Risk neutral, distributionally robust and risk averse multistage stochastic programming

A. Shapiro
School of Industrial and Systems Engineering,
Georgia Institute of Technology,
Atlanta, Georgia 30332-0205, USA

Grid Science Winter School
Center for Nonlinear Studies (CNLS)
Santa Fe, 2019

Consider optimization problem

$$\underset{x\in\mathcal{X}}{\text{Min}}\Big\{f(x)=\mathbb{E}[F(x,\xi)]\Big\},$$

where $\mathcal{X}\subset\mathbb{R}^n$, $F:\mathbb{R}^n\times\mathbb{R}^d\to\mathbb{R}$ and ξ is a d-dimensional random vector. In case of two-stage linear stochastic programming with recourse, $\mathcal{X}=\{x\in\mathbb{R}^n_+ : Ax=b\}$ and $F(x,\xi)$ is the first stage cost $c^\top x$ plus the optimal value of the second stage problem

$$\underset{y\in\mathbb{R}^m}{\text{Min}}\, q^\top y \;\; subject\, to\, Tx+Wy=h,\; y\geq 0,$$

with ξ formed from random components of q,T,W,h.

For fixed $x\in\mathcal{X}$ the expectation $\mathbb{E}[F(x,\xi)]$ is given by the integral

$$\mathbb{E}[F(x,\xi)]=\int F(x,z)dP(z),$$

where P is the probability distribution of ξ.

A standard approach to solving such stochastic programs is to discretize distribution P, i.e., to construct scenarios ξ_k, $k = 1, ..., K$, with assigned probabilities $p_k > 0$, and hence to approximate $\mathbb{E}[F(x, \xi)]$ by $\sum_{k=1}^{K} p_k F(x, \xi_k)$. In the two-stage linear case this leads to the linear program

$$\begin{array}{ll} \underset{x, y_1, ..., y_K}{\text{Min}} & c^\top x + \sum_{k=1}^{K} p_k q_k^\top y_k \\ s.t. & T_k x + W_k y_k = h_k,\ k = 1, ..., K, \\ & Ax = b,\ x \geq 0,\ y_k \geq 0,\ k = 1, ..., K. \end{array}$$

In order to have an accurate approximation of the 'true' distribution P the number K of required scenarios typically growths exponentially with dimension d of the vector of random parameters.

Even crude discretization of the distribution of the random vector ξ leads to an exponential growth of the number of scenarios with increase of its dimension d.

Could stochastic programming problems be solved numerically?

What does it mean to solve a stochastic program?

How do we know the probability distribution of the random data vector?

Why do we optimize the expected value of the objective (cost) function?

Computational complexity of solving two-stage linear stochastic programs (deterministic point of view): *the approximate solutions, with a sufficiently high accuracy, of linear two-stage stochastic programs with fixed recourse are $\#P$-hard even if the random problem data is governed by independent uniform distributions* (Dyer and Stougie, 2006, Hanasusanto, Kuhn and Wiesemann, 2016).

Sample complexity of solving stochastic programs

Generate a sample ξ^j, $j = 1, ..., N$, of random vector ξ and approximate the expectation $\mathbb{E}[F(x, \xi)]$ by the respective sample average. This leads to the following so-called Sample Average Approximation (SAA) of the 'true' problem

$$\operatorname*{Min}_{x \in X} \left\{ \widehat{f}_N(x) = \frac{1}{N} \sum_{j=1}^{N} F(x, \xi^j) \right\}.$$

Slow convergence of the sample average $\widehat{f}_N(x)$ to the expectation $f(x)$. By the Central Limit Theorem, for fixed x the error

$$\widehat{f}_N(x) - f(x) = O_p(N^{-1/2}).$$

Let $\widehat{v}_N$ be the optimal value of the SAA problem and v^0 and $\mathcal{S}^0$ be the optimal value and set of optimal solutions of the true problem. Then under mild regularity conditions

$$\widehat{v}_N = \min_{x \in \mathcal{S}^0} \widehat{f}_N(x) + o_p(N^{-1/2}).$$

In particular, if $\mathcal{S}^0 = \{x^0\}$, then

$$N^{1/2}[\widehat{v}_N - v^0] \Rightarrow N(0, \sigma^2(x^0))$$

(Shapiro, 1991).

Large Deviations type bounds. Suppose that: $\varepsilon > \delta \geq 0$, the set X is of finite diameter D, there is a constant $\sigma > 0$ such that

$$M_{x',x}(t) \leq \exp\{\sigma^2 t^2/2\}, \quad t \in \mathbb{R}, \ x', x \in X,$$

where $M_{x',x}(t)$ is the moment generating function of the random variable $F(x',\xi) - F(x,\xi) - \mathbb{E}[F(x',\xi) - F(x,\xi)]$, there exists $\kappa(\xi)$ such that its moment generating function is finite valued in a neighborhood of zero and

$$\left|F(x',\xi) - F(x,\xi)\right| \leq \kappa(\xi)\|x' - x\|, \quad x', x \in X \text{ and } a.e.\ \xi.$$

Then for $L = \mathbb{E}[\kappa(\xi)]$ and sample size

$$N \geq \frac{8\sigma^2}{(\varepsilon - \delta)^2}\left[n \log\left(\frac{O(1)DL}{(\varepsilon - \delta)^2}\right) + \log\left(\frac{2}{\alpha}\right)\right],$$

we are guaranteed that $\Pr\left(\widehat{\mathcal{S}}_N^{\delta} \subset \mathcal{S}^{\varepsilon}\right) \geq 1 - \alpha$. Here $\widehat{\mathcal{S}}_N^{\delta}$ and $\mathcal{S}^{\varepsilon}$ are the sets of δ-optimal and ε-optimal solutions of the SAA and true problems respectively.

Distributionally robust approach to stochastic programming

$$\operatorname*{Min}_{x\in\mathcal{X}} \left\{ g(x) := \sup_{Q\in\mathfrak{M}} \mathbb{E}_Q[F_x(\omega)] \right\}, \tag{1}$$

where $\mathcal{X} \subset \mathbb{R}^n$, $F_x(\omega) = F(x, \xi(\omega))$, $F : \mathbb{R}^n \times \Xi \to \mathbb{R}$, $\xi : \Omega \to \Xi$ is a measurable mapping from Ω into $\Xi \subset \mathbb{R}^d$ and $\mathfrak{M}$ is a (nonempty) set of probability measures (distributions) on the sample space $(\Omega, \mathcal{F})$.

Let $\mathcal{Z}$ be a linear space of measurable functions $Z : \Omega \to \mathbb{R}$. We assume that $F_x \in \mathcal{Z}$ for all $x \in \mathcal{X}$. Consider

$$\rho(Z) := \sup_{Q\in\mathfrak{M}} \mathbb{E}_Q[Z(\omega)] = \sup_{Q\in\mathfrak{M}} \int_\Omega Z(\omega) dQ(\omega), \quad Z \in \mathcal{Z}.$$

The functional $\rho : \mathcal{Z} \to \overline{\mathbb{R}}$ has the following properties:

(A1) Convexity:

$$\rho(\alpha Z_1 + (1-\alpha)Z_2) \leq \alpha\rho(Z_1) + (1-\alpha)\rho(Z_2)$$

for all $Z_1, Z_2 \in \mathcal{Z}$ and $\alpha \in [0,1]$.
(A2) Monotonicity: If $Z_1, Z_2 \in \mathcal{Z}$ and $Z_2 \geq Z_1$, then $\rho(Z_2) \geq \rho(Z_1)$.
(A3) Translation Equivariance: If $a \in \mathbb{R}$ and $Z \in \mathcal{Z}$, then $\rho(Z+a) = \rho(Z) + a$.
(A4) Positive Homogeneity:

$$\rho(\alpha Z) = \alpha\rho(Z), \quad Z \in \mathcal{Z},\ \alpha > 0.$$

Functionals $\rho : \mathcal{Z} \to \overline{\mathbb{R}}$ satisfying axioms (A1)–(A4) are called *coherent risk measures* (Artzner, Delbaen, Eber, Heath (1999)).

Examples

Space $\mathcal{Z} := L_p(\Omega, \mathcal{F}, P)$, where P is a (reference) probability measure on $(\Omega, \mathcal{F})$ and $p \in [1, \infty)$. That is, $\mathcal{Z}$ is the space of random variables $Z(\omega)$ having finite p-th order moment.

For $\zeta = dQ/dP$, space $\mathcal{Z}$ is paired with its dual space $\mathcal{Z}^* = L_q(\Omega, \mathcal{F}, P)$, where $1/p + 1/q = 1$, and the scalar product

$$\langle \zeta, Z \rangle := \int_\Omega \zeta(\omega) Z(\omega) dP(\omega), \ \ \zeta \in \mathcal{Z}^*, \ Z \in \mathcal{Z}.$$

We also consider space $\mathcal{Z} := L_\infty(\Omega, \mathcal{F}, P)$, of essentially bounded (measurable) functions $Z : \Omega \to \mathbb{R}$, paired with space $L_1(\Omega, \mathcal{F}, P)$.

