Chapter 6: The Pricing of American Put and Call Options

DISCLAIMER: THIS CHAPTER SHOULD NOT BE CONSTRUED AS INVESTMENT ADVICE. I DON'T WANT TO GET IN TROUBLE WITH THE SECURITIES-AND-EXCHANGE COMMISSION. OPTIONS CAN BE VERY RISKY; RECENTLY, A MULTI-BILLION-DOLLAR INVESTMENT BANK WITH A 200-YEAR HISTORY WENT BANKRUPT BECAUSE OF BAD CHOICES IN DERIVATIVE SECURITIES.¹

Stocks and Options

They sought it with thimbles, they sought it with care;
They pursued it with forks and hope;
They threatened its life with a railway share;
They charmed it with smiles and soap.

To raise money for expansion, a company may sell shares, each representing partial ownership. Anyone may subsequently buy or sell shares on a secondary market, such as a stock exchange. The price at which these shares sell fluctuates on a timescale of minutes or less and is influenced by (among other factors) the company’s recent performance, its ability to pay dividends to shareholders, the public perception of what it’s likely to do in the next months or years, interest rates, psychology, and mass hysteria.

Institutions, such as pension funds, and individuals who invest in stocks are not always comfortable with the idea that the future value of their investments may be a function of mass hysteria. Other people like volatility and seek a way of realizing higher gains than possible with a stock (at the risk of greater losses). Options and similar derivatives provide a means to manage risks and potential gains. Options and futures contracts are probably even more important in commodity (e.g., cotton, copper, pork bellies) and currency markets, where large fluctuations would otherwise seriously threaten a farmer’s, supplier’s, or international company’s ability to carry on business. Recently, they have received the attention of electric utilities.² We consider stocks in this course, because data on stock options are easily accessible.

Perhaps the simplest derivative investment is a futures contract, in which two parties agree to carry out a transaction (for instance, the sale of 100 metric tonnes of pork bellies) on a particular date at a particular price. A European put or call is similar, except it is optional on one party. The bearer of a call contract has the right, but not the obligation, on the expiration date to purchase 100 shares of a particular stock at a previously-agreed strike price. Obviously, he will do so only if the strike price is less than the market price on that day. The bearer of a put contract has the right, but not the obligation, on the expiration date to sell 100 shares of a particular stock at a previously-agreed strike price. He will do so only if the strike price is greater than the market. Like stocks and bonds, options trade on the secondary market. Unlike securities, an individual may issue (write) an option, which then becomes binding on him.

An American option permits exercise on the same terms any time on or before the expiration date. An option that can be exercised is termed “in the money;” for a call, this means that the

¹ http://www.numa.com/ref/barings/index.htm . I understand that Nick Leeson was one of the few derivatives traders without a Ph.D. in math or physics.

² I wrote that sentence in 1999, well before the California power crisis of spring 2001. One of the peculiar features of California electric deregulation was a prohibition on the use of futures contracts by the utility companies.
exercise (or strike) price $E$ is less than the current market value $S$ of the stock, while for an in-the-money put, $E > S$. The geographical names appear to be historic; both kinds of options trade on all continents. However, stock options (not index options) trading on the Chicago Board and other options exchanges in the United States are all of the American variety. Other kinds of options exist; one that’s half-way between the two is called “Bermudan,” while exotic options involving complicated functions of a stock’s history are generically referred to as “Asian.”

My treatment below mostly follows that of Wilmott et al., *Option Pricing* (see the reference list toward the end).

Why this is physics

I am told that a majority of the people trading derivative securities for a living have Ph.D.s in either physics or mathematics. Beside the obvious necessity for quantitative analysis, the reason becomes clear once we consider the fundamental model used in the industry. As with most physical models, we begin with a simplification: all the many factors going in to the price of a stock are inherently unpredictable and so constitute essentially noise. (If I could reliably predict the future price of a stock, I would quickly make enough money to buy out the National Science Foundation and to rename USF). On the other hand, over time, stocks tend to appreciate in value. So we model the return on a stock as a biased random walk; on average, the rate of return (i.e., percentage per year) might be $\mu$, but on top of this we add noise.

Thus if we know the current stock price, $S$, we model the infinitesimal increment $dS$ over an infinitesimal period of time $dt$ as

$$dS(t) = \mu S(t) dt + \sigma S(t) dX(t) \quad ,$$

where $dX(t)$ represents a random variable, positive or negative, drawn from a Gaussian distribution of variance $dt$ and completely uncorrelated with $dX(t')$ for $t \neq t'$. I will show that under these assumptions, $\ln S$ follows a biased random walk.

The simplest investment: compound interest

As the simplest example of an investment, consider a deposit $S(0)$ left in a bank or United-States Treasury bill for a year, an investment carrying negligible risk to principal. Let the interest rate be $r$, for example 6% per annum, which we assume does not change. If the deposit pays simple interest, in one year it will be worth

$$S(1\text{year}) = (1 + r)S(0) = 1.06S(0).$$

If, however, for the same quoted annual rate, the bank compounds the interest monthly, each month for twelve months I will multiply the balance from the end of the previous month by the factor $(1 + r/12)$, or 1.005 if $r = 6\%$. Thus

$$S(1\text{year}) = (1 + r/12)^{12} S(0).$$

More generally,

$$S(t) = (1 + rt/n)^n S(0),$$

where $n$ is the number of periods in a time $t$.

The most liquid institutional money-market instruments might compound interest daily (multiplying each previous day’s balance by $1 + r/365.25$). It is convenient to take the limit as $n \to \infty$. By a simple binomial expansion and Taylor series, we find

$$S(t) = e^{rt} S(0) \quad ,$$

which solves the differential equation

$$\frac{dS}{dt} = rS(t) \quad (3)$$

for the initial condition $S(0)$. Equation (3) states that the money $S(t)$ increases proportionately at the rate $r$.

Money can be negative: that represents taking out a loan, which one might do in order to invest in stocks (not necessarily recommended). For the purposes of this course, we will assume that the
same interest rate $r$ governs both bank (or Treasury) investments and loans. We’ll pretend that $r$
does not change and so is known in advance and ignore any dependence on maturity (e.g., for a
Treasury instrument), assuming we can easily redeem our investment (or pay back our loan) at will.
Assume there are no costs (e.g., broker’s commissions) for any of the stock or other transactions
we make; this is not a bad approximation for large, institutional investors. We also ignore inflation
and all consequences of tax law.\footnote{By assuming the spherical horse, we have firmly established
this as physics. Of course, such considerations can be put in later.}

### Arbitrage and European put-call parity

Stocks, unlike bank deposits, involve risk; it is difficult to calculate the “right” value for a
stock. However, the Black-Scholes analysis provides a way of calculating the value of an option.
The analysis relies on the concept of arbitrage. If someone in New York is willing to sell a stock
at $1.125$, and someone in Philadelphia is willing to buy the same stock at $1.133$, an arbitrageur
will buy in New York and simultaneously sell in Philadelphia, making a riskless profit (we ignore
transaction costs). Similarly, if an option is selling at the “wrong” price, an arbitrageur can make
risk-free money by borrowing money and buying (if the price is low) or selling short (if it’s high)
and investing the proceeds.

