AutoEncoders for Step Functions

IA311 Project

Hélène Maxcici
Télécom Paris

February 20, 2022

Overview

1. Introduction
2. One Layer Encoder
3. One Layer Decoder
4. Two Layers Decoder

Autoencoders

- Autoencoders compress high dimensional data into lower dimensional code in an unsupervised manner.
- Autoencoders succeeded in many applications such as image denoising, dimensionality reduction, anomaly detection...
- However, the code is not always understandable

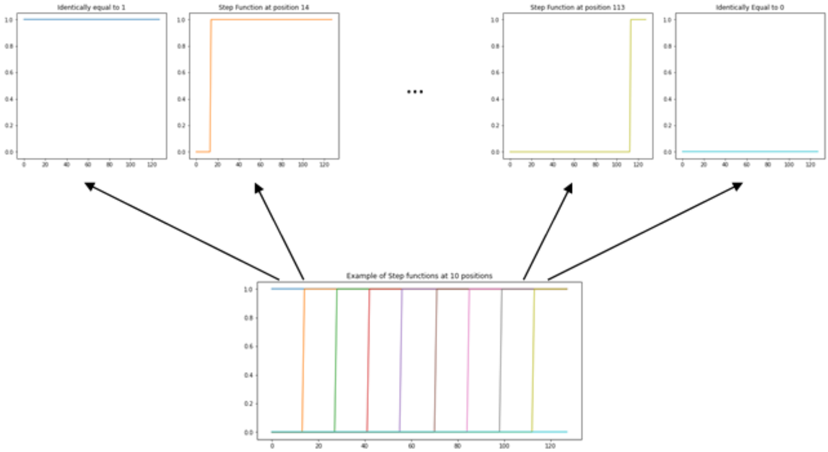
Goal

The goal is to build autoencoder for step functions with one varying parameter s.t.

- The reconstruction is perfect
- The code is scalar and linear with the parameter
- The output activation is a ReLU

Step Functions

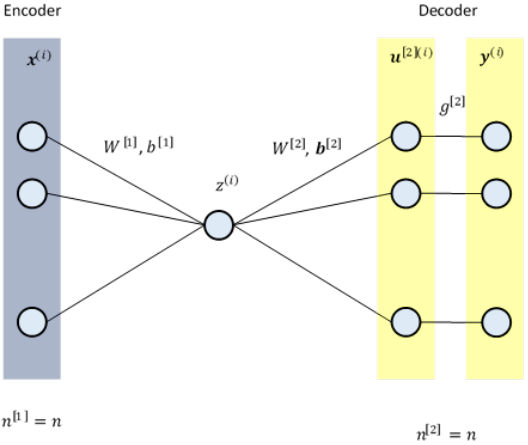
We will consider $n + 1$ step functions of dimension n.



One Layer Encoder

- A one MLP layer encoder is enough to generate linear codes with respect to the positions.
- A non-linearity is not required in the encoder.

One Layer Decoder



Necessary Condition

Lemma

A one layer decoder is able to generate the $n+1$ different step functions out of their scalar codes z only if there exist two non-overlapping proper intervals A and B such that

$$g^{[2]}(u) = \begin{cases} 0, & \text{if } u \in A \\ C, & \text{if } u \in B \end{cases}$$

