6.12.

原来强化学习效果不好是因为ELM的神经元个数没有调整对，现在调到100个左右就没有warning的问题，但学习效果还是不好，和手动的策略比，结果还差很多。

我在考虑可能是因为强化学习的信息不够，向看看到底通过ELM能不能够学习出手动的策略，下一步就是把手动的策略作为每一步的输入，这样来模拟最优的强化学习策略，看看ELM能不能够把手动的策略学到。如果可以，说明Q——learning需要改进，如果不行，则要改进ELM。

6.13

今天发现如果使用手动策略，ELM的结果能够大幅增加，但还没到手动策略的高度，主要原因可能就是训练不够充分，再调整一下应该能到位，对比一下手动策略和ELM策略的输出应该就能看出来了。

但从手动策略的结果看，基本上就是3次执行任务，1次维修，和定时维修差不多，而且似乎没有体现出变化的波动性和PHM的价值，所以想定需要重新设计，要稍微复杂一些，策略也要多一些，重新建一个文件来搞。

6.28.2019

通过这两天的调试，基本解决了简单想定的问题，使用ELM作为q learning可以最后逼近人工设定的最优值，之前主要问题集中在learning rate的设定上，这个值应该是在0.9左右，意思是每个周期，未来的收益在当前的折算是0.9，这个值应该根据要考虑的时间长度来设定，在0.9的时候，是最佳的。今天过了一下程序，也更深一层理解的Q learning的本质，就是一个迭代算法，将当前收益和未来收益结合。

后面还得进一步把程序优化，按照标准的qlearning的流程写，再把想定复杂化一下，定义清楚任务奖励和维修过程等。

场景建模

任务收益，参考任毅的论文，维修成本建模，参考zio，michele论文

长期收益，随机出现的高价值任务收益

今天看了kamal关于用LSTM预测，然后建立维修规则的论文，还是很有收获的，我明天好好把这些场景设计一下，把任务规则、维修成本细化一下，形成一个想定，当前的想定还是继续保持，做两个案例。

为什么在维修优化中用强化学习，列举一些强化学习的特长，见今日头条收藏中足球的帖子

2019.7.2

今天完成了复杂案例的编制，考虑了库存成本，将决策变量扩展到6维，考虑订货的影响。但初步结果不理想，收益值波动很大，考虑原因应该是强化学习的效率太低，没有累积正向的学习结果，导致正向的结果不能叠加，后面需要仔细过一遍程序，逐步发现问题。

其实本质上还是没有明白为什么第一个案例，把learning rate调为0.9了之后就解决了问题，把这个问题想明白，估计第二个案例也能解决。

2019.7.7

今天进一步完善了复杂案例的调试，发现了一个低级错误，就是在for循环内反复训练ELM模型，改正后至少出现了强化学习效能的逐步递增结果。但后面调试还是效果不理想，效能始终不能超过人工规则。后来发现，如果把order delay和库存成本提高，会对ELM方法有利，但从结果仍然不能解读。

下一步要讲每个输入输出都对应上，观察下ELM是否能选择高回报任务，再测试ELM性能和库存成本，任务回报之间的关系

2019.7.9

今天调试成功了，发现在任务收益波动比较大，库存成本较高，且订货周期比较长的情况下，elm的效果能超过手动策略，主要是因为手动策略在开始的几个周期内不停订货，导致后面的库存成本一直比较高。但手动策略完成的任务数是最多的，只是库存成本过高。再有一个改进就是不用随机样本进行在训练，而是只用上一个周期的样本，这个使结果有一定好转

state\_data.length=random('uniform',1,10,[nums,1]);

state\_data.mission\_timepoint=cumsum(state\_data.length);

state\_data.mission\_strengh=random('uniform',0.8,1.5,[nums,1]);

state\_data.mission\_reward=random('uniform',1,40,[nums,1]);

for i=1:nums %% randomly generate the very high mission reward with probability thre

if rand<thre

state\_data.mission\_reward(i)=random('uniform',50,100,[1,1]);

end

end

state\_data.corr\_m\_time=random('uniform',12,18,[nums,1]);

state\_data.corr\_m\_cost=random('uniform',8,12,[nums,1]);

state\_data.prev\_m\_time=random('uniform',4,6,[nums,1]);

state\_data.prev\_m\_cost=random('uniform',2.5,4,[nums,1]);

order\_t=15;

store\_cost\_rate=2;

initial health是否需要重新设置？定时维修需要设置