IE 495 – Lecture 4

Stochastic Programming – Recourse Models

Prof. Jeff Linderoth

January 22, 2003

Outline

- Formal notation for recourse models
  - Second-stage value function
  - Expected value function
- Forming the determinstic equivalent
  - An example
- A (famous) modeling example...
  - The NewsVendor Problem. (Complete with fancy math).

Please don't call on me!

- What is the EVPI?
- What is the VSS?

Random LP's

- Consider the following linear program $LP(\omega)$ that is parameterized by the random vector ω:

minimize

$$c^T x$$

subject to

$$\begin{aligned} Ax &= b \\ T(\omega)x &= h(\omega) \\ x &\in X \end{aligned}$$

- $X = \{x \in \Re^n : l \leq x \leq u\}$

Example – From Lecture #2

minimize

$$x_1 + x_2$$

subject to

$$\begin{aligned} \omega_1 x_1 + x_2 &\geq 7 \\ \omega_2 x_1 + x_2 &\geq 4 \\ x_1 &\geq 0 \\ x_2 &\geq 0 \end{aligned}$$

Random LPs

- Again, we deal with decision problems where the decision x must be made before the realization of ω is known.
- We do, however, know the distribution of ω on Ω.
- In recourse models, the random constraints are modeled as "soft" constraints. Possible violation is accepted, but the cost of violations will influence the choice of x.
- In fact, a *second-stage* linear program is introduced that will describe how the violated random constraints are dealt with.

The New $LP(\omega)$

- In the simplest case, we may just count penalize deviation in the constraints by penalty coefficient vectors q_+ and q_-

minimize

$$c^T x + q_+^T s(\omega) + q_-^T t(\omega)$$

subject to

$$\begin{aligned} Ax &= b \\ T(\omega)x + s(\omega) - t(\omega) &= h(\omega) \\ x &\in X \end{aligned}$$

The New Optimization Problem

- So then, a reasonable problem to solve (to deal with the randomness) is...

minimize

$$c^T x + \mathbb{E}_\omega \left[q_+^T s(\omega) + q_-^T t(\omega) \right]$$

subject to

$$\begin{aligned} Ax &= b \\ T(\omega)x + s(\omega) - t(\omega) &= h(\omega) \qquad \forall \omega \in \Omega \\ x &\in X \end{aligned}$$

Recourse

- In general, we can *react* in an intelligent (or optimal) way.

⋆ We have some *recourse!*

- A recourse structure is provided by three items
  - ◇ A set $Y \in \Re^p$ that describes the feasible set of recourse actions.

  Ex. $Y = \{y \in \Re^p : y \geq 0\}$

  - ◇ q : a vector of recourse costs.
  - ◇ W : a $m \times p$ matrix, called the *recourse matrix*

A Recourse Formulation

minimize

$$c^T x + \mathbb{E}_\omega \left[q^T y\right]$$

subject to

$$\begin{aligned} Ax &= b \\ T(\omega)x + Wy(\omega) &= h(\omega) \qquad \forall \omega \in \Omega \\ x &\in X \\ y(\omega) &\in Y \end{aligned}$$

- Right now, (and in nearly all problems we will see), we have only one W.

⇒ Our recourse does not change with the scenario.

- This is called *Fixed recourse*.

Some Definitions

$$\min_{x \in X : Ax = b} \left\{ c^T x + \mathbb{E}_\omega \left[\min_{y \in Y} \{ q^T y : Wy = h(\omega) - T(\omega)x \} \right] \right\}$$

- *Second stage value function*, or *recourse (penalty) function* $v : \Re^m \mapsto \Re$.
  - ◇ $v(z) \equiv \min_{y \in Y} \{q^T y : Wy = z\}$,
  - ◇ For any vector z of "deviations in the random constraints $T(\omega)x = h(\omega)$", it describes the corresponding cost.
- *Expected Value Function*, or *Expected minimium recourse function* $\mathcal{Q} : \Re^n \mapsto \Re$.
  - ◇ $\mathcal{Q}(x) \equiv \mathbb{E}_\omega[v(h(\omega) - T(\omega)x)]$
  - ◇ For any policy $x \in \Re^n$, it describes the expected cost of the recourse.

