STOCHASTIC PROGRAMMING

December 14, 2023

Today:

- → Exercise from the previous lecture
- → Risk measures
- → Probabilistic constraints (time permitting)

Exercise from lecture 3

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

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

Feasible set is nonempty
($\max(1-x,0)$ is feasible)

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

$\rightarrow \xi = 0$

$$Q(x,0) = \min_{y \in \mathbb{R}} \left\{ 0 \text{ s.t. } \begin{array}{l} y \geq 1-x \\ y \geq 0 \end{array} \right\} = 0$$

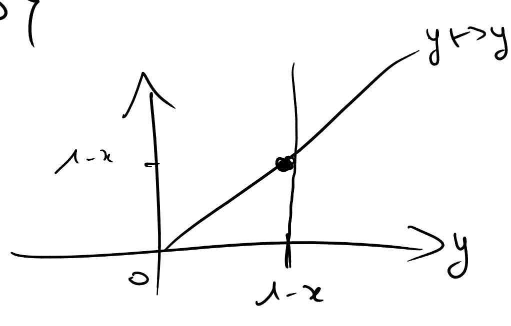
$\rightarrow \xi > 0$

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

(More generally, $\min_{x \in \mathbb{R}^n} \{a \times f(x) \text{ s.t. } x \in \mathcal{X}\} = a \min_{x \in \mathbb{R}^n} \{f(x) \text{ s.t. } x \in \mathcal{X}\}$)
$\underline{\forall a > 0}$

$$\min_{y \in \mathbb{R}} \left\{ y \text{ s.t. } y \geq 1-x,\ y \geq 0 \right\}$$

$$= \max(1-x, 0)$$



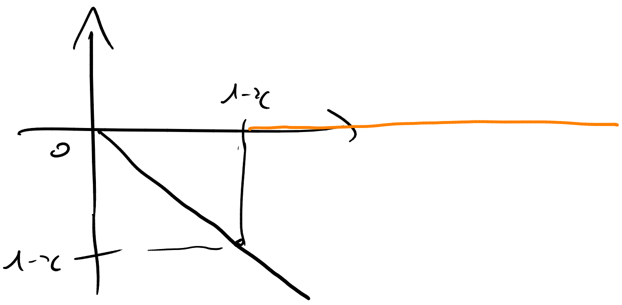
Hence $$Q(x,\xi) = \xi \max(1-x,0)$$

$\rightarrow \xi < 0$

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

$$\min_{y\in\mathbb{R}} \left\{ -y \text{ s.t. } \begin{array}{l} y \geq 1-x \\ y \geq 0 \end{array} \right\} = -\infty$$

because $y \mapsto -y$ is not lower bounded on $[\max(1-x,0), +\infty[$



Hence $Q(x,\xi) = -\infty$

Overall,
$$Q(x,\xi) = \begin{cases} \xi \max(1-x,0) & \text{if } \xi \geq 0 \\ -\infty & \text{if } \xi < 0 \end{cases}$$

b) Suppose that ξ has a discrete distribution in $\{1,3\}$

$$\mathbb{P}(\xi = 1) = \frac{3}{4} \qquad \mathbb{P}(\xi = 3) = \frac{1}{4}$$

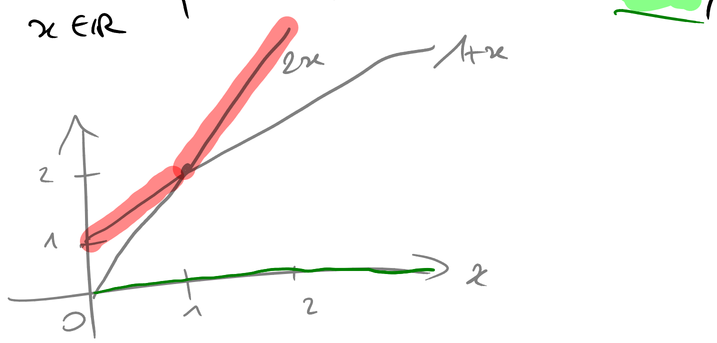
- (Expected) Wait-and-See solution (EWS / WS depending on the literature)

