Financial Signal Processing

Time Value of Money, NPV

Bonds

Return - Variance Estimation and Decomposition

Prof Danilo Mandic

Department of Electrical and Electronic Engineering

Imperial College London

February, 2020

Financial Signal Processing

Time Value of Money, NPV

Bonds

Return - Variance Estimation and Decomposition

Prof Danilo Mandic

Department of Electrical and Electronic Engineering

Imperial College London

February, 2020

Table of Contents

Table of Contents

- What would you prefer? \$100 now, in one year, 10 years or 100 years?
- OR same question different phrasing : How much \$$R_1$ would you want in one year to forgo \$100 now?
- You would want more because you sacrifice the pleasure of spending the money for one year.
- Assuming r_f: annual risk free rate constant at 5%, 100 is the net present values of r_1 in one year.
  - r_f paid annually $\rightarrow r_1 = 100 \times (1 + r_f) = 100 \times (1 + .05)$
  - r_f paid semiannually $\rightarrow r_1 = 100 \times (1 + r_f/2)^2 = 100 \times (1 + .025)^2$
  - r_f paid k times a year $\rightarrow r_1 = 100 \times (1 + r_f/k)^k$
  - As k goes to infinity $\rightarrow r_1 = 100 \times \exp(r_f)$
- For n years, constant r_f

$$R_n = 100 \times \exp(n \times r_f)$$

Forward pricing

Forward pricing formula:

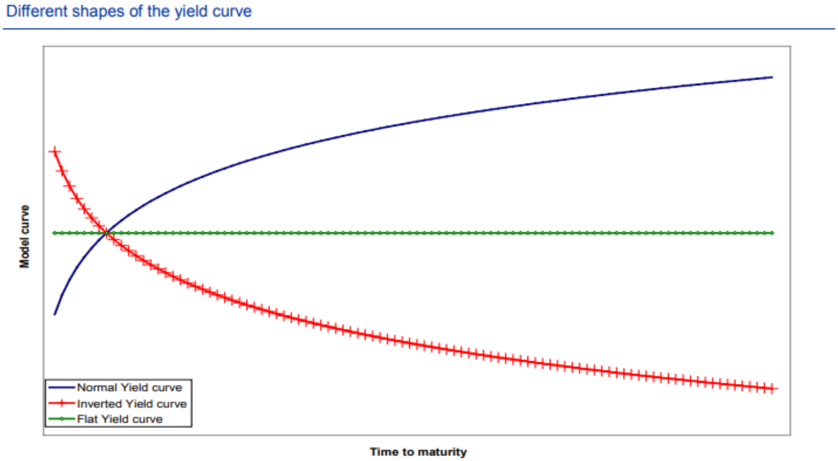
If the underlying asset is tradable and a divided exists, the forward price is given by:

$$F = S_0 e^{(r-q)T} - \sum_{i=1}^{N} D_i e^{(r-q)(T-t_i)}$$

- F is the forward price to be paid at time T;
- e^x is the exponential function (used for calculating continuous compounding interests)
- r is the risk-free interest rate
- q is the cost of carry
- S_0 is the spot price of the asset (i.e., what it would sell for at time 0)
- D_i is a dividend that is guaranteed to be paid at time t_i where $0 < t_i < T$.

https://en.wikipedia.org/wiki/Forward_price

Variable rate curves



https://www.ecb.europa.eu/pub/pdf/scpsps/ecb.sps27.en.pdf

- In real world rates not constant through time, typically upward sloping
- The investor would want more reward (per year) because he/she sacrifices the pleasure of spending the money for 10 years (not one).
- The value $V(t)$ at time t of an investment $V(0)$ earning continuously compounding fixed rate r for time t is given

$$V_t = V_0 e^{rt}, \ t \in \mathbb{R}_+$$

- For a variable rate function $r(s)$ (generally a stochastic process, since interest rates not predictable)

$$V_t = V_0 \exp\left(\int_0^t r_s ds\right)$$

Table of Contents

Bond Pricing

- The price of a default free bond depends on the interest curve rate (r_f curve)
- Since in reality the curve is stochastic, there is uncertainty (risk) in the bond value.
- So we need metrics that define that risk
- These metrics are additive they can also be applied to a set of bonds (portfolio)
- Metric of the portfolio is the weighted sum of the metric of the bond
- We will focus on duration and convexity

Bond Price Formula Intuition

$$
\begin{aligned}
P &= \left(\frac{C}{1+i} + \frac{C}{(1+i)^2} + \ldots + \frac{C}{(1+i)^N}\right) + \frac{M}{(1+i)^N} \\
&= \left(\sum_{n=1}^{N} \frac{C}{(1+i)^n}\right) + \frac{M}{(1+i)^N} \\
&= C\left(\frac{1-(1+i)^{-N}}{i}\right) + M(1+i)^{-N}
\end{aligned}
$$

where

- F is the face value
- i_F is the contractual interest rate
- $C = F \times i_F$ is the coupon payment (periodic interest payment)
- N is the number of payment
- i is the market interest rate, or required yield, or observed/appropriate yield to maturity
- M is the the value at maturity, usually equals to the face value
- P is the market price of the bond

https://en.wikipedia.org/wiki/Bond_valuation

Bond Price Formula Second Order Approximation

Yield: IR so curr price
= PV of future cash flows
Yield ↑, price ↓ IR ↑ price ↓

$$P_0 = \sum_{t=1}^{T} [C_t \times (1+K)^{-t}]$$

Provided that estimated future cash flow does not change, bond price will change with the passage of time (t) and with changes in yield (k). Using a Taylor Seires Approximation, the change in bond price, ΔP, over some small interval, Δt is give by the following partial differential equation:

