The Mathematics of Financial Risk Management

Luis A. Seco

Univ. of Toronto

To The Reader

These are the notes for the course on *The Mathematics of Risk Management*, given at the Annual SIAM meeting for 1998 in Toronto.

Comments, suggestions, etc. will be welcome at any time. They can be forwarded to

seco@math.toronto.edu

Acknowledgements

Thanks to the RiskLab team at the University of Toronto, who provided expert comments, errors, examples, etc. RiskLab is a joint initiative of the University of Toronto and Algorithmics Inc.

Parts of this course are based on joint work with C. Albanese, R. Brummelhuis, A. Córdoba, P. Fernández and M. Quintanilla.

Trivia

First known case of a mathematician in finance?

Thales bought options on the use of mills, in a year when a huge olive harvest was expected. Made a ton of money.

Risk:

If you know that a certain investment is going to lose 10%, there is no risk.

If you know a certain investment is guaranteed to earn between 10 and 20 percent, there is risk.

Besides banks, Math-Finance theories are also relevant in

- Government finance programs
- utility companies (debt refinancing).
- Insurance companies (environmental risk)
- Medicine (Vaccines, biotechnology risk)
- UNESCO: ecological risk.

Example 1.

In 1986, the Olympic Committee announces that Barcelona will hold the 1992 Games. The local organizing committee needs to prepare a budget (in USD) to submit for approval. The expenses will take place over a multi-year period, and will be in Spanish Pesetas (ESP).

In 1986, the USD exchanges at a rate of

$$1 \text{ USD} = 142 \text{ ESP}$$

The OC obtains a contract to purchase USD at a fixed rate of 114 ESP over the next six years. In 1992, USD has dropped to 90 ESP.

- Q1 (Pricing). How much does it cost to purchase such a contract?.
- Q2 (Pricing). At what rate does it come free? (Is it 114?).
- Q3 (Hedging). Did the financial institution that sold the contract loose money?
- Q4 (Risk Management). How can the financial institution know, back in 1986, how much money they could loose on the deal?

Financial Instruments – Equity

European Options Expire at a preset future time. Their pay-off f depends on the price of the underlying S at expiration.

Call options with strike K have pay-off given by

$$f(S) = (S - K)_+.$$

Put options have pay–off given by

$$F(S) = (K - S)_+.$$

American Options Can be exercised at any time in the future. Their pay-off is a function of the value of the underlying at that time.

Call options with strike K have pay-off given by

$$f(S, t) = (S(t) - K)_+.$$

Put options have pay–off given by

$$f(S, t) = (K - S(t))_+.$$

Asian Options Their price depends on the average value of the underlying. Can be issued with a European or American style.

Bermudan Options They are American options that can be exercised only at prescribed discrete future times.

Financial Instruments – Monetary

Bonds They pay a fixed amount (e.g., $1) at a future time. They are sold at a discount; their price determines interest rates. They usually pay coupons every few months or every year.

Bond Options Bonds can be bought or sold any time before they expire. Their price will fluctuate. As a consequence, they can be used as financial underlying for options. They are quite similar to equity, except for the fact that at the time of expiry of the bond, options make no sense. This in fact have very important implications.

Caps They are contracts that offer protection against time dependent interest rates rising over a certain ceiling, by paying the corresponding exceeding interest on a fixed notional.

Floors They charge the corresponding missing interest on a fixed notional. They have negative value.

Collars A combination of a cap and a floor. By setting the ceiling and floor appropriately, they can be issued for free.

Swaps They exploit the different interest rates that different parties will be charged for fixed and floating rate loans; a swap is a contract that exchanges future payments at fixed and floating rates.

Swaptions When a swaps is viewed as an underlying, options are issued on them.

Cross Currency swaps Same as swaps, but the exchange is between payments in two currencies.

Many other financial instruments are available for trade. Most of the time, they are designed with the objective of removing risk from uncertain future situations. They also offer risky speculative alternatives.

Common Terminology

Long Positions. Term used when the number of units of a certain instrument is positive.

Short Positions. Term used when the number of units of a certain instrument is negative.

Hedge.

Arbitrage.

Pricing Theories

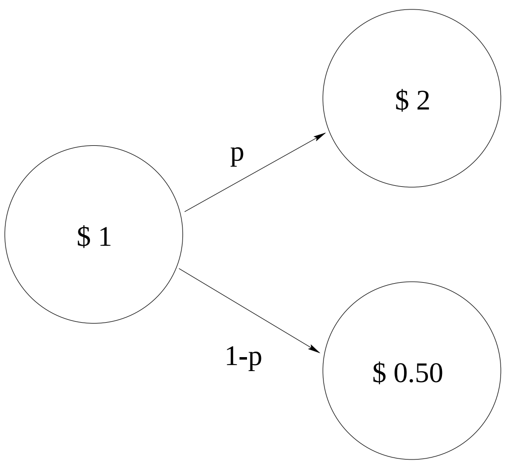
Prices should not be based on probabilistic expectations; that belongs in a casino. Instead, prices of instruments should be consistent with market prices of the instruments used in their hedging strategies. If one uses probabilistic considerations, the prices that mirror them should be in harmony with the observed market prices. One needs to search for the mathematical theory that supports this.

Contents

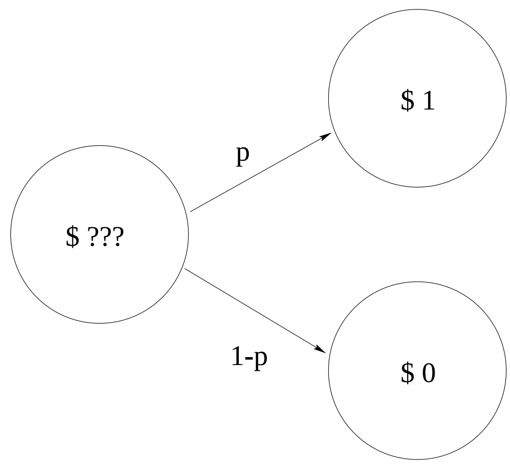
- Heuristic considerations
    - One period
    - Multiperiod
    - Continuous
- Pricing Theory
    - One Period
    - Stochastic Calculus
- Numerical Methods
    - Monte Carlo
    - Principal Components
    - Low Discrepancy
- Hedging
    - Implied volatilities
    - Greeks

Example 2

Example. (Ignore interest rates). Call Option. Pays $f_0(S) = (S - \$1)_+$.



1–Period Stock Tree



Option Value $= V$

Assume $p = 95\%$. Is $V = 0.95$?

Answer: No!.

$$V = 1/3\$.$$

Problems:

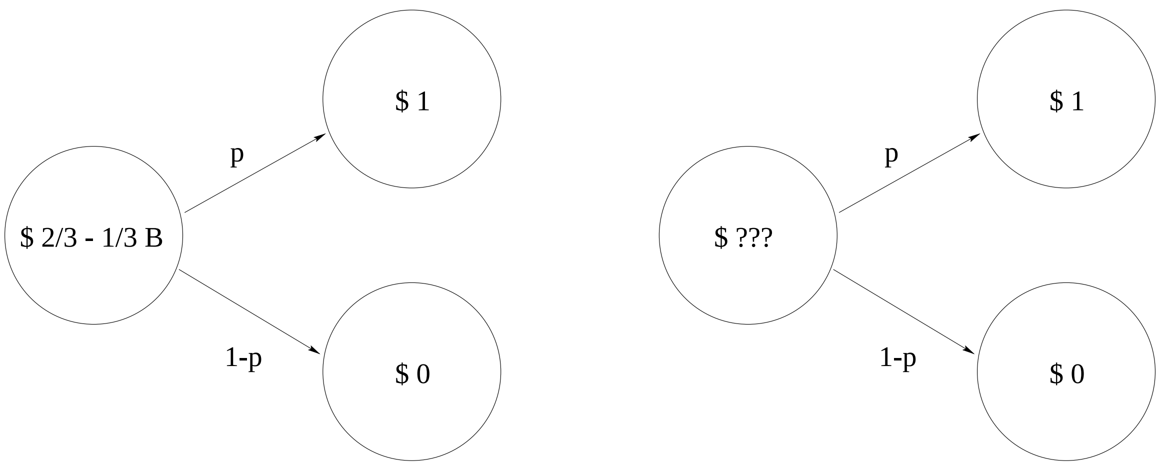
- How do we guess p?.
- Do we care?.

Discounted Values

Time is money. Assume the existence of a bond with constant interest rate r.

We build the following portfolio Π:

$$\Pi = \tfrac{2}{3}\ \text{Stock units} + (-\tfrac{1}{3})\ \text{bond}$$



Portfolio Values Option Value

No matter what p is, absence of arbitrage implies

$$\begin{aligned}\text{Option Price} &= \tfrac{2}{3} - \tfrac{1}{3}B \\ &= \tfrac{2}{3} - \tfrac{1}{3}e^{-rT}.\end{aligned}$$

where T is the time to expiration and r is the (constant) interest rate.

Implied Probabilities

We can still achieve

$$\begin{aligned} \text{Option Price} &= \mathbb{E}\left(e^{-rT} f_0\right) \\ &= p\, e^{-rT}, \end{aligned}$$

by selecting

$$p = \tfrac{2}{3}\, e^{rT} - \tfrac{1}{3}.$$

In other words, we can construct a probability measure $\mathbb{P}$ for the stock process, such that

$$\text{Option Price} = \mathbb{E}_{\mathbb{P}}\left(B_T^{-1}\, f_0\right).$$

More generally, if we define the (arbitrage-free) price to equal the discounted pay-off

$$V = B_T^{-1}\, f_0,$$

then, there exists a measure $\mathbb{P}$ under which V is a martingale: its value today is its expected future value.

Implied Market Data

Example: Assume the previous call option is sold for \$0.50.

$$\frac{2}{3} - \frac{1}{3}\, e^{-r} = 0.5.$$

Hence, the risk-free rate must equal

$$r = -\ln 2.$$

Example. Assume the stock valued at \$1 today, can be worth

$$S = \begin{cases} \$2 \\ \$1 \\ \$0.5 \end{cases}$$

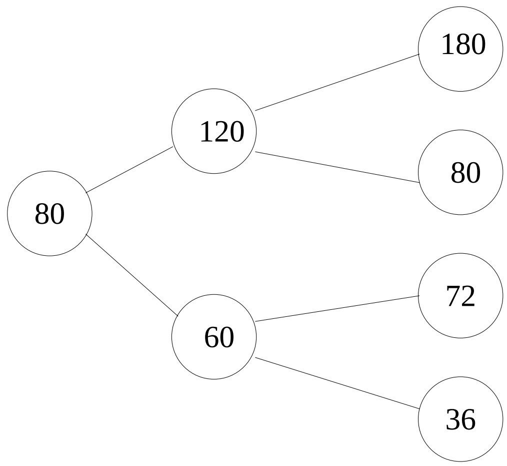
after a year. How can we price the call option with strike 1?.

Two possibilities:

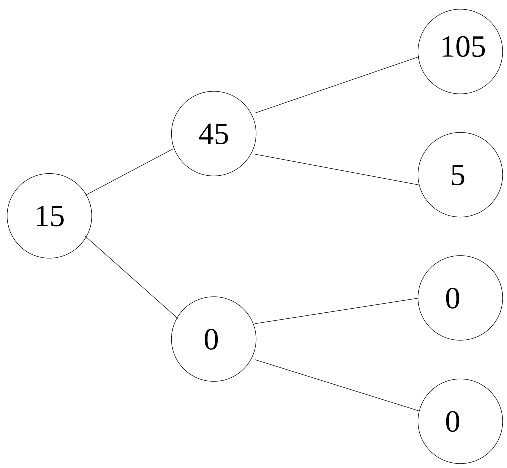
- Another derivative price is known
- We can re-balance our hedge once before maturity.

Multiperiod Pricing

Assume



A call option with strike \$75 can be priced as follows (r=0):



So its value today is \$15.

This is the arbitrage-free price. Implied probabilities can be obtained as usual.

Passage to the continuum

We think of infinitesimal time intervals dt.

Brownian motion moves up or down with probability $\frac{1}{2}$, by an amount of $\sqrt{dt}$:

$$dW = \pm\sqrt{dt}, \qquad \mathbb{E}(dW) = 0.$$

It is distributed at time t according to

$$P(x,t) \;=\; \frac{1}{\sqrt{2\pi t}}\;\exp\left(-\frac{x^2}{2t}\right).$$

Infinitesimal stock movements will be

$$dS = S\cdot(\mu\,dt + \sigma\,dW)\,.$$

Note that

$$d\log S = (\mu - \tfrac{1}{2}\sigma^2)\,dt + \sigma\,dW.$$

Ito's Lemma:

$$\partial_t f(S,t) = \partial_S f(S,t)\,dS + \partial_t f(S,t)\,dt - \tfrac{1}{2}\sigma^2\,dt.$$

Expected Discounted Payoff

European call option with payoff $f_0(S)$, expiration at time T:

$$f(S,t) \;=\; e^{-r(T-t)} \int_{-\infty}^{\infty} f_0\left(S_0 e^{(\mu - \frac{\sigma^2}{2})(T-t)\,+\,x} \right) P_\sigma(x, T-t)\, dx$$

$$P_\sigma(x,t) \;=\; \frac{1}{\sqrt{2\pi t \sigma^2}}\;\exp\left(-\,\frac{x^2}{2t\sigma^2} \right).$$

Problems

- What is μ?.
- What is σ?
- Can we replicate the price?

Bachelier (1900) worked out similar formulas.

The Black-Scholes Formulas

The price of a European call option on a stock S, valued today at S_0, maturing at time T with strike K, (constant) volatility σ and interest rate r is given by

$$V(t, K, \sigma, r) = S_0 \cdot N(d_1) - K \cdot e^{-r(T-t)} N(d_2),$$

where $N(d)$ is the cumulative normal

$$N(d) = \int_{-\infty}^{d} e^{-x^2/2} \frac{dx}{\sqrt{2\pi}},$$

and

$$d_1 = \frac{\ln(S_0/K) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}},$$

$$d_2 = \frac{\ln(S_0/K) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}.$$

The Black-Scholes Theory

Assume an option has price $f(t, S)$, at any given point in time, and any possible value of the underlying.