Another example. Let Ω be a metric space equipped with its Borel sigma algebra $\mathcal{F}$, and $\mathcal{Z} := C(\Omega)$ be the space of continuous functions $Z : \Omega \to \mathbb{R}$ with the max-norm $\|Z\| = \sup_{\omega\in\Omega} |Z(\omega)|$. Its dual space $\mathcal{Z}^*$ is the space of finite signed measures on $(\Omega, \mathcal{F})$ with the scalar product

$$\langle \mu, Z \rangle := \int_\Omega Z(\omega) d\mu(\omega), \;\; \mu \in \mathcal{Z}^*, \; Z \in \mathcal{Z}.$$

This framework is suitable when the uncertainty set $\mathfrak{M}$ is defined by moment constraints.

We mainly consider the first example with the reference probability space $(\Omega, \mathcal{F}, P)$ and paired spaces $\mathcal{Z} = L_p(\Omega, \mathcal{F}, P)$ and $\mathcal{Z}^* = L_q(\Omega, \mathcal{F}, P)$.

In case the functional ρ is law invariant, it can be considered as a function of the cdf $F_Z(z) = P(Z \leq z)$. Given a random sample $Z_1, ..., Z_N$ of the random variable Z, defined on $(\Omega, \mathcal{F}, P)$, we can approximate F_Z by the empirical cdf

$$\widehat{F}_N(z) := \frac{1}{N} \sum_{i=1}^{N} 1_{(-\infty, z]}(Z_i).$$

Consequently we can approximate $\rho(Z) = \rho(F_Z)$ by $\rho(\widehat{F}_N)$.

Suppose now that $\xi_1, ..., \xi_N$ is a sample of the random vector $\xi = \xi(\omega)$. Then we can estimate distributionally robust problem (1) by the SAA problem:

$$\operatorname*{Min}_{x \in \mathcal{X}} \rho(\widehat{F}_{x,N}). \tag{2}$$

Examples

Example 1 Consider $\mathcal{Z} := L_1(\Omega, \mathcal{F}, P)$ and

$$\mathfrak{A} := \left\{ \zeta : 1 - \beta_1 \leq \zeta(\omega) \leq 1 + \beta_2,\ \omega \in \Omega,\ \int_\Omega \zeta dP = 1 \right\},$$

where $\beta_1 \in (0, 1]$ and $\beta_2 \geq 0$. Clearly the set $\mathfrak{A}$ is invariant with respect to measure preserving transformations. The corresponding functional ρ is

$$\rho(Z) = (1 - \beta_1)\mathbb{E}_P[Z] + \beta_1 \mathsf{AV@R}_\alpha(Z),$$

where $\alpha = \beta_1/(\beta_1 + \beta_2)$ and

$$\mathsf{AV@R}_\alpha(Z) = \frac{1}{\alpha} \int_{1-\alpha}^1 F_Z^{-1}(t) dt = \inf_{t \in \mathbb{R}} \left\{ t + \alpha^{-1} \mathbb{E}_P[Z - t]_+ \right\}.$$

Example 2 (ϕ-divergence) Consider a convex continuous function $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\phi(1) = 0$ and $\phi(x) > 0$ for $x > 1$. Let

$$\mathfrak{A} := \left\{ \zeta \geq 0 : \int_\Omega \phi(\zeta(\omega)) dP(\omega) \leq c, \int_\Omega \zeta(\omega) dP(\omega) = 1 \right\}$$

for some $c > 0$. For example we can take $\phi(x) := |x - 1|^p$, $p \in [1, \infty)$. In that case it will be natural to use the space $\mathcal{Z} = L_p(\Omega, \mathcal{F}, P)$ and

$$\mathfrak{A} = \left\{ \zeta \geq 0 : \|\zeta - 1\|_p \leq c^{1/p}, \int_\Omega \zeta dP = 1 \right\}.$$

For $\phi(x) := x \ln x - x + 1$ we have that $\int_\Omega \phi(\zeta(\omega)) dP(\omega)$ defines the Kullback-leibler divergence, denoted $D_{KL}(\zeta \| P)$.

It is possible to show that in case of ϕ-divergence the corresponding functional can be written in the form

$$\rho(Z) = \inf_{\lambda \geq 0, \mu} \Big\{ \lambda c + \mu + \mathbb{E}_P[(\lambda\phi)^*(Z - \mu)] \Big\},$$

where $(\lambda\phi)^*$ is the conjugate function of $\lambda\phi$.

In particular for the Kullback-Leibler divergence,

$$\rho(Z) = \inf_{\lambda > 0} \Big\{ \lambda c + \lambda \mathbb{E}_P[e^{Z/\lambda}] \Big\}.$$

Multistage stochastic programming

Consider a multistage decision process of the from

$$\begin{array}{c}\text{decision } (x_1) \rightsquigarrow \text{observation } (\xi_2) \rightsquigarrow \text{decision } (x_2) \rightsquigarrow \\ \ldots\ldots \rightsquigarrow \text{observation } (\xi_T) \rightsquigarrow \text{decision } (x_T).\end{array} \tag{3}$$

Here $\xi_t \in \mathbb{R}^{d_t}$, $t = 1, ...,$ is a sequence of vectors with $\xi_{[t]} := (\xi_1, ..., \xi_t)$ representing history of the data process up to time t. At time period $t \in \{1, ..., T\}$ we observe the past, $\xi_{[t]}$, but future observations $\xi_{t+1}, ...,$ are uncertain. So our decision at time t should only depend on information available at that time, i.e., $x_t = x_t(\xi_{[t]})$ should be a function of $\xi_{[t]}$ and should not depend on future observations. This is the basic requirement of *nonanticipativity* of the decision process. A sequence $x_1, x_2(\xi_{[2]}), ...$ of such decisions is called a *policy* or a decision rule.

Risk neutral multistage stochastic programming

$$\operatorname*{Min}_{\pi\in\Pi} \ \mathbb{E}\Big[F_1(x_1) + F_2(x_2(\xi_{[2]}),\xi_2) + \cdots + F_T\left(x_T(\xi_{[T]}),\xi_T\right)\Big]$$

where Π is the set of policies $\pi = (x_1, x_2(\xi_{[2]}), ..., x_T(\xi_{[T]})$ satisfying the feasibility constraints

$$x_1 \in \mathcal{X}_1,\ x_t(\xi_{[t]}) \in \mathcal{X}_t(x_{t-1}(\xi_{[t-1]}),\xi_t),\ t = 2,\ldots,T,$$

for almost every (a.e.) realization of the random process, $F_t : \mathbb{R}^{n_t} \times \mathbb{R}^{d_t} \to \mathbb{R}$ are real valued functions and $\mathcal{X}_t : \mathbb{R}^{n_{t-1}} \times \mathbb{R}^{d_t} \rightrightarrows \mathbb{R}^{n_t}$, $t = 2,\ldots,T$, are multifunctions. For example

$$F_t(x_t,\xi_t) := c_t^{\mathsf{T}} x_t,$$

$$\mathcal{X}_t(x_{t-1},\xi_t) := \{x_{t-1} : B_t x_{t-1} + A_t x_t \le b_t\},$$

$t = 2, ..., T$, $\mathcal{X}_1 := \{x_1 : A_1 x_1 \le b_1\}$, with $\xi_t = (c_t, A_t, B_t, b_t)$, corresponds to linear multistage stochastic programming.

Note that it is assumed here that the probability distribution of the random process ξ_t does not depend on our decisions. Note also that

$$\mathbb{E}[Z] = \mathbb{E}_{|\xi_1}\Big[\mathbb{E}_{|\xi_{[2]}}\big[\cdots \mathbb{E}_{|\xi_{[T-1]}}[Z]\big]\Big].$$

This decomposition property of the expectation allows to write the multistage stochastic programming problem in the following nested form

$$\begin{aligned}\min_{x_1\in\mathcal{X}_1} \; & F_1(x_1) + \mathbb{E}_{|\xi_1}\Big[\inf_{x_2\in\mathcal{X}_2(x_1,\xi_2)} F_2(x_2,\xi_2) + \ldots \\ & + \mathbb{E}_{|\xi_{[T-2]}}\big[\inf_{x_{T-1}\in\mathcal{X}_T(x_{T-2},\xi_{T-1})} F_{T-1}(x_{T-1},\xi_{T-1}) \\ & + \mathbb{E}_{|\xi_{[T-1]}}[\inf_{x_T\in\mathcal{X}_T(x_{T-1},\xi_T)} F_T(x_T,\xi_T)]\big]\Big].\end{aligned}$$

This formulation assumes that: (i) the probability distribution of the data process is known (specified), (ii) the optimization is performed *on average* (both, with respect to different realizations of the random process, and with respect to time).

Numerical difficulties in solving multistage problems.

From a modeling point of view typically it is natural to assume that the random data process has a *continuous* distribution. This raises the question of how to compute the involved expectations (multivariate integrals). A standard approach is to discretize the random process by generating a finite number of possible realizations (called scenarios). These scenarios can be represented by the corresponding *scenario tree*.

How many scenarios are needed in order to approximate the "true" distribution of the random data process?

Note that solving the deterministic equivalent for the constructed scenario tree does not produce by itself an implementable policy for the "true" problem (with continuous distributions). This is because an actual realization of the data process could, and with probability one (w.p.1) will, be different from scenarios used in the constructed tree. In that case policy constructed for scenarios of the tree does not tell what decision to make. Of course, one can use only the first stage solution which is deterministic (does not depend on future observations) and update it as new observations become available - this is a rolling horizon approach. Such a rolling horizon approach requires resolving the corresponding multistage problem at every stage as new realization of the data becomes available.

