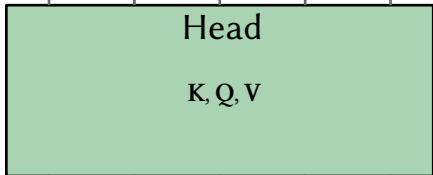


$$\mathbf{q}_0^\top \mathbf{k}_0 \sqrt{d} \mathbf{v}_0$$

$$\mathbf{q}_4^\top \mathbf{k}_0 \sqrt{d} \mathbf{v}_0 + \mathbf{q}_4^\top \mathbf{k}_1 \sqrt{d} \mathbf{v}_1 + \dots + \mathbf{q}_4^\top \mathbf{k}_4 \sqrt{d} \mathbf{v}_4$$



$$\sigma \left( (\mathbf{XQ})(\mathbf{XK})^\top / \sqrt{d} \right)$$

0.6	<del>0.1</del>	<del>0.1</del>	<del>0.1</del>	<del>0.1</del>
0.2	0.4	<del>0.1</del>	<del>0.1</del>	<del>0.2</del>
0.2	0.1	0.5	<del>0.1</del>	<del>0.1</del>
0.3	0.2	0.1	0.3	<del>0.1</del>
0.1	0.1	0.2	0.2	0.4

 $\mathbf{x}_0$ 
 $\mathbf{x}_1$ 
 $\mathbf{x}_2$ 
 $\mathbf{x}_3$ 
 $\mathbf{x}_4$