Let's set up the following arbitrage free argument:

At any point in time t, build a portfolio Π consisting of

$$a = -\partial_S f(S, t)$$

units of stock, and the option.

Using Ito's formula,

$$\begin{aligned} d_t\Pi &= d_t f + a\, dS \\ &= \tfrac{1}{2}\sigma^2\, S^2\, \partial_S^2 f + \partial_t f + \partial_S f\, dS + a\, dS \\ &= \tfrac{1}{2}\sigma^2\, S^2\, \partial_S^2 f + \partial_t f. \end{aligned}$$

This is a risk-free investment. Hence, it must earn risk-free interest and we obtain:

$$\begin{cases} \dfrac{\partial f}{\partial t} \;=\; -\dfrac{1}{2}\sigma^2 S^2 \dfrac{\partial^2 f}{\partial S^2} \;-\; r\, S\, \dfrac{\partial f}{\partial S} \;+\; r\, f, \\ f(S, T) = f_0(S). \end{cases}$$

It is a backward parabolic equation.

Pricing Theory (One Period)

Implied probabilities can be obtained, not only from prices dictated by arbitrage arguments, but also from market prices.

The implications of this is that a probabilistic approach to pricing is more useful than might have seemed from the considerations above.

In this section we assume there is a probability space for the payoffs of N securities available for trading,

A security is characterized by its *cost* now, and its *payoff* after one unit of time.

The **cost of the i-th security**, $i = 1, \ldots, N$, is q_i.

The **payoff** is given by the random variable $D_i(\omega)$.

The **expected payoff** of a security is $E(D_i(\omega))$.

A **portfolio** is a vector $\theta = (\theta_1, \ldots \theta_N) \in \mathbb{R}^N$, which represents the holdings of each security. θ_i can be positive or negative. If θ_i is positive, our position is said to be long. If θ_i is negative, our position is said to be short.

The **payoff** of the portfolio θ is $\theta \cdot D(\omega)$.

A market is said **complete** if

$$\text{Span } \left\{ \theta \cdot D(\omega)\ ,\ \theta \in \mathbb{R}^N \right\} \ = \ L^2(\mu).$$

and markets are usually assumed to be complete. In a complete market, for any payoff there is a portfolio with that payoff.

The **cost** of a portfolio θ is $q \cdot \theta$.

If a portfolio has nonzero cost, i.e. $q \cdot \theta \neq 0$, one defines its **return** to be

$$R_\theta(\omega) \ = \ \frac{\theta \cdot D(\omega)}{q \cdot \theta}.$$

In a real market, there are hedgers (people trying to minimize risk), speculators (people trying to maximize return) and arbitrageurs (people detecting market inefficiencies).

We say that there is an **arbitrage opportunity** if there is a portfolio θ such that

$$q \cdot \theta \leq 0, \text{ and } D \cdot \theta \geq 0 \quad a.e.,$$

and $D \cdot \theta > 0$ with non–zero probability.

The **Efficient Market Hypothesis (EMH)** states that there is no arbitrage and there are no transaction costs.

Theorem. (Riesz representation) *If p_i are linear functionals of the payoffs $L^2(\mu)$, then there exists a random variable $\pi(\omega)$ such that*

$$p \cdot \theta \;=\; E(\theta\pi \cdot D), \qquad \text{all } \theta \in \mathbb{R}^N. \tag{1}$$

If markets are complete, π is unique. If there are no arbitrage opportunities, $\pi > 0$.

In the case that we consider the cost as that linear functional, we obtain that the cost of a portfolio is the expectation of its payoff with probabilistic weight $\pi(\omega)$, which is called the **state–price deflator**. The name comes from the fact that

$$E(R_\theta \pi) \;=\; 1 \tag{2}$$

for all portfolios θ.

We always assume that $D_0(\omega)$ is constant for all $\omega \in \Omega$. This is a **savings account**.

A **riskless bond** is a portfolio θ_0 of constant payoff i.e. such that $\theta \cdot D(\omega) \;=\; \theta \cdot D(\omega')$ for all $\omega, \omega' \in \Omega$. It always exists: put $\theta = (1, 0, \ldots, 0)$. Then from (2) we find

$$R^0 \;\equiv\; E(R_{\theta_0}) \;=\; \frac{1}{E(\pi)}.$$

The **riskless interest rate** is given by

$$r = -\frac{1}{T}\ln \mathbb{E}\left(R_{\theta_0}\right).$$

Theorem. *A price deflator exists if and only if there is no arbitrage.*

Proof. If a price deflator exists, then $\Pi(0) = E(\pi\, \Pi(T))$. Since π is positive as a functional on L, if $\Pi(T) > 0$ then $\Pi(0) > 0$ and if $\Pi(T) = 0$ then $\Pi(0) = 0$.

On the other hand, let us suppose that there is no arbitrage. Let us consider the price-payoff vector space $V = \mathbb{R} \times L$. The (cost , pay-off) hyperplane is

$$M = \{(-\theta \cdot q, \theta \cdot P) : \theta \in \mathbb{R}^N\}.$$

The cone $K = \mathbb{R}_+ \times L_+$ contains all securities of non-positive price and non-negative payoff. If there is no arbitrage, then $K \cap M = \{0\}$.

By the separating hyperplane theorem, there exists a functional

$$F : V \to \mathbb{R}$$

such that $F(x) = 0$ for all $x \in M$ and $F(x) > 0$ for all $x \in K \setminus \{0\}$.

The Riesz representation of $F(x)$ is

$$F(v, c) = \alpha\, v + E(\phi \cdot c).$$

In terms of α and ϕ, we have that

$$-\alpha\, \theta \cdot q + \mathbb{E}(\phi \cdot (\theta \cdot P)) = 0$$

for all $\theta \in \mathbb{R}^N$. Hence

$$\pi \equiv \frac{\phi}{\alpha}$$

is a price deflator.

1-Period Summary

There is a measure $\mathbb{P}$, given by the up-ward probability

$$p = \frac{e^{rT}\, S_{\mathrm{now}} - S_{\mathrm{down}}}{S_{\mathrm{up}} - S_{\mathrm{down}}}.$$

An option with pay-off $f_0(S)$ at time T has a price f given by

$$f = \mathbb{E}_{\mathbb{P}}(B_T^{-1}\, f_0) \qquad \text{Martingale condition.}$$

The option can be replicated with a portfolio consisting of

$$a = \frac{f_0(S_{\mathrm{up}}) - f_0(S_{\mathrm{down}})}{S_{\mathrm{up}}) - S_{\mathrm{down}}} \text{units of stock}$$

$$B^{-1}\,(f - a \cdot S)\, \text{units of bonds.}$$

Moreover, the random variable $B_T^{-1} \cdot S$ is also a martingale:

$$B^{-1} \cdot S_{\mathrm{now}} = p \cdot S_{\mathrm{up}} + (1-p) \cdot S_{\mathrm{down}}.$$

Binomial Trees

General facts:

- $\{\mathcal{F}_i\}$, $i = 0, 1, 2 = n$. Filtration (σ-algebras).
- $\{\phi_i\}_{i=0}^n$ is a process if ϕ_i is $\mathcal{F}_i$–measurable for all i.
- The conditional expectation is

  $$\mathbb{E}_{\mathbb{P}}\left(\phi_j | \mathcal{F}_i\right),$$

  is the projection of ϕ_j onto $L^2(\mathcal{F}_i)$.
- Given a measure $\mathbb{P}$, ϕ is a martingale when

  $$\mathbb{E}_{\mathbb{P}}\left(\phi_j | \mathcal{F}_i\right) = \phi_i, \qquad i \leq j.$$

Trivial Fact: Given $\mathbb{P}$, and *any* function X on $L^2(\mathcal{F}_n)$, the process given by $\mathbb{E}_{\mathbb{P}}(X|\mathcal{F}_i)$ is a martingale.

Theorem Binomial Representation Theorem: *Given two processes $\{S_i\}$ and $\{P_i\}$ a binomial tree which are martingales with respect to the same measure $\mathbb{P}$, there exists a process $\{\phi_i\}$ such that*

$$P_i = P_0 + \sum_{k=1}^{i} \phi_k \cdot (S_i - S_{i-1}).$$

Arbitrage Pricing (Multi-Period)

- An option is an $\mathcal{F}_T$–measurable function.
- The stock process generates a measure $\mathbb{P}$ under which it is a martingale.
- The arbitrage-free option price is then given by the new martingale

$$P_i = \mathbb{E}_{\mathbb{P}}\left(B_T^{-1} \cdot X | \mathcal{F}_i\right).$$

- The replication strategy is given by the Binomial Representation Theorem.

$$P_i = P_0 + \sum_{k=1}^{i} \phi_k \cdot (S_i - S_{i-1}).$$

Continuous Pricing

Stock prices follow the Ito Process

$$\frac{dS}{S} = \mu\, dt + \sigma dW_t.$$

Pricing theories will be an extension of the multi-period case.

Agenda

- Elementary review of stochastic process.
- Brownian motion
- Stochastic integrals – Stochastic differential equations.
- Martingale representation
- Martingale pricing and hedging.

Stochastic Processes

A Filtration is $\{\mathcal{F}_t\}$,

- $\mathcal{F}_t$ is a σ–algebra for all t.
- $\mathcal{F}_t \subset \mathcal{F}_s$ when $t < s$.

Take a function $\phi \in L^2(\mathcal{F}_t)$. Its conditional expectation at time $s < t$ defined as

$$\mathbb{E}_s\phi = P_{L^2(\mathcal{F}_s)}\phi,$$

where the right hand side denotes its projection onto $L^2(\mathcal{F}_s)$.

A Stochastic Process ϕ_t is such that

- ϕ_t is $\mathcal{F}_t$–measurable

A Stochastic Process $\phi_t \in L^1(\mathcal{F}_t)$ is a martingale when

$$\mathbb{E}_t(\phi_s) = \phi_t, \quad \text{for } s \geq t.$$

We define the quadratic variation to be adapted process

$$|\phi|_t^2 = \lim_{n\to\infty} \sum_{j=0}^{2^n-1} \left(\phi_{2^{-n}(j+1)t} - \phi_{2^{-n}jt}\right)^2.$$

$\mathcal{M}^2$ denotes the class of martingales with finite quadratic variation.

Predictable Processes

An adapted process ϕ_t is left continuous if

$$\lim_{s \uparrow t} \phi_s = \phi_t, \quad \text{almost surely.}$$

Consider the filtration $\mathcal{F}'$ generated by all left–continuous adapted process.

A process is *predictable* if it is adapted to $\mathcal{F}'$.

Equivalently, it is approximated by processes ϕ_t which are constant on intervals $(t_i, t_{i+1}]$.

Brownian Motion

A process W_t is a $\mathbb{P}$–Brownian motion when

- W_t is continuous in t and $W_0 = 0$.
- $W_t - W_s$ is distributed, under $\mathbb{P}$, as a normal distribution with mean 0 and variance $t - s$.
- For $0 \leq t_0 < \cdots < t_n < \infty$, we have that B_{t_0} and all $B_{t_k} - B_{t_{k-1}}$ are all independent.

If $s < t$,

$$W_t = W_t - W_s + W_s,$$

Hence, since $\mathbb{E}_s(W_t - W_s) = 0$ and $\mathbb{E}_s W_s = W_s$,

$$\mathbb{E}_s W_t = W_s,$$

and Brownian motion is a martingale.

Moreover, $|W_t|^2 = t$, so $W \in \mathcal{M}^2$.

Stochastic Integrals

Given a stochastic process ϕ_t on $(\Omega, \mathcal{F}, \mathcal{T}, \mathbb{P})$, and a function $f(t)$, the stochastic integral

$$\int_0^t f(s)\, d\phi_s,$$

is simply defined as a random variable on $(\Omega, \mathcal{F}, \mathcal{T}, \mathbb{P})$, whose value for any $\omega \in \Omega$ is given by the Stiljes integral

$$\int_0^t f(s)\, d\phi_s(\omega),$$

Appropriate conditions on the paths $\phi_s(\omega)$ are required, such as $\phi_s(\omega)$ to be an increasing function of s. Note that Brownian motion does **not** satisfy this.

For general processes, we first note that, if f is piecewise constant, on intervals $(t_i, t_{i+1}]$, then

$$\int_0^t f(s)\, d\phi_s = \sum_{i=1}^{k} f(t_i) \left(\phi_{t_i} - \phi_{t_{i-1}} \right).$$

Brownian Integrals

Let $W \in \mathcal{M}^2$. Since $|W|_t^2$ is increasing, we can define

$$\|\phi\|_S^2(t) = \int_0^t \phi_s^2 \, d|W|_s^2,$$

the norm

$$\|\phi\|_S^2 = \mathbb{E} \int_0^T \phi_t^2 \, d|W|_t^2,$$

and the corresponding Hilbert Space $L^2|S|$.

If ϕ_t is piecewise constant

$$\begin{aligned}
\text{Variance} \int_0^T \phi \, dW &= \mathbb{E} \left| \sum_{i=1}^k \phi_{t_i} \left(W_{t_i} - W_{t_{i-1}} \right) \right|^2 \\
&\leq \mathbb{E} \sum_{i=1}^k \left| \phi_{t_i} \left(W_{t_i} - W_{t_{i-1}} \right) \right|^2 \\
&\leq \mathbb{E} \sum_{i=1}^k |\phi_{t_i}|^2 \cdot \left(|W|_{t_i}^2 - |W|_{t_{i-1}}^2 \right) \\
&= \|\phi\|_S^2.
\end{aligned}$$

Hence, $\int \phi_t \, dW_t$ can be defined by extension for all $\phi \in L^2|W|$.

