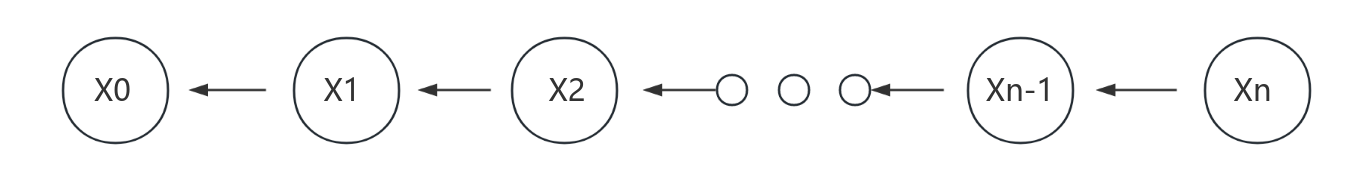
问题形式



求解思路：动态规划



由贝尔曼最优性原理定义cost-to-go funtion



并定义action funtion



V（xk）函数为在xk状态下到最终状态xN所需的最小代价

Q（xk,uk）函数为在xk状态和控制uk的作用下到最终状态xN的代价

易得



把Q（x,u）在标称轨迹点（xk,uk）附近二阶泰勒展开

可以得到δQ



令



得到最优控制扰动δu\*



由得



把δu\*代入上式得到



把V在标称轨迹点处（xk）二阶泰勒展开得到δV表达式



两式对比可得到Vx，和Vxx的表达式



要求得δu\*和Vx，Vxx需要求得Qx,Qu,Qxx,Quu,Qux(Qux和Qxu互为转置)





忽略动力学方程f的二阶导数则



从第N-1步开始迭代，因为第N步的V(Xn)已知，可求得Vx,Vxx，则可以更新第N-1步的Q的各阶导数，然后求得δu\*并更新第N-1步的Vx,Vxx，则可以继续求第N-2步的Q的各阶导数，如此不断迭代到初状态X0，则可以得到一个控制序列的最优扰动序列，把最优扰动序列和最初的控制序列相加，得到局部最优控制序列，然后用动力学方程前向rollout出一条新轨迹。新轨迹可以继续求最优扰动序列更新轨迹，如此不断迭代直到代价函数收敛。