Suppose that the data process is stagewise independent, i.e., random vector ξ_{t+1} is independent of $\xi_{[t]}$, $t = 1, ..., T-1$.

Discretization by Monte Carlo sampling Independent of each other random samples $\xi_t^j = (c_t^j, B_t^j, A_t^j, b_t^j)$, $j = 1, ..., N_t$, of respective ξ_t, $t = 2, ..., T$, are generated and the corresponding scenario tree is constructed by connecting every ancestor node at stage $t-1$ with the same set of children nodes $\xi_t^1, ..., \xi_t^{N_t}$. In that way the stagewise independence is preserved in the generated scenario tree. We refer to the constructed problem as the Sample Average Approximation (SAA) problem.

The total number of scenarios of the SAA problem is given by the product $\mathcal{N} = \prod_{t=2}^{T} N_t$ and quickly becomes astronomically large with increase of the number of stages even for moderate values of sample sizes N_t.

If we measure computational complexity, of the "true" problem, in terms of the number of scenarios required to approximate true distribution of the random data process with a reasonable accuracy, the conclusion is rather pessimistic. In order for the optimal value and solutions of the SAA problem to converge to their true counterparts all sample sizes $N_2, ..., N_T$ should tend to infinity. Furthermore, available estimates of the sample sizes required for a first stage solution of the SAA problem to be ε-optimal for the true problem, with a given confidence (probability), sums up to a number of scenarios which grows as $O(\varepsilon^{-2(T-1)})$ with decrease of the error level $\varepsilon > 0$. This indicates that from the point of view of the number of scenarios, complexity of multistage programming problems grows exponentially with increase of the number of stages.

Because of the exponential growth of the number of scenarios $\mathcal{N}$ it is hopeless to try to solve multistage stochastic programs by enumerating all scenarios.

An alternative approach is suggested by the dynamic programming.

"Any model is wrong but some are useful".

From a modeling point of view it is natural to assume that the random data process has a *continuous* distribution. We refer to such model as "true" or "continuous".

Dynamic Programming Equations.

For the last period T we have

$$Q_T(x_{T-1}, \xi_T) := \inf_{x_T \in \mathcal{X}_T(x_{T-1}, \xi_T)} F_T(x_T, \xi_T),$$

$$\mathcal{Q}_T(x_{T-1}, \xi_{[T-1]}) := \mathbb{E}_{|\xi_{[T-1]}}[Q_T(x_{T-1}, \xi_T)],$$

and for $t = T-1, \ldots, 2$,

$$Q_t\left(x_{t-1}, \xi_{[t]}\right) = \inf_{x_t \in \mathcal{X}_t(x_{t-1}, \xi_t)} \left\{F_t(x_t, \xi_t) + \mathcal{Q}_{t+1}\left(x_t, \xi_{[t]}\right)\right\},$$

where

$$\mathcal{Q}_{t+1}\left(x_t, \xi_{[t]}\right) := \mathbb{E}_{|\xi_{[t]}}\left\{Q_{t+1}\left(x_t, \xi_{[t+1]}\right)\right\}.$$

Finally, at the first stage we solve the problem

$$\mathrm{Min}_{x_1 \in \mathcal{X}_1} F_1(x_1) + \mathbb{E}[Q_2(x_1, \xi_2)].$$

In case of stagewise independence, by induction in $t = T, ...$, it is possible to show that cost-to-go functions $\mathcal{Q}_t(x_{t-1})$ do not depend on the data process. The dynamic programming equations take the form

$$Q_t\left(x_{t-1}, \xi_t\right) = \inf_{x_t \in \mathcal{X}_t(x_{t-1}, \xi_t)} \left\{ F_t(x_t, \xi_t) + \mathcal{Q}_{t+1}\left(x_t\right) \right\},$$

where

$$\mathcal{Q}_{t+1}\left(x_t\right) = \mathbb{E}\left[Q_{t+1}\left(x_t, \xi_{t+1}\right) \right].$$

In some cases stagewise dependent problems can be reformulated in a stagewise independent form at the price of increasing number of state variables. For example, suppose that only the right hand side vectors b_t are random and can be modeled as a (first order) autoregressive process

$$b_t = \mu + \Phi b_{t-1} + \varepsilon_t,$$

where μ and Φ are (deterministic) vector and regression matrix, respectively, and the error process ε_t, $t = 1, ..., T$, is stagewise independent. The corresponding feasibility constraints can be written in terms of x_t and b_t as

$$B_t x_{t-1} + A_t x_t \leq b_t, \;\; \Phi b_{t-1} - b_t + \mu + \varepsilon_t = 0.$$

That is, in terms of decision variables (x_t, b_t) this becomes a linear multistage stochastic programming problem governed by the stagewise independent random process $\varepsilon_1, ..., \varepsilon_T$.

Curse of dimensionality

One of the main difficulties in solving the dynamic programming equations is how to represent the cost-to-go functions in a computationally feasible way.

For dimension of x_t say greater than 3 and large number of stages it is practically impossible to solve the dynamic programming equations with high accuracy. Several alternatives were suggested in recent literature.

Risk averse and distributionally robust multistage programming.

Average Value-at-Risk (also called *Conditional Value-at-Risk*)

$$\mathsf{AV@R}_\alpha(Z) := \inf_{t\in\mathbb{R}} \left\{ t + \alpha^{-1}\mathbb{E}[Z-t]_+ \right\}$$

Note that the minimum in the above is attained at $t^* = F_Z^{-1}(1-\alpha)$, where $F_Z(t) := P(Z \le t)$ is the cdf of Z and $F_Z^{-1}(1-\alpha) = \mathsf{V@R}_\alpha(Z) = \inf\{t : F_Z(t) \ge 1-\alpha\}$.

Also

$$\mathsf{AV@R}_\alpha(Z) = \frac{1}{\alpha}\int_{1-\alpha}^1 F_Z^{-1}(t)dt.$$

If $F_Z(z)$ is continuous at $z = F_Z^{-1}(1-\alpha)$, then

$$\mathsf{AV@R}_\alpha(Z) = \mathbb{E}\Big[Z | Z \ge F_Z^{-1}(1-\alpha)\Big].$$

With every law invariant risk measure $\rho(Z)$ we can associate the respective *conditional* risk measure, denoted $\rho(Z|Y)$ or $\rho_{|Y}(Z)$, conditional on random variable Y, by employing conditional distribution of Z given Y.

Note that $\rho(Z|Y)$ is a function of Y and we can consider the composite risk measure $\rho(\rho(Z|Y))$. For example, for $\rho(Z) := \mathbb{E}[Z]$ and $\rho(Z|Y) = \mathbb{E}[Z|Y]$ we have $\mathbb{E}[\mathbb{E}[Z|Y]] = \mathbb{E}[Z]$.

Conditional version of the Average Value-at-Risk:

$$\mathsf{AV@R}_{\alpha|Y}(Z) = \inf_{t\in\mathbb{R}} \mathbb{E}_{|Y}\left\{t + \alpha^{-1}[Z-t]_+\right\}.$$

The minimum in the above is attained at

$t^* = \Big\{$ $(1-\alpha)$-quantile of the conditional distribution of Z given $Y\Big\}$.

Of course, t^* is a function of Y here.

Risk averse multistage programming.

Nested formulation of risk averse multistage programming problem:

$$\begin{aligned}\min_{x_1\in\mathcal{X}_1} \; & F_1(x_1)+\rho_{|\xi_1}\Big[\inf_{x_2\in\mathcal{X}_2(x_1,\xi_2)} F_2(x_2,\xi_2)+\ldots \\ & +\rho_{|\xi_{[T-2]}}\Big[\inf_{x_{T-1}\in\mathcal{X}_T(x_{T-2},\xi_{T-1})} F_{T-1}(x_{T-1},\xi_{T-1}) \\ & +\rho_{|\xi_{[T-1]}}[\inf_{x_T\in\mathcal{X}_T(x_{T-1},\xi_T)} F_T(x_T,\xi_T)]\Big]\Big],\end{aligned}$$

where $\rho_{|\xi_{[t]}}(\cdot)$, $t=1,...,T-1$, are conditional law invariant coherent (convex) risk measures. For example

$$\rho_{|\xi_{[t]}}(\cdot):=\lambda\mathbb{E}_{|\xi_{[t]}}[\cdot]+(1-\lambda)\mathsf{AV@R}_{\alpha|\xi_{[t]}}[\cdot]$$

is a convex combination of the conditional expectation and conditional Average Value-at-risk measure.