In particular, $\int_0^t \phi_t \, dW_t$ is normal with mean 0 and variance

$$\mathbb{E}\sigma^2 = \int_0^t \phi_s^2 \, ds.$$

Hence, we can derive the following formula, which will be of use later:

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}} \exp\left(\int_0^t \phi_t \, dW_t\right) &= \int_{\mathbb{R}} e^{x - x^2/(2\sigma^2)} \frac{dx}{\sqrt{2\pi\sigma^2}}, \\
&= e^{\sigma^2/2}, \\
&= e^{\frac{1}{2}\mathbb{E}\int_0^t \phi_s^2 \, ds}.
\end{aligned}$$

SDE

A process of the type

$$X_t = X_0 + \int_0^t \sigma_s \, dW_s + \int_0^t \mu_s \, ds$$

will be written down as

$$dW_t = \sigma_t \, dW_t + \mu_t \, dt.$$

SDE appear when the terms σ and μ above are made X dependent, as

$$dX_t = \sigma(X_t, t) \, dW_t + \mu(X_t, t) \, dt.$$

They are also called Ito processes.

Ito's Lemma

Ito's Lemma, in this context, reads as follows:

$$f(X_t) = f(X_0) + \int_0^t \left(f'(X_t)\, \mu_t \, dt + \tfrac{1}{2} f''(X_t)\, \sigma_t^2 \right) \, dt$$
$$+ \int_0^t f'(X_t)\, \sigma_t \, dW_t .$$

We can rewrite it as

$$d_t f(X_t) = \left(f'(X_t) + \tfrac{1}{2} \sigma_t^2 \, f''(X_t) \right) \, dt + f'(X)t)\, dW_t .$$

The product rule: if

$$dX_i = \mu_i \, dt + \sigma_i \, dW, \quad i = 1, 2,$$

then

$$d(X_1 \cdot X_2) = X_1 \, dX_2 + X_2 \, dX_1 + \sigma_1 \cdot \sigma_2 \, dt.$$

Feynman-Kac

The Feynman-Kac formula provides a solution to parabolic PDE's in terms of a stochastic integral. It can be viewed as the inverse of the Black–Scholes equation.

Consider the following simple formulation. Let's try to solve

$$\begin{aligned}\partial_t\psi(x,t) &= \tfrac{1}{2}\partial_x^2\psi(x,t) - V(x)\,\psi(x,t)\\ &= -H\psi.\end{aligned}$$

Define

$$X_t = \exp\left(-\int_0^t V(x+W_s)\,ds\right)\, f(x+W_t).$$

Using the product rule and Ito's formula,

$$\begin{aligned}dX_t &= \exp\left(-\int_0^t V\,ds\right)\,\left(-V\,f + f'\cdot dW_t + \tfrac{1}{2}\partial_x^2 f\,dt\right)\\ &= \exp\left(-\int_0^t V\,ds\right)\,\left(-H\,f + f'\cdot dW_t\right).\end{aligned}$$

Hence

$$d_t\mathbb{E}X = -\mathbb{E}\left[\exp\left(-\int_0^t V\,ds\right)\,H\,f\right].$$

Therefore, the linear operator defined by

$$P_t f = -\mathbb{E}\left[\exp\left(-\int_0^t V(x+W_s)\,ds\right)\, f(x+W_t)\right].$$

satisfies the equation

$$\begin{aligned}\partial_t P_t \psi &= -P_t(H\psi)\\ P_0\psi &= \psi.\end{aligned}$$

This yields

$$P_t = e^{-tH}.$$

and

$$\psi(x,t) = P_t\psi(x,0).$$

Girsanov Theorem

Assume a measure $\mathbb{P}$ and a filtration $\mathcal{F}$. If W_t is $\mathbb{P}$–Brownian and γ_t is $\mathcal{F}$–adapted, with the property

$$\mathbb{E}_{\mathbb{P}} e^{\left(\frac{1}{2}\int_0^T \gamma_t^2\, dt\right)} < \infty,$$

define $\mathbb{Q}$ such that

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(-\int_0^T \gamma_t \, dW_t - \frac{1}{2}\int_0^T \gamma_t^2 \, dt \right)$$

Then, $\tilde{W}_t = W_t + \int_0^t \gamma_s \, ds$ is a $\mathbb{Q}$–Brownian motion.

Martingale Representation Theorem

If X_t is a martingale, and ϕ_t is predictable and bounded, then $\int \phi \cdot dX$ is a martingale. The converse:

Theorem : *Assume M_t is a $\mathbb{Q}$–martingale with non–zero volatility σ_t. If N_t is any other $\mathbb{Q}$-martingale, there exists a predictable ϕ_t such that*

$$N_t = N_0 + \int_0^t \phi_s \, dM_s \, ds.$$

Furthermore,

$$\phi_t = \frac{\sigma(N_t)}{\sigma(M_t)}.$$

Corollary : *If $\mathbb{P}$ admits a Brownian motion, all martingales M_t are of the form*

$$dM_t = \sigma_t \, dW_t.$$

Arbitrage Pricing (Multi-Period)

- An option is an $\mathcal{F}_T$–measurable function.
- The stock process generates a measure $\mathbb{P}$ under which it is a martingale.
- The arbitrage-free option price is then given by the new martingale

$$P_t = \mathbb{E}_{\mathbb{P}} \left(B_T^{-1} \cdot X | \mathcal{F}_t \right).$$

- The replication strategy is given by the Binomial Representation Theorem.

$$P_t = P_0 + \int_0^t \phi_s \, dS.$$

or

$$\phi_t = \frac{\partial P_t}{\partial S}.$$

- The Black-Scholes equation follows from the Feymnan-Kac formula.

Continuous Pricing

Stock price process

$$\frac{dS}{S} = r\,dt + \nu\,dW.$$

European call option with payoff $f_0(S)$

$$f(S,t) \;=\; e^{-r(T-t)} \int_{-\infty}^{\infty} f_0\left(S_0 e^{(r-\frac{\nu^2}{2})(T-t)\,+\,x} \right) P_\nu(x, T-t)\,dx$$

where

$$P_\nu(x,t) \;=\; \frac{1}{\sqrt{2\pi\nu^2 t}}\;\exp\left(-\frac{x^2}{2\nu^2 t} \right).$$

Multifactor options involve higher dimensional integrals. They are studied using MonteCarlo methods, or Low–discrepancy sequences.

Also

$$\begin{cases} \dfrac{\partial f}{\partial t} \;=\; -\dfrac{1}{2}\nu^2 S^2 \dfrac{\partial^2 f}{\partial S^2} \;-\; r\,S\,\dfrac{\partial f}{\partial S} \;+\; r\,f, \\ f(S,T) = f_0(S). \end{cases}$$

American options give rise to free boundaries.

For Interest rate options, the underlying $S \in \mathbb{R}$ is replaced by a yield curve, $r(t,T)$.

The Real World.

- Stock prices are not log-normal.
- Transaction costs make continuous trading infinitely expensive.
- No one knows future volatility.
- Markets are not liquid, complete or efficient.
- The counterparty you are dealing with may not exist tomorrow.

Greeks

Provide an intuitive basis for understanding price changes, and risk.

Delta. Change in price as a function of the underlying.

$$\delta = \frac{\partial \Pi}{\partial S}.$$

Gamma. Change of delta as a function of the underlying.

$$\Gamma = \frac{\partial^2 \Pi}{\partial S^2}.$$

Vega. Change of price as a function of volatility

$$\mathcal{V} = \frac{\partial \Pi}{\partial \sigma}.$$

Theta. Essentially:

$$\theta = \frac{\partial \Pi}{\partial t}.$$

Numerical Methods

Consider stocks with price processes given by

$$d(\log S_j) \;=\; \mu_j\, dt \;+\; \sigma_j\, dW_j,$$

where the Brownian motions dW_j have correlation coefficients $V = (\rho_{i,j})$.

If the time to expiration is T, the interest rate is r, and the pay–off of an option is $f_0(S_1, \ldots, S_n)$, its price is

$$\frac{e^{-rT}}{\sqrt{(2\pi)^n\, \det V}} \int_{\mathbb{R}^n} f_0\left(e^{S_1}, \ldots, e^{S_n}\right)\, e^{-(\mathbf{S}-\mu)^{\dagger}\, V^{-1}\, (\mathbf{S}-\mu)}\, d\mathbf{S}.$$

With a change of variables, the problem then reduces to computing integrals of the type

$$\int_{[0\,,\,1]^n} f(x_1, \ldots, x_n)\, dx_1\, \cdots\, dx_n,$$

for appropriate integrands f.

Multidimensional integrals are not easy to compute. There are three basic methods:

1. Grid Methods.

2. Monte–Carlo Methods.

3. Low discrepancy methods.

Grid points. We put N evenly spaced points x_i in the unit cube, and approximate

$$\int_Q f\,dx = \frac{1}{N}\sum f(x_i).$$

The error in this approximation is like $N^{-1/n}$.

It can be improved to the trapezoidal rule

$$\int_Q f\,dx = \frac{1}{N}\sum \chi(x_i)\,f(x_i),$$

where χ is the *characteristic* function of the cube,

$$\chi(x) = \begin{cases} 1 & \text{if } x \in \operatorname{Int} Q, \\ 2^{-k} & \text{if } x \text{ belongs to } k \text{ faces of } Q, \end{cases}$$

which gives an error comparable to $N^{-2/n}$.

Pros–Cons

- As a function of n, it require an exponentially large number of points. (Compare with *Numerical Recipes*).
- It can easily be turned into an adaptative grid process. (Calderón–Zygmund. Compare with *Numerical Recipes*).

Monte–Carlo.

Monte–Carlo methods appeared officially in 1949, but they had been used by the U. S. Defense Department in secret projects for several years before that. The name was a code name used by Von Neumann and Ulam at Los Alamos in projects related to The Bomb (simulation of random neutron diffusion in fissionable material).

Applied to integration,

$$\int_Q f(x)\,dx = \frac{1}{N}\sum_{i=1}^{N} f(x_i) \;\pm\; \mathcal{O}\Big(N^{-1/2}\Big)\,.$$

if the x_i are uniformly distributed.

Pros–Cons

- Reasonable speed is the same in all dimensions.
- Not easily turned into an adaptative scheme.
- Ignores smoothness of the integrand.
- It's hard to generate random numbers, but it is otherwise easy to implement.
- The birthdate effect (clustering).

Higher Dimensions

Generating random numbers in higher dimensions is a difficult task. For the case of a multivariate normal, with given variance/covariance matrix $\mathbb{V}$, one proceeds as follows:

Consider the Cholesky decomposition of $\mathbb{V}$,

$$\mathbb{V} = \mathbb{H}^\dagger \cdot \mathbb{H}.$$

Let $y = (y_1, \ldots, y_n)$ be independent normally-(0,1) distributed random variables.

Then

$$x = y \cdot \mathbb{H}$$

produces random vectors which are normally distributed with covariance given by $\mathbb{V}$.

Indeed,

$$\begin{aligned} \text{Cov } x &= \mathbb{E}\left(x^\dagger \cdot x\right) \\ &= \mathbb{E}\left(\mathbb{H}^\dagger \, y^\dagger \, y \, \mathbb{H}\right) \\ &= \mathbb{V}. \end{aligned}$$

Principal Components

The biggest limitation in the algorithm for generation of random numbers in higher dimensions is the inability to produce clean univariate normally distributed numbers, when the dimension is very high.

In practice, covariance matrices have directions that will happen with high probability, and others that are very unlikely.

Set

$$\mathbb{V} = \mathbb{P}\mathbb{D}\mathbb{P}^{\dagger},$$

with $\mathbb{D}$ diagonal and $\mathbb{P}$ orthogonal.

$$D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_{n_1} \end{pmatrix}$$

Assume $\lambda_1 \geq \lambda_2 > \cdots$. One may choose to trim the covariance matrix, by selecting only the first few eigenvalues of $\mathbb{D}$. These are the principal components.

The significance of a selection $\lambda_1, \ldots, \lambda_k$ is given by

$$\frac{\lambda_1 + \cdots + \lambda_k}{\text{Trace } \mathbb{V}}.$$

Sub–random methods. Uses other limit theorems (i.e., ergodic theorems and number theory). Consider an example in dimension 1:

If γ is irrational, and $\{t\}$ denotes the fractional part of t,

$$\int_0^1 f(x)\,dx = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} f\left(\{n\gamma\}\right).$$

Moreover, we can estimate the remainder:

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^{N} f\left(\{n\gamma\}\right) &= \frac{1}{N} \sum_{i=1}^{N} \sum_{k=-\infty}^{\infty} \hat{f}(k)\, e^{2\pi i n k \gamma} \\ &= \int_0^1 f(x)\,dx \\ &\qquad + \sum_{k\neq 0} \hat{f}(k) \frac{1}{N} \sum_{n=1}^{N} e^{2\pi i n k \gamma}. \end{aligned}$$

The speed of convergence is therefore closely related to the diophantic properties of γ, and the smoothness of f.

Theorem : *If f is of bounded variation, and γ is an algebraic number, then*

$$\left| \int_0^1 f(x)\,dx - \frac{1}{N} \sum_{i=1}^{N} f\left(\{n\gamma\}\right) \right| \leq C\, N^{-1} \log N.$$

The constant C depends on f.

Higher Dimensions. The speed of convergence is based on the concept of *discrepancy* of a finite set of points in the unit cube in dimension d, $P \in [0,1]^d$:

Consider sets of the form

$$B = \prod_{i=1}^{d} [0, u_i).$$

The discrepancy is defined as

$$D(P) = \sup_B \left| |B| - \frac{|P \cap B|}{|P|} \right|.$$

Theorem :

$$\left| \frac{1}{N} \sum_{i=1}^{N} f(x_n) \; - \; \int f \right| \leq V(f) \cdot D(x_1, \ldots, x_N).$$

A sequence $\{x_n\}$ is usually referred to as *low discrepancy* if

$$D(x_1, \ldots, x_N) = \mathcal{O}\big(N^{-1+\varepsilon}\big), \qquad \text{for all } \varepsilon > 0.$$

Let $\gamma = (\gamma_1, \ldots, \gamma_d)$ such that $1, \gamma_1, \ldots, \gamma_d$ are linearly independent over the rationals, and the γ_i are algebraic, Then, $\{k \cdot \gamma\}$ is a low–discrepancy sequence.

HEDGING

Problem: to generate a buy/sell strategy that replicates pay-offs.

Used to remove risk from option writing.

