Advanced Topics in Control Systems

Exercises and Project Ideas

Pantelis Sopasakis

KU Leuven, Department of Electrical Engineering (ESAT),
STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics,
Kasteelpark Arenberg 10, 3001 Leuven, Belgium.

Last Updated: January 6, 2018

Abstract

For the final evaluation for the PhD course "Advanced Topics in Control Systems" you either have to do a project related to the course, or solve a set of exercises. In this document you will find a few ideas for projects and the exercises which you, alternatively, have to solve. Address any questions to Pantelis Sopasakis.

Contents

1 Project ideas

Any project related to the course is acceptable. It can be either a *literature survey*, or a *computational project*, or a *theoretical study*. In either case, you have to prepare a report with the structure of a conference or journal article, 6 to 9 pages long and up to 12 pages for literature surveys in the IFAC conference double-column format.

1.1 Literature review

If you choose the literature survey, you can look into the following topics:[1]

Economic MPC

i. Economic MPC and relevant applications, ii. Economic vs conventional RTO/ MPC methodologies, iii. Open problems and future directions in economic MPC.

Markovian and stochastic systems

i. MPC for constrained stochastic systems, ii. Scenario - based stochastic MPC and applications (e.g., supply chains, portfolio selection, robotics, energy management and dispatch, water networks, automotive applications, etc), iii. Scenario generation algorithms, iv. Chance-constrained MPC, v. Almost sure stabilisation of stochastic systems, vi. Stochastic tube methods, vii. Open problems and future research directions in MPC for stochastic systems.

Risk measures

i. Risk measures and their properties (strict/strong monotonicity and other theoretical properties) — several novel risk measures have been proposed in the literature such as the *expectiles*. A survey on risk measures would attempt to summarise the merits of the various risk measures found in the literature and discuss their potential applications, ii. Statistical estimation of risk measures, iii. The Kusuoka theorem and its applications.

Risk-averse optimisation and control

i. Risk-averse vs risk-neutral problem formulations, ii. Two-state vs multistage risk-averse formulations: differences, properties, numerical methods, iii. Dualisation of non-anticipativity constraints and applications, iv. Risk-averse optimal control and distrubutionally robust problem formulations, v. The problem of moments in risk-averse optimisation, vi. New formulations of optimal control problems involving risk measures and their properties such as formulations based on *separable expected conditionals*, vii. Information monotonicity of risk-averse formulations, viii. Discussion of the property of time consistency of multistage risk-averse problems.;

1.2 Application projects

Application projects can showcase the application of (i) economic MPC or (ii) stochastic MPC or (iii) risk-averse optimal control for the solution of a practical problem. You need to provide a short literature background indicating the state of the art, introduce the theoretical tools you used and demonstrate with simulations the advantages of the solution you propose. You may also combine any of the ideas listed above with simulations.

1.3 Theoretical studies

You may address one or more theoretical questions in any of the above fields.

[1] The topics listed here are only indicative. Feel free to study any different problem which is related to the course.

1.4 Requirements

Exercises. Your solutions to these exercises have to be typeset in LaTeX(or similar typesetting programmes like LuaTeX, XeTeX, etc). Acceptable solutions must be rigorous, explain all involved steps and state clearly all assumptions — these will be the criteria for their evaluation. Figures must include proper axis titles and captions and they need to be properly typeset (do not use screenshots).

Projects. Similarly, all projects must be typeset in LaTeXand have an abstract and a "Conclusions" section. If your project is a literature review, it will be graded on the basis of how complete it is and — *primarily* — how good understanding of the literature you have.

Submission. Send your final submission by email to Pantelis Sopasakis CCing Prof. Alberto Bemporad with subject "[ATCS] Final project" or "[ATCS] Exercises".

2 Exercises

This is a set of 27 exercises on three thematic sections: *economic model predictive control*, *Markovian stochastic processes and systems* and *risk measures*. In each exercise, its contribution to the total mark is shown in parentheses. The maximum mark is 100, while all marks sum up to 134. This allows you to choose which exercises you will solve. Unless otherwise stated, subquestions give equal points. These exercises are an alternative to doing your own project.

