Lecture 13

Midterm Review

Derivation of Black-Scholes-Merton Differential Equation

We are now in position to derive the Black-Scholes or Black-Scholes-Merton differential equation. We build the model via a riskless portfolio, as we did for binomial trees. As for binomial trees, we carry some stock along with shorting the option. The amount of stock changes instantaneously. Special assumptions are required:

1. The stock price follows the process defined earlier for μ and σ:

$$\frac{dS}{S} = \mu dt + \sigma dz$$

2. Short selling of securities with full use of proceeds is permitted
3. There are no transactions costs or taxes. All securities are perfectly divisible
4. There are no dividends during the life of the derivative
5. There are no riskless arbitrage opportunities
6. Security trading is continuous
7. The risk-free rate of interest, r, is constant and the same for all maturities

Derivation of Black-Scholes-Merton Differential Equation

Recall our process for a continuous stock movement modeled on an Itô process with expected gain μ and volatility σ.

$$dS = \mu S dt + \sigma S dz$$

Let f be the price of a call option that depends on S. The variable f depends, then S and t. Then

$$df = \left(\frac{\partial f}{\partial S} \mu S + \frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S^2 \right) dt + \frac{\partial f}{\partial S} \sigma S dz$$

Derivation of Black-Scholes-Merton Differential Equation

We now build a portfolio that will eliminate the stochasticity of the process. The appropriate portfolio (as we will see) is

- -1 option
- $\frac{\partial f}{\partial S}$ shares ($\Delta = \frac{fu - f_d}{S_0 u - S_0 d}$ is the Delta hedge found in binomial trees)

which changes continuously over time. Let Π be the value of the portfolio then

$$\Pi = -f + \frac{\partial f}{\partial S}$$

and $\Delta\Pi$ be the value of the portfolio in the time interval Δt then

$$\Delta\Pi = -\Delta f + \frac{\partial f}{\partial S}\Delta S$$

Derivation of Black-Scholes-Merton Differential Equation

Then portfolio has no stochastic part! Simple arbitrage argument and risk-neutral growth implies

$$\frac{\partial f}{\partial t} + rS\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} = rf \tag{1}$$

Equation (1) is the Black-Scholes partial differential equation. Any solution corresponds to the price of a derivative overlying a particular stock.

In order to specify further what the derivative is, we use a boundary condition to constrain it.

Boundary conditions for European call options:

$$f = \max\{S - K, 0\}$$

when $t = T$. Boundary conditions for European put options:

$$f = \max\{K - S, 0\}$$

when $t = T$. The portfolio created is riskless only for infinitesimally short periods.

Black-Scholes Pricing Formulas

The Black-Scholes formulas for the price at time 0 of a European call option on a non-dividend-paying stock and for a European put option on a non-dividend paying stock are

$$c = S_0 N(d_1) - K e^{-rT} N(d_2)$$

and

$$p = K N(-d_2) - S_0 e^{-rT} N(-d_1)$$

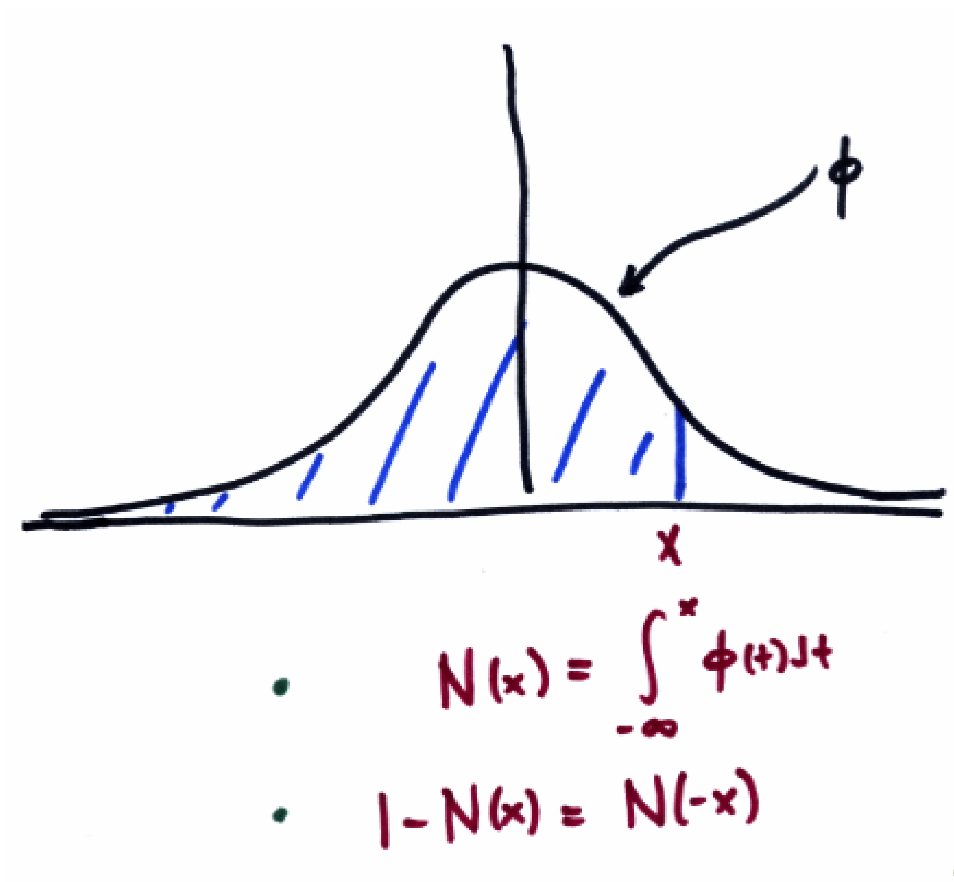
where

$$d_1 = \frac{\ln \frac{S_0}{K} + \left(r + \frac{\sigma^2}{2}\right) T}{\sigma \sqrt{T}}$$

$$d_2 = \frac{\ln \frac{S_0}{K} + \left(r - \frac{\sigma^2}{2}\right) T}{\sigma \sqrt{T}} = d_1 - \sigma \sqrt{T}$$

and $N(x)$ is the cumulative probability distribution function.

Black-Scholes Pricing Formulas



The variables c and p are the European call and put prices, S_0 is the current stock price at time 0, K is the strike price, r is the continuously compounded risk-free rate, σ is the stock price volatility, and T is the time to maturity of the option. Why?

Black-Scholes Formula for Option on stock with dividend yield

The price c of a European call and price p of a European put on a stock providing a dividend yield at rate q as

$$c = S_0 e^{-qT} N(d_1) - Ke^{-rT} N(d_2)$$

and

$$p = Ke^{-rT} N(-d_2) - S_0 e^{-qT} N(-d_1).$$

Since $\ln \frac{S_0 e^{-qT}}{K} = \ln \frac{S_0}{K} - qT$ then d_1 and d_2 are

$$d_1 = \frac{\ln \frac{S_0}{K} + \left(r - q + \frac{\sigma^2}{2}\right) T}{\sigma\sqrt{T}}$$

$$d_2 = \frac{\ln \frac{S_0}{K} + \left(r - q - \frac{\sigma^2}{2}\right) T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}$$

and $N(x)$ is the cumulative probability distribution function.

The associated Black-Scholes equation can be derived and is

$$\frac{\partial f}{\partial t} + (r - q)\, S \frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} = rf$$

This can be solved in a similar fashion as we solved the Black-Scholes equation last time...

Currency Options

Let S_0 denote the spot exchange rate. S_0 is the value of one unit of the foreign currency in US dollars.

Black-Scholes formulas with dividend yield

$$c = S_0 e^{-r_f T} N(d_1) - K e^{-rT} N(d_2)$$

and

$$p = K e^{-rT} N(-d_2) - S_0 e^{-r_f T} N(-d_1).$$

with d_1 and d_2 as

$$d_1 = \frac{\ln \frac{S_0}{K} + \left(r - r_f + \frac{\sigma^2}{2}\right) T}{\sigma \sqrt{T}}$$

$$d_2 = \frac{\ln \frac{S_0}{K} + \left(r - r_f - \frac{\sigma^2}{2}\right) T}{\sigma \sqrt{T}} = d_1 - \sigma \sqrt{T}$$

Note that $F_0 = S_0 e^{(r-r_f)T}$ then we can rewrite the equations as:

$$c = e^{-rT}\left[F_0 N(d_1) - KN(d_2)\right]$$

and

$$p = e^{-rT}\left[KN(-d_2) - S_0 N(-d_1)\right].$$

with d_1 and d_2 as

$$d_1 = \frac{\ln\frac{F_0}{K} + \frac{\sigma^2}{2}T}{\sigma\sqrt{T}}$$

$$d_2 = \frac{\ln\frac{F_0}{K} + -\frac{\sigma^2}{2}T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}$$

Naked & Covered Positions

- In a **naked position** the investor does nothing to hedge against losses.
  In our example, this approach does well so long as the stock remains below \$50. Then
- Alternatively the investor house can take a **covered position**. This involves buying 100,000 shares as soon as the option has been sold. If the option is exercised, the strategy works well. If the stock drops then there is a large loss. By the put-call parity this is similar to

$$c + Ke^{-rT} = p + S_0 \Longrightarrow -c + S_0 = -p + Ke^{-rT}$$

  so it is the same as writing a put option. Therefore, the covered position is bad if the stock price goes down.

