

# **Quantitative Investment Handbook**

Xinhe Liu

2018-2-28

# Contents

<b>1</b>	<b>Basic Financial Concepts</b>	<b>1</b>
1.1	Corporate Finance and Fundamental Analysis . . . . .	1
1.2	Asset Management . . . . .	1
<b>2</b>	<b>Portfolio Optimization Study</b>	<b>3</b>
2.1	Estimating Return and Co-variance Matrix . . . . .	3
2.2	Transaction Cost . . . . .	3
2.3	Tax . . . . .	3
2.4	Dynamic Portfolio Choice . . . . .	3
2.5	Covariance Matrix . . . . .	3
2.6	Easy . . . . .	4
2.7	Medium . . . . .	4
2.8	Hard . . . . .	4
<b>3</b>	<b>Statistical Arbitrage</b>	<b>5</b>
3.1	mean-reversion . . . . .	5
<b>4</b>	<b>Volatility Trading and Dispersion Trading</b>	<b>6</b>

<i>CONTENTS</i>	3
4.1 Volatility Models and Time-Series Models . . . . .	6
<b>5 Index and Smart Beta</b>	<b>7</b>
5.1 Smart beta and Smart Alpha . . . . .	7
5.2 Key things to Notice in Index Making . . . . .	7
5.3 factor analysis and risk premia . . . . .	7
5.4 Market Anomalies . . . . .	8
<b>6 Strategies Implementation/Platform Building</b>	<b>9</b>
6.1 Trading System . . . . .	10
6.2 Strategy Development Pipeline(Single/Linear Strategy) . . .	10
6.2.1 Basic Architecture/System . . . . .	10
6.2.2 Data . . . . .	13
6.2.3 Backtester/Simulator . . . . .	13
6.2.4 Trade Record and Money Management . . . . .	14
6.2.5 Analytics . . . . .	14
6.2.6 Research Team . . . . .	14
6.3 Alternative Data . . . . .	19



# Chapter 1

## Basic Financial Concepts

This chapter summarizes some basic financial concepts you should know about.

### 1.1 Corporate Finance and Fundamental Analysis

1. NPV and IRR
2. Discounted Cash Flow Model
  - Free Cashflow
  - Required Rate of Return = Cost of Capital = Risk-adjusted discount rate : Usually from the CAPM
3. Valuation using Multiples - P/E, P/B

### 1.2 Asset Management

1. Efficient Market Hypothesis (Weak, Semi-strong, Strong forms)
2. Markowitz Portfolio Optimization
  - Minimum Variance Portfolio / Tangency Portfolio

- Jensen's alpha

3. Capital Asset Pricing Model(CAPM) Model

$$r = \beta(r_m - r_f) + r_f$$

$$\beta = \frac{cov(r, r_M)}{var(r_M)}$$

4. APT(Arbitrage-Free-Pricing) Model

5. No-arbitrage(weak, strong) and Law-of-one-price

6. Metrics

- Sharpe Ratio
- Jensen's alpha
- Required Rate of Return = Cost of Capital = Risk-adjusted discount rate : Usually from the CAPM

7. Valuation using Multiples - P/E, P/B

## Chapter 2

# Portfolio Optimization Study

### 2.1 Estimating Return and Co-variance Matrix

### 2.2 Transaction Cost

### 2.3 Tax

### 2.4 Dynamic Portfolio Choice

### 2.5 Covariance Matrix

*Table 5.4*

**Covariance estimators**

Equally weighted	Exponentially weighted
$\sigma_{12}^2 = \frac{1}{T} \sum_{t=1}^T (r_{1t} - \bar{r}_1) (r_{1t} - \bar{r}_2)$	$\sigma_{12}^2 = (1-\lambda) \sum_{j=1}^T \lambda^{j-1} (r_{1t} - \bar{r}_1) (r_{1t} - \bar{r}_2)$

## 2.6 Easy

1. Two coins, one is double tail, you see one
  - three pancakes: golden-golden, golden-brunt, brunt-brunt, see one golden, probability that the other side is golden two
  -
2. Two coins - one toss  $n+1$ , one toss  $n$ , probability that  $n+1$  gets more tail than  $n$  man?

