École MACS 2019

Notions sur la robustesse

A. Aubry, C. Bloch, S. Himmiche
D. Lemoine, P. Marange, S. Norre

Soit le programme linéaire en variables mixtes suivant :

$$\mathcal{P} \left| \begin{array}{ll} \max\limits_{x} & \langle c, x \rangle \\ & \sum\limits_{j=1}^{n} a_{ij} x_j \leq b_i, i = 1, \cdots, m. \\ & x_j \in \mathbb{N}, j = 1, \cdots, p, \\ & x_j \in \mathbb{R}_+, j = p+1, \cdots, n. \end{array} \right.$$

On suppose que les incertitudes portent uniquement sur les coefficients a_{ij} . Plus précisémment, nous supposons que chaque coefficient a_{ij} prend ses valeurs dans un **intervalle borné** de la forme $[\bar{a}_{ij} - \hat{a}_{ij}, \bar{a}_{ij} + \hat{a}_{ij}]$, autrement dit, il existe une variable aléatoire ζ_{ij} à valeur dans $[-1, 1]$ tel que

$$a_{ij} = \bar{a}_{ij} + \zeta_{ij} \hat{a}_{ij}.$$

On peut donc reformuler le programme linéaire de la façon suivante :

$$\mathcal{P} \left| \begin{array}{ll} \max\limits_{x} & \langle c, x \rangle \\ & \sum\limits_{j=1}^{n} \bar{a}_{ij} x_j + \sum\limits_{j=1}^{n} \zeta_{ij} \hat{a}_{ij} x_j \leq b_i, i = 1, \cdots, m. \\ & x_j \in \mathbb{N}, j = 1, \cdots, p, \\ & x_j \in \mathbb{R}_+, j = p+1, \cdots, n. \end{array} \right.$$

L'idée est donc de se « protéger raisonnablement » de l'incertitude modélisée par la quantité $\sum\limits_{i=1}^{n} \zeta_{ij} \hat{a}_{ij} x_j$: on va donc chercher à se prémunir de ce risque grâce à une « fonction déterministe de protection » $\beta_i(x)$ telle que résoudre

$$\mathcal{P}' \left| \begin{array}{ll} \max\limits_{x} & \langle c, x \rangle \\ & \sum\limits_{j=1}^{n} \bar{a}_{ij} x_j + \beta_i(x) \leq b_i, i = 1, \cdots, m. \\ & x_j \in \mathbb{N}, j = 1, \cdots, p, \\ & x_j \in \mathbb{R}_+, j = p+1, \cdots, n. \end{array} \right.$$

donnera une solution acceptable pour le programme linéaire $\mathcal{P}$ sous une certaine acceptation du risque engendré par l'incertitude sur les données.

La suite du document présente les trois approches classiques pour déterminer de telles fonctions.

1 L'approche de Soyster

Le point de vue adopté dans cette méthode est de modéliser le pire des cas [1]. En reprenant les notations précédentes, l'idée est donc de poser

$$\beta_i(x) = \sum_{j=1}^{n} \hat{a}_{ij} x_j, i = 1, \cdots, m.$$

On résoud donc le modèle :

$$\mathcal{P}' \left| \begin{array}{ll} \max\limits_{x} & \langle c, x \rangle \\ & \sum\limits_{j=1}^{n} (\bar{a}_{ij} + \hat{a}_{ij})\, x_j \leq b_i, i = 1, \cdots, m. \\ & x_j \in \mathbb{N}, j = 1, \cdots, p, \\ & x_j \in \mathbb{R}_+, j = p+1, \cdots, n. \end{array} \right.$$

Les avantages d'une telle approche sont :

— une grande facilité de mise en œuvre,

— on est sûr que la solution obtenue sera toujours faisable pour le programme linéaire $\mathcal{P}$, quelque soit la réalisation de a_{ij}.

