Stochastic Calculus, Courant Institute, Fall 2021
http://www.math.nyu.edu/faculty/goodman/teaching/StochCalc2021/index.html

Week 5

Jonathan Goodman, November, 2021

1 Active strategies for diffusions

Dynamic stochastic models may be used to design active strategies for interacting with or controlling them. An active trading or control strategy makes a decision at time t based on information available at time t. This is sometimes called decision making *under uncertainty*.

In a dynamic setting, there is a diffusion process X_t. We use u_t to represent a *decision variable* or *control variable* at time t. Often *decision variable* refers to the case when u_t does not influence X, as in Section 3. A control variable would enter into the SDE governing X. The decision variable must be non-anticipating, which means u_t can be a function of $X_{[0,t]}$ but cannot use values X_s for $s > t$. A function $u_t = U(X_{[0,t]})$ is a *strategy*. The thing (person or algorithm) choosing u might be called an *agent* or a *controller*.

An *optimal strategy* or *optimal control* is one that optimizes (either minimizes or maximizes) some measure of merit, or *objective function*. The actual outcome is random, so the objective function should depend on expectations or probabilities. Optimization is a systematic say to find good strategies or controls. In many cases the objective function is a little arbitrary. The structure of the objective function can influence the qualitative nature of the optimal strategy in ways that may be unwanted. Section 2 argues for a specific kind of objective function (expected utility) that makes sense from a philosophical point of view, which is that the objective function should involve only "simple" expectations (expected values of functions of the outcome) rather than more complex functions of the probability distribution of the outcome. There are analytical and computational approaches to optimizing such objective functions, including *dynamic programming* (an example in Section 3) and stochastic gradient descent (Section 4).

2 Utility and choice theory

Utility theory is a philosophy of how people should make decisions under uncertainty. It is an opinion about what people should optimize, if they use optimization for decision making. Specifically, it is the view you should optimize the expected value a function of your wealth called the *utility* function.

In some groups, utility theory is so conventional that it is hardly questioned. Most of micro-economics is based on it. But utility theory is still controversial. For one thing, the utility function is somewhat arbitrary. The theorem below says you should use a utility function, but it does not tell you what function to use. Also, many people (ordinary people and sophisticated data-driven financial "rocket scientists") use criteria that are not equivalent to expected utility. The fields of behavioral economics and behavioral finance are dedicated to studying how people make decisions inconsistent with utility theory. My opinion is that showing that people do something in a bad way is not an argument for doing it that way. For example, the fact that Americans eat unhealthy food is not an argument for eating unhealthy food.

The *von Neumann Morgenstern theorem* is a strong argument for using expected utility. It starts with some natural axioms about choice under uncertainty. The theorem states that any choice system that satisfies these axioms is given by optimizing the expected value of utility function. If your optimization objective function is not equivalent to an expected utility, then you are violating axioms most people do not want to violate.

The theorem is about *preference systems*. The random outcome is modeled as a single random variable X that is thought of as wealth. This is specified by giving a probability density $p(x)$, for a continuous random variable, or probabilities P_i if X is equal to one of the discrete values x_i. A preference system is an ordering of random variables, which is modeled as an ordering of probability distributions. We write $X \prec Y$ if the agent prefers Y to X, and $X \sim Y$ if the agent is *indifferent* (does not prefer X over Y or Y over X). The axioms imply that an agent will sometimes be indifferent. An agent who is always indifferent satisfies the axioms in a trivial way. As was just said, a preference system really is about probability distributions, p or P. We should write $p_X \prec p_Y$ instead of $X \prec Y$.

Under the axioms, there is an objective function $v(p)$ so that $p_X \prec p_Y$ if and only if $v(p_X) < v(p_Y)$, and $p_X \sim p_Y$ if and only if $v(p_X) = v(p_X)$. This objective function is not unique. Clearly $\widetilde{v}(p) = 2v(p)$ and even $\widetilde{v} = e^v$ determine the same preference system. The conclusion of von Neumann Morgenstern theory is that any "rational" preference system has an objective function $v(p)$ that is determined by a *utility function* $V(x)$ (properties given below), which means

$$\begin{aligned} v(p) &= \mathrm{E}_p[\,V(X)] \\ &= \int V(x)\,p(x)\,dx \;, \quad \text{(continuous } X) \\ &= \sum_i V(x_i) P_i \qquad , \quad \text{(discrete } X)\;. \end{aligned}$$

This representation of of the objective function is *linear* in the probabilities p or P. An objective function that is not (equivalent to) a linear function of the probabilities must, therefore, violate the von Neumann Morgenstern axioms.

