

Optimization under Uncertainty

[Subtitle]

Vitor Luiz P. de P. Ferreira

École Nationale des Ponts et Chaussées

October 31, 2022

Outline

- 1 Introduction
- 2 Optimization under uncertainty
- 3 Optimization over time

Game

Option 1

Gain or lose between 10 and -10 euros.

Option 2

Gain or lose between 20 and -20 euros.

Game

Option 1

Gain or lose between 10 and -10 euros.

Gain 7 euros.

Option 2

Gain or lose between 20 and -20 euros.

Gain 3 euros.

Introduction

Decision $x \rightarrow$ Observation $\xi \rightarrow$ Decision y

Introduction

Decision x \rightarrow Observation ξ \rightarrow Decision y

- First-stage decision x ; (here-and-now)

Challenge: Evaluate consequences of the decision x .

Introduction

Decision $x \rightarrow$ Observation $\xi \rightarrow$ Decision y

- First-stage decision x ; (here-and-now)
- Uncertain quantity ξ ;

Challenge: How to model the uncertainty ξ .

Introduction

Decision $x \rightarrow$ Observation $\xi \rightarrow$ Decision y

- First-stage decision x ; (here-and-now)
- Uncertain quantity ξ ;
- Second-stage or recourse decision y . (wait-and-see)

Challenge: Recourse decision is a function of uncertainty $y = y(\xi)$.

Motivation

Decision-making over time:

- Generating electrical power using thermal engines or uncertain renewable resources (solar, wind, hydro);
- Planning factory and warehouse installation before knowing demand for products;
- Assigning jobs to people in order to meet a deadline.

Motivation

Decision-making over time:

- Generating electrical power using thermal engines or uncertain renewable resources (solar, wind, hydro);
- Planning factory and warehouse installation before knowing demand for products;
- Assigning jobs to people in order to meet a deadline.

What is the best decision you can make before all information is available?

How to find that decision?

Outline

- 1 Introduction
- 2 Optimization under uncertainty
- 3 Optimization over time

The (deterministic) Newsvendor Problem

Suppose you want to buy x newspapers in the morning to sell them along the day.

Each newspaper costs you c , while each sale earns you p .

The number of people interested in buying a newspaper is d .

We assume $0 < c < p$.

Newspapers sold: $\min \{x, d\}$;

Operational costs: $cx - p \min \{x, d\}$;

Operational constraints: $x \geq 0$;

$$\begin{aligned} \min_x \quad & cx - p \min \{x, d\} \\ \text{s.t.} \quad & x \geq 0. \end{aligned}$$

The (deterministic) Newsvendor Problem

Suppose you want to buy x newspapers in the morning to sell them along the day.

Each newspaper costs you c , while each sale earns you p .

The number of people interested in buying a newspaper is d .

We assume $0 < c < p$.

Newspapers sold: $\min \{x, d\}$;

Operational costs: $cx - p \min \{x, d\}$;

Operational constraints: $x \geq 0$;

$$\begin{aligned} \min_x \quad & cx - p \min \{x, d\} \\ \text{s.t.} \quad & x \geq 0. \end{aligned}$$

The (deterministic) Newsvendor Problem

Suppose you want to buy x newspapers in the morning to sell them along the day.

Each newspaper costs you c , while each sale earns you p .

The number of people interested in buying a newspaper is d .

We assume $0 < c < p$.

Newspapers sold: $\min \{x, d\}$;

Operational costs: $cx - p \min \{x, d\}$;

Operational constraints: $x \geq 0$;

$$\begin{aligned} \min_x \quad & cx - p \min \{x, d\} \\ \text{s.t.} \quad & x \geq 0. \end{aligned}$$

The (deterministic) Newsvendor Problem

Suppose you want to buy x newspapers in the morning to sell them along the day.

Each newspaper costs you c , while each sale earns you p .

The number of people interested in buying a newspaper is d .

We assume $0 < c < p$.

Newspapers sold: $\min \{x, d\}$;

Operational costs: $cx - p \min \{x, d\}$;

Operational constraints: $x \geq 0$;

$$\begin{aligned} \min_x \quad & cx - p \min \{x, d\} \\ \text{s.t.} \quad & x \geq 0. \end{aligned}$$

The (deterministic) Newsvendor Problem

Suppose you want to buy x newspapers in the morning to sell them along the day.