Cependant, il y a un inconvénient majeur, inconvénient dû à l'hyperconservatisme de cette approche : on risque de donner une solution très dégradée au programme linéaire $\mathcal{P}$.

Pour synthétiser, cette approche offre une couverture maximale du risque au détriment de la qualité de la solution obtenue.

Les deux approches suivantes tentent de pallier le problème en offrant néanmoins une garantie de performance : **pour cela, on supposera désormais que les variables aléatoires ζ_{ij} sont indépendantes et symétriquement distribuées dans** $[-1, 1]$.

2 Les approches « en ligne »

L'idée de ces approches est de borner, pour chaque contrainte i, l'ensemble des variations dûes aux incertitudes par un paramètre spécifique à la contrainte (d'où le vocable d'en ligne). Dans la suite nous noterons $\zeta_i = (\zeta_{ij})_{j=1,\cdots,n}$ le vecteur aléatoire d'incertitudes lié à la contrainte i et c'est donc ce vecteur qu'on va borner en norme.

Dans l'approche de Ben-Tal et Nemirovski [2], on considère la norme euclidienne de $\mathbb{R}^n$: $\|x\|_2 = \sqrt{x_1^2 + \cdots + x_n^2}$. Dans l'approche de Bertsimas et

Sim [3], c'est la norme $\|x\|_1 = \sum_{i=1}^{n} |x_i|$ qui est prise en compte.

On pose donc

$$\beta_i\left(x\right) = \max_{\|\zeta_i\|_l \leq \Omega_i} \left(\sum_{j=1}^{n} \hat{a}_{ij} \zeta_{ij} x_j \right), i = 1, \cdots, m$$

où $l = 1, 2$ selon la norme utilisée.

$$\mathcal{P}_l \left| \begin{array}{ll} \max_{x} & \langle c, x \rangle \\ & \sum_{j=1}^{n} \bar{a}_{ij} x_j + \max_{\|\zeta_i\|_l \leq \Omega_i} \left(\sum_{j=1}^{n} \hat{a}_{ij} \zeta_{ij} x_j \right) \leq b_i, i = 1, \cdots, m. \\ & x_j \in \mathbb{N}, j = 1, \cdots, p, \\ & x_j \in \mathbb{R}_+, j = p+1, \cdots, n. \end{array} \right.$$

2.1 L'approche de Ben-Tal et Nemirovski

Dans cette approche, les auteurs utilisent donc la norme $\|\|_2$ (la norme euclidienne) : son utilisation aboutit à l'élaboration d'un modèle quadratique, aussi nous ne l'aborderons pas dans le cadre de ce document. Le lecteur pourra se référer à [4].

2.2 L'approche de Bertsimas et Sim

Dans cette approche, les auteurs utilisent donc la norme $\|\|_1$. Considérons un vecteur x^* de $\mathbb{R}_+^n$ et regardons $\beta_i\left(x^*\right)$.

Trouver $\beta_i\left(x^*\right)$ revient à résoudre le programme linéaire suivant :

$$\begin{array}{lll} \beta_i\left(x^*\right) = & \max & \sum_{j=1}^{n} \left(\hat{a}_{ij} x_j^*\right) \zeta_{ij} \\ & & \sum_{j=1}^{n} \zeta_{ij} \leq \Omega_i. \\ & & \zeta_{ij} \leq 1, j = 1, \cdots, n. \\ & & \zeta_{ij} \geq 0, j = 1, \cdots, n. \end{array}$$

Sous réserve que $\Omega_i \geq 0$, ce programme linéaire admet une solution optimale. Il est également trivial qu'il existe une solution optimale ζ_i^* possédant la structure suivante, en posant $J = \{1, \cdots, n\}$:

— $\exists J_i^* \subset J$ tel que $\forall j \in J_i^*, \zeta_{ij}^* = 1$ et $\sum_{j \in J_i^*} \zeta_{ij}^* = \lfloor \Omega_i \rfloor$

— $\exists! k_i^* \in J - J_i^*$ tel que $\zeta_{ik_i^*}^* = \Omega_i - \lfloor \Omega_i \rfloor$

— $\forall j \in J - \left(J_i^* \cup \{k_i^*\}\right), \zeta_{ij}^* = 0$

Nous noterons par la suite $S_i^* = J_i^* \cup \{k_i^*\}$ l'ensemble associé à cette solution optimale ζ_i^*.