Definitions and Preliminaries

In what follows, unless otherwise specified, $(\Omega, \mathscr{F}, \mathrm{P})$ is a probability space, $\Omega \neq \varnothing$, $\mathcal{Z} = \mathcal{L}_p(\Omega, \mathscr{F}, \mathrm{P})$ with $p \in [1, \infty)$, $\mathcal{Z}^* = \mathcal{L}_q(\Omega, \mathscr{F}, \mathrm{P})$ is the topological dual of $\mathcal{Z}$ with q such that $q^{-1} + p^{-1} = 1$. For $X \in \mathcal{Z}$ and $Y \in \mathcal{Z}^*$ we define the product of the two random variables to be $\langle X, Y \rangle = \int_\Omega XY \mathrm{dP} = \int_\Omega X(\omega)Y(\omega)\mathrm{dP}(\omega)$. The abbreviation a.s. stands for *almost surely*. The *characteristic function* of a set A is defined as

$$1_A(x) = \begin{cases} 1, & \text{if } x \in A, \\ 0, & \text{otherwise} \end{cases}$$

The average value-at-risk of $Z \in \mathcal{Z}$ is defined as

$$\mathrm{AV@R}_\alpha[Z] = \inf_{t \in \mathbb{R}} \{t + \alpha^{-1}\mathbb{E}[Z - t]_+\},$$

where $[\,Z(\omega)\,]_+ = \max\{Z(\omega), 0\}$. The *essential supremum* of a random variable Z on $(\Omega, \mathscr{F}, \mathrm{P})$ is defined as

$$\operatorname{ess\,sup} Z = \inf_{\alpha \in \mathbb{R}} \{\mathrm{P}[Z > \alpha] = 0\}.$$

A risk measure is called *coherent* if it satisfies the following assumptions for $Z, V \in \mathcal{Z}$, $\lambda \geq 0$ and $a \in \mathbb{R}$

1. Subadditivity $\rho(Z + V) \leq \rho(Z) + \rho(V)$
2. Positive homog. $\rho(\lambda Z) \leq \lambda \rho(Z)$
3. Monotonicity $Z \leq V \Rightarrow \rho(Z) \leq \rho(V)$
4. Translation invariance $\rho(Z + a) = a + \rho(Z)$

We say that two random variables X and Y on $(\Omega, \mathscr{F}, \mathrm{P})$ are equal in distribution if for all $A \in \mathscr{F}$, $\mathrm{P}[X \in A] = \mathrm{P}[Y \in A]$. This does not imply that the two random variables are P-a.s. equal.

Let $\mathcal{Z}_i = \mathcal{L}_p(\Omega, \mathscr{F}_i, \mathrm{P})$ for $i = 0, \dots, N$ be a sequence of probability spaces with $\mathcal{Z}_N = \mathcal{Z}$ ($\mathscr{F}_N = \mathscr{F}$) and $\mathscr{F}_0 = \{\varnothing, \Omega\}$ is the trivial σ-algebra. If a random variable Z is $\mathcal{Z}_0$-measurable, then, in fact, it is deterministic. We say that a mapping $\rho : \mathcal{Z}_0 \times \mathcal{Z}_1 \times \ldots \times \mathcal{Z}_N$ is a multiperiod risk measure if it satisfies the following axioms for $Z = (Z_0, Z_1, \dots, Z_N)$ and $Z' = (Z'_0, Z'_1, \dots, Z'_N)$ in $\mathcal{Z}_0 \times \mathcal{Z}_1 \times \ldots \times \mathcal{Z}_N$

1. Subadditivity $\rho(Z + Z') \leq \rho(Z) + \rho(Z')$
2. Positive homog. $\rho(\lambda Z) \leq \lambda \rho(Z)$
3. Monotonicity $Z \leq Z' \Rightarrow \rho(Z) \leq \rho(Z')$
4. Translation invariance: for every random variable Y which is $\mathscr{F}_i$-measurable, $i = 0, \dots, N-1$, it is $\rho(Z_0, \dots, Z_k, Z_{k+1} + Y, \dots, Z_N) = \rho(Z_0, \dots, Z_k + Y, Z_{k+1}, \dots, Z_N)$.

Notice that the first three axioms are very similar to the ones stated above for coherent risk measures.

2.1 Economic MPC

Max. points in this section: 8.

1–1. Dissipativity and economic MPC (8). Consider the following nonlinear discrete-time system with 2 states and 1 input

$$x_{k+1} = x_k \left(1 + \|x_k\|\right) u_k$$

and the cost functional $\ell : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ given by

$$\ell(x, u) = \|x\| \cdot |3 - u|.$$

The system is subject to the constraints $x_k \in [-5, 5] \times [-5, 5]$ and $u_k \in [-2, 2]$. The norm $\| \cdot \|$ is the Euclidean norm.