Including variance in the objective function is a common way to violate the axioms and therefore be "irrational". The variance seems natural in that investments are often chosen to balance value against risk, measured by $\mathrm{E}[X]$ and $\mathrm{var}(X)$ respectively. What's wrong with $v(p) = \mathrm{E}[X] - \mathrm{var}(X)$? The variance is (discrete case, for simplicity)

$$\mathrm{var}(X) = \sum_i \left(X_i - \sum_j X_j P_j \right)^2 P_i$$

This is a quadratic function of P, since it involves products $P_i P_j$. The variance is a natural, but not a "rational" way to measure risk.

The von Neumann Morgenstern "rationality" axioms seem simple and natural, but they have surprisingly specific consequences. The first is *transitivity*. If you prefer Y to Z and you prefer Z to Y, then you prefer Z to X. In the language of probability distributions, this is

1. If $p_X \prec p_Y$ and $p_Y \prec p_Z$, then $p_X \prec p_Z$. If $p_X \prec p_Y$ and $p_Y \sim p_Z$, then $p_X \prec p_Z$.

The *monotonicity* axiom is that "more is better". If $X \leq Y$ and $\Pr(X < Y) > 0$, then $X \prec Y$. These conditions may be expressed as $Y = X > W$, where W is non-negative and has a non-zero probability of being positive. Such a random W is an *arbitrage*. The monotonicity axiom says that an agent prefers to include any arbitrate. The condition $X < Y$ depends on the joint distribution of X and Y, but the choice system should not depend on the joint distribution – you have to choose one distribution only, not a joint distribution. Therefore the informal $X < Y$ should be replaced by conditions on X and Y separately. The *cumulative distribution function* is

$$C_X(a) = \Pr(X \leq a) = \int_{-\infty}^{a} p(x)\, dx \ .$$

The condition $X < Y$ "really means" that $C_X(a) \geq C_Y(a)$ for any a, and $C_X(a) > C_Y(a)$ for some a.

2. If $C_X(a) \geq C_Y(a)$ for all a and $C_X(a) > C_Y(a)$ for some a, then $p_X \prec p_Y$.

The *risk aversion* axiom is that you prefer a definite $\overline{x}$ to a random variable with expected value $\mathrm{E}[X] = \overline{x}$. An English country saying for this is: A bird in the hand is worth two in the bush.[1] For this axiom, let $\delta_{\overline{x}}$ be the probability distribution of the non-random variable $\overline{x}$. In this statement, $\mathrm{var}(X) > 0$ is just a way of saying that X is truly random – $\Pr(X \neq \overline{x}) > 0$.

3. If $\overline{x} = \mathrm{E}[X]$ and $\mathrm{var}(X) > 0$, then $p_X \prec \delta_{\overline{x}}$.

The final *interpolation* axiom seems purely technical, but the theory collapses without it. Informally, it says that if $X \prec Y \prec Z$, then there is some W that "interpolates" between X and Z with $W \sim Y$. That is, you cannot go in a

[1]This is about shooting birds to eat. A bird "in the hand" is one that you already have shot and can take home to eat. A bird "in the bush" is one that you can see but might not succeed in shooting. If your "hit rate" is 50%, then two birds in the bush have the same expected value as one bird in the hand.

continuous way from being worse than Y to being better than Y without being equivalent to Y at some point between. The "interpolation" is in the sense of probability distributions. For any q with $0 \leq q \leq 1$, there is a probability density (I apologize for the less-than-ideal notation.)

$$p_{q,X,Y}(x) = q\, p_X(x) + (1-q)\, p_Y(x) \ .$$

This is p_X when $q = 1$ and p_Y when $q = 0$. It is a probability density (non-negative, integral equals 1) as long as $q \geq 0$ and $q \leq 1$.

Interpolating the probability distributions may be interpreted as choosing X with probability q and Y with probability $1-q$. The interpolated random variable is

$$W_q = \begin{cases} X & \text{with probability } q \\ Y & \text{with probability } 1-q \ . \end{cases}$$

The interpolation axiom is

4. If $p_X \prec p_Y \prec p_Z$, then there is $q \in (0,1)$ so that $p_Y \sim q\, p_X + (1-q)\, p_Z$.

This is like the *intermediate value theorem* of ordinary calculus. If $f(q)$ represents the desirability of W_q, then it cannot go from less than the desirability of Y to being above the desirability of Y without, at some q in between, being equal to the desirability of Y. Note that the distribution of W_q is not the distribution of Y even though the agent is indifferent between them.

The von Neumann Morgenstern theorem is that a preference system that satisfies these axioms is determined by a *utility function* V. A function $V(x)$ is a utility function if it is monotone increasing, $V(y) \geq V(x)$ if $y > x$, and concave, $V''(x) \leq 0$. A utility function determines a preference system if

$$X \prec Y \iff \int V(x)\, p_X(x)\, dx < \int V(x)\, p_Y(x)\, dx \ .$$

A probability distribution is optimal if it is better than any other available distribution. Therefore, you optimize by finding the maximum of the expected utility. Section refsec:sgd explains one reason this might be easier in practice than optimizing an objective function that depends on p_X in a non-linear way.

