Qualitative/Quantitative Financial Analysis

Juan Flores **Federico González** **Beatriz Flores**

Fac. Ing. Eléctrica **Fac. Contabilidad y Administración**

{juanf, fsantoyo}@zeus.umich.mx
Universidad Michoacana de San Nicolás de Hidalgo
Morelia, México

Abstract

We present a methodology to perform Financial Analysis with as much information as available. The problem we solve is the assessment of investment projects; we are to determine whether or not an investment project is profitable. We perform that task at the qualitative, semi-quantitative, and quantitative levels, depending on how much information there is at hand.
The main advantage of the method presented here is that it can deal with very little information about quantities, still yielding some results. The least piece of information the system can work with is a set of Order of Magnitude Relations (omrs) among the model variables. If those omrs are sufficient to disambiguate the results of our model, we can tell whether an investment project is worthwhile or not. Our system remembers what has been said, and takes any additional information to refine its results. If all we have is a few omrs it may or may not determine the outcome of our investment project analysis. Later on, we provide the system with some more information (perhaps imprecise), and the results will be refined. The more precise the provided information is, the more accurate the results will be. If at the end, all provided variables are precise, the results are the same as performed via traditional analysis.
This methodology takes ideas from the field of Qualitative Reasoning, some algorithms from graph theory and information structures, and the traditional interval computation arithmetic to perform operations under uncertainty.

Keywords: order of magnitude reasoning, approximate reasoning, qualitative reasoning, interval computation, financial analysis.

Introduction

In the area of Financial Analysis, people (see [González99]) often use the Benefit-Cost method to determine whether or not an investment project will be profitable. This financial indicator measures the relation of the benefit to the cost (brought to present value) in a given project. The analysis is performed in a given planning horizon.
The formula for the benefit-cost rate is shown in Equation (1).

$$\frac{B}{C} = \frac{\sum_{t=1}^{n} FF_t^{+} \Big/ (1+i)^t}{\sum_{t=1}^{n} FF_t^{-} \Big/ (1+i)^t} \qquad \textbf{(1)}$$

where:

FF_t^{+} = Positive Cash Flow
FF_t^{-} = Negative Cash Flow
i = Capital Cost
t = Planning Horizon (time); 0, 1, 2, …, n

The decision rules are:

$$B/C = \begin{cases} >1 & OK \\ =1 & Equilibrium \\ <1 & Loses \end{cases} \qquad (2)$$

To compute FF_t we will use the criteria established by González ([González85]). In his paper, González establishes that:

$$FF_t = NUL_t + D_t + A_t + RV_t - I_t \qquad (3)$$

where:

NUL ^{t} = Net utility or loss
D ^{t} = Depreciation
A ^{t} = Amortization
RV ^{t} = Rescue Value
I ^{t} = Investment

Under some approximate representation of quantities (e.g. fuzzy numbers or intervals – see [Gotz83], [Kosko96], and [Pedrycz98]), the evaluation of the discrimination criteria (Eq. 2) does not represent any problem. Nevertheless, dealing with omrs the propagation procedures does find it difficult (see section 4). So, instead of evaluating a ratio, we will use the equivalent subtraction shown in (4).

$$B-C \begin{cases} >0 & OK \\ =0 & Equilibrium \\ <0 & Loses \end{cases} \qquad (4)$$

The system we present in this paper is capable of performing Financial Analysis at various levels of granularity. Indeed, since its main representation scheme is intervals, the number of levels of granularity is infinite. At the coarsest level of abstraction, Order of Magnitude Reasoning (OMR) has to be performed to achieve useful results.
The rest of the paper is organized as follows. Section 2 shows how approximate representation of quantities can help us solve the problem under uncertainty. We use an alternative representation, intervals and contrast solutions. Section 3 introduces Order of Magnitude Reasoning, the propagation algorithms, proposes a novel representation for omrs, and contrasts the results of this representation with previous ones. Section 4 proposes an algorithm to evaluate the discrimination criteria using omrs. Section 5 shows how to combine qualitative and quantitative information to produce solutions whose quality improves with the arrival of new information. Section 6 concludes the article, highlighting pros and conses, and proposing future research directions.

