$$\underset{W}{\text{minimize}} \quad \frac{1}{2}\sum_{j=0}^{N} ||W_H z_H^j - y^j||^2$$

$$\text{subject to} \quad 0 = z_{k+1}^j - \sigma(W_k x_k^j + b_k), \quad k = 0, \ldots, H-1, j = 1, \ldots, N$$

$$\begin{aligned} \min \quad & \frac{1}{2}||F(x)||_2^2 \\ \text{s. t.} \quad & h(x) = 0 \end{aligned}$$

$$\begin{aligned} \mathcal{L}_c(x, \lambda) &= \frac{1}{2}||F(x)||_2^2 + < \lambda, h(x) > + \frac{c}{2}||h(x)||_2^2 \\ &= \frac{1}{2}||F(x)||_2^2 + \frac{c}{2}||h(x) + \lambda/c||_2^2 - \frac{1}{2c}||\lambda||_2^2 \\ &= \frac{c}{2}\left|\left|\begin{bmatrix} F(x)/\sqrt{c} \\ h(x) + \lambda/c \end{bmatrix}\right|\right|^2 \end{aligned}$$

Find

$$x^k \text{ s. t. } \quad ||\nabla_k \mathcal{L}_{c_k}(x^k, \lambda^k)||_2 \leq \epsilon \tag{1}$$

$$\lambda^{k+1} = \lambda^k + c_k h(x^k) \tag{2}$$

$$\begin{aligned} F &= y - (W_3 z_2 + b_3) \\ h_1 &= z_1 - \sigma(W_1 x + b_1) \\ h_2 &= z_2 - \sigma(W_2 z_1 + b_2) \end{aligned} \tag{3}$$

	$\nabla\mathcal{L}$	W_{1_1}	W_{1_2}	W_{1_3}	b_1
		1	1	1	3
F	N	0	0	0	0
h_1	N	$-x\sigma'(W_{1_1}x + b_{1_1})$	0	0	$-\sigma'(W_{1_1}x + b_{1_1})$
	N	0	$-x\sigma'(W_{1_2}x + b_{1_2})$	0	$-\sigma'(W_{1_2}x + b_{1_2})$
	N	0	0	$-x\sigma'(W_{1_3}x + b_{1_3})$	$-\sigma'(W_{1_3}x + b_{1_3})$
h_2	N	0	0	0	0
	N	0	0	0	0
	N	0	0	0	0

	$\nabla\mathcal{L}$	W_{2_1}	W_{2_2}	W_{2_3}	b_2
		3	3	3	3
F	N	0	0	0	0
h_1	N	0	0	0	0
	N	0	0	0	0
	N	0	0	0	0
h_2	N	$-z_1\sigma'(W_{2_1}z + b_{2_1})$	0	0	$-\sigma'(W_{2_1}x + b_{2_1})$
	N	0	$-z_1\sigma'(W_{2_2}z + b_{2_2})$	0	$-\sigma'(W_{2_2}x + b_{2_2})$
	N	0	0	$-z_1\sigma'(W_{2_3}z + b_{2_3})$	$-\sigma'(W_{2_3}x + b_{2_3})$

	$\nabla\mathcal{L}$	W_3	/	/	b_3
		3	0	0	1
F	N	$-z_2/\sqrt{c}$	0	0	$-1/\sqrt{c}$
h_1	N	0	0	0	0
	N	0	0	0	0
	N	0	0	0	0
h_2	N	0	0	0	0
	N	0	0	0	0
	N	0	0	0	0

Square Diagonal Matrixes

	$\nabla\mathcal{L}$	z_{1_1} N	z_{1_2} N	z_{1_3} N
F	N	0	0	0
h_1	N	1	0	0
	N	0	1	0
	N	0	0	1
h_2	N	$-W_{2_{1,1}}\sigma'(W_{2_1}z_1 + b_{2_1})$	$-W_{2_{1,2}}\sigma'(W_{2_1}z_1 + b_{2_1})$	$-W_{2_{1,3}}\sigma'(W_{2_1}z_1 + b_{2_1})$
	N	$-W_{2_{2,1}}\sigma'(W_{2_2}z_1 + b_{2_2})$	$-W_{2_{2,2}}\sigma'(W_{2_2}z_1 + b_{2_2})$	$-W_{2_{2,3}}\sigma'(W_{2_2}z_1 + b_{2_2})$
	N	$-W_{2_{3,1}}\sigma'(W_{2_3}z_1 + b_{2_3})$	$-W_{2_{3,2}}\sigma'(W_{2_3}z_1 + b_{2_3})$	$-W_{2_{3,3}}\sigma'(W_{2_3}z_1 + b_{2_3})$

	$\nabla\mathcal{L}$	z_{2_1} N	z_{2_2} N	z_{2_3} N
F	N	$-W_{3_1}/\sqrt{c}$	$-W_{3_2}/\sqrt{c}$	$-W_{3_3}/\sqrt{c}$
h_1	N	0	0	0
	N	0	0	0
	N	0	0	0
h_2	N	1	0	0
	N	0	1	0
	N	0	0	1