Each newspaper costs you c , while each sale earns you p .

The number of people interested in buying a newspaper is d .

We assume $0 < c < p$.

Newspapers sold: $\min \{x, d\}$;

Operational costs: $cx - p \min \{x, d\}$;

Operational constraints: $x \geq 0$;

$$\begin{aligned} \min_x \quad & cx - p \min \{x, d\} \\ \text{s.t.} \quad & x \geq 0. \end{aligned}$$

The (stochastic) Newsvendor Problem

Now, consider that the demand d is unknown when you buy newspapers.

A possible approach is to treat it as a random variable d with known law.

$$\begin{aligned} \min_x \quad & cx - p \min\{x, d\} \quad ? \\ \text{s.t.} \quad & x \geq 0. \end{aligned}$$

The (stochastic) Newsvendor Problem

Now, consider that the demand d is unknown when you buy newspapers.

A possible approach is to treat it as a random variable d with known law.

$$\begin{array}{ll} \min_x & cx - p \min\{x, d\} \\ \text{s.t.} & x \geq 0. \end{array} \quad ? \quad \longrightarrow \quad \begin{array}{ll} \min_x & \mathbb{E}[cx - p \min\{x, d\}] \\ \text{s.t.} & x \geq 0. \end{array}$$

The (stochastic) Newsvendor Problem

Now, consider that the demand d is unknown when you buy newspapers.

A possible approach is to treat it as a random variable d with known law.

$$\begin{array}{ll} \min_x & cx - p \min\{x, d\} \\ \text{s.t.} & x \geq 0. \end{array} \quad ? \quad \longrightarrow \quad \begin{array}{ll} \min_x & \mathbb{E}[cx - p \min\{x, d\}] \\ \text{s.t.} & x \geq 0. \end{array}$$

How to justify the expectation?

The (stochastic) Newsvendor Problem

Now, consider that the demand d is unknown when you buy newspapers.

A possible approach is to treat it as a random variable d with known law.

$$\begin{array}{ll} \min_x & cx - p \min \{x, d\} \\ \text{s.t.} & x \geq 0. \end{array} \quad ? \quad \longrightarrow \quad \begin{array}{ll} \min_x & \mathbb{E} [cx - p \min \{x, d\}] \\ \text{s.t.} & x \geq 0. \end{array}$$

How to justify the expectation?

If you repeat this day after day, the **Law of Large Numbers** guarantees the sample mean (for fixed $x = x$) converges:

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n cx - \min \{x, d_i\}}{n} \xrightarrow{\text{a.s.}} \mathbb{E} [cx - p \min \{x, d\}].$$

The (stochastic) Newsvendor Problem

Often it is possible to make a correction to your initial decision after observing the uncertainty.

Now, assume that each day you can buy an additional y newspapers at noon, when you have certainty of the day's demand. However, at noon newspapers cost you b , such that $c < b < p$.

$$\begin{aligned} \min_{x,y} \quad & \mathbb{E} [cx + by - p \min \{x + y, d\}] \\ \text{s.t.} \quad & x \geq 0, \\ & y \geq 0. \end{aligned}$$

The (stochastic) Newsvendor Problem

Often it is possible to make a correction to your initial decision after observing the uncertainty.

Now, assume that each day you can buy an additional y newspapers at noon, when you have certainty of the day's demand. However, at noon newspapers cost you b , such that $c < b < p$.

$$\begin{aligned} \min_{x,y} \quad & \mathbb{E} [cx + by - p \min \{x + y, d\}] \\ \text{s.t.} \quad & x \geq 0, \\ & y \geq 0, \\ & y = y(d). \end{aligned}$$

The second decision y can depend on the observed demand!
By $y \geq 0$ we mean $\mathbb{P}[y \geq 0] = 1$.

