AMH4 - ADVANCED OPTION PRICING

ANDREW TULLOCH

1. Theory of Option Pricing

Definition 1.1 (Brownian motion)**.** A process W_t is a $\mathbb{P}$-Brownian motion if it satisfies

(1) W_t is continuous with $W_0 = 0$ (a.s.)
(2) W_t has stationary and independent increments.
(3) For any $t > 0$, $W_t \sim N(0, t)$ under the probability measure $\mathbb{P}$.

Theorem 1.2 (Properties of conditional expectation)**.** *Assume we have a probability space $(\Omega, \mathbb{P})$ and σ-algebras $\mathcal{G}, \mathcal{G}_1, \mathcal{G}_2$. Assume that $\mathcal{G}_2 \subset \mathcal{G}_1$. Then*

(1) If X is a random variable, then

$$\mathbb{E}(X \mid \mathcal{G}_2) = \mathbb{E}(\mathbb{E}(X \mid \mathcal{G}_1) \mid \mathcal{G}_2)$$

(2) If Y is a $\mathcal{G}$-measurable random variable, then

$$\mathbb{E}(XY \mid \mathcal{G}) = Y\mathbb{E}(X \mid \mathcal{G})$$

Definition 1.3 (Martingale)**.** A stochastic process X_t is a $\mathcal{F}_t$-martingale if $\mathbb{E}(|X_t|) < \infty$ and

$$X_s = \mathbb{E}(X_t \mid \mathcal{F}_s)$$

for all $s \leq t$.

Theorem 1.4 (Itô's lemma)**.** *If $F(X_t, t)$ is $C_{2,1}$ and $dX_t = \alpha_t\, dt + \beta_t\, dW_t$, then*

$$dF = (F_t + \alpha F_x + \frac{1}{2}\beta^2 F_{xx})\, dt + \beta F_x\, dW_t$$

Lemma 1.5 (Product and Quotient rule)**.** *Let X_t be an Itô processes, so that*

$$dX_t = \alpha dt + \beta dW_t.$$

Let $F(X_t, t), G(X_t, t)$ be $C_{2,1}$. Then

$$d(FG) = (F\, dG + G\, dF) + \beta^2 F_x G_x\, dt$$

$$d(F/G) = \frac{G\, dF - F\, dG}{G^2} + \frac{\beta^2 G_x}{G^3}(FG_x - GF_x)\, dt$$

Lemma 1.6 (Itô isometry). *If $\sigma_s \in L^2$, then*

$$\mathbb{E}(\int_0^t \sigma_s \, dW_s)^2 = \mathbb{E}(\int_0^t \sigma^2 \, ds)$$

Definition 1.7 (Local martingale). X_t is a local martingale if there exists a sequence of stopping times ν_n such that for every n, the process $X_t^n = X_{\min(\nu_n, t)}$ is a martingale.

Theorem 1.8 (Martingale representation theorem). *Let $\mathcal{F}_t$ be the natural filtration of a Brownian motion.*

(1) Any progressively measurable process σ_t satisfying

$$\mathbb{P}(\int_0^t \sigma_s^2 \, ds) < \infty = 1 \quad \forall t$$

the process

$$t \mapsto \int_0^t \sigma_s \, dW_s$$

is a local martingale.

(2) If X_t is an L^2 martingale, then there exists a progressively measurable process σ_s such that

$$X_t = \int_0^t \sigma_s \, dW_s$$

Hence the Brownian martingales (martingales with respect to the Brownian filtration) are essentially the Itô integrals.

Theorem 1.9 (Girsanov). *Let λ_t be progressively measurable with*

$$\mathbb{E} \exp(\frac{1}{2} \int_0^T \lambda^2(t) \, dt) < \infty$$

Then there exists a measure $\mathbb{P}^\star$ such that

(1) $\mathbb{P}^\star$ is equivalent to $\mathbb{P}$,

(2)

$$\frac{dP^\star}{dP} = \exp(-\int_0^t \lambda_t \, dW_t - \frac{1}{2} \int_0^t \lambda_t^2 \, dt)$$

(3) $W_t^\star = W_t + \int_0^t \lambda_s \, ds$ is a $\mathbb{P}^\star$-Brownian motion

As a partial corollary, if $\mathbb{P}^\star$ is equivalent to $\mathbb{P}$ then there exists a progressively measurable process λ_t such that

$$W_t^\star = W_t + \int_0^t \lambda_s \, ds$$

is a Brownian motion under $\mathbb{P}^\star$.

