

**LINEAR OPTIMIZATION
WITH
MICROSOFT EXCEL[®]**

June 30, 2005

Peter Wood

Faculty of Mathematics
University of Waterloo
Waterloo, Ontario
Canada

© 2005

INTRODUCTION

The purpose of these notes is to provide a basic introduction to Solver, Microsoft Excel's tool for solving linear optimization problems.

These notes are not an introduction to linear optimization. In particular, it is assumed that the reader is familiar with the basic concepts in linear programming, and how to formulate a practical problem into a linear program.

As is true of many software applications, the best way to learn is by example. It is highly recommended that the reader try these examples in Excel while working through these notes.

EXAMPLE 1: A PRODUCT MIX PROBLEM

Suppose a company can manufacture four different alloys, numbered 1 through 4, by combining differing amounts of two constituent components. In particular, suppose we are given the following information:

Metal	Proportion of metal in alloy				Availability per day
	1	2	3	4	
1	0.5	0.6	0.3	0.1	25 tons
2	0.5	0.4	0.7	0.9	40 tons
Alloy price (\$/ton)	750	800	1200	1400	

Table 1: Problem Data

From the table, we see that alloy #2 consists of 60% metal 1 and 40% metal 2, and it sells for \$800 per ton. Additionally, we are given that there is a maximum of 25 tons of metal 1 available and a maximum of 40 tons of metal 2. Our goal is to find a production program that will maximize revenues. (From: Murty, K.G., "Chapter 2 Modeling Linear Programs" from Operations Research: Deterministic Optimization Models, p 22-28 Prentice-Hall 1995)

Solution

It is helpful to clearly identify all the decision variables, the objective function and the constraints before we attempt to apply Excel.

Decision Variables:

Clearly there are four decision variables, namely x_i where $i = 1 \cdots 4$, and where x_i = the amount of alloy i produced.

Objective Function:

Since our goal is to maximize total revenue, we have the objective function

$$\text{Maximize } 750x_1 + 800x_2 + 1200x_3 + 1400x_4$$

Constraints:

In this problem, there are only two types of constraints: non-negativity and the supply availability.

Non-negativity. Since production cannot be negative, we have

$$x_i \geq 0 \text{ for all } i$$

Supply constraints: Since we have a limited supply of raw materials, we have

$$\text{Metal 1 constraint} \quad 0.5x_1 + 0.6x_2 + 0.3x_3 + 0.1x_4 \leq 25$$

$$\text{Metal 2 constraint} \quad 0.5x_1 + 0.4x_2 + 0.7x_3 + 0.9x_4 \leq 40$$

To solve this linear program in Excel, we begin by building the following spreadsheet.

Inputs - Costs, Product Mix, Capacity					
	Alloy Mix				Maximum
Metal	1	2	3	4	Availability
1	0.5	0.6	0.3	0.1	25
2	0.5	0.4	0.7	0.9	40
Price (\$/ton)	750	800	1200	1400	

Decision Variables				
	1	2	3	4
Amount Produced	0	0	0	0

Constraints	
	Unused Supply
1	25.00
2	40.00

Objective Function	
	Revenue
	0

Initially, the decision variables are set to zero.

Cell B14 contains the formula

=F4-SUMPRODUCT(B4:E4,B\$10:E\$10). The SUMPRODUCT function has as its input two arrays of the same size. The function then multiplies the corresponding entries component-wise, then takes the sum of all these values. For example the formula
 =SUMPRODUCT(B4:E4,B\$10:E\$10) is equivalent to =B4*B10+C4*C10+D4*D10+E4*E10

Cut and paste the contents of cell B14 into B15 to get
`=F5-SUMPRODUCT(B5:E5,B$10:E$10)`

(Notice that we used the *absolute reference* B\$10:E\$10 instead of the *relative reference* B10:E10 in cell B14 – this is the purpose of the \$ – had we not done this, we would not be able to simply “cut and paste” the formula. If you are unfamiliar with cell referencing in Excel, consult the Excel help for more detailed information)

Notice now that the original supply constraint on Metal 1 is now equivalent to the constraint that Cell B14 is greater than or equal to zero, and the original supply constraint on Metal 2 is now equivalent to the constraint that Cell B15 is greater than or equal to zero. In Excel, it is generally easier to encode (as well as easier to follow) if all the constraints are simply a matter of ensuring various cells are greater than or equal to zero, as opposed to some other constant.

