Robust combinatorial optimization with variable uncertainty

Michael Poss

Heudiasyc UMR CNRS 7253, Université de Technologie de Compiègne

17th Aussois Combinatorial Optimization Workshop

Outline

Outline

Combinatorial optimization under uncertainty

$$\begin{aligned} \min \quad & \sum_{i=1}^{n} c_i x_i \\ \text{s.t.} \quad & \sum_{j=1}^{n} a_{ij} x_j \leq b_i, \quad i = 1, \ldots, m \\ & x \in \{0,1\}^n \end{aligned}$$

Suppose that the parameters (a, b, c) are uncertain:

- They vary over time
- They must be predicted from historical data
- They cannot be measured with enough accuracy
- ...

Let's do something clever (and useful)!

How much do we know?

$$\underbrace{\text{Stochastic programming}}_{\text{A lot}} \quad \Leftrightarrow \quad \underbrace{\text{Robust programming}}_{\text{A little}}$$

How much do we know?

$$\underbrace{\text{Stochastic programming}}_{\text{A lot}} \Leftrightarrow \underbrace{\text{Robust programming}}_{\text{A little}}$$

Robust pr. Uncertain parameters are merely assumed to belong to an uncertainty set $\mathbf{U} \Rightarrow$ one wishes to optimize some worst-case objective over the uncertainty set

Stochastic pr. Uncertain parameters are precisely described by probability distributions $\Rightarrow$ one wishes to optimize some expectation, variance, Value-at-risk, ...

Intermediary models exist: distributionally robust optimization, ambiguous chance-constrained

When do we take decisions?

Now All decisions must be taken before the uncertainty is known with precision $\Rightarrow$ probability constraints, (static) robust optimization

Delayed Some decisions may be delayed until the uncertainty is revealed $\Rightarrow$ multi-stage stochastic programming, adjustable robust optimization

Robust combinatorial optimization

$$
\begin{aligned}
\min \quad & \sum_{i=1}^{n} c_i x_i \\
\text{s.t.} \quad & \sum_{j=1}^{n} a_{ij} x_j \leq b_i, \quad i = 1, \ldots, m, \quad \forall a_i \in \mathbf{U}_i \\
& x \in \{0, 1\}^n,
\end{aligned}
$$

The linear relaxation of this problem is tractable if $\mathbf{U}_i$ is defined by conic constraints:

$$\mathbf{U}_i = \{a_i \in \mathbb{R}^n : u\, a_i - v \in K\}.$$

In particular, polyhedrons and polytopes are nice ($K = \mathbb{R}^n_+$).

Feasibility set

$$\sum a_i x_i \leq b, \quad \forall a \in \mathbf{U} \qquad \Leftrightarrow \qquad \begin{array}{l} \sum_j c_j \alpha_j \leq b \\ \sum_j u_{ji} \alpha_j \geq x_j \\ \alpha_j \geq 0 \end{array}$$

The feasibility set of the constraint is a polyhedron (thus, convex) !

Feasibility set

$$\sum a_i x_i \leq b, \quad \forall a \in \mathbf{U} \qquad \Leftrightarrow \qquad \begin{array}{l} \sum_j c_j \alpha_j \leq b \\ \sum_j u_{ji} \alpha_j \geq x_j \\ \alpha_j \geq 0 \end{array}$$

The feasibility set of the constraint is a polyhedron (thus, convex) !
A (very) popular polyhedral uncertainty set is (Bertsimas and Sim, 2004):

$$\mathbf{U}^\Gamma := \left\{ a \in \mathbb{R}^n : a_i = \overline{a}_i + \delta_i \hat{a}_i, -1 \leq \delta_i \leq 1, \sum |\delta_i| \leq \Gamma \right\}.$$

Main reasons for popularity:

- Nice computational properties for MIP and combinatorial problems.
- Intuitive interpretation.
- Probabilistic interpretation.

- Consider a vehicle routing problem with uncertain travel times t and time limit T. For simplicity, we suppose $t \in \mathbf{U}^{\Gamma}$ for $\Gamma = 5$.

- Consider a vehicle routing problem with uncertain travel times t and time limit T. For simplicity, we suppose $t \in \mathbf{U}^\Gamma$ for $\Gamma = 5$.
- Consider two robust-feasible routes x^1 and x^2 : $\|x^1\|_1 = 3$ and $\|x^2\|_1 = 10$.