As an example of the kind of arbitrage argument that later leads to Black-Scholes, consider
buying one unit of stock currently trading at price $S(t=0)$ and a European put with exercise price $E$
due to expire in time $T$. At the same time, write (sell) a European call with the same exercise
price and expiration.

While the correct price may not be apparent now, the future correct price at expiration is
very easy to calculate. If $S(T) < E$, the call expires worthless, but I can exercise the put, selling
the stock for $E$. On the other hand, if $S(T) > E$, my put is worthless, but the holder of the call
can exercise it, forcing me to sell my stock, again for $E$. If $S(T) = E$ exactly, both options are
worthless, and I sell my stock for $E$. Thus the payoff at expiration is $E$ in all cases, without any
risk.\footnote{Even if the company that issued $S$ declares bankruptcy, $S(T) = 0$ is just a special case of $S(T) < E$.}

Thus the portfolio $\Pi = S + P - C$ resembles a bank deposit, so it had better cost the same at
time $t = 0$ as does a bank deposit paying $E$ after time $T$; from (2), this is $Ee^{-rT}$. If $\Pi < Ee^{-rT}$,
an arbitrageur can borrow money at rate $r$ to buy more $\Pi$, risk free. Conversely, if $\Pi > Ee^{-rT}$, the
arbitrageur sells the stock short, writes a put, and buys a call, investing the free money at rate $r$.

Professionals with huge credit lines constantly look for such investments, and by taking advan-
tage of any almost immediately drive the market prices in the direction to make further arbitrage
impossible. Thus, if the markets operate efficiently, there should be no arbitrage opportunities left;
in particular (for European options),

$$S + P - C = Ee^{-rT}$$

This gives a relation between the three instruments (stock, put, call) and the (assumed) fixed
interest rate; Black-Scholes will provide a mechanism to price $P$ and $C$ individually.

### Random processes

Before explaining the noisy part of a stock price modeled in (1), let me review a more familiar
random process, the random walk. Starting at position $Y_0 = 0$, at each time step $i$, I add to
$Y$ a random variable $X_i$, which is $+1$ with probability $1/2$ and $-1$ with probability $1/2$. The
random variable is unbiased, \( \langle X_i \rangle = 0 \), and I assume that all the time steps are uncorrelated, \( \langle X_i X_j \rangle = \delta_{ij} \). Then at the time given by \( n \) timesteps, the position \( Y_n = \sum_{i=1}^{n} X_i \). The mean position \( \langle Y_n \rangle = \langle X_1 \rangle + \langle X_2 \rangle + \ldots = 0 \). but the mean square position \( \langle Y_n^2 \rangle = n \), which since the mean vanishes is also the variance. In other words, we go nowhere in particular, but we end up a distance \( \sqrt{\text{time}} \) (on average) from the origin.

The deviate, or random variable, \( X_i \) stands for an unknown value +1 or –1. If we wish to compute something definite, we can calculate the moments of \( X_i \), or equivalently the probability distribution of \( Y_n \). If \( n \) is even, the probability that \( Y_n = 2m \) is

\[
\text{Prob}_Y(2m) = \frac{1}{2^n} \binom{n}{n/2 - m}.
\]

Consider the large-\( n \) limit. By taking the logarithm, applying Stirling’s approximation, and Taylor expanding the resulting logarithms (being careful to expand to a consistent order in \( 1/n \)), we recover the well-known result that the binomial distribution reduces to a Gaussian with variance \( n/4 \):

\[
\lim_{n \to \infty} \text{Prob}_Y(2m) = \sqrt{2\pi ne^{-2m^2/n}}.
\]

This is just a special case of the central-limit theorem, which states that the sum of many independent random processes is Gaussian, provided the processes’ second moments don’t diverge. In particular, a sum of Gaussian processes is Gaussian. If we imagine the stock price receiving a tiny kick at every tiny time step, we’re going to end up with a Gaussian process anyway, so we might as well model the small kicks as also Gaussian. Otherwise the exact form of our noise will depend on the time scale.

The Gaussian process is conveniently scale invariant. Consider a random kick \( \delta X \) every time step \( \delta t \) (which need not be small) with probability density

\[
P_{X,\delta t}(\delta X) = \frac{1}{\sqrt{2\pi \delta t}} e^{-\delta X^2/(2\delta t)}.\]

This normalized Gaussian has variance \( \delta t \). What if we look only every \( 2\delta t \)? The probability distribution for the noise on the longer time scale (i.e., two time steps) is

\[
P(\delta X) = \int du P_{X,\delta t}(u) P_{X,\delta t}(\delta X - u),
\]

which after a little algebra comes out to exactly \( P_{X,2\delta t}(\delta X) \) in (7), thus establishing the scale invariance of uncorrelated Gaussian noise.

We see explicitly in (7) that \( \delta X \sim (\delta t)^{1/2} \), just as in the binary random walk.

**Estimating volatility**

If we assume Gaussian-distributed uncorrelated noise as in (1) and (7), we can measure the historical volatility of a stock over some period as the suitably normalized standard deviation of the population of periodic returns on the stock.\(^5\)

\(^5\) The convergence is non-uniform, which may account for the fact that experiments often show larger tails than would be predicted.

\(^6\) Insofar as dividends represent return of stock-holder equity, stock prices should be adjusted for them. The Yahoo Web site referred to toward the end of the chapter does this, or you can avoid the subtlety by picking stocks that pay no dividends. Such stocks have tended to be more interesting lately anyway. I will provide some scripts and programs to help out.
6. The Pricing of American Put and Call Options

As a practical matter, daily stock closing prices are more readily available than hourly or more frequent measures, so I'll assume you have a list of dates and closing values. The rate of return between day \( t_i \) and day \( t_{i-1} \) is

\[
R_i = \frac{S(t_i) - S(t_{i-1})}{(t_i - t_{i-1})S(t_{i-1})},
\]

indicating the fractional profit or loss in the period \( t_i - t_{i-1} \). I write \( t_i - t_{i-1} \) in the denominator instead of 1 day, because stocks do not trade publicly on weekends or holidays; intervals over these times will be longer. If you want an annualized rate of return, you will need to measure time in years (days divided by 365.25, roughly); you can do the division later. The units of \( R_i \) are \( \text{year}^{-1} \).

Before scaling, the volatility \( \bar{\sigma} \) is simply the standard deviation, estimable as usual from the population by

\[
\bar{\sigma} = \sqrt{\frac{1}{N-1} \sum (R_i - R)^2},
\]

where \( N \) is the number of returns and \( R \) the average periodic return. We have already discussed robust one-pass and two-pass algorithms for the numerical computation of (10).

