IE 495 – Stochastic Programming Problem Set #2 — Solutions

1 Math Time

Consider the problem (P):

subject to

$$z^* = \min 2y_1 + y_2$$

$$y_1 + y_2 \ge 1 - x_1$$

 $y_1 \ge \xi - x_1 - x_2$
 $y_1, y_2 \ge 0$

1.1 Problem

Show that P has complete recourse.

Answer:

Write P as

 $\min 2y_1 + y_2$

subject to

$$y_1 + y_2 - y_3 = 1 - x_1$$

$$y_1 - y_4 = \xi - x_1 - x_2$$

$$y_1, y_2, y_3, y_4 \ge 0$$

So the recourse matrix is

$$W = \left[\begin{array}{rrrr} 1 & 1 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{array} \right].$$

To show that P has complete recourse, we must show that $Wy = t, y \ge 0$ has a solution for every $t = (t_1, t_2)$. This is fairly clear. One solution is

$$\begin{array}{rcl} y_1 & = & \max\{0,t_1,t_2\} \\ y_2 & = & 0 \\ y_3 & = & y_1-t_1 \\ y_4 & = & y_1-t_2. \end{array}$$

Note that Birge is misleading. It doesn't matter that $\mathbb{E}[\xi] < \infty$, for any realization $\hat{\xi}$, we can certainly find a feasible y. Q.E.D.

1.2 Problem

Assume that $0 \le x_1 \le 1, 0 \le x_2 \le 1$ show that the following are optimal solutions:

$$\begin{cases} y_1^* = 0, y_2^* = 1 - x_1 & \xi \le x_1 + x_2 \\ y_1^* = \xi - x_1 - x_2, y_2^* = (1 - \xi + x_2)^+ & \xi \ge x_1 + x_2 \end{cases}$$

So that...

$$Q(x,\xi) = \begin{cases} 1 - x_1 & 0 \le \xi \le x_1 + x_2\\ \xi + 1 - 2x_1 - x_2 & x_1 + x_2 \le \xi \le 1 + x_2\\ 2(\xi - x_1 - x_2) & 1 - x_2 \le \xi \end{cases}$$

Answer:

First note that the solution y_1^*, y_2^* is primal feasible. Write the dual (D) of (P) as the following:

$$w^* = \max(1 - x_1)\pi_1 + (\xi - x_1 - x_2)\pi_2$$

subject to

$$\begin{array}{rcl}
\pi_1 + \pi_2 &\leq & 2\\
\pi_1 &\leq & 1\\
\pi_1, \pi_2 &\geq & 0
\end{array}$$

We will demonstrate (in every case) that there is a feasible dual solution π_1^*, π_2^* such that $w^* = z^*$. This suffices to prove the claim.

Case 1. $\xi \le x_1 + x_2$. (So that $y_1^* = 0, y_2^* = 1 - x_1$).

Consider dual solution $\pi_1^* = 1, \pi_2^* = 0$. This solution is feasible to D. Since $w^* = z^* = 1 - x_1$, $y_1^* = 0, y_2^* = 1 - x_1$ is an optimal second stage solution.

Case 2. $\xi \ge x_1 + x_2 \cdot 1 - \xi + x_2 \le 0$, (So that $y_1^* = \xi - x_1 - x_2, y_2^* = 0$) Consider the dual solution $\pi_1^* = 0, \pi_2^* = 2$. This solution is feasible to *D*. Since $w^* = z^* = 2(\xi - x_1 - x_2), y_1^* = \xi - x_1 - x_2, y_2^* = 0$) is an optimal second stage solution.

Case 3. $\xi \ge x_1 + x_2, 1 - \xi + x_2 > 0$, (So that $y_1^* = \xi - x_1 - x_2, y_2^* = 1 - \xi + x_2$).

Consider the dual solution $\pi_1^* = 1, \pi_2^* = 1$. This solution is feasible to D, and $w^* = z^* = 1 + \xi - 2x_1 - x_2$, so $y_1^* = \xi - x_1 - x_2, y_2^* = 1 - \xi + x_2$ is an optimal second stage solution. Q.E.D.

1.3 Problem

Show that $Q(x,\xi)$ is piecewise linear convex function in ξ .

