金融数学ビジュアライゼーションを 実現するためのデータベース構築

†学習院大学 経済学部経営学科〒171-8588 東京都豊島区目白 1-5-1‡千葉商科大学 〒272-8512 千葉県市川市国府台 1-3-1*学習院大学 計算機センター〒171-8588 東京都豊島区目白 1-5-1

E-mail: † yukari.shirota@gakushuin.ac.jp, ‡ takako@cuc.ac.jp, *kuboyama@tk.cc.gakushuin.ac.jp

あらまし 本研究は、債券数学における多様な知識をデータベースに蓄積し、電子教材作成支援、学習支援、および、ビジネスにおける意思決定支援を実現することを目的とする。債券数学の問題解決においては、解にいたる解法プランが複雑かつ代数学的であるため、直観的に理解しにくいという一般的課題がある。直観的理解の実現のため、我々は2種類のビジュアライゼーションを提案する。そのためのマルチメディア知識データベースの構築について論じる。

キーワード 債券数学,金融数学,ビジュアライゼーション,知識データベース

1. はじめに

本研究は,債券数学における多様な知識をデータベ ースに蓄積し,電子教材作成支援,学習支援,および, ビジネスにおける意思決定支援を実現することを目的 とする. 債券数学の問題解決においては、解にいたる 解法プランが複雑かつ代数学的であるため, 直観的に 理解しにくいという一般的課題がある. 直観的理解の 実現のため、我々は2種類のビジュアライゼーション を提案する. 1 番目は解法プランのビジュアライゼー ションであり,これは演繹推論過程,つまり問題解決 アルゴリズムを視覚化する. フローチャート, ビジュ アル・プログラミングなどでは、手続きが要素となる が,我々の解法プランでは,数学公式が要素となる.2 番目は, 具体的な数値を入れて, 解法プランを実行す るプロセスの視覚化で、世の中で、グラフを使って問 題を解く、と呼ばれる解の求め方である(図1に、制約 付き最適化問題を視覚的に解く様子を示した). これを 解法プラン実行過程のビジュアライゼーションと呼ぶ. この2種類のビジュアライゼーション機能を各種のシ ステム(学習支援システムなど)が有するようになれば, 解法プランおよび, その解法プラン実行過程が見るだ けで直観的に理解できる非常に使いやすいシステムと なる.

ビジュアライゼーション機能実現のためには、核となるマルチメディア知識ベースも、それに特化したものが必要となる。我々は、現在、<u>債券数学分野における学習支援システム</u>をターゲットとして、2種類のビジュアライゼーション機能実現のためのマルチメディア知識ベースの構築を研究している[1]. この知識データベースの用途は、教育分野だけに限定されるものではなく、広く、実際のビジネスの場面での意思決定過程のビジュアライゼーションにも利用可能な汎用性を

重視している.

知識データベースの要件を考察するために, 金融分 野を含む経済数学教育の事情について言及する. 経済 学部で数学を教えてきた経験から感じることは,多く の経済学部学生が経済数学を難しいと考えていること である. その理由として, (1)経済に関する知識及び, 数学に関する知識の不足,(2)知識からの演繹推論力の 不足, があげられる. 経済数学の場合, 数学の知識と, 経済に関する知識が複雑に交じり合うため、演繹推論 プロセスが複雑になるので, 問題解法のための演繹推 論が正しく行えないのである. 本研究で扱う推論の種 類は演繹推論のみであるので,以下,単に推論と言っ た場合は演繹推論のこととする. さらに、学生の問題 解決を困難にさせる理由としては,業界用語の不統一 性問題が存在する. 特に金融の分野で用語の不統一性 の問題は大きい. これに関しては, 我々は, 債券数学 の文章題の日本語処理において, 用語をユニークなセ マンティクス・ラベルに変換することを, 文章題内の 文脈だけを用いて行うことが可能であるか否か, 研究 中である. 経済数学教育のため, 白田は, 知識ベース に基づく推論エンジンシミュレーション法(以下,推論 エンジン法と略す)という数学教授法を創案し,実際の 講義においてこの教授法を実践してきた[2-4]. 推論エ ンジン法の特長は、学生に2種類の知識ベースを、頭 の中に構築させる点である. 経済の知識ベースには, 経済の公式を,数学の知識ベースには数学公式を覚え させる. 学生アンケートの結果からも, この教授法の 効果が高いことが示された[2,4].

