Stochastic optimization
Two-stage stochastic programming with recourse

Fabian Bastin
fabian.bastin@umontreal.ca
Université de Montréal – CIRRELT – IVADO – Fin-ML

Formalization

Uncertainty: representation by means of random elements.
The realizations are denoted by ω, and they are drawn form the sample space Ω.

A event A is a subset of Ω; the collection of random of random events is denoted by $\mathcal{A}$. The event $A \in \mathcal{A}$ occurs if the output of the experiment is an element from A.

A random linear program

Consider the linear program (LP), parametrized by the random variable (r.v.) $\xi : \Omega \to \mathbb{R}^2$:

$$
\begin{aligned}
\min_x \ & c^T x \\
\text{s.t. } & Ax = b \\
& T(\xi)x = h(\xi) \\
& x \in X,
\end{aligned}
$$

with $X = \{x \in \mathbb{R}^n | l \leq x \leq u\}$. Example:

$$
\begin{aligned}
\min_x \ & x_1 + x_2 \\
\text{s.t. } & \xi_1 x_1 + x_2 \geq 7 \\
& \xi_2 x_1 + x_2 \geq 4 \\
& x_1, x_2 \geq 0,
\end{aligned}
$$

where $\xi_1 \sim U[1, 4]$, $\xi_2 \sim U[1/3, 1]$.

What to do?

- How to solve this problem?
- What is the *meaning* of solving this problem?
- Can we decide on x *after* having observed the realization of the r.v. ξ?
  We then talk of an wait-and-see approach. The problem is then easier to solve (we have here a simple linear program).
- But this approach is rarely appropriate!!! We usually have to decide on x before we know the realizations of ξ!
- Usually, the "wait-and-see" approach is not appropriate to model the reality behavior: we have to decide on x before we know the realizations from ξ.
- Three suggestions:
  1. try to estimate, predict, the uncertainty;
  2. chance constraints;
  3. penalties on deviations.

Remove the randomness?

A popular approach consists to look for reasonable values for ξ_1 and ξ_2. How?

Propositions:

- unbiased: choose the mean values for each random variable;
- pessimistic: choose the worst-case values for ξ;
- optimistic: choose the best-case values for ξ.

Each approach will deliver a different optimal solution!

Penalization of violations

Again, we have to deal with decision problems where the decision x has to be taken before we know the realization of ξ. In the simplest case, we can simply penalize the constraints deviations by vectors of penalty coefficients q_+ and q_-.

$$
\begin{aligned}
\min\ & c^T x + q_+^T s(\xi) + q_-^T t(\xi) \\
\text{s.t. } & Ax = b, \\
& T(\xi)x + s(\xi) - t(\xi) = h(\xi), \\
& x \in X.
\end{aligned}
$$

But it is still not possible to solve the problem!

The new optimization problem

A reasonable, and solvable, problem is then

$$
\begin{aligned}
\min\ & c^T x + E_{\boldsymbol{\xi}}[q_+^T s(\xi) + q_-^T t(\xi)] \\
\text{s.t. } & Ax = b, \\
& T(\xi)x + s(\xi) - t(\xi) = h(\xi), \quad \text{for a.e. } \xi \\
& x \in X,
\end{aligned}
$$

where *a.e.* stands for "almost every".

- In general, we can react in a correct (and maybe optimal) way: we have a recourse to "correct" the first decision once the uncertainty is removed.
- A LP recourse structure is provided by 3 elements:
  - a set $Y \subset \mathbb{R}^p$ that describes the feasible set of recourse actions, for instance $Y = \{y \in \mathbb{R}^p \mid y \geq 0\}$;
  - q: a vector of recourse costs;
  - $W \in \mathbb{R}^{m \times p}$: recourse matrix.

Recourse formulation

The previous considerations lead us to formulate the more general following program:

$$
\begin{aligned}
\min\ & c^T x + E_{\boldsymbol{\xi}}[q^T(\boldsymbol{\xi})y(\boldsymbol{\xi})] \\
\text{s.t. } & Ax = b, \\
& T(\xi)x + Wy(\xi) = h(\xi), \quad \text{for a.e.} \\
& x \in X, \\
& y(\xi) \in Y, \quad \text{for a.e. } \xi.
\end{aligned}
$$

We could have W varying with the realization ξ. If W is unique, as in the previous formulation, we speak of fixed recourse: the recourse does not change with the scenario.
But how to decide on y?

The two-stage linear stochastic problem (SP)

Using the previous definitions, we can rewrite the stochastic programming problem with recourse in terms of x only:

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

It is a (nonlinear) mathematical programming problem in $\mathbb{R}^n$. The properties of $\mathcal{Q}(x)$ influence the solution techniques.

Is $\mathcal{Q}(x)$

- linear?
- convex?
- continuous?
- differentiable?

Expression in terms of y's

$$
\begin{aligned}
\min_{x,\ y(\boldsymbol{\xi})} \quad & E_{\boldsymbol{\xi}}[c^T x + q^T y(\boldsymbol{\xi})] & \\
\text{s.t.} \quad & Ax = b & \text{first-stage constraints} \\
& T(\xi)x + Wy(\xi) = h(\xi), \ \ \text{for a.e. } \xi & \text{second-stage contraints} \\
& x \in X,\ y(\xi) \in Y \ \text{ for a.e. } \xi. &
\end{aligned}
$$

Consider the (discrete) case where $\Omega = \{\omega_1,\ \omega_2, \ldots, \omega_S\} \subset \mathbb{R}^r$.

$$
\begin{aligned}
P(\omega = \omega_s) = p_s,\ s = 1, 2, \ldots ..., S \\
T_s = T(\xi(\omega_s)),\ h_s = h(\xi(\omega_s))
\end{aligned}
$$

Deterministic equivalent

Develop along the S scenarios.

$$\min_{x,y_1,\ldots,y_S} \; c^T x + p_1 q^T y_1 + p_2 q^T y_2 + \ldots p_S q^T y_S$$

s.t.

$$\begin{array}{lllll} Ax & & & & = b \\ T_1 x & +Wy_1 & & & = h_1 \\ T_2 x & & +Wy_2 & & = h_2 \\ \vdots & & & \ddots & \\ T_S x & & & +Wy_s & = h_s \\ \multicolumn{5}{l}{x \in X, \; y_1 \in Y, \; y_2 \in Y, \ldots, y_s \in Y.} \end{array}$$

Deterministic equivalent (cont'd)

- $y_s = y(\xi(\omega_s))$ is the recourse action to take if the scenario s occurs.
- Advantage: it is a linear program.
- Drawback: it is a linear program of (very) large dimension:
  - $n + pS$ variables;
  - $m_1 + mS$ constraints.
- Advantage: the linear program matrix has a special structure (stairway shape).
  Can we exploit it?

