Insper

Computer Vision

Class 12: The Simplicity and Power of Non-Linearity

Dense layer as affine transformation:

matrix multiplication and array addition

$$\underbrace{\begin{bmatrix} \alpha_{0,0} & \alpha_{0,1} & \alpha_{0,2} & \cdots \\ \alpha_{1,0} & \alpha_{1,1} & \alpha_{1,2} & \cdots \\ \alpha_{2,0} & \alpha_{2,1} & \alpha_{2,2} & \cdots \\ & \vdots & & \\ \alpha_{n-1,0} & \alpha_{n-1,1} & \alpha_{n-1,2} & \cdots \end{bmatrix}}_{A} \cdot \underbrace{\begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_{m-1} \end{bmatrix}}_{x} + \underbrace{\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{n-1} \end{bmatrix}}_{b} = \underbrace{\begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_{n-1} \end{bmatrix}}_{y}$$

Affine transformation as linear transformation:

matrix multiplication

$$
\begin{bmatrix}
\alpha_{0,0} & \alpha_{0,1} & \alpha_{0,2} & \cdots & \beta_0 \\
\alpha_{1,0} & \alpha_{1,1} & \alpha_{1,2} & \cdots & \beta_1 \\
\alpha_{2,0} & \alpha_{2,1} & \alpha_{2,2} & \cdots & \beta_2 \\
 & \vdots & & & \vdots \\
\alpha_{n-1,0} & \alpha_{n-1,1} & \alpha_{n-1,2} & \cdots & \beta_{n-1}
\end{bmatrix}
\cdot
\begin{bmatrix}
x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_{m-1} \\ 1
\end{bmatrix}
=
\begin{bmatrix}
y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_{n-1}
\end{bmatrix}
$$

Linear transformations

A function f is *linear* if…

- $f(x_1 + x_2) = f(x_1) + f(x_2)$; *(additivity)*
- $f(\lambda x) = \lambda f(x)$ *(homogeneity)*

…and this is true if there exists a matrix M such that $f(x) = Mx$.

Example *(Class 3)*

The gray level conversion $f(c) = 0.299\, c_r + 0.587\, c_g + 0.114\, c_b$ is linear:

- $$\begin{aligned} f(c + d) &= (0.299\,(c_r + d_r) + 0.587\,(c_g + d_g) + 0.114\,(c_b + d_b)) \\ &= (0.299\, c_r + 0.587\, c_g + 0.114\, c_b) + (0.299\, d_r + 0.587\, d_g + 0.114\, d_b) \\ &= f(c) + f(d); \end{aligned}$$

- $$\begin{aligned} f(\lambda c) &= (0.299\, \lambda c_r + 0.587\, \lambda c_g + 0.114\, \lambda c_b) \\ &= \lambda(0.299\, c_r + 0.587\, c_g + 0.114\, c_b) \\ &= \lambda f(c). \end{aligned}$$

Examples:

- gray level conversion; *(Class 3)*
- brightness adjustment; *(Class 4)*
- contrast adjustment; *(Class 4)*
- probability dark or bright; *(Class 5)*
- convolution; *(Classes 6-8)*
- geometric transformations. *(Class 10)*

Counterexample *(Class 3)*

The HSV conversion, which uses $f(c) = \max(c_r, c_g, c_b)$, is not linear:

- $f((1, 2, 3) + (5, 3, 1)) = \max(1 + 5, 2 + 3, 3 + 1) = \max(6, 5, 4) = 6;$
- $f((1, 2, 3)) + f((5, 3, 1)) = \max(1, 2, 3) + \max(5, 3, 1) = 3 + 5 = 8.$

Counterexamples:

- boolean dark or bright; *(Class 1)*
- template matching; *(Class 1)*
- HSL conversion; *(Class 3)*
- Harris-Stephens detector. *(Class 9)*

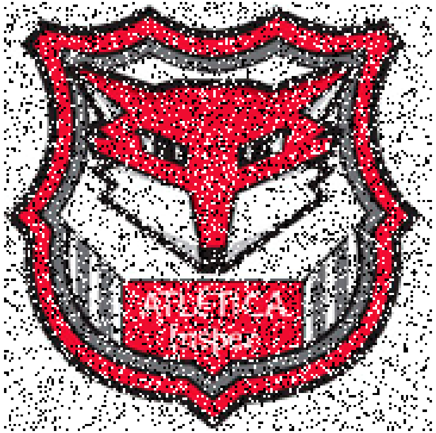
Non-linearity can also give some alternatives to linear algorithms.