We can write the risk averse multistage programming problem as

$$\begin{array}{cl} \underset{x_1, x_2(\cdot), \ldots, x_T(\cdot)}{\text{Min}} & \bar{\rho}\Big[F_1(x_1) + F_2(x_2(\xi_{[2]}), \xi_2) + \cdots + F_T\left(x_T(\xi_{[T]}), \xi_T\right)\Big] \\ \text{s.t.} & x_1 \in \mathcal{X}_1,\ x_t(\xi_{[t]}) \in \mathcal{X}_t(x_{t-1}(\xi_{[t-1]}), \xi_t),\ t = 2, \ldots, T, \end{array}$$

where

$$\begin{aligned} \bar{\rho}(Z_1 + \ldots + Z_T) &= \rho_{|\xi_1}\Big(\rho_{|\xi_{[2]}}\Big(\cdots \rho_{|\xi_{[T-1]}}(Z_1 + \ldots + Z_T)\Big)\Big) \\ &= Z_1 + \rho_{|\xi_1}\Big(Z_2 + \rho_{|\xi_{[2]}}\Big(+ \cdots \rho_{|\xi_{[T-1]}}(Z_T)\Big)\Big) \end{aligned}$$

is the corresponding composite risk measure. The optimization is performed over (nonanticipative) policies $x_1, x_2(\xi_{[2]}), \ldots, x_T(\xi_{[T]})$ satisfying the feasibility constraints.

If $\rho_{|\xi_{[t]}}(\cdot) := \mathbb{E}_{|\xi_{[t]}}(\cdot)$ are conditional expectations, then $\bar{\rho}(\cdot) = \mathbb{E}(\cdot)$. In that case this becomes the risk neutral stochastic programming. If

$$\rho_{|\xi_{[t]}}(\cdot) := \operatorname{ess\,sup}(\cdot) = \mathsf{AV@R}_{0|\xi_{[t]}}(\cdot),$$

then $\bar{\rho}(\cdot) = \mathsf{AV@R}_0(\cdot)$. This case corresponds to multistage *robust* optimization.

Let ρ be a law invariant coherent risk measure. It turns out that only $\rho(\cdot) := \mathbb{E}(\cdot)$ and $\rho(\cdot) := \operatorname{ess\,sup}(\cdot)$ risk measures have the decomposition property

$$\rho(\rho_{|Y}(Z)) = \rho(Z),\ Z \in \mathcal{Z}.$$

Distributionally robust multistage stochastic programming.

We can write the risk neutral multistage problem as

$$\operatorname*{Min}_{\pi \in \Pi} \mathbb{E}_P[Z^\pi], \tag{4}$$

where P is the probability distribution of random vector $\xi_{[T]} = (\xi_1, ..., \xi_T)$, Π is a set of policies satisfying the feasibility constraints

$$x_1 \in \mathcal{X}_1,\ x_t(\xi_{[t]}) \in \mathcal{X}_t(x_{t-1}(\xi_{[t-1]}), \xi_t),\ t = 2, ..., T-1,$$

and $Z^\pi = Z^\pi(\xi_{[T]})$ is defined as

$$Z^\pi := F_1(x_1) + F_2(x_2(\xi_{[2]}), \xi_2) + \cdots + F_T(x_T(\xi_{[T]}), \xi_T).$$

It looks natural to formulate the following distributionally robust analogue of problem (4). Consider a set $\mathfrak{M}$ of probability distributions of $\xi_{[T]}$ supported on a set $\Xi \subset \mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_T}$ equipped with its Borel sigma algebra $\mathcal{B}$, and the problem

$$\operatorname*{Min}_{\pi \in \Pi} \sup_{Q \in \mathfrak{M}} \mathbb{E}_Q[Z^\pi]. \tag{5}$$

However, there is a problem with formulation (5).

The expectation operator has the following property (recall that ξ_1 is deterministic)

$$\mathbb{E}_Q[Z] = \mathbb{E}_Q\left[\mathbb{E}_{Q|\xi_{[2]}}\left[\cdots \mathbb{E}_{Q|\xi_{[T-1]}}[Z]\right]\right].$$

Therefore for $Z = Z(\xi_{[T]}) \in \mathcal{Z}$ and $Q \in \mathfrak{M}$ we have that

$$\sup_{Q \in \mathfrak{M}} \mathbb{E}_Q[Z] \leq \mathcal{R}(Z),$$

where

$$\mathcal{R}(Z) := \sup_{Q \in \mathfrak{M}} \mathbb{E}_Q \left[\sup_{Q \in \mathfrak{M}} \mathbb{E}_{Q|\xi_{[2]}} \left[\cdots \sup_{Q \in \mathfrak{M}} \mathbb{E}_{Q|\xi_{[T-1]}}[Z] \right] \right]. \quad (6)$$

The functional $\mathcal{R}(\cdot)$ satisfies the axioms of coherent risk measures and hence can be represented in the dual form

$$\mathcal{R}(Z) = \sup_{Q \in \widehat{\mathfrak{M}}} \mathbb{E}_Q[Z]$$

for some set $\widehat{\mathfrak{M}}$ of probability measures. Note that in general $\mathfrak{M} \neq \widehat{\mathfrak{M}}$ even in the rectangular case when

$$\mathfrak{M} = \{Q = Q_1 \times \cdots \times Q_T : Q_t \in \mathfrak{M}_t,\ t = 1, ..., T\}$$

where $\mathfrak{M}_t$ is a set of marginal distributions of ξ_t.

Dynamic Programming Equations.

For the last period T we have

$$Q_T(x_{T-1}, \xi_T) := \inf_{x_T \in \mathcal{X}_T(x_{T-1}, \xi_T)} F_T(x_T, \xi_T),$$

$$\mathcal{Q}_T(x_{T-1}, \xi_{[T-1]}) := \rho_{|\xi_{[T-1]}}[Q_T(x_{T-1}, \xi_T)],$$

and for $t = T-1, \ldots, 2$,

$$Q_t\left(x_{t-1}, \xi_{[t]}\right) = \inf_{x_t \in \mathcal{X}_t(x_{t-1}, \xi_t)} \left\{F_t(x_t, \xi_t) + \mathcal{Q}_{t+1}\left(x_t, \xi_{[t]}\right)\right\},$$

where

$$\mathcal{Q}_{t+1}\left(x_t, \xi_{[t]}\right) := \rho_{|\xi_{[t]}}\left\{Q_{t+1}\left(x_t, \xi_{[t+1]}\right)\right\}.$$

Finally, at the first stage we solve the problem

$$\mathrm{Min}_{x_1 \in \mathcal{X}_1} F_1(x_1) + \rho_{|\xi_1}[Q_2(x_1, \xi_2)].$$

In case of stagewise independence, the cost-to-go functions $\mathcal{Q}_t(x_{t-1})$ do not depend on the data process, and dynamic programming equations take the form

$$Q_t\left(x_{t-1}, \xi_t\right) = \inf_{x_t \in \mathcal{X}_t(x_{t-1}, \xi_t)} \left\{ F_t(x_t, \xi_t) + \mathcal{Q}_{t+1}\left(x_t\right) \right\},$$

$t = T, \ldots, 2$, where

$$\mathcal{Q}_{t+1}\left(x_t, \xi_{[t]}\right) := \rho\left\{ Q_{t+1}\left(x_t, \xi_{t+1}\right) \right\},$$

with $\mathcal{Q}_{T+1}(\cdot) \equiv 0$. Finally, at the first stage we solve the problem

$$\operatorname*{Min}_{x_1 \in \mathcal{X}_1} F_1(x_1) + \rho[Q_2(x_1, \xi_2)].$$

Time consistency

An optimal policy $\pi \in \Pi$ designed at the first stage should be optimal at every stage $t = 2, ..., T$ of the decision process conditional on an observed realization of the data process up to time t. In order to formalize this we need to give a precise definition of optimality at every stage $t = 2, ..., T$ conditional on a realization of the data process up to the considered time period.

For the nested formulation it is natural to use the respective *nested* conditional expectation criterion

$$\sup_{Q \in \mathfrak{M}} \mathbb{E}_{Q|\xi_{[t]}} \Big[\sup_{Q \in \mathfrak{M}} \mathbb{E}_{Q|\xi_{[t+1]}} \Big[\cdots \sup_{Q \in \mathfrak{M}} \mathbb{E}_{Q|\xi_{[T-1]}} [\, \cdot \,] \Big] \Big].$$

Note that this criterion is not the same as taking

$$\sup_{Q \in \mathfrak{M}} \mathbb{E}_{Q|\xi_{[t]}} [\, \cdot \,].$$

Interchangeability principle

Let $\mathcal{R} : \mathcal{Z} \to \mathbb{R}$ be a *monotone* functional. Consider

$$F(\omega) := \inf_{y \in Y} f(y, \omega), \tag{7}$$

where Y is an abstract set and $f : Y \times \Omega \to \mathbb{R} \cup \{+\infty\}$ is an extended real valued function. Let $\mathcal{Y}$ be the set of mappings $\eta : \Omega \to Y$ such that $f_\eta \in \mathcal{Z}$, where $f_\eta(\cdot) := f(\eta(\cdot), \cdot)$.

Suppose that the minimum in (7) is attained at $\bar{y}(\omega) \in Y$ for $\omega \in \Omega$, and hence $F(\omega) = f(\bar{y}(\omega), \omega)$. Then by monotonicity of $\mathcal{R}$, assuming that $F \in \mathcal{Z}$, we have that

$$\mathcal{R}(F) = \inf_{\eta \in \mathcal{Y}} \mathcal{R}(f_\eta).$$

That is the minimization operator and functional $\mathcal{R}$ can be interchanged. For monotone functionals this interchangeability holds in general (without assuming existence of minimizers). Moreover, the following implication holds

$$\bar{\eta}(\cdot) \in \arg\min_{y \in Y} f(y, \cdot) \Rightarrow \bar{\eta} \in \arg\min_{\eta \in \mathcal{Y}} \mathcal{R}(f_\eta).$$

It is possible to give simple examples showing that the converse implication

$$\bar{\eta} \in \arg\min_{\eta \in \mathcal{Y}} \mathcal{R}(f_\eta) \Rightarrow \bar{\eta}(\cdot) \in \arg\min_{y \in Y} f(y, \cdot)$$

may not hold, unless the functional $\mathcal{R}$ is *strictly* monotone.

