NASA Technical Memorandum 102662

Overcoming the Bellman's "Curse of Dimensionality" in Large Optimization Problems

Jaroslaw Sobieszczanski-Sobieski

April 1990

(NASA-TM-102662) OVERCOMING THE BELLMAN'S CURSE OF DIMENSIONALITY IN LARGE OPTIMIZATION PROBLEMS (NASA) 6 p CSCL 01C

N90-23591

Unclas

G3/05 0287437

National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23665-5225

OVERCOMING THE BELLMAN'S "CURSE OF DIMENSIONALITY" IN LARGE OPTIMIZATION PROBLEMS

Jaroslaw Sobieszczanski-Sobieski
NASA Langley Research Center
Hampton, Virginia U.S.A.

ABSTRACT

Decomposition of large problems into a hierarchic pyramid of subproblems was proposed in the literature as a means for optimization of engineering systems too large for "all-in-one" optimization. This decomposition was established heuristically. The paper shows that the dynamic programing (DP) method due to Bellman when augmented with an optimum sensitivity analysis provides a mathematical basis for the above decomposition, and overcomes the "curse of dimensionality" that limited the original formulation of DP. Numerical examples are cited.

INTRODUCTION

Engineering optimization problems, e.g., maximization of aircraft performance, are usually computationally expensive and depend on the correct simulation of interaction of many parts and disciplines. Therefore these problems are natural candidates for optimization by decomposition that converts one large problem into a set of coordinated smaller problems. This paper shows that the linear decomposition method, whose applications recently appeared in the literature, may be viewed as an extension and generalization of the well-established algorithm of Dynamic Programing due to Bellman (ref. 1). In particular, this method alleviates to a large extent the algorithm limitation known as "the curse of dimensionality".

DYNAMIC PROGRAMING AS OPTIMIZATION BY DECOMPOSITION

Dynamic Programing, originally due to Bellman (ref. 1), is perhaps the best mathematically established method for optimization by decomposition. Briefly summarized, the method applies to a system that may be represented by a train of the "black boxes" as in figure 1 (called the stages in ref. 1). The black boxes are mathematical models of the physical parts or conceptual aspects (engineering disciplines) of a large problem, and they are unified by a flow of information from the n-th box through the intermediate boxes ending with the 1-st box. An i-th box receives an input vector S_{i+1} from its predecessor, an input vector of design variables X_i, and outputs a vector S_i that becomes an input to the successor box i-1. The i-th box also outputs a quantity R_i interpreted as a component of the objective function of the entire system. Each $\{S_i\}$ must be reducible to a function of a single variable, s_i, $S_i = f(s_i)$, and, because $S_i = f(S_{i+1}, X_i)$ and $R_i = f(S_{i+1}, X_i)$, it follows that $S_i = f(s_{i+1}, X_i)$, $R_i = f(s_{i+1}, X_i)$, and $s_i = f(s_{i+1}, X_i)$.

The problem of finding a set of vectors X_i, $i = 1 \ldots n$, that minimizes the sum of $R_i, i = 1 \ldots n$, is solved by starting with the 1-st box at the end of the train. The variable s_2 governing the input S_2 is assumed to vary within an interval of interest. Several values of s_2 distributed over that interval are set and an optimization problem, constrained or unconstrained, is solved to find $\{X_1\}$ so as to minimize R_1 at each value. That solution yields $\{X_1\}_{\text{opt}} = f(s_2)$ and $R_{1\min} = f(s_2)$, either in a discrete (a look-up table) form or in a continuous form interpolated between the s_1 values assumed above, dependently on the nature of the problem.

Moving up to the 2-nd box, one seeks for each of the several values of s_3 in an interval of interest an optimal $\{X_2\}$ (denoted $\{X_2\}_{\text{opt}}$) that minimizes the sum $R_2+R_{1\min}$. One must consider that $s_2 = f(s_3, X_2)$ and that for each value of s_2, there are $\{X_1\}_{\text{opt}}$ and $R_{1\min}$ already known from the optimizations that have been executed for box 1. This operation generates the values of $\{X_2\}_{\text{opt}} = f(s_3)$ and $(R_2 + R_{1\min})_{\min}$.

