Introduction to Model Building

Taek-Ho Lee
Department of Industrial Engineering, SeoulTech
Mail: taekho.lee@seoultech.ac.kr

Contents

Introduction to Modeling

- Operations Research (management science) is a scientific approach to decision making that seeks to best design and operate a system, usually under conditions requiring the allocation of scarce resources.

- A system is an organization of interdependent components that work together to accomplish the goal of the system.

Introduction to Modeling

- The term <u>operations research</u> was first introduced during World War II when leaders asked scientists and engineers <u>to analyze several military problems.</u>

- The scientific approach to decision making requires the use of one or more mathematical models. A <u>mathematical model</u> is a <u>mathematical representation of the actual situation</u> that may be used to make better decisions or clarify the situation.

Introduction to Modeling

- Modeling perspective in a military case
  - Radar deployment and defense
    - ✓ Background: To defend against the threat of the German Air Force, radar technology was developed, but it was uncertain where and how many radar stations should be installed to most effectively detect enemy aircraft.
    - ✓ Objective: Detect enemy aircraft as quickly as possible to maximize defense efficiency.
    - ✓ Constraints
      - Possible radar station locations (installable for some terrain or coastline)
      - Limited resources: number of radar stations available, personnel, installation cost, etc.
    - ✓ Decision variables
      - Where to install radar stations and how many to install

Introduction to Modeling

- Modeling perspective in a military case
  - The Atlantic convoy operation
    - ✓ Background: The allied forces had to transport weapons and supplies from the United States to the United Kingdom. However, many cargo ships were attacked from German submarines. When individual ships sailed independently, they were attacked as soon as they were detected. By grouping multiple ships into convoys, escort ships could be deployed more efficiently, but if the convoy became too large, its speed decreased, causing delays in arrival.
    - ✓ Objective: Minimize the loss of cargo ships while satisfying the required delivery time of supplies.
    - ✓ Constraints
      - The number of escort ships and cargo ships available.
      - Speed limitations depending on convoy size
      - Strict deadlines for supply delivery (no delays allowed)
    - ✓ Decision variables: Convoy size, sailing routes, sailing speed, and escort deployment

Introduction to Modeling

- Modeling perspective in a military case
  - Bomber formations and route selection
    - ✓ Background: Allied bombers aim to strike targets in Germany. They were exposed to attacks from enemy anti-aircraft artillery and planes. Operating individually reduced the probability of detection but made them more vulnerable. Flying in formations increased their defense but also raised detection probability. The level of damage also varied depending on the chosen route.
    - ✓ Objective: Maximize bombing success rates while minimizing the probability of being shot.
    - ✓ Constraints
      - Range limitations due to fuel capacity (limits on detour)
      - Mission time restrictions
      - Weather conditions
    - ✓ Decision variables: Whether to maintain formations, choice of route, and flight altitude.

Introduction to Modeling

- A modeling example
  - Eli Daisy produces the drug Wozac in huge batches by heating a chemical mixture in a pressurized container. Each time a batch is produced, a different amount of Wozac is produced. The amount produced is the process *yield*.

  - Daisy is interested in understanding the factors that influence the yield of Wozac production process.

  - The solution on subsequent slides describes a model building process for this situation.

Introduction to Modeling

- A modeling example
  - Daisy is interested in determining the factors that influence the process yield. This would be referred to as a descriptive model since it describes the behavior of the actual yield as a function of various factors.