Cell B18 contains the objective function `=SUMPRODUCT(B10:E10,B6:E6)`, which represents total revenue.

Now we invoke Solver. Click on **Tools/Solver** to get the following dialogue box (If **Solver** is not listed under the **Tools** menu, click **Add-ins** under **Tools**, scroll through the list until you find “Solver Add-In” and select it.)

The screenshot shows a Microsoft Excel window titled "Example 1". The spreadsheet contains the following data:

Inputs - Costs, Product Mix, Capacity					
	Alloy Mix				Maximum
Metal	1	2	3	4	Availability
1	0.5	0.6	0.3	0.1	25
2	0.5	0.4	0.7	0.9	40
Price (\$/ton)	750	800	1200	1400	

Decision Variables				
	1	2	3	4
Amount Produced	0	0	0	0

Constraints	
Unused Supply	
1	25.00
2	40.00

Objective Function	
Revenue	0

The Solver Parameters dialog box is open, showing the following settings:

- Set Target Cell:** \$B\$18
- Equal To:** ☒ Max ☐ Min ☐ Value of: 0
- By Changing Cells:** \$B\$10:\$E\$10
- Subject to the Constraints:**
 - \$B\$10:\$E\$10 >= 0
 - \$B\$14:\$B\$15 >= 0

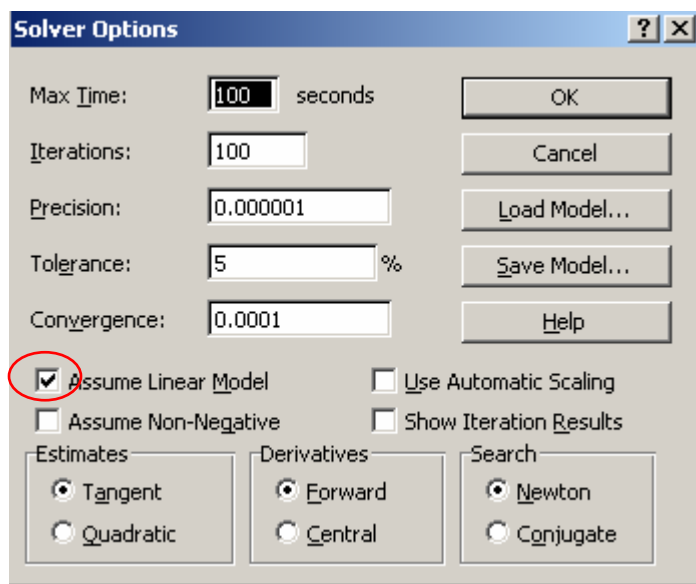
The Solver Parameters dialog box also includes buttons for Solve, Close, Options, Add, Change, Delete, Reset All, and Help.

In the **Set Target Cell** we select the cell containing our Objective Function (select it by placing the cursor in the **Set Target Cell** area, then simply click on Cell B18). Ensure the **Max** button is

selected (since this is a maximize problem). Place the cursor in the **By Changing Cells** area and select the region B10 to E10 (these are our decision variables).

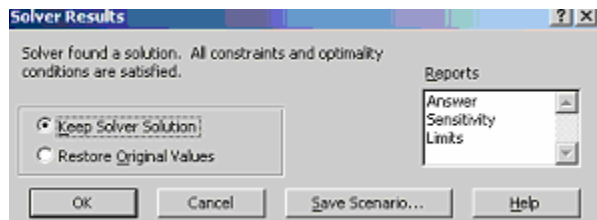
To add constraints, click on **Add**. In the **Cell Reference** box, highlight our decision variables B10 to E10, select \geq from the drop down list, and put 0 in the **Constraint** box. (This is the non-negativity constraint). Now click **Add** to add the next constraint. Select cells B14 and B15 and set them ≥ 0 . (This is our re-formulated supply constraint). Click **OK** to return to the first dialogue box.

It is important to ensure Solver assumes the model we are optimizing is a linear model. This will help both the computation speed as well as its accuracy. To do this, click **Options** in the Solver dialogue box. Then check the box next to **Assume Linear Model** in the Options dialogue box, as shown below.



Now click OK.

