Why Bankers Should Learn Convex Analysis

Jim Zhu
Western Michigan University
Kalamazoo, Michigan, USA

Part 1: Stochastic market model

March 3, 2011
A tale of two financial economists

Edward O. Thorp and Myron Scholes

Influential works:

- “Beat the Dealer (1962)” and “Beat the Market (1967)”
- The Black-Scholes formula (1973).
A tale of two financial economists

Edward O. Thorp and Myron Scholes

- Influential works:
  - “Beat the Dealer (1962)” and “Beat the Market (1967)”
  - The Black-Scholes formula (1973).

- Place the ideas were conceived: MIT
Edward O. Thorp and Myron Scholes

- Influential works:
  - “Beat the Dealer (1962)” and “Beat the Market (1967)”
  - The Black-Scholes formula (1973).

- Place the ideas were conceived: MIT

- Investment practice:
  - Managing partner of Princeton/Newport Partners and the President of Edward O. Thorp & Associates. Annualized return of 20% over 28.5 years
  - Partner of Long Term Capital Management: essentially bankrupted in less than two years and almost causing a crisis.
A tale of two financial economists

Edward O. Thorp and Myron Scholes

Influential works:
- “Beat the Dealer(1962)” and “Beat the Market(1967)”
- The Black-Scholes formula(1973).

Place the ideas were conceived: MIT

Investment practice:
- Managing partner of Princeton/Newport Partners and the President of Edward O. Thorp & Associates. Annualized return of 20% over 28.5 years
- Partner of Long Term Capital Management: essentially bankrupted in less than two years and almost causing a crisis.
A tale of two financial economists

Edward O. Thorp and Myron Scholes

Recognition:

- One of the story in Poundstone’s 2005 book “Fortune’s formula: The untold story....”
- Noble Price in Econ 1997 and main stream financial economics.

What is going on?
Bankers don’t know convex analysis!
A tale of two financial economists

Edward O. Thorp and Myron Scholes

Recognition:

- One of the story in Poundstone’s 2005 book “Fortune’s formula: The untold story....”
- Noble Price in Econ 1997 and main stream financial economics.

What is going on?
Bankers don’t know convex analysis!
Summary

- **Convex analysis** used to play a central role in economics and finance via (concave) utility functions.
- A ‘new’ paradigm was emerged since the 1970’s after Black-Scholes introduced the replicating portfolio pricing method for option pricing, and
Summary

- **Convex analysis** used to play a central role in economics and finance via (concave) utility functions.

- A ‘new’ paradigm was emerged since the 1970’s after Black-Scholes introduced the replicating portfolio pricing method for option pricing, and

- Cox and Ross developed the risk neutral measure pricing formula.
Summary

- **Convex analysis** used to play a central role in economics and finance via (concave) utility functions.
- A ‘new’ paradigm was emerged since the 1970’s after Black-Scholes introduced the replicating portfolio pricing method for option pricing, and
- **Cox and Ross** developed the risk neutral measure pricing formula.
- This new paradigm marginalized many time tested empirical rules.
Summary

- **Convex analysis** used to play a central role in economics and finance via (concave) utility functions.
- A ‘new’ paradigm was emerged since the 1970’s after Black-Scholes introduced the replicating portfolio pricing method for option pricing, and
- Cox and Ross developed the risk neutral measure pricing formula.
- This new paradigm marginalized many time tested empirical rules.
- It brought about unprecedented prosperity in financial industry yet also led to the 2008 crisis.
Convex analysis used to play a central role in economics and finance via (concave) utility functions.

A ‘new’ paradigm was emerged since the 1970’s after Black-Scholes introduced the replicating portfolio pricing method for option pricing, and Cox and Ross developed the risk neutral measure pricing formula.

This new paradigm marginalized many time tested empirical rules.

It brought about unprecedented prosperity in financial industry yet also led to the 2008 crisis.
We will show that the ‘new paradigm’ is a special case of the traditional utility maximization and its dual.