However, since the standard deviation of \( \delta X \) in (7) depends on the time scale \( \delta t \), we must multiply by the square root of the time scale we measured:

\[
\sigma = \bar{\sigma} \sqrt{\delta t}.
\]

As required in (1), \( \sigma \) has units of \( \text{year}^{-1/2} \). The typical time scale \( \delta t \approx 1 \) day.

One might argue in favor of using \( \delta t = \langle t_i - t_{i-1} \rangle \) and weighting daily returns differently after weekends and holidays; I found reasonably close agreement with published volatility estimates using the procedure just described. Although our model (1) assumes a fixed \( \sigma \), the estimate of \( \sigma \) from any sample population will vary depending on the days included. Furthermore, the model may not apply well to stocks that undergo long periods of stasis before suddenly becoming volatile. Many sources uses a trailing period of 30 days to estimate volatility; in the assignment, you will compare this estimate to that inferred by fitting the Black-Scholes model to actual option trades.

**Stochastic differential equations and Itô’s lemma**

Equities do not trade on a discrete time grid; they trade whenever a buyer and a seller agree on a price, and in any case it is more convenient analytically to deal with continuous than with discrete time. The noise term \( dX(t) \) in (1) is supposed to represent the limit of \( \delta X \) for infinitesimal time, \( \delta t \rightarrow dt \). This sort of equation is common in hydrodynamics, statistical mechanics, and plasma physics; dividing through by \( dt \), we write

\[
\dot{S}(t) = \mu S(t) + \sigma S(t)L(t),
\]

immediately recognizable as a Langevin equation, with \( L(t) = dX/dt \) a noise term satisfying \( \langle L(t) \rangle = 0 \) and \( \langle L(t)L(t') \rangle = \delta(t - t') \).

Because we imagine our kicks to occur instantaneously and infinitely often, they are also unbounded; their density makes them behave in some sense worse than isolated Dirac-\( \delta \) functions. Put another way, the Wiener process \( X(t) = \int_0^t dX(t')dt' \) is continuous, but its derivative \( L(t) \) exists nowhere. It is not surprising, then, that the usual rules of calculus fail to apply to symbols such as \( dX \) and \( dS \) and to equations such as (1) and (12).

First, one must note that these two equations are in fact ill-posed without additional information. In integrating the last term of (1), are we to write

\[
\int_0^t S(t)dX(t) = \lim_{n \rightarrow \infty} \sum_{j=0}^{n-1} S(jt/n)dX(jt/n)
\]

\[7\] The higher moments factor; see van Kampen, Stochastic Processes in Physics and Chemistry, Elsevier, 1981.
or

\[ \int_0^t S(t) dX(t) = \lim_{n \to \infty} \sum_{j=0}^{n-1} (1/2)[S(jt/n) + S((j+1)t/n)] dX(jt/n) \quad ? \]

We have discussed such choices numerically in the context of quadrature, but mathematically we have always assumed them equivalent in the limit. For stochastic variables, however, they lead to different results. The second choice is due to Stratonovich, the first to Ito; see the reference of footnote 7. In (14), the instantaneous change in the asset \( S \) is determined (via (12)) in part by its future value. To avoid such a non-physical (or advanced) interpretation, we use (13) exclusively.

To make use of equations such as (1), we require the transformation laws that replace the chain rule of classical calculus. A proper treatment begins from one of two starting points. The first very carefully Taylor expands \( dX \) and multiply by \( dt \) to get

\[ df = \left( \mu S \frac{\partial f}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} + \frac{\partial f}{\partial t} \right) dt + \sigma S \frac{\partial f}{\partial S} dX \quad ; \]

as a particular example, take \( f(S) = \ln S \). Then \( df \) obeys the equation

\[ df = (\mu - \frac{1}{2} \sigma^2) dt + \sigma dX \quad . \]

This is simpler than (1), stating that the logarithmic asset price, \( f \), follows a normally distributed, biased random walk. For this reason, \( S \) in (1) is called a log-normal walk.

Note that (19) is not what one would get naively. For example, if we were allowed to write \( \dot{f} = \dot{S}/S \) (we’re not), we could divide both sides of (12) by \( S \) and multiply by \( dt \) to get

\[ df = \mu dt + \sigma dX \quad , \]

differing from (19) by the absence of the extra drift piece \( \frac{1}{2} \sigma^2 \). At first this seems paradoxical: setting \( \mu = 0 \), we see from (19) that, on average, \( f = \ln S \) decreases with time. (Don’t worry about this as a model of stock price: we can always restore the positive \( \mu \).) However, doesn’t (1) (or (12)) appear to be neutral with \( \mu = 0 \)? No: consider the discrete-time version. Imagine we start with $1 and that the first two random noises are \( \delta X_1 = +.1 \) and

6. The Pricing of American Put and Call Options
δX₂ = −.1. We end up with 0.9 · 1.1 · $1 = $0.99. Conversely, if δX₁ = −.1 and δX₂ = +.1, we again hold at the end 1.1 · 0.9 · $1 = $0.99 (obviously the same, since, as Tom Lehrer sings, multiplication is commutative). Clearly the slow average drift continues to hold for any number of multiplied factors so long as we obey the choice (13); even in the limit, the unbiased dX leads to a biased df. This helps make the result (19), and the need for (18), more plausible.⁸

Black-Scholes

We’re now ready to apply an arbitrage argument to the pricing of options. We require one result from the previous section, (18), which replaces the chain rule of ordinary calculus when applied to the specific equation (1).

The price V(S, t) of a put or call option (consider European options at first) plausibly depends on its exercise price, the time T − t to expiration, the price S(t) of the underlying asset, and the model parameters µ and σ that describe the (expected, or guessed) long-time drift of the stock and the strength of stochastic noise. I’ll delay consideration of a dividend yield on the stock.

We now construct a simpler portfolio than (4): we purchase some derivative V (put or call) and sell short a relative fraction of shares ∆ in the underlying, so that the entire portfolio is

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

(21)

Linear operations on stochastic variables are safe, so we can also write dΠ = dV − ∆dS, assuming we hold ∆ constant. Engaging in the arbitrage technique (21) is called “Delta hedging;” in practice, the arbitrager has to re-calculate and change ∆ quite often. In three lines, we’ll decide what ∆ should be to yield a portfolio instantaneously without risk.