Once you have returned to the Solver dialogue box, click **Solve**. Solver will now calculate the optimal solution and present the following box:



Click OK and Solver will place the solutions in the various boxes (as shown below)

The screenshot shows an Excel spreadsheet titled "Microsoft Excel - Example 1". The spreadsheet is organized into several sections for a linear optimization problem:

- Inputs - Costs, Product Mix, Capacity:**

Metal	1	2	3	4	Maximum Availability
1	0.5	0.6	0.3	0.1	25
2	0.5	0.4	0.7	0.9	40
- Price (\$/ton):**

1	750	800	1200	1400
---	-----	-----	------	------
- Decision Variables:**

	1	2	3	4
Amount Produced	0	18.33333	46.66667	0
- Constraints:**

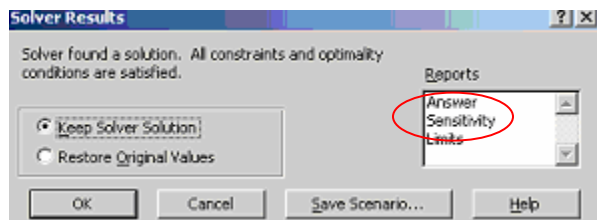
	Unused Supply
1	0.00
2	0.00
- Objective Function:**

Revenue	70666.67
---------	----------

The Solver Parameters dialog box is open, showing the objective function cell (F6) and the variable cells (B9:E10). The Solver has found a solution, and the "Reports" section is visible, with "Answer Sensitivity Links" selected.

Notice that the optimal revenue is \$70,666.67, reached when we produce 18.3 tons of alloy 2, 46.7 tons of alloy 3 and none of alloys 1 or 4. Additionally, we have no unused supply of either of our input metals.

Solver can produce additional information if desired. For example if we click on **Sensitivity** in the **Reports** section (as shown below).



Excel will now generate a new sheet called Sensitivity report 1, detailing various pieces of information including for example the Shadow Prices of the Metals 1 and 2. (As shown below). Solver can provide other information about the solution, however we will not explore that further here.

Microsoft Excel 9.0 Sensitivity Report

Report Created: 8/30/2004 1:07:10 PM

Adjustable Cells

Cell	Name	Final Value	Reduced Cost	Objective Coefficient	Allowable Increase	Allowable Decrease
\$B\$10	Amount Produced Alloy Mix	0	-183.3333334	750	183.3333334	1E+30
\$C\$10	Amount Produced	18.33333333	0	800	99.99999996	114.2857143
\$D\$10	Amount Produced	46.66666667	0	1200	200.0000001	39.99999998
\$E\$10	Amount Produced	0	-66.66666664	1400	66.66666664	1E+30

Constraints

Cell	Name	Final Value	Shadow Price	Constraint R.H. Side	Allowable Increase	Allowable Decrease
\$B\$14	Unused Supply	0.00	-266.67	0	7.857142858	35
\$B\$15	Unused Supply	0.00	-1600.00	0	23.33333333	18.33333334

It is worth mentioning that since Solver keeps the parameters, it is easy to make changes to the original problem and to see the effects. For example, suppose the supply of Metal 1 was increased to 50 tons. We can simply change the contents of Cell F4 to 50 and re-run Solver. We do not need to re-enter or make any changes to the constraints. All we have to do is select **Solver** and click on **Solve**. This helps explain why we encoded the supply constraint in the fashion shown, as opposed to trying to enter the constraint constants in the Solver dialogue box. With the method presented, all the changes to the model can be entered into the spreadsheet; we should not need to change any of the Solver Parameters.

As a final observation, notice how the data relating to the alloys was entered in columns, thus the decision variables (how much of each alloy to produce) were also entered in columns. Additionally, the data on the two metals appeared in rows, thus the supply constraints also appeared on two rows. While this is not necessary, it does make the formula entry *much* easier, as we can simply use the SUMPRODUCT function in one easy step. You will find that spending a few minutes of careful thought on how to set up the spreadsheet will save a lot of time and will make your spreadsheet much more readable.

EXAMPLE 2: A MIXED 0-1 PROBLEM

In a mixed 0-1 problem, some of the decision variables may only assume the value 0 or 1 (called binary variables) while others may take on non-integer values. The following is an example of such a problem.

A company has two currently operating plants, plants 1 and 2, for manufacturing a product which they sell through four distribution centers. Due to increasing demand for the product, the company has decided to set up a new plant that can produce 1100 units of finished product per week. Data on the unit transportation costs and expected weekly demand are given below.