Approximate Solutions

Throughout the paper, we present part of an example to illustrate the different concepts and issues that rise in each section. Consider a consortium that wants to start a new company. For that matter, a feasibility study is performed. From financial analysis, marketing research, etc., we estimated the cash flow for several subsequent months. The goal is to determine the B/C indicator. Table 1 shows the input data, expressed as fuzzy numbers (each triplet represents a Triangular Fuzzy Number). For each monthly period, the Cash Flow (FF), and interest rate (i) are shown.
For the previous example, [González99] has used a fuzzy representation of quantities, performing all operations in the fuzzy domain. We propose to use intervals instead. The reason is because intervals provide a uniform representation to deal with all types of information. omrs can be represented as intervals (see next section); sign values and uncertain information can be seen as intervals as well; real numbers can be seen as point intervals.
For the fuzzy approach, the three numbers represent the extreme and middle points of a triangular fuzzy number. For interval computations, we disregard the middle number and take the extremes to define the interval. For real numbers, we eliminate the extremes and consider the middle point of the fuzzy number as the real magnitude of the involved quantity.
The ratio B/C was evaluated using each of the described formalisms. The results are shown in Table 2.
The wider the fuzzy numbers and intervals, the more uncertainty there is; If we make those numbers wide enough, we will get to a point where the results would include zero. In those cases we could not answer the target question (i.e. the evaluation of B-C). Taking the situation to the extreme, it could be the case that we do not have any knowledge at all of the values of those quantities. We could still use some other kind of information to try to resolve the ambiguity, and provide results even in those cases. omrs are the keyword to our problem. The next section introduces omrs and Order of Magnitude Reasoning (OMR), and section 4 proposes an algorithm to use omrs in the determination of B-C.

Period t	**FF_t**	**i_t**
0	-25000	0
1	(12000, 15419, 16500)	(0.20, 0.25, 0.26)
2	(15000, 14891, 20000)	(0.25 0.26, 0.28)
3	(16000, 18335, 19500)	(0.29, 0.30, 0.31)
4	(16000, 18820, 19000)	(0.31, 0.32, 0.33)
5	(16000, 17389, 18000)	(0.36, 0.37, 0.38)
6	(16000, 18761, 19000)	(0.38, 0.39, 0.40)
7	(15500, 18759, 19500)	(0.38, 0.39, 0.40)
8	(18000, 18666, 20000)	(0.38, 0.39, 0.42)
9	(16000, 18640, 19000)	(0.40, 0.41, 0.42)
10	(16000, 18555, 19000)	(0.40, 0.41, 0.42)
11	(6000, 9980, 10000)	(0.42, 0.42, 0.42)

Table 1. Expected Cash Flows and Interest Rates for Example

	Real	**Fuzzy**	**Intervals**
B/C	2.0191	(1.707, 2.0191, 2.0327)	(1.6, 2.16)

Table 2. Results of Financial Analysis applied to Example of Table 1

Order of Magnitude Reasoning

Ordering constraints impose a partial order on the quantity space. Examples of this kind of constraints are A = B, A < B, and A > B, where A and B are two quantities that belong to the model. Given the constraints A < B and B < C, the system should be able to infer A < C (all valid inferences need to be sanctioned, of course).
The example in the preceding paragraph mentions only ordering constraints. When people use common sense reasoning, or when engineers deal with numeric expressions, often use order of magnitude relations to simplify expressions and models. It is very common to find physics books that drop a term from an expression to simplify a given model, justifying their assumption saying that the dropped term was negligible with respect to other terms in the same expression. To cope with those situations, we define a set of om operators and their semantics as shown in table 3.

Operator	**Meaning**
<<	Much smaller
-<	Moderately smaller
~<	Slightly smaller
=	Equal
>~	Slightly greater
>-	Moderately greater
>>	Much greater

Table 3. Order of Magnitude Operators

Several semantics for drawing inferences about omrs have been defined [Raiman86, Mavrovouniotis87, Flores96]. The main difference among them is their aggressiveness. That is, some of them sanction inferences like A ~< B and B ~< C → A ~< C, which is error prone, but correct enough to make a person happy with it.