  - Daisy might determine the following factors influence yield:
    - ✓ Container volume in liters (V)
    - ✓ Container pressure in milliliters (P)
    - ✓ Container pressure in degrees centigrade (T)
    - ✓ Chemical composition of the processed mixture

Introduction to Modeling

- A modeling example
  - Letting A, B, and C be the percentage of the mixture made up of chemical A, B, and C, then Daisy might find, for example, that
  - $Yield\ =\ 300\ +\ 0.8 \cdot V\ +\ 0.01 \cdot P\ +\ 0.06 \cdot T\ +\ 0.001 \cdot T \cdot P - 0.01 \cdot T^2 - 0.001 \cdot P^2 + 11.7 \cdot A + 9.4 \cdot B + 16.4 \cdot C + 19 \cdot A \cdot B + 11.4 \cdot A \cdot C - 9.6 \cdot B \cdot C$

Introduction to Modeling

- A modeling example
  - To determine this relationship, the yield of the process would have to be measured for many different combinations of the previously listed factors. Knowledge of this equation would enable Daisy to describe the yield of the production process once volume, pressure, temperature, and chemical composition were known.

Introduction to Modeling

- Prescriptive or Optimization Models
  - Prescriptive models “prescribes” behavior for an organization that will enable it to best meet its goals. Components of this model include:
    - ✓ Objective function(s)
    - ✓ Decision variables
    - ✓ Constraints
  - An optimization model seeks to find values of the decision variables that optimize (maximize or minimize) an objective function among the set of all values for the decision variables that satisfy the given constraints.

Introduction to Modeling

- A modeling example – the objective function
  - The Daisy example seeks maximize the yield for the production process. In most models, there will be function we wish to maximize or minimize. This function is called the model's objective function. To maximize the process yield we need to find the values of V, P, T, A, B, and C that make the yield equation (below) as large as possible.
  - $Yield = 300 + 0.8 \cdot V + 0.01 \cdot P + 0.06 \cdot T + 0.001 \cdot T \cdot P - 0.01 \cdot T^2 - 0.001 \cdot P^2 + 11.7 \cdot A + 9.4 \cdot B + 16.4 \cdot C + 19 \cdot A \cdot B + 11.4 \cdot A \cdot C - 9.6 \cdot B \cdot C$

Introduction to Modeling

- In many situations, an organization may have more than one objective.
- For example, in assigning students to the two high schools in Bloomington, Indiana, the Monroe Count School Board stated that the assignment of students involve the following objectives:
  - Equalize the number of students at the two high schools
  - Minimize the average distance students travel to school
  - Have a diverse students body at both high schools

Introduction to Modeling

- The decision variables
  - Variables whose values are under our control and influence system performance are called decision variables. In the Daisy example, V, P, T, A, B, and C are decision variables.

- Constraints
  - In most situations, only certain values of the decision variables are possible. For example, certain volume, pressure, and temperature conditions might be unsafe. Also, A, B, and C must be non-negative numbers that sum to one. These restrictions on the decision variable values are called constraints.

Introduction to Modeling

- The modeling example
  - Suppose the Daisy example has the following constraints:
    - ✓ Volume must be between 1 and 5 liters
    - ✓ Pressure must be between 200 and 400 milliliters
    - ✓ Temperature must be between 100 and 200 degrees centigrade
    - ✓ Mixture must be made up entirely of A, B, and C.
    - ✓ For the drug to perform properly, only half the mixture at most can be product A.

Introduction to Modeling

- The modeling example
  - Mathematically these constraints can be expressed as:
    - ✓ $V \leq 5$ and $V \geq 1$
    - ✓ $P \leq 400$ and $P \geq 200$
    - ✓ $T \leq 200$ and $T \geq 100$
    - ✓ $A \geq 0, \mathrm{B} \geq 0, \mathrm{C} \geq 0$
    - ✓ $A + B + C = 1.0$
    - ✓ $A \leq 0.5$

