Presessional course
Mathematical finance
and
Financial engineering & MORSE

Andrea Pachera

axp649@student.bham.ac.uk

23-27.09.19

Outline

1. Calculus
2. Probability
3. Taylor series
4. Non-linear problems
5. Finite difference schemes
6. Chain rule

Limits

Definition

$$\lim_{x\to x_0} f(x) = l \Leftrightarrow \forall\varepsilon > 0\ \exists\delta > 0 \text{ s.t. } |f(x) - l| < \varepsilon\ \forall|x - x_0| < \delta$$

$$\lim_{x\to x_0} f(x) = +\infty \Leftrightarrow \forall M > 0\ \exists\delta > 0 \text{ s.t. } f(x) > M\ \forall|x - x_0| < \delta$$

$$\lim_{x\to +\infty} f(x) = l \Leftrightarrow \forall\varepsilon > 0\ \exists N > 0 \text{ s.t. } |f(x) - l| < \varepsilon\ \forall x > N$$

$$\lim_{x\to +\infty} f(x) = +\infty \Leftrightarrow \forall M > 0\ \exists N > 0 \text{ s.t. } f(x) > M\ \forall x > N$$

And similarly for $-\infty$.

Rules of thumb

Continuity

If $f(x)$ is continuous at x_0, then $\lim_{x \to x_0} f(x) = f(x_0)$.

Growth

For $x \to +\infty$, $e^x >> x^n >> \log x$.
For polynomials, everything is decided by the leading term.

Quotient

How to deal with $\lim \frac{f(x)}{g(x)}$:

- if $f(x) \to a$, $g(x) \to b \neq 0$, then $\lim = a/b$
- if $f(x) \to a$, $g(x) \to 0$, then $\lim = \pm\infty$ (check the sign)
- if $f(x) \to a$, $g(x) \to \pm\infty$, then $\lim = 0$
- if $f(x) \to \pm\infty$, $g(x) \to b$, then $\lim = \pm\infty$ (check the sign)
- otherwise ($\pm\infty / \pm\infty$, $0/0$), apply L'Hôpital

L'Hôpital

Theorem: L'Hôpital

Let $x_0 \in \widetilde{\mathbb{R}} = \mathbb{R} \cup \pm\infty$, and f, g differentiable functions such that $\lim_{x \to x_0} \frac{f(x)}{g(x)}$ is of type $\frac{\pm\infty}{\pm\infty}$ or $\frac{0}{0}$.
If $\lim_{x \to x_0} \frac{f'(x)}{g'(x)}$ exists, and if there exists a neighbourhood U of x_0 such that $g'(x) \neq 0$ in $U \setminus \{ x_0 \}$, then

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \lim_{x \to x_0} \frac{f'(x)}{g'(x)}.$$

Differentiation

Definition

$$\frac{d}{dx}f(x) = f'(x) := \lim_{h\to 0}\frac{f(x+h)-f(x)}{h}$$

Rules

- Product rule: $(f(x)g(x))' = f'(x)g(x) + f(x)g'(x)$
- Chain rule: $(f(g(x)))' = (f'(g(x))) \cdot g'(x)$
- Quotient rule: $\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x)g(x) - f(x)g'(x)}{(g(x))^2}$
- Inverse: $(f^{-1}(x))' = \frac{1}{\frac{d}{dx}f(f^{-1}(x))}$

Common differentiations

$f(x)$	$f'(x)$
x^k	kx^{k-1}
$\ln x$	$\frac{1}{x}$
e^x	e^x
$\sin x$	$\cos x$
$\cos x$	$-\sin x$
$\tan x$	$\frac{1}{\cos^2 x}$
$\arcsin x$	$\frac{1}{\sqrt{1-x^2}}$
$\arccos x$	$-\frac{1}{\sqrt{1-x^2}}$
$\arctan x$	$\frac{1}{1+x^2}$

Multiple dimensions

Multivariate function

If $f : \mathbb{R}^n \to \mathbb{R}$, then

$$\frac{\partial}{\partial x_i} f(x) = \lim_{h \to 0} \frac{f(x_1, \ldots, x_{i-1}, x_i + h, x_{i+1}, \ldots, x_n) - f(x)}{h}.$$

Gradient, Hessian

$$\nabla f(x) = \left(\frac{\partial}{\partial x_1} f(x), \ldots, \frac{\partial}{\partial x_n} f(x) \right)$$

$$\nabla^2 f(x) = H(f(x)) = \left(\frac{\partial^2}{\partial x_i \partial x_j} f(x) \right)_{i,j}$$