For frictionless ideal markets:

$$\frac{\partial f}{\partial S}(S(t), t)$$

units of the stock replicates the option price at all times.

In real situations:

- Transaction costs: viscosity solutions.
- Incomplete markets: incomplete hedges.
- Discrete time. Dynamic programming.

INVERSE PROBLEMS (implied parameters)

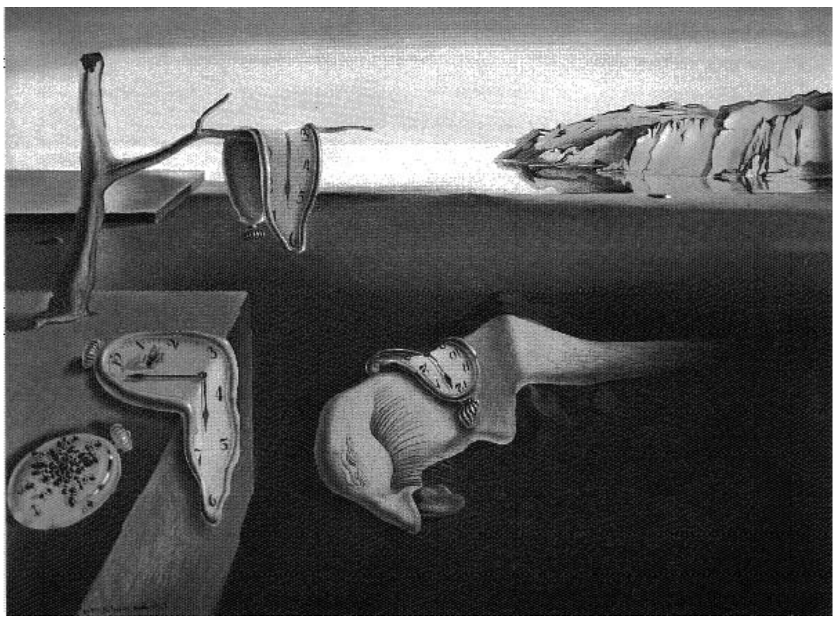
If certain options are liquid enough,

$$f(S,t) \;=\; e^{-r(T-t)} \int_{-\infty}^{\infty} f_0 \left(S e^{(r-\frac{\nu^2}{2})(T-t)\,+\,x} \right) P_\nu(x, T-t)\, dx$$

is known. This implies a certain value for the volatility.

- Interest rates: term structure
- American options: inverse scattering.
- Time dependent volatilities. Volatility smiles.
- Implied volatility surfaces.
- Calibration risk.

Interest Rate Theory.



Time is money. The implications of this fact in pricing theories are tremendous. There is no established way to analyze this. We have to content ourselves with a botanical theory of interest rate models with different properties.

Bonds, Yields

Consider a bond that, with a payment of $P(t,T)$ at time t, pays \$1 at time T (and has no other intermediate payments). If the interest rate r is assumed constant, then we would have

$$P(t,T) = e^{-r(T-t)}.$$

Hence,

$$r = -\frac{\log P(t,T)}{(T-t)}.$$

This is useful since it is P, not r that is observed (not quite: see next section). When r is not constant, we simply define the *yield rate*

$$r(t,T) = -\frac{\log P(t,T)}{(T-t)}.$$

Since r determines P, r determines the entire term structure.

As a function of T, r is smooth. As a function of t, it is a random.

Forward Rates

Short rate: the cost of instantaneous borrowing:

$$r_t = r(t,t) = -\left.\frac{\partial}{\partial T}\log P(t,T)\right|_{T=t}.$$

- Similarity with stock prices
- loss of information.

Forward Rate: $f(t,T)$: "our prediction at t of r_T."

Consider the following futures contract:

- Agreement date: now (time t).
- Product to deliver: a zero-coupon bond B issued at T_1, paying \$1 at T_2.
- Delivery date: T_1.
- Price: $P(t,T_1,T_2)$ (Unknown).
- Payment date: T_1.

It has the only two cash flows:

- At time T_1: $P(t, T_1, T_2)$ (Unknown).
- At time T_2: \$1.

Consider a portfolio Π of 1 bond unit worth $P(t, T_2)$ each, and $-x$ bond units worth $P(t, T_1)$ each, with

$$x = \frac{P(t, T_2)}{P(t, T_1)}.$$

Because it costs nothing now, it has only two cash flows:

- At time T_1, an cash out-flow equal to $-x$ (or in-flow of x).
- At time T_2, a cash in-flow of \$1.

Hence,

$$P(t, T_1, T_2) = \frac{P(t, T_2)}{P(t, T_1)}.$$

The corresponding forward yield is, by definition

$$\begin{aligned} r(t, T_1, T_2) &= -\frac{\log P(t, T_1, T_2)}{T_2 - T_1} \\ &= -\frac{\log P(t, T_2) - \log P(t, T_1)}{T_2 - T_1}. \end{aligned}$$

The *forward rate* is defined to be

$$
\begin{aligned}
f(t,T) &= r(t,T,T) \\
&= -\frac{\partial}{\partial T}\log P(t,T).
\end{aligned}
$$

The term structure is reconstructed from f as follows:

$$
P(t,T) = \exp\left(-\int_t^T f(t,u)\,du\right).
$$

r **volatility matrix**, if these are taken at discrete times.

Bootstrapping

In practice, bonds pay coupons. Hence, the calculation of the yield needs to be modified.

Example (J. Hull) Consider the following set of bonds, with $100 principal and associated prices, with coupons paid every six months as described below:

Time to Maturity	annual coupon	price
3 mo.	$0	$97.5
6 mo.	$0	$94.9
1 yr.	$0	$90.0
1.5 yr.	$8	$96.0
2.0 yr.	$12	$101.6
2.75 yr.	$10	$99.8

Solving the obvious set of (non–linear equations), we can arrive at the yield curve with values given by

Term	annualized rate
3 mo.	10.12%
6 mo.	10.47%
1 yr.	10.54%
1.5 yr.	10.68%
2 yr.	10.81%
2.75 yr.	10.87%

Black 76

Consider a caplet: pays at t_{i+1} the difference between the excess between the yield rate $r(t_i, t_{i+1})$, and a strike K.

Absence of arbitrage implies that, for valuation purposes, we should treat $r(t_i, t_{i+1})$ as $f(t_0, t_i, t_{i+1})$.

If we assume this is log-normally distributed, we are just looking at a call option with $f(t, t_i, t_{i+1})$ as the underlying risk factor, whose price is given by the usual Black–Scholes formula

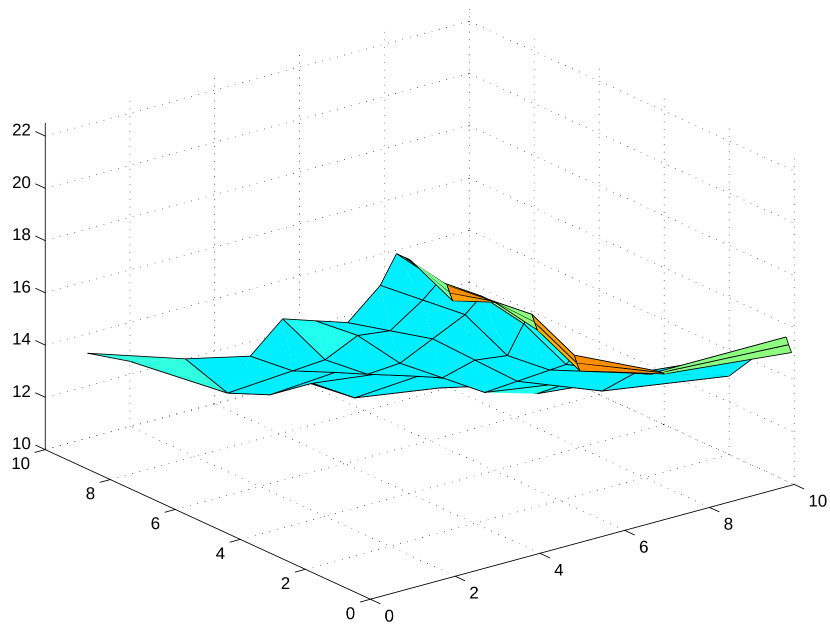
$$\text{Caplet} = (F(t_0, t_i, t_{i+1}) \cdot N(d_+) - K\, N(d_-)) \cdot P(t_0, t_1),$$

with

$$d_\pm = \frac{\ln(F/K) \pm \frac{1}{2}\sigma^2(t_i - t_0)}{\sigma\sqrt{t_i - t_0}}.$$

In practice, caps prices are fixed by the market, and the formula above is used in an attempt to extract the implied volatility from the market.

Since caplets involve two future dates, implied vols are a two parameter function, the **volatility surface** or **volatility matrix**, when these are a discrete set of points.



Implied Vol Surface

Decaf Heath–Jarrow–Morton

$$d_t f(t,T) = \alpha(t,T)\,dt + \sigma\,dW_t$$

This SDE has a trivial explicit solution

$$f(t,T) = f(0,T) + \int_0^t \alpha(s,T)\,ds + \sigma\,W_t.$$

Model properties:

- At a fixed time t, $f(t,T)$ is smooth in T.
- $f(t,T) - f(t,S)$ is deterministic.
- The only source of randomness comes from t.
- f (and r_t) can become negative.
- σ is independent of the term and time.

The Money Market Account

We invest \$1 permanently in the short rate:

$$dB_t = r_t \, B_t \, dt, \qquad B_0 = 1.$$

Since

$$r(t) = f(0,t) + \int_0^t \alpha(s,t)\, ds + \sigma \, W_t.$$

we have

$$\begin{aligned} B_t &= \exp\left(\int_0^t r_s \, ds \right) \\ &= \exp\left(\sigma \int_0^t W_s \, ds + \int_0^t f(0,u)\, du + \int_0^t \int_s^t \alpha(s,u)\, du \, ds \right) \end{aligned}$$

Replicating Strategies

We have an option X on a bond. The option expires at S, while the bond matures at a later time T.

We will use the money market account for the discounting factor.

Hence, The discounted bond is given by

$$Z_t = B_t^{-1}\, P(t,T).$$

A replicating strategy will be given by:

- Finding a measure $\mathbb{P}$ so Z_t is a martingale.

In this way,

$$E_t = \mathbb{E}_{\mathbb{P}} \left(B_S^{-1} X | \mathcal{F}_t \right).$$

is a $\mathbb{P}$–martingale. By the martingale representation theorem, there is a process ϕ_t such that

$$dE_t = \phi_t\, dZ_t,$$

which gives us a replicating strategy.

The Measure

We had

$$B_t = \exp\left(\sigma \int_0^t W_s\, ds + \int_0^t f(0,u)\, du + \int_0^t \int_s^t \alpha(s,u)\, du\, ds\right)$$

$$P(t,T) = \exp\left(-\int_t^T f(0,u)\, du - \int_t^T \int_0^t \alpha(s,u)\, ds\, du - (T-t)\,\sigma\, W_t\right)$$

Hence, Z_t is given by

$$\exp\left(-\int_0^T f(0,u)\, du - \int_0^t \int_s^T \alpha(s,u)\, du\, ds - (T-t)\,\sigma\, W_t - \sigma \int_0^t W_s\, ds\right)$$

Using Ito's Lemma, we get

$$\frac{dZ_t}{Z_t} = -\int_t^T \alpha(t,u)\, du - \sigma(T-t)\, dW_t + \tfrac{1}{2}\,\sigma^2 (T-t)^2\, dt.$$

In order to use Girsanov's theorem, set

$$dW_t = d\tilde{W}_t + \gamma(t,T),$$

$$\gamma(t,T) = -\tfrac{1}{2}\,\sigma\,(T-t) + \frac{1}{\sigma(T-t)} \int_t^T \alpha(t,u)\, du.$$

This yields

$$dZ_t = \sigma\, Z_t\,(T-t) d\tilde{W}_t.$$

and Z_t is a martingale.

Absence of Arbitrage

If different durations T give different drifts $\gamma(t,T)$. we would have arbitrage. Therefore, We must have

$$\frac{\partial \gamma}{\partial T} = 0.$$

Equivalently, we have the following condition on the drift:

$$\alpha(t,T) = \sigma^2(T-t) + \sigma\gamma(t,T).$$

Heath–Jarrow–Morton

For $0 \le t \le T$,

$$f(t,T) = f(0,T) + \int_0^t \sigma(s,T)\, dW_s + \int_0^t \alpha(s,T)\, ds,$$

Equivalently,

$$d_t f(t,T) = \sigma(t,T)\, dW_t + \alpha(t,T)\, dt.$$

Features:

- σ and α are t-adapted processes.
- $f(0,T)$ is deterministic.

Short Rate Models

$$dr(t) = \nu(r_t, t)\,dt + \rho(r_t, t)\,dW_t.$$

Bond Prices:

$$\begin{aligned} -\log P(t,T) &= \int_t^T f(t,u)\,du \\ &= g(r_t, t, T). \\ &= \log \mathbb{E}_{\mathbb{Q}} \left(e^{-\int_t^T \, r(s)\,ds} \mid r_r = x \right). \end{aligned}$$

The Hull–White Model

Given a, σ (assumed constant), and a time-dependent reversion level $\theta(t)$, we write the SDE for the spot rate:

$$dr_t = (\theta(t) - a\, r_t)\; dt + \sigma\, dW_t.$$

Assuming bond prices are given by $P(r, t, T)$, Ito's lemma tells us

$$\begin{aligned} d_t P &= \left[\partial_r P \cdot (\theta(t) - a\, r_t) + \partial_t P + \tfrac{1}{2}\,\sigma^2\,\partial_r^2 P\right]\; dt + \sigma\,\partial_r P\, dW_t \\ &= \mu(t, T)\, dt + \nu(t, T)\, dW_t. \end{aligned}$$

The portfolio

$$\Pi(t) = P(t, T) + \alpha\, P(t, T_2),$$

where

$$\begin{aligned} \alpha &= -\frac{\partial_r P(t, T)}{\partial_r P(t, T_2)} \\ &= -\frac{\nu(t, T)}{\nu(t, T_2)} \end{aligned}$$

is deterministic. Hence,

$$d_t P(t, T) - \alpha(t, T)\, d_t P(t, T_2) = r(r)\, \Pi(t)\, dt$$

which implies

$$\lambda(t,r) = \frac{\mu(t,T) - r(t)P(r,t,T)}{\nu(t,T)}.$$

must be independent of T. Subbing μ above, we obtain

$$\begin{aligned}\partial_r P \cdot (\theta(t) - a\, r_t) + \partial_t + \tfrac{1}{2}\partial_r^2 \sigma^2 &= rP + \lambda\, \nu(t,T) \\ &= rP + \lambda \sigma \partial_r P.\end{aligned}$$

or

$$\partial_t P + (\phi(t) - a\, r)\ \partial_r P + \tfrac{1}{2}\sigma^2\, \partial_r^2 P = r\, P,$$

with

$$\phi(t) = \theta(t) - \lambda(t)\, \sigma.$$

The quantity λ is the market price of risk.