Stop-Loss Strategy

The **stop-loss strategy** involves the following:

- Consider a bank that has written a call option with strike price K.
- The bank buys one unit of stock as soon as the price rises above K and selling it as soon as its price is less than K.
- Point is to hold a naked position whenever the stock is less than K and a covered position whenever the stock price is greater than K.
- The scheme is designed to ensure that at time T the institution owns the stock if the option closes in the money and does not own it if the option closes out of the money.
- Strategy seems to produce payoffs that are the same as the payoffs on the option.

Two problems:

- Cash flows to the hedger occur at different times and must be discounted
- Purchases and sales cannot be made at exactly the same price K. Crucial point...If the stock purchases are made at $K+\varepsilon$ and sold at $K-\varepsilon$ then every purchase and sale incurs a loss of 2ε.

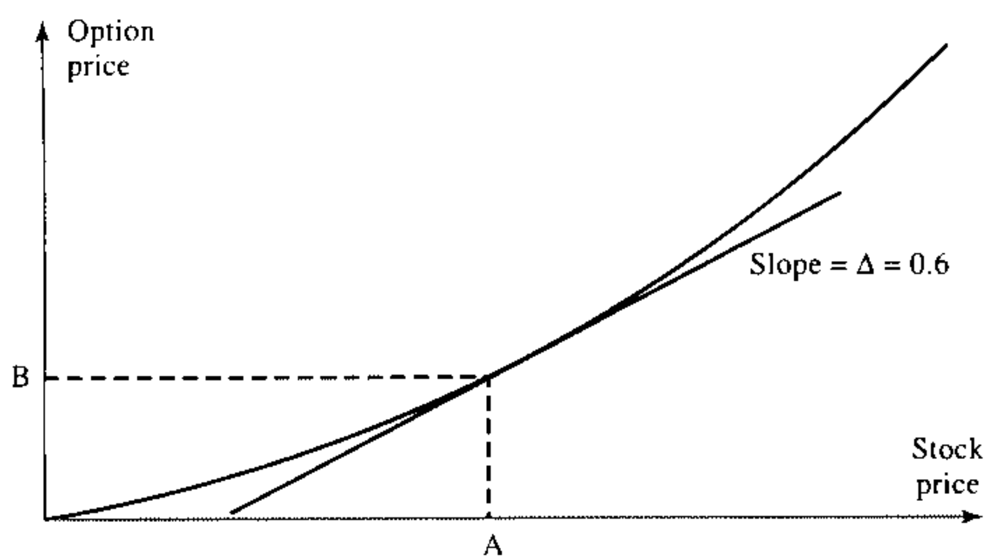
  If the stock prices change continuously (as it is modeled on a Brownian motion) then we expect the curve S to cross our line $S=K$ an **infinite** number of times! Our profit will go away due to excessive number of transactions.

Delta Hedging

Instead of designing a portfolio with a stop-loss strategy, a different strategy is to design a **delta hedge**.

- Recall that Δ of an option is the rate of change of the option price with respect to the price of the underlying asset.
- We have that $\Delta = \frac{\partial c}{\partial S}$ where c is the price of the call option and S is the stock price.

Figure 15.2 Calculation of delta.



Delta Hedging, cont.

Since delta changes over time, the investor's position remains delta hedged (delta neutral) for relatively short periods of time.

A hedge is **rebalanced**, or adjusted periodically, to remain delta neutral.

We will describe a **dynamic-hedging scheme** that rebalances the portfolio periodically to ensure a delta-neutral portfolio.

This is in contrast to **static hedging schemes** where the hedge is set up and left alone. Such schemes are called **hedge and forget** schemes.

We will use Black-Scholes analysis to help devise a good delta hedge scheme. Recall that the Black-Scholes portfolio that is riskless is

$$\begin{aligned} -1 &: \text{ option} \\ +\Delta &: \text{ shares of stock} \end{aligned}$$

Delta of European Stock Options

A European call option on a non-dividend-paying stock is

$$\Delta(call) = N(d_1)$$

and a European put option on a non-dividend-paying stock is

$$\Delta(put) = N(d_1) - 1$$

Keeping a delta hedge for a long position in a European call option involves maintaining a short position of $N(d_1)$ shares at any given time.

Note that the Δ for a European put option is negative,so that the a long position in a put option should be hedged with a long position in the underlying stock, and a short position in a put option should be hedged with a short position in the underlying stock.

Delta of Other European Options

For European call options on an asset paying a yield q,

$$\Delta(call) = e^{-qT} N(d_1)$$

where d_1 is defined by

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

and for European puts

$$\Delta(put) = e^{-qT} \left[N(d_1) - 1\right]$$

for the same d_1.

If the asset is a currency, we replace q with r_f, the foreign risk-free interest rate. If the asset is a futures contract, they are correct with q equal to the risk-free interest rate r and $S_0 = F_0$ in the definition of d_1.

Dynamic Delta Hedging

Consider the operation of a delta hedging for our first example.

$$S_0 = 49 \quad K = 50 \quad r = 0.05 \quad \sigma = 0.20 \quad T = 0.3846 \quad \mu = 0.13$$

on 100,000 shares of stocks. The European call option has been written for \$300,000. The initial Δ is calculated:

$$d_1 = \frac{\ln \frac{S_0}{K} + \left(r + \frac{\sigma^2}{2}\right) T}{\sigma \sqrt{T}}$$

$$= \frac{\ln \frac{49}{50} + \left(0.05 + \frac{0.02^2}{2}\right) 0.3846}{0.2\sqrt{0.3846}}$$

and

$$N(d_1) = 0.522$$

- Once option has been written, the investor has to buy

$$100,000 \times 0.522 = 52,200 \text{ shares}$$

  for a cost of $52,200 \times 49 = 2,557,800$. The interest rate is 5%, so after one week the interest costs

$$2,557,800e^{0.05 \times \frac{1}{52}} = 2,500$$

- Suppose now that the stock drops to \$48.12. The delta declines to 0.458. The hedge needs to be

$$100,000 \times 0.458 = 45,800 \text{ shares}$$

  Therefore, the bank needs to sell 6400 shares. The strategy realizes \$308,000 in cash, and the borrowings become $2,557,800 + 2,500 - 308,000 = 2,252,300$. The interest over this period is

$$2,252,300 \times e^{0.05 \times \frac{1}{52}} = 2,200$$

- And on...

Table 15.2 Simulation of delta hedging. Option closes in the money and cost of hedging is $263,300.

Week	*Stock price*	*Delta*	*Shares purchased*	*Cost of shares purchased ($000)*	*Cumulative cost including interest ($000)*	*Interest cost ($000)*
0	49.00	0.522	52,200	2,557.8	2,557.8	2.5
1	48.12	0.458	(6,400)	(308.0)	2,252.3	2.2
2	47.37	0.400	(5,800)	(274.7)	1,979.8	1.9
3	50.25	0.596	19,600	984.9	2,966.6	2.9
4	51.75	0.693	9,700	502.0	3,471.5	3.3
5	53.12	0.774	8,100	430.3	3,905.1	3.8
6	53.00	0.771	(300)	(15.9)	3,893.0	3.7
7	51.87	0.706	(6,500)	(337.2)	3,559.5	3.4
8	51.38	0.674	(3,200)	(164.4)	3,398.5	3.3
9	53.00	0.787	11,300	598.9	4,000.7	3.8
10	49.88	0.550	(23,700)	(1,182.2)	2,822.3	2.7
11	48.50	0.413	(13,700)	(664.4)	2,160.6	2.1
12	49.88	0.542	12,900	643.5	2,806.2	2.7
13	50.37	0.591	4,900	246.8	3,055.7	2.9
14	52.13	0.768	17,700	922.7	3,981.3	3.8
15	51.88	0.759	(900)	(46.7)	3,938.4	3.8
16	52.87	0.865	10,600	560.4	4,502.6	4.3
17	54.87	0.978	11,300	620.0	5,126.9	4.9
18	54.62	0.990	1,200	65.5	5,197.3	5.0
19	55.87	1.000	1,000	55.9	5,258.2	5.1
20	57.25	1.000	0	0.0	5,263.3	

In this simulation, the stock price climbs.

As it becomes evident that the option will be exercised at the maturity date, and delta approaches 1.0.

By week 20, the hedger has a fully covered position. The hedger receives $5 million for the stock held, so that the total cost of writing the option and hedging it is $263,300.

On the other hand consider a sequence of events such that the option closes out of the money. As it becomes clear that the option will not be exercised, delta approaches zero.

Delta of a Portfolio

The delta of a portfolio of options or other derivatives dependent on a single asset whose price is S is

$$\frac{\partial \Pi}{\partial S}$$

where Π is the value of the portfolio.

- The delta of the portfolio can be calculated from the deltas of the individual options in the portfolio. If a portfolio consists of a quantity w_i of option, the delta of the portfolio is given by

$$\Delta = \sum_{i=1}^{n} w_i \Delta_i$$

  where Δ_i is the delta of the ith option.
- The formula can be used to calculate the position in the underlying asset necessary to make the delta of hte portfolio zero. When this position has been taken, the portfolio is referred to as being **delta neutral**

Gamma

The **gamma**, Γ of a portfolio of options on an underlying asset is the rate of change of the portfolio's delta with respect to the price of the underlying asset.

It is the second partial derivative of the portfolio with respect to asset price:

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

If gamma is small, delta changes slowly, and adjustments to keep a portfolio delta neutral need to be made only relatively infrequently.