Plant	Unit transp. costs to dist. centre				Weekly Prod Capacity
	1	2	3	4	
1	0.4	0.6	0.3	0.5	1500
2	0.9	1.0	0.3	0.8	1800
New, to be built					1100 (planned)
<i>Expected weekly demand</i>	1400	1100	900	1000	

Table 2: Production Data

It takes 0.54 units of labour, 0.10 units of raw material R1, 0.04 units of raw material R2 and 0.20 units of coal to manufacture one unit of product at the new plant.

Two sites, sites 1 and 2, are available to locate the new plant 3. The cost of shipping the product from a new plant at site 1 (resp. 2) to the four distribution centers is 0.3, 0.5, 0.4, 0.2 (resp. 0.4, 0.6, 0.1, 0.3) per unit.

The cost of setting up a new plant at site 1 or 2 is the same.

The costs of the various inputs at the two sites are given below.

Site	Cost per unit of			
	Labour	R1	R2	Coal
1	4	0.05	0.03	0.04
2	4.5	0.04	0.03	0.03

Table 3: Cost Data

The company's plants 1 and 2 are well established and have been running for some time. It wants to determine whether to set up the new plant at site 1 or 2, in order to minimize the production costs at this plant plus the weekly shipping cost. (From: Murty, K.G., "Chapter 2 Modeling Linear Programs" from Operations Research: Deterministic Optimization Models, p 22-28 Prentice-Hall 1995)

Solution:

Decision Variables:

To find our total costs, we need to know the number of units produced at each plant, and how many of those are shipped to the various distribution centers. Thus we let

x_{ij} = number of units produced at plant i and shipped to distribution centre j for $i, j = 1 \cdots 4$ (here, plant 3 refers to the production at the new site #1, and plant 4 refers to the production at the new site #2)

Since we will open only one of the two sites, we introduce the binary variables to capture this. Thus, let

$y_i = 1$ if plant i is operating and 0 otherwise (for $i = 1 \dots 4$). Notice from the problem that plant 1 and 2 will remain open, regardless of where we build the new plant. Therefore, it seems that we need only 2 binary variables, one for each of the new sites. However, as we shall see, it will be easier to encode all this in the spreadsheet if we have one binary variable for each plant.

Objective Function:

Our goal is to minimize cost. The cost consists of production costs and transportation costs to the various distribution centers. Thus we have

Total Cost = Production costs + Shipping costs.

Production Costs: Note that at site 1, the production cost is the sum of labour, cost of raw materials 1 and 2, and coal, which is

$$(.54)(4) + (.1)(.05) + (.04)(.03) + (.2)(.04) = \$2.1742 \text{ per unit}$$

A similar calculation yields production cost at site 2 of \$2.3872 per unit.

Notice that we are not given the production cost data for the two existing plants. Since it has been decided that those two plants will remain open, regardless of whether we build at the new site #1 or #2, and that the production costs at the existing plants are independent of where we build the new plant, these data are not necessary for the model.

Thus,

$$\begin{aligned} \text{Production Costs} &= \$2.1742 * \text{plant 3 production} + \$2.3872 * \text{plant 4 production} \\ &= 2.1742 * \sum_j x_{3,j} + 2.3872 * \sum_j x_{4,j} \end{aligned}$$

(In practice, it is likely better to include the production cost data as part of the spreadsheet. This way, it is very easy to adjust some of the cost parameters and re-run the model. However, in this example, we choose not to do this)

Shipping Costs: If we let c_{ij} = cost to ship one unit from plant i to distribution centre j , then the total shipping costs are clearly given by $\sum_{i,j} c_{ij} x_{ij}$.

Thus the total costs are: $2.1742 * \sum_j x_{3,j} + 2.3872 * \sum_j x_{4,j} + \sum_{i,j} c_{ij} x_{ij}$

Constraints:

Let D_j = Demand at distribution centre j , and let C_i = Capacity at plant i , then we have

$$\sum_i x_{ij} = D_j \quad (\text{demand constraint})$$

$$\sum_j x_{ij} \leq y_i C_i \quad (\text{Capacity constraint})$$

Note in the capacity constraint, the maximum capacity is C_i if the plant is open, and 0 otherwise (this is the point of the binary variables).