We now make the following assumption:

$$P(r,t,T) = A(t,T)\,\exp\left(-B(t,T)\,r\right).$$

This implies

$$\begin{cases} \partial_t B - a(t)\,B + 1 = 0 \\ B(T,T) = 0 \end{cases} \qquad \begin{cases} \partial_t A - \phi(t)\,A\,B + \frac{1}{2}\sigma^2\,A\,B^2 = 0 \\ A(T,T) = 0. \end{cases}$$

where

$$\begin{aligned} B(t,T) &= a^{-1}\left(1 - e^{-a(T-t)}\right), \\ \log A(t,T) &= \log\frac{P(0,T)}{P(0,t)} + B(t,T)\cdot f(0,t) \\ &\qquad - \frac{\sigma^2\left(e^{-aT} - e^{-at}\right)^2\left(e^{2at} - 1\right)}{4\,a^3}. \end{aligned}$$

Other short rate models

Ho-Lee

$$dr = \theta_t \, dt + \sigma \, dW_t$$

Vasicek

$$dr = (\theta - \alpha \, r) \, dt + \sigma \, dW_t$$

Cox-Ingersoll-Ross

$$dr = (\theta_t - \alpha_t \, r) \, dt + \sigma_t \sqrt{r} \, dW_t$$

Black-Derman-Toy

$$d(\log r) = \left(\theta'(t) + \frac{\partial \log \sigma(t)}{\partial t} \, \log r \right) \, dt + \sigma(t) \, dz.$$

Black-Karasinski

$$d(\log r) = (\theta_t - \alpha_t \, \log r) \, dt + \sigma_t \, dW_t$$

Calibration

All these models assume a certain set of parameters, such as the volatility, mean reversion parameters, etc.

There is no well established way of doing this.

A popular method is to observe market prices for standard instruments, and work out the parameters that best fit market data. This process is called **calibration**.

While probably inadequate for risk management purposes, it is probably the right attitude for hedging (and hence pricing) purposes, if one uses observed market prices for the hedging instruments in the calibration exercise.

Portfolio theory.

Prices have been established in the comfort of the indifference provided by arbitrage-free arguments.

The investor needs to select the investment that will maximize returns with a given level of risk, or will minimize risk with a target rate of return.

Markowitz Mean Variance Model

Assume n instruments available for trade, with stochastic returns given by S_i.

Return = Mean: $R_i = \mathbb{E}(S_i)$.

Risk = Variance: $\sigma_i^2 = \mathbb{E}(S_i^2) - R_i^2$.

Consider also the variance/covariance matrix

$$\mathbb{V} = \{\sigma_{i,j}\}, \qquad \sigma_{i,j}^2 = \mathbb{E}(S_i \cdot S_j) - R_i \cdot R_j.$$

An investment choice is given by a weight vector $w = (w_1, \ldots, w_n)$. Such investment will have a return equal to

$$\sum_{i=1}^{N} w_i \cdot R_i,$$

and a variance given by

$$w^t \cdot \mathbb{V} \cdot w.$$

(Static) Investment decisions

Usual investment choices:

- For a given level of return α , minimize risk:
- For a given level of risk α , maximize return:

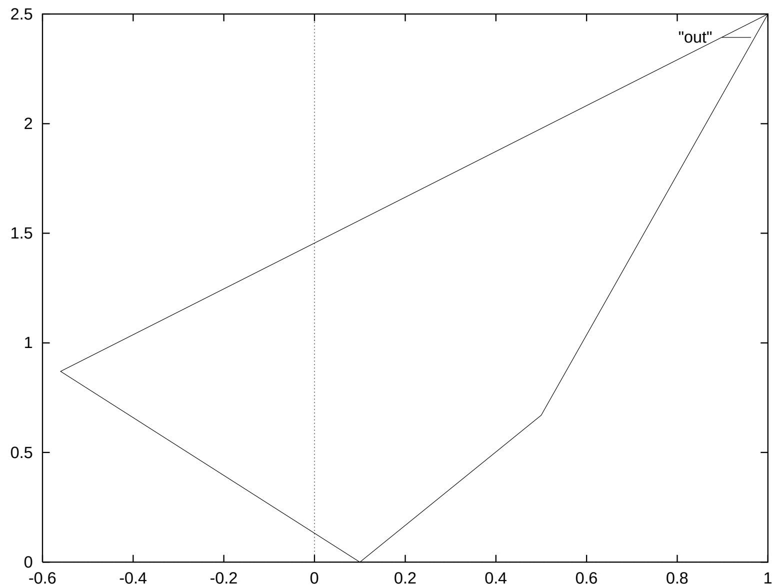
This can be generalized to quantifying the relative importance of risk into a parameter λ in order that we now seek

$$\max_w \left(w \cdot R - \lambda \, w^t \, \mathbb{V} \, w \right),$$

subject to restrictions, which may include a cap the amount of money invested $(\sum w_i)$, and taking only long positions $(w_i \geq 0)$.

All these are quadratic programming problems, and due to the interaction between risk and return, they always involve two-dimensional choices.

They can be summarized in a two dimensional risk/return graph.



All possible investments will give rise to a convex subset of $\mathbb{R}^2_+$. Its boundary is the Efficient Frontier.

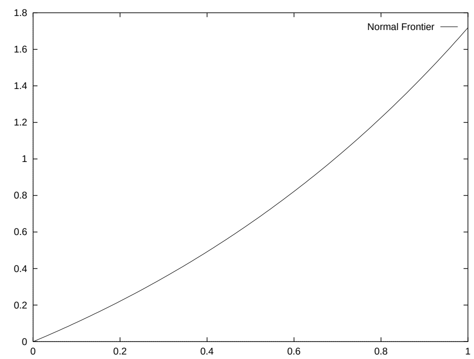
Theorem : *The Efficient Frontier is convex.*

It will be convenient to subtract the risk–free rate from returns. In this way, the efficient frontier will go through the origin if the risk-free rate is available.

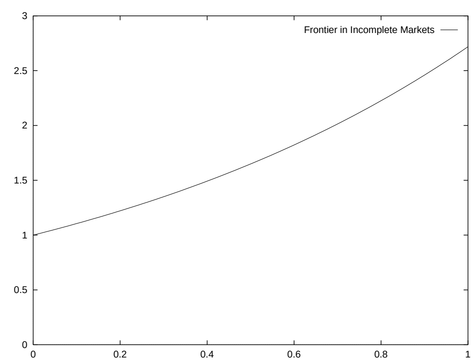
The Efficient Frontier

The efficient frontier carries a lot of graphical information. As a sample:

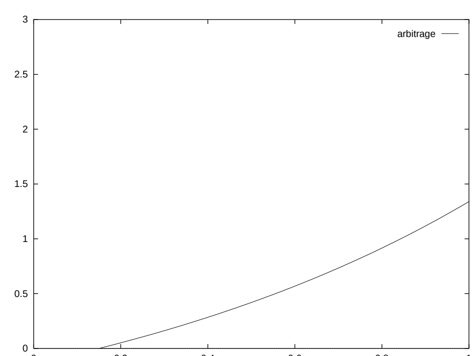
Normal Markets:



Incomplete Markets:



Arbitrage Opportunities:



Sharpe's ratio

It provides a one-dimensional approach to performance evaluation. It is defined as

$$\text{Sharpe's ratio} = \frac{(\text{Return}) - (\text{Riskfree rate})}{\text{risk}}.$$

For pay-off distributions that are not symmetric, downside-risk

$$\sqrt{\mathbb{E}(S_i - R_i)_-^2},$$

can be used to replace the standard deviation.

Regret/Reward Model

(R. Dembo, D. Rosen, D. Saunders.)

Regret. Replaces the concept of risk.

Let D be the pay-out matrix, and τ a benchmark portfolio. Set

$$Dx - \tau = y^+ - y^-$$

denote the positive and negative tracking errors with respect to the benchmark.

Regret is the expected underperformance:

$$R_\tau^- = p^T \cdot y^- .$$

Reward. It is the expected excess profit:

The expected profit is

$$p^T(Dx) - q^T x,$$

and the expected profit earned by the benchmark is

$$p^T \tau - c.$$

Reward is

$$
\begin{aligned}
R_\tau^+ &= p^T(Dx) - q^T x - p^T \tau + c \\
&= p^T(y^+ - y^-) - q^T x + c.
\end{aligned}
$$

We restrict our attention to portfolios with cost less than or equal to that of the benchmark (i.e., $q^T x \le c$).

The efficient frontier can then be defined in a similar manner, and all previous considerations hold.

Financial Risk

"Even in this age of high-tech computing, the basic architecture of risk management remains primitive — it is as if all that fancy technology is stored in the intellectual equivalent of a wooden shack."

— *Seeing Tomorrow*, R. Dembo and A. Freeman.

History

Orange County. Bob Citron (County Treasurer), built a leveraged portfolio of short term loans with long term notes. This proved to be a profitable investment, since short term rates were low and long term rates were high.

When interest rates began to rise in 1994, losses occurred; a snow ball effect happened when investors found out about the losses, and attempted to withdraw the money.

Such portfolios are usually perfectly hedged against parallel shifts in the yield rate, but show big exposure to tilts in the yield curve.

The snow ball effect is a usual device through which liquidity risk shows up.

Barings Bank.

Nick Leeson had derivative positions on the Nikkei, that bet on a stable market. The decline of 15% of the Nikkei, which followed the earthquake of Kobe in 1995, produced large losses. The situation lead to a total loss of $1.3 Billion, after the positions were further increased after the initial losses. Some of the positions were unauthorized. although his superiors had approved a cash infusion of $1billion to cover margin calls.

The price of shares dropped to $0. Baring's Bank was worth about $1 billion in terms of market capitalization. Bond holders received 5c. for each $1. ING offered to purchase Baring's for 1 British Pound.

The Bassel convention (1991).

Risk Management

Problem: to determine potential losses.

- Calibration risk
- Market risk
- Credit risk
- Model risk

J.P. Morgan introduced RiskMetrics about 3 years ago, and CreditMetrics last summer. They are industrial standards now.

Banks must meet Market-Risk calculation criteria following from the Bassel convention in 1991.

Estimating Volatility

Volatilities are the quantification of the risk of the market.

Portfolios will have a volatility adapted to them, but estimating volatility is at the heart of analyzing the risk of any portfolio.

Three ways of estimating volatility:

Implied volatility methods. Used for hedging and trading. Possibly inadequate measure of future market movements.

Averaging Methods. Provides a simple estimation based on historical market movements.

Stochastic volatility models. Accounts for the possibility of volatility jumps. Useful for the study of longer-time horizons.

Averaging Methods

Consider a historical series of risk factors (stock prices, yield rates, etc.) given by $\{s_i\}$, for $i = 0, \ldots, n$. We adopt the convention that s_0 is today's observation, and the values move backward in time with increasing i.

If they are log-normally distributed, we can define

$$\mu = \frac{\sum_{i=1}^n \lambda^i \log\left(\frac{s_{i-1}}{s_i}\right)}{\sum_{i=1}^n \lambda^i},$$

for a parameter $\lambda < 1$ (0.95, for example), which gives more weight to recent observations. Similarly, we set

$$\sigma^2 = \frac{\sum_{i=1}^n \lambda^i \left(\log\left(\frac{s_{i-1}}{s_i}\right) - \mu\right)^2}{\sum_{i=1}^n \lambda^i},$$

For multifactor series $\{s_i^{(k)}\}$, covariances are found in a similar way:

$$\sigma_{k,l}^2 = \frac{\sum_{i=1}^n \lambda^i \left(\log\left(\frac{s_{i-1}^{(k)}}{s_i^{(k)}}\right) - \mu_k\right)\left(\log\left(\frac{s_{i-1}^{(k)}}{s_i^{(k)}}\right) - \mu_k\right)}{\sum_{i=1}^n \lambda^i},$$

GARCH models

It is a stochastic volatility model. It is popular because it accounts for volatility jumps. The general GARCH process is given by the coupled system of equations

$$\begin{cases} r_i = \xi \cdot \sigma_i \\ \sigma_i = \omega + \alpha \cdot \sigma_{i-1} + \beta \cdot r_{i-1}. \end{cases}$$

where the label i now goes forward in time.

The model is used as follows: given a known series of log-returns (normalized to mean 0),

$$r_i = \log(s_i) - \log(s_{i-1}), \quad i = 1, \ldots, n,$$

and a choice of parameters, the likelihood of that given choice is given by the expression

$$\ln(\sigma_t) + \sum_{i=1}^{n} \frac{r_t^2}{\sigma_t^2}$$

We calibrate the GARCH parameters ω, α and β to this series of returns by maximizing this likelihood.

The parameters we obtain, together with today's return, will tell us the volatility today; they can also make predictions about future volatilities.