Interval Representation

In performing OMR, we have chosen to allow the operators shown in Table 1. If A << B, it must be that A/B << 1, which implies A/B < e, for some small e. That way, the different OM relations can be expressed in terms of the quotient of the quantities they compare. Given that, the OM operators can be expressed by the intervals they cover in the real line. For example, if A << B, then A/B ∈ [0, e]. The set of operators can be defined as in Figure 1.

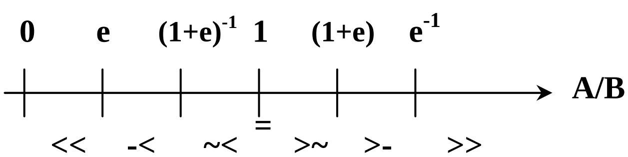

Figure 1. Order of Magnitude Operators as Intervals

All operators in Figure 1 have the intuitive semantics, expressed in [Mavrovouniotis87]. A = B has the standard meaning; A ~< B means A is slightly less than B; A -< B means A is significantly less than B; A << B means A is negligible with respect to B.
With the interval representation, we allow composite OM relations. For example, A < B means A is less than B, and represents the union of the three intervals between operators ~< and <<.

Constraint Propagation

The problem of inferring new constraints from a set of given ones is equivalent to that of computing the transitive closure of a labeled graph, where the vertices are variables and the edges represent relations among any two vertices. The Floyd-Warshall algorithm [CLR91] can be used to solve this problem in time bounded by O(n3), where n is the number of variables involved in all constraints in the model. This algorithm is easier to code if the constraint set is represented as a matrix M, indexed by the variables and where MA,B=r iff A r B for any A, B variables in the model. Figure 2 shows the algorithm, where n is the order of the matrix (i.e. number of variables), and T and Told are temporal variables.
The main idea of the algorithm is to compute the relation between two variables i and j, based on an auxiliary variable k. Returning to the example in the previous paragraph, let i=A, j=B, k=C, T[i,j]=< (i.e. A < B), and T[j,k]=< (i.e. B < C) at the time of the assignment statement in the algorithm. AND(<, <) = < represents the

```
Transitive_closure(M)
 T ← M
 For k = 1 to n
    T[k, k] ← =
 For k = 1 to n
    Told ←T
    For i = 1 to n
       For j = 1 to n
          if i ≠ j then
             T[i,j] ← OR(Told [i,j],
                          AND(T[i,k], T[k,j]))
 Return T
```

Figure 2. The Floyd-Warshall Algorithm

intuitive knowledge A < B, B < C → A < C. OR(?,<) = < means that if we did not know anything about the relation between A and C, and we discover that A < C, we can consider the discovered relation a valid one.
In other implementations (e.g. [Flores97, Flores96, Mavrovouniotis87]) the sanctioned semantics for the AND and OR functions are expressed as relations (i.e. triplets). For example, AND(~<, <<, <<) means that given A~<B and B << C, we can infer A << C. OR(<, <<, <<) means that if we previously knew that A < C, and just inferred that A << C, which is more restrictive, we can update our knowledge about the relationship between A and C to A << C. On the other hand, OR(~<, <<, fail) represents that if we knew that A ~< C, and discover that A << C, our discovery is inconsistent with our previous knowledge.
As mentioned in section 1.1, omrs are represented as intervals. Under this representation, inferences can be done using interval operations, which can be coded in a more concise way than qualitative (symbolic) inference rules (less intuitive, though).
Given that X=A/B ∈ [x1, x2] and Y=B/C ∈ [y1, y2], if we want to infer the relation between A and C, XY = (A/B) (B/C) = A/C produces the right relation. If X ranges in interval [x1, x2], and Y ranges in interval [y1, y2], XY ranges in the product of the two intervals. Thus, the function AND, mentioned in the Floyd-Warshall algorithm, can be implemented using interval product, as defined in the standard bibliography. That is, Z=X AND Y = [x1, x2] [y1, y2].
Now let us say we know A/B ∈ [x1, x2] and we discover A/B ∈ [x3, x4]. If x1< x3 we just discovered a tighter left limit for the ratio A/B. If x1>= x3, we already knew something at least as good or better (tighter); the old limit is kept. We proceed similarly for the right limit. So the update function, OR, can be implemented using interval intersection.
omrs can be seen as constraints over the ratios of quantities. Seen that way, we can see the OR function as an update to those constraints, where interval intersection updates both, the left and right limits of those intervals. Under those considerations, when we derive two intervals that do not intersect, we can say that we have derived contradictory constraints. For example, if we knew that

