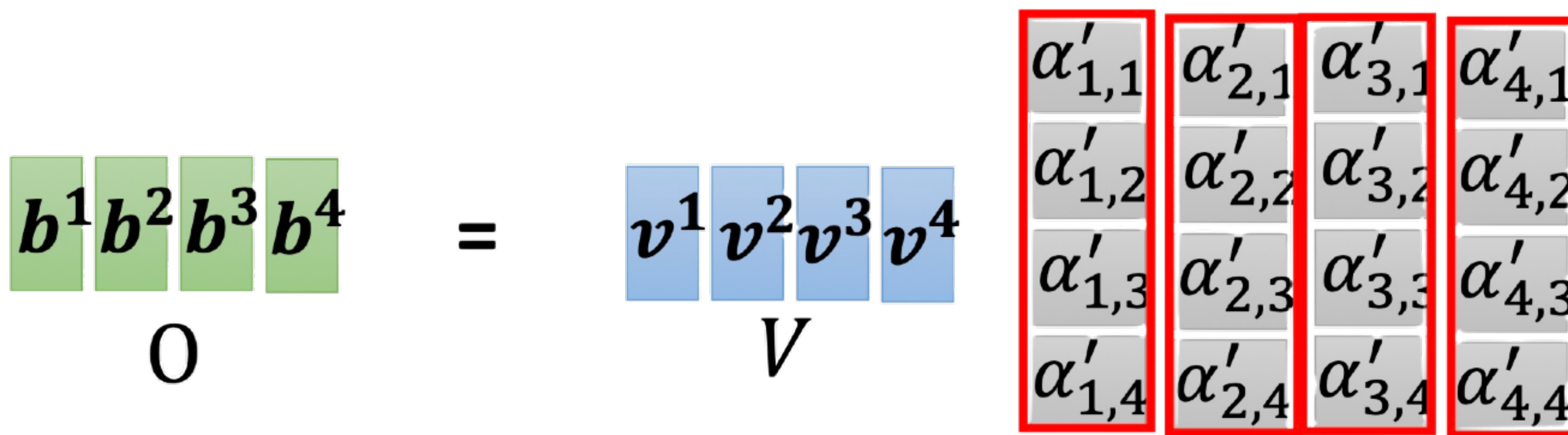
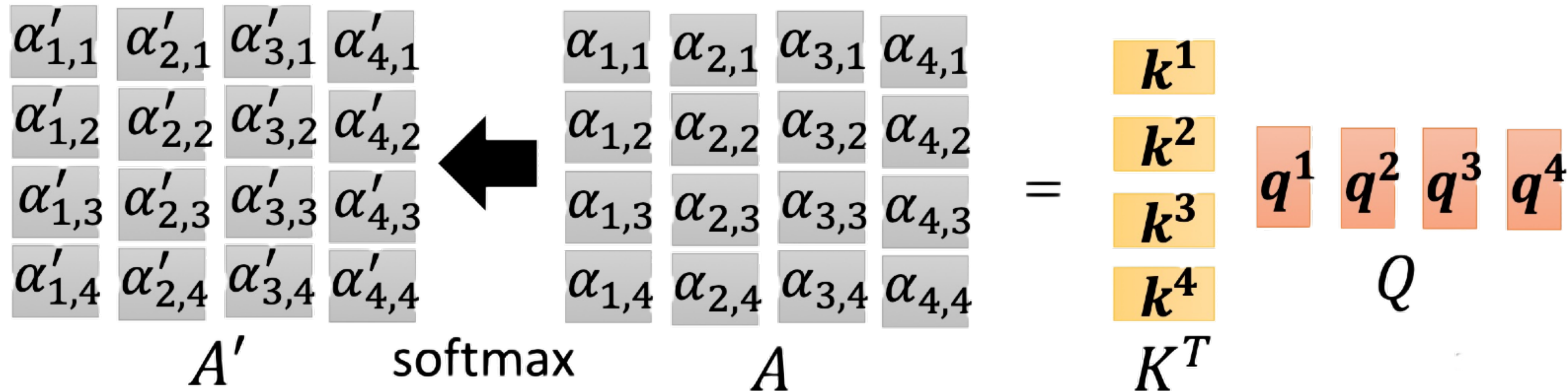


## Attention



# Self-Attention

## Attention:

一般会对两个序列进行注意力计算，具体来说就是Query来源于Target，KV来源于Source。比如下面两句话

target: "你很美"

source: "you are beautiful"

要计算你和source中词的关联性，需要分别计算

score("你", "you")

score("你", "are")

score("你", "beautiful")

**Self-Attention:** source和target是一个序列，比如melonkid是一个小孩，他喜欢吃瓜。他指代的是谁？这里就需要用到自注意力机制了。