GARCH list

ARCH(q).

$$\sigma_i = \omega + \sum_{k=1}^{q} \alpha_k \cdot r_{i-k}.$$

GARCH(p,q).

$$\sigma_i = \omega + \sum_{k=1}^{q} \alpha_k \cdot r_{i-k} + \sum_{\ell=1}^{p} \beta_\ell \cdot r_{i-\ell}.$$

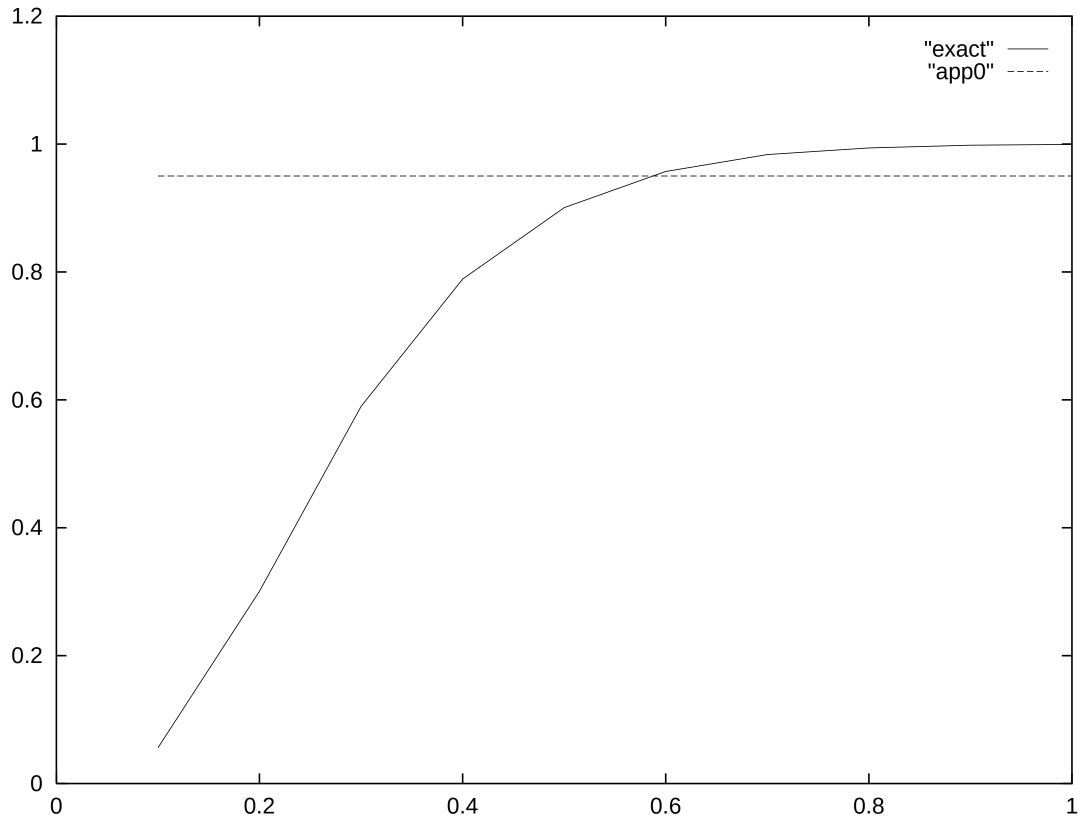
AGARCH.

$$\sigma_i = \omega + \alpha \cdot (r_{i-1} - \xi) + \beta \cdot r_{i-1}.$$

Value at Risk

VaR of a portfolio Π is what the portfolio can loose overnight with a certain probability:

$$\textbf{Prob } \{\Pi(0) - \Pi(t) > \textbf{VaR} \} = 5\%.$$



95%

Methodologies.

Three methods.

- Historical Methods.
- Monte Carlo.
- Analytic. RiskMetrics.

Caution:

- Lottery ticket
- Identical short and long positions

Historical Methods

Profit and Loss statistics are computed using the value of the actual portfolio under a set of historical scenarios.

Optimization problems:

Portfolio Compression Replace a portfolio with a smaller one that preserves certain characteristics (price, VaR, sensitivities, etc.)

Portfolio Replication Replace exotic or "unwanted" instruments by a (probably larger) portfolio of standardized instruments, in a way that certain properties are preserved.

Basically, given a target portfolio Π, and a base basket of instruments π_i, for $i = 1, \ldots, n$, we try to choose the position numbers θ_i that minimize the expression

$$\left\| \Pi - \sum_{i=1}^{n} \theta_i \, \pi_i \right\| ,$$

where $\|\cdot\|$ can denote a variety of things, such as absolute value of price, VaR, downside risk, etc.

Monte Carlo

Historical scenarios are replaced by "random" ones. Let the variance/covariance matrix given by $\mathbb{V}$:

Consider the portfolio $\Pi(\mathbf{S})$, as a function of n risk factors $\mathbf{S}$ distributed normally according to $\mathbb{V}$, with present value given by $\mathbf{S}_0$. Let $\mathbb{V}^{1/2}$ be the Cholesky decomposition of $\mathbb{V}$. Let $\xi_1, \ldots, \xi_N$ i.i.d. normal random vectors in $\mathbb{R}^n$, with mean 0 and var/covar equal to the identity.

Future scenarios are then given by

$$\xi_i \cdot \mathbb{V}^{1/2} + \mathbf{S}_0, \qquad i = 1, \ldots, N.$$

The P&L of the portfolio is given by

$$\Pi_i \;=\; \Pi\left(\xi_i \cdot \mathbb{V}^{1/2} + \mathbf{S}_0\right) - \Pi(\mathbf{S}_0), \qquad i = 1, \ldots N.$$

VaR can then be computed in two ways:

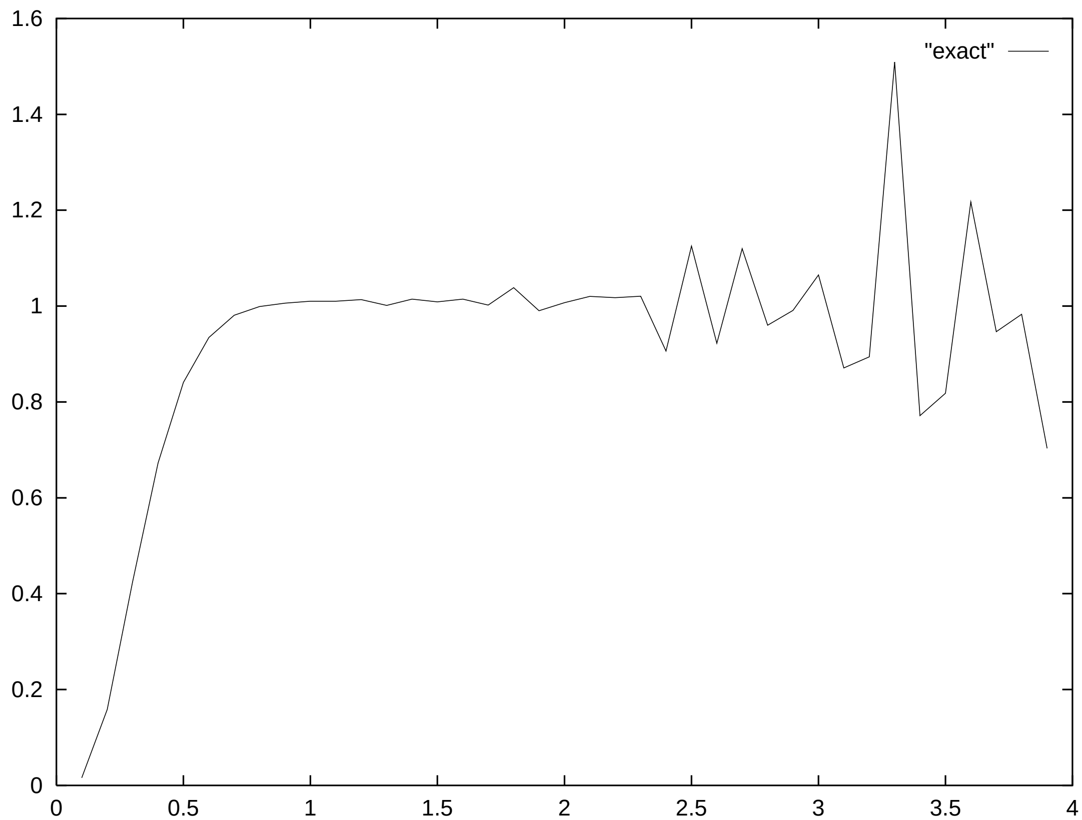
Parametric VaR: fit the distribution of Π_i to a normal distribution and compute its standard deviation σ. Then

$$\mathbf{VaR} \;= 1.65 \cdot \sigma.$$

Non–parametric VaR: order the values of Π_i in an increasing way; VaR is the 95% percentile.

Watch out!

- Cancellation problems
- Dimensionality problems



Actual calculation

Analytic Methods

Delta Normal VaR (RiskMetrics)

Approximate

$$\Pi(t) - \Pi(0) \approx \left. \frac{\partial \Pi}{\partial t} \right|_{t=0} + \sum_{i=1}^{n} \delta_i \cdot \left[S_i(t) - S_i(0) \right],$$

where

$$\delta_i = \left. \frac{\partial \Pi}{\partial S_i} \right|_{t=0}.$$

VaR is then given by

$$\text{Prob} \left\{ \sum_{i=1}^{n} \delta_i \cdot [S_i(t) - S_i(0)] < -\mathbf{VaR} \right\} = 0.05.$$

If $S_i(t) - S_i(0)$ is normally distributed, with 0 mean and variance/covariance matrix given by $\mathbb{V}$, this is equivalent to

$$\int_{\delta \cdot x < -\mathbf{VaR}} e^{-x^{\dagger} \mathbb{V}^{-1} x/2} \frac{dx}{\sqrt{\det(2\pi \mathbb{V})}} = 0.05.$$

We now introduce the Cholesky decomposition of $\mathbb{V} = \mathbb{H}\,\mathbb{H}^{\dagger}$, and change variables

$$x\mathbb{H}^{-1} = y,$$

to obtain

$$\int_{\delta\mathbb{H}y^{\dagger}<-\mathbf{VaR}} e^{-|y|^2/2}\,\frac{dy}{(2\pi)^{n/2}} = 0.05.$$

Let $\mathbb{A}$ be the rotation that sends $\delta\mathbb{H}$ into $(|\delta\mathbb{H}|, 0, \ldots, 0)$, and change variables $y = z\mathbb{A}$, to obtain

$$\int_{|\delta\mathbb{H}|z_1<-\mathbf{VaR}} e^{-|z|^2/2}\,\frac{dz}{(2\pi)^{n/2}} = 0.05.$$

Since

$$\int_{|\delta\mathbb{H}|z_1<-\mathbf{VaR}} e^{-|z|^2/2}\,\frac{dz}{(2\pi)^{n/2}} = \int_{-\infty}^{-\mathbf{VaR}} e^{-x^2/2}\,\frac{dx}{\sqrt{2\pi}},$$

we conclude that

$$\begin{aligned}\mathbf{VaR} &= -z\sqrt{|\delta\mathbb{H}|}\\ &= -z\sqrt{\delta^T \cdot \mathbb{V}^{-1} \cdot \delta},\end{aligned}$$

where $\mathbb{V}$ is the variance/covariance matrix, and z is the 95% percentile of the univariate normal distribution.

- Simple
- Inaccurate for non–linear instruments.

Example

Let's attempt to recreate the situation at Orange County.

We would have a portfolio $\Pi(r_1, \ldots, r_n)$ of interest rate instruments, depending therefore on yield rates r_i, for i ranging from 1 (overnight lending rate) to 14 (the 30 year rate).

The yield curve will have a covariance matrix $\mathbb{V}$ that measures the market volatility with respect to motions of the interest rate curve.

We can find the principal components of $\mathbb{V}$, which will give us the directions of movement in the market that will tend to be more pronounced.

In doing this, we may find that parallel shifts of the yield curve will be most likely, with a 60% chance. Tilts will be next, with a 25% chance.

Our portfolio is perfectly hedged against parallel shifts; it is exposed to tilts. In other words, if we map the risk factors r_i into the principal components, and keep only the first two ones (then most likely ones), we may find

$$\delta = (0, 50\%),$$

expressed in percentage points.

It would be a mistake to think that the portfolio is risk-free because it is insensitive to parallel shifts. Similarly, it would be a mistake to think (as the OC officials did) that the portfolio is risk-free because if we hold to maturity, it will make money with certainty.

In fact, Delta-Normal-VaR will, in our situation, yield

$$1.65 \cdot \sqrt{\delta^\dagger \mathbb{V} \delta} \approx 40.$$

This means that we should expect our portfolio to lose 40% of its value once a month.

Bond – VaR

(joint with P. Fernández.; A. Kreinin)

Zero coupon bond with notional N_0. Its price is

$$P(R(t)) = N_0\, e^{-R(t)},$$

where $R(t)$ is the one-year interest rate.

$$R = R_0\, e^{\zeta},$$

where R_0 is today's interest rate and ζ is a normal distributed random variable with mean zero and variance σ^2 (the daily volatility of the interest rate).

R_0 and R today's and tomorrow's interest rates, respectively.

α-Var is defined through

$$\mathrm{Prob}\left\{N_0 e^{-R_0} - N_0 e^{-R} > P_\alpha\right\} = \alpha.$$

We solve this by setting $\lambda_\alpha = P_\alpha N_0^{-1} e^{R_0}$:

$$\begin{aligned}
\alpha &= \mathrm{Prob}\ \left\{1 - e^{R_0(1-e^{\zeta})} > \lambda_\alpha\right\} \\
&= \mathrm{Prob}\left\{e^{\zeta} > 1 - \frac{1}{R_0}\log(1-\lambda_\alpha)\right\} \\
&= \mathrm{Prob}\left\{\zeta > \log\left(1 - \frac{1}{R_0}\log(1-\lambda_\alpha)\right)\right\}.
\end{aligned}$$

As ζ is a $N(0, \sigma^2)$ random variable, we obtain that

$$\log\left(1 - \frac{1}{R_0}\log(1-\lambda_\alpha)\right) = \sigma q_\alpha,$$

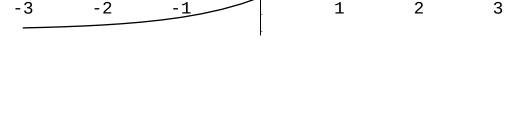
where q_α is the α-quantile of the standard normal distribution, that is

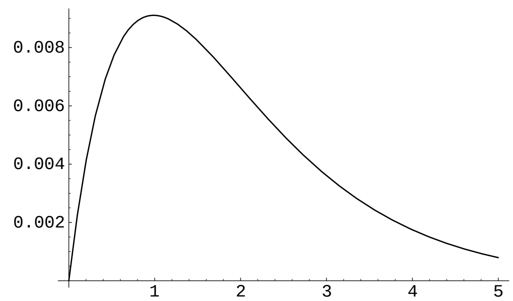
$$\alpha = \frac{1}{\sqrt{2\pi}}\int_{q_\alpha}^{\infty} e^{-y^2/2}dy.$$

With the above definition of λ_α, we can solve for P_α to obtain

$$P_\alpha = N_0 e^{-R_0}\left(1 - e^{R_0(1-e^{\sigma q_\alpha})}\right).$$

It is important to note that this exact calculation is possible because the increment in the value of the bond $P(0) - P(1)$ is an increasing function of the random variable ζ. And this allows us to invert the relation between both of them.