The properties of a utility function are related to the axioms of Von Neumann Morgenstern theory. Informally, $V(x)$ tells you how much "value" the money x has for you. $V'(x)$, which is called *marginal utility* tells you how much happier one dollar would make you. The condition $V' > 0$ is related to the "more is better" monotonicity axiom. The second derivative V'' is supposed to be negative. This says that the marginal utility decreases as x increases: $\frac{d}{dx}V'(x) < 0$. If I have ten dollars, I might be very happy to have one more. But if I have a million dollars, one more dollar would not make me much happier.

The concavity $V''(x) < 0$ is also related to risk aversion. The relationship is from a mathematical fact called *Jensen's inequality*. This says that if V is an concave function and X is any random variable, then

$$\mathrm{E}[\, V(X)] \leq V(\mathrm{E}[X]) \ .$$

Thus, a concave utility function satisfies the risk aversion axiom 3 of von Neuman Morgenstern theory.

3 Optimal dynamic investment

Here is an example of a dynamic stochastic optimization problem. We would call the decision variable a "policy" rather than a "control" because it does not influence the stochastic market price in the the theory. This example illustrates the dynamic programming approach to dynamic stochastic policy/control problems. It uses a value function that satisfies a partial differential equation like the backward equations we saw in earlier classes. PDEs that arise through stochastic optimization in this way are sometimes called *Hamilton Jacobi Bellman* equations, or *HJB* equations. Hamilton and Jacobi had nothing to do with diffusion processes or optimization. Bellman found that the value function satisfies a PDE with properties similar to the ones used by Hamilton and Jacobi for other purposes. In this problem, the value function and the optimal policy are determined together.

Note that the notation and terminology for this problem is different from the general terminology for dynamic stochastic optimization in the introduction. Any applied mathematician has to learn to "speak" different forms of notation for the same thing to different people in different settings.

In this problem there is a wealth Z_t at time t. At any time t, some of this is invested in a risky asset S_t and the rest in a risk free asset M_t (for "money", or "money market account"). The goal is to buy and sell these assets to optimize the outcome at time T. As we saw in Section 2, that there is a utility function $V(z)$ and we seek to maximize $f = \mathrm{E}[\,V(Z_T)]$.

The decision variable is X_t, which is the amount invested in the risky asset. The rest, which is $Z_t - X_t$ is invested in the risk free asset. The risky asset is taken to be a geometric Brownian motion

$$dS_t = \mu S_t dt + \sigma S_t dW_t \; .$$

The risk free asset has a fixed and deterministic growth rate r (the risk-free rate)

$$dM_t = rM_t dt \; .$$

The parameters μ (expected return of the risky asset), σ (volatility of the risky asset), and r (risk-free rate of return) are assumed constant and known.

Investing X_t in the risky asset means owning $n_t = X_t/S_t$ "shares" of the risky asset. If you don't trade in a time interval $dt > 0$, then $dn = 0$ for that interval. This means that

$$\begin{aligned} X_{t+dt} &= n_t S_{t+dt} \\ &= n_t\left(\mu S_t dt + \sigma S_t dW_t\right) \\ &= \mu n_t S_t dt + \sigma n_t S_t dW_t \\ &= \mu X_t dt + \sigma X_t dW_t \; . \end{aligned}$$

Similarly, if you allocate Y_t to the risk-free asset, then the value changes to

$$Y_{t+dt} = rY_t dt \; .$$

Since Z_{t+dt} is the sum of the risky and risk free parts, we have

$$Z_{t+dt} = Z_t + \mu X_t dt + \sigma X_t dW_t + r(Z_t - X_t)dt \; .$$

This may be written

$$dZ_t = rZ_t dt + (\mu - r)X_t dt + \sigma X_t dW_t \; . \qquad (1)$$

The quantity $\mu - r$ is the *excess return*. It is natural to think the excess return is positive. Otherwise, investing in the risky asset would increase your risk while decreasing your expected return.

There are many simplifying assumptions in the Merton analysis. Obviously there is continuous time, but also continuous allocation (X_t can be any real number) and continuous trading: n_t, the number of shares, can be any continuous (adapted) function of t. More serious are the assumption of free and frictionless trading without market impact. At any time you can buy or sell as much of the risky asset as you want at the price S_t. The price is the same for buying or selling. In real markets, there is a price for buying (the *ask* or *offer* price) and a slightly lower price for selling (the *bid* price). Look at your favorite stock on Yahoo Finance and you will see that these differ by a few cents – not a huge number but enough to make certain trading strategies impraactical. Also, there are a limited number of shares available at the bid price – typically a few hundred. If you want more than that, you have to pay a little more. That means your trade changes the market price. You have had *market impact* and you are no longer a pure *price taker*. Finally, there are no constraints on the sign of X_t or $Y_t = Z_t - X_t$. If $X_t < 0$, you "own" a negative amount of risky asset and have more than Z_t in cash (risk free asset). This corresponds to borrowing shares, selling them, and putting the money into the risk free asset. If $X_t > Z_t$, then Y_t (the cash *position*) is negative. You have borrowed money to buy more stock than you can pay for with your wealth. The Merton theory says the interest rate you pay for borrowing is the same as the interest rate you get for lending (investing in the risk free asset is lending).