For European call options on non-dividend-paying stocks, we have

$$\Gamma = \frac{N'(d_1)}{S\sigma\sqrt{T-t}}$$

Making a Portfolio Gamma Neutral

- Suppose that a delta-neutral portfolio has a gamma equal to Γ, and a traded option has a gamma equal to Γ_T.
- If the number of traded options added to the portfolio is w_T, the gamma of the portfolio is

$$w_T\Gamma_T + \Gamma$$

- Including the traded option is likely to change the delta of the portfolio, so the position in the underlying asset cannot be changed continuously when delta hedging is used.
- Delta neutrality provides protection against larger movements in this stock price between hedge rebalancing

Vega

We set **vega** of a portfolio of derivatives, $\mathcal{V}$, is the rate of change of the value of the portfolio with respect to the volatility of the underlying asset:

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

- If vega is high in absolute terms, then the portfolio's value is very sensitive to small changes in volatility.
- If vega is low in absolute terms, then volatility changes have little impact on the value of the portfolio.

Note that a position in the underlying asset has zero vega, but the vega of a portfolio can be changed by adding a position in a traded option.

- If $\mathcal{V}$ is the vega of the portfolio and $\mathcal{V}_T$ is the vega of a traded option, a position of $-\frac{\mathcal{V}}{\mathcal{V}_T}$ in the traded option makes the portfolio instantaneously vega neutral.
- However, a portfolio that is gamma neutral will not generally be vega neutral and vice-versa.
- If a hedger requires a portfolio to be both gamma and vega neutral, at least two traded derivatives dependent on the underlying asset must usually be used.

Calculating vega

For a European call or put on a non-dividend-paying stock, vega is given by

$$\mathcal{V} = S_0\sqrt{T}N'(d_1)$$

where $d_1 = \frac{\ln\frac{S}{K}+(r+\frac{\sigma^2}{2})T}{\sigma\sqrt{T}}$.

For a European call or put on a dividend-paying stock with yield q, the vega is

$$\mathcal{V} = S_0\sqrt{T}N'(d_1)e^{-qT}$$

where $d_1 = \frac{\ln\frac{S}{K}+(r-q+\frac{\sigma^2}{2})T}{\sigma\sqrt{T}}$.

When the asset is a stock index, q is the dividend yield. When it is a currency contract, then set q to be the risk-free foreign rate r_f. When it is a futures contract, $S_0 = F_0$ and $q = r$.

Synthetic Puts

In general a portfolio manager wishes to acquire a put option to protect against large declines while achieving gains if the market appreciates.

One approach is to buy put options on a market index. Another approach is to create the put **synthetically**.

To create a synthetic put option, one maintains a position in the underlying asset so that the delta of the position is equal to the delta of the required option.

This can be more attractive than buying the put from the market:

- Options markets do not always have the liquidity to absorb trades that managers of large funds would like to have access to.
- Fund managers often require strike prices and exercise dates that are different from those available from the exchange-traded markets.

How to synthetically create the put?

The option can be created by trading the portfolio or by trading in index futures contracts.

- To create the put option synthetically, a fund manager should ensure that at any give time a proportion

$$e^{-qT}\left[N(d_1) - 1\right]$$

  of the stocks in the original portfolio has been sold and the proceeds invested in riskless assets. As the value of the original portfolio declines, the delta of the put becomes more negative and the proportion of the original portfolio sold must be increased.

  As the value of the original portfolio increases, the delta of the put becomes less negative and the proportion of the original portfolio sold must be decreased (and shares purchased)
- This strategy to create portfolio insurance entails dividing funds between the stock portfolio on which the insurance is required and riskless assets.
- As the value of the stock portfolio increases, riskless assets are sold and the position in the stock portfolio is increased.
- As the value of the stock portfolio decreases, the position in the stock portfolio is decreased and riskless assets are purchased.
- Insurance costs arise as the fact that selling occurs after a decline in the market and buying occurs after a rise in the market.

Volatility Smiles

How close are market prices to those predicted by Black-Scholes? Are Black-Scholes formulas used to price options?

Not entirely. Traders typically allow for volatility to depend on **price strike price and time to maturity**.

Plot of implied volatility of an option as a function of strike price is known as a **volatility smile**

Implied Volatility

There are two ways to think about volatility:

- From price changes, we can compute the volatility via standard deviation.
- Another method is to consider data used in Black-Scholes:

$$c = S_0 N(d_1) - K e^{-rT} N(d_2)$$

with

$$d_1 = \frac{\ln \frac{S_0}{K} + \left(r + \frac{\sigma^2}{2}\right) T}{\sigma \sqrt{T}}$$

$$d_2 = \frac{\ln \frac{S_0}{K} + \left(r - \frac{\sigma^2}{2}\right) T}{\sigma \sqrt{T}} = d_1 - \sigma \sqrt{T}$$

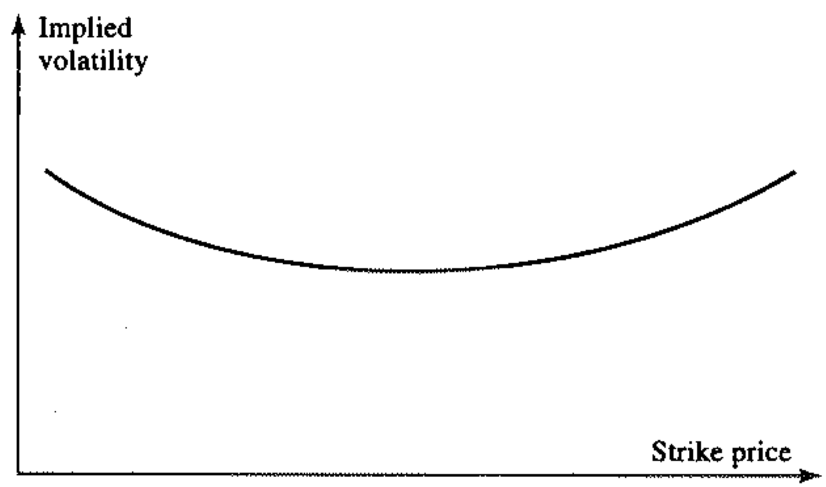
If we know S_0, K, r, and c, then we can solve, implicitly, for σ.
The resulting σ is called the **implied volatility**.

Foreign Currency Options

We now consider our first **volatility smile**. This is a graph of volatility as a function of strike price. We assumed for Black-Scholes that this is a constant function...

On the other hand traders use the following volatility smile

Figure 16.1 Volatility smile for foreign currency options.



- Volatility is relatively low for at-the-money options.
- Volatility is relatively high for the more in-the-money or out-of-the-money the strike price is.

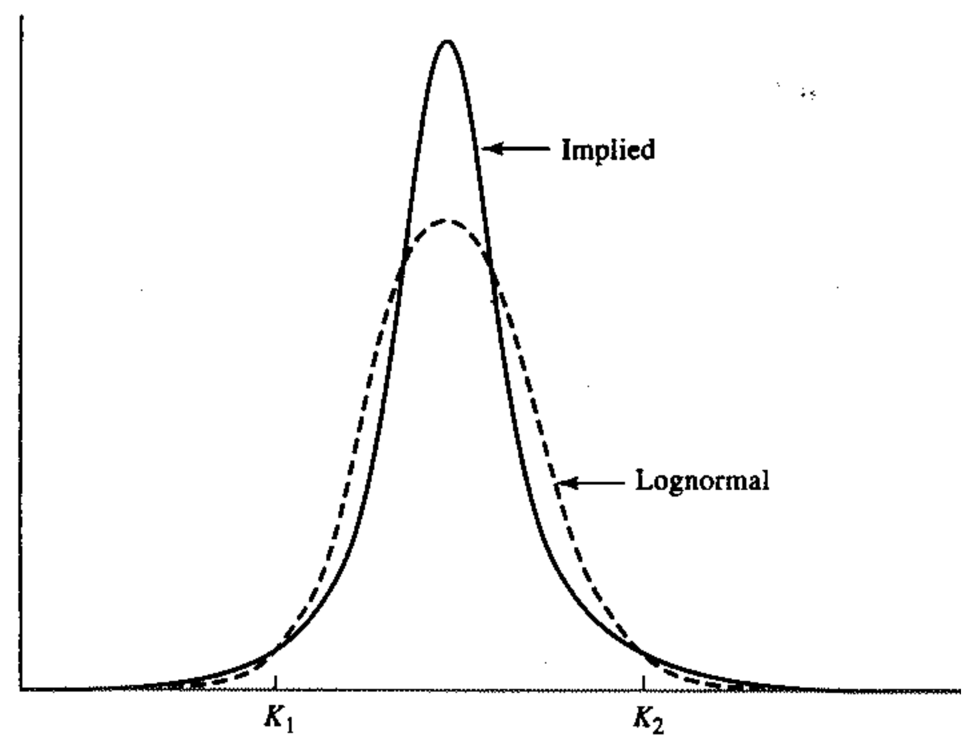
Volatility Smile for Foreign Currency Options

The associated probability distribution should no longer be lognormal, since the crucial ingredient to S being lognormal was

$$\frac{dS}{S} = \mu dt + \sigma \epsilon dz$$

Now σ is a function....

