$$\nabla_{\theta} L \qquad \nabla \ell \qquad Dg_k \qquad Dg_{k-1} \qquad Dg_2 \qquad Dg_1$$

$$1 \times p \qquad = \qquad 1 \times d_k \qquad d_k \times d_{k-1} \qquad d_{k-1} \times d_{k-2} \qquad \cdots \qquad d_2 \times d_1 \qquad d_1 \times p$$

$$1 \times p \qquad = \qquad 1 \times d_k \qquad d_k \times d_{k-1} \qquad d_{k-1} \times p \qquad \mathcal{O}(kp) \qquad \text{Forward}$$

$$1 \times p \qquad = \qquad \mathcal{O}(k+p) \qquad \qquad \mathcal{O}$$