

基于反向传播的网格交易参数设计

代码位于 <https://github.com/JennyJiang118/grid.git>

一、主要原理

反向传播即误差传播算法，在神经网络中用于根据历史训练结果调整权重等参数。过程为：求导->求梯度->更新反向->损失更小->（局部）收敛->局部极值。

实际处理中，需注意以下几个问题：

1.反向传播需要设置初始值作为传播的起始点，这里使用人工设置的参数作为初始值。人工设置中，调整参数并运行 `human set grids(line118)` 可获得该组参数下收益情况，然后根据收益情况调整参数大小。由于大多数参数的取值与实际资金有关，且多为离散值，因此不纳入反向传播自动调参中。

2.这里的局部极值，是在人工设置的初始参数的周围形成“局部”，因此人工设置的初始参数十分重要。由于本问题中，参数的设置效果基本为凸函数或渐进函数，因此人工设置的参数基本在全局最优的周围。

3.实际操作中根据历史价格作出参数调整，因此 BP 调整参数有效的前提在于，未来价格波动与历史价格波动有一定程度的一致性

二、代码解读

1.human set grids:

这里处理的是人为给定的 `grids`，使用自动调参时可以不运行

2.BP:

#目标

获取最优网格 `grids`

根据已有信息，自动调整 `grids` 密度，代替人工设置或调整

#设计思路

使用逆向传播 `bp`，斜率代替导数

#实际操作步骤

1.使用 `human set grids` 找出相对较好的 `grids` 和其他参数，作为初步参数

2.实际操作中，BP 中的初始 `grids` 参数，初始化为 `human set grids` 里的最优值

#超参：

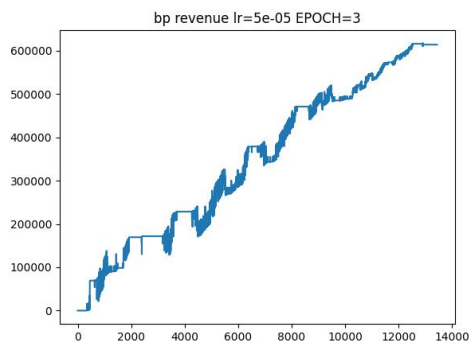
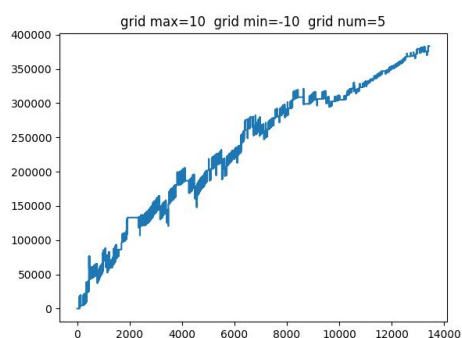
`lr` 学习率

`EPOCH` 总训练轮数

三、实验结果

1. IF

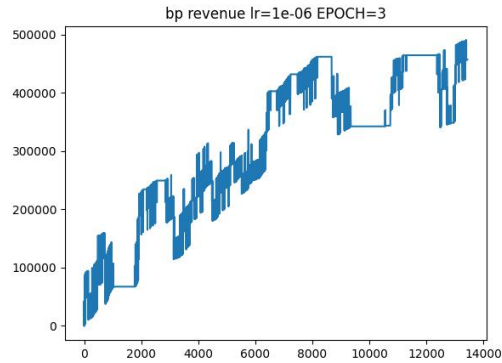
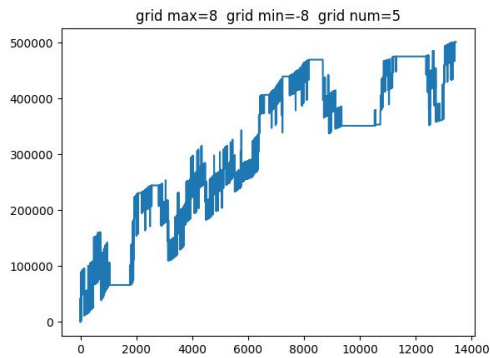
IF03-12 合约中，固定网格数为 5（保证金限制），手工初步调参最优结果，与 BP 算法所得（局部）最优结果如下：



参数设置为:

参数名称	数值	解释	来源
网格与均线的最高差值 grid_max	10	超过该值后，价格再涨也不再买入。由风险承受能力决定，承受能力高则提高该值	人工设置
网格与均线的最低差值 grid_min	-10	同上	人工设置
网格数 grid_num	5	决定网格密度	人工设置
仓位上限 max_hold	4	持仓超过该值后，不再买入。由资金充裕程度决定	人工设置
仓位下限 min_hold	-4	同上	人工设置
基准线 base_line	avg3	可选: avg2,avg5	人工设置
训练轮数 EPOCH	3	对网格密度迭代更新的次数。轮数越多，网格密度更新得越好，过拟合风险、时间开销也越大	人工设置
学习率 lr	5e-5	对网格密度迭代更新的速度	人工设置
网格设置 grids	[-2.90328236 0.55807827 1.55807827 2.55807827 5.24845667]	在人工给定较优初值后，反向传播（超参为 EPOCH, lr）所得。收益比人工设置的最优收益高 50%	BP 迭代更新所得，自动调参

2. IC



参数设置为:

参数名称	数值	解释	来源
网格与均线的最高差值 grid_max	8	超过该值后，价格再涨也不再买入。由风险承受能力决定，承受能力高则提高该值	人工设置
网格与均线的最低差值 grid_min	-8	同上	人工设置
网格数 grid_num	5	决定网格密度	人工设置
仓位上限 max_hold	4	持仓超过该值后，不再买入。由资金充裕程度决定	人工设置
仓位下限 min_hold	-4	同上	人工设置
基准线 base_line	avg5	可选: avg2,avg5	人工设置
训练轮数 EPOCH	3	对网格密度迭代更新的次数。轮数越多，网格密度更新得越好，过拟合风险、时间开销也越大	人工设置
学习率 lr	1e-6	对网格密度迭代更新的速度	人工设置
网格设置 grids	[-7.88087025 -3.88087025 -0.3111267 4.84135106 7.64536759]	在人工给定较优初值后，反向传播（超参为 EPOCH, lr）所得。收益与人工设置相比变化不明显	BP 迭代更新所得，自动调参

四、后续工作

1. 保证金当前以仓位上下限的形式作出限制，后续可以加入保证金对可用资金的影响-->修改网格设置：资金限制，参数减小
- 2*. 价格的变动是离散值，当前使用持有手数进行标记，后续可插值离散化，使结果更精确
3. BP 仅处理了网格密度的自动调参，后续可将其他参数，如网格数量、仓位上下限等纳入调参（仓位上下限受保证金影响，调参意义不大，主要内容会集中在网格数量的纳入上）
- 4*. 损失过大时，及时切除（纳入训练）
- 5*. 同一种类，使用同套参数-->多数据验证，防止过拟合
6. 加权数据训练