The **implied distribution** turns out to be



The distribution with the same mean and same standard deviation has

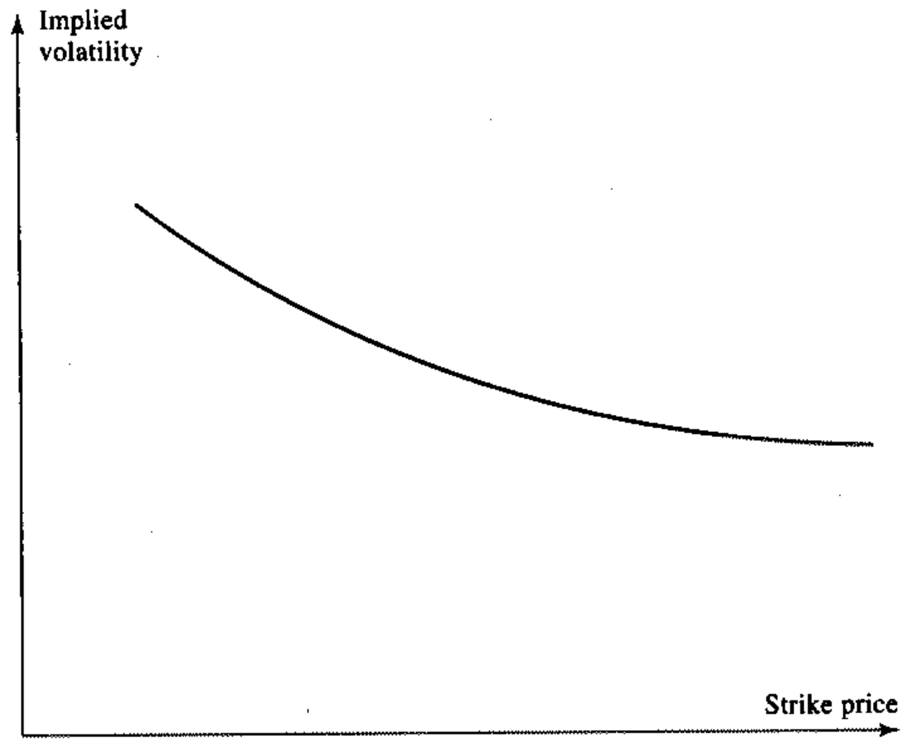
- **fat tails**
- **steeper**

Volatility Smiles for Equity Options

Before the crash of 1987, stocks were generally assumed to follow the lognormal distribution.

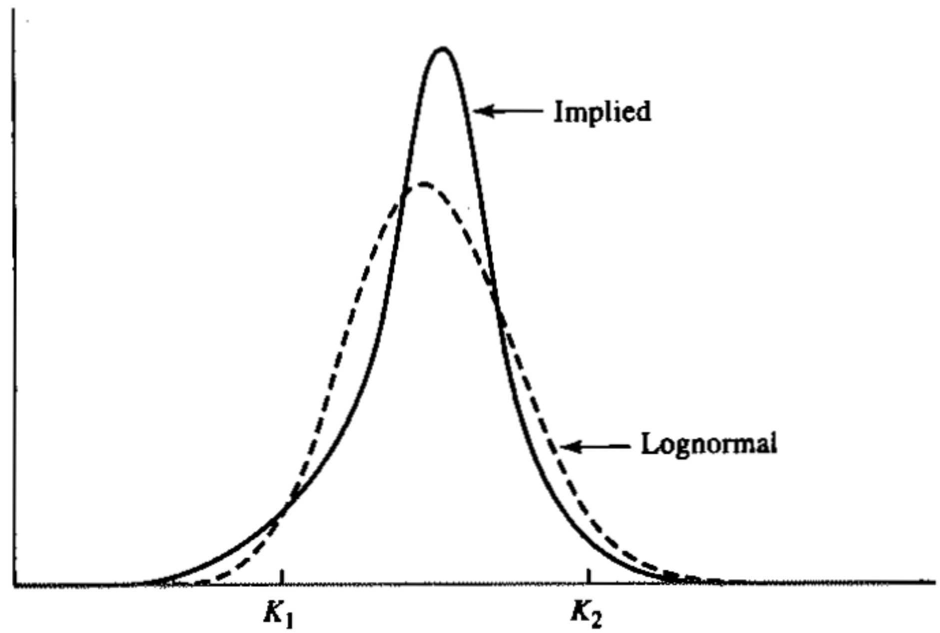
After the crash, a volatility smile for equity options was introduced by Rubinstein and Jackwerth-Rubinstein.

Figure 16.3 Volatility smile for equities.



The volatility smile or **volatility skew**, has the form of a downward sloping parabola.

- Volatility to price a **low-strike-price** option (deep-out-of-the-money put or deep-in-the-money call) is significantly higher than that used to price a **high-strike-price** option (deep-in-the-money put or deep-out-of-the-money call).
- The volatility smile for equity options corresponds to the implied probability distribution given by below:



compared to the corresponding lognormal distribution.

Volatility Surfaces

Traders also consider the volatility term structure when pricing options.

In other words the volatility used to price an at-the-money option depends on the maturity of the option.

- Volatility tends to be an increasing function of maturity when short-dated volatilities are historically low, since there is expectation that volatility will increase.
- Volatility tends to be a decreasing function of maturity when short-dated volatilities are historically high, since there is expectation that volatility will decrease.

Basic Numerical Procedures: Semester II

- Generalized Binomial Trees
- Monte Carlo Methods for Black-Scholes
- Finite Difference Methods

Binomial Trees

Black-Scholes theory provides exact formulas for the pricing of European options under ideal situations.

American options do not have such a nice representation. Binomial trees are very useful for such derivatives:

Recall the Binomial Tree setup:

- Assume that in a short period of time Δt the stock either rises to Su or drops to Sd.
- Thus $u > 1$ and $d < 1$. The probability of the up movement is p and the probability of a down movement is $1 - p$.
- Assume that the world is **risk neutral**:
  - Assume that the expected return from all traded assets is the risk-free interest rate
  - Value payoffs from the derivative by calculating their expected values and discounting at the risk-free interest rate.

- The values p, u, d must give correct values for the mean and variance of asset price changes during the time interval Δt.
- Since we assume the risk-neutral world hypothesis holds, the expected return from the asset is the risk-free interest rate, r.
- Suppose that the asset provides a yield of q, then the expected return of the capital gains must be $r - q$. Therefore, the expected value of the asset price at the end of the time interval of length Δt becomes

  $$Se^{(r-q)\Delta t}$$

  where S is the value of the stock at the start of the time period.

Therefore, we find the **expected value of capital gains increase** is

$$Se^{(r-q)\Delta t} = pSu + (1-p)Sd$$

or

$$e^{(r-q)\Delta t} = pu + (1-p)d \tag{2}$$

independent of the stock price. Here we used the Mean Growth to determine a relationship between p, u, d.

- We now use the volatility. The variance of a variable $\operatorname{var} Q = E(Q^2) - E(Q)^2$.
- The variance over a time interval Δt then $\sigma^2 \Delta t$. We compute

$$\sigma^2 \Delta t = pu^2 + (1-p)d^2 - e^{2(r-q)\Delta t}$$

From (2) we find $p = \frac{a-d}{u-d}$ where $a = e^{(r-q)\Delta t}$, we see

$$\sigma^2 \Delta t = e^{(r-q)\Delta t}(u+d) - ud - e^{2(r-q)\Delta t} \tag{3}$$

We get two conditions on u, d, p from (2) and (3)

- Finally we impose a simplifying condition $d = \frac{1}{u}$. Ignoring higher-order terms in Δt we get

$$p = \frac{a-d}{u-d} \tag{4}$$

$$u = e^{\sigma\sqrt{\Delta t}} \tag{5}$$

$$d = e^{-\sigma\sqrt{\Delta t}} \tag{6}$$

$$a = e^{(r-q)\Delta t} \tag{7}$$

a is the growth factor.

- For a one step tree we evaluate

$$f = e^{-r\Delta t}\left[p f_u + (1-p) f_d\right]$$

where $e^{-r\Delta t}$ is the discounting factor on the time-step.

- Recursively, we check at at the end time $f_{N,j} = \max\{K - S_0 u^j d^{N-j}, 0\}$. We then evaluate the worth recursively.

- More after an example:

Expressing Binomial Trees Algebraically

- Recursively, we check at at the end time $f_{N,j} = \max\{K - S_0 u^j d^{N-j}, 0\}$. We then evaluate the worth recursively.
- At some node (i, j) there is a probability p of moving from (i, j) at time $i\Delta t$ to $(i + 1, j + 1)$ at time $(i + 1)\Delta t$, and a probability $(1 - p)$ of moving from $(i + 1, j - 1)$ at time $(i + 1)\Delta t$.
- If no early exercise then the risk-neutral value is

$$f_{i,j} = e^{-r\Delta t}\left[p f_{i+1,j+1} + (1 - p) f_{i+1,j-1}\right]$$

  for $0 \leq 1 \leq N - 1$ and $0 \leq j \leq i$.
- If early exercise is taken into account then we get

$$f_{i,j} = \max\{K - S_0 u^j d^{i-j}, e^{-r\Delta t}\left[p f_{i+1,j+1} + (1 - p) f_{i+1,j-1}\right]\}$$

In the limit as $\Delta t \to 0$, we get the true price for an American put option.

Estimating the Greek letters

We now compute the Delta from the Binomial Tree:

Recall that

$$\Delta = \frac{\partial f}{\partial S} \approx \frac{\Delta f}{\Delta S}$$

Consider a two node tree, then the change in price at the next time step is $\Delta S = S_0 u - S_0 d$. The change in the price of the option is $f_{1,1} - f_{1,0}$.