Definition 1 *It is said that a functional* $\mathcal{R} : \mathcal{Z} \to \mathbb{R}$ *is strictly monotone, if* $Z \succeq Z'$ *and* $Z \neq Z'$ *imply that* $\mathcal{R}(Z) > \mathcal{R}(Z')$.

For example

$$\mathsf{AV@R}_\alpha(Z) = \frac{1}{1-\alpha} \int_\alpha^1 F_Z^{-1}(t) dt,$$

is not strictly monotone for $\alpha \in (0, 1)$.

Dynamic equations and time consistency

Consider two stage risk averse stochastic program

$$\min_{x\in\mathcal{X}_1, y(\cdot)\in\mathcal{X}_2(x,\cdot)} \mathcal{R}\Big(g(x, y(\omega), \omega)\Big), \tag{8}$$

where $\mathcal{X}_1 \subset \mathbb{R}^n$, $g : \mathbb{R}^n \times \mathbb{R}^k \times \Omega \to \mathbb{R}$ and $\mathcal{X}_2 : \mathbb{R}^n \times \Omega \rightrightarrows \mathbb{R}^k$ is a multifunction. An alternative formulation is

$$\min_{x\in\mathcal{X}_1} \mathcal{R}\Big(\underbrace{\min_{y\in\mathcal{X}_2(x,\omega)} g(x, y, \omega)}_{f(x,\omega)}\Big), \tag{9}$$

where $f_x(\omega) = f(x, \omega)$ is the optimal value of the second stage problem.

An optimal solution $(\bar{x}, \bar{y}(\cdot))$ of problem (8) is time consistent if $\bar{y}(\cdot)$ is an optimal solution of the second stage program, i.e.,

$$\bar{y}(\omega) \in \arg\min_{y\in\mathcal{X}_2(\bar{x},\omega)} g(\bar{x}, y, \omega),\ \omega \in \Omega.$$

For the two stage programs the above means that the optimal values of problems (8) and (9) are the same and if $\bar{x}$ is an optimal solution of the first stage problem and $\bar{y}(\omega)$, $\omega \in \Omega$, is an optimal solution of the second stage problem

$$\min_{y \in \mathcal{X}_2(\bar{x},\omega)} g(\bar{x}, y, \omega),$$

then $(\bar{x}, \bar{y}(\cdot))$ is an optimal solution of problem (8). The converse of that is true if $\mathcal{R}$ is *strictly* monotone.

If $\mathcal{R}$ is not strictly monotone, then the first stage problem (8) could have an optimal solution $(\bar{x}, \bar{y}(\cdot))$ such that conditional on $x = \bar{x}$, the solution $\bar{y}(\omega)$ is not optimal for the second stage problem for some $\omega \in \Omega$. That is for some $\omega \in \Omega$ the corresponding value $g(\bar{x}, \bar{y}(\omega), \omega)$ is strictly bigger than the minimal value

$$\min_{y \in \mathcal{X}_2(\bar{x}, \omega)} g(\bar{x}, y, \omega).$$

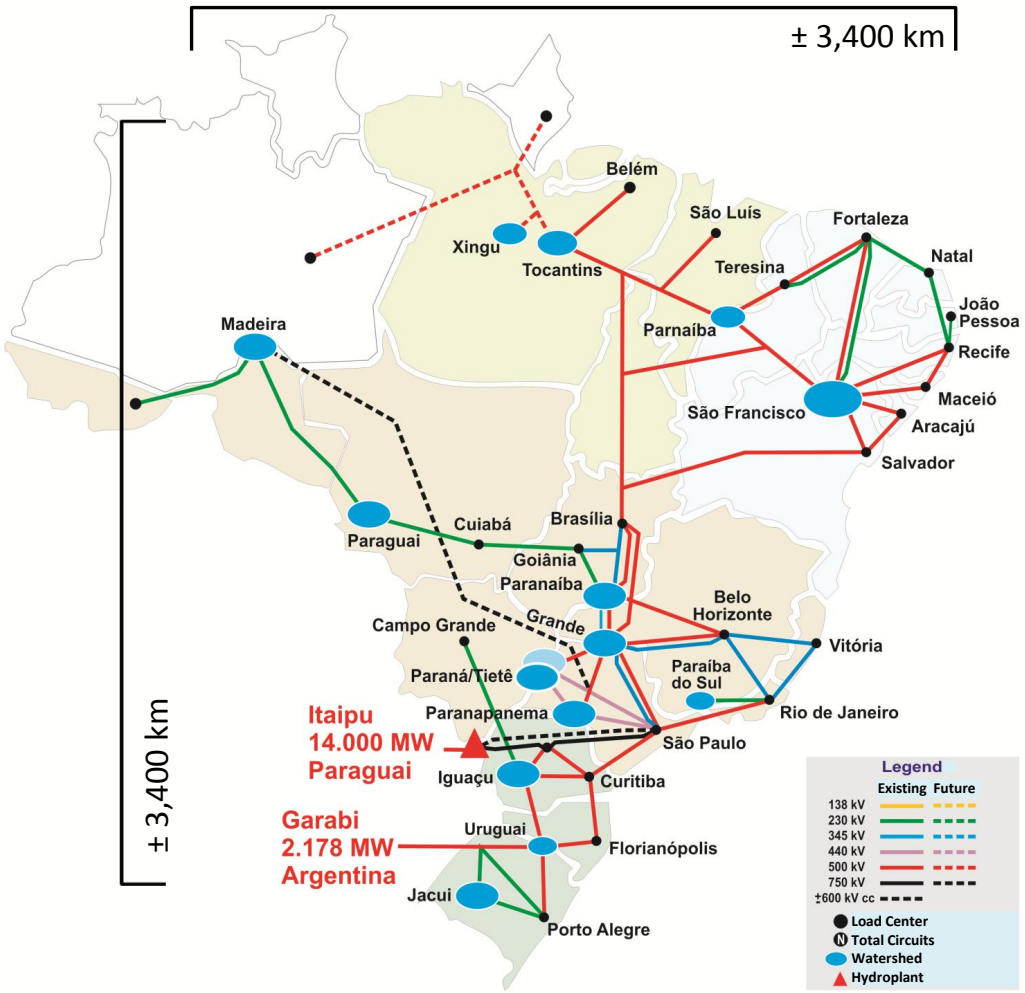
Such solutions are not time consistent. That is, without strict monotonicity time inconsistent optimal policies could exist already for two stage problems and finite number of scenarios.

The Brazilian hydro power operation planning problem

The Brazilian power system generation is hydro dominated (about 75% of the installed capacity) and characterized by large reservoirs presenting multi-year regulation capability, arranged in complex cascades over several river basins. The hydro plants use store water in the reservoirs to produce energy in the future, replacing fuel costs from the thermal units. Since the water inflows depend on rainfalls, the amount of future inflows is uncertain and cannot be predicted with a high accuracy.

The purpose of hydrothermal system operation planning is to define an operation strategy which, for each stage of the planning period, given the system state at the beginning of the stage, produces generation targets for each plant.

The Brazilian hydro power operation planning problem is a multistage, large scale (more than 200 power plants, of which 141 are hydro plants), stochastic optimization problem. On a high level, planning is for 5 years on monthly basis together with 5 additional years to smooth out the end of horizon effect. This results in 120-stage stochastic programming problem. Four energy equivalent reservoirs are considered, one in each one of the four interconnected main regions, SE, S, N and NE. The resulting policy obtained with the aggregate representation can be further refined, so as to provide decisions for each of the hydro and thermal power plants.


± 3,400 km
± 3,400 km
Belém
São Luís
Fortaleza
Xingu
Tocantins
Teresina
Natal
João Pessoa
Madeira
Parnaíba
Recife
Maceió
São Francisco
Aracajú
Salvador
Cuiabá
Brasília
Paraguai
Goiânia
Paranaíba
Belo Horizonte
Campo Grande
Grande
Vitória
Paraná/Tietê
Paraíba do Sul
Itaipu
14.000 MW
Paraguai
Paranapanema
Rio de Janeiro
São Paulo
Iguaçu
Curitiba
Garabi
2.178 MW
Argentina
Uruguai
Florianópolis
Jacui
Porto Alegre
Legend
Existing Future
138 kV
230 kV
345 kV
440 kV
500 kV
750 kV
±600 kV cc
Load Center
Total Circuits
Watershed
Hydroplant

Approximate dynamic programming

Basic idea is to approximate the cost-to-go functions by a class of computationally manageable functions. Since functions $\mathcal{Q}_t(\cdot)$ are convex it is natural to approximate these functions by piecewise linear functions given by maximum of cutting hyperplanes.

Stochastic Dual Dynamic Programming (SDDP) method (Pereira and Pinto, 1991).