The procedure continues recursively from box i to box $i+1$, carrying forward $\{X_j\}_{\text{opt}}$, and $(R_j + (R_{j-1} + (R_{j-2} \ldots + (R_2 + R_{1\min})_{\min})_{\min})_{\min} \cdots)_{\min}$, $j = 1 \ldots i$, through the initial box in the train, $i = n$, whereby the minimum sum of all R_i's and a complete set of $\{X_i\}_{\text{opt}}$'s gets established. The procedure rests on the fundamental principle formulated by Bellman which asserts that the set of $\{X_j\}$, $j = 1 \ldots n$, is optimal when its subset for $j = 1 \ldots i$ taken for any i minimizes the sum of R_j, $j = 1 \ldots i$, for S_{i+1} input given from the remainder of the train.

The procedure computational cost heavily depends on the aforementioned assumptions of $S_i = f(s_i)$ where s_i is a single variable. Indeed, if s_i were a vector of m elements, $R_{i\min}$ would grow from a line plot into a hypersurface in m dimensions. Assuming a quadratic representation of that hypersurface (the lowest order nonlinear approximation), the number of discrete points at which optimizations would have to be performed would grow proportional to the square of m, thus quickly destroying advantages of the procedure as a computational cost saver. Bellman called this the "curse of dimensionality" and regarded it as a barrier limiting applicability of the method.

OVERCOMING THE CURSE OF DIMENSIONALITY

Optimum sensitivity analysis formulated in ref. 2 provides a means for generalization of the above procedure to include s_i defined as a vector of m elements. The optimum sensitivity analysis algorithm yields derivatives of the optimal $\{X\}$ and R with respect to the parameters of the optimization problem (unconstrained or constrained). Taking box 1 as an example, $\{S_2\}$ may now be defined as $\{S_2\} = f(\{s_2\})$, where $\{s_2\}$ is a vector of elements s_{2k}, $k = 1 \ldots m$. For $\{s_2\}$ given, one may find $\{X_1\}_{\text{opt}}$ and minimum of R_1 and their derivatives with respect to each s_{2k}, regarded as an optimization parameter. Using the notation $D(X_1, s_{2k})$ and $D(R_1, s_{2k})$ for these derivatives, the linear part of the Taylor series enables one to express $\{X_1\}_{\text{opt}}$ and $R_{1\min}$ as continuous, albeit approximate, functions of $\{s_2\}$:

$$\{X_1\}_{\text{opt}} = f(\{s_2\}) = (\{X_1\}_{\text{opt}})_o + [D(X_1, s_{2k})]\Delta\{s_2\} \tag{1}$$

$$R_{1\min} = f(\{s_2\}) = (R_{1\min})_o + \{D(R_1, s_{2k})\}'\Delta\{s_2\} \tag{2}$$

The Bellman's Dynamic Programing procedure may now be executed using the above approximations in place of $\{X_1\}_{\text{opt}} = f(s_2)$ and $R_{1\min} = f(s_2)$, otherwise the procedure remains unchanged. The new component in the modified procedure is the optimum sensitivity analysis to be executed after each optimization involving boxes 1, $(2+1)$, $(3+2+1), \ldots n$, recursively. Because the linear relationships, eqs. 1 and 2, introduce errors whose control requires move limits on design variables in each optimization, the entire procedure has to be repeated p times until satisfactory convergence is attained. In this case, the number p depends on the nonlinearities of the problem at hand. Consequently, because there is only one optimization in each box in one pass, the number of optimizations required to converge the procedure is pn. This is in contrast to nmm, which is necessary for the original procedure. The curse of dimensionality with respect to m is removed. The ratio $pn/nmm = p/mm$ tends to be very small for large m and renders the modified procedure usable where the original one would be prohibitively expensive.

However, unlike the original procedure, the modified procedure relies on the continuity of the approximation function in eqs. 1 and 2; hence, it cannot accommodate discrete design variables.

HIERARCHIC DECOMPOSITION

Further generalization of the modified procedure is possible if the boxes in the train may be partitioned internally as shown in figure 2. This figure shows the boxes split internally into smaller ones. In this scheme, the train of boxes that was horizontal in figure 1 is depicted vertically to form a pyramid whose levels correspond to the boxes in figure 1. A typical level is populated by several boxes that formed a single box in figure 1. The pyramidal arrangement emphasizes the hierarchic dependence of the boxes in level i ("children") on the information transmitted from a box located at the level above $j > i$ ("parent"), with the underlying assumption that the boxes at the same level ("siblings") do not exchange information with each other directly. Similar to the system shown in figure 1, the behavior information from each box flows in figure 2 from the parent to the children, or from the top level n down to level 1. The optimization information from each box flows in the opposite direction. This information includes the optimum sensitivity derivatives that enable optimization in each parent box to be performed taking into account the effect of its $\{X\}$ on the optimization results in all boxes descendent from that parent.