Thus

$$\Delta = \frac{f_{1,1} - f_{1,0}}{S_0 u - S_0 d} \tag{8}$$

At a later node, we compute the discrete Δ hedge parameter

$$\Delta_{j,k} = \frac{f_{j+1,k+1} - f_{j+1,k}}{S_0 u - S_0 d} \tag{9}$$

The next Greek letter we can compute is Γ. Since Γ is a second derivative, we use the following finite difference scheme:

$$\Gamma = \frac{\partial^2 c}{\partial S^2} \approx \frac{\Delta_{j+1,k+1} - \Delta_{j+1,k}}{h}$$

where h is the difference in stock price at the second level.

Expanding out yields:

$$\Gamma = \frac{\left[\frac{f_{2,2}-f_{2,1}}{S_0u^2-S_0}\right] - \left[\frac{f_{2,1}-f_{2,0}}{S_0-S_0d^2}\right]}{h} \tag{10}$$

with

$$h = \frac{S_0u^2 - S_0d^2}{2}. \tag{11}$$

The next Greek letter we can compute is Θ. Since $\Theta = \frac{\partial f}{\partial t}$ is a first derivative, it is not to difficult to compute:

$$\Theta = \frac{\partial f}{\partial t} \approx \frac{f_{2,1} - f_{0,0}}{2\Delta t} \tag{12}$$

since the stock price is the same at nodes $0, 0$ and $2, 1$.

To compute Vega $\mathcal{V}$, we need to compute $\frac{\partial f}{\partial \sigma}$. First compute the option price f with using a binomial tree with volatility σ and then compute the option price f^* using a binomial tree with volatility $\sigma + \Delta\sigma$. Then compute

$$\mathcal{V} = \frac{f^* - f}{\Delta\sigma} \tag{13}$$

Binomial trees for dividend-paying stock

- We consider the case of a known dollar dividend. This is harder than a known dividend yield.
- Assuming that the volatility is constant over the life of the option, the nodes will not recombine at later times:
- Suppose that there is one dividend at some date τ which is between $k\Delta t$ and $(k+1)\Delta t$ with a dollar amount D.
- When $i \leq k$ the nodes at time $i\Delta t$ correspond to stock prices

$$S_0 u^j d^{i-j} \qquad j = 0, \ldots, i$$

When $i = k+1$, the nodes on the tree correspond to stock prices

$$S_0 u^j d^{i-j} - D, \qquad j = 0, \ldots, i$$

When $i = k+2$, the nodes on the tree correspond to stock prices

$$\left(S_0 u^j d^{i-j} - D\right) u \text{ and } \left(S_0 u^j d^{i-j} - D\right) d$$

- We note that $Du \neq Dd$, so the nodes do not recombine in the middle. Therefore, there are $2i$ nodes instead of $i+1$.
- When $i = k+m$, there are $m(k+2)$ nodes instead of $k+m+2$.

We can simplify the analysis by separating the uncertain, stochastic part from the certain dividend part.

- Suppose that there is one dividend date τ during the life of the option. and

$$k\Delta t \leq \tau \leq (k+1)\Delta t$$

- The value of the uncertain component S^* at time $i\Delta t$ is
  - $S^* = S - De^{-r(\tau - i\Delta t)}$ when $i\Delta t < \tau$
  - $S^* = S$ when $i\Delta t > \tau$

  where D is the dividend. Define σ^* to be the volatility of S^* which is assumed to be constant.
- We compute p, u, d using σ^* and S^*. We add the present value of any future dividends (if any) to determine S from S^*.

- Suppose S_0^* is the value of S^* at time zero. At time $i\Delta t$, the nodes on the tree correspond to

$$S_0^* u^j d^{i-j} + De^{-r(\tau - i\Delta t)}, \qquad j = 0, \dots, i$$

  when $i\Delta t < \tau$ and

$$S_0^* u^j d^{i-j}$$

- The tree at time $i\Delta t$ has $i+1$ nodes.

Monte Carlo and Derivatives

We can use Monte Carlo to offer a risk-neutral valution by computing sample paths. Consider a derivative dependent on a single market variable S that provides a payoff at time T.

1. Sample a random path for S in a risk-neutral world.
2. Calculate the payoff from the derivative
3. Repeat steps 1 and 2 to get many sample values of the payoff from the derivative in a risk-neutral world.
4. Calculate the mean of the sample payoffs to get an estimate of the expected payoff in a risk-neutral world.
5. Discount the expected payoff at the risk-free rate to get an estimate of the value of the derivative

Checking Black-Scholes

One can numerically check the veracity of the Black-Scholes formula. How?

We are given constants S_0, K, r, σ, T that we can use in the Black-Scholes formula.

- Compute the stock process via a Monte Carlo method:

$$S(T) = S(0)\exp\left[\left(\mu - \frac{\sigma^2}{2}\right)T + \sigma\epsilon\sqrt{T}\right]$$

  by choosing a sample from the standard normal distribution.
- Given $S(T)$ we evaluate the option value as

$$e^{-rT}\max\{S(T) - K, 0\}$$

- We repeat the procedure a number of trials and average the value.

Number of Trials

- Accuracy of the result given by Monte Carlo simulation depends on the number of trails.
- Usually one calculates the standard deviation and the mean of the discounted payoffs given by the simulation trials.
- Denote μ and ω to be the mean and standard deviation, and we assume μ is the price of the derivative at the end of the simulated trial.
- The standard error of the estimate of the price of the derivative is given by

$$\frac{\omega}{\sqrt{M}}$$

  where M is the number of trials.
- A 95% confidence interval for the price f of a derivative is given by

$$\mu - 1.96\frac{\omega}{\sqrt{M}} < f < \mu + 1.96\frac{\omega}{\sqrt{M}}$$

Greek letters

In order to compute the Greek letters from a Monte Carlo simulation, one needs to compute the partial derivative of f with respect to a derivative. We consider the approximate derivative of f with respect to x.

- First compute the Monte Carlo simulation in the usual way to calculate $\hat{f}$ with a fixed value of x.
- Second compute the value of the derivative $\hat{f}^*$ with a new $x + \Delta x$.
- Third compute

$$\frac{\hat{f}^* - \hat{f}}{\Delta x}$$

- In order to minimize the standard error of the estimate, the number of intervals N, the number of random streams, and the number of trials M should be the same for calculating both $\hat{f}$ and $\hat{f}^*$.

Sampling through a Tree

Instead of implementing Monte Carlo simulation by randomly sampling from the stochastic process for the underlying variable, we can use an N-step binomial tree and sample from the 2^N paths that are possible.

- Suppose we have a binomial tree here the probability of an up-movement is 0.6. The procedure for sampling a random path through the tree is as follows.
- At each node, we sample a random number between 0 and 1. If the number is less than 0.4, we take the down path.
- Once we have a complete path from the initial node to the end of the tree, we can calculate a payoff.
- This completes a first trial. Similar procedure is used to complete more trials. The mean of the payoffs is discounted at the risk-free rate to get an estimate of the value of the derivative.

Finite Difference Schemes

Finite difference methods are useful for solving partial differential equations. The differential equation is converted into a set of **difference equations** that are solved iteratively.

Consider how we might value an American put option on a stock paying a dividend yield of q. The differential equation that the option must satisfy is the associated Black-Scholes equation

$$\frac{\partial f}{\partial t} + (r - q)\, S\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2\frac{\partial^2 f}{\partial S^2} = rf$$

- Suppose the life of the option is T. Divide this into N equally spaced intervals of length $\Delta t = \frac{T}{N}$. A total of $N + 1$ times are therefore considered

$$0, \Delta t, 2\Delta t, \ldots, T$$

- Suppose that S_{max} is a stock price sufficiently high that (if reached) the put has virtually no value. Define $\Delta S = \frac{S_{max}}{M}$ and consider a total of $M + 1$ equally spaced stock price:

$$0, \Delta S, 2\Delta S, \ldots, S_{max}$$

Choose S_{max} so that a $k\Delta S$ is the current stock price.

- The time points and stock price points define a grid consisting of $(N+1)\times(M+1)$ points. The point (i,j) on the grid is the point that corresponds to time $i\Delta t$ and stock price $j\Delta S$.
- Use the discrete variable $f_{i,j}$ to denote the value of the option on the grid point (i,j).

We can write down the finite difference scheme for the Black-Scholes, using $S = j\Delta S$, to get

$$\begin{aligned}
f_{i+1,j} &= \left(\frac{1}{2}(r-q)j\Delta t - \frac{1}{2}\sigma^2 j^2 \Delta t\right) f_{i,j-1} \\
&\quad + \left(1 + \sigma^2 j^2 \Delta t + r\Delta t\right) f_{i,j} \\
&\quad + \left(-\frac{1}{2}(r-q)j\Delta t - \frac{1}{2}\sigma^2 j^2 \Delta t\right) f_{i,j+1} \\
&= a_j f_{i,j-1} + b_j f_{i,j} + c_j f_{i,j+1}
\end{aligned}$$

We now choose boundary conditions for our problem.

Next the value of the put at time T is $\max\{K - S_T, 0\}$, where S_T is the stock price at time T.

$$f_{N,j} = \max\{K - j\Delta S, 0\} \qquad j = 0, \ldots, M$$

The value of the put option when the stock price is zero is K. Hence

$$f_{i,0} = K \qquad i = 0, \ldots, N$$

The value of the put option when the stock price is S_{max} is zero. Hence

$$f_{i,M} = 0 \qquad i = 0, \ldots, N$$

Finite Difference Cont.