Itô’s lemma (18) tells us

$$dV = \left( \mu S \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \frac{\partial V}{\partial t} \right) dt + \sigma S \frac{\partial V}{\partial S} dX ,$$  

(22)

so that the differential portfolio value is given by

$$d\Pi = \left( \sigma S \frac{\partial V}{\partial S} - \Delta \sigma S \right) dX + \left( \mu S \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \frac{\partial V}{\partial t} - \mu S \Delta \right) dt .$$  

(23)

(I’ve plugged in (1).) By setting ∆ = ∂V/∂S (evaluated at the current time and price), I can eliminate the entire stochastic term (dX), yielding a completely deterministic equation:

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

(24)

Because we know (within the model) exactly what Π will be worth in a short time, the portfolio carries no risk. Appealing again to the principle that all risk-free investments must bear the same yield (else make George Soros very wealthy), we decide that

$$d\Pi = r\Pi dt ,$$  

(25)

⁸ It’s unfortunate the financial people didn’t write

$$dS(t) = (\mu + \frac{1}{2} \sigma^2)S(t) + \sigma S(t)dX(t)$$

instead of (1), but one can always think of the drift absorbing the extra term.
where \( r \) is the rate on short-term United-States Treasury bills. Combining (21), (24), and (25), we get

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

the Black-Scholes equation.

Not only has the choice of \( \Delta \) between (23) and (24) eliminated randomness: it’s also got rid of the drift term, \( \mu \). Amazingly, while the price of an option may depend on many things, it does not depend on whether the overall trend of the underlying stock is up or down. This contradicts common sense, since if a stock (say AMZN) falls, puts will increase in value and calls fall. For someone who holds on to an option for a long time (or until expiration), \( \mu \) certainly does matter. So why is it absent in the Black-Scholes equation?

The first answer is that the Black-Scholes analysis involves only an instant of time. From our review of random walks, in which \( dX \sim \sqrt{dt} \), it is clear that for short enough time scales, the stochastic term in \( dS \), proportional to \( dX \), completely dominates the drift term proportional to \( dt \), so \( \mu \) doesn’t matter. A second answer is that Black-Scholes argues from arbitrage. Since any arbitrage opportunity yields infinite profits, at least in the model, it mustn’t exist. An infinite constraint overcomes other influences. Finally, the model (1) fits stock prices moderately well after the fact, and one may even be able to estimate \( \sigma \) in advance, if it doesn’t change too quickly (we’ve assumed it doesn’t change at all). However, if anyone could predict \( \mu \) accurately, she’d have a free-money machine quite independent of options.

Note that Black-Scholes uses a self-consistent argument: we assume we can calculate the value \( V(S,t) \) of an option and its derivatives in order to derive an equation that does let us calculate them. The argument most likely needs some modification in the real world, particularly when the options markets have limited liquidity (few people interested in buying or selling).

Any physicist or applied mathematician should be able to find numerous extensions to the Black-Scholes model, and a few appear in the books cited in the reference section. However, here is one important difference between finance and physics: if you come up with a really good improvement, don’t publish! Unlike stocks, options are a zero-sum game: for every dollar someone makes, someone else loses a dollar. So as soon as everyone else catches on to your new trick, it ceases to be useful.

One improvement I’ll quote (see Wilmott for the derivation): if a stock pays a dividend at rate \( D_0 \), and if we imagine that the company disburses the dividend continuously instead of in one lump each quarter (more spherical horses), the modified Black-Scholes equation reads

\[
\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r - D_0)S \frac{\partial V}{\partial S} - rV = 0 .
\]

(27)

American options and free boundary conditions

For European options, Black-Scholes has a closed-form solution; see page 100 in Wilmott et al. The early-exercise feature of American puts and calls, however, ensures a minimum value for in-the-money options. As an extreme example, consider a stock with no volatility and no drift (maybe it’s been nationalized by an evil government, which promises to buy and sell at a fixed price). A European put option at \( E > S \) (now \( S \) is fixed for all time) is clearly worth money, but it’s worth less than \( E - S \), because it can’t be exercised for a time \( T \). Specifically, right now (time 0), it’s worth \((E - S)e^{-rT}\). However, the same option with American rules is worth \( E - S \), since the holder can exercise it immediately. (It isn’t worth any more than \( E - S \) if the price of \( S \) truly is fixed for all time; this is of course an artificial example.)
So long as they’re worth more than $E - S$ (put, or $S - E$, call), American options are subject to the same arguments that led to Black-Scholes, (27). However, whenever an in-the-money option would otherwise fall below $|E - S|$, that “payoff” function provides the floor for the option.

This does not mean that $C_{\text{American}} = \max(C_{\text{European}}, S - E)$ (call) or that $P_{\text{American}} = \max( P_{\text{European}}, E - S)$ (put). This is because a partial differential equation, such as Black-Scholes, determines the value of a function at some point in space and time based on the value of the function at neighboring points of space and time. However, these points affect their neighbors, et cetera, out to all regions of the time-price plane. Thus the conditions (I’ll drop geographic subscripts and assume everything is American)

$$C(S, t) \geq \max(S - E, 0)$$

and

$$P(S, t) \geq \max(E - S, 0)$$ (28)

affect the solution globally, just as the initial value determines the solution globally in an ordinary differential equation, and just as the boundary values determine the solution to Laplace’s equation in electrostatics. However, the boundary conditions (28) suffer from one complication: we don’t know in advance where in $(S, t)$ they apply. Such conditions, for which we must solve as we solve the differential equation, are called free boundaries. A closely related physical example is Stefan’s problem of the propagation of the solid-liquid front in melting ice.\(^9\)

Wilmott et al. suggest a simple example of a free-boundary problem: consider a taut string tied firmly to the $x$ axis at points ±1 that must pass over an obstacle $f(x) = \max(1 - (2x + 0.5)^2, 0) + \max(0.5 - 0.5(2x - .8)^2, 0)$ (see figure 1).

---

so we can also write \( y(x) \geq f(x) \). One particularly elegant solution to the problem combines the differential equation with the constraint to write

\[
y''(y - f) = 0 ,
\]

accompanied by the additional conditions \( y'' \geq 0 \) and \( y - f \geq 0 \) and the fixed boundary conditions \( y(-1) = y(0) = 0 \). Tautness also requires continuity of \( y' \). Since the possibility of early exercise of an American-style option turns Black-Scholes into a similar inequality, we can formulate our current problem similarly.

**Scaling the differential inequality**

It is a familiar theme in theoretical physics to rescale dynamical variables into dimensionless forms independent of irrelevant measuring conventions. In the Lennard-Jones problem, once we scaled the radial distance by the Lennard-Jones length \( \sigma \), the only physical length in the problem, it no longer mattered whether we measured distance in meters, yards, or parsecs. Similarly, a stock split\(^\text{11}\) should not change the scaled variables.

In finance, money plays the role of distance.\(^\text{12}\) The only known length (\$) scale in the problem is the exercise price, \( E \), of the option. Furthermore, we should take the logarithm of stock price \( S \); our model (1) of noise affects prices proportionately, not absolutely, so we exercise a random walk of scale independent of stock price only when we take the logarithm, as in (19), defining a scaled asset price \( x \) by

\[
S(t) = E e^x \\
x = \ln(S/E) .
\]

Note that the transformation preserves the sense of stock price increasing or decreasing and that it maps the interval \((0, \infty)\) to \((-\infty, \infty)\).

