Class 2: More Modeling Examples

1 Capacity Expansion

Capacity expansion models optimal choices of the timing and levels of investments to meet future demands of given product. Here we consider the case of power expansion for electricity generation: to find optimal levels of investments in various types of power plants to meet future electricity demand. We discuss three modeling here: *static deterministic model*, *dynamic model* and *stochastic model*.

1.1 Static Deterministic Model

Three properties of a given power plant i can be singled out: the investment cost r_i, the operating cost q_i and the availability factor a_i which indicates the percent of time the power plant can effectively be operated. Demands for electricity can be considered a single product, but the level of demand varies over time. Analysts usually use *load duration curve* to describe the demand over time in decreasing order of demand level.

The load duration curve can be approximated by a piecewise constant curve with m segments. Let $d_1 = D_1$, $d_j = D_j - D_{j-1}$, $j = 2, \cdots, m$ represent the additional power demand in mode j with a time τ_j. In the static situation, the problem consists of finding the optimal investment foe each mode, i.e.., to find a particular plant i such that the total cost of effectively producing electricity during the time τ_j:

$$i(j) = \arg \min_{i=1,\cdots,n} \left\{ \frac{r_i + q_i \tau_j}{a_i} \right\}.$$

The solution of the static model captures one essential feature of the problem: base load demand should be covered by equipment with low operating costs and peak load demand should be covered by equipment with low investment costs.

1.2 Dynamic Model

At least four elements justify using a dynamic or multistage model:

1. the long-term evolution of equipment costs;
2. the long-term evolution of the load curve;
3. the appearance of new technologies;
4. the obsolescence of currently available equipment.

Of significant importance is the evolution of demand in both the total energy demanded, and the peak level which determines the total capacity that must be available.

Hence, the following multistage model is proposed to handle this:

- $t = 1, \cdots, H$ index the periods or stages;
- $i = 1, \cdots, n$ index the available technologies;
- $j = 1, \cdots, m$ index the operating modes in the load duration curve;
- a_i = availability factor of i;
- L_i = life time of i;
- g_i^t = existing capacity of i at time t, decided before $t = 1$;
- r_i^t = unit investment cost for i at time t;
- q_i^t = unit production cost for i at time t;
- d_j^t = maximal power demanded in mode j at time t;
- τ_j^t = duration of mode j at time t;
- x_i^t = new capacity made available for technology i at time t;
- w_i^t = total capacity of i available at time t;
- y_{ij}^t = capacity of i effectively used at time t in mode j.

The electricity generation H-stage problem can be defined as

$$\begin{aligned}
&\min_{x,y,w} \sum_{i=1}^{H} \left(\sum_{i=1}^{n} r_i^t w_i^t + \sum_{i=1}^{n} \sum_{j=1}^{n} q_i^t \tau_j^t y^t ij \right) \\
\text{s.t.:} \quad & w_i^t = w_i^{t-1} + x_i^t - x_i^{t-L_i}, \ i = 1, \cdots, n, \ t = 1, \cdots, H, \\
& \sum_{i=1}^{n} y_{ij}^t = dt_j, \ j = 1, \cdots, m, \ t = 1, \cdots, H, \\
& \sum_{j=1}^{m} y_{ij}^t \leq a_i (g_i^t + w_i^t), \ i = 1, \cdots, m, \ t = 1, \cdots, H, \\
& x, y, w \geq 0.
\end{aligned}$$

The objective function is the sum of investment plus maintenance costs and operating costs. Compared to the static model, the factor a_i goes to the constraint.

1.3 Stochastic Model

In contrast to the dynamic model which assumes that the evolutions are all deterministic, we can also consider them as truly random. This leads to the stochastic model. The major difference here is x_i^t now represent the new capacity of i decided at time t, which becomes available at $x_i^{t+\Delta_i}$ (Δ_i is the construction delay). We use boldface to represent random variables:

- $\mathbf{x}_i^t$ = new capacity decided at time t for equipment i;
- $\mathbf{w}_i^t$ = total capacity of i available at time t;

- $\boldsymbol{\xi}$ = the vector of random parameters at time t;

