AI for Trading

Term 1 Notes by Pranjal Chaubey

https://www.linkedin.com/in/pranjallchaubey/

https://github.com/pranjalchaubey

UDACITY

3rd November 2018

Stocks — Limited Liability Companies

- Common Stock
  - Dividends
  - Vote on Decisions
  - Receive remainings assets in case of liquidation
- Preferred Stock
  - fixed income each year
  - receive dividends before common Stock holders
  - usually no voting rights

Security — Debt — money that is owed and must be repaid

Derivative — values depend on the prices of other assets

Equity → (Value of an Asset — All debt or liabilities on that asset)

└ net value of something owned

Financial Instrument that has some monetary value

Derivative Security — Options Contracts

└ buy/sell asset at a specified price on or by a specified date

Futures Contracts

└ buyers/sellers are obligated to buy/sell at a predetermined price at a specified time in the future.

Market Maker creates LIQUIDITY in the ~~market~~ market.

Large volume of buy orders — price goes up ↑

Large volume of sell orders — price goes down ↓

6 September 2018

STOCK SPLIT Adjustment

2:1 stock split

$100 ⟶ $50
old   new

to Normalize, divide share value before stock split by 2.

DIVIDENDS

Ex dividend date — last date to buy the stock to receive dividend

on dividend date (1$ Dividend)
↳ stock went from $50 to $49.50
  ↳ but 1$ Dividend implies ⟹ $0.50 profit

Normalizing stock prices BEFORE dividend, divide by Adjusted Price Factor

$$APF = 1 + \frac{D}{S}$$

D — Dividend
S — Stock Price at ex-dividend date

Simple Moving Average or Rolling Mean

- If the price deviates too much from the mean
  - Long or Short the stock
- +ve & -ve Standard Deviation around the Mean
  - BOLLINGER BANDS
    - 2 SDs above & below the Mean

Fundamental Analysis of a Company

- Every quarter — Balance Sheet & Cash Flow Statements
- Sales per Share → Revenue divided by no. of shares (cost of sales not included)
- Earnings per Share → Revenue − Cost of Sales
  - Change in equity over the past 3 months
    - Stock is the fractional ownership of a company's equity
- Dividends per Share — $\dfrac{\text{Earnings decided to be given away by the government}}{\text{Total number of Shares outstanding}}$
- Price to Earnings (PE) Ratio — $\dfrac{\text{Current Stock Price}}{\text{Most Recent Reported Earnings per Share}}$
  - Use to compare health of similar companies in ~~[illegible]~~ same geographic region.

ETF Compositional Data

- Composition of stocks in an ETF
  - Use it to find correlation b/w stock prices

Market Data
Corporate Actions
Fundamental Information
Compositional Data

} STANDARD INFORMATION that everyone uses.

7th September 2018

Stock Raw Return (r) $= \dfrac{P_t - P_{t-1}}{P_{t-1}}$

LOG RETURNS

$$R = \log_e\left(\frac{P_t}{P_{t-1}}\right) = \ln(P_t / P_{t-1})$$

Continuously Compounded Return!

$$\underbrace{\frac{P_t - P_{t-1}}{P_{t-1}}}_{r} + 1 = \frac{P_t}{P_{t-1}} \Rightarrow$$

Log Return

$$R = \ln(r+1)$$

$$r = e^R - 1$$ Raw Return

$$|r| \ll 1 \Rightarrow \ln(r+1) \approx r$$

4% interest rate compounded anually on \$100

- $\$100 \times (1+0.04) = \$104$
- Semi-annual compounding (every 6 months)
  - $\$100 \times (1+.04/2) \times (1+.04/2)$
  - $= \$100 \times (1+0.04/2)^2 = \$104.04$
- Quaterly compounded
  - $\$100 \times (1+.04/4)^4 = \$104.06$
- Daily $(n=252) \Rightarrow \$104.08$

Log Returns don't suffer from the problem of ARITHMETIC UNDERFLOW.

Compounding Formula

$$P_t = P_{t-1}\left(1+\frac{r}{n}\right)^n$$

$$\lim_{n\to\infty}\left(1+\frac{r}{n}\right)^n = e^r$$

r = Continuously Compounded Annual Return

Continuous Compounding

Additive over time

- Compounded continuously
  - $\$100 \times e^{0.04} = \$104.08$
- Annualized monthly compounded return
  - $\$100 \times e^{0.04 \times 12}$
- log return for Jan + log return for feb

$$\ln\left(\frac{P_{feb_1}}{P_{jan_1}}\right) + \ln\left(\frac{P_{mar_1}}{P_{feb_1}}\right)$$

$$\Rightarrow \boxed{\ln\left(\frac{P_{mar_1}}{P_{Jan_1}}\right)}$$

8th September 2018

Long-term prices and cumulative returns can be modeled as approximately lognormally distributed

- ∵ products of Independently Identically Distributed (IID) Random Variables.
- Log returns SUM over time
  - Sum of long-term log returns will be
    - 'NORMAL'
    - Even if the individual returns are not 'normal'
      - Long-Term sum will be approximately NORMAL.
        - CENTRAL LIMIT Theorem
          - Sum of Random Variables having the same distribution that are not related to each other approaches a NORMAL DISTRIBUTION

$$\text{for } \underbrace{n}_{\text{no. of Random Variables}} \longrightarrow \infty$$

$$\underbrace{Y \sim \text{normal}}_{Y \text{ distributed 'normally'}} \longrightarrow \underbrace{e^{Y} \sim \text{log-normal}}_{e^{Y} \text{ then distributed log-normally}}$$

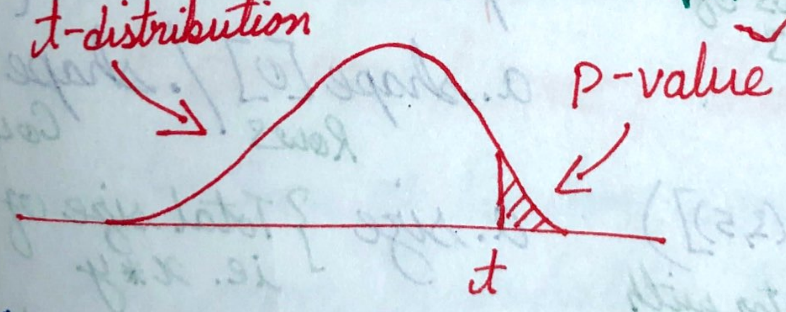
t-Statistic or t-test

$$t = \frac{\bar{x}}{SE_{\bar{x}}}$$

$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ — Mean of the strategy returns. Monthly returns in this case.

$SE_{\bar{x}} = \frac{\text{standard deviation}}{\sqrt{n}}$ — Standard Error of the Mean

n — No. of observations in the sample



Sample Mean $= \bar{x}$

Sample Standard Deviation, $s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$

Standard Error $s_{\bar{x}} = \frac{s}{\sqrt{n}}$

if, null hypothesis $= 0$ i.e. population mean $= \mu = 0$

t-statistic, $t = \frac{\bar{x}-\mu}{s_{\bar{x}}} = \frac{\bar{x}-0}{s_{\bar{x}}} = \frac{\bar{x}}{s_{\bar{x}}}$

$$t\text{-score} = \frac{\text{difference b/w two groups}}{\text{difference within the groups}}$$

- large number → more difference — This is what you want
- small number → more similarity
- p-value ∈ t-value
  - Indicates if the result happened by chance
    - Small p-value ⇒ results did NOT happen by chance
      - 0.05 ✓

14th September 2018

[redacted]

1. Single Asset Strategies
2. Pairwise Strategies
3. Cross-sectional strategies → HIGHEST CAPACITY
   (equity statistical arbitrage, equity market neutral investing)
4. Alternative Data based Strategies
   - Satellite Imagery
   - Social Media
   - Geolocation
   - Consumer Transaction Data

CROSS SECTIONAL EQUITY INVESTING

DATA

Universe Definition

Alpha 1 .. Alpha N

Alpha Combination → Portfolio Construction

Risk Model → Portfolio Construction

Objective Function

Constraints → Portfolio Construction

Prior Portfolio → Portfolio Construction

Portfolio Construction → Ideal Portfolio

Portfolio Construction $\xrightarrow{\text{Trade List}}$ Trading

ALPHA — An expression that outputs a vector where each component is a value indicative of future returns for an individual stock.

└ Several Alpha combined to have better performance than the best alpha

└ Combination
- Adding weights
- Averaging
- Lowest Possible variance of the combined alpha — finding weights
- Turn α vectors into features and use them as inputs in ML classifier to capture the relationships of individual alpha

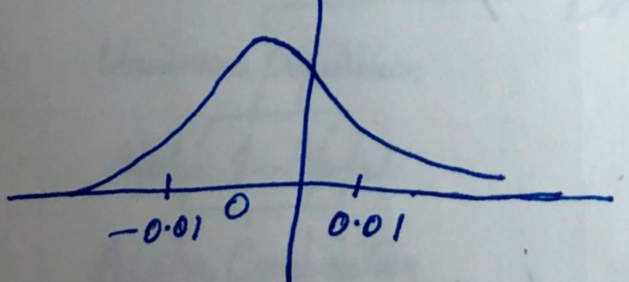
In Finance
└ Risk = uncertainty
└ Variability of returns

RISKS

- Systematic Risks
  - Inflation, recession, interest rates etc.
    - Inherent to entire market
  - Sector specific risks
- Idiosyncratic Risks
  - Inherent to individual stocks

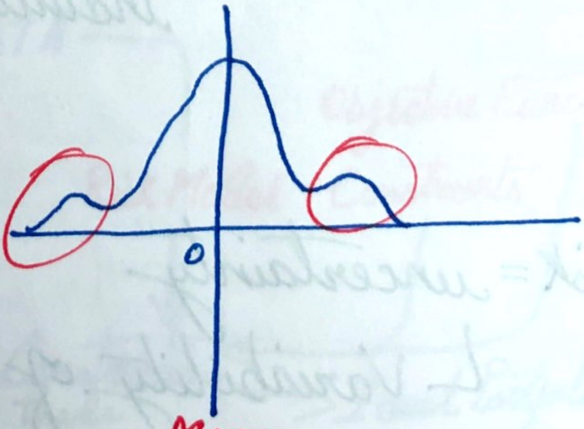
OUTLIERS

- Manual Errors due to data entry
  - FAT FINGER ERRORS
- Rules to filter out outliers
- Extremely skewed shapes or bumps in the tail of return distribution spell trouble



Mean

Too much positively skewed.



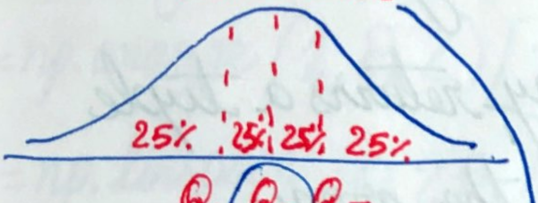
Mean

Q-Q plot

└ Plot of QUANTILES of the first dataset vs. the Quantiles of the second dataset

QUANTILES
- Quartiles – 4 divs
- Deciles – 10 divs
- Percentiles – 100 divs

Quartiles



Median

Divides the distribution in 4 parts

DECILES

Divides the data in any number of groups

$\frac{\text{SIGNAL}}{\text{NOISE}} = \text{LOW}$ } Predictive models tend to overfit the data they are trained on. } If SNR is LOW

RANDOM Variable's value is determined by its PROBABILITY DISTRIBUTION → Tennis Ball machine analogy.

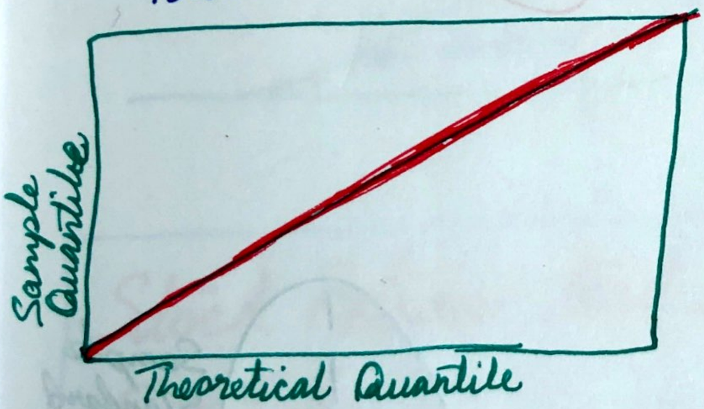
Normal Distributions are symmetric about their mean. → Median = Mean

└ Box PLOT to test for Normality

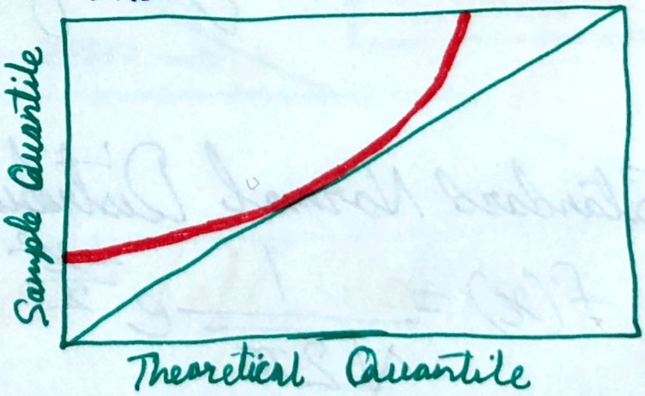
Q-Q Plots

6th November 2018

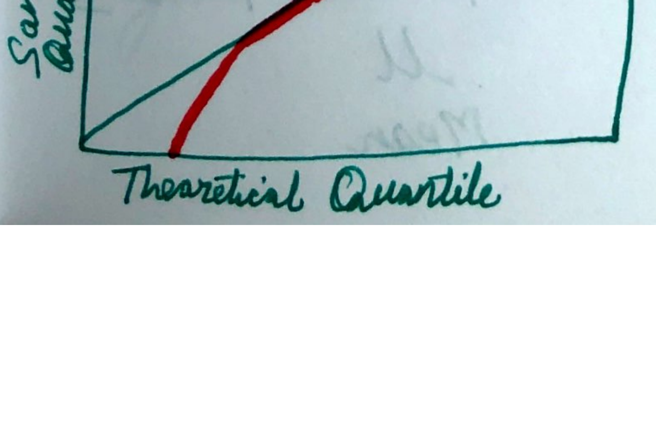
Symmetrical Plot – No Profit No Loss scenario



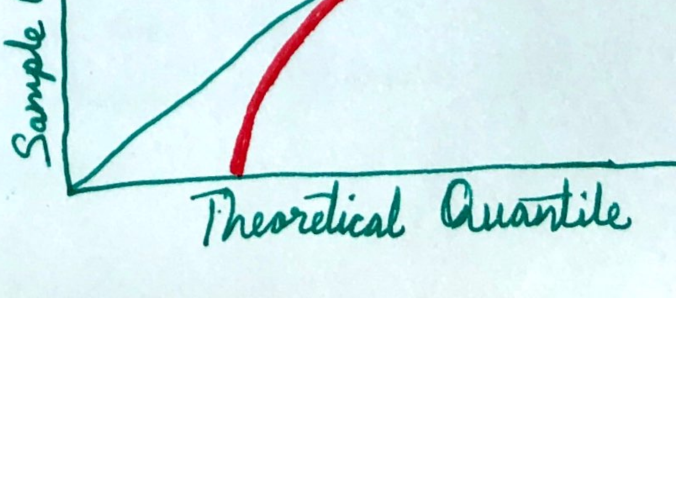
Positive Skew – Healthy Alpha / Profitable Strategy



Normal Distribution has fat tails – RED FLAG



Negative Skew – Strategy will run into losses.



9th November 2018

Probability Density Function (PDF)

$$X \sim D$$

Random Variable X → follows → Probability Distribution D

Probability Distribution

$$\underbrace{P(x \mid D)}_{\text{Probability of 'x', given 'D'}} = p(x)$$

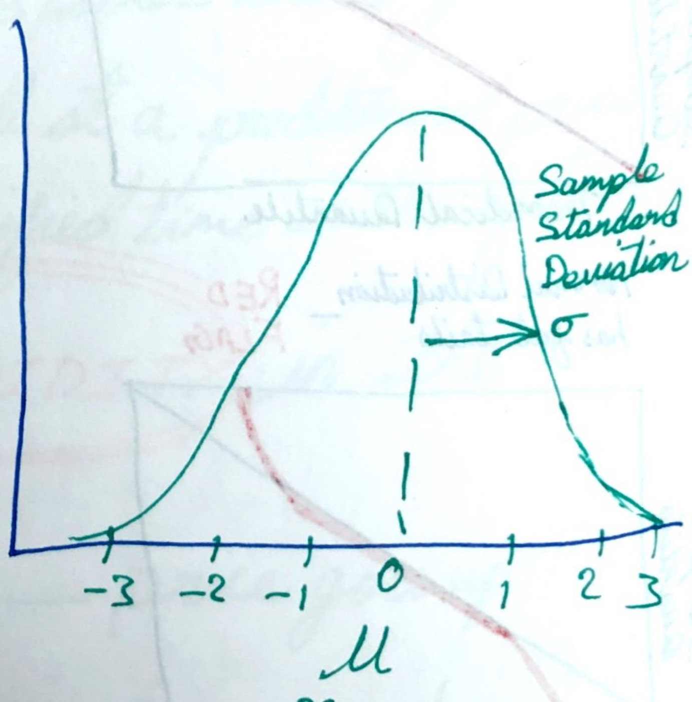
Probability of 'x', given 'D'

