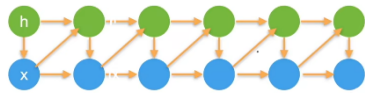


## 序列模型

- 实际中很多数据有时序结构：电影评价随时间变化而变化（评奖信息、季节性、演员导演负面报道、同样题材对比等）、音乐、语音、文本、视频、人的互动、股价预测都是连续的
- 在时间 $t$ 观察随机变量 $X$ ，可以得到 $T$ 个不独立时间变量，服从 $p(x)$ ，如何对 $p$ 建模？使用条件概率展开： $p(a, b) = p(a) \cdot p(b|a)$
- 因此  $p(x) = p(x_1) \cdot p(x_2|x_1) \cdot p(x_3|x_2) \cdots p(x_T|x_{T-1})$
- 可以反向计算，但在物理意义上不一定可行
- 最重要的是对条件概率建模： $p(x_t|x_1 \cdots x_{t-1}) = p(x_t|f(x_1 \cdots x_{t-1}))$ 对见过的数据建模，也称自回归模型
- 方案A 马尔可夫假设：假设当前数据只与过去 $a$ 个数据相关  $p(x_t|x_1 \cdots x_t) = p(x_t|f(x_{t-a}, \cdots, x_t))$  固定长度数据，可训练MLP
- 方案B 潜变量模型：引入潜变量 $h(t)$ 表示过去信息 $f(x_1, \dots, x_t)$ ，这样  $x_t = p(x_t|h_t)$



模型1：已知  $h$  与  $x \Rightarrow h'$

模型2：已知  $x$  与  $h' \Rightarrow x'$