

# 量子计算 —量子金融

# Quantum Finance

网址: [www.qubits.top](http://www.qubits.top)

Calvin Tang

179209347@qq.com



# 量子金融

## 什么是量子金融？

量子金融是一个跨学科的研究领域，是经济物理学的一个分支，主要是运用量子计算优势和经济学理论方法来解决金融问题。

## 量子金融应用的场景：

- 股票追踪
- 交易策略
- 投资组合优化：静态，动态
- 资产定价
- 风险分析
- 加速交易效率
- 实时定价
- 信用评分流程优化
- 货币套利优化
- 衍生工具估值与调整
- .....



# 量化金融 - Q Quant 和 P Quant

## 什么是Quant ?

Quant全称是Quantitative Analyst即量化研究员，利用量化的方法（编程和模型）解决金融问题。

## Quant分类：

- P Quant：寻找Alpha，其任务是“模拟未来”。
- Q Quant：主要做衍生品定价，其任务是“推断当前”。

量化金融领域主要分为两派：衍生品定价的Q Quant，他们的主要任务是“推断现在”；量化风险和投资组合管理的P Quant，他们的主要任务是“模拟未来”；

Q重模型而轻数据，P则重数据而轻模型。当然两者也都要即有模型也有数据，但从应用上来讲，Q是模型固定，用数据来精化模型的参数（calibration）；而P则可以有若干备选模型，由数据的计算结果来选择最佳的模型（estimation）。

这个区别也造成在业界的划分。Q可以让你在缺少数据的情况得出一些结论，从而可以凭空制造一些东西出来，所以卖方（投行）用来做衍生品定价，业务模式是开发新的衍生品出来卖出去。P则喜欢大数据量，这天然就是买方（Hedge Fund类）所需要的技术，因为他们本来就需要针对大量证券做出筛选和投资决策，业务是数据驱动的。

# Q Quant 和 P Quant 对比

	P-Quant	Q-Quant
测度	真实概率	风险中性
工具	时间序列、贝叶斯、机器学习等	随机过程，偏微分方程
业务	买方，基金，承担风险	卖方，投行，转嫁风险
偏重	数据驱动，让数据说话	建模，发挥人脑优势
方法	估计 (estimation)	校准 (calibration)
目标	对未来建模，预测未来	推断现在
应用	寻找alpha，风险与组合管理	衍生品定价
交叉应用	对冲	统计套利

## ➤ 解析解：

- 优点：解析解计算快，形式优美
- 缺点：需要添加限制条件，比较理想化，描述能力较弱

## ➤ 数值解：

- 优点：描述能力强，实现简单
- 缺点：需要消耗大量计算资源

# Q Quant

	描述
目标	推断现在
数学环境	风险中性概率测度Q
随机过程	连续时间鞅
数据维度	低
数学工具	随机微积分、偏微分方程
面对挑战	校准
业务定位	卖方 (sell-side)

量化衍生品定价由Bachelier于1900年提出，他在其学位论文中首次将最基本同时也是极具影响力的随机过程——布朗运动，应用于期权的定价。

但这一理论一直没有引起关注，直到Merton (1969) 以及Black和Scholes (1973) 将第二个极具影响力的随机过程——几何布朗运动，引入期权定价。而下一个推动Q Quant发展的里程碑，则是Harrison和Pliska在1981年提出的资产定价的基础理论。证券价格所服从的随机过程的概率测度被称为是“风险中性”的，通常用字母Q表示，这也是“Q Quant”这一名称的由来。

从事Q Quant——衍生品定价的金融工程师，对其建模的特定产品都有着极其深入的了解。每一个证券都是被单独定价的，因此本质上Q Quant中的问题都是低维的。校准 (calibration) 是Q Quant所面临的一个主要挑战。处理连续时间Q-过程的主要量化工具为随机微积分和偏微分方程。

# P Quant

	描述
目标	对未来建模
数学环境	真是概率测度P
随机过程	离散时间序列
数据维度	高
数学工具	多元统计分析
面对挑战	估计
业务定位	买方 (buy-side)

风险与组合管理定位于在某一给定的投资范围内，对市场价格建立概率分布模型。这一真实的概率分布通常用字母P表示，以区别于衍生品定价中的“风险中性”测度Q。

基于真实分布，买方的主要工作就是决定证券的仓位以改善组合的收益-风险特征。

风险与组合管理的量化理论起源于Markowitz (1952) 的均值-方差体系。随后，Treynor (1962)、Mossin (1966)、Sharpe (1964)、Lintner (1965) 和Ross (1976) 建立的资本资产定价模型 (CAPM) 和套利定价模型 (ATP) 更是让这一领域有了突破性的发展。

