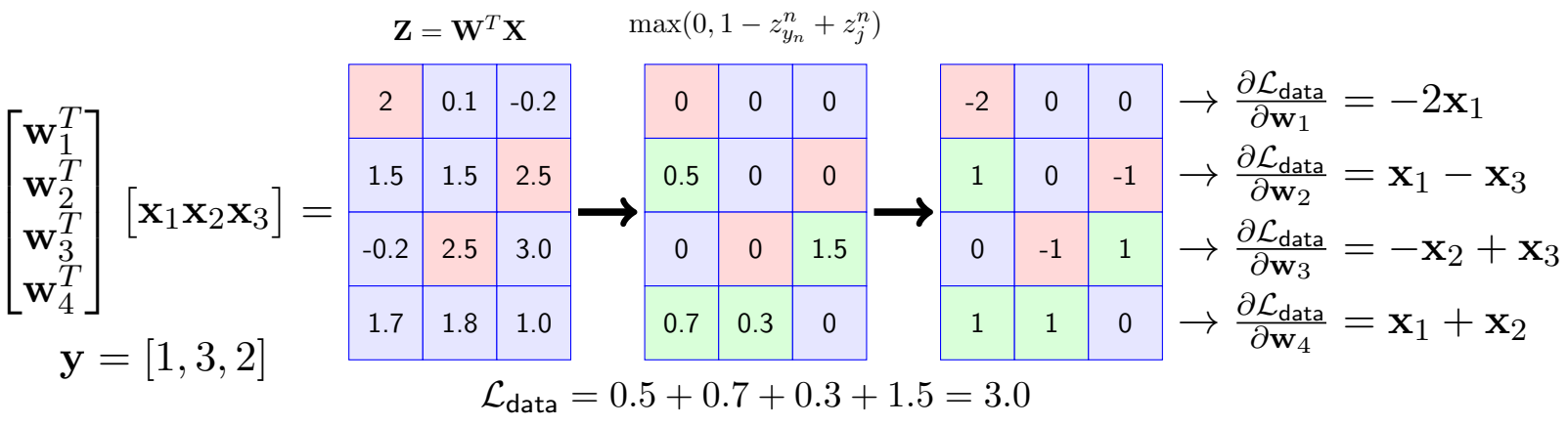
$\mathbf{Z} = \mathbf{W}^T\mathbf{X}$
$\max(0, 1 - z^n_{y_n} + z^n_j)$
$\begin{bmatrix} \mathbf{w}_1^T \\ \mathbf{w}_2^T \\ \mathbf{w}_3^T \\ \mathbf{w}_4^T \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \mathbf{x}_2 \mathbf{x}_3 \end{bmatrix} =$
2 0.1 -0.2
1.5 1.5 2.5
-0.2 2.5 3.0
1.7 1.8 1.0
0 0 0
0.5 0 0
0 0 1.5
0.7 0.3 0
-2 0 0
1 0 -1
0 -1 1
1 1 0
$\rightarrow \frac{\partial \mathcal{L}_{\text{data}}}{\partial \mathbf{w}_1} = -2\mathbf{x}_1$
$\rightarrow \frac{\partial \mathcal{L}_{\text{data}}}{\partial \mathbf{w}_2} = \mathbf{x}_1 - \mathbf{x}_3$
$\rightarrow \frac{\partial \mathcal{L}_{\text{data}}}{\partial \mathbf{w}_3} = -\mathbf{x}_2 + \mathbf{x}_3$
$\rightarrow \frac{\partial \mathcal{L}_{\text{data}}}{\partial \mathbf{w}_4} = \mathbf{x}_1 + \mathbf{x}_2$
$\mathbf{y} = [1, 3, 2]$
$\mathcal{L}_{\text{data}} = 0.5 + 0.7 + 0.3 + 1.5 = 3.0$