

Programmation stochastique

Fabian Bastin

IFT-6512 – Hiver 2011

Objectif principal : introduction à la programmation stochastique, qui est une manière d'optimiser, au sens mathématique du terme, tout en tenant compte de l'incertitude.

Plan

- incertitude et optimisation ;
- optimisation stochastique avec recours ;
- programmes stochastiques linéaires à deux étapes, formulations compactes et explicites, méthode L-Shaped ;
- problèmes à plusieurs étapes ;
- problèmes nonlinéaire convexes et non-convexes ;
- méthodes échantillonnales et étude de consistance.

Prérequis

- solide bagage mathématique et statistique ;
- fortement conseillé d'avoir suivi le cours IFT-6504 sur la programmation mathématique ;
- connaissance de base du C++.

Horaire

- lundi de 12h30 à 14h30 ;
- mercredi de 12h30 à 14h30.

Livre de base

- Peter Kall et Stein W. Wallace, *Stochastic Programming*, John Wiley & Sons, 1994.

Disponible en téléchargement à l'adresse

<http://www.iro.umontreal.ca/~bastin/books/KallWall194.pdf>

Livres conseillés

- John R. Birge et François Louveaux, *Introduction to Stochastic Programming*, Springer-Verlag, 1997.
- Andrzej Ruszczyński et Alexander Shapiro, *Stochastic Programming*, Elsevier, 2003.

Logiciels

- GAMS (<http://www.gams.com>),
- COIN-OR (<http://www.coin-or.org>),
- ORATIO (<http://www.slashbin.net/amlet>).

Les modalités d'évaluation précises seront discutées en classe, mais elle se répartira comme suit :

- 4 devoirs, chacun comptant pour 10% de la note finale ;
- un examen intra, comptant pour 20% ;
- une présentation d'article (récent), comptant pour 40%.

Objectifs : introduction à la programmation stochastique, i.e.

- apprendre les termes, possibilités et limitations de base des modèles de programmation stochastique ;
- apprendre à formuler des modèles analytiques incorporant de l'incertitude sous forme de programmes stochastiques ;
- apprendre la théorie de base ;
- apprendre les techniques algorithmiques principales mise en oeuvre pour résoudre les problèmes résultants.

Disponible sur rendez-vous à mon bureau (3367).

courriel : `bastin@iro.umontreal.ca`

Les transparents servant de support au cours sont disponibles à l'adresse `http:`

`//www.iro.umontreal.ca/~bastin/Cours/IFT6512` et sont susceptibles d'être mis à jour.

Lecture suggérée : chapitre 1 de Kall & Wallace

Partons du cas déterministe, avec le programme général

$$\begin{aligned} \min & g_0(x) \\ \text{t.q.} & g_i(x) \leq 0, i = 1, \dots, m \\ & x \in X \subset \mathbb{R}^n. \end{aligned}$$

Tout les paramètres sont supposés être connus exactement.

Réaliste ?

- erreurs de mesure ;
- incertitudes sur le futur ;
- données non disponibles. . .

Programmation mathématique et programmation stochastique

- **Programmation mathématique** (optimisation) :
typiquement : problème de prise de “décision” (le mot décision est à prendre au sens large).
- La **programmation stochastique** concerne la prise de décision sous incertitude.
- On peut voir la programmation stochastique comme de la programmation mathématique avec des paramètres aléatoires.

$$\begin{aligned} & \text{” min ” } g_0(x, \xi) \\ & \text{t.q. } g_i(x, \xi) \leq 0, \quad i = 1, \dots, m, \end{aligned}$$

où ξ est un vecteur aléatoire. Sens de “min” ?