change from time

change from yield

$$\Delta P = \frac{\delta P}{\delta t}\Delta t + \frac{\delta P}{\delta k}\Delta k + \frac{1}{2}\frac{\delta^2 P}{\delta k^2}(\Delta k)^2$$

$$\frac{\Delta P}{P} = \frac{1}{P}\frac{\delta P}{\delta t}\Delta t + \frac{\delta P}{\delta k}\Delta k + \frac{1}{2}\frac{\delta^2 P}{\delta k^2}(\Delta k)^2$$

http://www.appliedbusinesseconomics.com/files/gvsbdc01.pdf

Duration I

The Macaulay duration is one measure of the approximate change in price for a small change in yield.

$$Macaulay\ \ duration = \frac{\frac{1C}{(1+y)^1} + \frac{2C}{(1+y)^2} + \ldots + \frac{nC}{(1+y)^n} + \frac{nM}{(1+y)^n}}{P}$$

where P is the price of the bond, C is the semiannual coupon interest (in dollars), y is one-half the yield to maturity or required yield, n is the number of semiannual periods (number of years times 2), and M is the maturity value (in dollars).

In general, if the cash flow occur m times per year, the durations are adjusted by dividing by m, that is,

$$duration\ \ in\ \ year = \frac{duration\ \ in\ \ m\ \ periods\ \ per\ \ year}{m}$$

Duration II

Investors refer to the ratio of Macaulay duration to $1 + y$ as the modified duration.

The equation is

$$\textit{modified duration} = \frac{\textit{Macaulay duration}}{1 + y}$$

$y = \frac{1}{2} \times YTM$

YTM = yield to maturity (required yield)

The Modified duration is related to the approximate percentage change in price for a given change in yield as given by

$$\frac{dP}{dy} = -\textit{modified duration}$$

$$\textit{modified duration} = \frac{\frac{C}{y^2}\left[1 - \frac{1}{(1+y)^n}\right] + \frac{n(100 - C/y)}{(1+y)^{n+1}}}{P}$$

https://stanford.edu/class/msande247s/2009/summer%2009%20week%205/Fabozzi_BMAS7_CH04_Bond_Price_Volatility_Solutions.pdf

Duration III

Duration assumes a linear relationship between bond prices and changes in interest rates.

Duration IV

In actuality, however, prices fall at an increasing rate as interest rates rise; similarly, prices rise at an increasing rate as interest rates fall. This disparity associated with large upward move in interest rate. Conversely, duration will consistently under estimate the amount of price increase associated with a large drop in interest rates.

https://www.blackrock.com/fp/documents/understanding_duration.pdf

Convexity I

The percentage price change due to convexity is:

divide by P for % price change

$$\frac{dP}{P} = \frac{1}{2}(convexity\ \ measure)(dy)^2$$

y: YTM

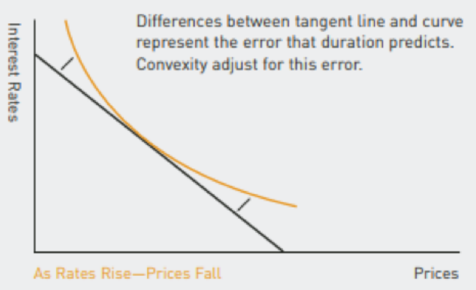
Convexity measure is defined as:

$$convexity\ \ measure = \left[\frac{2C}{y^3}[1 - \frac{1}{(1+y)^n}] - \frac{2Cn}{y^2(1+y)^{n+1}} + \frac{n(n+1)(100 - C/y)}{(1+y)^{n+2}}\right]\left[\frac{1}{P}\right]$$

$$convexity\ \ measure\ \ in\ \ year = \frac{convexity\ \ measure\ \ in\ \ m\ \ period\ \ per\ \ year}{m^2}$$

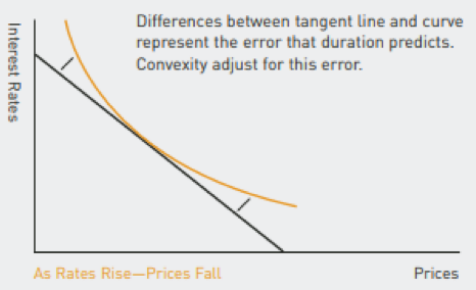
https://stanford.edu/class/msande247s/2009/summer%2009%20week%205/Fabozzi_BMAS7_CH04_Bond_Price_Volatility_Solutions.pdf

Convexity II



In order to compensate for this disparity, the concept of 'convexity' was developed. Convexity corrects for the error that duration produces in anticipating price changes given large movements in interest rate. As such, accounting for the dynamic relationship between prices and rates.