Large scale,... and?

Assume that we have r random variables ($\Omega \subset \mathbb{R}^r$).

- Consider the following problem (source: Linderoth). A Telecom company want to expand its network in order to meet an unknown (random) demand.
- There are 86 unknown demands. Each demand is independant and take a value in a set of 7 values. Consequently

$$S = |\Omega| = 7^{86} \approx 4.77 \times 10^{72}.$$

  ...number of subatomic particles in the universe!
- It can be even worse...
  If Ω is not finite, but holds an infinite number of elements? It is especially true with continuous random variables. Our "deterministic equivalent" would have an infinite number of variables and constraints!
- We can solve an approximate problem, obtained by sampling over the random vector.

An example (cont'd)

Consider again our toy problem

$$\begin{aligned} \min_{x} \; & x_1 + x_2 \\ \text{s.t. } & \xi_1 x_1 + x_2 \geq 7 \\ & \xi_2 x_1 + x_2 \geq 4 \\ & x_1, x_2 \geq 0, \end{aligned}$$

where $\xi_1 \sim U[1, 4]$, $\xi_2[1/3, 1]$.

How to build the deterministic equivalent?

Example: recourse formulation

Assume for now that Ω is finite, with S scenarios.

$$
\begin{aligned}
\min_x \; & x_1 + x_2 + \sum_{s \in S} p_s \lambda (y_{1s} + y_{2s}) \\
\text{s.t. } & \xi_{1s} x_1 + x_2 + y_{1s} \geq 7 \\
& \xi_{2s} x_1 + x_2 + y_{1s} \geq 4 \\
& x_1, x_2 \geq 0, \\
& y_{1s}, y_{2s} \geq 0.
\end{aligned}
$$

A difficulty is therefore to decide how to construct the deterministic equivalent. How to choose λ?

How to construct the scenarios? We can proceed with Monte Carlo sampling, with $p_s = 1/N$, $\forall s$. We will explore this approach in more details later.

Example: recourse formulation (cont'd)

More generally, we can build the program

$$\begin{aligned} \min_x \; & x_1 + x_2 + E_\xi[Q(x,\xi)] \\ \text{s.t. } & x_1, x_2 \geq 0, \end{aligned}$$

and

$$\begin{aligned} Q(x, xi) = \min_y \; & q_1 y_1 + q_2 y_2 \\ \text{s.t. } & \xi_1 x_1 + x_2 + y_1 \geq 7, \\ & \xi_2 x_1 + x_2 + y_2 \geq 4. \end{aligned}$$

Two-stage linear programming problem, fixed recourse

More generally, consider the problem

$$\min c^T x + E_{\boldsymbol{\xi}}[q(\boldsymbol{\xi})^T y(\boldsymbol{\xi})]$$

subject to the constraints

$$\begin{aligned} Ax &= b, \\ T(\xi)x + Wy(\xi) &= h(\xi) \qquad \text{for a.e. } \xi \in \Xi, \\ x &\in X, \\ y(\xi) &\in Y, \end{aligned}$$

where $\boldsymbol{\xi}$ is a random vector defined on the random space $(\Omega, \mathcal{F}, P)$, and Ξ is the support of $\boldsymbol{\xi}$.

Let

$$Q(x, \xi) = \min_{y \in Y} \left\{ q(\xi)^T y : Wy = h(\xi) - T(\xi)x \right\}.$$

Reformulation(s)

$$\min_{x \in X \mid Ax = b} \left\{ c^T x + E_{\xi} \left[\min_{y \in Y} \{ q(\xi)^T y \mid Wy = h(\xi) - T(\xi)x \} \right] \right\}$$

Second-stage function, or recourse function, $v : \Xi \times \mathbb{R}^m \to \mathbb{R}$:

$$v(\xi, z) \stackrel{def}{=} \{ q(\xi)^T y \mid Wy = z \}.$$

Given a “policy” x and a realization of the random vector ξ, z measures the deviation of the first stage, i.e. $z = h(\xi) - T(\xi)x$, $v(\xi, z)$ is the minimum cost to “correct” the decision in order to satisfy the constraints again.

Recourse function

The expected recourse function, or the function of minimum expected recourse, $\mathcal{Q} : \mathbb{R}^n \to \mathbb{R}$, for any policy $x \in \mathbb{R}^n$:

$$\mathcal{Q} \stackrel{def}{=} E_{\boldsymbol{\xi}}[Q(x, \boldsymbol{\xi})],$$

describes the recourse cost expectation, with

$$Q(x, \xi) = v(\xi, h(\xi) - T(\xi)x).$$

With these definitions, the problem can be rewritten as:

$$\min_{x \in X} c^T x + \mathcal{Q}(x) \text{ such that } Ax = b.$$

It is a nonlinear program over $\mathbb{R}^n$. Properties?

