IE 495 – Lecture 3

Stochastic Programming Modeling

Prof. Jeff Linderoth

January 20, 2003

Outline

- Review convexity
- Review Farmer Ted
- *Expected Value of Perfect Information*
- *Value of the Stochastic Solution*
- Building the Deterministic Equivalent
  - ◇ In an algebraic modeling language
- Formal notation
- More examples

Please don't call on me!

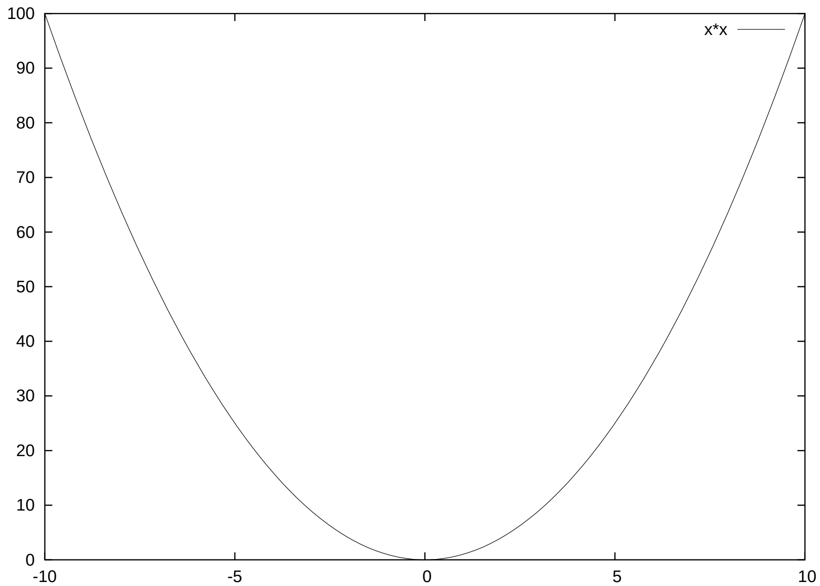
- Name one way in which to deal with randomness in mathematical programming problems.
- Name another way.
- Name *another way*
- A set C is convex if and only if...
- A function f is convex if and only if...
- What does Farmer Ted like to grow?

For the Math Lovers Out There...

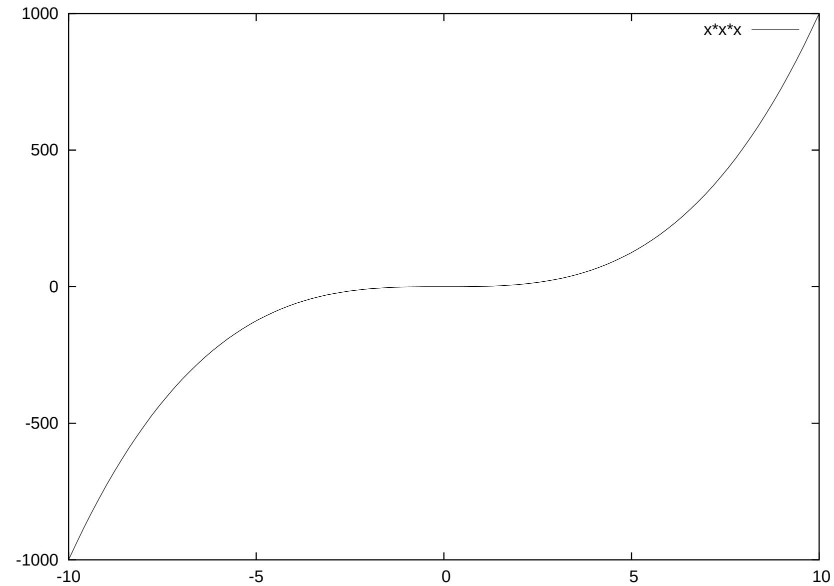
- It is *extremely* important to understand the convexity properties of a function you are trying to optimize.

