Décomposition par prédiction

08 décembre 2020

Rappel du cadre de la décomposition...

On étudie un problème d'optimisation classique :

$$\min_{u \in U^{\mathrm{ad}} \subset \mathcal{U}} J(u) \quad \text{sous la contrainte} \quad \Theta(u) = \theta \in \mathcal{V} \ ,$$

mais qui présente les caractéristiques suivantes :

1. $\mathcal{U}$ est un produit cartésien d'espaces $\mathcal{U}_1 \times \cdots \times \mathcal{U}_N$,
2. tel que $U^{\mathrm{ad}} = U_1^{\mathrm{ad}} \times \cdots \times U_N^{\mathrm{ad}}$ avec $U_i^{\mathrm{ad}} \subset \mathcal{U}_i$,
3. les fonctions J et Θ sont additives suivant ces espaces :

$$J(u_1, \ldots, u_N) = \sum_{i=1}^{N} J_i(u_i) \quad , \quad \Theta(u_1, \ldots, u_N) = \sum_{i=1}^{N} \Theta_i(u_i) \ .$$

Le point 1 fournit la trame de la décomposition, alors que les points 2 et 3 caractérisent la structure additive du problème.

...et des méthodes par les prix et les quantités

- Le principe de la méthode de décomposition par les prix est de former le Lagrangien du problème et d'appliquer l'algorithme d'Uzawa. La structure additive du problème fait que l'étape de minimisation en u à $p = p^{(k)}$ fixé se décompose.

  La contrainte initiale $\sum_i \Theta_i(u_i) = \theta$ est dualisée et n'apparaît plus en tant que contrainte dans les sous-problèmes.

- Le principe de la méthode de décomposition par les quantités est d'ajouter des variables v_i et d'écrire la contrainte initiale sous la forme $\Theta_i(u_i) - v_i = 0 \ \forall i$, avec $\sum_i v_i - \theta = 0$, de telle sorte que la minimisation en u à $v = v^{(k)}$ fixé se décompose.

  La contrainte initiale $\sum_i \Theta_i(u_i) - \theta = 0$ apparaît alors dans chacun des sous-problèmes sous la forme : $\Theta_i(u_i) - v_i^{(k)} = 0$.

Plan du cours

1. Décomposition par prédiction
   - Principe de la méthode par prédiction
   - Les questions et leurs réponses
   - Analyse de la méthode

2. Comparaison des méthodes et conclusions
   - Comparaison des 3 méthodes de décomposition
   - Conclusion sur la première partie du cours

3. Travaux dirigés sur la décomposition par les quantités
   - Le cas des contraintes inégalités
   - Formes spéciales d'allocation de ressources
   - Réseau de distribution d'eau

Idée de base de la décomposition par prédiction

Pour pallier les inconvénients des méthodes de décomposition par les prix et par les quantités, on va introduire la méthode de décomposition par prédiction, qui fera en sorte que chaque sous-problème ne voit apparaître qu'une partie des contraintes. Autrement dit, on va partitionner l'ensemble des contraintes entre les sous-problèmes.

Pour cela, on ajoute un ingrédient dans la structure du problème, à savoir qu'il faut, pour chacune des contraintes du problème, **choisir** à quel sous-problème elle est rattachée. Ce **choix** est caractéristique de la méthode de décomposition par prédiction, et c'est ce choix qui induira les qualités (ou les inconvénients. . .) de la méthode de décomposition par prédiction.

Mécanisme d'affectation des contraintes

Mécanisme de base. La contrainte $\Theta : \mathcal{U} \rightarrow \mathcal{V} = \mathbb{R}^m$ se compose de m contraintes scalaires $\Theta_j : \mathcal{U} \rightarrow \mathbb{R}$, et on choisit pour chaque contrainte scalaire j_0 le sous-problème i_0 auquel sera affectée cette contrainte. Dans ce but, on isole la partie de la contrainte j_0 qui dépend de u_{i_0} et qui sera traitée par le sous-problème i_0 :

$$\Theta_{j_0}(u) = \Theta_{j_0,i_0}(u_{i_0}) + \sum_{i \neq i_0} \Theta_{j_0,i}(u_i) \; .$$

Mécanisme général. On partitionne l'espace $\mathcal{V}$ d'arrivée des contraintes en $\mathcal{V}_1 \times \cdots \times \mathcal{V}_N$, le sous-espace $\mathcal{V}_i$ correspondant aux indices des contraintes qui sont affectées au sous-problème i. Dans ces contraintes $\Theta^i : \mathcal{U} \rightarrow \mathcal{V}_i$, on isole la partie qui dépend de u_i et qui sera traitée par le sous-problème i :

$$\Theta^i(u) = \Theta^i_i(u_i) + \sum_{j \neq i} \Theta^i_j(u_j) \; .$$

Le cas d'une contrainte scalaire I

On commence par le cas où la contrainte du problème est scalaire, et on **choisit** d'affecter cette contrainte au sous-problème 1 :

$$\Theta_1(u_1) + \sum_{i=2}^{N} \Theta_i(u_i) - \theta = 0 \in \mathbb{R} ,$$

- Le sous-problème 1, qui « voit » la contrainte, s'écrit :

$$\min_{u_1 \in U_1^{\mathrm{ad}}} J_1(u_1) \quad \text{sous} \quad \Theta_1(u_1) - v^{(k)} = 0 ,$$

  où $v^{(k)}$ est une production à déterminer.

- Les sous-problèmes $2, \dots, N$ « ne voient pas » la contrainte, mais on y ajoute un terme d'incitation à produire :

$$\min_{u_i \in U_i^{\mathrm{ad}}} J_i(u_i) + \left\langle p^{(k)} , \Theta_i(u_i) \right\rangle ,$$

  où $p^{(k)}$ est un prix à déterminer.

Le cas d'une contrainte scalaire II

Le sous-problème 1 est paramétré par la variable $v^{(k)}$: on note $u_1^{k+1} \in \widetilde{U}_1(v^{(k)})$ une solution et $\lambda_1^{(k+1)}$ un multiplicateur associé.

