STOCHASTIC PROGRAMMING

December 7, 2023

Today:

→ More on stochastic programming with recourse (metrics of interest, multi-stage)
→ Beyond expected value: risk measures

① Assessing the value of stochastic programming

↳ Illustrated using two-stage stochastic programming

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \left\{ f(x) + \mathbb{E}_{\xi}\left[Q(x,\xi)\right] \text{ s.t. } x \in X \right\}$$

where $\quad Q(x,\xi) = \underset{y \in \mathbb{R}^{n_1}}{\min} \left\{ g(y,\xi) \text{ s.t. } y \in Y(x,\xi) \right\}$

↳ Originally, the problem was

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \left\{ f(x) + Q(x,\xi) \text{ s.t. } x \in X \right\} \quad \text{with } \xi \text{ being uncertain}$$

↳ Could we have considered

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \left\{ f(x) + Q(x, \mathbb{E}_{\xi}[\xi]) \text{ s.t. } x \in X \right\} \; ?$$

↳ How does $\mathbb{E}_{\xi}\left[\underset{x \in \mathbb{R}^n}{\min} \left\{ f(x) + Q(x,\xi) \text{ s.t. } x \in X \right\} \right]$

compare to

$$\underset{x \in \mathbb{R}^n}{\min} \left\{ f(x) + \mathbb{E}_{\xi}\left[Q(x,\xi)\right] \text{ s.t. } x \in X \right\} \; ?$$

a) Expected value of perfect information (EVPI)

- Define the value of the recourse problem as

$$VRP = \underset{x \in \mathbb{R}^n}{\min} \left\{ f(x) + \mathbb{E}_{\xi}\left[Q(x,\xi)\right] \text{ s.t. } x \in X \right\}$$

(optimal value of the two-stage stochastic program)

$\rightarrow$ Deciding x before observing ξ

- We want to compare VRP with the value that one can get on average by waiting until ξ is observed

$\Rightarrow$ We define the expected wait-and-see cost as

$$EWS = \mathbb{E}_{\xi}\left[\min_{x \in \mathbb{R}^n} \{ f(x) + Q(x,\xi) \text{ s.t. } x \in \mathcal{X} \}\right]$$

$\rightarrow$ We wait until ξ is revealed to decide x

$\rightarrow$ EWS is the average performance/cost of that approach

Def: The expected value of information (or EVPi) for the problem is given by $EVPi = VRP - EWS$

$\rightarrow$ $EVPi \geq 0$ by property of the expected value

$\rightarrow$ $EVPi = 0$: no interest in making the decision before observing ξ

$\rightarrow$ $EVPi > 0$: represents the price one would pay to know the distribution of ξ in advance

b) Value of the stochastic solution

$\hookrightarrow$ An expected value decision for the problem is defined as

$$\bar{x} \in \operatorname*{argmin}_{x \in \mathbb{R}^n} \{ f(x) + Q(x, \mathbb{E}_{\xi}[\xi]) \text{ s.t. } x \in \mathcal{X} \}$$

→ Remove the uncertainty by fixing ξ to its mean

→ the optimal value of the resulting problem is

$$\underbrace{f(\bar{x}) + Q(\bar{x}, \mathbb{E}_\xi[\xi])}_{\text{This value is called the EEV}} = \min_{x \in \mathbb{R}^n} \{ f(x) + Q(x, \mathbb{E}_\xi[\xi]) \text{ s.t. } x \in X \}$$

$\underbrace{\text{EV}}_{\text{Expected Value}}$

Def: the value of the stochastic solution (VSS) is defined by

$$VSS = EEV - VRP$$

→ $VSS \geq 0$

→ Represents the cost of ignoring uncertainty by fixing ξ to its mean

There are no generic rules on what the VSS and the EVPI are equal to:

→ Depends on the distribution of ξ

→ Depends on the set of solutions of the problems involved (recourse problem, expected value problem, $\underset{x \in \mathbb{R}^n}{\text{minimize}} \{ f(x) + Q(x, \xi) \text{ s.t. } x \in X \}$ for fixed ξ)