Summary

Summarize our formulations.

$$\min_{x \in \mathbb{R}^n_+ \mid Ax=b} \left\{ c^T x + E_{\xi} \left[\min_{y \in \mathbb{R}^p_+} \{ q(\xi)^T y \mid Wy = h(\xi) - T(\xi)x \} \right] \right\}$$

$$\min_{x \in \mathbb{R}^n_+ \mid Ax=b} \left\{ c^T x + E_{\xi} \left[v(\xi, h(\xi) - T(\xi)x) \right] \right\}$$

$$\min_{x \in \mathbb{R}^n_+ \mid Ax=b} \left\{ c^T x + E_{\xi} \left[Q(x, \xi) \right] \right\}$$

$$\min_{x \in \mathbb{R}^n_+} \left\{ c^T x + \mathcal{Q}(x) \mid Ax = b \right\}$$

Notations

- First-stage feasible set:

$$K_1 = \{x \in \mathbb{R}^n_+ \mid Ax = b\}.$$

- Second-stage strong feasible set:

$$K_2^s = \{x \mid \mathcal{Q}(x) < \infty\}.$$

Therefore we can rewrite the problem as

$$\min_x \{c^T x + \mathcal{Q}(x) \mid x \in K_1 \cap K_2^s\}.$$

Weak feasible set

See Walkup, D. W. and Wets, R. J.-B., *Stochastic programs with recourse*, SIAM Journal on Applied Mathematics, 15(5):1299–1314, 1967.

Positive hull (or conical hull)

$$\operatorname{pos} W = \{z \mid z = Wy,\ y \in \mathbb{R}^p_+\}$$

Weak second-stage feasible set

$$\begin{aligned} K_2 &= \{x \in \mathbb{R}^n \mid Q(x, \xi) < +\infty \text{ a.s.}\} \\ &= \{x \in \mathbb{R}^n \mid (h(\xi) - T(\xi)x) \in \operatorname{pos} W \text{ a.s.}\} \end{aligned}$$

Relatively complete recourse

A problem is said to have a relatively complete recourse if $K_1 \subseteq K_2$.

Advantage: the second-stage problem is feasible $\forall x$ feasible in the first stage, almost surely.

Issue: $K_2^s \subseteq K_2$. We would like $K_2^s = K_2$.

Complete recourse

- The relatively complete recourse is very useful in practice and on a theoretical point of view, but it can be difficult to identify.
- A particular case of relatively complete recourse can however often be identified from the structure of W.
- Complete recourse: $\text{pos}\, W = \mathbb{R}^m$.
- The complete recourse property implies that $\forall x$, $T(\xi)$, $h(\xi)$, $Q(x, \xi) < \infty$, as $z = h(\xi) - T(\xi)x$.
- Complete recourse $\Rightarrow$ relatively complete recourse.

$K_2 \neq K_2^s$

Consider the second-stage problem

$$
\begin{aligned}
\min_y \;& 2y_1 + y_2 \\
\text{s.t. } & y_1 + y_2 \geq 1 - x_1, \\
& y_1 \geq \xi - x_1 - x_2, \\
& y_1, y_2 \geq 0,
\end{aligned}
$$

with

$$P\left[\xi = 2^n\right] = \frac{1}{2^{n+1}}, \; n = 0, 1, 2, \ldots$$

Note that $P\left[\xi = 2^n\right] \in (0, 1), \; n = 0, 1, 2, \ldots, \Xi = \{2^n, \; n \in \mathbb{N}_+\}$, and

$$\sum_{n=0}^{+\infty} P\left[\xi = 2^n\right] = \sum_{n=0}^{+\infty} 2^{-n-1} = \frac{1}{2}\frac{1}{1 - \frac{1}{2}} = 1.$$

$K_2 \neq K_2^s$ (cont'd)

Given $x \in \mathbb{R}^2$, we have

$$
\begin{aligned}
\mathcal{Q}(x) &= \mathbb{E}_{\boldsymbol{\xi}}\left[\min_{y \geq 0} 2y_1 + y_2 \,\middle|\, y_1 + y_2 \geq 1 - x_1,\ y_1 \geq \xi - x_1 - x_2\right] \\
&\geq \mathbb{E}_{\boldsymbol{\xi}}\left[\min_{y_1 \geq 0} 2y_1 \,\middle|\, y_1 \geq \xi - x_1 - x_2\right] \\
&= 2\mathbb{E}_{\boldsymbol{\xi}}\left[\max\{0, \boldsymbol{\xi} - x_1 - x_2\}\right] \\
&= 2\sum_{\xi \in \Xi} P[\boldsymbol{\xi} = \xi] \max\{0, \xi - x_1 - x_2\} \\
&= 2\sum_{n=0}^{+\infty} 2^{-n-1} \max\{0, 2^n - x_1 - x_2\} \\
&= \sum_{n=0}^{+\infty} 2^{-n} \max\{0, 2^n - x_1 - x_2\} = +\infty
\end{aligned}
$$

$K_2 \neq K_2^s$ (cont'd)

Thus, $K_2^s = \emptyset$.

The problem, set under standard form, can be rewritten as

$$\begin{aligned} \min_{y,u}\ & 2y_1 + y_2 \\ \text{s.t. } & y_1 + y_2 - u_1 = 1 - x_1, \\ & y_1 - u_2 = \xi - x_1 - x_2, \\ & y_1, y_2, u_1, u_2 \geq 0, \end{aligned}$$

and

$$W = \begin{pmatrix} 1 & 1 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix}.$$

$\text{pos}\, W = ?$

$K_2 \neq K_2^s$ (cont'd)

Complete recourse if $\operatorname{pos} W = \mathbb{R}^2$, i.e.

$$\forall z \in \mathbb{R}^2, \exists y \geq 0, u \geq 0 \text{ s.t. } W \begin{pmatrix} y \\ u \end{pmatrix} = z.$$

We have that

$$W \begin{pmatrix} y \\ u \end{pmatrix} = \begin{pmatrix} y_1 + y_2 - u_1 \\ y_1 - u_2 \end{pmatrix} = z$$

$$\Leftrightarrow \begin{cases} y_1 + y_2 - u_1 & = z_1, \\ y_1 - u_2 & = z_2. \end{cases}$$

Thus, we can take

$$\begin{cases} y_1 = z_2,\ u_2 = 0, & \text{if } z_2 \geq 0, \\ y_1 = 0,\ u_2 = -z_2, & \text{otherwise.} \end{cases}$$

$K_2 \neq K_2^s$ (cont'd)

Then, writing $y_2 - u_1 = z_1 - y_1$, we take

$$\begin{cases} y_2 = z_1 - y_1,\ u_1 = 0, & \text{if } z_1 - y_1 \geq 0, \\ y_2 = 0,\ u_1 = y_1 - z_1, & \text{otherwise.} \end{cases}$$

Thus,

- $\operatorname{pos} W = \mathbb{R}^2$, and the recourse is complete,
- $K_2 = \mathbb{R}^2$, and the recourse is relatively complete,
- $K_2 \neq K_2^s$.