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

$$= \mathbb{P}(\xi=1) \min_{x\in\mathbb{R}} \{ 2x + Q(x,1) \text{ s.t. } x \geq 0 \} + \mathbb{P}(\xi=3) \min_{x\in\mathbb{R}} \{ 2x + Q(x,3) \text{ s.t. } x \geq 0 \}$$

$$= \frac{3}{4} \times \min_{x\in\mathbb{R}} \{ 2x + \max(1-x,0) \text{ s.t. } x \geq 0 \} + \frac{1}{4} \times \min_{x\in\mathbb{R}} \{ 2x + 3\max(1-x,0) \text{ s.t. } x \geq 0 \}$$

$$\rightarrow \min_{x\in\mathbb{R}} \{ 2x + \max(1-x,0) \text{ s.t. } x \geq 0 \} = \min_{x\in\mathbb{R}} \{ \max(1+x, 2x) \text{ s.t. } x \geq 0 \}$$



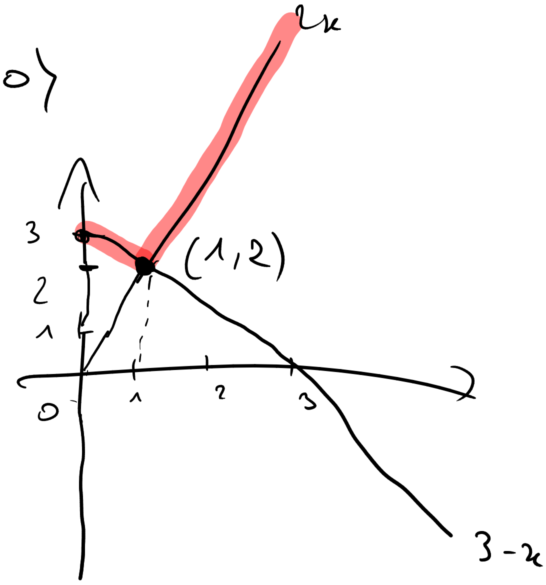
Minimized for $x = 0$

$$\min_{x\in\mathbb{R}} \{ 2x + \max(1-x,0) \text{ s.t. } x \geq 0 \} = 1$$

$$\rightarrow \min_{x\in\mathbb{R}} \{ 2x + 3\max(1-x,0) \text{ s.t. } x \geq 0 \}$$

$$= \min_{x\in\mathbb{R}} \{ \max(3-x, 2x) \text{ s.t. } x\geq 0 \}$$

$$= 2$$



$$EWS = \frac{3}{4}\times 1 + \frac{1}{4}\times 2 = \frac{5}{4}$$

Some authors call this EV and use EEV for "Expectation of Expected value"

$\bar{x} = \operatorname{argmin}_{x\in\mathbb{R}} \{ 2x + Q(x, \mathbb{E}_\xi[\xi]) \text{ s.t. } x\geq 0 \}$

$EEV = \mathbb{E}_\xi[2\bar{x} + Q(\bar{x}, \xi)]$

EEV / EV : Expected value solution

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

$$\mathbb{E}_\xi[\xi] = \mathbb{P}(\xi=1)\times 1 + \mathbb{P}(\xi=3)\times 3 = \frac{3}{4}\times 1 + \frac{1}{4}\times 3 = \frac{3}{2}$$

$$\min_{x\in\mathbb{R}} \{ 2x + Q(x, \tfrac{3}{2}) \text{ s.t. } x\geq 0 \}$$

$$= \min_{x\in\mathbb{R}} \{ 2x + \frac{3}{2}\max(1-x,0) \text{ s.t. } x\geq 0 \}$$

$$= \min_{x\in\mathbb{R}} \{ \max(\frac{3}{2} + \frac{x}{2}, 2x) \text{ s.t. } x\geq 0 \}$$