Since plants 1 and 2 will be open regardless, we have

$$y_1, y_2 = 1$$

Since only one of plants 3 or 4 can be open, we have

$$y_3 + y_4 = 1$$

Finally, since production cannot be negative, we have

$$x_{ij} \geq 0 \text{ for all } i, j$$

Solution:

We begin by building the following spreadsheet.

Microsoft Excel - 0-1 Example

FileEditViewInsertFormatToolsMegaStatDataWindowHelpAcrobat

</

Cells F12 to F15 are the binary variables. Note that F12 and F13 are initially set to 1 (since those plants will be open). All other variables are set to zero.

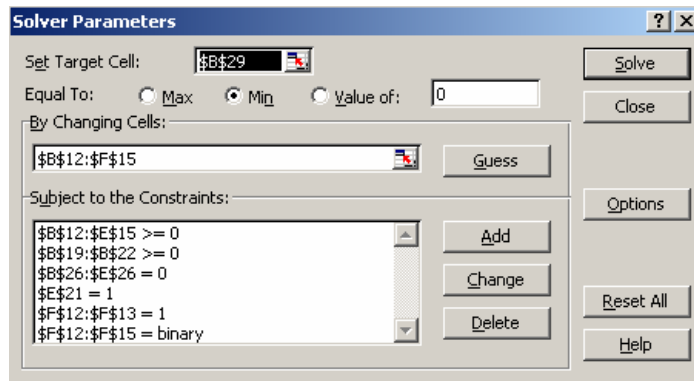
B19 contains the formula `=F4*F12-SUM(B12:E12)`. Note that $B19 \geq 0$ is equivalent to the capacity constraint on plant 1. We cut and paste this formula into B20 to B22. (After the "cut and paste", always check the new formulas to ensure they are correct.)

B26 contains the formula `=B8-SUM(B12:B15)`. Thus $B26=0$ is equivalent to the demand constraint. We cut and paste this into C26 to E26.

E21 contains the formula $=F14+F15$. Thus $E21=1$ is equivalent to the statement that exactly one of site #1 and site #2 has a new plant.

B29 contains the objective function $=2.1742*\text{SUM}(B14:E14)+2.3872*\text{SUM}(B15:E15)+\text{SUMPRODUCT}(B4:E7,B12:E15)$. The first two terms are the production costs, and the last term is the shipping costs.

Now we enter the constraints into solver.



Set Target Cell is our objective function B29. **By Changing Cells** are the decision variables B12 to F15. (Note that the decision variables are a rectangular array. Therefore, if we tried to use only 2 binary variables – one for each of the new plants – instead of one for each plant, it would be very hard to enter the decision variables into solver. Additionally, the entry of the capacity constraint is only one formula cut and pasted instead of two different formulas – one for the old plants and another one for the new plants.)

The first constraint is the capacity constraint, the second is non-negativity, the third is the demand constraint, the fourth ensures exactly one new plant is opened, the fifth ensures plants 1 and 2 remain open, and the last constraint tells solver that we want cells F12 to F15 to be binary values only.

Running Solver yields (remember to select Linear Model under “Options”)

Microsoft Excel - 0-1 Example_1

File Edit View Insert Format Tools MegaStat Data Window Help Acrobat

100%

B29 =2.1742*SUM(B14:E14)+2.3872*SUM(B15:E15)+SUMPRODUCT(B4:E7,B12:E15)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Inputs - Costs, Capacities, Demands														
2		Transportation Costs to centre				Capacity									
3	Plant	1	2	3	4										
4	1	0.4	0.6	0.3	0.5	1500									
5	2	0.9	1	0.3	8	1800									
6	3	0.3	0.5	0.4	0.2	1100									
7	4	0.4	0.6	0.1	0.3	1100									
8	DEMAND	1400	1100	900	1000										
9															
10	Decision Variables														
11	Plant	Production for Centre				Open(1=YES, 0=No)									
12	1	1300	200.0000005	0	0	1									
13	2	0	900	900	0	1									
14	3	100	0	0	1000	1									
15	4	0	0	0	0	1.03E-12									
16															
17	Constraints														
18	Plant	Excess capacity													
19	1	-3.79E-06													
20	2	-2.27E-06													
21	3	-2.52E-06													
22	4	1.14E-09				Open?	1								
23															
24		Distribution Centre													
25		1	2	3	4										
26	Unmet Demand	-3.28E-06	-5.05001E-07	-2.27E-06	-2.53E-06										
27															
28	Objective Function														
29	Costs	4431.62													

Sheet1 Sheet2 Sheet3

Draw AutoShapes

Ready

NUM

3:00 PM

Inbox - Outlook Ex... Wood, Peter - Jan... Notes Solver Notes_2 - Mi... 0-1 Example_1

As this result clearly demonstrates, Excel is not working symbolically with arbitrary precision. To get rid of the ugly numbers like 1.03E-12 (which is basically zero), we can format the cells to 0 decimal places of accuracy, to get the better-looking sheet below.

