INFORMATICA, 1996, Vol. 7, No. 3, 337–348

ON THE DEGREE OF DECOMPOSITION IN LINEAR PROGRAMMING

James K. HO

Information and Decision Sciences, University of Illinois at Chicago
m/c 294, 601 South Morgan, Chicago, IL 60607 USA
E-mail: jim.ho@uic.edu

Etienne LOUTE

Center for Operations Research and Econometrics
B-1348, Louvain-la-Neuve, Belgium

Abstract. In the application of Dantzig–Wolfe decomposition to block-angular linear programming problems with R natural blocks, it is possible to have from 1 to R subproblems structurally while solving all R independent subproblems computationally. Early literature on the topic was inconclusive regarding the relative merits of such formulations. This paper attempts clarification by characterizing the significance of the degree of decomposition as well as presenting extensive empirical results.

Key words: linear programming, decomposition, computational mathematical programming.

1. Introduction. Many resource allocation problems involve semi-autonomous subsystems (e.g., divisions in a firm, regions in an economy) competing for limited, shared resources. When such problems can be modeled as a linear program to optimize some system-wide objective, we have what is known as the block-angular structure with the resource sharing expressed as coupling constraints. The Decomposition Principle of Dantzig and Wolfe (1960) provides both an algorithmic approach to solving such problems by a sequence of smaller subproblems, and an economic procedure of price-directive decentralization. Essentially, a system coordinator has the job of setting prices on the shared resources. At any given level of prices, each subsystem is asked to determine its own optimal activities. Such activities can be translated into actual utilization of the shared resources and submitted to the coordinator as

proposals. Weighing the proposals from all the subsystems, the coordinator can attempt to improve on the system-wide objective by determining the next round of prices. The Decomposition Principle shows that this is a finite process converging to equilibrium prices on the shared resources. It is well known that decentralization cannot be achieved by price-direction alone because the unilaterally optimal activities of the subsystems based on the equilibrium prices are not feasible system-wide in general. System-wide feasibility can only be enforced by the coordinator. For this reason, early literature on the topic settled for letting the coordinator determine the final course of actions for the subsystems, making the interpretation of decentralization somewhat awkward. It is perhaps less well known (see e.g. Dantzig 1963, Ch. 23) that a better alternative is for the coordinator simply to compute the optimal allocation of the shared resources based on the equilibrium prices. The subsystems can then determine their own activities with the given allocation. This way, the subsystems are told what their shares of the common resources are rather than what they must do, which should convey a better sense of autonomy and decentralization. It turns out that this scheme is also key to the analysis of the degree of decentralization. As it is always possible to aggregate subsystems, the question is how more or less decentralization affects the process of coordination. Early literature (Labro, 1964; Lasdon 1970, Ch. 3; Madsen, 1973) was inconclusive. This paper attempts to clarify this issue and to draw some conclusions from extensive empirical evidence.

2. Block-angular linear programs. We consider decision processes in resource allocation that can be modeled as block-angular linear programs (BLP). Such a problem with R blocks has the following form in which vectors and matrices are denoted in bold-face.

$$(P)\qquad \begin{aligned} &\text{maximize} \sum_{r=1,\ldots,R} \mathbf{c}_r\mathbf{x}_r, && (1)\\ &\text{subject to} \sum_{r=1,\ldots,R} \mathbf{A}_r\mathbf{x}_r = \mathbf{b}_0, && (2)\\ &\qquad\qquad \mathbf{B}_r\mathbf{x}_r = r, \quad r = 1,\ldots,R, && (3)\\ &\qquad\qquad \mathbf{x}_r \geqslant \mathbf{0}, \quad r = 1,\ldots,R, && (4) \end{aligned}$$

where $\mathbf{c}_r$ is $1 \times n_r$, $\mathbf{b}_r$ is $m_r \times 1$ and all other vectors and matrices are of compatible dimensions. Each index r from 1 to R represents one of the R

subsystems with the corresponding vector of activities (decision variables) $\mathbf{x}_r$. The system-wide objective function, given by (1), is the sum of linear objective functions $\mathbf{c}_r\mathbf{x}_r$ for the R subsystems. $\mathbf{b}_0$ is the vector of shared resources. The consumption matrix $\mathbf{A}_r$ transforms activities in subsystem r into consumptions of shared resources. The balance of system-wide consumption of shared resources is expressed in (2). Each subsystem has its own technology matrix $\mathbf{B}_r$ and demand vector $\mathbf{b}_r$, defining local constraints through (3). Finally, the activity levels are required to be non-negative. The problem is then to determine activity levels for all the subsystems that are feasible within the respective subsystem, that collectively satisfy the balance of shared resources, and that maximize the system-wide objective function. Or, letting each subsystem optimize its own objective while vying for the shared resources, an optimal allocation of the latter is sought to maximize the sum total of the objectives.