- 趋势跟踪：均线策略，MACD策略
- 均值回归
- 统计套利：距离函数法，协整检验法，随机控制法，时间序列法
- 多因子模型
- 高频交易
- 机器学习策略



# P-Quant - 趋势跟踪 - 均线型指标

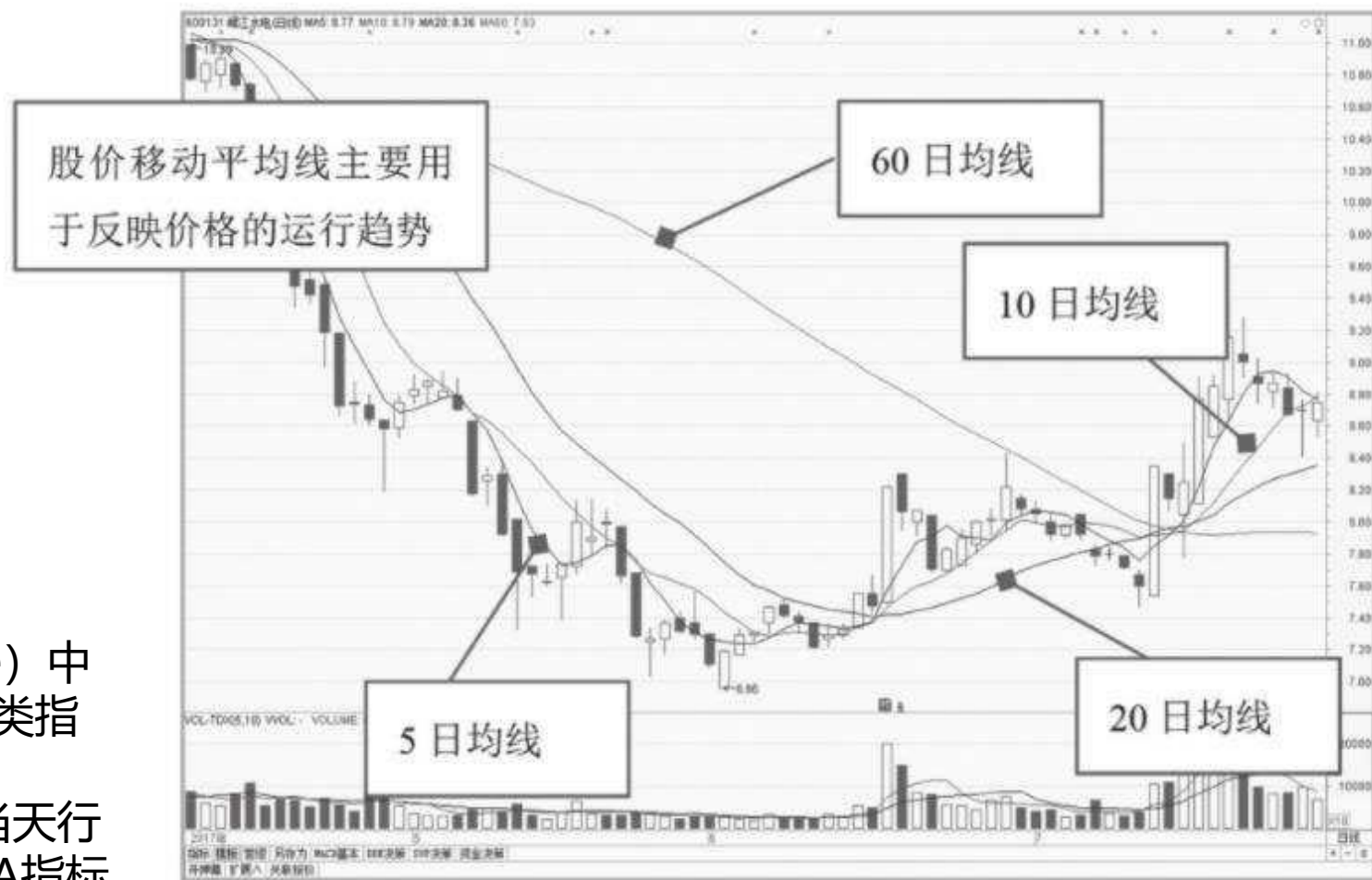
## MA——移动平均线

就是在某一段时间内，投资者持筹价格的平均值。不同周期的移动平均线可以为您表明短线、中线和长线的大趋势。

## EXPMA——指数移动平滑平均线

EXPMA指标 (Exponential Moving Average) 中文名称叫作指数平均数指标，它也是一种趋向类指标。

EXPMA指标由于计算公式中着重考虑了价格当天行情的权重，因此指标自身的计算公式克服了MA指标的滞后性。



# P-Quant - 趋势跟踪 - 趋势型指标

## MACD——平滑移动平均线

MACD经过多次平滑,淡化了单根K线的跳动,指示了当前趋势方向.其主要特点是稳健性,可以过滤掉市场的无序噪音,使其在周期较大, K线数目较多的行情中给出相对稳定的趋势指向。

## MACD研究的前提

- 趋势在一段时间内是可以把握的;
- 每个指标都有有效的时候, 没有指标会始终有效;





## P-Quant - 统计套利

在金融市场的交易中，统计套利算法是非常重要的一类交易算法，它的核心是对两种或多个资产你关系的建模与检验，如：距离函数法，协整检验法，随机控制法，时间序列法等。

**其中非常重要的一类时间序列方法，叫做协整检验法：**

- 协整的作用检验它们的回归方程所描述的因果关系是否是伪回归的。
