METHODES

La programmation mathématique positive dans les modèles d'exploitation agricole

Principes et importance du calibrage

Alexandre GOHIN
Frédéric CHANTREUIL

Cahiers d'économie et sociologie rurales, n° 52, 1999

Alexandre GOHIN*, Frédéric CHANTREUIL*

Positive mathematical programming in agricultural economics. Principles and importance of calibrating

Key-words:
positive mathematical programming, agricultural production, land allocation, calibration procedure

Summary *– The modelling of agricultural producer behaviour using mathematical programming has a long tradition in agricultural economics. The linear mathematical programming approach has been prevalent in this field for a long time. But linear programming models, that are tightly constrained to reproduce agricultural producers' choices observed at the base period, are often unacceptable and also inappropriate under policy changes. Several researchers have alluded to this problem in the past and came up with several solutions such as incorporating risk or considering "flexibility" constraints. Furthermore, to solve this problem, new methodological developments also occurred, including Positive Mathematical Programming (PMP). PMP emerged more than ten years ago but its formal presentation is relatively recent. In this article, we first present the principles of PMP, using a simple example of an arable crop producer. It appears that the two main advantages of PMP are its perfect calibration to base period levels of endogenous variables and its derivation of smooth simulation results, both resulting from the incorporation of non linear terms in the objective function. Empirical implications of the standard PMP's parameter calibration process are then discussed.*

La programmation mathématique positive dans les modèles d'exploitation agricole. Principes et importance du calibrage

Mots-clés:
programmation mathématique positive, production agricole, allocation de la terre, processus de calibrage

Résumé – La modélisation du comportement des agriculteurs par la programmation mathématique a une longue tradition en économie agricole. La Programmation mathématique linéaire (PML) a longtemps prédominé mais la Programmation mathématique positive (PMP) s'impose maintenant de plus en plus. La différence essentielle de la PMP par rapport à la PML réside dans la spécification de fonctions non linéaires, permettant ainsi de calibrer de manière exacte les modèles et d'éviter les discontinuités caractérisant les modèles PML et leurs résultats. Cet article présente tout d'abord les principes et la mise en œuvre standard de la PMP. La discussion est ensuite centrée sur un aspect de cette approche, à savoir le calibrage des paramètres de comportement.

* *INRA, Unité d'économie et sociologie rurales, rue Adolphe Bobierre, CS 61103, 35011 Rennes cedex.*
e-mail : gohin@roazhon.inra.fr ; chantreuil@roazhon.inra.fr

Les auteurs tiennent à remercier ici la Direction générale de l'Agriculture de la Commission européenne pour son soutien financier dans le cadre du projet FAIR 5-CT97-3481. Ils remercient également les deux lecteurs anonymes de la revue pour leurs remarques et suggestions. Ils sont cependant seuls responsables du contenu de l'article.

LA modélisation du comportement des agriculteurs à l'offre de produits et/ou à la demande dérivée d'intrants par la programmation mathématique est une longue tradition en économie agricole[1]. De manière très générale, les modèles de programmation mathématique appliqués au niveau de l'exploitation agricole individuelle consistent à déterminer les niveaux des variables de décision de cette exploitation qui maximisent une variable économique sous des contraintes techniques. La variable économique maximisée est généralement le profit et la marge brute de l'exploitation, plus rarement les recettes brutes ou l'opposé des coûts. Les contraintes techniques définissent implicitement un ensemble de production convexe par rapport aux variables de décision de l'exploitation. Ces modèles de programmation mathématique permettent alors de représenter le fonctionnement technico-économique des exploitations agricoles et de simuler les impacts de chocs exogènes (un changement de politique agricole par exemple) sur leurs variables de décision.