Answer:

Since

$$Q(x,\xi) = \begin{cases} 1 - x_1 & 0 \le \xi \le x_1 + x_2 \\ \xi + 1 - 2x_1 - x_2 & x_1 + x_2 \le \xi \le 1 + x_2 \\ 2(\xi - x_1 - x_2) & 1 - x_2 \le \xi \end{cases},$$

 $Q(x,\xi)$ is piecewise linear in ξ . To show that it is convex, examine its graph in Figure 1

Q.E.D.

1.4 Problem

Show that $Q(x,\xi)$ is a piecewise linear convex function in $x, \forall x \in K$





Answer:

Consider the graph of the "regions" of $Q(x,\xi)$ in Figure 2.

 $Q(x,\xi)$ has different shapes on each of the regions R_I, R_{II} , and R_{III} . Consider two points $p = (p_1, p_2), q = (q_1, q_2)$. The goal will be to show that the "convexity relationship" holds for these two arbitrary points. That is, $\forall \lambda \in [0, 1]$,

$$\lambda Q(p,\xi) + (1-\lambda)Q(q,\xi) - Q(\lambda p + (1-\lambda)q,\xi) \ge 0.$$

Unfortunately, we must do this on a case-by-case basis. The first thing to note is that if p and q both are in the same region, the convexity relationship obviously holds, since $Q(x,\xi)$ is linear on each region. Now we will do the remaining cases...

Case 1. $p \in R_I, q \in R_{II}$.

By the definition of $Q(x,\xi)$, we know that

$$\lambda Q(p,\xi) + (1-\lambda)Q(q,\xi) = \lambda(1-p_1) + (1-\lambda)(\xi+1-2q_1-q_2).$$
(1)

Case 1-a. $\lambda p + (1 - \lambda)q \in R_I$.

By the definition of $Q(x,\xi)$, we know that

$$Q(\lambda p + (1 - \lambda)q, \xi) = 1 - (\lambda p_1 + (1 - \lambda)q_1).$$
(2)

Subtracting (2) from (1) yields

$$\lambda Q(p,\xi) + (1-\lambda)Q(q,\xi) - Q(\lambda p + (1-\lambda)q,\xi) = \lambda(1-p_1) + (1-\lambda)(\xi+1-2q_1-q_2) - (1-(\lambda p_1 + (1-\lambda)q_1)) = (1-\lambda)(\xi-q_1-q_2) \ge 0.$$

In the last inequality, we have explicitly used the fact that $q \in R_{II}$, so that $\xi \ge q_1 + q_2$.

An Aside—Maple Rules!

If you are not using or have never used Maple, you definitely should. How did I know that

$$\lambda(1-p_1) + (1-\lambda)(\xi+1-2q_1-q_2) - (1-(\lambda p_1+(1-\lambda)q_1)) = (1-\lambda)(\xi-q_1-q_2)$$

you ask? I used Maple, of course. Here is all you have to do...

> simplify(e1);

> factor(e1);

$$-(-1 + L) (z - q1 - q2)$$

Case 1-b. $\lambda p + (1 - \lambda)q \in R_{II}$.

By the definition of $Q(x,\xi)$, we know that

$$Q(\lambda p + (1 - \lambda)q, \xi) = \xi + 1 - 2(\lambda p_1 + (1 - \lambda)q_1) - (\lambda p_2 + (1 - \lambda)q_2)$$
(3)

Subtracting (3) from (1) yields

$$\lambda Q(p,\xi) + (1-\lambda)Q(q,\xi) - Q(\lambda p + (1-\lambda)q,\xi) = \lambda(1-p_1) + (1-\lambda)(\xi+1-2q_1-q_2) - (\xi+1-2(\lambda p_1 + (1-\lambda)q_1) - (\lambda p_2 + (1-\lambda)q_2)) = \lambda(p_1+p_2-\xi) \ge 0$$

In the last inequality, we have used the fact that $p \in R_I$, so $p_1 + p_2 \ge \xi$.

Case 2. $p \in R_{II}, q \in R_{III}$.

By the definition of $Q(x,\xi)$, we know that

$$\lambda Q(p,\xi) + (1-\lambda)Q(q,\xi) = \lambda(\xi+1-2p_1-p_2) + 2(1-\lambda)(\xi-q_1-q_2).$$
(4)

Case 2-a. $\lambda p + (1 - \lambda)q \in R_{II}$.

