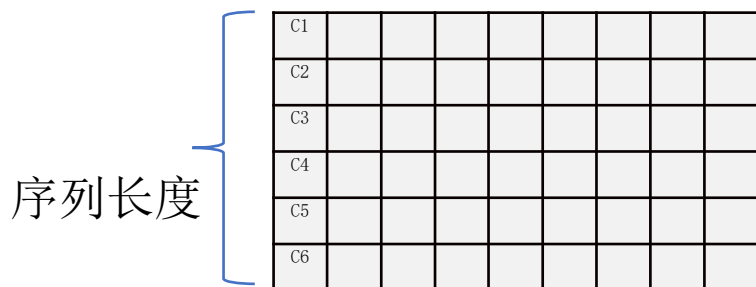
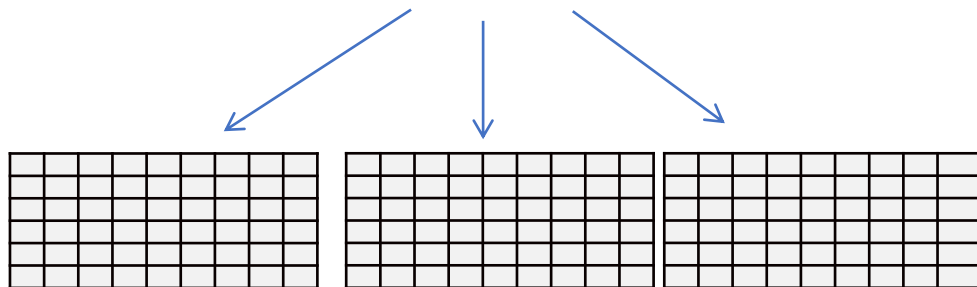


嵌入向量



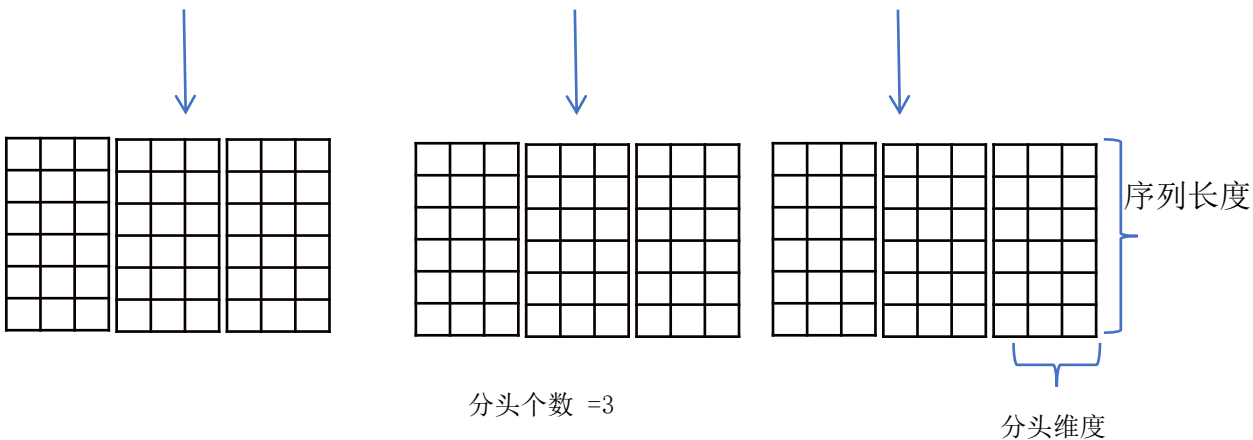
embedding dimension = hidden dimension



Q

K

V



分头维度 = 嵌入维度 / 人头个数