Integration

Fundamental theorem of calculus

Let f be a continuous function on $[a, b]$, and F such that $\frac{d}{dx}F(x) = f(x)$ on $[a, b]$. Then

$$\int_a^b f(x)dx = [F(x)]_a^b = F(b) - F(a).$$

Facts

$$\int_a^a f(x)dx = 0$$

$$\int_a^b f(x)dx + \int_b^c f(x)dx = \int_a^c f(x)dx$$

$$\int_a^b (uf(x) + vg(x))dx = u\int_a^b f(x)dx + v\int_a^b g(x)dx$$

$$\left|\int_a^b f(x)dx\right| \leq \int_a^b |f(x)|dx$$

Improper integrals

Limit ∞

$$\int_a^{+\infty} f(x)dx = \lim_{t\to+\infty}\int_a^t f(x)dx$$

$$\int_{-\infty}^{b} f(x)dx = \lim_{t\to-\infty}\int_t^b f(x)dx$$

$$\int_{-\infty}^{+\infty} f(x)dx = \lim_{t_1\to-\infty}\int_{t_1}^a f(x)dx + \lim_{t_2\to+\infty}\int_a^{t_2} f(x)dx$$

Function not defined at the limit

If $f(a)$ is not defined, then

$$\int_a^b f(x)dx = \lim_{t\to a^+}\int_t^b f(x)dx.$$

Differentiating integrals

General rule

Let $f(x,t)$ be a continuous function s.t. $\frac{\partial}{\partial t} f(x,t)$ exists and is continuous. Then

$$\frac{d}{dt}\int_{a(t)}^{b(t)} f(x,t)dx = \int_{a(t)}^{b(t)} \frac{\partial}{\partial t} f(x,t)dx + f(b(t),t)\frac{d}{dt}b(t) - f(a(t),t)\frac{d}{dt}a(t).$$

Analytical integration

Antiderivative

If you know the antiderivative, apply the FTC to compute the integral.

Integration by parts

$$\int_a^b f'(x)g(x)dx = [f(x)g(x)]_a^b - \int_a^b f(x)g'(x)dx$$

Integration by substitution

Let $x = g(t)$, then

$$\int_a^b f(x)dx = \int_{g^{-1}(a)}^{g^{-1}(b)} f(g(t))g'(t)dt = [F(g(t))]_{g^{-1}(a)}^{g^{-1}(b)}.$$

Numerical integration

What if we cannot compute $\int_a^b f(x)dx$ analytically?
Approximation:

- divide $[a, b]$ into N subintervals $[a_i, a_{i+1}]$, with $a = a_0 < a_1 < \ldots < a_N = b$
- $h_i = a_i - a_{i-1}$, $x_i = \frac{a_i + a_{i-1}}{2} = a_{i-1} + \frac{h_i}{2}$ (length and midpoint)
- $$\int_a^b f(x)dx = \sum_{i=1}^{N} \int_{a_{i-1}}^{a_i} f(x)dx$$

Common approximations

Uniform subdivision: $h = \frac{b-a}{N}$

- midpoint rule: $\sim h \sum f(x_i)$
- trapezoidal rule: $\sim \frac{h}{2}(2 \sum f(a_i) - f(a) - f(b))$
- Simpson's rule: $\sim \frac{h}{6}(2 \sum f(a_i) - f(a) - f(b) + 4 \sum f(x_i))$

Financial applications

Zero rate

The zero rate $r(0, t)$ is the rate of return of a cash deposit made at time 0 and maturing at time t.

Instantaneous rate

The instantaneous rate $r(t)$ is the rate of return of deposits made at time t and maturing at time $t + dt$:

$$r(t) = \lim_{dt \to 0} \frac{1}{dt} \frac{B(t + dt) - B(t)}{B(t)} = \frac{B'(t)}{B(t)}.$$

Types of interest

Interest can be compounded at discrete time intervals (annually, monthly, etc.) or continuously, in which case $B(t) = e^{t \cdot r(0,t)} B(0)$.

Integrating the identity

$$\frac{B'(\tau)}{B(\tau)} = r(\tau)$$

gives

$$\int_0^t r(\tau) d\tau = \int_0^t \frac{B'(\tau)}{B(\tau)} d\tau = \log \frac{B(t)}{B(0)}.$$

By exponentiating both sides we obtain

$$B(t) = B(0) \exp\left(\int_0^t r(\tau) d\tau \right),$$

hence

$$r(0,t) = \frac{1}{t} \int_0^t r(\tau) d\tau.$$

Constant rate

If $r(0,t) = c$ for $t \in [0,T]$, equiv. $r(t) = c$ for $t \in [0,T]$, then for $t_1 < t_2$

$$B(t_2) = e^{c(t_2 - t_1)} B(t_1).$$

Financial applications

Derivative

A derivative is a financial asset that derives its value from another asset, called underlying asset.
e.g. shares of a stock, options, future, forwards.

Arbitrage

An arbitrage opportunity is an investment opportunity that is guaranteed to earn money without any risk involved.

Arbitrage-free pricing

"There is no such thing as a free lunch": no arbitrage principle.

The law of one price

In an arbitrage-free market, the following holds:

The law of one price

Let V_1 and V_2 be two portfolios. If there exists a future time $\tau > t$ such that $V_1(\tau) = V_2(\tau)$ regardless of the state of the market at the time τ, then $V_1(t) = V_2(t)$.

Analogous statements hold for $V_1(\tau) < V_2(\tau)$ and $V_1(\tau) > V_2(\tau)$.

Corollary

If the value $V(T)$ of a portfolio at time $T > t$ in the future is independent of the state of the market at time T, then

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

where r is the constant risk-free rate.