With dynamic trading, the investor chooses X_t and Y_t at time t subject only to the constraint $X_t + Y_t = Z_t$. Once this allocation is made, we "watch" the markets for time dt. The result is (1). The agent knows $S_{[0,t]}$ and $X_{[0,t]}$, and therefore knows Z_t. The process is Markov, so the past does not influence the future. If you know Z_t, the rest of the information is irrelevant. This means that the allocation decision X_t is a function of Z_t alone.

The value function is the best expected utility an agent can achieve starting at time t with wealth $Z_t = z$. In order to achieve this best expected utility, the agent must follow the optimal policy from time t to time T. The value function, therefore, is

$$f(z,t) = \max_{\text{policy}} \mathrm{E}[\, V(Z_T) \mid Z_t = z] \; . \qquad (2)$$

The subscript policy means that the agent maximizes over all adapted policies.

The value function (2) satisfies a PDE called the *Hamilton Jacobi Bellman* equation. This is based on the *dynamic programming principle*, which is the idea that you make a decision at time t assuming that all decisions after that will be optimal. The derivation of the HJB equation starts like one of the derivations of backward equations from earlier classes. The new thing here is that X_t is not known. Instead, you choose the optimal X_t, which is the one that gives you the largest expected utility at time $t + dt$.

$$f(z,t) = \max_{X_t} \mathrm{E}[\, f(z + dZ, t + dt)] \ . \tag{3}$$

This formula represents the dynamic programming principle just stated: you choose the best X_t at time t and then receive the expected utility of the optimal path from $Z_{t+dt} = z + dZ$ starting at time $t + dt$.

We calculate the expectation on the right side using Taylor series as before. We use subscripts for partial derivatives and leave out arguments (z,t). The conditional expectation assumes $Z_t = t$, so f and its derivatives evaluated at (z,t) come out of the expectation:

$$\begin{aligned} &\mathrm{E}[\, f(z + dZ, t + dt) \mid Z_t = z] \\ &\quad = \mathrm{E}\left[f + f_z dZ + \frac{1}{2} f_{zz} (dZ)^2 + f_t dt \mid Z_t = z \right] \\ &\quad = f + f_z \mathrm{E}[\, dZ \mid Z_t = z] + \frac{1}{2} f_{zz} \mathrm{E}\big[\, dZ^2 \mid Z_t = z \big] + f_t dt \ . \end{aligned}$$

The expectations on the right are found from the stochastic dynamics (1):

$$\begin{aligned} \mathrm{E}[\, dZ \mid Z_t = z] &= rz dt + (\mu - r) X_t dt \\ \mathrm{E}\big[\, dZ^2 \mid Z_t = z \big] &= \sigma^2 X_t^2 dt \ . \end{aligned}$$

Assembling these calculations to evaluate the right side of (3) gives

$$f = f + \max_{X_t} (rz + (\mu - r) X_t) f_z dt + \frac{1}{2} \sigma^2 X_t^2 f_{zz} dt + f_t dt \ .$$

This simplifies to (taking out the parts that don't depend on X_t)

$$0 = rz f_z + f_t + \max_{X_t} \left[(\mu - r) f_z X_t + \frac{1}{2} \sigma^2 X_t^2 f_{zz} \right] \ .$$

Suppose the agent chooses to allocate x to the risky asset at this time. Then at time $t + dt$, the wealth would be given by (1)

$$Z_{t+dt} = z + r Z_t dt + (\mu - r) X_t dt + \sigma X_t dW_t \ .$$

The optimal expected utility starting from $t + dt$ would be

$$f(z + dZ_t, t + dt) = f(z,t) + \partial_z f(z,t) dZ_t + \frac{1}{2} \partial_z^2 f(z,t) (dZ_t)^2 + \partial_t f(z,t)\, dt \ . \tag{4}$$

We do the maximization by differentiating with respect to X_t, setting the derivative to zero, and solving for X_t. That is

$$\begin{aligned}(\mu - r) f_z + \sigma^2 X_t f_{zz} &= 0 \\ X_t &= -\frac{(\mu - r) f_z}{\sigma^2 f_{zz}} \; . \end{aligned} \tag{5}$$

We substitute this optimal allocation into the HJB equation and get the result

$$0 = \partial_t f(z,t) + rz\partial_z f(z,t) - \frac{(\mu - r)^2}{2\sigma^2} \frac{\left(\partial_z f(z,t)\right)^2}{\partial_z^2 f(z,t)} \; . \tag{6}$$

This is the non-linear backward equation for Merton's optimal dynamic investment problem.