我々は、推論エンジン法に基づき、学生を支援する 学習支援システムを構築したい、学生は、解にいたる までの、公式を要素とする推論プロセス(以下、解法プ ランと呼ぶ)を求めようと試みるが、それが困難な際に、 システムが正しく演繹推論するための知識を提示し、学生が演繹推論の最適なパスを得るまで支援する. 我々は、現在、「債券ポートフォリオの管理」分野を対象として研究している.

第2節では、この債券数学学習支援システム構築のための研究課題と、2010年度前半の活動で明らかになった問題点の報告を行う。第3節では、ビジュアライゼーションのためのマルチメディア知識データベースについて考察する。第4節では、公式データベースにおける知識の体系化について論じる。第5節では、学習支援システムにおける知識データベースの役割ついて論じる。最終章はまとめである。

2. 学習支援システム構築における課題

現在,我々は「債券ポートフォリオの管理」分野を対象とする債券数学学習支援システムをターゲットとして知識データベース構築の研究を行っている.2010年度前半は以下のような研究活動を行なった。その中で,明らかになった研究課題について論じてみる.

▶ コーパスの作成:

債券数学の「債券ポートフォリオの管理」の分野のための、コーパスの第 1 版を作成し、XML で記述した(図 2 参照). ユニークなセマンティクスのラベルは、約 40 個となった. 現在、ラベル属性として、デフォルト値、変数に付随する可能性のある単位(年、四半期、月など)、ベースとなる単位、ベースとなる単位への変換規則、などを用意している. 単位変換とは、例えば、システム内部では、金利は小数点で表しているが、文章題で%表示された場合、内部表示に変換するため、100 で割る必要がある. 今後、実際に日本語処理を進めながら、属性の拡充を行う必要がある.

▶ 文章題のタグ付け:

自然言語処理の分野において、日本語文書へのタグ付けの研究は広く行われているが、我々のように数学文章題へのタグ付けを試みている研究事例は、我々の知る限り、この研究が初めてである。図3に、文章題および解法プランのXML例を示した。この<word>タグの部分が、文章題のタグ付け部分である。日本語処理機能を用いてこのようなタグ付けを行うことで、その後、言語処理部が文章題からGiven Data およびUnknown を抽出することが容易になる。この部分についての研究はこれから、開始する.

タグ付けの利点について論じる. 文章題のタグ付けが実現できたならば、パラメータ変更により、違う文章題を作成することが容易に行えるようになる. これは、類似文章題の自動生成に役立つ. また、ここで追記されたタグ情報は、解法プラン実行過程のビジュアライゼーションにおいても利点がある. 例えば、軸のタイトル付け("利回り"、"債券価格"など)や、GUI で

のスライダーによる、パラメータの対話的変動を考えた場合、スケールの条件範囲指定などにおいて、コーパスで定義されたラベルの意味情報(例えば、金利は 0 から 1 の間である)が利用できる. また、コーパスには、"半年複利"のような業界専門用語も定義しておく必要がある. 我々のユニーク・セマンティク・ラベルでは、例えば、"半年複利"は <var label="Times" value=2 unit="year">半年複利</var> と表される.

➤ Given Data の解釈:

言語処理部の役割として、Given Data および Unknown の抽出がある. 抽出された Given Data の解釈は、文脈依存である. 例えば、"デュレーション"という用語が出てきた場合、多くは、"修正デュレーション"を指すが、ポートフォリオのイミュニゼーションの場合だけは"マコーレー・デュレーション"を使う、というような慣例がある. 抽出した変数を、ユニークなセマンティクのラベルに変換可能か否かは、今後、研究していく.

▶ 解法プランの定義:

現在,代表的文章題に対して,解法プランを XML で定義する作業を進めている(図3参照).現仕様では、 解法プランの XML は自動生成せず, 数学教員が書く ことを仮定している.理由は、推論をシステムに任せ るためには, 文章題の記述能力を限定する必要がある ためである.これは避けるべきである、と我々は考え た. 推論自動化による手間の削減が図れたとしても, 作りたい教材が作れないのでは意味がない. 例えば, 公式データベースに無い関係式を, その文章題内でロ ーカルに定義したい場合がある. そのような一時的な 追加式を入れるためには、人手による XML 記述のほ うが,実現が容易である.他方,「良い教材を作りたい. そのためには、教材の記述能力(表現能力)を重視する. つまり, 少数の良い教材のほうが, 多数の自動生成さ れた教材よりも教育効果が高い」という考えがある*. こうした議論の結果、現仕様として人手による解法プ ラン記述を選択した. XML を書く数学教員を大いに支 援するのは、公式および連立方程式への Given Data のチェックである.推論過程において,現在,与えら れている変数値だけで, 方程式が解けるか否かのチェ ック, つまり, 与えられた変数の過不足チェックをシ ステムが代行してくれる. これは非常に数学教員の手 間を削減する.