The important dynamical time in the problem is time remaining until expiration of the option, \( T - t \). Interest rate, dividend yield, and volatility all set time scales. (So does drift \( \mu \), but it plays no role in the Black-Scholes equation (27).) Note from (1) or from (11) that \( \sigma \) has units of \((\text{time})^{-1/2}\); an additional factor of two turns out to be convenient, so we write

\[
t = T - 2\tau/\sigma^2 .
\]

Finally, we scale the option price \( V \) by the only length scale in the problem, \( E \), so

\[
V(S,t) = Ev(x,\tau) .
\]

In the variables transformed by (30)–(32), Black-Scholes (27) appears as

\[
\frac{\partial v}{\partial \tau} = \frac{\partial^2 v}{\partial x^2} + \frac{\partial v}{\partial x}(k_2 - 1) - k_1 v ,
\]

\(^10\) Except if the obstacle has a sharp edge, which possibility we exclude to avoid cutting our string. Without the continuity condition, we could propose \( y(x) = \max(0, f(x)) \), which obviously violates tautness.

\(^11\) A company may declare a stock split, at which (typically) every share of outstanding stock is replaced by two. The market value per share immediately falls in half, and option exercise prices are automatically adjusted. Companies sometimes do this if their stock price has risen so much that the average investor can no longer afford an even lot of 100 shares. More often, the motivation is purely psychological.

\(^12\) This contradicts the folk wisdom that “time is money.” In fact, since \( dX \sim \sqrt{dt} \), one might more properly say “time is money squared.”
where
\[ k_1 = \frac{2r}{\sigma^2} \]
and
\[ k_2 = \frac{2(r - D_0)}{\sigma^2} \]  
(34)
are dimensionless constants describing interest and dividend rates in terms of the time scale set by volatility.\(^\text{13}\)

Equation (33) resembles a diffusion equation with additional source terms. We can make these vanish by one additional transformation of the scaled option value \(v\): plugging
\[ v(x, \tau) = e^{\alpha x + \beta \tau} u(x, \tau) \]  
(35)
in to (33), with
\[ \alpha = -\frac{1}{2}(k_2 - 1) \]
\[ \beta = -\left(\frac{1}{4}(k_2 - 1)^2 + k_1\right) \]  
(36)
we get the pure diffusion equation
\[ \frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial x^2} \]  
(37)

Since the early-exercise feature of an American option can only make the value higher, never lower, with time than it would be otherwise, we should really write
\[ \frac{\partial u}{\partial \tau} \geq \frac{\partial^2 u}{\partial x^2} \]  
(38)

To set boundary conditions, first note that the option value at expiration, \(\tau = 0\), is the payoff \(\max(S - E, 0)\) for the call or \(\max(E - S, 0)\) for the put. Transformed into dimensionless variables, this takes the less pleasing form
\[ g(x, \tau) = \begin{cases} 
  e^{-\alpha x - \beta \tau} \max(e^x - 1, 0) & \text{call} \\
  e^{-\alpha x - \beta \tau} \max(1 - e^x, 0) & \text{put} 
\end{cases} \]  
(39)

The transformation (35) introduced the explicit time variable, \(\tau\), which we can set equal to zero for the “final” condition analogous to the initial condition we studied for ordinary differential equations. We leave it in (39), because the free boundary condition for American options takes the identical form evaluated at (backward) time \(\tau > 0\). Thus both the free and the initial boundary conditions are expressed by (39) and
\[ u(x, \tau) \geq g(x, \tau) \]  
(40)

We know two other fixed boundaries: if the stock price \(S\) falls to zero, the holder should certainly exercise a put, since the stock cannot fall below zero. Thus \(P(0, t) = E\). Similarly, in the limit of infinite stock price, the holder of a call should exercise: \(\lim_{S \to \infty} C(S, t) = S - E\). Conversely, the put becomes worthless in the limit \(S \to \infty\), and the call worthless if the company ever declares bankruptcy.\(^\text{14}\)

Expressed in dimensionless variables, for either type of option
\[ \lim_{x \to \pm \infty} u(x, \tau) = \lim_{x \to \pm \infty} g(x, \tau) \]  
(41)

\(^{13}\) I borrow the notation used by Wilmott et al.

\(^{14}\) The model (1) does not permit the stock ever to recover from \(S = 0\).
In addition, there is a somewhat subtle condition of continuity. I refer the reader to Wilmott et al., who demonstrate an arbitrage opportunity at the free boundary (the point on one side of which \( V(S(t), t) \) satisfies (28) as an equality) unless \( \partial V / \partial S \) is continuous across the boundary and so equal to \(-1\). For the payoff functions of our model (the right-hand sides of (28)), \( \partial V / \partial S \), and hence \( \partial u / \partial x \), must simply be continuous everywhere.\(^{15}\)

By analogy to the obstacle problem, (29), we express the differential inequality with free boundaries as

\[
\left( \frac{\partial u}{\partial \tau} - \frac{\partial^2 u}{\partial x^2} \right) \left( u(x, \tau) - g(x, \tau) \right) = 0
\]

(42)

with the conditions (38), (40), and (41).

In the context of Brownian motion, we should contrast the diffusion equation for option value, (37), with that eventually derivable for the probability distribution of underlying asset price \( S \) from the Fokker-Planck equation (16), for while the boundary conditions for \( u(x, \tau) \) in the price-time plane are completely known, (16) instead describes the diffusive fuzziness of our future knowledge of stock prices, and \( P \) can be reset to a Dirac-\( \delta \) function every time we check the actual market.

**Discretization and stability**

To solve a partial differential equation such as (37) or (42), we must first discretize space \( x \) and time \( \tau \) and the spatial and temporal derivatives of \( u \). Depending on the method, we may need additionally to solve a system of linear equations.

We first consider numerical solution of (37) for European options and thus fixed boundary conditions. The discretization scheme, if not the solution method, will carry straightforwardly to the American case. For any method, we must obviously replace the continuous variables \( x \) and \( \tau \) with discrete versions, \( x = n\delta x \), \( \tau = m\delta \tau \). The goal is to find the scaled option prices \( u_{nm} = u(n\delta x, m\delta \tau) \) on the grid of integers \((m, n)\).

In numerically approximating a first derivative, we have the three obvious choices of figure 2 (see Numerical Recipes, §5.7 for details):

\[
\frac{\partial u}{\partial \tau}(x, \tau) \approx \begin{cases} 
\frac{u(x, \tau + \delta \tau) - u(x, \tau)}{\delta \tau} + O(\delta \tau) & \text{forward difference} \\
\frac{u(x, \tau) - u(x, \tau - \delta \tau)}{\delta \tau} + O(\delta \tau) & \text{backward difference} \\
\frac{u(x, \tau + \delta \tau/2) - u(x, \tau - \delta \tau/2)}{\delta \tau} + O((\delta \tau)^2) & \text{central difference.}
\end{cases}
\]

(43)

One normally prefers the central difference for analyzing numerical data.\(^{16}\) For reasons of numerical stability, sometimes the forward or backward difference may be better or more applicable in the solution of partial differential equations.