By the definition of $Q(x,\xi)$, we know that

$$Q(\lambda p + (1 - \lambda)q, \xi) = \xi + 1 - 2(\lambda p_1 + (1 - \lambda)q_1) - (\lambda p_2 + (1 - \lambda)q_2).$$
(5)

Subtracting (5) from (4) yields

$$\lambda Q(p,\xi) + (1-\lambda)Q(q,\xi) - Q(\lambda p + (1-\lambda)q,\xi) = \lambda(\xi+1-2p_1-p_2) + 2(1-\lambda)(\xi-q_1-q_2) - (\xi+1-2(\lambda p_1+(1-\lambda)q_1) - (\lambda p_2+(1-\lambda)q_2)) = (1-\lambda)(\xi-(1+q_2)) \ge 0.$$

In the last inequality, we have used the fact that $q \in R_{III}$, so $\xi \ge 1 + q_2$.

Case 2-b. $\lambda p + (1 - \lambda)q \in R_{III}$.

By the definition of $Q(x,\xi)$, we know that

$$Q(\lambda p + (1 - \lambda)q, \xi) = 2(\xi - (\lambda p_1 + (1 - \lambda)q_1) - (\lambda p_2 + (1 - \lambda)q_2))$$
(6)

Subtracting (6) from (4) yields

$$\lambda Q(p,\xi) + (1-\lambda)Q(q,\xi) - Q(\lambda p + (1-\lambda)q,\xi) = \lambda(\xi + 1 - 2p_1 - p_2) + 2(1-\lambda)(\xi - q_1 - q_2) - (2(\xi - (\lambda p_1 + (1-\lambda)q_1) - (\lambda p_2 + (1-\lambda)q_2))) = \lambda(1 - \xi + p_2) \geq 0.$$

In the last inequality, we have used the fact that $p \in R_{II}$, so $\xi \leq 1 + p_2$.

If $p \in R_I, q \in R_{III}$, the convexity relationship follows from what has already been whon here. Q.E.D.

If $\xi \approx \mathcal{U}[0,2]$, then $\mathcal{Q}(x) = 1/4(x_1^2 + 2x_2^2 + 2x_1x + 2 - 8x_1 - 6x_2 + 9)$. Check the following:

1.5 Problem

 $\mathcal{Q}(x)$ is differentiable on K_2 and $\mathcal{Q}(x)$ is Lipschitz convex on K_2 .

Answer:

Here, we assume $K_2 = \{x = \Re^2 : 0 \le x_1 \le 1, 0 \le x_2 \le 1\}$, since that is the region over which $Q(x,\xi)$ was calculated.

The gradient of the expected recourse function is

$$\nabla \mathcal{Q}(x) = \begin{bmatrix} 1/2x_1 + 1/2x_2 - 8\\ 1/2x_1 + x_2 - 6 \end{bmatrix},$$

Problem 1

so $\mathcal{Q}(x)$ is differentiable. The Hessian of $\mathcal{Q}(x)$ is

$$\nabla^2 \mathcal{Q}(x) = \left[\begin{array}{cc} 1/2 & 1/2 \\ 1/2 & 1 \end{array} \right]$$

Since $\nabla^2 \mathcal{Q}(x)$ is positive definite, $\mathcal{Q}(x)$ is convex.

There are a number of valid ways to show that $\mathcal{Q}(x)$ is Lipschitz continuous.

- Convex functions are continuous on the interior of their domain, and the function is obviously continuous on the boundary of its domain.
- Since Q(x) is differentiable, use the Mean Value Theorem.

Q.E.D.

Interlude

You will need to know the following stuff to answer problem #2.