(i) Find all equilibrium points of the system,

(ii) verify that $x_s = 0$, $u_s = 0$ is an optimal steady state with respect to ℓ and

(iii) show that the system is dissipative with respect to the supply rate $s(x, u) = \ell(x, u) - \ell(x_s, u_s)$[2],

(iv) implement an economic model predictive controller using ℓ as a stage cost, with horizon $N = 15$ and using the terminal constraint $x_N = 0$. Demonstrate with simulations that your formulation is recursively feasible.

2.2 Markovian processes and systems

Max. points in this section: 33.

2–1. Markovianity: History does not matter (3). Does it? Often, a rough description of what a Markov process runs

> "a process where the past (times $k-1, k-2, \ldots$) has no effect on the future (time $k+1$) once the present is known (time k)."

This said, one would infer that the following is correct

$$\mathrm{P}[x_{k+1} \in A \mid x_k \in B, x_{k-1} \in C] = \mathrm{P}[x_{k+1} \in A \mid x_k \in B],$$

for all $k \in \mathbb{N}$ and for measurable sets A, B and C. Although Fukuyama would probably find this correct, it is not. Can you see why? Show this with a counterexample.

2–2. MJLS oddity (2+2). Construct a Markovian system $x_{k+1} = A_{\theta_k} x_k$ with $x_k \in \mathbb{R}^2$ so that A_{θ_k} are all stable matrices, $\mathbb{E}[x_k] \to 0$, but the covariance of x_k diverges.

Construct an MJLS $x_{k+1} = A_{\theta_k} x_k + B_{\theta_k} u_k$ with $\theta_k \in \{1, 2\}$ so that (A_1, B_1) is not stabilisable, but there are K_i so that $x_{k+1} = (A_{\theta_k} + B_{\theta_k} K_{\theta_k}) x_k$ is mean square stable.

Find an example of a system with $x_{k+1} = A_{\theta_k} x_k$ with $x_k \in \mathbb{R}^2$ and $\theta_k \in \{1, 2\}$ which is almost surely convergent, but not mean-stable, that is $\mathbb{E}[x_k]$ diverges.

2–3. Stabilization of MJLS (2+3). Consider a MJLS with two modes described by the system dynamics $x_{k+1} = A_{\theta_k} x_k + B_{\theta_k} u_k$ and system matrices

$$A_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, A_2 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, B_1 = \begin{bmatrix} 1 \\ 0.1 \end{bmatrix}, B_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

and probability transition matrix

$$P = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}.$$

[2] The system is not dissipative with respect to the storage function $\lambda(x) = \alpha\|x\|$. However, if you start working with this function, you might get some intuition that will help you solve the problem. Note that there is no unique continuous function $\lambda : \mathbb{R}^2 \to \mathbb{R}$ so that $\lambda(f(x, u)) - \lambda(x) \leq s(x, u)$. Do not forget to take into account the state and input constraints.

Suppose that x_k and θ_k can be observed at time k. (i) Design a linear feedback $u(x,i) = K_i x$ which stabilises the closed-loop system in the mean-square sense. (ii) Assume that θ_k is not known at time k, but, instead, an estimate $\hat{\theta}_k$ of it is available. We know that

$$\begin{aligned} \mathrm{P}[\hat{\theta} = 1 \mid \theta = 1] &= 0.97 \\ \mathrm{P}[\hat{\theta} = 2 \mid \theta = 2] &= 0.95. \end{aligned}$$

Design a feedback controller of the form $u(x, \hat{\theta}) = K_{\hat{\theta}} x$ so that the closed-loop system, that is

$$x_{k+1} = (A_{\theta_k} + B_{\theta_k} K_{\hat{\theta}_k}) x_k,$$

is mean-square stable. Provide simulation results.

2–4. Limit probability (3). Consider a Markov chain with 3 modes and transition matrix

$$P = \begin{bmatrix} 0.1 & 0.2 & 0.8 \\ 0.2 & 0.1 & 0.7 \\ 0.9 & 0 & 0.1 \end{bmatrix}$$

(i) Compute its limit probability. (ii) Assume that the initial distribution is $v = \begin{bmatrix} 0.2 & 0.1 & 0.7 \end{bmatrix}$ — that is, $\mathrm{P}[\theta_0 = 1] = 0.2$, $\mathrm{P}[\theta_0 = 2] = 0.1$ and $\mathrm{P}[\theta_0 = 2] = 0.7$. Determine the probabilities $\mathrm{P}[\theta_5 = i]$, $\mathrm{P}[\theta_{20} = i]$, $\mathrm{P}[\theta_{50} = i]$ for $i = 1, 2, 3$.

