Euler Approximation

- Define $t_n \stackrel{\Delta}{=} n\Delta t$.
- The following approximation follows from Eq. (75),

$$\begin{aligned} & \widehat{X}(t_{n+1}) \\ =& \widehat{X}(t_n) + a(\widehat{X}(t_n), t_n)\,\Delta t + b(\widehat{X}(t_n), t_n)\,\Delta W(t_n). \end{aligned} \tag{76}$$

- It is called the Euler or Euler-Maruyama method.
- Recall that $\Delta W(t_n)$ should be interpreted as

$$W(t_{n+1}) - W(t_n),$$

not $W(t_n) - W(t_{n-1})$!

Euler Approximation (concluded)

- With the Euler method, one can obtain a sample path

$$\widehat{X}(t_1), \widehat{X}(t_2), \widehat{X}(t_3), \ldots$$

  from a sample path

$$W(t_0), W(t_1), W(t_2), \ldots .$$

- Under mild conditions, $\widehat{X}(t_n)$ converges to $X(t_n)$.

More Discrete Approximations

- Under fairly loose regularity conditions, Eq. (76) on p. 579 can be replaced by

$$
\begin{aligned}
&\widehat{X}(t_{n+1}) \\
=&\widehat{X}(t_n) + a(\widehat{X}(t_n), t_n)\,\Delta t + b(\widehat{X}(t_n), t_n)\sqrt{\Delta t}\,Y(t_n).
\end{aligned}
$$

  – $Y(t_0), Y(t_1), \ldots$ are independent and identically distributed with zero mean and unit variance.

More Discrete Approximations (concluded)

- An even simpler discrete approximation scheme:

$$\begin{aligned}&\widehat{X}(t_{n+1})\\ =&\widehat{X}(t_n)+a(\widehat{X}(t_n),t_n)\,\Delta t+b(\widehat{X}(t_n),t_n)\sqrt{\Delta t}\,\xi.\end{aligned}$$

  - $\text{Prob}[\,\xi=1\,]=\text{Prob}[\,\xi=-1\,]=1/2$.
  - Note that $E[\,\xi\,]=0$ and $\text{Var}[\,\xi\,]=1$.
- This is a binomial model.
- As Δt goes to zero, $\widehat{X}$ converges to X.[a]

[a] He (1990).

Trading and the Ito Integral

- Consider an Ito process

$$d\boldsymbol{S}_t = \mu_t \, dt + \sigma_t \, dW_t.$$

  - $\boldsymbol{S}_t$ is the vector of security prices at time t.

- Let $\boldsymbol{\phi}_t$ be a trading strategy denoting the quantity of each type of security held at time t.
  - Hence the stochastic process $\boldsymbol{\phi}_t \boldsymbol{S}_t$ is the value of the portfolio $\boldsymbol{\phi}_t$ at time t.

- $\boldsymbol{\phi}_t \, d\boldsymbol{S}_t \stackrel{\Delta}{=} \boldsymbol{\phi}_t(\mu_t \, dt + \sigma_t \, dW_t)$ represents the change in the value from security price changes occurring at time t.

Trading and the Ito Integral (concluded)

- The equivalent Ito integral,

$$G_T(\boldsymbol{\phi}) \stackrel{\Delta}{=} \int_0^T \boldsymbol{\phi}_t \, d\boldsymbol{S}_t = \int_0^T \boldsymbol{\phi}_t \mu_t \, dt + \int_0^T \boldsymbol{\phi}_t \sigma_t \, dW_t,$$

  measures the gains realized by the trading strategy over the period $[\,0, T\,]$.

Ito's Lemma[a]

A smooth function of an Ito process is itself an Ito process.

Theorem 18 *Suppose* $f : R \to R$ *is twice continuously differentiable and* $dX = a_t\, dt + b_t\, dW$. *Then* $f(X)$ *is the Ito process,*

$$\begin{aligned} & f(X_t) \\ = \; & f(X_0) + \int_0^t f'(X_s)\, a_s\, ds + \int_0^t f'(X_s)\, b_s\, dW \\ & + \frac{1}{2} \int_0^t f''(X_s)\, b_s^2\, ds \end{aligned}$$

for $t \geq 0$.

---

[a]Ito (1944).

Ito's Lemma (continued)

- In differential form, Ito's lemma becomes

$$df(X) = f'(X)\, a\, dt + f'(X)\, b\, dW + \frac{1}{2}\, f''(X)\, b^2\, dt. \tag{77}$$

- Compared with calculus, the interesting part is the third term on the right-hand side.

- A convenient formulation of Ito's lemma is

$$df(X) = f'(X)\, dX + \frac{1}{2}\, f''(X)(dX)^2. \tag{78}$$

Ito's Lemma (continued)

- We are supposed to multiply out $(dX)^2 = (a\,dt + b\,dW)^2$ symbolically according to

$\times$	dW	dt
dW	dt	0
dt	0	0

  - The $(dW)^2 = dt$ entry is justified by a known result.

- Hence $(dX)^2 = (a\,dt + b\,dW)^2 = b^2\,dt$ in Eq. (78).

- This form is easy to remember because of its similarity to the Taylor expansion.

Ito's Lemma (continued)

Theorem 19 (Higher-Dimensional Ito's Lemma) *Let $W_1, W_2, \ldots, W_n$ be independent Wiener processes and $X \stackrel{\Delta}{=} (X_1, X_2, \ldots, X_m)$ be a vector process. Suppose $f : R^m \to R$ is twice continuously differentiable and X_i is an Ito process with $dX_i = a_i\, dt + \sum_{j=1}^n b_{ij}\, dW_j$. Then $df(X)$ is an Ito process with the differential,*

$$df(X) = \sum_{i=1}^m f_i(X)\, dX_i + \frac{1}{2} \sum_{i=1}^m \sum_{k=1}^m f_{ik}(X)\, dX_i\, dX_k,$$

where $f_i \stackrel{\Delta}{=} \partial f / \partial X_i$ and $f_{ik} \stackrel{\Delta}{=} \partial^2 f / \partial X_i \partial X_k$.

Ito's Lemma (continued)

- The multiplication table for Theorem 19 is

$\times$	dW_i	dt
dW_k	$\delta_{ik}\, dt$	0
dt	0	0

in which

$$\delta_{ik} = \begin{cases} 1, & \text{if } i = k, \\ 0, & \text{otherwise.} \end{cases}$$

Ito's Lemma (continued)

- In applying the higher-dimensional Ito's lemma, usually one of the variables, say X_1, is time t and $dX_1 = dt$.
- In this case, $b_{1j} = 0$ for all j and $a_1 = 1$.
- As an example, let

$$dX_t = a_t \, dt + b_t \, dW_t.$$

- Consider the process $f(X_t, t)$.