Si les fondements de la programmation mathématique sont relativement simples et cohérents avec la théorie micro-économique néoclassique du producteur, la mise en œuvre de cette approche est nettement plus délicate. Deux problèmes majeurs se posent pour le modélisateur qui dispose souvent d'un ensemble limité d'informations : la spécification de la fonction objectif et des contraintes techniques, d'une part, le calibrage des paramètres introduits dans ces fonctions, d'autre part. La programmation mathématique linéaire (PML) est un cas particulier de cette modélisation où la fonction objectif et les contraintes techniques sont spécifiées de manière linéaire par rapport aux variables de décision. Cette approche linéaire, longtemps prédominante en économie agricole, offre l'avantage d'être plus facile à résoudre que les approches non linéaires[2]. De plus, la spécification linéaire permet de borner le nombre de paramètres à calibrer à $(m + 1)$ fois le nombre de variables de décision, où m représente le nombre de contraintes techniques. Enfin et surtout, la PML représente des technologies de production de type Leontief ; les paramètres des contraintes techniques de production s'interprètent alors comme des coefficients inputs-outputs (Just *et al.*, 1983). Le calibrage déterministe de ces derniers peut alors être réalisé à partir d'informations techniques extérieures à la modélisation. Mais cette relative simplicité de

[1] A notre connaissance, les premiers modèles de programmation mathématiques en économie agricole datent du début des années 50 avec les travaux de King (1953) et de Heady (1954).

[2] Cet argument en faveur de l'utilisation de la PML est aujourd'hui moins pertinent.

mise en œuvre de la PML a plusieurs revers. En particulier, cette approche ne permet généralement pas de reproduire exactement les prises de décision observées des agriculteurs, sauf si des contraintes techniques très sévères « figeant » le modèle sont introduites. Par ailleurs, la simulation de scénarios avec la PML entraîne soit aucun changement, soit des « basculements » importants dans les décisions des agriculteurs. En d'autres termes, les modèles de PML produisent des résultats discontinus et peuvent conduire à des spécialisations extrêmes des exploitations agricoles dans certaines productions.

Ces problèmes de la PML ont déjà été mentionnés à plusieurs reprises (par exemple, McCarl, 1982 ; Hazell et Norton, 1986, Heckelei, 1997). De nombreux développements de la modélisation par la programmation mathématique ont alors cherché à dépasser ces problèmes, et ont donné progressivement naissance à la Programmation mathématique positive (PMP)[3]. Bien qu'étant appliquée depuis plus de dix ans, la présentation formelle de cette méthode par Howitt est relativement récente (Howitt, 1995a) et depuis, de nombreuses études ou articles finalisés s'appuient sur cette méthode[4]. La PMP permet de calibrer de manière exacte les modèles d'exploitation agricole en utilisant un ensemble de données restreint tout en ne figeant pas le modèle. La différence essentielle de la PMP par rapport à la PML réside dans la spécification de fonctions non linéaires qui permettent alors de reproduire une situation observée et de « lisser » les résultats de scénarios. La non-linéarité a été principalement introduite dans la fonction objectif du profit au niveau de la fonction de production (Howitt, 1995a) ou des coûts de production (Arfini et Paris, 1995).

Dans cet article à vocation méthodologique, les objectifs sont doubles. Le premier objectif est de préciser les fondements théoriques et d'expliciter la mise en œuvre « standard » de cette approche PMP. Ceci nous permet en particulier de souligner les apports de celle-ci en économie de la production agricole. A partir de cette présentation, le second objectif de cet article est de souligner et de discuter une des limites de cette mise en œuvre standard. Cette limite provient de la méthode de calibrage des paramètres de comportement. Les principales solutions proposées dans la littérature pour résoudre ce problème sont expliquées avant d'en suggérer une nou-

[3] Une autre extension de la programmation mathématique concerne l'introduction du risque au niveau des rendements par hectare et/ou des prix des cultures et/ou des aides directes. Dans ce cadre, l'exploitation agricole ne maximise pas son profit mais l'utilité espérée de son profit (voir Boussard *et al.*, 1997). Une discussion de cette approche dépasse le cadre de cet article.