Introduction to Modeling

- The modeling example
  - The complete Daisy optimization model
    - ✓ Letting z represents the value of the objective function (the yield), the entire optimization model may be written as:
      - Maximize $z = 300 + 0.8 \cdot V + 0.01 \cdot P + 0.06 \cdot T + 0.001 \cdot T \cdot P - 0.01 \cdot T^2 - 0.001 \cdot P^2 + 11.7 \cdot A + 9.4 \cdot B + 16.4 \cdot C + 19 \cdot A \cdot B + 11.4 \cdot A \cdot C - 9.6 \cdot B \cdot C$
      - Subject to
        - $V \le 5,\ \ V \ge 1,\ \ P \le 400,\ \ P \ge 200,\ \ T \le 200,\ \ T \ge 100$
        - $A \ge 0,\ \ B \ge 0,\ \ C \ge 0,\ \ A + B + C = 1.0,\ \ A \le 0.5$

Introduction to Modeling

- The modeling example
  - Any specification of the decision variables that satisfies all the model's constraints is said to be in the feasible region. For example, V=2, P=300, T=150, A=0.4, B=0.3, C=0.3 is in the feasible region.
  - An optimal solution to an optimization model is any point in the feasible region that optimizes (in this case, maximizes) the objective function.

Introduction to Modeling

- Static and Dynamic Models
  - A static model is one in which the decision variables do not involve sequences of decisions over multiple rounds.
  - A dynamic model is a model in which the decision variables do involve sequences of decisions over multiple periods.
  - In a static model, we solve a one-shot problem whose solutions are prescribe optimal values of the decision variables at all points in time. The Daisy problem is an example of a static model.

Introduction to Modeling

- Static and Dynamic Models
  - For a dynamic model, consider a company (SailCo) that must determine how to minimize the cost of meeting (on-time) the demand for sail boats it produces during the next year. SailCo must determine the number of sail boats to produce during each of the next four quarters. SailCo's decisions must be made over multiple periods and thus posses a dynamic model.

Introduction to Modeling

- Linear and non-linear models
  - Suppose that whenever decision variables appear in the objective function and in the constraints of an optimization model, the decision variables are always multiplied by constants and then added together. Such a model is a linear model.
  - The Daisy example is a nonlinear model. While the decision variables in the constraints are linear, the objective function is nonlinear since the objective terms are non-linear:
    - ✓ $0.001 \cdot T \cdot P, \quad -0.01 \cdot T^2, \quad -0.001 \cdot P^2, \quad 19 \cdot A \cdot B, \quad 11.4 \cdot A \cdot C, \quad -9.6 \cdot B \cdot C$
  - In general, nonlinear models are much harder to solve.

Introduction to Modeling

- Integer and non-integer models
  - If one or more the decision variables must be integer, then we say that an optimization model is an integer model. If all the decision variables are free to assume fractional values, then an optimization model is a non-integer model.

  - The Daisy example is a non-integer example since volume, pressure, temperature, and percentage compositions are all decision variables which may assume fractional values.

  - If decision variables in a model represent the number of workers starting during each shift, then we clearly have an integer model.
  - Integer models are much harder to solve then non-integer models.

Introduction to Modeling

- Deterministic and Stochastic Models
  - Suppose that, for any value of the decision variables, the value of the objective function and whether the constraints are satisfied or not is known with certainty. We then have a deterministic model. If this is not the case, then we have a stochastic model.

  - If we view the Daisy example as a deterministic model, then we are making the assumption that for given values of V, P, T, A, B, and C, the process yield will always be the same. Since this is unlikely, the objective function can be viewed as the average yield of the process given decision variables values.

The Seven-step Model-building Process

- Operations research used to solve an organization’s problem follows a seven-step model building procedure:

  - 1. Formulate the problem
    - ✓ Define the problem
    - ✓ Specify objectives
    - ✓ Determine parts of the organization to be studied

  - 2. Observe the system
    - ✓ Determine parameters affecting the problem
    - ✓ Collect data to estimate values of the parameters

The Seven-step Model-building Process

- Operations research used to solve an organization's problem follows a seven-step model building procedure:
  - 3. Formulate a mathematical model of the problem
  - 4. Verify the model and use the model for prediction
    - ✓ Does the model yield results for values of decision variables not used to develop the model?
    - ✓ What eventualities might cause the model to become invalid?
  - 5. Select a suitable alternative
    - ✓ Given a model and a set of alternative solutions, determine which solution best meets the organizations objectives.

