Mathematical finance: basic models and unsolved problems

Nikolai Dokuchaev

Department of Mathematics and Statistics,

University of Limerick, Ireland.

email Nikolai.Dokuchaev@ul.ie

http://www.staff.ul.ie/nikolaid/

April 22th, 2005

Abstract

Mathematical finance is a relatively new mathematical field. It was in a phase of explosive growth last 10-15 years, and there is very indication it will continue growing for a while yet. The growth is due to a combination of demand from financial institutions and a breakthrough in the mathematical theory of option pricing. The talk will outline basic mathematical theorems and ideas used here, some unsolved problems, and author's results for optimal investment problem in maximin setting.

The main parts of MF are

(1) Option pricing based on Black-Scholes concept, i.e., on the principle of "No Arbitrage", on attainability of contingent claims, and on assumptions about the stochastic stock price process.

    Related mathematical fields: stochastic processes (martingales, Itô's calculus), PDEs (nonlinear Bellman equations, Stefan problem, numerical methods), optimal stopping.

(2) Optimal portfolio selection.

    Related mathematical fields: stochastic optimal control, optimization, game theory, Bellman equations, filtering and parameters estimation.

(3) Statistical finance and parameters estimation.

    Related fields: statistics, econometrics, filtering and parameters estimation.

(4) Modelling of financial instruments and markets.

(5) Problems related to mathematical economics (equilibrium, agents behavior, demand-supply relationship)

1 Stochastic market models

Consider a risky asset (stock, bond, foreign currency unit, etc.) with time series of prices $s_1, s_2, s_3, \ldots$, for example, daily prices.

1.1 Discrete time model

Consider historical discrete time prices s_k at times t_k, where $s_k = S(t_k)$, $k = 1, 2, ..., n$. Let

$$\xi_{k+1} \triangleq \frac{s_{k+1}}{s_k} - 1.$$

Clearly,

$$s_{k+1} = s_k(1 + \xi_{k+1})$$

Bachelier (1900) was first who discovered the Square Root Law: the prices of real stocks are such that

$$\mathrm{Var}\, \xi_k \sim \mathrm{const}\, /n,$$

i.e.,

$$\xi_k \sim \mathrm{const}\, \sqrt{1/n} \sim \mathrm{const}\, \sqrt{t_k - t_{k-1}}.$$

Therefore,

$$\mathrm{Var}\, (S(t_2) - S(t_1)) \sim \mathrm{const}\, |t_2 - t_1|,$$

i.e.,

$$S(t_2) - S(t_1) \sim \mathrm{const}\, \sqrt{t_2 - t_1}.$$

This property of $\{s_k\}$ matches with the one for the Ito's processes.

Clearly, the discrete time model does not reqiure such advanced theory as the continuous model, and it describes the real market immediately, hence much better. However, it is not really popular in Mathematical Finance. The reason is that theoretical results are more difficult for discrete time. Some powerful theorems are not valid for a general discrete-time model. For example, the discrete time market is *incomplete* if ξ_k can have more than two values.

1.2 Continuous time model

The premier model of price evolution is such that $s_k = S(t_k)$, where

$$S(t) = S(0)e^{\int_0^s \mu(s)ds + \xi(t)}, \tag{1.1}$$

where $\xi(t)$ is a continuous time *martingale*, i.e., $\mathbf{E}\{\xi(T)|\xi(\cdot)|_{[0,t]}\} = \xi(t)$ for any t and $T > t$. For the simplest model, $\xi(t)$ is a Gaussian process such that $\xi(t+\Delta) - \xi(t)$ does not depend on $\xi(\cdot)|_{[0,t]}$ for any $t \geq 0.$,

$$\mathbf{E}\{\xi(t+\Delta t)|\xi(\cdot)|_{[0,t]}\} = \xi(t),$$
$$\mathrm{Var}\,[\xi(t+\Delta t) - \xi(t)] \sim \sigma^2 \cdot \Delta t \quad \forall t > 0,\ \Delta t > 0,$$

$\sigma \in \mathbf{R}$ is a parameter.

Eq. (1.1) can be replaced for the following *Itô's equation* :

$$dS(t) = S(t)[a(t)dt + \sigma(t)dw(t)], \tag{1.2}$$

where

$$\left.\begin{array}{lll} a(t) & - & \text{appreciation rate} \\ \sigma(t) & - & \text{volatility} \end{array}\right\} \text{market parameters}$$

Here $w(t)$ is a Brownian motion (Wiener process).

Let us discuss some basic properties of Itô's equation (1.2). The solution $S(t)$ of this equation is such that

- sample paths maintain continuity;
- paths are non-differentiable;
- paths are not absolutely continuous, and any path forms a fractal of a fractional dimension;
- if a, σ are deterministic, then the relative-increments $[S(t)-S(\tau)]/S(\tau)$ are independent of the σ-algebra $\sigma(S(\cdot)|_{[0,\tau]})$, $0 \leq \tau < t$.
- if a, σ are deterministic and constant, then

$$\log \frac{S(t+\Delta t)}{S(t)} \sim N(a\Delta t, \sigma^2 \Delta t).$$

It will be shown below that the diffusion model is the only possible continuous model for the risky asset prices.