[4] Voir, entre autres, Heckelei et Britz (1999) ; Garvey et Steele (1998) ; Paris et Howitt (1998) ; Barkaoui et Butault (1998) ; Judez *et al.* (1999, 1998) ; Rhöm *et al.* (1997) ; Carles *et al.* (1998) ; Britz et Heckelei (1997) ; Arfini (1996) ; Arfini et Paris (1995) ; Howitt (1995b).

velle. A l'inverse des premières, cette dernière est basée sur l'utilisation de la seule information disponible dans les bases de données. Précisons à ce stade qu'il ne s'agit pas ici de fournir une présentation exhaustive de la littérature sur la PMP, ni de discuter de tous les problèmes inhérents à cette méthodologie mais plutôt d'en détailler un aspect.

L'article est organisé de la façon suivante : les principes et la mise en œuvre classique de la PMP sont présentés à l'aide d'un exemple simple représentant le comportement d'un agriculteur à l'offre de produits de grandes cultures dans une première partie. Cet exemple est utilisé dans la suite de l'article pour illustrer nos propos. La seconde partie est consacrée au problème du calibrage des paramètres de comportement.

PRINCIPES ET MISE EN ŒUVRE DE LA PROGRAMMATION MATHÉMATIQUE POSITIVE

Pour présenter les principes et la mise en œuvre classique de la PMP, nous considérons le cas réel d'une exploitation agricole française spécialisée en grandes cultures[(5)]. Les données de cette exploitation, reportées dans l'annexe 1, sont issues du Réseau d'information comptable agricole (RICA) et sont relatives à l'année 1997. Les informations qu'il est possible d'obtenir aisément à partir de cette base de données sont limitées aux surfaces consacrées à chaque culture (l_i), aux quantités produites (y_i) et aux valeurs des productions ($p_i.y_i$) de chaque culture. Peuvent être alors déduits de ces informations les rendements à l'hectare (r_i) et les prix à la production (p_i) pour chaque culture. Les aides directes à l'hectare (s_i) sont supposées également facilement observables. Par contre, les coûts de production imputables à chaque culture ne sont pas renseignés dans cette base de données mais seulement des données sur les différents coûts variables de production pour l'ensemble des cultures (par exemple, les coûts de production en engrais ou en produits phytosanitaires). Enfin, un coût total de production de végétaux inclut l'ensemble de ces coûts variables de production. Nous supposons dans le reste de cet article que le modélisateur ne dispose que de ces données pour élaborer un modèle de PMP. Nous supposons par ailleurs que l'objectif de ce modèle est de simuler le comportement de cette exploitation agricole dans ses choix d'allocation de sa surface totale et ses niveaux de production suite à un choc exogène, de politique agricole

(5) Pour une application de la PMP au secteur de l'élevage, voir Bauer et Kasnakoglu (1990).

par exemple. Par souci de simplification, nous supposons enfin que cette simulation est effectuée à court terme, *i.e.* à structure constante de l'exploitation (les niveaux de capital, de main-d'œuvre et de terre sont fixés). Même si elles peuvent paraître réductrices pour présenter la PMP, il convient de noter que ces deux dernières hypothèses sont adoptées dans de nombreuses études utilisant cette approche.

Les principes de la programmation mathématique positive

Dans le cadre usuel de la théorie micro-économique néoclassique du producteur, l'exploitation agricole choisit une allocation de sa surface totale et des niveaux de production qui maximisent son profit, étant donné les contraintes techniques de production, le système de prix et les instruments de politique agricole. Les prix sont supposés exogènes et les anticipations sur ces prix rationnelles. Au niveau de la politique agricole, les aides directes sont supposées parfaitement couplées au facteur terre. En ce qui concerne les contraintes techniques de production, l'exploitation agricole est tout d'abord limitée par sa surface totale ; la somme des surfaces cultivées est inférieure ou égale à cette surface totale.

