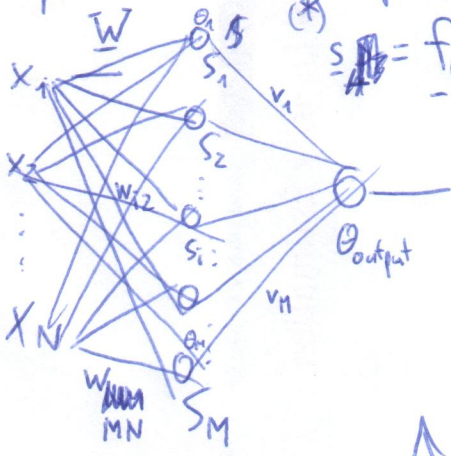
Multilayer Perceptron (MLP) Intro

d) What is a feedforward MLP?

Horizontally stacked vertical columns of connectionist neurons.

Example: scalar output, N-dim. input, 1 hidden layer



(*) $\underline{s} = f_{hidden}(\underline{W}\underline{x} - \underline{\theta})$ $\quad \hat{y} = f_{output}(\underline{v}^T \underline{s} - \theta_{output})$

$= f_{output}(\underline{v}^T f_{hidden}(\underline{W}\underline{x} - \underline{\theta}))$

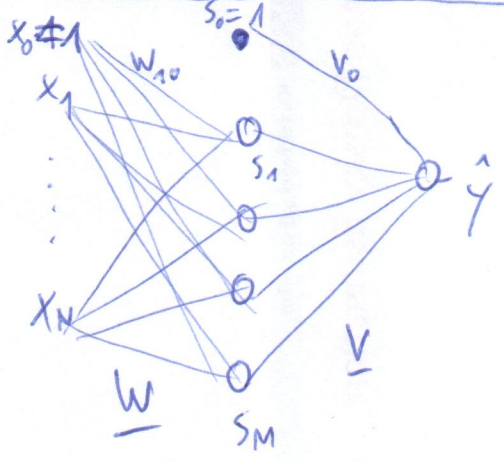
with $\underline{f}_{hidden}(\underline{h}) = \begin{pmatrix} f_{hidden}(h_1) \\ \vdots \\ f_{hidden}(h_M) \end{pmatrix}$

input | hidden layer | output layer $\hat{=}$ 2 layers

Deep neural net $\hat{=}$ MLP with many hidden layers (and typically sparse weight matrices)

(*) ~~$s_i = f_{hidden}(\sum_{j=1}^{N} w_{ij} x_j - \theta_i)$~~ $\quad s_i = f_{hidden}\left(\sum_{j=1}^{N} w_{ij} x_j - \theta_i\right)$

Lecture: bias parameter absorbed into weights $\Rightarrow$ dimension increased by 1 per hidden layer & input



with $w_{i0} = -\theta_i$

$v_0 = -\theta_{output}$

$\hat{y} = f_{output}(\underline{v}^T \underline{s})$

$= f_{output}(\underline{v}^T f_{output}(\underline{W}\underline{x}))$

$\underline{x} \in \mathbb{R}^{N+1}$, $\underline{W} \in \mathbb{R}^{M,N+1}$, $\underline{s} \in \mathbb{R}^{M+1}$

e.g. $\underline{x} = (1, x_1, \dots, x_N)^T$ "extended" data point