VaR in terms of the interest rate R_0

An approximation Approach

Taylor series:

$$P(R) \approx P(R_0) + \underbrace{\left.\frac{\partial P}{\partial R}\right|_{R=R_0}}_{\delta}(R - R_0) + \underbrace{\frac{1}{2}\left.\frac{\partial^2 P}{\partial R^2}\right|_{R=R_0}}_{\Gamma}(R - R_0)^2$$

Linear approx.

$$\delta = \left.\frac{\partial(N_0 e^{-R})}{\partial R}\right|_{R=R_0} = -N_0 e^{-R_0}.$$

The P&L:

$$f(\zeta) = P(R_0) - P(R) = -\delta(R - R_0) = N_0 R_0 e^{-R_0}\left(e^{\zeta} - 1\right).$$

And, as is an increasing function of ζ,

$$P_\alpha = N_0 e^{-R_0} R_0 (e^{\sigma q_\alpha} - 1).$$

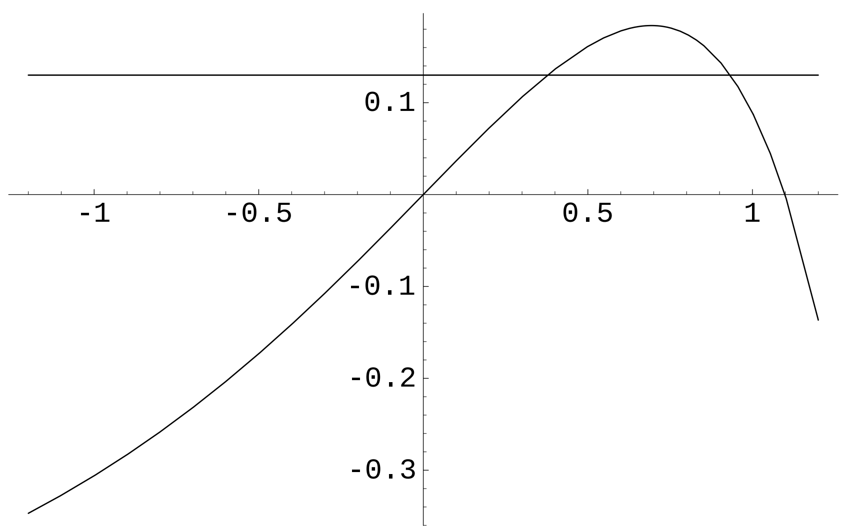
Quadratic attempt.

$$\Gamma = \frac{1}{2} \frac{\partial^2 (N_0 e^{-R})}{\partial R^2}\bigg|_{R=R_0} = N_0 e^{-R_0}.$$

P&L:

$$g(\zeta) = N_0 e^{-R_0} R_0 \left[e^{\zeta} - 1 - \frac{R_0}{2}(e^{\zeta} - 1)^2 \right].$$

And this not an increasing function any more. In fact, it attains a maximum at $\zeta = \log(1 + 1/R_0)$.



Picture of the function $g(\zeta)$.

RiskMetrics

The principal assumption of RiskMetrics methodology is that the log-returns of the underlying are small, and so, we can estimate the difference between two consecutive values keeping only the first term of the corresponding Taylor series. In our case, that means

$$R - R_0 = R_0 e^{\zeta} - R_0 \sim R_0 \zeta.$$

With this assumption, we can perform the same calculation as before:

$$\begin{aligned} P(R_0) - P(R) &= -\delta(R - R_0) - \Gamma(R - R_0)^2 - \ldots \\ &= -\delta R_0(\zeta + \zeta^2 + \ldots) - \Gamma R_0^2(\zeta + \ldots)^2 + \ldots \end{aligned}$$

For the δ-approximation, we should only keep the first order terms in ζ, that is,

$$P(R_0) - P(R) = -\delta\zeta = N_0 e^{-R_0} R_0 \zeta.$$

Again, this is an increasing function of ζ, and the VaR is easily calculated:

$$\alpha - \mathrm{VaR}_\delta = N_0 e^{-R_0} R_0 \sigma q_\alpha.$$

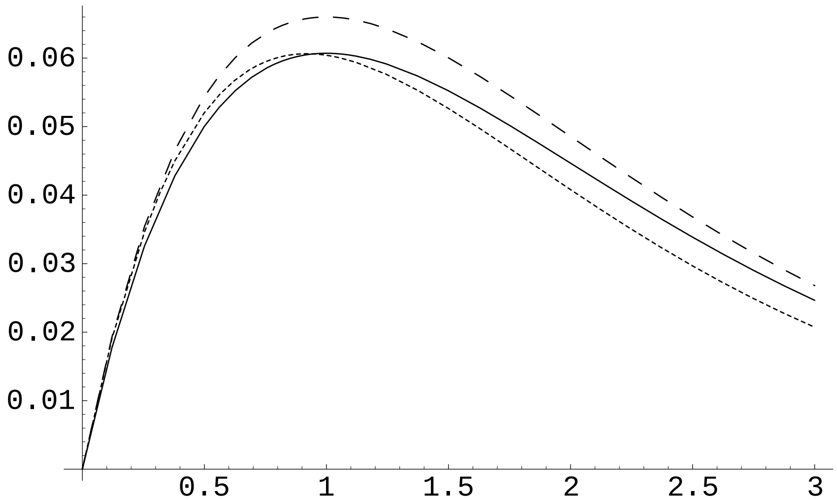
For the second order approximation, we should keep all the terms up to second order in ζ, that is,

$$\begin{aligned} P(R_0) - P(R) &= -\delta R_0 \zeta - \delta R_0 \zeta^2 - \Gamma R_0^2 \zeta^2 \\ &= N_0 e^{-R_0} R_0 \left[\zeta + \zeta^2 \left(1 - \frac{R_0}{2}. \right) \right]. \end{aligned}$$

And the $\alpha - \mathrm{Var}_{\delta-\Gamma}$, P_α is now calculated as:

$$\alpha = \mathrm{Prob} \left(\zeta + \left(1 - \frac{R_0}{2} \right) \zeta^2 > \frac{P_\alpha}{N_0 e^{-R_0} R_0} \right)$$

Again, this can be solved numerically for P_α.



Comparison of the different VaRs obtained. In continuous line, the RiskMetrics approach. In small-dotted line, the exact value. In big-dotted, the first order approximation.

Quadratic Finance

Consider portfolio of price Π with $S_1, \ldots, S_n$ as underlyings.

Underlying vector

$$\mathbf{S} = (S_1, \ldots, S_n).$$

Portfolio parameters:

$$\text{Delta}: \qquad \Delta = \nabla_{\mathbf{S}} \Pi = \left(\frac{\partial \Pi}{\partial S_1}, \; \cdots \;, \frac{\partial \Pi}{\partial S_n} \right)$$

$$\text{Gamma}: \qquad \Gamma = [\text{Hessian}]_{\mathbf{S}} \Pi = \left\{ \frac{\partial^2 \Pi}{\partial S_i \, \partial S_j} \right\}$$

	Finance	*Time*
Δ	Linear	Short Term
Γ	Non–linear	Long Term

Quadratic approximation:

$$\Pi(t) \approx \Pi(0) + \Delta \cdot \xi + \tfrac{1}{2}\, \xi \cdot \Gamma \cdot \xi^t, \qquad \xi = \mathbf{S}(t) - \mathbf{S}(0)$$

Quad–VaR

Under the quadratic approximation, this becomes

$$\textbf{Prob}\ \left\{ \Delta \cdot \xi + \tfrac{1}{2}\,\xi \cdot \Gamma \cdot \xi^t < -\textbf{VaR}\ \right\} = 0.05$$

for

$$\xi = \mathbf{S}(t) - \mathbf{S}(0),$$

log–normally distributed. After some elementary analysis, VaR changes into a related quantity K, given by

$$\boxed{I_0(K) \ = \ \int_{x\cdot\Delta + \frac{1}{2}\,(x,\Gamma\,x) \le K} e^{-\pi(x,\mathbb{A}\,x)}\,dx = 0.05}$$

- Complexity is independent of number of instruments
- Complexity $\rightarrow$ number of underlying risk factors.
- Monte Carlo friendly

  (full instrument valuation replaced by algebraic formula).

Two Analytical Issues

Asymptotic Analysis. Regard VaR as the solution of

$$I_0(K) = \alpha,$$

for $\alpha = 0.05$, and solve in the limit $\alpha \to 0$.

- Explicit algebraic expressions.
- Vega friendly.

Visualization of Risk. Display the dependence of VaR as a function of the delta, gamma and $\mathbb{V}$ of the portfolio.

Portfolio Volatility.

Lemma :

$$I_0(K) = \int_{\Delta\cdot x + \frac{1}{2}x^t\Gamma x \le -K} \exp(-x\mathbb{V}^{-1}x^t/2)\ \frac{dx}{\sqrt{\det 2\pi\mathbb{V}}}$$

$$= \int_{\frac{1}{2}\, x\,\mathbb{D}\, x^t \le R} \exp(-|x-v|^2/2)\ \frac{dx}{(2\pi)^{n/2}}$$

where v replaces delta, $\mathbb{D}$ replaces $\mathbb{V}$ and Γ, and R replaces VaR, given by $-K$.

PROOF:

Change variables for $x = y\mathbb{V}^{1/2}$,

$$I_0(K) = \int_{\Delta'\cdot y + \frac{1}{2}y^t\Gamma' y \le -K} \exp(-|y|^2/2)\ \frac{dy}{(2\pi)^{n/2}},$$

where

$$\Delta' = \Delta\mathbb{V}^{1/2}, \qquad \Gamma' = \mathbb{V}^{1/2}\Gamma\mathbb{V}^{1/2}.$$

Put

$$\Gamma' = \mathbb{S}^{-1}\mathbb{D}\mathbb{S}$$

with $\mathbb{D}$ diagonal and $\mathbb{S}$ orthogonal. Change variables again, $z = \mathbb{S}y$

$$I_0(K) = \int_{\Delta''\cdot z + \frac{1}{2}z\mathbb{D}z^t \le -K} \exp(-|z|^2/2)\ \frac{dz}{(2\pi)^{n/2}},$$

with

$$\Delta'' = \Delta' \mathbb{S}^{-1} = \Delta \mathbb{V}^{1/2}\, \mathbb{S}^{-1}.$$

Finally, put

$$z = (x - v), \quad v = \Delta'' \cdot \mathbb{D}^{-1},$$

which yields

$$I_0(K) = \int_{\frac{1}{2} z\mathbb{D}z^t \leq -K + \Delta''\mathbb{D}^{-1}\Delta''^t} \exp(-|z-v|^2/2)\, \frac{dz}{(2\pi)^{n/2}},$$

QED

Remark Putting

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{n_1} \geq -\mu_1 \geq \cdots \geq -\mu_{n_2}$$

the risk factor largest responsible for VaR is the one corresponding to μ_{n_2}.

Reduction to positive Gamma

"The eleventh commandment was 'Thou Shalt Compute' or 'Thou Shalt Not Compute' . . . I forget which."

– Epigrams in Programming, ACM SIGPLAN Sept. 1982

For $\mathbb{D}$ as before (real diagonal),

$$I(R) = \int_{x\cdot\mathbb{D}\cdot x^t \leq R} e^{-|x-\nu|^2}\, dx.$$

$$\mathbb{D} = \begin{pmatrix} \mathbb{D}_1^{-1} & 0 \\ 0 & -\mathbb{D}_2^{-1} \end{pmatrix}$$

with $\mathbb{D}_1$, $\mathbb{D}_2$ positive diagonal.

$$\begin{aligned} I(R) &= \int_{|x-\nu_1|^2-|y-\nu_2|^2\leq R} e^{-(x,\mathbb{D}_1 x)-(y,\mathbb{D}_2 y)}\, dx\, dy \\ &= \int_0^\infty \int_{|y-\nu_2|^2=r^2} e^{-(y,\,\mathbb{D}_2 y)} \int_{|x-\nu_1|^2\leq r^2+K} e^{-(x,\mathbb{D}_1 x)}\ r^{n_2-1}\ dr \\ &= \int_0^\infty I_{n_1}(\sqrt{r^2+K})\ \frac{\partial}{\partial r} I_{n_2}(r^2)\ r^{n_2-1}\ dr \end{aligned}$$

Harmonic VaR

(joint work with C. Albanese).

$$\int_{|x|\le R} e^{-\pi(x-v)\mathbb{D}^{-1}(x-v)^t}\,dx$$

$$= R^n \int_{\mathbb{R}^n} \frac{J_{\frac{n}{2}}(2\pi R\,|\xi|)}{|R\,\xi|^{\frac{n}{2}}} e^{-\pi(\xi,\mathbb{D}\xi)}\cos(2\pi\xi\cdot v)\frac{d\xi}{\sqrt{\det\pi\mathbb{D}^{-1}}}$$

$$= \sum_{k,j}^{\infty} R^{n+2k}\,\pi^{k+j}\,a_{kj}\int_{\mathbb{R}^n}|\xi|^{2k}(\xi\cdot v)^{2j}\,e^{-\pi\xi\mathbb{D}\xi}\frac{d\xi}{\sqrt{\det\pi\mathbb{D}^{-1}}},$$

The a_{kj} are basically the Taylor coefficients of the Bessel functions and the cosine function.

