

Optimization under Uncertainty

Robert M. Freund

April 27, 2004

©2004 Massachusetts Institute of Technology.

1 Motivation

The development of the concepts of linear and nonlinear optimization models presumes that all of the data for the optimization model are known with certainty. However, uncertainty and inexactness of data and outcomes pervade many aspects of most optimization problems. As it turns out, when the uncertainty in the problem is of a particular (and fairly general) form, it is relatively easy to incorporate the uncertainty into the optimization model.

2 Brief History of Optimization under Uncertainty

- 1950s, Dantzig and Beale started to work on linear optimization under uncertainty
- 1962, solution method developed by Benders
- Many and varied applications in the linear, nonlinear, and discrete form
 - Electric utility capacity planning
 - Financial Planning and Control
 - Supply chain optimization
 - Airline Planning (fleet assignment)
 - Water Resource Modeling
 - Forestry Planning
 - Many others . . .

3 Gemstone Tool Company

Gemstone Tool Company (GTC) is a privately held company that competes in the consumer and industrial market for construction tools. In addition to its main manufacturing facility in Seattle, Washington, GTC also operates several other manufacturing plants located in the United States, Canada, and Mexico. For the sake of simplicity, let us suppose that the Winnipeg, Canada plant only produces wrenches and pliers. Wrenches and pliers are

made from steel, and the process involves molding the tools on a molding machine and then assembling the tools on an assembly machine. The amount of steel used in the production of wrenches and pliers and the daily availability of steel is given in the first line of the Table 1 below. On the next four lines are the machine utilization rates needed in the production of wrenches and pliers and the capacity of these machines as well. Finally, the last two rows of the table indicate the daily market demand for these tools, and their variable (per unit) contribution to earnings.

	Wrenches	Pliers	Availability (Capacity)
Steel (lbs.)	1.5	1.0	27,000 lbs./day
Molding machine (hours)	1.0	1.0	21,000 hours/day
Assembly machine (hours)	0.3	0.5	9,000 hours/day
Demand limit (tools/day)	15,000	16,000	
Contribution to earnings (\$/1,000 units)	\$130	\$100	

Table 1: Data for the Gemstone Tool Company

GTC would like to plan for the daily production of pliers and wrenches at its Winnipeg plant so as to maximize the contribution to earnings. The mathematical formulation of GTC's linear programming problem is as follows:

$$\text{Maximize contribution} = 130W + 100P$$

subject to:

$$\begin{array}{rcll}
 \text{Wrench demand:} & W & & \leq 15 \\
 \text{Plier demand:} & & P & \leq 16 \\
 \text{Steel availability:} & 1.5W & + & P \leq 27 \\
 \text{Molding machine usage:} & W & + & P \leq 21 \\
 \text{Assembly machine usage:} & 0.3W & + & 0.5P \leq 9 \\
 & W & & \geq 0 \\
 & & P & \geq 0
 \end{array}$$

The solution of this linear program is given by $W^* = 12$, $P^* = 9$, with a contribution to earnings of $\$2,460 = 130W^* + 100P^*$.

4 GTC Planning under Uncertainty

For the current quarter, GTC had contracted with a steel supplier for the delivery of 27,000 lbs. of steel per day. This restriction was incorporated into the GTC linear optimization model as the steel availability constraint:

$$1.5W + 1.0P \leq 27.$$

Now suppose that GTC is planning for next quarter, and that they would like to determine how much steel to contract for with local suppliers for the next quarter. Suppose that steel contracts are typically arranged for daily deliveries over the entire quarter, and that the market price for such contracts is $\$58.00/1,000$ lbs. of steel. Let us define the decision variable:

S = the amount of steel to contract for, for next quarter, in 1,000 lbs./day.

GTC would like to determine the optimal value of S . We suppose that the following aspects of the problem are known:

- The market price of steel is $\$58.00/1,000$ lbs.
- The utilization of steel, assembly machine hours, and molding machine hours in wrenches and pliers is the same as given in the original problem.
- The molding machine capacity is the same as in the original problem, namely 21,000 hours/day.
- The demand for wrenches and pliers is the same as in the original problem, namely 15,000 wrenches per day and 16,000 pliers per day.
- The unit contribution to earnings of production of pliers is the same as in the original problem, namely $\$100/1,000$ units.

However, we also suppose that the following aspects of the problem are uncertain:

- The assembly machine capacity for next quarter is uncertain. GTC has ordered new assembly machines to replace as well as to augment their existing assembly machines, but it is not known if these new machines will be delivered in time to be used next quarter. Let us suppose that the assembly machine capacity for next quarter will either be 8,000 hours/day (with probability 0.5) or 10,000 hours/day (with probability 0.5).
- The unit contribution to earnings of production of wrenches next quarter is uncertain, due to fluctuations in the market for wrenches. Suppose that GTC estimates that the unit contribution to earnings of wrenches will be in the range between \$90 and \$160. For the sake of simplicity, let us suppose that this unit earnings contribution will be either \$90 (with probability 0.5) or \$160 (with probability 0.5).

The data for this problem is summarized in Table 2.

