

$W_K W_Q^T$ in Layer 1, Head 1 of GPT-2

