

Introduction

Dans ce chapitre d'introduction, on tente d'abord de définir le contour de ce qui a émergé dans les années 1960-1970 comme le champ des « grands systèmes » (*Large Scale Systems* en Anglais LASDON (1970); WISMER (1971), aussi qualifiés de « systèmes hiérarchisés » ou « multi-niveaux », LASDON et SCHOEFLER (1965); MESAROVIC et collab. (1970); FINDEISEN et collab. (1980)) et les problématiques associées. On restreint ensuite ces problématiques à la théorie de la décomposition-coordination en optimisation qui constitue l'objet de ce cours, préoccupation apparue dès le début des années 60 ARROW et HURWICZ (1960); DANTZIG et WOLFE (1961). On termine par un aperçu du contenu de cet ouvrage et des notions que le lecteur est supposé avoir acquises en optimisation pour en comprendre les développements.

1.1 Sur les « Grands Systèmes »

Au milieu des années 60, dans le domaine des mathématiques de la décision, commencent à apparaître diverses préoccupations qui gravitent autour de la notion de « grand système ». On donne ici un aperçu de cette notion et des diverses problématiques apparues autour de ce thème.

1.1.1 Tentative de caractérisation

On parle de « grand système » ou encore de « système complexe » dans l'un des cas suivants.

- Le système est décrit par un *grand nombre de variables et de contraintes*, ce qui entraîne beaucoup de calculs et/ou de stockage lorsqu'on s'attaque à un problème d'optimisation impliquant ce système. Une situation typique est celle où l'on doit utiliser la *programmation dynamique* pour résoudre un problème de commande optimale : la croissance du volume des calculs étant exponentielle par rapport au nombre

de variables d'état, on parle alors de « malédiction de la dimension » (« curse of dimensionality » BELLMAN (1957)).

- Le système global est constitué de sous-systèmes interconnectés BRO-SILOW et collab. (1965) : il s'agit donc d'une *structure spatiale complexe* (voir Figure 1.1), ce qui s'accompagne parfois d'une certaine *hétérogénéité* dans la nature des sous-systèmes. En effet, la plupart des grands systèmes (industriels ou économiques par exemple) se sont constitués progressivement par interconnexion de systèmes plus petits, et parfois de nature différente.



Fig. 1.1. Sous-systèmes interconnectés

- Le système (dynamique) met en jeu des phénomènes à *plusieurs échelles de temps* ou changeant de façon brutale au cours du temps : dans ce cas, c'est la *structure temporelle* qui est *complexe*. Un exemple est donné par une fusée larguant ses étages successifs le long de sa trajectoire (voir Figure 1.2), ce qui occasionne des changements abrupts du modèle mathématique correspondant.

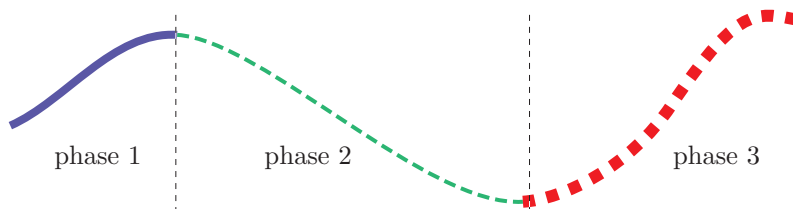


Fig. 1.2. Trajectoire composite

- Le système est commandé simultanément par *plusieurs décideurs* agissant sur l'ensemble du système (voir Figure 1.3) ou sur des systèmes interconnectés avec
 - des informations différentes : en effet, la centralisation des informations peut être impossible ou économiquement prohibitive pour des systèmes très étendus géographiquement ;
 - des objectifs concordants, ou bien différents, voire conflictuels. Dans le premier cas, on entre dans le champ de la *théorie des équipes* MARSCHAK et RADNER (1971), et dans le second, on tombe plutôt

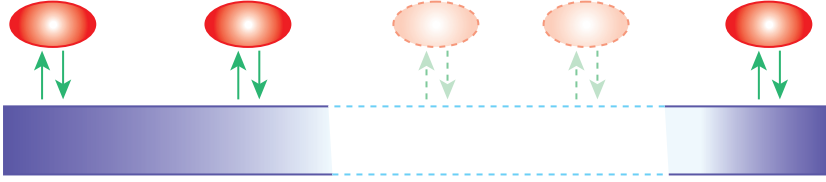


Fig. 1.3. Décision multi-agents

dans le domaine de la *théorie des jeux* VON NEUMANN et MORGENTHAU (1953), aspects qui ne seront plus abordés dans la suite (on se restreindra au cadre de l'optimisation avec une fonction objectif et un décideur uniques).

