Ch 10.1: Neural Nets

Lecture 28 - CMSE 381

Prof. Elizabeth Munch

Michigan State University
::
Dept of Computational Mathematics, Science & Engineering

Mon, Nov 28, 2022

Announcements

Last time:

- SVM

This lecture:

- Feed Forward Neural Nets

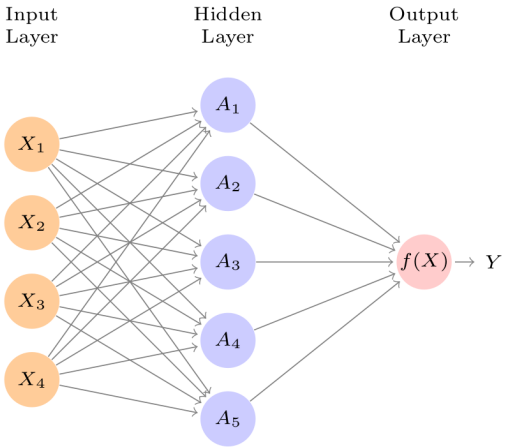
Announcements:

Section 1

Neural Nets

The idea

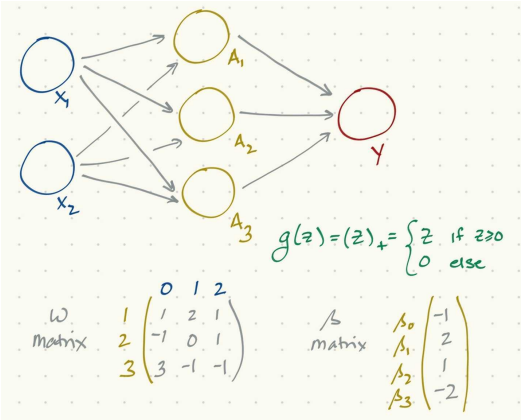
Feed Forward Neural Network: The cartoon



A very simple example

Computing A_k for $(1, 0)$

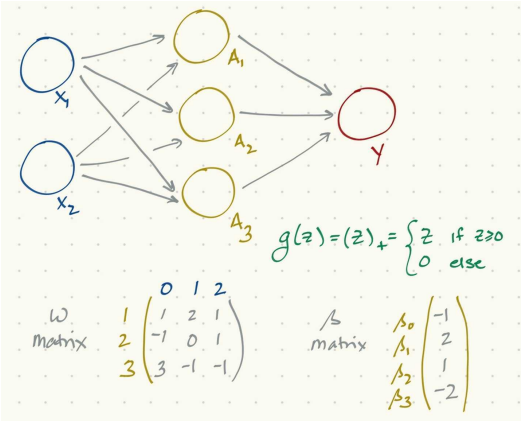
$$A_k = h_k(X) = g(w_{k0} + \sum_{j=1}^{p} w_{kj} X_j),$$



A very simple example

Computing Y for $(1, 0)$

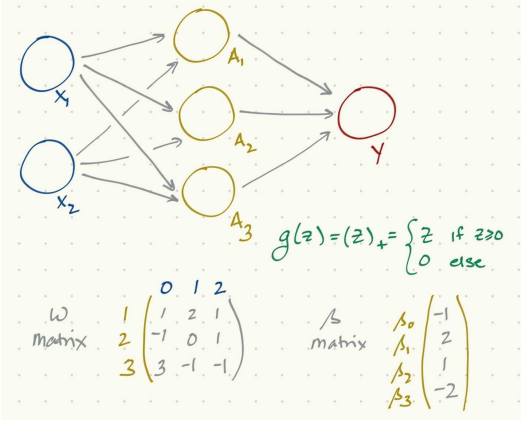
$$f(X) = \beta_0 + \sum_{k=1}^{K} \beta_k A_k$$



A very simple example

Computing Y for $(0, 1)$

$$A_k = h_k(X) = g(w_{k0} + \sum_{j=1}^{p} w_{kj} X_j), \qquad f(X) = \beta_0 + \sum_{k=1}^{K} \beta_k A_k$$



A different example

- Draw the diagram for a neural net with input data points with $p = 3$ (i.e., (X_1, X_2, X_3)) and two units in the hidden layer.
- Using the ω and β matrices, what is the output predicted Y for the point $(2, 0, 1)$?

$$\omega = \begin{pmatrix} 1 & 0 & -2 \\ 3 & 1 & 0 \end{pmatrix} \qquad \beta = \begin{pmatrix} -1 \\ 2 \\ 0 \\ 1 \end{pmatrix}$$

- Use the activation function

$$g(z) = (z)_+ = \begin{cases} 0 & \text{if } z < 0 \\ z & \text{else.} \end{cases}$$

Extra space

Choices for activation function

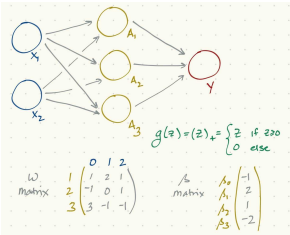
Sigmoid:

$$g(z) = \frac{e^z}{1+e^z} = \frac{1}{1+e^{-z}}$$

ReLU: Rectified linear unit

$$g(z) = (z)_+ = \begin{cases} 0 & \text{if } z < 0 \\ z & \text{else.} \end{cases}$$

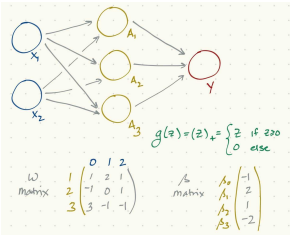
Matrix version: First layer



$$A_k = h_k(X) = g(w_{k0} + \sum_{j=1}^{p} w_{kj} X_j),$$

$$A^{(1)} = g(\mathbf{W}^{(1)} \cdot \mathbf{X}) \qquad \mathbf{X}^T = (1\ X_1\ X_2\ \cdots\ X_p)$$

Matrix version: Output



$$f(X) = \beta_0 + \sum_{k=1}^{K} \beta_k A_k$$

$$Y = \beta \cdot \mathbf{A}^{(1)} \qquad (\mathbf{A}^{(1)})^T = (1 \; A_1^{(1)} \; A_2^{(1)} \; \cdots \; A_K^{(1)})$$

Now what?

Choose parameters by minimizing RSS, $\sum_{i=1}^{n}(y_i - f(x_i))^2$

Chosen in advance:

Tuned by the model:

Coding

Next time

20	F	Nov 4	Polynomial & Step Functions.	7.1,7.2	
21	M	Nov 7	Step Functions	7.2	
22	W	Nov 9	Basis functions, Regression Splines	7.3,7.4	
23	F	Nov 11	Decision Trees	8.1	HW #7 Due
24	M	Nov 14	Random Forests	8.2.1, 8.2.2	
25	W	Nov 16	Maximal Margin Classifier	9.1	
26	F	Nov 18	SVC	9.2	HW #8 Due
27	M	Nov 21	SVM	9.3, 9.4, 9.5	
28	W	Nov 23	Extended virtual office hours		
	F	Nov 25	No class - Thanksgiving		
29	M	Nov 28	Single layer NN	10.1	HW #9 Due
30	W	Nov 30	Multi Layer NN	10.2	
31	F	Dec 2	CNN	10.3	
32	M	Dec 5	Unsupervised Learning & Clustering	12.1, 12.4	HW #10 Due
	W	Dec 7	Review		
	F	Dec 9	Midterm #3	Bring your cheat sheet and a non-internet-connected calculator	