

source code:
sum_val += multi_head_out[i][k] * W_attn[k][j]