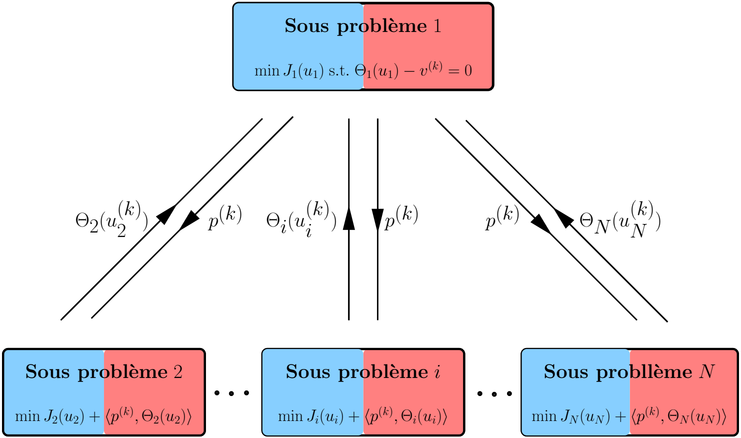
Les sous-problèmes $2, \ldots, N$ sont paramétrés par la variable $p^{(k)}$: leur résolution fournit des solutions $u_i^{k+1} \in \widehat{U}_i(p^{(k)})$, $i = 2, \ldots, N$.

Pour faire évoluer les variables $(v^{(k)}, p^{(k)})$, on procède comme suit.

- la différence entre la demande θ et la somme des productions des unités 2 à N est un bon candidat pour la production v :
$$v^{(k+1)} = \theta - \sum_{i=2}^{N} \Theta_i(u_i^{(k+1)}) \ .$$
- le multiplicateur $\lambda_1^{(k+1)}$ représente la sensibilité du coût de l'unité 1 et est un bon candidat pour le prix p :
$$p^{(k+1)} = \lambda_1^{(k+1)} \ .$$

Le cas d'une contrainte scalaire III

On obtient le schéma de décomposition/coordination suivant :



où chaque unité tient lieu de coordination pour d'autres unités...

Questions sur la méthode de décomposition

1. **Bien-fondé de la méthode**

   Supposons que l'algorithme converge : il existe un couple $(v^\sharp, p^\sharp)$ tel que $u_1^\sharp \in \widetilde{U}_1(v^\sharp)$ et $\lambda_1^\sharp = p^\sharp$, $u_i^\sharp \in \widehat{U}_i(p^\sharp)$ pour $i \geq 2$ et $v^\sharp = \theta - \sum_{i \geq 2} \Theta_i(u_i^\sharp)$.
   Peut-on dire que $(u_1^\sharp, \dots, u_N^\sharp)$ est solution du problème global ?

2. **Existence du couple d'équilibre**

   À quelles conditions existe t'il un tel couple $(v^\sharp, p^\sharp)$?

3. **Algorithmes de calcul**

   Comment calculer $(v^\sharp, p^\sharp)$?

4. **Non unicité des solutions**

   Que se passe t'il si les ensembles $\widetilde{U}_1(v^\sharp)$ et $\widehat{U}_i(p^\sharp)$, $i = 2, \dots, n$ ne sont pas réduits à un singleton ?

5. **Interruptibilité de la méthode**

   De quel résultat dispose-t-on si l'on interrompt l'algorithme avant qu'il n'ait convergé ?

Réponse aux questions sur le bien-fondé et l'existence

Lemme

Soit $(u_1^\sharp, p^\sharp)$ un point-selle de $\min_{u_1} J_1(u_1)$ sous $\Theta_1(u_1) - v^\sharp = 0$, et soit $u_i^\sharp$ une solution de $\min_{u_i} J_i(u_i) + \langle p^\sharp , \Theta_i(u_i) \rangle$, $i = 2, \dots, N$. Si on suppose de plus que la condition $v^\sharp = \theta - \sum_{i=2}^N \Theta_i(u_i^\sharp)$ est vérifiée, alors $(u_1^\sharp, \dots, u_N^\sharp, p^\sharp)$ est un point-selle du problème global.

Réciproquement, si $(u_1^\sharp, \dots, u_N^\sharp, p^\sharp)$ est un point-selle du problème global, alors $(u_1^\sharp, p^\sharp)$ est un point-selle du sous-problème 1 et les $u_i^\sharp$ sont solutions des sous-problèmes $i = 2, \dots, N$.

Preuve. La preuve consiste à écrire et à sommer les conditions d'optimalité des sous-problèmes, et à spécialiser celle du problème global. $\square$

La réponse à la question 1 est donnée par la partie directe du lemme. La réponse à la question 2 est l'existence d'un point-selle du Lagrangien du problème global.

Réponse à la question sur la non unicité

Tout d'abord, la non unicité de la solution du sous-problème 1 ne pose pas de difficulté car toute solution $u_1^\sharp \in \widetilde{U}_1(v^\sharp)$ vérifie par construction :

$$\Theta_1(u_1^\sharp) = v^\sharp \ ,$$

et donc toute solution $u_1^\sharp$ du sous-problème convient.

Mais si l'un des ensembles $\widehat{U}_i(p^\sharp)$ des sous-problèmes i, pour $i \geq 2$, n'est pas réduit à un point unique, on retrouve les difficultés liées à la non stabilité du Lagrangien : certaines solutions $u_i^\sharp$ prises dans les ensembles $\widehat{U}_i(p^\sharp)$ ne vérifient pas la condition :

$$\sum_{i=2}^{N} \Theta_i(u_i^\sharp) = \theta - \Theta_1(u_1^\sharp) \ ,$$

et l'algorithme ne fournit pas une solution du problème global.

Réponse à la question sur l'interruptibilité

L'ensemble des solutions $(u_1^{(k+1)}, \dots, u_N^{(k+1)})$ obtenu à la fin de l'itération k est telle que :

$$\Theta_1(u_1^{(k+1)}) = v^{(k)} = \theta - \sum_{i=2}^{N} \Theta_i(u_i^{(k)}) \neq \theta - \sum_{i=2}^{N} \Theta_i(u_i^{(k+1)}) \; .$$

et donc ne vérifie pas la contrainte du problème global.

Dans le cas particulier étudié, il est cependant facile d'obtenir une solution admissible que l'on construit à partir des solutions des deux dernières itérations, à savoir $(u_1^{(k+1)}, u_2^{(k)}, \dots, u_N^{(k)})$.

Dans le cas général avec plusieurs contraintes, reconstruire une solution admissible à partir des solutions disponibles à une itération k n'est pas toujours possible. Il faut alors attendre la convergence de l'algorithme pour disposer d'une solution vérifiant toutes les contraintes du problème.

