Multidisciplinary Design Optimization with Mixed Integer Quasiseparable Subsystems

RAPHAEL T. HAFTKA
Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, FL 32611-6250

LAYNE T. WATSON
Departments of Computer Science and Mathematics, Virginia Polytechnic Institute & State University, Blacksburg, VA 24061-0106

Abstract. Numerous hierarchical and nonhierarchical decomposition strategies for the optimization of large scale systems, comprised of interacting subsystems, have been proposed. With a few exceptions, all of these strategies have proven theoretically unsound. Recent work considered a class of optimization problems, called quasiseparable, narrow enough for a rigorous decomposition theory, yet general enough to encompass many large scale engineering design problems. The subsystems for these problems involve local design variables and global system variables, but no variables from other subsystems. The objective function is a sum of a global system criterion and the subsystems' criteria. The essential idea is to give each subsystem a budget and global system variable values, and then ask the subsystems to independently maximize their constraint margins. Using these constraint margins, a system optimization then adjusts the values of the system variables and subsystem budgets. The subsystem margin problems are totally independent, always feasible, and could even be done asynchronously in a parallel computing context. An important detail is that the subsystem tasks, in practice, would be to construct response surface approximations to the constraint margin functions, and the system level optimization would use these margin surrogate functions. The present paper extends the quasiseparable necessary conditions for continuous variables to include discrete subsystem variables, although the continuous necessary and sufficient conditions do not extend to include integer variables.

Keywords: decomposition, global/local optimization, mixed integer programming, multidisciplinary design, response surface approximation, separable

1. Introduction

Many discrete optimization problems are NP-hard, and thus intractable for even a moderately large number of design variables. All known algorithms for continuous global optimization have exponential computational complexity, and thus continuous global optimization is similarly intractable. Consequently, decomposition methods may be particularly useful for such problems.

Numerous hierarchical and nonhierarchical decomposition strategies have been proposed—Dantzig-Wolfe decomposition, concurrent subspace optimization (CSSO) (Rodríguez et al., 1998a), Sobieski's nonhierarchical decomposition (Shankar et al., 1993), Kroo's collaborative optimization (CO) (Kroo, 1997), multilevel decomposition (Alexandrov, 1997), Renaud and Watson's augmented Lagrangian approach (Rodríguez et al., 1998b), to name but a few. With but a few exceptions (such as Dantzig-Wolfe and Renaud-Watson), these strategies have not been rigorously proven to converge.

Previous work by Haftka and Watson (2004) focused on a class of quasiseparable optimization problems narrow enough for a rigorous decomposition theory, yet general enough to encompass many large scale engineering design problems. The subsystems for these problems involve local design variables and global system variables, but no variables from other subsystems. The objective function is a sum of a global system criterion and the subsystems' criteria. Haftka and Watson (2004) developed a quasiseparable decomposition (QSD) algorithm with global search at the lower level, and proved that the QSD algorithm does not generate any spurious solutions. Furthermore, Liu et al. (2003) demonstrated the power of QSD to perform most of the global optimization computation in low dimensions at the subsystem level. The objective of the present paper is to extend QSD to include discrete subsystem variables.

The essential idea of QSD is to give each subsystem a budget and global system variable values, and then ask the subsystems to independently maximize their constraint margins. (A constraint margin is the amount of slack a constraint has at a particular design point.) Using these constraint margins, a system optimization then adjusts the values of the system variables and subsystem budgets. The subsystem margin problems are totally independent, always feasible, and could even be done asynchronously in a parallel computing context.

An example of the type of optimization problem targeted by the decomposition theory presented here is vehicle structural design. Vehicle structural design may be viewed as a multilevel optimization problem. At the highest level the entire vehicle is designed; at a lower level major components, such as the wing of an aircraft, or the chassis of a car, are designed; at an even lower level subcomponents, such as individual panels, are designed. Unfortunately, because of computational limitations, it has not been possible to perform this multilevel optimization rigorously. Instead, various ad hoc design procedures are used to decouple the three levels (e.g., Harte et al., (2004)).

For aircraft structural design, previous decomposition schemes focused on the use of response surface approximations, fit to the values of multiple lower level optima, as a way to integrate the various design levels (e.g., Giunta et al. (1997b), Balabanov et al. (1999), Ragon et al. (1997), and Liu et al. (2000)). This involved performing hundreds of subsystem optimizations, and fitting the optimal response surface with a polynomial for use by the system optimization algorithm. Work using these ad hoc decomposition schemes occasionally verified that the two-level procedure led to the same design as a single level optimization for some examples (Liu and Haftka, 2003).

Even though the application of these decomposition procedures was successful, usually no attempt was made to explore their theoretical properties. Previous work established these theoretical properties for the case of continuous design variables and specific decomposition strategies, e.g., variable fidelity augmented Lagrangian (Rodríguez et al., 1998b) and QSD (Haftka and Watson, 2004). However, in many of these system design problems, some or all of the design variables at the component level are discrete. For example, in Liu and Haftka (2003), lower level variables are the ply angles of composite laminates that can take only discrete values, such as $0°$, $90°$, or $\pm 45°$. The existence of discrete variables in realistic design problems is the motivation for this paper, whose purpose is to extend the QSD theory to permit discrete subsystem design variables.