1 Definition

A function $f: D(f) \mapsto R(f)$ is Lipschitz continuous on its domain D(f) if and only if there exists a constant L (called the Lipschitz constant) such that

$$|f(a) - f(b)| \le L|a - b|$$

for all $a, b \in D(f)$

2 Definition

The *left-derivative* of a function $f : \Re \mapsto \Re$ is the quantity

$$f'_{-}(x) = \lim_{x \to t_{-}} \frac{f(x) - f(t)}{x - t} = \sup_{x < t} \frac{f(x) - f(t)}{x - t}.$$

3 Definition

The *right-derivative* of a function $f : \Re \mapsto \Re$ is the quantity

$$f'_{+}(x) = \lim_{x \to t_{+}} \frac{f(x) - f(t)}{x - t} = \inf_{x > t} \frac{f(x) - f(t)}{x - t}.$$

1 Theorem

If $f: \Re \mapsto \Re$ is a (proper) convex function, then $\partial f(x) = [f'_{-}(x), f'_{+}(x)].$

2 Shortage Properties

Let ω be a one-dimensional random variable with cumulative distribution F. Define the *expected* shortage function as

$$H(x) = \mathbb{E}_{\omega} \left[(\omega - x)^{-} \right].$$

Assume that $\mu = \mathbb{E}_{\omega} \left[(\omega)^{-} \right] < \infty$.

2.1 Problem

Prove H(x) is convex.

Problem 2

Answer:

The function $h(x, \omega) = (\omega - x)^{-} = \max\{0, x - \omega\}$ is convex in x for a fixed ω , as it is the maximum of two convex functions. Since $H(x) = \mathbb{E}_{\omega}[h(x, \omega)]$, the claim follows, since the expectation operator preserves convexity. Q.E.D.

2.2 Problem

Prove H(x) is Lipschitz continuous. What is the Lipschitz constant?

Answer:

First, note that for all $x, y \in \Re$,

$$|(t-x)^{-} - (t-y)^{-}| \le |x-y| \quad \forall t \in \Re.$$

Now,

$$|H(x) - H(y)| = |\mathbb{E}_{\omega} [(\omega - x)^{-}] - \mathbb{E}_{\omega} [(\omega - y)^{-}]|$$

$$\leq \mathbb{E}_{\omega} [|(\omega - x)^{-} - (\omega - y)^{-}|]$$

$$\leq |x - y|$$

This proves that H(x) is Lipschitz continuous with the Lipschitz constant 1.

Q.E.D.

2.3 Problem

What is the left derivative of H at x?

Answer:

First, I will show that¹

$$H(x) = \int_{u=-\infty}^{x} F(u) du.$$

$$H(x) = \mathbb{E}_{\omega}[\max\{0, x - \omega\}]$$

= $\int_{-\infty}^{x} (x - \omega) dF(\omega)$
= $\int_{-\infty}^{x} x dF(\omega) - \int_{-\infty}^{x} \omega dF(\omega)$
= $xF(x) - \left[xF(x) - \int_{-\infty}^{x} F(u) du\right]$
= $\int_{-\infty}^{x} F(u) du.$

Now we continue using this relation.

$$H'_{-}(x) = \lim_{t \to x^{-}} \frac{H(x) - H(t)}{x - t}$$
$$= \lim_{t \to x^{-}} \frac{\int_{u = -\infty}^{x} F(u) du - \int_{u = -\infty}^{t} F(u) du}{x - t}$$

 $^{^1\}mathrm{It}\xspace$ s bit like the newsvendor problem we saw in class

The numerator and denominator in this ratio both $\rightarrow 0$, as $t \rightarrow x$, so let's apply L'Hôpital's rule.

$$H'_{-}(x) = \lim_{t \to x^{-}} \frac{\left[\int_{u=-\infty}^{x} F(u)du - \int_{u=-\infty}^{t} F(u)du\right]'}{[x-t]'}$$
$$= \lim_{t \to x^{-}} \left[\int_{u=-\infty}^{x} F(u)du\right]'.$$

The Fundamental Theorem of Calculus tells us that

$$\left[\int_{u=-\infty}^{x} F(u)du\right]' = F(x) - F(-\infty) = F(x)$$

So we have

$$H'_{-}(x) = \lim_{t \to x^{-}} F(x) = \mathcal{P}(\omega < x).$$

The strict inequality here is because we must account for "jumps" in the probability distribution. Q.E.D.

2.4 Problem

What is the right derivative of H at x?

