

遗传算法和遗传规划

—— 寻找不一样的Alpha



证券分析师

陈杰 A0230513080006

2016.9.6

主要内容

1. 遗传算法
2. 遗传算法的实现和检验
3. 遗传规划的实现和检验

1.1 遗传算法的核心逻辑：物竞天择

- 遗传算法（Genetic Algorithm，GA）是模拟生物进化过程的仿生计算机算法。把实际遇到的问题，用程序语言进行基因编码，并建立选择、交叉、变异、和淘汰的机制，通过不断迭代来搜寻全局最优的解决方案。
- 遗传算法已经在多个领域体现价值：管理决策、自动控制、路径规划、图像处理、人工生命、机器学习等等。

1.2 遗传算法的算法流程

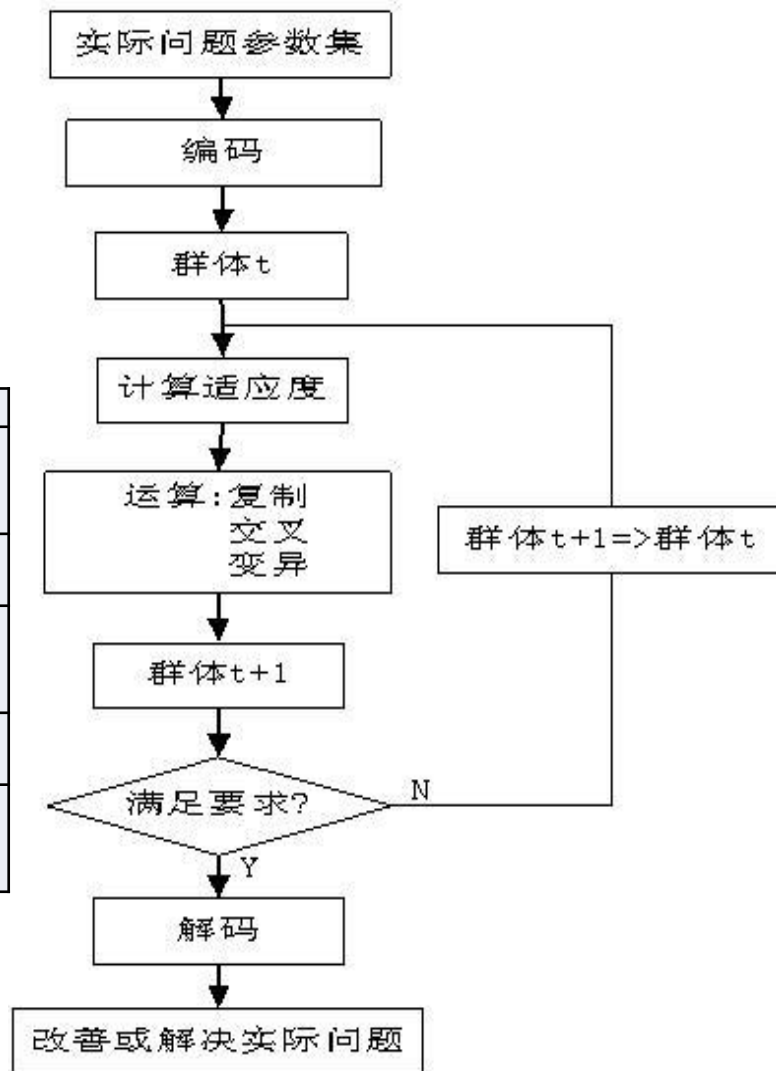
■ 编码类型

- 指标编码[D]：1，0，-1
- 逻辑编码[T]：AND，OR，NOR
- 实数编码[R]：0.3415，-1.415

■ 5个基因座编码的例子

	1	2	3	4	5
基因座	价格创5日最高[D]	与或非[T]	MACD方向指标[D]	与或非[T]	最近10日波动率<[R]买入，否则卖出
基因编码（个体1）	1	00	1	01	0.015
翻成交易规则	5日创新高买入	并且	按MACD指示买卖	或	最近10日波动率<0.015买入，否则卖出
基因编码（个体2）	-1	01	0	00	0.008
翻成交易规则	5日创新高卖出	或	不按MACD指示买卖	并且	最近10日波动率<0.008买入，否则卖出