Convexity III



Convexity can help you anticipate how quickly the price of your bonds are likely to change given a change in interest rates. Everything else being equal, you may find issues with greater convexity more attractive, as greater convexity may translate into greater price gains as interest rates fall and lessened price declines as interest rate rise.

https://www.blackrock.com/fp/documents/understanding_duration.pdf

Table of Contents

- C: Coupon \$40, n: semi annual periods 4, P and M price and face value same \$1000, y:rf 0.04
- Basic formula

$$\text{Macaulay duration (half years)} = \frac{\frac{1C}{(1+y)^1} + \frac{2C}{(1+y)^2} + \cdots + \frac{nC}{(1+y)^n} + \frac{nM}{(1+y)^n}}{P}$$

$$= \frac{\frac{1(\$40)}{(1.04)^1} + \frac{2(\$40)}{(1.04)^2} + \cdots + \frac{4(\$40)}{(1.04)^4} + \frac{4(\$1{,}000)}{(1.04)^4}}{\$1,000} = \frac{\$3,775.09}{\$1,000} = 3.77509.$$

$$\text{Macaulay duration (years)} = \frac{\text{Macaulay duration (half years)}}{2} = \frac{3.77509}{2} = \mathbf{1.8875}$$

$$\text{Modified duration} = \frac{\text{Macaulay duration}}{(1+y)} = \frac{1.8875}{1.04} = \mathbf{1.814948}$$

https://stanford.edu/class/msande247s/2009/summer%2009%20week%205/Fabozzi_BMAS7_CH04_Bond_Price_Volatility_Solutions.pdf

- C: Coupon \$4, n: semi annual periods 4, P and M price and face value same \$100, y:rf 0.04
- Shortcut formula

$$\frac{\frac{C}{y^2}[1-\frac{1}{(1+y)^n}]+\frac{n(100-C/y)}{(1+y)^{n+1}}}{P}$$

$$=\frac{\frac{\$4}{0.04^2}[1-\frac{1}{(1.04)^4}]+\frac{4(\$100-\$4/0.04)}{(1.04)^5}}{\$100}=\frac{(\$362.98952+\$0)}{\$100}=3.6298952.$$

Converting to annual number by dividing by two gives a modified duration for **bond A** of **1.814948** which is the same answer shown above.

https://stanford.edu/class/msande247s/2009/summer%2009%20week%205/Fabozzi_BMAS7_CH04_Bond_Price_Volatility_Solutions.pdf

- C: Coupon \$4, n: semi annual periods 4, P and M price and face value same \$100, y:rf 0.04
- Convexity: Convexity measure (half years)=

$$[\frac{2(\$4)}{(0.04)^3}[1-\frac{1}{(1.04)^4}]-\frac{2(\$4)4}{(0.04)^2(1.04)^5}+\frac{4(5)(100-\$4/0.04)}{(1.04)^6}][\frac{1}{\$100}] =$$

$$[\$125,000(0.14519581)-\$16,438.542135+\$0][\frac{1}{\$100}] =$$

$$[\$18,149.4761-\$16,438.5421+0][\frac{1}{\$100}] = 1,710.934[0.01] = \mathbf{17.10934}$$

$$\text{Convexity measure (years)} = \frac{\text{convexity measure in m period per year}}{\mathrm{m}^2} = \frac{17.10934}{2^{(2)}} = \mathbf{4.277}$$

https://stanford.edu/class/msande247s/2009/summer%2009%20week%205/Fabozzi_BMAS7_CH04_Bond_Price_Volatility_Solutions.pdf

- C: Coupon \$40, n: semi annual periods 4, P and M price and maturity value same \$1000, y:rf 0.04. Since rf and coupon the same the bond is at par.
- Bond price if yearly rf is 0.05 (+ 100 bps), new duration 1.805159

$$P = C\left[\frac{1-\frac{1}{(1+r)^n}}{r}\right] = \$40\left[\frac{1-\frac{1}{(1.045)^4}}{0.045}\right] = \$143.501.$$

- The present value of the par or maturity value of \$1,000 is:

$$\frac{M}{(1+r)^n} = \frac{\$1,000}{(1.045)^4} = \$838.561.$$

- Thus, the value of a bond with a yield of 9%, a coupon rate of 8%, and a maturity of 2 years is: $P = \$143.501 + \$838.561 = \$982.062$. Thus, we get a bond quote of \$98.2062.
- The actual change in price is: $(\$982.062 - \$1,000) = -\$17.938$ and the actual percentage change in price is: $-\$17.938/\$1,000 =$ **-0.017938**

$$\frac{dP}{P} = -(\text{modified duration})(dy) = -1.805159(0.01) = -0.01805159$$

https://stanford.edu/class/msande247s/2009/summer%2009%20week%205/Fabozzi_BMAS7_CH04_Bond_Price_Volatility_Solutions.pdf

Table of Contents

DEFAULT RISK I

- In real world the bank that was supposed to give the interest might default (or the country whose government bonds we hold)
- Same is obviously true for companies
- Entities DO default, so the risk free rate is an abstraction. Typically when pricing bonds we take into account the probability of default. In general to estimate probability of default requires an advanced background in finance (1)
- We will give some intuition on the credit spread (additive to the risk free) to compensate for that probability.
  We assume that a bond paying 1 has a probability of default λ per period ($\Rightarrow$ survival 1-λ), loss given default L ($\Rightarrow$recovery $R = 1 - L$), r is the risk free rate and s is the spread over risk free. For simplicity we assume the bond is a one period bullet. Then we have

$$NPV = \frac{1}{1+r+s} = \frac{1(1-\lambda)+R\lambda}{1+r} \rightarrow s \approx (1-R)\lambda \rightarrow \lambda = \frac{s}{1-R}$$