	$\nabla\mathcal{L}$	W_{0_1}	W_{0_2}	...	W_{0_W}	b_1
		I	I	...	I	W
F	O*N	0	0	...	0	0
h_1	N	$-x\sigma'(W_{0_1}x + b_{0_1})$	0	...	0	$-\sigma'(W_{0_1}x +$
	N	0	$-x\sigma'(W_{0_2}x + b_{0_2})$	...	0	$-\sigma'(W_{0_2}x +$
	...	...	...	...	...	...
	N	0	0	...	$-x\sigma'(W_{0_W}x + b_{0_W})$	$-\sigma'(W_{0_W}x +$
h_2	W*N	0	0	...	0	0
...	...	...	...	...	...	...
h_D	W*N	0	0	...	0	0

	$\nabla\mathcal{L}$	W_{i_1}	W_{i_2}	...	W_{i_W}	b_2
		W	W	...	W	W
F	O*N	0	0	...	0	0
h_1	W*N	0	0	...	0	0
...	...	...	...	...	...	...
h_{i+1}	N	$-z_1\sigma'(W_{i_1}z + b_{i_1})$	0	...	0	$-\sigma'(W_{i_1}x +$
	N	0	$-z_1\sigma'(W_{i_2}z + b_{i_2})$	...	0	$-\sigma'(W_{i_2}x +$
	...	...	...	...	...	...
	N	0	0	...	$-z_1\sigma'(W_{i_W}z + b_{i_W})$	$-\sigma'(W_{i_W}x +$
...	...	...	...	...	...	...
h_D	W*N	0	0	...	0	0

	$\nabla\mathcal{L}$	W_{D_1}	W_{D_2}	...	W_{D_O}	b_3
		W	W	...	W	O
F	N	$-z_D\sigma'_O(W_{D_1}x + b_{D_1})/\sqrt{c}$	0	...	0	$-\sigma'_O(W_{D_1}x + b_I$
	N	0	$-z_D\sigma'_O(W_{D_2}x + b_{D_2})/\sqrt{c}$	...	0	$-\sigma'_O(W_{D_2}x + b_I$
	...	...	...	...	...	...
	N	0	0	...	$-z_D\sigma'_O(W_{D_O}x + b_{D_O})/\sqrt{c}$	$-\sigma'_O(W_{D_O}x + b_I$
h_1	W*N	0	0	...	0	0
...	...	...	...	...	...	...
h_D	W*N	0	0	...	0	0

Square Diagonal Matrixes

	$\nabla\mathcal{L}$	z_{i_1}	z_{i_2}	...	z_{i_W}
		N	N	...	N
F	O*N	0	0	...	0
h_1	W*N	0	0	...	0
...	...	...	...	...	...
h_i	N	1	0	...	0
	N	0	1	...	0
	...	...	...	...	...
	N	0	0	...	1
h_{i+1}	N	$-W_{i_{1,1}}\sigma'(W_{i_1}z_i + b_{i_1})$	$-W_{i_{1,2}}\sigma'(W_{i_1}z_i + b_{i_1})$	...	$-W_{i_{1,W}}\sigma'(W_{i_1}z_i + b_{i_1})$
	N	$-W_{i_{2,1}}\sigma'(W_{i_2}z_i + b_{i_2})$	$-W_{i_{2,2}}\sigma'(W_{i_2}z_i + b_{i_2})$	...	$-W_{i_{2,W}}\sigma'(W_{i_2}z_i + b_{i_2})$
	...	...	...	...	...
	N	$-W_{i_{W,1}}\sigma'(W_{i_W}z_i + b_{i_W})$	$-W_{i_{W,2}}\sigma'(W_{i_W}z_i + b_{i_W})$	...	$-W_{i_{W,W}}\sigma'(W_{i_W}z_i + b_{i_W})$
...	...	...	...	...	...
h_D	W*N	0	0	...	0

	$\nabla\mathcal{L}$	z_{D_1}	z_{D_2}	...	z_{D_W}
		N	N	...	N
F	N	$-W_{D_{1,1}}\sigma'_O(W_{D_1}z_D + b_{D_1})$	$-W_{D_{1,2}}\sigma'_O(W_{D_1}z_D + b_{D_1})$	...	$-W_{D_{1,W}}\sigma'_O(W_{D_1}z_D + b_{D_1})$
	N	$-W_{D_{2,1}}\sigma'_O(W_{D_2}z_D + b_{D_2})$	$-W_{D_{2,2}}\sigma'_O(W_{D_2}z_D + b_{D_2})$	...	$-W_{D_{2,W}}\sigma'_O(W_{D_2}z_D + b_{D_2})$
	...	...	...	...	...
	N	$-W_{D_{O,1}}\sigma'_O(W_{D_O}z_D + b_{D_O})$	$-W_{D_{O,2}}\sigma'_O(W_{D_O}z_D + b_{D_O})$	...	$-W_{D_{O,W}}\sigma'_O(W_{D_O}z_D + b_{D_O})$
h_1	W*N	0	0	...	0
...	...	...	...	...	...
h_D	N	1	0	...	0
	N	0	1	...	0
	...	...	...	...	...
	N	0	0	...	1