Note that

$$\mathbb{E}[\xi] = \sum_{n=0}^{+\infty} P\left[\xi = 2^n\right] 2^n = \sum_{n=0}^{+\infty} \frac{2^n}{2^{n+1}} = \sum_{n=0}^{+\infty} \frac{1}{2} = +\infty.$$

$K_2 \neq K_2^s$ (cont'd)

$\mathbb{E}[\xi]$ not finite does not necessarily imply that $\mathcal{Q}(x) = +\infty$. Consider

$$\begin{aligned} \min_y \; & y_1 - y_2 \\ \text{s.t. } & y_1 \geq \xi \\ & y_2 \leq \xi. \end{aligned}$$

We have that $\forall \xi$, $Q(x, \xi) = 0$, and therefore $\mathcal{Q}(x) = 0$.

$K_2 = K_2^s$

Theorem

If ξ has finite second order moments, then

$$P[\xi \mid Q(x,\xi) < \infty] = 1 \Longrightarrow \mathcal{Q}(x) < \infty,$$

and consequently

$$K_2 = K_2^s.$$

Reminder: almost surely, or with probability one. An event A is said to occur almost surely if $P[A] = 1$.

Elementary feasible set

- Given a realization ξ, the elementary second-stage feasible set is defined as:

$$K_2(\xi) = \{x \mid Q(x,\xi) < \infty\}.$$

- Define possibility interpretation of the second-stage feasibility as

$$K_2^P = \cap_{\xi \in \Xi} K_2(\xi).$$

- Clearly $K_2 = K_2^P$ if ξ has a finite support. Is it still the case when ξ follows a continuous distribution?

Theorem

For a stochastic program with fixed recourse, where ξ has finite second order moments,

$$K_2 = K_2^s = K_2^p.$$

Simple recourse

A particular case of complete recourse is the simple recourse, for which we have

$$W = \begin{pmatrix} I & -I \end{pmatrix},$$

with I the identiy matrix, of order m.

It is easy to see that $\operatorname{pos} W = \mathbb{R}^m$ as, writing

$$y = \begin{pmatrix} y^+ \\ y^- \end{pmatrix},$$

where $y \in \mathbb{R}^{2m}$, $y^+, y^- \in \mathbb{R}^m$, for $z \in \mathbb{R}^m$, we can solve the system

$$Wy = z, \ y \geq 0,$$

by chosing $y^+ = \max\{z, 0\}$, $y^- = \max\{-z, 0\}$.

Simple recourse (cont'd)

The second stage program can be read as

$$\begin{aligned} Q(x,\xi) = \min_y \; & q^+(\xi)^T y^+ + q^-(\xi)^T y^- \\ \text{s.t. } & y^+ - y^- = h(\xi) - T(\xi)x, \\ & y^+,\ y^- \geq 0. \end{aligned}$$

That is, for $q^+(\xi) + q^-(\xi) \geq 0$, the recourse variables y^+ and y^- can be chosen to measure the absolute violations in the stochastic constraints.

Simple recourse (cont'd)

- Since the recourse is complete, it is feasible for any first-stage decision x.
- The following results establishes conditions to have a finite recourse (i.e. $|\mathcal{Q}| < \infty$).

Theorem

Assume that the two-stage (linear) stochastic program is feasible and has a simple recourse, and that ξ has finite second-order moments. Then $\mathcal{Q}(x)$ is finite if and only if, for any component i, $\boldsymbol{q}_i^+ + \boldsymbol{q}_i^- \geq 0$ with probability one.

Simple recourse (cont'd)

Proof.

($\Rightarrow$) Assume by contradiction that $\mathcal{Q}$ is finite, but for some component i, $q_i^+(\xi(\omega)) + q_i^-(\xi(\omega)) < 0$ for $\omega \in \Omega_1$ with $P[\Omega_1] > 0$. Then, for any feasible x, for all $\omega \in \Omega_1$ with $h_i(\xi(\omega)) - T_i(\xi(\omega))x > 0$, define

$$y_i^+(\xi(\omega)) = h_i(\xi(\omega)) - T_i(\xi(\omega))x + u,\ y_i^-(\xi(\omega)) = u.$$

Therefore,

$$y_i^+(\xi(\omega)) - y_i^-(\xi(\omega)) = h_i(\xi(\omega)) - T_i(\xi(\omega))x,\ y_i^+ \geq 0,\ y_i^- \geq 0.$$

Moreover, since $\mathcal{Q}$ is finite, $Q(x, \xi(\omega))$ is feasible almost surely, so, almost surely, we can choose y_j^+ and y_j^- feasible, $j \neq i$.

□

Simple recourse (cont'd)

Proof.

($\Rightarrow$)

When $u \to \infty$, $Q(x, \xi(\omega)) \to -\infty$ since $q_i^+(\xi(\omega))y_i^+ + q_i^-(\xi(\omega))y_i^- \to -\infty$.

A similar argument can be applied if $h_i(\xi(\omega)) - T_i(\xi(\omega))x \leq 0$, by taking

$$y_i^+(\xi(\omega)) = u, \; y_i^-(\xi(\omega)) = -h_i(\xi(\omega)) + T_i(\xi(\omega))x + u.$$

By componing these two cases, we conclude that $\mathcal{Q}$ is not finite. $\square$

Simple recourse (cont'd)

Proof.