4 Stochastic Gradient Descent

Stochastic gradient descent is a general way to optimize an objective function that that is the expected value of a random variable and control. It is a stochastic version of ordinary gradient descent (reviewed below) that does not need accurate estimates of the objective function or its gradient. What is presented here should be called the *Robbins Monro* algorithm, or, as Robbins and Monro called it, *stochastic approximation*. The term *stochastic gradient descent* is more properly used for a closely related method machine learning people use to train neural networks when there are many independent samples.

Ordinary gradient descent is a way to find the minimum of a function $v(u)$. If you want to maximize v, just minimize $-v$. The algorithm might then be called "gradient ascent", but people don't use that term much. Suppose v is an objective function whose gradient can be computed:

$$\nabla v(u) = \begin{pmatrix} \frac{\partial v}{\partial u_1} \\ \vdots \\ \frac{\partial v}{\partial u_n} \end{pmatrix} \; .$$

Gradient descent is the iteration

$$u_{k+1} = u_k - s_k \nabla v(u_k) \; . \tag{7}$$

The parameter s_k is the *learning rate*. It should be positive. It either is taken to be a fixed somewhat small number, or it goes to zero as $k \to \infty$.

The negative gradient direction in (7) is a *descent direction*. That means that if $\nabla v(u_k) \neq 0$ and if s_k is small enough, then

$$v(u_{k+1}) < v(u_k) \; .$$

This is clear from the Taylor series calculation

$$\begin{aligned} v(u_{k+1}) &= v(u_k - s_k \nabla v(u_k)) \\ &= v(u_k) + \nabla v(u_k)^t \left(-s_k \nabla v(u_k)\right) + O(s_k^2) \\ &= v(u_k) - s_k \left\| \nabla v(u_k) \right\|^2 + O(s_k^2) \end{aligned}$$

If $\nabla v(u_k) \neq 0$ and s_k is small enough, then $v(u_{k+1}) < v(u_k)$. In practical gradient descent algorithms, it would be hard to know what s_k is good. It is common to do some kind of *line search*, which means decreasing s_k if $v(u_{k+1} > v(u_k)$ and increasing s_k to see whether that lowers $v(u_{k+1})$. A simple version of this might use *binary search*, which means cutting s_k in half if it's too small and doubling it to see whether u goes down more. Strategies like these are hard to apply to stochastic gradient descent.

The Robbins Monro algorithm is for a version of the optimization problem in which the objective function is the expected value of a random variable. It is used in case there is no analytic formula for the expected value. One approach would be to estimate the expected value many independent samples. Stochastic gradient descent is a striking variation on this idea in that uses a small number of samples, often just one, per iteration. The algorithm finds the optimal u (almost surely) despite the fact that it uses inaccurate estimates of u and ∇u at each iteration.

Here is a technical explanation. In the language of this class, suppose W is a random variable and we have a way to generate *samples* of W – independent random variables with the same distribution that W has. For example, W could be a diffusion process described by an SDE. Another possibility is that W is the noise driving an SDE where the components of u are parameters. Either way, independent simulations would produce independent sample paths. In a simple version of the Robbins Monro formalism, the objective function is the expected value of a random variable we could think of as a utility function:

$$v(u) = \mathrm{E}[\, V(W,u)] \ . \tag{8}$$

We seek algorithms to optimize $v(u)$ for situations where it is feasible to generate independent samples $W_n \sim W$ and evaluate $V(W,u)$, but it is harder to evaluate $v(u)$ directly.

Gradient based algorithms use the *stochastic gradient*, which is

$$\nabla_u V(W,u) \ .$$

The basic stochastic approximation algorithm uses independent $W_n \sim W$ and uses them to create iterates u_k using the stochastic gradient instead of the simple gradient:

$$u_{k+1} = u_k - s_k \nabla_u V(W_k, u_k) \ . \tag{9}$$

Here, there is no guarantee that $v(u_{k+1}) < v(u_k)$ even if s_k is very small. The *search direction*, which is $\nabla_u(W_k, u_u)$ need not be a descent direction. It is a descent direction if $\nabla v \neq 0$ "in the mean", because

$$\mathrm{E}\left[\, \nabla v(u_k)^t \nabla_u V(W_k, u_k)\right] = \nabla v(u_k)^t \nabla v(u_k) = \left\| \nabla v \right\|^2 \ .$$

But the actual random variable $\nabla v(u_k)^t \nabla_u V(W_k, u_k)$ does not have to be positive.

5 Option hedging in continuous time

A *stock option* it the right to buy or sell a specific stock at a specific price at a specific time or until a specific time. The right to buy is a *call* option. The right to sell is a *put* option. If the right exists only at time T, it is a *European style* option. If the right exists at any time up to time T, it is an *American style* option. Options are traded in public exchanges and their price is determined in the market. However, the *Black Scholes* theory of option pricing says what the option price should be, in an economic model. Market prices disagree with this simple theory, but the theory nevertheless provides an important way to think about buying and selling options.