De nombreuses autres contraintes techniques interviennent également dans le programme économique de l'exploitation agricole : contrainte agronomique, pédologique, climatique, de capital, de disponibilité de main-d'œuvre, etc. Introduire ces contraintes dans le modèle suppose des informations précises dont le modélisateur ne dispose généralement pas. Ces contraintes doivent cependant être prises en compte pour deux raisons principales. D'une part, sans ces contraintes, le modèle ne pourra pas reproduire les choix initiaux de production de l'exploitation. D'autre part, lors de la phase de simulation, un choc donné peut produire des changements importants dans les assolements, voire des « basculements » d'une culture à une autre ; par exemple, une baisse modérée de la marge d'une culture peut conduire à sa suppression. Ces problèmes sont rencontrés avec les modèles de PML.

La première idée de la PMP par rapport à la PML pour résoudre ce problème de spécification des « autres » contraintes techniques est de représenter indirectement ces dernières dans la fonction objectif. Dans la pratique, une fonction de coût de production (Arfini et Paris, 1995) ou des fonctions de rendement à l'hectare (Howitt, 1995a) ou encore des fonctions de coût et de rendement (Barkaoui et Butault, 1999) sont spécifiées pour « résumer » dans un cadre statique ces autres contraintes techniques, qui peuvent être de nature dynamique. Nous supposons ici que les rendements à l'hectare sont exogènes et de ce fait, les décisions d'allocation de la terre sont identiques aux décisions de production. Les autres contraintes techniques sont donc incorporées dans une fonction de coût de production[6][7]. Notons à ce stade que cette hypothèse ne fait que déplacer le problème de spécification des autres contraintes techniques vers la fonction de coût de production. Elle n'assure, a priori, ni la reproduction de la situation initiale, ni un comportement lisse du modèle. En outre, cette procédure « dénature » en quelque sorte la programmation mathématique dans la mesure où l'ensemble convexe des productions possibles n'est plus uniquement décrit de manière primale par les contraintes techniques mais également de manière duale à travers la fonction de coût de production.

La deuxième idée de la PMP consiste à considérer que l'allocation choisie par l'exploitation agricole et observée par le modélisateur est une allocation optimale, c'est-à-dire qui maximise son profit étant donné ces contraintes (techniques, de prix et de politique agricole). Ainsi, les données observées peuvent servir de base au calibrage des paramètres spécifiés dans la fonction de coût de production. Généralement, seules les données d'une année de référence sont prises en compte par le modélisateur pour calibrer de manière déterministe ces paramètres. Cette procédure de calibrage garantit que le modèle reproduit l'allocation de la terre et les volumes produits de l'année de référence.

Enfin, la troisième et dernière idée de la PMP réside dans la spécification non linéaire, par rapport aux variables de décision, de la fonction de coût de production. Le profit de l'exploitation agricole est alors également non linéaire par rapport aux variables de décision. Dans ce cas, les profits marginaux à l'hectare, qui déterminent l'allocation optimale de la surface totale, sont fonction des surfaces allouées à chaque culture. Cette spécification non linéaire assure un comportement « lisse » du modèle. Dans la pratique, afin de limiter le nombre de paramètres à calibrer, la fonction de coût de production est souvent supposée être égale à la somme des fonctions de coûts variables de production par culture, ces dernières étant généralement spécifiées comme des fonctions quadratiques des volumes produits[8].

[6] L'autre solution adoptée dans certaines applications et qui consiste à supposer des coûts de production constants et à spécifier des fonctions de rendements à l'hectare n'est pas retenue ici car les coûts de production par culture ne sont pas observables pour le modélisateur.

[7] Cette fonction de coût n'inclut pas les coûts fixes car le raisonnement est mené à court terme.