($\Leftarrow$) Assume $\boldsymbol{q}_i^+ + \boldsymbol{q}_i^- \geq 0$ with probability one, $\forall i$. Any feasible solution satisfies

$$y^+(\xi(\omega)) - y^-(\xi(\omega)) = h(\xi(\omega)) - T(\xi(\omega))x, \; y^+(\xi(\omega)) \geq 0, \; y^-(\xi(\omega)) \geq 0.$$

Therefore for a.e. ω, $Q(x, \xi(\omega))$ is bounded below, as, for a.e. ω, $\forall i$,

$$\begin{aligned}
& q_i^+(\xi(\omega)) y_i^+(\xi(\omega)) + q_i^-(\xi(\omega)) y_i^-(\xi(\omega)) \\
&= q_i^+(\xi(\omega)) \left(y_i^+(\xi(\omega) - y_i^-(\xi(\omega)) \right) + q_i^+(\xi(\omega)) y_i^-(\xi(\omega)) + q_i^-(\xi(\omega)) y_i^-(\xi(\omega)) \\
&= q_i^+(\xi(\omega)) \left(h(\xi(\omega)) - T(\xi(\omega))x \right)_i + \left(q_i^+(\xi(\omega)) + q_i^-(\xi(\omega)) \right) y_i^-(\xi(\omega)) \\
&\geq q_i^+(\xi(\omega)) \left(h(\xi(\omega)) - T(\xi(\omega))x \right)_i .
\end{aligned}$$

$\square$

Simple recourse (cont'd)

Proof.

($\Leftarrow$) From the fundamental theorem of linear programming, we can choose as the optimal solution

$$
\begin{aligned}
y^+(\xi(\omega)) &= (h(\xi(\omega)) - T(\xi(\omega))x)^+ \,, \\
y^-(\xi(\omega)) &= (-h(\xi(\omega)) + T(\xi(\omega))x)^+ \,,
\end{aligned}
$$

where $a^+ = \max\{0, a\}$. Thus,

$$
\begin{aligned}
Q(x, \xi(\omega)) = \sum_{i=1}^{m} (q_i^+(\xi(\omega))(h_i(\xi(\omega)) - T_i(\xi(\omega))x)^+ + \\
q_i^-(\xi(\omega))(-h_i(\xi(\omega)) + T_i(\xi(\omega))x)^+).
\end{aligned}
$$

$\square$

Simple recourse (cont'd)

Proof.

($\Leftarrow$) Consequently $Q(x, \xi(\omega))$ is finite for a.e. ω and $\mathcal{Q}(x) < \infty$. $\mathcal{Q}(x) > -\infty$ follows from convexity of $Q(x, \xi(\omega))$ (to be shown) and Jensen's inequality: if X is a random variable with $\mathbb{E}[|X|] < \infty$, and g is a convex function, then

$$g(\mathbb{E}[X]) \leq \mathbb{E}[g(X)].$$

$\square$

Exercise

Consider the second stage program

$$Q(x, \xi) = \min_y \{ y \mid \xi y = 1 - x, y \geq 0 \}.$$

We assume that $\boldsymbol{\xi}$ follows a triangular distribution on $[0, 1]$, with $P[\boldsymbol{\xi} \leq u] = u^2$.

(a) Is the recourse fixed? Why?

The recourse is not fixed, as $W \equiv \boldsymbol{\xi}$, and therefore, W is random. Moreover, as ξ can take the value 0, the transformation

$$y = 1/\xi - x/\xi,$$

is not properly defined on $\Xi = [0, 1]$; this also means that

$$W = \begin{cases} 0 & \text{si } \xi = 0; \\ 1 & \text{si } \xi \neq 0. \end{cases}$$

Exercise (cont'd)

(b) Express $K_2(\xi)$ for all ξ in $[0, 1]$.

We have to consider two cases: $\xi = 0$ or $\xi \in (0, 1]$.

1. $\xi \in (0, 1]$ In this case, as $y, \xi \geq 0$, $1 - x$ has to be non-negative in order to have a well-defined problem:

$$K_2(\xi) = \{x \mid x \leq 1\}.$$

The value and optimal solutions are

$$Q^*(x, \xi) = (1 - x)/\xi, \quad y^* = (1 - x)/\xi.$$

2. $\xi = 0$ There exists no y such that $0.y = 1 - x$, except if $x = 1$, so

$$K_2(0) = \{1\}.$$

Exercise (cont'd)

(c) Express K_2, K_2^P and $\mathcal{Q}$.

From the previous point, we have

$$K_2^P = \{x \,|\, x \leq 1\} \cap \{1\} = \{1\}.$$

We also have, as $P[\xi = 0] = 0$,

$$\mathcal{Q}(x) = \int_0^1 \frac{1-x}{\xi} 2\xi d\xi = 2(1-x), \forall x \leq 1.$$

Consequently $K_2^P \subset K_2 = \{x \leq 1\}$.

The difference comes from the fact that a point is not in K_2^P as soon as it is not feasible for a given value of ξ, but K_2 does not consider unfeasible situations that occur with a null probability.

Recourse function

Let y_1^* and y_2^* be two optimal solutions of $v(\xi, z)$, associated to $z = z_1$ and $z = z_2$, respectively. Then, the convex combination $y_\alpha \stackrel{def}{=} \alpha y_1^* + (1-\alpha)y_2^*$, $\alpha \in [0, 1]$, is feasible with respect to $z_\alpha = \alpha z_1 + (1-\alpha)z_2$, as $\alpha y_1^* + (1-\alpha)y_2^* \geq 0$, and

$$W(\alpha y_1^* + (1-\alpha)y_2^*) = \alpha W y_1^* + (1-\alpha)W y_2^* = \alpha z_1 + (1-\alpha)z_2 = z_\alpha.$$

Moreover,

$$\begin{aligned} v(\xi, z_\alpha) &= q(\xi)^T y_\alpha^* \leq q(\xi)^T(\alpha y_1^* + (1-\alpha)y_2^*) \\ &= \alpha q(\xi)^T y_1^* + (1-\alpha)q(\xi)^T y_2^* \\ &= \alpha v(\xi, z_1) + (1-\alpha)v(\xi, z_2). \end{aligned}$$

In other words, v is a convex function w.r.t. $z \in \mathbb{R}^m$.

