Scenario Grouping and Decomposition Algorithms for Chance-constrained Programs

Siqian Shen
Dept. of Industrial and Operations Engineering
University of Michigan

Joint work with
Yan Deng (UMich, Google)
Shabbir Ahmed (Georgia Tech)
Jon Lee (UMich)
(supported by NSF grants CMMI-1433066, 1633196 & ONR N00014-14-0315)

INFORMS Optimization Soceity Conference
March 24, 2018

Outline

Introduction

Optimization-based Scenario Grouping
- Base Model
- MILP Reformulation
- Branch-and-Cut

Heuristic-based Scenario Grouping

Numerical Studies
- Experimental Design
- Results of Group-based Bounds
- Results of Scenario Decomposition with Grouping

Chance-constrained Program

$$\begin{aligned} \min \quad & c^\top x && \text{(1a)} \\ \text{s.t.} \quad & \mathbb{P}\{x \in \mathcal{F}(\xi)\} \geq 1 - \epsilon && \text{(1b)} \\ & x \in \mathcal{X} \subseteq \mathbb{R}^d, && \text{(1c)} \end{aligned}$$

- x: a d-dimensional decision vector; $c \in \mathbb{R}^d$: cost parameter
- ξ: a multivariate random vector. (W.l.o.g., we consider a finite support $\Xi = \{\xi^1, \ldots, \xi^K\}$, and each scenario is realized with equal probability.)
- $\mathcal{F}(\xi) \subseteq \mathbb{R}^d$: a region parameterized by ξ. Let $\mathcal{F}_k = \mathcal{F}(\xi^k)$.
- $\mathcal{X}$: a deterministic feasible region; either continuous or discrete.

Literature Review

Chance-constrained programs have wide applications in energy, healthcare, transportation problems, but in general nonconvex and intractable to solve. We present some main literature below.

- **Convex approximations**: Prékopa (1970), Nemirovski and Shapiro (2006)
- **SAA and MILP-based algorithms**: Luedtke and Ahmed (2008),Pagnoncelli et al. (2009), Luedtke et al. (2010), Küçükyavuz (2012), Luedtke (2014), Song et al. (2014), Ahmed et al. (2016)
- **Scenario decomposition**: Watson et al. (2010), Ahmed (2013), Carøe and Schultz (1999), Dentcheva and Römisch (2004), Miller and Ruszczyński (2011), Collado et al. (2012), Deng et al. (2016)

Quantile Bounds I

An equiavlent formulation of model (1) is:

$$\begin{aligned} \text{CCP}: \quad v^* := \min \quad & c^\top x \\ \text{s.t.} \quad & \sum_{k=1}^{K} \mathbb{I}(x \notin \mathcal{F}_k) \leq K' \\ & x \in \mathcal{X}, \end{aligned}$$

where $\mathbb{I}(\cdot)$ is an indicator function and $K' := \lfloor \epsilon K \rfloor$. Using binary variables to model outcomes of the indicator function in all the scenarios, we can further reformulate CCP as an MILP with a knapsack constraint (see Ahmed et al. (2016)) and solve it by branch-and-cut (see Luedtke (2014))

Quantile Bounds II

- To satisfy the probabilistic constraint, x must fall in sufficiently many subregions $\mathcal{F}_k$'s.

Quantile Bounds II

- To satisfy the probabilistic constraint, x must fall in sufficiently many subregions $\mathcal{F}_k$'s.
- The optimal objective values of the K scenario subproblems:

$$\psi_k := \min\left\{c^\top x : x \in \mathcal{F}_k,\ x \in \mathcal{X}\right\},\ \forall k = 1, \ldots, K. \quad (2)$$

Quantile Bounds II

- To satisfy the probabilistic constraint, x must fall in sufficiently many subregions $\mathcal{F}_k$'s.
- The optimal objective values of the K scenario subproblems:

$$\psi_k := \min\left\{c^\top x : x \in \mathcal{F}_k,\ x \in \mathcal{X}\right\},\ \forall k = 1, \ldots, K. \qquad (2)$$

- Then order them to obtain a permutation σ of the set $\{1, \ldots, K\}$ such that $\psi_{\sigma_1} \geq \cdots \geq \psi_{\sigma_K}$.

Quantile Bounds II

- To satisfy the probabilistic constraint, x must fall in sufficiently many subregions $\mathcal{F}_k$'s.
- The optimal objective values of the K scenario subproblems:

$$\psi_k := \min\left\{c^\top x : x \in \mathcal{F}_k,\ x \in \mathcal{X}\right\},\ \forall k = 1, \ldots, K. \qquad (2)$$

- Then order them to obtain a permutation σ of the set $\{1, \ldots, K\}$ such that $\psi_{\sigma_1} \geq \cdots \geq \psi_{\sigma_K}$.
- Given $K' = \lfloor \epsilon K \rfloor$, the $(K'+1)^{\text{th}}$ quantile value, $\psi_{\sigma_{K'+1}}$, is a valid lower bound for CCP, due to that x will fall in at least one $\mathcal{F}_k$ with $k \in \{\sigma_1, \ldots, \sigma_{K'+1}\}$.