	Wrenches	Pliers	Availability
Steel (lbs.)	1.5	1.0	<i>S</i> (to be determined)
Molding Machine (hours)	1.0	1.0	21,000 hours/day
Assembly Machine (hours)	0.3	0.5	either 8,000 hours/day or 10,000 hours/day
Demand Limit (tools/day)	15,000	16,000	
Contribution to Earnings (\$/1,000 units)	either \$160 or \$90	\$100	

Table 2: Data for the Gemstone Tool Company steel supply planning problem.

GTC must soon decide how much steel per day to contract for, for next quarter. At the beginning of next quarter, the assembly machine capacity will become known. Also, at the beginning of next quarter, the unit earnings

contribution of wrenches will become known. This sequence of events is shown in Table 3.

Time	Event or Action
Today:	GTC must decide how much steel per day to contract for for the next quarter.
Soon thereafter:	<ul style="list-style-type: none"> • GTC will discover the actual assembly machine availability for next quarter (either 8,000 or 10,000 hours/day). • GTC will discover the actual unit earnings contribution of wrenches for next quarter (either \$160 or \$90/1,000 units).
Next quarter:	GTC must decide the production quantities of wrenches and pliers.

Table 3: The sequence of events in the GTC steel supply planning problem.

4.1 Stage-one and Stage-two

We will divide up the flow of time in our problem into two stages, which we refer to as “stage-one” and “stage-two”, and where today (that is, the current quarter) is stage-one, and next quarter is stage-two. In our steel supply planning problem, there is only one decision to make in the stage-one, namely the amount of steel S to contract for, for next quarter. The decisions that must be made next quarter are the stage-two decisions. The stage-two decisions for our steel supply planning problem are the quantities of wrenches and pliers to produce next quarter. Note that these decisions do not have to be made until next quarter.

Note also that in this framework, there is uncertainty in stage-one about what the data for the problem will be in stage-two. That is, in stage-one, we do not yet know what the assembly machine capacity will be next quarter (either 8,000 hours/day or 10,000 hours per day), and we also do not know what the wrenches unit earnings contribution will be next quarter (either \$160/1,000 units or \$90/1,000 units). However, this uncertainty will be resolved prior to the start of stage-two.

4.2 Formulation of the Problem as a Linear Optimization Model

We begin the formulation of the steel supply planning problem by identifying the decision variables for stage-one. Recall that S is the amount of steel per day to contract for, for next quarter. The first-stage decision that GTC needs to make is the amount of steel per day to contract for, for next quarter, which is S .

The next step in the formulation of the model is to identify the decision variables for stage-two of the problem. In order to identify these decision variables, we first need to enumerate all of the possible “states of the world” that might transpire next quarter. Table 4 shows the four possible states of the world that might transpire next quarter, with their associated probabilities. There are four possible states of the world for next quarter, corresponding to the two possible assembly machine capacity values, and the two possible unit earnings contributions of wrenches. For example, in Table 4, the first state of the world that might transpire is that the assembly machine capacity will be 8,000 hours/day and the unit earnings contribution of wrenches will be \$160/1,000 units. If we presume that the assembly machine capacity uncertainty and the wrenches earnings contribution uncertainty are independent, then the probability that the first state of the world will transpire is simply

$$0.25 = 0.50 \times 0.50 ,$$

because there is a 50% probability that the assembly machine capacity will be 8,000 hours/day and a 50% probability that the unit earnings contribution of wrenches will be \$160. The probabilities of each of the four possible states of the world are shown in the fourth column of Table 4.

We then proceed by creating a decision variable for each of next quarter’s decisions, for each possible state of the world. We therefore define:

State of the World	Assembly Machine Capacity	Unit Earnings Contribution of Wrenches	Probability
1	8,000 hours/day	\$160/1,000 units	0.25
2	10,000 hours/day	\$160/1,000 units	0.25
3	8,000 hours/day	\$90/1,000 units	0.25
4	10,000 hours/day	\$90/1,000 units	0.25

Table 4: The four possible states of the world for next quarter.

W_1 = the number of wrenches per day to produce next quarter, in 1,000s,
= if state – of – the – world 1 transpires,

P_1 = the number of pliers per day to produce next quarter, in 1,000s,
= if state – of – the – world 1 transpires,

W_2 = the number of wrenches per day to produce next quarter, in 1,000s,
= if state – of – the – world 2 transpires,

P_2 = the number of pliers per day to produce next quarter, in 1,000s,
= if state – of – the – world 2 transpires,

W_3 = the number of wrenches per day to produce next quarter, in 1,000s,
= if state – of – the – world 3 transpires,

P_3 = the number of pliers per day to produce next quarter, in 1,000s,
= if state – of – the – world 3 transpires,

W_4 = the number of wrenches per day to produce next quarter, in 1,000s,
= if state – of – the – world 4 transpires,

P_4 = the number of pliers per day to produce next quarter, in 1,000s,
= if state – of – the – world 4 transpires.

For example, the interpretation of P_2 is that P_2 is the quantity of pliers that GTC will produce next quarter if state-of-the-world 2 transpires, that

is, if assembly machine capacity is 10,000 hours and the wrench unit earnings contribution is \$160.