For a multistock market model, $S(t) = \{S_i(t)\}$, $a = \{a_i\}$, $w = \{w_i\}$ are vectors, and $\sigma = \{\sigma_{ij}\}$ is a matrix.

There are the following key problems:

- *Optimal investment problem: To find a strategy of buying and selling stocks*
- *Pricing problem: To find a "fair" price for derivatives (i.e. options, futures, etc.)*

There is an auxiliary problem:

- *To estimate the parameters* $(a(t), \sigma(t))$ *from market statistics.*

In fact, $\sigma(t)\sigma(t)^\top$ is an explicit function of $S(\cdot)$. The estimation of $a(\cdot)$ is much more difficult.

If a and σ are constant and deterministic, then the process $S(t)$ is log-normal (i.e., the process $\log S(t)$ is Gaussian). Empirical research has shown that the real distribution of stock prices is not exactly log-normal. The imperfection of the log-normal hypothesis on the prior distribution of stock prices can be taken into account by assuming that a and σ are random processes. This more sophisticated model is much more challenging: for example, the market is *incomplete* (i.e., an arbitrary random claim cannot be replicated by an adapted self-financing strategy).

2 Stochastic calculus

2.1 Itô's integral

We are given a standard complete probability space $(\Omega, \mathcal{F}, \mathbf{P})$, $\Omega = \{\omega\}$.

Let $w(t)$ be a Wiener process (or Brownian motion). Let $\mathcal{F}_t$ be the filtration generated by w. Let $L_{22}(0,T)$ be the (Hilbert) space of $\mathcal{F}_t$-adapted functions f such that $\|f\|^2 = \mathbf{E}\int_0^T |f(t,\omega)|^2 dt < +\infty$.

Let $I : L_{22} \to L_2(\Omega, \mathcal{F}_T, \Omega)$ be such that

$$I(f) = \sum f(t_i)[w(t_{i+1}) - w(t_i)]$$

for piece-wise constant f (with jumps at t_i only). Then Itô's integral

$$I(f) \triangleq \int_0^T f(t,\omega) dw(t)$$

is defined as continuous mapping

$$I : L_{22} \to L_2(\Omega, \mathcal{F}_T, \mathbf{P}).$$

This mapping is isometric (up to affine transform):

$$\mathbf{E}I(f) = 0, \quad \mathbf{E}I(f)^2 = \mathbf{E}\int_0^T f(t,\omega)^2 dt.$$

Warning: Itô's integral $I(f) = I(f,T)$ is not defined as function of T for a given $\omega \in \Omega$!

2.2 Itô's formula

Let

$$y(t) = y(s) + \int_s^t a(r)dr + \int_s^t b(r)dw(r),$$

i.e.

$$dy(t) = a(t)dt + b(t)dw(t).$$

Then

$$d_t V(y(t),t) = \Big[\frac{\partial V}{\partial t}(y(t),t) + \mathcal{L}V(y(t),t)\Big]dt + \frac{\partial V}{\partial y}(y(t),t)b(t)dw(t),$$

where

$$\mathcal{L}V = \frac{\partial V}{\partial y}(y(t),t)a(t) + \frac{1}{2}\frac{\partial^2 V}{\partial y^2}(y(t),t)b(t)^2.$$

Proof is based on Teylor series and the estimate

$$\Delta y \triangleq y(t+\Delta t) - y(t) \sim a\Delta t + b\Delta w,$$

where $(\Delta w)^2 \sim \Delta t$.

2.3 Most important theorems

Clark-Haussmann formula

Let $F : C([0,T];\mathbf{R}) \to \mathbf{R}$ be a mapping, then there exists a $w(t)$-adapted process $f(t) = f(t,\omega)$ such that

$$F(w(\cdot)) = \mathbf{E}F(w(\cdot)) + \int_0^T f(t)dw(t).$$

One may say that this theorem claims that the mapping $I : L_{22}(0,T) \to L_2^{(0)}(\Omega, \mathcal{F}_T, \mathbf{P})$ is an isometric bijection, where $L_2^{(0)}(\Omega, \mathcal{F}_T, \mathbf{P}) = \{\xi \in L_2(\Omega, \mathcal{F}_T, \mathbf{P}) : \ \mathbf{E}\xi = 0\}$.

There is no analog of this in the deterministic calculus!

Girsanov's Theorem

Any Ito's process

$$y(t) = y(s) + \int_s^t a(r)dr + \int_s^t b(r)dw(r)$$

with drift a is an Ito's process with zero drift (i.e, a marttingale) for some another probability measure. In other words,

$$dy(t) = b(t)dw_*(t),$$

where $w_*(t)$ is a process such that it is a Wiener process and some new probability measure $\mathbf{P}_*$ and such that

$$b(t)dw_*(t) = a(t)dt + b(t)dw(t).$$

There are many other amazing results without analogs of this in deterministic calculus. They looks unusual, but in fact these properties ensures that Itô's processes represent the ultimate model for stock prices.