▶ 意味モデルの作成:

解法プランを求めるためのメソッドとして,我々は債券数学用意味モデルを作成した.これは,債券数学に係る公式をエンティティとした,

^{*}日本経営数学会,2010年度秋季大会における議論.2010年10月 23日.神田,専修大学.

Entity-Relationship ダイアグラムである[5](図 4 参照). 公式間の関係の意味をモデル化したものである. 各公式には、使われる変数群が定義されている. n 個の変数のうち、(n-1)個の値が確定すれば、残る 1 個の変数の実数解は一意に決まる. 債券数学では、虚数解は使わない. どの変数を Given Data とし、どの変数を Unknown とするかは、その使われるオカレンスに依存するので、公式としては、単に内包する変数集合を持っていればよい. この E-R ダイアグラムは、学生の推論を支援するツールである. ダイアグラムな見ながら、文章題の Given Data である変数をチェックしていくと、使うべき公式が発見できる. このツールを使うことで、公式を組み合わせる演繹推論ができない学生も、機械的作業で、解法プランが作れるようになる. 非常に学習効果の高いツールである[5].

フリー数式処理ソフトウェアの選択: 推論エンジ ン部内には, 実際の数式処理計算を行う機能と, 解法 プラン実行過程のビジュアライゼーションのための数 式描画機能が必要である. 我々は, 本システムの無料 配布を計画しているので,数式処理システムなどは, フリーソフトウェアであることが必要条件であった. そして数あるフリーソフトウェアの中から MAXIMA† を選択した. これは数式処理ソフトウェアを使う日本 の高校の数学教員の中で、MAXIMA が普及している ことも大きな要因であった.一般に製品の数式処理シ ステムのほうが, グラフィクス描画機能が高いが, MAXIMA の機能でもプログラムを追加することで、 我々の要件を満たすことが確認できた. 実際に方程式 を解く作業も,数式処理システムの役目である. MAXIMA は機能が豊富で、債券数学の数学に関して は、 殆どの連立方程式が問題なく解けることを確認し た. 債券ポートフォリオの管理の分野では、微分方程 式は出てこないため、微分方程式は現在の我々の研究 対象には含めていない. 例外としては, あまりに高次 な方程式(100 次以上)は、解けない場合がある. 我々 は、既に、経済数学に関するグラフィクス教材サイト Graphics10 を公開し、MAXIMA のプログラム付きで 教材ポワーポイントを公開している[6,7]. 内容は,代 表的な債券数学の問題,ポートフォリオ,債券の価格・ 利回りなどに関する 10 種類の 3 次元グラフィクスで ある. グラフィクス描画については, XML 化は未着手 の状態で、これは完全に人手で MAXIMA プログラム を書いたものである. グラフィクスの XML 化につい ては,次節で考察する.

3. ビジュアライゼーションの重要性

我々は、債券数学学習支援システムにおいて、2 種類

† Maxima, a Computer Algebra System, http://maxima.sourceforge.net/.

のビジュアライゼーションを実現したい. 1 番目は解法プランのビジュアライゼーションであり、これは演繹推論過程、つまり問題解決アルゴリズムを視覚化する. 2 番目は、解法プラン実行過程のビジュアライゼーションであり、具体的な数値を入れて、解法プランを実行するプロセスを視覚化するものである.

3.1 解法プランのビジュアライゼーション

我々は、学習支援システムにおいて、解法プランを視 覚化したい. 図 5 に、解法プランのビジュアライゼー ションの例を示す. XML で記述された解法プランをも とに, 使う公式と, そこに与えられる Given Data が, 有向グラフで表示される. このグラフを解法プラング ラフと呼ぶ. 解法プランの各ステップは, 四角の枠で 示される. 各ステップには, 公式(一般には, 公式から なる連立方程式)が存在し、そこに Given Data が与えら れる. 各ステップには, Unknown が指定され, 数式処 理システムが連立方程式を解き、その Unknown を計算 する. その Unknown は, 以降のステップの入力として 使われる. 最後のステップの場合, Unknown は, 問題 全体の Unknown, つまり, 解となる. 解法プラン XML から解法プラングラフを作成するフェーズは, 自動化 可能であると考えている. 解法プランの XML を解析 する際,システムは、入力変数の過不足チェック,連 立方程式の式の過不足チェックを行うようにしたい. 上記の要求仕様を満たすためには、数学公式データベ ースには,以下の属性が必要である.