1.1.2 Quelques exemples typiques de grands systèmes

On les rencontre notamment à travers les grands ensembles de production, de distribution et de services : réseaux d'eau, de chauffage urbain, d'électricité, de gaz, de téléphone et autres moyens de télécommunications, réseaux de transport, ainsi qu'en économie (entreprises à filiales multiples, usines composées de plusieurs ateliers), etc..

1.1.3 Problématiques en optimisation de grands systèmes

On se restreint à la formulation et à la résolution de problèmes d'*optimisation*. D'autres préoccupations, comme par exemple la *stabilisation* de sous-systèmes dynamiques interconnectés, ont également fait l'objet de nombreux travaux, mais ces sujets ne seront pas abordés dans ce texte.

Optimisation déterministe et optimisation stochastique, structures d'information

En matière d'optimisation, il faut d'abord distinguer entre

- le cadre *déterministe*, dans lequel aucune notion d'incertitude n'est explicitement présente dans la formulation,
- et le cadre *stochastique* dans lequel un modèle des incertitudes (généralement basé sur des distributions de probabilité) fait partie intégrante de la formulation du problème.

Il faut de plus, et *dans le contexte stochastique uniquement*, distinguer

- les situations *dynamiques* où les décisions à prendre doivent être basées sur des informations révélées *progressivement* au fur et à me-

- sure de l'écoulement de l'horizon d'optimisation (information dite « en ligne »)¹,
- des situations *statiques* où toute l'information est rendue disponible immédiatement avant la prise de toutes les décisions (c'est-à-dire « hors ligne »).

Cette distinction entre situations dynamique et statique n'a pas lieu d'être faite dans le cadre déterministe, par définition même de la notion de déterminisme qui suppose la connaissance a priori de toutes les informations utilisables pour élaborer les décisions.

Décomposition-coordination en optimisation déterministe

Dans le contexte de l'optimisation déterministe, où, comme on vient de le dire, la notion d'information (sous-entendue « en ligne ») n'est pas une notion pertinente, la préoccupation essentielle est donc la résolution hors ligne du problème d'optimisation.

Pour contourner les difficultés liées à la taille et exploiter les caractéristiques évoquées ci-dessus, on fait appel aux idées suivantes :

- la **décomposition** du problème initial (dit « problème global ») qui consiste à formuler des sous-problèmes relatifs, chacun, à l'un des sous-systèmes composant le grand système ;
- la **coordination** dont le but ultime est de faire en sorte que chaque solution de sous-problème (dite « solution locale ») fournisse une partie de la solution du problème global. Ce but ne sera généralement atteint qu'à l'issue d'un processus itératif.

Cette structure à deux niveaux est illustrée par la Figure 1.4.

Optimisation stochastique de grands systèmes

Comme mentionné plus haut, lorsque la performance finale dépend à la fois des actions dont le ou les décideurs ont la maîtrise, mais aussi de la réalisation d'aléas (que l'on peut qualifier d'« action de la nature »), on doit préciser dans la formulation du problème d'optimisation quelle connaissance le ou les décideur(s) a/ont de la réalisation de ces aléas.

1. Dans CARPENTIER et collab. (2015), on explique que les situations « dynamiques » en optimisation stochastique ne sont pas exclusivement ou nécessairement liées à l'écoulement du temps et à la présence d'un « système dynamique », généralement représenté par des équations différentielles dans un modèle en temps continu, ou par des équations récurrentes en temps discret, ce qui donne lieu à des problèmes de « commande optimale » (déterministes ou stochastiques). La notion de « dynamique » en optimisation stochastique est plus essentiellement caractérisée par la « structure d'information » entrant dans la définition du problème. Cependant, les problèmes de commande optimale stochastique sont les plus typiques de cette situation (voir §6.2.2).