The (stochastic) Newsvendor Problem

Remarks:

- The decision x is not random, thus

$$\mathbb{E}[cx + \dots] = cx + \mathbb{E}[\dots];$$

- For each fixed x and d , define

$$\hat{V}(x, d) = \min_{y \geq 0} by - \min \{x + y, d\};$$

The (stochastic) Newsvendor Problem

Remarks:

- The decision x is not random, thus

$$\mathbb{E}[cx + \dots] = cx + \mathbb{E}[\dots];$$

- For each fixed x and d , define

$$\hat{V}(x, d) = \min_{y \geq 0} by - \min\{x + y, d\};$$

Substituting in our previous formulation, we obtain

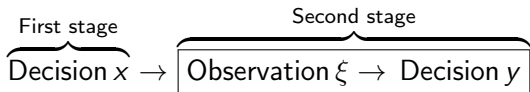
$$\begin{aligned} \min_x \quad & cx + \mathbb{E}[\hat{V}(x, d)] \\ \text{s.t.} \quad & x \geq 0. \end{aligned}$$

Decomposition

Decision $x \rightarrow$ Observation $\xi \rightarrow$ Decision y

Often, there is a decomposable structure: useful in modeling and algorithms.

Decomposition



Often, there is a decomposable structure: useful in modeling and algorithms.

The pieces are called **stages**. Commonly related to time—but not necessarily.

Thus, we also say that x is the **first-stage variable** while y is the **second-stage variable**.

The (stochastic) Newsvendor Problem

First-stage problem:

$$\begin{aligned} \min_x \quad & cx + \mathbb{E} \left[\hat{V}(x, d) \right] \\ \text{s.t.} \quad & x \geq 0. \end{aligned}$$

Second-stage problem:

$$\hat{V}(x, d) = \min_{y \geq 0} by - \min \{x + y, d\};$$

The (stochastic) Newsvendor Problem

First-stage problem:

$$\begin{aligned} \min_x \quad & cx + \mathbb{E} \left[\hat{V}(x, \mathbf{d}) \right] \\ \text{s.t.} \quad & x \geq 0. \end{aligned}$$

Second-stage problem:

$$\hat{V}(x, d) = \min_{y \geq 0} by - \min \{x + y, d\};$$

Solution techniques often employ:

- 1 **Approximation** of $\mathbb{E} \left[\hat{V}(x, \mathbf{d}) \right]$, improved iteratively by...

The (stochastic) Newsvendor Problem

First-stage problem:

$$\begin{aligned} \min_x \quad & cx + \mathbb{E} \left[\hat{V}(x, \mathbf{d}) \right] \\ \text{s.t.} \quad & x \geq 0. \end{aligned}$$

Second-stage problem:

$$\hat{V}(x, \mathbf{d}) = \min_{y \geq 0} by - \min \{x + y, \mathbf{d}\};$$

Solution techniques often employ:

- 1 Approximation of $\mathbb{E} \left[\hat{V}(x, \mathbf{d}) \right]$, improved iteratively by...
- 2 **Monte Carlo sampling** of a (tentative) second-stage solution. (Even when solving the second-stage exactly for given x , the first-stage decision can change after approximation improvement.)

The (robust) Newsvendor Problem

An alternative formulation is to assume the demand d belongs to a set D . For example: take $D = [\underline{d}, \bar{d}]$.

In this setting, we can't take the expectation— d isn't a random variable!

Instead, look at $d \in D$ that results in the worst outcome. This is a *pessimistic* approach.

The (robust) Newsvendor Problem

An alternative formulation is to assume the demand d belongs to a set D . For example: take $D = [\underline{d}, \bar{d}]$.

In this setting, we can't take the expectation— d isn't a random variable!

Instead, look at $d \in D$ that results in the worst outcome. This is a *pessimistic* approach.

The (robust) Newsvendor Problem

An alternative formulation is to assume the demand d belongs to a set D . For example: take $D = [\underline{d}, \bar{d}]$.

In this setting, we can't take the expectation— d isn't a random variable!

Instead, look at $d \in D$ that results in the worst outcome. This is a *pessimistic* approach.