- Consider a vehicle routing problem with uncertain travel times t and time limit T. For simplicity, we suppose $t \in \mathbf{U}^\Gamma$ for $\Gamma = 5$.
- Consider two robust-feasible routes x^1 and x^2 : $\|x^1\|_1 = 3$ and $\|x^2\|_1 = 10$.
- Because x^1 and x^2 are robust-feasible:

$$\sum_{i:x_i^1=1} t^i \leq T, \quad \forall a \in \mathbf{U}^\Gamma, \qquad \text{and} \qquad \sum_{i:x_i^2=1} t^i \leq T, \quad \forall a \in \mathbf{U}^\Gamma$$

- Consider a vehicle routing problem with uncertain travel times t and time limit T. For simplicity, we suppose $t \in \mathbf{U}^\Gamma$ for $\Gamma = 5$.
- Consider two robust-feasible routes x^1 and x^2 : $\|x^1\|_1 = 3$ and $\|x^2\|_1 = 10$.
- Because x^1 and x^2 are robust-feasible:

$$\sum_{i:x_i^1=1} t^i \leq T, \quad \forall a \in \mathbf{U}^\Gamma, \qquad \text{and} \qquad \sum_{i:x_i^2=1} t^i \leq T, \quad \forall a \in \mathbf{U}^\Gamma$$

which becomes

$$\sum_{i:x_i^1=1} t^i \leq T, \qquad \text{and} \qquad \sum_{i:x_i^2=1} t^i \leq T, \quad \forall a \in \mathbf{U}^\Gamma$$

- Consider a vehicle routing problem with uncertain travel times t and time limit T. For simplicity, we suppose $t \in \mathbf{U}^{\Gamma}$ for $\Gamma = 5$.
- Consider two robust-feasible routes x^1 and x^2 : $\|x^1\|_1 = 3$ and $\|x^2\|_1 = 10$.
- Because x^1 and x^2 are robust-feasible:

$$\sum_{i:x_i^1=1} t^i \leq T, \quad \forall a \in \mathbf{U}^{\Gamma}, \qquad \text{and} \qquad \sum_{i:x_i^2=1} t^i \leq T, \quad \forall a \in \mathbf{U}^{\Gamma}$$

which becomes

$$\sum_{i:x_i^1=1} t^i \leq T, \qquad \text{and} \qquad \sum_{i:x_i^2=1} t^i \leq T, \quad \forall a \in \mathbf{U}^{\Gamma}$$

For any probability distribution for t:

$$P\left(\sum_{i:x_i^1=1} t^i > T\right) = 0.$$

- Consider a vehicle routing problem with uncertain travel times t and time limit T. For simplicity, we suppose $t \in \mathbf{U}^\Gamma$ for $\Gamma = 5$.
- Consider two robust-feasible routes x^1 and x^2 : $\|x^1\|_1 = 3$ and $\|x^2\|_1 = 10$.
- Because x^1 and x^2 are robust-feasible:

$$\sum_{i:x_i^1=1} t^i \leq T, \quad \forall a \in \mathbf{U}^\Gamma, \qquad \text{and} \qquad \sum_{i:x_i^2=1} t^i \leq T, \quad \forall a \in \mathbf{U}^\Gamma$$

which becomes

$$\sum_{i:x_i^1=1} t^i \leq T, \qquad \text{and} \qquad \sum_{i:x_i^2=1} t^i \leq T, \quad \forall a \in \mathbf{U}^\Gamma$$

For any probability distribution for t:

$$P\left(\sum_{i:x_i^1=1} t^i > T\right) = 0.$$

If x^2 is not robust-feasible for $\Gamma = 10$, there exists probability distributions:

$$P\left(\sum_{i:x_i^2=1} t^i > T\right) > 0$$

Outline

1. Robust optimization
2. Variable budgeted uncertainty
3. Cost uncertainty

Robust optimization and probabilistic constraint

Let $\tilde{a}_i$ be random variables and $\epsilon > 0$. The chance constraint

$$P\left(\sum \tilde{a}_i x_i > b\right) \leq \epsilon \qquad (1)$$

leads to very difficult optimization problems in general.