We are now ready to construct the linear optimization model of the steel supply planning problem. The objective will be to maximize the expected contribution to earnings, over all possible states of the world that might transpire next quarter. The expression for the objective function is:

$$\begin{aligned}\text{Objective} &= 0.25 \cdot (160W_1 + 100P_1) + 0.25 \cdot (160W_2 + 100P_2) \\ &= +0.25 \cdot (90W_3 + 100P_3) + 0.25 \cdot (90W_4 + 100P_4) - 58.00S.\end{aligned}$$

The constraints of the model will be the steel availability, assembly and molding machine capacity, and demand constraints, for each possible state of the world that might transpire next quarter. The resulting linear optimization model is as follows:

$$\begin{aligned}
&\text{maximize} && 0.25 \cdot (160W_1 + 100P_1) + 0.25 \cdot (160W_2 + 100P_2) + \\
&&& 0.25 \cdot (90W_3 + 100P_3) + 0.25 \cdot (90W_4 + 100P_4) - 58.00S \\
&\text{subject to :} && \\
&\text{Steel1 :} && 1.5W_1 + 1.0P_1 - S \leq 0 \\
&\text{Molding1 :} && 1.0W_1 + 1.0P_1 \leq 21 \\
&\text{Assembly1 :} && 0.3W_1 + 0.5P_1 \leq 8 \\
&\text{W - demand1 :} && W_1 \leq 15 \\
&\text{P - demand1 :} && P_1 \leq 16 \\
&\text{Steel2 :} && 1.5W_2 + 1.0P_2 - S \leq 0 \\
&\text{Molding2 :} && 1.0W_2 + 1.0P_2 \leq 21 \\
&\text{Assembly2 :} && 0.3W_2 + 0.5P_2 \leq 10 \\
&\text{W - demand2 :} && W_2 \leq 15 \\
&\text{P - demand2 :} && P_2 \leq 16 \\
&\text{Steel3 :} && 1.5W_3 + 1.0P_3 - S \leq 0 \\
&\text{Molding3 :} && 1.0W_3 + 1.0P_3 \leq 21 \\
&\text{Assembly3 :} && 0.3W_3 + 0.5P_3 \leq 8 \\
&\text{W - demand3 :} && W_3 \leq 15 \\
&\text{P - demand3 :} && P_3 \leq 16 \\
&\text{Steel4 :} && 1.5W_4 + 1.0P_4 - S \leq 0 \\
&\text{Molding4 :} && 1.0W_4 + 1.0P_4 \leq 21 \\
&\text{Assembly4 :} && 0.3W_4 + 0.5P_4 \leq 10 \\
&\text{W - demand4 :} && W_4 \leq 15 \\
&\text{P - demand4 :} && P_4 \leq 16 \\
&\text{Nonnegativity :} && S, W_1, P_1, W_2, P_2, W_3, P_3, W_4, P_4 \geq 0.
\end{aligned}$$

This linear optimization model is called a **two-stage linear optimization model under uncertainty**, or more simply a **two-stage model**. This is because the model is constructed based on there being two time-stages (today and next quarter), and because there is uncertainty about the data for stage-two (the assembly machine capacity next quarter will be either 8,000 hours/day or 10,000 hours per day next quarter, and the wrenches unit earnings contribution will be either \$160/1,000 units or \$90/1,000 units

next quarter).

4.3 Observations on the Two-Stage Model

Let us make several observations about the two-stage model that we have just constructed. First, notice in this linear optimization model that the objective function consists of the expected contribution to earnings from the daily production of wrenches and pliers in each state of the world, minus the cost of steel. Secondly, for each state of the world, we have our usual constraints on steel utilization, molding machine capacity and assembly machine capacity, and demand for wrenches and pliers. However, the model uses different values of assembly machine capacity (either 8,000 or 10,000 hours/day) corresponding to the different possible states of the world, consistent with the description of the four different states of the world in Table 4. Similarly, the model uses different unit earnings contributions of wrenches (either \$160 or \$90) corresponding to the different possible states of the world, also consistent the description of the four states of the world in Table 4.

Notice as well that the model expresses the constraint that GTC cannot use more steel than it has contracted for delivery, in the four constraints:

$$\begin{aligned}1.5W_1 + 1.0P_1 - S &\leq 0 \\1.5W_2 + 1.0P_2 - S &\leq 0 \\1.5W_3 + 1.0P_3 - S &\leq 0 \\1.5W_4 + 1.0P_4 - S &\leq 0.\end{aligned}$$

These four constraints state that regardless of which state of the world will transpire next quarter, GTC cannot utilize more steel next quarter than they have contracted for.

4.4 Interpreting the Solution of the Two-Stage Model

The optimal solution of the two-stage model is shown in Table 5. According to Table 5, the optimal value of S is $S = 27.25$. This means that GTC should contract today for the purchase of 27,250 lb./day of steel for next

quarter. In order to interpret the optimal solution values of the decision variables $W_1, P_1, W_2, P_2, W_3, P_3, W_4,$ and P_4 , let us re-organize the optimal values of these eight decision variables into the format shown in Table 6.