The stochastic model is then:

$$\begin{aligned}
&\min_{\mathbf{x},\mathbf{y},\mathbf{w}} \mathbb{E}_{\boldsymbol{\xi}} \sum_{i=1}^{H} \left(\sum_{i=1}^{n} \mathbf{r}_i^t \mathbf{w}_i^t + \sum_{i=1}^{n} \sum_{j=1}^{n} \mathbf{q}_i^t \boldsymbol{\tau}_j^t \mathbf{y}^t ij \right) \\
\text{s.t.:}\quad & \mathbf{w}_i^t = \mathbf{w}_i^{t-1} + \mathbf{x}_i^t - \mathbf{x}_i^{t-L_i},\ i = 1, \cdots, n,\ t = 1, \cdots, H, \\
& \sum_{i=1}^{n} \mathbf{y}_{ij}^t = \mathbf{d}t_j,\ j = 1, \cdots, m,\ t = 1, \cdots, H, \qquad (1) \\
& \sum_{j=1}^{m} \mathbf{y}_{ij}^t \le a_i(g_i^t + \mathbf{w}_i^{t-\Delta_i}),\ i = 1, \cdots, m,\ t = 1, \cdots, H, \\
& \mathbf{x}, \mathbf{y}, \mathbf{w} \ge 0.
\end{aligned}$$

Notice here, the decision $\mathbf{w}^t$, $\mathbf{y}^t$ only depend on the realization of the random vector up to time t, but can not depend on future realizations of the random vector. That is, it is non-anticipative.

If the decision variable ($\mathbf{w}^t$, $\mathbf{y}^t$) were not dependent on $\boldsymbol{\xi}^t$, the objective function can be written as

$$\sum_t \sum_i \left(\mathbb{E}_{\boldsymbol{\xi}} \mathbf{r}_i^t w_i^t + \sum_j \mathbb{E}_{\boldsymbol{\xi}} \mathbf{q}_i^t \tau_i^t y_{ij}^t \right) = \sum_t \sum_i \left(\bar{r}_i^t w_i^t + \sum_j \overline{(q_i \tau_j)} y_{ij}^t \right).$$

That is, it becomes a deterministic formulation.

Problem 1 is a multistage stochastic liner program with a special property: *block separable recourse*. This property stems from a separation that can be made between the aggregate level decision ($\mathbf{x}^t, \mathbf{w}^t$) and the detailed-level decisions $\mathbf{y}^t$. Suppose future demands are always independent of the past. In this case, the decision on capacity to install in the future at some t only depends on available capacity and does not depend on the outcomes up to time t. The same $\mathbf{x}^t$ must then be optimal for any realization of $\boldsymbol{\xi}$. The only stochastic decision is in the operation-level: $\mathbf{y}^t$, which now depends separately on each period's capacity. Thus, this multi-period problem becomes a less complex two-period problem.

Consider the following example: we have a two-period three-operating-mode problem, with $n = 4$, $\Delta_i = 1$, $a = 1$; $g = 0$. The only random variable is $\mathbf{d}_1 = \boldsymbol{\xi}$. The other demands are $d_2 = 3$ and $d_3 = 2$. The investment costs are $r^1 = (10, 7, 16, 6)^\top$ with production costs $q^2 = (4, 4.5, 3.2, 5.5)^\top$ and load duration $\tau^2 = (10, 6, 1)^\top$. We also add a budget constraint to keep all investment below 120. The resulting stochastic program is:

$$\begin{aligned}
\min :\ & 10x_1^1 + 7x_2^1 + 16x_3^1 + \mathbb{E}_{\boldsymbol{\xi}}[\sum_{j=1}^{3} \tau_j^2 (4\mathbf{y}_{1j}^2 + 4.5\mathbf{y}_{2j}^2 + 3.2\mathbf{y}_{3j}^2 + 5.5\mathbf{y}_{4j}^2)] \\
\text{s.t.:}\ & 10x_1^1 + 7x_2^1 + 16x_3^1 \le 120, \\
& -x_i^1 + \sum_{i=1}^{3} \mathbf{y}_{ij}^2 \le 0,\ i = 1 \cdots, 4, \\
& \sum_{i=1}^{y} \mathbf{y}_{i1}^2 = \boldsymbol{\xi}, \\
& \sum_{i=1}^{y} \mathbf{y}_{ij}^2 = d_j^2,\ j = 2, 3 \\
& x, y \ge 0.
\end{aligned}$$

Assuming that $\boldsymbol{\xi}$ takes on the values 3, 5 and 7 with probabilities 0.3, 0.4 and 0.3 respectively, an optimal solution includes $x^1 = (2.67, 4.00, 3.33, 2.00)^\top$ with optimal value of 381.85. If we consider the expected parameter formulation, the optimal solution is $(0.00, 3.00\, 5.00\, 2.00^\top)$ with an objective value 365. However, if we use this solution in the stochastic problem, then with probability 0.3 the demands are not satisfied.