## 2.7 Medium

1. Use coins to create probabilities : fair coin to create  $1/3$  probability - toss twice, one result is retoss unfair coin to create  $1/3$  probability - combine two toss as one

## 2.8 Hard



## Chapter 3

# Statistical Arbitrage

### 3.1 mean-reversion

intraday mean-reversion

## Chapter 4

# Volatility Trading and Dispersion Trading

### 4.1 Volatility Models and Time-Series Models

Volatility Clustering and Leverage Effect

## Chapter 5

# Index and Smart Beta

### 5.1 Smart beta and Smart Alpha

### 5.2 Key things to Notice in Index Making

### 5.3 factor analysis and risk premia

value, momentum, quality, volatility, betaRM-RF The return spread between the capitalization weighted stock market and cash.

QQuality Minus Junk (quantity)

SMB The return spread of small minus large stocks (i.e., the size effect).

HML The return spread of cheap minus expensive stocks (i.e., the value effect).

RMW The return spread of the most profitable firms minus the least profitable.

CMA The return spread of firms that invest conservatively minus aggressively.

UMD (momentum/trend) UMD is long winners and short losers and also

from Ken French's website)

Carry Vol Carry

Special: liquidity premia

Market Inefficiency Analysis: funding constraint of financial institutions,  
grand move of large funds constraints

## 5.4 Market Anomalies

Behavioral Market Anomalies/bias

other: liquidity risk premia

1. Risk Management and Hedging
2. Leverage
3. Correlation
4. Strategy replacements, leverage rebalancing and rebalancing frequency, leverage reset

The following sectioning commands are available:

part  
chapter  
section  
subsection  
subsubsection  
paragraph  
subparagraph

But note that—unlike the `book` and `report` classes—the `article` class does not have a “chapter” command.