- 协整是指若两个或多个非平稳的变量序列，其某个线性组合后的序列呈平稳性。
- 协整理论的作用在于正确地解释了经济现象和预测现象。
- 协整回归的所有变量必须是同阶单整的，但是满足同整单阶的不一定是协整的。

**常用的协整检验：**

- Engel-Granger 两步协整检验法
- Johansen 协整检验法

它们二者的区别在于 Engler-Granger 采用的是一元方程技术，而 Johansen 则是多元方程技术，所以 Johansen 协整检验法受限更小。

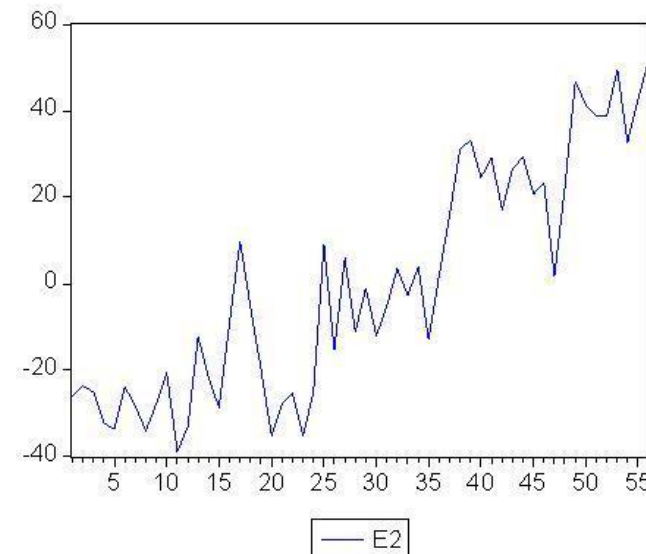
# P-Quant - 统计套利 - 协整检验法

## 基于协整性检验法的统计套利流程:

1. 从数据中找到具有运动协同性的组合
2. 对找到的时间序列组合, 进行Engle-Granger二步法检验
3. 对找到的复合协整性条件的股票组合, 设计进出场信号并进行模拟和实盘

## 数据特征:

1. 数据规模巨大
2. 搜索空间增长快
3. 算法复杂度高: 需要多次线性回归, 每次回归都需要进行大矩阵求逆过程。在经典算法里, 这个过程对数据规模N的依赖是非常高的。



## 量子优势:

1. 快速探测共线性  
条件数探测算法可以不要求方程的逆直接的到特征向量对应的特征值, 进而估算矩阵的条件数。
2. 加速协整性检验  
利用HHL算法为核心的量子线性回归算法, 可以快速计算多列数据的线性回归系数, 从而计算残差和进行后续统计假设检验。