- 个体1的程序编码为 1 00 1 01 0.015，按照这个编码进行交易得到一个交易结果
- 对种群中所有个体计算交易结果，并转换成种群的适应度
- 基因座个数，种群中个体个数取决于计算力

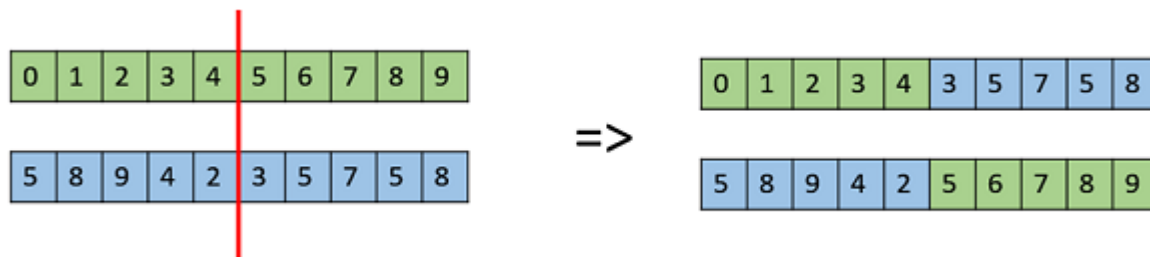


资料来源：申万宏源研究

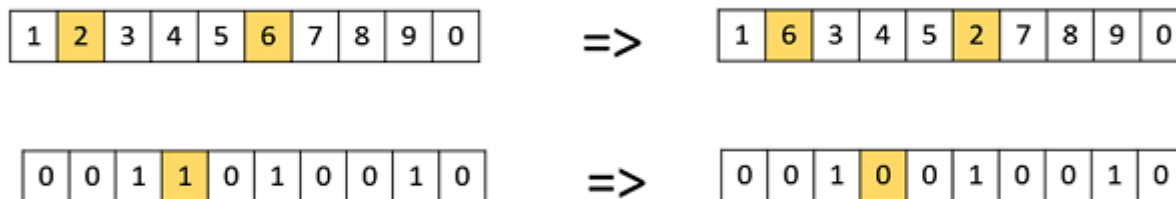
1.2 复制、交叉、变异

- **复制**：每一代种群中，按照适应度复制保留生存下来的个体，用来生成新的种群，并执行交叉和变异计算。

- **交叉**：随机地选择原有两个个体部分基因，进行交换，生成新个体



- **变异**：随机地在原有个体上的基因进行改变，生成新的个体



资料来源：申万宏源研究

1.2 遗传算法的优势

- 算法的迭代取决于种群的效果，而非单一个体的结果，稳健性强
- 种群目标可以个性化设置，目标可以是方向、权重、分类筛选等实际问题
- 算法的搜索路径具备多样化，非线性，随机性
- 不依赖于模型假设，基因座可扩展性强，可以处理不同类型信息
- 计算并行度高，潜在计算力强，可以处理大数据集
- 全局搜索，容易跳出局部优化区域，迭代效率高
- 结果具有显示表达，非黑箱，易纠错