The SP Problem

- Using these definitions, we can write our recourse problem in terms only of the x variables:

$$\min_{x \in X} \{c^T x + \mathcal{Q}(x) : Ax = b\}$$

- This is a (nonlinear) programming problem in $\Re^n$.

⇒ The ease of solving such a problem depends on the properties of $\mathcal{Q}(x)$.

? Does anyone know what $\mathcal{Q}(x)$ is?

  - ◇ Linear?
  - ◇ Convex?
  - ◇ Continuous?
  - ◇ Differentiable?

Writing With the y's

$$\min_{x \in \Re^n, y(\omega) \in \Re^p} \mathbb{E}_\omega \left[c^T x + q^T y(\omega) \right]$$

subject to

$$\begin{array}{lclcll} Ax & & & = & b & & \text{First Stage Constraints} \\ T(\omega)x & + & Wy(\omega) & = & h(\omega) & \forall \omega \in \Omega & \text{Second Stage Constraints} \\ x \in X & & y(\omega) \in Y & & & & \end{array}$$

- Imagine the case where $\Omega = \{\omega_1, \omega_2, \ldots \omega_S\} \subseteq \Re^r$.
- $\mathrm{P}(\omega = \omega_s) = p_s, \forall s = 1, 2, \ldots, S$
- $T_s \equiv T(\omega_s), h_s = h(\omega_s)$

Deterministic Equivalent

- We can then write the *deterministic equivalent* as:

$$
\begin{array}{cccccccccccc}
c^T x & + & p_1 q^T y_1 & + & p_2 q^T y_2 & + & \cdots & + & p_s q^T y_s & & \\
\text{s.t.} & & & & & & & & & & \\
Ax & & & & & & & & & = & b \\
T_1 x & + & W y_1 & & & & & & & = & h_1 \\
T_2 x & & & + & W y_2 & & & & & = & h_2 \\
\vdots & & & + & & & \ddots & & & & \\
T_S x & & & & & & & + & W y_s & = & h_s \\
x \in X & & y_1 \in Y & & y_2 \in Y & & & & y_s \in Y & &
\end{array}
$$

About the DE

- $y_s \equiv y(\omega_s)$ is the recourse action to take if scenario ω_s occurs.
- Pro: It's a linear program.
- Con: It's a BIG linear program.
  - ◇ $n + pS$ variables
  - ◇ $m_1 + mS$ constraints.
- Pro: The matrix of the linear program has a very special (staircase) structure.
  - ? Has anyone heard of Bender's Decomposition?

What is BIG

We have r random variables (That is why $\Omega \in \Re^r$)

- Imagine the following (real) problem. A Telecom company wants to expand its network in a way in which to meet an unknown (random) demand.
- There are 86 unknown demands. Each demand is independent and may take on one of seven values.
- $S = |\Omega| = \Pi_{k=1}^{86}(5) = 5^{86} = 4.77 \times 10^{72}$
  - ◇ The number of subatomic particles in the universe.

? How do we solve a problem that has more variables and more constraints than the number of subatomic particles in the universe?

But Its Even Worse!