Ito's Lemma (continued)

- Then

$$
\begin{aligned}
df &= \frac{\partial f}{\partial X_t}\, dX_t + \frac{\partial f}{\partial t}\, dt + \frac{1}{2}\frac{\partial^2 f}{\partial X_t^2}\,(dX_t)^2 \\
&= \frac{\partial f}{\partial X_t}\,(a_t\, dt + b_t\, dW_t) + \frac{\partial f}{\partial t}\, dt \\
&\quad + \frac{1}{2}\frac{\partial^2 f}{\partial X_t^2}\,(a_t\, dt + b_t\, dW_t)^2 \\
&= \left(\frac{\partial f}{\partial X_t}\, a_t + \frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial X_t^2}\, b_t^2 \right) dt \\
&\quad + \frac{\partial f}{\partial X_t}\, b_t\, dW_t. \qquad (79)
\end{aligned}
$$

Ito's Lemma (continued)

Theorem 20 (Alternative Ito's Lemma) *Let* $W_1, W_2, \ldots, W_m$ *be Wiener processes and* $X \triangleq (X_1, X_2, \ldots, X_m)$ *be a vector process. Suppose* $f : R^m \to R$ *is twice continuously differentiable and* X_i *is an Ito process with* $dX_i = a_i\, dt + b_i\, dW_i$*. Then* $df(X)$ *is the following Ito process,*

$$df(X) = \sum_{i=1}^{m} f_i(X)\, dX_i + \frac{1}{2} \sum_{i=1}^{m} \sum_{k=1}^{m} f_{ik}(X)\, dX_i\, dX_k.$$

Ito's Lemma (concluded)

- The multiplication table for Theorem 20 is

$\times$	dW_i	dt
dW_k	$\rho_{ik}\, dt$	0
dt	0	0

- Above, ρ_{ik} denotes the correlation between dW_i and dW_k.

Geometric Brownian Motion

- Consider geometric Brownian motion

$$Y(t) \stackrel{\Delta}{=} e^{X(t)}.$$

  - $X(t)$ is a (μ, σ) Brownian motion.
  - By Eq. (72) on p. 555,

$$dX = \mu\, dt + \sigma\, dW.$$

- Note that

$$\begin{aligned} \frac{\partial Y}{\partial X} &= Y, \\ \frac{\partial^2 Y}{\partial X^2} &= Y. \end{aligned}$$

Geometric Brownian Motion (continued)

- Ito's formula (77) on p. 586 implies

$$\begin{aligned} dY &= Y\,dX + (1/2)\,Y\,(dX)^2 \\ &= Y\,(\mu\,dt + \sigma\,dW) + (1/2)\,Y\,(\mu\,dt + \sigma\,dW)^2 \\ &= Y\,(\mu\,dt + \sigma\,dW) + (1/2)\,Y\sigma^2\,dt. \end{aligned}$$

- Hence

$$\frac{dY}{Y} = \left(\mu + \sigma^2/2\right) dt + \sigma\,dW. \tag{80}$$

- The annualized *instantaneous* rate of return is $\mu + \sigma^2/2$ (not μ).[a]

[a] Consistent with Lemma 9 (p. 289).

Geometric Brownian Motion (concluded)

- Similarly, suppose

$$\frac{dY}{Y} = \mu\, dt + \sigma\, dW.$$

- Then $X(t) \stackrel{\Delta}{=} \ln Y(t)$ follows

$$dX = \left(\mu - \sigma^2/2\right) dt + \sigma\, dW.$$

Product of Geometric Brownian Motion Processes

- Let

$$\frac{dY}{Y} = a\,dt + b\,dW_Y,$$
$$\frac{dZ}{Z} = f\,dt + g\,dW_Z.$$

- Assume dW_Y and dW_Z have correlation ρ.
- Consider the Ito process

$$U \stackrel{\Delta}{=} YZ.$$

Product of Geometric Brownian Motion Processes (continued)

- Apply Ito's lemma (Theorem 20 on p. 592):

$$
\begin{aligned}
dU &= Z\,dY + Y\,dZ + dY\,dZ \\
&= ZY(a\,dt + b\,dW_Y) + YZ(f\,dt + g\,dW_Z) \\
&\quad + YZ(a\,dt + b\,dW_Y)(f\,dt + g\,dW_Z) \\
&= U(a + f + bg\rho)\,dt + Ub\,dW_Y + Ug\,dW_Z.
\end{aligned}
$$

- The product of correlated geometric Brownian motion processes thus remains geometric Brownian motion.

Product of Geometric Brownian Motion Processes (continued)

- Note that

$$\begin{aligned} Y &= \exp\left[\left(a - b^2/2\right) dt + b\, dW_Y\right], \\ Z &= \exp\left[\left(f - g^2/2\right) dt + g\, dW_Z\right], \\ U &= \exp\left[\,\left(a + f - \left(b^2 + g^2\right)/2\right) dt + b\, dW_Y + g\, dW_Z\,\right]. \end{aligned}$$

  – There is no $bg\rho$ term in U!

Product of Geometric Brownian Motion Processes (concluded)

- $\ln U$ is Brownian motion with a mean equal to the sum of the means of $\ln Y$ and $\ln Z$.
- This holds even if Y and Z are correlated.
- Finally, $\ln Y$ and $\ln Z$ have correlation ρ.

Quotients of Geometric Brownian Motion Processes

- Suppose Y and Z are drawn from p. 597.
- Let
$$U \triangleq Y/Z.$$
- We now show that[a]
$$\frac{dU}{U} = (a - f + g^2 - bg\rho)\, dt + b\, dW_Y - g\, dW_Z. \tag{81}$$
- Keep in mind that dW_Y and dW_Z have correlation ρ.

[a]Exercise 14.3.6 of the textbook is erroneous.

Quotients of Geometric Brownian Motion Processes (concluded)

- The multidimensional Ito's lemma (Theorem 20 on p. 592) can be employed to show that

$$
\begin{aligned}
& dU \\
= \;& (1/Z)\, dY - (Y/Z^2)\, dZ - (1/Z^2)\, dY\, dZ + (Y/Z^3)\, (dZ)^2 \\
= \;& (1/Z)(aY\, dt + bY\, dW_Y) - (Y/Z^2)(fZ\, dt + gZ\, dW_Z) \\
& -(1/Z^2)(bgYZ\rho\, dt) + (Y/Z^3)(g^2 Z^2\, dt) \\
= \;& U(a\, dt + b\, dW_Y) - U(f\, dt + g\, dW_Z) \\
& -U(bg\rho\, dt) + U(g^2\, dt) \\
= \;& U(a - f + g^2 - bg\rho)\, dt + Ub\, dW_Y - Ug\, dW_Z.
\end{aligned}
$$

Forward Price

- Suppose S follows

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

- Consider $F(S,t) \stackrel{\Delta}{=} Se^{y(T-t)}$ for some constants y and T.
- As F is a function of two variables, we need the various partial derivatives of $F(S,t)$ with respect to S and t.
- Note that in partial differentiation with respect to one variable, other variables are held constant.[a]

---

[a]Contributed by Mr. Sun, Ao (`R05922147`) on April 26, 2017.