Corollary. *We can then use Girsanov's theorem to transform a Brownian motion with drift to a martingale. e.g. Under $\mathbb{P}$,*

$$\begin{aligned} dX_t &= \mu_t\,dt + \sigma_t\,dW_t \\ &= \sigma_t d(W_t + \int_0^t \sigma_s^{-1}\mu_s\,ds) \\ &= \sigma_t dW_t^\star \end{aligned}$$

where we set $\lambda_s = \sigma_s^{-1}\mu_s$ in Girsanov's theorem.

Theorem 1.10 (Multivariate Itô's lemma)**.** *Let $dX_{i,t} = \alpha_i\,dt + \beta_i\,dW_{i,t}$ with $W_{i,t}$ correlated Brownian motions. Then if $F(X_{1,t},\dots,X_{n,t},t)$ is $C_{2,1}$, then*

$$dF = \left(F_t + \sum_{i=1}^n \alpha_i F_i + \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n \beta_i\beta_j\rho_{ij}F_{ij}\right)dt + \sum_{i=1}^n \beta_i F_i\,dW_i(t)$$

2. Black-Scholes PDE Method

Theorem 2.1 (Black-Scholes PDE)**.** *Let $f(X_t,t)$ represent the price of a contingent claim on an asset X_t, where X_t is assumed to follow geometric Brownian motion. Under certain assumptions, we can derive the Black-Scholes PDE,*

$$f_t = rf - rxf_x - \frac{1}{2}\sigma^2x^2f_{xx}$$

Solving the Black-Scholes PDE along with initial conditions and payoff at expiration yields the function $f(X_t,t)$ which gives the option value at any time t and any underlying value X_t.

3. Martingale method

Consider a market with risky security X_t and riskless security B_t.

Definition 3.1 (Contingent claim)**.** A random variable $C_T : \Omega \to \mathbb{R}$, $\mathcal{F}_T$-measurable is called a contingent claim. If C_T is $\sigma(X_T)$-measurable it is **path-independent.**

Definition 3.2 (Strategy)**.** Let α_t represent number of units of X_t, and β_t represent number of units of B_t. If α_t, β_t are $\mathcal{F}_t$-adapted, then they are strategies in our market model. Our strategy value V_t at time t is

$$V_t = \alpha_t X_t + \beta_t B_t$$

Definition 3.3 (Self-financing strategy)**.** A strategy (α_t, β_t) is self financing if

$$dV_t = \alpha_t\,dX_t + \beta_t\,dB_t$$

The intuition is that we make one investment at $t = 0$, and after that only rebalance between X_t and B_t.

Definition 3.4 (Admissible strategy)**.** (α_t, β_t) is an **admissible strategy** if it is self financing and $V_t \geq 0$ for all $0 \leq t \leq T$.

Definition 3.5 (Arbitrage)**.** An arbitrage is an admissible strategy such that $V_0 = 0, V_T \geq 0$ and $\mathbb{P}(V_T > 0) > 0$.

Definition 3.6 (Attainable claim)**.** A contingent claim C_T is said to be attainable if there exists an admissible strategy (α_t, β_t) such that $V_T = C_T$. In this case, the portfolio is said to replicate the claim. By the law of one price, $C_t = V_t$ at all t.

Definition 3.7 (Complete)**.** The market is said to be **complete** if every contingent claim is attainable

Theorem 3.8 (Harrison and Pliska)**.** *Let $\mathbb{P}$ denote the real world measure of the underlying asset price X_t. If the market is arbitrage free, there exists an equivalent measure $\mathbb{P}^\star$, such that the discounted asset price $\hat{X}_t$ and every discounted attainable claim $\hat{C}_t$ are $\mathbb{P}^\star$-martingales. Further, if the market is complete, then $\mathbb{P}^\star$ is unique. In mathematical terms,*

$$C_t = B_t \mathbb{E}^\star(B_T^{-1} C_T \,|\, \mathcal{F}_t).$$

$\mathbb{P}^\star$ is called the equivalent martingale measure (EMM) or the risk-neutral measure.

4. Monte Carlo methods

4.1. **Method of antithetic variances.** Instead of simulating X, also simulate a random variable Z with the same variance and expectation as X, but is negatively correlated with X. Then take as Y the random variable

$$Y = \frac{X+Z}{2}$$

Obviously $\mathbb{E}(Y) = \mathbb{E}(X)$. On the other side, we have

$$\begin{aligned}\mathrm{Var}(Y) &= \mathrm{Cov}\left(\frac{X+Z}{2}, \frac{X+Z}{2}\right) \\ &= \frac{1}{4}\mathrm{Var}(X) + 2\mathrm{Cov}(X,Z) + \mathrm{Var}(Z) \leq \frac{1}{2}\mathrm{Var}(X)\end{aligned}$$

So we can reduce variance by a factor of two.