Standard Normal Distribution

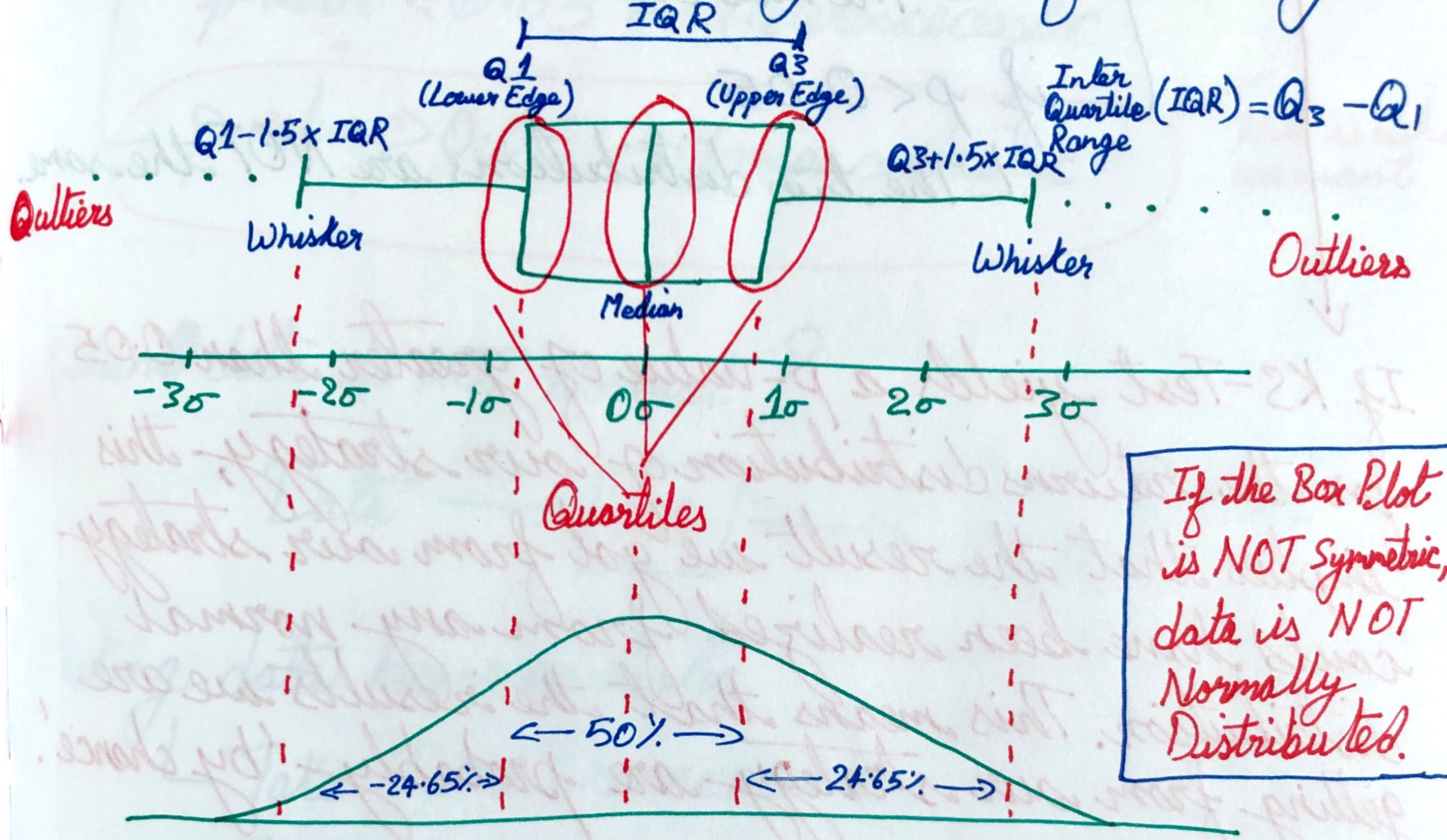
$$f(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$$

Normal Distribution

$$f(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



Normal Distributions are symmetric about their mean.
└ Use Box Plots to visually check for normality



If the Box Plot is NOT Symmetric, data is NOT Normally Distributed.

Stock Returns tend to exhibit Left Skew and Fat Tails.

Shapiro-Wilk
D'Agostino-Pearson
} p-value > 0.05 → Normal Distribution
p-value ≤ 0.05 → Not Normal Distribution

Kolmogorov-Smirnoff — compares if two distributions are similar or not. We will compare our returns distribution against the Normal Distribution.

KS Test — if $p > 0.05$
└ The two distributions are the same.

if $p < 0.05$
└ The two distributions are NOT the same

If KS-Test yields a p-value of greater than 0.05 for the returns distribution of our strategy, this implies that the result we got from our strategy could have been realized from any normal distribution. This means that the results we are getting from our strategy are probably 'by chance'.

Stationary Data?

- Mean
- Variance
- Covariance

} Should be same over time

in particular, variance of the data should be stable over time.

Homoscedasticity — Constant Variance over Time

Heteroscedasticity — Changing variance over time

Breusch - Pagan Test

- p-value < 0.05 — Heteroscedastic
- p-value > 0.05 — Homoscedastic

This is what we want.

When data is NOT Normal?

Data → $\boxed{\text{LOG}}$ → 'More' Normal Data

Making data Homoscedastic

- Take Time Difference

$$\frac{\text{Friday}}{\text{Thursday}} \rightarrow \text{Log} \rightarrow \text{Log Returns!}$$

Box-Cox Transformation — makes data both normally distributed, and homoscedastic.

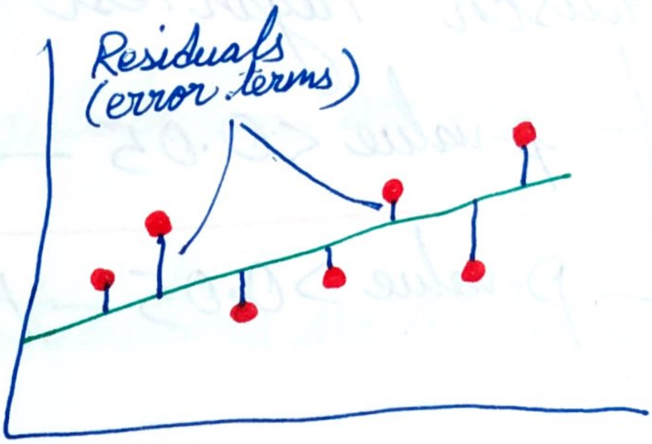
$$T(x) = \frac{(x^{\lambda} - 1)}{\underbrace{\lambda}_{\text{Play around with this value}}}$$

if $\lambda = 0$

$\Rightarrow T(x) = \ln(x)$

Regression

$$y = \beta x + \alpha$$



Residual = $y_{actual} - y_{predicted}$

follows a Normal Distribution

└ RANDOM — model's predicted value is equally likely to be higher/lower than the actual value

Multiple Regression — more than one independent variable is used to predict the dependent variable.

$$y = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \alpha$$

To evaluate regression model,

- R-squared — 0 to 1 → 1 ⇒ All variation in the dependent variable can be explained by variations in the independent variables.
- Adjusted R-Squared — Minimum combination of independent variables most relevant for our model.
- F-Test — p-value ≤ 0.05 → Our model describes a meaningful relationship.
  - └ parameters are not 0

Multivariate Multiple Regression

- Multiple Independent variables predicting multip[le] Dependent variables.

× ———————————— ×

10th November 2018

Time Series

- Autoregression
- Moving Averages
- Autoregressive Moving Averages
- Autoregressive Integrated Moving Averages
- Kalman and Particle filters
- Recurrent Neural Networks

- Use past data to predict future values
- if properties of data change over time
  - Past data becomes less useful
- ∴ We use stock returns
  - LOG OF STOCK RETURNS

} NEED STATIONARY DATA

Autoregressive (AR) Model

Coefficients

$$y_t = \alpha + B_1 y_{t-1} + B_2 y_{t-2} + \ldots + \epsilon_t$$

Dependent Variable

Intercept

Independent Variable

Error (noise)

use Adjusted R^2 values to measure how well the model is predicting

Wednesday Tuesday Monday

└ Lag 2 AR model ⟹ AR(p) ⟶ AR(2)

$$y_{1,t} = \alpha_1 + B_{1,1} y_{1,t-1} + B_{1,2} y_{2,t-2} + \ldots + \epsilon_{1,t}$$

$$y_{2,t} = \alpha_2 + B_{2,1} y_{1,t-1} + B_{2,2} y_{2,t-2} + \ldots + \epsilon_{2,t}$$

Multiple Time Series

if price of the stock is related to price of another stock?

└ Vector Autoregressive Model (VAR)

$$\begin{tikzcd}
y_{1,t} & & y_{1,t-1} \arrow[ll] \arrow[dll] \\
y_{2,t} & & y_{2,t-2} \arrow[ll] \arrow[ull]
\end{tikzcd}$$

Moving Average (MA) Model [Moth & Lantern Analogy]

$$y_t = \mu + \epsilon_t + \underbrace{\theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2}}_{} + \dots\dots + \theta_q \epsilon_{t-q}$$

Average

Linear combination of $\boxed{\text{residuals}}$ from previous time periods

Represent new unpredictable information that cannot be modeled

$MA(q) \longrightarrow$ q is the LAG

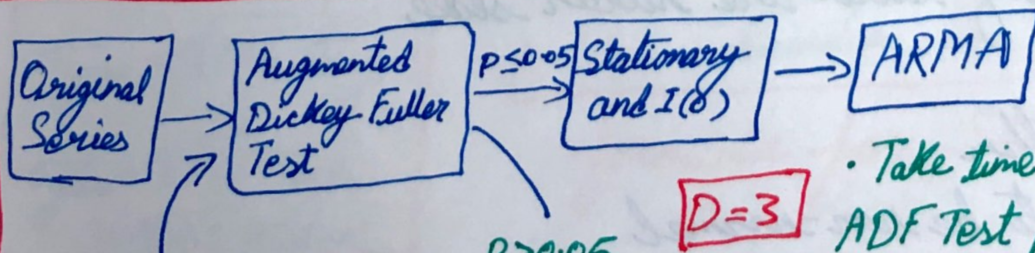
Use autocorrelation plot to determine the best value for 'q'

Autoregressive Moving Average ARMA(p, q)

$$y \sim AR(p) + MA(q)$$

$$y_t = \alpha + B_1 y_{t-1} + B_2 y_{t-2} + \dots$$
$$+ \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots$$

Autoregressive Integrated Moving Average (ARIMA)

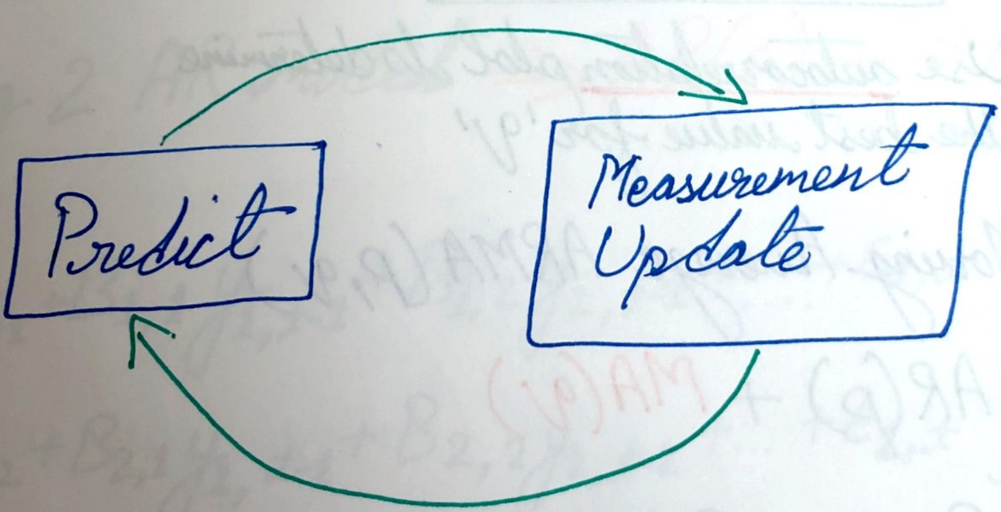


$\boxed{D=3}$

- Take time difference 'D' till ADF Test p-value ≤ 0.05
- Then current time difference is integrated of order 0 → I(0)
- Original Time Series was integrated of order D.

KALMAN FILTER

- in ARMA, we manually set p & q values
- in Kalman Filter, we have a set of variables at T-1 to represent the past.
  - in layman terms, KF automatically keeps the best p & q values.
- KF excels in predicting values in very noisy environments

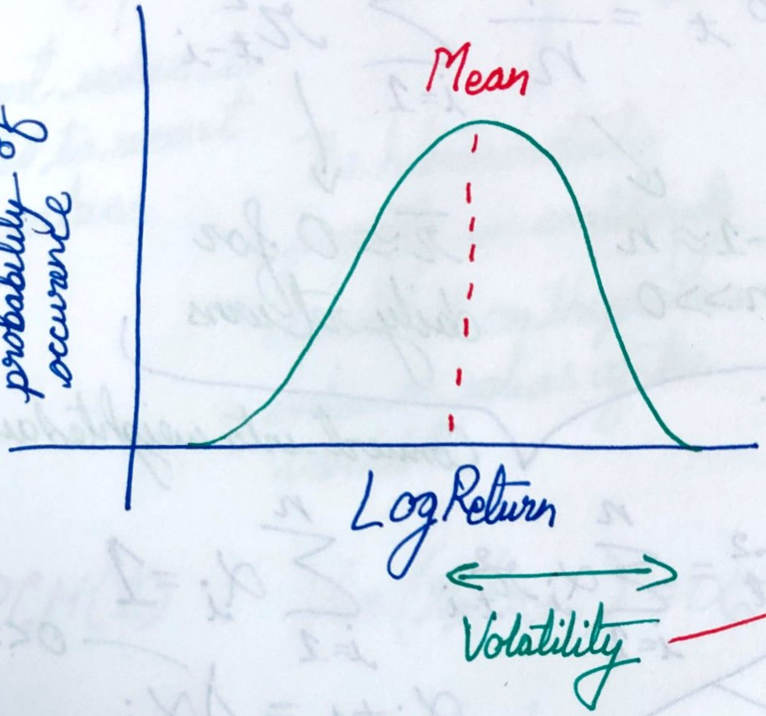


Previous time period state
+
Current measurement
↓ predicts
Next State

1. KF predicts the stock return as a probability distribution
2. Measures actual stock returns
3. Updates its belief about the hidden state

Particle Filters
- Genetic Algorithm
- Assumes data to be non-normal
- Assumes data to be non-linear

VOLATILITY — variability of returns.



Volatility is the measure of the spread of returns distribution.

Standard Deviation

Log Returns SUM over time.

Calculating Volatility

$$r_i = \ln\left(\frac{P_i}{P_{i-1}}\right)$$

price at time i

log return at time i

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\bar{r} - r_i)^2}$$

mean log return

Volatility

number of log return observations

log return at time i

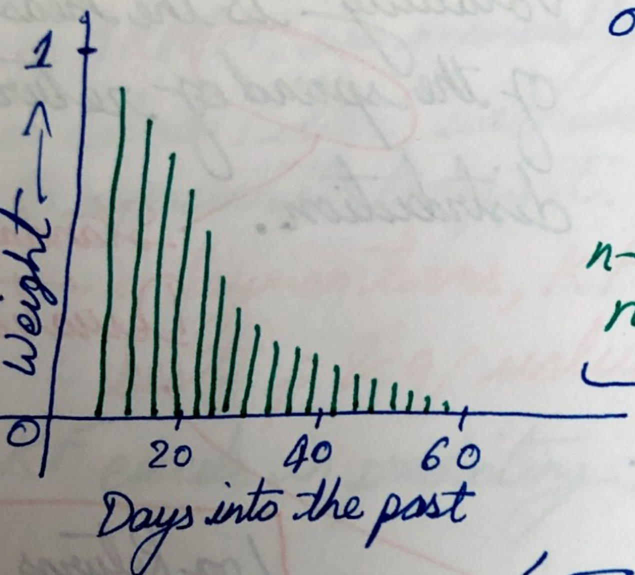
formula for Standard Deviation

$\sigma_{year} = \sqrt{252}\,\sigma_{day}$

$\sigma_{year} = \sqrt{12}\,\sigma_{day}$

$= \sqrt{52}\,\sigma_{week}$

Exponential Moving Average



$$\sigma_t^2 = \frac{1}{n}\sum_{i=1}^{n} r_{t-i}^2$$

$n-1 \approx n$
$n >> 0$

$\bar{r} \approx 0$ for daily returns

Convert into weighted average

$$\sigma_t^2 = \sum_{i=1}^{n} \alpha_i r_{t-i}^2 \qquad \sum_{i=1}^{n} \alpha_i = 1$$

$0 < \lambda < 1$

$$\alpha_{i+1} = \lambda \alpha_i$$

$$\sigma_t^2 = \frac{r_{t-1}^2 + \lambda r_{t-2}^2 + \lambda^2 r_{t-3}^2 + \ldots + \lambda^{n-1} r_{t-n}^2}{1 + \lambda + \lambda^2 + \lambda^3 + \ldots + \lambda^{n-1}}$$

Exponentially weighted, moving average estimate of the variance of the log returns.

Volatility is easier to predict than price.

Autoregressive Conditionally Heteroscedastic (ARCH)

Autoregressive — Current value is relates to recent past values

Conditionally — The heteroscedastic property is conditionally dependent on the previous value or values of the variable.

Heteroscedastic — Variable being modeled might have variability that changes with time. Variability is measured using variance.

ARCH(1) $\quad Var(r_t | r_{t-1}) = \alpha_0 + \alpha_1 r_{t-1}^2$

ARCH(m) $\quad Var(r_t | r_{t-1}, r_{t-2}, \ldots, r_{t-m})$

$$= \alpha_0 + \alpha_1 r_{t-1}^2 + \alpha_2 r_{t-2}^2 + \cdots + \alpha_m r_{t-m}^2$$

GARCH — current estimate also depends on the previous estimates of the variance.

GARCH — Generalized ARCH model

$$\sigma_t^2 = Var(y_t | y_{t-1}, y_{t-2}, \ldots, y_{t-m}, \sigma_{t-1}^2, \ldots \sigma_{t-n}^2)$$

GARCH(m, n) — m: no. of Log Return terms; n: no. of variance terms

$$= \alpha_0 + \alpha_1 y_{t-1}^2 + \alpha_2 y_{t-2}^2 + \cdots + \alpha_m y_{t-m}^2$$

$$+ \beta_1 \sigma_{t-1}^2 + \beta_2 \sigma_{t-2}^2 + \cdots + \beta_n \sigma_{t-n}^2$$

Position Sizing based on volatility

$$\text{Position Size} = \frac{R}{\sigma \times M \times \text{Last Close}}$$

R — Amount the trader is willing to lose

σ — Annualized volatility of the security or strategy in question

M — trader defined integer

Last Close — Last Closing price of the security

Position Size — The $\boxed{\text{number of shares}}$ to trade

MEAN REVERSION

└ MR time series moves back and forth around some constant value.

Drift and Volatility Model

$$dP_t = \underbrace{P_t\, \mu\, dt}_{\text{Drift Term}} + \underbrace{P_t\, \sigma\, \epsilon \sqrt{dt}}_{\text{Volatility Term}}$$

dP_t — Change in price over time 't'

Drift Term
- P_t – Current Price
- μ – Constant Average Term
- dt – Change in time

Volatility Term
- P_t – Current Price
- σ – Standard Deviation
- ϵ – Random Noise Factor
- $\sqrt{dt}$ – Square root of change in time

Relative difference between Pairs of Stocks

- Spread and Hedge Ratio

Stock A	Stock B
Long	Short

} Combined value of both the positions is constant over time

Neutral Position

11th November 2018

HEDGE RATIO — creates the 'Neutral Position'.

