PSO and Neural Network training

Francesco Marson

[1] Department of Computer Science
University of Geneva

Metaheuristics for Optimization

Goal

Goal

- Train the Neural Network (NN)
  - find the weights of the links to have the following behavior:
    1. input: image of a "2" $\Rightarrow$ output: $h \to 1$
    2. input: image of a "3" $\Rightarrow$ output: $h \to 0$
- Use the Particle swarm optimization (PSO)
  - the optimal position of a PSO particle in the research space should tell us the optima coefficients for the NN

The Neural Network

Graphical representation

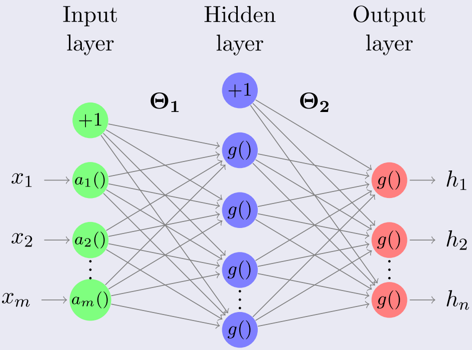

FIGURE 1 – A 3-layers neural network. This one has several outputs at the last level, while the one considered in the exercise has only one. In our case we also set $a_i(x_i) = x_i$.

The Neural Network

Matrix representation of the NN

$$\boldsymbol{x} = \begin{bmatrix} x_1 \in [0,1] \\ \vdots \\ x_m \end{bmatrix} \rightarrow \text{one line of X.dat}$$

$$\boldsymbol{z} := \boldsymbol{g}\left(\underbrace{\overbrace{\begin{bmatrix} \theta^{(1)}_{1,1} & \cdots & \\ \vdots & \ddots & \\ & & \theta^{(1)}_{s,(m+1)} \end{bmatrix}}^{\boldsymbol{\Theta_1}} \cdot \begin{bmatrix} 1 \\ x_1 \\ \vdots \\ x_m \end{bmatrix}}_{\boldsymbol{c}} \right) = \begin{bmatrix} g\left(c_1\right) \\ \vdots \\ g\left(c_s\right) \end{bmatrix} = \begin{bmatrix} z_1 \\ \vdots \\ z_s \end{bmatrix}$$

$$h_{\boldsymbol{\Theta_1},\boldsymbol{\Theta_2}}(\mathbf{x}) := \boldsymbol{g}\left(\overbrace{\begin{bmatrix} \theta^{(2)}_{1,1} & \cdots & \theta^{(2)}_{1,(s+1)} \end{bmatrix}}^{\boldsymbol{\Theta_2}} \cdot \begin{bmatrix} 1 \\ z_1 \\ \vdots \\ z_s \end{bmatrix} \right)$$

From coefficients to PSO particle

consider the particle i of the PSO, we want to connect its position $\boldsymbol{s}_t^i$ to the coefficients of $\boldsymbol{\Theta_1}, \boldsymbol{\Theta_2}$ that are made as follows

$$\theta_{i,j}^{(1)} \in \boldsymbol{\Theta_1} \quad \text{and} \quad \theta_{1,k}^{(2)} \in \boldsymbol{\Theta_2}$$

then the suggestion (but not the unique way) is to construct $\boldsymbol{s}_t^i$ as follow:

$$\boldsymbol{s}^i = [\theta_{1,1}^{(1)}, \ldots, \theta_{1,m+1}^{(1)}, \theta_{2,m+1}^{(1)}, \ldots, \ldots, \theta_{s,(m+1)}^{(1)}, \theta_{1,1}^{(2)}, \ldots, \theta_{1,(s+1)}^{(2)}]$$

where in our case $m = 400$ $s = 25$ and $\boldsymbol{s}_t^i \in R^{s\times(m+1)+s+1}$

PSO optimization

Move the PSO particle

$$
\begin{aligned}
\mathbf{v}^i_{t+1} &= \omega \mathbf{v}^i_t + c_1 r_1 (\mathbf{b}^i_t - \mathbf{s}^i_t) + c_2 r_2 (\mathbf{b}^G_t - \mathbf{s}^i_t) \\
\mathbf{s}^i_{t+1} &= \mathbf{s}^i_t + \mathbf{v}^i_{t+1}
\end{aligned}
$$

Compute the fitness to find the best positions

$$J_k(\mathbf{\Theta_1}, \mathbf{\Theta_2}) := (y_k - h_{\mathbf{\Theta_1},\mathbf{\Theta_2}}(\mathbf{x}_k))^2$$

$$J(\mathbf{\Theta_1}, \mathbf{\Theta_2}) := \frac{1}{m} \sum_{k=1}^{m} (y_k - h_{\mathbf{\Theta_1},\mathbf{\Theta_2}}(\mathbf{x}_k))^2$$

The continuum research space and its boundaries

Limit the range of values for the weights

- The matrix coefficients are *weights* $\rightarrow$ it is not important their absolute value but the relative magnitude
- fix boundaries in the continuum space $\rightarrow \quad \theta_i \in [-0.5, 0.5]$
- $\Rightarrow \quad S \subset R^{s\times(m+1)+s+1}$
- $\Rightarrow$ the research space is a hypercube!

Problems

- define what to do when the progression formula of the PSO "pushes" a particle out of the boundaries
- define velocities cutoff for the particles compatible with the defined S