Overlooking sensitivity analysis in the ‘new paradigm’ is one of the main problem.
Summary

- We will show that the ‘new paradigm’ is a special case of the traditional utility maximization and its dual.
- Overlooking sensitivity analysis in the ‘new paradigm’ is one of the main problem.
- The recent financial crisis is a wake up call that it is time again for bankers to learn convex analysis.
Summary

- We will show that the ‘new paradigm’ is a special case of the traditional utility maximization and its dual.
- Overlooking sensitivity analysis in the ‘new paradigm’ is one of the main problems.
- The recent financial crisis is a wake up call that it is time again for bankers to learn convex analysis.
The talk is divided into two parts. In the first part we discuss

- A discrete model for financial markets.
- Arbitrage and martingale (risk neutral) measure.
- Fundamental theorem of asset pricing.
- Utility functions and risk measures.
- Markowitz portfolio theory
The second part focuses on the financial derivatives.

- The new paradigm of financial derivative pricing.
- A Convex Analysis Perspective.
- Sensitivity and Financial Crisis.
- Alternative methods and an illustrative example using real historical market data.
Uncertainty is ubiquitous in the financial world

- Stock price is unpredictable.
- Financial derivatives can bring about prosperity and disaster.
- Bond is considered safe but that is when interest rate is stable.
- Cash is better if only there is no inflation.
- To model financial markets one has to model uncertainty.
Uncertainty is ubiquitous in the financial world

- Stock price is unpredictable.
- Financial derivatives can bring about prosperity and disaster.
- Bond is considered safe but that is when interest rate is stable.
- Cash is better if only there is no inflation.
- To model financial markets one has to model uncertainty.
Model Uncertainty

- For problem involving only one decision such as analyzing a portfolio we need random variables.
- For problem involving multiple decisions such as trading we need stochastic process.
Model Uncertainty

- For problem involving only one decision such as analyzing a portfolio we need random variables.
- For problem involving multiple decisions such as trading we need stochastic process.
- The process of information becomes available also need to be modeled.
For problem involving only one decision such as analyzing a portfolio we need random variables.

For problem involving multiple decisions such as trading we need stochastic process

The process of information becomes available also need to be modeled.
The game of tossing a coin

- Bet on flipping a fair coin.
  - **Head**: the house will double your bet.
  - **Tail**: you lose your bet to the house.
A random variable

- Suppose we play the game only once and bet 1.
- Denote the outcome of the game by $X$.
- Then $X$ is a random variable taking only 1 or $-1$ as its value and $P(X = 1) = P(X = -1) = 1/2$. 
A discrete stochastic process

- Play the game $i$ times and always bet 1.
- Denote the outcome of the $i$th game by $X_i$. 
A discrete stochastic process

- Play the game $i$ times and always bet 1.
- Denote the outcome of the $i$th game by $X_i$.
- Then $X_i$ is a random variable and $P(X_i = 1) = P(X_i = -1) = 1/2$. 
A discrete stochastic process

- Play the game $i$ times and always bet 1.
- Denote the outcome of the $i$th game by $X_i$.
- Then $X_i$ is a random variable and
  \[ P(X_i = 1) = P(X_i = -1) = 1/2. \]
- If we start with an initial endowment of $w_0$ then our total wealth after the $i$th game is
  \[ w_i = w_0 + X_1 + \ldots + X_i. \]  \(\text{(1)}\)
A discrete stochastic process

- Play the game \( i \) times and always bet 1.
- Denote the outcome of the \( i \)th game by \( X_i \).
- Then \( X_i \) is a random variable and \( P(X_i = 1) = P(X_i = -1) = 1/2 \).
- If we start with an initial endowment of \( w_0 \) then our total wealth after the \( i \)th game is

\[
    w_i = w_0 + X_1 + \ldots + X_i.
\]

(1)

- Now \( (w_i)_{i=1}^{n} \) is an example of a discrete stochastic process.
A discrete stochastic process

- Play the game \( i \) times and always bet 1.
- Denote the outcome of the \( i \)th game by \( X_i \).
- Then \( X_i \) is a random variable and 
  \[ P(X_i = 1) = P(X_i = -1) = 1/2. \]
- If we start with an initial endowment of \( w_0 \) then our total wealth after the \( i \)th game is 
  \[ w_i = w_0 + X_1 + \ldots + X_i. \] (1)

- Now \((w_i)_{i=1}^n\) is an example of a discrete stochastic process.
• Will knowing $X_1, \ldots, X_i$, help us to play the $(i + 1)$th game?