To approximate the second spatial derivative in (37), we start with the central difference \( \frac{\partial u}{\partial x} \) and apply a central difference to that, giving a symmetric central second derivative,

\[
\frac{\partial^2 u}{\partial x^2} \approx \frac{u(x + \delta x, \tau) - 2u(x, \tau) + u(x - \delta x, \tau)}{\delta x^2} + O(\delta x^2) .
\]

(44)

---

\(^{15}\) Other models may introduce discontinuities in the derivative of \( V \) if they have discontinuities, other than at \( S = E \), in the derivative of the payoff function.

\(^{16}\) One of my current projects has me estimating a numerical derivative evaluated in the limit \( \tau \to 0 \). On this particular project, there is a kink exactly at \( \tau = 0 \), but the derivative exists everywhere else. Since I cannot straddle \( \tau = 0 \), I am stuck extrapolating a limiting series of forward differences.
6. The Pricing of American Put and Call Options

Figure 2. Comparison of three methods for numerically calculating the derivative with the same \( \delta \tau \) (horizontal spread): the left chord represents a backward difference, the right chord a forward difference, and the chord straddling the point in question the central difference. Clearly the last most closely approximates the derivative.

(I again refer the reader to *Numerical Recipes* for an analysis of the order of accuracy.)

Perhaps the simplest method for integrating the diffusion equation sets the forward difference for time (43) equal to the central spatial difference (44), yielding (with the notation \( u_m^n = u(n\delta x, m\delta \tau) \))

\[
\frac{u^{m+1}_n - u_m^n}{\delta \tau} = \frac{u^{m+1}_{n+1} - 2u_m^n + u^{m+1}_{n-1}}{\delta x^2},
\]

or

\[
u^{m+1}_n = u_m^n + \alpha(u^{m+1}_{n+1} - 2u_m^n + u^{m+1}_{n-1}),
\]

where the parameter \( \alpha = \delta \tau / (\delta x)^2 \) measures how long a time step we’re taking for a given spatial step. Subscripts are price, superscripts time, so (46) represents an explicit formula for \( u(x, \tau + \delta \tau) \) given \( u(y, \tau) \) for three separate prices \( y \). Figure 3a illustrates the method.

![Figure 3a](image)

Figure 3. (a) The explicit method solves for \( u^{m+1}_n \) (open circle) in terms of the already-known points \( u^{-1}_n, u_n^m, \) and \( u^{m+1}_{n+1} \) (closed circles). The horizontal axis is scaled backward time, the vertical axis scaled logarithmic option value (“space”). This method can be unstable, depending on the grid spacing. (b) The implicit method solves for \( u^{m+1}_n \), \( u^{m+1}_{n+1} \), and \( u^{m+1}_{n+1} \) in terms of the already-known point, \( u_n^m \). This method is stable.

As we saw with the Euler method for ordinary differential equations, explicit methods can suffer from instabilities, although in this case of a somewhat different nature. I follow the von-Neumann stability analysis from *Numerical Recipes*, §19.2. Imagine that roundoff has resulted in
6. The Pricing of American Put and Call Options

a small error in \( u \), so that we calculate not \( u \), but

\[
\hat{u}^m_n = u^m_n + E^m_n ,
\]

(47)

where I can expand the noise \( E^m \) at time \( m\delta \tau \) in its spatial Fourier components,

\[
E^m_n = \sum_k E^m(k) e^{ikn} .
\]

(48)

The correct \( u \) is certainly a solution of (46), as is the noisy solution (47) by construction. By linearity so is the noise or any one of its Fourier components. Thus I can analyze the noise one wavenumber \( k \) at a time.

Von Neumann asks whether a particular noise component will tend to be amplified or tend to be damped out as a function of time. In the former case, the method is unstable. Plugging a single term from the right-hand side of (48) in to (46), I find

\[
E^{m+1}(k) = E^m(k)(1 + 2\alpha[\cos k - 1]) .
\]

(49)

The term in parentheses represents the factor by which the \( k \)th Fourier component of error increases at each time step. Its absolute value had better be less than unity for all non-zero values of \( k \), for which I require (for positive \( \alpha \))

\[
0 < \alpha < 1/2 .
\]

(50)

This means that once we have chosen a spatial step size \( \delta x \), we are obliged to pick a sufficiently small \( \delta \tau \), or the error terms will quickly blow up. Small \( \delta \tau \) is desirable anyway from the point of view of accuracy, but it can also slow the calculation considerably. The refinement developed below allows larger time steps, speeding the calculation, while simultaneously improving accuracy.

As with ordinary differential equations, we are led to an implicit method. Replacing the left-hand side of (45) with the backward time difference, I find instead of (46)

\[
(1 + 2\alpha)u^{m+1}_n - \alpha u^{m+1}_{n+1} - \alpha u^{m+1}_{n-1} = u^m_n .
\]

(51)

Repeating the analysis that led to (49), I compute

\[
E^{m+1} = (1 + 2\alpha(1 - \cos k))^{-1} E^m .
\]

(52)

The coefficient multiplying the error is now less than 1 in magnitude for all \( k \) (except 0, for which it is marginal) and all \( \alpha > 0 \).\(^{17}\)

As figure 3b illustrates, a single known point at time \( m \) sets a relation among three unknown points at time \( m + 1 \). If we think of \( \mathbf{u}^m \) at fixed time as a vector whose components are the scaled values of the option at that time for the various scaled strike prices, equation (51) takes the form of a linear system of equations,

\[
T\mathbf{u}^{m+1} = \mathbf{u}^m + \mathbf{b}^m ,
\]

(53)

where \( T \) is a symmetric, tridiagonal “Toeplitz” matrix with entries \( 1 + 2\alpha \) all along the diagonal and \(-\alpha \) all along the two subdiagonals. (Toeplitz just means that the entries in each subdiagonal are all equal: \( T_{i,i+j} = T_{i+1,i+j+1} \) within the range of indices.) The vector \( \mathbf{b}^m \) wouldn’t be necessary

\(^{17}\) I needn’t worry about the zero-wavenumber (infinite-wavelength) error mode, which the boundary conditions at \( x \to \pm \infty \) suppress.
if our system were infinite (in “space” $x$). However, since we have to stop for some large spatial subscripts $n = \pm N$, we absorb the boundary conditions into $b$: \(^{18}\)

\[
\begin{align*}
    b^m_N &= \alpha u^m_{N+1} \\
    b^m_{-N} &= \alpha u^m_{-N-1} \\
    b^m_j &= 0, \quad j \neq \pm (N + 1).
\end{align*}
\]  \(^{(54)}\)

We can never reach the limits $x = n\delta x \to \pm \infty$ in (41); instead, we go as far out as we can afford (in computer time or memory), and approximate $b_{\pm N} = u_{\pm(N+1)} \approx g_{\pm(N+1)}$ (compare equation (41), where $u$ isn’t set equal to $g$ until $\pm \infty$, although the free boundary may make it equal for finite $x$).