The terminology of this section is that T is the *expiration time* of the option. An option has a *strike price*, written K, that is the price at which the stock will be bought or sold. An option is the right to buy or sell, but not a requirement. Consider a European style option. Suppose you own a put option (option to sell) at price K and the price is S_T. If $S_T > K$, then you can sell a share of stock for S_T on the market or for price K to the *counterparty* (the person who sold you the option). If you have a share of stock, you get more by selling on the market, so you don't exercise the option. We say the option is *out o the money*. If $S_T < K$, then you can buy a share for S_T and sell for K. This gives you a profit of $K - S_T$. For European options that are traded on exchanges, the option is *settled in cash*, which means that the exchange gives you the cash value of the option. For a put, this is

$$V(S_T) = \max\{K - S_T, 0\} = (K - S_T)_+ \ .$$

For a call, similar reasoning gives

$$V(S_T) = \max\{S_T - K, 0\} = (S_T - K)_+ \ .$$

American style options with the *early exercise* feature present the owner with a dynamic optimization problem – finding the optimal strategy for exercising the option.

The Black Scholes pricing theory was first developed by Fisher Black and Miron Scholes using reasoning similar to that of Section 3. This section explains the reasoning, as I have come to understand it. Later the *binomial tree* model was invented as a way to explain option pricing to people who are not familiar with stochastic calculus. This is explained in Section 6. The explanations here may be a little quick, because the course *Derivative Securities* covers that material more deeply.

The Black Scholes model of the trading world is this. There is a risk free asset, cash, with rate of return r. There is a risky asset, the stock, whose price is S_t that is a geometric Brownian motion with parameters μ (expected rate of return) and σ (volatility). The market is full of "agents" (traders) who can buy or sell the option or the stock without market "frictions". The amount of cash or stock can be positive or negative. For cash, this is written as "borrowing = lending"; the interest rate you get for your cash is the same as the interest you pay if you borrow. It might seem surprising that this is approximately true for big agents. Owning a negative amount of stock is called having a *short position*. Selling stock you don't own (to get a negative amount of stock) is *short selling*. The Black Scholes theory allows all this, and with zero transaction cost. For European style options, the theory assumes that if you own the option at time T, then

You can get into the Black Scholes theory by asking about *dynamic replication* of an option without buying or selling the option itself. The time t, the agent has a wealth Z_t. The agent allocates X_t to the stock and the rest to cash. The agent seeks a trading strategy so that $Z_T = V(S_T)$. We say the trading strategy *replicates* the option. The strategy satisfies eqrefdZ, which means that it is *self financing*. The trader starts with some wealth and then only takes that wealth to the market. The equivalent of the value function is the wealth you need at time t with stock price s to replicate the option:

$$f(s,t) = Z_t \text{ so that } Z_T = V(S_T) \ . \tag{10}$$

This f is the *Black Scholes arbitrage price* of the option. The idea is that if the option price is different from the arbitrage price, then you can make a guaranteed profit by replicating the option. For example, if the option price is $P > f(S_t,t)$, then the trader can sell one option and receive P. The trader then uses $f < P$ of the money to replicate the option and keeps the rest. At time T, the trader has $Z_T = V(S_T)$, so he/she can satisfy the person he/she sold the option to. The rest is risk free profit. Basic finance theory (economic philosophy) is that arbitrage opportunities like this cannot exist. If they did exist, smart traders would jump on them and they would quickly be sold out.

The technical argument of Black and Scholes is ingenious no matter how you say it. In this version, the trading strategy will have the effect that

$$Z_t = f(S_t,t) \ , \quad \text{for all } t \leq T \ .$$

This means that if you follow the strategy (details in the next paragraph) and if you start with the right wealth at time t_0, then at all later times up to time T, you still have exactly the wealth to replicate the option.

For the calculation, we use the terminology of Black and Scholes by writing the stock component of the portfolio as $X_t = \Delta_t S_t$. That means that Δ_t is the number of shares of the stock that you own. As in the Merton theory, this can be any real number. The cash position (amount of wealth in cash) is $Y_t = Z_t - \Delta_t S_t$. We calculate $dZ = f(S_t,t)$ using the market formula (1) and using Ito's lemma. The resulting equation gives a PDE for f, which is the *Black Scholes equation*. First,

$$dZ = rY_t dt + \Delta_t dS_t \ .$$

This is the "tricky" part of this approach to Black Scholes theory. You imaging that you trade (choose Δ_t and Y_t), and then keep them for time dt while the market moves. We used the same idea in the Merton theory. We then use the "budget constraint" to eliminate Y_t and write

$$dZ_t = r(Z_t - \Delta_t S_t) + \Delta_t dS_t \ .$$

The desired replication formula $Z_t = f(S_t, t)$ allows this to be written as

$$dZ_t = r(f(S_t, t) - \Delta_t S_t) + \Delta_t dS_t \ .$$

We compare this to what you get from Ito's lemma, which is (using the Ito rule, $(dS)^2 = \sigma S_t^2 dt$)

$$dZ_t = df(S_t, t) = \partial_s f(S_t, t) dS_t + \partial_t f(S_t, dt) dt + \frac{1}{2} \partial_s^2 f(S_t, t) \sigma S_t^2 dt \ .$$