Scenario Grouping I

We partition $\{1, \ldots, K\}$, into N disjoint subsets $G_1, \ldots, G_N$, and obtain a relaxation of CCP as:

$$\begin{aligned} \text{SGM}: \quad v_{\text{SGM}} := \min \quad & c^\top x \\ \text{s.t.} \quad & \sum_{n=1}^{N} \mathbb{I}\left(x \notin \bigcap_{k \in G_n} \mathcal{F}_k \right) \leq K' \\ & x \in \mathcal{X}. \end{aligned}$$

- SGM is not a relaxation, if $\{G_1, \ldots, G_N\}$ does not form a partition of scenarios $\{1, \ldots, K\}$.

Scenario Grouping II

We consider the quantile bound of SGM as:

$$v^{\mathrm{Q}}_{\mathrm{SGM}} := \max\left\{\rho : \sum_{n=1}^{N} \mathbb{I}\left(\rho \leq \phi_n\right) \geq K' + 1\right\}, \text{ where we solve} \tag{3}$$

$$\phi_n := \min\left\{c^\top x : x \in \bigcap_{k \in G_n} \mathcal{F}_k,\ x \in \mathcal{X}\right\},\ n = 1, \ldots, N. \tag{4}$$

- We call Model (4) *group subproblems*.

Scenario Grouping II

We consider the quantile bound of SGM as:

$$v_{\text{SGM}}^{\text{Q}} := \max \left\{ \rho : \sum_{n=1}^{N} \mathbb{I}\left(\rho \leq \phi_n\right) \geq K' + 1 \right\}, \text{ where we solve} \tag{3}$$

$$\phi_n := \min \left\{ c^\top x : x \in \bigcap_{k \in G_n} \mathcal{F}_k,\ x \in \mathcal{X} \right\},\ n = 1, \ldots, N. \tag{4}$$

- We call Model (4) *group subproblems*.
- The grouping-based quantile bound $v_{\text{SGM}}^{\text{Q}}$ is a valid lower bound for CCP (i.e., $v_{\text{SGM}}^{\text{Q}} \leq v_{\text{SGM}} \leq v^*$).

Scenario Grouping II

We consider the quantile bound of SGM as:

$$v^{\mathrm{Q}}_{\mathrm{SGM}} := \max\left\{\rho : \sum_{n=1}^{N} \mathbb{I}\left(\rho \leq \phi_n\right) \geq K' + 1\right\}, \text{ where we solve} \tag{3}$$

$$\phi_n := \min\left\{c^\top x : x \in \bigcap_{k \in G_n} \mathcal{F}_k,\ x \in \mathcal{X}\right\},\ n = 1, \ldots, N. \tag{4}$$

- We call Model (4) *group subproblems*.
- The grouping-based quantile bound $v^{\mathrm{Q}}_{\mathrm{SGM}}$ is a valid lower bound for CCP (i.e., $v^{\mathrm{Q}}_{\mathrm{SGM}} \leq v_{\mathrm{SGM}} \leq v^*$).
- Related work: Escudero et al. (2013), Crainic et al. (2014), Ryan et al. (2016) (all for expectation-based stochastic programs)

Outline

Introduction

Optimization-based Scenario Grouping
- Base Model
- MILP Reformulation
- Branch-and-Cut

Heuristic-based Scenario Grouping

Numerical Studies
- Experimental Design
- Results of Group-based Bounds
- Results of Scenario Decomposition with Grouping

Notation and Parameter

Recall the following parameter:

- K: the number of scenarios, N: the number of groups
- $K' = \lfloor \epsilon K \rfloor$ where ϵ is the risk tolerance level in Model (1).

Decision variables:

- $y_{kn} \in \{0, 1\}, \ k = 1, \ldots, K, \ n = 1, \ldots, N$: whether scenario k is assigned to group G_n, such that $y_{kn} = 1$ if yes, and $= 0$ o.w.

Procedures:

- solve the ordered objective values $\phi_1, \ldots, \phi_N$ of group subproblems (4) and maximize $\phi_{K'+1}$ for $K' = \lfloor \epsilon K \rfloor$.