Price Ratio

$\frac{B}{A}$

Regression

$B = \beta A + \alpha$

Hedge Ratio

$\text{Spread} = \underbrace{B_{actual}}_{} - \underbrace{B_{estimate}}_{\beta A + \alpha}$

Remember error term from linear regression?

Spread is constant

'MIGHT' be a good pair to trade

There should be some economic link among the stocks.

When a Time Series is Stationary, it is also integrated of order 0.

Log Return Series : $I(0)$

Find some linear combination of a pair of stocks

$y_t \sim I(1)$

$x_t \sim I(1)$

Series integrated of order 1

$$y_t \approx \alpha + \beta x_t \Big\} \text{ Hedge Ratio}$$

$$\text{Spread} = y_t - (\alpha + \beta x_t)$$

If,

Spread is Stationary → Spread is $I(0)$
└→ $\boxed{x \text{ and } y \text{ are cointegrated}}$

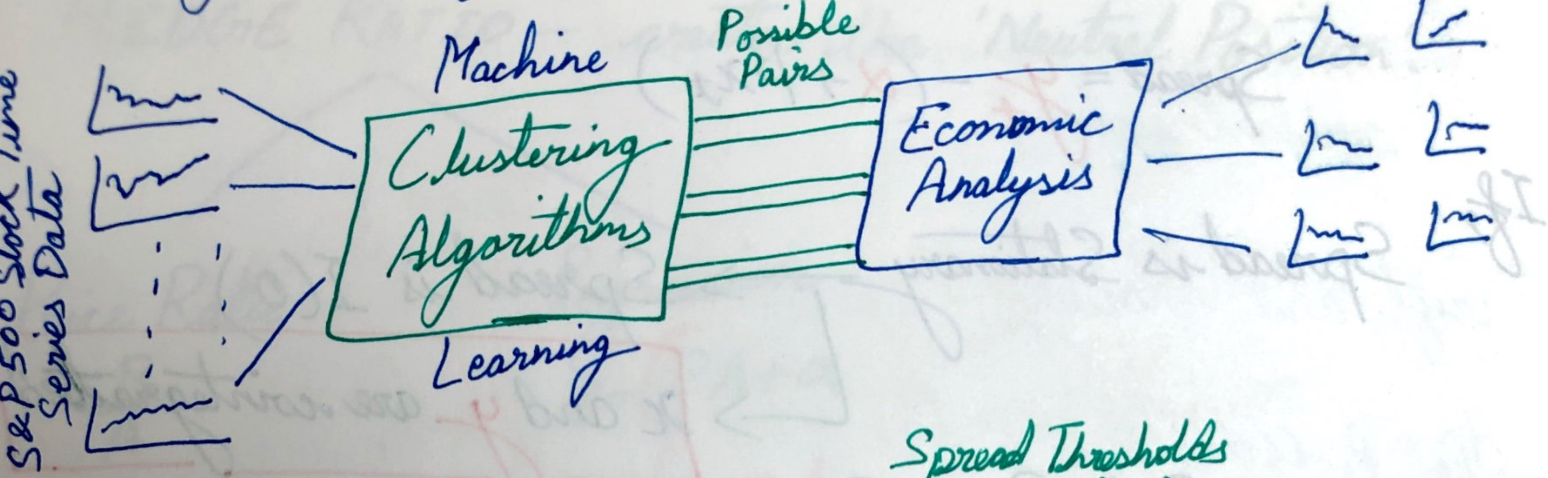
Hedge Ratio β is Coefficient of Cointegration

$$\boxed{\text{Correlation} \neq \text{Cointegration}}$$

Augmented Dickey-Fuller Test — Tests Stationarity
└ p-value ≤ 0.05 → Series is Stationary
└ ~~Stocks are~~ may be cointegrated

Engle-Granger Test — Tests time series cointegration
└ Calculate hedge ratio by running regression
$$y_t = \beta x_t$$ Hedge Ratio
└ $z_t = y_t - \beta x_t$
└ Run ADF on z_t
└ If z_t is stationary, x & y are cointegrated

Finding pairs of stocks to trade



Spread Widens → Short the spread

Short higher asset | Long Lower Asset

Spread Narrows → Go Long the spread

Short higher Asset | Long Lower Asset

Spread thresholds are defined by the z-score.

$$\text{z-score} = \frac{(x - \text{mean})}{\text{Standard Deviation}}$$

Stocks, Indices, Funds

7th December 2018

$\boxed{\text{Equity} = \text{Assets} - \text{Liabilities}}$

Indices help us keep track of 1000s of companies at once.

- Aggregated value of a group of stocks.
- Virtual Portfolio, NOT an actual fund.
- Specific to
  - Stock Exchange
  - Country
  - Sector
- used for
  - Market Information
  - Benchmarks by Portfolio Managers
- Grouped by
  - Large Cap
  - Mid Cap
  - Small Cap

Growth stock

Value stock

S&P : Standard & Poor's 500 index } USA
Dow : Dow Jones Industrial Average

IBOVESPA : Ibovespa Brazil Sau Paulo Stock Exchange BAZIL

MERVAL : Buenos Aires SE Merval Index ARGENTINA

NIKKEI : Nikkei 225 Index JAPAN

Hang Seng : Hang Seng Composite Index HONG KONG

FTSE 100 : Financial Times SE100 Index UK

EURO STOXX : EURO STOXX 50 EUROPE

CSI 300 : Stocks on Shanghai & Shenzhen SE CHINA

NASDAQ 100 Technology Index : Tech Sector } USA
NASDAQ Financial 100 : Finance Sector

S&P MidCap 400 : 400 Growth + 400 Value } USA
S&P SmallCap 600 : 600 Growth + 600 Value

Fidelity Contrafund : Seeks to outperform S&P 500

Vanguard 500 Index Investor Fund : Tracks S&P 500

Vanguard Equity Income Fund
T. Rowe Price Bluechip Growth Fund
BlackRock Technology Opportunities Fund

BlackRock iShares China Large-Cap ETF
L FXI - Top 50 Stocks on Hong Kong Stock Exchange
State Street Global Advisors SPDR S&P 500 ETF
Fidelity MSCI Information Tech ETF

Growth or Value stock?

- Look at valuation metrics
  - Price to Earnings Ratio
  - Price to Sales Ratio
  - Price to Book Ratio
  - Rank Stocks accordingly
    - Growth Stocks have higher ratios
    - Value stocks usually have lower ratios

Price Weighting

- Add up all the prices

$$\sum_1^n price_i = \text{some index number}$$

Market Cap Weighting

- Add up the MARKET CAP of all the stocks in the index

$$\sum_1^n (price_i \times \text{number of shares}_i)$$

- S&P500, IBOVESPA, MERVAL, HANG SENG, FTSE, EURO STOXX

During an index add or delete event
- Index needs to be rebalanced
  - Perform the index update as usual

HANG SENG Index
- Capped Free Float Adjusted Market Cap Weighted Index
  - Index uses only the shares available in the market for trading to calculate the Market Cap
    - Founders' share might be locked for some time period
    - Employees with unvested ESOPs might not be able to sell etc. etc.

Funds
- Professionally managed portfolios
- Diversification improves risk vs. return (SHARPE Ratio)
- Target is to outperform index
  - Actively Managed Fund
- Target is to match performance of an index
  - Passively Managed Fund (Index Funds)

Active Funds → Alpha Funds (Generate excess return above market return)
└ Active Returns

Passive Funds → Beta Funds
└ Tracking Error

$$\boxed{\text{Excess Return} = \text{Portfolio Return} - \text{Risk Free Rate}}$$

$\beta = 1$
- ↳ +2% market
  - ↳ +2% portfolio
- ↳ −2% market
  - ↳ −2% portfolio

Active + Passive Funds → SMART BETA
└ Control portfolio volatility & Risk

Mutual Funds
- for everyday investors
- Long Only No Shorting
- No Lock-up Period

Hedge Funds — High Networth Individuals or Institutions (Pension Funds for eg.)
- Long and Short
- Can trade Derivatives (Options, Futures)
- Lockup Period

Fund Performance
- RELATIVE RETURNS ⇒ Fund Returns - Benchmark Returns
  - Evaluated Relative to the Benchmark
  - Active and Passive Mutual Funds
- ABSOLUTE RETURNS ⇒ Measured by fund performance alone.

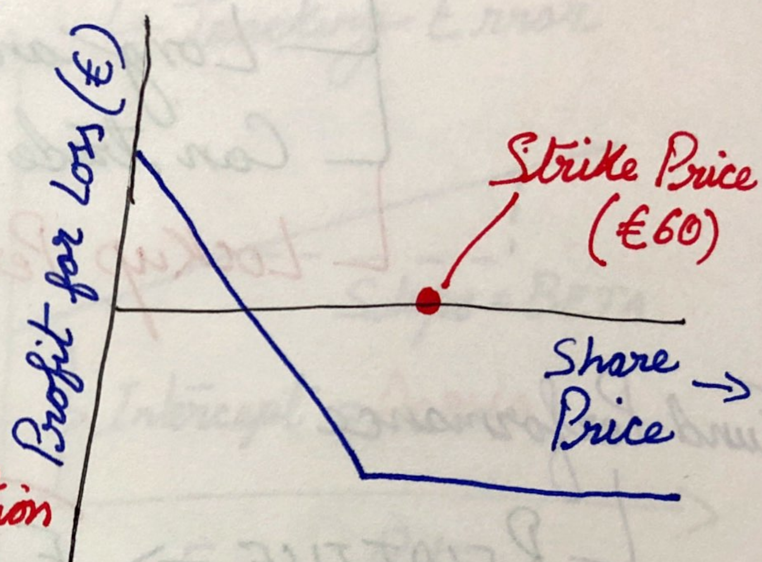
$$\boxed{\text{TRACKING ERROR} = \sqrt{252} * \text{Sample.Stdev}\left(r_{portfolio} - r_{benchmark}\right)}$$

Hedging — Entering into a transaction in order to reduce exposure to price fluctuations.
- Market NEUTRAL Returns

8th December 2018

Hedging using PUT OPTIONS

- Longed 100 shares of Diamler
- Buy 100 put options as well at FIXED price of €60
  - Exercise the put option if the Diamler share price drops below €60, let's say €50. → Sell put options for €60 and keep €10 profit.



A Market Neutral Strategy

Fair Value of a fund,

$$\text{Net Asset Value (NAV)} = \frac{\text{Assets under Management (AUM)} - \text{Expenses}}{\text{number of shares}}$$

$$\text{Gross Expense Ratio} = \frac{\text{Expenses}}{\text{AUM}}$$ } This is what investors pay in the long run

$$\text{Net Expense Ratio} = \frac{\text{Expenses} - \text{Discounts}}{\text{AUM}}$$ } Short-term discounts usually when the fund starts off

OPEN-END Mutual Funds

- New investments AFTER the fund starts.
- Withdraw money DIRECTLY from the fund.
- Fund creates NEW Shares for the investors when they want to buy.
  - They sell by turning-in their shares and receiving the cash back.
    - These shares are also removed from the total shares outstanding.

→ OEMF need to keep some %age of AUM in CASH, not invested.

{ Optimum utilization of cash is NOT HAPPENING.

→ OEMF might be forced to ~~withdraw~~ SELL AUM if too many investors WITHDRAW money at once.

CLOSED-END Mutual Funds

- Accept investor money INITIALLY
- No NEW Investments, Direct Withdrawals after the fund STARTS operating
- Exiting Investors can sell fund shares to new investors like a stock.

FAIR VALUE of shares is not possible, as the value varies based on SUPPLY & DEMAND.

Transaction Costs

- Brokerage Fees
- Moving the price due to a large trade
- Frequent portfolio rebalancing can be expensive in terms of Transaction Costs.
  - Rebalancing is skipped if Transaction costs outweigh the projected benefits.
- Funds in the same financial institution sometimes trade internally to reduce transaction costs.

EXCHANGE TRADED FUNDS (ETFs)

9th December 2018

- Lower fees
- Lower operational costs
- Lower Tax costs
- Shares are tradable like stocks
- Share price follow $\boxed{\text{FAIR VALUE}}$ of funds
- Commodities — energy, precious metals, agriculture
  - International Stocks
  - ~~Heg~~ Hedging

Futures contract:

agreement to buy/sell the asset

- on a future date
- at a fixed price

} buyer is "long" the future; seller is "short" the future

- ~~exit~~ close the position or 'roll-over' on the fixed date.
  - if you don't close the position, you are OBLIGATED to BUY the asset.

ETFs that track commodities save you from all the hassle.

Long Position Portfolio is S&P 500 } Hedging Strategy in case S&P500 goes down
- Short position in ETF (SPDR S&P 500 ETF)

ETF Sponsors
- Financial Institutions that issue ETFs
  - BlackRock iShares
    - $1.75 trillion AUM!!!
- Earn fee as %age of AUM
- Designs and maintains the underlying ETF portfolio
- Sponsors only deal with APs (Authorized Participants)
  - Meryll Lynch, Morgan Stanley, Goldman Sachs et. al.

ETF Create Process (Increases ETF shares)
- AP buys stock and gives to the ETF sponsor
- ETF sponsor CREATES ETF Shares and gives to AP
- AP SELLS ETF shares to Investors

ETF Redeem Process (Decreases ETF Shares)
- AP buys ETF shares from Investors in the Stock Market
- AP trades ETF shares with the sponsor for Original Stocks
- AP SELLS the original stocks in the Stock Market.

ETF Sponsors and APs exchange ETF shares and original of the same value

- ∴ No profit or loss — 0% Capital Gains
  - Hence, NO Capital Gains TAX !!!

ETF Create/Redeem Process

- APs can MAKE money
- C/R process aligns ETF price to the FAIR value of the ~~basket~~ stocks it represents.

ARBITRAGE is SIMULTANEOUS buying and selling of a security in different markets in order to benefit from the price difference of the SAME ASSET.

ETF trading at a Premium

- AP buys stocks — increases share prices ⇓ ETF fair value achieved
- exchanges them with ETF shares with the Sponsor
- Sells the newly minted ETF shares in the market — decreases ETF price ⇑ ETF fair value achieved
  - Profits from the Premium

} ETF Create Process

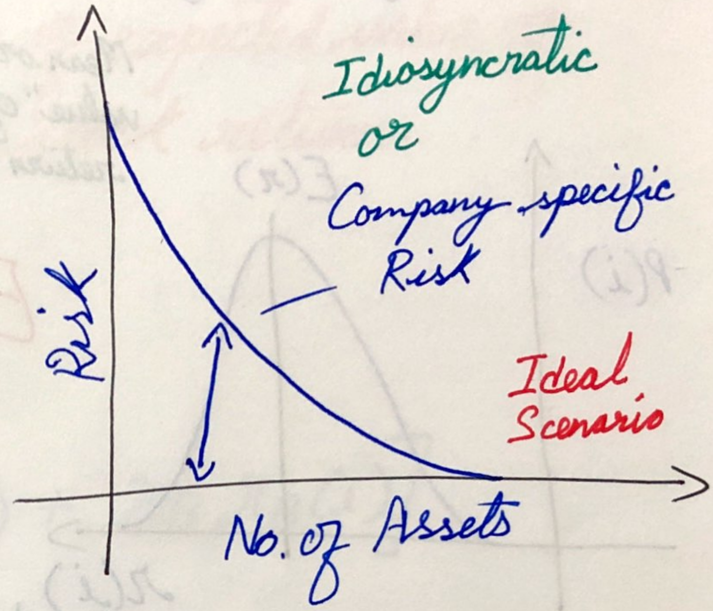
29th November 2018

Portfolio Risk and Return

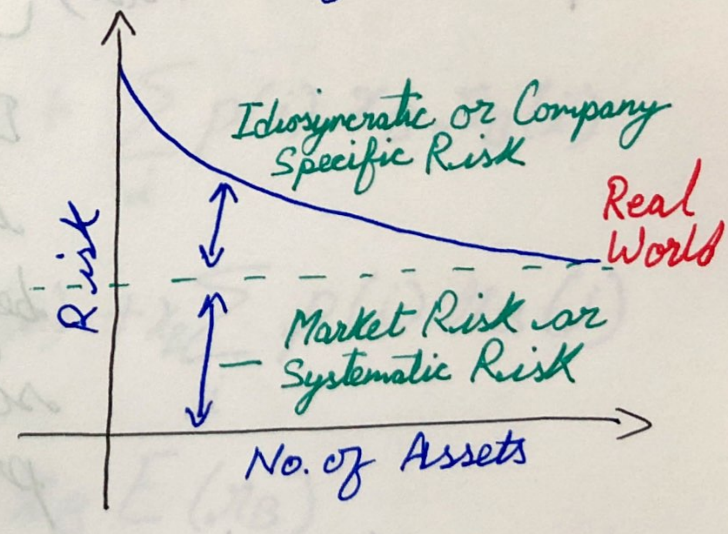
If all the sources of Risk are INDPENDENT

└ Reduce risk to ZERO

  └ Spread money in a large number of stocks



Can't reduce the overall portfolio risk to ZERO

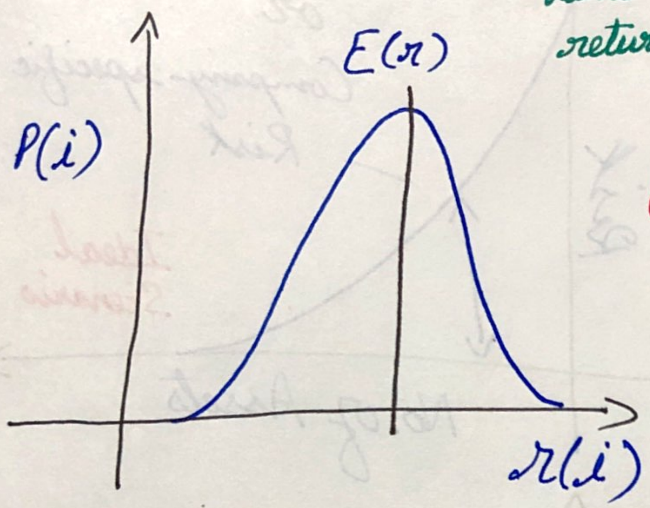


$\boxed{\text{Stock A}}$ $\boxed{\text{Stock B}}$

Weights $\{ \; x_A + x_B = 1$

Portfolio Mean and Variance

└ Model future log returns as Random Variables



Mean or "expected value" of log return distribution

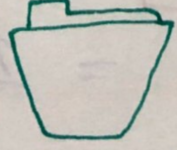
Log return in scenario 'i'

$$E(r) = \sum_{i=1}^{n} p(i)\, r(i)$$

probability of scenario 'i'