- Hypothèse fondamentale : nous connaissons une distribution de probabilité (jointe).
- Cela peut paraître limitatif, et nous n'aborderons pas ici en détails les techniques pour sélectionner cette distribution, mais
 - il n'est pas toujours nécessaire de connaître l'entièreté de la distribution jointe (on s'intéresse généralement seulement à l'impact dû à certaines variables) ;
 - une spécification "subjective", imparfaite, de la distribution jointe peut déjà donner une information utile.
 - il faut par conséquent être capable d'identifier l'origine de l'incertitude.
- Si vraiment rien ne peut être dit sur la probabilité, on peut essayer de traiter l'incertitude de manière différente, comme la programmation floue (qui ne sera pas abordée ici).

Un premier exemple

Pour ce faire, introduisons l'**approche par scénarios**.

- L'approche par scénarios est une approche d'apparence raisonnable pour gérer l'incertitude dans de nombreux cas, bien que ce ne soit pas la seule manière de faire (nous en verrons d'autres).
- L'approche par scénarios suppose qu'il y a un ensemble fini d'observations qui peuvent survenir, autrement dit les événements aléatoires se présentent suivant des distributions discrètes. Chacune des réalisations possibles est un scénario. Exemples :
 - la demande pour un produit est "basse, moyenne, ou élevée" ;
 - la temps est humide ou sec ;
 - la bourse va monter ou baisser.
- Même si la nature agit d'une manière continue, une approximation discrète est souvent utile.

Le problème du fermier

Source : Birge et Louveaux, Section 1.1.

Un fermier européen dispose de 500 acres de terre et cultive du blé, du maïs et des betteraves sucrières.

Au moins 200T de blé et 240T de maïs sont nécessaires pour nourrir son bétail. Tout surplus peut être vendu, mais en cas de sous-production, le complément doit être acheté, avec un prix d'achat est 40% supérieur au prix de vente. Le fermier peut vendre la betterave pour \$36T pour les 6000 premières tonnes, et \$10T au delà, en raison des quotas européens.

Comment décider des surfaces à affecter à chaque plante ?

Le problème du fermier (2)

Culture	Blé	Maïs	Betterave
Rendement moyen	2.5	3	20
Coût de plantation (\$/acre)	150	230	260
Prix de vente (\$/T)	170	150	36 ($\leq 6000T$), 10
Prix d'achat (\$/T)	238	210	-
Minimum requis (T)	200	240	-

Notations :

- x_1, x_2, x_3 : acres pour le blé, le maïs, la betterave ;
- y_1, y_2 : tonnes de blé, maïs achetées ;
- w_1, w_2 : tonnes de blé, maïs vendues ;
- w_3, w_4 : tonnes de betteraves vendues, à prix favorable et à bas prix.

Le problème du fermier (3)

Problème (linéaire) :

$$\begin{aligned} \min & 150x_1 + 230x_2 + 260x_3 + \\ & 238y_1 - 170w_1 + 210y_2 - 150w_2 - 36w_3 - 10w_4 \\ \text{t.q.} & x_1 + x_2 + x_3 \leq 500; \\ & 2.5x_1 + y_1 - w_1 \geq 200; \\ & 3x_2 + y_2 - w_2 \geq 240; \\ & w_3 + w_4 \leq 20x_3; \\ & w_3 \leq 6000; \\ & x_1, x_2, x_3, y_1, y_2, w_1, w_2, w_3, w_4 \geq 0. \end{aligned}$$

Résolvons ce programme.

Le problème du fermier (3)

Profit total : \$118600. Détails :

Culture	Blé	Maïs	Betterave
Surface (acres)	120	80	300
Rendement (T)	300	240	6000
Ventes (T)	100	-	6000
Achats (T)	-	-	-

Les rendements réels sont cependant dépendants de la situation climatique. Suivant le climat, des rendements allant de 20% à 25% au-dessus ou en-dessous du rendement moyen ne sont pas inhabituels.

De manière (très) simplifiée, supposons 3 cas possibles : des années bonnes (rendement supérieur de 20% pour toutes les semences), des années moyennes et des années mauvaises (rendement inférieur de 20%). Les prix ne changent pas.