$= \frac{3}{2}$

$EV = \frac{3}{2} > \underbrace{EWS = \frac{5}{4}}$ $\Rightarrow$ Better on average to wait-and see than deciding from the average value of ξ

NB: $\underset{x \in \mathbb{R}}{\operatorname{argmin}} \left\{ 2x + Q(x, \frac{3}{2}) \text{ s.t. } x \geq 0 \right\} = \{0\}$

$$\mathbb{E}_\xi \left[2 \times 0 + Q(0, \xi) \right] = \frac{3}{4} \times Q(0,1) + \frac{1}{4} Q(0,3)$$
$$= \frac{3}{4} \max(1-0,0) + \frac{1}{4} \times 3 \times \max(1-0,0)$$
$$= 3/2$$

In that example, the two definitions of EEV/EV coincide.

BACK TO RISK-AVERSE STOCHASTIC PROGRAMMING

Motivation: Finance

→ Setup: $x \in \mathbb{R}^n$ portfolio of investments

$\xi \in \mathbb{R}^n$ vector of random returns

→ classical approach: maximize the expected return

Linear program $\left\{ \underset{x \in \mathbb{R}^n}{\text{maximize}} \; \underbrace{\mathbb{E}_\xi[\xi^T x]}_{= \mathbb{E}_\xi[\xi]^T x} \quad \text{s.t. } x \in \mathcal{X} \right.$

↑ Typically $\sum x_i = 1$, $x \geq 0$

$\Rightarrow$ Risk neutral approach: assets with very different values but same average return are treated identically

$\longrightarrow$ Markowitz portfolio allocation problem:

Typically a quadratic program

$$\left\{ \begin{array}{l} \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad \underbrace{-\mu^T x}_{\text{(-)Expected returns}} + \overset{\gamma>0:\ \text{tradeoff parameter}}{\frac{\gamma}{2}} \underbrace{x^T V x}_{\text{volatility / "risk" of the portfolio}} \quad \text{s.t. } \underbrace{x \in \mathcal{X}}_{\sum x_i = 1,\ x \geq 0} \end{array} \right.$$

where $\mu = \mathbb{E}_\xi[\xi]$ and $V = \underbrace{\mathbb{E}_\xi\left[(\xi-\mu)(\xi-\mu)^T\right]}_{\text{covariance matrix of } \xi}$

↳ Combination of the expected return and a risk/volatility estimation term
- γ expresses the importance of risk over expected return

↳ One example of a risk-averse optimization problem (take into account other properties of ξ than its expected value)

⊕ Yields a quadratic program, even when scenarios are used to approximate μ and V

⊖ Can produce very conservative solutions (low investment in assets with large variance, even when they yield good average returns)

Other risk measures

↳ Informal definition: any function of a random variable that contains information about its values away from its expectation or its extreme values

Examples: • Variance / Volatility $var[Z]$, Standard deviation $\sqrt{var[Z]}$

$\underbrace{\qquad\qquad\qquad\qquad}_{\text{Dispersion measures}}$

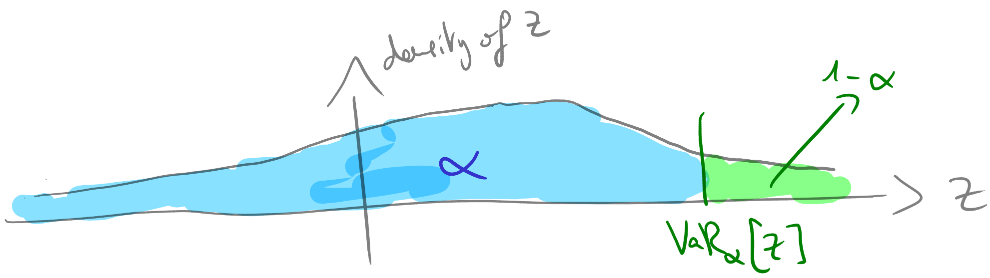
Z random variable

$var[Z] = \mathbb{E}[(Z-\mathbb{E}[Z])^2]$ • $\mathbb{E}[Z] + var[Z]$, $\mathbb{E}[Z] + \sqrt{var[Z]}$