2–5. Stochastic MPC (10). Consider a Markovian system describing an industrial process with two modes corresponding to proper (1) and faulty (2) operation. The transition matrix between the two modes is

$$P = \begin{bmatrix} 0.98 & 0.01 \\ 0.1 & 0.9 \end{bmatrix}$$

The system data are

$$A_1 = \begin{bmatrix} 1.0 & 0.7 \\ -0.75 & 0.4 \end{bmatrix}, A_2 = \begin{bmatrix} 1.1 & 0.9 \\ -0.9 & 0.3 \end{bmatrix}, B_1 = B_2 = \begin{bmatrix} 0.8 \\ 0.1 \end{bmatrix}.$$

The system is subject to the state and input constraints

$$\begin{aligned} -\begin{bmatrix} 5 \\ 5 \end{bmatrix} \leq & x_k \leq \begin{bmatrix} 5 \\ 5 \end{bmatrix} \\ -3 \leq & u_k \leq 3 \end{aligned}$$

Design and implement (e.g., in MATLAB) a stochastic model predictive controller with horizon N so that the controlled system satisfies the constraints and renders the closed-loop system mean-square stable. Simulate the closed-loop behaviour of the system starting from various initial states.

2–6. MSS sequence with diverging higher moment (2). Take a random variable ξ_n^p with parameter p which takes values in $\{0, n\}$, with probability distribution $\mathrm{P}[\xi_n^p = n] = p \cdot n$. Let $X = 0$ and X_k, $k \in \mathbb{N}$, be defined as $X_k = \xi_k^{p_k}$ where $p_k = 1/(k^2 \ln k)$. Show that X_k converges to 0 in the mean-square sense, but $\mathbb{E}[|X_k|^s]$ diverges for all $s > 2$.

2–7. Diverging higher moment in MJLS (6). Consider a Markovian linear system $x_{k+1} = A_{\theta_k} x_k$, where θ_k is a Markov process. Suppose that the system is mean-square stable. Does $\|x_k\|_p$ always converge for $p \in (2, \infty)$? Give an example of an MJLS which is mean-square stable, but not stable in a higher mean (e.g., $\mathbb{E}[\|x_k\|^4]$ diverges).

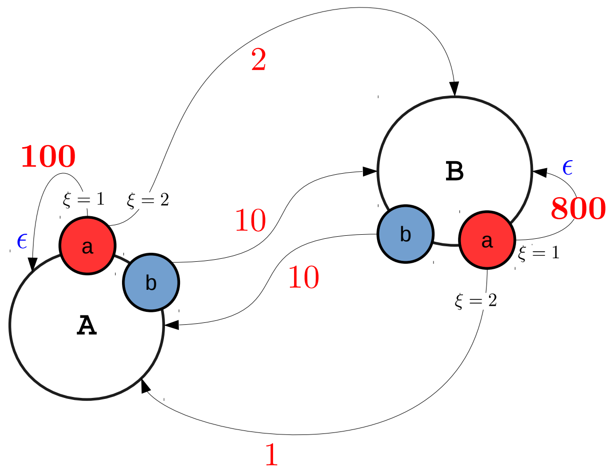
2.3 Stochastics and Risk measures

Max. points in this section: 93.

3–1. Average value-at-risk of normally distributed variables (2). Take a normally distributed real-valued random variable $Z \sim \mathcal{N}(\mu, \sigma^2)$. Show that

$$\mathrm{AV@R}_\alpha[Z] = \mu + \frac{\sigma}{\alpha\sqrt{2\pi}} \exp\left(-\frac{\Phi^{-1}(1-\alpha)}{2}\right)$$

3–2. Markov decision processes and risk-averse decision making (4). Consider the problem shown in the following figure which shows a *Markov decision process*. This is an one-player game which works as follows: assume we are at node `A`. We have two possible actions: a and b. If we choose b we know that the next node will be `B`; this decision will cost us €10. Otherwise, if we choose to make the decision a we will either land in node `A` and pay €100, or land in `B` and pay €2. Clearly, not going for the deterministic choice b incurs a risk: we will either pay less (with probability $1-\epsilon$), or, we risk to have to pay €100 (with probability ϵ). When at node `B`, we are facing an even more extreme high-gain-high-loss situation: we will either choose to pay €10 by choosing b, or choose a and risk to pay €800 with probability $1-\epsilon$.