We now solve for the rest of the $f_{i,j}$'s. We know $f_{N,j}$ then our equation yields equations

$$a_j f_{N-1,j-1} + b_j f_{N-1,j} + c_j f_{N-1,j+1} = f_{N,j}$$

for $j = 1, \ldots, M-1$. The right hand sides are known from the boundary condition:

$$f_{N-1,0} = K \qquad f_{N-1,M} = 0$$

- Therefore, we have $M-1$ **linear** equations for $M-1$ unknowns. This can be solved easily to get
$$f_{N-1,1}, f_{N-1,2}, \ldots, f_{N-1,M-1}$$
- We now check whether $f_{N-1,j}$ is optimal. If $f_{N-1,j} < K - j\Delta S$ then we should exercise early and $f_{N-1,j}$ is reassigned the value $K - j\Delta S$.
- Once $T - \Delta t$ has been evaluated at points $(N-1, j)$, we move to the points on the grid referring to $T - 2\Delta t$.
- Finally, we get the grid points $f_{0,1}, f_{0,2}, \ldots, f_{0,M-1}$. We choose the point that is

Explicit Finite Difference Method

The implicit finite difference scheme is very robust, and as Δt and $\Delta S \to 0$ then the solution goes to the solution of the Black-Scholes.

On the other hand implicit finite difference requires the solution of a set of equations at each fixed time.

We can do a simpler method that doesn't require solving the system of equations at each time step.

- Assume that $\frac{\partial f}{\partial S}$ and $\frac{\partial^2 f}{\partial S^2}$ is the same at (i, j) as at $(i+1, j)$. Then

$$\frac{\partial f}{\partial S} \approx \frac{f_{i+1,j+1} - f_{i+1,j-1}}{2\Delta S}$$

  and

$$\frac{\partial^2 f}{\partial S^2} \approx \frac{f_{i+1,j+1} - 2f_{i+1,j} + f_{i+1,j-1}}{(\Delta S)^2}$$

This yields a finite difference scheme

$$\begin{aligned} f_{i,j} &= \frac{1}{1+r\Delta t}\left(-\frac{1}{2}(r-q)j\Delta t + \frac{1}{2}\sigma^2 j^2 \Delta t\right) f_{i+1,j-1} \\ &\quad + \frac{1}{1+r\Delta t}\left(1-\sigma^2 j^2 \Delta t\right) f_{i,j} \\ &\quad + \frac{1}{1+r\Delta t}\left(\frac{1}{2}(r-q)j\Delta t + \frac{1}{2}\sigma^2 j^2 \Delta t\right) f_{i+1,j+1} \\ &= a_j^* f_{i+1,j-1} + b_j^* f_{i+1,j} + c_j^* f_{i+1,j+1} \end{aligned}$$

Since we know the information at the previous time, we can directly compute the value of $f_{i,j}$ without solving a system of equations.

Change of Variables Improvement

We can improve the efficiency of the finite difference methods by using $\ln S$ rather than S as the underlying variable. Setting $Z = \ln S$ then Black-Scholes becomes

$$\frac{\partial f}{\partial t} + \left(r - q - \frac{\sigma^2}{2}\right)\frac{\partial f}{\partial Z} + \frac{1}{2}\sigma^2\frac{\partial^2 f}{\partial Z^2} = rf$$

We discretize in equal Z steps, rather than for S steps. The difference equation for the implicit method becomes

$$\frac{f_{i+1,j} - f_{i,j}}{\Delta t} + \left(r - q - \frac{\sigma^2}{2}\right)\frac{f_{i,j+1} - f_{i,j-1}}{2\Delta Z} + \frac{1}{2}\sigma^2\frac{f_{i,j+1} - 2f_{i,j} + f_{i,j-1}}{(\Delta Z)^2} = rf_{i,j}$$

So:

We get the finite difference scheme

$$\alpha_j f_{i,j-1} + \beta_j f_{i,j} + \gamma_j f_{i,j+1} = f_{i+1,j}$$

where

$$\alpha_j = \frac{\Delta t}{2\Delta Z}\left(r - q - \frac{\sigma^2}{2}\right) - \frac{\Delta t}{2\left(\Delta Z\right)^2}\sigma^2$$

$$\beta_j = 1 + \frac{\Delta t}{\left(\Delta Z\right)^2}\sigma^2 + r\Delta t$$

$$\gamma_j = -\frac{\Delta t}{2\Delta Z}\left(r - q - \frac{\sigma^2}{2}\right) - \frac{\Delta t}{2\left(\Delta Z\right)^2}\sigma^2$$

The explicit finite difference scheme

$$\alpha_j^* f_{i+1,j-1} + \beta_j^* f_{i+1,j} + \gamma_j^* f_{i+1,j+1} = f_{i,j}$$

where

$$\alpha_j^* = \frac{1}{1+r\Delta t}\left[-\frac{\Delta t}{2\Delta Z}\left(r-q-\frac{\sigma^2}{2}\right)+\frac{\Delta t}{2\left(\Delta Z\right)^2}\sigma^2\right]$$

$$\beta_j^* = \frac{1}{1+r\Delta t}\left[1-\frac{\Delta t}{\left(\Delta Z\right)^2}\sigma^2\right]$$

$$\gamma_j^* = \frac{1}{1+r\Delta t}\left[\frac{\Delta t}{2\Delta Z}\left(r-q-\frac{\sigma^2}{2}\right)+\frac{\Delta t}{2\left(\Delta Z\right)^2}\sigma^2\right]$$

The change of variables approach has the property that α_j, β_j, γ_j as well as α_j^*, β_j^*, γ_j^*.

Most efficient if

$$\Delta Z = \sigma\sqrt{3\Delta t}$$

Value at Risk

We now look for a quantity that gives a measure of the total risk of a portfolio

Value at Risk (VaR) - attempts to provide a single number that summarizes the total risk in a portfolio of financial assets.

This is widely used by corporate treasurers and fund managers as well as by financial institutions. Central bank regulators use VaR in determining the capital the bank is required to keep to reflect the market risks it is bearing.

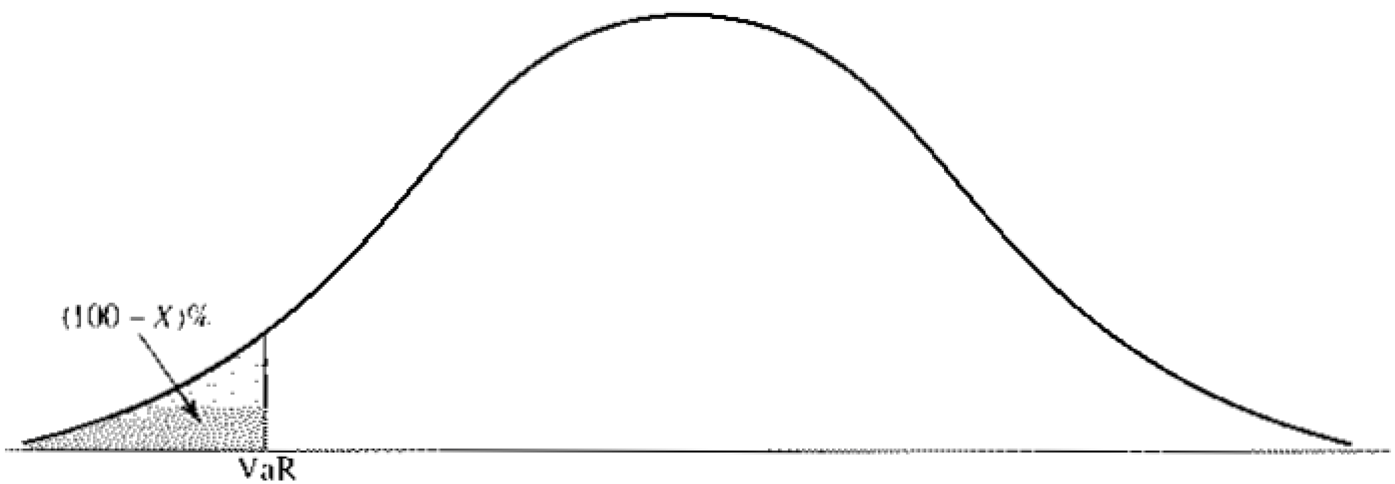
- The value-at-risk measure, we are interested is of the form

  We are X percent certain that we will not lose more than V dollars in the next N days.
- The variable V is the VaR of the portfolio
- The VaR is a function of two parameters - the time horizon N and the confidence level X. It measure the loss level over N days that we are X certain will not be exceeded.
- Bank regulators require that banks calculate VaR with $N = 10$ and $X = 99$.