The matrix $T$ is very easily inverted\(^{19}\), lettings us write $u^{m+1} = T^{-1}(u^m + b^m)$. Since the matrix is the same at every time step, we need perform the inversion but once. However, while the sparse matrix $T$ takes up practically no memory storage, its inverse $T^{-1}$ is dense. Thus if we have many spatial steps, it may be preferable to use one of the linear-equations algorithms described in Numerical Recipes chapter 2 or in Wilmott et al. or in Golub and van Loan.

The discussion of the fully explicit and fully implicit methods sets the stage for the Crank-Nicolson discretization, which is simply the average of (46) and (51):

\[
(1 + \alpha) u^{m+1}_n - \frac{\alpha}{2}(u^{m+1}_{n+1} + u^{m+1}_{n-1}) = (1 - \alpha) u^m_n + \frac{\alpha}{2}(u^m_{n+1} + u^m_{n-1}).
\]  \(^{(55)}\)

This equation can be cast in the form

\[
Cu^{m+1} = Du^m + b^m
\]  \(^{(56)}\)

for tridiagonal Toeplitz matrices $C$ and $D$ and a suitable definition of $b$ and solved as before (see Wilmott et al. for details). While no more difficult than the implicit method, Crank-Nicolson achieves superior convergence by averaging the forward and backward time differences to achieve a central difference for the time derivative.

**Methods for free-boundary problems**

The methods of the previous section all work on problems with fixed boundary conditions. To quote Wilmott et al., “the chief problem with free boundaries, from the point of view of numerical analysis, is that we do not know where they are.” Therefore, we can’t apply them until we’ve found them. This sits well enough with the explicit method, (46)—one simply replaces locally a solution violating (40) with (39)—but there’s no obvious way of doing this with the two implicit methods, (53) and (56).

Our reference (Wilmott et al.; see also Crank from footnote 9) recommends an iterative approach, the Projected Successive Over-Relaxation method (PSOR), which can be shown to converge to the correct result in the presence of free boundaries.

Just as the previous section needed to consider first one algorithm, then its antithesis, and finally their synthesis in order to make sense of the final result, so in order to explain PSOR must we first understand Jacobi’s iterative method, then its refinement, Gauss-Seidel relaxation, in order to combine them into the Successive Over-Relaxation (SOR) algorithm. PSOR is a slight modification of SOR to take account of the condition (40).

---

\(^{18}\) Alternatively, we could place non-zero terms in the lower-left and upper right corners of the matrix $T$. However, this would destroy most of its wonderful tridiagonal properties.

6. The Pricing of American Put and Call Options

Discretization of Relation (42)

In the solution of equation (42), the accompanying conditions, (40), (41), and the continuity requirement \( \frac{\partial V}{\partial S} = -1 \), are as essential as the analogous conditions were to the obstacle problem of figure 1. The proposed solution algorithm must therefore not only solve (42) but also must satisfy the conditions. I will not prove that the following method accomplishes this, but the fact is easy enough to verify numerically. (See Crank for a proof.)

I follow Wilmott (chapter 21) in applying Crank-Nicolson discretization (55) to the first set of parentheses in (42). The second set of parentheses evaluated at time \( \tau = (m + 1)\delta \tau \) is represented simply by \( u_{m+1}^n - g_{m+1}^n \). Collecting all the time-\( m \) terms from the right-hand side of (55) in a vector

\[
 b_m = (1 - \alpha)u_n^m + \frac{\alpha}{2}(u_{n+1}^m + u_{n-1}^m) ,
\]

we have for the approximation of (42)

\[
 (u_{n+1}^m - \frac{\alpha}{2}(u_{n+1}^m - 2u_n^m + u_{n-1}^m) - b_n^m) (u_n^m - g_n^m) = 0 .
\]

As with the example of the implicit method (see (54)), the general form of (57) must differ for the extreme spatial indices \( n \). Here, we take \( n \) to range from \(-N + 1\) to \( N - 1\); the extremal values replacing (57) are

\[
 b_{\pm(N-1)} = u_{\pm(N-1)}^m + \frac{\alpha}{2}(g_{\pm N}^m - 2u_{\pm(N-1)}^m + u_{\pm(N-2)}^m + g_{\pm(N-1)}^m) .
\]

This defines \( b \) at time \( m \) partly in terms of \( g \) at time \( m + 1 \), but it presents no difficulty, as the scaled payoff function \( g \) is known from (39) in advance for all times.

In terms of the Toeplitz matrix

\[
 C = \begin{pmatrix}
 1 + \alpha & -\alpha/2 & 0 & \ldots & 0 \\
 -\alpha/2 & 1 + \alpha & -\alpha/2 & 0 & \ldots \\
 0 & -\alpha/2 & 1 + \alpha & \ddots & 0 \\
 \vdots & \ddots & \ddots & \ddots & \ddots \\
 0 & 0 & -\alpha/2 & 1 + \alpha & -\alpha/2 \\
 0 & 0 & 0 & -\alpha/2 & 1 + \alpha \\
\end{pmatrix}
\]

with diagonal \( 1 + \alpha \) and \(-\alpha/2 \) subdiagonals, (58) becomes

\[
 (u_{m+1}^n - g_{m+1}^n) \cdot (Cu_{m+1}^n - b_m^m) = 0 .
\]

The boundary conditions are succinctly expressed by the requirement that no component of either of the two orthogonal vector factors on the left-hand side of (61) may go negative.

At the beginning of each time step \( m \), we know the value of \( u_m^m \), and \( \text{via} \ (57) \) and \( (59) \) that of \( b_m^m \). Of course we also know the payoff \( g_{m+1}^m \); there remains only to solve (61) for \( u_{m+1}^m \).

Iterative solutions for linear systems

Solutions of linear-algebra problems such as (53), (56), and (61) fall into two general categories. The more familiar algorithms, such as, for example, the technique one would use to solve \( Au = b \)
with pen and paper for small \( A \) (say \( 2 \times 2 \)), are deterministic: one knows in advance exactly how many steps are required for a solution, and that solution is exact except for roundoff error.\(^{20}\)

Alternatively, one can solve systems of equations iteratively. Beginning with a guess for the solution \( u \), one applies some procedure that refines the guess. One repeats the process with the refinement from the previous step as the guess for the next step until \( u \) stops changing appreciably. Of course, we employed iteration in the zero-finding subroutine for the first mechanics project. One would like some theoretical assurance first that the sequence of guesses converges and second that it converges to the right answer.\(^{21}\)

Jacobi’s method and the Gauss-Seidel method (Numerical Recipes §19.5) solve the linear system
\[
Au = b
\]
for \( u \) by first writing
\[
A = L + D + R
\]
where \( D \) is diagonal, \( L \) has non-zero components only below the diagonal (i.e., lower-left, or column index less than row index), and \( R \) has non-zero components only above the diagonal. Jacobi notes that \( D \) is very easy to invert, while \( L + R \) is not so easy. Specifically, the inverse of a diagonal matrix is another diagonal matrix whose entries are the inverses of the original diagonal, \((D^{-1})_{ii} = (D_{ii})^{-1}\).