Optimal Scenario Grouping Model

To obtain the tightest quantile bound $v^{\mathrm{Q}}_{\mathrm{SGM}}$ that is equal to $\phi_{K'+1}$:

$$
\begin{aligned}
\mathrm{QGP}: \max \quad & \phi_{K'+1} && \text{(5a)}\\
\text{s.t.} \quad & \phi_n \le \min\{c^\top x : \mathbb{I}(x \in \mathcal{F}_k) \ge y_{kn},\ \forall k,\ x \in \mathcal{X}\} \quad \forall n = 1, \dots, N && \text{(5b)}\\
& \phi_n - \phi_{n+1} \ge 0 \quad \forall n = 1, \dots, N-1 && \text{(5c)}\\
& \sum_{n=1}^{N} y_{kn} = 1 \quad \forall k = 1, \dots, K && \text{(5d)}\\
& \sum_{k=1}^{K} y_{kn} \le P \quad \forall n = 1, \dots, N && \text{(5e)}\\
& y_{kn} \in \{0, 1\} \quad \forall n = 1, \dots, N,\ k = 1, \dots, K. && \text{(5f)}
\end{aligned}
$$

(5c) are to avoid symmetric solutions. We restrict each group size by an integer parameter P in (5e) with $P \ge K/N$. Without this, i.e., if $P = K$, the model will allocate scenarios densely into $K' + 1$ groups and make some subproblems hard to solve.

Outline

Introduction

Optimization-based Scenario Grouping
- Base Model
- MILP Reformulation
- Branch-and-Cut

Heuristic-based Scenario Grouping

Numerical Studies
- Experimental Design
- Results of Group-based Bounds
- Results of Scenario Decomposition with Grouping

Special Cases and MILP Reformulation

Consider chance-constrained linear programs with $\mathcal{X} = \mathbb{R}^d_+$ and $\mathcal{F}_k = \{x : A_k x \geq r_k\}$. We can reformulate (5b) as

$$\phi_n \leq \min\left\{c^\top x : A_k x \geq r_k - M_k(1 - y_{kn}),\ \forall k = 1, \ldots, K,\ x \in \mathbb{R}^d_+\right\}, \tag{6}$$

Let $\lambda_{kn} \in \mathbb{R}^{m_k}_+$ be the dual of the k^{th} set of constraints in the minimization problem in (6). The dual problem is:

$$D_n(y) := \quad \max \quad \sum_{k=1}^{K} \left(r_k^{\mathsf{T}} \lambda_{kn} - M_k^{\mathsf{T}} \lambda_{kn} (1 - y_{kn}) \right) \tag{7a}$$

$$\text{s.t.} \quad \sum_{k=1}^{K} A_k^{\mathsf{T}} \lambda_{kn} \leq c \tag{7b}$$

$$\lambda_{kn} \in \mathbb{R}^{m_k}_+. \tag{7c}$$

Linearize Dual Formulation

Further define $w_{kn} \equiv \lambda_{kn} y_{kn},\ \forall k = 1, \dots, K,\ n = 1, \dots, N$ and use the McCormick inequalities, we can linearize the dual and derive an equivalent MILP to replace the right-hand side of (6) in QGP:

$$
\begin{aligned}
\max_{y,\lambda,w} \quad & \sum_{k=1}^{K} \left(r_k^\top \lambda_{kn} - M_k^\top \lambda_{kn} + M_k^\top w_{kn} \right) && \text{(8a)} \\
\text{s.t.} \quad & \text{(7b)} \\
& w_{kn} \le \lambda_{kn},\ w_{kn} \le \overline{\lambda}_{kn} y_{kn}, \quad \forall k = 1, \dots, K && \text{(8b)} \\
& w_{kn} \ge \lambda_{kn} - \overline{\lambda}_{kn}(1 - y_{kn}), \quad \forall k = 1, \dots, K && \text{(8c)} \\
& \lambda_{kn}, w_{kn} \in \mathbb{R}_+^{m_k}, y_{kn} \in \{0, 1\}, \quad \forall k = 1, \dots, K. && \text{(8d)}
\end{aligned}
$$

The overall MILP for optimal scenario grouping is:

$$
\max_{\phi,\lambda,w,y} \quad \Bigg\{ \phi_{K'+1} : \phi_n \le \sum_{k=1}^{K} \left(r_k^\top \lambda_{kn} - M_k^\top \lambda_{kn} + M_k^\top w_{kn} \right),\ n = 1, \dots, N, \\
\text{(5c)–(5e), (7b), (8b), (8d)} \Bigg\}.
$$

Outline

Introduction

Optimization-based Scenario Grouping
- Base Model
- MILP Reformulation
- Branch-and-Cut

Heuristic-based Scenario Grouping

Numerical Studies
- Experimental Design
- Results of Group-based Bounds
- Results of Scenario Decomposition with Grouping