Réponse à la question sur le calcul I

L'itération k de l'algorithme de prédiction se déroule comme suit.

- **Phase de décomposition** :

$$\min_{u_1 \in U_1^{\mathrm{ad}}} J_1(u_1) \ \textit{sous} \ \Theta_1(u_1) - v^{(k)} = 0 \quad \rightsquigarrow (u_1^{(k+1)}, \lambda_1^{(k+1)}) \ ,$$

$$\min_{u_i \in U_i^{\mathrm{ad}}} J_i(u_i) + \left\langle p^{(k)} \, , \Theta_i(u_i) \right\rangle \quad \rightsquigarrow u_i^{(k+1)} \ , \quad i = 2, \ldots, N \ .$$

- **Phase de coordination** :

$$v^{(k+1)} = \theta - \sum_{i=2}^{N} \Theta_i(u_i^{(k+1)}) \ ,$$

$$p^{(k+1)} = \lambda_1^{(k+1)} \ .$$

L'algorithme est donc entièrement spécifié !

On va chercher à mieux comprendre à quoi correspond ce calcul.

Réponse à la question sur le calcul II

L'algorithme de décomposition par prédiction peut se représenter de manière formelle sous la forme suivante :

$$\begin{array}{lll} \text{Sous-problème } 1 & : & v^{(k)} \xrightarrow{\Psi} p^{(k+1)} \, , \\ \text{Sous-problèmes } (2, \ldots, N) & : & p^{(k)} \xrightarrow{\Xi} v^{(k+1)} \, . \end{array}$$

Une itération s'écrit donc

$$\begin{pmatrix} v^{(k+1)} \\ p^{(k+1)} \end{pmatrix} = \begin{pmatrix} 0 & \Xi \\ \Psi & 0 \end{pmatrix} \begin{pmatrix} v^{(k)} \\ p^{(k)} \end{pmatrix} \, ,$$

et l'algorithme de décomposition par prédiction correspond donc à la recherche d'un point-fixe de l'opérateur $\begin{pmatrix} 0 & \Xi \\ \Psi & 0 \end{pmatrix}$:

$$\begin{pmatrix} v^{\sharp} \\ p^{\sharp} \end{pmatrix} = \begin{pmatrix} 0 & \Xi \\ \Psi & 0 \end{pmatrix} \begin{pmatrix} v^{\sharp} \\ p^{\sharp} \end{pmatrix} \, .$$

Réponse à la question sur le calcul III

Contrairement aux deux algorithmes de décomposition vus à la séance précédente (prix et quantités), qui correspondent à des méthodes variationnelles (Uzawa, gradient projeté) pour lesquelles on disposent de théorèmes de convergence classiques, l'algorithme de décomposition par prédiction correspond à une méthode de point-fixe. Un théorème de convergence existe (voir le poly, §4.3, page 100 et suivantes), sous des hypothèses assez restrictives.

Dans le cas étudié ici d'une contrainte scalaire, on a vu que la décomposition par prédiction ne donne pas à chaque itération une solution admissible. Cependant, si on reformule l'algorithme sous forme séquentielle plutôt que parallèle, à savoir :

$$p^{(k)} \xrightarrow{\Xi} v^{(k+1)} \xrightarrow{\Psi} p^{(k+1)} ,$$

la solution de l'itération est admissible pour le problème global. On rappelle que ceci n'est pas généralisable au cas général.

Formalisation mathématique de la prédiction I

Le principe de la méthode de prédiction consiste à remplacer la contrainte initiale par un jeu de deux contraintes équivalentes faisant intervenir une nouvelle variable v :

$$\sum_{i=1}^{N} \Theta_i(u_i) - \theta = 0 \quad \Longleftrightarrow \quad \begin{cases} \Theta_1(u_1) - v = 0 \; , \\ \displaystyle\sum_{i=2}^{N} \Theta_i(u_i) + v - \theta = 0 \; . \end{cases}$$

Cet éclatement de la contrainte correspond au choix que l'on a fait d'isoler dans la contrainte le terme dépendant de u_1.

On choisit alors de traiter la première contrainte en tant que telle, et de dualiser la seconde contrainte à l'aide d'un multiplicateur p. Supposant l'existence d'un point-selle, le problème s'écrit :

$$\min_{(u_1,\ldots,u_N)} \min_{v\in\mathcal{V}} \max_{p\in\mathcal{V}} \left\{ \sum^{N} J_i(u_i) + \left\langle p \, , \sum^{N} \Theta_i(u_i) + v - \theta \right\rangle \text{ sous } \Theta_1(u_1) - v = 0 \right\} .$$

Formalisation mathématique de la prédiction II

$$\max_{p \in \mathcal{V}} \min_{v \in \mathcal{V}} \Big\{ \underbrace{\min_{(u_1,\ldots,u_N)} \sum_{i=1}^{N} J_i(u_i) + \Big\langle p \,, \sum_{i=2}^{N} \Theta_i(u_i) + v - \theta \Big\rangle \ \text{ sous } \ \Theta_1(u_1) - v = 0}_{\mathcal{L}(v,p)} \Big\} \,.$$

À $(v,p) = (v^{(k)}, p^{(k)})$ fixés, le problème de minimisation en $(u_1, \ldots, u_N)$ se décompose en N sous-problèmes indépendants :

- $$\min_{u_1 \in U_1^{\mathrm{ad}}} J_1(u_1) \ \text{ sous } \ \Theta_1(u_1) - v^{(k)} = 0$$

  $\rightsquigarrow$ point-selle $\left(\widetilde{u}_1(v^{(k)}), \widetilde{\lambda}_1(v^{(k)})\right)$,

- $$\min_{u_i \in U_i^{\mathrm{ad}}} J_i(u_i) + \left\langle p^{(k)} \,, \Theta_i(u_i) \right\rangle \,, \quad i = 2, \ldots, N$$

  $\rightsquigarrow$ solution $\widehat{u}_i(p^{(k)})$.

C'est la phase de décomposition de la méthode par prédiction.