- ▶ 内部計算用 MAXIMA としての数式記述
- 解法プラングラフでの表示用数式記述(例えば、TeX など)
- 3.2 解法プラン実行過程のビジュアライゼーション 解法プランのビジュアライゼーションが、XML から自動生成できる可能性がある一方、解法プラン実行過程 のビジュアライゼーションの自動化には、なお一層の研究を要する. その要因は大きく分けて 2 つある.
 - (1) グラフィクス描画機能が成熟していない.
 - (2) グラフィクス教材は、教員の創意工夫が重要で現段階では機械任せにできない.

まず,単純な技術問題である(1)から解説する. 例えば, ある数式処理システムでは, 曲面を複数重ねて表示することができない. 偏微分をグラフィクスで示すためには, 平面で, 曲面を切断し, インターセクション部を見せたい場合がよくある. 図 6 は Maple で描かせた, 債券価格の利回り曲面である[6]. 満期の違いによる, 利回り曲線の違いを比較する際には, このように, 平面による切断操作が必要である. フリーのソフトウェアでは, 現段階では, このようなインターセクションをとる操作が簡単に実現できない. 補助プログラムが必要となる.

次に、(2)の要件の解説をする. グラフィクス教材を作る上での、教員の創意工夫とは何かを説明したい. スケールの取り方をテーマに説明する. 同じ数式処理システムを使っても、数学の知識が無いと見やすいグラフィクスは描けない. 演習のクラスで、学生がきれいなグラフィクスを描けない多くの原因は、スケールの取り方である. 見たい部分を選択するには、解のあたりをつける必要があるが、それを考えようとしないからである. 例えば、解が x=100000 のにある場合、x を 1 から 10 までのスケールで描かせても、何も描画されない.

複数の式を描画させて、その連立方程式の解を、式の間で表示する、というのであれば、代数学的に解を求めて、そこを中心としてスケールを決めれば、自動化はできそうである。しかし、スケール自動化いい問題は、何を見たいかが、単純に決められないである。3次元グラフィクスが特長である数学概念を理解したい数学概念を理解したい数学概念を理解したい数学である。

数学教師がどのような選択を行っているのか, 現時 点で分かる範囲でまとめてみる.

- (1) 2次元グラフィクス/3次元グラフィクスのどちらに すべきか.
- (2) <u>軸の変数として何をとるべきか</u>. 扱う関係式は,多変数関数である. それを 3 次元以下の次元に落とすためには,変数を選択しなくてはならない. ユーザが注目している変数を残すことが必要である. 図 6 は,漸化式によって債券価格を時間を逆向きに計算する手法を,表したものであるが,価格,残存期間のほかに,3番目の変数として,スポットレートを入れて,スポットレートのイールドカーブと合わせて見ることができるようにした[6].
- (3)<u>軸のスケール範囲をどうとるか</u>. 適切なスケールを 決めることは、一般に難しい. 数学知識の無い学生 がグラフィクスを描くと、不適切なスケールのため、 見るべき部分が描かれないことが多い.
- (4)分かりやすさのために、制約条件をいかに重ね合わせて描画するか。制約条件は、曲面の場合、平面の場合、カーブ(直線含む)の場合、点の場合がある。平面での、曲面の切断操作もある(図6参照)。
- (5)<u>3次元平面から2次元平面への投影をいかに行うか</u>. 2次元平面での関数シフトなど,1変数増やした関

数式を用いて、3次元曲面を描き、その等高線を2次元平面に投影すると、関数シフトの理由がよく理解できる. 国民所得決定問題における関数シフトは、この手法で非常に理解しやすくなる[3].

(6)色や,線の太さなどの物理的な設定をいかにするか. 非常に物理的なことであるが、3次元の曲面の色付 けの問題は自動化が難しい. 見えないことには意味 がないので、教材を作る際に、非常に時間をかける 部分である. また白黒印刷で資料を配布した場合に, 見えるように配慮する必要も場合によってはある. この他, 金融数学では, 主成分分析, 重回帰分析な どの多変量解析を用いる.これらの特長は,今まで見 えなかったデータの特徴を, ある数学的処理をするこ とで、見えるようにすることである. 多変量解析の可 視化テクニックについては, それぞれの解析手法につ いて、研究がなされている[10]. これは研究分野とし て既に存在しており, 我々が研究をしようとしている 解法プラン実行過程のビジュアライゼーションとは目 的が異なる. 我々は、特定文章題の解法プランを理解 させることを目的とする.