- A function $f : \Re^n \rightarrow \Re$ is *convex* if for any two points x and y, the graph of f lies below or on the straight line connecting $(x, f(x))$ to $(y, f(y))$ in $\Re^{n+1}$.
  - $f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y) \qquad \forall 0 \leq \alpha \leq 1$
- A function $f : \Re^n \rightarrow \Re$ is *concave* if for any two points x and y, the graph of f lies above or on the straight line connecting $(x, f(x))$ to $(y, f(y))$ in $\Re^{n+1}$.
  - $f(\alpha x + (1-\alpha)y) \geq \alpha f(x) + (1-\alpha)f(y) \qquad \forall 0 \leq \alpha \leq 1$
- A function that is neither convex nor concave, we will call *nonconvex*.

CONVEX



NONCONVEX

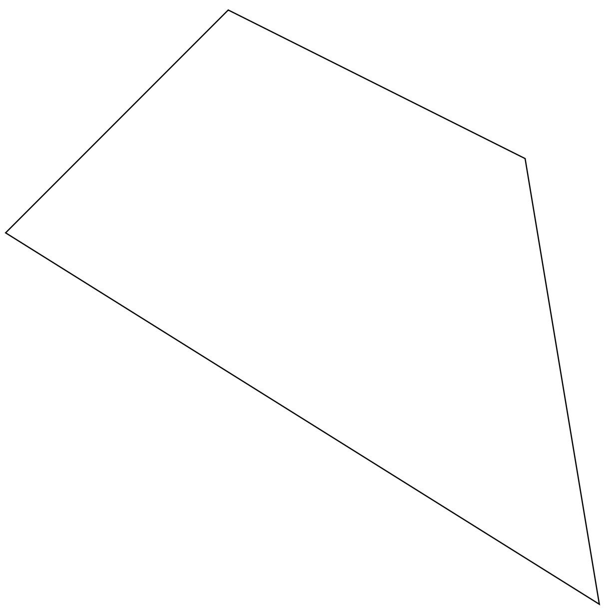


Convexity – Again. Ugh!

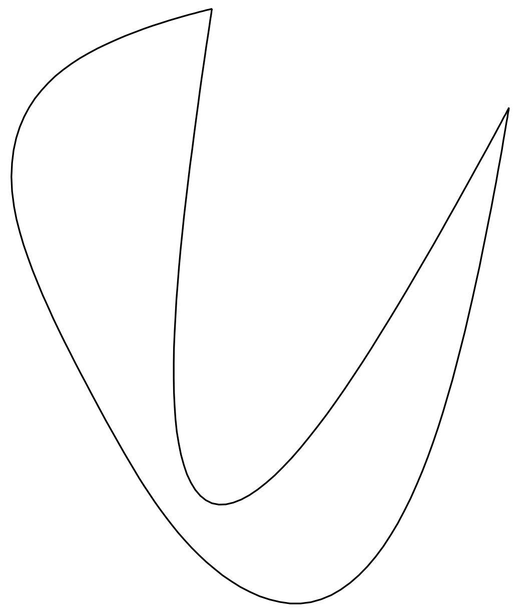
- A set S is *convex* if the straight line segment connecting any two points in S lies entirely inside or on the boundary of S.
  - ◇ $x, y \in S \Rightarrow \alpha x + (1-\alpha)y \in S \qquad \forall 0 \leq \alpha \leq 1$

- A Confusing Point...
  - ◇ Why do they have a *convex function* and a *convex set*? How are they related?
  - ◇ f is convex if and only if the *epigraph*, or "over part" of f is a convex set.

CONVEX



NONCONVEX



True or False

- Discrete Constraint Sets are convex?
- Empty Constraint Sets are convex?
- Discontinuous functions are convex?