Sigmoid Activation Function

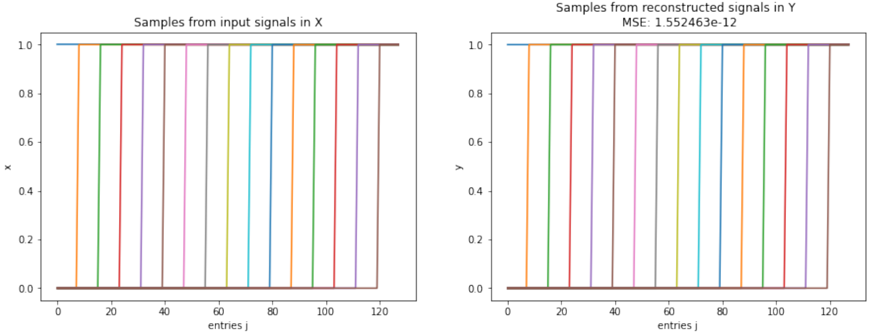
Proposition

Let the output activation be a *sigmoid* defined by $g^{[2]}(u) = \frac{e^u}{1+e^u}$. And let the encoding be linear. Then, we can find a solution $W^{[2]}$, $\mathbf{b}^{[2]} \in \mathbb{R}^n$ for the decoder, for which there exists an $\epsilon \in \mathbb{R}^+$ arbitrarily small such that

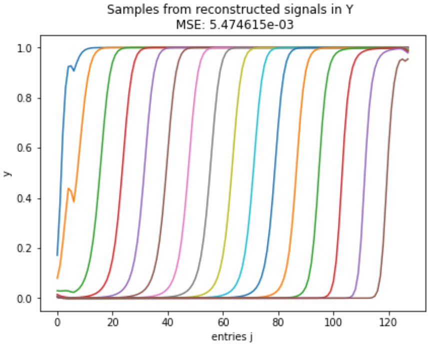
$$MSE = \frac{1}{n(n+1)} \sum_{i=1}^{n+1} ||\mathbf{y}^{(i)} - \mathbf{x}^{(i)}||_2^2 < \epsilon$$

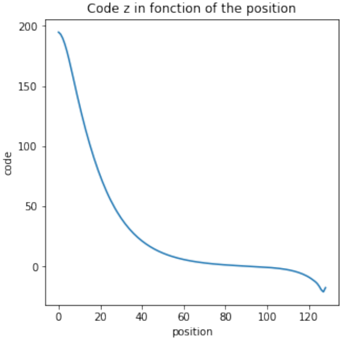
Handcrafted Solution

- Linear encoder : $z^{(i)} = i$
- MSE upper bound : $\epsilon = 10^{-5}$

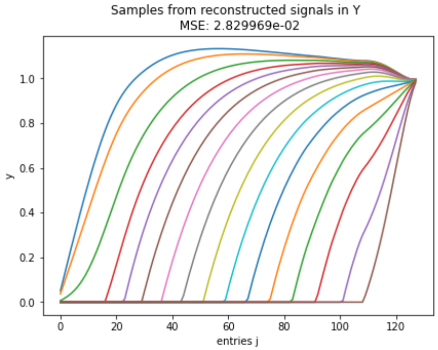


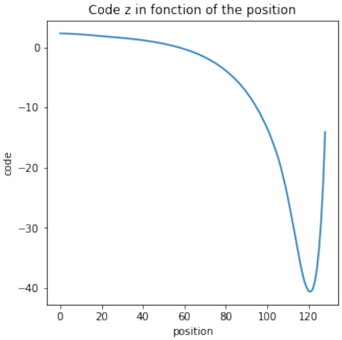
Training with no constraints



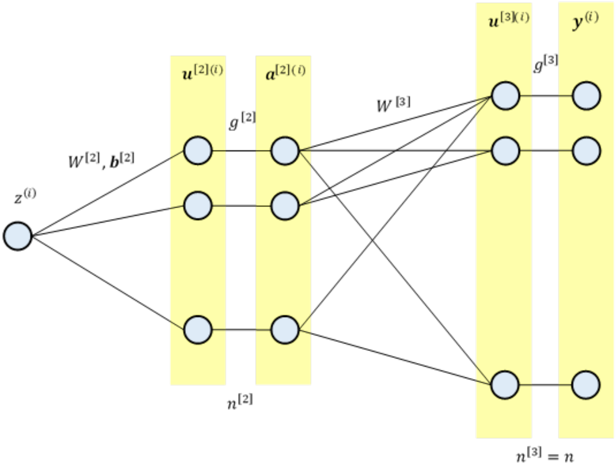


Training with no constraints





Two Layers Decoder



Layer 3

- Perfect Reconstruction:

$$\underset{(n+1)\times n}{Y} = \underset{(n+1)\times n}{X} = \begin{bmatrix} C & C & \dots & C & C \\ 0 & C & \dots & C & C \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & C \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix}$$

