Optimización bajo Incertidumbre
2. Modelado

Carlos Testuri – Germán Ferrari

Departamento de Investigación Operativa – Instituto de Computación
Facultad de Ingeniería – Universidad de la República

2003-2022

Contenido

1. Decisiones y Etapas

2. Formulaciones

Formalización del modelado: decisiones y etapas

La incertidumbre es representada con eventos aleatorios, $\omega \in \Omega$.
Se tienen variables aleatorias definidas sobre los eventos, $\xi(\omega) \in \Xi$.
Estas son parámetros inciertos de los problemas.

Las decisiones se clasifican en relación a cuando tiene lugar el experimento:

- antes del experimento: decisiones de *primer etapa*; denotadas usualmente mediante x.
- luego del experimento: decisiones de *segunda etapa*; denotadas usualmente por y, $y(\omega)$, o $y(\omega, x)$.

El modelo incluye todos los elementos: x, $\xi(\omega)$, $y(\omega, x)$.
La implementación los ordena según: 1) x, 2) $\xi(\omega)$, y 3) $y(\omega, x)$.

Programa de dos etapas con recurso fijo
Formulación clásica (Dantzig 1955, Beale 1955)

El programa estocástico lineal de dos etapas con *recurso* fijo W, decisiones de primer etapa x, eventos aleatorios $\omega \in \Omega$, parámetro aleatorio $\xi(\omega) := [q(\omega), h(\omega), T(\omega)]$ y decisiones de segunda etapa $y(\omega)$, es

$$\begin{array}{cl} \min & c^\tau x + \mathbb{E}_\xi q(\omega)^\tau y(\omega) \\ \text{s.a} & Ax = b, \\ & T(\omega)x + Wy(\omega) = h(\omega), \\ & x \geq 0, y(\omega) \geq 0. \end{array}$$

Las decisiones y son elegidas de tal forma que las restricciones (donde participan) tienen lugar casi seguramente (almost surely, a.s.), es decir para todo $\omega \in \Omega$ excepto para un conjuntos de ω con probabilidad cero. Todas las restricciones aleatorias se asumirán de esta forma excepto que se establezca una probabilidad dada.

Ejemplo: Problema de localización

Una empresa puede abrir plantas en ciertos lugares $j = 1, ..., n$ que operan con costos fijo c_j y variable v_j para atender la demanda d_i de un producto a un precio unitario r_i, por parte de sus clientes $i = 1, ..., m$.

Cada cliente puede ser suministrado desde una planta que se decide abrir, a un costo unitario de transporte t_{ij}.
El beneficio obtenido cuando el cliente i es suministrado por la planta j es $q_{ij} = (r_i - v_j - t_{ij})d_i$.

El objetivo es elegir los lugares donde abrir plantas, y atender la demanda de los clientes por parte de las plantas de forma de maximizar los beneficios.

Problema de localización: modelo determinista

- x_j : 1, si la planta j es abierta, 0, en otro caso,
- y_{ij} : fracción de la demanda del cliente i suministrada desde la planta j.

$$
\begin{array}{ll}
\max & -\sum_{j=1}^{n} c_j x_j + \sum_{i=1}^{m} \sum_{j=1}^{n} q_{ij} y_{ij} \\
\text{s.a} & \sum_{j=1}^{n} y_{ij} \leq 1, \quad i = 1, ..., m, \\
& 0 \leq y_{ij} \leq x_j, \quad i = 1, ..., m, \, j = 1, ..., n, \\
& x_j \in \{0, 1\}, \quad j = 1, ..., n.
\end{array}
$$

Problema de localización: modelo con datos aleatorios

Se tienen costos de producción, costos de distribución y precios aleatorios. Sean la demanda y el esquema de distribución deterministas (primer etapa). El nuevo beneficio es $q_{ij}(\omega) := (r_i(\omega) - v_j(\omega) - t_{ij}(\omega))d_i$.

Entonces el nuevo objetivo es

$$-\sum_{j=1}^{n} c_j x_j + \mathbb{E}_\xi \sum_{i=1}^{m} \sum_{j=1}^{n} q_{ij}(\omega) y_{ij},$$

que al intercambiar las sumatorias con la esperanza queda como

$$-\sum_{j=1}^{n} c_j x_j + \sum_{i=1}^{m} \sum_{j=1}^{n} \left(\mathbb{E}_\xi q_{ij}(\omega)\right) y_{ij}.$$

Equivalente a un modelo determinista.