[8] Cette spécification a certes le mérite de la parcimonie mais impose de fortes hypothèses sur les jointures entre les différentes productions. En particulier, elle implique que la jointure provienne uniquement de l'existence d'un facteur fixe allouable (la terre) (Guyomard *et al.*, 1996). Il est plus rarement spécifié une fonction de coût de production multi-cultures. En effet, l'introduction de termes croisés dans la fonction de coût total augmente considérablement le nombre de paramètres et rend plus difficile le calibrage de ceux-ci (Paris et Howitt, 1998). D'autres formes fonctionnelles ont été testées (Arfini et Paris, 1995).

La mise en œuvre standard de la programmation mathématique positive

L'écriture mathématique d'un modèle de programmation mathématique positive

Le modèle « standard » de PMP pour l'exploitation agricole décrite dans l'annexe 1 s'écrit de la manière suivante :

$$\begin{aligned}&\max_{l_i, y_i} \sum_{i=b,m,p} \left(p_i y_i + s_i l_i - C_i(y_i)\right) \\ &= \max_{l_i} \sum_{i=b,m,p} \left(p_i r_i l_i + s_i l_i - C_i(r_i l_i)\right)\end{aligned} \tag{1}$$

sous la contrainte :

$$\sum_{i=b,m,p} l_i \leq \bar{l} \tag{2}$$

L'équation (1) définit le profit de l'exploitation agricole. Il est égal à la somme des productions en valeur plus les aides directes aux surfaces moins la somme des coûts variables de production par culture. Ce coût variable par culture est supposé être une fonction non linéaire de la surface allouée à celle-ci. L'inégalité (2) correspond à la contrainte de terre disponible. Cette contrainte implique que la somme des surfaces allouées aux différentes cultures ne peut être supérieure à la surface totale de l'exploitation.

Les conditions du premier ordre de ce programme d'optimisation sont les suivantes :

$$\begin{aligned}&p_i r_i + s_i - \frac{\partial C_i(r_i l_i)}{\partial l_i} - \alpha \leq 0, \\ &l_i \geq 0, \left(p_i r_i + s_i - \frac{\partial C_i(r_i l_i)}{\partial l_i} - \alpha\right) l_i = 0, \quad \forall\ i = b,\ m,\ p\end{aligned} \tag{3}$$

$$\sum_{i=b,m,p} l_i - \bar{l} \leq 0, \quad \alpha \geq 0, \left(\sum_{i=b,m,p} l_i - \bar{l}\right) \alpha = 0 \tag{4}$$

où α désigne le multiplicateur de Lagrange associé à la contrainte de disponibilité de terre. Ce multiplicateur mesure l'impact sur le profit d'une variation marginale de la surface totale de l'exploitation ; il peut s'interpréter comme le prix d'opportunité de la terre pour l'exploitation considérée.

Les équations (3) signifient que la surface allouée à la culture i est strictement positive uniquement lorsque le profit marginal à l'hectare dégagé par cette culture est égal au prix d'opportunité de la terre. Cette surface peut être également nulle dans ce cas. La surface allouée à une culture est par contre nulle lorsque ce profit marginal est strictement inférieur au prix d'opportunité de la terre. De la même manière, d'après les équations (4), le prix d'opportunité de la terre est strictement positif uniquement lorsque la somme des surfaces allouées aux différentes cultures est égale à la surface totale disponible.

Pour résoudre ce modèle, il reste à spécifier la forme des fonctions de coût variable de production par culture et calibrer les paramètres de celles-ci. L'éventail des formes possibles est bien évidemment large. Pour des raisons de commodité, la forme quadratique est usuellement utilisée. Nous supposons ici que ces fonctions s'écrivent :

$$C_i(y_i) = b'_i y_i = C_i(r_i l_i)\ b_i l_i^2, \quad \forall i = b, m, p \tag{5}$$

Selon cette spécification, le coût marginal de production est une fonction croissante du volume produit et également de la surface allouée à la culture considérée. Le problème qui se pose alors est de calibrer les paramètres b_i ou b'_i, sachant que les valeurs des coûts variables de production par culture ne sont pas observées.