Decision Variable	Optimal Solution Value
S	27.25
W_1	15.00
P_1	4.75
W_2	15.00
P_2	4.75
W_3	12.50
P_3	8.50
W_4	5.00
P_4	16.00

Table 5: The optimal solution of the linear optimization model of the GTC steel supply planning problem.

State of the world Next Quarter	Production of Wrenches (units/day)		Production of Pliers (units/day)	
	Decision Variable	Value	Decision Variable	Value
1	W_1	15,000	P_1	4,750
2	W_2	15,000	P_2	4,750
3	W_3	12,500	P_3	8,500
4	W_4	5,000	P_4	16,000

Table 6: The optimal production plan for next quarter for the GTC steel supply planning problem.

We can interpret the optimal solution values in Table 5 and Table 6 as follows:

- Today, GTC should contract for 27,250 lb./day of steel for next quarter.
- Next quarter, if assembly hour availability is 8,000 hours/day and the contribution of wrenches is \$160, then GTC should produce 15,000 wrenches per day and 4,750 pliers per day.
- Next quarter, if assembly hour availability is 10,000 hours/day and the contribution of wrenches is \$160, then GTC should produce 15,000 wrenches per day and 4,750 pliers per day.
- Next quarter, if assembly hour availability is 8,000 hours/day and the contribution of wrenches is \$90, then GTC should produce 12,500 wrenches per day and 8,500 pliers per day.
- Next quarter, if assembly hour availability is 10,000 hours/day and the contribution of wrenches is \$90, then GTC should produce 5,000 wrenches per day and 16,000 pliers per day.

4.5 Flexibility of the Two-Stage Linear Optimization Modeling Paradigm

The modeling framework for a two-stage linear optimization under uncertainty allows considerable flexibility in modeling uncertainty. Here we indicate how we can model a variety of different issues that might arise in this context:

Modeling different probabilities. We could have modeled different probabilities for different possible states of the world. For example, suppose that we presume that the following probabilities hold for the problem:

$$\begin{aligned} P(\text{assembly machine hours} = 8,000) &= 0.8, \\ P(\text{assembly machine hours} = 10,000) &= 0.2 \end{aligned}$$

and

$$\begin{aligned} P(\text{wrench contribution} = \$160) &= 0.7, \\ P(\text{wrench contribution} = \$90) &= 0.3. \end{aligned}$$

State of the World	Assembly Machine Capacity	Unit Earnings Contribution of Wrenches	Probability
1	8,000 hours/day	\$160/1,000 units	$0.56 = 0.8 \times 0.7$
2	10,000 hours/day	\$160/1,000 units	$0.14 = 0.2 \times 0.7$
3	8,000 hours/day	\$90/1,000 units	$0.24 = 0.8 \times 0.3$
4	10,000 hours/day	\$90/1,000 units	$0.06 = 0.2 \times 0.3$

Table 7: The four possible states of the world for next quarter, with different probabilities of transpiring.

Under these scenarios, the states of the world and their associated probabilities would be as shown in Table 7.

Modeling different numbers of states of the world. Suppose that there are seven different assembly machine capacity levels that might transpire next quarter, and that there are six different wrench unit earnings contributions levels that might transpire next quarter. Then we would have $42 = 7 \times 6$ possible states of the world, and would need $1 + 42 \times 2 = 85$ decision variables, and $42 \times 5 = 210$ constraints in the model. Therefore, the number of distinct states of the world can increase the size of the model quite a lot.

Modeling different numbers of stages. We might want to model more than two stages: today, next quarter, the next quarter after that, etc. The same modeling principles illustrated here would then apply, but the resulting linear optimization model can become much more complicated, as well as much larger.

Block ladder structure. Notice that the model formulation has the block ladder structure depicted in Figure 1. There are special techniques that make use of this kind of structure to reduce the time and storage needed to solve large-scale problems of this type. One of these techniques is Benders' decomposition method.

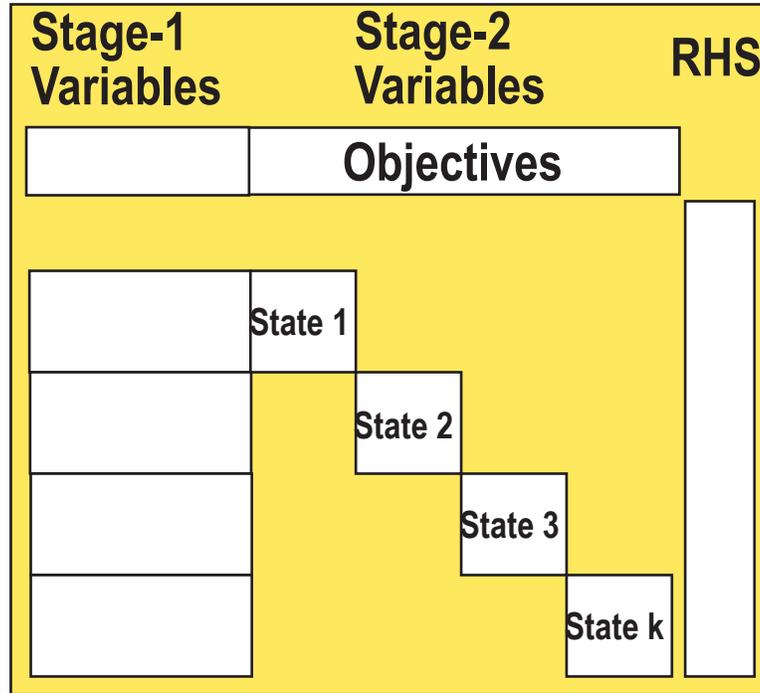


Figure 1: Block ladder structure of two-stage stochastic linear optimization.