Let $\epsilon = 0.01$. How would a risk-neutral decision maker (who decides using $\mathbb{E}$) behave? How would a decision maker behave if they account for the worst-case outcome? How would one play if they used the average value-at-risk with some confidence level $\alpha \in (0,1)$?

3–3. New risk measures from old ones (2). Let $\rho : \mathcal{Z} \to \bar{\mathbb{R}}$ be a coherent risk measure. Then, show that $\tilde{\rho} = \lambda\mathbb{E} + (1-\lambda)\rho$, with $\lambda \in [0,1]$, is a coherent risk measure too.

3–4. Coherency (4). Show that the following functional

$$\mathcal{H}_\alpha(X) = \frac{\log \mathbb{E}(e^{\alpha X})}{\alpha},$$

with $\alpha > 0$, which is known as the *exponential premium*, in not a coherent risk measure.

3–5. Zero Risk (5). Let $(\Omega, \mathscr{F}, \mathrm{P})$ be a nonatomic probability space and $\rho : \mathcal{Z} \to \mathbb{R}$ be a law-invariant risk measure. Assume that ρ is proper, lower semicontinuous and coherent. Let Z be an almost-everywhere nonnegative random variable on $(\Omega, \mathscr{F}, \mathrm{P})$. Then, prove that $\rho[Z] = 0$ if and only if $Z = 0$ almost everywhere.

3–6. Subgradient (7). Let $\rho : \mathcal{Z} \to \bar{\mathbb{R}}$ be a lower semicontinuous coherent risk measure. The subgradient of ρ at $X_0 \in \mathcal{Z}$ is defined as

$$(\partial\rho)[X_0] = \left\{ Y \in \mathcal{Z}^* \;\middle|\; \begin{array}{l} \rho(X) - \rho(X_0) \geq \langle Y, X - X_0 \rangle \\ \text{for all } X \in \mathcal{Z} \end{array} \right\}$$

Given that ρ is given as $\rho[Z] = \sup_{Y \in \mathfrak{A}} \langle Y, Z \rangle$, show that

$$(\partial\rho)[X_0] = \operatorname*{arg\,max}_{Y \in \mathfrak{A}} \langle Y, X_0 \rangle.$$

3–7. Polytopic risk measure (5). Let ρ be a *polytopic risk measure*, that is, let $Y_1, \ldots, Y_\kappa \in \mathcal{Z}^*$ with $Y_i \geq 0$ a.s. and $\mathbb{E}[Y] = 1$ and let

$$\mathfrak{A} = \operatorname{conv}\{Y_i\}_{i=1,..,\kappa}$$

that is, $\mathfrak{A}$ is the convex hull of $Y_1, \ldots, Y_\kappa$. Define a risk measure $\rho : \mathcal{Z} \to \mathbb{R}$ by $\rho[Z] = \max_{Y \in \mathfrak{A}} \langle Y, Z \rangle$. Show that ρ is a coherent risk measure, compute its subgradient and show that for all $Z \in \mathcal{Z}^*$, $\partial\rho(Z)$ is nonempty, convex and bounded. Show that polytopic risk measures are continuous.

3–8. Strict monotonicity of $\mathrm{AV@R}_\alpha$ **(5).** We know that $\mathrm{AV@R}_\alpha$ is monotone in the sense that $\mathrm{AV@R}_\alpha[Z_1] \leq \mathrm{AV@R}_\alpha[Z_2]$ whenever $Z_1 \leq Z_2$. Assume $\alpha \in (0,1)$. Then show that $\mathrm{AV@R}_\alpha$ is also strictly monotone, i.e., $\mathrm{AV@R}_\alpha[Z_1] < \mathrm{AV@R}_\alpha[Z_2]$ whenever $Z_1 < Z_2$. Show with a counterexample that the essential supremum operator does not possess this property[3].

3–9. Strict monotonicity (5). Let ρ be strictly monotone. Show that $\tilde{\rho} = \lambda \mathbb{E} + (1-\lambda)\rho$, with $\lambda \in [0,1]$, is also strictly monotone.