• The answer should be NO but how do we clearly describe this conclusion?
Will knowing \(X_1, \ldots, X_i\), help us to play the \((i + 1)\)th game?

The answer should be NO but how do we clearly describe this conclusion?

Let us look at the game with \(n = 3\) to get some feeling.

We use \(H\) to represent a head and \(T\), tail.

The information we can get at each stage can be illustrated with the following binary tree.
Will knowing $X_1, \ldots, X_i$, help us to play the $(i + 1)$th game?

The answer should be NO but how do we clearly describe this conclusion?

Let us look at the game with $n = 3$ to get some feeling.

We use $H$ to represent a head and $T$, tail.

The information we can get at each stage can be illustrated with the following binary tree.
A tale of two financial economists

Summary

Model Financial Market
A discrete model for the financial markets
No Arbitrage and Martingale
Preference
Example: Markowitz Portfolio Theory
A Question

Uncertainty
Model Uncertainty
An example
A random variable
A discrete stochastic process
Information
Example of Filtration
Definition of filtration

Part 1: Stochastic market model

Why Bankers Should Learn Convex Analysis
Filtration for 3 coin tosses

- All the information are represented by $F_3 = 2^\Omega, \Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}.$

- Similarly, after 2 tosses $F_2 = 2^{\{HH,H,T,TH,TT\}},$ where $\{HH, H, T, TH, TT\} = \{\{HHH, HHT\}, \{HTH, HTT\}, \{THH, THT\}, \{TTH, TTT\}\}.$

- $F_2$ has less information than $F_3.$

- Similarly, $F_1 = 2^{\{H,T\}},$ where $\{H, T\} = \{\{HHH, HHT, HTH, HTT\}, \{THH, THT, TTH, TTT\}\}.$

- $F_0 = \{\emptyset, \{\Omega\}\}.$
The sequence

$$\mathcal{F} : \mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}_3$$

is a filtration for $$(w_i)_{i=0}^3$$.

For each $i$, $\mathcal{F}_i$ is a set algebra, i.e., its elements as sets are closed under union, intersection and compliment.
General filtration

Let $\Omega$ be a sample space (representing possible states of a chance event).

A sequence of algebra ($\sigma$-algebra when $\Omega$ is infinite)
$\mathcal{F} : \mathcal{F}_i, i = 0, 1, \ldots, n$ satisfying

$$\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2 \subset \ldots \subset \mathcal{F}_n$$

is called a filtration.

If $\mathcal{F}_0 = \{\Omega\}$ and $\mathcal{F}_n = \Omega$ then $\mathcal{F}$ is called an information structure.
All possible economic states is represented by a finite set $\Omega$. Probability of each state is described by a probability measure $P$ on $2^{\Omega}$.

Let $RV(\Omega)$ be the space of all random variables on $\Omega$, with inner product

$$\langle \xi, \eta \rangle = \mathbf{E}[\xi \eta] = \int_{\Omega} \xi \eta dP = \sum_{\omega \in \Omega} \xi(\omega)\eta(\omega)P(\omega).$$

$0 < \xi \in RV(\Omega)$ means $\xi(\omega) \geq 0$ for all $\omega \in \Omega$ and at least one of the inequality is strict.
Information system

- Suppose that actions can only take place at \( t = 0, 1, 2, \ldots \).
- Use \( \mathcal{F} = \{\mathcal{F}_t \mid t = 0, 1, \ldots\} \) to represent an information system of subsets of \( \Omega \), that is,

\[
\sigma(\{\Omega\}) = \mathcal{F}_0 \subset \mathcal{F}_1 \subset \ldots \subset \mathcal{F}_t \subset \ldots \text{ and } \bigcup_{t=0}^{\infty} \mathcal{F}_t = \sigma(\Omega).
\]

- Here, algebra \( \mathcal{F}_t \), represents available information at time \( t \).
- Implied in the definition is that we never loss any information and our knowledge increases with time \( t \).
- If action is finite \( t = 0, 1, \ldots, T \), we assume \( \mathcal{F}_T = \sigma(\Omega) \).
- The triple \((\Omega, \mathcal{F}, \mathcal{P})\) models the gradually available information.
Let \( \mathcal{A} = \{a^0, a^1, \ldots, a^M\} \) be \( M + 1 \) assets.