Plain vanilla European options

Call option on an underlying asset:

- contract between two parties
- the buyer pays $C(t)$
- a price K (strike) is fixed for a unit of the asset
- a time $T > t$ (maturity) is fixed
- at time T, the buyer has the right to buy a unit of the asset for K

Put option:

- same concept
- at time T, the buyer has the right to sell a unit of the asset for K

American option: the option can be exercised before maturity.
The buyer of the option is long the option, the seller is short the option.

Payoffs

Let $S(t)$ be the price of the underlying asset at time t.

	$S(T) < K$	$S(T) = K$	$S(T) > K$
Call	OTM	ATM	ITM
Put	ITM	ATM	OTM

Payoff:

$$C(T) = \max\{S(T) - K, 0\}$$

$$P(T) = \max\{K - S(T), 0\}$$

Put-Call parity

Put-Call parity

In an arbitrage-free market, the following holds:

$$P(t) + S(t) - C(t) = Ke^{-r(T-t)}.$$

If the underlying asset pays dividends continuously, then

$$P(t) + S(t)e^{-q(T-t)} - C(t) = Ke^{-r(T-t)}.$$

Proof: consider the portfolio $V(t) = P(t) + S(t) - C(t)$, evaluate at maturity and apply the corollary of the law of one price.

Probability

Discrete probability

Ω is the sample space, i.e. (countable) set of possible outcomes.
Random variable: $X : \Omega \to \mathbb{R}$.
Probability: $P : \Omega \to [0, 1]$ s.t. $\sum_{\omega \in \Omega} P(\omega) = 1$.

Continuous probability

Ω is the sample space. Let $\mathcal{A}$ be a family of subsets of Ω s.t.

1. $\Omega \in \mathcal{A}$;
2. if $A \in \mathcal{A}$ then $\Omega \setminus A \in \mathcal{A}$;
3. if $A_1, A_2, \ldots \in \mathcal{A}$ then $\bigcup_{k=1}^{\infty} A_k \in \mathcal{A}$;

then $A \in \mathcal{A}$ is an event, i.e. collection of possible outcomes.
Probability: $P : \mathcal{A} \to [0, 1]$ s.t. $P(\Omega) = 1$ and
$P(\bigcup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} P(A_k)$ if $A_i \cap A_j = \emptyset$ for $i \neq j$.
Random variable: $X : \Omega \to \mathbb{R}$ s.t. $\{ \omega \in \Omega \mid X(\omega) \leq t \} \in \mathcal{A}$ for all $t \in \mathbb{R}$.

Cumulative distribution function

Discrete

$$F_X(x) = \sum_{y \leq x} P(X = y)$$

Continuous

$$F_X(x) = \int_{-\infty}^{x} f_X(y)dy$$

Probability density function: f nonnegative s.t. $\int_{-\infty}^{+\infty} f(x)dx = 1$ and

$$P(a \leq X \leq b) = \int_a^b f(x)dx.$$

Expected value

Discrete

$$E[X] = \sum_{\omega \in \Omega} P(\omega) X(\omega)$$

Continuous

$$E[X] = \int_{-\infty}^{+\infty} x f(x) dx$$

Other important values

Variance

$$\mathrm{Var}(X) := E[(X - E[X])^2] = E[X^2] - E[X]^2$$

Standard deviation

$$\sigma(X) := \sqrt{\mathrm{Var}(X)}$$

Covariance

$$\mathrm{Cov}(X, Y) := E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y]$$

Correlation

$$\mathrm{Corr}(X, Y) = \rho_{XY} := \frac{\mathrm{Cov}(X,Y)}{\sigma(X)\sigma(Y)}$$

Properties

X, Y, U, V are random variables on the same probability space

The expected value is linear:

$$E[aX + bY] = aE[X] + bE[Y]$$

$$\sigma(cX) = |c|\sigma(X)$$

$$\mathrm{Var}(aX + bY) = a^2\,\mathrm{Var}(X) + b^2\,\mathrm{Var}(Y) + 2ab\,\mathrm{Cov}(X, Y)$$

The covariance is symmetric bilinear:

$$\begin{aligned}\mathrm{Cov}(X + aU, Y + bV) = \mathrm{Cov}(X, Y) &+ a\,\mathrm{Cov}(U, Y) \\ &+ b\,\mathrm{Cov}(X, V) + ab\,\mathrm{Cov}(U, V)\end{aligned}$$

$$\mathrm{Corr}(X, Y) \in [-1, 1]$$

Independent random variables

Definition

X_1, X_2 over Ω are independent iff, for any $a_1 \leq b_1, a_2 \leq b_2$,

$P((a_1 \leq X_1 \leq b_1) \cap (a_2 \leq X_2 \leq b_2)) = P(a_1 \leq X_1 \leq b_1)P(a_2 \leq X_2 \leq b_2)$.