Robust optimization and probabilistic constraint

Let $\tilde{a}_i$ be random variables and $\epsilon > 0$. The chance constraint

$$P\left(\sum \tilde{a}_i x_i > b\right) \leq \epsilon \tag{1}$$

leads to very difficult optimization problems in general.
In some situations, we know that (1) can be approximated by

$$\sum a_i x_i \leq b \qquad \forall a \in \mathbf{U} \tag{2}$$

for a properly chosen $\mathbf{U}$.
These approximations are conservative: any x feasible for (2) is feasible for (1).
We must balance conservatism and protection cost $\Rightarrow$ devise good protection sets $\mathbf{U}$.

Robust optimization and probabilistic constraint

What about $\mathbf{U}^{\Gamma}$?

Robust optimization and probabilistic constraint

What about $\mathbf{U}^\Gamma$?
Let $\tilde{a}_i$ be random variables independently and symmetrically distributed in $[\overline{a}_i - \hat{a}_i, \overline{a}_i + \hat{a}_i]$.
Bertsimas and Sim (2004) prove that if a vector x satisfies the robust constraint

$$\sum a_i x_i \leq b \quad \forall a \in \mathbf{U}^\Gamma,$$

then it satisfies also the probabilistic constraint

$$P\left(\sum \tilde{a}_i x_i > b\right) \leq \exp\left(-\frac{\Gamma^2}{2n}\right).$$

Something is wrong ...

From

$$P\left(\sum \tilde{a}_i x_i > b\right) \leq \exp\left(-\frac{\Gamma^2}{2n}\right),$$

we see that choosing $\Gamma = (-2\ln(\epsilon))^{1/2}\, n^{1/2}$ yields

$$P\left(\sum \tilde{a}_i x_i > b\right) \leq \epsilon.$$

Something is wrong ...

From

$$P\left(\sum \tilde{a}_i x_i > b\right) \leq \exp\left(-\frac{\Gamma^2}{2n}\right),$$

we see that choosing $\Gamma = (-2\ln(\epsilon))^{1/2}\, n^{1/2}$ yields

$$P\left(\sum \tilde{a}_i x_i > b\right) \leq \epsilon.$$

For many problems, $\|x\|_1 < n^{1/2}$ for optimal (or feasible) vectors x (network design, assignment, ...)
$\Rightarrow \Gamma > n^{1/2}$ already for $\epsilon = 0.5$
$\Rightarrow$ for these problems, protecting with probability 0.5 yields protection with probability 0!
$\Rightarrow$ overprotection !

Multifunctions

It is easy to see that the bound from Bertsimas and Sim can be adapted to

$$P\left(\sum \tilde{a}_i x_i > b\right) \leq \exp\left(-\frac{\Gamma^2}{2\|x\|_1}\right).$$

Multifunctions

It is easy to see that the bound from Bertsimas and Sim can be adapted to

$$P\left(\sum \tilde{a}_i x_i > b\right) \leq \exp\left(-\frac{\Gamma^2}{2\|x\|_1}\right).$$

- Γ can be reduced when x is small
- Let's use multifunctions !

Multifunctions

It is easy to see that the bound from Bertsimas and Sim can be adapted to

$$P\left(\sum \tilde{a}_i x_i > b\right) \leq \exp\left(-\frac{\Gamma^2}{2\|x\|_1}\right).$$

- Γ can be reduced when x is small
- Let's use multifunctions !

Define

$$\alpha_\epsilon(x) = \left(-2\ln(\epsilon)\|x\|_1\right)^{1/2}.$$

Consider x^* be given. If

$$\sum a_i x_i^* \leq b \quad \forall a \in \mathbf{U}^{\alpha_\epsilon(x^*)} \qquad ,$$

then

$$P\left(\sum \tilde{a}_i x_i^* > b\right) \leq \exp\left(-\frac{\alpha_\epsilon(x^*)^2}{2\|x^*\|_1}\right) = \epsilon.$$

New robust model

Let $\gamma : \{0,1\}^n \to \mathbb{R}_+$ be a non-negative function.

$$\mathcal{U}^\gamma(x) := \left\{ a \in \mathbb{R}^n : a_i = \overline{a}_i + \delta_i \hat{a}_i, -1 \leq \delta_i \leq 1, \sum |\delta_i| \leq \gamma(x) \right\}.$$

New robust model

Let $\gamma : \{0,1\}^n \to \mathbb{R}_+$ be a non-negative function.