- If Ω is not a countable set (say if it is made up of continuous-valued random variables, our "deterministic equivalent" would have ∞ variables and constraints. :-)
- The answer is we can't!
- We solve an approximating problem obtained through sampling.
  - ◇ We'll talk more about this later in the course

An Example

Let's solve a deterministic equivalent version of our little problem...
minimize

$$x_1 + x_2$$

subject to

$$\begin{aligned} \omega_1 x_1 + x_2 &\geq 7 \\ \omega_2 x_1 + x_2 &\geq 4 \\ x_1 &\geq 0 \\ x_2 &\geq 0 \end{aligned}$$

- $\omega_1 \sim \mathcal{U}[1, 4]$
- $\omega_2 \sim \mathcal{U}[1/3, 1]$

A Recourse Formulation

- Imagine for a moment that Ω was countable, with a finite set of scenarios S.

minimize

$$x_1 + x_2 + \sum_{s \in S} p_s \lambda (y_{1s} + y_{2s})$$

subject to

$$\begin{aligned}
\omega_{1s} x_1 + x_2 + y_{1s} &\geq 7 \qquad \forall s \in S \\
\omega_{2s} x_1 + x_2 + y_{2s} &\geq 4 \qquad \forall s \in S \\
x_1 &\geq 0 \\
x_2 &\geq 0 \\
y_{1s} &\geq 0 \\
y_{2s} &\geq 0
\end{aligned}$$

AMPL – 1

```
param n := 50;
set S := 1 .. n;
param p{s in S} default 1/card(S);
param w1{S} := Uniform(1,4);
param w2{S} := Uniform(1/3,1);

param PENALTY := 5;

var x1 >= 0;
var x2 >= 0;

var y1{S} >= 0;
var y2{S} >= 0;
```

AMPL – 2

```
minimize ObjPlusRecourse:
 x1 + x2 + sum{s in S} p[s] * PENALTY * (y1[s] + y2[s]);

subject to c1{s in S}:
w1[s] * x1 + x2 + y1[s] >= 7;

subject to c2{s in S}:
w2[s] * x1 + x2 + y2[s] >= 4;
```

Hot Off the Presses

- Since many of you are interested in supply chain, I would be derelict if I didn't mention the newsvendor problem.
- A paperboy (newsvendor) needs to decide how many papers to buy in order to maximize his profit.

⋆ He doesn't know at the beginning of the day how many papers he can sell (his demand).

  ◇ Each newspaper costs c.

  ◇ He can sell each newspaper for a price of q.

  ◇ He can return each unsold newspaper at the end of the day for r.

? Given only knowledge of the probability distribution $F(t) = \mathrm{P}(\omega \leq t)$, how may papers should the newsvendor buy to maximize his profits?

Newsvendor, Cont.

- According to our recourse definitions, the newsvendor would like to solve the following optimization problem.

$$\max_{x \geq 0} \{-cx + \mathcal{Q}(x)\}$$

- $\mathcal{Q}(x)$ is the expected amount of money the newsvendor can make if he purchases x newspapers:

$$\mathcal{Q}(x) = \mathbb{E}_\omega Q(x, \omega)$$

- This is some more notation. You will often see $Q(x, \omega) = v(h(\omega) - T(\omega)x)$.

Newsvendor, Cont.

- Here $Q(x, \omega)$ is the amount of money the newsvendor makes if he purchases x papers and demand is ω.
- For this problem, we don't need to formulate a linear program (although you can see how in BL). Let's just reason it out...

Let's convince ourselves that...

$$Q(x, \omega) = \begin{cases} qx & x \leq \omega \\ q\omega + r(x - \omega) & x \geq \omega \end{cases}$$

Calculating $\mathcal{Q}(x)$ – Ugly Math

$$\begin{aligned}
\mathcal{Q}(x) \equiv \mathbb{E}_\omega Q(x,\omega) &= \int_{-\infty}^{\infty} Q(x,\omega) dF(\omega) \\
&= \int_{\omega=-\infty}^{x} (q\omega + r(x-\omega) dF(\omega) + \int_{\omega=x}^{\infty} qx dF(\omega)
\end{aligned}$$

$$\begin{aligned}
&= (q-r)\int_{\omega=-\infty}^{x} \omega dF(\omega) + rx\int_{\omega=-\infty}^{x} dF(\omega) + qx\int_{\omega=x}^{\infty} dF(\omega) \\
&= (q-r)\int_{\omega=-\infty}^{x} \omega dF(\omega) + rxF(x) + qx(1-F(x))
\end{aligned}$$

All About $\int$

? What the heck is $\int g(x)dF(x)$?