General Case and Branch-and-Cut

Consider a master problem of QGP as:

$$\max\left\{\phi_{K'+1} : \text{(5c)–(5e)},\ (y,\phi) \in \mathcal{A},\ \phi_n \in \mathbb{R},\ y_{kn} \in \{0,1\},\ \forall k, n\right\}.$$

For any $(\hat{y}, \hat{\phi})$ (where $\hat{y}$ could be fractional), consider and define a group set $G_n^* := \{k \in \{1, \ldots, K\} : \hat{y}_{kn} > 0\}$ for each $n = 1, \ldots, N$.

$$\phi_n^* = \min\left\{c^\top x : x \in \bigcap_{k \in G_n^*} \mathcal{F}_k,\ x \in \mathcal{X}\right\}.$$

If $\phi_n^* < \hat{\phi}_n$, following integer L-shaped method, we add a cut

$$\phi_n \leq (U - \phi_n^*)\left(\sum_{k:\hat{y}_{kn}=0} y_{kn} - 1\right) + U \tag{9}$$

where $U = \max\{c^\top x : x \in \mathcal{X}\}$ (an upper bound).

Anchored Grouping

- Solve scenario subproblems in (2) to obtain $\psi_1, \dots, \psi_K$, and sort their objective values such that $\psi_{k_1} \geq \cdots \geq \psi_{k_K}$.
- Clearly, $\psi_{k_{K'+1}}$ is a valid quantile bound for CCP.
- We construct N ($N \geq K/P$ and $N \geq K' + 1$) *non-empty* groups such that scenario k_n is in G_n for $n = 1, \dots, N$.
- Then distribute the remaining scenarios into different groups and meanwhile make sure that all the group sizes do not exceed P.
- Following this, the resulting SGM has a quantile bound that is at least $\psi_{k_{K'+1}}$. (Proved in Proposition 2 in our paper.)

Next, we group scenarios based on their similarity or dissimilarity. let $v^1, \dots, v^K$ be the vectors characterizing features of scenarios $1, \dots, K$, and measure the distance between two scenarios by $d(k, k') := \|v^k - v^{k'}\|$.

Greedy Scenario-grouping Approach

We evenly distribute K scenarios across N groups, as a result of which the sizes of the maximal and the minimal groups differ at most by 1.

- Randomly pick an ungrouped scenario to start a group.
- Then repeatedly include a scenario closest to the center of the incumbent group until we reach the size limit of that group.
- The center of any group $G \subseteq \{1, \ldots, K\}$, denoted $\overline{v}$, is defined as the arithmetic mean of the characterizing vectors of the contained scenarios, i.e.,

$$\overline{v} = (1/|G|) \sum_{k \in G} v^k.$$

This way we derive N groups with similar sizes and scenarios. As an alternative, we also employ the K-means clustering (Lloyd, 1982) in machine learning to group similar scenarios.

Grouping Dissimilar Scenarios

If dissimilar scenarios are grouped together, each group subproblem may become harder to solve, but can potentially produce tighter quantile bounds since the solution of each subproblem needs to satisfy constraints across dissimilar scenarios.

Grouping Dissimilar Scenarios

If dissimilar scenarios are grouped together, each group subproblem may become harder to solve, but can potentially produce tighter quantile bounds since the solution of each subproblem needs to satisfy constraints across dissimilar scenarios.

- First group similar scenarios to form Ω groups as $G'_1, \ldots, G'_\Omega$.
- Then collect one scenario from each group G'_ω ($\omega \in \{1, \ldots, \Omega\}$) to form a new group G_n, which then consists of Ω "dissimilar scenarios", each from a different group obtained from the previous similar scenario grouping.
- Repeat the above process until all the scenarios are grouped.
- To make this method comparable, we fix the total number of groups at N.

Outline

Introduction

Optimization-based Scenario Grouping
- Base Model
- MILP Reformulation
- Branch-and-Cut

Heuristic-based Scenario Grouping

Numerical Studies
- Experimental Design
- Results of Group-based Bounds
- Results of Scenario Decomposition with Grouping

Experimental Setup and Test Instances I

We test instances of the following two problems on the Stochastic Integer Programming Test Problem Library (SIPLIB):

- Problem (i) chance-constrained portfolio optimization: contains only linear variables and constraints, and its optimal grouping model QGP can be solved directly as an MILP.

$$\begin{aligned} \min \quad & c^\top x \\ \text{s.t.} \quad & \mathbb{P}\left\{(a(\xi))^\top x \geq r\right\} \geq 1-\epsilon \\ & x \in \mathcal{X} = \left\{x \in \mathbb{R}^d_+ : e^\mathsf{T} x = 1\right\}, \end{aligned}$$

$d = 20$ assets and the number of scenarios $K = 200$;
$a_k \sim U(0.8, 1.5)$; $r = 1.1$; $c \sim U(1, 100)$; $\epsilon = 0.075$.