$$\mathcal{U}^{\gamma}(x) := \left\{ a \in \mathbb{R}^n : a_i = \overline{a}_i + \delta_i \hat{a}_i, -1 \leq \delta_i \leq 1, \sum |\delta_i| \leq \gamma(x) \right\}.$$

We have shown that the new model

$$\sum a_i x_i \leq b \quad \forall a \in \mathcal{U}^{\alpha_\epsilon}(x),$$

should be considered instead of the classical model

$$\sum a_i x_i \leq b \quad \forall a \in \mathbf{U}^{\Gamma}.$$

Better bound

The previous bound is bad. Bertsimas and Sim propose a better bound:

$$P\left(\sum a_i x_i^* > b\right) \leq B(n,\Gamma) = \frac{1}{2^n}\left((1-\mu)\binom{n}{\lfloor\nu\rfloor} + \sum_{l=\lfloor\nu\rfloor+1}^{n}\binom{n}{l}\right),$$

where $\nu = (\Gamma + n)/2$, $\mu = \nu - \lfloor\nu\rfloor$.

- We can make this bound dependent on x by considering $B(\|x\|_1, \Gamma)$.
- $\beta_\epsilon(x)$ is the solution of the equation

$$B(\|x^*\|_1, \Gamma) - \epsilon = 0 \tag{3}$$

  in variable Γ.
- We solve (3) numerically.

$$\begin{aligned} \max \quad & \sum_{i=1}^{n} c_i x_i \\ \text{s.t.} \quad & \sum_{i=1}^{n} a_i x_i \leq b, a \in \mathcal{U}^{\gamma} \\ & x \in \{0,1\}^n, \end{aligned}$$

which can be rewritten as

$$\begin{aligned} \max \quad & \sum_{i=1}^{n} c_i x_i \\ \text{s.t.} \quad & \sum_{i=1}^{n} \overline{a}_i x_i + \max_{\substack{0 \leq \delta_i \leq 1 \\ \sum \delta_i \leq \gamma(x)}} \sum_{i=1}^{n} \delta_i \hat{a}_i x_i \leq b, \\ & x \in \{0,1\}^n, \end{aligned}$$

Knapsack problem

Using the dualization approach:

$$
\begin{aligned}
\max \quad & \sum_{i=1}^{n} c_i x_i \\
\text{s.t.} \quad & \sum_{i=1}^{n} \bar{a}_i x_i + z\gamma(x) + \sum_{i=1}^{n} p_i \leq b, \\
& z + p_i \geq \hat{a}_i x_i, \quad i = 1, \dots, n \\
& z, p \geq 0, \\
& x \in \{0,1\}^n.
\end{aligned}
$$

- Non-convex reformulation.
- x binary may help.

Dualization

Theorem

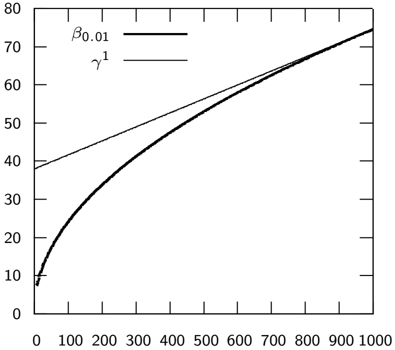
Consider robust constraint

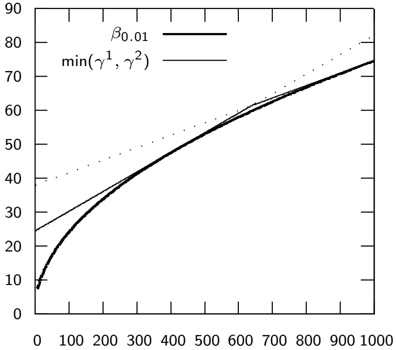
$$\begin{aligned} a^T x &\leq b, \qquad \forall a \in \mathcal{U}^\gamma(x), \\ x &\in \{0,1\}^n, \end{aligned} \tag{4}$$

and suppose that $\gamma = \gamma_0 + \sum \gamma_i x_i$ *is an affine function of* x*, non-negative for* $x \in \{0,1\}^n$*. Then,* (4) *is equivalent to*

$$\begin{aligned} \sum_{i=1}^n \overline{a}_i x_i + \gamma_0 z + \sum_{i=1}^n \gamma_i w_i + \sum_{i=1}^n p_i &\leq b \\ z + p_i &\geq \hat{a}_i x_i, \quad i = 1, \ldots, n, \\ w_i - z &\geq -\max_j(\hat{a}_j)(1 - x_i), \quad i = 1, \ldots, n, \\ p, w, z \geq 0, \quad & x \in \{0,1\}^n. \end{aligned}$$

Non-affine functions γ





Computational results

Objective

1. Is there a benefit in using $\mathcal{U}^{\beta}$ instead of $\mathbf{U}^{\Gamma}$?
2. Computational "complexity" of solving the robust counterparts.