Recall Farmer Ted

- Farmer Ted can grow Wheat, Corn, or Beans on his 500 acres.
- Farmer Ted requires 200 tons of wheat and 240 tons of corn to feed his cattle
  - These can be grown on his land or bought from a wholesaler.
  - Any production in excess of these amounts can be sold for \$170/ton (wheat) and \$150/ton (corn)
  - Any shortfall must be bought from the wholesaler at a cost of \$238/ton (wheat) and \$210/ton (corn).
- Farmer Ted can also grow beans
  - Beans sell at \$36/ton for the first 6000 tons
  - Due to economic quotas on bean production, beans in excess of 6000 tons can only be sold at \$10/ton

Formulate the LP – Decision Variables

- $x_{W,C,B}$ Acres of Wheat, Corn, Beans Planted
- $w_{W,C,B}$ Tons of Wheat, Corn, Beans sold (at favorable price).
- e_B Tons of beans sold at lower price
- $y_{W,C}$ Tons of Wheat, Corn purchased.

Formulation

maximize

$$-150x_W - 230x_C - 260x_B - 238y_W + 170w_W - 210y_C + 150y_C + 36w_B + 10e_B$$

subject to

$$\begin{aligned}
x_W + x_C + x_B &\leq 500 \\
2.5x_W + y_W - w_W &= 200 \\
3x_C + y_C - w_C &= 240 \\
20x_B - w_B - e_B &= 0 \\
w_B &\leq 6000 \\
x_W, x_C, x_B, y_W, y_C, e_B, w_W, w_C, w_B &\geq 0
\end{aligned}$$

Randomness

- Farmer Ted knows he doesn't get the yields Y all the time.
- Assume that three yield scenarios $(1.2Y, Y, 0.8Y)$ occur with equal probability.
- Maximize *Expected* Profit
- Attach a scenario subscript $s = 1, 2, 3$ to each of the purchase and sale variables.
  - ◇ 1: Good, 2: Average, 3: Bad

Ex. w_{C2} : Tons of corn sold at favorable price in scenario 2

Ex. e_{B3} : Tons of beans sold at unfavorable price in scenario 3.

Expected Profit

- An expression for Farmer Ted's Expected Profit is the following:

$$
\begin{aligned}
& -150x_W - 230x_C - 260x_B \\
& +1/3(-238y_{W1} + 170w_{W1} - 210y_{C1} + 150y_{C1} + 36w_{B1} + 10e_{B1}) \\
& +1/3(-238y_{W2} + 170w_{W2} - 210y_{C2} + 150y_{C2} + 36w_{B2} + 10e_{B2}) \\
& +1/3(-238y_{W3} + 170w_{W3} - 210y_{C3} + 150y_{C3} + 36w_{B3} + 10e_{B3})
\end{aligned}
$$

Expected Value Problem – Constraints

$$
\begin{aligned}
x_W + x_C + x_B &\leq 500 \\
3x_W + y_{W1} - w_{W1} &= 200 \\
2.5x_W + y_{W2} - w_{W2} &= 200 \\
2x_W + y_{W3} - w_{W3} &= 200 \\
3.6x_C + y_{C1} - w_{C1} &= 240 \\
3x_C + y_{C2} - w_{C2} &= 240 \\
2.4x_C + y_{C3} - w_{C3} &= 240 \\
24x_B - w_{B1} - e_{B1} &= 0 \\
20x_B - w_{B2} - e_{B2} &= 0 \\
16x_B - w_{B3} - e_{B3} &= 0 \\
w_{B1}, w_{B2}, w_{B3} &\leq 6000 \\
\text{All vars} &\geq 0
\end{aligned}
$$

Optimal Solution

	Wheat	Corn	Beans	
s	Plant (acres)	100	25	375
1	Production	510	288	6000
1	Sales	310	48	6000
1	Purchase	0	0	0
2	Production	425	240	5000
2	Sales	225	0	5000
2	Purchase	0	0	0
3	Production	340	192	4000
3	Sales	140	0	4000
3	Purchase	0	48	0

- (Expected) Profit: $108390

DE

- Congratulations, we've just solved our first stochastic program.
- What we've done is known as forming (and solving) the *deterministic equivalent* of a stochastic program
- Note that you can always do this when...
  - ◇ Ω is a finite set. (There are a finite number of scenarios $\omega_1, \omega_2, \ldots \omega_K \in \Omega$)
  - ◇ We are interested in optimizing an expected value.