• Value-at-Risk (VaR $\neq$ var)

Def: If Z is a random variable in $\mathbb{R}$ and $\alpha \in (0,1)$, the value-at-risk of level α, denoted by $VaR_\alpha[Z]$, is the value γ such that

$$\mathbb{P}(Z \geq \gamma) = 1-\alpha$$



→ Typically α is chosen close to 1 (0.9, 0.95, 0.99) and the value-at-risk is then an upper bound on Z with high probability (0.1, 0.05, 0.01)

→ Initially proposed by J-P Morgan in the late 2000s, still used today (in the constraints of the portfolio allocation problems, see later)

Proposition: If Z is a discrete random variable $Z \in \{z_1, \dots, z_K\}$ with $\mathbb{P}(Z = z_k) = p_k \; \forall k = 1..K$

$z_1 \leq \dots \leq z_K$

Then $\forall \alpha \in (0,1)$, $\quad VaR_\alpha[Z] = z_{k_\alpha}$, where k_α is the largest value in $\{1, \dots, K\}$ such that

$$\sum_{k=k_\alpha}^{K} p_k \geq 1-\alpha$$

→ The Value-at-risk has several drawbacks, including that it does not characterize how large Z can be with probability $1-\alpha$ ($Z \geq VaR_\alpha[Z]$), and that $VaR_\alpha[Z+Y] \not\leq VaR_\alpha[Y] + VaR_\alpha[Z]$

⇒ Mathematically, these are a consequence of VaR not being a <u>coherent</u> risk measure

Def: $\rho : \mathbb{R} \to \mathbb{R}$ is a coherent risk measure if for any random variables Z, Y, the following properties hold:

- $Z \geq 0 \Rightarrow \rho(Z) \geq 0$
- $\rho(Y+Z) \leq \rho(Y) + \rho(Z)$ ("subadditivity")
- $\rho(cZ) = c\rho(Z) \quad \forall c > 0$
- $\rho(Z+c) = \rho(Z) + c \quad \forall c \in \mathbb{R}$

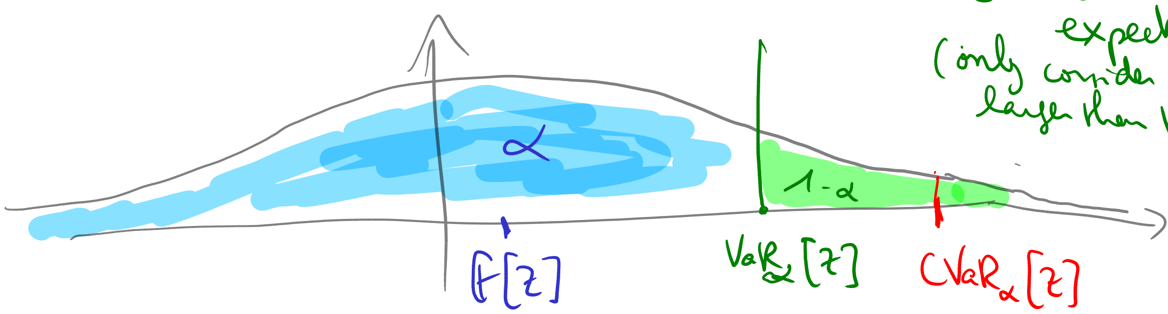
↳ There exists a coherent risk measure based on the value-at-risk, called the conditional value-at-risk or average value-at-risk

(Notation for this course: CVaR
others: AVaR, AV@R, CV@R, ...)