The (robust) Newsvendor Problem

First, let us consider the problem without recourse. For a fixed x , the worst outcome is:

$$\max_{d \in D} \left(-p \min \{x, d\} \right)$$

The problem then is

$$\begin{aligned} \min_x \quad & cx + \max_{d \in D} \left(-p \min \{x, d\} \right) \\ \text{s.t.} \quad & x \geq 0. \end{aligned}$$

The (robust) Newsvendor Problem

First, let us consider the problem without recourse. For a fixed x , the worst outcome is:

$$\max_{d \in D} \left(-p \min \{x, d\} \right)$$

The problem then is

$$\begin{aligned} \min_x \quad & cx + \max_{d \in D} \left(-p \min \{x, d\} \right) \\ \text{s.t.} \quad & x \geq 0. \end{aligned}$$

The (robust) Newsvendor Problem

With recourse:

$$V(x) = \max_{d \in D} \min_{y \geq 0} \left(by - p \min \{x + y, d\} \right)$$

$$\min_x \quad cx + \max_{d \in D} \min_{y \geq 0} \left(by - p \min \{x + y, d\} \right)$$

$$\text{s.t. } x \geq 0.$$

The (robust) Newsvendor Problem

With recourse:

$$V(x) = \max_{d \in D} \min_{y \geq 0} \left(by - p \min \{x + y, d\} \right)$$

$$\begin{aligned} \min_x \quad & cx + \max_{d \in D} \min_{y \geq 0} \left(by - p \min \{x + y, d\} \right) \\ \text{s.t.} \quad & x \geq 0. \end{aligned}$$

Note that in the inner minimization, $y = y(d)$ can depend on d .
What does $y \geq 0$ mean?

The (robust) Newsvendor Problem

With recourse:

$$V(x) = \max_{d \in D} \min_{y \geq 0} \left(by - p \min \{x + y, d\} \right)$$

$$\begin{aligned} \min_x \quad & cx + \max_{d \in D} \min_{y \geq 0} \left(by - p \min \{x + y, d\} \right) \\ \text{s.t.} \quad & x \geq 0. \end{aligned}$$

Note that in the inner minimization, $y = y(d)$ can depend on d .
What does $y \geq 0$ mean? It means $y(d) \geq 0 \quad \forall d \in D$.

The (robust) Newsvendor Problem

With recourse:

$$V(x) = \max_{d \in D} \min_{y \geq 0} \left(by - p \min \{x + y, d\} \right)$$

$$\begin{aligned} \min_x \quad & cx + \max_{d \in D} \min_{y \geq 0} \left(by - p \min \{x + y, d\} \right) \\ \text{s.t.} \quad & x \geq 0. \end{aligned}$$

Note that in the inner minimization, $y = y(d)$ can depend on d .
What does $y \geq 0$ mean? It means $y(d) \geq 0 \quad \forall d \in D$.

When D is infinite, as in our case $D = [\underline{d}, \bar{d}]$, this results in a *semi-infinite problem*!

The (robust) Newsvendor Problem

Solving a robust optimization problem often requires a **reformulation** to merge min and max problems and represent an infinite number of constraints in a tractable manner. It can also involve an **approximation** of V instead of or in addition to.

Comparison

Robust optimization can be **more resilient** to bad outcomes. Also, it can employ **a simpler uncertainty model**. However, in comparison with stochastic optimization, it can **overestimate the cost**.

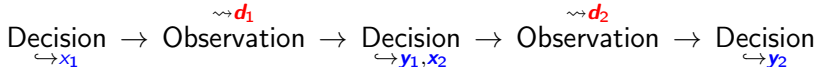
Outline

- 1 Introduction
- 2 Optimization under uncertainty
- 3 Optimization over time**

The (two-stage) Newsvendor Problem

Consider that your true problem isn't to buy-and-sell newspapers in a single day, but rather over multiple days.

For example, assume that you consider the stochastic problem over two days.



Newspapers bought in the morning of day t : x_t ;

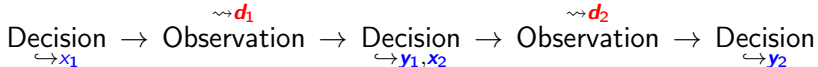
Newspapers bought in the afternoon of day t : y_t ;

Demand for newspapers over day t : d_t .