Le problème du fermier (4)

Nouvelles solutions optimales ?

Bonne année. Profit total : \$167667. Détails :

Culture	Blé	Maïs	Betterave
Surface (acres)	183.33	66.67	250
Rendement (T)	550	240	6000
Ventes (T)	350	-	6000
Achats (T)	-	-	-

Mauvaise année. Profit total : \$59950. Détails :

Culture	Blé	Maïs	Betterave
Surface (acres)	100	25	375
Rendement (T)	200	240	6000
Ventes (T)	-	-	6000
Achats (T)	-	180	-

Le problème du fermier (5)

Les décisions sont fortement sujettes au climat, mais comment le connaître à long terme, en particulier au moment de l'ensemencement ?

Les décisions (x_1, x_2, x_3) sur l'affectation des terres doivent être faites maintenant, mais les ventes et les achats $(w_i, i = 1, \dots, 4, y_j, j = 1, 2)$ dépendent des rendements.

Scenarios.

Index $s = 1, 2, 3$, désignant des rendements meilleurs que la moyenne, égaux à la moyenne, ou inférieurs à la moyenne, respectivement.

Nouvelles variables w_{is} et y_{is} .

Le problème du fermier (6)

On veut maximiser le **profit espéré**. En supposant que les 3 scénarios sont équiprobables, cela conduit au programme

$$\min 150x_1 + 230x_2 + 260x_3 +$$

$$+ \sum_{s=1}^3 \frac{1}{3} (238y_{1s} - 170w_{1s} + 210y_{2s} - 150w_{2s} - 36w_{3s} - 10w_{4s})$$

$$\text{t.q. } x_1 + x_2 + x_3 \leq 500;$$

$$3x_1 + y_{11} - w_{11} \geq 200; 2.5x_1 + y_{12} - w_{12} \geq 200; 2x_1 + y_{13} - w_{13} \geq 200;$$

$$3.6x_2 + y_{21} - w_{21} \geq 240; 3x_2 + y_{22} - w_{22} \geq 240;$$

$$2.4x_2 + y_{23} - w_{23} \geq 240;$$

$$w_{31} + w_{41} \leq 24x_3; w_{31} + w_{41} \leq 20x_3; w_{31} + w_{41} \leq 16x_3;$$

$$w_{31} \leq 6000; w_{32} \leq 6000; w_{33} \leq 6000;$$

$$x, y, w \geq 0.$$

Problème sous **forme étendue**.

Le problème du fermier (7)

Les décisions d'ensemencement seront dites **décisions de première étape**. Les décisions de vente et d'achat seront appelées **décisions de seconde étape**.

Profit total : \$108390. Détails :

	Culture	Blé	Maïs	Betterave
Première étape	Surface (acres)	170	80	250
$s = 1$	Rendement (T)	510	288	6000
	Ventes (T)	310	48	6000
	Achats (T)	-	-	-
$s = 2$	Rendement (T)	425	240	5000
	Ventes (T)	225	-	5000
	Achats (T)	-	-	-
$s = 3$	Rendement (T)	340	192	4000
	Ventes (T)	140	-	4000
	Achats (T)	-	48	-

Le problème du fermier : observations

La décision optimale n'est plus la même !!!

Si le fermier pouvait connaître à l'avance quel scénario se produirait, le profit moyen optimal annuel serait \$115406. On parle de décision sous **information parfaite**. La différence avec la décision optimale sous incertitude est dite **valeur espérée de l'information parfaite** : perte de profit due à l'incertitude.

Si au contraire, le fermier n'utilise que l'information moyenne, i.e. il remplace les valeurs aléatoires par leur espérance (ce qui se fait souvent en optimisation), le profit moyen serait de \$107240 (solution en valeur espérée), soit une perte de \$1150 par rapport à la solution du problème stochastique. Cette différence est connue sous le nom de **valeur de la solution stochastique**.