1.http://www.gfigroup.com/portal/pdfs/CredDefSw1.pdf

S: Spread

payment of 1 → 1 given prob survive + recovery given default

DEFAULT RISK II

- Not all companies (and countries) have the same credit risk. Some are riskier than others.
- Rating agencies express views on the credit worthiness of entities : IG (investment grade, less risk), HY (high yield, more risk)
- Ratings IG (AAA,AA,A,BBB), HY (BB,B,CCC,CC,C)*
- Typical ratings agencies Moody's, S&P, Fitch

https://en.wikipedia.org/wiki/Credit_rating_agency#The_Big_Three_agencies

Estimated spreads and default rates by rating grade

Rating	Basis point spread [80][81][82]	Default rate[83][84]
AAA/Aaa	43	0.18%
AA/Aa2	73	0.28%
A	99	n/a
BBB/Baa2	166	2.11%
BB/Ba2	299	8.82%
B/B2	404	31.24%
CCC	724	n/a

Sources: Basis spread is between US treasuries and rated bonds over a 16-year period;[23][81] Default rate over a 5-year period, from a study by Moody's investment service[83][84]

DEFAULT RISK II

- Typical determinants of corporate credit risk are market cap, stock volatility, and debt

  Intuition (for well behaved companies):

  - Market cap 60m, std of market cap 20m $\Rightarrow$ so market cap is 3 std away from 0
  - When market cap reaches 0 the company defaults
  - The probability of that happening is (normality) is 0.15%, the probability of default of that company
- In reality we should not assume normality, as this is the most pronounced case of fat tails (and tales) in finance

DEFAULT RISK II

- The real calculation involves solving a system of stochastic differential equations (Merton model*)
- It involves quantities that are not directly observable, based on mathematical assumptions

$$DD(t) = \frac{\log(\frac{V_A}{D}) + (r - \frac{1}{2}\delta_A^2)(T - t)}{\sigma_A \sqrt{T - t}}$$

$$PD(t) = P[V_A \leq D] = \cdots = \Phi(-DD)$$

- V_A: value of the company, Debt + Equity
- Sigma: asset volatility
- T: time to maturity of the debt
- R: risk free

*https://en.wikipedia.org/wiki/Robert_C._Merton

Table of Contents

- In real world we might hold other assets, apart from non defaultable bonds, e.g. stocks.
- They do not guarantee the return over one year, rather the return value follows a distribution.
- Assume historical daily returns for a stock $r_1, r_2, \cdots, r_{256}$ (working days in a year). Assume that daily returns identically independently distributed.
- Then yearly stock return is $E(r) = \mathrm{mean}(r_1, r_2, \cdots, r_{256}) \times 256$
- Yearly volatility: $Vol(r) = std(r_1, r_2, \cdots, r_{256}) \times 16 = \sqrt{var(r_1, r_2, \cdots, r_{256}) \times 256}$
- Quite often it is more convenient to work with *var* since it is additive
- The above logic is distribution free, quite often we assume normal distribution (strong caveats, but much easier calculations)
- Assume prices at $P_0 = 100, P_1 = 102$, two consecutive days, Daily return can be defined either:
  - arithmetic returns: $ret = (101 - 100)/100 = 0.01$
  - log returns: $ret = ln(101/100) = .0099$, so that $P_1 = P_0 \times exp(ret)$
  - Unless you want to be very precise (pricing) the two methods yield similar results

- The *Vol* of the return distribution is extremely important
- Would you prefer to hold a stock $(E, Vol) = (5\%, 10\%)$ or $(3\%, 5\%)$
- The ratio E/Vol - Sharpe ratio: a key criterion (Signal to Noise ratio)   E = exp. return
- *Vol* generally not constant through time, returns can vary dramatically
- So given historic returns up to t, how do we calculate an estimate for *Vol* at $t + 1$?

- In general we impose some structure $Vol(t+1) = F(Vol(t), \cdots)$. A simpler model is autoregressive conditional heteroscedasticity model ARCH:
  - Estimate the best fitting autoregressive model AR(q):

$$y_t = a_0 + a_1 y_{t-1} + \cdots + a_q y_{t-q} + \epsilon_t = a_0 + \sum_{i=1}^{q} a_i y_{t-i} + \epsilon_t$$

  - Obtain the squares of the error $\hat{\epsilon}^2$ and regress them on a constant and q lagged values:

$$\hat{\epsilon}^2 = \hat{\alpha}_0 + \sum_{i=1}^{q} \hat{\alpha}_i \hat{\epsilon}_{t-i}^2$$

    where q is the length of ARCH lags.

- Typical models are the generalised ARCH - GARCH(p,q):

$$h_t = w + \sum_{i=1}^{p} \alpha_i (R_{t-i} - \mu)^2 + \sum_{j=1}^{q} \beta_j h_{t-j}$$

- Here w, α and β are parameters to be estimated (MLE), R is the return and h is the historic vol.
- One could say that p and q are the hyper-parameters.

http://www.stern.nyu.edu/rengle/EnglePattonQF.pdf

https://en.wikipedia.org/wiki/Autoregressive_conditional_heteroskedasticity

- Volatility of vol is an important metric, shows how volatile the vol is
- Typically vol is mean reverting
- Volatility clustering
- Exogenous (non asset) factors can affect volatility
- Assuming returns is the signal vol can be considered its power
- Typically one can estimate vol of a period by adding var of the subperiods as a first approximation

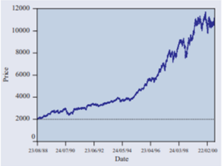


Figure 1. The Dow Jones Industrial Index, 23 August 1988 to 22 August 2000.

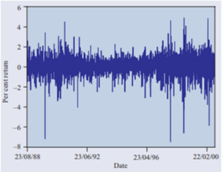


Figure 2. Returns on the Dow Jones Industrial Index.