Forward Prices (continued)

- Now,

$$\begin{aligned} \frac{\partial F}{\partial S} &= e^{y(T-t)}, \\ \frac{\partial^2 F}{\partial S^2} &= 0, \\ \frac{\partial F}{\partial t} &= -ySe^{y(T-t)}. \end{aligned}$$

- Then

$$\begin{aligned} dF &= e^{y(T-t)}\, dS - ySe^{y(T-t)}\, dt \\ &= Se^{y(T-t)}\, (\mu\, dt + \sigma\, dW) - ySe^{y(T-t)}\, dt \\ &= F(\mu - y)\, dt + F\sigma\, dW. \end{aligned}$$

Forward Prices (concluded)

- One can also prove it by Eq. (79) on p. 591.
- Thus F follows
$$\frac{dF}{F} = (\mu - y)\, dt + \sigma\, dW.$$
- This result has applications in forward and futures contracts.
- In Eq. (55) on p. 469, $\mu = r = y$.
- So
$$\frac{dF}{F} = \sigma\, dW,$$
a martingale.[a]

[a] It is also consistent with p. 545.

Ornstein-Uhlenbeck (OU) Process

- The OU process:

$$dX = -\kappa X\, dt + \sigma\, dW,$$

where $\kappa, \sigma \geq 0$.

- For $t_0 \leq s \leq t$ and $X(t_0) = x_0$, it is known that

$$\begin{aligned}
E[\, X(t)\,] &= e^{-\kappa(t-t_0)}\, E[\, x_0\,], \\
\text{Var}[\, X(t)\,] &= \frac{\sigma^2}{2\kappa} \left(1 - e^{-2\kappa(t-t_0)}\right) + e^{-2\kappa(t-t_0)}\, \text{Var}[\, x_0\,], \\
\text{Cov}[\, X(s), X(t)\,] &= \frac{\sigma^2}{2\kappa}\, e^{-\kappa(t-s)} \left[\, 1 - e^{-2\kappa(s-t_0)}\, \right] \\
&\quad + e^{-\kappa(t+s-2t_0)}\, \text{Var}[\, x_0\,].
\end{aligned}$$

Ornstein-Uhlenbeck Process (continued)

- $X(t)$ is normally distributed if x_0 is a constant or normally distributed.
- X is said to be a normal process.
- $E[\,x_0\,] = x_0$ and $\text{Var}[\,x_0\,] = 0$ if x_0 is a constant.
- The OU process has the following mean reversion property.
  - When $X > 0$, X is pulled toward zero.
  - When $X < 0$, it is pulled toward zero again.

Ornstein-Uhlenbeck Process (continued)

- A generalized version:

$$dX = \kappa(\mu - X)\, dt + \sigma\, dW,$$

where $\kappa, \sigma \geq 0$.

- Given $X(t_0) = x_0$, a constant, it is known that

$$\begin{aligned} E[\, X(t)\,] &= \mu + (x_0 - \mu)\, e^{-\kappa(t-t_0)}, \quad (82) \\ \text{Var}[\, X(t)\,] &= \frac{\sigma^2}{2\kappa} \left[1 - e^{-2\kappa(t-t_0)} \right], \end{aligned}$$

for $t_0 \leq t$.

Ornstein-Uhlenbeck Process (concluded)

- The mean and standard deviation are roughly μ and $\sigma/\sqrt{2\kappa}$, respectively.
- For large t, the probability of $X < 0$ is extremely unlikely in any finite time interval when $\mu > 0$ is large relative to $\sigma/\sqrt{2\kappa}$.
- The process is mean-reverting.
  - X tends to move toward μ.
  - Useful for modeling term structure, stock price volatility, and stock price return.[a]

[a]See Knutson, Wimmer, Kuhnen, & Winkielman (2008) for the biological basis for mean reversion in financial decision making.

Square-Root Process

- Suppose X is an OU process.
- Consider
$$V \stackrel{\Delta}{=} X^2.$$
- Ito's lemma says V has the differential,
$$\begin{aligned} dV &= 2X\,dX + (dX)^2 \\ &= 2\sqrt{V}\,(-\kappa\sqrt{V}\,dt + \sigma\,dW) + \sigma^2\,dt \\ &= \left(-2\kappa V + \sigma^2\right)dt + 2\sigma\sqrt{V}\,dW, \end{aligned}$$

  a square-root process.

Square-Root Process (continued)

- In general, the square-root process has the stochastic differential equation,

$$dX = \kappa(\mu - X)\, dt + \sigma\sqrt{X}\, dW,$$

where $\kappa, \sigma \geq 0$ and $X(0)$ is a nonnegative constant.

- Like the OU process, it possesses mean reversion: X tends to move toward μ, but the volatility is proportional to $\sqrt{X}$ instead of a constant.

Square-Root Process (continued)

- When X hits zero and $\mu \geq 0$, the probability is one that it will not move below zero.
  - Zero is a reflecting boundary.
- Hence, the square-root process is a good candidate for modeling interest rates.[a]
- The OU process, in contrast, allows negative interest rates.[b]
- The two processes are related.[c]

---

[a] Cox, Ingersoll, & Ross (1985).
[b] Some rates did go negative in Europe in 2015.
[c] Recall p. 610.

Square-Root Process (concluded)

- The random variable $2cX(t)$ follows the noncentral chi-square distribution,[a]

$$\chi\left(\frac{4\kappa\mu}{\sigma^2}, 2cX(0)\, e^{-\kappa t}\right),$$

where $c \stackrel{\Delta}{=} (2\kappa/\sigma^2)(1-e^{-\kappa t})^{-1}$.

- Given $X(0) = x_0$, a constant,

$$\begin{aligned} E[\,X(t)\,] &= x_0 e^{-\kappa t} + \mu\left(1-e^{-\kappa t}\right), \\ \mathrm{Var}[\,X(t)\,] &= x_0\,\frac{\sigma^2}{\kappa}\,\left(e^{-\kappa t} - e^{-2\kappa t}\right) + \mu\,\frac{\sigma^2}{2\kappa}\,\left(1-e^{-\kappa t}\right)^2, \end{aligned}$$

for $t \geq 0$.

[a] William Feller (1906–1970) in 1951.

Modeling Stock Prices

- The most popular stochastic model for stock prices has been the geometric Brownian motion,

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

- The continuously compounded rate of return $X \triangleq \ln S$ follows

$$dX = \left(\mu - \frac{\sigma^2}{2}\right) dt + \sigma\, dW$$

by Ito's lemma.[a]

---

[a] Recall Eq. (80) on p. 595. Consistent with Lemma 9 (p. 289).

Local-Volatility Models

- The more general deterministic-volatility model posits

$$\frac{dS}{S} = (r_t - q_t)\, dt + \sigma(S, t)\, dW,$$

  where instantaneous volatility $\sigma(S, t)$ is called the local volatility function.[a]

- A (weak) solution exists if $S\sigma(S, t)$ is continuous and grows at most linearly in S and t.[b]

- One needs to recover the local volatility surface $\sigma(S, t)$ from the implied volatility surface.

---

[a] Derman & Kani (1994); Dupire (1994).
[b] Skorokhod (1961).

Local-Volatility Models (continued)

- Theoretically,[a]

$$\sigma(X,T)^2 = 2\,\frac{\frac{\partial C}{\partial T} + (r_T - q_T)X\,\frac{\partial C}{\partial X} + q_T C}{X^2\,\frac{\partial^2 C}{\partial X^2}}. \qquad (83)$$

- C is the call price at time $t = 0$ (today) with strike price X and time to maturity T.
- $\sigma(X,T)$ is the local volatility that will prevail at *future time* T and *stock price* $S_T = X$.