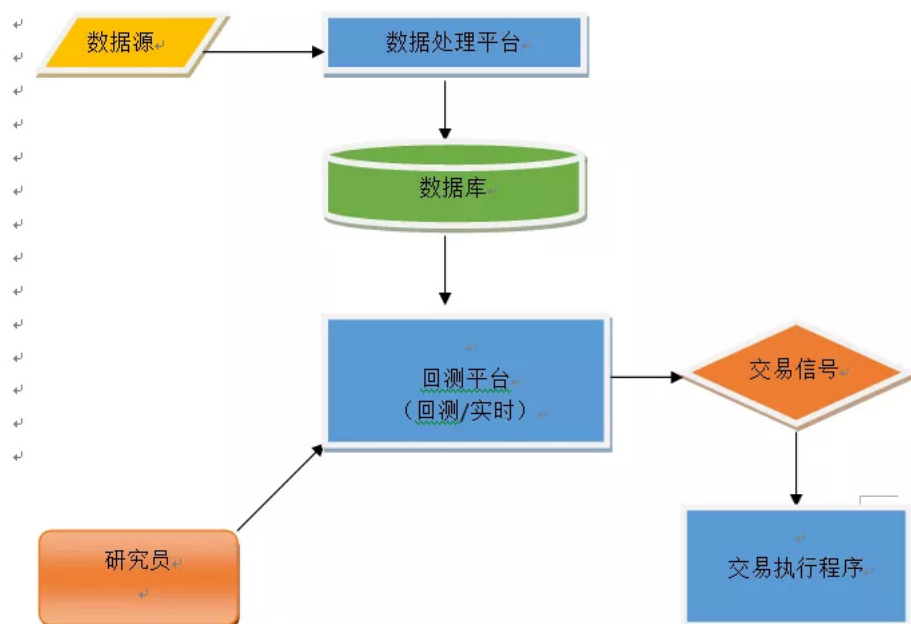
## Chapter 6

# Strategies Implementation/Platform Building

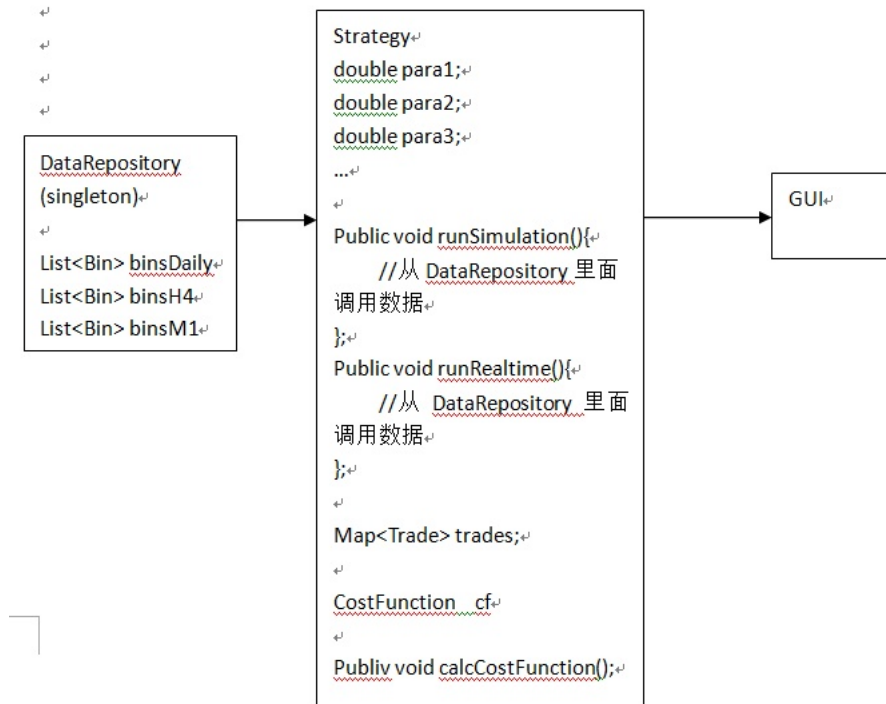
### 6.1 Trading System

### 6.2 Strategy Development Pipeline(Single/Linear Strategy)

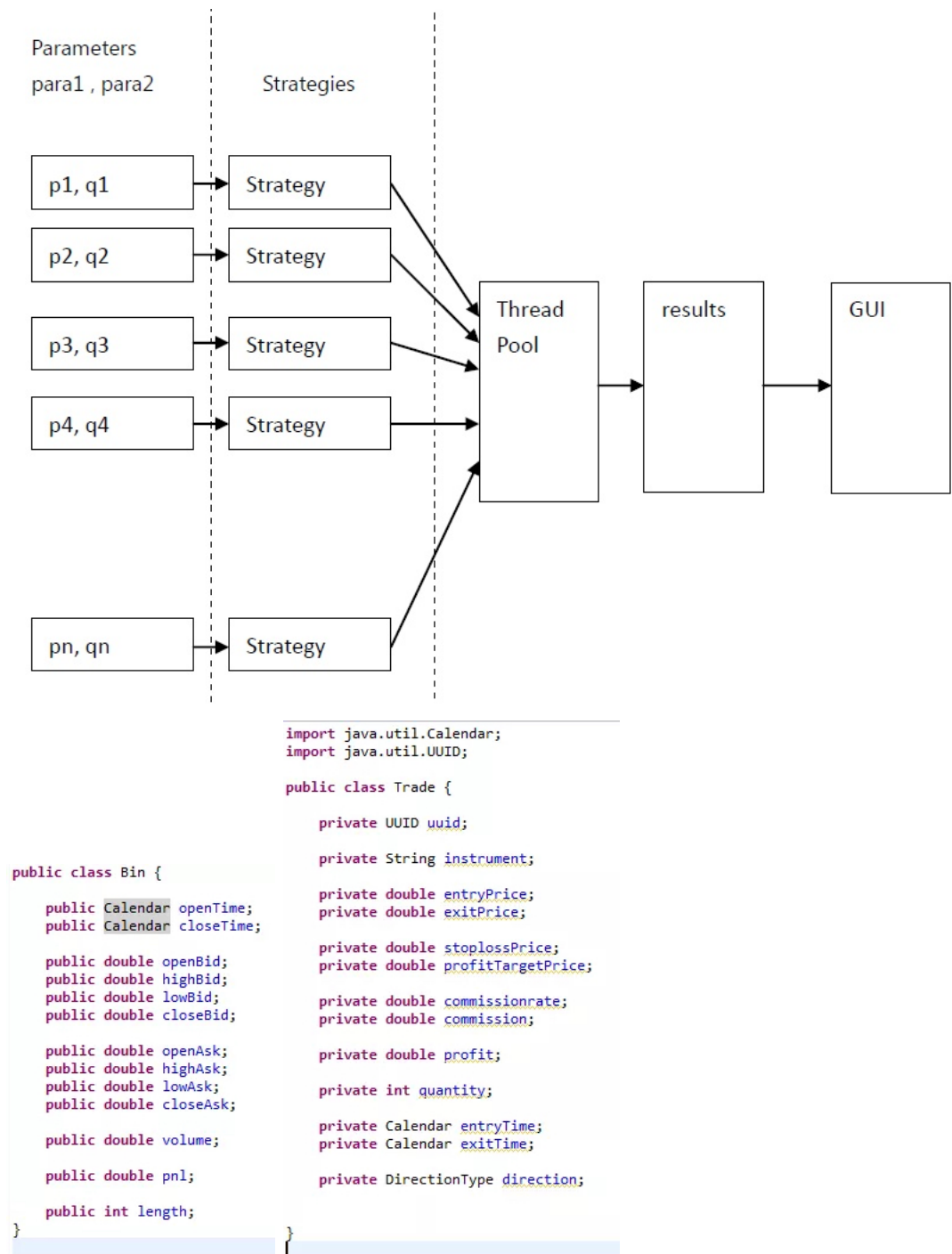
#### 6.2.1 Basic Architecture/System



## 6.2. STRATEGY DEVELOPMENT PIPELINE(SINGLE/LINEAR STRATEGY) 11



## 12CHAPTER 6. STRATEGIES IMPLEMENTATION/PLATFORM BUILDING





## 6.2. STRATEGY DEVELOPMENT PIPELINE(SINGLE/LINEAR STRATEGY) 13

```
for binsDaily {  
  for (Bin bin:binsM1) {  
    if(达到入场价){  
      //入场一单  
      Trade trade = new Trade()  
      Trades.put(trade);  
    }  
    For (trades){  
      If(打止损/打止损){  
        Trade.close()  
      }  
    }  
  }  
}
```

Historical Data is a singleton. All Market Data (eg. a candle stick ) should be organized to feed the researcher to program strategies on the backtester (eg. like quantopian). All back-testing should be parallized (ideally on GPU) to display parameter-profit relationship. (heatmap, stock charts, etc ) Ideally the optimization process could be visualized (like Tensorflow)

All like a research facility feedback cycle.

Key Details: API Design, Module Separation etc.

### 6.2.2 Data

Key is a real-time listener. Technical Considerations: KDB, Hadoop and HDFI(?), SQL Like, Mongo Db to store Archive Data. Market Data Providers consideration buying from Wind, BBG, Reuters, Etc.  
teams: platform operation engineer, analytics builder, strategy control/management and risk management, data team, execution team, researcher team ( 3 x tech )  
data licensing and data quality insurance  
data base, text file archive, big data issue  
cheap data: brokerage: interative brokers.

### 6.2.3 Backtester/Simulator

Key Components

\*Send Singal to Quoting/Trading/Exection Tool(Real Time) \*Market

Data Objects (eg. loop for every time bins) \*stop loss/risk control system integration \*parameter-backtest profit/statistics result: optimization and loss function set function to tune the parameters \*multi-thread: Java backter (Java thread pool\*) \*human selection of parameters: parameter table and visualization

#### 6.2.4 Trade Record and Money Management

record every trade, summarize execution shortfall, statistical trends and information (shortcomings of strategy executions) and market information ( learning material) build statistics and storage

More: order book and trade book level data handling

#### 6.2.5 Analytics

##### Strategy Management

, Sharpe Anslysis, Holding Period, Slippage visualization to better assistant strategic allocation