Comme le programme linéaire précédent admet une solution optimale sous réserve que $\Omega_i \geq 0$, il en est de même pour son dual et la valeur des deux fonctions objectifs coincident. Le dual s'écrit donc

$$\begin{array}{lll} \beta_i\left(x^*\right) = & \min & \displaystyle\sum_{j=1}^{n} p_{ij} + z_i \Omega_i \\ & & z_i + p_{ij} \geq \hat{a}_{ij} x_j^*, j = 1, \cdots, n. \\ & & p_{ij} \geq 0, j = 1, \cdots, n. \\ & & z_i \geq 0. \end{array}$$

En substituant dans $\mathcal{P}_1$, on obtient le programme linéaire équivalent suivant :

$$\mathcal{P}_1' \left| \begin{array}{ll} \displaystyle\max_x & \langle c, x \rangle \\ & \displaystyle\sum_{j=1}^{n} \bar{a}_{ij} x_j + \sum_{j=1}^{n} p_{ij} + z_i \Omega_i \leq b_i, i = 1, \cdots, m. \\ & z_i + p_{ij} \geq \hat{a}_{ij} x_j^*, i = 1, \cdots, m, j = 1, \cdots, n. \\ & p_{ij} \geq 0, i = 1, \cdots, m, j = 1, \cdots, n. \\ & z_i \geq 0. \\ & x_j \in \mathbb{N}, j = 1, \cdots, p, \\ & x_j \in \mathbb{R}_+, j = p+1, \cdots, n. \end{array} \right.$$

qu'on peut alors résoudre avec un solveur commercial.

Reste alors à qualifier la performance de la solution optimale obtenue.

Théorème 1

Soit x^* une solution optimale du programme linéaire $\mathcal{P}_1$. En reprenant les notations précédentes et sous réserve que les variables aléatoires ζ_{ij} soient indépendantes et symétriquement distribuées dans $[-1, 1]$, alors

$$\mathbb{P}\left(\sum_{j=1}^{n} \bar{a}_{ij} x_j^* + \sum_{j=1}^{n} \zeta_{ij} \hat{a}_{ij} x_j^* > b_i\right) \leq \exp\left(-\frac{{\Omega_i}^2}{2n}\right)$$

Démonstration.

Dans la démonstration, nous aurons besoin du résultat suivant, qui est un corollaire de l'inégalité de Markov :

Proposition 1 (corollaire de l'inégalité de Markov)

Soit ϕ une fonction croissante et positive ou nulle sur un intervalle I. Soit Y une variable aléatoire réelle définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ telle que $\mathbb{P}(Y \in I) = 1$. Alors $\forall b \in I$, tel que $\phi(b) > 0$, $\mathbb{P}(Y \geq b) \leq \dfrac{\mathbb{E}[\phi(Y)]}{\phi(b)}$