- How many people know what a Lebesgue-Stieltjes integral is?
  - (Me neither!)
- Interpret the integral that you see here (and likely in any of the papers you will read) in the following way...
- If F is continuous
  - Which means $F(x) = \int f(x)dx$, then
  - $\int g(x)dF(x) = \int g(x)f(x)dx$

All About $\int$

- If F is discrete.
  - So there exists "atoms" a_i and "weights" w_i so that $F(x) = \sum_{i:a_i \leq x} w_i$
  - $\int g(x) dF(x) = \sum_i g(a_i) w_i$

⋆ You can also combine the two if F is a combination of a continuous and discrete function.

Integrate by Parts – I Learned That LONG Ago

- If $F(t)$ is "nice"
  - $(\lim_{t \to -\infty} tF(t) = 0)$
- We can integrate by parts to get...

$$\begin{aligned}\int_{\omega=-\infty}^{x} \omega dF(\omega) &= \omega F(\omega)|_{\omega=-\infty}^{x} - \int_{\omega=-\infty}^{x} F(\omega) d\omega \\ &= xF(x) - \int_{\omega=-\infty}^{x} F(\omega) d\omega\end{aligned}$$

Putting it All Together

$$\mathcal{Q}(x) = qx - (q - r) \int_{\omega=-\infty}^{x} F(\omega) d\omega$$

- Why did we do this exercise?
  - ◇ "To get to the other side"
  - ◇ Also, to help out the newsvendor
- So we need to optimize $-cx + \mathcal{Q}(x)$.

? How many people know what the KKT-conditions are?

- ◇ They are conditions under which we can ensure that a given solution $\hat{x}$ is an optimal solution.

Helping Out the Newsvendor

- The KKT conditions for this problem are especially simple.

⋆ We take the first derivative of the objective function and set it equal to 0

$$\mathcal{Q}'(x) = q - (q - r)F(x)$$

- So, the optimal solution satisfies...

$$-c + q - (q - r)F(x) = 0$$

- x^* is optimal when $F(x) = \frac{q-c}{q-r}$

$$x^* = F^{-1}\left(\frac{q - c}{q - r}\right)$$

An Example

- $c = 0.15$
- $q = 0.25$
- $r = 0.02$
- $\omega \sim \mathcal{N}(650, 80)$.

$$x^* = \mathcal{N}^{-1}(0.1/0.23) = 636.8631378336536954520852304995 05$$

All That Math for Nothing?!?!?!

- Just to show you that math is useless (just kidding), let's arrive at the same formula arguing from a more intuitive approach.
- Let's ask the question (for the newsvendor), suppose we have bought t newspapers, what is the expected marginal revenue of buying one more?
- From an economic viewpoint, we would like this marginal revenue ($MR(t)$) to be 0. (Just like the KKT conditions say).

A Can't Believe He Made Me Do All Those Integrals

$$\begin{aligned} MR(t) &= -c + q\mathrm{P}(\omega \geq t) + r\mathrm{P}(\omega \leq t) \\ &= -c + q(1 - F(t)) + rF(t) \end{aligned}$$

- Doing the math, we see that

$$MR(t) = 0 \Leftrightarrow F(t) = \left(\frac{q-c}{q-r}\right)$$

- So the optimal solution is to buy newspapers until $t = x^* = F^{-1}(\frac{q-c}{q-r})$

Next time

- Two more modeling examples
- Then that's it for modeling (for the time being).
- You should definitely have read most if not all of the first two chapters. (Come see me if you have questions).
  - ◇ In particular, 2.5 and 2.8 have interesting material that I probably won't cover (at least now)
  - ◇ 1.3 and 1.4 are other modeling examples I won't cover explicitly either.