For trial decisions $\bar{x}_t$, $t = 1, ..., T-1$, at the backward step of the SDDP algorithm, piecewise linear approximations $\mathfrak{Q}_t(\cdot)$ of the cost-to-go functions $\mathcal{Q}_t(\cdot)$ are constructed by solving problems

$$\operatorname*{Min}_{x_t \in \mathbb{R}^{n_t}} (c_t^j)^{\mathsf{T}} x_t + \mathfrak{Q}_{t+1}(x_t) \text{ s.t. } B_t^j \bar{x}_{t-1} + A_t^j x_t = b_t^j, \ x_t \geq 0,$$

$j = 1, ..., N_t$, and their duals, going backward in time $t = T, ..., 1$.

Denote by v^0 and $\widehat{v}_N$ the respective optimal values of the true and SAA problems.

By construction

$$\mathcal{Q}_t(\cdot) \geq \mathfrak{Q}_t(\cdot),\ t = 2, ..., T.$$

Therefore the optimal value of

$$\underset{x_1 \in \mathbb{R}^{n_1}}{\text{Min}}\ c_1^{\mathsf{T}} x_1 + \mathfrak{Q}_2(x_1)\ \text{ s.t. } A_1 x_1 = b_1,\ x_1 \geq 0$$

gives a lower bound for the optimal value $\widehat{v}_N$ of the SAA problem.

We also have that

$$v^0 \geq \mathbb{E}[\widehat{v}_N].$$

Therefore *on average* $\widehat{v}_N$ is also a lower bound for the optimal value of the true problem.

The approximate cost-to-go functions $\mathfrak{Q}_2, ..., \mathfrak{Q}_T$ and a feasible first stage solution $\bar{x}_1$ define a feasible policy. That is for a realization (sample path) $\xi_1, ..., \xi_T$ of the data process, $\bar{x}_t = \bar{x}_t(\xi_{[t]})$ are computed recursively in $t = 2, ..., T$ as a solution of

$$\underset{x_t}{\text{Min}}\, c_t^\top x_t + \mathfrak{Q}_{t+1}(x_t) \text{ s.t. } B_t \bar{x}_{t-1} + A_t x_t \leq b_t.$$

In the *forward step* of the SDDP algorithm M sample paths (scenarios) are generated and the corresponding $\bar{x}_t$, $t = 2, ..., T$, are used as trial points in the next iteration of the backward step.

It is essential for convergence of this algorithm that at each iteration in the forward step the paths (scenarios) are *resampled*, i.e., generated independently of the previous iteration.

Note that the functions $\mathfrak{Q}_2, ..., \mathfrak{Q}_T$ and $\bar{x}_1$ define a feasible policy also for the *true* problem.

Convergence of the SDDP algorithm

It is possible to show that, under mild regularity conditions, the SDDP algorithm converges as the number of iterations go to infinity. That is, the computed optimal values and generated policies converge w.p.1 to their counterparts of the considered SAA problem. However, the convergence can be very slow and one should take such mathematical proofs very cautiously.

Moreover, it should be remembered that the SAA problem is just an approximation of the "true" problem. It is possible to show that, in a certain probabilistic sense, the SAA problem converges to the "true" problem as all sample sizes N_t, $t = 2, ..., T$, tend to infinity.
It was found in our numerical experiments that optimal solutions of the SAA problems started to stabilize for sample sizes of about $N_t = 100$, $t = 2, ..., T$.

Stopping criteria

The policy value $\mathbb{E}\left[\sum_{t=1}^{T} c_t^{\mathsf{T}} \bar{x}_t(\xi_{[t]})\right]$ can be estimated in the forward step of the algorithm. That is, let $\xi_2^i, ..., \xi_T^i$, $i = 1, ..., M$, be sample paths (scenarios) generated at a current iteration of the forward step, and

$$\vartheta_i := \sum_{t=1}^{T} (c_t^i)^{\mathsf{T}} \bar{x}_t^i, \ i = 1, ..., M,$$

be the corresponding cost values. Then $\mathbb{E}[\vartheta_i] = \mathbb{E}\left[\sum_{t=1}^{T} c_t^{\mathsf{T}} \bar{x}_t(\xi_{[t]}^i)\right]$, and hence

$$\bar{\vartheta} = \frac{1}{M} \sum_{i=1}^{M} \vartheta_i$$

gives an unbiased estimate of the policy value.

Also

$$\widehat{\sigma}^2 = \frac{1}{M-1}\sum_{i=1}^{M}(\vartheta_i - \bar{\vartheta})^2$$

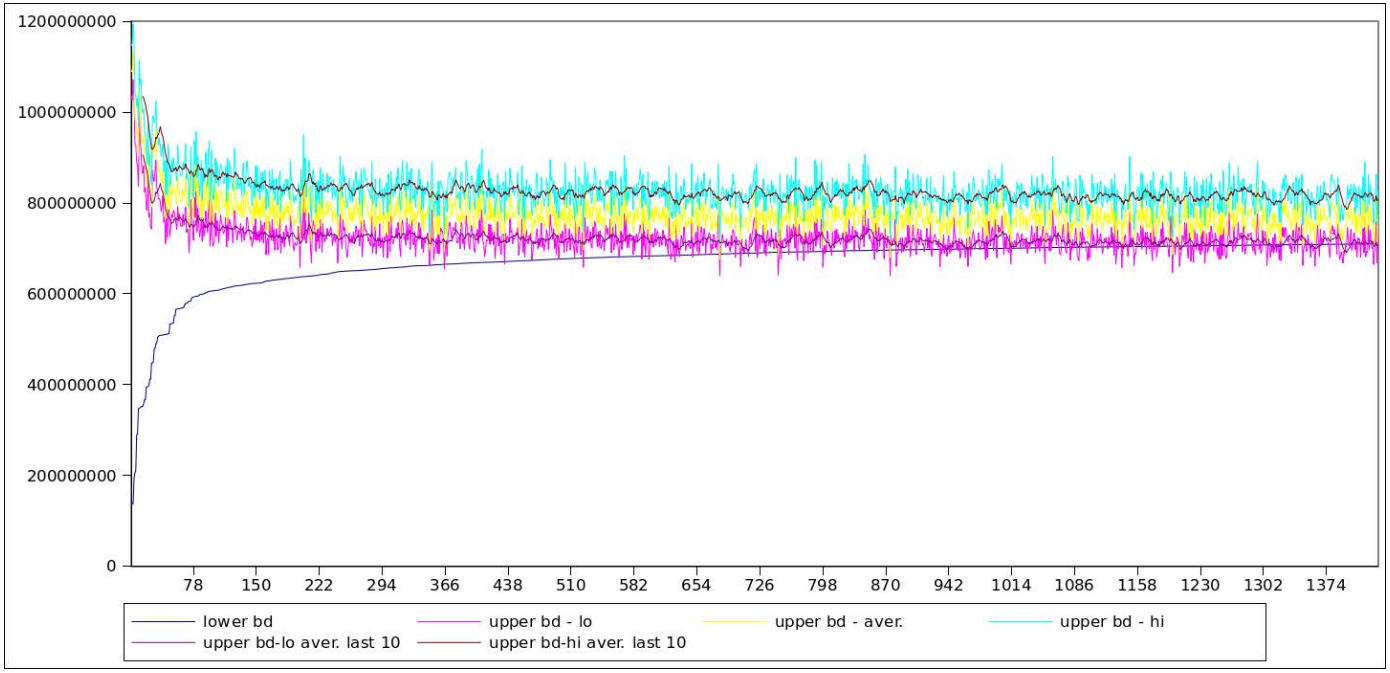
estimates variance of the sample $\vartheta_1, ..., \vartheta_M$. Hence

$$\bar{\vartheta} + z_\alpha \widehat{\sigma}/\sqrt{M}$$

gives an *upper* bound for the policy value with confidence of about $100(1-\alpha)\%$. Here z_α is the corresponding critical value.

At the same time this gives an upper bound for the optimal value of the corresponding multistage problem, SAA or the "true" problem depending from what data process the random scenarios were generated.