It is worthwhile to discuss first the intuitive basis of quasiseparable decomposition (QSD). The main concept involved here is arranging subtasks, with assigned goals and constraints, that can be pursued independently and concurrently. Before presenting a mathematical formulation, it may help to provide a real life scenario that illustrates the basic concept of the decomposition formulation. The decomposition approach is similar to the fairly common standard practice of a manager assigning budget and performance requirements to each department or team that reports to her. The decomposition procedure may be viewed as a negotiation session between the manager and the team or department leaders, where she presents each with sets of performance requirements and budgets. For each set the team leader provides a measure of how well he can satisfy the requirements and the budget. This allows the manager to intelligently divide resources between teams. Discrete design variables within each department may involve decisions on how many people to hire or how many computers to buy.

The decomposition method described in the next two sections is based on this concept. Each subsystem designer is provided with a budget and a set of performance constraints, and his task is to maximize the margin of safety for satisfying both.

This paper is the result of collaboration between an engineer and a mathematician. Therefore, mathematical theorems and proofs are accompanied by notes intended to provide engineering intuition into some aspects of the basis of these theorems and proofs. In particular, a key concept in this intuitive understanding is that two distant locally optimal solutions of the original problem may become close in the decomposed problem. In the context of the management problem discussed above, this represents the viewpoint of the manager. A department may use its budget to achieve exactly or approximately the same goals or outputs in several very different ways. However, the manager, who sees only the results, considers these solutions to be identical or close.

The management analogy also illustrates one important advantage of the decomposition over a single level solution, due to the fact that both the system level problem and the subsystem level problems are of lower dimensions than the single level (original) problem. A manager would be interested in a global rather than a local solution to the problem. By delegating part of the global optimization to each department, both the manager and the departments seek this global optimum in a lower dimensional space than if the manager micromanaged the entire (global) search for the optimum. (Technically, though, even if each department were to find the global optimum for its assigned budget and requirements, the manager still has the option of finding assignments that are only locally optimal over all possible assignments to the departments.)

The mathematical theory follows, with the special case of all subsystem variables being discrete addressed in Section 2, and both continuous and discrete subsystem variables addressed in Section 3. Section 4 concludes with some interesting numerical examples.

2. Theory for purely discrete quasiseparable subsystems

This section considers a special case where the global system variables are real, and all the local subsystem variables are discrete (integer). The proofs for this case are identical to those for subsystems with mixed real and integer local variables, but the notation for the general case is somewhat complicated, and the proofs are nontrivial. Hence to avoid becoming mired in notational difficulties, the definitions and proofs are first done for the case of purely discrete subsystem (local) variables here, and then repeated for the general case of mixed integer subsystem variables in the next section.

Denote real n-dimensional Euclidean space by E^n, the set of n-tuples of integers by Z^n, let g_j denote the jth component of a vector g, and let $B(x, \delta) \subset E^n$ denote the open ball of radius δ centered at x. At some point $\bar{x}$, the *constraint margin* for a (scalar) constraint $g(x) \leq 0$ is the amount $-\mu$ by which the constraint is slack: $g(\bar{x}) - \mu = 0$. The margin $-\mu$ for a vector constraint $G(x) \leq 0$ is the smallest component margin: $\max_j G_j(\bar{x}) - \mu = 0$. Writing the (positive at feasible points) constraint margin as $-\mu$ rather than μ leads to elegant decomposed problem formulations.

Denote the system variables by $s \in E^{n_0}$, the local subsystem variables by $\ell^{(i)} \in Z^{n_i}$, for $i = 1$, $\ldots$, N, and the total number of variables by

$$n = n_0 + n_1 + \cdots + n_N.$$

Let $\ell = \big(\ell^{(1)}, \ldots, \ell^{(N)}\big)$ denote all the local variables, let $f_0(s)$, $f_1\big(s, \ell^{(1)}\big)$, $\ldots$, $f_N\big(s, \ell^{(N)}\big)$ be C^2 (in the real variables) real-valued functions, and let $g^{(0)}(s)$, $g^{(1)}\big(s, \ell^{(1)}\big)$, $\ldots$, $g^{(N)}\big(s, \ell^{(N)}\big)$ be C^2 (in the real variables) real vector-valued functions of dimensions m_0, m_1, $\ldots$, m_N, respectively.

The quasiseparable subsystem problem (Haftka and Watson, 2004) is

$$\min_{s,\ell} f_0(s) + \sum_{i=1}^{N} f_i\big(s, \ell^{(i)}\big) \qquad \text{subject to} \qquad \begin{array}{c} g^{(0)}(s) \le 0, \\ g^{(i)}\big(s, \ell^{(i)}\big) \le 0,\ i = 1, \ldots, N. \end{array} \tag{SSP}$$