##### Execution Analysis and Cost

quantitative trading/systematic trading strategies: \* equity long/short

#### 6.2.6 Research Team

Key problems: \* Optimization and Combination of Sub-Strategies (Eg. factors) \* Market Regime Change Detection(problem not solved): Distinguish between trend and oscillation market \* market supply/demand imbalance analysis (risk-premia) \* volatility trading, dispersion trading - 2nd and 3rd degree trading, (vol model, vol clustering effect, vol leverage effect) \* hedging/overlay strategy research: hedging cost and hedging risk management, how to adjust hedge according to market condition. \* common ideas: market imbalance, mean-reversion, autocorrelation

## 6.2. STRATEGY DEVELOPMENT PIPELINE(SINGLE/LINEAR STRATEGY) 15

patterns etc (find patterns and trade) - based on statistics. Risk factors, implied arbitrage - based on math.

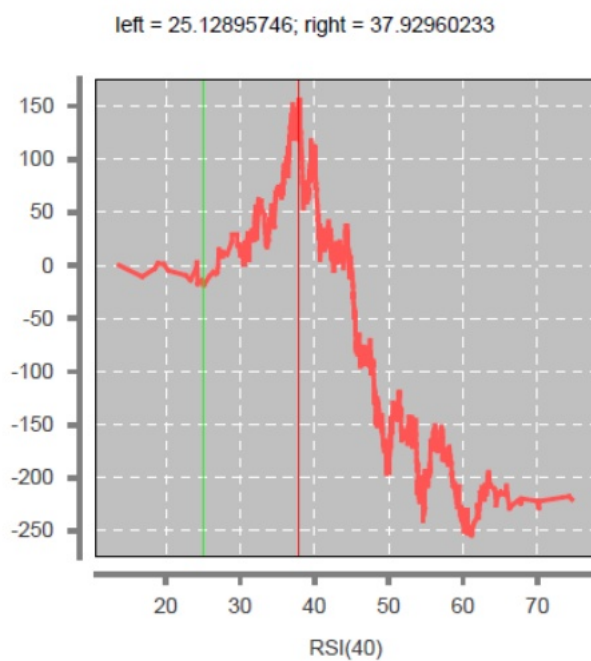
### Parameter Optmization and Control

Rely on GUI - parameter distribution and selection optmization  
methdologies from machine learning ( see optimization chapter) robustness  
anslysis and out-of sample test \*\* ( random cut the universe of rolling  
window on selection period )

### Signal Indicator Design

For example, based on fundamental ratio and technical indicators - design  
a formula. And check the level of prediction power (if any)

1.seeking stationarity: find a stationary time series use difference,  
integration, and normalize with volatility 2.find signal level, plot  
cumulative back-test return against different signal level (use own quotes,  
and use signal level to filter quotes)



3.Check stability of

customized indicators



4. check overfitting and type-II error in all settings, apply noise filtering if possible
5. design a interface to input indicator(math formula parser to read string) and visualize information using GUI.(HTML/XML Render)