Convexity of $Q(x, \xi)$?

$$Q(x, \xi) = v(\xi, h(\xi) - T(\xi)x).$$

$$\begin{aligned}
\lambda Q(x_1, \xi) &+ (1-\lambda)Q(x_2, \xi) \\
&= \lambda v(\xi, h(\xi) - T(\xi)x_1) + (1-\lambda)v(\xi, h(\xi) - T(\xi)x_2) \\
&\geq v(\xi, \lambda(h(\xi) - T(\xi)x_1) + (1-\lambda)(h(\xi) - T(\xi)x_2)) \\
&= v(\xi, h(\xi) - T(\xi)(\lambda x_1 + (1-\lambda)x_2)) \\
&= Q(\lambda x_1 + (1-\lambda)x_2, \xi).
\end{aligned}$$

Therefore $Q(x, \xi)$ if convex w.r.t. x, given ξ. More generally

Theorem

If A if a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$, and $f(x)$ is a convex function on $\mathbb{R}^m$, the composite function $(fA)(x) \stackrel{def}{=} f(Ax)$ is a convex function on $\mathbb{R}^n$.

Convexity of second-stage function

We have the following result (Birge and Louveaux, Chapter 3, Theorem 5).

Theorem

For a stochastic program with fixed recourse, $Q(x, \xi)$ is

(a) *a piecewise convex linear function in (h, T),*

(b) *a piecewise concave linear function in q,*

(c) *a piecewise convex linear function in x, for all x in $K = K_1 \cap K_2$.*

Convexity of second-stage function (cont'd)

Proof.

In order to show convexity in (a) and (c), it is sufficient to prove that $v(\xi, z) = \min\{q(\xi)^T y \mid Wy = z\}$ is convex, which has already been done. We can proceed similarly to show concavity w.r.t. q.

The piecewise linearity follows from the fact that the number of different optimal bases for a linear program is finite. $\square$

Convexity of the recourse?

$$\mathcal{Q}(x) = E_{\boldsymbol{\xi}}[Q(x, \xi)].$$

Suppose for now that ξ has a finite support, i.e. $\Xi = \{\xi_1, \xi_2, \dots, \xi_m\}$. Then

$$\mathcal{Q}(x) = \sum_{i=1}^{m} P[\boldsymbol{\xi} = \xi_i] Q(x, \xi_i).$$

Convexity of the recourse

Theorem

If $f(x)$ is convex, and $\alpha \geq 0$, $g(x) = \alpha f(x)$ is convex.

Theorem

If $f_k(x)$, $k = 1, 2, \ldots, K$, are convex functions, then $g(x) = \sum_{k=1}^{K} f_k(x)$ is convex.

$\mathcal{Q}(x)$ is therefore a convex function w.r.t. x.

What is happening in the continuous case?

We have the following result: if $g(x, y)$ is convex w.r.t. x, then $\int g(x, y)dy$ is convex w.r.t. x. Since

$$\mathcal{Q}(x) = \int_{\Xi} Q(x, t)dF(t),$$

$\mathcal{Q}(x)$ is convex.

An example…

Consider the second-stage function $Q(x, \xi)$ defined as:

$$\min y^+ + y^- \text{ s.t. } y^+ - y^- = \xi - x,\ y^+ \geq 0,\ y^- \geq 0.$$

In other terms:

$$y = \begin{pmatrix} y^+ \\ y^- \end{pmatrix} \quad q = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, W = \begin{pmatrix} 1 & -1 \end{pmatrix} \quad h(\xi) = \xi \quad T(\xi) = 1.$$

Relying on the fundamental theorem of linear programming, we are looking from an optimal basis solution, implying that $y^+ = 0$ or $y^- = 0$.

An example...

We immediately see that

$$y^+ = \begin{cases} \xi - x & \text{if } \xi - x \geq 0, \\ 0 & \text{otherwise,} \end{cases}$$

and

$$y^- = \begin{cases} -\xi + x & \text{if } \xi - x < 0, \\ 0 & \text{otherwise,} \end{cases}$$

Alternative approach

Dual:

$$\max\ (\xi - x)\pi$$
$$\text{s.t. } \pi \leq 1,\ -\pi \leq 1$$

or

$$\begin{aligned} \max\ & (\xi - x)\pi \\ \text{s.t. } & \pi + s_1 = 1 \\ & -\pi + s_2 = 1 \\ & s_1 \geq 0,\ s_2 \geq 0 \end{aligned}$$

Consequently,

$$Q(x, \xi) = \begin{cases} \xi - x & \text{si } x \leq \xi, \\ x - \xi & \text{si } x \geq \xi. \end{cases}$$

An example: optimality conditions

The recourse is simple, and the primal-dual/KKT conditions give

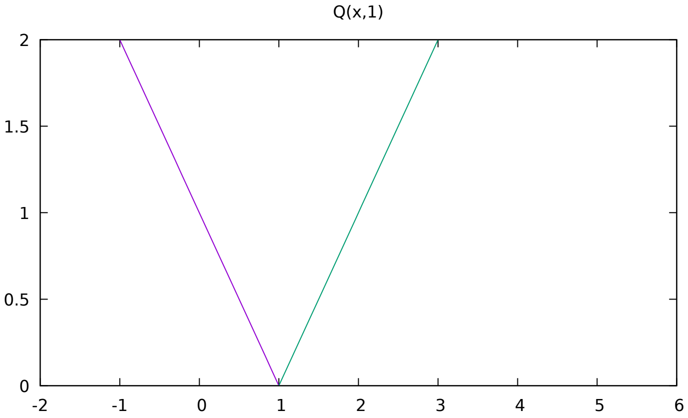
$$
\begin{aligned}
\begin{pmatrix} 1 \\ 1 \end{pmatrix} &= \begin{pmatrix} 1 \\ -1 \end{pmatrix} \pi + \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \\
y^+ - y^- &= \xi - x \\
y^+ &\geq 0, \;\; y^- \geq 0 \\
s_1 &\geq 0, \;\; s_2 \geq 0 \\
s_1 y^+ &= 0, \;\; s_2 y^- = 0.
\end{aligned}
$$