3. Dantzig–Wolfe decomposition. First, we summarize the Dantzig–Wolfe (1961) decomposition algorithm (see Ho (1987) for recent development in this approach). Then its interpretation as decentralized planning will be given. Let $F_r = \{\mathbf{x}_r \mid \mathbf{B}_r\mathbf{x}_r = \mathbf{b}_r, \mathbf{x}_r \geqslant 0\}$ be the set of all feasible solutions to subsystem r; and $X_r = \{\mathbf{x}_r^k;\ k = 1, \ldots, Kr \mid \mathbf{x}_r^k$ is an extreme point of $F_r\}$ be the set of extreme points of F_r. For simplicity, we assume that F_r is nonempty and bounded. It is then a bounded polyhedral convex set. Using the fact that any point in such a set can be represented by a nonnegative convex combination of its extreme points, one can rewrite (P) in the following equivalent extremal form:

$$
(E)\qquad
\begin{aligned}
&\text{Max} \sum_{r=1,\ldots,R} \sum_{i=1,\ldots,Kr} (\mathbf{c}_r\mathbf{x}_r^i)\lambda_{ri},\\
&\text{s.t} \sum_{r=1,\ldots,R} \sum_{i=1,\ldots,Kr} (\mathbf{A}_r\mathbf{x}_r^i)\lambda_{ri} = \mathbf{b}_0,\\
&\qquad \sum_{i=1,\ldots,Kr} \lambda_{ri} = 1, \quad r = 1, \ldots, R,\\
&\qquad \lambda_{ri} \geqslant 0, \quad i = 1, \ldots, Kr,\ r = 1, \ldots, R, \quad \text{where} \quad \mathbf{x}_r^i \varepsilon X_r.
\end{aligned}
$$

Thus solving (E) is equivalent to solving (P). Since K is often very large and not known a priori, a restriction strategy is applied to solve (E). The Dantzig–Wolfe decomposition algorithm uses a subset of Jr extreme points in X_r to formulate a (Restricted) Master Problem, say,

$$
(M^k)\qquad
\begin{aligned}
&\text{Max } \mathbf{z}^k = \sum_{r=1,\ldots,R}\ \sum_{i=1,\ldots,Jr} (\mathbf{c}_r \mathbf{x}_r^i)\lambda_{ri}, && \text{Dual Variables}\\
&\text{s.t.} \quad \sum_{r=1,\ldots,R}\ \sum_{i=1,\ldots,Jr} (\mathbf{A}_r \mathbf{x}_r^i)\lambda_{ri} = \mathbf{b}_0, && \pi^k\\
&\qquad \sum_{i=1,\ldots,Jr} \lambda_{ri} = 1, \quad r = 1,\ldots,R, && \sigma_r^k\\
&\qquad \lambda_{ri} \geqslant 0, \quad i = 1,\ldots,Jr,\ r = 1,\ldots,R.
\end{aligned}
$$

During the k-th cycle of the algorithm, (M^k) is solved. Let $(\pi^k, \sigma_1^k, \ldots, \sigma_R^k)$ be the optimal dual solution. The vector π^k is known as the *price vector* corresponding to the coupling constraints. To see if the objective can be further improved by introducing any extreme point in X not yet included in M^k, the following subproblem is solved. This is essentially an implicit simplex pricing step using the prices π^k.

$$
(S_r^k)\qquad
\begin{aligned}
&\max \quad v_r^k = (\mathbf{c}_r - \pi^k \mathbf{A}_r)\mathbf{x}_r - \sigma_r^k,\\
&\text{s.t.} \quad \mathbf{B}_r \mathbf{x}_r = \mathbf{b}_r,\\
&\qquad \mathbf{x}_r \geqslant \mathbf{0}.
\end{aligned}
$$

If $v_r^k \leqslant 0$, then nothing in F_r can improve the objective. If this holds for all r, then the solution of (M^k) is optimal for (E). Otherwise a new extreme point, called a proposal, with $v_r^k > 0$ can be included in (M^{k+1}) for an improved solution in the next cycle. Convergence is finite as each X_r is a finite set (see Dantzig (1963)). Furthermore, if ρ^k is the dual optimal solution to (S_r^k), then $(\pi^k, \rho_1^k \ldots, \rho_R^k)$ is a dual feasible solution to (P) and hence provides an upper bound Z^k on the maximum value of $\mathbf{z}$. Therefore, in practice, we can stop the decomposition algorithm if $|Z^k - z^k| < \varepsilon$ for some $\varepsilon \geqslant 0$.