VaR

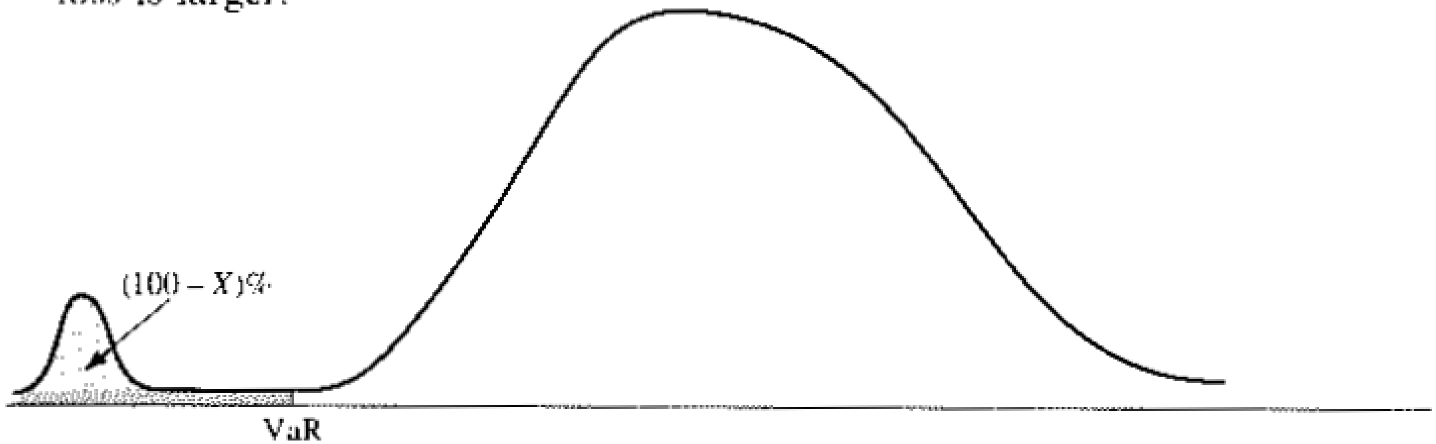
- VaR is attractive since it simply asks
  How bad can things get?
- When the value of the portfolio is normally distributed we see VaR looks like:

Figure 18.1 Calculation of VaR from the probability distribution of the change in the portfolio value; confidence level is $X\%$.



- There is a problem if the portfolio is not normally distributed. Consider a case where there is a larger probability of a very large down movement and less otherwise:

Figure 18.2 Alternative situation to Figure 18.1. VaR is the same, but the potential loss is larger.



- The two graphs have the same VaR, but the second is riskier, since there is a larger probability of a very large loss.

Some traders may look for a different measure of total risk, **Conditional VaR**.

- C-VaR asks

  If things do get bad, how much can we expect to lose?
- C-VaR is the expected loss during an N-day period conditional that we are in the $100 - X\%$ left tail of the distribution.
- VaR is a most popular measure than other such risk measures.

The Time Horizon

VaR has two parameters - the N-day time horizon, and the X confidence interval.

- In practice $N = 1$, since there is usually not enough data for a longer period.
- Usually one assumes
$$N - \text{day VaR} = 1 - \text{day VaR} \times \sqrt{N}$$
- Since bank's are required to have capital at least three times the 10-day 99% VaR, then it is required to have
$$3 \times \sqrt{10} = 9.49$$
times the 1-day 99% VaR.

Historical Simulations

Historical simulation is one approach of estimating VaR.

- Use past data in a direct way as a guide to what might happen in the future.
- Suppose that we wish to calculate the 99% confidence level with a 1-day horizon using 500 days of data.
  1. Identify the market variables affection the portfolio (typically **exchange rates, equity prices, interest rates,** etc).
  2. Collect data on the movements in these markets variables over the most recent 500 days.
  3. We have 500 alternative scenarios for what can happen between today and tomorrow
  4. Scenario 1 is where the percentage changes in the values of all variables are the same as they were on the first day for which we have collected data.
  5. Scenario 2 is where the percentage changes in the values of all variables are the same as they were on the second day for which we have collected data.
  6. Etc...
  7. For each scenario we calculate the dollar change in the value of the portfolio between today and tomorrow.
  8. This yields a probability distribution for daily changes in the value of our portfolio.

Model-Building Approach

A major approach outside of the historical approach is to use model approach.

Daily Volatilities

- Usually we measure volatilities in years. Model-building approach to VaR, we measure time in days and the volatility of an asset is usually quoted as "volatility per day"
- Define σ_{year} to be the volatility per year of a certain asset and σ_{day} as the equivalent volatility per day of the asset. Assuming 252 trading days per year, we use

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

or

$$\sigma_{day} = \frac{\sigma_{year}}{\sqrt{252}} \approx 0.063\sigma_{year}$$

- σ_{day} is approximately equal to the standard deviation of the percentage change in the asset price in one day. For purposes of calculating VaR we assume exact equality. We define the daily volatility of an asset price as equal to the standard deviation of the percentage change in one day.

Linear Model

To calculate the standard deviation of ΔP, we define σ_i as the daily volatility of the ith asset and ρ_{ij} as the coefficient of correlation between returns on asset i and asset j. This means that σ_i is the standard deviation of Δx_i and ρ_{ij} is the coefficient of correlation between Δ_i and Δx_j.

The variance of ΔP, denoted σ_P^2 is given by

$$\sigma_P^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_{ij} \alpha_i \alpha_j \sigma_i \sigma_j$$

or

$$\sigma_P^2 = \sum_{i=1}^{n} \alpha_i^2 \sigma_i^2 + 2 \sum_{i=1}^{n} \sum_{j<i} \rho_{ij} \alpha_i \alpha_j \sigma_i \sigma_j \quad (14)$$

The standard deviation of the change over N days is $\sigma_P \sqrt{N}$, and the 99% VaR for an N-day time horizon is $2.33 \sigma_P \sqrt{N}$.

Linear Model and Options

- Consider how a linear model can be used when there are options. Consider first a portfolio consisting of options on a single stock whose current price is S.
- Suppose that the delta of the position is δ. Since δ is the rate of change of the value of the portfolio with S, it is approximately true that

$$\delta = \frac{\Delta P}{\Delta S}$$

  or

$$\Delta P = \delta \Delta S$$

  where ΔS is the dollar change in the stock price in 1 day and ΔP is the dollar change in the portfolio in 1 day.
- We define Δx as the percentage change in the stock price in 1 day, so that

$$\Delta x = \frac{\Delta S}{S}$$

- It follows that an approximate relationship between ΔP and Δx is

$$\Delta P = S\delta\Delta x$$

- When we have several underlying market variables that includes options, we can derive an approximate linear relationship between ΔP and Δx similarly.

- The relationship is

$$\Delta P = \sum_{i=1}^{n} S_i \delta_i \Delta x_i$$

where S_i is the value of the ith market variable and δ_i is the delta of the portfolio with respect to the ith market variable. This yields

$$\Delta P = \sum_{i=1}^{n} \alpha_i \Delta x_i$$

with $\alpha_i = S_i \delta_i$.

Comparison of Approaches

Two methods

- **Historical Simulation**: historical data determines the joint probability distribution of the market variables. Advantage - avoids the need for cash-flow mapping. Disadvantages - computationally expensive and does not easily allow volatility updating schemes.
- **Model Building**: Advantages - results can be produced very quickly and can be used in conjunction with volatility updating schemes. Disadvantages - assumes that the market variables have a multivariate normal distribution (daily changes are usually not normally distributed) and gives poor results for small delta portfolios.

Estimating Volatilities and Correlations

We discuss how to use historical data to extract estimates on the current and future levels of volatilities and correlations.

Estimating Volatility: Define σ_n as the volatility of a market variable on day n, as estimated at the end of day $n-1$. The square of the volatility, σ_n^2, on day n is the **variance rate**. We described the standard approach to estimating σ_n from historical data.

Suppose that the value of the market variable at the end of day i is S_i. The variable u_i is defined as the continuously compounded return during day i

$$u_i = \ln \frac{S_i}{S_{i-1}}$$

An unbiased estimate of the variance rate per day, σ_n^2 using the most recent m observations on the u_i is

$$\sigma_n^2 = \frac{1}{m-1} \sum_{i=1}^{m} (u_{n-i} - \overline{u})^2$$

where $\overline{u}$ is the mean of the u_i's:

$$\overline{u} = \frac{1}{m} \sum_{i=1}^{m} u_{m-i}$$

For the purposes of monitoring daily volatility, the formula is usually changed in a number of ways

1. u_i is defined as the percentage change in the market variable between the end of the day $i-1$ and the end of day i so that

$$u_i = \frac{S_i - S_{i-1}}{S_{i-1}} \approx \ln(1 + \frac{S_i - S_{i-1}}{S_{i-1}}) \tag{15}$$

2. $\overline{u}$ is assumed to be zero
3. $m-1$ is replaced by m

These three changes make very little difference to the estimates that are calculated, but they allow us to simplify the formula for the variance rate to

$$\sigma_n^2 = \frac{1}{m} \sum_{i=1}^{m} u_{n-i}^2 \tag{16}$$

where u_i is given by (15).

Weighting Schemes

The sigma given by (16) gives equal weight to

$$u_{n-1}^2, u_{n-2}^2, \ldots, u_{n-m}^2$$

Our objective is to estimate the current level of volatility σ_n. Therefore, it makes sense to give more wieght to recent data. One such model is

$$\sigma_n^2 = \sum_{i=1}^{m} \alpha_i u_{n-i}^2 \tag{17}$$

The variable α_i is the amount of weight given to the observation i days ago. The α's are positive.

If we choose them so that $\alpha_i < \alpha_j$ when $i > j$, less weight is given to older observations. The weights must sum to unity, so we have

$$\sum_{i=1}^{m} \alpha_i = 1$$

Weighting Schemes, cont.

