OPER 732: Optimization under Uncertainty
Lecture 1: Introduction

Department of Statistical Sciences and Operations Research
Virginia Commonwealth University

Aug. 27, 2017

Today's Outline

- Course logistic
- What the course is about
- Some sample applications
- About me
- About you
- My expectations
- Basic concepts

Course logistic

- Class time: Monday and Wednesday, 9:30am-10:45am
- Office hour: Monday and Wednesday, 11:00am-11:59am, and by appointment
- Grade percentage:
    - Homework: 30%
    - Mid-term: 30% (Oct. 18th, in class)
    - Final Project: 30%
    - Participation: 10%
- My first gift for you: 5 free late days!
    - Optimize the way you use them!
- Slides and handouts are available before each lecture

Final Project

- Implementation-based (I believe that the best way of learning is through implementation)
  - I have a long list of potential projects
  - Best case: incorporate stochastic programming modeling into your current line of research
  - If you are a PhD student, and do not have a research topic yet, this could be your first research topic
- Presentation at the end of semester

Your logistic

- Textbooks:
  - Introduction to Stochastic Programming, second edition, by J. Birge and F. Louveaux. (You may have found an electronic version online, but a hard copy is recommended)
  - Robust optimization, by A. BenTal et.al., can be found at `http://www2.isye.gatech.edu/~nemirovs/FullBookDec11.pdf`
- Optimization: linear programming, modeling
- Statistics: random variables, expectation, statistical tests, sampling, etc.
- Modeling/Implementation skills: Julia (see folder on the course website: Course Documents/Julia Tutorials/)

Other references

1. Books (both available online)
   - *Lectures in Stochastic Programming, Modeling and Theory*, by Shapiro, Dentcheva, and Ruszczynski
   - *Stochastic Programming*, Kall and Wallace
2. Websites
   - Stochastic Programming Community (a little out of date, but a useful gallary): http://stoprog.org
   - A collection of videos on stochastic programming tutorials: search "ICSP2016" on Youtube

About me...

My primary research area is stochastic programming, including:

- Two-stage stochastic programs
- Multi-stage stochastic programs
- Stochastic integer programs
- Chance-constrained stochastic programs

Good news: I am the principal investigator of a research project on stochastic programming supported by NSF. I am looking for a PhD student to work with me as a research assistant.

- This means $>$ 2 years of full financial support + tuition support!
- Let me know if you are interested.

About you...

My expectations

1. This is a PhD level course, please be prepared to DO SOME STUFF
2. We only have ξ people ($\xi \sim \mathcal{N}(3, 1)$, estimated according to the enrollment at 9:34pm on Aug. 27th), so please stop me ANY time you have a question
3. This is not a math class, this is (mostly) about modeling and algorithms
4. You are **encouraged** to work together on homework assignments, but you have to write up your solutions independently
5. You can ask me about the homework questions, and I will try to share my answer with the whole class

Let's get started!

My expectations

1. This is a PhD level course, please be prepared to DO SOME STUFF
2. We only have ξ people ($\xi \sim \mathcal{N}(3,1)$, estimated according to the enrollment at 9:34pm on Aug. 27th), so please stop me ANY time you have a question
3. This is not a math class, this is (mostly) about modeling and algorithms
4. You are **encouraged** to work together on homework assignments, but you have to write up your solutions independently
5. You can ask me about the homework questions, and I will try to share my answer with the whole class

Let's get started!

My expectations

1. This is a PhD level course, please be prepared to DO SOME STUFF
2. We only have ξ people ($\xi \sim \mathcal{N}(3, 1)$, estimated according to the enrollment at 9:34pm on Aug. 27th), so please stop me ANY time you have a question
3. This is not a math class, this is (mostly) about modeling and algorithms
4. You are **encouraged** to work together on homework assignments, but you have to write up your solutions independently
5. You can ask me about the homework questions, and I will try to share my answer with the whole class

Let's get started!

Optimization: three components

1. Objective: profit, time, energy, cost...
2. Decision variables: objective is a function of decision variables , which are to be determined by decision makers (you)
3. Constraints: certain rules these variables should satisfy

Huge assumption: The data involved in the model is known/given/deterministic

- Too good to be true in many cases!

Optimization: three components

1. Objective: profit, time, energy, cost...
2. Decision variables: objective is a function of decision variables , which are to be determined by decision makers (you)
3. Constraints: certain rules these variables should satisfy

Huge assumption: The data involved in the model is known/given/deterministic

- Too good to be true in many cases!

Optimization: three components

1. Objective: profit, time, energy, cost...
2. Decision variables: objective is a function of decision variables , which are to be determined by decision makers (you)
3. Constraints: certain rules these variables should satisfy

Huge assumption: The data involved in the model is known/given/deterministic

- Too good to be true in many cases!

Optimization: three components

1. Objective: profit, time, energy, cost...
2. Decision variables: objective is a function of decision variables , which are to be determined by decision makers (you)
3. Constraints: certain rules these variables should satisfy

Huge assumption: The data involved in the model is known/given/deterministic

- Too good to be true in many cases!