Different scenarios are indexed by 'i', with 'r(i)' being the log return of that scenario and p(i) being the probability of that scenario

portfolio return in each scenario i,

$$x_A \times \$ + x_B \times \$ =$$ 

$$\boxed{r_p(i) = x_A\, r_A(i) + x_B\, r_B(i)}$$

Total portfolio return in scenario 'i'.

$$E(r_P) = \underbrace{x_A E(r_A) + x_B E(r_B)}$$

Expected value of Portfolio Return = Weighted sum of expected value of individual stock returns

$$\Downarrow$$

$$E(r_P) = \sum_{i=1}^{n} p(i)\left[x_A r_A(i) + x_B r_B(i)\right]$$

$$= \sum_{i} p(i)\, x_A r_A(i) + \sum_{i} p(i)\, x_B r_B(i)$$

$$= x_A \sum_{i} p(i)\, r_A(i) + x_B \sum_{i} p(i)\, r_B(i)$$

$$= x_A E(r_A) + x_B E(r_B)$$

Metric for total risk inherent to the portfolio

└ Risk is measured with Volatility

└ specifically, Variance (= Volatility^2)

Variance of log return distribution

Log return in Scenario 'i'

$$\sigma^2(r) = \sum_{i=1}^{n} p(i)\left[r(i) - E(r)\right]^2$$

Probability of scenario 'i' occuring

Expected value of log return distribution

Variance of the portfolio,

$$\sigma_P^2 = \sum_i p(i)\left[r_P - E(r_P)\right]^2$$

$$\Downarrow$$

$$= x_A^2 \sigma_A^2 + x_B^2 \sigma_B^2 + 2 x_A x_B \, Cov(r_A, r_B)$$

Covariance is the measure of Joint Variability of 2 random variables

$$Cov(r_A, r_B) = \underbrace{\rho_{r_A r_B}}_{\text{Correlation coefficient}} \sigma_A \sigma_B$$

Standard Deviations of A & B

$$\boxed{\sigma_P^2 = x_A^2\sigma_A^2 + x_B^2\sigma_B^2 + 2x_A x_B \sigma_A \sigma_B \rho_{r_A r_B}}$$

Portfolio Variance

if $\rho_{r_A r_B} = 1$ (perfect correlation)

$$\sigma^2_P = (x_A\sigma_A + x_B\sigma_B)^2$$

$$\Rightarrow \sigma_P = x_A\sigma_A + x_B\sigma_B$$

$$\boxed{0 < \rho_{r_A r_B} < 1}$$

in reality

if $\rho_{r_A r_B} = -1$ (perfect negative correlation)

$$\sigma_P^2 = (x_A\sigma_A - x_B\sigma_B)^2$$

$$\Rightarrow \sigma_P = \text{abs}(x_A\sigma_A - x_B\sigma_B)$$

Results in a perfectly hedged portfolio with variance = 0

when $\rho < 1$,

$$\sigma_P < x_A\sigma_A + x_B\sigma_B$$

Portfolio SD ~~variance~~ is less than the weighted SD of individual stocks

You could see the benefits of Diversification right here!

if, $\sigma_{P1} < \sigma_{P2} \Rightarrow \sigma_{P1}^2 < \sigma_{P2}^2$

σ_P of perfectly correlated stocks > σ_P of imperfectly correlated stocks

Adding uncorrelated stocks helps in reducing overall portfolio risk

The Covariance Matrix

. zip-ing all the maths in the form of vectors and matrices.

$$\sigma_P^2 = x_A^2\sigma_A^2 + x_B^2\sigma_B^2 + 2x_A x_B \, Cov(r_A, r_B)$$

also, $\sigma_A^2 = Cov(r_A, r_A) \Rightarrow \underbrace{\rho_{r_A r_A}}_{=1} \sigma_A \sigma_A = \sigma_A^2$

$$\therefore \sigma_P^2 = x_A^2 \, Cov(r_A, r_A) + x_B^2 \, Cov(r_B, r_B) + 2x_A x_B \, Cov(r_A, r_B)$$

covariance Matrix $\mathbf{P} = \begin{bmatrix} Cov(r_A, r_A) & Cov(r_A, r_B) \\ Cov(r_B, r_A) & Cov(r_B, r_B) \end{bmatrix}$ Symmetric Matrix

weight vector $\mathbf{x} = \begin{bmatrix} x_A \\ x_B \end{bmatrix}$

since, portfolio variance σ_p^2 is an example of the quadratic form

$\therefore$ it can be represented as,

$$\boxed{\sigma_p^2 = \mathbf{x}^T \mathbf{P} \mathbf{x}}$$

$$r^T r = \begin{bmatrix} r_A^T r_A & r_A^T r_B \\ r_B^T r_A & r_B^T r_B \end{bmatrix}$$

r is a matrix containing vectors r_A & r_B as its columns.

if each vector of operation in the data matrix has mean 0, the covariance matrix can be calculated as,

$$\boxed{\frac{1}{n-1} r^T r}$$

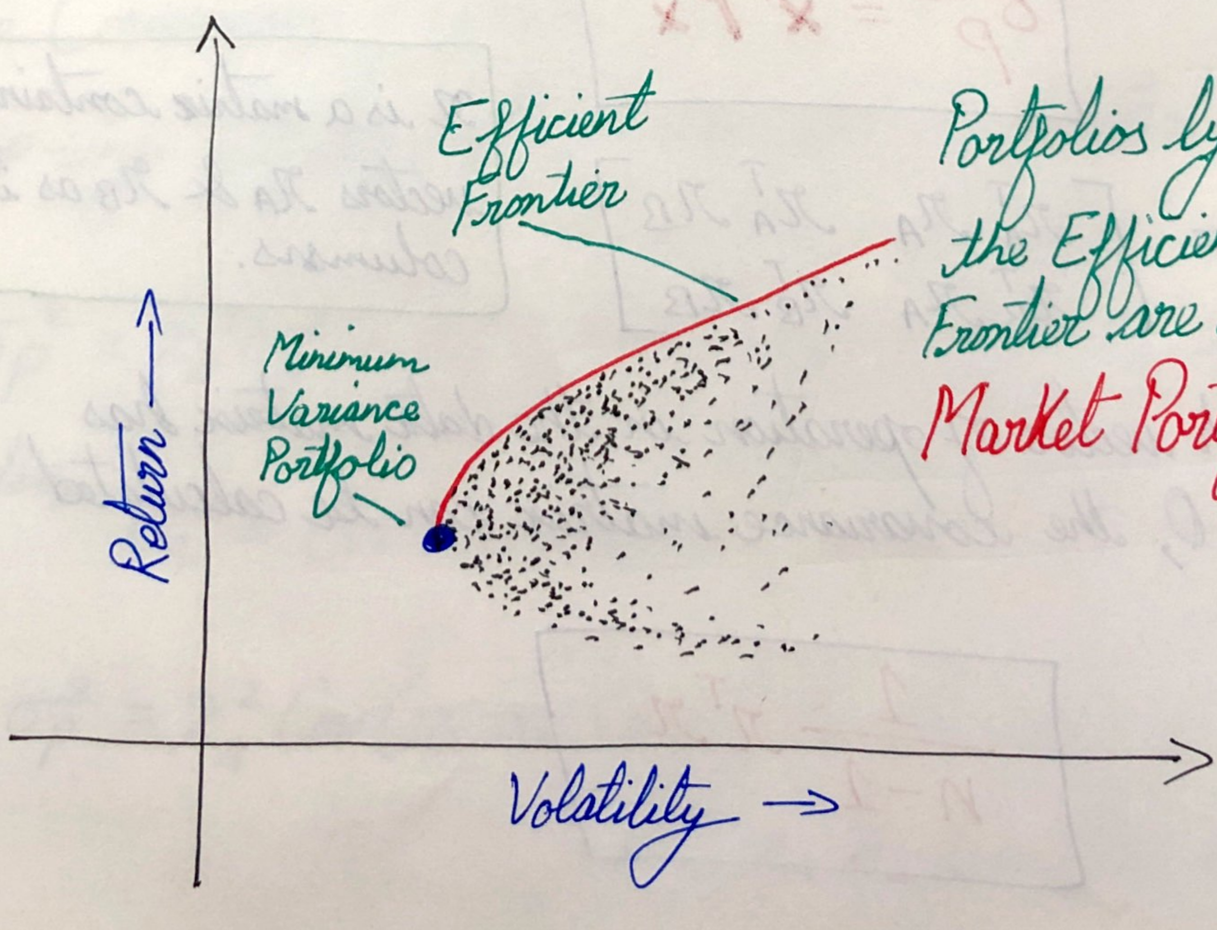
The Efficient Frontier

$$\mu_p = \sum_{i=1}^{N} x_i \mu_i = x^T \mu$$

Portfolio expected return
L weighted sum of each stock's expected return

$$\sigma_p^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} x_i \, Cov(r_i, r_j) \, x_j$$
$$= x^T P x$$

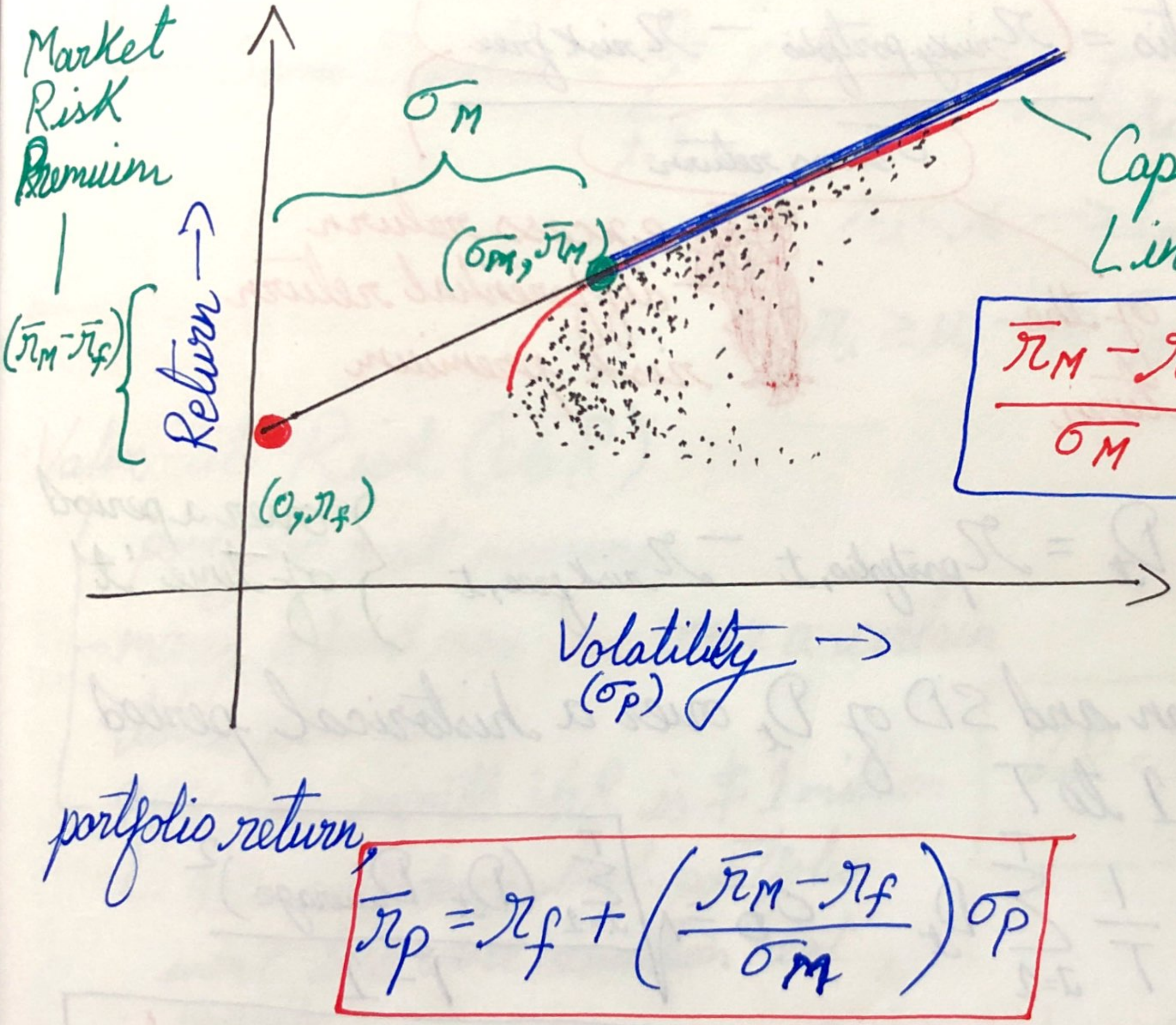
Portfolio Variance
L Sum of pairwise covariances weighted by the products of the weights



Portfolios lying on the Efficient Frontier are called Market Portfolios

Risk-free asset ($r_f = 0.03, \sigma_f = 0$)
└ Guaranteed Rate of Return
└ Uncertainty Level = 0

Rate of return on a 3-month treasury bill (2.37%)



$$\frac{\bar{r}_M - r_f}{\sigma_M} = \text{Sharpe Ratio}$$

portfolio return

$$\boxed{\bar{r}_p = r_f + \left(\frac{\bar{r}_M - r_f}{\sigma_M}\right)\sigma_P}$$

Any level of risk or return beyond the portfolio along the tangent Capital Market Line can be manufactured using LEVERAGE.

The Sharpe Ratio

- ratio of reward to volatility
  - performance of an asset relative to its risk.

$$\text{Sharpe Ratio} = \frac{r_{\text{risky portfolio}} - r_{\text{risk free}}}{\sigma_{\text{excess return}}}$$

volatility of the excess return

excess return
differential return
risk premium

$$\text{Risk Premium, } D_t = r_{\text{portfolio}, t} - r_{\text{risk free}, t}$$ } over a period of time 't'

Calculate mean and SD of D_t over a historical period from $t=1$ to T

$$D_{\text{average}} = \frac{1}{T}\sum_{t=1}^{T} D_t \qquad \sigma_D = \sqrt{\frac{\sum_{t=1}^{T}\left(D_t - D_{\text{average}}\right)^2}{T-1}}$$

$$\boxed{\text{Sharpe Ratio} = \frac{D_{\text{average}}}{\sigma_D}}$$

$$\text{Sharpe Ratio}_{\text{year}} = \sqrt{252}\ \text{Sharpe Ratio}_{\text{day}}$$

- SR depends on the time period
- Allows comparing stocks with different returns
  - SR adjusts the returns by their level of risk

Semi-Deviation

- measures the Downside Risk, rather than VOLATILITY
- calculation includes observations that are LESS than the MEAN only!
- Semi Deviation $= \sum_{t=1}^{n} (\mu - r_i)^2 \times I_{r_i < \mu}$

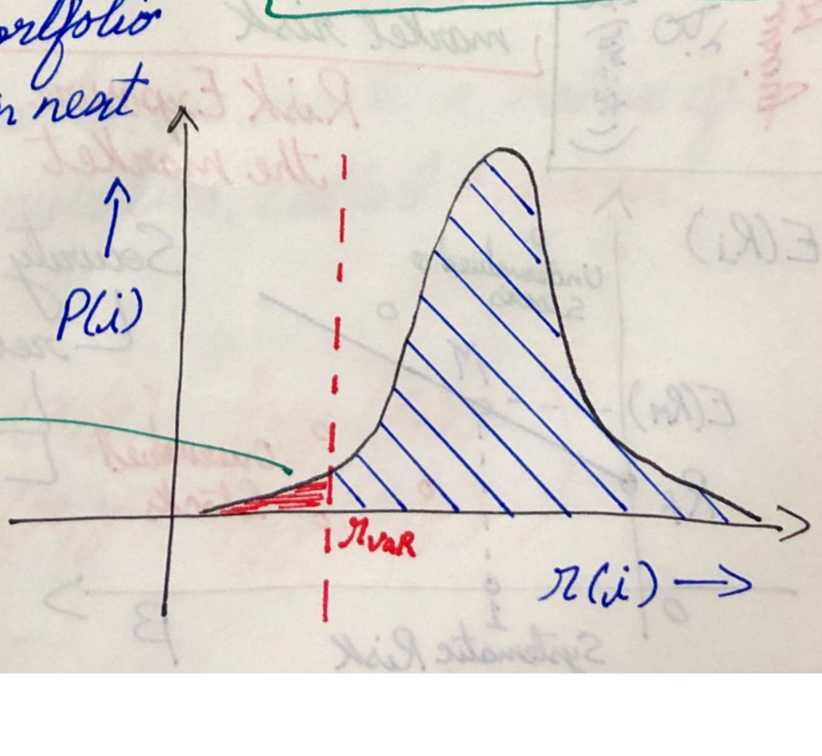
$$r_i < \mu \longrightarrow I_{r_i<\mu} = 1$$

$$r_i \geq \mu \longrightarrow I_{r_i<\mu} = 0$$

Value-at-Risk (VaR)

- portfolio risk measure
- money a fund may lose over a certain time period
- 95% one month VaR is \$1 million
  - 95% confidence that portfolio won't lose over \$1 million next month

$$\boxed{VaR = r_{VaR} \times \text{Amount invested in the stock}}$$



5% of the total area under the curve, represents VaR for this particular stock.