Problema de localización: demanda aleatoria

El esquema de distribución se asume determinista.
Se redefine la variable y_{ij} como la cantidad suministrada.
Sean p_i^+ la penalidad por unidad de demanda d_i no satisfecha, y p_i^- la penalidad por exceso sobre d_i.

Se definen las variables de segunda etapa:

- $z_i^+(\omega)$ cantidad de demanda en i no satisfecha, y
- $z_i^-(\omega)$ cantidad en exceso enviada a i, en estado ω.

$$
\begin{array}{ll}
\max & -\sum_{j=1}^n c_j x_j + \sum_{i=1}^m \sum_{j=1}^n (-v_j - t_{ij}) y_{ij} + \\
 & \mathbb{E}_\xi \left(-\sum_{i=1}^m p_i^+ z_i^+(\omega) - \sum_{i=1}^m p_i^- z_i^-(\omega) \right) + \mathbb{E}_\xi \sum_{i=1}^m r_i d_i(\omega) \\
\text{s.a} & \sum_{i=1}^m y_{ij} \leq M x_j, \quad j = 1, ..., n, \\
 & z_i^+(\omega) - z_i^-(\omega) = d_i(\omega) - \sum_{j=1}^n y_{ij}, \quad i = 1, ..., m, \\
 & x_j \in \{0, 1\},\ y_{ij}(\omega), z_i^+(\omega), z_i^-(\omega) \geq 0\ i = 1, ..., m, j = 1, ..., n.
\end{array}
$$

Problema de localización: demanda y distribución aleatorios

La incertidumbre corresponde a escenarios de largo plazo.
Se redefine $y_{ij}(\omega)$ como la fracción de la demanda $d_i(\omega)$, y se definen las variables w_j, capacidad de la planta j con costo unitario de inversión g_j. El nuevo beneficio es $q_{ij}(\omega) := (r_i - v_j - t_{ij})d_i(\omega)$.

$$
\begin{array}{ll}
\max & -\sum_{j=1}^{n} c_j x_j - \sum_{j=1}^{n} g_j w_j + \mathbb{E}_\xi \sum_{i=1}^{m} \sum_{j=1}^{n} q_{ij}(\omega) y_{ij}(\omega) \\
\text{s.a} & \sum_{j=1}^{n} y_{ij}(\omega) \leq 1, \quad i = 1, ..., m, \\
& \sum_{i=1}^{m} d_i(\omega) y_{ij}(\omega) \leq w_j, \quad j = 1, ..., n, \\
& 0 \leq y_{ij}(\omega) \leq x_j, \quad i = 1, ..., m, j = 1, ..., n, \\
& x_j \in \{0, 1\}, w_j \geq 0, \quad j = 1, ..., n.
\end{array}
$$

Problema de localización y distribución

Relación entre etapas y períodos

Sea el caso anterior con un horizonte temporal de 36 meses.
Donde la localización de las plantas (decisión de primer etapa) dura 6 meses, y la distribución (la decisión de segunda etapa) dura 30 meses.
Independientemente de los períodos el problema sigue siendo de dos etapas.

¿Cómo se modelaría la condición de que se pueden abrir nuevas plantas en el tercer semestre de planificación (meses 12-18)?

Se tendrían tres etapas:
1era. con localización de plantas (meses 1-6),
2da. con distribución (meses 7-18) y localización de plantas (meses 12-18), y
3era. con distribución (meses 19-36).

Problema de localización y distribución
Modelo

Sean ω_2 y ω_3 los eventos aleatorios que definen las etapas 2 y 3, respect. Las variables x^1 y $x^2(\omega_2)$ representan las decisiones de localizaciones en las etapas 1 y 2, respectivamente. Las variables $y^2(\omega_2)$ y $y^3(\omega_3)$ representan las decisiones de distribución en etapas 2 y 3, respectivamente. La capacidad de cada planta es M.