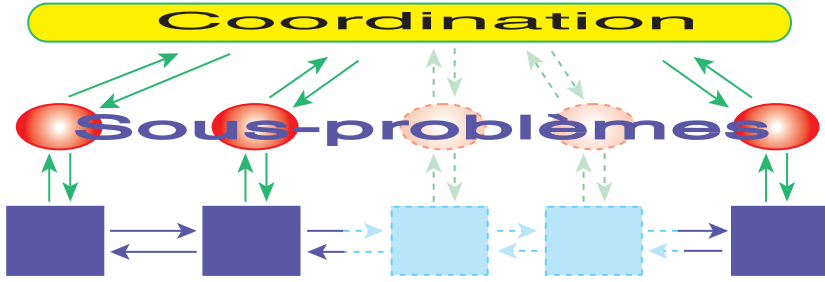


Fig. 1.4. Coordination

Dans la situation *statique*, on suppose que la seule connaissance est une connaissance *statistique* modélisée par une distribution de probabilité. On ne connaît donc pas à l'avance à quelle réalisation des aléas on sera confronté ; la fonction objectif (ou fonction coût) est évaluée en *espérance mathématique* sur toutes les réalisations possibles des aléas compte tenu de leur distribution de probabilité. Même si le calcul de cette espérance mathématique constitue en soi une difficulté supplémentaire (dont on verra dans cet ouvrage comment elle est contournée par la technique du gradient stochastique et sa généralisation aux algorithmes de décomposition-coordination), la situation n'est somme toute pas très différente du cas déterministe du point de vue de la structure d'information et on traitera donc aussi dans ce livre la généralisation des techniques de décomposition-coordination à ce cadre stochastique.

La situation stochastique *dynamique* se caractérise, elle, par une alternance de prises de décisions et de réalisations d'aléas au cours du déroulement d'un horizon d'optimisation (pour nous limiter à nouveau au cas le plus courant de la commande optimale stochastique — voir la note 1). Dans ce processus, chaque décision instantanée est possiblement prise en connaissant les réalisations *passées* des aléas, mais doit aussi prendre en compte les aléas *futurs* sous la forme de leur distribution de probabilité. Non seulement la notion de « passé » et de « futur » est relative au positionnement de chaque décision dans l'horizon d'optimisation, mais de surcroît, chaque « agent » ou décideur représenté à la Figure 1.3 n'a éventuellement pas à sa disposition les mêmes informations que les autres agents au même instant. Ces situations, dites « structures d'information non classiques », donnent lieu à des problèmes d'optimisation stochastique qui peuvent être d'une extrême complexité (voir par exemple CARPENTIER et collab. (2015)) et qui vont bien au delà du cadre de cet ouvrage.

En résumé, dans la suite de cet ouvrage, on se restreindra aux problèmes d'optimisation déterministe dans une première partie, et aux problèmes d'optimisation stochastique avec structure d'information statique dans la seconde partie.

1.2 Avantages de la décomposition-coordination en optimisation

Dans cette section, on discute de quelques idées sur l'apport des méthodes de décomposition-coordination en optimisation.

1.2.1 Intérêt de l'optimisation des grands systèmes

Du fait de leur taille, les grands systèmes sont soumis à une *économie d'échelle* : pour de tels systèmes, 1% de gain sur la valeur de la fonction objectif atteinte peut représenter de grandes économies en valeur absolue. C'est pourquoi il importe, malgré la difficulté de résolution de ces problèmes de grande taille, de ne pas renoncer à atteindre la solution optimale du problème global.

1.2.2 Bénéfices de la décomposition

Le premier bénéfice de la décomposition d'un grand problème d'optimisation en sous-problèmes, bénéfice auquel il a déjà été fait allusion, est évidemment la taille réduite des sous-problèmes à résoudre par rapport au problème globalement posé. Les volumes et donc les temps de calcul requis pour la résolution d'un problème d'optimisation croissent généralement de façon *superlinéaire* avec la taille du problème, et parfois même de façon *exponentielle* (comme dans l'approche par programmation dynamique), de sorte que la somme des calculs requis pour résoudre l'ensemble des sous-problèmes obtenus par décomposition du problème initial sera largement inférieure à celle requise pour la résolution de ce dernier par une approche globale. Certes, la coordination qui, comme on le verra, est un processus itératif, nécessitera la résolution *répétée* des sous-problèmes, mais cette répétition jusqu'à obtenir une convergence satisfaisante, ne viendra pas en général effacer le bénéfice en volume de calculs obtenu par décomposition.

Cependant, ce gain en volume de calculs n'est pas le seul avantage procuré par la décomposition. Du fait de la formulation de sous-problèmes rendus indépendants par l'action de la coordination à chaque itération de ce processus, il devient possible d'exploiter de façon immédiate et naturelle les possibilités du *calcul parallèle* de plus en plus répandues sur les machines modernes. C'est une autre source éventuelle de compression du *temps* de calcul.