Ito's processes and PDE

If $a(t) = f(y(t), t)$, $b(t) = \beta(y(t), t)$, then $y(t)$ is a Markov diffusion process. Let $H(x, s)$ be the solution of Cauchy problem for the backward parabolic equation

$$\begin{cases} \frac{\partial H}{\partial s}(x,s) + \frac{\partial H}{\partial x}(x,s)a(x,s) + \frac{1}{2}\frac{\partial^2 H}{\partial y^2}(x,s)b(x,s)^2 = 0, \\ H(x,T) = \Phi(x). \end{cases}$$

Let $\xi = \Phi(y(T))$. In this case, Clark-Haussmann formula can be rewritten as the following.

Theorem 2.1

$$\xi = \mathbf{E}\xi + \int_0^T \frac{\partial H}{\partial x}(y(t), t)b(t)dw(t),$$

and $H(x, s) = \mathbf{E}\{\Phi(y(T))|y(s) = x\}$.

Proof. Let $y(s) = x$. By Ito's formula,

$$\begin{aligned} &\mathbf{E}\Phi(y(T)) - H(x,s) = \mathbf{E}H(y(T),T) - H(y(s),s) \\ &= \mathbf{E}\textstyle\int_s^T [\frac{\partial H}{\partial t} + \mathcal{L}H](y(t),t)dt + \mathbf{E}\int_s^T \frac{\partial H}{\partial y}(y(t),t)b(t)dw(t) = 0. \end{aligned}$$

□

Remark. It follows that the diffusion process $y(t)$ can be considered as characteristics of the parabolic equation. It is known from physical models that the speed of heat propagation is infinite, and that the wave propagation described by the first order hyperbolic equations has bounded speed. That means that "physical" diffusion processes has unlimited speed.

In particular, if $f \equiv ay(t)$, $b(t) \equiv \sigma y(t)$, then the equation for $y(t)$ is the equation for the stock price $dS(t) = S(t)[adt + \sigma dw(t)]$.

3 Market model: wealth and portfolio strategies

Assume that there are n stocks with prices $S_1, ..., S_n$. We usually assume also that there is a riskless asset (bond, or bank account) with price

$$B(t) = B(0)\exp\Big(\int_0^t r(s)ds\Big),$$

where $r(t)$ is a process of risk-free interest rates.

The portfolio is a process $(\gamma(\cdot), \beta(\cdot))$ with values in $\mathbf{R}^n \times \mathbf{R}$, $\gamma(\cdot) = (\gamma_1(t), \dots, \gamma_n(t))$, where $\gamma_i(t)$ is the quantity of the ith stock; and $\beta(t)$ is the quantity of the bond.

A portfolio $(\gamma(\cdot), \beta(\cdot))$ is said to be *self-financing* if there is no income from or outflow to external sources. In that case,

$$dX(t) = \sum_{i=1}^n \gamma_i(t)dS_i(t) + \beta(t)dB(t).$$

It can be seen that

$$\beta(t) = \frac{X(t) - \sum_{i=1}^n \gamma_i(t)S_i(t)}{B(t)},$$

and the equation for the self-financing $X(t)$ became given $\gamma(t)$.

Let

$$\pi_0(t) \triangleq \beta(t)B(t),$$
$$\pi_i(t) \triangleq \gamma_i(t)S_i(t), \quad \pi(t) = (\pi_1(t), \dots, \pi_n(t))^\top.$$

By the definitions, the process $\pi_0(t)$ is the investment in the bond, and $\pi_i(t)$ is the investment in the ith stock. We have shown above that the vector π alone suffices to specify the self-financing portfolio. We shall use the term *self-financing strategy* for a vector process $\pi(\cdot) = (\pi_1(t), \dots, \pi_n(t))$, where the pair $(\pi_0(t), \pi(t))$ describes the self-financing portfolio at time t:

$$X(t) = \sum_{i=1}^n \pi_i(t) + \pi_0(t).$$

3.1 Why the diffusion model is the only meaningful continuous model

Let $n = 1$, $r(t) \equiv 0$, i.e., $dB(t) \equiv 0$, then it follows that

$$X(T) = X(0) + \int_0^T \gamma(t)dS(t).$$

Suppose that there exists $t_1 < t_2$ such that there exists $\frac{dS}{dt}(t)$ on $[t_1, t_2]$. Then one can take

$$\gamma(t) \triangleq \begin{cases} M\text{Sign}\,\frac{dS}{dt}(t) & t \in [t_1, t_2] \\ 0 & \text{otherwise} \end{cases}$$

and

$$X(T) = X(0) + \int_{t_1}^{t_2} \gamma(t)dS(t) = M \int_{t_1}^{t_2} \left| \frac{dS}{dt}(t) \right| dt.$$

Thus, one can have a risk free positive gain (which can be arbitrarily large for large enough M). Therefore, continuous time functions can model price effectively only if they are nowhere differentiable.