3–10. Strong monotonicity (7). Let $\rho : \mathcal{Z} \to \bar{\mathbb{R}}$ be a coherent risk measure. For $Z, Z' \in \mathcal{Z}$ we denote $Z \prec Z'$ whenever $Z(\omega) \leq Z'(\omega)$ for almost all $\omega \in \Omega$ and $\mathrm{P}(Z < Z') > 0$. We say that ρ is *strongly monotone* if $Z \prec Z'$ implies that $\rho(Z) < \rho(Z')$. Show that

(i) The expectation $\mathbb{E}$ is strongly monotone,

(ii) $\mathrm{AV@R}_\alpha$ is not strongly monotone (you can do that by counterexample),

(iii) if ρ is monotone and $\lambda \in (0,1]$, then $\tilde{\rho} = \lambda \mathbb{E} + (1-\lambda)\rho$ is strongly monotone,

(iv) For a coherent risk measure ρ, if for all Z there is $Y \in \partial\rho(Z)$, which is $Y > 0$ a.e., then ρ is strongly monotone.

(v) (*) Let ρ be a coherent risk measure. Then, ρ is strongly monotone if and only if for all $Z \in \mathcal{Z}$ and $Y \in \partial\rho(Z)$, we have $Y > 0$ a.s.

3–11. Dual representation of certain risk measures (3+8). Recall that the p-norm (the norm of $\mathcal{L}_p(\Omega, \mathscr{F}, \mathrm{P})$) is defined as

$$\|Z\|_p := \left(\int_\Omega Z^p \mathrm{dP} \right)^{1/p}$$

for $p \in [1, \infty)$ and

$$\|Z\|_\infty = \operatorname{ess\,sup}[\, Z \,],$$

where $\operatorname{ess\,sup} Z$ is the *essential supremum* of Z. Show that (i) $\rho_p : \mathcal{L}_p(\Omega, \mathscr{F}, \mathrm{P}) \to \mathbb{R}$ defined by $\rho_p[Z] = \|Z\|_p$ are coherent risk measures; (ii) show that for $p \in [1,\infty)$ ρ_p is strongly monotone; (iii) show that for $p \in [1,\infty)$ derive the dual representation of ρ_p, that is, find sets $\mathfrak{Q}_p \subseteq \mathcal{L}_q(\Omega, \mathscr{F}, \mathrm{P})$ so that

$$\rho_p[Z] = \sup_{\zeta \in \mathfrak{Q}_p} \langle \zeta, Z \rangle.$$

(iv*)(optional, 8 points) Write ρ_∞ in the form

$$\rho_\infty[Z] = \sup_{\mu \in \mathfrak{Q}_\infty} \mathbb{E}_\mu[Z].$$

Hint: For (iv), keep in mind that the dual of $\mathcal{L}_\infty(\Omega, \mathscr{F}, \mathrm{P})$ is not an $\mathcal{L}_p$ space, but the space $\mathfrak{ba}(\Omega, \mathscr{F}, \mathrm{P})$ — the space of *bounded* and *finitely additive signed measures* on $(\Omega, \mathscr{F})$ which are *absolutely continuous* with respect to P.

3–12. $\mathrm{AV@R}_\alpha$ **of sum (3).** Let $\mathcal{Z} = \mathcal{Z}_0 \times \ldots \times \mathcal{Z}_{N-1}$ and $\varrho : \mathcal{Z} \to \mathbb{R}$ be defined as follows

$$\varrho[Z] = \varrho[Z_0, Z_1, \ldots, Z_{N-1}] = \mathrm{AV@R}_\alpha \left[\sum_{i=0}^{N-1} Z_i \right],$$

for $Z \in \mathcal{Z}$ ($Z_i \in \mathcal{Z}_i$ for $i = 0, \ldots, N-1$). However, this risk measure is not used in practice in the contect of multistage risk-averse optimisation because it suffers from certain pathodologies. Can you spot them?

3–13. Conditional risk mappings (7). Consider two spaces of random variables $\mathcal{Z}_i = \mathcal{L}_p(\Omega, \mathscr{F}_i, \mathrm{P})$, $i = 1, 2$ with $\mathscr{F}_1 \subseteq \mathscr{F}_2$. Let us recall the following definition: A mapping $\rho : \mathcal{Z}_2 \to \mathcal{Z}_1$ is called a *conditional risk mapping* if it satisfies the following axioms

1. $\rho(\lambda Z + (1-\lambda)Z') \leq \lambda\rho(Z) + (1-\lambda)\rho(Z')$ for all $\lambda \in [0,1]$ and $Z, Z' \in \mathcal{Z}_2$

[3] Note that in finite probability spaces (when $\mathscr{F}$ is finite), the essential supremum becomes strictly monotone. We need to look for a counterexample in an infinite dimensional space. Try to find such an example on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$.