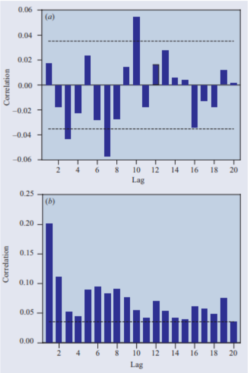


Figure 3. Correlograms of returns and squared returns.

Table of Contents

$R(t): r_t\,dt + r_s\,dw$

deterministic   stochastic

- We assume that the return of an asset at t, $R(t)$ is random.
- However there is intuition that there are exogenous factors affecting it:
  - If the market (all stocks) rally a given stock will probably also rally
  - if the oil price rallies, an oil's producer stock price will probably increase
  - if a country is in economic expansion probably its stocks will perform well
- We model the returns using a linear model, $R_{factor}(t)$ is the return of a factor
  - So $R(t) = \overset{\text{const}}{a} + b \times R_{factor}(t) + e$, a and b to be estimated, e the error term
  - b is the exposure of the asset to the factor b = beta a = alpha e = epsilon
  - e is the specific return (after accounting for systematic factor return)
  - a small in magnitude otherwise arbitrage (could be set to 0)   alpha should be small
- How do you estimate those?
- Why a linear model?

CAPM

Market portfolio could be e.g. S&P500 for simplicity

$$E(R_i) = R_f + \beta_i(E(R_m) - R_f)$$

- $E(R_i)$ is the expected return on the capital asset
- R_f is the risk-free rate of interest such as interest arising from government bonds
- β_i (the beta) is the sensitivity of the expected excess asset returns to the expected excess market returns, or also $\beta_i = \frac{Cov(R_i,R_m)}{Var(R_m)} = \rho_{i,m}\frac{\sigma_i}{\sigma_m}$
- $E(R_m)$ is the expected return of the market
- $E(R_m) - R_f$ is sometimes known as the market premium (the difference between the expected market rate of return and the risk-free rate of return).
- $E(R_i) - R_f$ is also known as risk premium
- $\rho_{i,m}$ denotes the correlation coefficient between the investment i and the market m
- σ_i is the standard deviation for the investment i
- σ_m is the standard deviation for the market m.

$n = 128$

$\frac{1}{2}$ yr

$$\left\{\begin{aligned}
r_1 &= r_{f1} + b(R_1 - r_{f1}) \\
r_2 &= r_{f2} + b(R_2 - r_{f2}) \\
&\vdots \\
r_n &= r_{fn} + b(R_n - r_{fn})
\end{aligned}\right.$$

Fama–French model I

$y = \alpha + bx + \varepsilon$ - linear reg.

$var(y) = b^2 var(x) + 1$ ← if $\varepsilon \sim N(0,1)$

$$r = \underset{\text{know}}{R_f} + \beta(\underset{\text{know}}{R_m} - R_f) + b_s \cdot SMB + b_v \cdot HML + \alpha$$

- r is the portfolio's expected rate of return, R_f is the risk-free return rate, and R_m is the return of the portfolio.
- The "three factor" β is analogous to the classical β but not equal to it, since threre are now two additional factors to do some of the work.
  - SMB stands for "Small [market capitalization Minus Big" usually true. small companies can (usually) outperform big ones
  - HML for "High [book-to-market ratio] Minus Low" book val ≠ mkt val. mkt price takes future cash flows into account, book val can't

They measure the historic excess returns of small caps over big caps and of value stocks over growth stocks.

book val: e.g. 500$ in assets, 400$ in liabilities
↳ ratio = $\frac{400}{500} = 0.8$.
higher ratio is better

return of mkt | SMB | HML

return of mkt	SMB	HML
R_{M_1}	$P_{S_1} - P_{B_1}$	$P_{H_1} - P_{C_1}$
R_{M_2}	$P_{S_2} - P_{B_2}$	$P_{H_2} - P_{C_2}$
$\vdots$	$\vdots$	$\vdots$
R_{M_n}	$P_{S_n} - P_{B_n}$	

: return of stock w/ highest book to mkt ratio - return ... w/ smallest book to mkt ratio

---

size of companies

C_{S_1}
C_{S_2}
} top 20% = P_S ← return of port. based on top 20%

$\vdots$

C_{S_k} } bottom 20% = P_B ← return of port. based on bottom 20%

returns : $P_S - P_B$

rank companies based on size for SMB,
use a f^n like size = log (mkt cap + debt)

↑ to compact vals - would otherwise be companies in millions vs billions

---

Fama–French model II

- What are the factors?
- What is the market?
- Factor Time invariance?

APT I

arbitrage pricing theory

want small set of uncorrelated factors ideally

- Risky asset returns are said to follow a factor structure if they can be expressed as:
  - $r_j = a_j + b_{j1}F_1 + b_{j2}F_2 + \ldots + b_{jn}F_n + \epsilon_j$

    linear comb of factors + error

    rule of thumb: divide no of stocks by 30

    eg. S&P 500 - 500 factors max (as 500 stocks)

    ↳ $\frac{500}{30} \approx 15$ factors needed

    where

    factor could be eg. SMB, HML etc.

    - a_j is a constant for asset j
    - F_n is a systematic factor
    - b_{jn} is the sensitivity of the jth asset to factor n, also called factor loading,
    - and ϵ_j is the risky asset's idiosyncratic random shock with mean zero.
- Idiosyncratic shocks are assumed to be uncorrelated across assets and uncorrelated with the factors.