We compare these expressions and see that we can eliminate the dS term (the term with dW) if we take

$$\Delta_t = \partial_t f(S_t, t) \ . \tag{11}$$

Finally, we equate the remaining terms and drop the dt from both sides. We get

$$r \left[f(S_t, t) - S_t \partial_s (f(S_t, t) \right] = \partial_t f(S_t, t) + \frac{1}{2} \sigma_2 S_t^2 \partial_s^2 f(S_t, t) \ .$$

Some algebra puts this into a more standard form

$$0 = \partial_t f + rs \partial_s f + \frac{1}{2} \sigma^2 s^2 \partial_s^2 f - rf \ . \tag{12}$$

This is the Black Scholes equation.

6 Hedging in discrete time

7 Black Scholes formula

The *Black Scholes formula* is a formula for the solution of the Black Scholes equation with final condition $f(s, T) = V(s) = (s - K)_+$ or $f(s, T) = (K - s)_+$. One way to find the Black Scholes formula is to use the fact that the Black Scholes equation is a backward equation for a geometric Brownian motion

$$dS_t = rS_t dt + \sigma S_t dW_t \ . \tag{13}$$

This is different from the geometric Brownian motion model used to derive the PDE (12) in that the expected rate of return is r instead of μ. More precisely, f is the value function

$$f(s, t) = \mathrm{E}\Big[V(S_T) e^{-r(T-t)} \mid S_t = s \Big] \ . \tag{14}$$

To be clear, (14) with process (13) is a formula for the solution of (12), but it is not the derivation. Nevertheless, (14) says that the option price is the expected payout if S is the *risk-free* process (13). An investor is *risk free* or *risk neutral* (as we said before) if he/she makes the price of a risky asset equal to its discounted expected value. The conclusion of the Black Scholes theory may be stated as giving the price as the discounted expected value using the risk free process.

The Black Scholes formula may be derived using the "risk free representation" above. The solution of the SDE is

$$S_T = S_0 e^{\sigma W_T + (r - \frac{\sigma^2}{2})T} \, .$$

We can get the distribution of W_T using $\sqrt{T}Z$, where $Z \sim \mathcal{N}(0,1)$. For a put, we get

$$f(S_0, 0) = e^{-rT} \mathrm{E}\Big[\, (K - S_0 e^{\sigma\sqrt{T}Z + (r - \frac{\sigma^2}{2})T})_+ \Big] \, .$$

As an integral, this is

$$f(S_0, 0) = \frac{1}{\sqrt{2\pi}} e^{-rT} \int_{-\infty}^{z_0} \left(K - S_0 e^{\sigma\sqrt{T}Z + (r - \frac{\sigma^2}{2})T} \right) e^{-\frac{1}{2}z^2} \, dz \, .$$

The endpoint of integration is the value of z that makes $S_T = K$. The result is

$$z_0 = \frac{\log(K/S_0) - \left(r - \frac{\sigma^2}{2} \right)}{\sigma\sqrt{T}} \, .$$

The part of this formula involving just K is

$$\frac{1}{\sqrt{2\pi}} e^{-rT} K \int_{-\infty}^{z_0} e^{-\frac{1}{2}z^2} dz = e^{-rT} N(z_0) \, .$$

8 Exercises

1. Give a proof of Jensen's inequality. Hint: Show that if V is concave then

$$V(x) \leq V(\overline{x}) + V'(\overline{x})(x - \overline{x}) \, .$$

   Integrate this inequality over x with PDF $p(x)$ and use the fact that $\overline{x}$ is the expected value of X.

2. Show that the mean-variance analysis fails in the following simple way. Define $v(p) = \mathrm{E}_p[X] - c \, \mathrm{var}_p(X)$. Here, we allow any penalty parameter $c > 0$ for the variance of X in the p probability distribution. Suppose $X > 0$. The von Neumann Morgenstern axioms suggest that you should prefer nX to X for any $n > 1$. That is, instead of one copy of X, we prefer n copies (n shares) for any $n > 1$. Show that the monotonicity axiom implies this. For that, you need to invent notation for the CDF of nX, etc. Show that the variance penalty implies that there is an optimal n, beyond which we do not want more X. This shows that variance penalization is not "rational".