Typical example of behavior of the lower and upper bounds produced by the SDDP algorithm for an SAA problem

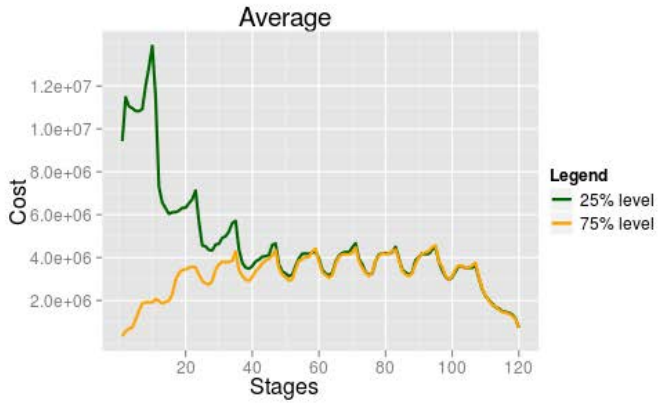


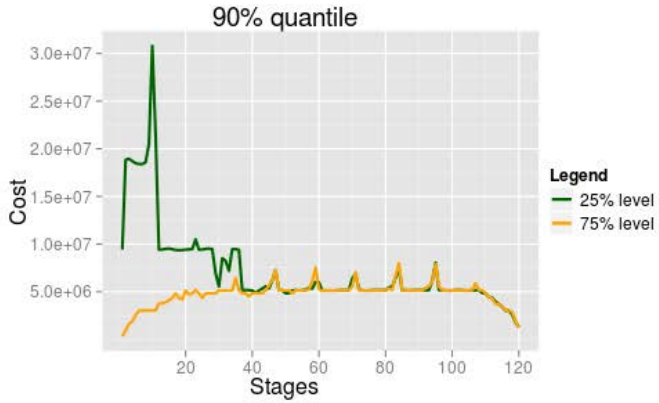
8 state variables, 120 stages, 1 cut per iteration

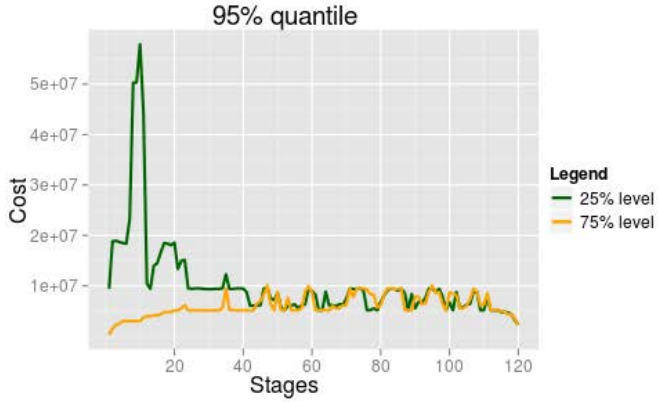
Theoretical analysis and numerical experiments indicate that computational complexity of the SDDP algorithm grows fast with increase of the number of state variables. The optimality gap jumped from 4% to 20% when the number of state variables was increased from 4 to 8 as a result of considering an autoregressive model.

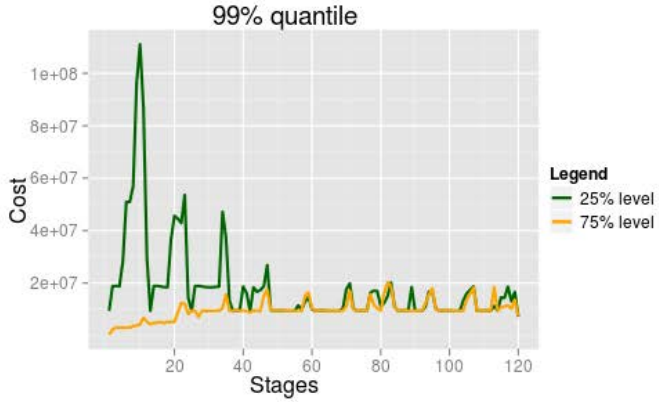
Sensitivity to initial conditions

Individual stage costs for the risk neutral approach in two cases: all the reservoirs start at 25% or at 75% of the maximum capacity. The yellow curve denotes the 75% initial reservoir level and the dark green denotes the 25% initial level.


Average
Cost
1.2e+07
1.0e+07
8.0e+06
6.0e+06
4.0e+06
2.0e+06
20
40
60
80
100
120
Stages
Legend
25% level
75% level


90% quantile
Cost
3.0e+07
2.5e+07
2.0e+07
1.5e+07
1.0e+07
5.0e+06
20
40
60
80
100
120
Stages
Legend
25% level
75% level


95% quantile
Cost
5e+07
4e+07
3e+07
2e+07
1e+07
20
40
60
80
100
120
Stages
Legend
25% level
75% level


99% quantile
Cost
1e+08
8e+07
6e+07
4e+07
2e+07
20
40
60
80
100
120
Stages
Legend
25% level
75% level

Variability of SAA problems

Table shows the 95% confidence interval for the lower bound and average policy value at iteration 3000 over a sample of 20 SAA problems. Each of the policy value observations was computed using 2000 scenarios. The last 2 columns of the table shows the range divided by the average of the lower bound (where the range is the difference between the maximum and minimum observation) and the standard deviation divided by the average value. This problem has relatively low variability (approx. 4%) for both of the lower bound and the average policy value.

	95% C.I. left ($\times 10^9$)	Average ($\times 10^9$)	95% C.I. right ($\times 10^9$)	range average	sdev. average
Lower bound	22.290	22.695	23.100	15.92%	4.07%
Average policy	27.333	27.836	28.339	17.05%	4.12%

SAA variability for risk neutral SDDP

Risk averse approach

How to control risk, i.e., to reduce chances of extreme costs, at every stage of the time process.

Value-at-Risk of a random outcome (variable) Z at level $\alpha \in (0,1)$:

$$\text{V@R}_\alpha(Z) = \inf\{t : F_Z(t) \geq 1-\alpha\},$$

where $F_Z(t) = \Pr(Z \leq t)$ is the cdf of Z. That is, $\text{V@R}_\alpha(Z)$ is the $(1-\alpha)$-quantile of the distribution of Z.

Note that $\text{V@R}_\alpha(Z) \leq c$ is equivalent to $\Pr(Z > c) \leq \alpha$. Therefore it could be a natural approach to impose constraints (chance constraints) of $\text{V@R}_\alpha(Z) \leq c$ for $Z = cost$, chosen constant c and significance level α at every stage of the process.

There are two problems with such approach. It is difficult to handle chance constraints numerically and could lead to infeasibility problems.

Average Value-at-Risk (also called *Conditional Value-at-Risk*)

$$\mathsf{AV@R}_\alpha(Z) = \inf_{t\in\mathbb{R}} \left\{ t + \alpha^{-1}\mathbb{E}[Z-t]_+ \right\}$$

Note that the minimum in the above is attained at $t^* = \mathsf{V@R}_\alpha(Z)$. If the cdf $F_Z(z)$ is continuous, then

$$\mathsf{AV@R}_\alpha(Z) = \mathbb{E}\Big[Z | Z \geq \mathsf{V@R}_\alpha(Z) \Big].$$

It follows that $\mathsf{AV@R}_\alpha(Z) \geq \mathsf{V@R}_\alpha(Z)$. Therefore the constraint $\mathsf{AV@R}_\alpha(Z) \leq c$ is a conservative approximation of the chance constraint $\mathsf{V@R}_\alpha(Z) \leq c$.

In the problem of minimizing expected cost $\mathbb{E}[Z]$ subject to the constraint $\mathsf{AV@R}_\alpha(Z) \leq c$, we impose an infinite penalty for violating this constraint. This could result in infeasibility of the obtained problem. Instead we can impose a finite penalty and consider problem of minimization of $\mathbb{E}[Z] + \kappa \mathsf{AV@R}_\alpha(Z)$ for some constant $\kappa > 0$. Note that this is equivalent to minimization of $\rho(Z)$, where

$$\rho(Z) = (1 - \lambda)\mathbb{E}[Z] + \lambda \mathsf{AV@R}_\alpha(Z)$$

for $\lambda \in (0, 1)$ and $\kappa = \frac{\lambda}{1-\lambda}$.

This leads to the following (nested) formulation of risk averse multistage problem.

$$\begin{aligned}\min_{A_1x_1\leq b_1} \quad & c_1^{\mathsf{T}}x_1 + \rho_{2|\xi_1}\Big[\inf_{\substack{B_2x_1+A_2x_2=b_2\\ x_2\geq 0}} c_2^{\mathsf{T}}x_2 + \ldots \\ & + \rho_{T-1|\xi_{[T-2]}}\Big[\inf_{\substack{B_{T-1}x_{T-2}+A_{T-1}x_{T-1}=b_{T-1}\\ x_{T-1}\geq 0}} c_{T-1}^{\mathsf{T}}x_{T-1} \\ & + \rho_{T|\xi_{[T-1]}}\big[\inf_{\substack{B_Tx_{T-1}+A_Tx_T=b_T\\ x_T\geq 0}} c_T^{\mathsf{T}}x_T\big]\Big]\Big],\end{aligned}$$

with

$$\rho_{t|\xi_{[t]}}(\cdot) := (1-\lambda)\mathbb{E}_{|\xi_{[t]}}[\cdot] + \lambda \mathsf{AV@R}_{\alpha|\xi_{[t]}}(\cdot)$$

being conditional analogue of $\rho(\cdot)$.

We can write the risk averse multistage programming problem as

$$\begin{array}{cl} \underset{x_1, x_2(\cdot), \ldots, x_T(\cdot)}{\text{Min}} & \bar{\rho}\Big[F_1(x_1) + F_2(x_2(\xi_{[2]}), \xi_2) + \cdots + F_T\left(x_T(\xi_{[T]}), \xi_T\right)\Big] \\ \text{s.t.} & x_1 \in \mathcal{X}_1,\ x_t(\xi_{[t]}) \in \mathcal{X}_t(x_{t-1}(\xi_{[t-1]}), \xi_t),\ t = 2, \ldots, T, \end{array}$$

where $F_t(x_t, \xi_t) = c_t^{\mathsf{T}} x_t$ and

$$\mathcal{X}_t(x_{t-1}, \xi_t) = \{x_t : B_t x_{t-1} + A_t x_t = b_t,\ x_t \geq 0\}.$$

$$\begin{aligned} \bar{\rho}(Z_1 + \ldots + Z_T) &= \rho_{|\xi_1}\Big(\rho_{|\xi_{[2]}}\Big(\cdots \rho_{|\xi_{[T-1]}}(Z_1 + \ldots + Z_T)\Big)\Big) \\ &= Z_1 + \rho_{|\xi_1}\Big(Z_2 + \rho_{|\xi_{[2]}}\Big(+ \cdots \rho_{|\xi_{[T-1]}}(Z_T)\Big)\Big) \end{aligned}$$

is the corresponding composite risk measure. The optimization is performed over (nonanticipative) policies $x_1, x_2(\xi_{[2]}), \ldots, x_T(\xi_{[T]})$ satisfying the feasibility constraints.