APT II

$E(r_j) = E(\lambda + b_1 F_1 + \dots + b_n F_n + \varepsilon)$

$= \lambda + E[b_1 F_1 + \dots + b_n F_n] + 0 \leftarrow E[\varepsilon] = 0$ as uncorrelated & $\varepsilon \sim N(0,1)$

$= \lambda + b_1 E(F_1) + \dots$

- The APT states that if asset returns follow a factor structure then the following relation exists between expected returns and the factor sensitivities:
  - $E(r_j) = r_f + b_{j1} RP_1 + b_{j2} RP_2 + \ldots + b_{jn} RP_n$ exp ret. of asset = rf ret. + Σ sensitiv. factor risk premium
  - where
    - RP_n is the risk premium of the factor, how much to be paid for risk of factor
    - r_f is the risk-free rate,
- That is, the expected return of an asset j is a linear function of the asset's sensitivities to the n factors.

Table of Contents

Assume for 1 asset CAPM holds

no cover between as e is error

- $r = b \times f + e$
- $var(r) = var(b \times f + e) = var(b \times f) + var(e) = b^2 \times var(f) + var(e)$
- where: $r \sim$ asset return, $b \sim$ market exposure, $f \sim$ market return, $e \sim$ specific return
- f and e independent by construction (time series regression)
- So the variance of an asset is a linear combination of the variance of the market factor and its specific variance

Assume *APT* holds for 100 assets (portfolio) and 5 factors (no constant term)

- $\mathbf{R} = \mathbf{B} \times \mathbf{F} + \mathbf{E}$

for 1 day

  - $\mathbf{R}$ (100x1) returns, B (100x5) betas, $\mathbf{F}$ (5x1) factor returns, $\mathbf{E}$ (100x1) specific returns
- the equation holds in vector form for one period (cross sectional regression)
- The portfolio variance follows:
- $var(\mathbf{R}) = var(\mathbf{B} \times \mathbf{F} + \mathbf{E}) = var(\mathbf{B} \times \mathbf{F}) + var(\mathbf{E}) = \mathbf{B} \times \overbrace{\mathbf{VCV}}^{var(f)} \times \mathbf{B}' + \mathbf{D}$

B^2

  - $\mathbf{VCV}$ (5x5) factor covariance matrix $B\ VCV\ B'$ $(100\times5)\times(5\times5)\times(5\times100)$
  - $\mathbf{D}$(100x100) diagonal E^2

$r_1 = \Sigma b_i \cdot F_i^1$

$r_2 = \Sigma b_i F_i^2$

$\vdots$

you can do this many times

- Advantages
  - Mathematical stability (**VCV** low dimensionality)
  - Intuition
  - Easier to estimate with factor portfolios
  - Less data needed

Choosing factors – estimating exposures (betas) I

- Fundamental models
  - Calculate beta, regress for factor exposure (100 equations 5 unknowns)
  - Fama-French factors
  - Macro factors (oil price), exposures?
  - Industry
  - Geography ...
- Statistical models
  - *PCA* on asset by asset **VCV** (100x100)
  - Get e.g. 5 factors, 5 first components, or explaining 50% of variance
  - Factor **VCV** diagonal
  - Timeseries regression for exposures

Choosing factors – estimating exposures (betas) II

- Issues
  - Intuition
  - Interpretabililty - In *PCA* fctor definition changes, *PCA* on Corr or **VCV**?
  - In fundamental collinearity
  - Thin industries

Table of Contents

Outlier vs Anomaly I

- An **outlier** is an acceptable data point(s) within the model; it may be far away from the bulk of the data, but it is still explainable (in the tails of the distribution).
  - Generated from the same process;
  - Normally used in risk management, portfolio, return analysis, etc.
  - A fertile ground for data scientists
- An **anomaly**refers to unacceptable data given the model; it violates the model assumption (comes from a different distribution).
  - Generated from another process;
  - Normally used in detecting: Credit card fraud, cyber-intrusion, terrorist activity, system breakdown, etc.

Outlier vs Anomaly II

Outlier vs Anomaly III

- assuming a Gaussian model (process), the probability of data points outside 3σ is 0.07% ($\approx$ impossible). Thus, when this happens, we treat them as anamalous points. However, for heavily-tailed distributions, data points outside 3σ are likely to happen. We may treat these points as outliers to calculate the risk (e.g., copulas in finance).

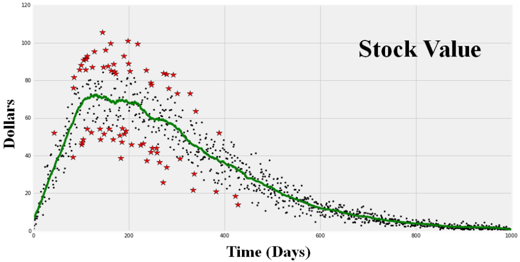


Figure: Red dots: Points outside the 2σ

Some heavy-tailed distributions (Black Monday 1987, Dot-com bubble)

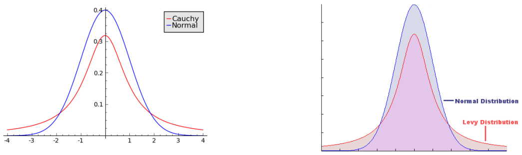


Figure: The heavy-tailed Cauchy (random mean) and Levy distributions

- Point anomalies: Detecting credit card fraud based on "amount spent"
- Contextual anomalies (common in time-series data): Spending 100 GBP on food every day during the holiday season is normal, but may be odd otherwise.
- **Finance**: Heavy tails imply additional risk, e.g. investor excessive optimism or pessimism leading to large market moves. In marketing, the 80-20 rule (20% of customers account for 80% of the revenue) is a manifestation of a fat tail distribution.

