



# ALGORITHMIC TRADING

## IMPERIAL COLLEGE LONDON

# AlgoCourses Notebook

Summarised by Guanzhong Yang, CID 02655923



# 方向单

## ● 动量 Momentum

(1) 简单移动平均线 SMA: 作为动量的滞后指标。

$$SMA_t = \frac{1}{n} \sum_{i=t-n+1}^t P_i$$

(2) 指数移动平均线 EMA:

通过权重调整使其对价格变动更敏感, 从而更早识别趋势, 但也可能增加错误信号

$$EMA_t = \frac{2}{n+1} P_t + \frac{n-1}{n+1} EMA_{t-1}$$

- 短期均线 EMA 上穿长期均线: 买入信号, 上涨动量确认
- 短期均线 EMA 下穿长期均线: 卖出信号, 下跌动量确认

## ● Black-Scholes (BS) 模型

其能算出一个期权理论上应该值多少钱。

$$C = N(d_1)S_t - N(d_2)Ke^{-rt}$$

其中  $d_1 = \frac{1}{\sigma\sqrt{t}} \left( \ln \frac{S_t}{K} + \left( r + \frac{\sigma^2}{2} \right) t \right)$ ,  $d_2 = d_1 - \sigma\sqrt{t}$ ,  $S$  是标的价格,  $K$  是行权价,  $t$  是剩余到期时间,

$r$  是无风险利率,  $\sigma$  是波动率(唯一不知道也最重要的参数)。

我们可以通过公式反推出隐含波动率  $\sigma$ , 即市场对未来波动的预期价格。

其对期权方向单的指导为:

(1) 现在 IV 低 / 你认为 IV 未来会涨:

期权便宜, 适合买期权做方向. 方向对了 + IV 上升 = 赚得更多

(2) 现在 IV 高 / 你认为 IV 未来会跌:

哪怕方向对了, IV 一下跌, 收益会被吃掉甚至亏钱

# 做市单

## ● 欧拉-马鲁山方法

其用于模拟资产价格随时间演变的通用离散化:

$$\Delta X_t = \mu_t \Delta t + \sigma_t \Delta W_t$$

其中  $\Delta X_t$  表示当日价格变化;  $\mu_t$  为自残预期收益率, 可以用平均日收益率表示;  $\sigma_t$  为扩散系数, 可以用资产价格的历史波动率确定,  $\Delta W_t$  是布朗运动的随机增量, 服从正态分布  $N(0, \Delta t)$

简单来说, 这个公式把价格变动分成了“确定的趋势收益”和“不确定的随机波动”两部分, 用来模拟真实市场中价格的连续变化, 常用于回测

## ● Avellaneda-Stoikov (AS) 做市模型

(1) 保留价格 reservation price: 做市商认为的资产“公允”价格, 考虑了当前库存风险敞口:

$$r(s, q, t) = s - q\gamma\sigma^2(T-t)$$

其中  $s$  是当前中间价,  $q$  是当前库存(多头用正数, 空头用负数),  $\gamma$  是风险厌恶系数(自己确定, 也可以通过回测扫描进行确定),  $\sigma$  是波动率,  $T-t$  是距离收盘的时间

(2) 最优买卖价差 optimal bid & ask spread:

$$\delta^a + \delta^b = \gamma\sigma^2(T-t) + \frac{2}{\gamma} \ln\left(1 + \frac{\gamma}{k}\right)$$

其中  $k$  是市场流动性参数. 在拿到盘口数据(不同价位的挂单, 成交速率, 撤单率)后, 对成交速率做指数拟合  $\lambda(\delta) = Ae^{-k\delta}$ , 即可回归出  $k$  的值.

$k$  很大说明市场对价差很敏感, 稍微报远一点, 成交概率直接暴跌.  $k$  很小说明哪怕你价差报大一点, 也依然能成交.