With some modifications the SDDP algorithm can be applied to the above multistage problem. Assuming the stagewise independence, the dynamic programming equations for the adaptive risk averse problem take the form

$$Q_t\left(x_{t-1},\xi_t\right) = \inf_{x_t\in\mathbb{R}^{n_t}} \left\{c_t^{\mathsf{T}} x_t + \mathcal{Q}_{t+1}(x_t) : B_t x_{t-1} + A_t x_t = b_t,\ x_t \geq 0\right\},$$

$t = T, ..., 2$, where $\mathcal{Q}_{T+1}(\cdot) \equiv 0$ and

$$\mathcal{Q}_{t+1}\left(x_t\right) := \rho_{t+1|\xi_{[t]}}\left[Q_{t+1}\left(x_t,\xi_{t+1}\right)\right].$$

Since ξ_{t+1} is independent of $\xi_{[t]}$, the cost-to-go functions $\mathcal{Q}_{t+1}\left(x_t\right)$ do not depend on the data process. In order to apply the backward step of the SDDP algorithm we only need to know how to compute subgradients of the cost-to-go functions.

The value of this problem corresponds to the total objective

$$\begin{aligned}\bar{\rho}(Z_1 + ... + Z_T) &= \rho_{|\xi_{[1]}}\Big(\cdots \rho_{|\xi_{[T-1]}}(Z_1 + ... + Z_T)\Big) \\ &= Z_1 + \rho_{|\xi_{[1]}}\Big(Z_2 + \cdots + \rho_{|\xi_{[T-1]}}(Z_T)\Big)\end{aligned}$$

The dynamic programming equations of the risk averse formulation of the SAA program take the form

$$Q_t^j(x_{t-1}) = \inf_{x_t}\Big\{(c_t^j)^{\mathsf{T}}x_t + \mathcal{Q}_{t+1}(x_t) : B_t^j x_{t-1} + A_t^j x_t = b_t^j,\ x_t \geq 0\Big\},$$

$j = 1, ..., N_t$, $t = T, \ldots, 2$, and

$$\mathcal{Q}_{t+1}(x_t) = \rho\left(Q_{t+1}^1(x_t), ..., Q_{t+1}^{N_{t+1}}(x_t)\right),$$

with $\mathcal{Q}_{T+1}(\cdot) \equiv 0$ and the first stage problem

$$\operatorname*{Min}_{A_1 x_1 \leq b_1} c_1^{\mathsf{T}}x_1 + \rho\left(Q_2^1(x_1), ..., Q_2^{N_2}(x_1)\right).$$

For $\rho(\cdot) = (1-\lambda)\mathbb{E}[\cdot] + \lambda \mathsf{AV@R}_\alpha(\cdot)$, and $(Z_1, ..., Z_N) = \left(Q^1_{t+1}(x_t), ..., Q^N_{t+1}(x_t)\right)$ we have that

$$\mathcal{Q}_{t+1}(x_t) = \frac{1-\lambda}{N_{t+1}} \sum_{j=1}^{N_{t+1}} Z_j + \lambda \left(Z_\iota + \frac{1}{\alpha N_{t+1}} \sum_{j: Z_j > Z_\iota} \left[Z_j - Z_\iota \right] \right),$$

where Z_ι is the $(1-\alpha)$-quantile of $Z_1, ..., Z_{N_{t+1}}$. Note that if $N_{t+1} < (1-\alpha)^{-1}$, then $Z_\iota = \max\{Z_1, ..., Z_{N_{t+1}}\}$.

A subgradient of $\mathcal{Q}_{t+1}(x_t)$ is given by

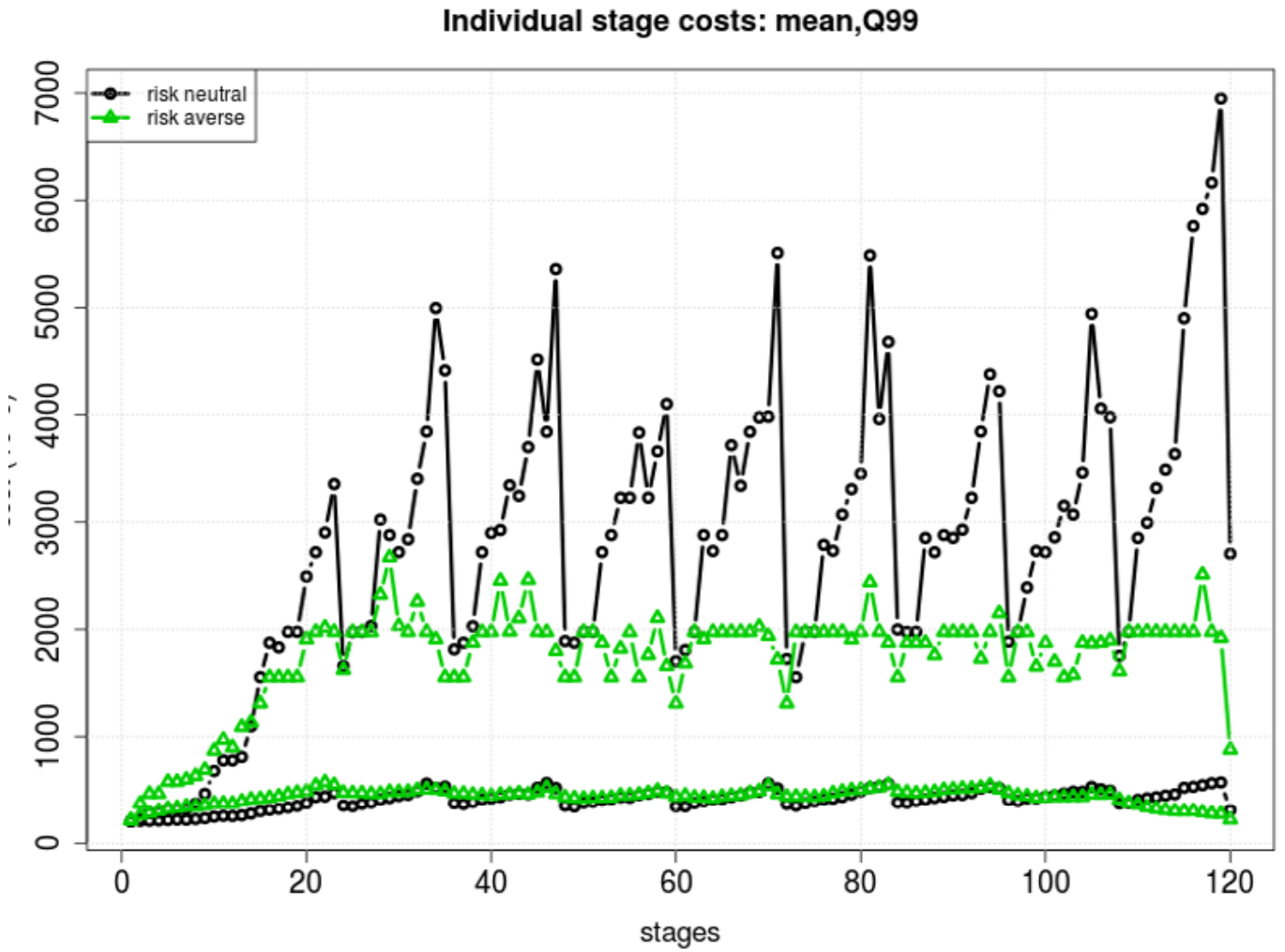
$$\nabla \mathcal{Q}_{t+1}(x_t) = \frac{1-\lambda}{N} \sum_{j=1}^{N_{t+1}} \nabla Q^j_{t+1}(x_t) +$$

$$\lambda \left(\nabla Q^\iota_{t+1}(x_t) + \frac{1}{\alpha N_{t+1}} \sum_{j: Z_j > Z_\iota} \left[\nabla Q^j_{t+1}(x_t) - \nabla Q^\iota_{t+1}(x_t) \right] \right).$$

These formulas allow construction of cuts in the backward step of the SDDP algorithm. In the forward step trial points are generated in the same way as in the risk neutral case.

Remarks

Unfortunately there is no easy way for evaluating value of the risk objective of generated policies, and hence constructing a corresponding upper bound. Some suggestions were made in the recent literature. However, in larger problems the optimality gap (between the upper and lower bounds) never approaches zero in any realistic time. Therefore stopping criteria based on stabilization of the lower bound (and may be optimal solutions) could be reasonable. Also it should be remembered that there is no intuitive interpretation for the risk objective $\bar{\rho}(cost)$ of the total cost. Rather the goal is to control risk at every stage of the process.


Individual stage costs: mean,Q99
risk neutral
risk averse
0
1000
2000
3000
4000
5000
6000
7000
0
20
40
60
80
100
120
stages