$$
\begin{array}{ll}
\max & -\sum_{j=1}^{n} c_j x_j^1 + \mathbb{E}_{\xi_2} \max\{\sum_{i=1}^{m}\sum_{j=1}^{n} q_{ij}^2(\omega_2) y_{ij}^2(\omega_2) \\
& -\sum_{j=1}^{n} c_j^2(\omega_2) x_j^2(\omega_2) + \mathbb{E}_{\xi_3|\xi_2} \max\{\sum_{i=1}^{m}\sum_{j=1}^{n} q_{ij}^3(\omega_3) y_{ij}^3(\omega_3)\}\} \\
\text{s.a} & \sum_{j=1}^{n} y_{ij}^2(\omega_2) \leq 1, \quad i = 1, \dots, m, \\
& \sum_{i=1}^{m} d_i(\omega_2) y_{ij}^2(\omega_2) \leq M x_j^1, \quad j = 1, \dots, n, \\
& \sum_{j=1}^{n} y_{ij}^3(\omega_3) \leq 1, \quad i = 1, \dots, m, \\
& \sum_{i=1}^{m} d_i(\omega_3) y_{ij}^3(\omega_3) \leq M(x_j^1 + x_j^2(\omega_2)), \quad j = 1, \dots, n, \\
& x_j^1 + x_j^2(\omega_2) \leq 1, \quad j = 1, \dots, n, \\
& x_j, x_j^2(\omega_2) \in \{0,1\}, y_{ij}^2(\omega_2), y_{ij}^3(\omega_3) \geq 0, \quad i = 1, \dots, m, j = 1, \dots, n.
\end{array}
$$

Programación probabilística

Algunas restricciones o la función objetivo son expresadas en términos probabilísticos con respecto a las decisiones de primer etapa.

Las restricciones no tienen por que cumplirse casi-seguramente, es decir podrían cumplirse con cierto nivel de confiabilidad o probabilidad.
Se establecen restricciones probabilísticas $i = 1, ..., I$ que se cumplen en forma conjunta

$$P\{A^i(\omega)x \geq h^i(\omega)\} \geq \alpha^i,$$

con nivel de confianza $\alpha^i \in (0, 1)$.

Manejarlas directamente implica trabajar con funciones discontinuas.
El objetivo es encontrar equivalentes deterministas de ellas que tengan propiedades de convexidad.

Programación probabilística

Ejemplo de equivalente directo

Dado el problem del "Canilla", se sabe que la demanda ξ tiene distribución normal $\xi \sim \mathcal{N}(\mu = 30, \sigma^2 = 16)$.
Se requiere que la demanda sea satisfecha con un nivel de probabilidad de 0,95, por lo que tiene la restricción probabilística

$$P\{x \geq \xi\} \geq 0,95$$

La distribución normal es invariante con respecto a transformaciones lineales. Se tiene que para una probabilidad de 0,95 el cuantil correspondiente de la distribución normal estándar es 1,645. Por lo que su transformación equivalente cumple $(x - \mu)/\sigma = (x - 30)/4 \geq 1,645$, con lo que se tiene la restricción equivalente

$$x \geq 36,58$$

Programación probabilística

Ejemplo de equivalente indirecto (1/2)

Dado el problema de cobertura, con ubicaciones potenciales $j = 1, ..., n$, donde x_j indica si se abre o no servicio en localización j a costo c_j.
Dados clientes $i = 1, ..., m$, que son atendidos si existe un servicio a distancia t_i, donde la distancia entre i y j es t_{ij}. Se definen las ubicaciones posibles para el cliente i como $N_i = \{j | t_{ij} < t_i\}$.

$$\begin{array}{cl} \min & \sum_{j=1}^{n} c_j x_j \\ \text{s.a} & \sum_{j \in N_i} x_j \geq 1, \quad i = 1, ..., m, \\ & x_j \in \{0, 1\}, \quad j = 1, ..., n. \end{array}$$

Programación probabilística

Ejemplo de equivalente indirecto (2/2)

Una ubicación puede cubrir más de un cliente.
Sea q la probabilidad de que no haya servicio disponible en cada ubicación j.
La formulación puede modificarse con el requerimiento de que la probabilidad de servicio disponible de una ubicación posible sea superior a cierto nivel de confianza α.

Entonces, para cada cliente la restricción probabilística de obtener servicio es

$$1 - q^{\sum_{j \in N_i} x_j} \geq \alpha, \quad i = 1, ..., m,$$

para la cual luego de aplicar logaritmos se obtiene el equivalente lineal

$$\sum_{j \in N_i} x_j \geq \left\lceil \frac{\ln(1-\alpha)}{\ln q} \right\rceil, \quad i = 1, ..., m.$$