02/14/2023 Autoencoder Continued

- reconstruction error

for 1 sample and M features:

$$\text{Error } (L) = \frac{1}{M}\sum_{i=1}^{M}(x_i - \hat{x}_i)^2$$

for n samples and M features:

$$\text{Error } (L) = \frac{1}{n}\sum_{j=1}^{n}\left(\frac{1}{M}\sum_{i=1}^{M}(x_i - \hat{x}_i)^2\right)$$

- we want to minimize error

$\min_{\theta} L$ gradient descent:

$$\theta_{j+1} \leftarrow \theta_j - \alpha \nabla_\theta L$$

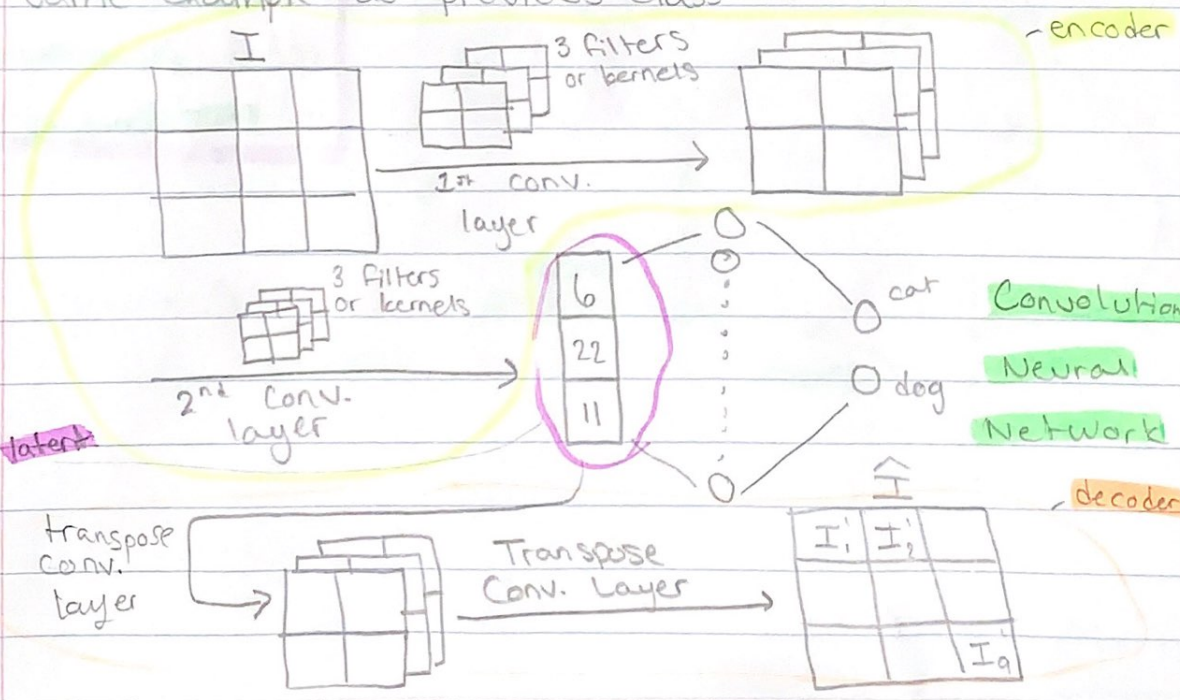
Where

$$\theta = [W_1, b_1, W_2, b_2, W_3, b_3, W_4, b_4]$$

basic example of gradient descent:

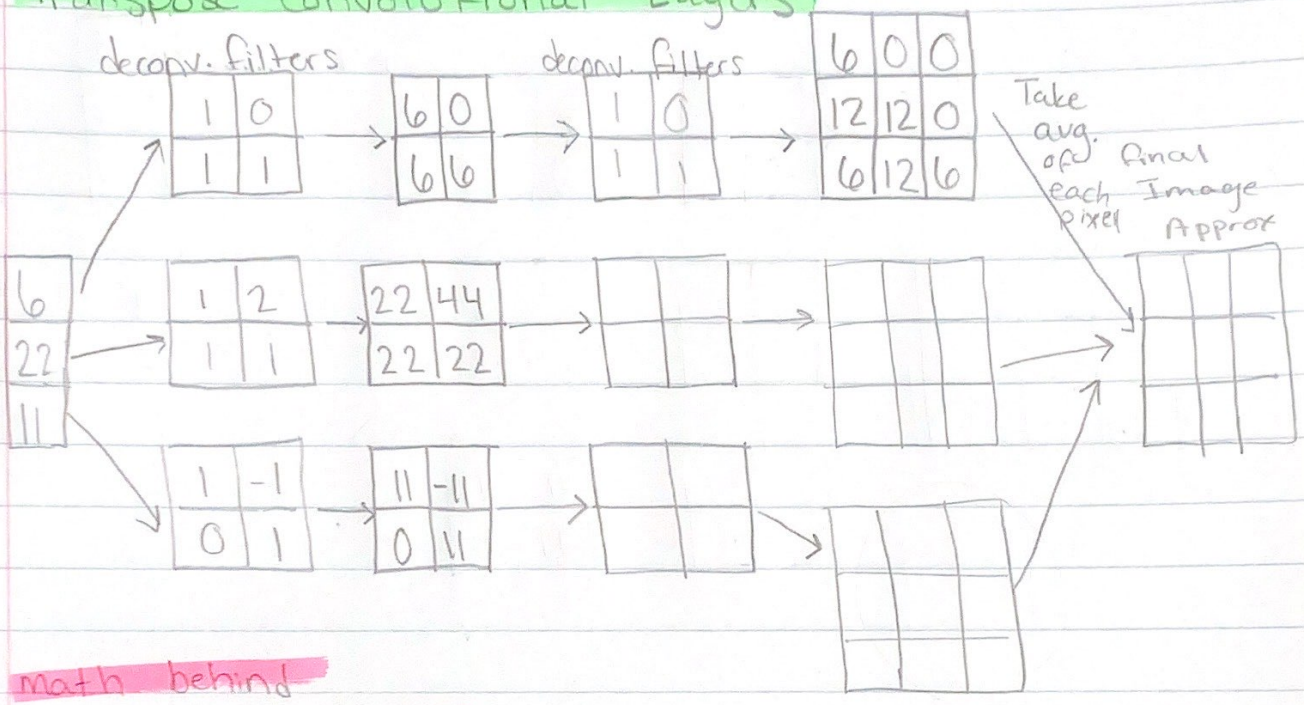
$$\theta = \begin{bmatrix}\theta_1\\ \theta_2\\ \theta_3\end{bmatrix} \qquad \nabla_\theta L = \begin{bmatrix}\frac{dL}{d\theta_1}\\ \frac{dL}{d\theta_2}\\ \frac{dL}{d\theta_3}\end{bmatrix}$$

- same example as previous class



therefore $L = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{9}\sum_{i=1}^{9}\left(I_i - I_i'\right)^2\right)$

Transpose Convolutional Layers



Math behind

Second deconv. layer