Optimization under uncertainty?

When parameters in the model is unknown/unreliable/uncertain...

- 100 year disaster planning: (almost) everything is unknown
- Transportation: airline schedules may be unreliable, traffic pattern is uncertain
- Power system: wind power is naturally uncertain
- Others?

In summary, we need to deal with randomness in the model

How to deal with randomness?

You may deal with randomness by:

- Sensitivity analysis: first fix the parameters, then see how solutions change as these parameters change
  - You can get some idea how “stable” a solution is
  - But you may not be able to do anything about it
- Try different realizations of the random variables, and get different solutions
  - But what if you have to commit to one decision?

This course focuses on stochastic programming and robust optimization models, but there are many different models you may use

A syllabus problem (aka the newsvendor problem)

- Prof. Song needs to decide how many syllabus to print out in order to maximize his happiness.
- He does not know before the class how many students are going to show up (a random number).
  - Each syllabus costs c.
  - Each syllabus handed out will add $s > c$ to Prof. Song's happiness.
  - Prof. Song can use the remaining syllabus to fold a paper plane for Christopher, which is added to his happiness $r < c$.
  - The class size is D, a random variable, suppose its cdf is $H(t) = \mathbb{P}[D \leq t]$.

Song's happiness function

$$F(x, D) = \begin{cases} (s-c)x & \text{if } x \leq D \\ sD + r(x-D) - cx & \text{if } x > D. \end{cases}$$

A syllabus problem (aka the newsvendor problem)

- Prof. Song needs to decide how many syllabus to print out in order to maximize his happiness.
- He does not know before the class how many students are going to show up (a random number).
  - Each syllabus costs c.
  - Each syllabus handed out will add $s > c$ to Prof. Song's happiness.
  - Prof. Song can use the remaining syllabus to fold a paper plane for Christopher, which is added to his happiness $r < c$.
  - The class size is D, a random variable, suppose its cdf is $H(t) = \mathbb{P}[D \leq t]$.

Song's happiness function

$$F(x, D) = \begin{cases} (s-c)x & \text{if } x \leq D \\ sD + r(x-D) - cx & \text{if } x > D. \end{cases}$$

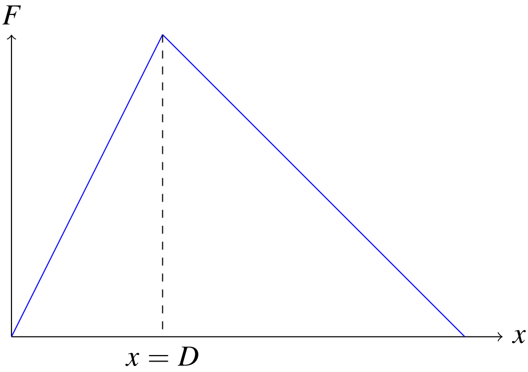
A syllabus problem (aka the newsvendor problem)

- Prof. Song needs to decide how many syllabus to print out in order to maximize his happiness.
- He does not know before the class how many students are going to show up (a random number).
  - Each syllabus costs c.
  - Each syllabus handed out will add $s > c$ to Prof. Song's happiness.
  - Prof. Song can use the remaining syllabus to fold a paper plane for Christopher, which is added to his happiness $r < c$.
  - The class size is D, a random variable, suppose its cdf is $H(t) = \mathbb{P}[D \leq t]$.

Song's happiness function

$$F(x, D) = \begin{cases} (s - c)x & \text{if } x \leq D \\ sD + r(x - D) - cx & \text{if } x > D. \end{cases}$$

Objective: maximize the happiness



- $s = 4, c = 2, r = 1$
- $D = 2$

- If I knew what D is, of course I will print out exactly D syllabus!
- But $\max_x F(x, D)$ does not make sense! How to maximize a random number?

An average person proposes an average idea

Plan for the average

- Let $\mu = \mathbb{E}[D]$, and maximize $\max_x F(x, \mu)$
- Solution $x^* = \mu$, I should print the mean value

- This solution may be far from the "optimal" solution

A pessimistic person proposes a pessimistic idea

Plan for the worst case

- Suppose $D \in [l, u]$
- After I make a decision x, I assume that the outcome will always turn out to be against me (after knowing this decision x)
- $\max_x \min_{D \in [l,u]} F(x, D)$

Note that $F(x, D) = \min\{(s - c)x, D(s - r) + (r - c)x\}$

$$\max_x \min_{D \in [l,u]} F(x, D) =$$

After this course, you will propose the following idea

Treat D as a random variable

Stochastic programming: incorporate the entire distribution information of D, e.g., its cdf, into the model, either in the objective or in the constraint

- The "average" idea only uses the first-order moment (expectation) information of D

---

1. Optimize the expectation:
   - $\max_x \mathbb{E}[F(x, D)]$ (Note: different than $\max_x F(x, \mathbb{E}[D])$!)
   - Optimize the happiness in the long run

2. Control the probability of getting a certain level b of profit:
   - $\min\{x \mid \mathbb{P}[F(x, D) \geq b] \geq 1 - \epsilon\}$, where ϵ is the risk tolerance, typically small, e.g., 0.05
   - Chance-constrained stochastic program