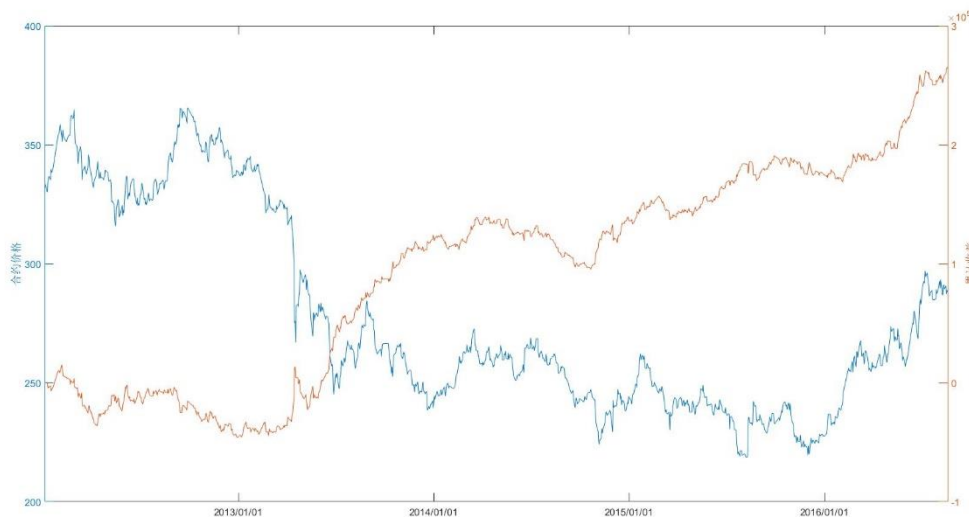
2. 遗传算法构建多技术指标交易策略

- 单个技术指标有钝化，实践中往往运用一组技术指标。
- 不同技术指标之间具有复杂性，如何把控指标之间的逻辑、顺序、重要程度需要很强的经验。
- 代码展示

2. 遗传算法在黄金期货 (AU) 上的测试效果

■ 寻优结果：

- 买入信号：(Short Term RSI(20) > 50 Or ((Close(11) Closes Below Keltner_Up(52, 2.2)) Or (Keltner_Down(46, 1.4) < Open(19))))
- 卖出信号：(Short Term RSI(20) < 50 Or ((Close(11) Closes Above Keltner_Down(52, 2.2)) Or (Keltner_Up(46, 1.4) > Open(19))))
- 止损信号：进入价格反向触及 $0.94 \times \text{ATR}(20)$



资料来源：申万宏源研究

期货种类	累计收益	交易次数	获胜次数	失败次数	胜率	单次均收益	单次获胜均收益	单次失败均收益	最大回撤金额
AU	264220	897	447	450	49.83%	294.56	2368.21	-1765.27	60840

3.1 如此复杂的Alpha因子是如何被挖掘的？

- 《101 formulative Alphas》披露了由世坤投资咨询（WorldQuant）公司提供的Alpha因子，世坤公司在2015年管理市值规模约为40亿美金。
- 文章中披露的Alpha因子的复杂程度几乎不可能由传统的优化方法得到，从因子的构造类型看，很像遗传规划搜索得来的结果。

```
Alpha#63: ((rank(decay_linear(delta(IndNeutralize(close, IndClass.industry), 2.25164), 8.22237))  
- rank(decay_linear(correlation(((vwap * 0.318108) + (open * (1 - 0.318108))), sum(adv180,  
37.2467), 13.557), 12.2883)))) * -1)
```

```
Alpha#64: ((rank(correlation(sum(((open * 0.178404) + (low * (1 - 0.178404))), 12.7054),  
sum(adv120, 12.7054), 16.6208)) < rank(delta((((high + low) / 2) * 0.178404) + (vwap * (1 -  
0.178404))), 3.69741))) * -1)
```

```
Alpha#65: ((rank(correlation(((open * 0.00817205) + (vwap * (1 - 0.00817205))), sum(adv60,  
8.6911), 6.40374)) < rank((open - ts_min(open, 13.635)))) * -1)
```

```
Alpha#66: ((rank(decay_linear(delta(vwap, 3.51013), 7.23052)) + Ts_Rank(decay_linear((((low  
* 0.96633) + (low * (1 - 0.96633))) - vwap) / (open - ((high + low) / 2))), 11.4157), 6.72611)) * -1)
```

```
Alpha#67: ((rank((high - ts_min(high, 2.14593)))^rank(correlation(IndNeutralize(vwap,  
IndClass.sector), IndNeutralize(adv20, IndClass.subindustry), 6.02936))) * -1)
```

```
Alpha#68: ((Ts_Rank(correlation(rank(high), rank(adv15), 8.91644), 13.9333) <  
rank(delta(((close * 0.518371) + (low * (1 - 0.518371))), 1.06157))) * -1)
```