4.6 Summary of the Method for Constructing a Two-Stage Linear Optimization Model under Uncertainty

Although we have presented the two-stage linear optimization modeling technique in the context of a simple example, the methodology applies broadly for modeling linear optimization problems under uncertainty. Here we summarize the main steps in constructing a two-stage linear optimization model under uncertainty.

Procedure for Constructing a Two-Stage Linear Optimization Model under Uncertainty

1. Determine which decisions need to be made in stage-one (today), and which decisions need to be in stage-two (next period).

2. Enumerate the possible states of the world that might transpire next period, what the data will be in each possible state of the world, and what is the probability of each state of the world occurring.
3. **Creating the decision variables:** Create one decision variable for each decision that must be made in stage-one. Create one decision variable for each decision that must be made in stage-two, for each possible state of the world.
4. **Constraints:** Create the necessary constraints for each possible state of the world.
5. **Objective function:** Account for the contribution of each of today's decisions in the objective function. Account for the expected value of the objective function contribution of each of next period's possible states of the world.

In order to use two-stage models effectively, we must have a reasonably accurate estimate of the probabilities of the future states of the world. Also, in order to keep the size of the model from becoming too large, it is important to limit the description of the different possible future states of the world to a reasonably low number.

There is great modeling power in two-stage linear optimization under uncertainty. Indeed, to the extent that the most important decisions that we need to make are concerned with optimally choosing actions today in the face of uncertainty about tomorrow, then the two-stage modeling framework is a core modeling tool.

5 Two-Stage Linear Optimization in Matrix Form

There are two sets of decisions, one for each of the two consecutive stages.

The first-stage variables are x and are subject to constraints:

$$Ax = b \quad , \quad x \geq 0 \quad .$$

The direct contribution of the x decisions on the objective function is $c^T x$ for some vector c .

The second-stage variables are y and are subject to constraints:

$$Bx + Dy = d \quad , \quad y \geq 0 .$$

The direct contribution of the y decisions on the objective function is $f^T y$ for some vector f .

If there were no uncertainty, then the problem would look like:

$$\begin{aligned} \text{minimize}_{x,y} \quad & c^T x + f^T y \\ \text{s.t.} \quad & Ax = b \\ & Bx + Dy = d \\ & x \geq 0 \quad y \geq 0 \end{aligned}$$

Under uncertainty, we assume that the values of the data B, D, d, f are uncertain. There are $\omega = 1, \dots, K$ possible future scenarios, with scenario ω having a probability α_ω of being realized, for $\omega = 1, \dots, K$.

The data B, D, d, f takes on values $B_\omega, D_\omega, d_\omega, f_\omega$ with probability α_ω for $\omega = 1, \dots, K$. We only learn the values of this data *after* we have made our first-stage decisions x . Once the values of B, D, d, f are known, we then make our second-stage decisions y .

Let y_ω denote the decisions y under the condition that scenario ω is realized, for $\omega = 1, \dots, K$.

The objective is to choose x and $y_\omega, \omega = 1, \dots, K$, so as to solve the following optimization model:

$$\begin{array}{llll}
\text{minimize} & c^T x + \alpha_1 f_1^T y_1 + \alpha_2 f_2^T y_2 + \cdots + \alpha_K f_K^T y_K & & \\
x, y_1, \dots, y_K & \text{s.t.} & Ax & = b \\
& & B_1 x + D_1 y_1 & = d_1 \\
& & B_2 x + D_2 y_2 & = d_2 \\
& & \vdots & \vdots \\
& & B_K x + D_K y_K & = d_K \\
& & x, y_1, y_2, \dots, y_K & \geq 0
\end{array}$$

Notice that this problem has the block-ladder structure shown in Figure 1.

6 Extensions of the Two-Stage Model

- Multi-stage Stochastic Programs
- Nonlinear Stochastic Programs
- Integer Stochastic Programs

7 A Powerplant Investment Problem

The newly unified nation of Timoria must invest in a system of power plants to meet its current and future demand for electrical power. These plants are to be built for the first year only, and are expected to operate over the next 15 years. The budget for construction of power plants is \$10 billion, which is to be allocated for four different types of plants: gas turbine, coal, nuclear power, and hydroelectric. The objective is to find the power plant allocation which minimizes the sum of the investment cost and the expected value of the operating cost over 15 years. The operating cost is stochastic

due to uncertainty in future demand as well as fuel prices. The model used for this problem is based on [1], which explores the power plant investment problem in a non-stochastic framework.