Example

$$\underset{x \in \mathbb{R}^2}{\text{minimize}} \quad x_1 + 4x_2 + Q(x,\xi)$$

$$\text{s.t.} \quad x_1 + x_2 = 1$$
$$x_1, x_2 \geq 0$$

$$Q(x,\xi) = \min_{y \in \mathbb{R}^3} \{ y_1 + 10 y_2 + 10 y_3 \ \text{ s.t. } \ y_1 + y_2 - y_3 = \xi + x_1 - 2x_2, \ y_1 \leq 2, \ y_1, y_2, y_3 \geq 0 \}$$

- Suppose that ξ is uniformly distributed in $[1,3]$

  Then EVPI $= 0$ and VSS $= \frac{11}{4} > 0$

  $\rightarrow x^* = \begin{bmatrix} 1/3 \\ 2/3 \end{bmatrix}$ is a solution of the problem for any $\xi \in [1,3]$

  $\Rightarrow$ EVPI $= 0$

  $\rightarrow$ when ξ is fixed to $\mathbb{E}_\xi[\xi] = 2$, there are multiple solutions to $\underset{x \in \mathbb{R}^2}{\text{minimize}} \{ x_1 + 4x_2 + Q(x,2) \ \text{ s.t. } \ x_1 + x_2 = 1, \ x_1, x_2 \geq 0 \}$

  that do not give the same value of

  $$x_1 + 4x_2 + \mathbb{E}_\xi[Q(x,\xi)]$$

- Suppose now that $\xi \in \{0, \frac{3}{2}, 3\}$ with $\mathbb{P}(\xi = 0) = \mathbb{P}(\xi = 3/2) = \mathbb{P}(\xi = 3) = 1/3$

  Then EVPI $= 3 > 0$ and VSS $= 0$

$\rightarrow$ $\mathbb{E}_\xi[\xi] = 3/2$ and the fact that $\left(\frac{2}{3}, \frac{1}{3}\right)$ is a solution for every possible value give $VSS = 0$

$\rightarrow$ The solution sets for $\xi \in \{0, \frac{3}{2}, 3\}$ are not identical.

$\Rightarrow$ By picking different values $\bar{x}(0)$, $\bar{x}(\frac{3}{2})$ and $\bar{x}(3)$, we can show that

$$EWS = \mathbb{E}_\xi\left[f(\bar{x}(\xi)) + Q(\bar{x}(\xi), \xi) \right]$$

$$= \frac{1}{3} \times \left(f(\bar{x}(0)) + Q(\bar{x}(0), 0)\right) + \frac{1}{3} \times \left(f(\bar{x}(3/2)) + Q(\bar{x}(\tfrac{3}{2}), \tfrac{3}{2})\right)$$

$$+ \frac{1}{3} \times \left(f(\bar{x}(3)) + Q(\bar{x}(3), 3)\right)$$

$$< f(x^*) + \frac{1}{3}\left(Q(x^\circ, 0) + Q(x^\circ, \tfrac{3}{2}) + Q(x^\circ, 3)\right)$$

where x° solves the two-stage recourse problem

Exercise: Consider

$$\underset{x \in \mathbb{R}}{\text{minimize}} \left\{ 2x + \mathbb{E}_\xi[Q(x, \xi)] \text{ s.t. } x \geq 0 \right\}$$

where $Q(x, \xi) = \min_{y \in \mathbb{R}} \{ \xi \cdot y \text{ s.t. } y \geq 1 - x, y \geq 0 \}$

a) Give a closed-form expression for $Q(x, \xi)$

b) Suppose that $\xi \in \{1, 3\}$ with $\mathbb{P}(\xi = 1) = \frac{3}{4}$ and $\mathbb{P}(\xi = 3) = \frac{1}{4}$

show that $EWS < EEV$

② From two-stage to multi-stage stochastic programs

Generic form of a two-stage program

$$\underset{x_0 \in \mathbb{R}^{n_0}}{\text{minimize}} \left\{ f_0(x_0) + \mathbb{E}_{\xi_1}\left[Q_1(x_0, \xi_1) \right] \text{ s.t. } x_0 \in \mathcal{X}_0 \right\}$$

where $$Q_1(x_0, \xi_1) = \min_{x_1 \in \mathbb{R}^{n_1}} \left\{ f_1(x_1, \xi_1) \text{ s.t. } x_1 \in \mathcal{X}_1(x_0, \xi_1) \right\}$$