Experimental Setup and Test Instances II

- Problem (ii) chance-constrained multi-dimensional 0-1 knapsack: contains binary packing variables, and we need to implement branch-and-cut for optimal grouping.
- Test two sets of instances *mk-20-10* and *mk-39-5* (*mk-n-m* have n items and m knapsack constraints.)
- scenario size $K \in \{100, 500, 1000\}$; five replications for each instance
- risk parameter $\epsilon = 0.1, 0.2$.

Linux workstation with four 3.4 GHz processors and 16 GB memory; one thread; gap tolerance = 0.01%; C++, CPLEX 12.6; CPU time limit = 3600 seconds.

Outline

Introduction

Optimization-based Scenario Grouping
- Base Model
- MILP Reformulation
- Branch-and-Cut

Heuristic-based Scenario Grouping

Numerical Studies
- Experimental Design
- Results of Group-based Bounds
- Results of Scenario Decomposition with Grouping

Overall of Procedures and Results

For each instance of Problems (i) and (ii)

- We directly optimize the MILP reformulation of CCP based on the K scenarios, and report under v^*;
- We solve the quantile bound of CCP and report the results under v^{Q};
- For Problem (i) we directly optimize the MILP in (9), and for Problem (ii) we implement branch-and-cut. We obtain the quantile bound $v^{\mathrm{Q}}_{\mathrm{SGM}}$ and report under **OG**;
- We construct N groups by distributing scenarios to each group in a round-robin manner, and then compute the quantile bound $v^{\mathrm{Q}}_{\mathrm{SGM}}$ and report under **RG** (round-robin grouping);
- We construct N groups by applying the anchored grouping method and by following the heuristics to group similar or dissimilar scenarios. We then compute $v^{\mathrm{Q}}_{\mathrm{SGM}}$ of the corresponding SGM for each heuristic and report their results under **AG** (anchored grouping), **SG** (similar scenario grouping), **KG** (K-means clustering grouping), and **DG** (dissimilar scenario grouping), respectively.

Results of Problem (i) Instances

Inst.	v^*	v^Q	v^Q_{SGM} OG	RG	AG	SG	KG	DG
			Group Size: $N = 100, P = 2$					
1	**40.7**	9.31	**15.73**	9.74	9.40	9.40	8.89	9.82
2	15.36	8.42	9.75	8.52	8.47	8.49	8.28	8.59
3	29.13	20.56	22.80	20.57	20.61	20.57	20.57	20.80
4	36.91	4.37	9.34	4.56	4.51	4.48	4.47	4.66
5	23.91	7.86	11.45	8.55	8.16	7.86	7.89	8.20
6	35.18	9.78	13.32	9.96	9.96	9.85	9.81	10.02
7	41.59	12.06	17.36	12.20	12.25	12.20	12.24	12.66
8	22.52	7.34	10.63	7.47	7.68	7.49	7.53	7.66
9	43.98	13.94	21.67	14.68	14.11	14.06	14.56	15.09
10	33.50-35.73†	10.48	15.25	10.48	10.70	10.61	10.70	10.83
			Group Size: $N = 20, P = 10$					
1	**40.7**	9.31	**24.03**	9.38	11.81	12.51	12.50	13.15
2	15.36	8.42	11.55	8.83	9.09	9.38	9.18	9.05
3	29.13	20.56	25.56	20.56	21.33	20.93	20.62	21.20
4	36.91	4.37	16.88	5.86	5.88	5.99	6.71	6.55
5	23.91	7.86	17.49	8.59	8.47	8.30	9.31	9.66
6	35.18	9.78	21.86	11.67	12.30	11.44	11.08	11.85
7	41.59	12.06	24.81	14.83	15.17	14.41	15.26	15.78
8	22.52	7.34	15.85	8.20	9.80	8.46	8.45	9.22
9	43.98	13.94	27.79	16.92	17.76	17.47	17.12	17.96
10	33.50-35.73†	10.48	19.27	11.18	13.05	11.85	12.56	12.76
			Group Size: $N = 10, P = 20$					
1	**40.7**	9.31	**36.73**	11.04	14.84	16.65	17.57	17.60
2	15.36	8.42	13.69	9.15	9.75	10.36	10.17	9.54
3	29.13	20.56	28.65	20.56	22.07	21.30	20.67	21.60
4	36.91	4.37	30.48	7.53	7.67	8.01	10.06	9.21
5	23.91	7.86	22.72	8.63	8.80	8.76	10.99	11.38
6	35.18	9.78	33.88	13.68	15.19	13.30	12.52	14.02
7	41.59	12.06	35.47	18.04	18.80	17.03	19.02	19.66
8	22.52	7.34	21.64	9.01	12.52	9.57	9.47	11.09
9	43.98	13.94	35.63	19.50	22.36	21.70	20.12	21.38
10	33.50-35.73†	10.48	27.37	11.93	15.91	13.23	14.73	15.05