4. Decentralized resource allocation. The Dantzig–Wolfe decomposition algorithm can be interpreted as a process of price-directive coordination of the subsystems which are allowed to optimize their own objectives while vying for the shared resources. The master problem (M^k) plays the role of coordinator who sets the prices π^k on the shared resources $\mathbf{b}_0$. At such given price levels, each subsystem seeks to maximize its own objective while "paying" for the consumption of shared resources. This process is the subproblem (S_r^k)

whose solution gives rise to a proposal (proposed comsumption) $\mathbf{A}_r\mathbf{x}_r$. If the corresponding value of v_r^k is positive, it signals to the coordinator that the proposal can be used to improve on the system-wide objective. After collecting all such profitable proposals from this round of coordination, the coordinator maximizes the system-wide objective over all available proposals. By taking convex combinations of proposals coming from each subsystem and requiring such combinations to satisfy the balance of shared resources, the coordinator can maintain system-wide feasibility while improving on the objective. This is the process involved in the master problem (M^{k+1}).

Considering each sequence of Master and Subproblems as a cycle in the coordinating process, equilibrium prices π^* on the shared resources will be obtained after a finite number (k^*) of cycles. At such prices, no subsystem will be able to generate a profitable proposal. The optimal activity levels for each subsystem is given by the covex combination of proposed solutions:

$$\sum_{i=1,\ldots,k^*} (\mathbf{x}_r^i)\lambda_{ri}^*, \tag{5}$$

where λ_{ri}^* are the corresponding optimal weights in the master problem (M^{k*}). This illustrates the well known fact that decentralization cannot be achieved by price direction alone. Indeed, given only π^*, it is in general impossible for a subsystem to recover the optimal solution in (5) by itself. Nonetheless, having the subsystem activities dictated by the coordinator according to (5) obviously detracts from the spirit of decentralized planning. Fortunately, a better alternative exists. From (M^{k*}), the coordinator can compute the optimal allocation of the shared resources to each subsystem as:

$$\mathbf{b}_{0r}^* = \sum_{i=1,\ldots,k^*} (\mathbf{A}_r\mathbf{x}_r^i)\lambda_{ri}^*. \tag{6}$$

Given this allocation, each subsystem can then proceed to determine its own activities by solving:

$$(S_r^*) \qquad \begin{aligned} \max \quad & \mathbf{c}_r\mathbf{x}_r, \\ \text{s.t.} \quad & \mathbf{A}_r\mathbf{x}_r = \mathbf{b}_{0r}^* \\ & \mathbf{B}_r\mathbf{x}_r = \mathbf{b}_r, \\ & \mathbf{x}_r \geqslant 0. \end{aligned}$$

This way, the subsystems are accorded more autonomy procedurally. We show below that this concept is also key to a better understanding of the degree of decentralization.

5. Degree of decentralization. Since subsystems can be aggregated there is a choice of the degree of decentralization. Given a maximum of R subsystems as in the above development, it is possible to decompose with anywhere from 1 to R subproblems simply by various arbitrary grouping of the subsytems. Let us examine first the extreme case of using a single subsystem. The master and subproblems become:

$$(M1^k)\qquad \begin{aligned} \text{Max } \mathbf{z}^k &= \sum_{i=1,\ldots,J1}\Big[\sum_{r=1,\ldots,R}(\mathbf{c}_r\mathbf{x}_r^i)\Big]\lambda_{1i} && \text{Dual Variables}\\ \text{s.t.}\quad &\sum_{i=1,\ldots,J1}\Big[\sum_{r=1,\ldots,R}(\mathbf{A}_r\mathbf{x}_r^i)\Big]\lambda_{1i} = \mathbf{b}_0, && \pi^k\\ &\sum_{i=1,\ldots,J1}\lambda_{1i} = 1, && \sigma_1^k\\ &\lambda_{1i} \geqslant 0, \quad i = 1,\ldots,J1. \end{aligned}$$

$$(S1_r^k)\qquad \begin{aligned} \max\ v_1^k &= \Big[\sum_{r=1,\ldots,R}(\mathbf{c}_r - \pi^k\mathbf{A}_r)\mathbf{x}_r\Big] - \sigma_1^k,\\ \text{s.t.}\quad &\mathbf{B}_r\mathbf{x}_r = \mathbf{b}_r; \quad r = 1,\ldots,R,\\ &\mathbf{x}_r \geqslant \mathbf{0}; \quad r = 1,\ldots,R. \end{aligned}$$

