Markowitz 投資理論と株式市場

青森公立大学　田中寛
tanaka@nebuta.ac.jp

1　はじめに

1993 年に経済学分野の学術雑誌の名前が変更された。それは、Computer Science in Economics and Management から Computational Economics にである。商業出版社系の学術雑誌のこの名称変更は、雑誌に投稿する研究者層の変化の直接の影響であり、さらにいえばその背景にある学問自体の変革によるものである。

経済学の流れのもっとも大きな変化は、経済学としての理論面でも実用面でも数理化が一層進展したことであり、その結果として情報科学との境界が学問的に消滅した。これは、経済学が情報科学に統合されて存在しなくなったという訳ではなく、それが対象とする社会現象に対する研究手法が自然科学などと同じようになったことを意味している。このことは、先述した雑誌の名称変化からも直接に読み取ることができる。

経済学のもう一つの変化は、Management 学の基礎としての役割が増大してきたことである。経済学の数理化に伴い、Management 学も数理化を強めている。この欧米の Management 学を日本では経営学と訳することが多いが、その両者が全くの別物であることが経済学の数理化によって非常に鮮明に示されるようになった。日本の経営学研究者の多くは、自身の学問を経済学のように数理化することを拒否しているように見える。欧米の Management 学が 1990 年代以降の各国において実用的に成功しているという評価が与えられているというのにである。

経済学全体のこの数理化という流れによって、教育にも様々な影響を及ぼすことは当然避けられない。今までは、大学で経済学を学ぶのは文系ということが当たり前であったのが、数学の勉強を抜きにしては経済学が学べなくなっているのである。高校教育では、文系といわれるクラスで微分積分を使いこなせるまで勉強することは稀である。その為、大学入学段階の数学レベルと大学卒業時の経済学に必要とされる数学を比較すると、大学在学中に学ぶことが要求される数学の内容と程度が分かる。しかしながら、この落差の大きさは、他の勉強もしなければならないことを考えると絶望的といっても良い。

この悲観的といっても良い状況を確認することが、この論文の目的である。例題として取り上げるのは、ポートフォリオを組む際に必要となる投資理論を用いて現実のデータから計算できることを示すことである。今や、ポートフォリオによって貯蓄を行うことが、国民の自己責任とまでいわれるのであるから、大学で経済学を学ぶものにとっては必須事項であるのは明らかである。

まず Markowitz がそれによってノーベル賞を受賞した投資理論の概要を 2 節において述べる。3 節では、具体的に数値計算を行う方法を述べる。 4 節では、本研究を通して明らかとなった経済学教育において必要となる事項について述べる。5 節では、本研究のまとめを述べる。

2　Markowitz 投資理論

ポートフォリオを組む投資対象が n 個あり、その全ての投資対象のデータとして収益の時間に対する変化が順に M 個あるものとする。その時、各投資対象についての共分散行列を Σ、ポートフォリオの投資比率のベクトルを $\mathbf{w}$ とすると、

$$\mathbf{w}'\ \Sigma\ \mathbf{w} \tag{1}$$

というスカラー量が、期待する収益率 y に対して最小になるように、ベクトル $\mathbf{w}$ を求めれば良いというのが Markowitz [1] が確立した投資理論である。ただし、

$$\sum_{i=1}^{n} w_i = 1 \tag{2}$$

$$\sum_{i=1}^{n} x_i w_i = y \tag{3}$$

という条件が課される。ここで、w_i は投資比率ベク

トルの各要素であり、x_i は各投資対象の平均収益率である。Markowitz は、シンプレックス法を用いた独自の方法を用いた解法によって、上記の問題を解いた。しかし、この解き方は、多次元の複雑な問題に適用することは困難である。

