol{v})\right) = (0, 0, 0, 0, 0, 0, 1, \ldots)$

$\eta = 0.1$	PREDICTED 0	PREDICTED 1
ACTUAL 0	2,705	160
ACTUAL 1	1	11

$\lambda = (5, 11, 11, 20)$

$$\widehat{f}(\boldsymbol{v}) = (-51.4, -26.9, -62.9, -29.6, -25.2, -30.2, -2.1, \dots)$$

$$\widehat{\mathbb{HB}}(\boldsymbol{v}) = (4.4e{-}23, 1.9e{-}12, 4.6e{-}28, 1.3e{-}13, 1.0e{-}11, 7.0e{-}14, 1.0e{-}01, \dots)$$

$$\text{with } \eta = 0.1, \quad \mathrm{cutoff}\left(\widehat{\mathbb{HB}}(\boldsymbol{v})\right) = (0, 0, 0, 0, 0, 0, 1, \dots)$$

$\eta = 0.1$	PREDICTED 0	PREDICTED 1
ACTUAL 0	2,705	160
ACTUAL 1	1	11

$$\text{specificity (top row)} = \frac{\text{true negatives}}{\text{true negatives} + \text{false positives}} = 94\%$$

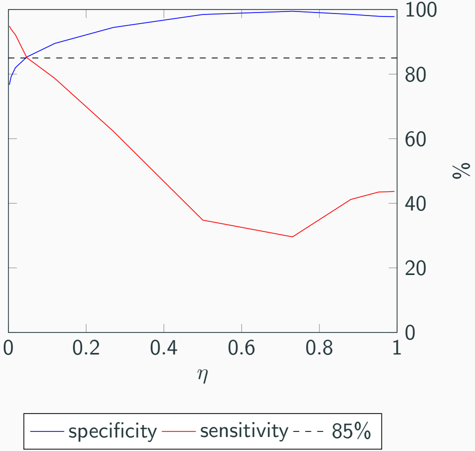
$$\text{sensitivity (bottom row)} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}} = 92\%$$

5,000 random samples aggregated

$\eta = 0.01$	PREDICTED 0	PREDICTED 1
ACTUAL 0	11,448,675	2,845,413
ACTUAL 1	6,572	92,971

specificity $= 80\%$ sensitivity $= 93\%$

Effect on $\widehat{\mathbb{HB}}$ of varying η



The resulting approximation $\widehat{\mathbb{IDP}}$

In the sample of 5,000 simplices, there were 112 IDP examples ($\approx 2.4\%$)

The resulting approximation $\widehat{\mathbb{IDP}}$

In the sample of 5,000 simplices, there were 112 IDP examples ($\approx 2.4\%$)

Result of applying $\widehat{\mathbb{IDP}}$ to the sample:

τ \ η	0.5	0.25	0.12	0.05
0	3/7 (42.9%)	3/4 (75.0%)	3/3 (100.0%)	3/3 (100.0%)
10	21/320 (6.6%)	11/38 (29.0%)	8/21 (38.1%)	6/12 (50.0%)
20	46/1026 (4.5%)	21/102 (20.6%)	11/45 (24.4%)	8/27 (29.6%)
30	65/1770 (3.7%)	35/196 (17.9%)	23/103 (22.3%)	16/64 (25.0%)

The IDP examples predicted by $\widehat{\mathbb{IDP}}$ ($\eta = 0.1$, $\tau = 65$)

1,1,1,1	1,1,3,9	1,1,21,24	1,2,14,10	1,2,14,10	24,2,1,16
1,3,16,3	1,3,24,1	1,4,2,16	1,4,20,20	1,8,1,1	24,4,2,4
1,10,10,8	1,10,24,24	1,12,4,12	1,15,3,1	1,18,1,6	24,24,23,12
1,21,1,4	1,24,1,9	1,24,14,2	1,24,17,1	1,24,18,1	24,24,6,24
1,24,18,4	1,24,24,20	2,2,2,7	2,3,12,18	2,8,8,4	23,24,24,12
2,10,1,16	2,20,10,5	3,1,1,9	3,6,12,1	3,12,2,24	23,18,3,24
3,14,21,3	3,19,3,1	3,23,15,3	4,1,1,4	4,8,2,16	23,2,2,6
4,20,1,14	4,20,10,20	4,23,4,12	4,24,1,16	6,1,2,12	22,22,20,1
6,2,6,3	6,2,18,9	6,6,6,3	6,14,6,15	6,17,9,18	22,16,22,1
7,3,21,7	7,7,1,7	7,7,16,16	8,1,8,2	8,2,12,24	22,16,4,1
8,16,4,2	9,1,1,9	9,6,18,2	9,9,4,4	9,18,4,4	22,2,2,22
9,18,18,6	9,22,1,11	10,1,5,22	10,5,10,9	10,24,4,1	21,21,16,4
11,22,5,5	12,1,2,6	12,1,24,19	12,2,3,12	12,2,18,3	20,22,1,22
12,3,2,6	12,3,11,6	12,6,1,1	12,6,1,3	12,12,4,12	20,20,4,20
12,16,1,16	12,24,2,24	12,24,6,1	13,2,2,20	14,6,14,7	20,20,4,1
14,7,2,24	14,7,12,1	15,1,13,15	15,15,1,1	16,1,6,6	20,20,1,20
16,4,2,16	16,7,16,16	16,8,4,2	16,16,12,3	16,24,1,22	20,14,24,1
17,1,7,1	17,17,8,4	17,17,17,1	18,1,1,15	18,2,6,6	
18,2,22,1	18,10,1,15	19,19,1,16	20,2,1,12	20,8,19,8	

The big search

We computed the value of $\widehat{\mathbb{IDP}}$ for all 390,625 $\Delta_{(1,q)}$ simplices with q-vector in $[1, 25]^4$ using $\eta = 0.007$ and $\tau = 50$.

The computation produced 3,773 predicted positives.

We then computed $\mathbb{IDP}$ for these examples and found that 856 were IDP.

This corresponds to a specificity of about 23%.

Remarks

1. If we train our approximation $\widehat{\mathbb{HB}}$ to have high sensitivity, then we can recover a set of lattice points containing the Hilbert basis. If the specificity is high, then reducing this set will require fewer steps than reducing the entire fundamental parallelepiped.
2. If we record the number of FPP points in each bin instead of the presence of Hilbert basis elements, then we have a model for predicting **unimodality** of the h^*-vector.