$A << B \rightarrow A/B \in [0, 0.2]$1, and derive $A \sim< B \rightarrow A/B \in [0.833, 1]$, we can say that those constraints are contradictory because their respective interval representations do not intersect. That is, $[0, 0.2] \cap [0.833, 1] = \phi \rightarrow \perp$.
Let us consider the following inference example:

$$\begin{gathered} A << B \rightarrow A/B \in [0, e] = [0, 0.2] \\ B -< C \rightarrow B/C \in [e, (1+e) -1]=[0.2, 0.833] \\ (A/B)\ (B/C) = A/C \in [0, e/(1+e)]=[0, 0.166] \end{gathered} \quad (1)$$

The resulting interval implies that A << C. Now, assuming that we previously knew that $A/C \in [0, 1]$ (i.e. A < C), we could update our knowledge by taking the intersection of the two intervals.

$$[0, 1] \cap [0, 0.166] = [0, 0.166] \quad (2)$$

By performing inferences using intervals, we can draw more aggressive inferences in a safer way than the implementations of [Flores96, Mavrovouniotis87, Raiman86]. For example,

$$\begin{gathered} A << B \rightarrow A/B \in [0, 0.2] \\ B >\sim C \rightarrow B/C \in [1, 1+e]=[1, 1.2] \\ A/C \in [0, 0.24] \rightarrow A << \ldots -< C \\ C \sim< D \rightarrow C/D \in [0.833, 1] \\ A/D \in [0, 0.24] \rightarrow A << \ldots -< D \end{gathered} \quad (3)$$

In [Raiman86] the notation A << … -< C stands for an OMR that ranges from << to -<. The composition of relations A << B and B>~ C produces the ambiguous relation A << … -< C. Nevertheless, our intuition indicates that since now C~<D, that takes D to be very close to B (see Figure 3). Interval computation of the omrs performs the right inference, while the semantics described in [Flores96, Mavrovouniotis87, Raiman86] do not. For example, Flores sanctions A << B, B >~ C → A < C and A < C, C ~< D → A < D, which is counterintuitive.

Notice that the inference process in the previous example took place without the need to declare the semantics for each possible combination of omrs. Comparing our results with those presented in [Mavrovouniotis87], his formalism defines 28 legitimate omrs; in our representation, we keep the omrs as intervals all the time, having an infinite number of possible omrs. The interval definitions shown in Figure 2 are used to convert the output interval omrs to qualitative ones.

[1] Throughout the paper, we assume **e**=0.02 (see [Mavrovouniotis87]).

Disambiguation Algorithm

A very common data class in the area of Qualitative Reasoning is sign algebra [DeKleer84]. If we use sign algebra to solve this problem, B-C would be ambiguous in most cases. We can use omrs to solve the ambiguity. Even if we have omrs among the model variables, it is impossible, or very costly in case it is possible, to derive an omr that relates B and C. Furthermore, there may be cases where we have just a few omrs, relating a subset of the model variables. We can still use that information to try to determine groups of variables in B dominating groups of variables in C, or vice versa.
The omrs in our model form a graph where the vertices are variables and the edges the omrs, relating pairs of variables. If, after propagation, the graph was full, we could refer each variable to the first one (could be any one), and determine whether the resulting interval for B-C is less than, contains to, or is greater than zero.
If the information we have does not produce a full graph after propagation, then we have islands of related variables. If we do not count with related omrs, all islands contain pairs of variables. If any variable does not take part in any omr, it will form an island by itself.
In the case the omr graph contains islands, we group the variables Vi, and for each island we refer the variables to the first positive one. If at the end all groups have the same sign, that sign will be the sign of B-C; otherwise, we cannot tell, using that information.
To determine the islands, we can use a data structure known as Union-Find. At the beginning each variable is stored in a heap. For every omr involving Vi and Vj, we union their respective heaps. When we finish with the omrs, the resulting heaps represent the islands in the graph. Figure 3 shows the algorithm to determine the islands in an omr graph and Figure 4 shows the algorithm to determine the discriminant B-C.

```
Find-Islands(omrs, V)
    For I=1 to |V|
        Make-heap(Vi)
    For each (Vi r Vj) ∈ omrs
        Union(Find(Vi), Find(Vj))
```