3.2 A paradox

Consider a diffusion market model with a single stock $S(t)$ with $r = 0$ (with zero interest bank account). Assume that $dS(t) = S(t)dw(t)$, where $w(t)$ is the scalar Wiener process (or, for simplicity, you can assume that $S(t) = S(0) + w(t)$). A (self-financing) strategy of an investor is the number of shares $\gamma(t)$, and the corresponding wealth is such that $X(t) = X(0) + \int_0^t \gamma(s)dS(s)$. John had initial wealth $X_0 = S(0)$ and he uses the following strategy: $\gamma(t) = \mathbb{I}_{\{S(t) \geq S(0)\}}$, where $\mathbb{I}$ denotes the indicator function.

This means that John keeps one share of stock when $S(t) \geq S(0)$ and keeps zero amount of shares if $S(t) < S(0)$, i.e., in that case he keeps all money in risk-free cash account).

John hopes to have the wealth $X_T = \max(S(0), S(T))$ at time T. Is this risk-free gain?

4 Pricing problem

The basic approach here is to use Ito's formula and the rule that the price P of an option is the minimal initial wealth that can be raised to the terminal derivative value ϕ by self-financing strategies:

$$Price = \inf\{X_0 : \ \exists\gamma(\cdot) : X(T) = X_0 + \int_0^T \gamma(t)dS(t) \geq \phi \quad \text{a.s}\}.$$

(We assume that the risk free rate r is zero).

Case of martingale price

Let $a = 0$. Let $\phi = \Phi(S(\cdot)|_{t\in[0,T]})$, i.e, it is an European type of an option. For instance, we allow $\phi = (S(T) - K)^+$. Then, by Clark theorem, there exists a process f such that

$$\phi = \mathbf{E}\phi + \int_0^T f(t)dw(t) = c + \int_0^T f(t)S(t)^{-1}dS(t).$$

Hence $\gamma(t) = S(t)^{-1}f(t)$ can be considered as a self-financing strategy that replicates the claim ϕ, and $Price = \mathbf{E}\phi$ is the initial wealth that ensures replicating.

Non-martingale case

This case can be effectively reduced to the martingale case with Girsanov Theorem and measure change. In that case, $Price = \mathbf{E}_*\phi$, where $\mathbf{E}_*$ is the expectation with respect to a measure such that $S(t)$ is a martingale.

Markov case

Let $y(t)$ be a Markov process, $\phi = \Phi(S(T))$.

Let $H(x, s)$ be the solution of Cauchy problem for the backward parabolic equation

$$\begin{cases} \frac{\partial H}{\partial s}(x, s) + \frac{1}{2}\frac{\partial^2 H}{\partial y^2}(x, s)\sigma^2 y^2 = 0, \\ H(x, T) = \Phi(x). \end{cases}$$

By Ito's formula,

$$\begin{aligned} &\Phi(S(T)) - H(x, 0) = H(S(T), T) - H(S(0), 0) \\ &= \mathbf{E}\int_s^T [\tfrac{\partial H}{\partial t} + \mathcal{L}H](S(t), t)dt + \int_s^T \tfrac{\partial H}{\partial y}(y(t), t)b(t)dw(t) \\ &= \int_s^T \tfrac{\partial H}{\partial y}(y(t), t)S(t)^{-1}dS(t). \end{aligned}$$

If $dS(t) = \sigma S(t)dw(t)$, i.e., the appreciation rate a is zero, then

$$\Phi(S(T)) - H(x,0) = \int_s^T \frac{\partial H}{\partial y}(y(t),t)S(t)^{-1}dS(t).$$

Hence $X(t) = H(S(t),t)$ is the wealth and

$$X(0) = H(S(0),0) = \mathbf{E}\{\Phi(S(T))|S(0), a \equiv 0\}.$$

If $\Phi(x) = \max(0, x-K)$, then $H(x,s)$ is the Black-Scholes price of a call option (with zero risk free rate, and the parabolic equation is said to be Black-Scholes equation. The explicit solution of this equation with this special Φ gives the celebrated Black-Scholes formula for option price. Clearly, change of variable $x = e^y$ can convert it to a heat equation.

5 Portfolio selection problem

5.1 Optimal investment problem

We can state a generic *optimal investment problem:*

$$\begin{aligned} &\text{Maximize} \quad \mathbf{E}U(X(T)) \\ &\text{over self-financing strategies} \quad \pi(\cdot). \end{aligned}$$

Here T is the terminal time, and $U(\cdot)$ is a given *utility function* that describes risk preferences. The most common utilities are log and power, i.e., $U(x) = \log x$ and $U(x) = \delta^{-1}x^{\delta}$, $\delta < 1$.

There are many modifications of the generic optimal investment problem:

- optimal investment-consumption problems
- optimal hedging of non-replicable claims
- problem with constraints
- $T = +\infty$
- etc.

Merton's strategy

We describe now strategies that are optimal for the generic model with $U(x) = \log x$ or $U(x) = \delta^{-1}x^{\delta}$:

$$\pi(t)^{\top} = \nu(a(t) - r(t)\mathbf{1})^{\top}[Q(t)X(t)],$$

where $\nu = \nu(\delta) = (1-\delta)^{-1}$, $Q(t) \triangleq (\sigma(t)\sigma(t)^{\top})^{-1}$, $r(t)$ is the interest rate for a risk-free investment, and $\mathbf{1}^{\top} \triangleq (1, 1, \ldots, 1)^{\top}$.