Def: Let Z be a random variable with values in $\mathbb{R}$, and $\alpha \in (0,1)$. The conditional value-at-risk of level α for Z is defined as:

$$\mathrm{CVaR}_\alpha[Z] = \mathbb{E}_Z\left[Z \mid Z \geq \mathrm{VaR}_\alpha[Z] \right]$$

conditional expectation (only consider values larger than $\mathrm{VaR}_\alpha[Z]$)



Special case: $Z \in \{z_1, \ldots, z_K\}$ $\quad z_1 \leq \ldots \leq z_K$

$$\mathbb{P}(Z = z_k) = p_k$$

k_α: largest integer such that $\sum_{k=k_\alpha}^{K} p_k \geq 1-\alpha$

$$\mathrm{CVaR}_\alpha[Z] = \frac{1}{1-\alpha} \sum_{k=k_\alpha}^{K} p_k z_k$$

Theorem

Let Z be a random variable with values in $\mathbb{R}$ and $\alpha \in (0,1)$.

Then

$$\mathrm{CVaR}_\alpha[Z] = \min_{\gamma \in \mathbb{R}} \left\{ \gamma + \frac{1}{1-\alpha} \mathbb{E}_Z\left[\max(Z-\gamma, 0)\right] \right\}$$

and

$$\{\mathrm{VaR}_\alpha[Z]\} = \operatorname{argmin}_{\gamma \in \mathbb{R}} \left\{ \gamma + \frac{1}{1-\alpha} \mathbb{E}_Z\left[\max(Z-\gamma, 0)\right] \right\}$$

→ Not a closed-form formula for $\mathrm{CVaR}_\alpha[Z]$ or $\mathrm{VaR}_\alpha[Z]$

→ Shows that $\mathrm{CVaR}_\alpha[Z]$ (and $\mathrm{VaR}_\alpha[Z]$) can be computed by solving a two-stage linear stochastic program

$$\underset{\gamma \in \mathbb{R}}{\text{minimize}} \quad \gamma + \frac{1}{1-\alpha} \mathbb{E}\left[\max(Z-\gamma, 0)\right]$$

can be reformulated as

$$\underset{\gamma \in \mathbb{R}}{\text{minimize}} \left\{ \gamma + \mathbb{E}_Z\left[Q(\gamma, Z)\right] \right\}$$

where $Q(\gamma, Z) = \underset{y \in \mathbb{R}}{\min} \left\{ \frac{1}{1-\alpha} y \quad \text{s.t.} \quad \begin{array}{l} y \geq Z - \gamma \\ y \geq 0 \end{array} \right\}$

$$\mathrm{CVaR}_\alpha[Z] = \mathrm{VaR}_\alpha[Z] + \frac{1}{1-\alpha} \mathbb{E}\left[\max(Z - \mathrm{VaR}_\alpha[Z], 0)\right]$$

$$= \mathrm{VaR}_\alpha[Z] + \frac{1}{1-\alpha}\Big[\mathbb{P}(Z \geq \mathrm{VaR}_\alpha[Z])\, \mathbb{E}\left[Z - \mathrm{VaR}_\alpha[Z] \mid Z \geq \mathrm{VaR}_\alpha[Z]\right] + \mathbb{P}(Z < \mathrm{VaR}_\alpha[Z]) \times 0\Big]$$

$$= \mathrm{VaR}_\alpha[Z] + \frac{1}{1-\alpha} \underbrace{\mathbb{P}(Z \geq \mathrm{VaR}_\alpha[Z])}_{=1-\alpha} \mathbb{E}\left[Z - \mathrm{VaR}_\alpha[Z] \mid Z \geq \mathrm{VaR}_\alpha[Z]\right]$$

$$= \mathrm{VaR}_\alpha(Z) + \mathbb{E}\left[Z - \mathrm{VaR}_\alpha[Z] \mid Z \geq \mathrm{VaR}_\alpha[Z]\right]$$