\( a^0 \) is reserved for the risk free assets.

The prices of these assets are represented by vector stochastic process \( S := \{S_t\}_{t=0,1,\ldots} \),

where \( S_t := (S^0_t, S^1_t, \ldots, S^M_t) \) is the discounted price vector of the \( M + 1 \) assets at time \( t \).

Using the discounted price, we have \( S^0_t = 1 \) for all \( t \).

Assume \( S_t \) is \( \mathcal{F}_t \)-measurable, i.e. determined up to the available information.

We say such an \( S \) is \( \mathcal{F} \)-adapted.

A described above is a financial market model.
A portfolio $\Theta_t$ on the time interval $[t - 1, t)$ is a $\mathcal{F}_{t-1}$ measurable random vector $\Theta_t = (\Theta^0_t, \Theta^1_t, \ldots, \Theta^M_t)$ where $\Theta^m_t$ indicates the weight of asset $a^m$ in the portfolio.

A portfolio $\Theta_t$ is always purchased at $t - 1$ and liquidated at $t$. The acquisition price is $\Theta_t \cdot S_{t-1}$ and the liquidation price is $\Theta_t \cdot S_t$. 
A trading strategy is a $\mathcal{F}$-predictable process of portfolios $\Theta = (\Theta_1, \Theta_2, \ldots)$, where $\Theta_t$ denotes the portfolio in the time interval $[t-1,t)$. A trading strategy is self-financing if at any $t$

$$\Theta_t \cdot S_t = \Theta_{t+1} \cdot S_t, \forall t = 1, 2, \ldots.$$ 

We use $T(A)$ to denote all the self-financing trading strategies for market $A$.

$\Theta = (\Theta_1, \Theta_2, \ldots)$ is $\mathcal{F}$-predictable means that $\Theta_t$ is $\mathcal{F}_{t-1}$ measurable.
Gain

For a trading strategy $\Theta$, the initial wealth is

$$w_0 = \Theta_1 \cdot S_0. \quad (3)$$

The net wealth at time $t = T$ is

$$w_T = \sum_{t=1}^{T} \Theta_t \cdot (S_t - S_{t-1}) + w_0. \quad (4)$$

Random variable

$$G_T(\Theta) = \sum_{t=1}^{T} \Theta_t \cdot (S_t - S_{t-1}) = w_T - w_0 \quad (5)$$

is the net gain.
A self-financing trading strategy is called an arbitrage if $G_t(\Theta) \geq 0$ for all $t$ and at least one of them is strictly positive.

Intuitively, an arbitrage trading strategy is a risk free way of making money. We note that we may always assume $G_T(\Theta) > 0$. 
No Arbitrage Principle

There is no arbitrage in a competitive financial market.
Fair game and martingale

- Toss a fair coin is a fair game in the sense that no player has an advantage.
- In other words, restricted to information at \((i - 1)\)th game, the expectation of \(w_i\) and \(w_{i-1}\) are the same.
- Mathematically,

\[
\mathbb{E}^P[w_i \mid \mathcal{F}_{i-1}] = w_{i-1}.
\] (6)

- A \(\mathcal{F}\)-adapted stochastic process satisfying (6) is called a \(\mathcal{F}\)-martingale.
- We will omit \(P\) and/or \(\mathcal{F}\) if it is clear in the context.
Examples

1. Let $X_i$ be independent with $E[X_i] = 0$ for all $i$. Then, $S_0 = 0$, $S_i = X_1 + \ldots + X_i$ defines a martingale.