Capital Asset Pricing Model (CAPM)

└ Describes the relationship between systematic risk and expected return for assets

└ risk inherent to the entire market

return of stock 'i', market return

$$r_i = r_f + \beta \times (r_m - r_f)$$

stock return — risk free rate

$$\beta_i = \frac{cov(r_i, r_m)}{\sigma_m^2}$$

Direction and by how much a stock or portfolio moves relative to the market

$$\Rightarrow r_i - r_f = \underbrace{\beta}_{\text{Exposure of an asset to the overall market risk}} \times \underbrace{(r_m - r_f)}_{\text{Risk Premium}}$$

Risk Exposure to the market

Portfolio return using CAPM

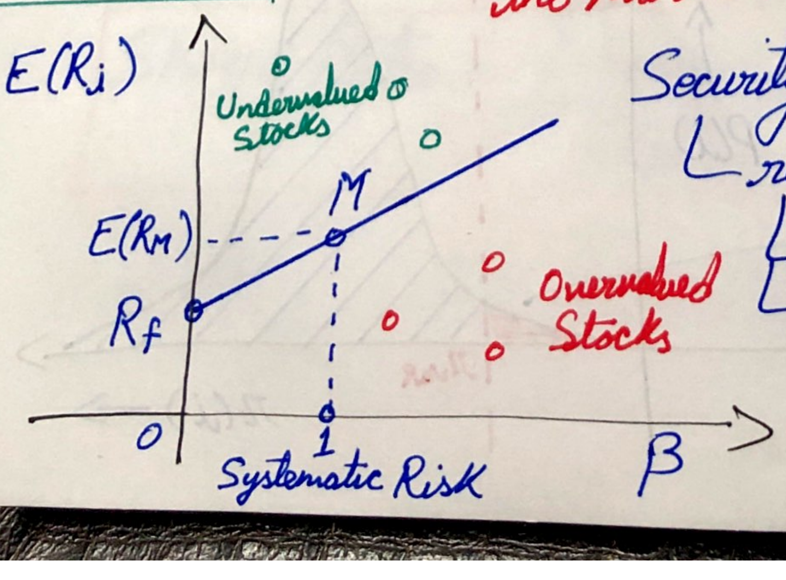
$$r_p = \sum_i^n \omega_i (r_f + \beta_i (r_m - r_f))$$

$$\beta_p = \sum_i^n \omega_i \beta_i$$

$\beta_i = 1$
└ Stock Movement ≡ Market movement

$\beta_i > 1$
└ Stock movement > Market Movement

$\beta_i < 1$
└ Stock Movement < Market Movement



Security Market Line
└ relation b/w risk & return of stocks
  └ x-axis is β
  └ Level of risk increases with β

30th November 2018

Optimization

└ process of finding the Maximum or Minimum of a func.

  └ slope of the function becomes ZERO

    ├ Take the derivative of the fuction, set it equal to 0 and solve for 'x'.

    └ Maxima, Minima or a Saddle Point

      └ Check the function's curvature around the point

if, $\frac{d^2y}{dx^2}(x_0) < 0 \Rightarrow f$ has local maximum at x_0

$\frac{d^2y}{dx^2}(x_0) > 0 \Rightarrow f$ has local minimum at x_0

$\frac{d^2y}{dx^2}(x_0) = 0 \Rightarrow$ the test is inconclusive.

for a function of two variables, we create a matrix of second order partial derivatives, called Hessian Matrix $H(x,y)$

$$H(x,y) = \begin{pmatrix} f_{xx}(x,y) & f_{xy}(x,y) \\ f_{yx}(x,y) & f_{yy}(x,y) \end{pmatrix}$$

determinant of a matrix,

$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$$

if the first order partial derivatives are 0 at point (a, b), that is

$f_x(a,b) = f_y(a,b) = 0$, then we apply the following rule,

$\det(H)(a,b) > 0$ and $f_{xx}(a,b) > 0$

└ (a,b) is the local minimum of f

$\det(H)(a,b) > 0$ and $f_{xx}(a,b) < 0$

└ (a,b) is a local maximum of f

$\det(H)(a,b) < 0$

└ (a,b) is a saddle point of f

$\det(H)(a,b) = 0$

└ the 2nd derivative test is inconclusive

└ (a,b) could be any of the minimum, maximum or saddle point

Optimization with Constraints

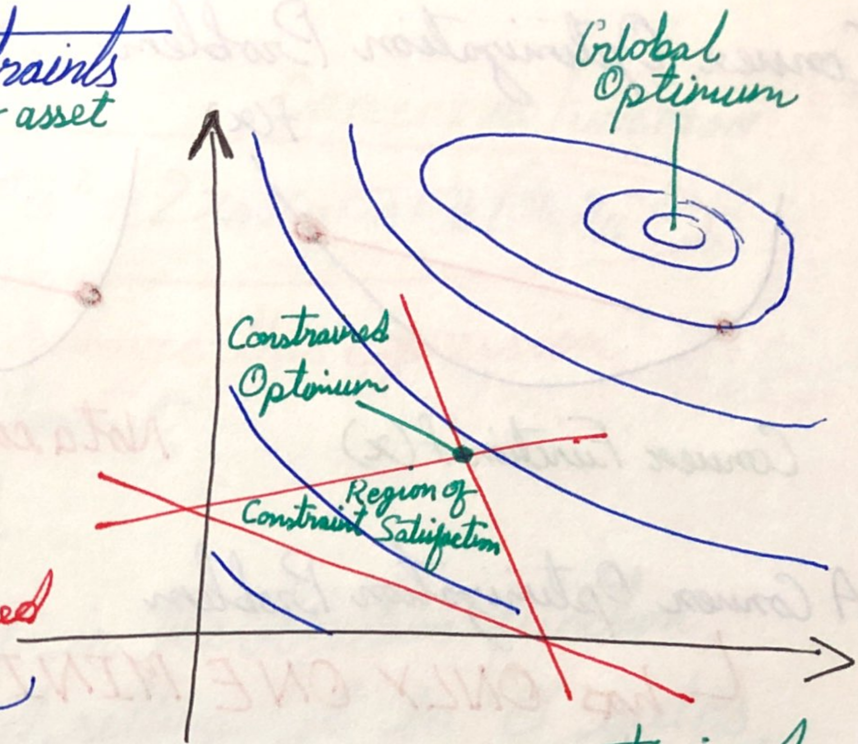
Objective function or Cost Function

Vector of asset weights

$$\text{minimize } f_0(x) \qquad \mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

We are optimizing this function

Variable being optimized

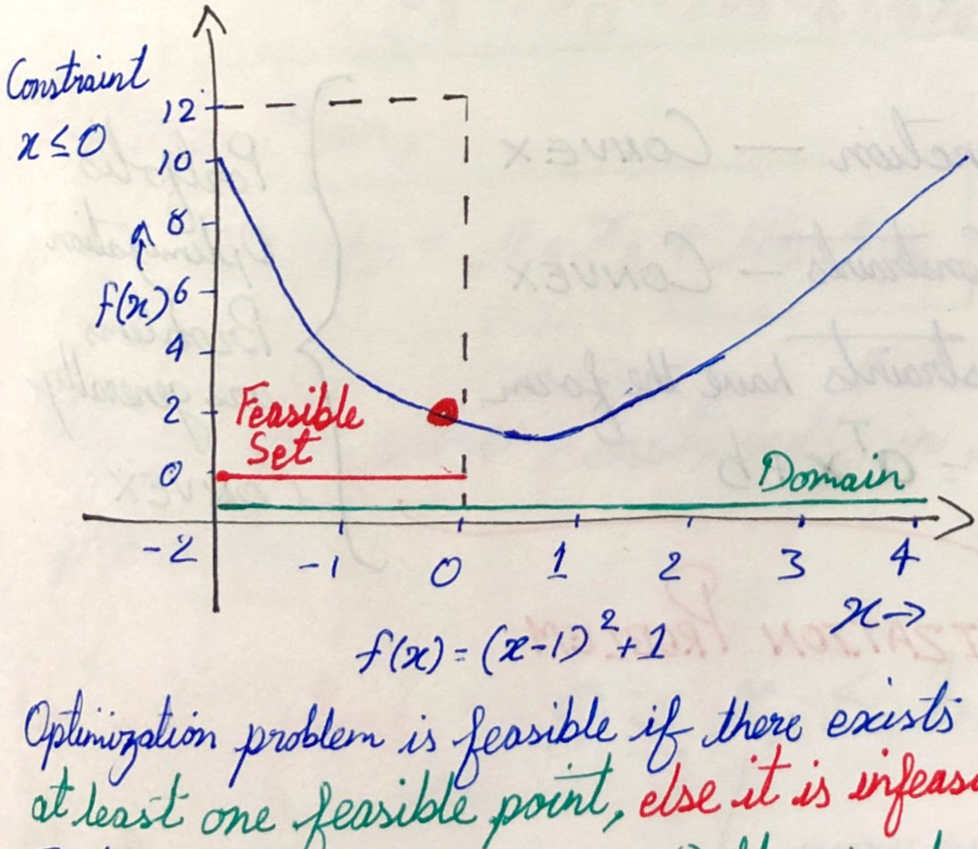


We may have constraints for this optimization → if there are NO constraints → **Unconstrained Problem**

$\mathbf{X}$ is an optimal or a solution, if: $\mathbf{X}$ has the smallest objective function value among all vectors that satisfy the constraints.

Domain: The set of points for which the objective and all constraint functions are defined.

Feasible Set: The set of points that satisfy all the constraints.

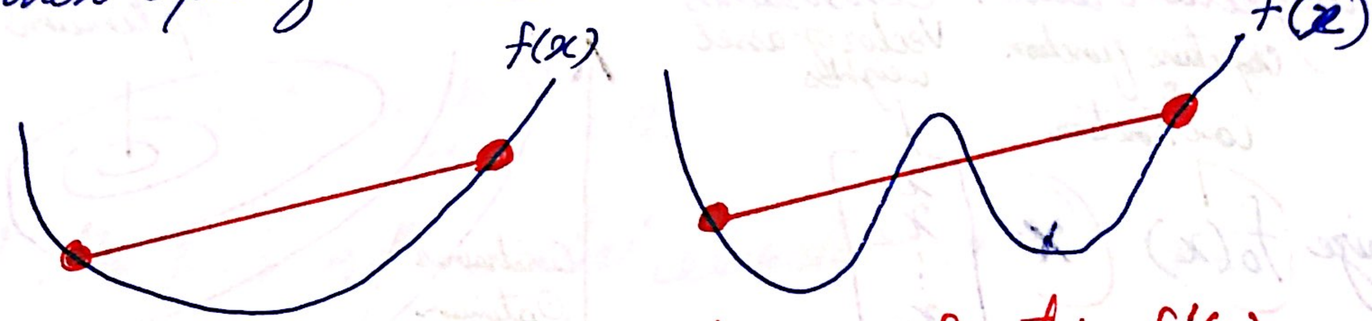


$$f(x) = (x-1)^2 + 1$$

Optimization problem is feasible if there exists at least one feasible point, else it is infeasible.

Feasible points for which $f_0(x) \rightarrow -\infty \rightarrow$ Problem is unbounded below e.g. $f(x) = -\ln(x)$

Convex Optimization Problem



Convex Function f(x)

Not a convex function f'(x)

A Convex Optimization Problem
- has ONLY ONE MINIMUM
  - Local Minimum $\Rightarrow$ Global minimum over the whole feasible set

when,
- Objective function — CONVEX
- Inequality Constraints — CONVEX
- Equality Constraints have the form
  $f(\mathbf{x}) = a^T\mathbf{x} + b$

CONVEX OPTIMIZATION PROBLEM

Portfolio Optimization Problems are generally CONVEX

For a two asset portfolio,

OBJECTIVE FUNCTION

$$\sigma_P^2 = x_A^2\sigma_A^2 + x_B^2\sigma_B^2 + 2x_Ax_B\sigma_A\sigma_B\rho_{r_Ar_B}$$

We seek to minimize this expression with the constraint,

$$x_A + x_B = 1$$

Substituting the constraint in the objective function, taking a derivative and setting it to '0' yields,

$$\boxed{x_A = \frac{\sigma_B^2 - \sigma_A\sigma_B\rho_{r_Ar_B}}{\sigma_A^2 + \sigma_B^2 - 2\sigma_A\sigma_B\rho_{r_Ar_B}} \qquad x_B = 1 - x_A}$$

Portfolio Mean,

$$\boxed{\mu_P = \mu_A x_A + \mu_B x_B}$$

Plugging some constraints in the optimization problem,

no short selling : $0 \leq x_i \leq 1 \qquad i = 1, 2, \ldots, n$

every weight in vector x

sector limits:

$$x_{biotech\,1} + x_{biotech\,2} + x_{biotech\,3} \leq M$$

M = percent of portfolio to invest in biotech companies

constraint on portfolio return:

$$\mathbf{x}^T \mu \geq r_{min}$$

r_{min} = minimum acceptable portfolio return

maximizing portfolio return:

Best returns possible, but with losses limited to 'p' percent.

objective: minimize $-\mathbf{x}^T \mu$

constraint: $\mathbf{x}^T P \mathbf{x} \leq p$, p = Maximum permissible portfolio variance

maximizing portfolio return AND minimizing variance:

objective: minimize $-\mathbf{x}^T \mu + b\, \mathbf{x}^T P \mathbf{x}$, b = tradeoff parameter

└ % return you are willing to give up for each unit of variance you take on

minimize distance to a set of target weights:

minimize $\|\mathbf{x} - \mathbf{x}^*\|_2$, $\mathbf{x}^*$ = a set of target portfolio weights

└ could be an alpha vector

Tracking an Index:

minimize $\mathbf{x}^T P \mathbf{x} + \lambda \|\mathbf{x} - q\|_2$

q = set of index weights

λ = a tradeoff parameter

CVXPY

└ solver for Convex Optimization Problems

x = cvx.Variable(2) Optimization vector of length 2

objective = cvx.Minimize('exp') Objective expression('exp') to be minimized

constraints = [x+y==1, x-y>=1] => $x+y=1$, and $x-y \geq 1$

cvx.quad_form(x, P) => Creates quadratic form $x^T P x$

cvx.norm(x-b, 2) => $\|x-b\|_2$

problem = cvx.Problem(objective, constraints)

└ immutable

└ Cannot modify 'objective' or 'constraints' after creation

problem.solve() Runs the optimization solver

problem.status() If the problem is UNFEASIBLE or UNBOUNDED

problem.value Optimal value of the Objective Function

x.value Optimal value of the Optimization Variable

Portfolio Rebalancing Costs
- Transaction Costs
- Taxes
- Time and Labor costs

Transaction Costs are hard to model.
- ∴ We make the assumption that

$$\text{TRANSACTIONAL COSTS} \propto \underbrace{\text{To the change in the magnitude of the holdings}}_{\text{Portfolio Turnover}}$$

To calculate portfolio turnover,

$$|X_{t_1} - X_{t_2}| = \left| \begin{bmatrix} x_{t_1,1} \\ x_{t_2,2} \\ \vdots \\ x_{t_1,n} \end{bmatrix} - \begin{bmatrix} x_{t_2,1} \\ x_{t_2,2} \\ \vdots \\ x_{t_2,n} \end{bmatrix} \right|$$

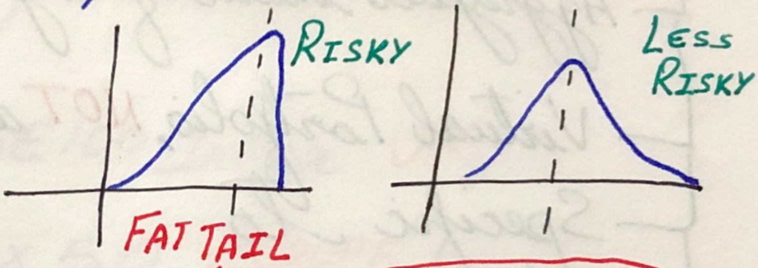
$$\boxed{\text{turnover} = |x_{t_1,1} - x_{t_2,1}| + |x_{t_1,2} - x_{t_2,2}| + \cdots + |x_{t_1,n} - x_{t_2,n}}$$

$$\text{Annulized Turnover} = \underbrace{\frac{\text{sum total turnover}}{\text{no. of total rebalancing events}}}_{\text{Average turnover per rebalancing event}} \times \text{No. of rebalancing events per year}$$

Limitations of Portfolio Optimization Techniques

(... that I just learned)

- Mean Returns are DIFFICULT to estimate
  - Avoid estimating returns directly
    - FACTOR MODELS :wink: :wink:
- Estimating Portfolio Variance (is a pain in the ....)
  - Variance MAY NOT capture risk
  - huge n×n matrix to optimize
    - Inefficient and slow
    - Estimation errors get AGGREGATED



Similar mean and variance but

  - Long historical data required
    - 50 assets require 5 years of DAILY historical data
- Estimates are inherently noisy
  - Trade less frequently?
- Conflicting predictions over different time horizons
  - MULTI-PERIOD OPTIMIZATION
    - Include Transaction Costs
- Transaction Costs
  - Difficult to model

PART - 4

FACTOR MODELS

10th December 2018

Models of portfolio return and risk
└ Combined into our Optimization Framework

ALPHA FACTORS: Drivers of MEAN Returns

RISK FACTORS: Drivers of VOLATILITY

Factors can be based on:
- Momentum
- Fundamental Information
- Signals from Social Media

Sum of,
- Demeaned Weights = ZERO
- Rescaled Weights = ZERO
- |Rescaled Weights| = ONE
- Rescaled Short Positions = −0·5

Standardized Factors:

DE-MEAN	RESCALE
$\sum_{i=1}^{N} x_i = 0$	$\sum_{i=1}^{N} \lvert x_i \rvert = 1$
Sum of Weights = 0	Sum of Absolute Values = 1
$\mu = \frac{1}{N} \sum_{i=1}^{N} x_i$	$\lambda = \sum_{i=1}^{N} \lvert x_i \rvert$
$x_i = x_i - \mu \quad \forall i = 1, 2 \ldots N$	$x_i = \frac{x_i}{\lambda} \quad \forall i = 1, 2 \ldots N$

$$x_i = \frac{x_i}{\sum_{i=1}^{N} \lvert x_i \rvert}$$

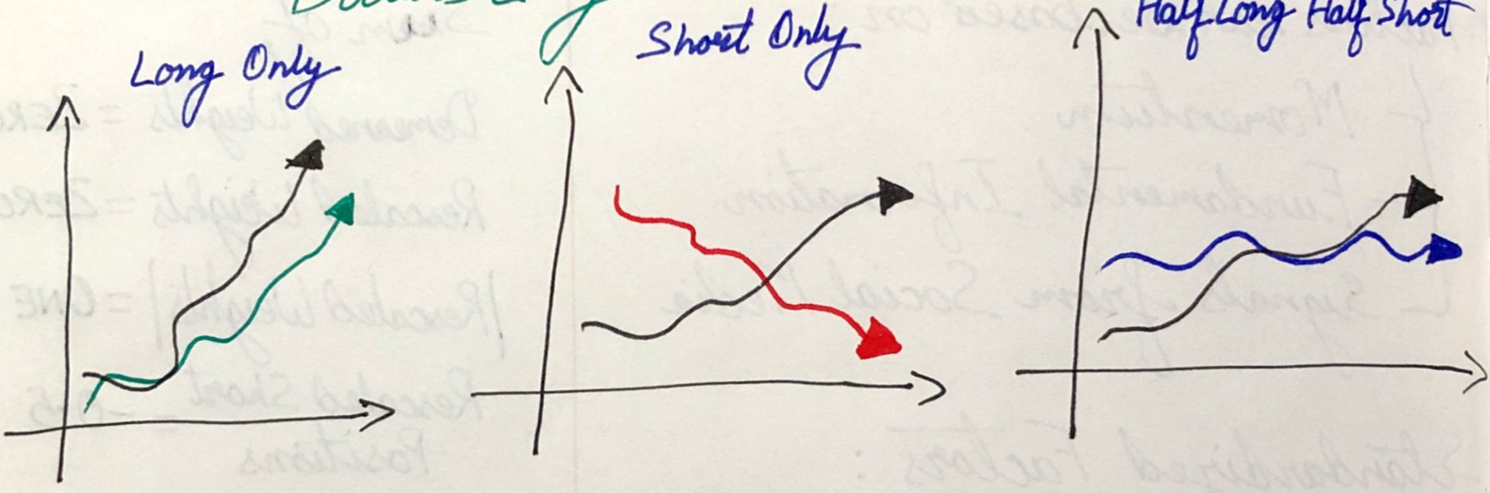
De-Mean

Factor Values will be used as PORTFOLIO WEIGHTS.