Average CPU Time for Problem (i)

v^*	v^Q	Group Size	v^Q_{SGM}					
			OG	RG	AG	SG	KG	DG
		$N = 100, P = 2$	36.83	11.69	11.57	12.02	15.92	12.17
3103.72	10.44	$N = 20, P = 10$	249.27	86.03	59.81	66.46	72.01	107.19
		$N = 10, P = 20$	909.48	182.61	117.73	160.02	131.88	292.41

- Quantile bounds are very fast to obtain.
- The optimal objective value given by the MILP reformulation is hard to compute.
- The CPU time of obtaining quantile bounds using the optimal grouping and other heuristic grouping methods drastically increases as we increse the size of each group (or decrease the number of groups).
- The CPU time of optimal grouping is longer than heuristic grouping.

Root gap closed after adding bounds for Problem (i)

Inst.	v^{Q}	$N = 100$ $(P = 2)$						$N = 20$	$N = 10$
		OG	RG	AG	SG	KG	DG	OG	OG
1	0%	12%	0%	0%	0%	0%	1%	34%	89%
2	3%	9%	3%	3%	3%	2%	3%	25%	74%
3	4%	8%	4%	4%	4%	4%	4%	26%	82%
4	0%	11%	0%	0%	0%	0%	0%	21%	67%
5	0%	9%	1%	1%	0%	0%	1%	36%	88%
6	0%	7%	0%	0%	0%	0%	1%	33%	86%
7	1%	7%	1%	1%	1%	1%	1%	29%	66%
8	0%	8%	0%	0%	0%	0%	0%	31%	85%
9	0%	10%	1%	0%	1%	1%	1%	19%	74%
10	1%	8%	1%	1%	1%	1%	1%	17%	54%

- The bounds are much more stronger for the LP relaxation at root node if we use larger-sized groups.
- The bounds based on heuristic grouping is not effective at all for closing the root node gap.

Bound comparison for Problem (ii) *mk-20-10* instances

Inst.	ϵ	K	v^{Q}	v^{Q}_{SGM} OG	OG-20%	OG-50%	RG	AG	SG	KG	DG
				Group Size: $N = 0.1K,\ P = 10$							
mk-20-10	0.1	100	1.6%	0.5%	0.9%	0.5%	1.1%	1.0%	1.0%	0.9%	1.0%
		500	1.8%	0.6%	1.3%	0.7%	1.5%	1.4%	1.4%	1.2%	1.4%
		1000	2.0%	0.8%	1.7%	0.8%	1.8%	1.8%	1.7%	1.6%	1.8%
	0.2	100	2.3%	0.7%	2.1%	0.8%	2.3%	2.3%	2.2%	2.1%	2.3%
		500	1.5%	0.3%	1.3%	0.4%	1.5%	1.5%	1.5%	1.4%	1.5%
		1000	2.2%	0.6%	1.9%	0.8%	2.2%	2.2%	2.1%	2.0%	2.2%
				Group Size: $N = 0.05K,\ P = 20$							
mk-20-10	0.1	100	1.6%	0.3%	0.8%	0.3%	1.1%	1.1%	1.0%	0.8%	1.0%
		500	1.8%	0.4%	1.2%	0.4%	1.4%	1.4%	1.3%	0.9%	1.3%
		1000	2.0%	0.5%	1.3%	0.5%	1.7%	1.7%	1.6%	1.3%	1.5%
	0.2	100	2.3%	0.3%	1.7%	0.4%	2.1%	2.2%	2.2%	1.8%	2.1%
		500	1.5%	0.2%	1.0%	0.3%	1.5%	1.4%	1.5%	1.1%	1.4%
		1000	2.2%	0.3%	1.7%	0.3%	2.2%	2.0%	2.0%	1.6%	2.0%

- Different from Problem (i), all the bounds including v^{Q} are very tight.
- Heuristic-based grouping bounds v^{Q}_{SGM} are slightly tighter than v^{Q}.
- The optimization grouping bound v^{Q}_{SGM} is still much tighter than the others, and can be strengthened if we increase P and decrease the number of groups.

Average Time for Problem (ii) *mk-20-10* Instances with $N = 0.1K$

Inst.	ϵ	K	v^{Q}	v^{Q}_{SGM} OG	RG	AG	SG	KG	DG
mk-20-10	0.1	100	24.09	29.67	6.38	6.51	6.64	7.72	6.49
		500	143.27	121.77	34.53	34.18	33.49	49.38	33.84
		1000	272.99	238.73	63.69	68.15	64.96	100.63	63.05
	0.2	100	24.01	29.56	6.77	6.91	7.04	8.80	6.57
		500	146.16	211.70	36.28	34.83	36.64	47.16	37.01
		1000	277.48	349.51	63.70	65.61	63.06	97.46	64.34

- The heuristic grouping based bounds are easier to obtain.
- The quantile bounds and optimal grouping bounds require relatively the same computational effort.
- The CPU time of all methods for obtaining the bounds increases as we increase scenario numbers.