Computational results

Objective

1. Is there a benefit in using $\mathcal{U}^{\beta}$ instead of $\mathbf{U}^{\Gamma}$?
2. Computational "complexity" of solving the robust counterparts.

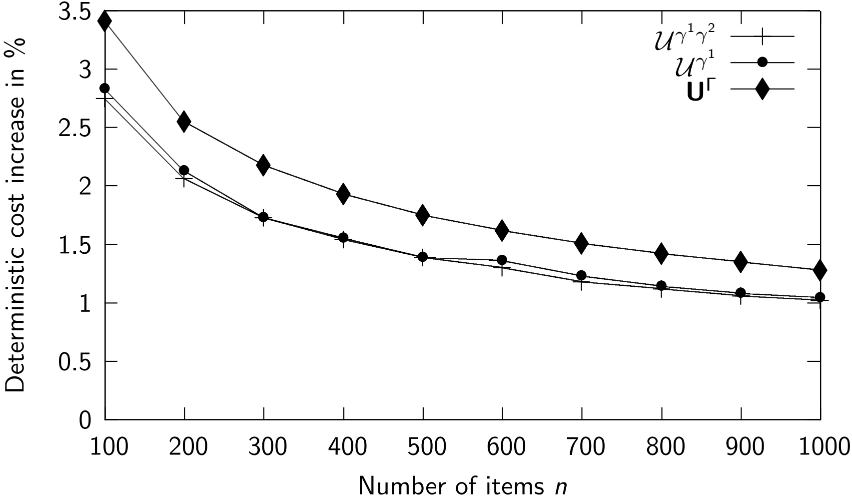
Models

We compare the following at $\epsilon = 0.01$:

- $\mathbf{U}^{\Gamma}$ The classical robust model with budget uncertainty.
- $\mathcal{U}^{\gamma^1}$ Our new model with variable budget uncertainty: γ^1 overapproximates β.
- $\mathcal{U}^{\gamma^1\gamma^2}$ Our new model with variable budget uncertainty: $\min(\gamma^1, \gamma^2)$ over-approximates β.

The price of robustness at $\epsilon = 0.01$

Instances from Bertsimas and Sim (2004)



Computational complexity

model	$\mathcal{U}^{\gamma^1}$	$\mathcal{U}^{\gamma^1\gamma^2}$	$\mathcal{U}^{\gamma^1\gamma^2\gamma^3}$
time model/time $\mathbf{U}^\Gamma$	1.7	3.4	6.1
gap model/gap $\mathbf{U}^\Gamma$	0.87	0.98	1.1

- Fixing M to $\max_j(\hat{a}_j)$ affects the LP relaxation.
- If $M = 1000$, gap $\mathcal{U}^{\gamma^1}$/gap $\mathbf{U}^\Gamma \rightarrow 3.9$!

Outline

Cost uncertainty

Suppose that only cost coefficient are uncertain

$$\begin{aligned} \min \max_{c \in \mathbf{U}} \quad & \sum_{i=1}^{n} c_i x_i \\ \text{s.t.} \quad & \sum_{j=1}^{n} a_{ij} x_j \leq b_i, \quad i = 1, \ldots, m \\ & x \in \{0,1\}^n, \end{aligned}$$

which can be rewritten

$$CO^\Gamma \equiv \min_{x \in X} \max_{c \in \mathbf{U}^\Gamma} c^T x.$$

The previous probabilistic approximation leads to a relation between CO^Γ and

$$\min_{x \in X} \text{VaR}_\epsilon \, c^T x.$$

Value-at-risk

Definition: $\text{VaR}_\epsilon(c^T x) = \inf\{t | P(c^T x \leq t) \geq 1 - \epsilon\}$.
We see easily that