Precisely, a nonlinear programming problem has *quasiseparable subsystems* (or is *quasiseparable*) if it has the form (SSP). The essential ideas are to introduce a budget $b_i \in E$ for each subsystem criterion $f_i\big(s, \ell^{(i)}\big)$, and to attack the subsystem constraints $g^{(i)}\big(s, \ell^{(i)}\big) \le 0$ by maximizing the constraint margin $-\mu_i$ for each subsystem. The global system variables s are real, but all the local subsystem variables $\ell^{(i)}$ are discrete here. The original problem (SSP) is decomposed into an upper level problem and N lower level problems, which are always feasible and are totally independent. The decomposed problem is (upper level)

$$\min_{s,b} f_0(s) + \sum_{i=1}^{N} b_i \qquad \text{subject to} \qquad \begin{array}{c} g^{(0)}(s) \le 0, \\ \mu_i(s, b) \le 0,\ i = 1, \ldots, N, \end{array} \tag{DSSP}$$

where $\mu_i(s, b)$ is the (global) solution to the ith lower level problem, for $i = 1$, $\ldots$, N, given by

$$\min_{\ell^{(i)}} \mu_i \qquad \text{subject to} \qquad \begin{array}{c} \max\limits_{1 \le j \le m_i} g_j^{(i)}\big(s, \ell^{(i)}\big) - \mu_i \le 0, \\ f_i\big(s, \ell^{(i)}\big) - b_i - \mu_i \le 0. \end{array} \tag{DSSP}$$

Let X and X_D denote the feasible sets for (SSP) and (DSSP), respectively. Assume that the set $\big\{\ell \mid (s, \ell) \in X\big\}$ of feasible discrete values is finite. Let $\theta(s, \ell)$ be the objective function of (SSP), and $\theta_D(s, b)$ the upper level objective function of (DSSP). The following lemma is straightforward to verify.

Lemma 2.1. If (s, ℓ) is feasible for (SSP), then

$$(s, b) = (s, b_1, \ldots, b_N) = \Big(s, f_1\big(s, \ell^{(1)}\big), \ldots, f_N\big(s, \ell^{(N)}\big)\Big)$$

is feasible for (DSSP).

It is also easy to check that at an optimal solution of (DSSP) there is no margin on the budget constraints, so that $b_i = f_i\big(s, \ell^{(i)}\big)$. In the simpler case considered by Haftka and Watson (2004) where $\theta(s, \ell)$ only depended on s, points that were close to locally optimal solutions for (SSP) corresponded to even closer points for the upper level of (DSSP) because the ℓ components were dropped (these problems were called (ISV) and (DISV), respectively, in (Haftka and Watson, 2004)). This guaranteed that the decomposition did not introduce spurious locally optimal solutions. Now, with the b variables and the discrete variables ℓ, this is not as trivial. Here the continuity of the f_i ensures that points that are close to locally optimal solutions of (SSP) correspond to nearby points in the upper level of (DSSP). The following definition and theorem imply that every locally optimal solution of (DSSP) corresponds to a locally optimal solution of (SSP).

Definition 2.1. $(s^*, \ell^*) \in X$ is a local solution for (SSP) if there exists an open ball $B(s^*, \delta)$ around s^* such that $(s, \ell) \in X$ and $s \in B(s^*, \delta)$ imply $\theta(s^*, \ell^*) \le \theta(s, \ell)$. (Note that this implies the minimum is global with respect to the integer variables.)

Theorem 2.1. If $(s^*, b^*) = \Big(s^*,\ f_1\big(s^*, \ell^{(1)*}\big),\ \ldots,\ f_N\big(s^*, \ell^{(N)*}\big)\Big) \in X_D$ is a local solution for (DSSP), then $(s^*, \ell^*) \in X$ is a local solution for (SSP).

Proof. The claim is that there exist $\delta > 0$, $\epsilon > 0$ such that

$$\begin{aligned}&\theta_D(s^*, b^*) \le \theta_D(s, b) \text{ for } (s, b) \in X_D \cap B\big((s^*, b^*), \delta\big) \Longrightarrow \\ &\theta\big(s^*, \ell^*\big) \le \theta(s, \ell) \text{ for } (s, \ell) \in X \text{ and } s \in B\big(s^*, \epsilon\big).\end{aligned}$$

Suppose the hypothesis holds for some $\delta > 0$, but the conclusion does not hold for any $\epsilon > 0$; this will be shown to lead to a contradiction. Then there exists a sequence $\big(s_{(k)}, \ell_{(k)}\big) \in X$ such that $s_{(k)} \to s^*$ and

$$\theta\big(s_{(k)}, \ell_{(k)}\big) < \theta\big(s^*, \ell^*\big)$$

for all k. Since the discrete part of X is assumed finite, it may be assumed that all $\ell_{(k)} = \ell'$, by reducing to a subsequence if necessary. With ℓ' as above let $b' = \Big(f_1\big(s^*, \ell^{(1)\prime}\big), \ldots, f_N\big(s^*, \ell^{(N)\prime}\big)\Big)$. By continuity,

$$\lim_{k \to \infty} \big(s_{(k)}, \ell_{(k)}\big) = (s^*, \ell') \in X.$$

Observe that for $\big\|s_{(k)} - s^*\big\|_2 < \delta/\sqrt{2}$,

$$b_{(k)} = \Big(f_1\big(s_{(k)}, \ell_{(k)}^{(1)}\big), \ldots, f_N\big(s_{(k)}, \ell_{(k)}^{(N)}\big)\Big)$$

is not within $\delta/\sqrt{2}$ of b^*, otherwise by the assumed valid hypothesis,

$$\theta\big(s_{(k)}, \ell_{(k)}\big) = \theta_D\big(s_{(k)}, b_{(k)}\big) \ge \theta_D(s^*, b^*) = \theta\big(s^*, \ell^*\big).$$