Thus, the above decomposition scheme first developed heuristically in ref. 3 is shown to be a generalization of the Bellman's Dynamic Programing. The scheme became known as hierarchic, linear decomposition.

APPLICATION EXAMPLES

Since its introduction in ref. 3, optimization by hierarchic linear decomposition has been demonstrated to be useful in several applications. For example, in ref. 4 it was used to develop structural optimization by substructuring. This case is illustrated by a portal framework (figure 3a) shown decomposed in figure 3b. The procedure histogram in figure 3c exhibits satisfactory convergence characteristics. Analytical information flowing down the pyramid consisted of the internal forces and cross-section stiffness properties as parameters of optimization. Optimal cross-sectional dimensions, minimal values of the cumulative constraints and their sensitivity derivatives with respect to the above parameters were transmitted in the opposite direction.

An example of the procedure application to a multidisciplinary problem of optimization of a transport aircraft for performance under constraints drawn from major contributing disciplines was described in ref. 5. The aircraft, its decomposition scheme, and a histogram of the optimization procedure are shown in figure 4, a, b, and c, respectively. The case featured over 1000 design variables and constraints and demonstrated a mathematical link from the design detail (e.g., wing panel cross-sectional dimensions) to the system performance (e.g., the mission fuel). The procedure convergence was smooth and rapid as seen in figure 4c.

CONCLUSIONS

It is shown that the Bellman's method for decomposition of large optimization problems known as Dynamic Programing may be generalized to encompass the cases when the information transmitted between the parts of the system is a function of many variables. The key component of the modified procedure is the derivatives of optimum with respect to the optimization parameters. Application examples illustrate and verify the procedure.

REFERENCES

1. Bellman, R.: Adaptive Control Processes: A Guided Tour; Princeton University Press, 1961.

2. Barthelemy, J. F.; and Sobieszczanski-Sobieski, J.: Optimum Sensitivity Derivatives of Objective Functions in Nonlinear Programing; AIAA J, Vol. 22, No. 6, June 1983, pp. 913-915.

3. Sobieszczanski-Sobieski, J.: A Linear Decomposition Method for Large Optimization Problems—Blueprint for Development; NASA TM 83248, February 1982.

4. Sobieszczanski-Sobieski, J.; James, B. B.; and Dovi, A. R.: Structural Optimization by Multilevel Decomposition; AIAA J., Vol. 23, No. 11, November 1985, pp. 1775-1782.

5. Wrenn, G. A.; and Dovi, A. R.: Multilevel Decomposition Approach to the Preliminary Sizing of a Transport Aircraft Wing; AIAA Journal of Aircraft, Vol. 25, No. 7, July 1988, pp. 632-638.