Note that these strategies require direct observation of (σ, a). But, in practice, the parameters $a(\cdot), \sigma(\cdot)$ need to be estimated from historical market data.

6 Examples of unsolved problem

6.1 Discrete time market

The real observed prices are given as time series, so the optimal solution for continuous model is optimal only for the continuous limit, and it is not optimal after application to the time series of prices.

The existing optimal solutions for discrete time processes requires solutions of Bellman's discrete time equations starting from terminal time that includes calculating of conditional densities at any step, and it is not so nice as Merton's strategies.

As far as I know, it is still unknown how to derive optimal Merton's strategy for multi-period discrete time model for $U(x) = x^\delta$ even for the number of steps equal two. (This problem is solved for $U(x) = kx - x^2$, but the solution is not real analog of Merton's strategy).

6.2 Stefan problem for American option

Similarly to European option pricing, the price of American option with payoff $F(S(\tau)$ (where τ is any (Markov random) time chosen by option's holder) satisfies Stefan problem

$$\begin{aligned}&\frac{\partial H}{\partial t}(t,x) + \frac{\sigma^2 x^2}{2}\frac{\partial^2 H}{\partial x^2}(t,x) = r\Big[H(t,x) - x\frac{\partial H}{\partial x}(t,x)\Big], \quad (x,t):\ H(x,t) > F(x),\\ &H(T,x) = F(x).\end{aligned}$$

This problem does not have explicit solution for $F(x) = (K - x)^+$, i.e., for the simplest cases. Moreover, it s very difficult to obtain a solution even for only slightly generalized problem.

6.3 Explicit formulas for replicating strategies

Assume that admissible strategies are using historical prices only. Typically, optimal claim can be found in that setting. However, the integrand in Clark's formula (or the optimal strategy) is unknown for a very generic cases, for instace, for $U(x) = x^{1/3}$.

6.4 Optimal investment problems for observable but unhedgeable parameters

Solution (i.e., the optimal investment strategy) is unknown if we include observable parameters such as trade volume (with unknown evolution law). The only known case if when $U(x) = \log x$.

7 Solution in minimax setting

(This is my result presented in 57th British Mathematical Colloquium, Liverpool, April 4=7, 2005; the paper accepted to "IMA Journal Management Mathematics").

For problems with uncertainty in prior distributions, the most popular and straightforward approach is solution in maximin setting: *Find a strategy which maximizes the infimum of expected utility over all admissible parameters from a given class.* The maximin setting has long history in optimization and optimal control theory. It is presented in robust control, in particular, in H^p-control. In economics, there is a large literature devoted to related investment problems. Uncertainty in prior probability measures is referred sometimes as the Knightian Uncertainty. Maximin setting in mathematical economics is presented in theory of problems with robust performance criteria.

We give a new solution for a general $U(\cdot)$ for an "incomplete market" and for the problem with uncertainty. We assume that the market parameters are observable but their future distribution is unknown and they are non-predictable.

Let $D \subset \mathbf{R}$ be a given interval. Let $T > 0$ and the initial wealth $X(0)$ be given. Let $\mathcal{M} = \{\mu(\cdot)\}$ be a class of parameters $\mu(\cdot) = [r(\cdot), a(\cdot), \sigma(\cdot)]$. Consider the following problem:

$$\text{Maximize} \quad \inf_{\mu \in \mathcal{M}} \mathbf{E} U[\tilde{X}(T)] \quad \text{over} \quad \pi(\cdot): \ X(T) \in D \text{ a.s.}$$

Let

$$R_\mu \triangleq \int_0^T |\theta(s)|^2 ds, \qquad R_{\min} = \inf_{\mu \in \mathcal{M}} R_\mu,$$

where $\theta(t) \triangleq \sigma(t)^{-1}[a(t) - r(t)\mathbf{1}]$ is the market price of risk, $\mathbf{1} = (1, 1, ..., 1) \in \mathbf{R}^n$.

Let the function $F(\cdot, \cdot)$ be such that

$$F(z, \lambda) \in \arg\max{}_x \{zU(x) - \lambda x\}. \qquad (*)$$

Let

$$\tau_\mu(t) \triangleq \frac{1}{R_{\min}} \int_0^t |\theta(s)|^2 \, ds. \qquad (**)$$