Therefore again by continuity, $\lim_{k\to\infty} b_{(k)} = b'$ is also not within $\delta/\sqrt{2}$ of b^*. This together with the inequality

$$\sum_{i=1}^{N} \big(b_i' - b_i^*\big) = \theta_D(s^*, b') - \theta_D(s^*, b^*) = \theta\big(s^*, \ell'\big) - \theta\big(s^*, \ell^*\big) \le 0$$

implies that some component $b_i' \le b_i^* - \delta/\big(\sqrt{2N}N\big)$. Now from the two points (s^*, ℓ^*) and (s^*, ℓ') in X build a new point $(s, \ell) \in X$ with $s = s^*$, $\ell^{(j)} = \ell^{(j)*}$ for $j \ne i$, and $\ell^{(i)} = \ell^{(i)\prime}$. By Lemma 2.1, this point (s, ℓ) corresponds to $\big(s,\ b_1^*,\ \ldots,\ b_{i-1}^*,\ b_i',\ b_{i+1}^*,\ \ldots,\ b_N^*\big) \in X_D$, from which can be deduced the new point

$$(s, b) = \Big(s, b_1^*, \ldots, b_{i-1}^*, b_i^* - \frac{\delta}{2\sqrt{N}N}, b_{i+1}^*, \ldots, b_N^*\Big) \in X_D$$

with $\|(s, b) - (s^*, b^*)\|_2 < \delta$ and $\theta_D(s, b) < \theta_D(s^*, b^*)$. This contradicts the assumed hypothesis, so the result follows. Q. E. D.

The proof of Theorem 2.1 immediately yields the following useful observation.

Corollary 2.1. $(s^*, b^*) = \Big(s^*,\ f_1\big(s^*, \ell^{(1)*}\big),\ \ldots,\ f_N\big(s^*, \ell^{(N)*}\big)\Big) \in X_D$ is a global solution for (DSSP) $\iff$ $(s^*, \ell^*) \in X$ is a global solution for (SSP).

3. Theory for mixed integer quasiseparable subsystems

As in the previous section, denote real n-dimensional Euclidean space by E^n, the set of n-tuples of integers by Z^n, let g_j denote the jth component of a vector g, and let $B(x, \delta) \subset E^n$ denote the open ball of radius δ centered at x. At some point $\bar{x}$, the *constraint margin* for a (scalar) constraint $g(x) \le 0$ is the amount $-\mu$ by which the constraint is slack: $g(\bar{x}) - \mu = 0$. The margin $-\mu$ for a vector constraint $G(x) \le 0$ is the smallest component margin: $\max_j G_j(\bar{x}) - \mu = 0$. Writing the (positive at feasible points) constraint margin as $-\mu$ rather than μ leads to more elegant notation later.

Consider first a trivial special case where the objective function $f(s)$ depends on the global system design variables s but not on any of the subsystem local design variables $\ell = \big(\ell^{(1)}, \ldots, \ell^{(N)}\big)$. Such a problem has the form

$$\min_{s,\ell} f(s) \qquad \text{subject to} \qquad \begin{array}{c} g^{(0)}(s) \le 0, \\ g^{(i)}\big(s, \ell^{(i)}\big) \le 0,\ i = 1, \ldots, N. \end{array} \tag{ISV}$$

(The name ISV derives from the objective function, which is *independent* of *subsystem variables.*) Here the system variables $s \in E^{n_0}$, the local subsystem variables $\ell^{(i)} = \big(\overline{\ell^{(i)}}, \widehat{\ell^{(i)}}\big) \in E^{\bar{n}_i} \times Z^{\hat{n}_i}$, $n_i = \bar{n}_i + \hat{n}_i$, for $i = 1, \ldots, N$, and to simplify the discussion assume that $f(s)$, $g^{(0)}(s)$, $g^{(1)}\big(s, \ell^{(1)}\big)$, $\ldots$, $g^{(N)}\big(s, \ell^{(N)}\big)$ are C^2 (in the real variables) real vector-valued functions of dimensions 1, m_0, m_1, $\ldots$, m_N, respectively. To simplify the separation of the discrete and real variables, write $\bar{\ell} = \big(\overline{\ell^{(1)}}, \ldots, \overline{\ell^{(N)}}\big)$ and $\hat{\ell} = \big(\widehat{\ell^{(1)}}, \ldots, \widehat{\ell^{(N)}}\big)$. In this trivial case, there is no optimization done at the subsystem level, since each subsystem consists merely of constraints $g^{(i)}\big(s, \ell^{(i)}\big) \le 0$. Thinking in terms of global system design optimization, the system optimization should have as much leeway as possible, so the subsystem tasks should not be merely to find feasible $\ell^{(i)}$, but to maximize the constraint margins $-\mu_i$. This idea leads to the decomposed (upper level) problem

$$\min_{s} f(s) \qquad \text{subject to} \qquad \begin{array}{c} g^{(0)}(s) \le 0, \\ \mu_i(s) \le 0,\ i = 1, \ldots, N, \end{array} \tag{DISV}$$

where $\mu_i(s)$ is the (global) solution to the ith lower level problem, for $i = 1, \ldots, N$, given by

$$\min_{\ell^{(i)}} \mu_i \qquad \text{subject to} \qquad \max_{1 \le j \le m_i} g_j^{(i)}\big(s, \ell^{(i)}\big) - \mu_i \le 0. \tag{DISV}$$

Denoting the feasible set for (ISV) by X, the feasible set for (DISV) is

$$X_D = \{s \mid (s, \ell) \in X \text{ for some } \ell\},$$

the projection of X onto the system design space. Note that distant feasible points $(s, \ell) \in X$ may project to the same feasible point $s \in X_D$, and that a disconnected open set in X may project to a connected open set in X_D. The fact that projections from X to X_D do not increase distances between points is important because it guarantees that the decomposition (DISV) does not introduce spurious local solutions. This is formalized in the next definition and theorem.

