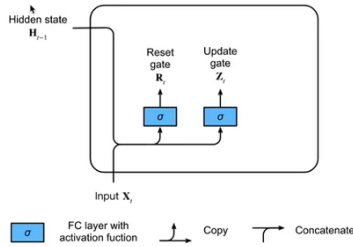


GRU

- 关注一个序列
 - 并不是每个观察值同等重要
 - 想只记住相关的观察需要：能关注的机制（更新门），能遗忘的机制（遗忘门）



$$R_t = \sigma(X_t W_{xr} + H_{t-1} W_{hr} + b_r) \text{ 重置}$$

> 长度与 H 相同

$$Z_t = \sigma(X_t W_{xz} + H_{t-1} W_{hz} + b_z) \text{ 更新}$$

候选隐藏状态

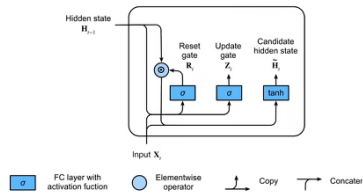
$$\tilde{H}_t = \tanh(X_t W_{xh} + (R_t \odot H_{t-1}) W_{rh} + b_h)$$

↗ 按元素相乘

R_t 控制 H_{t-1} 的权重, 即过去状态

若 $R_t \rightarrow 1$ 则保留全部 H_{t-1}

若 $R_t \rightarrow 0$ 则丢弃 H_{t-1}



隐状态

$$H_t = Z_t \odot H_{t-1} + (1 - Z_t) \odot \tilde{H}_t$$

Z_t 控制现有状态 X_t , 若 $Z_t \rightarrow 1$ 则 $H_t = H_{t-1}$, 丢弃了 X_t

若 $Z_t \rightarrow 0$ 则 $H_t = \tilde{H}_t$ 保留了 X_t 参与运算