Formalisation mathématique de la prédiction III

$$\max_{p\in\mathcal{V}}\min_{v\in\mathcal{V}}\Big\{\underbrace{\min_{(u_1,\ldots,u_N)}\sum_{i=1}^{N}J_i(u_i)+\Big\langle p\,,\sum_{i=2}^{N}\Theta_i(u_i)+v-\theta\Big\rangle \text{ sous } \Theta_1(u_1)-v=0}_{\mathcal{L}(v,p)}\Big\}\,.$$

Il faut ensuite effectuer la maxi-minimisation en (v,p) de $\mathcal{L}$. Pour cela, on dispose des gradients partiels de la fonction marginale $\mathcal{L}$:

$$\nabla_p\mathcal{L}(v,p)=\sum_{i=2}^{N}\Theta_i\big(\widehat{u}_i(p)\big)+v-\theta \quad , \quad \nabla_v\mathcal{L}(v,p)=p-\widetilde{\lambda}_1(v)\,.$$

Partant d'un point $(v^{(k)},p^{(k)})$, la remise à jour « classique » des variables (v,p) se ferait par l'algorithme de Arrow-Hurwicz :

$$v^{(k+1)}=v^{(k)}-\epsilon\nabla_v\mathcal{L}(v^{(k)},p^{(k)})\,,$$
$$p^{(k+1)}=p^{(k)}+\rho\nabla_p\mathcal{L}(v^{(k)},p^{(k)})\,.$$

Mais ce n'est pas ce que fait la méthode par prédiction...

Formalisation mathématique de la prédiction IV

Une autre possibilité est de résoudre les conditions d'optimalité :

$$\nabla_p \mathcal{L}(v, p) = 0 \quad , \quad \nabla_v \mathcal{L}(v, p) = 0 \, .$$

C'est un système implicite que l'on « déboucle » par relaxation : partant d'un point $(v^{(k)}, p^{(k)})$, on résoud :

$$\nabla_p \mathcal{L}(v, p^{(k)}) = 0 \quad , \quad \nabla_v \mathcal{L}(v^{(k)}, p) = 0 \, ,$$

d'où l'on déduit les équations de mise à jour suivantes :

$$v^{(k+1)} = \theta - \sum_{i=2}^{N} \Theta_i\big(\widehat{u}_i(p^{(k)})\big) \, ,$$

$$p^{(k+1)} = \widetilde{\lambda}_1(v^{(k)}) \, .$$

Ce sont celles utilisées dans la méthode par prédiction !

Esquisse du cas général

La formalisation faite sur le cas d'une contrainte scalaire permet de passer au cas général. Considérons une contrainte vectorielle :

$$\sum_{i=1}^{N} \Theta_i(u_i) - \theta = 0 \in \mathbb{R}^m \quad \Longleftrightarrow \quad \sum_{i=1}^{N} \Theta_{j,i}(u_i) - \theta_j = 0 \in \mathbb{R} \;, \;\; j = 1, \dots, m \;.$$

Pour chaque contrainte $j_0 \in \{1, \dots, m\}$, on choisit l'unité $i_0 \in \{1, \dots, N\}$ à laquelle on affecte la contrainte, que l'on écrit de manière équivalente en ajoutant une nouvelle variable v_{j_0} :

$$\sum_{i=1}^{N} \Theta_{j_0,i}(u_i) - \theta_{j_0} = 0 \quad \Longleftrightarrow \quad \begin{cases} \Theta_{j_0,i_0}(u_{i_0}) - v_{j_0} = 0 \;, \\ \sum_{i \neq i_0} \Theta_{j_0,i}(u_i) + v_{j_0} - \theta_{j_0} = 0 \;. \end{cases}$$

La première contrainte est gardée telle quelle et apparaît dans le sous-problème i_0, associée à un multiplicateur λ_{j_0}, tandis que la seconde contrainte est dualisée, d'où le terme supplémentaire $\langle p_{j_0} \,, \Theta_{j_0,i}(u_i) \rangle$ dans le critère de chaque sous-problème $i \neq i_0$.

Esquisse du cas général II

Ce traitement appliqué à toutes les contraintes conduit à la formulation de N sous-problèmes. Notant J_i l'ensemble des contraintes affectées à l'unité i, le sous-problème i s'écrit :

$$\min_{u_i \in U_i^{\mathrm{ad}}} \quad J_i(u_i) + \sum_{j \notin J_i} \left\langle p_j^{(k)} \, , \Theta_{j,i}(u_i) \right\rangle$$

$$\text{sous} \quad \Theta_{j,i}(u_i) - v_j^{(k)} = 0 \quad \forall j \in J_i \; .$$

Les solutions des sous-problèmes sont notées $(u_1^{(k+1)}, \ldots, u_N^{(k+1)})$, et on note $(\lambda_1^{(k+1)}, \ldots, \lambda_m^{(k+1)})$ les multiplicateurs associés. Pour tout $j = 1, \ldots, m$, la remise à jour de v et p est donnée par :

$$p_j^{(k+1)} = \lambda_j^{(k+1)} \; ,$$

$$v_j^{(k+1)} = \theta_j - \sum_{\ell \neq i} \Theta_{j,\ell}(u_\ell^{(k+1)}) \; ,$$

avec i tel que $j \in J_i$.

Comparaison des méthodes de décomposition

Dans le cas d'un problème d'optimisation à structure additive :

$$\min_{(u_1,\dots,u_N)\in U_1^{\mathrm{ad}}\times\cdots\times U_N^{\mathrm{ad}}} \sum_{i=1}^{N} J_i(u_i) \quad \text{sous} \quad \sum_{i=1}^{N} \Theta_i(u_i) - \theta = 0 ,$$

on a étudié 3 méthodes de décomposition/coordination :

- la **décomposition par les prix**, simple à mettre en œuvre, conduit à des sous-problèmes bien formulés, mais la contrainte couplante est satisfaite seulement après convergence ;
- la **décomposition par les quantités** produit une solution admissible à chaque itération, mais elle peut se bloquer en formulant des sous-problèmes définis sur l'ensemble vide ;
- la **décomposition par prédiction**, un peu plus complexe, peut parvenir, suivant le choix d'affectation des contraintes, à cumuler les avantages les deux méthodes précédentes.

Conclusions et suite du cours

Mettre en œuvre une méthode de décomposition/coordination pose des questions de choix et d'architecture d'algorithme :

- il faut choisir la manière de décomposer le problème,
- il faut choisir la méthode de décomposition à utiliser,
- il faut choisir l'affection des contraintes aux sous-problèmes (cas de la prédiction),
- il faut choisir la manière de résoudre des sous-problèmes.