(3) 局限性:

1. 现实中订单达到不符合泊松分布, 且模型未考虑队列位置, 价格冲击和延迟.
2. 由于信息不对称导致的亏损交易(可以通过订单簿失衡 Orderbook imbalance 等信号来估计真实价格并规避噪声)
3. 做市商的报价限制
4. 忽略交易手续费, 期权费用, 做市商的返佣以及平台的奖励计划等

## ● 市场冲击, 平方根模型与 Almgren-Chriss (AC) 模型

(1) 定义

做市商进行大宗交易时, 市场冲击会直接引发价格变动

总交易成本 = 买卖点差 + 临时冲击/机械冲击 + 永久冲击/信息冲击 + 持仓风险

机械冲击: 由于市价买入直接消耗了卖单, 导致价格在短期内上升

信息冲击: 代表市场参与者对资产的信念

(2) 平方根模型

仅估算点差 + 临时冲击成本, 无法计算永久冲击与持仓风险, 是做市商初步判断流动性冲击成本的基础模型

$$\Delta P = \text{spread cost} + \alpha\sigma\sqrt{\frac{Q}{V}}$$

其中  $\Delta P$  是价格冲击成本;  $\text{spread cost}$  是买卖点差成本;  $\alpha$  是风险厌恶系数;  $\sigma$  是标的资产波动率;  $Q$  是做市商订单数量;  $V$  是市场单位时间成交量

根据你能接受的成本, 对每个  $\alpha$ , 求解每次最大下单量  $Q$ , 再用  $Q$  进行回测, 测出哪种情况下总成本最低

(3) Almgren-Chriss (AC) 模型

其估算时同时包含临时冲击 + 永久冲击 + 时间风险, 且直接输出一条最优路径无需回测

$$x(t) = X \frac{\sinh(\kappa(T-t))}{\sinh(\kappa T)}, \quad \kappa = \sqrt{\frac{\gamma\sigma^2}{2\eta}}$$

其中  $x(t)$  为  $t$  时刻做市商剩余持仓量;  $X$  为初始总持仓量;  $T$  为总交易执行时间;  $\gamma$  为风险厌恶系数, 做市商根据自身风险偏好设定校准;  $\sigma$  是标的资产波动率;  $\eta$  为永久冲击系数, 通过历史交易数据回归拟合, 反映做市商交易对价格的永久性影响程度.

(4) 常见拆单方式

1. 交易量加权, 市场成交量大时多下单, 成交量小时少下单

2. 时间加权, 按时间均匀拆分大单, 每隔固定时间下单等量资产, 不考虑市场成交量变化

## ● 套利

(1) 定义:

简单来说就是利用不同市场间的价格差异获利, 例如在 A 交易所买入, 同时在 B 交易所卖出. 当两者的价差超过设定的最大值阈值或低于最小值阈值时, 就出现了套利机会

一般通过“买入 ETF, 卖出成分股”或“买入成分股, 卖出 ETF”来实现套利

(2) BS 模型对期权套利的指导

1. 隐含波动率 > 自己预期: 期权太贵, 可以卖期权 + 对冲, 赚波动率回落的钱

· 卖了看涨期权, 同时做多一点股票做对冲, 做成 Delta 中性, 即涨跌暂时都不亏不赚

· 未来波动变小. 期权跌, 组合偏空了不再中性, 低价买回一点股票对冲; 期权涨, 组合偏多了不再中性, 高价卖出一点股票对冲. 即高抛低吸

2. 隐含波动率 < 自己预期: 期权便宜, 可以买期权 + 对冲, 赚波动率涨回的钱

· 买了看涨期权, 同时做空一点股票做对冲做成 Delta 中性, 即涨跌暂时都不亏不赚

· 未来波动变大. 期权涨, 组合偏多了不再中性, 高价卖掉一点股票对冲; 期权跌, 组合偏空了不再中性, 低价买入一点股票对冲. 即高抛低吸

(3) 限制:

在实际操作中, 套利受到交易手续费, 套利机会持续时间以及市场竞争的限制

## ● 对冲

(1) 定义:

$$R_i = \beta_0 R_m + \beta_1 F_1 + \dots + \beta_n F_n + \varepsilon$$

其中  $R_i$  = asset returns,  $R_m$  = market returns,  $\beta_i = \frac{\text{Cov}(R_m, R_i)}{\text{Var}(R_m)}$ ,  $F_i$  = factors,  $\varepsilon$  = errors

常见的  $F_i$  有市场, 市值, 和账面市值比

我们要做的是通过对冲掉因子的影响, 使收益仅取决于资产自身的残差  $\varepsilon$

(2) Delta 对冲

$\Delta = \frac{\partial V}{\partial S}$ , 即期权价格相对于标的资产价格的变化率. 通过平衡各个货物的 Delta, 实现 Delta 中性.