Même en mettant de côté cette idée du parallélisme, une autre source d'économie sur le temps de calcul qui peut s'avérer très substantielle est liée au phénomène de « swapping », lui-même lié à la technologie de la *mémoire virtuelle*. En effet, dans un ordinateur ayant une certaine capacité de mémoire vive (RAM, pour « random access memory » en Anglais), lorsque cette capacité est insuffisante pour contenir l'ensemble des informations relatives à la résolution du problème, la quantité de mémoire utilisable est artificiellement

étendue en échangeant des « pages » d'information entre la mémoire vive et le disque dur au fur et à mesure de l'avancement des calculs. Ce phénomène de swapping est extrêmement coûteux en temps d'exécution par rapport à une exécution se bornant à utiliser la mémoire vive dans la mesure où les temps d'accès à un disque dur sont beaucoup plus longs qu'avec la RAM. Il y a donc un intérêt majeur à pouvoir faire tenir tous les calculs en RAM grâce à la résolution indépendante de petits problèmes se succédant en mémoire vive.

En dehors de ces raisons de nature technologique, il y a du côté de l'opérateur humain en charge de la résolution du problème de nombreux avantages à n'avoir à traiter que des sous-problèmes de taille modérée. Selon Descartes (*Discours de la méthode*, 1637), il y a intérêt à « diviser chacune des difficultés [que j'examinais] en autant de parcelles qu'il se pourrait et qu'il serait requis pour mieux les résoudre ». Il est en effet beaucoup plus facile de mettre au point des programmes de résolution et d'en contrôler intuitivement les résultats lorsque la taille du problème reste compatible avec ce que l'esprit humain peut appréhender.

Mais il y a un autre intérêt méconnu à pouvoir scinder la résolution d'un problème d'optimisation de grand système en plusieurs sous-problèmes relatifs aux sous-systèmes le composant. En effet, un grand système qui résulte souvent de l'interconnexion de sous-systèmes préexistants n'est pas toujours un système *homogène*, c'est-à-dire que chacune de ses composantes nécessite éventuellement l'utilisation d'un modèle mathématique de nature différente, et donc une méthode de résolution adaptée. C'est précisément ce que permet l'approche par décomposition : en formulant un sous-problème indépendant pour chaque sous-système, on ouvre ainsi la possibilité de choisir pour chacun d'eux une méthode ou un algorithme spécifique. C'est ce qu'on peut qualifier de « croisement d'algorithmes » au sein de la résolution d'un même problème global.

Chaque sous-système se caractérise généralement par un degré d'interconnexion *interne* plus intense que celui des interconnexions *externes* avec le reste des sous-systèmes. Cette caractéristique est souvent la source d'une certaine *homogénéité* interne par opposition à l'*hétérogénéité* du système dans son ensemble. Par « homogénéité », on entend en particulier une certaine cohérence du comportement dynamique. Considérons par exemple un réseau de distribution d'eau composé notamment de réservoirs et de canalisations. Chaque réservoir, ou du moins la hauteur d'eau dans ce réservoir, est une variable à *mémoire* du système dynamique (on appelle cela plus précisément une « variable d'état ») et, comme on l'a dit plus haut, c'est le nombre de ces variables qui est rapidement critique pour une résolution d'un problème de *commande optimale* par la programmation dynamique. Intuitivement, une forte interconnexion entre réservoirs implique, par le principe des vases communicants, une évolution concomitante des niveaux d'eau dans ces réservoirs, alors que deux réservoirs moins fortement connectés connaîtront des évolutions à peu près indépendantes. En conséquence, il sera plus facile de construire un modèle réduit avec un seul réservoir agrégé (donc une seule variable d'état)

représentant de façon approchée l'évolution dynamique d'un ensemble de réservoirs fortement connectés, chose impossible pour des réservoirs faiblement connectés. Autrement dit, les possibilités de construire des *modèles agrégés* satisfaisants augmentent avec l'homogénéité elle-même plus probable à l'échelle des sous-systèmes qu'à celle du système global. C'est une autre source de réduction du volume des calculs qui montre la complémentarité des idées de décomposition et d'agrégation.