The Seven-step Model-building Process

- Operations research used to solve an organization's problem follows a seven-step model building procedure:
  - 6. Present the results and conclusions of the study to the organization
    - ✓ Present the results to the decision makers
    - ✓ If necessary, prepare several alternative solutions and permit the organization to choose the one that best meets their need
    - ✓ Any non-approval of the study's recommendations may have stemmed from an incorrect problem definition or failure to involve the decision makers from the start of the project. In such a case, return to step 1, 2, or 3.

The Seven-step Model-building Process

- Operations research used to solve an organization's problem follows a seven-step model building procedure:
  - 7. Implement and evaluate recommendations
    - ✓ Upon acceptance of the study by the organization, the analyst:
      - Assists in implementing the recommendations
      - Monitors and dynamically updates the system as the environment and parameters change to ensure that recommendations enable the organization to meet its goals.

The CITGO Petroleum Problem

- Klingman et al. (1987) applied a variety management science techniques to CITGO Petroleum. Their work saved the company an estimated $70 per year. CITGO is an oil-refining and marketing company that was purchased by Southland Corporation (7-11 Stores).
- The following slides focus on two aspects of the CITGO's team's work:
  - A mathematical model to optimize the operation of CITGO's refineries.
  - A mathematical model – supply, distribution and marketing (SDM) – used to develop an 11-week supply, distribution and marketing plan for the entire business.

The CITGO Petroleum Problem

- Optimizing Refinery Operations
  - Step 1. (Formulate the Problem)
    Klingman et al. wanted to minimize the cost of CITGO's refineries.
  - Step 2. (Observe the System)
    The Lake Charles, Louisiana, refinery was closely observed in an attempt to estimate key relationships such as:
    - ✓ How the cost of producing each of CTIGO's products (motor fuel, no. 2 fuel oil, turbine fuel, naphtha, and several blended motor fuels) depends upon the inputs used to produce each product.
    - ✓ The amount of energy needed to produce each product (requiring the installation of a new metering system).

The CITGO Petroleum Problem

- Optimizing Refinery Operations
  - Step 2 (Continued)
    - ✓ The yield with each input-output combination. For example, if 1 gallon of crude oil would yield 0.52 gallons of motor fuel, then the yield would be 52%.
    - ✓ To reduce maintenance costs, data were collected on parts inventories and equipment breakdowns. Obtaining accurate data required the installation of a new data based-management system and integrated maintenance information system. Additionally, a process control system was also installed to accurately monitor the inputs and resources used to manufacture each product.

The CITGO Petroleum Problem

- Optimizing Refinery Operations
  - Step 3. (Formulate a Mathematical Model of the Problem)
    Using linear programming (LP), a model was developed to optimize refinery operations.

    ✓ The model:
      - Determines the cost-minimizing method for mixing blending together inputs to obtain desired results.
      - Contains constraints that ensure that inputs are blended to produce desired results.
      - Includes constraints that ensure inputs are blended so that each output is of the desired quality.
      - Other model constraints ensure plant capacities are not exceeded and allow for inventory for each end product.

The CITGO Petroleum Problem

- Optimizing Refinery Operations
  - Step 4. (Verify the Model and Use the Model for Prediction)
    To validate the model, inputs and outputs from the refinery were collected for one month. Given actual inputs used at the refinery during that month, actual outputs were compared those predicted by the model. After extensive changes, the model's predicted outputs were close to actual outputs.

The CITGO Petroleum Problem

- Optimizing Refinery Operations
  - Step 5. (Select Suitable Alternative Solutions)
    Running the LP yielded a daily strategy for running the refinery. For instance, the model might say, produce 400,000 gallons of turbine fuel using 300,000 gallons of crude 1 and 200,000 of crude 2.