Answer:

$$H'_{+}(x) = \lim_{t \to x^{+}} \frac{H(x) - H(t)}{x - t}$$
$$= \lim_{t \to x^{+}} \frac{\int_{u = -\infty}^{x} F(u) du - \int_{u = -\infty}^{t} F(u) du}{x - t}$$

The numerator and denominator in this ratio both $\rightarrow 0$, as $t \rightarrow x$, so let's apply L'Hôpital's rule.

$$H'_{+}(x) = \lim_{t \to x^{+}} \frac{\left[\int_{u=-\infty}^{x} F(u)du - \int_{u=-\infty}^{t} F(u)du\right]'}{[x-t]'}$$
$$= \lim_{t \to x^{+}} \left[\int_{u=-\infty}^{x} F(u)du\right]'.$$

The Fundamental Theorem of Calculus tells us that

$$\left[\int_{u=-\infty}^{x} F(u)du\right]' = F(x) - F(-\infty) = F(x).$$

So we have

$$H'_{+}(x) = \lim_{t \to x^{+}} F(x) = F(x) = P(\omega \le x).$$
 Q.E.D.

2.5 Problem

What is $\partial H(x)$?

Answer:

$$\partial H(x) = [P(\omega < x), P(\omega \le x)]$$

Note that if ω is a random variable from a continuous probability distribution, then $\partial H(x) = F(x)$. Q.E.D.

3 Network Design

This problem is concerned with a simple network planning model. The goal of the planning model is to decide how to allocate (limited) additional capacity resources in order to meet a forecast demand. We would like to minimize the expected number of unserved requests for services while satisfying the budgetary limitations on the total capacity expansion.

Specifically, consider the network shown in Figure 3. The installed capacities are shown on the figure and also in Table 2. We are given the point-to-point demands d shown in Table 1. The cost of installing one unit of additional capacity is also given in Table 2. You can install fractional capacity, and you may not install any new links in the network.



Figure 3: Network and Installed Capacity

Table 1: Demand between nodes

For each point to point pair in Table 1, there are a set of feasible routes that can carry demand between the pair. These feasible routes are listed in Table 3. There is a budget of 30 to install new capacity.

Link	Installed Capacity	Cost for New Capacity
AB	4	2
AC	7	3
AD	5	2
BE	3	4
CD	8	3
CE	3	1
DF	5	2
EF	5	2

Table 2: Capacity on Links and Cost for Additional Capacity

Demand	Routes
A-E	ABE, ACE, ADCE, ADFE
B-D	BAD, BACD, BECD, BEFD
C-F	CDF, CEF
B-F	BADF, BEF
A-F	ADF, ACDF, ACEF

Table 3: Feasible Routes

3.1 Problem

Formulate a linear program that decides how much additional capacity to install on each arc in a way that minimizes the total amount of unserviced demand subject to the budgetary restriction.

Hint: You will have variables x_j for additional capacity, and f_{ir} for the amount of demand i that flows on route r

Answer:

Make the following definitions:

- I : Set of point-point demands
- J : Set of links
- R_i : Set of routes for demand pair $i \in I$
- $\bullet \ b$: Budget
- c_j : Cost of installing one additional unit of capacity on link $j \in J$
- u_j : Installed capacity on link $j \in J$
- d_i : Demand between pair $i \in I$
- $a_{irj} = \begin{cases} 1 & \text{If route } r \in R_i \text{ for demand pair } i \text{ uses arc } j \in J \\ 0 & \text{Otherwise} \end{cases}$
- x_j : Amount of additional capacity to install on arc $j \in J$.
- f_{ir} : Amount of demand for demand pait *i* that is served along route $r \in R_i$
- t_i : Amount of unserviced demand for demand pair $i \in I$.

$$\min\sum_{i\in I}t_i$$

subject to

$$\sum_{j \in J} c_j x_j \leq b$$

$$\sum_{i \in I} \sum_{r \in R_i} a_{irj} f_{ir} \leq u_j + x_j \quad \forall j \in J$$

$$\sum_{r \in R_i} f_{ir} + t_i = d_i \quad \forall i \in I$$

$$x_j \geq 0 \quad \forall j \in J$$

$$t_i \geq 0 \quad \forall i \in I$$

$$f_{ir} \geq 0 \quad \forall i \in I, \forall r \in R_i$$

Q.E.D.

3.2 Problem

Create a valid MPS file representing your formulation in Problem 3.1, and email it to jtl3@lehigh.edu.