$$= \mathbb{E}\left[Z \mid Z \geq \mathrm{VaR}_\alpha[Z]\right]$$

↳ Another example of risk measure: mean absolute deviation

$\mathbb{E}[|Z|]$ for a real random variable Z

Exercise: Consider the portfolio allocation problem

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \ \mathbb{E}\left[|\xi^T x| \right] \quad \text{s.t.} \ \sum_{i=1}^{n} x_i = 1$$
$$x \geq 0$$
$$\left(\Leftrightarrow x_1 \geq 0, \ \vdots, \ x_n \geq 0 \right)$$

a) Show that this problem can be rewritten as a linear two-stage stochastic program,

b) Suppose that $\xi \in \{ \xi_1, \dots, \xi_K \}$ with $\mathbb{P}(\xi = \xi_i) = p_i$. Write the corresponding (deterministic) linear program

Using risk measures in stochastic programming

- First option: Use a risk measure in the objective ($\simeq$ Markowitz)

Idea: Replace the expected value in classical stochastic programming by a risk measure.

"classical stochastic programming"

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

$$\Downarrow$$

Risk-averse stochastic program
$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \ \rho\left(F(x, \xi) \right) \quad \text{s.t.} \ x \in X$$

ρ : risk measure ($CVaR_\alpha$, VaR_α, $\mathbb{E}[\cdot] + \text{var}(\cdot)$, ...)

⚠ $\rho(F(x,\xi))$ is not random

but it depends on the distribution of ξ

→ Risk-averse stochastic programs can be tackled using the same tools as risk-neutral (i.e. $\mathbb{E}[\cdot]$) stochastic programs:

- scenario decomposition
- multistage formulation
- progressive hedging / variants of the L-shaped method

→ BUT because most risk measures are nonlinear functions, solving these problems requires deterministic solvers for nonlinear optimization problems

Remark

For coherent risk measures ρ, we can reformulate $\rho(Z)$ as the optimal value of an optimization problem over probability distributions

$$\rho(F(x,\xi)) = \max_{D \in \mathbb{D}} \mathbb{E}_{\xi \sim D}[F(x,\xi)]$$

Family of probability distributions that is defined according to ρ

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

is called a distributionally robust optimization problem

⇒ modern perspective on risk-averse optimization, still developing (and hard to solve!)

Second approach: Use risk measures in constraints

↳ Back to finance: Markowitz's formulation is equivalent to

$$(*)\begin{cases} \underset{x\in\mathbb{R}^n}{\text{maximize}} \ \mathbb{E}_\xi[\xi]^T x \quad \text{s.t.} \quad x \in X \\ \qquad\qquad\qquad\qquad\qquad\qquad x^T V x \leq \sigma^2 \end{cases}$$

$\exists \gamma > 0$ such that

$$\underset{x\in\mathbb{R}^n}{\text{minimize}} \ -\mathbb{E}_\xi[\xi]^T x + \frac{\gamma}{2} x^T V x \quad \text{s.t. } x \in X$$

has the same solution than $(*)$

$\gamma = \gamma(\sigma^2)$

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

$(*)$ is not a quadratic program (i.e. quadratic objective + linear constraints) but a conic program, which is typically harder to solve.

⟶ In principle, we can add constraints on risk measures to any risk-neutral formulation to make it risk-averse.

$$(Ex)\quad \begin{cases} \underset{x\in\mathbb{R}^n}{\text{minimize}} \ \mathbb{E}_\xi[F(x,\xi)] \\ \text{s.t.} \quad VaR_\alpha[F(x,\xi)] \leq \gamma \\ \qquad\quad x \in X \end{cases}$$

for some $\gamma \in \mathbb{R}$

$$VaR_\alpha[F(x,\xi)] \leq \gamma$$

$$\Leftrightarrow \underbrace{\mathbb{P}(F(x,\xi) \geq \gamma) \leq 1-\alpha}_{} \qquad (\text{equality when } \gamma = VaR_\alpha[F(x,\xi)])$$

↑ Probabilistic/chance constraint