Equivalently, iff

$$P((a_1 \leq X_1 \leq b_1)|(a_2 \leq X_2 \leq b_2)) = P(a_1 \leq X_1 \leq b_1),$$

where $P(a_2 \leq X_2 \leq b_2) > 0$.
Equivalently, iff

$$f(x_1, x_2) = f_1(x_1)f_2(x_2),$$

where

$$P((a_1 \leq X_1 \leq b_1) \cap (a_2 \leq X_2 \leq b_2)) = \int_{a_2}^{b_2} \int_{a_1}^{b_1} f(x_1, x_2) dx_1 dx_2.$$

Independent random variables

Behaviour

X_1, X_2 over Ω independent. For any $h_1, h_2 : \mathbb{R} \to \mathbb{R}$ continuous functions we have

$$E[h_1(X_1)h_2(X_2)] = E[h_1(X_1)]E[h_2(X_2)].$$

Correlation

X_1, X_2 over Ω independent. Then

$$\mathrm{Cov}(X_1, X_2) = 0,$$

$$\mathrm{Corr}(X_1, X_2) = 0,$$

$$\mathrm{Var}(X_1 + X_2) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2).$$

Standard normal variable

Definition

$Z \sim N(0,1)$ is defined by the probability density function:

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

Cumulative distribution

$$N(t) := P(Z \leq t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{x^2}{2}} dx$$

Facts

$$E[Z] = 0, \qquad \mathrm{Var}(Z) = 1, \qquad E[Z^2] = 1, \qquad 1 - N(a) = N(-a).$$

Normal random variables

Definition

$X \sim N(\mu, \sigma^2)$ if $X = \mu + \sigma Z$. It has probability density function

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

Facts

$$E[X] = \mu, \qquad \mathrm{Var}(X) = \sigma^2, \qquad \sigma(X) = |\sigma|,$$
$$P(X \le t) = P\left(Z \le \frac{t-\mu}{\sigma}\right) = N\left(\frac{t-\mu}{\sigma}\right).$$

Lognormal distribution

Definition

Y is lognormal if $\log Y = X$ and X is normal, i.e. if $\log Y = \mu + \sigma Z$. It has probability density function

$$f(y) = \begin{cases} \frac{1}{y|\sigma|\sqrt{2\pi}} \exp\left(-\frac{(\log y - \mu)^2}{2\sigma^2}\right) & y > 0, \\ 0 & y \leq 0. \end{cases}$$

Facts

$$E[Y] = e^{\mu + \frac{\sigma^2}{2}}, \qquad \mathrm{Var}(Y) = e^{2\mu + \sigma^2}\left(e^{\sigma^2} - 1\right).$$

Independent (log)normals

Normal

Let $X_1 = \mu_1 + \sigma_1 Z \sim N(\mu_1, \sigma_1^2)$, $X_2 = \mu_2 + \sigma_2 Z \sim N(\mu_2, \sigma_2^2)$. If they are independent, then

$$X_1 + X_2 = (\mu_1 + \mu_2) + \sqrt{\sigma_1^2 + \sigma_2^2} Z \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$

Lognormal

Let $\log Y_1 = \mu_1 + \sigma_1 Z$, $\log Y_2 = \mu_2 + \sigma_2 Z$. If Y_1 and Y_2 are independent, then

$$\log(Y_1 Y_2) = \log Y_1 + \log Y_2 = (\mu_1 + \mu_2) + \sqrt{\sigma_1^2 + \sigma_2^2} Z.$$

Financial applications

Evolution of asset prices: lognormal model. Why?
The rate of return between t and $t + \delta t$ is given by

$$\frac{S(t + dt) - S(t)}{S(t)dt}.$$

If the rate was constant (say μ), then by taking the limit $dt \to 0$ we would obtain the ODE

$$S'(t) = \mu S(t),$$

with initial condition $S(0)$, and with solution $S(t) = S(0)e^{\mu t}$.
More realistic: add random oscillation around the mean:

$$\frac{S(t + dt) - S(t)}{S(t)dt} = \mu + \frac{\sigma}{\sqrt{dt}} Z.$$

μ is the drift, σ is the volatility of the asset.

Lognormal model

If $S(t)$ has the form

$$\frac{S(t+dt)-S(t)}{S(t)dt}=\mu+\frac{\sigma}{\sqrt{dt}}Z,$$

then it must satisfy the SDE

$$dS=(\mu-q)Sdt+\sigma SdX,$$

where $X(t)$ is a Weiner process, and q is the continuous rate at which the asset pays dividends.
In order to solve the SDE, $S(t)$ must satisfy the lognormal model

$$\log\frac{S(t_2)}{S(t_1)}=\left(\mu-q-\frac{\sigma^2}{2}\right)(t_2-t_1)+\sigma\sqrt{t_2-t_1}Z,$$

for $0\leq t_1<t_2$.

Risk-neutral valuation

Risk-neutral valuation refers to valuing derivative securities as discounted expected values of their payoffs at maturity, under the assumption that the underlying asset has lognormal distribution with drift μ equal to the constant risk-free rate r.

Warning

Risk-neutral valuation can be used for European options, both vanilla and with other payoffs at maturity.
Risk-neutral valuation cannot be used for path-dependent options (e.g. American options).