## 6.2. STRATEGY DEVELOPMENT PIPELINE(SINGLE/LINEAR STRATEGY) 17

```
<?xml version="1.0"?>
<indicators>
  <section name="priceDivEMA">
    <indicator>
      <f>OpenDivEMA(\i)</f>
      <i>2,3,4,5,7,10,14,20,24,26,30,36,40,42,46,48,50</i>
    </indicator>
    <indicator>
      <f>CloseDivEMA(\i)</f>
      <i>2,3,4,5,7,10,14,20,24,26,30,36,40,42,46,48,50</i>
    </indicator>
    <indicator>
      <f>HighDivEMA(\i)</f>
      <i>2,3,4,5,7,10,14,20,24,26,30,36,40,42,46,48,50</i>
    </indicator>
    <indicator>
      <f>LowDivEMA(\i)</f>
      <i>2,3,4,5,7,10,14,20,24,26,30,36,40,42,46,48,50</i>
    </indicator>
  </section>
  <section name="ReturnEMA">
    <indicator>
      <f>ReturnOfOpenEMA5(\i)</f>
      <i>2,3,4,5,7,10,14,20,24,26,30,36,40,42,46,48,50</i>
    </indicator>
    <indicator>
      <f>ReturnOfCloseEMA5(\i)</f>
      <i>2,3,4,5,7,10,14,20,24,26,30,36,40,42,46,48,50</i>
    </indicator>
    <indicator>
      <f>ReturnOfHighEMA5(\i)</f>
      <i>2,3,4,5,7,10,14,20,24,26,30,36,40,42,46,48,50</i>
    </indicator>
    <indicator>
      <f>ReturnOfLowEMA5(\i)</f>
      <i>2,3,4,5,7,10,14,20,24,26,30,36,40,42,46,48,50</i>
    </indicator>
  </section>
</indicators>
```

```

public static HashMap<String, List<String>> parseXML(File fXmlFile) {

    HashMap<String, List<String>> map = new HashMap<String, List<String>>();

    DocumentBuilderFactory dbFactory = DocumentBuilderFactory.newInstance();
    DocumentBuilder dBuilder = null;
    try {
        dBuilder = dbFactory.newDocumentBuilder();
    } catch (ParserConfigurationException e) {
        e.printStackTrace();
    }
    Document doc = null;
    try {
        doc = dBuilder.parse(fXmlFile);
    } catch (SAXException | IOException e) {
        e.printStackTrace();
    }

    doc.getDocumentElement().normalize();

    NodeList nList = doc.getElementsByTagName("section");

    for (int i = 0; i < nList.getLength(); i++) {
        Node nNode = nList.item(i);
        if (nNode.getNodeType() == Node.ELEMENT_NODE) {
            Element element = (Element) nNode;
            String sectionName = element.getAttribute("name");
            map.put(sectionName, new ArrayList<String>());
            NodeList childList = element.getElementsByTagName("indicator");

            for (int j = 0; j < childList.getLength(); j++) {
                Node childNode = childList.item(j);
                Element cElement = (Element) childNode;
                String indicatorNameOrigin = cElement.getElementsByTagName("f").item(0).getChildNodes().item(0)
                    .getNodeValue();
                if (cElement.getElementsByTagName("i").getLength() > 0) {
                    String indicatorParaList = cElement.getElementsByTagName("i").item(0).getChildNodes().item(0)
                        .getNodeValue();
                    String[] indicatorParas = indicatorParaList.split(",");
                    for (String indicatorPara : indicatorParas) {
                        String indicatorName = indicatorNameOrigin.replace("\\i", indicatorPara);
                        map.get(sectionName).add(indicatorName);
                    }
                } else {
                    map.get(sectionName).add(indicatorNameOrigin);
                }
            }
        }
    }

    return map;
}

```

6. aggregate all indicators( eg. macd, ead ). Aggregate all strategies using optimization framework or selection framework to gain statistical alpha

7. indicator effectiveness test

1. test correlation - the correlation between indicator and profit vs. the correlation between correlation and white noise(hypothesis test) \* use spearman correlation rather than pearson correlation\*
2. Use Monte Carlo Simulation to do permutation test of effectiveness of indicator
3. Very very hard - detect sensensitivity to market regime change(osicallation and trend) and identify market regime change.

Integration of single indicators and portfolio theory

Form indicator as factors: standardization to mean-0, normal/t-distributed scores. Select powerful ones (ones that passed the permutation test).

Optimize to maximize holdings exposure to factor with risk penalty. The key is still feature engineering.

$$h'f - ah'Vh$$

$h$ :列向量, portfolio 里各个品种的持仓

$f$ :列向量, expected return

$a$  常数: risk aversion

$V$ : covariance matrix

$'$  意思是矩阵转秩

For Covariance, See section "covariance matrix".

### Strategy Risk Management and Money management

small stop loss, big stop gain level on reversion strategies. bigger stop loss, smaller stop gains on volatile markets - based on experience, market analysis.

Choose symmetric/non-symmetric risk control based on market belief

Hedging and Market Exposure Management - Volatility Control and Automatic de-leveraging.

together with cost consideration.

## 6.3 Alternative Data