Power plants are priced according to their electric capacity, measured in gigawatts (GW). Table 8 shows the investment cost per GW of capacity for each type of plant.

Plant	Cost per GW capacity
Gas Turbine	\$110 million
Coal	\$180 million
Nuclear	\$450 million
Hydroelectric	\$950 million

Table 8: Investment cost per GW of capacity.

Since hydroelectric energy depends on the availability of rivers which may be dammed, the geography of the country constrains the hydroelectric power capacity. In the case of Timoria, no more than 5.0 GW of power may be produced by hydroelectric plants.

Demand for electric power is typically described using a *load duration curve*. Figure 2 shows the projected demand for power during the first year. The continuous curve is approximated by a quantized set of demands and durations, also shown in this figure. The quantized demand data is shown in Table 9.

Demand Block	Demand (GW)	Duration (hours)
#1	10.0	490
#2	8.4	730
#3	6.7	2,190
#4	5.4	3,260
#5	4.3	2,090

Table 9: Power demand in the first year.

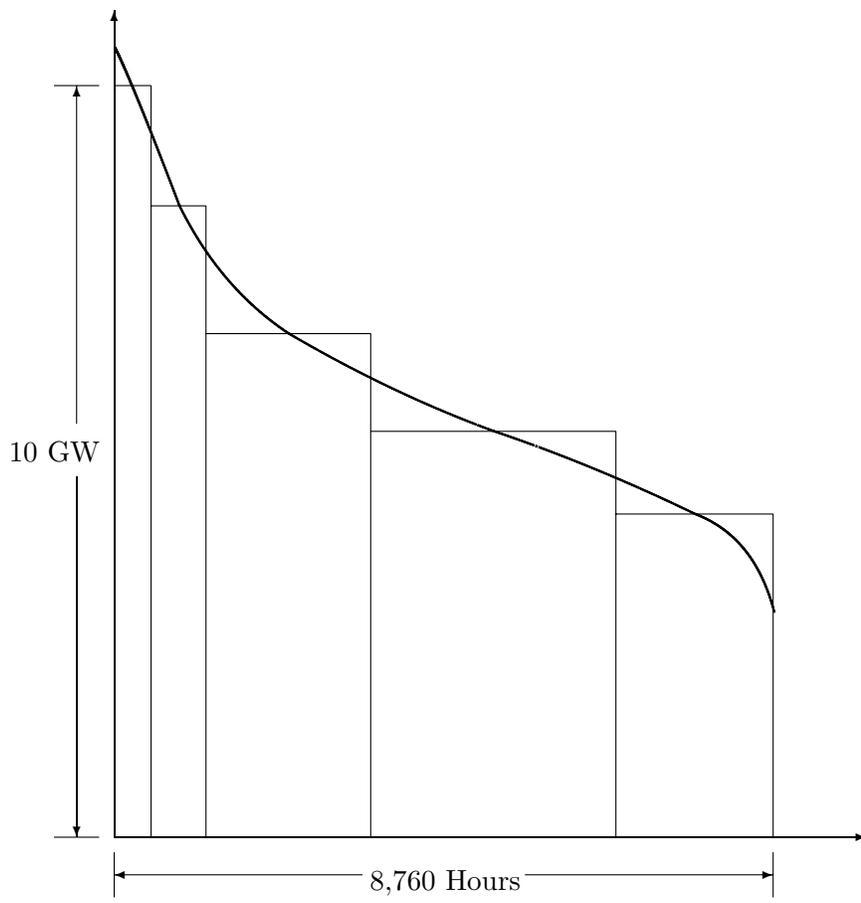


Figure 2: Projected load duration curve for the first year of demand.

As the country becomes further economically developed, the demand for electrical power is expected to grow each year. The value of the growth rate over the next 15 years is unknown, but economists have predicted the yearly growth rates shown in Table 10, with their corresponding probabilities. From this table, it is noted that the expected value of the growth rate is 3%.

Growth	Probability
-1%	20%
1%	20%
3%	20%
5%	20%
7%	20%

Table 10: Projected yearly growth in power demand

Along with the investment cost, the cost of operating the proposed power plants over the next 15 years must also be considered. Any demand which cannot be satisfied by the existing plants must be met by purchasing power from a neighboring country at a substantial cost. The operating costs of each type of power plant, as well as the cost of purchasing power from an external source, are shown in Table 11, where the units are in cents per kilowatt-hour (KWH). Due to fluctuations in coal and natural gas prices, the actual operating costs of the gas turbines and coal plants are not known in advance. The expected values of these costs are shown in Table 11. The probability distributions of these costs are shown in Table 12 and Table 13.

8 Stochastic Programming Formulation

The problem presented above can be modelled as a two-stage stochastic linear optimization model. In the first stage, we must determine the amount of capacity of each of the four types of powerplants to build. We must make these decisions first, before we know what the growth rate in electricity demand will be, and before we know what the prices of natural gas and coal will be. The first-stage decision variable is represented by a four-dimensional

Plant	Cost per KWH
Gas Turbine*	3.92 ¢
Coal*	2.44 ¢
Nuclear	1.40 ¢
Hydroelectric	0.40 ¢
External Source	15.0 ¢

Table 11: Operating cost of power generation. *Gas turbine and coal plant operating costs are expected values.