[a]Dupire (1994); Andersen & Brotherton-Ratcliffe (1998).

Local-Volatility Models (continued)

- For more general models, this equation gives the expectation as seen from today, under the risk-neural probability, of the instantaneous variance at time T given that $S_T = X$.[a]
- In practice, the $\sigma(S, t)^2$ derived by Dupire's formula (83) may have spikes, vary wildly, or even be negative.
- The term $\partial^2 C / \partial X^2$ in the denominator often results in numerical instability.

---

[a]Derman & Kani (1997).

Local-Volatility Models (continued)

- Denote the implied volatility surface by $\Sigma(X,T)$ and the local volatility surface by $\sigma(S,t)$.
- The relation between $\Sigma(X,T)$ and $\sigma(X,T)$ is[a]

$$\sigma(X,T)^2 = \frac{\Sigma^2 + 2\Sigma\tau\left[\frac{\partial\Sigma}{\partial T} + (r_T - q_T)X\frac{\partial\Sigma}{\partial X}\right]}{\left(1 - \frac{Xy}{\Sigma}\frac{\partial\Sigma}{\partial X}\right)^2 + X\Sigma\tau\left[\frac{\partial\Sigma}{\partial X} - \frac{X\Sigma\tau}{4}\left(\frac{\partial\Sigma}{\partial X}\right)^2 + X\frac{\partial^2\Sigma}{\partial X^2}\right]},$$

$$\tau \stackrel{\Delta}{=} T - t,$$

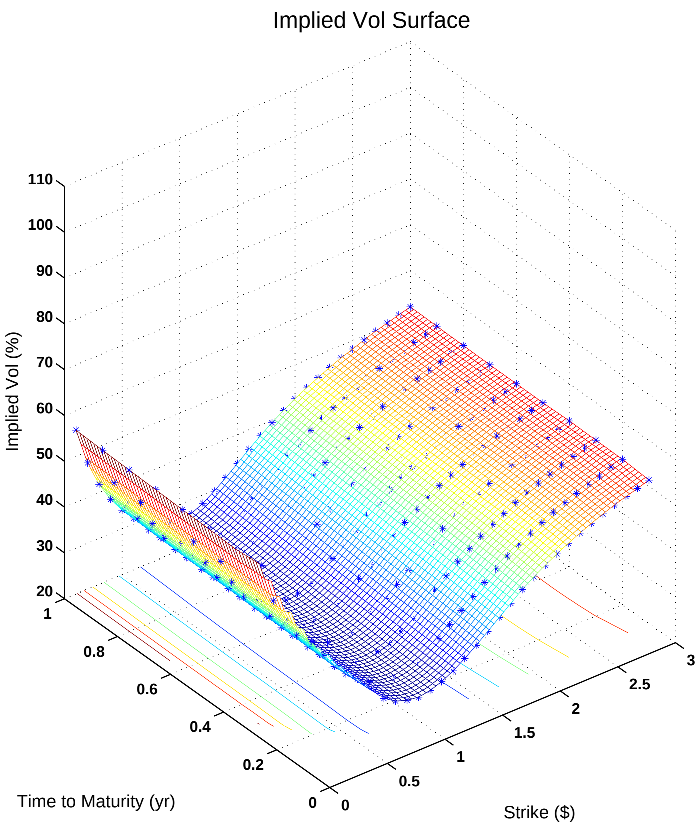
$$y \stackrel{\Delta}{=} \ln(X/S_t) + \int_t^T (q_s - r_s)\, ds.$$

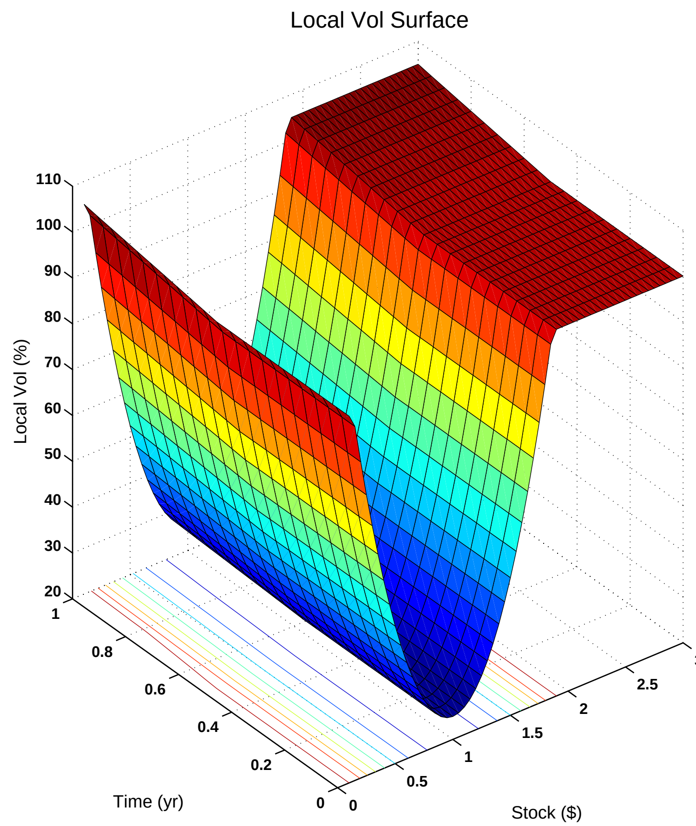
- Although this version may be more stable than Eq. (83) on p. 616, it is expected to suffer from similar problems.

---

[a]Andreasen (1996); Andersen & Brotherton-Ratcliffe (1998); Gatheral (2003); Wilmott (2006); Kamp (2009).

Implied and Local Volatility Surfaces[a]





---

[a] Contributed by Mr. Lok, U Hou (D99922028) on April 5, 2014.

Local-Volatility Models (continued)

- Small changes to the implied volatility surface may produce big changes to the local volatility surface.
- In reality, option prices only exist for a finite set of maturities and strike prices.
- Hence interpolation and extrapolation may be needed to construct the volatility surface.[a]
- But then some implied volatility surfaces generate option prices that allow arbitrage opportunities.[b]

---

[a]Doing it to the option prices produces worse results (Li, 2000/2001).

[b]See Rebonato (2004) for an example.

Local-Volatility Models (concluded)

- There exist conditions for a set of option prices to be arbitrage-free.[a]
- For some vanilla equity options, the Black-Scholes model seems "better than" the local-volatility model.[b]
- Hirsa and Neftci (2013), "most traders and firms actively utilize this [local-volatility] model."

---

[a] Davis & Hobson (2007).

[b] Dumas, Fleming, & Whaley (1998).

Implied Trees

- The trees for the local volatility model are called implied trees.[a]
- Their construction requires option prices at all strike prices and maturities.
  - That is, an implied volatility surface.
- The local volatility model does *not* require that the implied tree combine.
- An exponential-sized implied tree exists.[b]

[a] Derman & Kani (1994); Dupire (1994); Rubinstein (1994).
[b] Charalambousa, Christofidesb, & Martzoukosa (2007).