1.3 Un exemple en forme de contre-exemple

Dans cette section, on montre qu'une approche naïve dans la décomposition d'un problème d'optimisation ne fournira pas nécessairement la solution du problème initialement formulé, même après convergence d'un processus itératif ayant apparemment correctement « recollé les morceaux ». Cet exemple simple vise à illustrer qu'une démarche plus systématique est requise en matière de décomposition et coordination, ce qui est l'objet de la suite de cet ouvrage.

1.3.1 Formulation du problème

On considère le problème suivant où toutes les variables sont scalaires :

$$\min_{u_1, u_2, y_1, y_2} \left(\underbrace{(u_1 - 7)^2 + y_1^2}_{J_1(u_1, y_1)} + \underbrace{(u_2 - 8)^2 + y_2^2}_{J_2(u_2, y_2)} \right) \quad (1.1a)$$

sous

$$y_1 = \underbrace{u_1 + u_2}_{P_1(u_1, u_2)} \quad \text{et} \quad y_2 = \underbrace{u_1 + u_2}_{P_2(u_1, u_2)} . \quad (1.1b)$$

L'interprétation peut être la suivante : deux processus P_i d'entrées u_i et de sorties y_i , $i = 1, 2$, interagissent par le fait que leurs sorties dépendent aussi de l'entrée u_j , $j \neq i$, comme illustré par la partie gauche de Figure 1.5. On cherche à minimiser une fonction objectif constituée de la somme de deux termes $J_1 + J_2$ dépendant chacun exclusivement de l'entrée et de la sortie du sous-système correspondant.

1.3.2 Une tentative de résolution par décomposition-coordination

On peut essayer d'obtenir la solution par le processus itératif suivant : à l'étape k , la coordination communique (ou « *prédit* ») u_1^k au sous-système 2 et u_2^k au sous-système 1, et les sous-problèmes à résoudre sont alors :

$$\min J_i(u_i, y_i) \text{ avec } y_i = P_i(u_i, u_j^k), \quad i = 1, 2, \quad j \neq i, \quad (1.2)$$

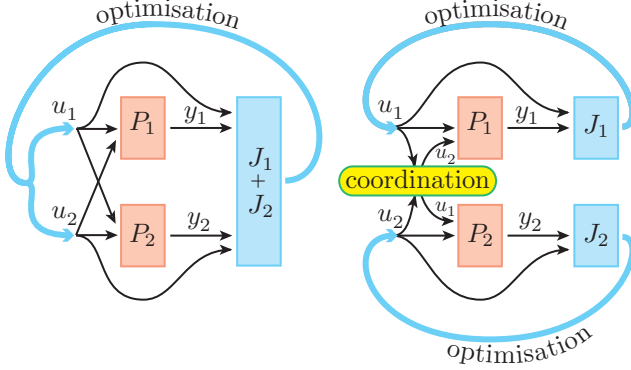


Fig. 1.5. Un exemple

ce qui définit deux nouvelles valeurs de u_1 et u_2 utilisées à l'étape suivante comme nouvelles « prédictions » u_1^{k+1} et u_2^{k+1} . Cette stratégie de décomposition-coordination est illustrée par la partie droite de la Figure 1.5.

La première question soulevée par cet algorithme est celle de la convergence de ce processus itératif. On y répond facilement étant donnée la simplicité de résolution des sous-problèmes qui donne lieu aux formules explicites suivantes pour ce processus itératif :

$$u_1^{k+1} = \frac{1}{2}(7 - u_2^k), \quad u_2^{k+1} = \frac{1}{2}(8 - u_1^k). \quad (1.3)$$

On constate immédiatement que cette récurrence est contractante et que son point fixe est

$$u_1^\infty = 2, \quad u_2^\infty = 3, \quad y_1^\infty = 5, \quad y_2^\infty = 5. \quad (1.4)$$

Exercice 1.1. Retrouver la récurrence (1.3), montrer qu'elle est contractante, et retrouver le point d'équilibre (1.4).

La seconde question à se poser est alors de savoir si on peut affirmer que le point d'équilibre du processus itératif (1.4) fournit la solution du problème global (1.1), étant donné qu'à l'issue du processus itératif, la coordination est capable de prédire à chaque sous-système la valeur correcte de l'entrée qui sera effectivement choisie par l'autre sous-système.

1.3.3 Trouver l'erreur

En fait, la solution du problème (1.1) est

$$u_1^\sharp = 1, \quad u_2^\sharp = 2, \quad y_1^\sharp = 3, \quad y_2^\sharp = 3, \quad (1.5)$$

qui n'a pas de rapport avec (1.4). Trouver l'erreur !