Rearranging some terms in (62) gives
\[
u = D^{-1}b - Ju
\]
where the Jacobi matrix
\[
J = D^{-1}(A - D)
\]
This inspires an iterative scheme wherein the \( k + 1 \)st estimate of \( u \) is given in terms of the \( k \)th by
\[
u^{(k+1)} = D^{-1}b - Ju^{(k)}
\]
The initial guess \( u^{(0)} \) hardly matters at all. By an argument similar to von Neumann’s earlier, (66) converges if and only if all eigenvalues of \( J \) fall inside the unit circle in the complex plane. The largest (in magnitude) such eigenvalue, \( \lambda_J \), determines how quickly \( u \) converges to the right answer.

Jacobi’s method converges too slowly for most purposes; Gauss-Seidel is a variant that treats the lower-triangular matrix \( L + D \) as “easy.” Lower-triangular matrices are nearly as easy as diagonal ones. Let \( B \) be a lower-triangular matrix; we wish to solve \( Bu = b \). The first row of the matrix equation reads \( B_{11}u_1 = b_1 \), which immediately gives us \( u_1 = b_1/B_{11} \). The second row reads \( B_{21}u_1 + B_{22}u_2 = b_2 \). Since we know \( u_1 \), we immediately get \( u_2 \), and so on. Thus one iterates
\[
(L + D)u^{(k+1)} = b - Ru^{(k)}
\]
Gauss-Seidel also converges slowly, but the clever numerical analysts have shown that the following can converge very rapidly for the right choice of \( 1 < \omega < 2 \):
\[
u^{(k+1)}_{\text{SOR}} = \omega u^{(k+1)}_{\text{GS}} + (1 - \omega)u^{(k)}_{\text{SOR}}
\]
\(^{20}\) The roundoff error isn’t always small in the most straightforward treatment of the linear system \( Au = b \); see the discussions in Numerical Recipes about matrix conditioning.

\(^{21}\) I refer the curious reader to the books already mentioned, especially Golub and van Loan for the unrestricted problems and Crank for (61) with its boundary conditions.
Here GS stands for “Gauss-Seidel” and SOR for “successive overrelaxation.” If we think of Gauss-Seidel as “relaxing” \( u \) from the guess toward the right answer, any \( 0 < \omega < 1 \) in (68) will cause the solution to “relax” less rapidly. Thus \( 1 < \omega < 2 \) “overrelaxes.”

It can be shown that the optimal value for \( \omega \) in the long-time limit is given in terms of the dominant eigenvalue of the Jacobi matrix by \( \omega = 2/(1 + \sqrt{1 - \lambda_J^2}) \). As the authors of Numerical Recipes point out, the calculation of \( \lambda_J \) is usually more involved than the solution of (62), so it’s difficult to apply this optimization. However, for our particular matrix (60), the largest eigenvalue of \( J \) is easily calculated, at least in the limit of infinite size of the matrix: \( \lambda_J = \alpha/(1 + \alpha) \). Although Numerical Recipes recommend an iterative scheme for \( \omega \) called Chebyshev acceleration, their analysis applies only to SOR, not to the projected SOR we consider next, and I got better results using the asymptotically optimal \( \omega \) throughout.

**Projected Successive Overrelaxation**

SOR solves the system (56), but the constrained system (61) requires only one additional step in the algorithm: we replace (68) with

\[
\mathbf{u}^{(k+1)}_{\text{SOR}} = \max(\mathbf{g}, \omega \mathbf{u}^{(k+1)}_{\text{GS}} + (1 - \omega) \mathbf{u}^{(k)}_{\text{SOR}})
\]

(69)

with the maximum applied component-wise. I have suppressed the time indices (too many superscripts!), evaluating \( \mathbf{g} \) and \( \mathbf{u} \) both at time \( m + 1 \) from (61) for all steps of the PSOR iteration.

**Anticlimax**

After all this work, the actual implementation of two subroutines to calculate the Black-Scholes value of an American option on pages 329 and 331 in Wilmott et al. must seem disappointingly simple. You should flesh out these two subroutines with a few subroutines of supporting code, paying attention to the copy of the e-mail I sent to the authors. The supporting code will have to allocate some arrays and convert between scaled and real-world variables by way of (30)–(36). Finally, embed that program in a larger program to calculate the volatility \( \sigma \) implied by a particular trading price for an option. There may be a better method, but what springs to mind is a zero-finding program for

\[
V(S, t; \sigma) - V_{\text{data}} = 0
\]

(70)

where \( V(S, t; \sigma) \) is the predicted value of the option, \( S \) the current underlying stock price, \( t \) the current time, and \( V_{\text{data}} \) the actual market price of the option.
6. The Pricing of American Put and Call Options

references and resources for this project
(*) indicate the most important for the course


Most of our project is based on this book, on reserve at the library.


This is the Nobel-prize-winning paper that founded the field. It is available from USF IP addresses through the library’s on-line journal collection.

Chicago Board Options Exchange, Characteristics and Risks of Standardized Options http://www.optionsclearing.com/publications/risks/download.jsp

* http://www.cboe.com

Place to get free current and historical data on the prices of options and stocks

* http://www.etrade.com

Another place for free current and historical data

http://finance.yahoo.com

Free current and historical data on stocks (no options)

Laloux et al., Phys. Rev. Lett. 83 (1999), 1467;

Recent pair of articles on financial modeling in Physical Review Letters


This week’s project

The average rate of return $\mu$ of a security drops out in pricing an option in the Black-Scholes model, but we still need to estimate the volatility $\sigma$. Often, one turns this around, and uses the current market premiums paid for options to estimate the market’s view of future volatility on a stock. Our best estimate of volatility on a particular date comes from calculating the variance in closing share prices for some length of time (60 days, say) with the target date in the center. The past 30 days are no problem. Unfortunately, traders do not have the luxury of downloading the next 30 days of future data.

Picking times in the past, compare volatility estimates based on the trailing $N$ days, $N$ days centered on the date in question, and $N$ days of future data to the volatilities implied by options trading that day. Use the resources indicated above for historical data; be sure to note that options listing 0 volume for a particular date did not trade at all, so their closing prices are meaningless. Note also that stock options in this country always expire on the third Friday of the month in question. Empirically, options of different expirations do not always imply the same volatility (see Wilmott et al.).

You will need to write a program to solve the Black-Scholes equation for an American option (this is thoroughly outlined in Wilmott et al.); depending on what you decide to emphasize, you may also find a linear-fitting program helpful (see Numerical Recipes). Other auxiliary programs or scripts may be useful for collecting data from Web addresses. I will provide an example program for calculating the number of days between two dates.