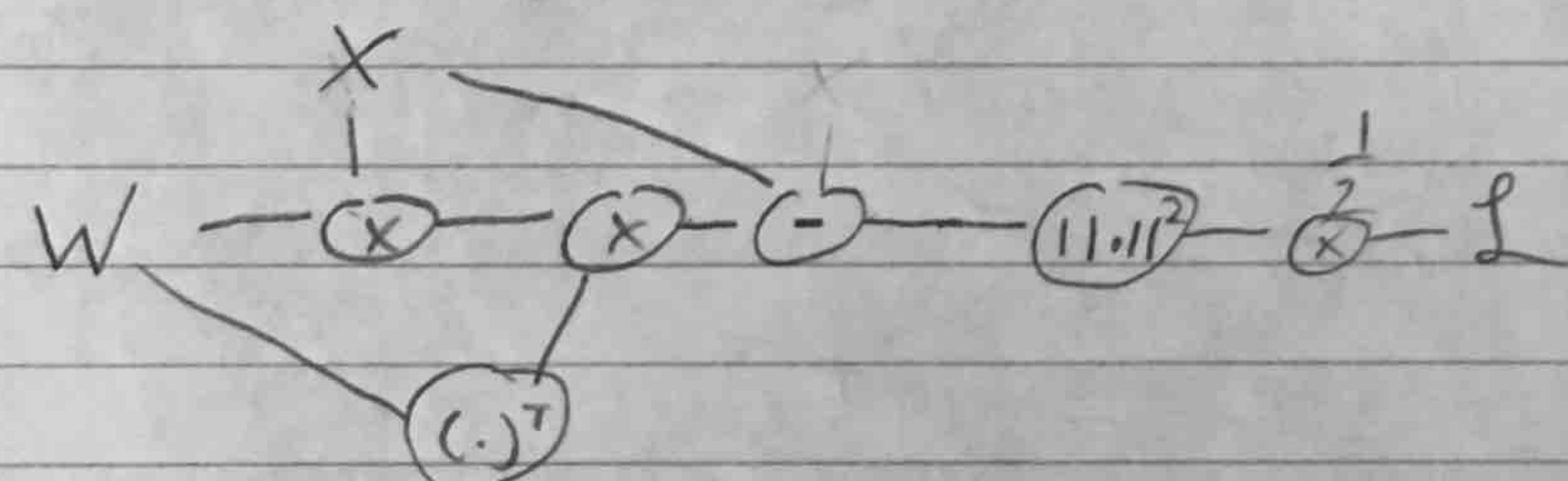
ECE 239AS Hw # 3

1. $\mathcal{L} = \frac{1}{2} \| W^T W x - x \|^2$

a.) If $W \in \mathbb{R}^{m \times n}$, then $W^T W \in \mathbb{R}^{n \times n}$.
Wx represents the dimensionally reduced input data. $W^T W x$ then represents the reconstructed data. Minimizing the difference between this term and the original vector will result in a matrix W that retains information within W.

b.)



