第一问

我们尝试建立SIR模型



由于只关心赶感染人数（在这里等价为确诊人数）。我们对上述微分方程前两式求解。易得



将此公式代入一式，

在带进去一。



这个方程没有解析解。我们选择差分迭代。

可画出图像

问题二模型

在讨论长短期记忆网络之前，需要先介绍以下循环神经网络的架构。

循环神经网络是通过带有子反馈的神经元，能够处理任意长度的时序数据。

给定一个输入序列，，循环神经网络通过下面公式更新带反馈边的隐藏层的活性值：



其中，为一个非线性函数。相比微分方程，循环神经网络也可以看作是一个动力系统，故常把隐藏层的活性值称为状态。

LSTM

普通的循环神经网络存在长程依赖、梯度爆炸或梯度消失的问题。所以针对本问题，我们采用长短期记忆网络（Long Short-Term Memory Network）。

长短期记忆网络是循环神经网络的一个变体。LSTM引入了一个新的内部状态，，专门进行线性的循环信息传递，同时再利用激活函数，非线性地输出，内部状态通过下面公式计算：



其中表示向量元素乘积，为双曲正切函数，其输出的值域区间为。，，为三个门，分别为遗忘门，（信息增强门）输入门，输出门。遗忘门控制当前状态更新是否需要参考过往的记忆（状态）单元。当遗忘门打开则表示需要记忆的状态。输入门控制了当前时刻的候选状态有多少信息需要被保留。输出门控制了当前时刻的内部状态有多少需要输出给外部状态。此外，为上一刻的记忆单元，，是通过非线性函数得到的候选状态：

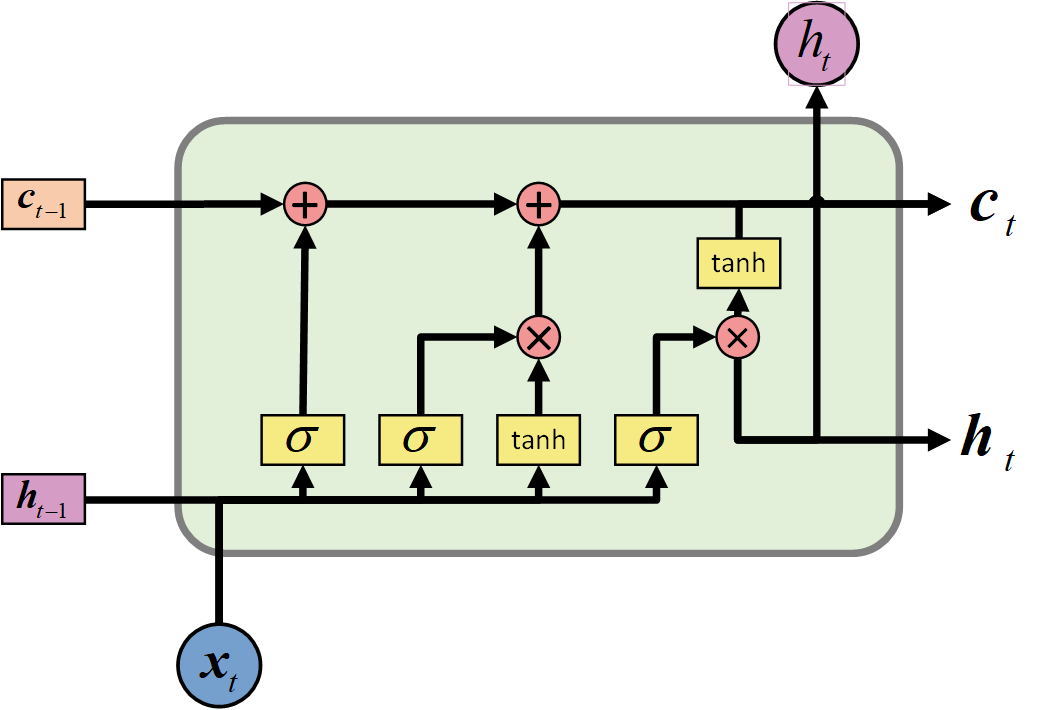


LSTM的门控制与数字电路中的门不同，它是一种软控制。也就是它在0-1之间是连续变化的。具体来说，通常由sigmoid函数实现“软门”。



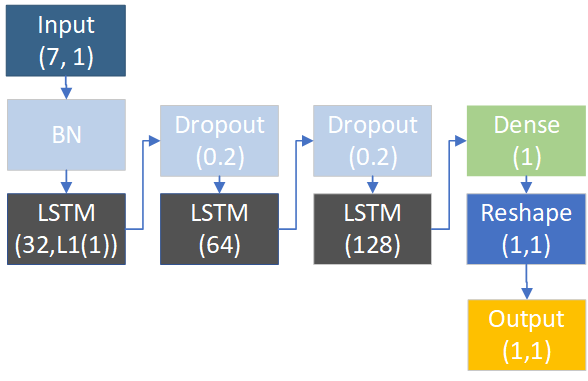
其中便为Sigmoid函数（Logistic函数），其输出区间为。为当前时刻的输入，为上一时刻的外部状态。

以下是LSTM单元的展示



此外，对LSTM网络引入噪声，通过在两个LSTM之间引入Dropout层实现。

具体LSTM网络结构如下



输入形状为(7,1)，表示一个时间窗口含盖7步，选取的特征为死亡数。输出为(1,1)，表示预测每个时间窗口的下一步的死亡数。上图中L1表示正则化约束。Dropout层括号里为丢弃率p，LSTM括号里为神经元按时间展开的步数。

由于LSTM预测必须要求有一个完整的序列，那么需要将每次预测的新时间序列并入原序列以进行未来的序列预测。为了组织数组形状方便，这里我们使用宽度为7的时间窗口，预测7步。也就是说用上一周的数据去预测下一周的数据。

经我们LSTM模型预测，从2021年8月21日后的42天，全世界死亡人数会突破502.3万。