同じ問題をラグランジの未定係数法によっても行列形式を保ったまま解析的に解くことができる [2]。ある y に対する式 (1) の最小値は、

$$\frac{\gamma\ y^2 - 2\ \beta\ y + \alpha}{\alpha\,\gamma\ -\ \beta^2}$$

であり、その時に得られる $\mathbf{w}$ は、

$$\frac{\Sigma^{-1}\{\mathbf{x}(\gamma\ y - \beta) + \mathbf{1}(\alpha - \beta\ y)\}}{\alpha\,\gamma\ -\ \beta^2}$$

である。つまり、式 (1) で与えられるスカラー量は、y の 2 次式である。すなわち、期待される収益率 y に対する分散 (1) の間の関係には、最小のリスクとなる値

$$\frac{\Sigma^{-1}\mathbf{1}}{(\mathbf{1}'\Sigma^{-1}\mathbf{1})}$$

があり、その時の y の値は、

$$\beta\,/\,\gamma$$

である。ここで、$\mathbf{1}$ はすべの要素が 1 であるベクトルであり、α・β・γ という量は、それぞれ

$$\alpha\ = \mathbf{x}'\Sigma^{-1}\mathbf{x}$$

$$\beta\ = \mathbf{x}'\Sigma^{-1}\mathbf{1}$$

$$\gamma\ = \mathbf{1}'\Sigma^{-1}\mathbf{1}$$

で定義されるスカラー量である。これら全ては、Σ^{-1} を求められれば計算できる。しかし、大次元の数値計算では行列の逆行列を直接に計算することは現実的でない。

ラグランジの未定係数法の式と式 (2) 及び (3) で与えられる制約条件を、異なる見地から解釈することができる [3]。未定係数を λ_1 及び λ_2 として、それらをまとめて記すと

$$\begin{array}{lllllll} \Sigma\,\mathbf{w} & + & \mathbf{1}\,\lambda_1 & + & \mathbf{x}\,\lambda_2 & = & 0 \\ \mathbf{1}'\mathbf{w} & & & & & = & 1 \\ \mathbf{x}'\mathbf{w} & & & & & = & y \end{array} \qquad (4)$$

となる。共分散行列 Σ は既知であり、さらに、ベクトル x の各要素もデータから計算できる。また、右辺は y が与えられれば全て確定する。従って、n 個の要素であるベクトル w と λ_1 及び λ_2 を未知数とする $(n+2)$ 次の連立一次方程式

$$\mathbf{A}\ \mathbf{b} = \mathbf{c} \qquad (5)$$

として式 (4) を見ることができる。ここで、

$$A = \begin{pmatrix} \Sigma & \mathbf{1} & \mathbf{x} \\ \mathbf{1}' & 0 & 0 \\ \mathbf{x}' & 0 & 0 \end{pmatrix} \qquad (6)$$

$$\mathbf{b}' = (\mathbf{w}'\ \ \lambda_1\ \ \lambda_2)$$

$$\mathbf{c}' = (\mathbf{0}'\ \ 1\ \ y)$$

である。共分散行列は対称行列であるので、連立一次方程式 (5) の行列 A も式 (6) により対称行列である。我々は、この連立方程式 (5) を解くことによって問題を解く。

3 東証株式市場に対する適用

3.1 データの入手

東証株式市場のデータは、インターネット上の株価情報サイト [4] で入手することができる。このサイトの特徴は、1999 年 1 月以降に東京証券取引所で取り引きされた全ての株（東証 1 部、東証 2 部、店頭、東証マザーズ）の情報が、一日単位のプレインテキストの形式でダウンロードできることである。そのファイルに含まれる情報は、取引日、銘柄コード、銘柄名、始値、高値、安値、終値、出来高、市場区分である。データとしては、銘柄コードと終値を利用する。例えば、2004 年 6 月 14 日のものでは、3183 種のデータがある。これだけの数のデータに対する数値計算は、決して小さいものではない。