Implied Trees (continued)

- How to construct a valid implied tree with efficiency has been open for a long time.[a]
  - Reasons may include: noise and nonsynchrony in data, arbitrage opportunities in the smoothed and interpolated/extrapolated implied volatility surface, wrong model, wrong algorithms, nonlinearity, instability, etc.
- Inversion is an ill-posed numerical problem.[b]

---

[a]Rubinstein (1994); Derman & Kani (1994); Derman, Kani, & Chriss (1996); Jackwerth & Rubinstein (1996); Jackwerth (1997); Coleman, Kim, Li, & Verma (2000); Li (2000/2001); Moriggia, Muzzioli, & Torricelli (2009); Rebonato (2004).

[b]Ayache, Henrotte, Nassar, & X. Wang (2004).

Implied Trees (concluded)

- It is finally solved for separable local volatilities σ.[a]
  - The local-volatility function $\sigma(S, V)$ is separable[b] if

$$\sigma(S,t) = \sigma_1(S)\,\sigma_2(t).$$

- A very general solution is recently obtained.[c]

---

[a] Lok (D99922028) & Lyuu (2015, 2016, 2017).
[b] Rebonato (2004); Brace, Gątarek, & Musiela (1997).
[c] Lok (D99922028) & Lyuu (2016, 2017).

The Hull-White Model

- Hull and White (1987) postulate the following model,

$$\frac{dS}{S} = r\,dt + \sqrt{V}\,dW_1,$$
$$dV = \mu_{\mathrm{v}} V\,dt + bV\,dW_2.$$

- Above, V is the instantaneous variance.
- They assume μ_{v} depends on V and t (but not S).

The SABR Model

- Hagan, Kumar, Lesniewski, and Woodward (2002) postulate the following model,

$$\begin{aligned}\frac{dS}{S} &= r\,dt + S^{\theta} V\,dW_1,\\ dV &= bV\,dW_2,\end{aligned}$$

for $0 \leq \theta \leq 1$.

- A nice feature of this model is that the implied volatility surface has a compact approximate closed form.

The Hilliard-Schwartz Model

- Hilliard and Schwartz (1996) postulate the following general model,

$$\begin{aligned} \frac{dS}{S} &= r\,dt + f(S)V^a\,dW_1, \\ dV &= \mu(V)\,dt + bV\,dW_2, \end{aligned}$$

for some well-behaved function $f(S)$ and constant a.

The Blacher Model

- Blacher (2002) postulates the following model,

$$\begin{aligned} \frac{dS}{S} &= r\,dt + \sigma \left[\, 1 + \alpha(S - S_0) + \beta(S - S_0)^2 \,\right] dW_1, \\ d\sigma &= \kappa(\theta - \sigma)\,dt + \epsilon\sigma\,dW_2. \end{aligned}$$

- The volatility σ follows a mean-reverting process to level θ.

Heston's Stochastic-Volatility Model

- Heston (1993) assumes the stock price follows

$$\frac{dS}{S} = (\mu - q)\, dt + \sqrt{V}\, dW_1, \tag{84}$$

$$dV = \kappa(\theta - V)\, dt + \sigma\sqrt{V}\, dW_2. \tag{85}$$

  - V is the instantaneous variance, which follows a square-root process.
  - dW_1 and dW_2 have correlation ρ.
  - The riskless rate r is constant.

- It may be the most popular continuous-time stochastic-volatility model.[a]

[a] Christoffersen, Heston, & Jacobs (2009).

Heston's Stochastic-Volatility Model (continued)

- Heston assumes the market price of risk is $b_2\sqrt{V}$.
- So $\mu = r + b_2 V$.
- Define

$$
\begin{aligned}
dW_1^* &= dW_1 + b_2\sqrt{V}\,dt, \\
dW_2^* &= dW_2 + \rho b_2\sqrt{V}\,dt, \\
\kappa^* &= \kappa + \rho b_2 \sigma, \\
\theta^* &= \frac{\theta\kappa}{\kappa + \rho b_2 \sigma}.
\end{aligned}
$$

- dW_1^* and dW_2^* have correlation ρ.

Heston's Stochastic-Volatility Model (continued)

- Under the risk-neutral probability measure Q, both W_1^* and W_2^* are Wiener processes.
- Heston's model becomes, under probability measure Q,

$$\begin{aligned} \frac{dS}{S} &= (r-q)\,dt + \sqrt{V}\,dW_1^*, \\ dV &= \kappa^*(\theta^* - V)\,dt + \sigma\sqrt{V}\,dW_2^*. \end{aligned}$$

Heston's Stochastic-Volatility Model (continued)

- Define

$$
\begin{aligned}
\phi(u,\tau) &= \exp\Big\{\, \imath u(\ln S + (r-q)\,\tau) \\
&\quad + \theta^* \kappa^* \sigma^{-2} \left[(\kappa^* - \rho\sigma u\imath - d)\,\tau - 2\ln\frac{1-ge^{-d\tau}}{1-g} \right] \\
&\quad + \frac{v\sigma^{-2}(\kappa^* - \rho\sigma u\imath - d)\left(1-e^{-d\tau}\right)}{1-ge^{-d\tau}} \Big\}, \\
d &= \sqrt{(\rho\sigma u\imath - \kappa^*)^2 - \sigma^2(-\imath u - u^2)}\,, \\
g &= (\kappa^* - \rho\sigma u\imath - d)/(\kappa^* - \rho\sigma u\imath + d).
\end{aligned}
$$

Heston's Stochastic-Volatility Model (continued)

The formulas are[a]

$$
\begin{aligned}
C &= S\left[\frac{1}{2}+\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(\frac{X^{-\imath u}\phi(u-\imath,\tau)}{\imath u S e^{r\tau}}\right)du\right] \\
&\quad -Xe^{-r\tau}\left[\frac{1}{2}+\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(\frac{X^{-\imath u}\phi(u,\tau)}{\imath u}\right)du\right], \\
P &= Xe^{-r\tau}\left[\frac{1}{2}-\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(\frac{X^{-\imath u}\phi(u,\tau)}{\imath u}\right)du\right], \\
&\quad -S\left[\frac{1}{2}-\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(\frac{X^{-\imath u}\phi(u-\imath,\tau)}{\imath u S e^{r\tau}}\right)du\right],
\end{aligned}
$$

where $\imath = \sqrt{-1}$ and $\mathrm{Re}(x)$ denotes the real part of the complex number x.

---

[a] Contributed by Mr. Chen, Chun-Ying (D95723006) on August 17, 2008 and Mr. Liou, Yan-Fu (R92723060) on August 26, 2008. See Cui, Rollin, & Germano (2017) for alternative formulas.

Heston's Stochastic-Volatility Model (concluded)

- For American options, we will need a tree for Heston's model.[a]
- They are all $O(n^3)$-sized.

---

[a] Nelson & Ramaswamy (1990); Nawalkha & Beliaeva (2007); Leisen (2010); Beliaeva & Nawalkha (2010); M. Chou (R02723073) (2015); M. Chou (R02723073) & Lyuu (2016).