⇒ We can write $\mathbb{E}_\omega f(x, \omega)$ as $\sum_{k=1}^{K} p_k f(x, \omega_k)$

Wait and See

- Recall from last time, that Farmer Ted also "ran some scenarios"
- *Given that he knew the yields*, what was his best policy?
  - We called these "Wait-and-see" solutions

	$0.8Y$	Y	$1.2Y$
Corn	25	80	66.67
Wheat	100	120	183.33
Beans	375	300	250
Profit	59950	118600	167667

Fortune Tellers

- Suppose Farmer Ted could *with certainty* tell whether or not the upcoming growing season was going to be wet, average, or dry (or what his yields were going to be).
  - ◇ His bursitits was acting up
  - ◇ Consulting the Farmer's Almanac
  - ◇ Hiring a fortune teller
- The real point here is how *much* Farmer Ted would be willing to pay for this "perfect" information.

⋆ In real-life problems, how much is it "worth" to invest in better forecasting technology?

- This amount is called *The Expected Value of Perfect Information*.

What is the EVPI?

- With perfect information, Farmer Ted's Long Run Profit/Year would be:
  - ◇ $(1/3)(167667) + (1/3)(118600) + (1/3)(59950) = 115406$
- Without perfect information, Farmer Ted can at best maximize his expected profit by solving the stochastic program.
- In this case, he would make 108390 in the long run
  - ⋆ EVPI = 115406 - 108390 = 7016.
- Is there any other important information that you would like to know?
  - ◇ What is the *value* of including the randomess?

The Value of the Stochastic Solution (VSS)

- Suppose we just replaced the "random" quantities (the yields) by their mean values and solved that problem.

? Would we get the same expected value for the Farmer's profit?

- How can we check?
  - ◇ Solve the "mean-value" problem to get a first stage solution x. (A "policy").
  - ◇ Fix the first stage solution at that value x, and solve all the scenarios to see Farmer Ted's profit in each.
  - ◇ Take the weighted (by probability) average of the optimal objective value for each scenario

AMPL, Everyone?

- To do this, we'll use AMPL
- You are welcome to solve problems anyway you can
  - ◇ Except for copying/cheating
  - ⋆ An algebraic modeling language will be quite useful!
- Average AMPL proficiency was around 7, and minimium was 3, so I am going to assume everyone comfortable with AMPL.
- There are some AMPL pointers on the web page.
- I have one copy of the AMPL book I can loan out for brief periods.
- AMPL is all about algebraic notation, so lets convert Farmer Ted to a more algebraic description...

Algebraic FT

- Sets...
  - C : Set of crops
  - $D \subseteq C$: Set of crops that have quotas
  - $Q \subseteq C$: Set of crops that FT can purchase.
- Variables...
  - $x_c, \ c \in C$: Acres to allocate to c
  - $w_c, \ c \in C$: Amount of c to sell (at high price)
  - $y_c, \ c \in C : (y_c = 0 \ \forall \ c \in C \setminus Q)$: Amount of c to purchase
  - $e_c, \ c \in C : (e_c = 0 \ \forall \ c \in D)$: Amount of c to sell (at low price)

AMPL

(Showing off AMPL here)

Great, but This Class is called *Stochastic* Programming

- Here's how to create the deterministic equivalent...
- For each possible state of nature (scenario), formulate an appropriate LP model
- Combine these submodels into one "supermodel" making sure
  - ◇ The first-stage variables are common to all submodels
  - ◇ The second-stage variables in a submodel appear only in that submodel
  - ⋆ Do this by attaching a "scenario index" to the second stage variables and to the parameters that change in the different scenarios

Deterministic Equivalent

- Combine these submodels into one "supermodel" making sure
  - The first-stage variables are common to all submodels
  - The second-stage variables in a submodel appear only in that submodel

$$
\begin{aligned}
x_W + x_C + x_B &\leq 500 \\
3x_W + y_{W1} - w_{W1} &= 200 \\
2.5x_W + y_{W2} - w_{W2} &= 200 \\
2x_W + y_{W3} - w_{W3} &= 200
\end{aligned}
$$

Constraints (cont.)