Outline

Introduction

Optimization-based Scenario Grouping
- Base Model
- MILP Reformulation
- Branch-and-Cut

Heuristic-based Scenario Grouping

Numerical Studies
- Experimental Design
- Results of Group-based Bounds
- Results of Scenario Decomposition with Grouping

Incorporating Grouping into Scenario Decomposition

We investigate the effectiveness of scenario grouping in scenario decomposition for solving chance-constrained 0-1 programs.

- u and ℓ: the upper and lower bounds of the optimal objective.
- Specifically, $u = c^\mathsf{T} x$ based on some best found solution x, and ℓ equals to the quantile bound of SGM, respectively.
- Until we close the gap between u and ℓ, we repeat
  - (i) finding a set of scenario groups;
  - (ii) optimizing group subproblems to identify temporary x-solutions and evaluate bounds;
  - (iii) eliminating the 0-1 x-solutions that have already been evaluated via no-good cuts included in the feasible region $\mathcal{X}$ in each group subproblem.

Scenario Decomposition Results for Problem (ii) instances with $N = 0.1K$ ($P = 10$)

Inst.	ϵ	K	Non-G		KG (ReG)		KG (w/o ReG)		OG (ReG)		OG (w/o ReG)
			Time (s)	Gap	Time (s)	Gap	Time (s)	Gap	Time (s)	Gap	Time (s)
mk-20-10	0.1	100	113.78	0.0% (5)	35.18	0.0% (5)	31.55	0.0% (5)	80.67	0.0% (5)	29.45
		500	377.22	0.0% (5)	134.64	0.0% (5)	120.97	0.0% (5)	419.92	0.0% (5)	86.20
		1000	1012.73	0.0% (5)	164.26	0.0% (5)	146.01	0.0% (5)	1217.47	0.0% (5)	71.33
	0.2	100	174.40	0.0% (5)	64.54	0.0% (5)	58.87	0.0% (5)	121.02	0.0% (5)	57.14
		500	367.88	0.0% (5)	173.29	0.0% (5)	158.59	0.0% (5)	325.24	0.0% (5)	91.24
		1000	1315.68	0.0% (5)	374.00	0.0% (5)	334.10	0.0% (5)	1361.49	0.0% (5)	257.56
mk-39-5	0.1	100	LIMIT	3.6% (0)	LIMIT	1.2% (0)	LIMIT	2.1% (0)	LIMIT	0.5% (0)	1955.93
		500	LIMIT	3.9% (0)	LIMIT	1.5% (0)	LIMIT	2.3% (0)	LIMIT	0.8% (0)	2306.17
		1000	LIMIT	4.0% (0)	LIMIT	1.9% (0)	LIMIT	2.2% (0)	LIMIT	0.8% (0)	2697.84
	0.2	100	LIMIT	3.4% (0)	LIMIT	1.7% (0)	LIMIT	2.6% (0)	LIMIT	0.5% (0)	2437.42
		500	LIMIT	3.2% (0)	LIMIT	1.9% (0)	LIMIT	2.7% (0)	LIMIT	0.7% (0)	2840.36
		1000	LIMIT	3.8% (0)	LIMIT	1.8% (0)	LIMIT	3.4% (0)	LIMIT	0.6% (0)	3258.73

- Scenario decomposition with optimal grouping has much shorter time on average as compared to no grouping.

Scenario Decomposition Results for Problem (ii) instances with $N = 0.1K$ ($P = 10$)