The (two-stage) Newsvendor Problem

Consider that your true problem isn't to buy-and-sell newspapers in a single day, but rather over multiple days.

For example, assume that you consider the stochastic problem over two days.



Newspapers bought in the morning of day t : x_t ;

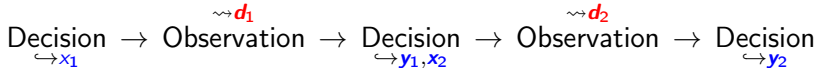
Newspapers bought in the afternoon of day t : y_t ;

Demand for newspapers over day t : d_t .

The (two-stage) Newsvendor Problem

Consider that your true problem isn't to buy-and-sell newspapers in a single day, but rather over multiple days.

For example, assume that you consider the stochastic problem over two days.



Newspapers bought in the morning of day t : x_t ;

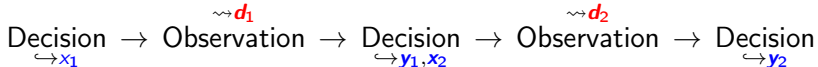
Newspapers bought in the afternoon of day t : y_t ;

Demand for newspapers over day t : d_t .

The (two-stage) Newsvendor Problem

Consider that your true problem isn't to buy-and-sell newspapers in a single day, but rather over multiple days.

For example, assume that you consider the stochastic problem over two days.



Newspapers bought in the morning of day t : x_t ;

Newspapers bought in the afternoon of day t : y_t ;

Demand for newspapers over day t : d_t .

The (two-stage) Newsvendor Problem

Problem formulation:

$$\min_{x_1, x_2, y_1, y_2} c x_1 + \mathbb{E} \left[b y_1 - p \min \{ x_1 + y_1, d_1 \} + \right. \\ \left. c x_2 + b y_2 - p \min \{ x_2 + y_2, d_2 \} \right]$$

$$\text{s.t. } x_1 \geq 0, x_2 \geq 0, y_1 \geq 0, y_2 \geq 0,$$

$$y_1 = y_1(d_1), x_2 = x_2(d_1), y_2 = y_2(d_1, d_2).$$

The (two-stage) Newsvendor Problem

Problem formulation:

$$\begin{aligned} \min_{x_1, x_2, y_1, y_2} \quad & c x_1 + \mathbb{E} \left[b y_1 - p \min \{ x_1 + y_1, d_1 \} + \right. \\ & \left. c x_2 + b y_2 - p \min \{ x_2 + y_2, d_2 \} \right] \\ \text{s.t.} \quad & x_1 \geq 0, x_2 \geq 0, y_1 \geq 0, y_2 \geq 0, \\ & y_1 = y_1(d_1), x_2 = x_2(d_1), y_2 = y_2(d_1, d_2). \end{aligned}$$

But why remember the demand d_1 when choosing how many newspapers to buy x_2 ?

The (two-stage) Newsvendor Problem

Problem formulation:

$$\begin{aligned} \min_{x_1, x_2, y_1, y_2} \quad & c x_1 + \mathbb{E} \left[b y_1 - p \min \{ x_1 + y_1, d_1 \} + \right. \\ & \left. c x_2 + b y_2 - p \min \{ x_2 + y_2, d_2 \} \right] \\ \text{s.t.} \quad & x_1 \geq 0, x_2 \geq 0, y_1 \geq 0, y_2 \geq 0, \\ & y_1 = y_1(d_1), x_2 = x_2(d_1), y_2 = y_2(d_1, d_2). \end{aligned}$$

But why remember the demand d_1 when choosing how many newspapers to buy x_2 ?

Past events can influence future (uncertain) events! (And our *prediction* of future events.)

Multistage problems

More generally, we can consider multistage problems with an arbitrary number of stages.

Decision \rightarrow Observation \rightarrow Decision \rightarrow Observation \rightarrow
Decision \rightarrow Observation $\rightarrow \dots \rightarrow$ Observation \rightarrow Decision

Multistage problems

More generally, we can consider multistage problems with an arbitrary number of stages.