Models a two-step decision process

$$x_0 \rightsquigarrow \xi_1 \rightsquigarrow x_1$$

x_0: Decision at time $t=0$

ξ_1: Uncertain parameter revealed at time $t=1$

x_1: Recourse / second-stage made at time $t=1$ according to x_0 and the value of ξ_1

Multi-stage stochastic programming

$T+1$ stages where $T \geq 1$ (here T is finite)

$$x_0 \rightsquigarrow \xi_1 \rightsquigarrow x_1 \rightsquigarrow \xi_2 \rightsquigarrow \cdots \rightsquigarrow \xi_T \rightsquigarrow x_T$$

Short notation

$$\underset{x_0 \in \mathcal{X}_0}{\text{minimize}}\ f_0(x_0) + \mathbb{E}_{\xi_1}\left[\min_{x_1 \in \mathcal{X}_1(x_0, \xi_1)} f_1(x_1, \xi_1) + \mathbb{E}_{\xi_2}\left[\cdots + \mathbb{E}_{\xi_T}\left[\min_{x_T \in \mathcal{X}_T(x_{T-1}, \xi_T)} f_T(x_T, \xi_T) \right] \cdots \right] \right]$$

"Long" version

$$\underset{x_0 \in \mathbb{R}^{n_0}}{\text{minimize}} \; f_0(x_0) + \mathbb{E}_{\xi_1}\left[Q(x_0, \xi_1)\right] \quad \text{s.t.} \; x_0 \in \mathcal{X}_0$$

where
$$Q_1(x_0, \xi_1) = \min_{x_1 \in \mathbb{R}^{n_1}} \left\{ f_1(x_1, \xi_1) + \mathbb{E}_{\xi_2}\left[Q_2(x_1, \xi_2)\right] \text{ s.t. } x_1 \in \mathcal{X}_1(x_0, \xi_1) \right\}$$

$$\vdots$$

$$Q_t(x_{t-1}, \xi_t) = \min_{x_t \in \mathbb{R}^{n_t}} \left\{ f_t(x_t, \xi_t) + \mathbb{E}_{\xi_t}\left[Q_{t+1}(x_t, \xi_{t+1})\right] \text{ s.t. } x_t \in \mathcal{X}_t(x_{t-1}, \xi_t) \right\}$$
$$\forall t = 2, \dots, T-1$$

$$Q_T(x_{T-1}, \xi_T) = \underset{x_T \in \mathbb{R}^{n_T}}{\text{minimize}} \left\{ f_T(x_T, \xi_T) \text{ s.t. } x_T \in \mathcal{X}_T(x_{T-1}, \xi_T) \right\}$$

↳ Two main approaches to multistage stochastic programming

- Scenario approach, essentially for linear multistage stochastic programs
- Dynamic programming

a) Scenarios and linear multistage programming

Linear multistage stochastic program

$$f_t(x_t, \xi_t) = c_t^T x_t$$

$$\mathcal{X}_t(x_{t-1}, \xi_t) = \left\{ x_t \mid B_t x_{t-1} + A_t x_t = b_t, \; x_t \geq 0 \right\}$$

$$\xi_t = (A_t, B_t, b_t, c_t)$$

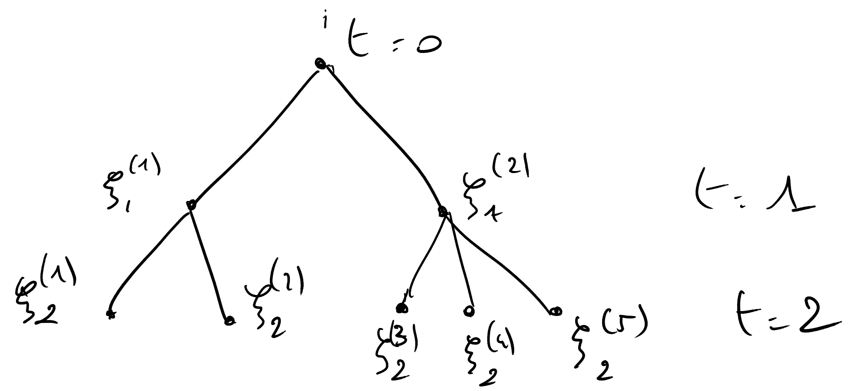
In that case, the problem can be reformulated as a linear program