Risk-neutral derivation of Black-Scholes formula

The value of a derivative with payoff $V(T)$ at maturity T given by risk-neutral valuation is

$$V(0) = e^{-rT} E_{RN}[V(T)],$$

where $S(t)$ is the one we discussed previously, with $\mu = r$, $t_1 = 0$, $t_2 = T$:

$$\log \frac{S(T)}{S(0)} = \left(r - q - \frac{\sigma^2}{2}\right) T + \sigma\sqrt{T} Z,$$

or equivalently

$$S(T) = S(0) e^{\left(r-q-\frac{\sigma^2}{2}\right) T + \sigma\sqrt{T} Z}.$$

So for plain vanilla European options we have

$$C(0) = e^{-rT} E_{RN}[\max\{\, S(T) - K, 0 \,\}],$$

$$P(0) = e^{-rT} E_{RN}[\max\{\, K - S(T), 0 \,\}],$$

and we will focus on C (the working for P is analogous).

$$S(T) \geq K \Leftrightarrow Z \geq \frac{1}{\sigma\sqrt{T}}\left(\ln\frac{K}{S(0)} - \left(r - q - \frac{\sigma^2}{2}\right)T\right) =: -d_2.$$

So

$$\max\{\, S(T) - K, 0 \,\} = \begin{cases} S(0)e^{\left(r-q-\frac{\sigma^2}{2}\right)T+\sigma\sqrt{T}Z} - K & \text{if } Z \geq -d_2, \\ 0 & \text{if } Z < -d_2. \end{cases}$$

$E_{RN}[\max\{S(T)-K,0\}]$ can be seen as an expectation over Z, which has probability density function $\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$, so

$$C(0) = e^{-rT}\frac{1}{\sqrt{2\pi}}\int_{-d_2}^{\infty}\left(S(0)e^{\left(r-q-\frac{\sigma^2}{2}\right)T+\sigma\sqrt{T}x} - K\right)e^{-\frac{x^2}{2}}dx$$
$$= \frac{S(0)e^{-qT}}{\sqrt{2\pi}}\int_{-d_2}^{\infty}e^{-\frac{1}{2}(x-\sigma\sqrt{T})^2}dx - \frac{Ke^{-rT}}{\sqrt{2\pi}}\int_{-d_2}^{\infty}e^{-\frac{x^2}{2}}dx.$$

Define $-d_1 := -d_2 - \sigma\sqrt{T}$, and substitute $y = x - \sigma\sqrt{T}$:

$$C(0) = \frac{S(0)e^{-qT}}{\sqrt{2\pi}}\int_{-d_1}^{\infty}e^{-\frac{y^2}{2}}dy - \frac{Ke^{-rT}}{\sqrt{2\pi}}\int_{-d_2}^{\infty}e^{-\frac{x^2}{2}}dx$$
$$= S(0)e^{-qT}N(d_1) - Ke^{-rT}N(d_2).$$

This is the Black-Scholes formula for $t = 0$.

Black-Scholes formula

$$C(t) = S(t)e^{-q(T-t)}N(d_1) - Ke^{-r(T-t)}N(d_2)$$

$$P(t) = Ke^{-r(T-t)}N(-d_2) - S(t)e^{-q(T-t)}N(-d_1)$$

Interpretation of $N(d_1)$

If $q = 0$ then

$$\frac{\partial C}{\partial S} = N(d_1).$$

Interpretation of $N(d_2)$

$N(d_2)$ represents the risk-neutral probability that the call option expires in the money.

Probability of expiring in the money

Recall

$$\log \frac{S(T)}{S(0)} = \left(\mu - q - \frac{\sigma^2}{2}\right) T + \sigma\sqrt{T} Z.$$

Then

$$\begin{aligned}
P(S(T) > K) &= P\left(\log \frac{S(T)}{S(0)} > \log \frac{K}{S(0)}\right) \\
&= P\left(Z > \frac{1}{\sigma\sqrt{T}}\left(\log \frac{K}{S(0)} - \left(\mu - q - \frac{\sigma^2}{2}\right) T\right)\right) \\
&= P(Z > a) = P(Z < -a) = N(-a).
\end{aligned}$$

Let $d_{2,\mu} = -a$, so that $P(S(T) > K) = N(d_{2,\mu})$.
If the drift of the underlying asset is equal to the constant risk-free rate, i.e. $\mu = r$, then $d_{2,\mu} = d_2$.

Non-linear problems

Problem

How to solve

$$f(x) = 0$$

when f is not a linear function.

Idea

Iterative approach: construct a sequence of x_n such that $\lim x_n = x^*$ satisfies $f(x^*) = 0$.

Warning

Note that for such methods to work, f needs to be "nice enough", for example continuous or with continuous derivatives, etc.
The exact requirements will depend on the method used.

Bisection method

Intermediate value theorem

Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then for any $y \in f([a, b])$ there exists $x \in [a, b]$ such that $f(x) = y$.

Idea

If $f(a)$ and $f(b)$ have opposite sign, then there exists $x \in [a, b]$ such that $f(x) = 0$.
So keep restricting $[a, b]$ until it is "close enough" to x, i.e. it is contained in $[x - \varepsilon, x + \varepsilon]$ for some ε.

Bisection method

Let $a_0 = a < b = b_0$ be such that $f(a)f(b) < 0$.