资料来源 <http://ssrn.com/abstract=2701346>

3.2 遗传规划：更强兵器

- 遗传规划（Genetic Programming，GP）用树状结构表示基因，而遗传算法的基因座是串行的。两者相比，遗传规划的扩展性，可变性更强。
- 遗传规划树状结构的节点是由“运算算子”构成的，这些算子包括了科学计算（+，-，log，sin，sum），自定义算子（例如 $z=f(x,y)$ 函数， $a=\text{rank}(b)$ 等）。
- 遗传规划的树状结构的层级和复杂度，随着迭代自动延伸，不需要对最终模型有先验判断。
- 理论上，基本数学和逻辑运算符，可以表述所有已知和未知的技术指标和交易经验。但基于现有的运算能力，限定运算和数据范围能有效提高交易系统开发效率。

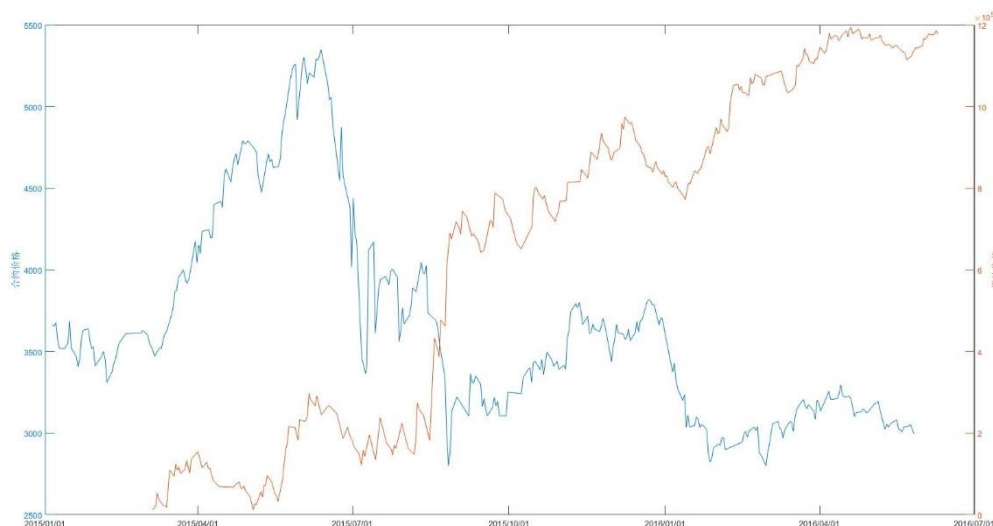
3.3 遗传算法的优势

- 遗传规划不需要预设基因座结构，突破固有思维局限。
- 迭代效率上，遗传规划的收敛速度是指数级别的，而遗传算法则停留在线性级别
- 遗传规划能描述层次化问题，树状结构的主干能提供更多有效信息。
- 代码演示

4.1 遗传算法实际测试

■ 寻优结果

- 买入条件：(((HighDaily(17) < (PivotR3 - (LowDaily(17) - LowDaily(8)))) Or ((Open(14) * (Volume(1) - -0.35)) > ((Volume(17) - -39.580002) * (Volume(6) * Volume(7)))) Or (((High(2) + PivotR3) - High(16)) < HighDaily(2)))
- 卖出条件：(((LowDaily(17) > (PivotR3 - (HighDaily(17) - HighDaily(8)))) Or ((Open(14) * (Volume(1) - -0.35)) < ((Volume(17) - -39.580002) * (Volume(6) * Volume(7)))) Or (((Low(2) + PivotR3) - Low(16)) > LowDaily(2)))
- 止损止盈条件：单边止损 50 个点止损



资料来源：申万宏源研究

期货种类	累计收益	交易次数	获胜次数	失败次数	胜率	单次均收益	单次获胜均收益	单次失败均收益	最大回撤金额
IF	1177748	330	146	184	44.2	3568.9	23741.5	-12437.6	202079.2