内容来源: 本源量子

# P-Quant – 多因子模型

多因子模型是关于资产定价的模型。该模型是量化投资领域应用最广泛也是最成熟的量化选股模型之一，建立在投资组合、资本资产定价（CAPM）、套利定价理论（APT）等现代金融投资理论基础。多因子模型假设市场是无效或弱有效的，通过主动投资组合管理来获取超额收益。多因子选股的核心思想在于，市场影响因素是多重的并且是动态的，但是总会有一些因子在一定的时期内能发挥稳定的作用。

## Alpha因子概念

### 多因子模型（MFM）

- 影响股票预期收益率的因素从CAPM的单一因素扩展到多个因素

$$\tilde{r}_j = \sum_{k=1}^K X_{jk} * \tilde{f}_k + \tilde{u}_j$$

$X_{jk}$ : 股票 $j$ 在因子 $k$ 上的因子暴露（因子载荷）  
 $\tilde{f}_k$ : 因子 $k$ 的因子收益  
 $\tilde{u}_j$ : 股票 $j$ 的残差收益率

我们要找到alpha因子来预测收益率

$R_i$ 是资产 $i$ 的回报，它是个连续性变量，这是个多元线性回归

我们要找到 $X$ （特征）作为解释变量

这里的 $X$ 就是alpha因子





# P-Quant – 多因子模型

## 多因子模型:

1. 宏观经济因子模型
2. 基础因子模型
3. 统计因子模型

## 因子选择:

1. 因子测试与统计检验: 量子VaR值模拟计算
2. 相关性
3. 因子打分回测 - back test

## 因子组合:

- 对于**相关性较高**的因子集合, 可以采取两种方式处理:
  1. **剔除因子**: 根据因子本身的有效性进行排序, 挑选最有效的因子进行保留
  2. **线性组合**: 对因子进行**合成**, 尽可能多的保留有效因子信息:
    1. 等权组合
    2. 历史收益率加权
    3. 历史信息比例加权
    4. 主成分分析
    5. 多因子组合优化: 量子优化算法

## 多因子对冲策略

收益率作为因变量  
寻找可能与因变量相关的因子



因子选择



因子组合



收益率预测

## 多元线性回归

确定因变量  
确定备选的解释变量



变量选择



验证基本假定



预测

# P-Quant – 高频交易

高频交易(High Frequency Trading, HFT)是算法交易(Algorithmic Trading)的一个分支，在算法交易中，买卖决策非常迅速。高频交易对买卖双方的有效匹配起着至关重要的作用。