- ReLU Activation Function:

$$\underset{(n+1)\times n}{Y} = ReLU(\underset{(n+1)\times n}{U^{[3]}}) = \begin{bmatrix} max(0, \mathbf{u}_1^{[3](1)}) & max(0, \mathbf{u}_2^{[3](1)}) & \dots & max(0, \mathbf{u}_n^{[3](1)}) \\ max(0, \mathbf{u}_1^{[3](2)}) & max(0, \mathbf{u}_2^{[3](2)}) & \dots & max(0, \mathbf{u}_n^{[3](2)}) \\ \vdots & \vdots & \ddots & \vdots \\ max(0, \mathbf{u}_1^{[3](n)}) & max(0, \mathbf{u}_2^{[3](n)}) & \dots & max(0, \mathbf{u}_n^{[3](n)}) \\ max(0, \mathbf{u}_1^{[3](n+1)}) & max(0, \mathbf{u}_2^{[3](n+1)}) & \dots & max(0, \mathbf{u}_n^{[3](n+1)}) \end{bmatrix}$$

Layer 3

Lemma

A perfect reconstruction $Y = X$, is achieved only if $U^{[3]}$ is **full rank** matrix of the following form,

$$U^{[3]} = \begin{bmatrix} C & C & \dots & C & C \\ \mathbf{u}_1^{[3](2)} & C & \dots & C & C \\ \mathbf{u}_1^{[3](3)} & \mathbf{u}_2^{[3](3)} & \dots & C & C \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{u}_1^{[3](n)} & \mathbf{u}_2^{[3](n)} & \dots & \mathbf{u}_{n-1}^{[3](n)} & C \\ \mathbf{u}_1^{[3](n+1)} & \mathbf{u}_2^{[3](n+1)} & \dots & \mathbf{u}_{n-1}^{[3](n+1)} & \mathbf{u}_n^{[3](n+1)} \end{bmatrix} \tag{1}$$

where, $\mathbf{u}_j^{[3](i)} \leq 0$ for all $i = 2, ..., n+1$ and $j = 1, ..., i-1$.

Layer 3

- Layer 3 linear projection:

$$U^{[3]} = A^{[2]} {W^{[3]}}^T$$

- Rank of a matrix product:

$$rank(U^{[3]}) \leq \min\left(rank(\underset{(n+1)\times n^{[2]}}{A^{[2]}}), rank(\underset{n\times n^{[2]}}{W^{[3]}}) \right)$$

Lemma

The preactivation matrix $U^{[3]}$ can be full rank under three necessary conditions:

1. $n^{[2]} = n$
2. $W^{[3]}$ is full rank
3. $A^{[2]}$ is full rank

Layer 2

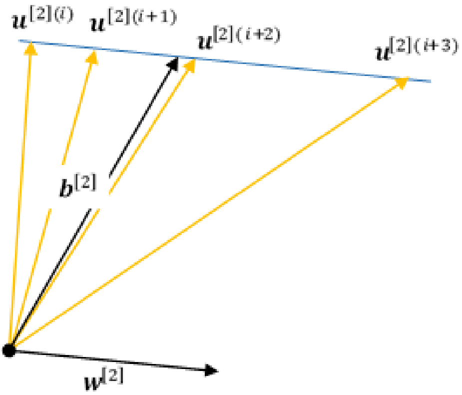
- Layer 2:

$$A^{[2]} = g^{[2]}\left(\begin{bmatrix} z^{(1)} \\ z^{(2)} \\ \vdots \\ z^{(n+1)} \end{bmatrix} {W^{[2]}}^{T} + \begin{bmatrix} \dots & \mathbf{b}^{[2]} & \dots \\ \dots & \mathbf{b}^{[2]} & \dots \\ & \vdots & \\ \dots & \mathbf{b}^{[2]} & \dots \end{bmatrix}\right)$$

Lemma

The decoder can learn to generate $A^{[2]}$ with full rank if and only if the activation is a non-linearity and in the presence of a bias.