Tout ce qu'on a fait jusqu'alors repose sur l'hypothèse cruciale (*en apparence*) que le problème étudié possède une structure additive. Dans le prochain cours, on verra que cette hypothèse n'est en fait pas utile et que l'on peut mettre en œuvre des méthodes de décomposition sur des problèmes généraux.

Rappel de la décomposition par les quantités I

$$\min_{(u_1,\dots,u_N)\in U_1^{\mathrm{ad}}\times\cdots\times U_N^{\mathrm{ad}}} \sum_{i=1}^{N} J_i(u_i) \quad \text{sous} \quad \sum_{i=1}^{N} \Theta_i(u_i) - \theta = 0 \ ,$$

Introduisant de nouvelles variables $(v_1,\dots,v_N)$, on a :

$$\sum_{i=1}^{N} \Theta_i(u_i) - \theta = 0 \iff \Theta_i(u_i) - v_i = 0 \ , \ \ i = 1,\dots,N \ \text{ et } \ \sum_{i=1}^{N} v_i - \theta = 0 \ ,$$

et le problème à résoudre s'écrit de manière équivalente :

$$\min_{(v_1,\dots,v_N)\in\mathcal{V}^N} \sum_{i=1}^{N} \Big\{ \underbrace{\min_{u_i\in U_i^{\mathrm{ad}}} J_i(u_i) \ \text{ sous } \ \Theta_i(u_i) - v_i = 0}_{G_i(v_i)} \Big\}$$

$$\text{sous la contrainte} \quad \sum_{i=1}^{N} v_i - \theta = 0 \ .$$

Rappel de la décomposition par les quantités II

- La minimisation en u à $v = v^{(k)}$ fixé se décompose :

$$\min_{u_i \in U_i^{\mathrm{ad}}} J_i(u_i) \;\text{ sous }\; \Theta_i(u_i) - v_i^{(k)} = 0 \;, \quad i = 1, \ldots, N \;,$$

  et on note $p_i^{(k+1)}$ le multiplicateur associé à la contrainte.

- la remise à jour de v par gradient projeté s'écrit :

$$\begin{pmatrix} v_1^{(k+1)} \\ \vdots \\ v_N^{(k+1)} \end{pmatrix} = \operatorname*{proj}_{\sum_{i=1}^N v_i = \theta} \begin{pmatrix} v_1^{(k)} + \rho p_1^{(k+1)} \\ \vdots \\ v_N^{(k)} + \rho p_N^{(k+1)} \end{pmatrix} .$$

  Après calcul analytique de l'opération de projection, on obtient :

$$v_i^{(k+1)} = v_i^{(k)} + \rho \Big(p_i^{(k+1)} - \frac{1}{N} \sum_{j=1}^N p_j^{(k+1)} \Big) \;, \quad \forall i = 1, \ldots, N \;.$$

Rappel de la décomposition par les quantités III

- Chaque sous-problème i est formulé sous les 2 contraintes $u_i \in U_i^{\mathrm{ad}}$ **et** $\Theta_i(u_i) - v_i^{(k)} = 0$. Ces contraintes peuvent être incompatibles, auquel cas l'algorithme se bloque.
- La solution de chaque itération k de l'algorithme vérifie :

$$\sum_{i=1}^{N} \Theta_i(u_i^{(k+1)}) = \sum_{i=1}^{N} v_i^{(k)} = \theta ,$$

  et est donc admissible pour le problème global.
- La philosophie de la méthode par les quantités est d'ajouter des variables et des contraintes au problème afin de pouvoir le décomposer. On verra que, dans certains cas, le choix des variables à ajouter peut être plus intéressant que le choix « canonique » $(v_1, \ldots, v_N)$.

Décomposition par allocation et contrainte inégalité E_1

On considère le cas d'une contrainte de type inégalité dans le problème d'optimisation à structure additive :

$$\min_{u_i \in U_i^{\mathrm{ad}} \subset \mathbb{R}^{n_i}} \sum_{i=1}^{N} J_i(u_i) \quad \text{sous} \quad \sum_{i=1}^{N} \Theta_i(u_i) - \theta \leq 0 \in \mathbb{R}^m \ ,$$

que l'on souhaite décomposer par les quantités avec une allocation $(v_1, \ldots, v_N)$.

1. Quelles sont les deux différentes possibilités de mise en œuvre de la décomposition par les quantités ?
2. Quelle est celle qui vous semble la plus proche de celle donnée en cours ?

Décomposition par allocation et contrainte inégalité R_1

La mise en œuvre de la méthode de décomposition par les quantités passe par la réécriture de la contrainte à l'aide de l'allocation. Dans le cas d'une contrainte inégalité, 2 choix sont possibles :

- $$\sum_{i=1}^{N} v_i - \theta \leq 0 \quad \text{et} \quad \Theta_i(u_i) - v_i = 0 \,, \quad i = 1, \ldots, N \,,$$
- $$\sum_{i=1}^{N} v_i - \theta = 0 \quad \text{et} \quad \Theta_i(u_i) - v_i \leq 0 \,, \quad i = 1, \ldots, N \,,$$

tous deux équivalents à la contrainte initiale :

$$\sum_{i=1}^{N} \Theta_i(u_i) - \theta \leq 0 \,.$$

Décomposition par allocation et contrainte inégalité R_2

Dans la première possibilité, les sous-problèmes sont inchangés :

$$\min_{u_i \in U_i^{\mathrm{ad}}} J_i(u_i) \quad \text{sous} \quad \Theta_i(u_i) - v_i^{(k)} = 0 \ , \quad i = 1, \dots, N \ ,$$

et la remise à jour de v par gradient projeté devient :

$$\begin{pmatrix} v_1^{(k+1)} \\ \vdots \\ v_N^{(k+1)} \end{pmatrix} = \operatorname*{proj}_{\sum_{i=1}^N v_i \leq \theta} \begin{pmatrix} v_1^{(k)} + \rho p_1^{(k+1)} \\ \vdots \\ v_N^{(k)} + \rho p_N^{(k+1)} \end{pmatrix} \ .$$

Le calcul de projection fait durant le cours ne s'applique alors pas car il correspond à la projection sur un hyperplan. La formule de projection sur un demi-espace n'a pas été donnée, et il faudrait donc la calculer !