Le calibrage des paramètres des fonctions de coût

La méthode de calibrage s'effectue en deux étapes. Dans une première étape, le modèle de PML constitué des équations (1), (2) est résolu en supposant que les paramètres des fonctions de coût variable de production sont nuls et en introduisant des contraintes spécifiques de calibrage. Ces contraintes assurent que les surfaces allouées à chaque culture sont inférieures ou égales aux surfaces initialement observées plus un petit terme positif. Ce dernier est introduit pour éviter la dépendance linéaire entre la contrainte de disponibilité de la terre et les contraintes spécifiques de calibrage. La résolution de ce modèle détermine des valeurs duales qui fournissent une mesure des coûts marginaux de production de chaque culture. Dans une seconde étape, ces valeurs duales sont utilisées pour calibrer les paramètres des fonctions de coût variable de production. La mise en œuvre de cette procédure de calibrage est plus amplement détaillée sur la base de l'exploitation décrite dans l'annexe 1. La première étape consiste à résoudre le programme suivant :

$$\max_{l_i} \sum_{i=b,m,p} (p_i r_i + s_i) l_i = \max 9745 l_b + 9555 l_m + 9585 l_p \tag{6}$$

sous les contraintes :

$$l_b + l_m + l_p \leq \bar{l} = 107 \tag{7}$$

$$l_b \leq l_b^0 + \varepsilon = 63{,}66 \tag{8.i}$$

$$l_m \leq l_m^0 + \varepsilon = 28{,}21 \tag{8.ii}$$

$$l_p \leq l_m + \varepsilon = 15{,}16 \tag{8.iii}$$

où l_i^0 désigne la surface initiale allouée à la culture i et $\varepsilon = 0{,}01$. Les variables duales associées aux contraintes (8.i), (8.ii) et (8.iii) sont respectivement λ_b, λ_m et λ_p. La résolution de ce modèle linéaire conduit à la solution optimale suivante :

$$l_b = l_b^0 + \varepsilon = 63{,}66\ ;\ l_m = l_m^0 - 2\varepsilon = 28{,}19\ ;\ l_p = l_p^0 + \varepsilon = 15{,}16\ ;$$
$$\alpha = p_m r_m + s_m = 9555\ ;$$
$$\lambda_b = p_b r_b + s_b - p_m r_m - s_m = 190\ ;\ \lambda_m = 0 \text{ et}$$
$$\lambda_p = p_p r_p + s_p - p_m r_m - s_m = 30.$$

Le prix d'opportunité de la terre est égal à la recette brute dégagée par la culture d'un hectare de maïs grain. En effet, si la surface totale varie d'un faible montant, alors la surface en maïs grain varie d'un même montant tandis que les autres surfaces restent inchangées à cause des contraintes spécifiques de calibrage. L'impact sur le profit de cette variation de la surface totale est par conséquent bien égal à la « recette brute maïs grain ». λ_b mesure l'intérêt, en terme de recette brute totale, de l'augmentation marginale de la surface allouée à la culture de blé tendre. Elle est égale à la différence entre la recette brute blé tendre et la recette brute maïs grain. De même, λ_p mesure l'intérêt de la culture de pois protéagineux et est égale à la différence entre la recette brute protéagineux et la recette brute maïs grain. Enfin, la dernière variable duale est nulle car la contrainte (8.ii) n'est pas saturée. Ces trois variables duales fournissent une mesure des coûts marginaux de production de chaque culture qui permettraient d'obtenir, à partir d'un modèle de programmation mathématique sans les contraintes spécifiques de calibrage ci-dessus, les allocations observées initialement. Ces variables duales fournissent donc de l'information pour calibrer les paramètres des fonctions de coût variable de production.