段階では,金融工学の本や論文には,コンピュー タ・グラフィクスが少ないことが問題である. 描く人 が少ないがゆえに, グラフィクスのためのノウハウが 十分に蓄積されていない. 株価のデータの変動などの グラフィクスは,単なるデータ自体のグラフ化であり, 未知数を求めるためのビジュアライゼーションではな い. スポットレートの利回りカーブを描いただけでは, 「未知数を求めるためのビジュアライゼーション」と は言えない. 我々の言う解法プラン実行過程のビジュ アライゼーションとは、解を求めるプロセスを、分か りやすく可視化することである. 今まで代数学的解法 で理解できなかった学生が, そのグラフィクスを見る と、解法プランが理解できるようになるという、効果 のあるものでなくてはいけない. 解法プラン実行過程 のビジュアライゼーションに関して現時点の研究すべ きことは、多くの実例をもとに、共通するパターンを 発見することであろう. 現段階は、まだ、多くのグラ フィクスを描画し,それを蓄積するフェーズと考える. そして,上述したように,良いグラフィクス教材を作 成するためには、数学の知識が必須である. そのため には, 金融数学の知識データベースをいかに構築する かが必要要件となる. 次節では, 数学公式の知識デー タベースについて論じる.

4. 数学公式データベース

数学公式データベースをどのように構築するかを 論じるためには、今の数学教育がどのように行われる べきかを考える必要がある. 我々は、数学の講義を通 して、学生の推論能力を養いたい. 推論を柔軟に行う ためには,公式データベースは,厳選された最小限の 公式のみが格納されるべきと考える. 最小限の公式集 合をもとに、演繹推論を行うことで、どのような問題 も解ける、そのような論理的思考能力を養うことが、 数学教育において最も重要と、考える. 推論および論 理的思考の育成と対局に位置するものとして, 冗長な 公式の丸暗記があり、世の中の数学問題集を見ると、 公式から導かれる2次的な式をすべて公式として覚え させるような,極めて非効率的な数学指導がなされて いる場合が多いように感じられる. 例えば、SPI 試験 の「速さ・時間・距離」の分野で、速度、距離、時間 の3つの公式を暗記させるという教育のやり方である. 速度,距離,時間の3つの変数間の関係はひとつの公 式だけで理解でき, あとは式の変形として必要に応じ て求めればよい. 良心的な SPI の教科書では,「3 つの うちのどれか1つを覚えれば、他は導き出せます」と 注意がきをしてくれている.

債券数学分野での類似例としては, 1年あたりの複 利計算の回数の違いによって,債券価格,修正デュレ ーション,マコーレー・デュレーションを求める公式 をそれぞれ丸暗記させる、というものが考えられる ‡. 限られた時間の中で,反射的に答えを求めるためには, 丸暗記が必要, という反論が考えられるが, 体系化さ れていない知識がフラットに並んでいる状況では、学 生の論理的思考力は養われない, かつ, 試験が終了す ると忘れてしまう確率が高いのではなかろうか.また, そのような数学の学習方法は, 学生にとっても知的に おもしろいとは言えないであろう. 推奨すべき学習方 法は,根幹の公式だけを暗記し,その公式の中で定義 された変数間の関係を深く理解することで,他の派生 的関係も容易に導出可能とするものである. そのメッ セージを学生に伝えるためには,公式データベースは 最小限の集合でなければならない. 債券数学用に我々 が選択した公式は,19個である[7].

本節では、我々の構築しようとする学習支援システムを、データベースシステムとして見た場合、以下のデータベースから構成される:公式データベース、コーパス(セマンティック・ラベル・データベース)、専門用語データベース、解法プランデータベース、グラフィクス教材データベース、MAXIMA 教材データベース.

5. まとめ

本稿では、債券数学の問題解決における解法プランの直観的理解の実現のため、2 種類のビジュアライゼーションを提案した. 1 番目は解法プランのビジュアライゼーションであり、これは演繹推論過程、つまり問題解決アルゴリズムを視覚化する. 2番目は、具体

的な数値を入れて,解法プランを実行するプロセスの

謝辞

本研究の一部は、平成 22 年度科研費基礎研究(C)一般「推論エンジン法による知識ベースの構築」(課題番号: 22500231, 代表者: 白田由香利)による,ここに記して謝意を表します.