The CITGO Petroleum Problem

- Optimizing Refinery Operations
  - Step 6. (Present the Results and Conclusions) and Step 7. (Implement and Evaluate Recommendations)

    Once the data base and process control were in place, the model was used to guide day-to-day operations. CITGO estimated the overall benefits of the refinery system exceeded $50 million annually.

The CITGO Petroleum Problem

- The Supply Distribution Marketing (SDM) System
  - Step 1. (Formulate the Problem)
    CITGO wanted a mathematical model that could be used to make supply, distribution, and marketing decisions such as
    - ✓ Where should the crude oil be purchased?
    - ✓ Where should products be sold?
    - ✓ What price should be charged for products?
    - ✓ How much of each product should be held in inventory?

The goal was to maximize profitability associated with these decisions.

The CITGO Petroleum Problem

- The Supply Distribution Marketing (SDM) System
  - Step 2. (Observe the System)
    A database that kept tack of sales, inventory, trades, and exchanges of all refined goods was installed. Also, regression analysis was used to develop forecasts of for wholesale prices and wholesale demand for each CITGO product.

  - Step 3. (Formulate a Mathematical Model of the Problem) and Step 5. (Select Suitable Alternative Solutions)
    A minimum-cost network flow model (MCNFM) is used to determine an 11-week supply, marketing and distribution strategy. The model makes all decisions discussed in Step 1. A typical model run involved 3,000 equations and 15,000 decision variables required only 30 seconds on an IBM 4381.

The CITGO Petroleum Problem

- The Supply Distribution Marketing (SDM) System
  - Step 4. (Verify the Model and Use the Model for Prediction)
    The forecasting models are continuously evaluated to ensure that they continue to give accurate forecasts.
  - Step 6. (Present the Results and Conclusions) and Step 6. (Implement and Evaluate Recommendations)
    Implementing the SDM required several organizational changes. A new vice-president was appointed to coordinate the operation of the SDM and refinery LP model. The product supply and product scheduling departments were combined to improve communications and information flow.

San Francisco Police Department Scheduling

- In the late 1970s, the San Francisco Police Department (SFPD) faced an operational challenge: <u>how to schedule patrol officers</u> efficiently across the city. The department had <u>a fixed number of officers</u> but <u>needed to cover a wide area</u> with <u>fluctuating demands for police services depending on time of day, day of week, and location</u>.
- At the time, SFPD was dealing with three major constraints:
  - Public safety and response time: citizens expected rapid response to calls, which means that <u>patrols had to be positioned strategically</u>.
  - Union and labor agreements: Officer schedules <u>had to respect shift lengths, days off, and other conditions</u>.
  - Budgetary limits: The city could <u>not simply hire more officers</u>; instead, it had to optimize deployment of the existing force.

San Francisco Police Department Scheduling

- Traditional scheduling was handled manually, relying on senior officers' judgement and historical patterns. This often led to mismatches between officer availability and actual call demand.
- To address these inefficiencies, researchers develop a mathematical model for police scheduling. The goal was to create a systematic, data-driven method that balanced service effectiveness with fairness and contractual obligations. This is one of the most notable early applications of OR in public sector workforce management.

San Francisco Police Department Scheduling

- Step 1. (Formulate Problem)
  - The SFPD needed an efficient way to schedule patrol officers across different areas. The existing manual scheduling was time-consuming, inconsistent, and often led to mismatches. The objective was to develop a system that could quickly generate weekly schedules that minimized both understaffing and overstaffing. The department also wants clear visualization of the schedules.
- Step 2. (Observe the System)
  - Data were collected on the actual demand for officers by time of day and day of the week. For example, on Sundays between 1 a.m. and 2 a.m., 38 officers might be required, while 4 a.m. only 14 officers would be needed. These demand curves reflected crime patterns, calls for service, traffic volume, and special events. This demand data served as the foundation for building the scheduling model.