1. Define $x_0 := \frac{a_0+b_0}{2}$.
2. If $f(x_0) \neq 0$ then:
   - if $f(x_0)f(a) > 0$ then set $a_1 := x_0$, $b_1 := b_0$,
   - if $f(x_0)f(b) > 0$ then set $b_1 := x_0$, $a_1 := a_0$.
3. Observe that $[a_1, b_1]$ is smaller than the previous interval (by half), and $f(a_1)f(b_1) < 0$.
4. Keep reiterating: when you have $[a_n, b_n]$ with $f(a_n)f(b_n) < 0$, define $x_n := \frac{a_n+b_n}{2}$, and set $[a_{n+1}, b_{n+1}]$ to be $[a_n, x_n]$ or $[x_n, b_n]$ accordingly.
5. Stop when $b_n - a_n < \delta$ and/or $\max\{\, |f(a_n)|, |f(b_n)| \,\} < \epsilon$.

Note that the bisection method converges to a solution.

Warning: multiple roots

If the starting interval $[a, b]$ contains more than one solution, you have no control over which one will obtained by bisection.

Newton's method

1. Let x_k be the approximation of the exact solution x^*.
2. Draw the tangent line to f from $f(x_k)$.
3. Let x_{k+1} be the x-intercept of said line.

This can be written as

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}.$$

Warning

- Newton's method does not converge for all starting values x_0.
- For certain f, the method will just not work!
- What if we do not know $f'(x)$?

Secant method

If we do not know an explicit expression to use for $f'(x)$ for Newton's method, we have to approximate it as well:

$$f'(x_k) \approx \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}.$$

So the recursion becomes:

$$x_{k+1} = x_k - \frac{(x_k - x_{k-1})f(x_k)}{f(x_k) - f(x_{k-1})}.$$

Warning

We need two starting points (x_{-1} and x_0) to initialise this. Note that we need $f(x_{-1}) \neq f(x_0)$.

This method has slower convergence than Newton's.

Taylor series

The Taylor series is used to approximate a function at a point using polynomials.

Assume f is differentiable n times, and define

$$P_n^f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k,$$

where $f^{(k)}$ denotes the k-th derivative ($f^{(0)} = f$, $f^{(1)} = f'$, and so on).

Note that

$$(P_n^f)^{(i)}(a) = f^{(i)}(a), \quad i = 0, \ldots, n.$$

Taylor series

Derivative form

Let $f \in \mathcal{C}^{n+1}$, i.e. it is $n+1$ times differentiable and $f^{(n+1)}(x)$ is continuous. Then there is a point η between a and x such that

$$f(x) - P_n^f(x) = \frac{(x-a)^{n+1}}{(n+1)!} f^{(n+1)}(\eta).$$

Integral form

Let $f \in \mathcal{C}^{n+1}$, i.e. it is $n+1$ times differentiable and $f^{(n+1)}(x)$ is continuous. Then

$$f(x) - P_n^f(x) = \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t) dt.$$

Big O and little o

Definition

Let $f, g \colon \mathbb{R} \to \mathbb{R}$. We write that

- $f(x) = O(g(x))$, as $x \to x_0 \in \widetilde{\mathbb{R}}$ iff $\limsup_{x \to x_0} \left| \frac{f(x)}{g(x)} \right| < \infty$.
- $f(x) = o(g(x))$, as $x \to x_0 \in \widetilde{\mathbb{R}}$ iff $\lim_{x \to x_0} \left| \frac{f(x)}{g(x)} \right| = 0$.

Note that by definition $O(cg(x)) = cO(g(x)) = O(g(x))$ and $o(cg(x)) = co(g(x)) = o(g(x))$ for any constant $c \neq 0$.

Taylor series in multiple variables

Let $f : \mathbb{R}^n \to \mathbb{R}$ and $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$. The Taylor approximation of f around a is:

- Second order approximation:

$$f(x) = f(a) + \sum_{i=1}^{n} (x_i - a_i) \frac{\partial}{\partial x_i} f(a) + \sum_{i=1}^{n} O(|x_i - a_i|^2).$$

- Third order approximation:

$$\begin{aligned} f(x) = f(a) &+ \sum_{i=1}^{n} (x_i - a_i) \frac{\partial}{\partial x_i} f(a) \\ &+ \sum_{1 \le i,j \le n} \frac{(x_i - a_i)(x_j - a_j)}{2} \frac{\partial^2}{\partial x_i \partial x_j} f(a) + \sum_{i=1}^{n} O(|x_i - a_i|^3). \end{aligned}$$

Example: Taylor in two variables

Let $f : \mathbb{R}^2 \to \mathbb{R}$:

- linear approximation (second order)

$$f(x,y) = f(a,b) + (x-a)\frac{\partial}{\partial x}f(a,b) + (y-b)\frac{\partial}{\partial y}f(a,b) + \\ + O(|x-a|^2) + O(|y-b|^2).$$

- quadratic approximation (third order)

$$f(x,y) = f(a,b) + (x-a)\frac{\partial}{\partial x}f(a,b) + (y-b)\frac{\partial}{\partial y}f(a,b) + \\ + \frac{(x-a)^2}{2}\frac{\partial^2}{\partial x^2}f(a,b) + (x-a)(x-b)\frac{\partial^2}{\partial x \partial y}f(a,b) + \\ + \frac{(y-b)^2}{2}\frac{\partial^2}{\partial y^2}f(a,b) + O(|x-a|^3) + O(|y-b|^3).$$