Theorem 1 *The optimal strategy (i.e. the strategy for the saddle point) is defined by*

$$\pi(t) = \theta(t) \frac{\partial u}{\partial x}(\mathcal{Z}_t, \tau_\mu(t)), \qquad (7.1)$$

$$d\mathcal{Z}(t) = \theta(t)^\top \mathcal{Z}(t) \, \sigma(t)^{-1} \mathbf{S}(t)^{-1} dS(t), \qquad (7.2)$$

where $\mathbf{S} = \operatorname{diag}(S_1, \ldots, S_n)$*, and where* $u(\cdot)$ *is the solution of the heat equation*

$$\frac{\partial u}{\partial t}(x,t) + R_{\min} \frac{1}{2T} \frac{\partial^2 u}{\partial x^2}(x,t) = 0, \tag{7.3}$$

$$u(x,T) = \tilde{F}(x, R_{\min}, \widehat{\lambda}), \tag{7.4}$$

Here $\widehat{F}(y,R,\lambda) \triangleq F(e^{y+R/2},\lambda)$, and $\widehat{\lambda}$ is a Lagrange multiplier defined from the equation $\mathbf{E}_* \tilde{X}(T) = X(0)$, where $\mathbf{E}_*$ corresponds the risk neutral measure $\mathbf{P}_*$ given σ (i.e., μ) such that $R_\mu = R_{\min}$; and this is the "worst" (*"saddle point"*) μ.

In fact, $\mathcal{Z}(t) = \mathcal{Z}(t,\mu)$, and $\mathcal{Z}(T)^{-1} = d\mathbf{P}_*/d\mathbf{P}$, where $\mathbf{P}$ is the original measure. The solution of the Cauchy problem can be expressed explicitly via integral with known kernel. Note that the heat equation is one dimensional, and the maximization is also one-dimensional, even for a case of a large number of stocks n.

The novelty of this result is that, we obtained explicitly (under certain conditions) the solution of the maximin problem even for a case when the solution is unknown for a given distribution of the random parameters. In other words, the solution in maximin setting with unknown prior distributions appears to be easier than for the problem with given prior distribution.

Proofs

7.1 Additional definitions

Without loss of generality, we describe the probability space as follows: $\Omega = \mathcal{T} \times \Omega'$, where $\Omega' = C([0,T];\mathbf{R}^n)$. We are given a σ-algebra $\mathcal{F}'$ of subsets of Ω' generated by cylindrical sets, and a σ-additive probability measure $\mathbf{P}'$ on $\mathcal{F}'$ generated by $w(\cdot)$. Furthermore, let $\mathcal{F}_{\mathcal{T}}$ be the σ-algebra of all Borel subsets of $\mathcal{T}$, and $\mathcal{F} = \mathcal{F}_{\mathcal{T}} \otimes \mathcal{F}'$. We assume also that each $\mu = \mu(\cdot)$ generates the σ-additive probability measure ν_μ on $\mathcal{F}_{\mathcal{T}}$ (this measure is generated by Θ which corresponds to μ).

Let $\overset{\circ}{\mathbf{R}}{}^n_+ \triangleq (0,+\infty)^n$.

For a function $\Gamma(t,\cdot) : C([0,t];\overset{\circ}{\mathbf{R}}{}^n_+) \times B([0,t];\mathbf{R}\times\mathbf{R}^n\times\mathbf{R}^{n\times n}) \to \mathbf{R}^n$, introduce the following norm:

$$\|\Gamma(\cdot)\|_{\mathbf{X}} \triangleq \mathbf{E} \sup_{\mu(\cdot)=\mu_\alpha(\cdot):\ \alpha\in\mathcal{T}} \Big(\sum_{i=1}^n \int_0^T \Gamma_i(t,[S(\cdot),\mu(\cdot)]|_{[0,t]})^2 dt\Big)^{1/2}. \tag{7.5}$$

Definition 7.1 *Let $\mathcal{C}_0$ be the set of all functions $\Gamma(t,\cdot) : C([0,t];\mathbf{R}^n) \times B([0,t];\mathbf{R}\times\mathbf{R}^n\times\mathbf{R}^{n\times n}) \to \mathbf{R}^n$ such that*

$$\begin{aligned}
&\|\Gamma(\cdot)\|_{\mathbf{X}} < +\infty,\\
&\pi(t) = \Gamma(t,[S(\cdot),\mu(\cdot)]|_{[0,t]}) \in \tilde{\Sigma}(\mathcal{F}^\mu_\cdot) \quad \forall \mu(\cdot),\\
&\tilde{X}(T,\Gamma(\cdot),\mu) \in D \quad \text{a.s.} \quad \forall\mu.
\end{aligned}$$

In fact, $\mathcal{C}_0$ is a subset of the linear space of functions with finite norm (7.5).

7.2 A duality theorem

We need the following duality theorem.

Theorem 7.1 *The following holds:*

$$\begin{aligned}
&\sup_{\Gamma(\cdot)\in\mathcal{C}_0} \inf_\mu \mathbf{E} U(\tilde{X}(T,\Gamma(\cdot),\mu))\\
&\qquad = \inf_\mu \sup_{\Gamma(\cdot)\in\mathcal{C}_0} \mathbf{E} U(\tilde{X}(T,\Gamma(\cdot),\mu)).
\end{aligned} \tag{7.6}$$

By this theorem, it follows that there exists a saddle point.

To prove Theorem 7.1, we need several preliminary results, which are presented below as lemmas.

Lemma 7.1 *The function* $\tilde{X}(T, \Gamma(\cdot), \mu)$ *is affine in* $\Gamma(\cdot)$.