Stochastic-Volatility Models and Further Extensions[a]

- How to explain the October 1987 crash?
  - The Dow Jones Industrial Average fell 22.61% on October 19, 1987 (called the Black Monday).
  - The CBOE S&P 100 Volatility Index (VXO) shot up to 150%, the highest VXO ever recorded.[b]
- Stochastic-volatility models require an implausibly high-volatility level prior to *and* after the crash.
  - Because the processes are continuous.
- Discontinuous jump models *in the asset price* can alleviate the problem somewhat.[c]

---

[a]Eraker (2004).
[b]Caprio (2012).
[c]Merton (1976).

Stochastic-Volatility Models and Further Extensions (continued)

- But if the jump intensity is a constant, it cannot explain the tendency of large movements to cluster over time.
- This assumption also has no impacts on option prices.
- Jump-diffusion models combine both.
  - E.g., add a jump process to Eq. (84) on p. 629.
  - Closed-form formulas exist for GARCH-jump option pricing models.[a]

[a]Liou (`R92723060`) (2005).

Stochastic-Volatility Models and Further Extensions (concluded)

- But they still do not adequately describe the systematic variations in option prices.[a]
- Jumps *in volatility* are alternatives.[b]
  - E.g., add correlated jump processes to Eqs. (84) and Eq. (85) on p. 629.
- Such models allow high level of volatility caused by a jump to volatility.[c]

---

[a] Bates (2000); Pan (2002).
[b] Duffie, Pan, & Singleton (2000).
[c] Eraker, Johnnes, & Polson (2000); Y. Lin (2007); Zhu & Lian (2012).

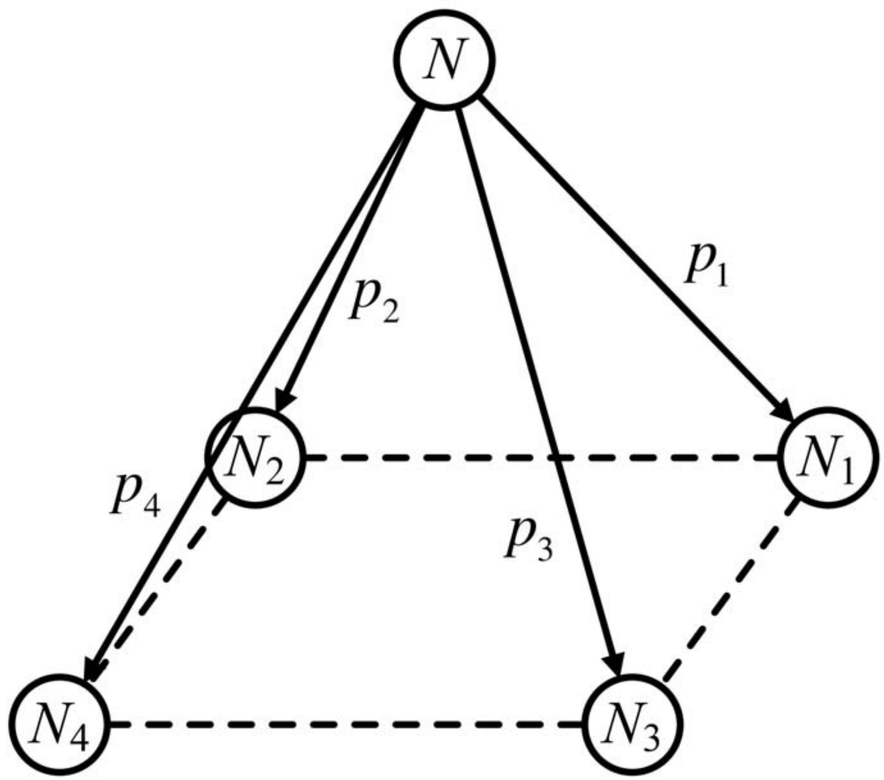
Why Are Trees for Stochastic-Volatility Models Difficult?

- The CRR tree is 2-dimensional.[a]
- The constant volatility makes the span from any node fixed.
- But a tree for a stochastic-volatility model must be 3-dimensional.
  - Every node is associated with a pair of stock price and a volatility.

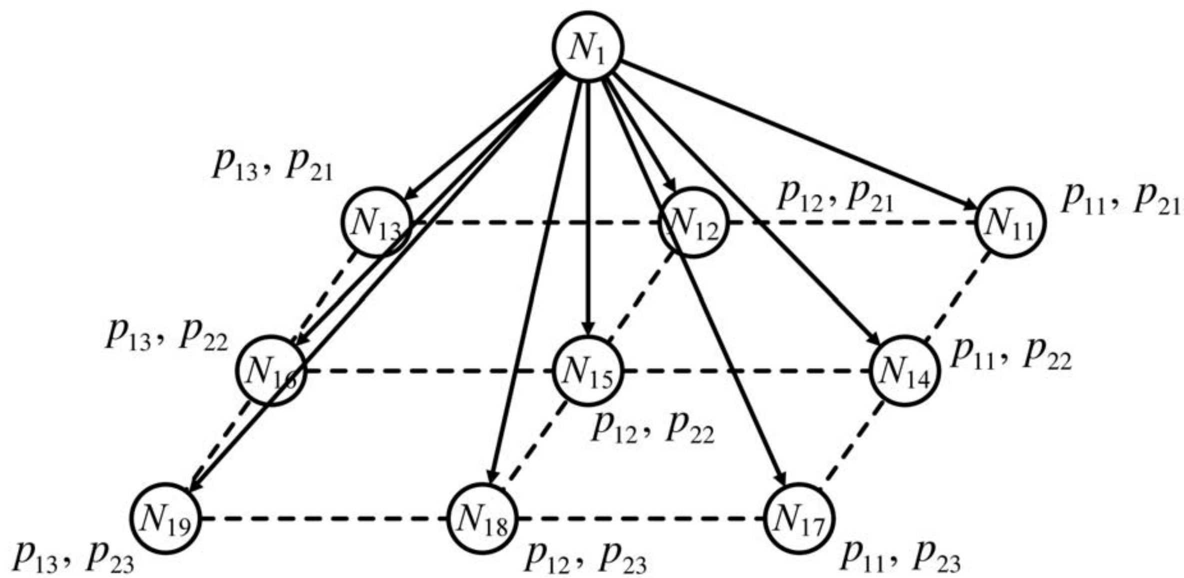
---

[a]Recall p. 286.

Why Are Trees for Stochastic-Volatility Models Difficult (Binomial Case)?



Why Are Trees for Stochastic-Volatility Models Difficult (Trinomial Case)?



Why Are Trees for Stochastic-Volatility Models Difficult? (concluded)

- Locally, the tree looks fine for one time step.
- But the volatility regulates the spans of the nodes on the stock-price plane.
- Unfortunately, those spans differ from node to node because the volatility varies.
- So two time steps from now, the branches will not combine!
- Smart ideas are thus needed.