Décomposition par allocation et contrainte inégalité R_3

Avec la seconde possibilité, les sous problèmes sont formulés sous contrainte d'inégalité :

$$\min_{u_i \in U_i^{\mathrm{ad}}} J_i(u_i) \text{ sous } \Theta_i(u_i) - v_i^{(k)} \leq 0 \ , \quad i = 1, \dots, N \ ,$$

et la remise à jour de v par gradient projeté :

$$\begin{pmatrix} v_1^{(k+1)} \\ \vdots \\ v_N^{(k+1)} \end{pmatrix} = \operatorname*{proj}_{\sum_{i=1}^N v_i = \theta} \begin{pmatrix} v_1^{(k)} + \rho p_1^{(k+1)} \\ \vdots \\ v_N^{(k)} + \rho p_N^{(k+1)} \end{pmatrix} \ ,$$

identique à celle du cours, se met sous la forme :

$$v_i^{(k+1)} = v_i^{(k)} + \rho \Big(p_i^{(k+1)} - \frac{1}{N} \sum_{j=1}^{N} p_j^{(k+1)} \Big) \ , \quad i = 1, \dots, N \ .$$

Compromis entre investissement et fonctionnement E_1

On considère le problème d'optimisation suivant :

$$\min_{(u_0,u_1,\ldots,u_N)\in\mathbb{R}^{N+1}} J_0(u_0)+\sum_{i=1}^{N} J_i(u_i) \text{ sous } \Omega_i(u_i)-u_0 \leq 0 \in \mathbb{R} \, , \quad i=1,\ldots,N \, ,$$

et sous les contraintes $u_i \in [\underline{u}_i, \overline{u}_i]$, $i = 0, \ldots, N$. Ce problème est celui de l'optimisation du fonctionnement de N unités dont la production est limitée par une variable d'investissement u_0.

1. Appliquer à ce problème la méthode de décomposition par les quantités telle qu'elle a été présentée dans le cours, c'est à dire en introduisant une allocation $(v_0, v_1, \ldots, v_N)$. Que dire des chances de succès de l'algorithme correspondant ?
2. Proposer une application plus subtile de la méthode par les quantités, qui en respecte l'esprit plutôt que la lettre.

Compromis entre investissement et fonctionnement R_1

Dans l'application directe de la décomposition par les quantités, on récrit les contraintes du problème :

$$\begin{pmatrix} \Omega_1(u_1) - u_0 \\ \vdots \\ \Omega_N(u_N) - u_0 \end{pmatrix} \leq 0 \in \mathbb{R}^N ,$$

sous la forme standard :

$$\sum_{i=0}^{N} \Theta_i(u_i) \leq 0 ,$$

avec

$$\Theta_0(u_0) = \begin{pmatrix} -u_0 \\ \vdots \\ -u_0 \end{pmatrix} , \quad \Theta_1(u_1) = \begin{pmatrix} \Omega_1(u_1) \\ 0 \\ \vdots \\ 0 \end{pmatrix} , \quad \ldots , \quad \Theta_N(u_N) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \Omega_N(u_N) \end{pmatrix} .$$

Compromis entre investissement et fonctionnement R_2

Introduisant alors l'allocation $(v_0, v_1, \ldots, v_N)$ telle que chaque $v_i \in \mathbb{R}^N$, l'algorithme de décomposition par les quantités est précisément celui étudié dans l'exercice précédent :

- **décomposition** : pour $i = 0, \ldots, N$,

$$\min_{u_i \in [\underline{u}_i, \overline{u}_i]} J_i(u_i) \quad \text{sous} \quad \Theta_i(u_i) - v_i^{(k)} \leq 0$$
$$\rightsquigarrow \left(u_i^{(k+1)}, p_i^{(k+1)}\right) ,$$

- **coordination** : pour $i = 0, \ldots, N$,

$$v_i^{(k+1)} = v_i^{(k)} + \rho\left(p_i^{(k+1)} - \frac{1}{N+1}\sum_{j=0}^{N} p_j^{(k+1)}\right) .$$

Compromis entre investissement et fonctionnement R_3

Cet algorithme, application brutale de la décomposition par les quantités, a toutes les chances de se bloquer. Considérons en effet la contrainte associé au sous-problème 1 :

$$\begin{pmatrix} \Omega_1(u_1) \\ 0 \\ \vdots \\ 0 \end{pmatrix} - \begin{pmatrix} v_{1,1}^{(k)} \\ v_{1,2}^{(k)} \\ \vdots \\ v_{1,N}^{(k)} \end{pmatrix} \leq \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^N .$$

- La première composante de cette contrainte impose que $\Omega_1(u_1) \leq v_{1,1}^{(k)}$, condition réalisable grâce à la décision u_1.
- Les conditions suivantes s'écrivent $v_{1,i}^{(k)} \geq 0$. Les quantités $v_{1,i}^{(k)}$ sont calculées par l'étape de coordination par un calcul de type gradient, et rien n'impose que ces quantités restent positives. Le calcul des multiplicateurs est alors inconsistant.

Compromis entre investissement et fonctionnement R_4

Un examen attentif du problème :

$$\min_{(u_0,u_1,\ldots,u_N)\in\mathbb{R}^{N+1}} J_0(u_0)+\sum_{i=1}^{N} J_i(u_i) \ \text{ sous } \ \Omega_i(u_i)-u_0 \leq 0 \in \mathbb{R} \ , \quad i=1,\ldots,N \ ,$$

montre que la variable u_0 peut en fait jouer à elle seule le rôle de l'allocation. En effet, le problème s'écrit de manière équivalente :

$$\min_{u_0\in[\underline{u}_0,\overline{u}_0]} J_0(u_0)+\sum_{i=1}^{N}\left\{\min_{u_i\in[\underline{u}_i,\overline{u}_i]} J_i(u_i) \ \text{ sous } \ \Omega_i(u_i)-u_0 \leq 0\right\} .$$

et l'on constate que, à u_0 fixé, le problème se décompose en N sous-problèmes ne dépendant chacun que d'une seule variable u_i, $i=1,\ldots,N$. La mise à jour de la variable u_0 est ensuite effectuée par un pas de gradient.