To compute each Gaussian moment, put

$$f(\alpha,\beta) = \left\{\det(\mathbb{D}+i\alpha+i\beta\,v^t\,v)\right\}^{-\frac{1}{2}}.$$

This can be easily computed, since $\mathbb{D}$ is diagonal, and

$$f(\alpha,\beta) = \left(\prod_{j=1}^{n}(\lambda_j+i\alpha)\right)^{-1/2}\cdot\left(1+\sum_{j=1}^{n}\frac{i\beta\,|v_j|^2}{\lambda_j+i\alpha}\right)^{-1/2}.$$

Then

$$\int_{\mathbb{R}^n} |\xi|^{2k} (\xi \cdot v)^{2j} e^{-\pi \xi \mathbb{D} \xi} \, d\xi = i^{k+j} \left. \frac{\partial^k}{\partial \alpha^k} \right|_{\alpha=0} \left. \frac{\partial^j}{\partial \beta^j} \right|_{\beta=0} f(\alpha, \beta)$$

Further,

$$\frac{\partial^{i+j} f}{\partial \alpha^i \, \partial \beta^j}(0,0) = \int_{\mathbb{R}^2} (2\pi i \hat{\alpha})^k \left(2\pi i \hat{\beta}\right)^j \widehat{f}(\hat{\alpha}, \hat{\beta}) \, d\hat{\alpha} \, d\hat{\beta}.$$

A technical point: f is not integrable in β, but $\partial_\beta f$ is. Hence,

$$F(\hat{\alpha}, \hat{\beta}) = \int_{\mathbb{R}^2} e^{-2\pi i (\hat{\alpha}\alpha + \hat{\beta}\beta)} \frac{\partial f}{\partial \beta}(\alpha, \beta) \, d\alpha \, d\beta$$

$$G(\hat{\alpha}) = \int_{\mathbb{R}} e^{-2\pi i \hat{\alpha}\alpha} f(\alpha, 0) \, d\alpha,$$

which yields

$$I(R) = R^{\frac{n}{2}} \left\{ \int_{-\infty}^{0} d\hat{\alpha} \frac{J_{\frac{n}{2}}\left(2R\pi\sqrt{|\hat{\alpha}|}\right) G(\hat{\alpha})}{(2\,|\hat{\alpha}|)^{\frac{n}{4}}} \right.$$

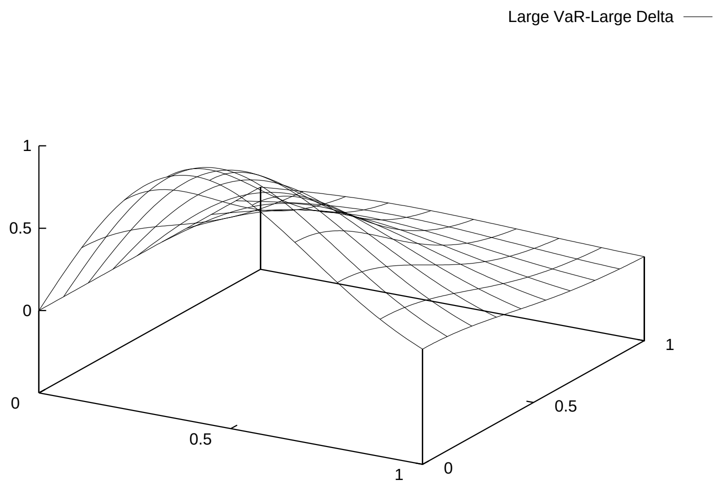
$$\left. + \pi i \iint \frac{\cos(2\pi\sqrt{2\hat{\beta}}) - 1}{2\pi^2 \hat{\beta}} \frac{J_{\frac{n}{2}}\left(2R\pi\sqrt{|\hat{\alpha}|}\right) F(\hat{\alpha}, \hat{\beta})}{(2\,|\hat{\alpha}|)^{\frac{n}{4}}} \, d\hat{\alpha} \, d\hat{\beta} \right\}$$

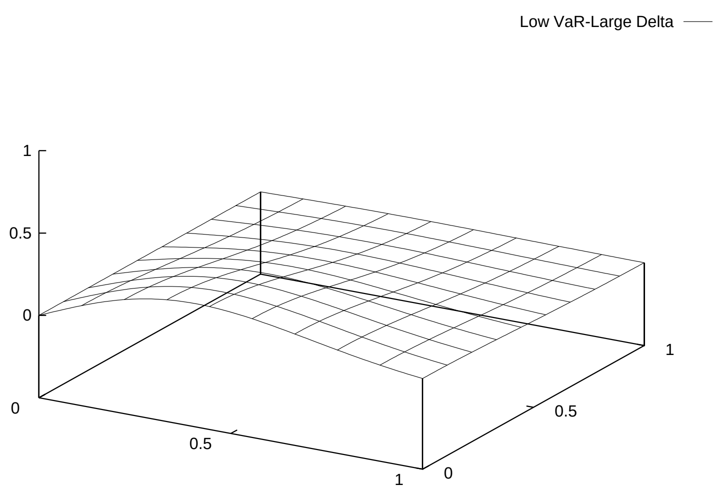
Lemma $F(\hat{\alpha}, \hat{\beta})$ is supported inside

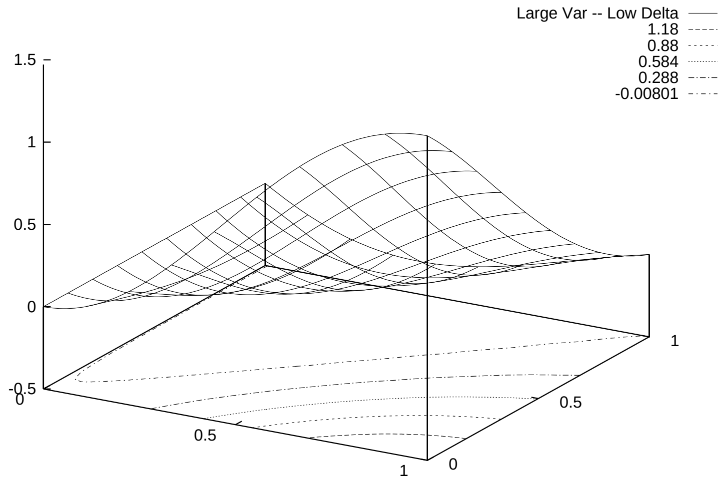
$$0 \leq \hat{\theta} \leq \tan^{-1} \|v\|^2,$$

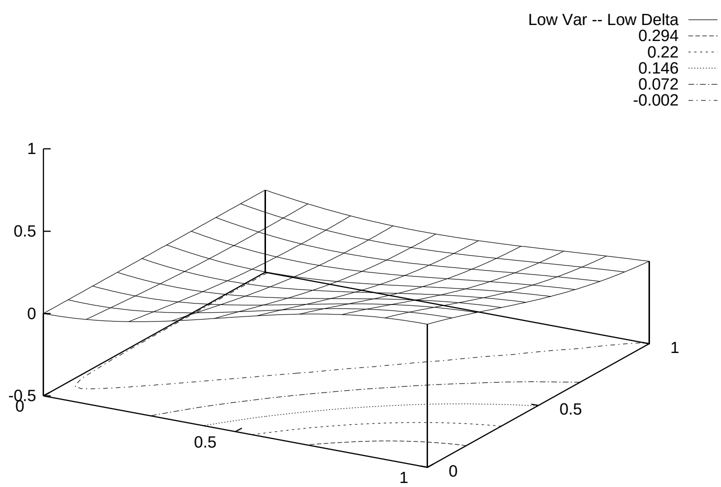
where $\hat{\theta}$ is the angle between $\hat{\alpha}$ and $\hat{\beta}$.

Visualizing Risk






Large Var -- Low Delta
1.18
0.88
0.584
0.288
-0.00801
1.5
1
0.5
0
-0.5
0
0.5
1
0
0.5
1


Low Var -- Low Delta
0.294
0.22
0.146
0.072
-0.002
1
0.5
0
-0.5
0
0.5
1
0
0.5
1

Asymptotic VaR

Joint work with R. Brummelhuis, A. Córdoba, M. Quintanilla

We want to solve

$$I(R) = \alpha,$$

for α near 0. Note that this is equivalent to the limit $R \to \infty$.

$$I(R) \approx \frac{\sqrt{2}\,\exp(-2R^2\lambda_1^{-1})}{\Pi_{i=2}^n(\lambda_i^{-1} - \lambda_1^{-1})^{1/2}\,R\,\lambda_1^{-1}}$$

Hence, VaR is approximated by the solution to the equation

$$\frac{\sqrt{2}\,\exp(-2R^2\lambda_1^{-1})}{\Pi_{i=2}^n(\lambda_i^{-1} - \lambda_1^{-1})^{1/2}\,R\,\lambda_1^{-1}} = 0.05.$$

Sketch or proof.

$$e^{-|x|^2/2} = (2\pi)^{n/2}\delta_0(x) + \frac{1}{2}\int_0^1 \Delta_x\left(\frac{e^{-|x|^2/2t}}{t^{n/2}}\right)dt$$

hence

$$I(R) = \frac{1}{2}\int_0^1 \frac{dt}{t^{n/2}} \int_{x\mathbb{D}x\geq R} \Delta_x(e^{-|x|^2/2})\,dx$$

The following lemma takes care of the rest.

Lemma *Let D be a manifold, with x_0 closest to then origin. Then,*

$$\int_D \Delta(\exp(-\lambda|x|^2/2))\,dx = e^{(-\lambda|x_0|^2/2)}$$

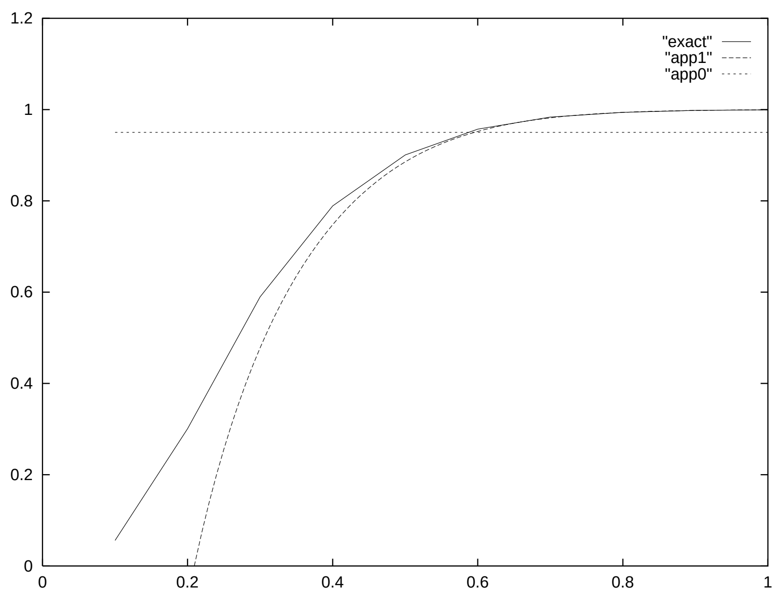
$$\cdot\left(\sum_{\nu<N} c_\nu \lambda^{-(n-3)/2-\nu} + O(\lambda^{-(n-3)/2-N})\right),$$

The first term is

$$c_0 \;=\; 2(2\pi)^{(n-1)/2}\cdot|x_0|\cdot\det(I+|x_0|K)^{-1/2},$$

where K is the principal curvature matrix at x_0.

Application to Value-at-Risk



95%

Credit risk

- Hedging credit risk is difficult, although Credit Derivatives provide a replicating approach to managing credit risk. Hence, arbitrage–free pricing is replaced by risk neutral pricing.

- Credit risk assumes an option contract with a party that is default prone. Credit exposure is based on default probabilities, which we assume to be independent of the underlying security of the option. This simplifies calculations but it is not realistic.

- When default occurs, a portion of the value of the asset can be usually recovered. This sill be modeled into the theory through the *recovery rate*, which we will simply assume to be constant.

Credit Premium

Losses due to credit risk will follow a certain probability distribution.

Expected Loss The expected losses under the default probability distribution.

Unexpected Loss The standard deviation of the losses.

As a trader, you want to charge for both.

Problems.

- The distribution of losses is not normal. And non-parametric approaches are hard.
- Discounting. A pricing scheme should discount to present value losses that will take place in the future.

Credit Spread

We define the credit spread s as follows:

let $P^*(t,T)$ be price of the bond issued by the default-prone party:

$$s = \frac{1}{T} \log \frac{P^*(0,T)}{P(0,T)}.$$

More generally, if we know the price V^* of a contract with the counterparty with cash flows c_i at times t_i,

$$V^* = \sum_i c_i \, P(0,t_i) \, e^{-s\, t_i}.$$

Note that the default-free price is

$$V = \sum_i c_i \, P(0,t_i).$$

Remark: Liquidity constraints will give rise to similar effects.

The Hazard Rate

Let T^* be the default time, and F its probability distribution:

$$F(t) = \text{Prob } \{t^* > t\} \, .$$

The process $h(t)$ is the hazard rate when

$$F(t) = \mathbb{E} \left(e^{-\int_0^t h(s)\, ds} \right) .$$

F can be calibrated through the observed price of coupon bearing bonds maturing at t_i, as follows:

$$P_j^* = \sum_i c_j^i \, P(0, t_i^j) \, F(t_i^j) + P(0, T_j) \, F(T^j)$$

$$+ R \cdot \sum_i P(0, t_i^j) \left(F(t_{i-1}^j - F(t_i^j) \right) .$$

R is the recovery rate.

This admits a solution in the form

$$F(t) = e^{-\int_0^t a(s)\, ds} ,$$

for a piecewise linear a.

Hazard Rate Models

We can now use the similarity between the hazard and the short rate, to translate the methodology for interest rate models into the modeling of the hazard rate evolution.

In a general setting, we can consider $h(t) = F(X_t)$, where

$$dX_t = \mu(X_t)\,dt + \sigma(X_t)\,dW_t,$$

It gives rise to a Markovian model.

The joint evolution of the interest rates and the hazard rate gives rise to two factor models.