Note that there is now a single convexity constraint (instead of R) in the master problem. Moreover, the subproblem $(S1_r^k)$ can still be decomposed and solved subsystem by subsytem as before because their constraints do not interact. The only difference is that once the subsystem activities $\mathbf{x}_r$ are determined, an aggregate proposal

$$\sum_{r=1,\ldots,R}(\mathbf{A}_r\mathbf{x}_r^i) \tag{7}$$

is formed and submitted to the master problem. This apparent decomposability of the aggregate subproblem may lead one to speculate that the formulations with 1 or R convexity constraints are totally equivalent. That this is not the case can be shown using the concept of resource allocation described above.

Indeed, with only one set of optimal weights λ_{1i}^*, it is not possible to allocate the shared resources to the subsystems as done with (6). Therefore, in this case there is actually no decentralization at all although algorithmically, it is still possible to solve the entire problem by a sequence of R subproblems and then applying the formula in (5) to construct a solution. Using this same argument to any other grouping of the subsystems implies that the number of convexity constraints is a direct measure of the degree of decentralization.

6. Optimal degree of decentralization. Having clarified its meaning, we can now address the question of what effect the degree of decentralization has on the process of price-directive resource allocation. Early literature (Labro, 1964; Lasdon, 1970, Ch. 3; Madsen, 1973) was inconclusive mainly because of a preponderance on algorithmic performance and the lack of tools and models for significant empirical results. In this work, we used an adaptation for Apple Macintosh computers of the advanced implementation of the Dantzig–Wolfe algorithm described in Ho & Sundarraj (1989). Ten test problems from diverse applications were included. Their origins and characteristics are summarized in Table 1.

Table 1. Characteristics of test problems

Problem	Blocks	Rows	Columns	Nonzeros		Application
		Coupling	Total			
EGD085	10	8	1443	680	3635	Electric Dispatch
FIXMAR	4	18	325	452	2601	Production Planning
FORESTRY	6	11	402	603	3794	Forestry Model
MEAT12	6	46	381	311	2992	Ingredient Mix
MEAT31	8	11	384	577	4284	Ingredient Mix
MEAT43	9	16	648	605	4726	Ingredient Mix
MEGE06	6	17	1011	687	6066	Electric Capacity Expansion
MEGE08	8	17	1343	908	8200	Electric Capacity Expansion
MRP3	3	31	301	522	2010	Material Requirements Planning
MRP5	5	61	961	1740	8450	Material Requirements Planning

As a measure of the amount of coordination necessary to achieve price-directive resource allocation, we focused on the number of cycles required by the Dantzig-Wolfe algorithm. To ensure comparability of the experiments, the option of generating only a single proposal at the optimality of each subproblem was chosen. Also, the tolerance for convergence (the ε discussed in §3) was set at the zero tolerance in all cases to require an exact optimal solution. Two sets of experiments were conducted. The first compared the effect of using a single subproblem to that of using the maximum number of subproblems for all the test problems. The results are shown in Table 2. Except in one case, complete decentralization required less coordination as measured by the number of price-proposal cycles than no decentralization. The exception was with a model for electric capacity expansion that required unusually few cycles and where the master problem is highly degenerate. While the difference between the 8 or 9 cycles observed may not be of significance, it does provide a counter example.