Complexities of Stochastic-Volatility Models

- A few stochastic-volatility models suffer from subexponential ($c^{\sqrt{n}}$) tree size.
- Examples include the Hull-White (1987), Hilliard-Schwartz (1996), and SABR (2002) models.[a]
- Future research may extend this negative result to more stochastic-volatility models.
  - We suspect many GARCH option pricing models entertain similar problems.[b]

---

[a] Chiu (R98723059) (2012).

[b] Y. C. Chen (R95723051) (2008); Y. C. Chen (R95723051), Lyuu, & Wen (D94922003) (2011).

Complexities of Stochastic-Volatility Models (concluded)

- Calibration can be computationally hard.
  - Few have tried it on exotic options.[a]
- There are usually several local minima.[b]
  - They will give different prices to options not used in the calibration.
  - But which one captures the smile dynamics?

---

[a] Ayache, Henrotte, Nassar, & X. Wang (2004).
[b] Ayache (2004).

Continuous-Time Derivatives Pricing

I have hardly met a mathematician
who was capable of reasoning.
— Plato (428 B.C.–347 B.C.)

Fischer [Black] is the only real genius
I've ever met in finance. Other people,
like Robert Merton or Stephen Ross,
are just very smart and quick,
but they think like me.
Fischer came from someplace else entirely.
— John C. Cox, quoted in Mehrling (2005)

Toward the Black-Scholes Differential Equation

- The price of any derivative on a non-dividend-paying stock must satisfy a partial differential equation (PDE).
- The key step is recognizing that the same random process drives both securities.
  - Their prices are perfectly correlated.
- We then figure out the amount of stock such that the gain from it offsets exactly the loss from the derivative.
- The removal of uncertainty forces the portfolio's return to be the riskless rate.
- PDEs allow many numerical methods to be applicable.

Assumptions[a]

- The stock price follows $dS = \mu S\,dt + \sigma S\,dW$.
- There are no dividends.
- Trading is continuous, and short selling is allowed.
- There are no transactions costs or taxes.
- All securities are infinitely divisible.
- The term structure of riskless rates is flat at r.
- There is unlimited riskless borrowing and lending.
- t is the current time, T is the expiration time, and $\tau \stackrel{\Delta}{=} T - t$.

[a]Derman & Taleb (2005) summarizes criticisms on these assumptions and the replication argument.

Black-Scholes Differential Equation

- Let C be the price of a derivative on S.
- From Ito's lemma (p. 588),

$$dC = \left(\mu S \, \frac{\partial C}{\partial S} + \frac{\partial C}{\partial t} + \frac{1}{2} \, \sigma^2 S^2 \, \frac{\partial^2 C}{\partial S^2} \right) dt + \sigma S \, \frac{\partial C}{\partial S} \, dW.$$

  - The same W drives both C and S.
- Short one derivative and long $\partial C / \partial S$ shares of stock (call it Π).
- By construction,

$$\Pi = -C + S(\partial C / \partial S).$$

Black-Scholes Differential Equation (continued)

- The change in the value of the portfolio at time dt is[a]

$$d\Pi = -dC + \frac{\partial C}{\partial S}\, dS.$$

- Substitute the formulas for dC and dS into the partial differential equation to yield

$$d\Pi = \left(-\frac{\partial C}{\partial t} - \frac{1}{2}\,\sigma^2 S^2\, \frac{\partial^2 C}{\partial S^2} \right) dt.$$

- As this equation does not involve dW, the portfolio is riskless during dt time: $d\Pi = r\Pi\, dt$.

---

[a] Bergman (1982) argues it is not quite right. But see Macdonald (1997).

Black-Scholes Differential Equation (continued)

- So
$$\left(\frac{\partial C}{\partial t} + \frac{1}{2}\, \sigma^2 S^2\, \frac{\partial^2 C}{\partial S^2} \right) dt = r \left(C - S\, \frac{\partial C}{\partial S} \right) dt.$$
- Equate the terms to finally obtain
$$\frac{\partial C}{\partial t} + rS\, \frac{\partial C}{\partial S} + \frac{1}{2}\, \sigma^2 S^2\, \frac{\partial^2 C}{\partial S^2} = rC.$$
- This is a backward equation, which describes the dynamics of a derivative's price *forward* in physical time.

Black-Scholes Differential Equation (concluded)

- When there is a dividend yield q,

$$\frac{\partial C}{\partial t} + (r - q)\, S\, \frac{\partial C}{\partial S} + \frac{1}{2}\, \sigma^2 S^2\, \frac{\partial^2 C}{\partial S^2} = rC. \tag{86}$$

- The local-volatility model (83) on p. 616 is simply the dual of this equation:[a]

$$\frac{\partial C}{\partial T} + (r_T - q_T) X \frac{\partial C}{\partial X} - \frac{1}{2}\, \sigma(X, T)^2 X^2 \frac{\partial^2 C}{\partial X^2} = -q_T C.$$

- This is a forward equation, which describes the dynamics of a derivative's price *backward* in maturity time.

---

[a]Derman & Kani (1997).

Rephrase

- The Black-Scholes differential equation can be expressed in terms of sensitivity numbers,

$$\Theta + rS\Delta + \frac{1}{2}\,\sigma^2 S^2 \Gamma = rC. \tag{87}$$

- Identity (87) leads to an alternative way of computing Θ numerically from Δ and Γ.
- When a portfolio is delta-neutral,

$$\Theta + \frac{1}{2}\,\sigma^2 S^2 \Gamma = rC.$$

  – A definite relation thus exists between Γ and Θ.

Black-Scholes Differential Equation: An Alternative

- Perform the change of variable $V \stackrel{\Delta}{=} \ln S$.
- The option value becomes $U(V, t) \stackrel{\Delta}{=} C(e^V, t)$.
- Furthermore,

$$\begin{aligned} \frac{\partial C}{\partial t} &= \frac{\partial U}{\partial t}, \\ \frac{\partial C}{\partial S} &= \frac{1}{S} \frac{\partial U}{\partial V}, && (88) \\ \frac{\partial^2 C}{\partial^2 S} &= \frac{1}{S^2} \frac{\partial^2 U}{\partial V^2} - \frac{1}{S^2} \frac{\partial U}{\partial V}. && (89) \end{aligned}$$

Black-Scholes Differential Equation: An Alternative (concluded)

- Equations (88) and (89) are alternative ways to calculate delta and gamma.[a]
- They are particularly useful for a tree of *logarithmic* prices.
- The Black-Scholes differential equation (86) on p. 651 becomes

$$\frac{1}{2}\,\sigma^2\,\frac{\partial^2 U}{\partial V^2} + \left(r - q - \frac{\sigma^2}{2}\right)\frac{\partial U}{\partial V} - rU + \frac{\partial U}{\partial t} = 0$$

subject to $U(V, T)$ being the payoff such as $\max(X - e^V, 0)$.

[a]See Eqs. (47) on p. 351 and (48) on p. 353.

[Black] got the equation [in 1969] but then was unable to solve it. Had he been a better physicist he would have recognized it as a form of the familiar heat exchange equation, and applied the known solution. Had he been a better mathematician, he could have solved the equation from first principles. Certainly Merton would have known exactly what to do with the equation had he ever seen it.