Common Taylor series

$$e^x = \sum_{k=0}^{+\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \dots \qquad x \in \mathbb{R}$$

$$\log(1+x) = \sum_{k=1}^{+\infty} (-1)^{k+1} \frac{x^k}{k} = x - \frac{x^2}{2} + \frac{x^3}{3} + \dots \qquad x \in (-1, 1]$$

$$\frac{1}{1+x} = \sum_{k=0}^{+\infty} (-1)^k x^k = 1 - x + x^2 - x^3 + \dots \qquad x \in (-1, 1)$$

$$\cos x = \sum_{k=0}^{+\infty} (-1)^n \frac{x^{2n}}{(2n)!} = 1 - \frac{x^2}{2} + \frac{x^4}{24} + \dots \qquad x \in \mathbb{R}$$

All of the above are expansions around 0.

Black-Scholes equation

Assumptions:

- For a portfolio Π, a long option $V(S,t)$ is hedged in derivative security, i.e. a short position in $\Delta := \frac{\partial V}{\partial S}$ units of S must be taken:

$$\Pi = V - \Delta S.$$

- The asset price follows a lognomal distribution, hence

$$(dS)^2 \sim \sigma^2 S^2 dt.$$

- No arbitrage opportunity.
- The risk free interest rate r and the asset volatility σ are known over time.
- Share price are perfectly divisible, there are no transactions fee and no dividends.

Over the time $[t, t+dt]$, the variation of the portfolio is

$$d\Pi = dV - \Delta dS.$$

No arbitrage principle: the rate of return on the portfolio is equal to the rate of return on any other riskless instrument:

$$\begin{aligned} d\Pi &= r\Pi dt \\ &= \big(V - \Delta S\big) rdt. \end{aligned}$$

Combining:

$$dV - \Delta dS = (V - \Delta S)rdt.$$

Taylor expansion of $V(S,t)$ around (S,t) evaluated at $(S+dS, t+dt)$:

$$\begin{aligned} V(S+dS, t+dt) = V(S,t) + dS\frac{\partial}{\partial S}V + dt\frac{\partial}{\partial t}V + \frac{(dS)^2}{2}\frac{\partial^2}{\partial S^2}(V) \\ + \frac{(dt)^2}{2}\frac{\partial^2}{\partial t^2}V + dSdt\frac{\partial^2}{\partial S \partial t}V + O((dS)^3) + O((dt)^3). \end{aligned}$$

Ignoring the terms of order larger than dt, we obtain

$$\begin{aligned} dV &= V(S+dS, t+dt) - V(S,t) \\ &\approx dS\frac{\partial}{\partial S}V + dt\frac{\partial}{\partial t}V + \frac{(dS)^2}{2}\frac{\partial^2}{\partial S^2}V \\ &\approx dS\frac{\partial}{\partial S}V + dt\frac{\partial}{\partial t}V + \frac{\sigma^2 S^2}{2}dt\frac{\partial^2}{\partial S^2}V. \end{aligned}$$

Recall

$$dV - dS\frac{\partial}{\partial S}V - Vrdt + Srdt\frac{\partial}{\partial S}V = 0$$

. Substituting dV

$$\left(\frac{\partial}{\partial t}V + \frac{1}{2}\sigma^2 S^2\frac{\partial^2}{\partial S^2}V + rS\frac{\partial}{\partial S}V - rV\right)dt = 0.$$

Blak-Scholes equation

$$\frac{\partial}{\partial t}V + \frac{1}{2}\sigma^2 S^2\frac{\partial^2}{\partial S^2}V + rS\frac{\partial}{\partial S}V - rV = 0$$

Finite difference schemes

This method is used to estimate solutions of ODE and PDE via approximations.

- first order ODE:
$$f'(x) = F(x, f(x))$$
- second order ODE:
$$f''(x) = F(x, f(x), f'(x))$$

We can split our interval $[a, b]$ where f is defined into subintervals and approximate values of the function at those.

First order approximation of f'

Taylor:

$$f(x) = f(a) + (x-a)f'(a) + O((x-a)^2).$$

Forward finite difference: use $x = a + h$, $h > 0$

$$f'(a) = \frac{f(a+h) - f(a)}{h} + O(h).$$

Backward finite difference: use $x = a - h$, $h > 0$

$$f'(a) = \frac{f(a) - f(a-h)}{h} + O(h).$$

Second order approximation of f'

Combine forward and backward:

$$f(a+h) = f(a) + hf'(a) + \frac{h^2}{2}f''(a) + O(h^3).$$

$$f(a-h) = f(a) - hf'(a) + \frac{h^2}{2}f''(a) + O(h^3).$$

Central finite difference for f': subtract the equations

$$f'(a) = \frac{f(a+h) - f(a-h)}{2h} + O(h^2).$$

Second order approximation of f''

Combine forward and backward:

$$f(a+h) = f(a) + hf'(a) + \frac{h^2}{2}f''(a) + \frac{h^3}{6}f'''(a) + O(h^4).$$

$$f(a-h) = f(a) - hf'(a) + \frac{h^2}{2}f''(a) - \frac{h^3}{6}f'''(a) + O(h^4).$$

Central finite difference for f'': add the equations

$$f''(a) = \frac{f(a+h) - 2f(a) + f(a-h)}{h^2} + O(h^2).$$

Approximating ODE

How to find a solution to the ODE?

$$\begin{cases} y'(x) & = f(x, y(x)), \quad x \in [a, b] \\ y(a) & = c_0 \end{cases}$$

A closed form solution may not even exist, or be unique.