Figure 3. Algorithm for the Determination of Islands in the omr graph

```
Qualitative-BC(omrs, V)
    Find-islands(omrs)
    Outcome = 0
    For each island in the graph
    Refer all variables to the first one in the island
    Determine result (sign) for that island
            outcome += result
    Return outcome
```

Figure 4. Algorithm for Qualitative Financial Analysis

	1	2	3	4	5	6	7	8
1	=	(1, 22.222)	?	(1, 21)	(1, 20)	?	<~	=~
2	(0.045, 1)	=	?	<~	(0.9063, 1)	?	(0.045, 0.952)	(0.42, 1.049)
3	?	?	=	?	?	>-	?	?
4	(0.047, 1)	>~	?	=	<~	?	(0.0476, 0..952)	(0.44, 1.05)
5	(0.05, 1)	(1, 1.103)	?	>~	=	?	<-	(0.0476, 1.05)
6	?	?	<-	?	?	=	?	?
7	>~	(1.05, 22.05)	?	(1.05, 21)	>-	?	=	(0.952, 1.1025)
8	=~	(0.0952, 23.809)	?	(0.952, 22.349)	(0.952, 21.008)	?	(0.907, 1.050)	=

Table 4. omr Matrix after Propagation

Consider an example with eight variables, where V1, V3, V4, V5, and V7 are positive (i.e. they contribute to B), and V2, V6, V5, V7, and V8 are negative. The initial omrs are shown in Figure 5 (dark nodes are negative variables).

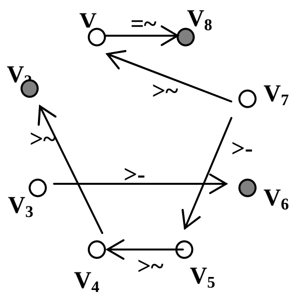

Figure 5. Constraint Graph for Example

After propagation, the final matrix still preserves the two islands. The resulting matrix is shown in Table 4. All operators have been printed in their shortest form, some of them match a qualitative relation, so they are printed in their symbolic form. ? stands for an unknown value.

After propagation, the matrix contains the best of our knowledge about the relation between every pair of variables. We can use the relation between variables (given in the matrix) to refer every variable in an island to the first positive one. If the result is positive, the addition of the group will be positive. For instance, V1 (1, 22.22) V2, means that V1/V2 $\in$ (1, 22.22), which means that V2 $\in$ V1/(1, 22.22). For our example, the first island, composed of V3 and V6, we have that:

$$V3 >\text{-}\ V6 \tag{4}$$
$$V3\ (1 - (0.05, 0.952)) = V3\ (0.048, 0.95) \rightarrow +$$

For the second island:

$$(V1+V4+V5+V7-V2-V3) \tag{5}$$
$$= V1((1 + (0.047, 1) + (0.05, 1) + (1, 1.05)) - ((0.045, 1) + (0.952, 1.05)))$$
$$= V1(0.047, 3.053) \rightarrow +$$

We can tell B-C will be positive under those conditions. The investment project can be accepted.

Mixed Propagation

As part of his dissertation, Flores [Flores97, Flores96] developed a framework to perform mixed propagation. That inference engine, called HRCP (Hybrid Representation Constraint Propagation), accepts a set of constraints of heterogeneous nature, and computes as many derived constraints and refines others. Those constraints can be omrs, algebraic constraints, value constraints, etc.

We can use that system and see it as a black box that computes exactly what we need. The kinds of derivations it can perform are:

- Omr propagation, using the Floyd-Warshall algorithm
- Value propagation, using algebraic and value constraints. If you know C=A+B, and know the values for any two variables, you can compute the third one.
- Discovery of omrs from numeric (interval values). If you know A=0.1 and B=50, you can say A<<B.
- Value refinement using value constraints and omrs. If you know A>~B, and A=(50, 75), and B=(30, 150), you can refine B's value to (50,90).
- Inconsistency checks. If you said A<B and later on you assert A>-B, something must have gone wrong.