$$\underset{\substack{x_0 \in \mathbb{R}^{n_0} \\ \vdots \\ x_T \in \mathbb{R}^{n_T}}}{\text{minimize}} \quad c_0^T x_0 + \mathbb{E}_{\xi_1}[c_1^T x_1] + \dots + \mathbb{E}_{\xi_T}[c_T^T x_T]$$

$$\begin{aligned} \text{s.t.} \quad & A_0 x_0 && = b_0 \\ & B_1 x_0 + A_1 x_1 && = b_1 \\ & \qquad B_2 x_1 + A_2 x_2 && = b_2 \\ & && \vdots \\ & \qquad\qquad B_T x_{T-1} + A_T x_T && = b_T \\ & \qquad\qquad x_0, x_1, \dots, x_T \geq 0 \end{aligned}$$

→ We approximate this linear program using scenarios

⇒ We build a scenario tree



⇒ 5 scenarios, each defining $(A_1, B_1, b_1, c_1, A_2, B_2, b_2, c_2)$

⇒ Can formulate the scenario problem by considering all scenarios with probabilities $(p^1, \dots, p^5)$ such that $\sum_{k=1}^{5} p^k = 1$

e.g. $\mathbb{E}[c_1^T x_1] \Rightarrow \sum_{k=1}^{5} p^k (c_1^k)^T x_1^k$

$$A_1 x_0 + B_1 x_1 = b_1$$

$$\Downarrow$$

$$\begin{cases} A_1^1 x_0 + B_1^1 x_1^1 = b_1^1 \\ \vdots \\ A_1^5 x_0 + B_1^5 x_1^1 = b_1^5 \end{cases}$$

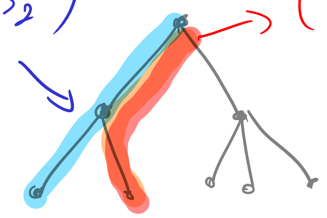
Every scenario k has its own variables $x_1^k, \ldots, x_T^k$

$\Rightarrow$ To guarantee consistency between the scenarios, we add non-anticipativity constraints

$$\forall\, t = 1..T, \quad x_t^k = x_t^\ell \quad \text{for every pair } (k,\ell)$$

$$\text{such that } \left(\xi_1^{(k)}, \ldots, \xi_t^{(k)}\right) = \left(\xi_1^{(\ell)}, \ldots, \xi_t^{(\ell)}\right)$$

$\left(\xi_1^{(1)} = \xi_1, \xi_2^{(1)}\right)$ $\left(\xi_1^{(2)} = \xi_1, \xi_2^{(2)}\right)$



NB: Not needed for two-stage stochastic program

Overall, the scenario approach gives a (very large) linear program, that we can decompose using variants of the L-shaped method (most common because it only requires a linear programming solver) or progressive hedging (requires a nonlinear programming solver)

Ex (J. Gondzio et al, 2005)
Linear six-stage stochastic program, scenario tree has 13 branches $\Rightarrow$ Asset management problem for finance

$\Rightarrow$ 4 million scenarios, 24 million variables, 12 million constraints

$\Rightarrow$ Every iteration of the L-shaped required to solve ~ 2000 linear programs in parallel, each involving ~ 2000 variables

b) Dynamic programming

$\rightarrow$ Idea: Break the multi-stage program into several two-stages, solve it "backwards"

$\rightarrow$ Problem formulation (from control theory)