DEFINITION 3.1. $(s^*, \ell^*) \in X$ is a local solution for (ISV) if there exists an open ball $B\big((s^*, \overline{\ell^*}), \delta\big)$ around $(s^*, \overline{\ell^*})$ such that $(s, \ell) \in X$ and $(s, \bar{\ell}) \in B\big((s^*, \overline{\ell^*}), \delta\big)$ imply $f(s^*) \le f(s)$.

THEOREM 3.1. If s^* is a local solution for (DISV), then there exists ℓ^* such that (s^*, ℓ^*) is a local solution for (ISV).

Proof. s^* being a local solution means $\exists\, \epsilon > 0$ such that $s \in X_D \cap B(s^*, \epsilon) \Longrightarrow f(s^*) \le f(s)$. Also $s^* \in X_D \Longrightarrow \exists\, \ell^*$ such that $(s^*, \ell^*) \in X$. Since clearly $\|s - s^*\|_2 \le \big\|(s, \bar{\ell}) - (s^*, \overline{\ell^*})\big\|_2$, it follows that $(s, \ell) \in X$ and $(s, \bar{\ell}) \in B\big((s^*, \overline{\ell^*}), \epsilon\big) \Longrightarrow s \in X_D \cap B(s^*, \epsilon) \Longrightarrow f(s^*) \le f(s) \Longrightarrow (s^*, \ell^*)$ is a local solution of (ISV). Q. E. D.

For the continuous case, with some convexity and quasiconvexity assumptions, it is possible to prove a theorem of the form "s^* is a local solution for (DISV) $\iff \exists\, \ell^*$ such that (s^*, ℓ^*) is a local solution for (ISV)." As shown by a counterexample in (Haftka and Watson, 2004), the convexity is essential, and thus such a theorem is not possible in the discrete case (any subset of Z^k with at least two points is not convex). However, it is easily seen that global minima of (ISV) and (DISV) do correspond without convexity assumptions.

COROLLARY 3.1. s^* is a global solution for (DISV) $\iff \exists\, \ell^*$ such that (s^*, ℓ^*) is a global solution for (ISV).

Consider next the more typical situation where the subsystem local design variables are involved in the system objective function. Denote the system variables by $s \in E^{n_0}$, the local subsystem variables by $\ell^{(i)} = \big(\overline{\ell^{(i)}}, \widehat{\ell^{(i)}}\big) \in E^{\bar{n}_i} \times Z^{\hat{n}_i}$, $n_i = \bar{n}_i + \hat{n}_i$, for $i = 1, \ldots, N$, and the total number of variables by

$$n = n_0 + n_1 + \cdots + n_N.$$

Let $\ell = \big(\ell^{(1)}, \ldots, \ell^{(N)}\big)$ denote all the local variables, $\bar{\ell} = \big(\overline{\ell^{(1)}}, \ldots, \overline{\ell^{(N)}}\big)$, $\hat{\ell} = \big(\widehat{\ell^{(1)}}, \ldots, \widehat{\ell^{(N)}}\big)$, let $f_0(s)$, $f_1\big(s, \ell^{(1)}\big)$, $\ldots$, $f_N\big(s, \ell^{(N)}\big)$ be C^2 (in the real variables) real-valued functions, and let $g^{(0)}(s)$, $g^{(1)}\big(s, \ell^{(1)}\big)$, $\ldots$, $g^{(N)}\big(s, \ell^{(N)}\big)$ be C^2 (in the real variables) real vector-valued functions of dimensions $m_0, m_1, \ldots, m_N$, respectively.

The quasiseparable subsystem problem is

$$\min_{s,\ell} f_0(s) + \sum_{i=1}^{N} f_i\big(s, \ell^{(i)}\big) \qquad \text{subject to} \qquad \begin{array}{c} g^{(0)}(s) \le 0, \\ g^{(i)}\big(s, \ell^{(i)}\big) \le 0,\ i = 1, \ldots, N. \end{array} \tag{SSP}$$