Inst.	ϵ	K	Non-G		KG (ReG)		KG (w/o ReG)		OG (ReG)		OG (w/o ReG)
			Time (s)	Gap	Time (s)	Gap	Time (s)	Gap	Time (s)	Gap	Time (s)
mk-20-10	0.1	100	113.78	0.0% (5)	35.18	0.0% (5)	31.55	0.0% (5)	80.67	0.0% (5)	29.45
		500	377.22	0.0% (5)	134.64	0.0% (5)	120.97	0.0% (5)	419.92	0.0% (5)	86.20
		1000	1012.73	0.0% (5)	164.26	0.0% (5)	146.01	0.0% (5)	1217.47	0.0% (5)	71.33
	0.2	100	174.40	0.0% (5)	64.54	0.0% (5)	58.87	0.0% (5)	121.02	0.0% (5)	57.14
		500	367.88	0.0% (5)	173.29	0.0% (5)	158.59	0.0% (5)	325.24	0.0% (5)	91.24
		1000	1315.68	0.0% (5)	374.00	0.0% (5)	334.10	0.0% (5)	1361.49	0.0% (5)	257.56
mk-39-5	0.1	100	LIMIT	3.6% (0)	LIMIT	1.2% (0)	LIMIT	2.1% (0)	LIMIT	0.5% (0)	1955.93
		500	LIMIT	3.9% (0)	LIMIT	1.5% (0)	LIMIT	2.3% (0)	LIMIT	0.8% (0)	2306.17
		1000	LIMIT	4.0% (0)	LIMIT	1.9% (0)	LIMIT	2.2% (0)	LIMIT	0.8% (0)	2697.84
	0.2	100	LIMIT	3.4% (0)	LIMIT	1.7% (0)	LIMIT	2.6% (0)	LIMIT	0.5% (0)	2437.42
		500	LIMIT	3.2% (0)	LIMIT	1.9% (0)	LIMIT	2.7% (0)	LIMIT	0.7% (0)	2840.36
		1000	LIMIT	3.8% (0)	LIMIT	1.8% (0)	LIMIT	3.4% (0)	LIMIT	0.6% (0)	3258.73

- Scenario decomposition with optimal grouping has much shorter time on average as compared to no grouping.
- Slightly faster for optimizing *mk-20-10* instances if we do not re-group scenarios in each iteration.

Scenario Decomposition Results for Problem (ii) instances with $N = 0.1K$ ($P = 10$)

Inst.	ϵ	K	Non-G		KG (ReG)		KG (w/o ReG)		OG (ReG)		OG (w/o ReG)
			Time (s)	Gap	Time (s)	Gap	Time (s)	Gap	Time (s)	Gap	Time (s)
mk-20-10	0.1	100	113.78	0.0% (5)	35.18	0.0% (5)	31.55	0.0% (5)	80.67	0.0% (5)	29.45
		500	377.22	0.0% (5)	134.64	0.0% (5)	120.97	0.0% (5)	419.92	0.0% (5)	86.20
		1000	1012.73	0.0% (5)	164.26	0.0% (5)	146.01	0.0% (5)	1217.47	0.0% (5)	71.33
	0.2	100	174.40	0.0% (5)	64.54	0.0% (5)	58.87	0.0% (5)	121.02	0.0% (5)	57.14
		500	367.88	0.0% (5)	173.29	0.0% (5)	158.59	0.0% (5)	325.24	0.0% (5)	91.24
		1000	1315.68	0.0% (5)	374.00	0.0% (5)	334.10	0.0% (5)	1361.49	0.0% (5)	257.56
mk-39-5	0.1	100	LIMIT	3.6% (0)	LIMIT	1.2% (0)	LIMIT	2.1% (0)	LIMIT	0.5% (0)	1955.93
		500	LIMIT	3.9% (0)	LIMIT	1.5% (0)	LIMIT	2.3% (0)	LIMIT	0.8% (0)	2306.17
		1000	LIMIT	4.0% (0)	LIMIT	1.9% (0)	LIMIT	2.2% (0)	LIMIT	0.8% (0)	2697.84
	0.2	100	LIMIT	3.4% (0)	LIMIT	1.7% (0)	LIMIT	2.6% (0)	LIMIT	0.5% (0)	2437.42
		500	LIMIT	3.2% (0)	LIMIT	1.9% (0)	LIMIT	2.7% (0)	LIMIT	0.7% (0)	2840.36
		1000	LIMIT	3.8% (0)	LIMIT	1.8% (0)	LIMIT	3.4% (0)	LIMIT	0.6% (0)	3258.73

- Scenario decomposition with optimal grouping has much shorter time on average as compared to no grouping.
- Slightly faster for optimizing *mk-20-10* instances if we do not re-group scenarios in each iteration.
- When solving *mk-39-5* instances, the re-grouping procedures can improve the optimality gaps if we cannot optimize the instances within the time limit.

Conclusions

We investigate:

- optimization driven scenario grouping for strengthening quantile bounds of general chance-constrained programs.
- solution methods for optimal grouping: MILP & branch-and-cut.
- heuristic-based scenario grouping methods
- improvements of bounds & performance of scenario decomposition

Conclusions

We investigate:

- optimization driven scenario grouping for strengthening quantile bounds of general chance-constrained programs.
- solution methods for optimal grouping: MILP & branch-and-cut.
- heuristic-based scenario grouping methods
- improvements of bounds & performance of scenario decomposition

Future research

- developing more efficient cutting-plane methods;
- implementing scenario grouping and decomposition algorithms in distributed computing frameworks;
- scenario grouping approaches broader classes of risk-averse stochastic programs.