San Francisco Police Department Scheduling

- Step 3. (Formulate a Mathematical Model)
  - An integer programming model was developed to determine schedules.
    - ✓ Decision variables specified when each officer’s shift should begin and where each officer should be deployed once on duty.
    - ✓ Constraints to ensure feasibility
      - Shift rules: officers could only work one shift at a time, with fixed length and mandatory days off.
      - Minimum coverages: each district had to maintain a required number of officers on duty at any given time.
      - Workforce limits: the total number of available officers could not be exceeded.
      - Additional rules such as contractual conditions (overtime restrictions, seniority rules, etc.)

San Francisco Police Department Scheduling

- Step 4. (Verify the Model and Use the Model for Prediction)
  - The model was tested against historical data to check whether its schedules matched actual demand patterns. By comparing predicted staffing levels with real call volumes, the department confirmed that the model produced realistic and reliable results.
- Step 5. (Select Suitable Alternative Solutions)
  - Several alternative schedules were generated and compared. The solutions balanced cost, officer workload, and response time. Management could choose among options that emphasized different priorities – for example, minimizing overtime versus maximizing coverage.

San Francisco Police Department Scheduling

- Step 6. (Present the Results and Conclusions)
  - The recommended schedules were presented to city officials. The model's outputs were communicated in clear terms – such as how many officers would be on duty in each district during each shift.

- Step 7. (Implement and Evaluate)
  - The new scheduling system was implemented, replacing the manual process. After adoption, the department monitored performance indicators like response time, overtime hours, and office satisfaction. The evaluations showed significant improvements in efficiency and service quality.

GE Capital

- During the 1980s, GE Capital, the financial services subsidiary (company) of General Electric, was rapidly expanding into areas such as equipment leasing, insurance, commercial lending, and real estate. With its diverse portfolio, the company faced a major challenge: how to evaluate investment opportunities and allocate capital efficiently.
- Traditional decision-making relied on managerial judgement, which often led to inconsistencies. To solve this, GE Capital designs quantitative models that could guide investment choices, balance risk and return, and improve portfolio management.

GE Capital

- Step 1. (Formulate the Problem)
  - GE Capital wanted to decide which investment opportunities to pursue and how to allocate limited capital among competing projects in order to maximize return while controlling risk.

- Step 2. (Observe the System)
  - Detailed financial and operational data were collected across multiple divisions. This included expected returns, risk measures, correlations between business units, and resource requirements. Historical records were studied to understand how prior investment decisions has performed.

GE Capital

- Step 3. (Formulate a Mathematical Model)
  - The investment allocation problem was modeled as a large-scale optimization problem. Decision variables represented how much capital to allocate to each project or unit.
  - Constraints ensured that
    - ✓ Total investment did not exceed available capital.
    - ✓ Risk exposure stayed within acceptable limits.
    - ✓ Regulatory and strategic guidelines were followed.

GE Capital

- Step 4. (Verify the Model and Use the Model for Prediction)
  - The model was tested by applying it to past data and comparing its recommendations to actual outcomes. The close alignment between predicted and real results gave confidence.

- Step 5. (Select Suitable Alternative Solution)
  - The model generated alternative allocation strategies reflecting different priorities – for instance, a conservative low-risk strategy, a balanced strategy, and a more aggressive high-return approach.

GE Capital

- Step 6. (Present the Results)
  - The results were presented in clear terms, showing expected returns, risk profiles, and capital allocations. This helped executives and board members understand the trade-offs and make informed choices.

- Step 7. (Implement and Evaluate)
  - The selected allocation strategy was implemented in GE Capital's investment planning process. Performance was tracked continuously, and the model was updated with new data over time.

General Comment on Model Building

- Model building is a prime example of so called
  - Interdisciplinary work that spans various engineering, scientific (primarily mathematics), and business fields.
  - It is a very productive and an exciting area to work in.

Thank you