Decision \rightarrow Observation \rightarrow Decision \rightarrow Observation \rightarrow
Decision \rightarrow Observation $\rightarrow \dots \rightarrow$ Observation \rightarrow Decision

We call a full sequence of observations a **scenario**. Each scenario has a corresponding probability of occurring (prob. density if continuous).

Multistage problems

Assume the random variables have finite support at each stage.
Thus, there are finitely many $S \in \mathbb{N}$ scenarios.

Multistage problems

Assume the random variables have finite support at each stage.
Thus, there are finitely many $S \in \mathbb{N}$ scenarios.
(How big is S if we have 2 options at each stage?)

Multistage problems

Assume the random variables have finite support at each stage.
Thus, there are finitely many $S \in \mathbb{N}$ scenarios.

(How big is S if we have 2 options at each stage?)

Let us look at what happens if we try to solve the problem directly.

Multistage problems

Assume the random variables have finite support at each stage.
Thus, there are finitely many $S \in \mathbb{N}$ scenarios.

(How big is S if we have 2 options at each stage?)

Let us look at what happens if we try to solve the problem directly.

Exploiting finiteness, we can construct a *deterministic* optimization problem equivalent to the original stochastic problem.

This is done by adding a decision variable for each scenario.

The (two-stage) Newsvendor Problem (revisited)

For example:

$$\begin{aligned} \min_{x,y} \quad & \sum_{s=1}^S p^s \left(cx_1^s + by_1^s - p \min \{x_1^s + y_1^s, d_1^s\} + \right. \\ & \left. cx_2^s + by_2^s - p \min \{x_2^s + y_2^s, d_2^s\} \right) \\ \text{s.t.} \quad & x \geq 0, y \geq 0. \end{aligned}$$

The (two-stage) Newsvendor Problem (revisited)

For example:

$$\begin{aligned} \min_{x,y} \quad & \sum_{s=1}^S p^s \left(cx_1^s + by_1^s - p \min \{x_1^s + y_1^s, d_1^s\} + \right. \\ & \left. cx_2^s + by_2^s - p \min \{x_2^s + y_2^s, d_2^s\} \right) \\ \text{s.t.} \quad & x \geq 0, y \geq 0. \end{aligned}$$

Is that all?

The (two-stage) Newsvendor Problem (revisited)

For example:

$$\begin{aligned} \min_{x,y} \quad & \sum_{s=1}^S p^s \left(cx_1^s + by_1^s - p \min \{x_1^s + y_1^s, d_1^s\} + \right. \\ & \left. cx_2^s + by_2^s - p \min \{x_2^s + y_2^s, d_2^s\} \right) \\ \text{s.t.} \quad & x \geq 0, y \geq 0. \end{aligned}$$

Is that all? No!

Nonanticipativity

We can't predict the future. Our decisions can only depend on past information.

Nonanticipativity

We can't predict the future. Our decisions can only depend on past information.

If two scenarios have the same **history** up to stage t , then they must make the same decision.

For example:

$$y_1^s = y_1^r \quad \forall s, r \text{ such that } d_1^s = d_1^r.$$

Nonanticipativity

We can't predict the future. Our decisions can only depend on past information.

If two scenarios have the same **history** up to stage t , then they must make the same decision.

For example:

$$y_1^s = y_1^r \quad \forall s, r \text{ such that } d_1^s = d_1^r.$$

These are **nonanticipativity** constraints.

(In particular, note that all x_1^s must be equal, because that decision is taken before any random variable is observed.)

Conclusion

Main aspects to consider.

- How to model uncertainty?

Conclusion

Main aspects to consider.

- How to model uncertainty?
- How to represent your attitude toward risk?

Conclusion

Main aspects to consider.

- How to model uncertainty?
- How to represent your attitude toward risk?
- What information structure to choose?

Questions?

Bibliography I

- [BEN09] Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski. *Robust optimization*. Vol. 28. Princeton university press, 2009.
- [Lec21] Vincent Leclère. *Two-stage stochastic program*. University lecture slides. 2021.
- [SDR21] Alexander Shapiro, Darinka Dentcheva, and Andrzej Ruszczyński. *Lectures on stochastic programming: modeling and theory*. SIAM, 2021.