$$
\begin{array}{lcl}
\mathbb{P}\left(\sum_{j=1}^{n} \bar{a}_{ij}x_j^* + \sum_{j=1}^{n} \zeta_{ij}\hat{a}_{ij}x_j^* > b_i\right) & = & \mathbb{P}\left(\sum_{j=1}^{n} \zeta_{ij}\hat{a}_{ij}x_j^* > \max_{\|\zeta_i\|_1 \leq \Omega_i} \left(\sum_{j=1}^{n} \hat{a}_{ij}\zeta_{ij}x_j\right)\right) \\
 & = & \mathbb{P}\left(\sum_{j=1}^{n} \zeta_{ij}\hat{a}_{ij}x_j^* > \sum_{j \in J_i^*} \hat{a}_{ij}x_j^* + \hat{a}_{ik_i^*}x_{k_i^*}^* \left(\Omega_i - \lfloor \Omega_i \rfloor\right)\right) \\
r = arg \min_{j \in S_i^*} \hat{a}_{ij}x_j^* & \leq & \mathbb{P}\left(\sum_{j=1}^{n} \zeta_{ij}\hat{a}_{ij}x_j^* > \hat{a}_{ir}x_r^* \left(\sum_{j \in J_i^*} 1 + \left(\Omega_i - \lfloor \Omega_i \rfloor\right)\right)\right) \\
 & \leq & \mathbb{P}\left(\sum_{j=1}^{n} \zeta_{ij}\hat{a}_{ij}x_j^* > \hat{a}_{ir}x_r^*\Omega_i\right) \\
\gamma_{ij} = \dfrac{\hat{a}_{ij}x_j^*}{\hat{a}_{ir}x_r^*} & \leq & \mathbb{P}\left(\sum_{j=1}^{n} \zeta_{ij}\gamma_{ij} \geq \Omega_i\right)
\end{array}
$$

Posons $\theta = \dfrac{\Omega_i}{n}$ et considérons la fonction $exp\,(\theta x)$ qui est croissante sur $\mathbb{R}$. On peut donc écrire, d'après le corollaire de l'inégalité de Markov :

$$
\begin{array}{lll}
\mathbb{P}\left(\sum_{j=1}^{n}\zeta_{ij}\gamma_{ij}\geq\Omega_i\right) & \leq & \dfrac{\mathbb{E}\left[exp\left(\theta\sum_{j=1}^{n}\zeta_{ij}\gamma_{ij}\right)\right]}{exp\left(\theta\Omega_i\right)}\\
 & \leq & \dfrac{\mathbb{E}\left[exp\left(\sum_{j=1}^{n}\theta\zeta_{ij}\gamma_{ij}\right)\right]}{exp\left(\theta\Omega_i\right)}\\
 & \leq & \dfrac{\mathbb{E}\left[\prod_{j=1}^{n}exp\left(\theta\zeta_{ij}\gamma_{ij}\right)\right]}{exp\left(\theta\Omega_i\right)}\\
\text{indépendance des } \zeta_{ij} & \leq & \prod_{j=1}^{n}\dfrac{\mathbb{E}\left[exp\left(\theta\zeta_{ij}\gamma_{ij}\right)\right]}{exp\left(\theta\Omega_i\right)}\\
 & \leq & \prod_{j=1}^{n}\dfrac{1}{exp\left(\theta\Omega_i\right)}\int_{\Omega}exp\left(\theta\gamma_{ij}\zeta_{ij}\left(\omega\right)\right)d\mathbb{P}(\omega)\\
 & \leq & \prod_{j=1}^{n}\dfrac{1}{exp\left(\theta\Omega_i\right)}\int_{-1}^{1}exp\left(\theta\gamma_{ij}x\right)d\mathbb{P}_{\zeta_{ij}}(x)\\
 & \leq & \prod_{j=1}^{n}\dfrac{1}{exp\left(\theta\Omega_i\right)}\int_{-1}^{1}\left(ch\left(\theta\gamma_{ij}x\right)+sh\left(\theta\gamma_{ij}x\right)\right)d\mathbb{P}_{\zeta_{ij}}(x)\\
\text{distribution symétrique} & \leq & \prod_{j=1}^{n}\dfrac{1}{exp\left(\theta\Omega_i\right)}2\int_{0}^{1}ch\left(\theta\gamma_{ij}x\right)d\mathbb{P}_{\zeta_{ij}}(x)\\
\text{ch est croissante sur } [0,1] & \leq & \prod_{j=1}^{n}\dfrac{ch\left(\theta\gamma_{ij}\right)}{exp\left(\theta\Omega_i\right)}2\int_{0}^{1}d\mathbb{P}_{\zeta_{ij}}(x)\\
\text{distribution symétrique} & \leq & \prod_{j=1}^{n}\dfrac{ch\left(\theta\gamma_{ij}\right)}{exp\left(\theta\Omega_i\right)}\\
 & \leq & \prod_{j=1}^{n}\dfrac{1}{exp\left(\theta\Omega_i\right)}\sum_{k=0}^{\infty}\dfrac{\left(\theta\gamma_{ij}\right)^{2k}}{(2k)!}
\end{array}
$$