Layer 2



Theoretical Conclusion on Architecture

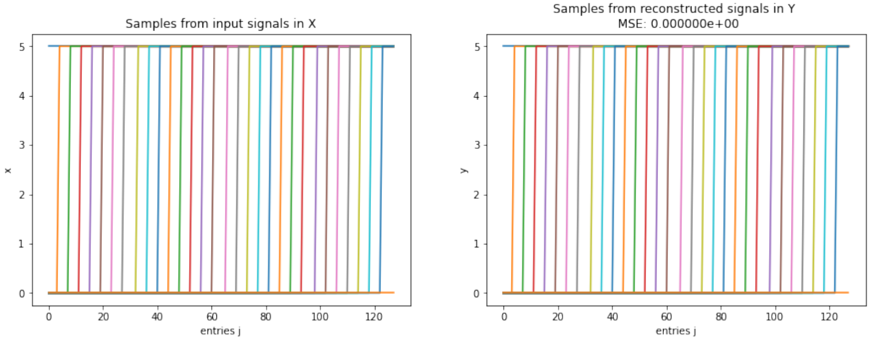
Proposition

Let's drop the signal identically equal to 0 from the dataset. A two layers decoder can perfectly reconstruction step functions from their scalar codes if and only if the layer 2 satisfies the following conditions:

1. Layer 2 has at least $n^{[2]} = n$ neurons, where n is the dimension of the signal.
2. The activation function $g^{[2]}$ of layer 2 is non-linear. In particular, $g^{[2]} = ReLU$.
3. Layer 2 has a bias vector.

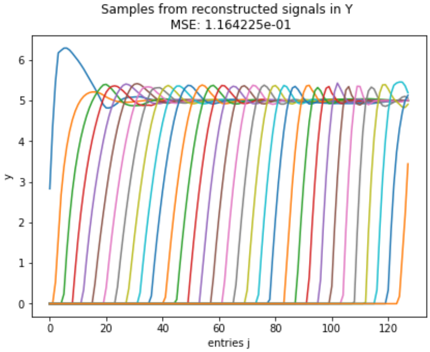
Handcrafted Solution - Linear Encoder

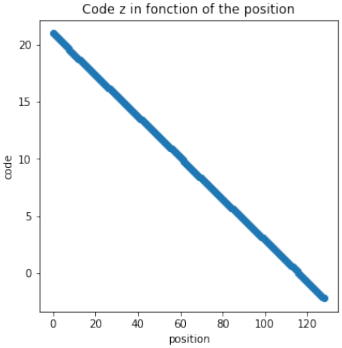
- Linear encoder : $z^{(i)} = i$
- $W^{[2]}$ and $\mathbf{b}^{[2]}$ makes $A^{[2]}$ full rank
- $U^{[3]}$ arbitrary of the form (1)
- $\implies W^{[3]} = \left(A^{[2]^{-1}} U^{[3]}\right)^T$



Trained AE - Constrained Linear Encoder

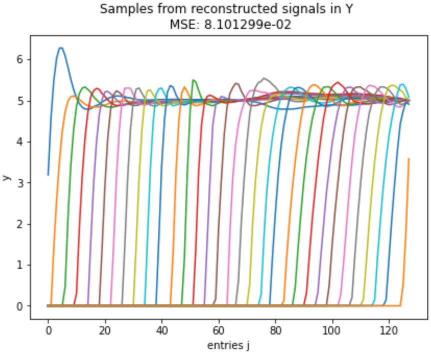
- Linear encoder : $z^{(i)} = \alpha i + \beta$
- RMSProp optimizer with a learning rate of 10^{-5}
- Number of epochs: 2500000