Random variable

- A random variable ξ on a probability space (Ω, Σ, P) is a real-valued function $\xi(w)$ that $\{w \mid \xi(w) \leq x\}$ is an event for all x
- Cumulative distribution function (cdf) $F_\xi(x) = \mathbb{P}[\xi \leq x]$
- Discrere random variables that have a finite support: $\{\xi_k\}_{k \in K}$
  - Density: $f(\xi_k) = \mathbb{P}[\xi = \xi_k]$
  - $\sum_{k \in K} f(\xi_k) = 1$
- Continuous random variables have density $f(\xi)$
  - $P(\xi = x) = 0$
  - The probability of ξ in an interval $[a, b]$ is given by:

$$\mathbb{P}[a \leq \xi \leq b] = \int_a^b f(\xi) d\xi = F(b) - F(a)$$

Expected value, variance

Expected value of ξ is:

- Discrete case: $\mathbb{E}[\xi] = \sum_{k \in K} \xi_k f(\xi_k)$
- Continuous case: $\mathbb{E}[\xi] = \int_{-\infty}^{\infty} \xi f(\xi) d\xi = \int_{-\infty}^{\infty} \xi dF(\xi)$

Variance of ξ is: $Var[\xi] = \mathbb{E}[\xi - \mathbb{E}[\xi]]^2$

Calculating $\mathbb{E}[F(x, D)]$

Assume D is a continuous random variable with cdf given by $H(t) = \mathbb{P}[D \leq t]$

Time to do some calculus: (not typical in this course! We will work primarily on discrete random variables or a finite number of sampled scenarios)

$$\mathbb{E}[F(x, D)] =$$

Finally....maximize $\mathcal{F}(x) = \mathbb{E}[F(x, D)]$

$$\max \ \mathcal{F}(x) = (s-c)x - (s-r)\int_{-\infty}^{x} H(D)dD$$

- This is an unconstrained smooth optimization problem
- Function $\mathcal{F}(x)$ is concave
- First order optimality condition: $\mathcal{F}'(x) = 0$

Let $\mathcal{F}'(x) = (s-c) - (s-r)H(x) = 0 \Rightarrow x^* = H^{-1}(\frac{s-c}{s-r})$
The optimal solution is the $(\frac{s-c}{s-r})$-percentile of distribution H

The scenario approach

- Encode the uncertainty using a finite discrete set of outcomes (scenarios)
  - In the newsvendor problem, we can just assume demand D follows $\mathbb{P}[D = d_k] = p_k$ with $\sum_{k \in N} p_k = 1$
  - Get rid of the integration! Hooray!
- Where are the scenarios from?
  - May come from a random sample from the true distribution! (In the newsvendor problem, sample from distribution H)
  - May come from historical data

Timeline of decisions

1. Some decisions have to be made now (first-stage), before the outcome reveals (here-and-now as oppose to wait-and-see)
   - e.g. Investment: you have to invest in a stock now, before you know if it goes up or down.
2. God throws a dice, some scenario (outcome) happens
3. After that (second-stage):
   - We do our best to repair the havoc brought by the nature
   - Or, nothing can be done to repair, we just evaluate the loss/shortfall/surplus... (simple recourse)

   The second-stage decision depends on the first-stage decision AND the outcome of randomness

---

A two-stage stochastic program with recourse

Optimize the first-stage cost/benefit + expected second-stage cost/benefit

The newsvendor problem is a two-stage stochastic program (with simple recourse)

1. First-stage decision: purchase x newspapers before customers come. Incur the first-stage benefit: $-cx$
2. Observe the randomness: D customers come
3. Second-stage decision: evaluate the first-stage decision
   - If $x \geq D$, second-stage benefit: $sD + r(x - D)$
   - If $x < D$, second-stage benefit: sx

Optimization problem:

$$\max_{x} \; -cx + \mathbb{E}[s \min\{x, D\} + r \max\{x - D, 0\}]$$

Fundemental assumption of stochastic programming

- Some data/parameters in the model is random, and is unknown before a solution is implemented
- The uncertainty is expressed by probability distribution
- The probability distribution of the random variables ξ in the model is known, e.g., from historical data, experts, physics
- The probability distribution is independent of the decisions
  - Otherwise, a very hard problem! Decision-dependent uncertainty

Reputation of stochastic programming

- Interesting, useful, nice theory, . . . , “the real problem” (Dantzig, father of linear programming)
- Difficult...(I agree!)
- Impractical...(depending on how you use it)
- Probability is just one model, but all we have is DATA!
  - A good argument for data-driven robust optimization

Summary

- This is not an easy course, but hopefully you will find it:
  - interesting when you are taking it;
  - and useful after you have taken it

- Reading:
  - Appendix math review (pp 94-97, 449-450 on textbook)
  - LP review handout
  - Setup JuliaBox account, and download IJulia notebook on your computer (following the instruction from *juliabook-preview.pdf* on the course website)

- Next time:
  - Stochastic programming modeling
  - Value of stochastic programs