Cost per KWH	Probability
3.1 ¢	10%
3.3 ¢	20%
3.9 ¢	40%
4.5 ¢	20%
4.9 ¢	10%

Table 12: Probability distribution of gas turbine operating costs.

vector $x = (x_1, x_2, x_3, x_4)$, representing the gigawatts of capacity to be built for each type of plant. For consistency, the plant types are always assumed to have the same order (1. Gas turbine, 2. Coal, 3. Nuclear, 4. Hydroelectric), and the subscript notation x_1, x_2, x_3, x_4 references these types respectively. The cost vector c is also a four-dimensional vector representing the investment costs shown in Table 8.

The second-stage decisions are the amount of electricity capacity used to produce electricity at each power plant for each demand block in each year. Specifically, let y_{ijk} denote the amount of electricity capacity used to produce electricity by power plant type i for demand block j in year k , for $i = 1, \dots, 5$, $j = 1, \dots, 5$, and $k = 1, \dots, 15$. Here y_{5jk} is the amount of electricity capacity purchased from the external source. The units of the y_{ijk} variables are in GW. The five demand blocks are as shown in the Table 9. To illustrate, the variable y_{312} represents the amount of nuclear power used at peak demand time during the second year.

Cost per KWH	Probability
1.7 ¢	10%
2.1 ¢	20%
2.4 ¢	40%
2.9 ¢	20%
3.1 ¢	10%

Table 13: Probability distribution of coal plant operating costs.

It is evident that the optimal value of the second-stage variables depends on the stochastic problem data. Each possible value of the problem data is referred to as a *scenario*, which will be indexed by ω . In actuality, there are an infinite number of possible future scenarios. However, in the interests of tractability, only a finite number of such scenarios are considered here. For the power plant investment problem, there are five different possible demand growth rates (see Table 10), five possible operating costs for gas turbines (see Table 12) and five possible operating costs for coal plants (see Table 13), for a total of $K = 5 \times 5 \times 5 = 125$ scenarios. We therefore can think of ω as an index taking on values in the range $\omega = 1, \dots, 125$. The second-stage variable is then written as a function of the scenario, as $y_{ijk\omega}$.

The operating cost for the various electricity sources, in cents/KWH, are represented by the scalars $f_i(\omega)$ (see Table 11, Table 12, and Table 13) for $i = 1, \dots, 5$ and $\omega = 1, \dots, 125$. The duration in hours of each demand block is the scalar h_j for $j = 1, \dots, 5$ (see the third column in Table 9). The total expected cost (in \$ million) is then written as:

$$\sum_{i=1}^4 c_i x_i + E \left(\sum_{i=1}^5 \sum_{j=1}^5 \sum_{k=1}^{15} (10^6 \text{ KW/GW}) \cdot (10^{-8} \text{ million } \$/\text{¢}) \cdot (f_i(\omega) \text{ ¢/KWH}) \cdot (h_j \text{ hours}) \cdot (y_{ijk\omega} \text{ GW}) \right)$$

where E is the expectation operator, taken over all possible scenarios ω .

The stochastic programming formulation of the problem is then written

as:

$$\begin{aligned}
\min_{x,y} \quad & \sum_{i=1}^4 c_i x_i + \sum_{\omega=1}^{125} \alpha_{\omega} \sum_{i=1}^5 \sum_{j=1}^5 \sum_{k=1}^{15} 0.01 f_i(\omega) h_j y_{ijk\omega} \\
\text{s.t.} \quad & \sum_{i=1}^4 c_i x_i \leq 10,000 \quad (\text{Budget constraint}) \\
& x_4 \leq 5.0 \quad (\text{Hydroelectric constraint}) \\
& y_{ijk\omega} \leq x_i \quad \text{for } i = 1, \dots, 4, \text{ all } j, k, \omega \quad (\text{Capacity constraints}) \\
& \sum_{i=1}^5 y_{ijk\omega} \geq D_{jk\omega} \quad \text{for all } j, k, \omega \quad (\text{Demand constraints}) \\
& x \geq 0, \quad y \geq 0
\end{aligned} \tag{1}$$

where α_{ω} is the probability of scenario ω , and $D_{jk\omega}$ is the power demand in block j and year k under scenario ω .

This problem is a standard linear program, although its size may present some difficulties for some LP solvers. There are 4 first-stage variables and $5 \times 5 \times 15 \times 125$ second-stage variables, for a total of 46,879 variables. Furthermore, there is one budget constraint, one hydroelectric constraint, $4 \times 5 \times 15 \times 125$ capacity constraints and $5 \times 15 \times 125$ demand constraints for a total of 46,877 constraints. A problem of this size is not unrealistic for many of today's computers and LP solvers. However, the addition of many more scenarios can quickly make this LP too unwieldy for even the most powerful computers currently available. For this reason, the problem is typically solved by Benders' decomposition, which exploits the problem structure by decomposing it into a sequence of smaller problems.