Formalisation du cadre de travail

Incertitude : représentation sous forme d'**éléments aléatoires**, dénotés individuellement par ω , et issus d'un ensemble Ω , l'espace échantillonal.

Un **événement** A est un sous-ensemble de Ω ; la collection des événements aléatoires est dénotée \mathcal{A} . L'événement $A \in \mathcal{A}$ survient si la sortie de l'expérience est un élément de A .

$\forall A \in \mathcal{A}$, nous associons une mesure de probabilité $P[A]$ t.q.

$$\begin{aligned}0 &\leq P[A] \leq 1, \\ P[\emptyset] &= 0, \quad P[\Omega] = 1, \\ P[A_1 \cup A_2] &= P[A_1] + P[A_2] \text{ if } A_1 \cap A_2 = \emptyset.\end{aligned}$$

Le triplet (Ω, \mathcal{A}, P) est appelé un espace de probabilité.

Une **variable aléatoire** ξ sur un espace de probabilité (Ω, \mathcal{A}, P) est une fonction à valeurs réelles $\xi(\omega)$ ($\omega \in \Omega$) telle que $\{\omega | \xi(\omega) \leq x\}$ est un événement pour tout x fini.
Dès lors, on peut associer une probabilité à $(\xi(\omega) \leq x)$.

ξ a une **fonction de répartition** $F_\xi(x) = P[\xi \leq x]$.

$\Xi \in \mathbb{R}^n$ est le **support** de ξ s'il est le plus petit sous-ensemble fermé dans \mathbb{R}^n tel que $P[\xi \in \Xi] = 1$.

Les **variables aléatoires discrètes** ne peuvent prendre qu'un nombre fini ou infini dénombrable de valeurs ξ_k , $k = 1, \dots, K$, avec une **masse de probabilité** $p_k = P[\xi = \xi_k]$, et $\sum_{k=1}^K p_k = 1$.

Variable aléatoire continue

Les **variables aléatoires continues** ont une densité $f(\xi)$.

- $P[\xi = x] = 0$.
- La probabilité pour ξ d'appartenir à l'intervalle $[a, b]$ est

$$\begin{aligned}P[a \leq \xi \leq b] &= \int_a^b f(\xi) d\xi \\ &= \int_a^b dF(\xi) \\ &= F(b) - F(a).\end{aligned}$$

- L'**espérance** de ξ est

$$\begin{aligned}E[\xi] &= \sum_{k=1}^K p_k \xi_k; \\ E[\xi] &= \int_{-\infty}^{\infty} f(\xi) d\xi = \int_{-\infty}^{\infty} dF(\xi).\end{aligned}$$

Un programme linéaire aléatoire

Considérons le problème (jouet)

$$\min_x x_1 + x_2$$

$$\text{t.q. } \omega_1 x_1 + x_2 \geq 7$$

$$\omega_2 x_1 + x_2 \geq 4$$

$$x_1, x_2 \geq 0,$$

où $\omega_1 \sim U[1, 4]$, $\omega_2 \sim U[1/3, 1]$.

Que faire ?

- Comment résoudre ce problème ?
- Que *signifie* résoudre ce problème ?
- Est-il possible de décider sur x *après* avoir observé la réalisation du vecteur aléatoire ω ?
On parle alors d'approche **wait-and-see**. Le problème est alors plus facile à résoudre (on a ici un simple programme linéaire).
- Mais cette approche est rarement adaptée !!! On doit généralement décider de x **avant** de connaître les valeurs de ω !

Que faire ? (2)

- Généralement, l'approche “wait-and-see” n'est pas appropriée pour modéliser le fonctionnement de la réalité : nous devons décider sur x avant de connaître les valeurs de ω .
- Dans le cadre de l'exemple, nous devons décider que faire sans connaître les réalisations ω_1, ω_2 .
- 3 suggestions :
 - 1 essayer d'estimer, de prédire l'incertitude ;
 - 2 contraintes en probabilité ;
 - 3 pénaliser les déviations.