L That's why you gotta De-Mean them!

Dollar Neutral Portfolio

Dollars Long = Dollars Short



Notional — Dollar amount associated with a Portfolio.

Dollars $s_i = x_i \times$ Notional

if Notional = \$100m

	Dollars s_i	x_i
Long	\$1m	0.01
Short	−\$1m	−0.01
Total	\$0	0

Dollar Neutral (Market Neutral)

RESCALE

$$\left.\begin{array}{l}\text{Maintain a leverage ratio equal to } \underline{1} \\ \quad \therefore \text{ Rescale the factors}\end{array}\right\}$$ To test the effectiveness of our factors.

$$\text{LEVERAGE RATIO} = \frac{\sum_{i=1}^{N} |x_i|}{\text{Notional}}$$

Example Portfolio = \$1m Notional	Leverage Ratio : $\frac{\text{Sum of Positions}}{\text{Notional}}$
\$1m Long	\$1m/\$1m = $\underline{1}$
leveraged \$2m Long	\$2m/\$1m = 2
\$1m long, \$1m Short	\$2m / \$1m = 2
\$2m long, \$2m Short	\$4m/\$1m = 4

Theoretical Portfolio = \$1	$\sum_{i=1}^{N} \lvert x_i \rvert$
−0.5, 0.3, 0.2	$\underline{1}$
−1.0, 0.6, 0.4	2

10th January 2019

Factor Model
- A Statistical model used to describe variability among observed correlated variables, in terms of potentially smaller number of unobserved variables called FACTORS.
  - (unobserved variables) LATENT VARIABLES
- Model asset returns and common variability of stocks

Generic Linear Factor Model

$\boxed{r = Bf + s}$

$$r_i = b_{i1} f_1 + b_{i2} f_2 + \cdots + b_{iK} f_K + s_i$$

r_i = the return on asset 'i'

f_1 = the value of factor return 1

b_{i1} = the change in the return on asset 'i' per unit change in factor return 1

OR

FACTOR EXPOSURE

K = the number of FACTORS

s_i = the portion of the returns on asset 'i' not related to the K Factors

"The return of any stock 'i' can be decomposed into the returns of factors times the stock's exposures to those factors, plus an unexplained portion."

Factor Returns are Latent Variables

∟ Their influence cannot be measured directly.

Residual Return s_i assumed to be uncorrelated with each of the Factor Returns.

$$\text{Corr}(s_i, f_K) = 0, \ \forall \ i \ \& \ K$$

Residual of one asset return is uncorrelated with residual of any other asset.

$Corr(s_i, s_j) = 0$, for every 'i' not equal to 'j'

Residual s_i is assumed to be specific or "idiosyncratic" to asset 'i'.

Covariance Matrix using a Factor Model

$$r_i = b_{i1} f_1 + b_{i2} f_2 + \dots + b_{iK} f_K + s_i$$

└ Single Stock Return

for multiple stock returns,

$$r = \begin{bmatrix} r_1 \\ \vdots \\ r_N \end{bmatrix} \quad B = \begin{bmatrix} B_{1,1} & \cdots & B_{1,K} \\ \vdots & \ddots & \vdots \\ B_{N,1} & \cdots & B_{N,K} \end{bmatrix} \quad f = \begin{bmatrix} f_1 \\ \vdots \\ f_K \end{bmatrix} \quad s = \begin{bmatrix} S_1 \\ \vdots \\ S_N \end{bmatrix}$$

r = Number of companies

B = Number of companies x Number of Factors

f = Number of factors

s = Number of companies

before calculating the Covariance Matrix the stock returns r_i are DEMEANED.

$$Cov(X,Y) = \frac{1}{n} \sum_{i=1}^{n} (X_i - \cancel{\bar{X}})(Y_i - \cancel{\bar{Y}})$$

Since r_i was demeaned

$$= E[(X - \cancel{\bar{X}})(Y - \cancel{\bar{Y}})] \Rightarrow E[XY]$$

Expectation value of a matrix

$$E(A) = \begin{bmatrix} E(A_{11}) & \cdots & E(A_{1P}) \\ \vdots & \ddots & \vdots \\ E(A_{N1}) & \cdots & E(A_{NP}) \end{bmatrix}$$

$$\boxed{\text{Covariance Matrix} = E[rr^T] \Rightarrow \text{Expectation value or } rr^T}$$

$$E(rr^T) = E[(Bf+s)(Bf+s)^T]$$

$$= E[(Bf+s)((Bf)^T + s^T)]$$

$$= E[Bf(Bf)^T + Bfs^T + s(Bf)^T + ss^T]$$

$$= E[Bff^TB^T + Bfs^T + sf^TB^T + ss^T]$$

$\because$ We assumed $Corr(s_i, f_k) = 0$, for every i & K

$$\therefore E(rr^T) = B\,E[ff^T]B^T + \underbrace{\cancel{B E[fs^T]}}_{0} + \underbrace{\cancel{E[sf^T]B^T}}_{0} + E(ss^T)$$

$$= B\,E[ff^T]B^T + E[ss^T]$$

$$\Rightarrow E(rr^T) = BFB^T + E[ss^T]$$

$$= BFB^T + S$$

$$\boxed{E(rr^T) = BFB^T + S} \qquad S = \begin{bmatrix} S_{11} & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & S_{NN} \end{bmatrix}$$

$\because$ We assumed $Corr(s_i, s_j) = 0$, for every 'i' not equal to 'j'

Factors that are predictive of Mean
└ ALPHA FACTORS

Factors that are predictive of Variance
└ RISK FACTORS

$$E(rr^T) = \underbrace{B\,F}_{\text{Explicitly describe risk}}\,B^T + \underbrace{S}_{\text{Says nothing about alpha}}$$

Practitioners usually buy B, F and S.

14th December 2018

Risk Factor Models

Motivation for Risk Factor Models

$\beta_{N,K}$ – Factor exposure of Stock N to Factor K

Volatility $\longrightarrow$ Risk

$$Var(r_p) = (Vol(r_p))^2$$

$\beta_{P,K}$ – Factor exposure of Portfolio to Factor K

Two Stock Portfolios

$$Var(r_p) = x_1^2 Var(r_1) + x_2^2 Var(r_2) + 2x_1 x_2 Cov(r_1, r_2)$$

$$\left.\begin{bmatrix} Var(r_1) & Cov(r_1, r_2) \\ Cov(r_2, r_1) & Var(r_2) \end{bmatrix}\right\} \text{Covariance Matrix}$$

Factor Model of Asset Return

1 Stock 1 Factor

$$r_i = \overbrace{\underbrace{\beta_{i,1}}_{\text{Factor Exposure}} \times \underbrace{f_1}_{\text{Factor Return}}}^{\text{Common Return}} + \underbrace{s_i}_{\text{Specific Return}}$$

Multi-Factor Model of Returns

$$\boxed{r_i = \sum_{k=1}^{K} (\beta_{i,k} \times f_k) + s_i}$$

$$s_i = r_i - \beta_{i,1} \times f_1$$

$\Rightarrow$ Specific Return = Asset Return − Common Return

Factor Model of Portfolio Return

Contribution of factors to Portfolio Return

$$r_p = (\beta_{p,1} \times f_1) + \ldots$$

$$\beta_{p,1} = \sum_{i=1}^{N} (x_i \times \beta_{i,1})$$

$$r_p = \underbrace{\sum_{k=1}^{K} (\beta_{p,k} \times f_k)}_{\text{Contribution of Factors}} + s_p \qquad s_p = \sum_{i=1}^{N} (x_i \times s_{i,k})$$

Specific Return

Factor Model of Portfolio Variance — A premiesu

$$Var(r_p) = X^T (BFB^T + S) X$$

B & B^T — Matrix of factor exposures (measuring spoons)

X — Matrix of Portfolio Weights (ice cream scoops)

F — Covariance Matrix of Factors (ingredients)

S — Matrix of Specific Variances "idiosyncratic variances"

Variance of 1 Stock

$$Var(r_i) = \beta_{i,1}^2 Var(f_1) + \beta_{i,2}^2 Var(f_2)$$

$$\underbrace{+ 2\beta_{i,1}\beta_{i,2} Cov(f_1, f_2)}_{\substack{\text{Systematic Variance} \\ \text{(main ingredients)}}} + \underbrace{Var(s_i)}_{\substack{\text{Specific Variance} \\ \text{(specific ingredients)}}}$$

Pairwise Covariance of 2 stocks

$$Cov(r_i, r_j) = Cov(\beta_{i,1} f_1 + \beta_{i,2} f_2 + s_i,$$
$$\beta_{j,1} f_1 + \beta_{j,2} f_2 + s_j)$$

$$\begin{bmatrix} Cov(\beta_{i,1} f_1, \beta_{j,1} f_1) & Cov(\beta_{i,1} f_1, \beta_{j,2} f_2) \\ Cov(\beta_{i,2} f_2, \beta_{j,1} f_1) & Cov(\beta_{i,2} f_2, \beta_{j,2} f_2) \end{bmatrix}$$

Covariance Matrix of Assets $\begin{bmatrix} Var(r_i) & Cov(r_i, r_j) \\ Cov(r_j, r_i) & Var(r_j) \end{bmatrix}$

Building-blocks for the Covariance Matrix of Assets

$$Var(r_i) = \beta_{i,1}^2 Var(f_1) + \beta_{i,2}^2 Var(f_2) + 2\beta_{i,1}\beta_{i,2} Cov(f_1, f_2) + Var(s_i)$$

$$Var(r_j) = \beta_{j,1}^2 Var(f_1) + \beta_{j,2}^2 Var(f_2) + 2\beta_{j,1}\beta_{j,2} Cov(f_1, f_2) + Var(s_j)$$

$$Cov(r_i, r_j) = \beta_{i,1}\beta_{j,1} Var(f_2) + \beta_{i,1}\beta_{j,2} Cov(f_1, f_2) + \beta_{i,2}\beta_{j,1} Cov(f_2, f_1) + \beta_{i,2}\beta_{j,2} Var(f_2)$$

Portfolio Variance

$$Var(r_p) = x_i^2 Var(r_i) + x_j^2 Var(r_j) + 2x_i x_j Cov(r_i, r_j)$$

$$\Downarrow$$

$$Var(r_p) = X^T (BFB^T + S) X$$

$$F = \begin{pmatrix} Var(f_1) & Cov(f_1, f_2) \\ Cov(f_2, f_1) & Var(f_2) \end{pmatrix} \qquad X = \begin{pmatrix} x_i \\ x_j \end{pmatrix}$$

$$B = \begin{pmatrix} \beta_{i,1}, & \beta_{i,2} \\ \beta_{j,1}, & \beta_{j,2} \end{pmatrix}$$

$$S = \begin{pmatrix} Var(s_i) & 0 \\ 0 & Var(s_j) \end{pmatrix}$$

Risk Factor Models with PCA

15th December 2018

Each new basis dimension or PC in the PCA represents the return of an individual company.

Don't use all of the PCs

- Use only those that describe most of the variance
  - Drop PCs that account for the least variance
- Total Variance in the New Basis = Total variance in the original basis
- in essense, used as a dimensional reduction algorithm.
  - Dimensionality Reduction

PCA as a Factor Model

- Reducing the data dimensionality
  - While capturing maximum amount of variance

$$\begin{array}{cc} & \text{Time (Observations, Samples)} \\ \text{Companies (Features, Original Dimensions)} & \begin{pmatrix} r_{1,1} & \cdots & r_{1,T} \\ \vdots & \ddots & \vdots \\ r_{N,1} & \cdots & r_{N,T} \end{pmatrix} \end{array}$$

16th December 2018

$$r = Bf + s$$

Asset Returns

Factor Exposures

Factor Returns

Idiosyncratic Risk

N = Number of Companies

K = Number of Factors

T = Number of Time Points

$$\begin{bmatrix} r_{1,1} & \cdots & r_{1,T} \\ \vdots & \ddots & \vdots \\ r_{N,1} & \cdots & r_{N,T} \end{bmatrix} = \begin{bmatrix} B_{1,1} & \cdots & B_{1,K} \\ \vdots & \ddots & \vdots \\ B_{N,1} & \cdots & B_{N,K} \end{bmatrix} \begin{bmatrix} f_{1,1} & \cdots & f_{1,T} \\ \vdots & \ddots & \vdots \\ f_{K,1} & \cdots & f_{K,T} \end{bmatrix} + \begin{bmatrix} s_{1,1} & \cdots & s_{1,T} \\ \vdots & \ddots & \vdots \\ s_{N,1} & \cdots & s_{N,T} \end{bmatrix}$$

Our Vector in old basis language = Principal Components × Our Vector in New Basis Language

New Basis Vectors written in old basis language

$$+ \begin{bmatrix} \vdots & \vdots & \vdots \\ s_1 & s_2 & s_3 \\ \vdots & \vdots & \vdots \end{bmatrix}$$

Compressed Representation of the Original Data

Matrix of Factor Exposures

Matrix of Factor Returns

Matrix of Factor Returns $(f) = B^T r$ — original returns data

Factor Returns

Factor exposures matrix

Factor Covariance Matrix $(F) = \frac{1}{T-1} f f^T$

$$F = \begin{bmatrix} F_{1,1} & 0 & \cdots & 0 \\ 0 & F_{2,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & F_{K,K} \end{bmatrix}$$

Factors are projections of the data on to the PCs, they are orthogonal.

True for PCA

$$F_{annual} = 252 \, F_{daily}$$

$s = r - Bf$ — Residual

calculate covariance matrix of the residuals

Idiosyncratic Risk Matrix

$$S = \frac{1}{T-1} s s^T$$

Set off diagonal elements to zero

$$S = \begin{pmatrix} s_{11} & 0 & \cdots & 0 \\ 0 & s_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & s_{K,K} \end{pmatrix}$$

Set to 0

17th December 2018

Alpha Processing Techniques

- Sector Neutralization
- Ranking
- Z-Scoring
- Smoothing
- Conditioning

Most significant risk factors — Dollar Neutral

- Market Risk — De-meaning alleviates MR
- Sector Risk — Make Alpha Sector Neutral
  - Subtract Sector Mean from each weight in the Sector
    - Longs = Shorts
      - SECTOR NEUTRAL

Simplifying Assumption

- Each Stock's $\beta = 1$

First, neutralize the alpha vector by MARKET.
- Then by SECTOR.

$\boxed{\text{Clip}}$ Alpha values that are very large
- To stay away from FREQUENT LARGE TRADES
  - Trading costs money!
- Clip 95^{th} & 5^{th} percentile values

WINSORIZING

"Winzorize an Alpha Vector"
- Numbers above the 95^{th} percentile are replaced by the number AT the 95^{th} percentile.
  - Similarly for numbers below the 5^{th} percentile.

We can also deal with outliers by setting the MAXIMUM allowed weight value.
- Might have to deal with outliers EACH DAY

We gotta find a way that transforms alpha signal to be
- less sensitive to noise and outliers
- avoid excessive trades

RANKING does it!

AAPL	MSFT	IBM	
0.39	0.32	0.28	} Alpha Values
3	2	1	} RANKINGS

- Highest Ranking (largest number) $\Longrightarrow$ Highest potential alpha Returns

We make the trade only when the rankings change (sort of!).

Normalizing Alpha Factors

- Subtracting mean
- Dividing by SD

} Gives the z-score

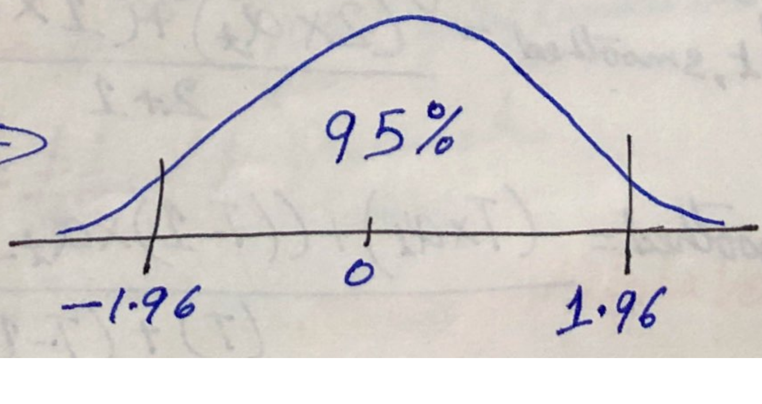
$$\boxed{z_i = \frac{x_i - \mu}{\sigma}}$$

- Makes it easier to compare and combine multiple alpha factors

Ranking helps compare different alpha vectors.

- Z-Scoring helps compare different alpha vectors even when the stock universes are differently sized.

Distribution of z-scores $\longrightarrow$



RANKING

- Makes alpha factors more robust against outliers and noise.
- Best to use when all alpha vectors are generated from the same stock universe.

Z-SCORING

- NOT ROBUST against outliers and noise
- Apply RANKING and then Z-SCORING, if alpha vectors are generated from different stock universes.
- Portfolio Managers in large asset companies might want a standardized vector from all quants for comparison.