- CO^Γ provides an upper bound of the optimization of VaR

Value-at-risk

Definition: $\text{VaR}_\epsilon(c^T x) = \inf\{t | P(c^T x \leq t) \geq 1 - \epsilon\}$.
We see easily that

- CO^Γ provides an upper bound of the optimization of VaR
- The upper bound is very bad for small cardinality vectors

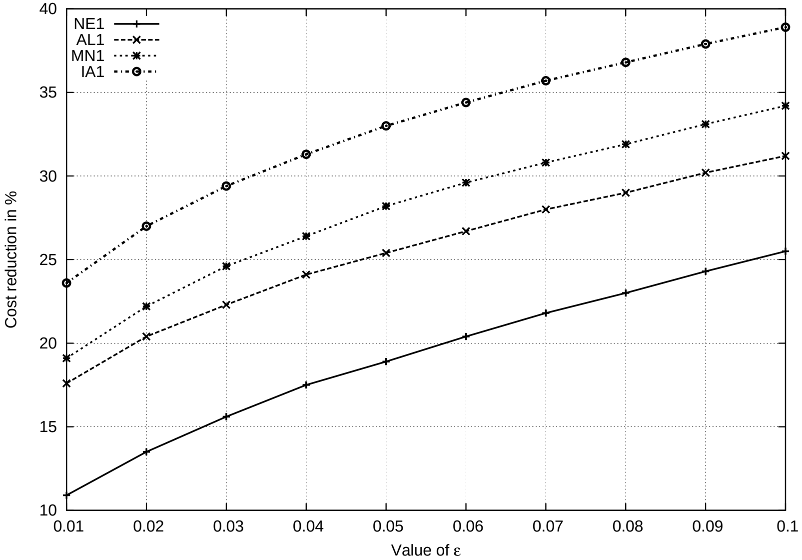
Value-at-risk

Definition: $\text{VaR}_\epsilon(c^T x) = \inf\{t | P(c^T x \leq t) \geq 1 - \epsilon\}$.
We see easily that

- CO^Γ provides an upper bound of the optimization of VaR
- The upper bound is very bad for small cardinality vectors
- Model CO^γ overcomes this flaw

$$CO^\gamma \equiv \min_{x \in X} \max_{c \in \mathcal{U}^\gamma} c^T x.$$

Shortest path problem



Complexity of the resulting problems

Theorem

When γ is an affine function, CO^{γ} can be solved by solving $n+1$ problems CO and taking the cheapest optimal solution.

Complexity of the resulting problems

Theorem

When γ is an affine function, CO^{γ} can be solved by solving $n+1$ problems CO and taking the cheapest optimal solution.

Theorem

When γ is a non-decreasing function of $\|x\|_1$, CO^{γ} can be solved by solving n cardinality constrained problems CO^{Γ} and taking the cheapest optimal solution.

Dynamic Programming

We use the notation $\Gamma' = \min(n, \max_{k=0,\dots,n} \gamma(k))$.

Theorem

Consider a combinatorial optimization problem that can be solved in $O(\tau)$ by using dynamic programming. If $\gamma(k) \in \mathcal{Z}$ for each $k = 0, \dots, n$, then CO^γ can be solved in $O(n\Gamma'\tau)$. Otherwise, CO^γ can be solved in $O(n^2\Gamma'\tau)$.

Dynamic Programming

We use the notation $\Gamma' = \min(n, \max_{k=0,\dots,n} \gamma(k))$.

Theorem

Consider a combinatorial optimization problem that can be solved in $O(\tau)$ by using dynamic programming. If $\gamma(k) \in \mathcal{Z}$ for each $k = 0, \dots, n$, then CO^γ can be solved in $O(n\Gamma'\tau)$. Otherwise, CO^γ can be solved in $O(n^2\Gamma'\tau)$.

Theorem

Consider a combinatorial optimization problem that can be solved in $O(\tau)$ by using dynamic programming. Then, CO^Γ can be solved in $O(\Gamma\tau)$.

If $\Gamma \sim n^{1/2}$, we get $O(n^{1/2}\tau)$, improving over the $O(n\tau)$ from Bertsimas and Sim.

Concluding remarks

- We introduce a new class of uncertainty models.
- They correct the flaw of Bertsimas and Sim model.
- The tractability of the new model is often comparable (or equal) to the traditional model.

Remark: The model can be extended to non-combinatorial problems but tractability becomes an issue.

Thank's for your attention