Compromis entre investissement et fonctionnement R_5

L'algorithme de décomposition par les quantités utilisant le fait que u_0 joue à elle seule le rôle d'allocation est alors :

- **décomposition** : pour $i = 1, \ldots, N$,

$$\min_{u_i \in [\underline{u}_i, \overline{u}_i]} J_i(u_i) \quad \text{sous} \quad \Theta_i(u_i) - u_0^{(k)} \leq 0$$
$$\rightsquigarrow \left(u_i^{(k+1)}, p_i^{(k+1)}\right) ,$$

- **coordination** :

$$u_0^{(k+1)} = \operatorname*{proj}_{[\underline{u}_0, \overline{u}_0]} \left(u_0^{(k)} - \rho \Big(\nabla J_0(u_0^{(k)}) - \sum_{i=1}^{N} p_i^{(k+1)} \Big) \right) .$$

On constate que ce nouvel algorithme ne peut jamais se bloquer !

Décomposition par les quantités d'un réseau d'eau E_1

On rappelle la formulation compacte du problème d'optimisation du grand réseau d'eau connecté :

$$\min_{(u_{i,1},u_{i,2})_{i=1,\dots,N+1}} \sum_{i=1}^{N+1} J_i(u_{i,1},u_{i,2}) \quad \text{sous} \quad \sum_{i=1}^{N} u_{i,t} - u_{N+1,t} = 0 \ , \quad t = 1,2 \ ,$$

avec les contraintes de bornes :

$$(u_{i,1},u_{i,2}) \in U_i^{\mathrm{ad}} = [0,\bar{v}_{i,1}] \times [0,\bar{v}_{i,2} - \bar{v}_{i,1}] \ , \quad i = 1,\dots,N \ .$$

Pour $i = 1,\dots,N$, l'expression détaillée de la fonction J_i est :

$$J_i(u_{i,1},u_{i,2}) = \min_{(v_{i,1},v_{i,2})} \frac{1}{2}\left(a_{i,1}v_{i,1}^2 + a_{i,2}v_{i,2}^2\right) \ ,$$

$$\text{sous} \quad \bar{v}_{i,1} - u_{i,1} - v_{i,1} \leq 0 \ ,$$

$$u_{i,1} + u_{i,2} + v_{i,1} + v_{i,2} - \bar{v}_{i,2} = 0 \ ,$$

et on a : $J_{N+1}(u_{N+1,1},u_{N+1,2}) = \frac{1}{2}\left(a_{N+1,1}u_{N+1,1}^2 + a_{N+1,2}u_{N+1,2}^2\right)$.

Décomposition par les quantités d'un réseau d'eau E_2

1. Écrire sur la formulation compacte du problème l'algorithme de décomposition par les quantités, en utilisant une allocation $(w_{i,1}, w_{i,2})_{i=1,\dots,N+1}$.
2. Donner la formulation détaillée de chaque sous-problème i, $i = 1, \dots, N+1$.
   - Discuter de la résolution des conditions de KKT de ces sous-problèmes ?
   - Montrer que l'ensemble admissible de certains sous-problèmes peut être vide, et donc que l'algorithme peut se bloquer.
3. Proposer d'autres manières de faire l'allocation de ressources qui permettent d'éviter le blocage mentionné ci-dessus.

Décomposition par les quantités d'un réseau d'eau $\quad R_1$

À l'itération k de la méthode par les quantités, avec une allocation $(w_{i,1}^{(k)}, w_{i,2}^{(k)})_{i=1,\dots,N+1}$ l'étape de décomposition s'écrit :

$$\min_{(u_{i,1},u_{i,2})\in U_i^{\mathrm{ad}}} J_i(u_{i,1},u_{i,2}) \ \text{ sous } \begin{cases} u_{i,1} - w_{i,1}^{(k)} = 0 \\ u_{i,2} - w_{i,2}^{(k)} = 0 \end{cases} \quad , \quad i = 1,\dots,N \ ,$$

$$\min_{(u_{N+1,1},u_{N+1,2})\in\mathbb{R}^2} J_{N+1}(u_{N+1,1},u_{N+1,2}) \ \text{ sous } \begin{cases} -u_{N+1,1} - w_{N+1,1}^{(k)} = 0 \\ -u_{N+1,2} - w_{N+1,2}^{(k)} = 0 \end{cases} \quad .$$

Notant $(p_{i,1}^{(k+1)}, p_{i,2}^{(k+1)})_{i=1,\dots,N+1}$ les multiplicateurs associés aux contraintes de ces sous-problèmes, la remise à jour de l'allocation est un pas de gradient projeté sur les hyperplans $\sum_{i=1}^{N+1} w_{i,t} = 0$:

$$w_{i,1}^{(k+1)} = w_{i,1}^{(k)} + \rho\Big(p_{i,1}^{(k+1)} - \frac{1}{N+1}\sum_{j=1}^{N+1} p_{j,1}^{(k+1)}\Big) \ ,$$

$$w_{i,2}^{(k+1)} = w_{i,2}^{(k)} + \rho\Big(p_{i,2}^{(k+1)} - \frac{1}{N+1}\sum_{j=1}^{N+1} p_{j,2}^{(k+1)}\Big) \ .$$

Décomposition par les quantités d'un réseau d'eau R_2

L'écriture détaillée du sous-problème $N+1$ est :

$$\min_{(u_{N+1,1},u_{N+1,2})\in\mathbb{R}^2} \frac{1}{2}\left(a_{N+1,1}u_{N+1,1}^2 + a_{N+1,2}u_{N+1,2}^2\right)$$

$$\text{sous} \quad \begin{cases} -u_{N+1,1} - w_{N+1,1}^{(k)} = 0 \\ -u_{N+1,2} - w_{N+1,2}^{(k)} = 0 \end{cases} ,$$

dont la résolution analytique donne la solution optimale :

$$u_{N+1,1}^{(k+1)} = -w_{N+1,1}^{(k)} \,, \quad p_{N+1,1}^{(k+1)} = a_{N+1,1}u_{N+1,1}^{(k+1)} \,,$$

$$u_{N+1,2}^{(k+1)} = -w_{N+1,2}^{(k)} \,, \quad p_{N+1,2}^{(k+1)} = a_{N+1,2}u_{N+1,2}^{(k+1)} \,.$$