2. $\rho(Z) \le \rho(Z')$ for all $Z, Z' \in \mathcal{Z}_2$ with $Z \le Z'$.
3. For all $Y \in \mathcal{Z}_1$ and $Z \in \mathcal{Z}_2$, $\rho(Y + Z) = Y + \rho(Z)$
4. For every $\lambda \ge 0$ and $Z \in \mathcal{Z}_2$, $\rho(\lambda Z) = \lambda \rho(Z)$.

where $\le$ are meant in the almost-sure sense. Show that for $Y \in \mathcal{Z}_1$, $Y \ge 0$ (a.s.) it is $\rho(YZ) = Y\rho(Z)$.

3–14. Conditional risk mappings from coherent risk measures (3). Let $\mathcal{Z}_i = \mathcal{L}_p(\Omega, \mathscr{F}_i, \mathrm{P})$ with $\mathscr{F}_1 \subseteq \mathscr{F}_2$. Given a coherent risk measure $\rho : \mathcal{Z}_2 \to \bar{\mathbb{R}}$ which has the dual representation

$$\rho(Z) = \sup_{\mu \in \mathfrak{A}} \mathbb{E}_\mu[Z],$$

let us define the mapping $\rho_{\mathscr{F}_1} : \mathcal{Z}_2 \to \mathcal{Z}_1$ as

$$\rho_{\mathscr{F}_1}(Z) = \sup_{\mu \in \mathfrak{A}} \mathbb{E}_\mu[Z \mid \mathscr{F}_1].$$

Show that $\rho_{\mathscr{F}_1}$ is a conditional risk mapping.

3–15. Coherent risk measures from conditional risk mappings (3). Let $\mathcal{Z}_i = \mathcal{L}_p(\Omega, \mathscr{F}_i, \mathrm{P})$ with $\mathscr{F}_1 \subseteq \mathscr{F}_2$. Given a conditional risk mapping $\rho : \mathcal{Z}_2 \to \mathcal{Z}_1$ and a set $A \in \mathscr{F}_1$ with $\mathrm{P}(A) > 0$ define the mapping ρ_A

$$\rho_A(Z) = \langle 1_A, \rho(Z) \rangle / \mathrm{P}(A).$$

Show that ρ_A is a coherent risk measure. Extend the above result by showing that for a $\mathscr{F}_1$-measurable random variable $\zeta \in \mathcal{Z}_1^*$ with $\int_\Omega \zeta \mathrm{dP} \neq 0$ and $\zeta \ge 0$ a.s., the mapping

$$\rho_\zeta(Z) = \frac{\int_\Omega \zeta \rho(Z) \mathrm{dP}}{\int_\Omega \zeta \mathrm{dP}}$$

is a coherent risk measure.

3–16. Interchangeability property (5). Let $(\Omega, \mathscr{F}, \mathrm{P})$ be a finite probability space with $\Omega = \{1, \ldots, n\}$, $\mathscr{F} = 2^\Omega$. Random variables on this space can be identified by vectors $X \in \mathbb{R}^n$. Let $\rho : \mathbb{R}^n \to \mathbb{R}$ be a convex and continuous risk measure. Let $F : \mathbb{R}^m \to \mathbb{R}^n$ be a mapping. Let $\mathcal{X} \subseteq \mathbb{R}^m$ and $\inf_{x \in \mathcal{X}} F(x)$ be interpreted coordinate-wise[4]. Show that

$$\rho\left(\inf_{x \in \mathcal{X}} F(x)\right) = \inf_{x \in \mathcal{X}} \rho(F(x)).$$

Additionally, show that

$$\operatorname*{arg\,min}_{x \in \mathcal{X}} F(x) \subseteq \operatorname*{arg\,min}_{x \in \mathcal{X}} \rho(F(x)).$$

Under what conditions on ρ is $\operatorname{arg\,min}_{x \in \mathcal{X}} F(x) = \operatorname{arg\,min}_{x \in \mathcal{X}} \rho(F(x))$?