## 高频交易特点：

### 效果

- 每笔收益率很低，但是总体收益稳定

### 持仓

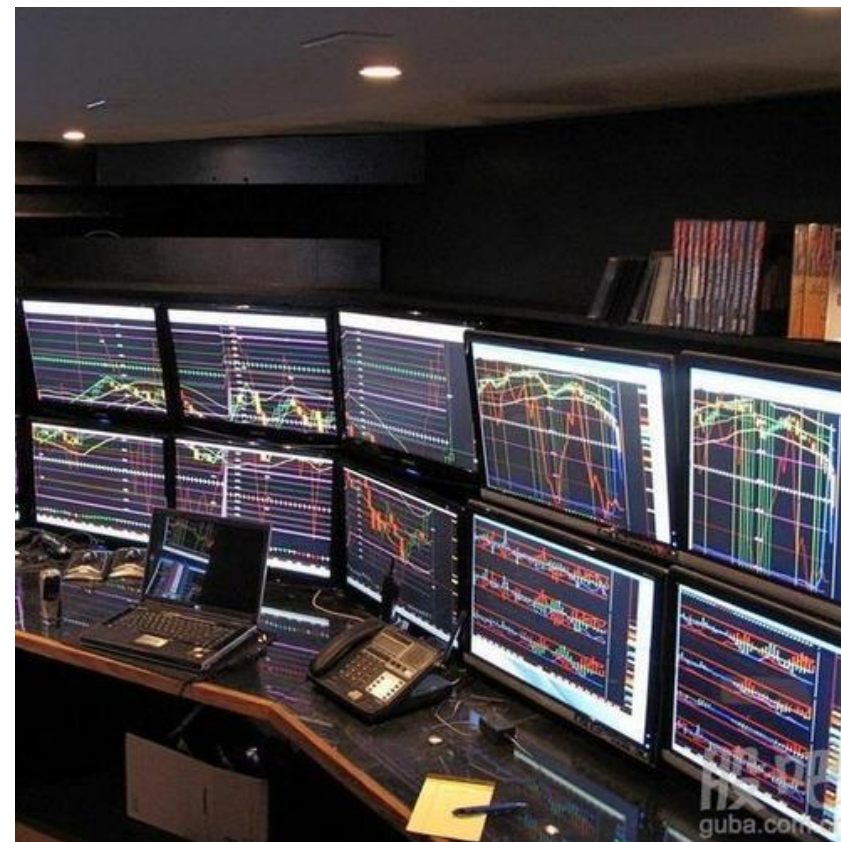
- 持仓时间极短

### 资金

- 流动性很高

### 基础设施

- 对系统的延时要求很高、模型的分析执行速度高



# 量化风控 - 金融欺诈识别

## 金融欺诈识别:

利用模式识别、异常点检测、聚类算法可以对金融欺诈进行识别。  
广义上主要包括：信贷欺诈，盗号盗刷，羊毛党三大类。

## 数据特征:

- 数据集不平衡
- 数据质量问题
- 数据标注问题

### 模型选择 1: 线性模型

包括：PCA、协方差分析技术、one class SVM

核心困难：需要做大规模矩阵的计算，特别是求矩阵的逆

### 量子计算:

- 能实现相同功能：qPCA, qSVM
- 计算速度的优势：基于HHL算法等，量子算法求矩阵的逆，相对于经典算法有巨大优势。

### 模型选择 2: 相似度模型

包括：distance-based, density-based, angle-based

### 量子计算:

- 基于Grover搜索算法，对非结构化搜索问题都能实现平方级的加速效果。

### 模型选择 3: 决策树集成

内容来源：本源量子



# 量化风控 - 信用分析

## 信用分析:

利用聚类算法可以对用户的信用数据进行分析。

## 数据特征:

- 数据规模大
- 数据质量参差不齐
- 稀疏度

## 建模流程:

1. 数据获取
2. 数据清洗
3. 数据探索
4. 特征提取

## 模型介绍:

1. 异常检测法
2. 拟合算法
3. 对抗生成网络
4. 多分类算法
5. PCA...

## 量子计算优势:

- 数据存储
- 模型训练
- 优化收敛

# 量化风控 – 风险预警

## 风险预警：

监管交易预防犯罪，或者利用蒙特卡罗算法模拟出投资组合的VaR值等。

## 数据特征：

- 老鼠仓：  
需要跟踪大量的交易账户，所有的股票类别，甚至所有相关的交易数据，并且对其中的资金，股票持有和交易情况都进行建模；
- 晃骗：  
需要跟踪各个股票，tick级别的交易数据，并且对其中的时间序列进行跟踪；

## 模型选择 1：图搜索

老鼠仓监管，可以跟踪一些关联账户中资金、股票的流动情况。有的流动网络比较隐蔽，并且涉及多个账户，需要进行图搜索。

## 量子计算：

- Grover算法对图搜索进行平方加速。

## 模型选择 2：神经网络与模式识别

关于晃骗的交易时间序列识别主要分为两种：一种是对已经出现过的交易犯罪，利用神经网络进行模式识别。

## 量子计算：

- 可以加速神经网络的训练过程，并且在数据表示上也有优势。

## 模型选择 3：异常检测

另外一种是利用异常检测对所有的交易进行监控预警。

## 量子计算：

- Grover搜索的相似度模型，基于HHL的线性模型都对异常检测算法有良好的加速效果。

内容来源：本源量子

# 量化金融算法

## 量子优化算法:

- 梯度下降法
- 模拟退火算法
- 量子退火算法

## 量子机器学习:

- 量子分类算法  
Grover搜索 → k-median算法,  
KNN算法, HHL → QSVM
- 量子回归算法  
制备量子态, HHL算法解线性方程组,  
利用相位估计算法求预测值...
- 量子主成分分析  
密度矩阵指数模拟, 振幅估计, qPCA
- 量子神经网络

## 量子蒙特卡洛模拟:

- 衍生品定价问题
- VaR值计算
- 投资组合优化

## 量子贝叶斯网络:

- 风控问题与复杂随机系统
- 经典贝叶斯网络
- 量子贝叶斯网络
- 量子贝叶斯网络应用

内容来源: 本源量子



A complex, abstract network of light blue lines and dots, resembling a neural network or a data visualization, is centered in the background. The lines connect various points, creating a web-like structure that fills the middle of the slide.

Thank

You