More heavy tailed distributions

- We now compare the following three distributions: Gaussian, Laplace, and the Cauchy (a special case of α-stable distributions)
- Alpha-stable distributions do not have a pdf, but they do have a characteristic function, and many standard pdf's are a special case of α-stable distributions.
- It is interesting to investigate how much of total data likelihood is captured by the width of σ, 2σ, 3σ, et. around the mean
- There are other neat ways of visualising the tails of these distributions, e.g. through the log which emphasises the differences in the small amplitude range.

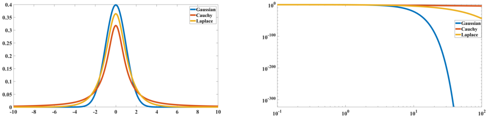


Figure: (a) 3 distributions (b) Tail properties in log-log axes

Table of Contents

Quadratic forms and positive–(semi)definite matrices

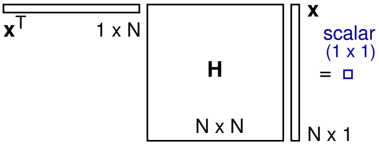
- Quadratic forms appear often in data analysis, and are expressed as $\mathbf{x}^T\mathbf{H}\mathbf{x}$, $\mathbf{x} \in \mathbb{R}^N, \mathbf{H} \in \mathbb{R}^{N\times N}$
- For simplicity, consider a 2nd order case, where

$$\text{variable vector } \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \text{ fixed matrix } \mathbf{H} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \tag{1}$$

- The quadratic form $Q_{\mathbf{H}}(\mathbf{x}) = Q_{\mathbf{H}}(x_1, x_2)$ of a matrix $\mathbf{H}$ is a scalar given by

$$Q_{\mathbf{H}}(x_1, x_2) = \mathbf{x}^T\mathbf{H}\mathbf{x} = \sum_{i=1}^{2}\sum_{j=1}^{2} h_{ij}x_ix_j = h_{11}x_1^2 + h_{22}x_2^2 + (h_{12} + h_{21})x_1x_2 \tag{2}$$

- If $Q_{\mathbf{H}}(x_1, x_2) \geq 0$, for any $\mathbf{x} \neq \mathbf{0}$, then the matrix $\mathbf{H}$ is called positive semidefinite ($\mathbf{H} \geq \mathbf{0}$)
- The matrix $\mathbf{H}$ is positive definite if $Q_{\mathbf{H}}(x_1, x_2) > 0$, $\forall\mathbf{x} \neq \mathbf{0}$



More on quadratic forms and covariance matrices

- Consider a vector of random variables $\mathbf{x} = [X_0, \ldots, X_{N-1}]^T$. Then, if these random variables are jointly Gaussian, their PDF is given by
$$p(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{N}{2}} \sqrt{det(\mathbf{C})}} e^{-\frac{1}{2}(\mathbf{x}-\mathbf{u})^T \mathbf{C}^{-1} (\mathbf{x}-\mathbf{u})} \text{ (quadratic form exponent)} \tag{3}$$
where $\mathbf{u} = E\{\mathbf{x}\}$ is mean vec. and $\mathbf{C} = E\{(\mathbf{x}-\mathbf{u})(\mathbf{x}-\mathbf{u})^T\}$ is covariance mat.
- For two jointly Gaussian random variables X_1 and X_2, the means $\mu_1 = E\{X_1\}$, $\mu_2 = E\{X_2\}$, variances $\sigma_1^2 = var(X_1)$, $\sigma_2^2 = var(X_2)$, covariance $\sigma_{12} = E\{(X_1 - u_1)(X_2 - u_2)\}$, and the correlation coefficient $\rho = \frac{\sigma_{12}}{\sigma_1 \sigma_2}$, then
$$p(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho)^2}\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right)} \tag{4}$$
Obviously, if X_1 and X_2 are uncorrelated, then $p(x_1, x_2) = p(x_1)p(x_2)$, and
$$p(x_1, x_2) = \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(x_1-\mu_1)^2}{2\sigma_1^2}} \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{(x_2-\mu_2)^2}{2\sigma_2^2}} \tag{5}$$

Quadratic forms, covariance matrices, and Gaussian PDF

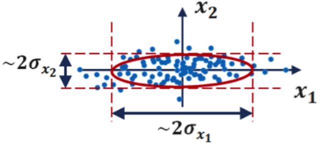
- For convenience, assume zero-mean data $\mathbf{x} = [x_1, x_2]^T \in \mathbb{R}^2$, then
$$p(\mathbf{x}) = \frac{1}{2\pi\sqrt{det(\mathbf{C})}} e^{-\frac{1}{2}\mathbf{x}^T\mathbf{C}^{-1}\mathbf{x}} \tag{6}$$
This is a quadratic form, as we can write $\mathbf{C}^{-1} = \mathbf{A}$ as another matrix. The "equi-potential" contours of this PDF are then determined by
$$\mathbf{x}^T\mathbf{A}\mathbf{x} = k \ (k \text{ is a constant}) \tag{7}$$
For this 2D case, the equi-potential contours of $\mathbf{x}^T\mathbf{A}\mathbf{x}$ are given by
$$a_{11}x_1^2 + a_{22}x_2^2 + (a_{12} + a_{21})x_1x_2 = k \ (\text{equation of an ellipse}) \tag{8}$$
Because $\mathbf{C}$ is symmetric, $\mathbf{C}^{-1}$ is symmetric too, so that $a_{12} = a_{21}$
- For uncorrelated x_1 and x_2, ellipse is aligned with the axes, since $a_{12} = a_{21} = 0$
$$\mathbf{C}^{-1} = \begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix} \Rightarrow \mathbf{C} = \begin{bmatrix} \sigma_1^2 = \frac{1}{a_{11}} & 0 \\ 0 & \sigma_2^2 = \frac{1}{a_{22}} \end{bmatrix} \tag{9}$$
- For correlated x_1 and x_2, ellipse is not aligned with the axes
$$\mathbf{C} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}, \sigma_{12} = \sigma_{21} \tag{10}$$