信息披露 证券分析师承诺

本报告署名分析师具有中国证券业协会授予的证券投资咨询执业资格并注册为证券分析师，以勤勉的职业态度、专业审慎的研究方法，使用合法合规的信息，独立、客观地出具本报告，并对本报告的内容和观点负责。本人不曾因，不因，也将不会因本报告中的具体推荐意见或观点而直接或间接收到任何形式的补偿。

与公司有关的信息披露

本公司隶属于申万宏源证券有限公司。本公司经中国证券监督管理委员会核准，取得证券投资咨询业务许可，资格证书编号为：ZX0065。本公司关联机构在法律许可情况下可能持有或交易本报告提到的投资标的，还可能为或争取为这些标的提供投资银行服务。本公司在知晓范围内依法合规地履行披露义务。客户可通过 compliance@swsresearch.com 索取有关披露资料或登录 www.swsresearch.com 信息披露栏目查询从业人员资质情况、静默期安排及其他有关的信息披露。

机构销售团队联系人

上海	陈陶	021-23297221	18930809221	chentao@swsresearch.com
北京	李丹	010-66500610	18930809610	lidan@swsresearch.com
深圳	胡洁云	021-23297247	13916685683	hujy@swsresearch.com
海外	张思然	021-23297213	13636343555	zhangsr@swsresearch.com
综合	朱芳	021-23297233	18930809233	zhufang@swsresearch.com

法律声明

本报告仅供上海申银万国证券研究所有限公司（以下简称“本公司”）的客户使用。本公司不会因接收人收到本报告而视其为客户。客户应当认识到有关本报告的短信提示、电话推荐等只是研究观点的简要沟通，需以本公司 <http://www.swsresearch.com> 网站刊载的完整报告为准，本公司并接受客户的后续问询。本报告首页列示的联系人，除非另有说明，仅作为本公司就本报告与客户的联络人，承担联络工作，不从事任何证券投资咨询服务业务。

本报告是基于已公开信息撰写，但本公司不保证该等信息的准确性或完整性。本报告所载的资料、工具、意见及推测只提供给客户作参考之用，并非作为或被视为出售或购买证券或其他投资标的的邀请或向人作出邀请。本报告所载的资料、意见及推测仅反映本公司于发布本报告当日的判断，本报告所指的证券或投资标的的价格、价值及投资收入可能会波动。在不同时期，本公司可发出与本报告所载资料、意见及推测不一致的报告。客户应当考虑到本公司可能存在可能影响本报告客观性的利益冲突，不应视本报告为作出投资决策的惟一因素。客户应自主作出投资决策并自行承担投资风险。本公司特别提示，本公司不会与任何客户以任何形式分享证券投资收益或分担证券投资损失，任何形式的分享证券投资收益或者分担证券投资损失的书面或口头承诺均为无效。本报告中所指的投资及服务可能不适合个别客户，不构成客户私人咨询建议。本公司未确保本报告充分考虑到个别客户特殊的投资目标、财务状况或需要。本公司建议客户应考虑本报告的任何意见或建议是否符合其特定状况，以及（若有必要）咨询独立投资顾问。在任何情况下，本报告中的信息或所表述的意见并不构成对任何人的投资建议。在任何情况下，本公司不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任。市场有风险，投资需谨慎。若本报告的接收人非本公司的客户，应在基于本报告作出任何投资决定或就本报告要求任何解释前咨询独立投资顾问。

本报告的版权归本公司所有，属于非公开资料。本公司对本报告保留一切权利。除非另有书面显示，否则本报告中的所有材料的版权均属本公司。未经本公司事先书面授权，本报告的任何部分均不得以任何方式制作任何形式的拷贝、复印件或复制品，或再次分发给任何其他人，或以任何侵犯本公司版权的其他方式使用。所有本报告中使用的商标、服务标记及标记均为本公司的商标、服务标记及标记。

简单金融 · 成就梦想

A Virtue of Simple Finance



申万宏源研究微信订阅号



申万宏源研究微信服务号

上海申银万国证券研究所有限公司
(隶属于申万宏源证券有限公司)

陈杰
chenjie@swsresearch.com