Exercices: zonotopes and constrained zonotopes

General properties and examples

We recall the definition of zonotopes below:

Definition 1 (Zonotope) *An n-dimensional zonotope $\mathcal{Z}$ with center $c \in R^n$ and a vector $G = \begin{bmatrix} g_1 \dots g_p \end{bmatrix} \in \mathbb{R}^{n,p}$ of p generators $g_j = (g_{ij})_{i=1,\dots,n} \in \mathbb{R}^n$ for $j = 1, \dots, p$ is defined as $\mathcal{Z} = \langle c, G \rangle = \{c + G\varepsilon \mid \|\varepsilon\|_\infty \leq 1\}$.*

In other words, for every dimension $1 \leq i \leq n$ we have the ith coordinate z_i of points $z \in \mathcal{Z}$ that belongs to the set:

$$z_i = \{c_i + \sum_{j=1}^{p} \boldsymbol{g}_{ij}\varepsilon_j \mid \varepsilon \in [-1, 1]^p\}$$

We now introduce constrained zonotopes, as zonotopes with linear constraints on the noise symbols ε_j:

Definition 2 (Constrained Zonotope) *An n-dimensional constrained zonotope $C\mathcal{Z}$ with center $c \in R^n$, a vector $G = \begin{bmatrix} g_1 \dots g_p \end{bmatrix} \in \mathbb{R}^{n,p}$ of p generators $g_j \in \mathbb{R}^n$ for $j = 1, \dots, p$ and q constraints given by $H \in \mathbb{R}^{q,p}$ and $d \in \mathbb{R}^q$ is defined as $C\mathcal{Z} = \langle c, G, H, d \rangle = \{c + G\varepsilon \mid \|\varepsilon\|_\infty \leq 1,\ H\varepsilon \leq d\}$.*

In other words, a constrained zonotope is a zonotope with q constraints on the p noise symbols. These constraints can be used to refine the precision of the abstraction. Given that $H = (h_{ij})_{i=1,\dots,q; j=1,\dots,p}$, these constraints can be written as, for all $1 \leq k \leq q$:

$$\sum_{j=1}^{p} h_{kj}\varepsilon_j \leq d_k$$

Question 1 *Represent geometrically the zonotope*

$$\mathcal{Z} = \langle c, G \rangle = \left\langle \begin{pmatrix} 0 \\ 1/4 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1/2 & 1/4 \end{pmatrix} \right\rangle$$

in the (z_1, z_2) plane.

Question 2 *Represent geometrically, also in the (z_1, z_2) plane, the constrained zonotope*

$$C\mathcal{Z} = \langle c, G, H, d\rangle = \left\langle \begin{pmatrix} 0 \\ 1/4 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1/2 & 1/4 \end{pmatrix}, \begin{pmatrix} 1/2 & -1/4 \\ -1/2 & -1/4 \end{pmatrix}, \begin{pmatrix} 1/4 \\ 1/4 \end{pmatrix} \right\rangle$$

Hint: you should translate the constraints $H\varepsilon \leq d$ on the noise symbols ϵ_1 and ϵ_2 into constraints on the coordinates (z_1, z_2) of points $z \in C\mathcal{Z}$.

Question 3 *Is a constrained zonotope a polyhedra ? If yes, what is the constraint representation of the constrained zonotope $C\mathcal{Z}$ of Question 2? In that case also, what is the generator representation, as a polyhedron, of the constrained zonotope $C\mathcal{Z}$?*

Question 4 *How can we compute, for a given constrained zonotope $C\mathcal{Z} = \langle c, G, H, d\rangle = \{c + G\varepsilon \mid \|\varepsilon\|_\infty \leq 1,\ H\varepsilon \leq d\}$, its projection onto coordinate z_i, $i = 1, \ldots, n$? This requires only a very brief answer.*

Consider the concretization γ of constrained zonotopes $C\mathcal{Z} = \langle c, G, H, d\rangle$ to be

$$\gamma(C\mathcal{Z}) = \{c + G\varepsilon \mid \|\varepsilon\|_\infty \leq 1,\ H\varepsilon \leq d\}$$

Question 5 *Can two different constrained zonotopes have the same concretization? If you think so, please provide a small counter-example, otherwise, please write down a short argument.*

Given a set S in $\mathbb{R}^n$, is there always an abstraction of S as a constrained zonotope with minimal concretization? If you think so, please write down a short argument, otherwise please provide a small counter-example.