The optimal expected cost of the power plant investment problem is turns out to be \$16.933 billion. The optimal capacity construction decisions are shown in the middle column of Table 14. The right-most column of Table 14 shows what the capacity construction decisions would have been if instead of using stochastic data, we instead had used expected demand data and expected cost data. In this case, the construction decision would yield total expected costs of \$17.794 billion, which is 5.1% higher than the optimal expected cost.

Plant	Optimal Construction Decision based on stochastic data	Optimal Construction Decision based on expected demand and costs
Gas Turbine	4.66 GW	1.92 GW
Coal	4.57 GW	3.33 GW
Nuclear	4.68 GW	4.0 GW
Hydroelectric	5.0 GW	5.0 GW
Expected Cost	\$16.933 billion	\$17.794 billion

Table 14: Power plant capacity construction decisions.

Table 15 shows a subset of the optimal second-stage variables, since it would not be practical to show all 46,875 variables. The table shows the actual power to be supplied by each plant during the 15th year, given the optimal capacity construction decisions were made, under a 7% growth rate scenario, 3.9¢ gas turbine operating cost, and 2.4¢ coal operating cost. The total operating cost for this year and scenario is \$2.81 billion for the optimal capacity allocation. Similarly, Table 16 shows the actual power to be supplied by each plant during the 15th year in the case when the capacity construction decisions were based on expected demand and costs, under a 7% growth rate scenario, 3.9¢ gas turbine operating cost, and 2.4¢ coal operating cost. The total operating cost for this year and scenario is \$4.41 billion with the sub-optimal capacity construction decisions.

9 Exercises

You should download the files `bender1.osc`, `eci_bender1_opl_master.mod`, `eci_bender1_opl_sub.mod`, and `eci_opl.dat` from the course website. These files are an OPLScript based model of the powerplant planning problem discussed in the lecture.

- The file `bender1.osc` is the script file that runs the problem.
- The file `eci_bender1_opl_master.mod` is the model file for the master problem.

Plant	Demand Block				
	#1	#2	#3	#4	#5
Gas Turbine	4.66	4.66	3.02	0.00	0.00
Coal	4.57	4.57	4.57	4.24	1.40
Nuclear	4.68	4.68	4.68	4.68	4.68
Hydroelectric	5.00	5.00	5.00	5.00	5.00
External Source	6.87	2.74	0.00	0.00	0.00
Cost (million \$)	689.6	575.2	685.4	610.5	249.0

Table 15: GW of power during 15th year, given the optimal capacity construction decisions, assuming a 7% growth rate, 3.9¢ gas turbine and 2.4¢ coal operating cost.

- The file `eci_bender1_opl_sub.mod` is the model file for the sub problem.
 - The file `eci_opl.dat` is the data file for the problem.
1. The current model uses a five-point probability distribution for the gas turbine operating costs as well as for the coal plant operating costs. Modify the model and/or data file so that the stochastic program instead uses only the one-point average values for the gas turbine and coal plant operating costs. How does use of stochastic operating costs versus average operating costs impact the capacity allocation? Can you explain why?
 2. The first column of Table 17 shows the probability distribution of the economic growth rates in the current version of the model. The second, third, fourth, and fifth columns portray ever-cruder approximations of this distribution. Compute the optimal capacity allocation for each one of these distributions. Using these allocations, compute the expected costs assuming that the actual growth rate distribution is the original one given in the first column of Table 17. Compare these costs to the optimal expected cost. What do you observe?
 3. List the different ways this model can be made more realistic (e.g., suppose the power plants are not all built the same year).

Plant	Demand Block				
	#1	#2	#3	#4	#5
Gas Turbine	1.92	1.92	1.92	1.59	0.00
Coal	3.33	3.33	3.33	3.33	2.08
Nuclear	4.00	4.00	4.00	4.00	4.00
Hydroelectric	5.00	5.00	5.00	5.00	5.00
External Source	11.53	7.40	3.02	0.00	0.00
Cost (million \$)	960.5	978.8	1497.5	710.5	263.2

Table 16: GW of power during 15th year, given the capacity construction decisions based on expected demand and costs, assuming a 7% growth rate, 3.9¢ gas turbine and 2.4¢ coal operating cost.

#1		#2		#3		#4		#5	
Growth	Prob.								
-1%	20%	-0.6%	25%	-0.2%	33.3%	0.6%	50%	3%	100%
1%	20%	1.8%	25%	3%	33.3%	5.4%	50%		
3%	20%	4.2%	25%	6.2%	33.3%				
5%	20%	6.6%	25%						
7%	20%								

Table 17: Original and ever-cruder growth rate distributions.

References

- [1] D. Anderson. Models for determining least-cost investments in electricity supply. *The Bell Journal of Economics*, 3:267–299, 1972.
- [2] D. Bertsimas and J. Tsitisklis. *Introduction to Linear Optimization*. Athena Scientific, 1997.
- [3] J. Birge and F. Louveaux. *Introduction to Stochastic Programming*. Springer-Verlag, 1997.