Convention (equivalent to a minimization problem)

$$\underset{x_0, \dots, x_T}{\text{maximize}} \quad \mathbb{E}_{\xi_1, \dots, \xi_T}\left[\sum_{t=0}^{T} f_t(s_t, x_t, \xi_t) + f_{T+1}(s_{T+1})\right]$$

$$\text{s.t.} \qquad s_{t+1} = \varphi_t(s_t, x_t, \xi_t) \quad \forall t = 0, \dots, T$$

s_t: state at time t, that obeys some dynamics

↳ Dynamic programming works by computing the best x_t given $(x_{t+1}, \dots, x_T)$

Motivation: Bellman's optimality principle

Define $$V(s_0) = \max_{x_0, \dots, x_T} \mathbb{E}\left[\sum_{t=0}^{T} f_t(s_t, x_t, \xi_t) + f_{T+1}(s_{T+1})\right]$$

and $$V_t(s_t) = \max_{x_t, \dots, x_T} \mathbb{E}\left[\sum_{u=t}^{T} f_u(s_u, x_u, \xi_u) + f_{T+1}(s_{T+1})\right]$$

$\forall t = 1, \dots, T$

Then

$$\forall t = 0, \dots, T \qquad V_t(s_t) = \max_{x_t} \left\{ \mathbb{E}_{\xi_t}\left[f_t(s_t, x_t, \xi_t) + V_{t+1}\left(\varphi_t(s_t, x_t, \xi_t)\right) \right] \right\}$$

Algorithmic framework:

- Solve
$$\begin{aligned} \underset{x_T}{\text{maximize}} \quad & \mathbb{E}_{\xi_T}\left[f_T(s_T, x_T, \xi_T) + f_{T+1}(s_{T+1}) \right] \\ \text{s.t.} \quad & s_{T+1} = \varphi_T(s_T, x_T, \xi_T) \end{aligned}$$

as a function of s_T.

- For $t = T-1, \dots, 0$

solve $\underset{x_t}{\text{maximize}} \ \mathbb{E}_{\xi_t}\left[f_t(s_t, x_t, \xi_T) + V_{t+1}\left(\varphi_t(s_t, x_t, \xi_t)\right) \right]$

Classical setup: f_t is quadratic in x_t, φ_t is linear in s_t and x_t

Main software/algorithm: SDDP (Stochastic Dual Dynamic Programming)

$\downarrow$

SDDP.jl

RISK - AVERSE STOCHASTIC PROGRAMMING

↳ Basic approach to stochastic programming

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad \mathbb{E}_\xi [F(x, \xi)] \quad \text{s.t. } x \in \mathcal{X}$$

⇒ risk-neutral model: does not account for the full distribution of $F(x, \xi)$

⇒ $\mathbb{E}_\xi[\cdot]$ can be replaced by other measures that are more risk-averse, i.e. they provide guarantees with respect to values of $F(x, \xi)$ far from the average

Motivation: Finance

→ Suppose that $x \in \mathbb{R}^n$ represents a portfolio of allocations among n assets

→ Suppose that $\xi \in \mathbb{R}^n$ is a random vector of returns on investment

$$\text{Invest } x \Rightarrow \text{Return } \xi^T x$$

Risk-neutral investment: obtained by solving

$$\underset{x \in \mathbb{R}^n}{\text{maximize}} \quad \underbrace{\mathbb{E}_\xi [\xi^T x]}_{= \mathbb{E}_\xi[\xi]^T x} \quad \text{s.t. } \underbrace{x \in \mathcal{X}}_{\substack{\text{suitable constraints} \\ (\text{typically } x \geq 0, \ \sum x_i = 1)}}$$

→ One way to take risk into account: Find the portfolio that minimizes the variance/dispersion

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad \underbrace{\mathbb{E}_{\xi}\left[\left(\left(\xi - \mathbb{E}_{\xi}[\xi]\right)^T x\right)^2\right]}_{= \, x^T V x} \qquad \text{s.t. } x \in \mathcal{X}$$

"minimum-risk portfolio"

where

$$V = \mathbb{E}_{\xi}\left[\left(\xi - \mathbb{E}_{\xi}[\xi]\right)\left(\xi - \mathbb{E}_{\xi}[\xi]\right)^T\right]$$

is the covariance matrix of ξ