An extension of the idea, called **ARCH(m)** or **Autoregressive Conditional Heteroscedasticity** in equation (17) is to assume that there is a long-run average variance rate and that this should be given some weight. This leads to the model that takes the form

$$\sigma_n^2 = \gamma V_L + \sum_{i=1}^{m} \alpha_i u_{n-i}^2 \tag{18}$$

where V_L is the long-run variance rate and γ is the weight assigned to V_L. Because the weights must sum to unity, we have

$$\gamma + \sum_{i=1}^{m} \alpha_i = 1$$

Define $\omega = \gamma V_L$, the model equation becomes

$$\sigma_n^2 = \omega + \sum_{i=1}^{m} \alpha_i u_{n-i}^2 \tag{19}$$

Exponentially Weighted Moving Average Model

The **Exponentially Weighted Moving Average Model** or **EWMA** is a particular case of (17) where the weights α_i decrease exponentially as we move back through time.

Specifically $\alpha_{i+1} = \lambda \alpha_i$, where λ is a constant between 0 and 1.

The formula is

$$\sigma_n^2 = \lambda \sigma_{n-1}^2 + (1 - \lambda)\, u_{n-1}^2 \tag{20}$$

The estimate σ_n of the volatility of day n is calculated from σ_{n-1} and u_{n-1}.

We can see why this corresponds to exponentially decreasing weights. Continuing yields

$$\sigma_n^2 = (1 - \lambda) \sum_{i=1}^{m} \lambda^{i-1} u_{n-i}^2 + \lambda^m \sigma_{n-m}^2 \tag{21}$$

For large m, the term $\lambda^m \sigma_{n-m}^2$ is sufficiently small to be ignored so that this is the same as a equation with

$$\alpha_i = (1 - \lambda)\, \lambda^{i-1}$$

The weights for the u_i decline at rate λ as we move back through time. Each weight is λ times the previous weight.

EWMA cont.

- The EWMA approach is attractive since only relatively little data needs to be stored.
- At any give time, we need to store only the current estimate of the variance rate and the most recent observation on the value of the market variable.
- When we get a new observation on the value of the market variable, we calculate a new daily percentage change and use equation

$$\sigma_n^2 = \lambda\sigma_{n-1}^2 + (1-\lambda)u_{n-1}^2$$

  to update our estimate of the variance rate.
- The value of λ governs how responsive the estimate of the daily volatility is to the most recent daily percentage change.
- A low value of λ leads to a great deal of weight being given to the u_{n-1}^2 when σ_n is calculated.
- A high value of λ (close to 1.0) produces estimates of the daily volatility that respond relatively slowly to new information provided by the daily percentage change

The GARCH(1,1) Model

A more generalized method to generate volatilities is the **GARCH(1,1)**, or **generalized autoregressive conditional heteroscedasticity**.

- GARCH(1,1) differs from EWMA by including a long-run variance rate. σ_n^2 is calculated from a long-rn average variance rate V_L as well as from σ_{n-1}^2 and u_{n-1}:

$$\sigma_n^2 = \gamma V_L + \alpha u_{n-1}^2 + \beta \sigma_{n-1}^2$$

  where γ is the weight assigned to V_L, α is the weight assigned to u_{n-1}^2, and β is the weight assigned to σ_{n-1}^2.
- The weights must sum as

$$\gamma + \alpha + \beta = 1$$

- EWMA is a particular case of GARCH(1,1) with $\gamma = 0$, $\alpha = 1 - \lambda$, and $\beta = \lambda$.
- The $(1, 1)$ in GARCH(1,1) implies that σ_n^2 is based on the most recent observation of u^2 and the most recent estimate of the variance rate

GARCH(1,1), cont.

- Setting $\omega = \gamma V_L$, the GARCH(1,1) model can also be written

$$\sigma_n^2 = \omega + \alpha u_{n-1}^2 + \beta \sigma_{n-1}^2$$

- This is the form of the model that is usually used for the purposes of estimating the parameters.
- Once ω, α, and β have been estimated, we can calculate γ as $1 - \alpha - \beta$. The long-term variance V_L can then be calculated as ω/γ.
- For a stable GARCH(1,1) process we require $\alpha + \beta < 1$. Otherwise the weight applied to the long-term variance is negative

GARCH(1,1), cont.

The Weights:
Substituting for σ_{n-1}^2 in the GARCH(1,1) model, we obtain

$$\sigma_n^2 = \omega + \alpha u_{n-1}^2 + \beta \left(\omega + \alpha u_{n-2}^2 + \beta \sigma_{n-2}^2 \right)$$

The weights decline exponentially at rate β.

The parameter β can be interpreted as a **decay rate**. It is similar to λ in the EWMA model.

It defines the relative importance of the observations on the u's in determining the current variance rate.

Mean Reversion
The GARCH(1,1) model recognizes that over time the variance tends to get pulled back to a long-run average level of V_L.

Choosing Between Models

In practice, variance rates ten to be mean reverting. The GARCH(1,1) model incorporates mean reversion, whereas EWMA model does not. The GARCH(1,1) model is therefore theoretically more appealing than the EWMA model.

In circumstances where the best-fit value of ω turns out to be negative, the GARCH(1,1) model is not stable and it makes sense to switch to EWMA model.

Maximum Likelihood Methods

Now appropriate to discuss how the parameters in the models we have been considering are estimated from historical data. We use the **maximum likelihood method**. We choose parameters that maximize the chance of the data occurring.

Consider the problem of estimating a variance of a variable X from m observations on X when the underlying distribution is normal with mean zero.

We assume that the the observations are $u_1, u_2, \ldots, u_m$ and that the mean of the underlying distribution is zero. Denote the variance by v. The likelihood of u_i being observed is the probability density function for X when $X = u_i$. This is

$$\frac{1}{\sqrt{2\pi v}} \exp\left[-\frac{u_i^2}{2v}\right]$$

The likelihood of m observations occurring in the order in which they are observed is

$$\prod_{i=1}^{m} \left[\frac{1}{\sqrt{2\pi v}} \exp\left[-\frac{u_i^2}{2v}\right]\right] \tag{22}$$

Using the maximum likelihood method, the best estimate of v is the value that maximizes this expression.

Maximizing an expression is equivalent to maximizing the logarithm of the expression. Taking logarithms of the expression in (22).

Ignoring constant multiplicative factors, it can be seen that we wish to maximize

$$\sum_{i=1}^{m} \left[-\ln(v) - \frac{u_i^2}{v} \right]$$

or

$$-m\ln(v) - \sum_{i=1}^{m} \frac{u_i^2}{v}$$

Differentiating this expression with respect to v and setting the result equation to zero, we see that the maximum likelihood estimator of v is

$$\frac{1}{m}\sum_{i=1}^{m} u_i^2$$

Correlations

Correlations are important, as seen from last week, for computing VaR. We show how correlation estimates can be updated in a similar way as volatility estimates.

The correlation between two variables X and Y can be defined by

$$\frac{\mathrm{cov}(X, Y)}{\sigma_X \sigma_Y}$$

where σ_X is the standard deviation of X, σ_Y is the standard deviation of Y, and $\mathrm{cov}(X, Y)$ is the covariance between X and Y. The covariance between X and Y is defined as

$$E\left[(X - \mu_X)(Y - \mu_Y)\right]$$

where μ_X and μ_Y are the means of X and Y. Easier to develop methods for the covariances as opposed to the correlations.

- Define x_i and y_i as the percentage changes in the values of X and Y between the end of day $i-1$ and the end of day i:

$$x_i = \frac{X_i - X_{i-1}}{X_{i-1}} \qquad y_i = \frac{Y_i - Y_{i-1}}{Y_{i-1}}$$

  where X_i and Y_i are the values of X and Y at the end day i.
- We also define the following
  - $\sigma_{x,n}$ - daily volatility of variable X, estimated for day n
  - $\sigma_{y,n}$ - daily volatility of variable Y, estimated for day n
  - cov_n - daily covariance between daily changes in X and Y, estimated for day n
- Then we estimate the correlation between X and Y on day n as

$$\frac{\mathrm{cov}_n}{\sigma_{x,n}\sigma_{y,n}}$$

- We use an equal-weighting scheme and assuming that the means of x_i and y_i are zero, then we can estimate the variance rates of X and Y from the most recent m observations as

$$\sigma_{x,n} = \frac{1}{m}\sum_{i=1}^{m} x_{n-i}^2 \qquad \sigma_{y,n} = \frac{1}{m}\sum_{i=1}^{m} y_{n-i}^2$$

  A similar estimate for the covariance between X and Y is

$$\mathrm{cov}_n = \frac{1}{m}\sum_{i=1}^{m} x_{n-i} y_{n-i} \tag{23}$$

One alternative for updating covariances is an EWMA model similar we find

$$\text{cov}_n = \lambda \, \text{cov}_{n-1} + (1 - \lambda) \, x_{n-1} y_{n-1}$$

A similar analysis to that presented for the EWMA volatility model shows that the weights given observations on the $x_i y_i$ decline as we move back through time. The lower the value of λ the greater the weight that is give to recent observations.

Consistency Condition for Covariances

Once all the variances and covariances have been computed, a variance-covariance matrix can be constructed. When $i \neq j$, the (i, j) element of the matrix shows the covariance between variable i and variable j. When $i = j$ it shows the variance of variable i.

Not all variance-covariance matrices are internally consistent. The condition for an $N \times N$ variance-covariance matrix Ω to be internally consistent is

$$w^T \Omega w \geq 0$$

for all $N \times 1$ vectors w, where w^T is the transpose of w. Such matrices are **positive semi-definite**.

To ensure that a positive-semidefinite matrix is produced, variances and covariances should be calculated consistently. For example, if variances are calculated by giving equal weight to the last m data items, the same should be done for covariances.

Midterm

Concentrate on:

- Chapter 13 - Chapter 19 with emphasis on Chapters 17-19.