3–17. Separable expected conditionals (5). Take the mapping $\varrho : \mathcal{Z}_0 \times \mathcal{Z}_1 \times \ldots \times \mathcal{Z}_N \to \mathbb{R}$ defined as

$$\varrho[Z_0, Z_1, \ldots, Z_{N-1}] = Z_0 + \rho(Z_1) + \sum_{k=2}^{N} \mathbb{E}[\rho(Z_k \mid \mathscr{F}_{k-1})],$$

where $\rho(\cdot \mid \mathscr{F}_{k-1}) : \mathcal{Z}_k \to \mathcal{Z}_{k-1}$, for $k = 2, \ldots, N$, is a conditional risk mapping (derived by some coherent risk measure ρ). Show that (i) this is a multiperiod risk measure and (ii) there is a coherent risk measure $\bar{\rho} : \mathcal{Z}_{N-1} \to \mathbb{R}$ so that $\varrho(Z_0, \ldots, Z_{N-1}) = \bar{\rho}(Z_0 + \ldots + Z_{N-1})$.

3–18. Different random variables, same distribution (5). Let $\Omega = \{0, 1\}$ be a finite sample space with two events. Give an example of two random variables X and Y on a probability space $(\Omega, 2^\Omega, \mathrm{P})$ — for some P — which are equal in distribution, but never equal (that is $X(\omega) \neq Y(\omega)$).

Let $\Omega = \mathbb{R}$ equipped with the Borel σ-algebra. Give an example of two random variables X and Y on $(\Omega, \mathscr{B}_\mathbb{R}, \mathrm{P})$ — for some probability measure P — which are equal in distribution, but almost surely not equal (that is $X(\omega) \neq Y(\omega)$ P-a.s.).

[4] That is, if $F(x) = (F_1(x), \ldots, F_n(x))$, $\inf_{x \in \mathcal{X}} F(x) := (\inf_{x \in \mathcal{X}} F_1(x), \ldots, \inf_{x \in \mathcal{X}} F_n(x))$

3–19. Information monotonicity of multiperiod risk measures (5). We say that a mutliperiod risk measure $\rho : \mathcal{Z}_0, \ldots, \mathcal{Z}_N \to \mathbb{R}$ is *information-monotone* if for two filtrations $\mathfrak{F} = \{\mathscr{F}_k\}_{k=0,\ldots,N}$ and $\mathfrak{G} = \{\mathscr{G}_k\}_{k=0,\ldots,N}$ with $\mathscr{F}_k \subseteq \mathscr{G}_k$ for all $k = 0, \ldots, N$, it is[5]

$$\rho(Z_0, \ldots, Z_N; \mathfrak{F}) \geq \rho(Z_0, \ldots, Z_N; \mathfrak{G}).$$

This means that more information (that is, $\mathfrak{G}$) cannot increase the total multiperiod risk[6].

Consider the following nested multiperiod risk functional

$$\rho(Z_0, Z_1, \ldots, Z_N) = Z_0 + \rho_1(Z_1 + \ldots + \rho_{N-1}(Z_{N-1} + \rho_N(Z_N))),$$

where Z_k is $\mathscr{F}_k$-measurable and ρ_k are conditional risk mappings.

Construct a simple example to show that such problems **may not be information-monotone** in the sense of the definition given above[7].

[5] If $Z_0, \ldots, Z_N$ is $\mathfrak{F}$-adapted (i.e., Z_k are $\mathscr{F}_k$-measurable), then it is also $\mathfrak{G}$-adapted. The converse is not true. Here we assume that $Z_0, \ldots, Z_N$ is $\mathfrak{F}$-adapted.

[6] Recall that we proved in class that for law invariant, lower semicontinuous and coherent risk measures, information does not increase risk (that is $\rho(\mathbb{E}[Z \mid \mathscr{G}]) \leq \rho(Z)$ for any σ-algebra $\mathscr{G}$).

[7] Hints: Assume that ρ_k have some simple form, e.g., they are conditional average value-at-risk mappings. Take $N = 2$ for simplicity. Construct a simple scenario tree which corresponds to a filtration $\mathfrak{F}$ (with a bew branches at each stage). Introduce a finer filtration $\mathfrak{G}$ (increase the branching). Evaluate $\rho(Z_0, \ldots, Z_N; \mathfrak{F})$ and $\rho(Z_0, \ldots, Z_N; \mathfrak{G})$.