2. Let $X_i$ be independent with $E[X_i] = 0$ and $\text{Var}[X_i] = \sigma^2$ for all $i$. Then, $M_0 = 0$, $M_i = S_i^2 - i\sigma^2$ gives a martingale.

3. Let $X_i$ be independent random variables with $E[X_i] = 1$ for all $i$. Then, $M_0 = 0$,

$$M_i = X_1 \times \ldots \times X_i$$

gives a martingale with respect to $\mathcal{F}_i$. 
Martingale for a financial market

1. Let $\mathcal{A}$ be a financial market.

2. We say that a probability measure $P$ is a martingale of $\mathcal{A}$ if $P(\omega) > 0$ for all $\omega \in \Omega$ and all the price process $S_t^m, m = 0, 1, \ldots, M$ are martingales with respect to $P$.

3. We use $\mathcal{M}(\mathcal{A})$ to denote the set of all martingale measures of $\mathcal{A}$. 
No Arbitrage and Martingale

Let $\mathcal{A}$ be a financial market model with finite period $T$. Then the following are equivalent

(i) there are no arbitrage trading strategies;
(ii) $\mathcal{M}(\mathcal{A}) \neq \emptyset$. 
Proof \((ii) \rightarrow (i)\)

Let \( Q \in \mathcal{M}(\mathcal{A}) \). If \( \Theta \in \mathcal{T}(\mathcal{A}) \) is an arbitrage, then, for some \( t \), \( G_t(\Theta) > 0 \) and consequently \( E^Q(G_t(\Theta)) > 0 \), a contradiction.
Proof (i)→(ii)

Observe that, $G_T(\mathcal{T}(\mathcal{A})) \cap \text{int}RV(\Omega)_+ = \emptyset$. Since $G_T(\mathcal{T}(\mathcal{A}))$ is a subspace, by the convex set separation theorem $G_T(\mathcal{T}(\mathcal{A}))^\perp$ contains a vector $q$ with all components positive. We can scale $q$ to a probability measure $Q$. Then it is easy to check $Q \in \mathcal{M}(\mathcal{A})$. 
Remark

(i) No arbitrage principle does not say one cannot make more than the risk free rate.

(ii) It says to do that one has to take risk.

(iii) Martingale probability measure is not the same as the real probability of economic events.
To discuss beating the risk free rate by taking risks we need measures for risk and reward. The preference of different market participants are different. Common way of modeling the preference are

(i) Utility functions;
(ii) Risk measures; and
(iii) The combination of the two.
Utility functions

- Experience tells us that mathematical expectation is often not what people use to compare payoffs with uncertainty.
- Lottery and insurance are typical examples.
- Economists explain this using utility functions: people are usually comparing the expected utility.
- Utility function is increasing reflecting the more the better and
- Concave: the marginal utility decreases as the quantity increases.
- Concavity is also interpreted as the tendency of risk aversion: the more we have the less we are willing to risk.
Examples of Utility functions

Several frequently used utility functions are

- Log utility \( u(x) = \ln(x) \) goes back to Bernoulli and the St. Petersburg wager problem.
- Power utility functions \( (x^{1-\gamma} - 1)/(1 - \gamma), \gamma > 0 \) and
- Exponential utility functions \( -e^{-\alpha x}, \alpha > 0 \).
- We note that \( \ln(x) = \lim_{\gamma \to 1} (x^{1-\gamma} - 1)/(1 - \gamma) \).
The following is a collection of conditions that are often imposed in financial models:

(\textit{u1}) (Risk aversion) $u$ is strictly concave,

(\textit{u2}) (Profit seeking) $u$ is strictly increasing and
\[
\lim_{t \to +\infty} u(t) = +\infty,
\]

(\textit{u3}) (Bankruptcy forbidden) For any $t < 0$, $u(t) = -\infty$ and
\[
\lim_{t \to 0^+} u(t) = -\infty,
\]

(\textit{u4}) (Standardized) $u(1) = 0$ and $u$ is differentiable at $t = 1$. 
Risk Measure