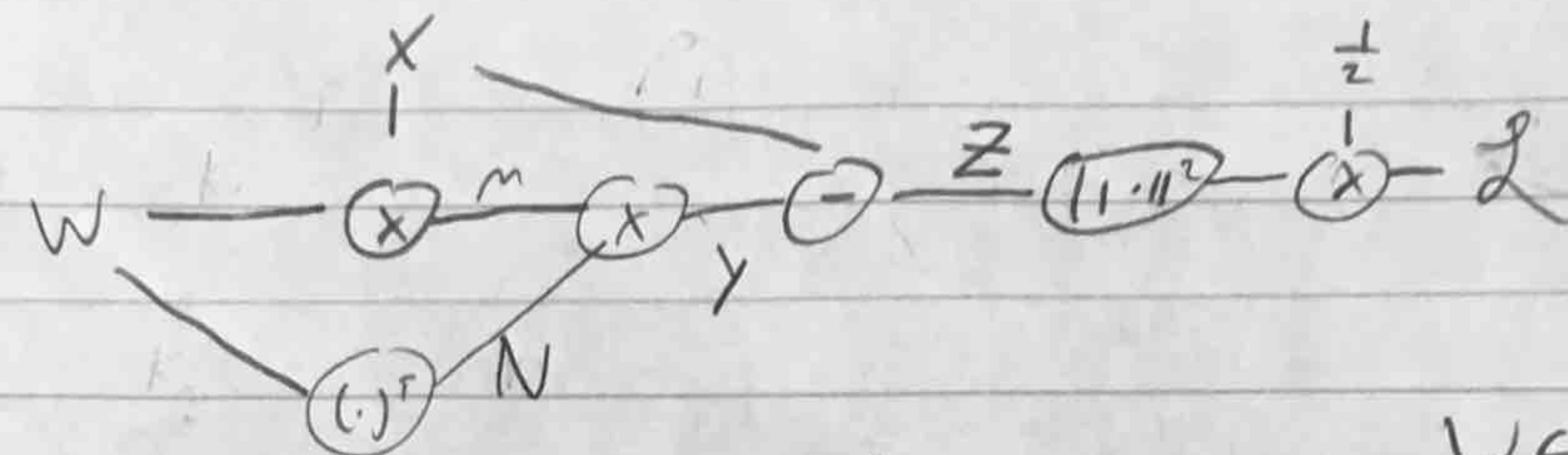
c.) The gradient of $\mathcal{L}$ w.r.t. W must be a summation of all paths leading from W. This is shown by the definition of a "total derivative" which is defined as:

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}$$

where x and y are variables dependent on t.

In this case, the derivative of $\mathcal{L}$ w.r.t W must include all back propagation paths leading to it.

d.



$$\mathcal{L} = \frac{1}{2}\|z\|^2$$
$$z = y - x$$
$$y = Nm$$
$$m = Wx$$
$$N = W^T$$

$$\frac{\partial \mathcal{L}}{\partial z} = z$$
$$\frac{\partial z}{\partial y} = 1$$
$$\frac{\partial y}{\partial m} = N^T, \quad \frac{\partial y}{\partial N} = * m^T$$
$$\frac{\partial m}{\partial W} = * x^T$$
$$\frac{\partial N}{\partial W} = (\cdot)^T$$

$W \in \mathbb{R}^{m\times n}$
$x \in \mathbb{R}^n$
$N = W^T \in \mathbb{R}^{n\times m}$
$z \in \mathbb{R}^n$
$y \in \mathbb{R}^n$
$m \in \mathbb{R}^m$

$$\frac{\partial L}{\partial z} = z$$

$$\frac{\partial L}{\partial y} = \frac{\partial z}{\partial y}\frac{\partial L}{\partial z} = z$$

$$\frac{\partial L}{\partial m} = \frac{\partial y}{\partial m}\cdot\frac{\partial L}{\partial y} = N^T z$$

$$\frac{\partial L}{\partial W_1} = \frac{\partial m}{\partial W_1}\cdot\frac{\partial L}{\partial m} = N^T z \cdot x^T$$

$$\frac{\partial L}{\partial N} = \frac{\partial y}{\partial N}\cdot\frac{\partial L}{\partial y} = z \cdot m^T$$

$$\frac{\partial L}{\partial W_2} = \frac{\partial N}{\partial W}\frac{\partial L}{\partial N} = (z m^T)^T = m z^T$$

$$\frac{\partial L}{\partial W} = \frac{\partial L}{\partial W_1} + \frac{\partial L}{\partial W_2} = N^T z \cdot x^T + m z^T$$

$$= W(y-x)x^T + Wx(y-x)^T$$
$$= W(Nm - x)x^T + Wx(Nm - x)^T$$
$$\boxed{\nabla_W L = W(W^T W x - x)x^T + Wx(W^T W x - x)^T}$$