$$
\begin{array}{lll}
\text{Or } (2k)! & = & \left(2\times4\times\cdots\times2k\right)\left(1\times3\times\cdots\times(2k-1)\right)\\
 & = & 2^k\left(1\times2\times\cdots\times k\right)\left(1\times3\times\cdots\times(2k-1)\right)\\
 & = & 2^k k!\left(1\times3\times\cdots\times(2k-1)\right)\\
 & \geq & 2^k k!
\end{array}
$$

$$
\begin{array}{lll}
\text{donc } \dfrac{\left(\theta\gamma_{ij}\right)^{2k}}{(2k)!} & \leq & \dfrac{\left(\theta\gamma_{ij}\right)^{2k}}{2^k k!}\\
 & \leq & \dfrac{\left(\frac{\theta^2\gamma_{ij}^2}{2}\right)^k}{k!}
\end{array}
$$

par suite, $ch\left(\theta\gamma_{ij}\right)\leq exp\left(\dfrac{\theta^2\gamma_{ij}^2}{2}\right)$. On obtient donc que

$$\begin{aligned}
\mathbb{P}\left(\sum_{j=1}^{n}\zeta_{ij}\gamma_{ij}\geq\Omega_i\right) &\leq \prod_{j=1}^{n}\frac{exp\left(\frac{\theta^2\gamma_{ij}^2}{2}\right)}{exp\left(\theta\Omega_i\right)}\\
&\leq \frac{exp\left(\sum_{j=1}^{n}\frac{\theta^2\gamma_{ij}^2}{2}\right)}{exp\left(\theta\Omega_i\right)}\\
&\leq exp\left(\sum_{j=1}^{n}\frac{\theta^2\gamma_{ij}^2}{2}-\theta\Omega_i\right)\\
&\leq exp\left(\sum_{j=1}^{n}\frac{\theta^2}{2}-\theta\Omega_i\right)
\end{aligned}$$

car par optimalité de $\beta_i\left(x^*\right)$, nous avons $\gamma_{ij}\leq 1$. Ainsi,

$$\begin{aligned}
\mathbb{P}\left(\sum_{j=1}^{n}\zeta_{ij}\gamma_{ij}\geq\Omega_i\right) &\leq exp\left(\sum_{j=1}^{n}\frac{\theta^2}{2}-\theta\Omega_i\right)\\
&\leq exp\left(n\frac{\theta^2}{2}-\theta\Omega_i\right)\\
\text{et } \theta=\frac{\Omega_i}{n} \qquad &\leq exp\left(-\frac{\Omega_i^2}{n}\right)
\end{aligned}$$

$\square$

Bibliographie

[1] A. L. Soyster. Technical Note : Convex Programming with Set-Inclusive Constraints and Applications to Inexact Linear Programming. *Operations Research*, 21(5) :1154–1157, October 1973.

[2] Aharon Ben-Tal and Arkadi Nemirovski. Robust solutions of Linear Programming problems contaminated with uncertain data. *Mathematical Programming*, 88(3) :411–424, September 2000.

[3] Dimitris Bertsimas and Melvyn Sim. The Price of Robustness. *Operations Research*, 52(1) :35–53, February 2004.

[4] J.-P. Vial F. Babonneau and R. Apparigliato. Robust optimization for environmental and energy planning. *Uncertainty and Environmental Decision Making, Springer's International Series in Operations Research and Management Science*, pages 79–126, 2010.