4.2. **Control variate method.**

Theorem 4.1. *Suppose we seek to estimate $\theta = \mathbb{E}(Y)$ where $Y = h(X)$ is the outcome of a simulation. Suppose that Z is also an output of the simulation, and assume that $\mathbb{E}(Z)$ is known. Let*

$$c = \frac{Cov(Y,Z)}{Var(Z)}. \tag{‡}$$

Then

$$\hat{\theta}_c = Y + c(\mathbb{E}(Z) - Z) \tag{$\dagger$}$$

is an unbiased estimator of θ*, and if Cov*$(Y, Z) \neq 0$*,* $\hat{\theta}_c$ *has a lower variance than* $\hat{\theta} = Y$*, and indeed has the lowest variance for all estimators of the form*

$$\hat{\theta}_\gamma = Y + \gamma(\mathbb{E}(Z) - Z)$$

Proof. We have

$$\mathrm{Var}(\hat{\theta}_c) = \mathrm{Var}(Y) + c^2\mathrm{Var}(Z) - 2c\,\mathrm{Cov}(Y, Z). \tag{$\star$}$$

From elementary methods of calculus, we see that $\mathrm{Var}\hat{\theta}_c$ is minimised at

$$c = \frac{\mathrm{Cov}(Y, Z)}{\mathrm{Var}(Z)}$$

Substituting in this value for c in ($\star$), we obtain

$$\begin{aligned}\mathrm{Var}(\hat{\theta}_c) &= \mathrm{Var}(Y) - \frac{\mathrm{Cov}(Y, Z)^2}{\mathrm{Var}(Z)} \\ &= \mathrm{Var}(\hat{\theta}) - \frac{\mathrm{Cov}(Y, Z)^2}{\mathrm{Var}(Z)}\end{aligned}$$

and thus we only need $\mathrm{Cov}(Y, Z) \neq 0$ to obtain our variance reduction.

In practice, we do not know $\mathrm{Cov}(Y, Z)$. Thus, we have to do a number of *burn-in* simulations to generate Y and Z, and then compute an estimate $\hat{c}$ to use in the full simulation. ☐

5. Numerical Simulation of Stochastic Differential Equations

Theorem 5.1. *Let*

$$dX_t = a(t, X_t)\, dt + b(t, X_t)\, dB_t$$

Assume $\mathbb{E}X_0 < \infty$*.* X_0 *is independent of* B_s *and there exists a constant* $c > 0$ *such that*

(1) $|a(t, x)| + |b(t, x)| \leq C(1 + |x|)$.

(2) $a(t, x), b(t, x)$ *satisfy the Lipschitz condition in* x*, i.e.*

$$|a(t, x) - a(t, y)| + |b(t, x) - b(t, y)| \leq C|x - y|$$

for all $t \in (0, T)$.

Then there exists a ***unique (strong) solution****.*

Definition 5.2 (Strong convergence)**.** A numberical scheme for solving an SE is said to converge with strong order γ, if for sufficiently small Δ, we have

$$\mathbb{E}(|X(T) - X_N|) \leq K_T \Delta^\gamma$$

This implies that the generated paths approximate the true paths of the SDE - and so one calls this path-wise convergence or strong convergence.

Definition 5.3 (Weak convergence)**.** A numerical scheme for solving an SDE is said to converge with weak order β if for sufficiently small Δ and each polynomial g, we have

$$|\mathbb{E}(g(X_T)) - \mathbb{E}(g(X_N))| \leq K_{g,T}\Delta^\beta$$

Note that strong convergence always implies weak convergence.

Note also that strong convergence implies pathwise convergence. This is true by Markov's inequality, we have

$$\begin{aligned}\mathbb{P}(|X_n - X(T)| \geq \Delta^{\beta/2})o &\leq \frac{\mathbb{E}(|X_n - X(T)|)}{\Delta^{\beta/2}} \\ &\leq C\frac{\Delta^\beta}{\Delta^{\beta/2}}\end{aligned}$$

Note.

1. Weak convergence is basically convergence in distribution, but it has no path-wise properties.
2. If terms like $\mathbb{E}(h(X_T))$ are computed via Monte Carlo, then the weak convergence concept is sufficient.
3. If the option is a path dependent option, then strong convergence is the right concept, as the payoff depends on the whole path, rather than the distribution of the terminal value of the stock.