Formulation of 1 requires that the demand is always satisfied, in practice it is often relaxed to a probability constraint:

$$P[\sum_{i=1}^{n-1} a_i(g_i^t + w_i^t) \geq \sum_{j=1}^{m} \mathbf{D}_j^t] \geq \alpha,\ \forall t,$$

for some $\alpha \in (0,\, 1)$.This is often called a *chance constraint* or *probability constraint* in stochastic programming. Notice this constraint is equivalent to a deterministic constraint:

$$\sum_{i=1}^{n-1} a_i(g_i^t + w_i^t) \geq (F^t)^{-1}(\alpha),\ \forall t,$$

where F^t is the distributed function of $\sum_{j=1}^{m} \mathbf{d}_j^t$.

2 Design for Manufacturing Quality

Consider a designer deciding various product specification to achieve some measure of product cost and performance. However, the specification may not completely determine the characteristics of each manufactured product. Key characteristics of the product are often random. For example, every item includes variations due to machining or other processing, and each customer also does not use the product in the same way. Thus, cost and performance become random variables, and stochastic programming can help in this case, because deterministic methods can yield costly results that are only discovered after production has begun.

In this section, we consider the design of a simple axle assembly for a bicycle cart.

The designer must determine the specified length w and the diameter ξ of the axle. Together, these quantities determine the performance characteristic of the product. The goal is to determine a combination to achieve the greatest expected profit.

In the production process, the actual dimensions are not exactly those that are specified. For this example, we suppose that the length w can be produced exactly but the diameter ξ is a random variable $\boldsymbol{\xi}(x)$ that depends on a specified value x that represents, for example the setting of a machine. We assume the following triangular distribution of $\boldsymbol{\xi}(x)$:

$$f_x(\xi) = \begin{cases} (100/x^2)(\xi - 0.9x) & \text{if } 0.9x \leq \xi \leq x, \\ (100/x^2)(1.1x - \xi) & \text{if } x \leq \xi \leq 1.1x, \\ 0 & \text{otherwise.} \end{cases}$$

The decision is to determine w and x subject to $w \leq w^{\max}$ and $x \leq x^{\max}$, in order to maximize the expected utility.

Further assume that we sell as many as we can produce under the maximum selling price, and the maximum selling price depends on the length and is expressed as

$$r(1 - e^{-0.1w}),$$

where r is the maximum possible for any such products.

Our production costs for labor and equipment are assumed to be fixed, so only material cost is variable. This cost is proportional to the mean value of the specified dimension because material is acquired before the production process. That is,

$$c(w\pi x^2)/4,$$

for some unit cost c.

We have certain quality constraint:

$$\frac{w}{\xi^3} \le 39.27,$$
$$\frac{w^3}{\xi^4} \le 63169.$$

When either of these constraints is violated, the axle deforms, which incurs a quality cost proportional to the square of the violation. That is:

$$Q(w,x,\xi) = \min_y\{qy^2 \big| w/\xi^3 - y \le 39.27;\ w^3/\xi^4 - 300y \le 63169\},$$

and the expected quality cost given w and x is

$$\begin{aligned}\mathcal{Q}(w,x) &= \int_\xi Q(w,x,\xi) f_x(\xi) d\xi \\ &= q\int_{0.9x}^{1.1x} (100/x^2)\min\{\xi - 0.9x, 1.1x - \xi\}[\max\{0, w/\xi^3 - 39.27, w^3/300\xi^4 - 210.56\}]^2 d\xi.\end{aligned}$$

The overall problem is to find

max (total revenue er item - manufacturing cost per item -expected quality cost).

Mathematically, we write this as

$$\begin{aligned}&\max z(w,x) = r(1 - e^{-0.1w}) - c(w\pi x^2)/4 - \mathcal{Q}(w,x)\\ &\text{s.t. } 0 \le w \le w^{max},\ 0 \le x \le x^{max}.\end{aligned}$$

Let $w^{max} = 36$, $x^{max} = 1.25$, $r = 10$, $c = 0.025$, $q = 1$. The optimal solution is $w^* = 33.6$ and $x^* = 1.038$ where the objective function is 8.94.

If we consider the expected parameter formulation, we get a solution $\bar{w} = 35.0719$ and $\bar{x} = 0.963$, with the optimal value 9.07. It seems that the expected profit is even better than the stochastic solution. That is because, the optimal value using the expected parameter *overestimate* the true expected profit. Substitute this solution into the stochastic formulation, the expected profit is -26.79. Hence we see that the VSS is 35.73.