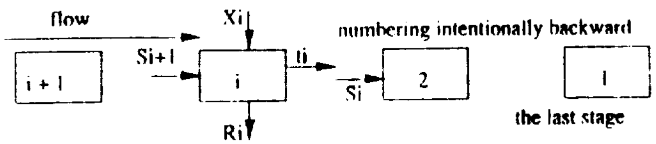


Fig. 1: Train of black boxes

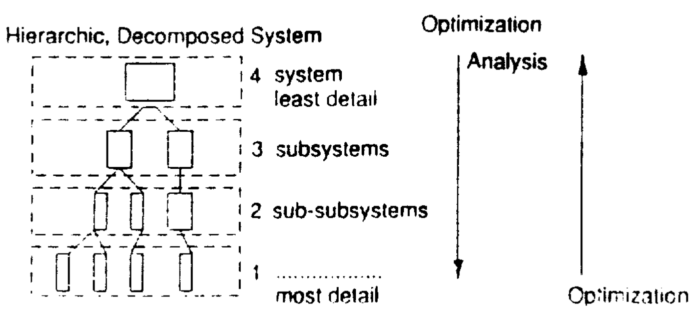


Fig. 2: Hierarchic decomposition

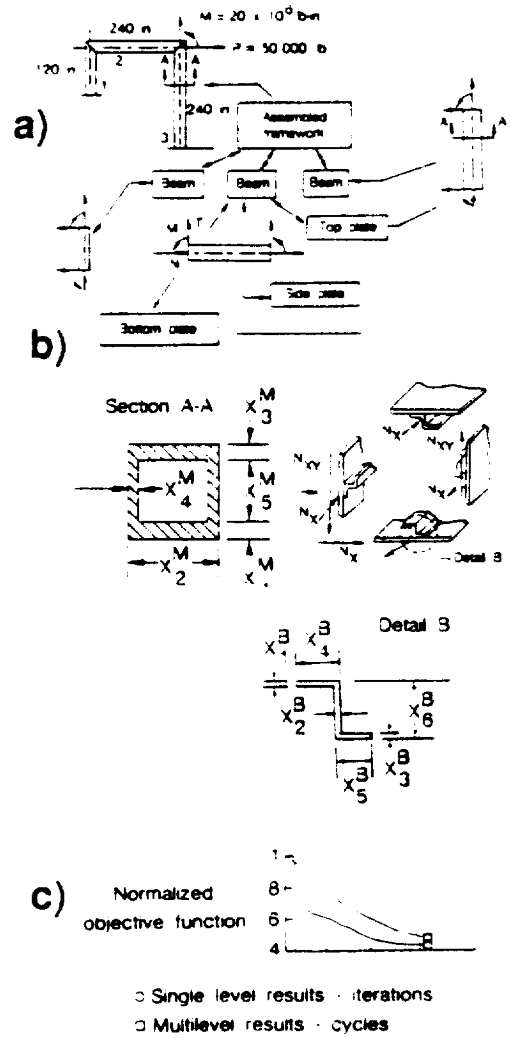
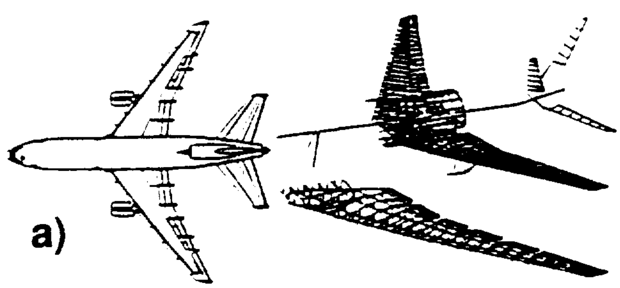
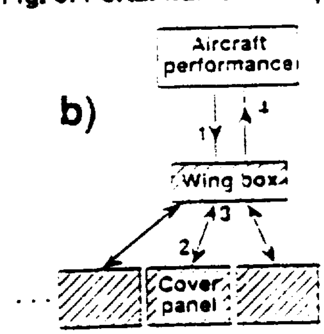


Fig. 3: Portal framework optimized by decomposition





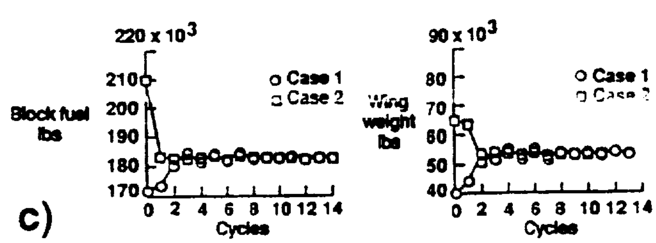


Fig. 4: Aircraft optimized by decomposition

ORIGINAL PAGE IS
OF POOR QUALITY

NASA
National Aeronautics and Space Administration

Report Documentation Page

1. Report No. NASA TM-102662	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Overcoming the Bellman's 'Curse of Dimensionality" in Large Optimization Problems		5. Report Date April 1990
		6. Performing Organization Code
7. Author(s) Jaroslaw Sobieszczanski-Sobieski		8. Performing Organization Report No.
		10. Work Unit No. 505-63-01
9. Performing Organization Name and Address NASA Langley Research Center Hampton, VA 23665		11. Contract or Grant No.
		13. Type of Report and Period Covered Technical Memorandum
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, DC 20546-0001		14. Sponsoring Agency Code

15. Supplementary Notes

To be presented at the Second Pan American Congress of Applied Mechanics, January 2-5, 1991, Valparaiso, Chile.

16. Abstract

Decomposition of large problems into a hierarchic pyramid of subproblems was proposed in the literature as a means for optimization of engineering systems too large for "all-in-one" optimization. This decomposition was established heuristically. The paper shows that the dynamic programming (DP) method due to Bellman was augmented with an optimum sensitivity analysis provides a mathematical basis for the above decomposition, and overcomes the "curse of dimensionality" that limited the original formulation of DP. Numerical examples are cited.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Decomposition Dynamic Programing optimization	Unclassified - Unlimited Subject Category 05

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of pages	22. Price
Unclassified	Unclassified	5	A02

NASA FORM 1626 OCT 86