Microsoft Excel - 0-1 Example_1

File Edit View Insert Format Tools MegaStat Data Window Help Acrobat

Formula Bar: B29 = 2.1742*SUM(B14:E14)+2.3872*SUM(B15:E15)+SUMPRODUCT(B4:E7,B12:E15)

Inputs - Costs, Capacities, Demands					
Plant	Transportation Costs to centre				Capacity
	1	2	3	4	
1	0.4	0.6	0.3	0.5	1500
2	0.9	1	0.3	8	1800
3	0.3	0.5	0.4	0.2	1100
4	0.4	0.6	0.1	0.3	1100
DEMAND	1400	1100	900	1000	

Decision Variables					
Plant	Production for Centre				Open(1=YES, 0=No)
	1	2	3	4	
1	1300	200	0	0	1
2	0	900	900	0	1
3	100	0	0	1000	1
4	0	0	0	0	0

Constraints					
Plant	Excess capacity				
	1	2	3	4	
1	0				
2	0				
3	0		Open?		1
4	0				

Distribution Centre					
	1	2	3	4	
Unmet Demand	0	0	0	0	

Objective Function	
Costs	4431.62

Sheet1 Sheet2 Sheet3

Ready

Taskbar: Start, Inbox - Outlook Ex..., Wood, Peter - Jan..., Notes, Solver Notes_2 - Mi..., 0-1 Example_1, 3:11 PM

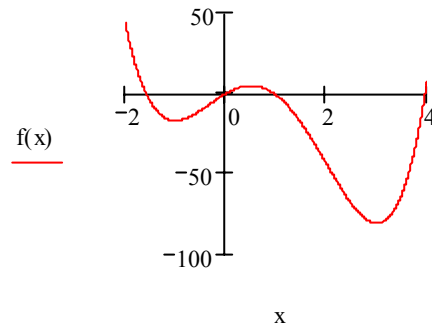
In particular, we see that we should build the new plant at site #1, and this will give total costs of about \$4,431.

ADDENDA

Corel Quattro Pro™ can handle optimization problems as well. The formulas can be entered the same way, and the linear optimizer can be run by selecting **Tools/Numeric Tools/Optimizer**. The general interface is similar to Solver, but with one small difference; Optimizer does not accept binary variables as a constraint. Instead, if you want the variable y to be restricted to binary, you must restrict it to be integer valued, then add the constraints $y \geq 0$ and $y \leq 1$. This has the same effect as restricting y to be binary.

There are lots of other practice problems if you wish to try more. Consult any of the texts in the reference section. The general spreadsheet format used in these notes comes from S. Chopra & P. Meindl's book. It has fewer problems than the Winston book, however, it does contain more information on using Solver to address optimization problems.

Solver can also handle non-linear problems as well, however, the user must be much more careful. The solution methods used in the Solver engine can come to an incorrect answer (or no answer). Try to find the minimum value of $3x^4 - 10x^3 - 12x^2 + 18x$ setting the variable initially to zero. Solver will converge to the solution -17 (found at $x = -1$) while the actual solution is -81 found at $x = 3$. The reason for this error is that Solver uses a "local" method, which converges to the local minimum near the initial value. Consider the graph of the given function.



Starting from $x = 0$, Solver uses a Newton's Method type argument to converge to the local minimum on the left, and not to the global minimum found at $x = 3$. As you may guess from looking at the graph, if we were to set the variable initially to $x = 2$, Solver will converge to the correct solution.

REFERENCES

Chopra, S. and Meindl, P. *Supply Chain Management, Strategy, Planning and Operation* second edition, Person Prentice-Hall, 2004

Murty, K.G., *Operations Research: Deterministic Optimization Models*, Prentice-Hall 1995

Winston, W., *Operations Research, Applications and Algorithms*, Thomson-Brooks/Cole, 2004