Exercice 1.2. Écrire et résoudre les conditions d'optimalité du problème d'optimisation sous contraintes (1.1).

Effectivement, en définissant

$$G_i(u_i, u_j) = J_i(u_i, P_i(u_i, u_j)) , \quad i = 1, 2, \quad j \neq i , \quad (1.6)$$

on réalise qu'on n'a pas résolu le problème d'optimisation posé, mais plutôt le système d'inéquations variationnelles suivant

$$\forall u_i , \quad G_i(u_i^\infty, u_j^\infty) \leq G_i(u_i, u_j^\infty) , \quad i = 1, 2, \quad j \neq i , \quad (1.7)$$

qui est en fait la définition d'un équilibre de Nash en théorie des jeux.

Exercice 1.3. Démontrer cette affirmation en traduisant la méthode de coordination ci-dessus avec les notations (1.6), puis en se plaçant à l'équilibre. Expliquer en quoi les conditions d'optimalité de (1.7) diffèrent de celles du problème (1.2), une fois celui-ci reformulé en termes des fonctions G_i .

1.4 Aperçu de l'ouvrage et prérequis

La suite de cet ouvrage est composée de deux parties, l'une concernant l'optimisation déterministe et l'autre traitant de l'optimisation stochastique avec structure d'information statique.

1.4.1 Partie « optimisation déterministe »

Dans la première partie, on donne d'abord, au Chapitre 2, une présentation simple de trois méthodes de décomposition en s'appuyant sur deux types de problèmes d'optimisation. En fait, comme on le montrera par un simple changement de notations, ces deux types de problèmes sont mathématiquement équivalents, mais chacune des deux formes met en évidence un mode spécifique de couplage entre sous-problèmes potentiels : la première forme peut être qualifiée de « couplage par les contraintes » alors que la seconde fait plutôt intervenir explicitement des « variables d'interaction ».

Dans ce chapitre, il n'y a pas de préoccupation de généralité² ni de rigueur mathématique pointilleuse. On cherchera plutôt à introduire quelques idées intuitives sur la décomposition-coordination (s'appuyant cependant sur des outils mathématiques classiques en optimisation comme la dualité) et à en développer des interprétations de type économique.

2. En particulier les fonctions objectif et les contraintes auront des formes additives (par rapport à la décomposition des vecteurs de variables en sous-vecteurs correspondant aux sous-systèmes), ou séparables, particularité dont on s'affranchira ultérieurement.

Dans le Chapitre 3, on introduit un formalisme général basé sur le « principe du problème auxiliaire » (PPA). Son objectif est de rendre compte de la plupart des méthodes de décomposition à partir d'un minimum de principes de base comme on le verra dans ce chapitre et dans le chapitre suivant. Ce faisant, on peut lever certaines restrictions (hypothèses de séparabilité ou d'additivité des fonctions objectif et des contraintes) qui semblaient pourtant essentielles dans le Chapitre 2. De plus, ce formalisme permet d'étudier la convergence des algorithmes itératifs de coordination dans un cadre unifié.

Alors que le Chapitre 3 est restreint à la décomposition de problèmes en sous-problèmes dont le couplage provient essentiellement d'une fonction objectif non séparable, le chapitre suivant (Chapitre 4) étend la technique du PPA à la situation où ce couplage provient aussi des contraintes en s'appuyant sur la théorie de la dualité et l'hypothèse de l'existence d'un point selle du Lagrangien associé au problème d'optimisation sous contraintes.

Le dernier chapitre de cette première partie (Chapitre 5) revient sur cette hypothèse d'existence d'un point selle du Lagrangien. Si cette existence est garantie dans un cadre convexe (moyennant quelques hypothèses techniques supplémentaires), elle peut plus facilement être garantie (au moins localement) par l'utilisation d'un *Lagrangien augmenté*. De plus, la convergence d'un algorithme primal-dual est notablement améliorée par cette technique. Cependant, le Lagrangien augmenté est une source de couplage supplémentaire entre sous-problèmes même dans le cas de fonctions séparables. Ce chapitre revient d'abord sur la théorie du Lagrangien augmenté et sa relation avec une technique de *régularisation* de la fonction duale. On montre ensuite comment l'utiliser pour la décomposition tout en conservant l'essentiel de ses avantages.