Smoothing

- Used when the data is $\boxed{\text{sparse}}$
- Rolling Window Average

$$\alpha_{t,\,smoothed} = \frac{\sum_{t=1}^{T}(\alpha_t)}{}$$

- Weighted Moving Average

$$\left.\alpha_{t,\,smoothed} = \frac{(2\times\alpha_t)+(1\times\alpha_{t-1})}{2+1}\right\}\text{Linear Decay}$$

$$\alpha_{t,\,smoothed} = \frac{(T\times\alpha_t)+((T-1)\times\alpha_{t-1})+\ldots+(1\times\alpha_{t-T+1})}{(T)+(T-1)+\ldots+(1)}$$

18th December 2018

Alpha Evaluation Metrics

- Factor Returns — Portfolio Returns due to alpha factor
- Sharpe Ratio
- Information Coefficient
- Information Ratio
- Quantile Analysis
- Turnover Analysis

- Standardized alpha factor
  - Demean + Rescaled
- Portfolio Weights
  - Each stock gets the same weight as prescribed by the Alpha Vector
- Calculate the stock returns for the day
- Factor Return (single day) $r_p = \sum_{i=1}^{N} (x_i \times r_i)$

x_i — stock weight

We repeat this multiple times and get a time series of Factor Returns

Universe $\Longleftrightarrow$ Universe Construction Rule
- free from Lookahead Bias
  - use Survivorship Bias free data

Return Denominator

$$R_D = \sum_{t=1}^{T} |\alpha_t|$$

Leverage is an act of borrowing money

Leverage Ratio $(L_R) = \dfrac{\text{Sum of all long-short positions}}{\text{Actual Capital Invested in Portfolio}}$

$$L_R = \frac{\$1}{\$1} = 1$$

Research Stage

Normalized Alpha Vector

$$\text{Sharpe Ratio } (S) = \frac{\text{mean}(f)}{\text{Stdev}(f)} = \frac{\text{daily factor return}}{\text{standard deviation of daily return}}$$

$$S_{annual} = \sqrt{252}\left(\frac{\text{mean}(f)}{\text{Stdev}(f)}\right)$$

$\boxed{\text{Always prefer alphas with higher Sharpe Ratios.}}$

$\boxed{\text{Sharpe Ratio is the key metric. Not the factor Returns.}}$

Higher S/R $\Rightarrow$ Higher Leverage $\Rightarrow$ Higher Returns

Rank Information Coefficient (Rank IC)

- ~~Kind of~~ correlation b/w asset ~~ranks~~ ~~alpha values~~ ranks and their future returns
  - if higher ~~alpha value~~ rank ~~rank~~ resulted in higher future returns $\Rightarrow$ Higher Rank IC
    - else it would be lower, or possibly negative
- Spearman Rank Correlation
  - it is like usual Pearson Correlation, but converts X & Y into RANKS before the calculation
- Correlation b/w ranked Alpha Vector and the ranked forward asset returns for a single time period, repeat and get a time series.

19th December 2018

INFORMATION RATIO

└ Special application of the Sharpe Ratio

return = systematic return + specific return

$$\boxed{\text{Information Ratio (IR)} = \sqrt{252} \times \text{Sharpe Ratio}_{\text{of Specific Return}}}$$

- Performance contributed by the fund manager to the portfolio
- When portfolio is Neutral to market and common factors
  - └ SR = IR

FUNDAMENTAL LAW OF ACTIVE MANAGEMENT

$$IR = IC * \sqrt{B}$$

IR — Information Ratio

IC — Information Coefficient — Great discretionary Investors

B — Breadth — Great Quants

BREADTH — Number of INDEPENDENT TRADING OPPORTUNITIES per year.

Example :

Long 30 oil Stocks
Short 30 semiconductor Stocks
└ 60 bets?
 └ Only 1 Bet!!!

} 1 Bet
└ Oil stocks will outperform Semiconductor Stocks.

To MAXIMIZE the number of INDEPENDENT BETS, remove exposure to RISK FACTORS.

└ Risk Factors indicate which stocks have strong COMMONALITY.
 └ Positions with STRONG COMMONALITY ≠ Holding INDEPENDENT positions.
 └ This is why we DEMEAN by Sector in the alpha research stage.

Ways to improve SHARPE RATIO

└ Improve alpha factor (IC) — SUPER Challenging [Warren Buffet, George Soros, Jhujhunwala]
└ Increase the number of independent trades [James Harris Simons]
 └ Cross sectional strategies with multiple stocks Maximizes Breadth.
 └ Combining multiple alpha factors Maximizes Breadth.

Real World Constraints
- Liquidity — Bid-Ask Spread
  - difference b/w the ask and bid price of a stock
  - Sell : $100
  - Buy : $100.05

  } Bid-Ask spread is 5 basis points
  - Good Alpha vector — 400 basis points
  - Illiquid stocks Bid Ask Spread — 20 to 30 basis points
    - Too much for a 400 point ~~to~~ alpha vector
- Transaction Costs
  - Explicit Cost
    - Commisions
  - Implicit Cost
    - Market Impact

Institutional players deal with huge implied transaction costs
- Portfolio Turnover is a good proxy to measure Transaction Costs
  - Quaterly Factor values based on fundamental data have lower transaction costs.
  - Factors that are updated daily have ~~a~~ higher transaction costs.
    - Turnover per day $= \dfrac{\text{Value of Trades per day}}{\text{Portfolio Value}} = \dfrac{\$2 \text{ million}}{\$100 \text{ million}} = 2\%$
  - Factor Rank Autocorrelation
    - Used when raw factor values have been converted into RANKS

FACTOR RANK AUTOCORRELATION

- value is close to 1
  - stock rankings don't change much from day to day
  - HIGH FRA ⟹ lower turnover
  - LOW FRA ⟹ higher turnover
- Find the Spearman Rank correlation between ranked alpha vector of previous day ($t-1$) and the ranked alpha vector of current day (t) over a time window and form a time series. $corr(\alpha_{t-1}, \alpha_t)$
- Two alpha factors have
  - Similar Sharpe Ratios
  - Similar Quintile Performance
  - Similar Factor Returns

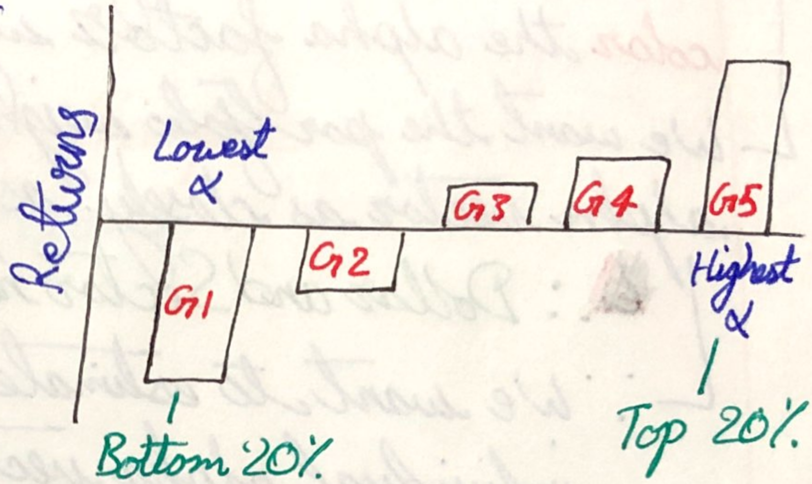
  } Always prefer the one with LOWER TURNOVER
- Alpha Factor with high Sharpe Ratio and high turnover requires careful evaluation and backtesting.
- We want higher Breadth $\sqrt{B}$ ⟹ Higher Turnover
  - ⟹ Higher Transaction Costs

  As usual, we have a TRADE-OFF
- Rank IC — Correlation of Alpha factor with FORWARD RETURNS.
- FRA — stability of the ranked alpha vectors from day to day.

Rank IC explains how the overall alpha vector performs

└ Quintile Analysis can explain which subset of stocks contributes the most or the least to the factor returns of the portfolio.

Ideally we would want to see a MONOTONIC Relationship among the quantiles.



└ Any deviation from this expected performance

- Alpha too risky or even Invalid
- Fat tails making the top and bottom perform extremely well while the G2/3/4 are insignificant
  - └ Factor Returns are relying on a smaller subset of the Alpha factor
- G2/3/4 are contributing significantly but G1 ~~are~~ and G5 are underperforming.

$\underbrace{\qquad\qquad\qquad\qquad}$

Examples of Deviation from the expected Monotonic Relationship

We take care of the risk from within Alpha Factors
- Do NOT RELY completely on the final Optimization Step
- This ensures that the final optimization step doesn't color the alpha factor weights too much.
- We want the portfolio weights to reflect the alpha vector as closely as possible
  - : Dollar and Sector neutral weights.
  - ∵ We want to estimate the performance of an individual alpha vector in isolation
    - ∴ Don't want the final optimization to alter the weights too much.

Transfer Coefficient
- Measures how closely the portfolio weights match the original Alpha Vector

Alpha Vector (α) → $\boxed{\text{Portfolio Optimizer}}$ → Portfolio Weights (x)

$$\boxed{\text{Transfer Coefficient} = corr(\alpha, x)}$$

- Should be as close to 1 as possible

Alpha Factor Lifecycle

Proposal and General Evaluation of an Alpha Factor

↓

Perform an Out-of-Sample testing of the Alpha Factor

↓

Conduct Paper Trading without using Real Money on Live Market Data

↓

Put the alpha in Production with real money. It will be blended with other alphas and pass through Portfolio Optimizer

↓

Start with giving a Small Weight. Give More Weight if Portfolio Performance is improved.

↓

REMOVE the alpha factor after its usefulness ERODES

Alpha Factor building blocks we are going to Learn in Term 1 20th December 2018

- Overnight Returns — %age change from the Close Price of the previous day to the Open Price of the next day.
  - Also called Close to Open Returns
- Accelerated/decelerated Gains/Losses
- Positive Skew
- Idiosyncratic Volatility

"PERSISTENCE" ⟹ MOMENTUM

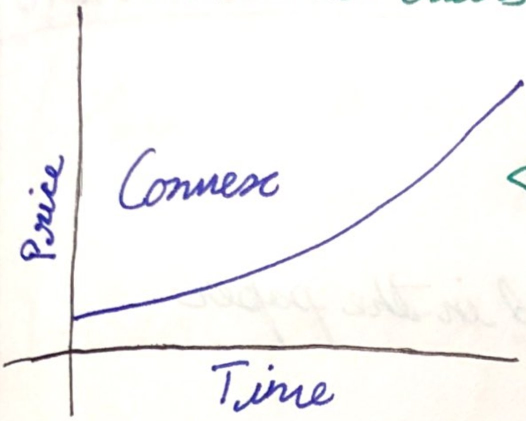
"High Overnight returns underperform" ⟹ MEAN REVERSION (REVERSAL)

"AND" ⟹ CONDITIONAL FACTOR

Mean Reversion, Momentum AND Harder to value firms

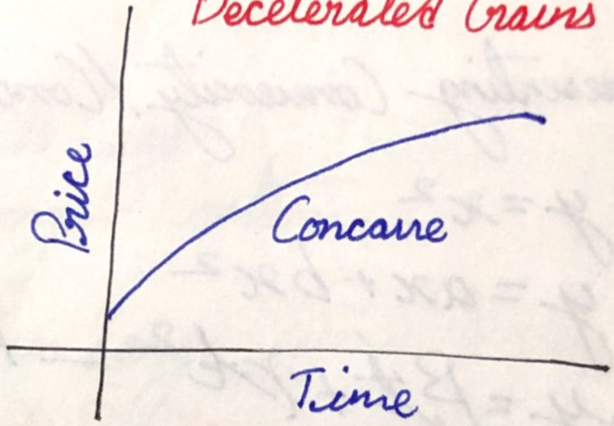
Paper 1 Strategy,
- Calculate Overnight Returns
- Aggregate Weekly Overnight Returns
- Momentum
  - Overweight stocks with higher weekly overnight returns
  - Underweight stocks that have low weekly overnight returns

Accelerated Gains

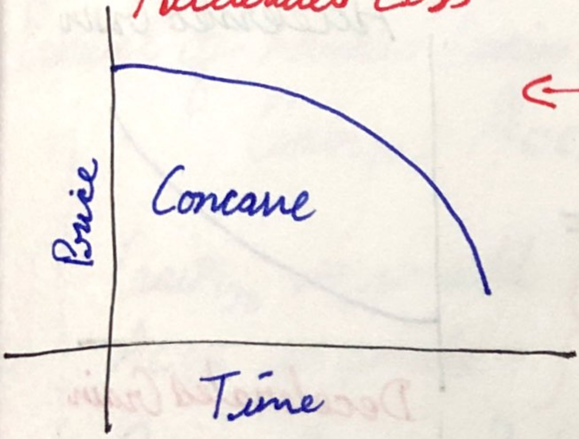


← Long
MORE

Decelerated Gains

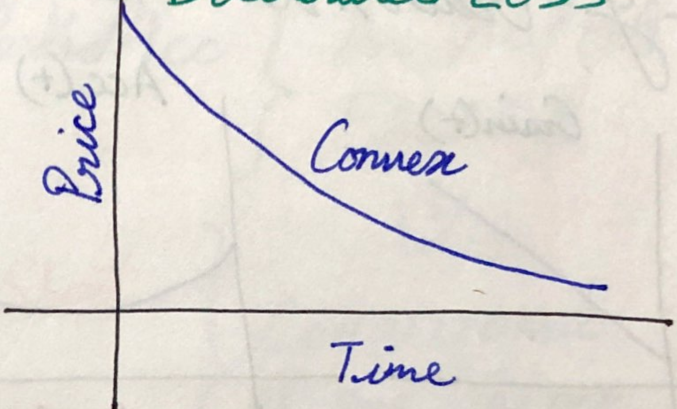


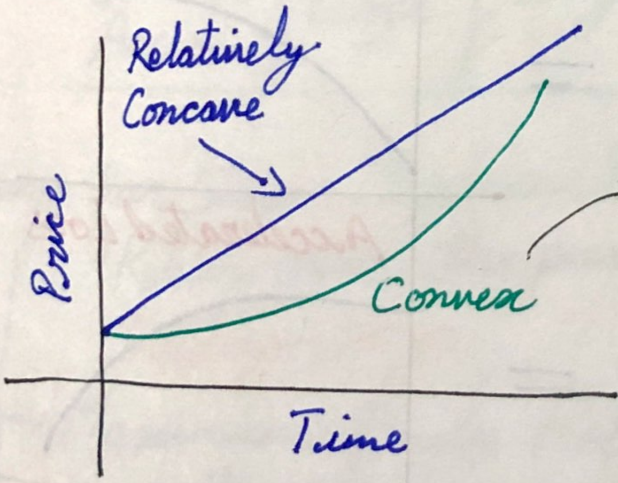
Accelerated Loss



← Short
MORE

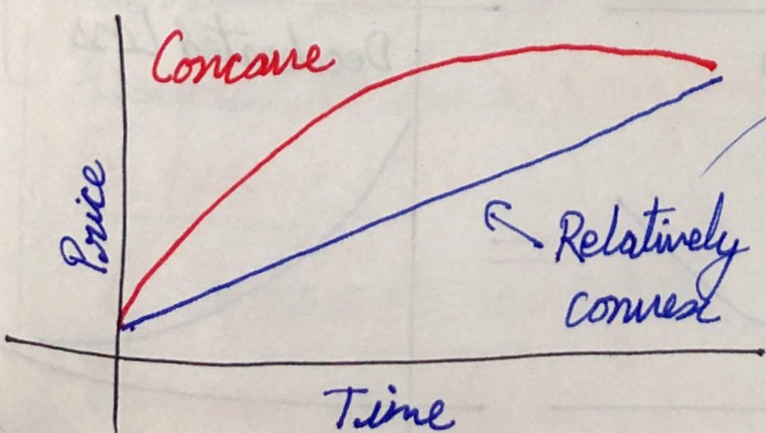
Decelerated Loss





→ Convex is a better choice in this case.

Long Convex MORE

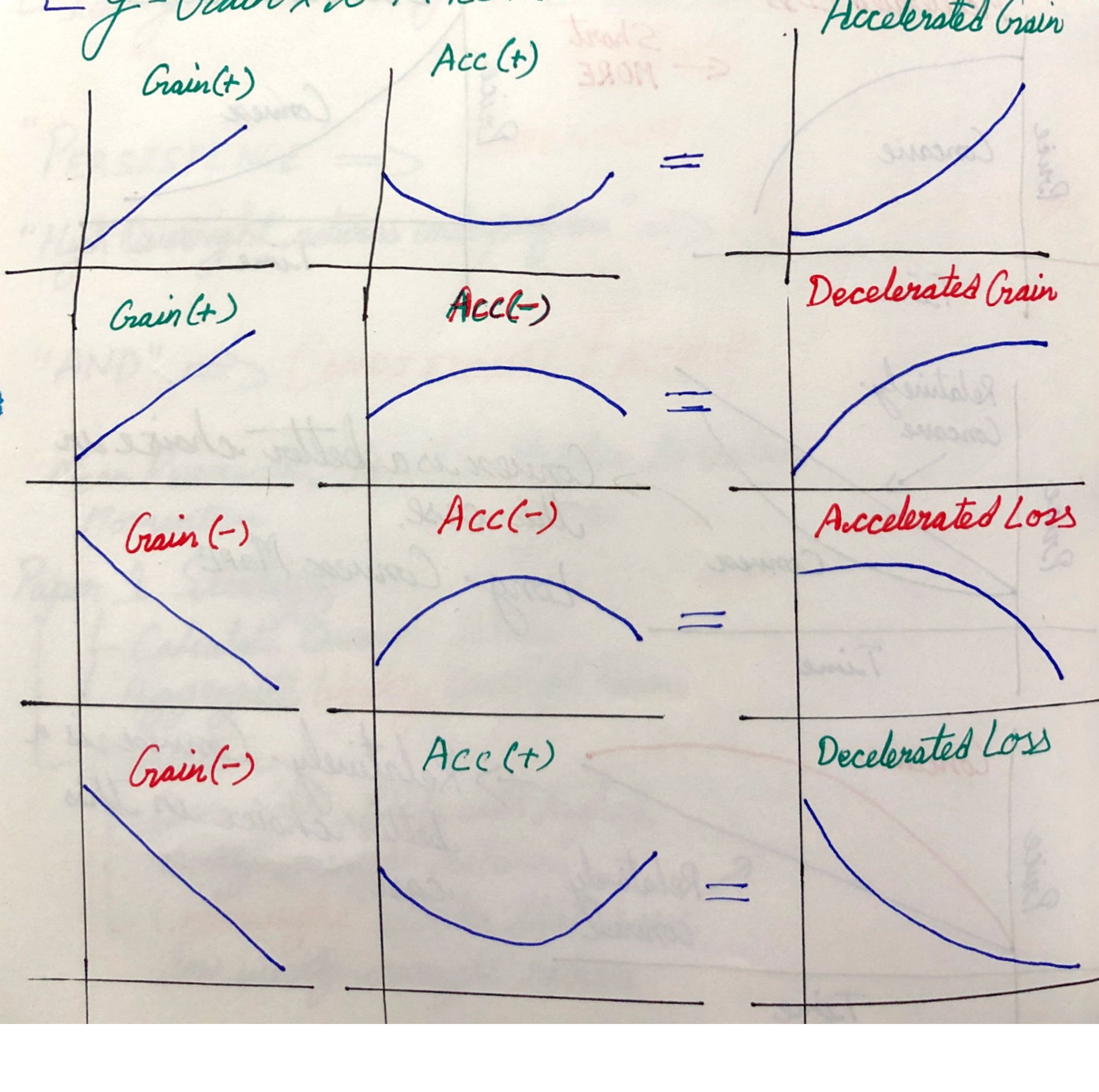


→ Relatively Convex is a better choice in this case.