Numerical solution

Solving a DE approximately means discretising the computational domain where the equation must be solved by choosing a finite number of points in the domain (usually equidistant), and finding numerical values at these discrete points (nodes) that approximate the values of the exact solution of the DE at those points.

1. Partition $[a, b]$ into n intervals $a = x_0 < x_1 < \ldots < x_n = b$. In general, $x_i = a + ih$, $h = \frac{b-a}{n}$.
2. Aim: find $y_i \approx y(x_i)$ for all i.
3. We have $y'(x_i) = f(x_i, y(x_i))$. (exact solution)
4. We have $y'(x_i) = \frac{y(x_{i+1}) - y(x_i)}{h} + O(h)$. (forward finite difference)
5. substitute and discretise (use the approximate y_i instead of the exact $y(x_i)$ to obtain
$$\frac{y_{i+1} - y_i}{h} = f(x_i, y(x_i)),$$
$$y_{i+1} = y_i + hf(x_i, y(x_i)).$$
6. boundary condition: $y_0 = y(x_0) = y(a) = c$.
7. recursively approximate y_{i+1} using y_i.

The scheme is convergent iff

$$\lim_{n\to\infty}\left(\max_{i=0,\ldots,n} |y_i - y(x_i)|\right) = 0.$$

Finite difference schemes in multiple variables

The Black-Scholes equation has two variables: $V(S,t)$. What to do in cases like this?
Idea: use a mesh on the (x,t)-plane, where points have the form (ndx, mdt); and approximate the value at a point using a finite number of "neighbour" points.

FTCS

FTCS: forward-time centred-space

- forward difference approximation of the time derivative
- centred approximation of the spatial derivative

Denote the points of the mesh by

$$(x_i, t^n) = (ih, n\delta), \quad 0 \leq i \leq N+1, \; n \geq 0,$$

$$y_i^n = y(x_i, t^n).$$

Forward difference for the time:

$$\frac{\partial}{\partial t} f = \frac{y_i^{n+1} - y_i^n}{\delta} + O(\delta).$$

Centred difference for the space:

$$\frac{\partial^2}{\partial x^2} f = \frac{y_{i+1}^n + y_{i-1}^n - 2y_i^n}{h^2} + O(h^2).$$

Chain rule in one variable

Statement

Let $f(x)$ be a differentiable function, and assume that $x = g(t)$, where g is a differentiable function. Then the composite function $f \circ g$ is also differentiable, and its derivative is given by

$$(f \circ g)'(t) = (f' \circ g)(t) \cdot g'(t).$$

Equivalently:

$$\frac{d}{dt}(f(g(t)) = \left(\frac{d}{dx}f\right)(g(t)) \cdot \frac{d}{dt}(g(t)).$$

The rule can be written as

$$\frac{df}{dt} = \frac{df}{dx}\frac{dx}{dt}.$$

Chain rule in two variables

Example 1

If $f(x, y) = f(x(t), y(t))$, then

$$\frac{d}{dt} f(x(t), y(t)) = x'(t) \cdot \left(\frac{\partial}{\partial x} f\right)(x(t), y(t)) + y'(t) \cdot \left(\frac{\partial}{\partial y} f\right)(x(t), y(t)).$$

Equivalently,

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}.$$

Chain rule in two variables

Example 2

If $f(x, y) = f(x(s,t), y(s,t))$, then

$$\frac{\partial f}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s}.$$

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t}.$$

Chain rule in multiple variables

Statement

Let $f(x_1, \ldots, x_n)$ be a differentiable function in n variables, and let each $x_i = x_i(t_1, \ldots, t_m)$ be a differentiable function in m variables for each i. Then

$$\frac{\partial}{\partial t_j} f = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \frac{\partial x_i}{\partial t_j}.$$

Change of variable to solve Black-Scholes

One way to solve the Black-Scholes PDE is to use a change of vavriables to reduce it to a heat equation, whose solution is known.
The change of variables is:

$$V(S,t) = e^{-ax-b\tau}u(x,\tau),$$

where

$$x = \log\frac{S}{K}, \quad \tau = \frac{(T-t)\sigma^2}{2}.$$

With an appropriate choice of the coefficient a and b, we can reduce

$$\frac{\partial}{\partial t}V + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2}{\partial S^2}V + (r-q)S\frac{\partial}{\partial S}V - rV = 0$$

to a heat equation, i.e. of the form

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial u}{\partial \tau},$$

which is much easier to solve.