Décomposition par les quantités d'un réseau d'eau R_3

La forme détaillée du sous-problème i, $i = 1, \ldots, N$ est :

$$
\begin{aligned}
\min_{(v_{i,1}, v_{i,2}, u_{i,1}, u_{i,2}) \in \mathbb{R}^4} \quad & \frac{1}{2}\left(a_{i,1} v_{i,1}^2 + a_{i,2} v_{i,2}^2\right) , \\
\text{sous} \quad & \overline{v}_{i,1} - v_{i,1} - u_{i,1} \leq 0 , \\
& v_{i,1} + v_{i,2} + u_{i,1} + u_{i,2} - \overline{v}_{i,2} = 0 , \\
& u_{i,1} - w_{i,1}^{(k)} = 0 , \\
& u_{i,2} - w_{i,2}^{(k)} = 0 , \\
& 0 \leq u_{i,1} \leq \overline{v}_{i,1} , \\
& 0 \leq u_{i,2} \leq \overline{v}_{i,2} - \overline{v}_{i,1} .
\end{aligned}
$$

Ce sous-problème comporte 5 contraintes inégalité : résoudre les conditions de KKT est donc de même nature que dans la décomposition par les prix, à savoir fastidieux mais faisable.

Décomposition par les quantités d'un réseau d'eau R_4

Considérons les 4 dernières contraintes du sous-problème i :

$$u_{i,1} - w_{i,1}^{(k)} = 0 \; ,$$
$$u_{i,2} - w_{i,2}^{(k)} = 0 \; ,$$
$$0 \leq u_{i,1} \leq \overline{v}_{i,1} \; ,$$
$$0 \leq u_{i,2} \leq \overline{v}_{i,2} - \overline{v}_{i,1} \; .$$

Les conditions portant sur $u_{i,1}$ sont facilement incompatibles : en effet, on demande à la variable $u_{i,1}$, d'une part d'appartenir à l'intervalle $[0, \overline{v}_{i,1}]$, et d'autre part de prendre la valeur $w_{i,1}^{(k)}$. Dès que les valeurs définissant le problème sont telles que :

$$w_{i,1}^{(k)} \notin [0, \overline{v}_{i,1}] \; ,$$

l'ensemble admissible de ce sous-problème est vide et l'algorithme se bloque. On a la même conclusion si $w_{i,2}^{(k)} \notin [0, \overline{v}_{i,2} - \overline{v}_{i,1}]$.

Décomposition par les quantités d'un réseau d'eau R_5

Pour résoudre ce problème de blocage, on peut « déplacer » les contraintes $(u_{i,1}, u_{i,2}) \in U_i^{\mathrm{ad}}$, $i = 1, \dots, N$, vers les variables $(w_{i,1}, w_{i,2})$ car l'allocation fait que ces variables sont égales.

Avec ce déplacement, le sous-problème $N+1$ n'est pas modifié, et les autres sous-problèmes i, qui deviennent :

$$\begin{aligned}
\min_{(v_{i,1}, v_{i,2}, u_{i,1}, u_{i,2}) \in \mathbb{R}^4} \quad & \frac{1}{2}\big(a_{i,1} v_{i,1}^2 + a_{i,2} v_{i,2}^2\big) \ , \\
\text{sous} \quad & \overline{v}_{i,1} - v_{i,1} - u_{i,1} \leq 0 \ , \\
& v_{i,1} + v_{i,2} + u_{i,1} + u_{i,2} - \overline{v}_{i,2} = 0 \ , \\
& u_{i,1} - w_{i,1}^{(k)} = 0 \ , \\
& u_{i,2} - w_{i,2}^{(k)} = 0 \ .
\end{aligned}$$

ne peuvent plus se bloquer et fournissent toujours une solution.

Décomposition par les quantités d'un réseau d'eau R_6

La remise à jour de l'allocation w par gradient projeté :

$$\begin{pmatrix} w_1^{(k+1)} \\ \vdots \\ w_{N+1}^{(k+1)} \end{pmatrix} = \operatorname*{proj}_{\Pi} \begin{pmatrix} w_1^{(k)} + \rho p_1^{(k+1)} \\ \vdots \\ v_{N+1}^{(k)} + \rho p_{N+1}^{(k+1)} \end{pmatrix} .$$

nécessite de calculer la projection sur l'ensemble Π, intersection des hyperplans $\big\{(w_{1,t}, \dots, w_{N+1,t}) \in \sum_{i=1}^{N+1} w_{i,t} = 0,\ t = 1, 2\big\}$, et des ensembles $U_i^{\mathrm{ad}} = \big\{(w_{i,1}, w_{i,2}) \in [0, \overline{v}_{i,1}] \times [0, \overline{v}_{i,2} - \overline{v}_{i,1}]\big\}$, $i = 1, \dots, N$.

Projeter sur chacun des ensembles constituant cette intersection est aisée, mais projeter sur l'intersection elle-même est plus difficile et nécessite l'utilisation d'un algorithme spécialisé. . .

Décomposition par les quantités d'un réseau d'eau R_7

Pour éviter la difficulté de projection sur une intersection d'ensembles, on peut enfin considérer une allocation réduite $(w_{i,1}, w_{i,2})_{i=1,\dots,N}$.[9] L'étape de décomposition s'écrit alors :

$$\min_{(u_{i,1},u_{i,2})\in\mathbb{R}^2} J_i(u_{i,1},u_{i,2}) \text{ sous } \begin{cases} u_{i,1} - w_{i,1}^{(k)} = 0 \\ u_{i,2} - w_{i,2}^{(k)} = 0 \end{cases} , \quad i = 1,\dots,N ,$$

$$\min_{(u_{N+1,1},u_{N+1,2})\in\mathbb{R}^2} J_{N+1}(u_{N+1,1},u_{N+1,2}) \text{ sous } \begin{cases} -u_{N+1,1} + \sum_{i=1}^N w_{i,1}^{(k)} = 0 \\ -u_{N+1,2} + \sum_{i=1}^N w_{i,2}^{(k)} = 0 \end{cases} .$$

La remise à jour de l'allocation réduite nécessite le calcul des gradients partiels par rapport aux variables $w_{i,t}$ qui apparaissent dans le sous-problème i ainsi que dans le sous-problème $N+1$. Mais la formule de mise à jour ne fait plus intervenir qu'une projection « simple » sur les ensembles U_i^{ad} !

9. L'allocation « canonique » est $(w_{i,1}, w_{i,2})_{i=1,\dots,N+1}$.