An example (cont'd)

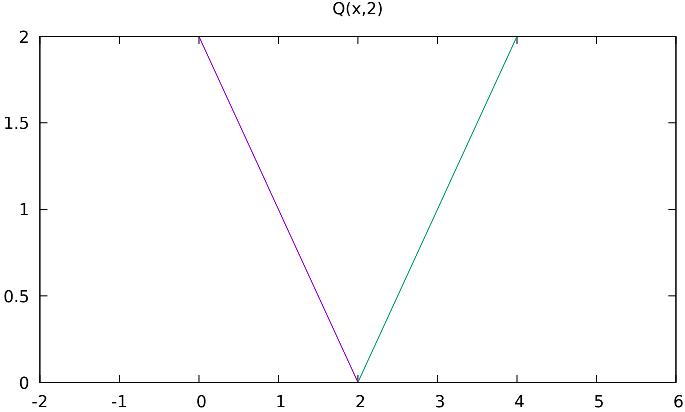
- The first condition implies that we cannot have $s_1 = s_2 = 0$.
- From the complementarity conditions, we have that $y^+ = 0$ or $y^- = 0$.
- We have to consider two cases:
    - $x \leq \xi$: in this situation, we have
    $$y^+ = \xi - x, \quad y^- = 0.$$
    - $x \geq \xi$: then,
    $$y^- = x - \xi, \quad y^+ = 0.$$

Graphically?

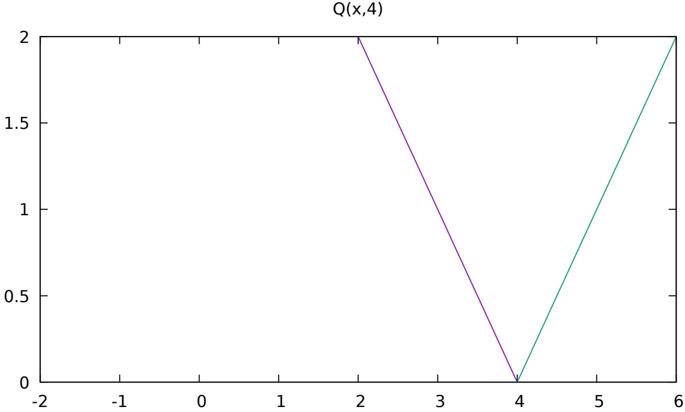
Assume that ξ can take the realizations 1, 2, 4.



Graphically (cont'd)



Graphically (cont'd)



$\mathcal{Q}(x)$

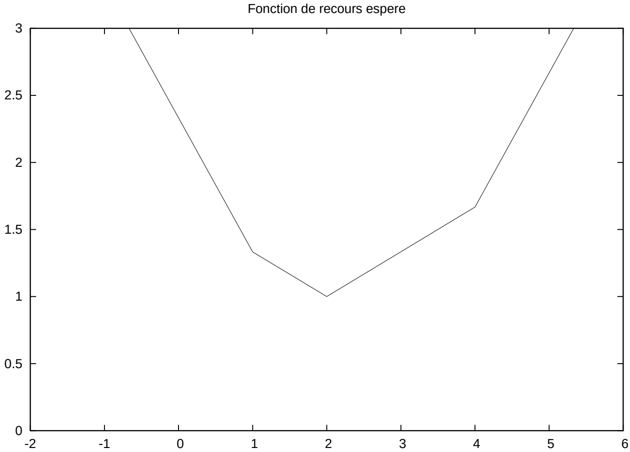
What about $\mathcal{Q}(x)$?

Assume that the three realizations have the same probability.

We have to consider 4 cases:

1. $x \leq 1$: $\mathcal{Q}(x) = 7/3 - x$;
2. $1 \leq x \leq 2$: $\mathcal{Q}(x) = 5/3 - x/3$;
3. $2 \leq x \leq 4$: $\mathcal{Q}(x) = x/3 + 1/3$;
4. $4 \leq x$: $\mathcal{Q}(x) = x - 7/3$;

Graphically



Properties of $\mathcal{Q}(x)$

We note that $\mathcal{Q}(x)$ is convex and piecewise linear. As $\mathcal{Q}(x)$ is a finite weighted sum of piecewise linear functions when the support of ξ is finite, we have the following result.

Theorem

For a stochastic program with fixed recourses where ξ has finite second-order moments,

(a) *$\mathcal{Q}(x)$ is a convex Lipschitz function and is finite over K_2;*

(b) *when ξ has a finite support, $\mathcal{Q}(x)$ is piecewise linear.*

Reminder: a function f is Lipschitz if there exists some $M < \infty$ such that for all x, y,

$$|f(x) - f(y)| \leq M\|x - y\|.$$

Differentiability of the recourse

Is $\mathcal{Q}(x)$ also differentiable?

The recourse function is partially differentiable with respect to x_j at $(\hat{x}, \hat{\xi})$ if the directional derivative exists for the direction e_j. In other terms, there exists a function $\frac{\partial Q(x,\xi)}{\partial x_j}$ such that

$$\frac{Q(\hat{x} + he_j, \hat{\xi}) - Q(\hat{x}, \hat{\xi})}{h} = \frac{\partial Q(x, \xi)}{\partial x_j} + \frac{\rho_j(\hat{x}, \hat{\xi}; h)}{h},$$

with

$$\frac{\rho_j(\hat{x}, \hat{\xi}; h)}{h} \to 0, \text{ as } h \to 0.$$

We will assume from now that $\nabla_x Q(x, \xi) = \left(\frac{\partial Q(x,\xi)}{\partial x_1}, \ldots, \frac{\partial Q(x,\xi)}{\partial x_n} \right)$ exists.