Precisely, a nonlinear programming problem has *quasiseparable subsystems* (or is *quasiseparable*) if it has the form (SSP). The essential ideas are to introduce a budget $b_i \in E$ for each subsystem criterion $f_i\big(s, \ell^{(i)}\big)$, and to attack the subsystem constraints $g^{(i)}\big(s, \ell^{(i)}\big) \le 0$ by maximizing the constraint margin $-\mu_i$ for each subsystem. The original problem (SSP) is decomposed into an upper level problem and N lower level problems, which are always feasible and are totally independent. The decomposed problem is (upper level)

$$\min_{s,b} f_0(s) + \sum_{i=1}^{N} b_i \qquad \text{subject to} \qquad \begin{array}{c} g^{(0)}(s) \le 0, \\ \mu_i(s, b) \le 0,\ i = 1, \ldots, N, \end{array} \tag{DSSP}$$

where $\mu_i(s,b)$ is the (global) solution to the ith lower level problem, for $i = 1, \ldots, N$, given by

$$\min_{\ell^{(i)}} \mu_i \qquad \text{subject to} \qquad \begin{aligned} \max_{1\le j\le m_i} g_j^{(i)}\big(s,\ell^{(i)}\big) - \mu_i &\le 0, \\ f_i\big(s,\ell^{(i)}\big) - b_i - \mu_i &\le 0. \end{aligned} \tag{DSSP}$$

Let X and X_D denote the feasible sets for (SSP) and (DSSP), respectively. Assume that the set $\big\{\hat\ell \mid (s,\ell) \in X\big\}$ of feasible discrete values is finite. Let $\theta(s,\bar\ell,\hat\ell)$ be the objective function of (SSP), and $\theta_D(s,b)$ the upper level objective function of (DSSP). The following lemma is straightforward to verify.

Lemma 3.1. If (s,ℓ) is feasible for (SSP), then

$$(s,b) = (s,b_1,\ldots,b_N) = \Big(s, f_1\big(s,\ell^{(1)}\big),\ldots,f_N\big(s,\ell^{(N)}\big)\Big)$$

is feasible for (DSSP).

It is also easy to check that at an optimal solution of (DSSP) there is no margin on the budget constraints, so that $b_i = f_i\big(s,\ell^{(i)}\big)$. In the simpler (ISV) case, points that were close to locally optimal solutions for (ISV) corresponded to even closer points for the upper level of (DISV) because the ℓ components were dropped. This guaranteed that the decomposition did not introduce spurious locally optimal solutions. Now, with the b variables and the discrete variables $\hat\ell$, this is not as trivial. Here the continuity of the f_i ensures that points that are close to locally optimal solutions of (SSP) correspond to nearby points in the upper level of (DSSP). The following definition and theorem imply that every locally optimal solution of (DSSP) corresponds to a locally optimal solution of (SSP).

Definition 3.2. $\big(s^*,\ell^*\big) \in X$ is a local solution for (SSP) if there exists an open ball $B\big(\big(s^*,\overline{\ell^*}\big),\delta\big)$ around $\big(s^*,\overline{\ell^*}\big)$ such that $(s,\ell)\in X$ and $(s,\bar\ell) \in B\big(\big(s^*,\overline{\ell^*}\big),\delta\big)$ imply $\theta\big(s^*,\overline{\ell^*},\widehat{\ell^*}\big) \le \theta\big(s,\bar\ell,\hat\ell\big)$.

Theorem 3.2. If $(s^*,b^*) = \Big(s^*,\ f_1\big(s^*,\ell^{(1)*}\big),\ \ldots,\ f_N\big(s^*,\ell^{(N)*}\big)\Big) \in X_D$ is a local solution for (DSSP), then $(s^*,\ell^*) \in X$ is a local solution for (SSP).

Proof. The claim is that there exist $\delta > 0$, $\epsilon > 0$ such that

$$\begin{aligned} &\theta_D(s^*,b^*) \le \theta_D(s,b) \text{ for } (s,b) \in X_D \cap B\big((s^*,b^*),\delta\big) \Longrightarrow \\ &\theta\big(s^*,\overline{\ell^*},\widehat{\ell^*}\big) \le \theta\big(s,\bar\ell,\hat\ell\big) \text{ for } (s,\ell)\in X \text{ and } \big(s,\bar\ell\big) \in B\big(\big(s^*,\overline{\ell^*}\big),\epsilon\big). \end{aligned}$$

Suppose the hypothesis holds for some $\delta > 0$, but the conclusion does not hold for any $\epsilon > 0$. Then there exists a sequence $\big(s_{(k)},\ell_{(k)}\big) \in X$ such that $\big(s_{(k)},\overline{\ell_{(k)}}\big) \to \big(s^*,\overline{\ell^*}\big)$ and

$$\theta\big(s_{(k)},\overline{\ell_{(k)}},\widehat{\ell_{(k)}}\big) < \theta\big(s^*,\overline{\ell^*},\widehat{\ell^*}\big)$$

for all k. Since the discrete part of X is assumed finite, it may be assumed that all $\widehat{\ell_{(k)}} = \widehat{\ell'}$, by reducing to a subsequence if necessary. Define ℓ' by $\overline{\ell'} = \overline{\ell^*}$, $\widehat{\ell'} = \widehat{\ell_{(k)}}$, and let $b' = \Big(f_1\big(s^*,\ell^{(1)\prime}\big),$ $\ldots,\ f_N\big(s^*,\ell^{(N)\prime}\big)\Big)$. By continuity,

$$\lim_{k\to\infty} \big(s_{(k)},\ell_{(k)}\big) = (s^*,\ell') \in X.$$

Observe that for $\|s_{(k)} - s^*\|_2 < \delta/\sqrt{2}$,

$$b_{(k)} = \Big(f_1\big(s_{(k)}, \ell^{(1)}_{(k)}\big), \ldots, f_N\big(s_{(k)}, \ell^{(N)}_{(k)}\big)\Big)$$

is not within $\delta/\sqrt{2}$ of b^*, otherwise by the assumed valid hypothesis,

$$\theta\big(s_{(k)}, \overline{\ell_{(k)}}, \widehat{\ell_{(k)}}\big) = \theta_D\big(s_{(k)}, b_{(k)}\big) \geq \theta_D(s^*, b^*) = \theta\big(s^*, \overline{\ell^*}, \widehat{\ell^*}\big).$$