3.2 **C 言語によるプログラミング**

数値計算は基本的に C 言語でプログラムを作成した。データを用いて得られる連立一次方程式 (5) を解く為に、まず、Crout 法 [5] というアルゴリズムを用いて行列 A を LU 分解 [6] を行った。そして右辺にある期待収益率 y を変化させて、n 個の要素がある投資比率 $\mathbf{w}$ とラグランジの未定系数 λ_1 及び λ_2 を求めた。ところが、標準の C 言語で扱うことができるもっとも精度が高い浮動小数点数である double データ型で計算しても、式 (1) で計算される量は常に正になるべきであるものが、y の値に対してプラスになったりマイナスになったりして安定した結果を得ることができなかった。この計算結果は、標準 C 言語で扱う数値計算精度が、今の場合に不足しているせいであることを示しているものとと考えられる。そこで、多倍長計算を行うことを検討した。

3.3 多倍長計算

数値計算の精度を計算させる側で指定できる仕組みは、一般的に多倍長計算と呼ばれている。しかし、何か基準となるものの任意の倍だけが計算できるという制約がある訳ではなく、C 言語をはじめとする多くの数値計算ができる言語の標準として単精度と倍精度のデータ型の浮動小数点数があることの類推から、自由に精度を指定できることがこのように呼ばれるようになったものと推測される。多倍長計算ができる機能を提供しているものがいくつかある。ここでは、標準の C 言語及び C++言語に対してインクルードファイルとライブラリという形で多倍長計算ができるものを検討する。すなわち、プログラム中にインクルードファイルを指定し、コンパイル時のパラメータとしてライブラリを検索するだけで良いものである。なお、C 言語で書かれたプログラムは、ほとんどそのままで C++コンパイラでコンパイルすることができる。

最初に検討したものは、以前に筆者が Gauss 型積分公式の分点と重みを任意の精度で計算できることを示した [7] 際に利用した C++言語のソースプログラムが提供されたもの [8] である。この場合は、3.2 で利用したプログラムに多倍長機能のソースプログラムを付け加え、さらに、データ型を書き換えるだけによって、あとの演算や関数がオーバライドされるという C++の機能が容易に利用できる。しかしながら、Intel Xeon 2.8GHz という CPU を 2 個装着している PC で一週間連続して計算しても、式 (1) の量が一個も得られることがなかった。従って、このプログラムは計算スピードの点で採用できない。

次に、同じ C++言語のものでもスピードの点で定評のある cln [9] について調べた。このソフトウェアは、この次に述べる gmp を利用することによってさらにスピードアップされるといわれている。cln は、元々C 言語で利用する gmp を C++言語で外皮として利用しようとするものである。この gmp を利用できるようにすることは、ダウンロードしたソースプログラムをコンフィギャする時に指定する。その後メイクしてルートユーザでインストールすれば利用できるようになる。しかしながら、先のと同じ PC でも、さらに、2.0GHz の CPU を 2 個搭載している Mac においても、実行させせようとするとエラーが出てきて実行できない。従って、このソフトウェア cln は今のところ計算に利用できない。

最後に我々の計算にとって利用の可能性があるものとして残されているのは、今のところ gmp[10] である。gmp は、C++言語ではなく C 言語のライブラリーである。それ故、C++言語のように演算子のオーバライドができないという問題がある。すなわち、データ型の定義を替えるだけでは済まなくて、全ての関係する演算を gmp の関数で置き換える必要がある。これをプログラミングの問題として考えると、正確を期すことはかなり複雑で困難な問題である。その結果として、数値計算の最終結果を未だ得ていない。

4 経済学教育に必要となる事項

インターネット上からダウンロードしたデータから、期待収益率に対する分散曲線 (フロンティア) を求めるには、どのような作業が必要かを具体的に見た。それらの作業を行う人にとって必要であったことを、教育現場で訓練されなければならないこととしてまとめて述べる。

4.1 マイクロ経済学

現在の経済学全体の数理化の核心がマイクロ経済学にあることは、否定しがたいであろう。ところが、逆にマイクロ経済学が万能であることに対して、実証済みという点に関して懐疑的である必要があるのであろう。科学には万能ということはあり得ないことが、教育される必要がある。