Answer:

You can download one from the web page. I use the AMPL write mnet command to create the file. AMPL model and data files are also downloaded from the web page. Q.E.D.

3.3 Problem

Solve your formulation from Problem 3.1.

Answer:

```
ampl: model net.mod;
ampl: data net.dat;
ampl: solve;
ILOG CPLEX 8.100, licensed to "lehigh university-bethlehem, pa", options: e m b
CPLEX 8.1.0: optimal solution; objective 3.75
11 dual simplex iterations (0 in phase I)
ampl: display x;
x [*] :=
AB 3.25
AC 0
AD 7.25
BE 0
CD 0
CE 7
DF
   1
EF 0
;
```

Q.E.D.

3.4 Problem

Unfortunately, the demands in Table 1 are random. Suppose there are demand scenarios $d_s, s = 1, 2, \ldots S$ each occuring with probability p_s . Formulate a stochastic programming version of Problem 3.1 that will minimize the *expected* amount of unserviced demand.

Answer:

Using the definitions given in the answer to Problem 3.1, and making the new following definitions:

- S : Set of scenarios
- p_s : Probability of scenario $s \in S$
- d_{is} : Demand between pair $i \in I$ in scenario $s \in S$.
- t_{is} : Amount of unserviced demand for demand pair $i \in I$ in scenario $s \in S$.
- f_{irs} : Amount of demand for demand pait *i* that is served along route $r \in R_i$ in scenario $s \in S$.

$$\min \sum_{i \in I} \sum_{s \in S} p_s t_{is}$$

subject to

$$\begin{split} \sum_{j \in J} c_j x_j &\leq b \\ \sum_{i \in I} \sum_{r \in R_i} a_{irj} f_{irs} &\leq u_j + x_j \quad \forall j \in J, \forall s \in S \\ \sum_{r \in R_i} f_{irs} + t_{is} &= d_{is} \quad \forall i \in I, \forall s \in S \\ x_j &\geq 0 \quad \forall j \in J \\ t_{is} &\geq 0 \quad \forall i \in I, \forall s \in S \\ f_{irs} &\geq 0 \quad \forall i \in I, \forall r \in R_i, \forall s \in S \end{split}$$

Q.E.D.

Q.E.D.

3.5 Problem

Let $d \in \Re^5$ be the demand in Table 1. Suppose that there are three are three demand scenarios $d_1 = 0.75d, d_2 = d, d_3 = 1.25d$, each occuring with probability 1/3. Create the proper SMPS files for this instance and mail them to jtl3@lehigh.edu.

Answer:

You can download proper SMPS files from the course web page:

3.6 Problem

Solve your formulation from Problem 3.5. What is the value of the stochastic solution in this case?

Answer:

You can either use the SMPS files and NEOS or formulate the deterministic equivalent in AMPL. The AMPL model and data files for the deterministic equivalent are on the course web page. Solving yields the following:

```
ampl: model nets.mod;
ampl: data nets.dat;
ampl: solve;
```

```
ILOG CPLEX 8.100, licensed to "lehigh university-bethlehem, pa", options: e m b
CPLEX 8.1.0: optimal solution; objective 4.8
24 dual simplex iterations (0 in phase I)
ampl: display x;
x [*] :=
AB 3
AC 0
AD 7
BE 0.2
CD 0
CE 8.8
DF 0.2
EF 0
;
```

The "mean value" solution was found as the answer to Problem 3.3. If we fix this solution, and evaluate the recourse function at this point, we get the following:

```
ampl: model nets-vss.mod;
ampl: data nets-vss.dat;
ampl: solve;
ILOG CPLEX 8.100, licensed to "lehigh university-bethlehem, pa", options: e m b
CPLEX 8.1.0: optimal solution; objective 5.083333333
18 dual simplex iterations (0 in phase I)
```

 \mathbf{So}

```
VSS = 5.083 - 4.8 = 0.283.
```

Q.E.D.

3.7 Problem

(Challenge – Bonus)

Now each of the point-to-point demands *independently* varies. For each point-to-point pair *i* the demand varies as $d_i \approx \mathcal{U}(0, 3d_i)$. Solve this instance as accurately as you can. The people with the two most accurate solutions win *FREE LUNCH* with me.