21st December 2018

Representing Convexity / Concavity

- $y = x^2$
- $y = ax + bx^2$
- $y = \beta t + \gamma t^2$ — As represented in the paper
- $y = \text{Gain} \times t + \text{Acc} \times t^2$



Product of Gain x Acc

- Gain(+), Acc(+) — Product is Large
  - Large Long Position
- Gain(−), Acc(−): If Product is Large
  - Large Short Position

Product of Ranked Gain times Ranked Acc } An ALPHA FACTOR ⇓

$Gain_r \times Acc_r$

- $Gain_r$ is small, Acc_r is small } Large Short
- $Gain_r$, Acc_r } Large } Large Long

} CONDITIONAL ALPHA FACTOR

Momentum AND Convexity

Proxy for Skew: Maximum Daily Returns over the past 20 trading days

- Maximum Daily Return over the trailing month
- Assumed to be a REVERSAL FACTOR
  - MEAN REVERSION
- Measure of "Positive" Skew

Paper 2 Strategy,

- Run a Multiple Regression with t & t^2 being independent variable
  - Dependent Variable = Stock PRICE
- Regression gives estimates for Gain & Acc
- Calculate $\underbrace{Gain_r \times Acc_r}_{\text{Ranked Gain and Acc coefficients}}$

Possible Combinations

- Momentum (+) SKew (+)
  - "Weakened Momentum"
- Momentum (+) SKew (less +)
  - "Enhanced Momentum"
- Momentum (-) SKew (+)
  - "Enhanced Momentum"

Strong Mean Reversion in this scenario

- Momentum (-) SKew (less +)
  - "Weakened Momentum"

Ranked Momentum

Ranked SKew — Mean Reversion Factor ⟹ Reverse Ranked Order for SKew

$\boxed{\text{Joint Factor: Momentum}_r \times \text{SKew}_r}$

- Smallest Alpha Value
  - Largest Short Position
- Largest Alpha Values
  - Largest Long Positions

Arbitrage
- Seeks profits due to mis-pricing of the assets
- Reduces the mispricing of the assets
- Buying underpriced assets and shorting overpriced assets also REDUCES mispricing of the assets.
  - a type of Arbitrage

Arbitrage opportunity in 2 stocks
- A – Low Risk – Already arbitraged due to low risk by HFTs.
- B – High Risk — Better choice due to less competition

Returns → Systematic + Idiosyncratic (specific)
Volatility of Returns → Systematic + Idiosyncratic (specific)
} from the Risk Model

Use idiosyncratic ~~risk~~ risk as measure of
- Arbitrage Risk
- Market participants are neutral to common factor risks
- But they are still exposed to idiosyncratic risk

} This is whe… you ma… $$$$

- Fit regression model with risk factors to isolate iVol
  - Residual = Actual Return – Estimated Return
  - idiosyncratic Risk "iVol" — Stdev (residual)
  - iVol — Arbitrage Risk

Quant + Fundamentals = QUANTAMENTAL INVESTING

Fundamentals driven Alpha
- Lower frequency (quaterly)
- Lower Turnover — Low Transaction Costs
- Less Responsive
- High Capacity
- Lower Sharpe Ratio

We need a combination of fundamentals and price data driven alpha
- Conditional Factor — Fundamental AND Other Factor
  - iVol AND Fundamental Factor
    - use as a Conditioning Factor

ADVANCED PORTFOLIO OPTIMIZATION

22nd December 2018

- Could be a portfolio from scratch
- Or changing the weights of an existing portfolios in production

alpha model $\longrightarrow$ objective

risk model $\longrightarrow$ constraint

$$\boxed{\text{minimize}: -\alpha^T x}$$

-ve ensures that we maximize alpha

$$\boxed{B^T F B + S \;\Big\}\; \text{Portfolio Covariance Matrix}}$$

- $\underbrace{x^T (B^T F B + S) x}_{\text{Portfolio Risk}} < C$
  - Usually given
  - Represents a business decision

Modified Objective Function

- minimize $-\alpha^T x + \lambda \|x\|_2$

Prevents the algorithm from putting too much weight on a single stock

- $\lambda = 0$
  - Full Confidenece in the alpha vector
- $\lambda = \infty$
  - No confidence in alpha vector
  - All the stocks get equal weightage
- λ = REGULARIZATION PARAMETER

Some real world Constraints

- Long only portfolio $x \geq 0$
- Market Neutral Portfolio $\sum_i x_i = 0$
- Limit the leverage to 1.
  - $\sum_i |x_i| \leq 1$
  - Not required in our case

    Since,

    Objective
    - minimize : $-\alpha^T x$

    Constraint
    - $x^T(B^TFB+S)x < C$
- Limiting the exposure to an individual factor

  $B^T x \leq K_{max}$

  $B^T x \geq K_{min}$
- Constraints on individual asset weights

  $$\left.\begin{array}{l} x \leq M_{max} \\ x \geq M_{min} \end{array}\right\} \text{Prevents taking extreme positions in individual assets}$$

if minimize: $-\alpha^T x + \lambda \|x\|^2$

DOES NOT produce satisfactory results

minimize: $\underbrace{\|x - x^*\|}$

Try to 'track' the new portfolio weight vector x^* as closely as possible

$$\boxed{x^* = \frac{\alpha - \bar{\alpha}}{\sum_i |\alpha_i|}}$$

objective: $\boxed{\text{minimize } (x - x^*)^T (B^T F B + S)(x - x^*)}$

Minimizing the distance between portfolio weights AND The Risk introduced by this distance.

objective: $\boxed{\text{minimize } -\alpha^T x + \lambda \left[x^T (B^T F B + S) x \right]}$

Maximize Alpha + Minimize Risk

We avoid the problem of an infeasible solution to the optimization due to conflicting constraints in this case

Infeasible Problems

Minimize:
$-\alpha^T x$
Constraints:
turnover $< t$

If a constraint yields an infeasible solution, move it to the minimize term and penalize it →

Minize:
$-\alpha^T x + p(\text{turnover})$
Constraints:

penalize this
(penalty term)

Objective Function:
Minimize: $-\alpha^T x$
Constraints: turnover $< t$

Decrease a little at every iteration

Keep iterating over the loop till the hard constraint of the turnover yields a Feasible Solution

NUMPY

20th September 2018

x = np.ndarray([1,2,3,4,5]) One dimensional array

x.dtype Datatype stored in the nd-array

x.shape Dimensions of the nd array, returns a tuple

x.size Total number of elements in the array

Numpy arrays must contain elements of the same datatype. Converts integers to strings.
L integers to floats if there are integers + floats

x = np.array ([1, 2, 3.5, 4.8], dtype = np.int64)
L [1, 2, 3, 4]

np.save('my_array', x)
y = np.load('my_array.npy')
} Saving and loading a numpy array on the disk

X = np.zeroes (3,4)
.ones (3,4)
} 3x4 matrix of either 0's or 1's

X = np.full ((4,3), 5) } 4x3 matrix with '5' as all the elements

X = np.eye(5) 5x5 identity matrix

X=np.diag([10,20,30]) $\begin{bmatrix} 10 & 0 & 0 \\ 0 & 20 & 0 \\ 0 & 0 & 30 \end{bmatrix}$ Diagonal matrix with the given values

x=np.arange(5) [0,1,2,3,4]

x=np.arange(2,6) [2,3,4,5]

x=np.arange(1,8,2) [1,3,5,7]

} End point is EXCLUSIVE

x=np.linspace(0,25,10) Creates 10 evenly spaced numbers b/w 0 & 25 } End point is INCLUSIVE

x=np.linspace(0,25,10,endpoint=False) → End point becomes exclusive.

x=np.reshape(x,(10,2)) Reshapes the array to the given size. No. of elements in both the arrays must be same.

X=np.random.random((3,3)) 3x3 matrix with random numbers

X=np.random.randint(4,15,(3,2)) 3x2 matrix with inclusive lower bound and exclusive upper bound

inclusive Lower Bound (4)

EXCLUSEVE Upper Bound (15)

$$X = np.random.normal(\underbrace{0}_{\text{mean}=0}, \underbrace{0.1}_{SD = 0.1}, size=\underbrace{(1000,1000)}_{1000\times1000 \text{ matrix}})$$

Creates a 1000x1000 matrix of elements drawn from the NORMAL distribution with a mean of 0 and standard deviation of 0.1

NP.arrays are MUTABLE.

x[3] = 28 Modifying the '4th' element of the array

x[0, 3] = 15 Modifying an element of a RANK 1 Array

x = np.delete(x, [0, 4]) List of indices that need to be deleted

Rank 1 array

0th element & 5th element

W = np.delete(~~Y~~, 0, axis=0) Delete the 1st Row of Y

Rank 2 Array

0th Row

Select 'Rows'

np.delete(Y, [0, 2], axis=1)

Rank 2 Array

Select Columns

Delete 1st & 3rd Column of Y

x = np.append(x, [7,8]) Add the numbers in the list (7&8) at the end of the array

Y = np.append(Y, [[10,11,12]], axis=0) Append a ROW

Y = np.append(Y, [[10],[11],[12]], axis=1) Append a COLUMN

The size of rows/columns should match.

x = np.insert(x, 2, [3,4]) Insert this at index this

W = np.insert(Y, 1, [4,5,6], axis=0) Insert ROW at this

with these elements

z=np.vstack((x,y)) Vertically stacking x, y arrays

w=np.hstack(($\underbrace{(y, x.reshape(2,1))}_{\text{Shape of } x \text{ \& } y \text{ must match}}$)) Horizontal Stacking

ndarray[start:end]
ndarray[start:]
ndarray[:end]

Three types of slicing.
'end' index is always EXCLUDED
'start' index is always INCLUDED

start/end could be ndarrays themselves

z=np.diag(X) Extracts the diagonal ~~items~~ elements of the array X

z=np.diag(X, K=1) Extracts the elements above the diagonal

z=np.diag(X, K=-1) Extracts the elements below the diagonal

np.unique() Extract all the unique elements from the ndarray

X[X>10] Extract elements greater than 10

Supports complex logical constructs

np.intersect1d(x,y)
np.setdiff1d(x,y)
np.union1d(x,y)

Standard set operations on ndarrays 'x' & 'y'

.sort() → used as a function()
└ Sorts out of place
└ Original ndarray is intact

→ used as a (method)
└ Sorts in-place
└ Original ndarray is manipulated

np.sort(x) Used as a function

$\underbrace{\text{x.sort()}}_{\text{Rank 1 array}}$ Used as a method

np.sort(x, axis=0) Sorts the 'rows'
Rank 2 array

Numpy allows elementwise operations as well as matrix operations on arrays.

x+y
x-y
x*y
x/y

Elementwise operations
Arrays should have the same size or be Broadcastable

np.split(X, [2, 3])

ary[:2] ary[2:3] ary[3:]

PANDAS

25th September 2018

groceries = pd.Series(data = [30, 6, 'Yes', 'No'],

index = ['eggs', 'apples', 'milk', 'bread'])

- contains different datatypes in a single instance
- index could be numeric or alphanumeric

groceries.shape	Same as numpy (size of each dimension of data)
.ndim	No. of dimensions of the data
.size	Total number of values in the array
.index	Prints the index labels
.values	Prints the values in the series

banana in groceries — FALSE

Checks if the index exists

groceries['eggs']
groceries[['milk', 'bread']]
groceries[0]
groceries[-1]
groceries[[0, 1]]

Indexing by index names or by numbers, like an array.

DEPRECATED (ambiguous)

groceries.loc[['eggs', 'apples']]

└ LOCATION — explicitly states that we are using a 'LABELED INDEX'.

groceries.iloc[[2, 3]] = ['More', 'Less'] } PD Series is mutable

└ INTEGER LOCATION — explicitly states that we are using a 'NUMERICAL INDEX'.

groceries.drop('apples') Drops element from the series 'Out of Place'

groceries.drop('apples', inplace=True) Drops elements from the series 'In Place'

Arithmetic Operations in a Pandas Series work exactly like ndarrays, given that the arithmetic operation exits for all the Datatypes present in the series.

PANDAS DATAFRAME is a two dimensional object that can hold multiple datatypes.

```
items = {'Bob' : pd.Series([245, 25, 55], index=['bike', 'pants', 'watch']),
'Alice' : pd.Series([40, 110, 500, 45], index = ['book', 'glasses',
                                                  'bike', 'pants'])}
```
} *

```
shopping_carts = pd.DataFrame(items)
```

	Alice	Bob
bike	500.0	245.0
book	40.0	NaN
glasses	110.0	NaN
pants	45.0	25.0
watch	NaN	55.0

Keys of the dictionary

NOT IN THE ORDER specified in the Dictionary

No data available, hence, NaN.

*Final Index is created by the 'UNION' of two indices supplied by pd.Series

```
data = {'Bob': pd.Series([245, 25, 55]),
        'Alice': pd.Series([40, 110, 500, 45])}
df = pd.DataFrame(data)
```

pd automatically assigned numerical index here.

	Alice	Bob
0	40.0	245.0
1	110	25.0
2	500	55.0
3	45	NaN

shopping_carts.index Gets indices [redacted]

.columns Gets column Labels

.values Gets df values

(5,2) ← .shape No. of rows & columns

2 ← .ndim Dimensions of df

10 ← .size Total no. of elements in df

bob_shopping_cart = pd.DataFrame(items, columns=['Bob'])

	Bob
bike	245
pants	25
watch	55

	Alice	Bob
pants	45	25.0
book	40	NaN

sel_shopping_cart = pd.DataFrame(items, index=['pants', 'book'])

	bikes	glasses	pants	watches
store 1	20	NaN	30	35
store 2	15	50.0	5	10

store_items[['bikes']]

	bikes
store 1	20
store 2	15

store_items[['bikes', 'pants']]

	bikes	pants
store 1	20	30
store 2	15	5

store_items.loc[['store 1']]

	bikes	glasses	pants	watches
store 1	20	NaN	30	35

store_items['bikes']['store 2'] 15

COLUMN name must always appear AHEAD of the ROW name when indexing df like this.

store_items['shirts'] = [15,2] Adds a new column 'shirts' with the new data

	bikes	glasses	pants	watches	shirts
store 1	20	NAN	30	35	15
store 2	15	50.0	5	10	2

```
new_items = [{'bikes': 20, 'pants': 30, 'watches': 35, 'glasses': 4}]
new_store = pd.DataFrame(new_items, index=['store 3'])
store_items = store_items.append(new_store)
```

Columns are in Alphabetical Order

	bikes 0	glasses 1	pants 2	shirts 3	suits 4	watches 5
store 1	20	NAN	30	15.0	45.0	35
store 2	15	50.0	5	2.0	7.0	10
store 3	20	4.0	30	NAN	NAN	35

store_items.insert(5, 'shoes', [8,5,0])

- 5 — insert before 'watches' column
- 'shoes' — Name of the Column
- [8,5,0] — Data

store_items.pop('suits') This method only removes columns

store_items.drop(['watches', 'shoes'], axis=1) Removes columns... (axis=1: Selects Columns)

.drop(['store1', 'store2'], axis=0) ...as well as rows (axis=0: Selects Rows) store_items = store_items.rename(columns={'bikes':'hats'})

└ Renames the column 'bikes' to 'hats'

Notice when renaming columns, 'columns' is used. When renaming rows, 'index' is used.

Renames the row index 'store 3' to 'last store'

store_items = store_items.rename(index={'store 3':'last store'})

store_items = store_items.set_index('pants') Changes the index of df

VERY IMPORTANT feature of PANDAS

	hats	glasses	shirts	suits
pants				
30	20	4.0	NAN	NAN

28th September 2018

	bikes	glasses	pants	shirts	shoes	suits	watches
store1	20	$\xrightarrow{*2}$ NAN	30	15.0	8	45.0	35
store2	15	50.0 ↑ *3	5	2.0 ↓ *1	5	7.0 ↓ *1	10
store3	20	4.0	30 $\xrightarrow{*2}$	NAN	10 $\xrightarrow{*2}$	NAN	35

x = store_items.isnull().sum().sum()

NAN in each column

Sums all the NANs

└ Calculates the total number of NAN values in the DataFrame store_items.count ( ) Counts the total number of Non-NAN values

store_items.dropna (axis=0) Removes Rows with NAN

└ Selects Rows

store_items.dropna (axis=1) Removes Columns with NAN

└ Selects Columns

Out of place modification, original df is still intact.

store_items.dropna (axis=1, inplace=TRUE) Inplace Modification

store_items.fillna (n) Replaces NAN values with 'n' value

store_items.fillna (method='ffill', axis=0) *1

└ NAN values get replaced with the values in the previous rows.

modifies Out of Place
Use Inplace modifier

store_items.fillna (method='ffill', axis=1) *2

└ NAN values get replaced with the values in the previous columns

store_items.fillna (method='backfill', axis=0) *3

└ NAN values get filled with the value in the rows AFTER them.

store_items.interpolate (method='linear', axis=0)

└ Linearly interpolate NAN values.

google_stock = pd.read_csv(FILEPATH) Read from CSV file

google_stock.head( ) Top 5 rows of the df
google_stock.head(10) Top 10 rows of the df

└ Use .tail( ) for bottom of the df

google_stock.isnull().any( ) Checks if any of the columns has NAN value

└ Returns boolean True/False

google_stock.describe( ) Statistics about the df
google_stock['Adj Close'].describe( ) Statistics about a particular Column

google_stock.max()
.mean()
.min( )
.corr( ) Correlation b/w different columns

	Year	Name	Department	Age	Salary
0	1990	Alice	HR	25	50000
1	1990	Bob	RD	30	48000
2	1990	Charlie	Admin	45	55000
3	1991	Alice	HR	26	52000
4	1991	Bob	RD	31	50000
5	1991	Charlie	Admin	46	60000
6	1992	Alice	Admin	27	60000
7	1992	Bob	RD	32	52000
8	1992	Charlie	Admin	28	62000

data.groupby(['Year'])['Salary'].sum() → .mean() would give the

Year	
1990	153000
1991	162000
1992	174000

data.groupby(['Name'])['Salary'].sum()

Name	
Alice	162000
Bob	150000
Charlie	177000

data.groupby(['Year','Department'])['Salary'].sum()

Year	Department	
1990	Admin	55000
	HR	50000
	RD	48000
1991	Admin	60000
	HR	52000
	RD	50000
1992	Admin	122000
	RD	52000

pd.set_option('precision', 1) 'Precision of the df will be set to one decimal place.