Similarly for $\frac{dL}{db}$, we get $\frac{dL}{db} = a^T @ \frac{dL}{dd}$

Grads w.r.t. layer's weights

For $\frac{dL}{dc}$, we get $\frac{dL}{dc} = \frac{dL}{dd}.sum(axis=0)$

Grads w.r.t. bias