Table 2. Number of cycles with R subproblems vs 1 subproblem

Problem	R (Bloks)	Number of cycles	
		R subsproblems	1 subproblem
EGD085	10	8	14
MEGE06	6	9	9
MEGE08	8	9	8
FORESTRY	6	16	55
MEAT12	6	39	136
MEAT31	8	13	36
MEAT43	9	11	43
MRP3	3	16	26
MRP5	5	45	89
FIXMAR	4	36	128

For the second set of experiments, three of the test problems that can be suitably decomposed into various number of subproblems were used. The number of price-proposal cycles required for the various degrees of decentralization

are summarized in Table 3. Again, the results support the general observation that the amount of coordination decreases with increasing degree of decentralization. Note that we are not concerned with the actual amount of computation or information involved in the process, but simply the number of times the coordinator would have to adjust prices on the shared resources and solicit proposals of activities from the subsystems. Naturally, fewer cycles should also imply less total effort on the part of each subsystems. That such effort is undertaken by more subsystems should not matter because in decentralization, parallel processes are *not additive*. Early literature on the topic was misleading by confounding the trade-off between work per subsystem and the number of subsystems which implied a need for the determination of an optimal degree of decentralization. Based on our perspective and observation, the empirically optimal degree of decentralization is the maximum number of subsystems.

Table 3. Number of cycles vs number of subproblems

Number of Subproblems	Number of cycles for problem		
	FIXMAR	FORESTRY	MEAT12
6	–	16	39
5	–	25	78
4	36	25	77
3	49	26	80
2	61	27	98
1	128	55	136

7. Conclusion. We have shown that the number of convexity constraints used in the coordinating master problem in the Dantzig-Wolfe approach to price-directive resource allocation provides a meaningful measure of the degree of decentralization. Also, the number of price-proposal cycles, i.e., the number of times the coordinator needs to adjust prices and solicit proposals of activities from the subsystems is an appropriate measure of the amount of coordination required by the process. Based on these arguments, empirical results are presented to suggest that in most cases, the optimal degree of decentralization is the maximum number of subsystems.

Acknowledgment. This investigation was prompted by a query from Prof. P.O. Lindberg of the Royal Institute of Technology in Stockholm, Sweden and undertaken while the first author was Visiting Scholar at the Center for Operations Research and Econometrics.

REFERENCES

Baumol, W.J., and T. Fabian (1964). Decomposition, pricing for decentralization and external economics. *Management Science,* **11.**

Dantzig, G.B. (1963). *Linear Programming and Extensions.* Princeton University Press, New Jersey.

Dantzig, G.B., and P. Wolfe (1961). The decomposition algorithm for linear programming. *Econometrica,* **29.**

Ho, J.K. (1987). Recent advances in the decomposition approach to linear programming. *Mathematical Programming Studies,* **31**, 119–128.

Ho, J.K., and E. Loute (1981). An advanced implementation of the Dantzig–Wolfe decomposition algorithm for linear programming. *Mathematical Programming,* **20**, 303-326.

Ho, J.K., and R.P. Sundarraj (1989). *DECOMP: an Implementation of Dantzig–Wolfe Decomposition for Linear Programming.* Springer–Verlag, New York.

Labro, C. (1964). Efficiency and degree of decomposition. *Working Paper No 98,* Center for Research in Management Science, University of California, Berkeley.

Lasdon, L.S. (1970). *Optimization Theory for Large Systems.* MacMillan, New York.

Madsen, O.B.G. (1973). The connection between decomposition algorithms and optimal degree of decomposition. In Himmelblau, D.M. (Ed.), *Decomposition of Large-Scale Problems.* North-Holland, Amsterdam. pp. 241–250.

Received February, 1996

J.K. Ho is a professor of Information and Decision Sciences at the University of Illinois at Chicago. He did his undergraduate work at Columbia University and obtained his Ph.D from Stanford University. His research interest is in the integration of information management and systems optimization. He is the author of *Prosperity in the Information Age: Creating value with technology – from mailrooms to boardrooms* (Infotomics, Wilmette, IL: 1994)

E. Loute is Dean of the Faculty of Economic, Social, and Political Sciences at the Facultes Universitaires Saint-Louis in Brussels, and a member of the Center for Operations Research and Econometrics at Louvain-la-Neuve in Belgium. He obtained his Ph.D from Cornell University. His work spans computer information systems and mathematical programming.

APIE DEKOMPOZICIJOS KRITERIJŲ TIESINIAME PROGRAMAVIME

James K. HO ir Etienne LOUTE

Staripsnyje nagrinėjami resursų paskirstymo tarp daugelio padalinių arba tarp daugelio regionų uždaviniai. Šie uždaviniai suvesti į blokinės-kampinės struktūros tiesinio programavimo uždavinius. Nagrinėjamas Dantzigo–Wolfo dekompozicijos algoritmas. Resursų paskirstymo uždavinį bandoma spręsti jį skaidant į dalinius uždavinius. Ieškoma optimalaus dekompozicijos kriterijaus.