4.2 投資理論

今の投資理論は、Markowitz 理論に基づくものに尽きる。ところが、共分散行列が与えられた後のことしか、教育現場では教えられていない。しかし、投資理論は元のデータから共分散行列を構成する過程を含めてもっと説得的に教えられなければならないのではないだろうか。少なくとも実用に耐えるものである必要がある。

4.3　統計学

数学の一分野としての統計学を、文系の学生に対して勉強させせなければならない。データを PC で処理した結果の説明ができる為には、数学としての厳密性を理解していなければならない。表面的なもっともらしさではなく、ことの本質が存在することを見させせないといけない。

4.4　線形代数学

統計学よりもっと基礎的な数学の分野であり、大学として数学の訓練として必須である。ある意味で、線形代数の勉強は、数学的な抽象思考の世界 (行列やベクトル) と現実世界 (データ) とをつなげる訓練として最適である。

4.5　アルゴリズム論

コンピュータでの処理に適しているやり方とそうでないやり方があることを理解しなければならない。特に、今の場合、連立一次方程式は、クラーメルの方法ではなく、Gauss 法を基礎にする LU 分解の方法で、解かなければならないことを教えられなければならない。

4.6　C 言語

コンピュータ上のアプリケーションソフトウェアを作ることができる C 言語は、コンピュータそのものを理解するうえでもキーでもある。情報社会の中で C 言語にこの両面があることを分かっていないと困ることが多い。また、新しいプログラム言語である C++言語や Java 言語の基礎でもある。

4.7　インターネットプログラミングリソース

情報社会が社会の一種である所以は、コミュニティの存在である。しかし、インターネットコミュニティは従来の概念ではなかなか理解できないものである。その中でも、オープンソースコミュニティはもっと特殊である。ところが、それがインターネット社会でのソフトウェア生産力の主力になろうとしている。そのほんの一端が、sln や gmp である。

4.8　インターネットデータ検索

情報社会ではデータという形の情報がインターネット上に溢れている。しかし、必要とする情報を探し当てようとすると、ポータルサイトという検索結果をまとめて表示してくれるサイトに頼らざるを得ない。我々の株価情報はグーグル [11] の検索から得られたものであることを明記しておく。

5　おわりに

経済学が顕著に数理化してきている一つの実例として、Markowitz 投資理論を数学的内容にまで立ち至って詳しく述べた。そして、それをコンピュータを用いて計算する為に行った、データの入手、採用したアルゴリズム、C 言語によるプログラミングの中身、そして、多倍長計算が必要になった事情を記した。さらに、それらの作業を行えるように学生が大学で学ぶべきことを列挙した。その結果、今の大学で学んでも 4 年間という時間内で到底到達できるものではないことが分かった。しかし、これが大学が社会から達成することを求められているのである。また、今世間で横行している大学改革と称するものが、この社会的要請に沿っているかのように見せかけているが、実は全くこれとは無縁であることは明らかである。では、我々大学で教育研究に携わっているものとして一体何ができるのであろうか。

謝辞

経済学についての様々の情報を教えて頂いた青森公立大学河野秀孝先生に感謝する。

参考文献

[1] H. M. Markowitz: Journal of Finance **7**, pp77-99(1952).

[2] J. H. Cochrane: ”Asset Pricing”, Princeton University Press, Priceton, PP83-86(2001).

[3] 田中寛: 青森公立大学経営経済学研究, **7**, no2, pp2-8(2002).

[4] http://www.bekkoame.ne.jp/ha/hahaha/.

[5] P. D. Crout: Trans. AIEE, **60**, pp1235-1240(1941).

[6] 戸川隼人: ”マトリックスの数値計算”, オーム社, p62(1971).

[7] 田中寛: 青森公立大学紀要, **8**, no2, pp2-11.

[8] http://hp.vector.co.jp/authors/VA000672/.

[9] http://http://packages.qa.debian.org/c/cln.html.

[10] http://www.swox.com/gmp/.

[11] http://www.google.co.jp/.