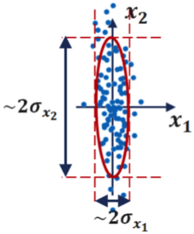
Correlation between RVs and error ellipsoids

- Consider a bivariate quadratic form, $\mathbf{x}^T\mathbf{C}\mathbf{x} = k$ (equi-potential ellipses):

Off-diagonal elements $= 0$

x_1 and x_2 are uncorrelated, $\sigma_1^2 > \sigma_2^2$

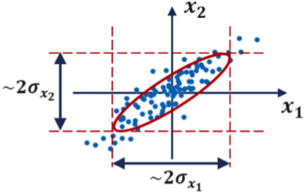


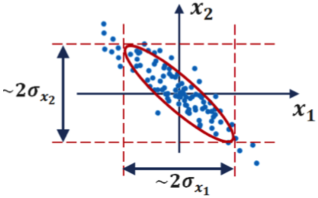


x_1 and x_2 are uncorrelated, $\sigma_1^2 < \sigma_2^2$

Off-diagonal elements $\neq 0$

x_1 and x_2 are positively correlated





x_1 and x_2 are negatively correlated

Table of Contents

An example with two assets I

- Main objective function:
  - argmax$\{expected_return - portfolio_variance\}$;
  - We want to maximise that by changing portfolio weights, vector **w**
- Expected return:

$$E(r_p) = \sum_i w_i E(r_i)$$

  where r_p is the return on the portfolio, r_i is the return on asset i and w_i is the weighting of component asset i (that is, the proportion of asset "i" in the portfolio).
- Portfolio return variance:

$$\sigma_p^2 = \sum_i w_i^2 \sigma_i^2 + \sum_i \sum_{j \neq i} w_i w_j \sigma_i \sigma_j p_{ij}$$

- For a **two asset** portfolio:
  - Portfolio return: $E(r_p) = w_A E(r_A) + w_B E(r_B) = w_A E(r_A) + (1 - w_A) E(r_B)$
  - Portfolio variance: $\sigma_p^2 = w_A^2 \sigma_A^2 + w_B^2 \sigma_B^2 + 2 w_A w_B \sigma_A \sigma_B \rho_{AB}$

An example with two assets II

In practical applications we would use

- factor model for risk
- between 5-20 signals for return forecasting
- Assume APT holds for 100 assets (portfolio) and 5 factors (no constant term)
  - $\text{argmax}\{return - l_f * factor\_risk - l_s * specific\_risk - tc * transaction\_cost\}$
  - We want to maximise that by changing portfolio weights, vector $\mathbf{w}$ (100×1)
  - Subject e.g. to total risk $<$ K
  - Alternative specifications might include return $>$ L, return/total_risk $>$ S

An example with two assets III

- Additional constraints might include:

$$sum(w_i) = 1, w_i > 0, \ (Long \ only),$$

$$sum(w_i) < L * fund \ size(leverage), \ if \ \ w_i > 0 \ long \ only$$

$$sum(w_i) < I \ \ for \ \ i \ \ in[...], \ factor \ \ exposure$$

An example with two assets IV

- Return: expected portfolio return
  - $\mathbf{w}' \times \mathbf{r}$
  - $\mathbf{r}$: expected asset return vector, (100x1)
- Transaction_cost
  - $(\mathbf{w} - \mathbf{w}_{t-1})' \times \mathbf{t}$
  - $\mathbf{t}$, transaction cost vector per asset, (100x1)
  - $\mathbf{w}_{t-1}$, holdings in previous period, (100x1)

An example with two assets V

- Factor_risk: the variance attributed to the factors
  - $\mathbf{w}' \times \mathbf{X} \times \mathbf{VCV} \times \mathbf{X}' \times \mathbf{w}$
  - $\mathbf{X}$: exposures to factors, (100x5)
  - $\mathbf{VCV}$: factor covariance, (5x5)
- Specific_risk: the variance attributed to specific asset return
  - $\mathbf{w}' \times \mathbf{D} \times \mathbf{w}$
  - $\mathbf{D}$: specific covariance matrix (100x100), can be assumed diagonal

An example with two assets VI

- Quadratic form
- Why break the variance in two?
- Constraints ...
- Same units!
- Forecasted risk – returns ...
  - Factor risk
  - Specific risk

An example with two assets VII

Forecasting returns:

- Where major research is happening
- Momentum
  - $s_t = \sum_{n=1}^{N} (\mathbf{w}'_{t-n} \times \mathbf{r}_{t-n})$
  - Trade s_t
- Mean reversion
- Growth – value
- Spread vs Probability of default
  - $Spr = a + b \times edf + e$
  - Trade e

Topics in optimisation

- Robust optimisation
- Different *VCV*
- Account for uncertainty in $E(r)$ estimation
- Adding terms (e.g. skew, kurtosis)