Using that inference engine, we can build a system that incrementally accepts information, runs propagation when new facts arrive, and improve the solutions it provides, as information is available. Perhaps with the information at hand we cannot tell much about a given investment project, but later on we acquire more information to be able to disambiguate the results. Later, as we refine our information (i.e. narrow the intervals), the solution becomes more accurate.

For instance, suppose the relation V3 >- V6, given in Eq. 4 was not known. In that case, the outcome of function Qualitative-BC would be unknown. Now assume the user asserts V3 = (2.1, 20), and V6 = (1, 2). The system determines V3 (0.05, 0.952) V6, and it is until then that it can determine a positive result for the investment project under analysis.

Conclusions

We have presented a framework capable of performing financial analysis at several levels of granularity in the uncertainty of its input data. The uncertainty grain may range from sign values and omrs to precise (real) values. This representation scheme solves the problem in situations where other approaches fail. That is, with the amount of uncertainty in the information we need to produce a decision, other systems would fail.

Another feature of our framework is that it works incrementally. That is, at some point the information the user provides may not be enough to produce a non-ambiguous result. Later on, the arrival of new information may be used to disambiguate the situation. Also, in the part of value propagation (computing the value of B-C), the result improves as new information arrives. As we narrow the intervals of the input data, the intervals of the result narrow as well. At the extreme, if all variables are specified precisely, the result becomes a real number. All this can happen in one single run of the program; it is not necessary to re-run the program to add new information.

In most cases, we perform Financial Analysis in a planning horizon of around one year, which means that the number of periods is 12, or at most a multiple of 12. Given the size of the problem, the propagation algorithms can be performed efficiently.

Bibliography

[CLR91] Cormen T, Leiserson C., Rivest R. Introduction to Algorithms. Mc Graw Hill. 1991

[deKleer84] de Kleer J. A Qualitative Physics Based on Confluences. Artificial Intelligence. 24:7-83. 1984.

[Flores96] Flores J., Farley A. Hybrid Representation Constraint Propagation. Cancún México, 1996.

[Flores97] Flores J. Qualitative Reasoning About Linear Circuits in Sinusoidal Steady State. Ph.D. Dissertation. University of Oregon. Eugene, Oregon, USA. 1997.

[Forbus84] Forbus K. Qualitative Process Theory. Artificial Intelligence. 24:85-168. 1984.

[Gonzalez99] González F., Flores B., Márquez N. Técnicas Financieras Fuzzy en la Toma de Decisiones. VI Congress of the International Society of Fuzzy Management and Economy. Pp 309-319. Morelia, Mexico. 1999.

[Gonzalez85] González F. Los Proyectos en la Industrialización Forestal. Ed. Universitaria. Universidad Michoacana de San Nicolás de Hidalgo. Morelia, México. 1985.

[Gotz83] Gotz and Herzberger, Jurgen Alefeld. Introduction to Interval Computations. Academic Press. 1983.

[Kaufmann93] Kaufmann A., Gil Aluja Jaime. Técnicas Especiales para Gestión de Expertos. Millaidoro. 1993.

[Kosko96] Bart Kosko. Fuzzy Engineering. Prentice Hall. 1996.

[Mavrovouniotis87] Mavrovouniotis M., Stephanopoulos G. Reasoning with Orders of Magnitude and Approximate Relations. 1987.

[Pedrycz98]Witold Pedrycz, Fernando Gomide. An Introduction to Fuzzy Sets : Analysis and Design (Complex Adaptive Systems). MIT Press. 1998.

[Raiman86] Raiman O. Order of Magnitude Reasoning. In proceedings of the American Association for Artificial Intelligence (AAAI-86). 100-104. 1986

[Terceño98] Terceño Gómez A. Instrumentos para análisis empresarial en un ambiente incierto. Facultat de Ciencies Economiques i Empresarials. Universitat Rivira i Virgili - Spain.

[Zollo00] Giuseppe Zollo et. al. The evaluations of customers satisfaction through fuzzy logic: A case - study. University of Naples Federico II -Italy.