# 全流程

## 通用量化交易算法框架

### (1) 信息采集与建模

市场内数据（从交易所实时获取）

- 价格:  $mid = (best\_bid + best\_ask) / 2$
- 流动性: 盘口价差、各档深度
- 成交量: 最近成交量均值, 用于判断市场活跃度
- 短期波动率:  $\sigma = std(recent\ mids)$ , 或用 realized vol、EWMA

市场外数据（alpha 信号来源）

- 任何影响资产真实价值的外部信息: 宏观数据、另类数据、衍生品隐含信息等
- 对每个信号建模, 输出  $(FV\_model, \sigma\_model)$

不确定度建模原则

- 线性资产: 方差可加,  $\sigma\_combined = \sqrt{\sum \sigma_i^2}$
- 非线性 payoff: 用 Monte Carlo, 直接模拟 payoff 分布取 std
- 希腊字母: 解析解优先; 无解析解时用有限差分:

$$\Delta = \frac{f(S+h) - f(S-h)}{2h}, \quad \Gamma = \frac{f(S+h) - 2f(S) + f(S-h)}{h^2}$$

### (2) 信号融合（贝叶斯更新）

把模型 FV 和市场 mid 融合, 核心思想是精度加权:

$$FV_{post} = \frac{\frac{1}{\sigma_{model}^2} FV_{model} + \frac{1}{\sigma_{market}^2} mid_{market}}{\frac{1}{\sigma_{model}^2} + \frac{1}{\sigma_{market}^2}}$$

关键设计决策:  $\sigma_{market}$  不是固定的, 而是根据模型历史偏差动态调整。模型持续偏离时, 增大对市场的信任权重。

通用原则: 任何信息源都有精度, 精度越高权重越大。这个框架可以融合任意多个信号源（不只是两个）。

### (3) 风险检查（执行任何交易前）

这步必须最先执行, 原因是后续所有交易都基于当前仓位状态。

检查项目:

- 单品种仓位上限
- 相关品种组合净敞口上限（防止相关资产同向押注叠加）
- 整体组合名义价值
- 止损线（unrealized PnL 触底则停止）

通用原则: 风险检查的顺序决定了你的最大损失边界。先减仓再交易, 而不是交易完再检查。

### (4) 主动信号执行（方向单 / 套利）

单资产方向单: 用 Z-score 衡量偏离程度

$$z = \frac{mid - FV_{post}}{\sigma_{post}}, \quad |z| > threshold \Rightarrow \text{开方向单}$$

多资产套利: 用 Basis 衡量组合定价偏差

basis = 价格组合 - 理论价值

阈值动态校准（取历史 basis 分布的高分位数），避免在正常波动范围内误触发。

通用原则:

- 方向单是在赌均值回归, 信号越强越可信, 但要有衰减机制防止单边过度押注
- 套利是在赌定价收敛, 阈值要覆盖交易摩擦成本（价差 + 滑点）

### (5) 对冲（主动交易后, 做市前）

为什么在做市前对冲: 对冲会改变仓位, 仓位直接影响做市的 Reservation Price, 必须先对冲到位再报价。

非线性资产: 需要显式 Delta/Gamma 对冲

$$target\_hedge = -pos \times \Delta$$

在 gamma 大的区域（折点、到期日附近）加快对冲频率。

线性资产: 不需要显式对冲, A-S 模型的 Reservation Price 偏移本身就是隐式对冲机制。

通用原则: 对冲频率和精度要和 gamma 大小匹配。gamma 越大, delta 变化越快, 对冲越频繁。

### (6) 被动做市（A-S 模型）

$$r = FV_{post} - \frac{pos}{pos\_limit} \cdot \gamma \cdot \sigma^2 \cdot T$$

$$half\_spread = \max\left(\text{spread}_{min}, \frac{\gamma \sigma^2 T}{2} + \frac{1}{\gamma} \ln\left(1 + \frac{\gamma}{k}\right)\right)$$

- $\gamma$ : 风险厌恶系数, 控制你愿意承担多少库存风险
- $k$ : 市场深度, 控制报价宽窄
- $T$ : 剩余时间, 越接近结算风险越高, 价差越宽

Reservation Price 的本质: 持多仓时  $r < FV$ , 报价整体下移, 主动吸引买单来消化库存。这是做市商的自动库存管理机制, 不需要额外下对冲单。

参数校准:  $\gamma$  和  $k$  不应该是固定的, 要根据实时市场状态（波动率、成交量、价差）动态调整。