$$
\begin{aligned}
3.6x_C + y_{C1} - w_{C1} &= 240 \\
3x_C + y_{C2} - w_{C2} &= 240 \\
2.4x_C + y_{C3} - w_{C3} &= 240 \\
24x_B - w_{B1} - e_{B1} &= 0 \\
20x_B - w_{B2} - e_{B2} &= 0 \\
16x_B - w_{B3} - e_{B3} &= 0 \\
w_{B1}, w_{B2}, w_{B3} &\leq 6000 \\
\text{All vars} &\geq 0
\end{aligned}
$$

Computing Farmer Ted's VSS

- Solve the "mean-value" problem to get a first stage solution x. (A "policy").
  - Mean yields $Y = (2.5, 3, 20)$
  - (We already solved this problem).
  - $x_W = 120, x_C = 80, x_B = 300$
- Fix the first stage solution at that value x, and solve all the scenarios to see Farmer Ted's profit in each.
- Take the weighted (by probability) average of the optimal objective value for each scenario

Fixed Policy – Average Yield Scenario

maximize

$$-150x_W - 230x_C - 260x_B - 238y_W + 170w_W - 210y_C + 150y_C + 36w_B + 10e_B$$

subject to

$$\begin{aligned}
x_W &= 120 \\
x_C &= 80 \\
x_B &= 300 \\
x_W + x_C + x_B &\leq 500 \\
2.5x_W + y_W - w_W &= 200 \\
3x_C + y_C - w_C &= 240 \\
20x_B - w_B - e_B &= 0 \\
w_B &\leq 6000 \\
x_W, x_C, x_B, y_W, y_C, e_B, w_W, w_C, w_B &\geq 0
\end{aligned}$$

Fixed Policy – Bad Yield Scenario

maximize

$$-150x_W - 230x_C - 260x_B - 238y_W + 170w_W - 210y_C + 150y_C + 36w_B + 10e_B$$

subject to

$$\begin{aligned}
x_W &= 120 \\
x_C &= 80 \\
x_B &= 300 \\
x_W + x_C + x_B &\le 500 \\
2x_W + y_W - w_W &= 200 \\
2.4x_C + y_C - w_C &= 240 \\
16x_B - w_B - e_B &= 0 \\
w_B &\le 6000 \\
x_W, x_C, x_B, y_W, y_C, e_B, w_W, w_C, w_B &\ge 0
\end{aligned}$$

Fixed Policy – Good Yield Scenario

maximize

$$-150x_W - 230x_C - 260x_B - 238y_W + 170w_W - 210y_C + 150y_C + 36w_B + 10e_B$$

subject to

$$\begin{aligned}
x_W &= 120 \\
x_C &= 80 \\
x_B &= 300 \\
x_W + x_C + x_B &\leq 500 \\
3x_W + y_W - w_W &= 200 \\
3.6x_C + y_C - w_C &= 240 \\
24x_B - w_B - e_B &= 0 \\
w_B &\leq 6000 \\
x_W, x_C, x_B, y_W, y_C, e_B, w_W, w_C, w_B &\geq 0
\end{aligned}$$

Profits

- If you solved those three problems, you would get

Yield	Profit
Average	118600
Bad	55120
Good	148000

⋆ Another trick – you don't need to solve all three. Just solve the DE with the first stage fixed.

- I'll show you this if we have time.

What's it Worth to Model Randomness?

- If Farmer Ted implemented the policy based on using only "average" yields, he would plant $x_W = 120, x_C = 80, x_B = 300$
- He would expect in the long run to make an average profit of...
  - ◇ $1/3(118600) + 1/3(55120) + 1/3(148000) = 107240$
- If Farmer Ted implemented the policy based on the solution to the stochastic programming problem, he would plant $x_W = 170, x_C = 80, x_B = 250$.
  - ◇ From this he would expect to make 108390

VSS

- The difference of the values 180390-107240 is the *Value of the Stochastic Solution* : \$1150.
  - It would pay off \$1150 per growing season for Farmer Ted to use the "stochastic" solution rather than the "mean value" solution.