Lemma 7.2 *The set* $\mathcal{C}_0$ *is convex.*

Let μ_α be μ that corresponds non-random $\alpha \in \mathcal{T}$.

Lemma 7.3 *There exists a constant* $c > 0$ *such that*

$$\mathbf{E}|\tilde{X}(T, \Gamma(\cdot), \mu_\alpha)|^2 \le c\,(\|\Gamma(\cdot)\|^2_{\mathbf{X}} + X_0^2) \quad \forall \Gamma(\cdot) \in \mathcal{C}_0, \quad \forall \alpha \in \mathcal{T}.$$

Lemma 7.4 *The function* $\mathbf{E}U(\tilde{X}(T, \Gamma(\cdot), \mu_\alpha))$ *is continuous in* $\Gamma(\cdot) \in \mathcal{C}_0$ *uniformly in* $\alpha \in \mathcal{T}$.

For $\alpha \in \mathcal{T}$, set

$$J'(\Gamma(\cdot), \alpha) \triangleq \mathbf{E}U(\tilde{X}(T, \Gamma(\cdot), \mu_\alpha(\cdot))).$$

Lemma 7.5 *For a given* $\Gamma(\cdot) \in \mathcal{C}_0$*, the function* $J'(\Gamma(\cdot), \alpha)$ *is continuous in* $\alpha \in \mathcal{T}$.

Let $\mathcal{V}$ be the set of all σ-additive probability measures on $\mathcal{F}_{\mathcal{T}}$. We consider $\mathcal{V}$ as a subset of $C(\mathcal{T}; \mathbf{R})^*$. Let $\mathcal{V}$ be equipped with the weak* topology in the sense that

$$\nu_1 \to \nu_2 \quad \Leftrightarrow \quad \int_{\mathcal{T}} \nu_1(d\alpha) f(\alpha) \to \int_{\mathcal{T}} \nu_2(d\alpha) f(\alpha) \quad \forall f(\cdot) \in C(\mathcal{T}; \mathbf{R}).$$

Lemma 7.6 *The set* $\mathcal{V}$ *is compact and convex.*

Proof. The convexity is obvious. It remains to show the compactness of the set $\mathcal{V}$. In our case, the set $\mathcal{T}$ is a compact subset of a finite-dimensional Euclidean space. Now we note that the Borel σ-algebra of subsets of $\mathcal{T}$ coincides with the Baire σ-algebra (see, e.g., Bauer (1981)). Hence, $\mathcal{V}$ is the set of Baire probability measures. By Theorem IV.1.4 from Warga (1972), it follows that $\mathcal{V}$ is compact. This completes the proof. $\square$

We are now in the position to give a proof of Theorem 7.1.

Proof of Theorem 7.1. For a $\Gamma(\cdot) \in \mathcal{C}_0$, we have $J'(\Gamma(\cdot), \cdot) \in C(\mathcal{T}; \mathbf{R})$ and

$$\begin{aligned} \mathbf{E}U(\tilde{X}(T, \Gamma(\cdot), \mu)) &= \textstyle\int_{\mathcal{T}} d\nu_\mu(\alpha) \mathbf{E}U(\tilde{X}(T, \Gamma(\cdot), \mu_\alpha)) \\ &= \textstyle\int_{\mathcal{T}} d\nu_\mu(\alpha) J'(\Gamma(\cdot), \alpha), \end{aligned}$$

where $\nu_\mu(\cdot)$ is the measure on $\mathcal{T}$ generated by Θ which corresponds $\mu(\cdot)$. Hence, $\mathbf{E}U(\tilde{X}(T, \Gamma(\cdot), \mu(\cdot)))$ is uniquely defined by ν_μ. Let

$$J(\Gamma(\cdot), \nu_\mu) \triangleq \mathbf{E}U(\tilde{X}(T, \Gamma(\cdot), \mu(\cdot))).$$

By Lemma 7.5, $J(\Gamma(\cdot), \nu)$ is linear and continuous in $\nu \in \mathcal{V}$ given $\Gamma(\cdot)$.

To complete the proof, it suffices to show that

$$\sup_{\Gamma(\cdot)\in\mathcal{C}_0} \inf_{\nu\in\mathcal{V}} J(\Gamma(\cdot), \nu) = \inf_{\nu\in\mathcal{V}} \sup_{\Gamma(\cdot)\in\mathcal{C}_0} J(\Gamma(\cdot), \nu). \tag{7.7}$$

We note that $J(\Gamma(\cdot), \nu) : \mathcal{C}_0 \times \mathcal{V} \to \mathbf{R}$ is linear in ν. By Lemmas 7.1 and 7.4-7.5, it follows that $J(\Gamma(\cdot), \nu)$ is either concave or convex in $\Gamma(\cdot)$ and that $J(\Gamma(\cdot), \nu) : \mathcal{C}_0 \times \mathcal{V} \to \mathbf{R}$ is continuous in ν for each $\Gamma(\cdot)$ and continuous in $\Gamma(\cdot)$ for each ν. Furthermore, the sets $\mathcal{C}_0$ and $\mathcal{V}$ are both convex, and the set $\mathcal{V}$ is compact. By the Sion Theorem, it follows that (7.7), and hence (7.6), are satisfied. This completes the proof of Theorem 7.1. $\square$