— Perry Mehrling (2005)

PDEs for Asian Options

- Add the new variable $A(t) \stackrel{\Delta}{=} \int_0^t S(u)\, du$.
- Then the value V of the Asian option satisfies this two-dimensional PDE:[a]

$$\frac{\partial V}{\partial t} + rS\, \frac{\partial V}{\partial S} + \frac{1}{2}\, \sigma^2 S^2\, \frac{\partial^2 V}{\partial S^2} + S\, \frac{\partial V}{\partial A} = rV.$$

- The terminal conditions are

$$\begin{aligned} V(T, S, A) &= \max\left(\frac{A}{T} - X, 0 \right) \quad \text{for call,} \\ V(T, S, A) &= \max\left(X - \frac{A}{T}, 0 \right) \quad \text{for put.} \end{aligned}$$

[a] Kemna & Vorst (1990).

PDEs for Asian Options (continued)

- The two-dimensional PDE produces algorithms similar to that on pp. 426ff.[a]
- But one-dimensional PDEs are available for Asian options.[b]
- For example, Večeř (2001) derives the following PDE for Asian calls:

$$\frac{\partial u}{\partial t} + r\left(1 - \frac{t}{T} - z\right)\frac{\partial u}{\partial z} + \frac{\left(1 - \frac{t}{T} - z\right)^2 \sigma^2}{2}\frac{\partial^2 u}{\partial z^2} = 0$$

  with the terminal condition $u(T, z) = \max(z, 0)$.

---

[a] Barraquand & Pudet (1996).

[b] Rogers & Shi (1995); Večeř (2001); Dubois & Lelièvre (2005).

PDEs for Asian Options (concluded)

- For Asian puts:

$$\frac{\partial u}{\partial t} + r\left(\frac{t}{T} - 1 - z\right)\frac{\partial u}{\partial z} + \frac{\left(\frac{t}{T} - 1 - z\right)^2 \sigma^2}{2}\frac{\partial^2 u}{\partial z^2} = 0$$

  with the same terminal condition.

- One-dimensional PDEs lead to highly efficient numerical methods.

Hedging

When Professors Scholes and Merton and I
invested in warrants,
Professor Merton lost the most money.
And I lost the least.
— Fischer Black (1938–1995)

Delta Hedge

- The delta (hedge ratio) of a derivative f is defined as

$$\Delta \triangleq \frac{\partial f}{\partial S}.$$

- Thus

$$\Delta f \approx \Delta \times \Delta S$$

  for relatively small changes in the stock price, ΔS.

- A delta-neutral portfolio is hedged as it is immunized against small changes in the stock price.
- A trading strategy that dynamically maintains a delta-neutral portfolio is called delta hedge.

Delta Hedge (concluded)

- Delta changes with the stock price.
- A delta hedge needs to be rebalanced periodically in order to maintain delta neutrality.
- In the limit where the portfolio is adjusted continuously, "perfect" hedge is achieved and the strategy becomes self-financing.

Implementing Delta Hedge

- We want to hedge N *short* derivatives.
- Assume the stock pays no dividends.
- The delta-neutral portfolio maintains $N \times \Delta$ shares of stock plus B borrowed dollars such that

$$-N \times f + N \times \Delta \times S - B = 0.$$

- At next rebalancing point when the delta is Δ', buy $N \times (\Delta' - \Delta)$ shares to maintain $N \times \Delta'$ shares.
- Delta hedge is the discrete-time analog of the continuous-time limit and will rarely be self-financing.

Example

- A hedger is *short* 10,000 European calls.
- $S = 50$, $\sigma = 30\%$, and $r = 6\%$.
- This call's expiration is four weeks away, its strike price is \$50, and each call has a current value of $f = 1.76791$.
- As an option covers 100 shares of stock, $N = 1{,}000{,}000$.
- The trader adjusts the portfolio weekly.
- The calls are replicated well if the cumulative cost of trading *stock* is close to the call premium's FV.[a]

[a] This example takes the replication viewpoint.

Example (continued)

- As $\Delta = 0.538560$

$$N \times \Delta = 538,560$$

shares are purchased for a total cost of

$$538{,}560 \times 50 = 26{,}928{,}000$$

dollars to make the portfolio delta-neutral.

- The trader finances the purchase by borrowing

$$B = N \times \Delta \times S - N \times f = 25{,}160{,}090$$

dollars net.[a]

---

[a]This takes the hedging viewpoint — an alternative. See Exercise 16.3.2 of the text.

Example (continued)

- At 3 weeks to expiration, the stock price rises to \$51.
- The new call value is $f' = 2.10580$.
- So the portfolio is worth

$$-N \times f' + 538{,}560 \times 51 - Be^{0.06/52} = 171,622$$

before rebalancing.

Example (continued)

- A delta hedge does not replicate the calls perfectly; it is not self-financing as \$171,622 can be withdrawn.
- The magnitude of the tracking error—the variation in the net portfolio value—can be mitigated if adjustments are made more frequently.
- In fact, the tracking error over *one* rebalancing act is positive about 68% of the time, but its expected value is essentially zero.[a]
- The tracking error at maturity is proportional to vega.[b]

---

[a] Boyle & Emanuel (1980).
[b] Kamal & Derman (1999).

Example (continued)

- In practice tracking errors will cease to decrease beyond a certain rebalancing frequency.
- With a higher delta $\Delta' = 0.640355$, the trader buys

$$N \times (\Delta' - \Delta) = 101,795$$

  shares for \$5,191,545.
- The number of shares is increased to $N \times \Delta' = 640,355$.

Example (continued)

- The cumulative cost is[a]

$$26{,}928{,}000 \times e^{0.06/52} + 5{,}191{,}545 = 32{,}150{,}634.$$

- The portfolio is again delta-neutral.

[a]We take the replication viewpoint here.

		Option value	Delta	Change in delta	No. shares bought	Cost of shares	Cumulative cost
τ	S	f	Δ		$N \times (5)$	$(1) \times (6)$	FV(8')+(7)
	(1)	(2)	(3)	(5)	(6)	(7)	(8)
4	50	1.7679	0.53856	—	538,560	26,928,000	26,928,000
3	51	2.1058	0.64036	0.10180	101,795	5,191,545	32,150,634
2	53	3.3509	0.85578	0.21542	215,425	11,417,525	43,605,277
1	52	2.2427	0.83983	−0.01595	−15,955	−829,660	42,825,960
0	54	4.0000	1.00000	0.16017	160,175	8,649,450	51,524,853

The total number of shares is 1,000,000 at expiration (trading takes place at expiration, too).

Example (concluded)

- At expiration, the trader has 1,000,000 shares.
- They are exercised against by the in-the-money calls for \$50,000,000.
- The trader is left with an obligation of

$$51{,}524{,}853 - 50{,}000{,}000 = 1{,}524{,}853,$$

which represents the replication cost.

- Compared with the FV of the call premium,

$$1{,}767{,}910 \times e^{0.06 \times 4/52} = 1{,}776{,}088,$$

the net gain is $1{,}776{,}088 - 1{,}524{,}853 = 251{,}235$.