Theorem 5.4 (Euler-Maruyama scheme)**.**

$$\begin{aligned}X_0 &= X(0) \\ X_{n+1} = X_n + a(t_n, X_n)\Delta t_n + b(t_n, X_n)\Delta W_n\end{aligned}$$

where

$$\begin{aligned}\Delta t_n &= t_{n+1} - t_n \\ \Delta W_n = W_{t_{n+1}} - W_{t_n} l\end{aligned}$$

Euler-Maruyama has ***strong convergence order*** $\gamma = \frac{1}{2}$ *and* ***weak convergence order*** $\beta = 1$*.*

Theorem 5.5 (Milstein scheme)**.** *Consider the homogenous scalar stochastic differential equation*

$$dX_t = a(X_t)\,dt + b(X_t)\,dW_t$$

$$X_0 = X(0)$$

$$X_{n+1} = X_n + a(X_n)\Delta t_n + b(X_n)\Delta W_n + \frac{1}{2}b'(X_n)b(X_n)((\Delta W_n)^2 - \Delta t_n)$$

One can prove that the Milsten scheme has ***strong and weak convergence order*** $\gamma = 1$.

6. Stochastic Optimal Control

Definition 6.1 (Controlled stochastic differential equation).

$$dx(t) = f(t, x(t), u(t))\,dt + \sigma(t, x(t), u(t))\,dW(t)$$

where $u(t, \omega) = u(t, x(t, \omega))$ is a stochastic process, known as the **control**.

Definition 6.2 (Admissible control). A control u is called admissible for the constraints if for every initial value $x_0 \in S$ the corresponding stochastic differential equation has a unique solution with $x(0) = x_0$ and $u(t, \omega) \in \mathcal{U}$ for all $t \in [0, \infty]$. We denote the set of admissible controls with $\mathcal{A}$.

Definition 6.3 (Stochastic optimal control problem). We seek to solve

$$\max_{u \in \mathcal{A}} \mathbb{E}\left[\int_0^T e^{-rt} B(t, x(t), u(t))\,dt + e^{-rT} S(x(T)) \cdot \mathbf{1}_{T<\infty}\,dt\right]$$

under the dynamic constraint

$$dx(t) = f(t, x(t), u(t))\,dt + \sigma(t, x(t), u(t))\,dW(t)$$

with initial condition $x(0) = x_0$, and discount rate $r > 0$.

B is called the benefit function, S is called the final payoff, and the control u is called the optimal control, and the optimal value is called the value of the problem.

Definition 6.4 (Value function).

$$\max_{u \in \mathcal{A}} \mathbb{E}\left[\int_t^T e^{-r(s-t)} B(s, x(s), u(s))\,ds + e^{-r(T-t)} S(x(T)) \cdot \mathbf{1}_{T<\infty}\,dt \,|\, x(t) = x\right]$$

subject to

$$\begin{aligned} dx(s) &= f(s, x(s), u(s))\,ds + \sigma(s, x(s), u(s))\,dW(s) \\ x(t) &= x \end{aligned}$$

Note that $V(0, x_0)$ is the value of the optimal control problem. $V(t, x)$ is the value of the problem, if we started at time t with initial state x.

Theorem 6.5 (Hamilton-Jacobi-Bellman equation). *Assume $T < \infty$. Let $V : [0, T] \times S \to R$ be a $C_{1,2}$ function and assume it satisfies the HJB equation*

$$\begin{aligned} rV(t, x) - V_t(t, x) &= \max_{u \in \mathcal{A}} \left(B(t, x, u) + V_x(t, x) f(t, x, u(t)) + \frac{1}{2} tr(V_{xx}(t, x)\sigma(t, x, u)\sigma(t, x, u)^T) \right) \\ V(T, x) &= S(x). \end{aligned}$$

Let $\varphi(t,x)$ be the set of maximisers of the right hand side and let $u^* \in \mathcal{A}$ such that $u^*(t,\omega) \in \varphi(t, x(t,\omega))$ for all $t \in [0,T], \omega \in \Omega$. Then u^* is the optimal control and V is the value function for the stochastic optimal control problem.

Theorem 6.6 (Hamilton-Jacobi-Bellman equation, infinite time)**.** *Consider the time homogenous, infinite time horizon problem*

$$\max_{u \in \mathcal{A}} \mathbb{E}\left[\int_0^\infty e^{-rt} B(x(t), u(t))\, dt\right]$$

subject to

$$dx(t) = f(x(t), u(t))\, dt + \sigma(x(t), u(t))\, dW_t.$$

Then the value function is independent of t, and so $V(t,x) = V(x)$, and the optimal control is of the type $u(t,x) = u(x)$. The HBJ equation in this case becomes the ODE

$$rV(x) = \max_u \left(B(x,u) + V'(x) f(x,u) + \frac{1}{2} V''(x) \sigma(x,u)^2 \right)$$