7.3 Proof of Theorem 1

Let $\widehat{\alpha} \in \mathcal{T}$ be such that $R_{\widehat{\mu}} = R_{min}$, where $\widehat{\mu}(\cdot) \triangleq \mu_{\widehat{\alpha}}(\cdot)$. Let $\lambda = \widehat{\lambda}$. It follows that

$$\mathbf{E}U(\tilde{X}(T, \widehat{\Gamma}_{\widehat{\alpha}}(\cdot), \widehat{\mu})) = \sup_{\Gamma(\cdot)\in\mathcal{C}_0} \mathbf{E}U(\tilde{X}(T, \Gamma(\cdot), \widehat{\mu})). \tag{7.8}$$

Let μ be arbitrary. Let

$$\tilde{T}_\mu(t) \triangleq \inf\{s : \int_0^s |\theta_\mu(r)|^2 dr > t\}, \quad T_\mu \triangleq \tilde{T}_\mu(R_{min}).$$

Clearly, $T_\mu \le T$, and $T_\mu = \inf\{t : \ \tau_\mu(t, R_{min}) = T\}$, where $\tau_\mu(t, R) \triangleq \frac{T}{R}\int_0^t |\theta_\mu(s)|^2 ds$.

Set $I_\mu(t) \triangleq \int_0^t \theta_\mu(s)^\top dw(s)$. This is a martingale. By the Dambis–Dubins–Schwarz theorem, $I'_\mu(t) \triangleq I_\mu(\tilde{T}_\mu(t))$ is a Brownian motion. We have that

$$Y(T, \widehat{\mu}) = \int_0^T |\theta_{\widehat{\mu}}(t)|^\top dw(s), \qquad \ln \mathcal{Z}(T, \widehat{\mu}) = \int_0^T |\theta_{\widehat{\mu}}(t)|^\top dw(s) + \tfrac{1}{2}R_{min},$$

$$Y(T_\mu, \mu) = I_\mu(T_\mu) = I'_\mu(R_{min}), \qquad \ln \mathcal{Z}(T_\mu, \mu) = I_\mu(T_\mu) + \tfrac{1}{2}R_{min} = I'_\mu(R_{min}) + \tfrac{1}{2}R_{min}.$$

These two random variables are Gaussian with mean $R_{min}/2$ and variance R_{min}. Therefore, the variables $Y(T, \widehat{\mu})$ and $Y(T_\mu, \mu)$ have the same probability distribution.

It is easy to see that the process

$$\tilde{X}'(t) = \begin{cases} u(Y(t,\mu), \tau_\mu(t, R_{min}), R_{min}, \widehat{\lambda}), & t \le T_\mu \\ \tilde{X}'(T_\mu), & t > T_\mu \end{cases}$$

is the normalized self-financing wealth for some admissible strategy $\Gamma'(\cdot) \in \mathcal{C}_0$, i.e., $\tilde{X}'(t) = \tilde{X}(t, \Gamma'(\cdot), \mu)$. Furthermore,

$$\tilde{X}'(T) = \tilde{X}(T, \Gamma'(\cdot), \mu) = F(\mathcal{Z}(T_\mu, \mu), \widehat{\lambda}) = \widehat{F}(Y(T_\mu, \mu(\cdot)), R_{min}, \widehat{\lambda}),$$

and this variable has the same distribution as

$$\tilde{X}(T, \widehat{\Gamma}_{\widehat{\alpha}}(\cdot), \widehat{\mu}) = F(\mathcal{Z}(T, \widehat{\mu}), \widehat{\lambda}) = \widehat{F}(Y(T, \widehat{\mu}), R_{min}, \widehat{\lambda}).$$

Hence

$$\mathbf{E}U(\tilde{X}(T, \widehat{\Gamma}_{\widehat{\alpha}}(\cdot), \widehat{\mu})) = \mathbf{E}U(\tilde{X}(T, \Gamma'(\cdot), \mu)).$$

Therefore,

$$\mathbf{E}U(\tilde{X}(T, \widehat{\Gamma}_{\widehat{\alpha}}(\cdot), \widehat{\mu})) \le \sup_{\Gamma(\cdot) \in \mathcal{C}_0} \mathbf{E}U(\tilde{X}(T, \Gamma(\cdot), \mu)) \quad \forall \mu. \tag{7.9}$$

By (7.8) and (7.9), the pair $(\widehat{\mu}(\cdot), \Gamma_{\widehat{\alpha}}(\cdot))$ solves the problem

$$\text{Minimize} \quad \sup_{\Gamma(\cdot) \in \mathcal{C}_0} \mathbf{E}U(\tilde{X}(T, \Gamma(\cdot), \mu)) \quad \text{over} \quad \mu. \tag{7.10}$$

By Theorem 7.1 it follows that the pair $(\widehat{\mu}(\cdot), \widehat{\Gamma}_{\widehat{\alpha}}(\cdot))$ is a saddle point for the problem. This completes the proof of Theorem 1. □