Affine transforms

We recall that zonotopes are closed under affine transformations: for $A \in \mathbb{R}^{m,n}$ and $b \in \mathbb{R}^m$ we can define $A\mathcal{Z} + b = \langle Ac + b, AG\rangle$ as the m-dimensional resulting zonotope.

Question 6 *Can affine transformations be also interpreted in an exact manner in contrained zonotopes? In that case, please define the affine transform of a constrained zonotope, otherwise give an short argument why this would not be the case.*

ReLU transforms

Different abstractions can be defined for the ReLU transform, among which the following one that we used in the course: let $[l_x, u_x]$ be the range reachable by component $\hat{x}$ of the input zonotope of the ReLU layer. When $l_x \leq 0$ and $u_x \geq 0$, we define the zonotope transformer for $\hat{y} = max(0, \hat{x})$ by

$$\hat{y} = \lambda\hat{x} - \frac{\lambda l_x}{2} - \frac{\lambda l_x}{2}\varepsilon_{new} \text{ with } \lambda = \frac{u_x}{u_x - l_x}. \quad (1)$$

Question 7 *Consider* $x_1 \in [-1,1]$*, what is the zonotope abstraction of* (x_1, x_2) *for* $x_2 = ReLU(x_1)$ *using the abstraction of Equation (1) ?*

Question 8 *Consider again the contrained zonotope of Question 2. Is it a correct abstraction for* $x_2 = ReLU(x_1)$ *for* $x_1 \in [-1,1]$*? Please give a short argument supporting your answer.*

Is it the best refinement, as a constrained zonotope, of the zonotope of Question 7? By refinement, we mean the following: $C\mathcal{Z}$ *is a refinement of* Z *if* $C\mathcal{Z}$ *has* Z *as underlying zonotope (hence just adding extra constraints).*

Question 9 *In view of the example of Question 8, define a ReLU transformer for constrained zonotopes refining the ReLU transformer for zonotopes by the addition of new constraints. Is it possible to make the transformer exact ?*

Analyzing a small network

Consider the toy network of Figure 1, where for simplicity all biases are taken equal to zero, and the weights are represented on the edges:

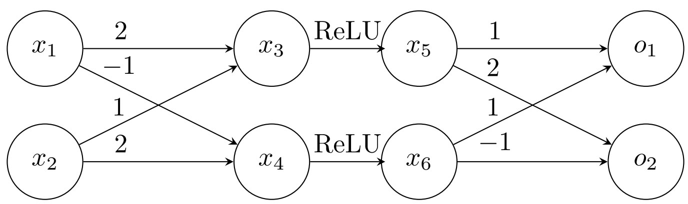

Figure 1: Toy network

Question 10 *We are interested in the local robustness of the network of Figure 1 around input* $(1,1)$*.*

Using interval computations, is $[1-1/8, 1+1/8] \times [1-1/8, 1+1/8]$ *a locally robust neighborhood of* $(1,1)$*? For this neighborhood, we say that its radius (around* $(1,1)$*, for the max distance) is 1/8.*

What is the maximal robustness radius around $(1,1)$ *that can be proved for this neural net, using the interval abstraction ?*

Question 11 *Compute the zonotope for each layer of the network of Figure 1 obtained using the zonotope abstraction with input domain* $(x_1, x_2) \in [2/3, 4/3] \times [2/3, 4/3] \wedge 3x_2 \leq 4x_1$*. As this input domain is not a zonotope, we are obliged to compute with, the input zonotope being given by the square* $[2/3, 4/3] \times [2/3, 4/3]$*.*

Can you use the zonotopic analysis to prove or disprove the property that for this input domain, we always have on the outputs $o_2 \geq o_1$ *?*

Question 12 *Now same question as Question 11, with constrained zonotopes instead of zonotopes.*