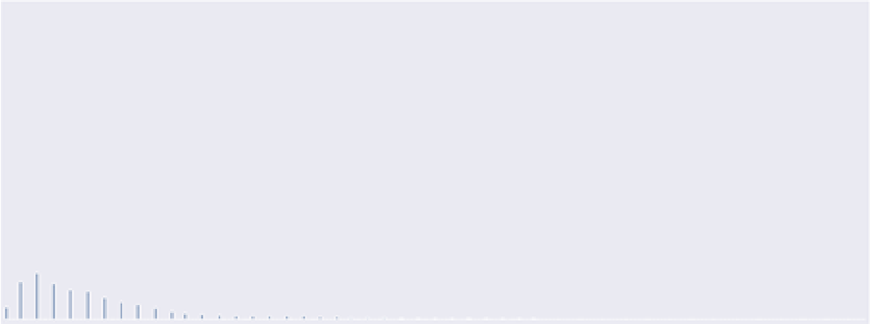
So let's revisit previous problems.

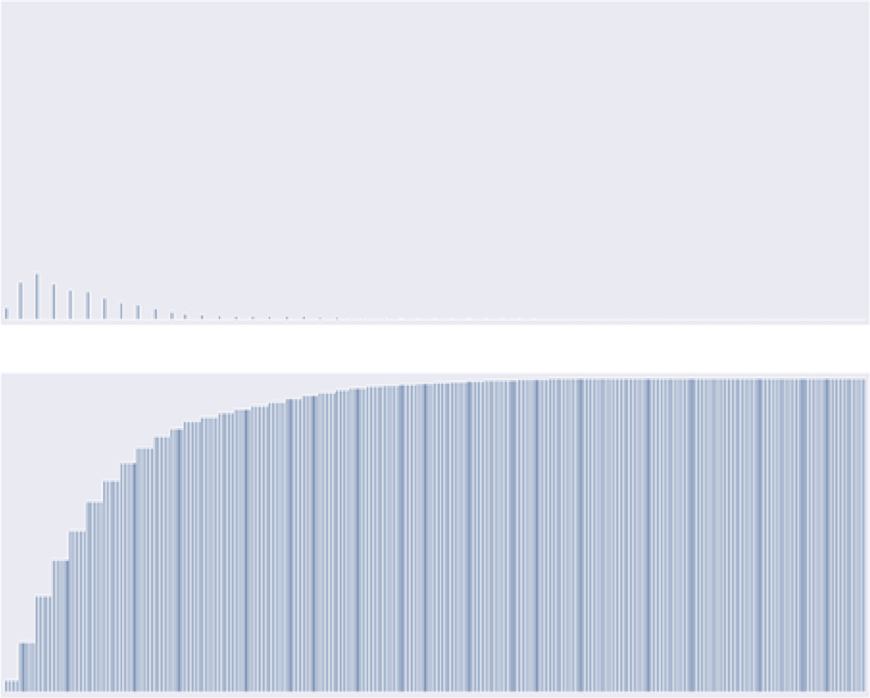
Gaussian noise:

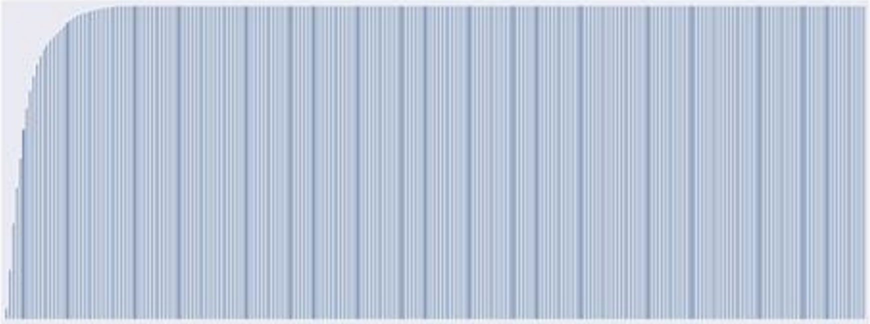
variations are rarely extreme

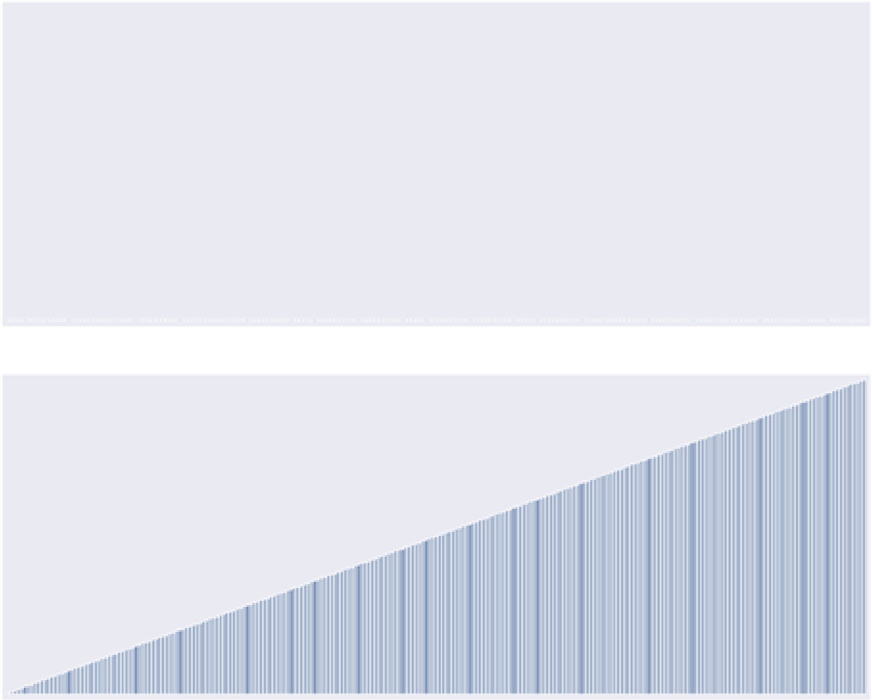
Impulse noise:
variations are always extreme

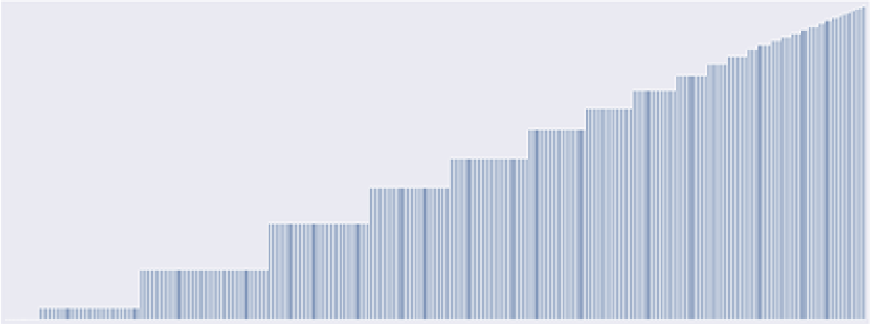












handout

Toolkit

- **Language:** Mathematics
- **Library:** None
- **Platform:** Pencil and Paper

Instructions

1. Organize in groups of 2 or 3 members. No more, no less.

1. Take a copy of the handout, read it, and do the activities.

1. Register the member names and return it to the professor.

Neural network mysteries

- How can we separate a dataset in training data and testing data?
- ~~What is a hidden layer?~~
- ~~Does it matter if the input tensors are 3D, 2D, or 1D?~~
- ~~What is a dense layer?~~
- ~~How do we calculate the number of parameters in a dense layer?~~
- ~~How do we calculate the number of steps in a training process?~~
- ~~How do we calculate the number of steps in a testing process?~~
- Does the number of layers matter?
- ~~Does the size of a dense layer matter?~~
- What exactly `activation='relu'` does?
- ~~Can classic vision algorithms be represented by dense layers?~~
- ~~In a dense layer, are all the weights and biases always necessary?~~
- ~~How do we calculate the number of parameters in a convolutional layer?~~

Deep learning is not simply about adding extra layers.

Deep learning is about ensuring that the extra layers are relevant.

Next class:

- non-linear classic vision.

Credits

This material was based on the work of other professors, listed below.

- Fabio Miranda (fabiomiranda@insper.edu.br)
- Raul Ikeda (RaulIGS@insper.edu.br)
- Fabio Ayres (FabioJA@insper.edu.br)
- Igor Montagner (IgorSM1@insper.edu.br)
- Andrew Kurauchi (AndrewTNK@insper.edu.br)
- Luciano Silva (LucianoS4@insper.edu.br)
- Tiago Sanches (tiagoss4@insper.edu.br)

Well, except for the errors. Any errors you might find are probably my fault.

Images

https://en.wikipedia.org/wiki/Histogram_equalization
https://acidcow.com/pics/15485-mindfuck-pictures-74-pics.html
https://en.wikipedia.org/wiki/Thresholding_(image_processing)