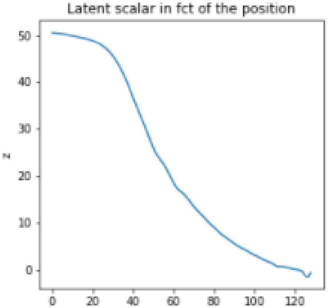




Trained AE - No constraints

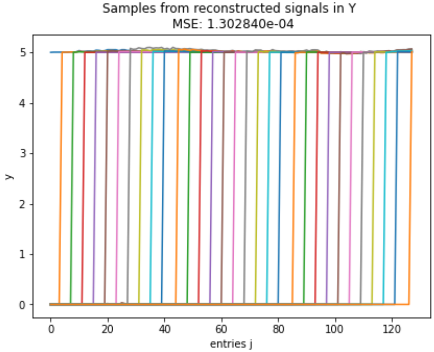
- RMSProp optimizer with a learning rate of 10^{-4}
- Number of epochs: 2500000

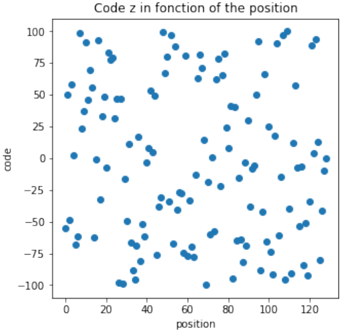




Handcrafted Solution - Random Encoder

- Random Encoder
- Sort $z^{(i)}$ then find a solution





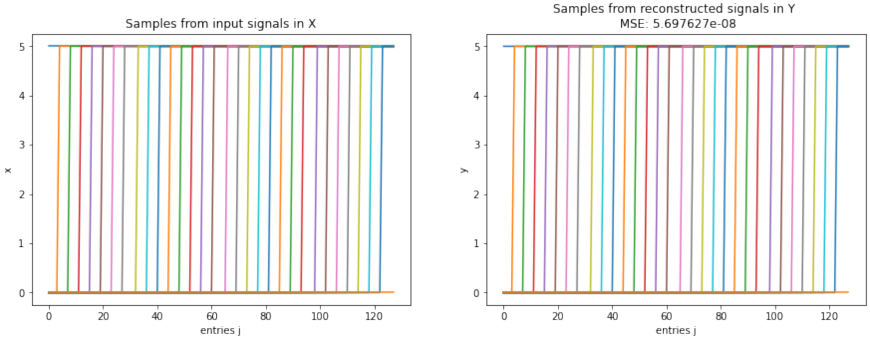
Set of solutions

Lemma

If $(W^{[2]}, \mathbf{b}^{[2]}, W^{[3]})$ is a solution for a decoder with a perfect reconstruction, then for all positive dioganal matrices D, $(DW^{[2]}, \mathbf{b}^{[2]}D, W^{[3]}D^{-1})$ is also a solution.

Generated Solutions

- $(W^{[2]}, \mathbf{b}^{[2]}, W^{[3]})$ from handcrafted solution with linear encoder
- D random positive diagonal matrix



Set of solutions

- $\mathbf{k} = [k_1, k_2, ..., k_n] = [\frac{\mathbf{b}_1^{[2]}}{W_1^{[2]}}, \frac{\mathbf{b}_2^{[2]}}{W_2^{[2]}}, ..., \frac{\mathbf{b}_n^{[2]}}{W_n^{[2]}}]^T$

Proposition

Let $\mathcal{U}$ be the set of all $U^{[3]}$ having the form (1). And let $\mathcal{K}$ be the set of all $\mathbf{k}$ that make the activation matrix of layer 2 full rank. For each fixed couple $\left(\mathbf{k}, U^{[3]}\right) \in \mathcal{K} \times \mathcal{U}$, there exists a set of infinitely many solutions described by the two following equations,

$$W^{[2]} = {W^{[3]}}^{-1} {W^{[2]}}^*$$

$$\mathbf{b}^{[2]} = {\mathbf{b}^{[2]}}^* {{W^{[3]}}^{-1}}^T$$

where ${W^{[2]}}^*$ is a constant matrix of dimension $n \times 1$ and ${b^{[2]}}^*$ is a constant vector of dimension n.

Conclusion

Thanks For Your Attention!

Any Questions?