1.4.2 Partie « optimisation stochastique »

Dans la seconde partie, on commence par donner, au Chapitre 6, une brève présentation des différentes problématiques que l'on peut rencontrer en optimisation stochastique, en distinguant le cas « statique » dans lequel les variables de décision ne sont pas des variables aléatoires, et le cas « dynamique » pour lequel les variables de décision dépendent de la réalisation des aléas présents dans le problème. On ne se préoccupera que du cas de l'optimisation « statique » dans la suite de cette seconde partie, et plus spécifiquement de la méthode du gradient stochastique dont le principe repose sur le fait de marier au sein d'un même algorithme itératif l'idée d'estimation à la Monte Carlo (pour calculer une espérance) et l'idée de descente de gradient (pour minimiser une fonction).

Dans le Chapitre 7, on présente la vision d'ensemble de la méthode du gradient stochastique. On s'intéresse donc au cadre probabiliste adapté à son étude, on donne les résultats classiques de convergence et de vitesse de convergence pour l'algorithme du gradient stochastique et sa version « moyennée », et on décrit certains aspects pratiques de mise en œuvre de ces méthodes.

Dans le Chapitre 8, on s'intéresse au mélange de l'idée du « principe du problème auxiliaire » (PPA) et de celle du gradient stochastique, dans le cas où le couplage provient de la fonction objectif et non des contraintes, et on donne un résultat de convergence dans un cadre unifié.

Puis, dans le Chapitre 9, on cherche à étendre la technique du PPA au cas des problèmes d'optimisation stochastique sous contraintes déterministes, en s'appuyant sur la technique de dualité. On montre alors que les algorithmes de type « Uzawa » n'ont pas d'équivalents dans le cadre stochastique, alors que les algorithmes de type « Arrow-Hurwicz » s'étendent bien dans ce cadre.

Enfin, dans le Chapitre 10, on montre comment utiliser les méthodes présentées au chapitre précédent dans le cas où les contraintes déterministes portent sur l'espérance d'une fonction aléatoire. On conclut ce chapitre en indiquant des pistes permettant de prendre en compte des contraintes en probabilité.

1.4.3 Prérequis du texte

La lecture de cet ouvrage suppose de la part du lecteur une certaine familiarité avec les notions de base en optimisation déterministe et stochastique.

Dans la première partie consacrée à l'optimisation déterministe, le cadre est celui de l'analyse convexe. En optimisation, il est fait usage de la théorie locale des conditions nécessaires d'optimalité du premier ordre dans le cas de fonctions différentiables (conditions de Lagrange-Kuhn-Tucker), et éventuellement de celles du second ordre, ainsi que de la théorie globale des conditions suffisantes (Lagrangien, dualité et fonction duale, point selle, etc.). L'interprétation marginaliste des multiplicateurs de Lagrange-Kuhn-Tucker et sa relation avec la « fonction perturbation » (coût optimal comme fonction du second membre des contraintes) joueront un rôle important dans la compréhension intuitive des méthodes de décomposition-coordination. Pour toutes ces notions, on pourra par exemple consulter le cours « *Convexité et Optimisation* » COHEN (2000) ainsi que les références bibliographiques qui y sont mentionnées. Des exercices sont proposés au §2.4 pour revisiter certaines d'entre elles. L'annexe 4.4 fournira également quelques rappels à ce sujet. Les annexes 3.7 à 3.9 introduiront quelques notions techniques importantes. Le lecteur peut éventuellement commencer la lecture des chapitres correspondants par celle de ces exercices et annexes.

Dans la seconde partie consacrée à l'optimisation stochastique, l'accent est essentiellement mis sur les aspects « optimisation », et l'on fait donc usage des mêmes théories que dans la première partie de l'ouvrage (conditions d'optimalité, dualité). Il est aussi nécessaire d'avoir une certaine maîtrise de la théorie des probabilités, en particulier en ce qui concerne les notions d'espérance et d'espérance conditionnelle, les notions de convergence des variables aléatoires, ainsi que les propriétés liées à la simulation et à l'estimation des variables aléatoires par la méthode de Monte Carlo. On pourra consulter la référence « *Probabilités de l'ingénieur* » BOULEAU (1986).

Décomposition-coordination en optimisation
déterministe et stochastique

Carpentier, P.; Cohen, G.

2017, XVII, 333 p. 26 ill., 23 ill. en couleurs., Softcover

ISBN: 978-3-662-55427-2