Therefore again by continuity, $\lim_{k\to\infty} b_{(k)} = b'$, and some component $b'_i \leq b^*_i - \delta/\big(\sqrt{2N}N\big)$. Now from the two points (s^*, ℓ^*) and (s^*, ℓ') in X build a new point $(s, \ell) \in X$ with $s = s^*$, $\overline{\ell} = \overline{\ell^*}$, $\widehat{\ell^{(j)}} = \widehat{\ell^{(j)*}}$ for $j \neq i$, and $\widehat{\ell^{(i)}} = \widehat{\ell^{(i)\prime}}$. By Lemma 3.1, this point (s, ℓ) corresponds to $\big(s,\ b^*_1,\ \ldots,\ b^*_{i-1},\ b'_i,\ b^*_{i+1},\ \ldots,\ b^*_N\big) \in X_D$, from which can be deduced the new point

$$(s, b) = \Big(s, b^*_1, \ldots, b^*_{i-1}, b^*_i - \frac{\delta}{2\sqrt{N}N}, b^*_{i+1}, \ldots, b^*_N\Big) \in X_D$$

with $\|(s, b) - (s^*, b^*)\|_2 < \delta$ and $\theta_D(s, b) < \theta_D(s^*, b^*)$. This contradicts the assumed hypothesis, so the result follows. Q. E. D.

Just as Corollary 3.1 followed from the proof of Theorem 3.1, the proof of Theorem 3.2 immediately yields the following analog of Corollary 3.1.

Corollary 3.2. $(s^*, b^*) = \Big(s^*,\ f_1\big(s^*, \ell^{(1)*}\big),\ \ldots,\ f_N\big(s^*, \ell^{(N)*}\big)\Big) \in X_D$ is a global solution for (DSSP) $\iff$ $(s^*, \ell^*) \in X$ is a global solution for (SSP).

4. Example

This example demonstrates that the theory presented here earlier depends critically on the form of the inequalities, and that it becomes invalid if strict inequalities are permitted in the definition of constraints. That is, with strict inequalities, it is possible to get a spurious solution from the decomposed problem (Theorem 2.1 is false).

Consider a simple budgeting problem where a manager has to decide on dividing work and allocating funds for computer workstations for two departments. The manager can decide what fraction s of a total workload of 10 units per day to assign to Department A, with the remainder $(1 - s)$ assigned to Department B. Each unit of work (or fraction thereof) requires a workstation. Let local subsystem variables $\ell^{(1)}$ and $\ell^{(2)}$ denote the number of workstations to be purchased for Departments A and B, respectively. Due to the nature of the work, the workstation required by Department A costs \$2,000 and the one for Department B costs \$8,000. Because of personnel availability it is desirable to have s as close to 0.65 as possible, so introduce a dollar penalty of $2{,}400{,}000(0.65 - s)^2$ for deviating from this value, reflecting expenses for temporary workers.

The objective is to minimize the cost (in units of \$1,000):

$$\min_{s,\ell} 2400(0.65 - s)^2 + 2\ell^{(1)} + 8\ell^{(2)}.$$

The workload constraints are

$$-\ell^{(1)} + 10s \leq 0,$$
$$-\ell^{(2)} + 10(1-s) < 0.$$

with the strict inequality constraint for Department B reflecting contractual requirements. Bounds on variables are

$$\ell^{(1)}, \ell^{(2)} \geq 0, \qquad \ell^{(1)}, \ell^{(2)} \in Z; \qquad 0 \leq s \leq 1, \qquad s \in E.$$

The (DSSP) decomposition formulation is

$$\min_{s,b} 2400(0.65-s)^2 + \sum_{i=1}^{2} b_i \qquad \text{subject to} \qquad \begin{aligned} -s &\leq 0, \\ s-1 &\leq 0, \\ \mu_i(s,b) &\leq 0, \end{aligned}$$

where (Department A) μ_1 solves

$$\min_{\ell^{(1)}} \mu_1 \qquad \text{subject to} \qquad \begin{aligned} -\ell^{(1)} + 10s - \mu_1 &\leq 0, \\ 2\ell^{(1)} - b_1 - \mu_1 &\leq 0, \end{aligned}$$

and (Department B) μ_2 solves

$$\inf_{\ell^{(2)}} \mu_2 \qquad \text{subject to} \qquad \begin{aligned} -\ell^{(2)} + 10(1-s) - \mu_2 &< 0, \\ 8\ell^{(2)} - b_2 - \mu_2 &\leq 0. \end{aligned}$$

By direct verification the only (local or global) solution of the original problem lies at $\left(s, \ell^{(1)}, \ell^{(2)}\right) = (0.65, 7, 4)$ for a total cost of exactly \$46,000. This corresponds to expenditures of \$14,000 by Department A, \$32,000 by Department B, and no system-wide penalty for deviation from the optimal allocation of manpower.