3. The *Markowitz mean variance allocation* theory involves random variables R_j, which represent the return (profit) from investing one unit of money on asset j. These have expected returns

$$\mu_j = \mathrm{E}[R_j] \ .$$

The covariance is

$$C_{ij} = \mathrm{cov}(X_i, X_j) \ .$$

Suppose you have a unit amount of money and invest w_j of that in asset j. Then your total return is

$$X = \sum_{j=1}^{n} w_j R_j = w^t R \ .$$

The vector notation in the last version on the right is $w \in \mathbb{R}^n$ with components w_j and $R \in \mathbb{R}^n$ with components R_j. The expected return is

$$r = \mathrm{E}[X] = \sum_{j=1}^{n} w_j \mu_j = w^t \mu \ . \tag{15}$$

The covariance matrix C has entries C_{jk}. The variance of the return is

$$\sigma^2 = w^t C w \ . \tag{16}$$

The *budget constraint* involves the vector $\mathbf{1} \in \mathbb{R}^n$ with all components equal to one:

$$\sum_{j=1}^{n} w_j = w^t \mathbf{1} = 1 \ . \tag{17}$$

An *allocation* (or *portfolio*) is w that satisfies the budget constraint. An allocation is *efficient* if it maximizes r with a fixed σ^2 and budget constraint (17). An allocation is *inefficient* if it is not efficient. Suppose $X = w^t R$ is an efficient allocation and $\widetilde{X} = \widetilde{w}^t R$ is an inefficient allocation with the same variance. Show that, *if R is a multi-variate Gaussian*, then the expected utility of X is larger than the expected utility of $\widetilde{X}$. *Hint*. Just think of X and $\widetilde{X}$ as two Gaussian random variables with the same variance and different means. [We have emphasized that using variance to guide investment can violate the von Neumann Morgenstern axioms and lead to poor investment choices. This exercise shows that the Markowitz mean-variance analysis is OK even though it uses variance, but only because the returns are assumed to be Gaussian.]

According to von Neumann Morgenstern choice theory, any rational investor would prefer an efficient allocation to an inefficient allocation with the same variance. *Harder, attempt only after the rest is finished*: Show that this may not be true for non-Gaussian returns. *Hint*. If Y and Z are Gaussian with the same variance, then you can think of Z as larger than Y in the sense of the arbitrage axiom if the mean of Z is larger. However, there are non-Gaussian random variables Y and Z have $\mu_Y < \mu_Z$ and $\sigma_Y^2 = \sigma^2 Z$ but Z is not an arbitrage from Y in the sense that $\Pr(Z < a) > \Pr(Y < a)$ for some a. This can happen if Z has *fatter tails* than Y. [*My opinion.* Mean variance analysis is popular even though it can lead to "irrational" allocations. You might excuse this by saying it's only supposed to apply to Gaussian returns. Yet, nobody thinks returns are anything like Gaussian.]

4. Suppose the utility is a power law $V(z) = z^\gamma$.. Take the ansatz $f(z,t) = A(t)z^\gamma$.

   (a) Show that a power law utility is increasing and concave if and only if $0 < \gamma < 1$.

   (b) Substitute the ansatz into the Merton PDE (6) to show that the ansatz works. Find the differential equation and then the formula for $A(t)$.

   (c) Show that the optimal allocation has the form $x_* = mz$ and find a formula for the *Merton proportion* as a function of the parameters γ, σ, and r.

   (d) An investor is *risk neutral* if they maximize expected wealth rather than expected utility. How does the Merton strategy break down in the risk neutral limit $\gamma \to 1$?

   (e) We saw that in geometric Brownian motion, it can happen that the expected value grows exponentially but the median value goes to zero exponentially. Can this happen for this Merton problem? Can the expected utility grow exponentially while the median utility decays? What does this say about how the utility function z^γ captures risk aversion?

5. (*Extra credit, do only after everything else is done, and if you're interested in economics.*) Here is the optimal policy problem that includes consumption. The rate of consumption at time t will be C_t. You "consume" money, so the wealth dynamics with consumption are

$$dZ_t = rZ_t dt + (\mu - r)X_t dt + \sigma X_t dW_t - C_t dt \ .$$

   As with wealth, we use the utility of consumption rather than consumption itself. The reasoning is similar. You might be very happy to consume two cookies rather than one cookie, but you may not care as much for cookie 101 if you already have 100 of them. There is a *discount rate*, ρ, in addition to the risk free rate. If you consume c at time t, the utility "today" (time $t = 0$) is reduced by $e^{-\rho t}$. The agent chooses X_t and C_t at time t in a way that seeks to maximize

$$H = \mathrm{E}\left[\int_0^T e^{-\rho t} U(C_t)\, dt \right] \ .$$

The constraint is $Z_T \geq 0$. Formulate a value function, the dynamic programming principle, and the Hamilton Jacobi Bellman equation appropriate for this problem. Describe the solution when the utility function has the form $U(c) = c^\gamma$.