参考文献

- 1.白田由香利,橋本隆子,久保山哲二,金融工学の数学のための知識データベース構築の提案.情報処理学会研究報告,データベースシステム(DBS), Vol.2010-DBS-150 No.5, 2010.
- 2.白田由香利, *推論エンジンをベースとした経営数学* 教授法とそのアンケート評価. 学習院大学経済論集, 2010. **Vol.46**(No.3).
- 3.白田由香利, 悩める学生のための経済・経営数学入 門 3 つの解法テクニックで数学アレルギーを克 服!. 2009、東京: 共立出版.
- 4.白田由香利, 推論エンジンをベースとした経済数学 教授法. 情報処理学会研究報告, 情報学基礎(FI), 2009. Vol.2009-FI-95(No.10): p. 1-7.
- 5.橋本隆子, 久保山哲二, 白田由香利, *意味モデルによる債券数学の解法*. 日本経営数学会, 秋季大会, 2010.
- 6.白田由香利,橋本隆子,久保山哲二. MAXIMA による 経済数学グラフィクス 10 個 -Bond Mathematics by Graphics- (略称 Graphics10) 2010; Available from: http://www-cc.gakushuin.ac.jp/~20010570/private/MAX IMA/.
- 7.白田由香利,橋本隆子,久保山哲二,金融工学数学 のためのグラフィクス教材データベース --Bond Mathematics by Graphics-,学習院大学経済論集,2011: 印刷中.
- 8.Bittinger, M.L., Calculus and Its Applications(Eighth Edition). 2004: Pearson, Addisonn Wesley.
- 9. Strauss, M.J., G.L. Bradley, and K.J. Smith, *Calculus*. 3rd ed. 2002, New Jersey: Prentice-Hall, Inc.
- 10.佐川昌弘 , 廣岡康雄, マーケティング・データ解析, 木島正明, 中川慶一郎, 生田目崇(Ed). 2003, 朝倉書店.

視覚化で、<u>解法プラン実行過程のビジュアライゼーション</u>である. そのビジュアライゼーションのために必要な知識データベースについて述べた. 現在, これらのデータベースシステムを構築中である.

[‡] 実際にそのような教科書があるわけではない.

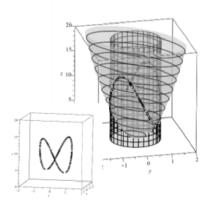


図 1: 解法プランのビジュアライゼ ーションの例 (制約付き最適化問題 の視覚的解法)

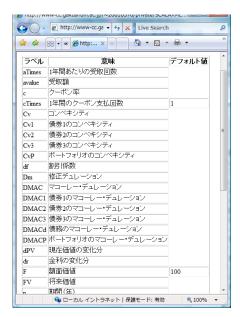


図 2:XML で記述したコーパス(部分)

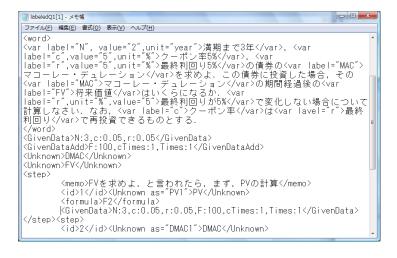


図 3:文章題および解法プランの例

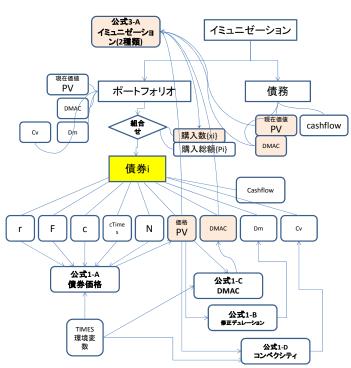


図 4:公式をエンティティとした E-R ダイアグラム

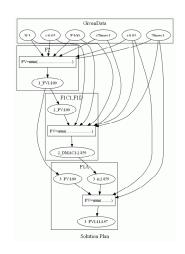


図 5:解法プラングラフ

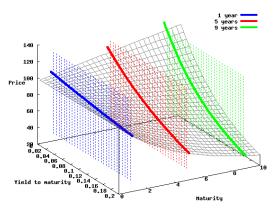


図 6: 価格―利回り曲面