For the decomposed problem, the corresponding local solution is $(s, b_1, b_2) = (0.65, 14, 32)$. However, the point $(s, b_1, b_2) = (0.5, 10, 48)$ is another local solution of the decomposed problem with a total cost of \$112,000. For this latter point, the lower level optimization yields $\ell^{(1)} = 5$ and $\ell^{(2)} = 6$. Decreasing s from 0.5 leaves the workstation cost unchanged and increases the manpower cost. Increasing s by an arbitrary small positive value ϵ violates the workload constraint for Department A, forcing $\ell^{(1)}$ to increase by 1 and the budget for Department A to increase to \$12,000. Simultaneously $\ell^{(2)}$ can be decreased by 1 and the budget for Department B can be reduced to \$40,000 giving a total cost of ≈\$106,000. However, there is no continuous feasible path from $(s, b_1, b_2) = (0.5, 10, 48)$ to $(s, b_1, b_2) = (0.5+\epsilon, 12, 40)$, so there is a local minimum at $(s, b) = (0.5, 10, 48)$. This shows that there is a local solution of the decomposed problem (DSSP) that does *not* correspond to a local solution of (SSP), violating Theorem 2.1. (In fact, (DSSP) has a local minimum for each $s = 0.1j$, $j = 0, 1, \ldots, 6$, by the same reasoning as for $s = 0.5$.)

Acknowledgement

This work was supported in part by Air Force Research Laboratory Grant F30602-01-2-0572, NSF Grant EIA-0103660, and AFOSR Grant F49620-02-1-0090.

References

N. M. Alexandrov, "Multilevel methods for MDO", in *Multidisciplinary Design Optimization*, N. M. Alexandrov and M. Y. Hussaini (eds.), SIAM, Philadelphia, PA, 1997, pp. 79–89.

V. O. Balabanov, A. A. Giunta, O. Golovidov, B. Grossman, W. H. Mason, L. T. Watson, and R. T. Haftka, "Reasonable design space approach to response surface approximation", *J. Aircraft* vol. 36, pp. 308–315, 1999.

A.A. Giunta, V. Balabanov, D. Haim, B. Grossman, W.H. Mason, L.T. Watson, and R.T. Haftka, "Multidisciplinary optimisation of a supersonic transport using design of experiments theory and response surface modelling", *Aeronautical J.* vol. 101, no. 1008, pp. 347–356, 1997a.

A. A. Giunta, V. Balabanov, M. Kaufman, S. Burgee, B. Grossman, R. T. Haftka, W. H. Mason, and L. T. Watson, "Variable-complexity response surface design of an HSCT configuration", in *Multidisciplinary Design Optimization*, N. M. Alexandrov and M. Y. Hussaini (eds.), SIAM, Philadelphia, PA, 1997b, pp. 348–367.

R. T. Haftka and L. T. Watson, "Multidisciplinary design optimization with quasiseparable subsystems", *Optim. Engrg.* to appear, 2004.

A.M. Harte, J.F. McNamara, and J.D. Roddy, "A multilevel approach to the optimisation of a composite light rail vehicle body shell", *Composite Structures* vol. 63(3–4), pp. 447–453, 2004.

S. Hosder, L. T. Watson, B. Grossman, W. H. Mason, H. Kim, R. T. Haftka, and S. Cox, "Polynomial response surface approximations for the multidisciplinary design optimization of a high speed civil transport", *Optim. Engrg.* vol. 2, pp. 431–452, 2001.

I. Kroo, "MDO for large-scale design", in *Multidisciplinary Design Optimization*, N. M. Alexandrov and M. Y. Hussaini (eds.), SIAM, Philadelphia, PA, 1997, pp. 22–44.

B. Liu and R.T. Haftka, "Single-level composite wing optimization based on flexural lamination parameters", *Structural Multidisciplinary Optimization* to appear, 2003.

B. Liu, R.T. Haftka, and M.A. Akgün, "Two-level composite wing structural optimization using response surfaces", *Structural Optimization* vol. 20, no. 2, pp. 87–96, 2000.

B. Liu, R. T. Haftka, and L. T. Watson, "Global-local structural optimization using response surfaces of local optimization margins", *Structural Multidisciplinary Optimization* to appear, 2004.

S.A. Ragon, Z. Gürdal, R.T. Haftka, and T.J. Tzong, "Global/local structural wing design using response surface techniques", in *38th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conf.*, AIAA Paper 97-1051, Kissimmee, FL, April 7–10, 1997, Part 2, pp. 1204–1214..

J. F. Rodríguez, J. E. Renaud, and L. T. Watson, "Trust region augmented Lagrangian methods for sequential response surface approximation and optimization", *ASME J. Mech. Design* vol. 120, pp. 58–66, 1998a.

J. F. Rodríguez, J. E. Renaud, and L. T. Watson, "Convergence of trust region augmented Lagrangian methods using variable fidelity approximation data", *Structural Optim.* vol. 15, pp. 141–156, 1998b.

J. Shankar, C. J. Ribbens, R. T. Haftka, and L. T. Watson, "Computational study of a nonhierarchical decomposition algorithm", *Comput. Optim. Appl.* vol. 2, pp. 273–293, 1993.