- An alternative to maximizing utility functions is to minimize risks.
- Pioneering work: Markowitz’s portfolio theory measures the risks using the variation.
- Modeling the risk control of market makers of exchanges, Artzner, Delbaen, Eber and Heath introduced the influential concept of coherent risk measure.
Common properties

Here are some common properties of risk measures:

(r1) (Convexity) for $X_1, X_2 \in RV(\Omega)$ and $\lambda \in [0, 1],$

$$\rho(\lambda X_1 + (1 - \lambda) X_2) \leq \lambda \rho(X_1) + (1 - \lambda) \rho(X_2),$$

diversification reduces the risk.

(r2) (Monotone) $X_1 - X_2 \in RV(\Omega)_+$ implies $\rho(X_1) \leq \rho(X_2).$ A dominate random variable has a smaller risk.

(r3) (Translation property) $\rho(Y + c\bar{1}) = \rho(Y) - c$ for any $Y \in RV(\Omega)$ and $c \in R,$ one may measure the risk by the minimum amount of additional capital to ensure not to bankrupt.

(r4) (Standardized) $\rho(0) = 0.$
Convexity and diversity

- Convexity is essential in characterizing the preference.
- Diverse in choosing particular preference is intrinsic.
Use \( \hat{S} \) and \( \hat{\Theta} \) to denote risky part of the price process and the portfolio. Giving the expected payoff \( r_0 \) and an initial wealth \( w_0 \), Markowitz’s problem is

\[
\begin{align*}
\text{minimize} & \quad \text{Var}(\hat{\Theta} \cdot \hat{S}_1) \\
\text{subject to} & \quad E[\hat{\Theta} \cdot \hat{S}_1] = r_0 \\
& \quad \hat{\Theta} \cdot \hat{S}_0 = w_0.
\end{align*}
\]

(7)
The portfolio problem is equivalent to the entropy maximization problem

\[
\begin{align*}
\text{minimize} & \quad f(x) := \frac{1}{2} x^\top \Sigma x \\
\text{subject to} & \quad Ax = b.
\end{align*}
\]

Here \( x = \hat{\Theta}^\top \), \( \Sigma = (E[(S_i^1 - E[S_1^i])(S_i^j - E[S_1^j])])_{i,j=1,...,M} \) and

\[
A = \begin{bmatrix} E[\hat{S}_1] \\ \hat{S}_0 \end{bmatrix}, \quad \text{and} \quad b = \begin{bmatrix} r_0 \\ w_0 \end{bmatrix}.
\]
Dual problem

Assuming $\Sigma$ positive definite the dual problem is

$$\text{maximize} \quad b^T y - \frac{1}{2} y^T A \Sigma^{-1} A^T y. \quad (9)$$
Solving the dual problem we derive the following relationship

$$\sigma(r_0, w_0) = \sqrt{\frac{\gamma r_0^2 - 2\beta r_0 w_0 + \alpha w_0^2}{\alpha \gamma - \beta^2}},$$

(10)

where $\alpha = \mathbb{E}[\hat{S}_1]\Sigma^{-1}\mathbb{E}[\hat{S}_1]^\top$, $\beta = \mathbb{E}[\hat{S}_1]\Sigma^{-1}\hat{S}_0^\top$ and $\gamma = \hat{S}_0\Sigma^{-1}\hat{S}_0^\top$. The corresponding minimum risk portfolio is

$$\Theta(r_0, w_0) = \frac{\mathbb{E}[\hat{S}_1](\gamma r_0 - \beta w_0) + \hat{S}_0(\alpha w_0 - \beta r_0)}{\alpha \gamma - \beta^2} \Sigma^{-1}$$

(11)
Markowitz bullet

Draw this function on the $\sigma\mu$-plan we get

\[ \text{which is commonly known as a Markowitz bullet for its shape.} \]
Markowitz portfolio theory became popular largely due to its simple linear-quadratic problem with explicit solutions. Well known extensions and applications include Capital Asset Pricing Model and Sharpe ratio for mutual fund performances. Are there other risk-utility function pairings that can lead to convenient explicit solutions?