Differentiability of the recourse (cont'd)

What about the differentiability of $\mathcal{Q}(x)$?

$$\frac{\mathcal{Q}(\hat{x}+he_j)-\mathcal{Q}(\hat{x})}{h} = \int_{\Xi} \frac{Q(\hat{x}+he_j,\xi)-Q(\hat{x},\xi)}{h} dP$$
$$= \int_{\Xi\setminus N_\delta} \frac{\partial Q(\hat{x},\xi)}{\partial x_j} dP + \int_{\Xi\setminus N_\delta} \frac{\rho_j(\hat{x},\xi;h)}{h} dP,$$

where N_δ is measurable and $P[N_\delta]=0$. Therefore, we have

Theorem

If $Q(x,\xi)$ if partially differentiable almost everywhere, if its partial derivative $\frac{\partial Q(\hat{x},\xi)}{\partial x_j}$ is integrable and if the residual satisfies $(1/h)\int_{\Xi\setminus N_\delta} \rho_j(\hat{x},\xi;h)dP \overset{h\to 0}{\to} 0$, then $\frac{\partial \mathcal{Q}(\hat{x})}{\partial x_j}$ exists and

$$\frac{\partial \mathcal{Q}(\hat{x})}{\partial x_j} = \int_{\Xi} \frac{\partial Q(\hat{x},\xi)}{\partial x_j} dP.$$

Differentiability of the recourse: discrete case

But how to prove $(1/p)\int_{\Xi\setminus N_\delta} \rho_j(\hat{x},\xi;h)dP \overset{h\to 0}{\to} 0$?

If we stay in the linear framework with fixed recourse, and vectors $\boldsymbol{\xi}$ with finite second-order moments, we have seen that for $\boldsymbol{\xi}$ with finite support, $\mathcal{Q}(x)$ is piecewise linear. Therefore $\mathcal{Q}(x)$ is not differentiable.

Differentiability of the recourse: continuous case

If $\boldsymbol{\xi}$ is continuous, $\mathcal{Q}(x)$ is obtained as an integral over the $Q(x, \xi)$'s, that are not differentiable as they are piecewise linear given ξ. However, it is x that has to be fixed, not ξ. It is possible to show that (the proof is quite technical)

Theorem

For a stochastic program with fixed recourse where $\boldsymbol{\xi}$ has finite second-order moments, if $\boldsymbol{\xi}$ is continuous, $\mathcal{Q}(x)$ is differentiable over K_2.

Intuitively, the function $\mathcal{Q}(x)$ is "smoothed" by the superposition of an infinite number of functions $Q(x, \xi)$.

Two-stage non-linear problems

Now consider the general program

$$\min_{x \in X} E_{\boldsymbol{\xi}}[f_0(x, \boldsymbol{\xi})] = \min_{x \in X} E_{\boldsymbol{\xi}}[g_0(x, \boldsymbol{\xi}) + Q(x, \boldsymbol{\xi})].$$

Theorem

If $g_0(\cdot, \xi)$ and $Q(\cdot, \xi)$ are convex with respect to x, $\forall\, \xi \in \Xi$, and if X is a convex set, the aforementioned program is convex.

Proof.

For x, $y \in X$, $\lambda \in (0, 1)$ and $z = \lambda x + (1 - \lambda)y$, we have

$$\begin{aligned} & g_0(z, \xi) + Q(z, \xi) \\ & \qquad \leq \lambda(g_0(x, \xi) + Q(x, \xi)) + (1 - \lambda)(g_0(y, \xi) + Q(y, \xi)). \end{aligned}$$

The result follows by taking the expectation. $\square$

In a more standard form

Inspired from Birge et Louveaux, Section 3.4.

We consider the problem

$$\begin{aligned} \inf z = {} & f^1(x) + \mathcal{Q}(x), \\ \text{s.t. } & g_i^1(x) \le 0,\ i = 1, \ldots, \overline{m}_1, \\ & g_i^1(x) = 0,\ i = \overline{m}_1 + 1, \ldots, m_1, \end{aligned}$$

where $\mathcal{Q}(x) = E_\omega[Q(x, \omega)]$ and

$$\begin{aligned} Q(x, \omega) = {} & \inf f^2(y(\omega), \omega), \\ \text{s.t. } & t_i^2(x, \omega) + g_i^2(y(\omega), \omega) \le 0,\ i = 1, \ldots, \overline{m}_2, \\ & t_i^2(x, \omega) + g_i^2(y(\omega), \omega) = 0,\ i = \overline{m}_2 + 1, \ldots, m_2, \end{aligned}$$

We say that the recourse is additive (why?).

In a more standard form (cont'd)

The functions $f^2(\cdot,\omega)$, $t_i^2(\cdot,\omega)$, and $g_i^2(\cdot,\omega)$ are continuous for any given ω, and measurable w.r.t. ω for any given first argument. This allows to prove that $Q(x,\omega)$ is measurable, and therefore that $\mathcal{Q}(x)$ is well defined.

Reintroduce K_1, $K_2(\omega)$ and K_2.

$$
\begin{aligned}
K_1 &= \{x \mid g_i^1(x) \leq 0,\ i = 1, \ldots, \overline{m}_1, \\
&\qquad\quad g_i^1(x) = 0,\ i = \overline{m}_1 + 1, \ldots, m_1\}, \\
K_2(\omega) &= \{x \mid \exists\, y(\omega) \text{ t.q. } t_i^2(x,\omega) + g_i^2(y(\omega),\omega) \leq 0,\ i = 1, \ldots, \overline{m}_2, \\
&\qquad\quad t_i^2(x,\omega) + g_i^2(y(\omega),\omega) = 0,\ i = \overline{m}_2 + 1, \ldots, m_2\}, \\
K_2 &= \{x \mid \mathcal{Q}(x) < \infty\}.
\end{aligned}
$$

Remarks

The formulation is not yet totally general. We will consider more general forms when we will discuss sampling techniques.

Here, there is no more fixed recourse, but the first-stage decision x acts separately in the constraints. Goal: extend the previous results.

Questions: convexity, differentiability, optimality. We should also consider the concept of lower semi-continuity.