Analytical and computational properties of distributed approaches to MDO

Natalia Alexandrov
Multidisciplinary Optimization Branch

Michael Lewis
ICASE

NASA Langley Research Center
Hampton, Virginia

This work was supported by the National Aeronautics and Space Administration under NASA Contract NAS1-97046.

Problem formulation and computation

Analytical properties of MDO problem formulations have a direct and powerful influence on the practical solution of the resulting computational optimization problem.

Difficulties may be introduced by attempts to achieve desirable goals.

Outline

1. Two-discipline example: a canonical formulation

2. Some representative distributed formulations:

   - Collaborative optimization (CO).
   - Optimization by linear decomposition (OLD).

3. Analytical features of approaches based on discrepancy functions:

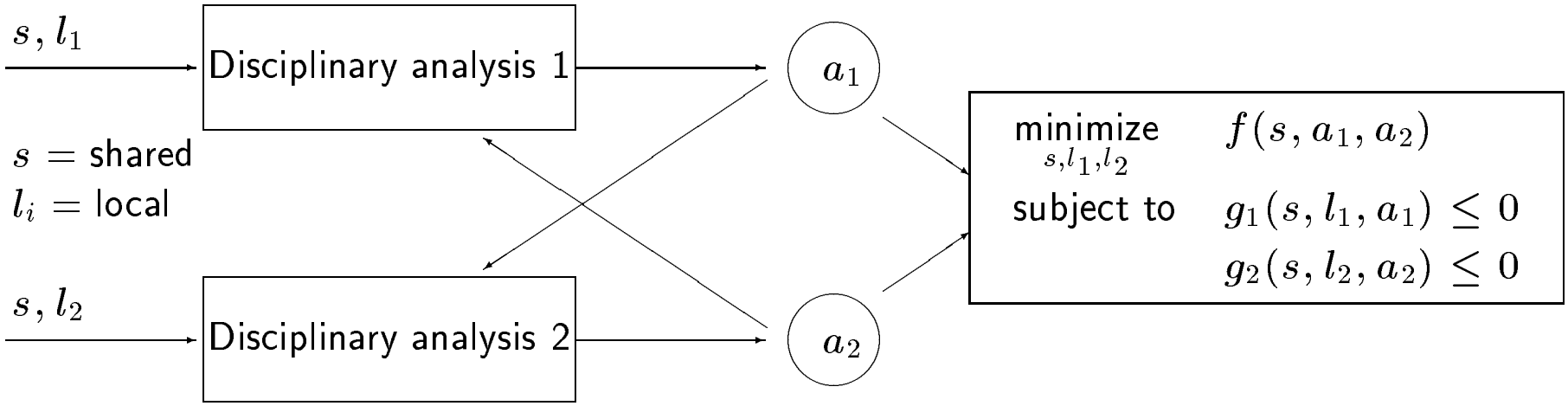
   - Breakdown of the KKT conditions.
   - Non-smoothness of the constraints.

4. Possible remedies:

   - Trilevel methods

5. Conjectures on bilevel approaches to MDO

Two discipline example: Fully integrated optimization



Simplifying assumption: no design constraints g_i involve both a_1 and a_2.

Multidisciplinary analysis:
$$\begin{aligned} a_1 &= A_1(s, l_1, a_2) \\ a_2 &= A_2(s, l_2, a_1) \end{aligned}$$

Motivation for distributed formulations

- Single-level formulations:
  - Fully integrated optimization: multidisciplinary analysis + optimization.
  - Simultaneous analysis and design: multidisciplinary analysis treated as equality constraints in optimization.
  - Distributed analysis optimization: intermediate to preceding (later talk).
- Features of single-level formulations:
  - Conventional nonlinear programming approaches—no analytical or computational surprises.
  - Extensive disciplinary autonomy in implementation.
- May wish to dispense with MDA and maximize disciplinary autonomy in execution.

Distributed bilevel formulations

Idea: Eliminate disciplinary design variables l_1, l_2 via disciplinary optimization problems.

Representative approaches:

Optimization by linear decomposition (OLD):

- Maintain interdisciplinary consistency.
- Disciplinary level: minimize violation of the disciplinary design constraints.

Sobieski (1982); Sobieski, James, and Dovi (1983); Barthelemy (1983); Sobieski, James, and Riley (1987); Sobieski (1993); Balling and Sobieski (1994)

Collaborative optimization (CO):

- Satisfy disciplinary design constraints.
- Disciplinary level: minimize violation of interdisciplinary consistency.

Ramanathan and Schmit (1978); Schmit and Mehrinfar (1982); Schmit and Chang (1984); Thareja and Haftka (1986);

Adelman, Walsh, and Pritchard (1992); Walsh, Young, Pritchard, Adelman, and Mantay (1995)

CO: Braun (1996); Braun and Kroo (1995); Braun, Moore, and Kroo (1997); I. Sobieski and Kroo (1996); Manning (1999)

Mathematical structure of methods using discrepancy functions

System-level problem:

$$\begin{array}{ll} \underset{(s,l_1,l_2)}{\text{minimize}} & f(s,t_1,t_2) \\ \text{subject to} & c_1(s,t_1,t_2) = 0 \\ & c_2(s,t_1,t_2) = 0. \end{array}$$

c_i: discrepancy function associated with Discipline i.

c_i is derived from the optimal solution of the disciplinary optimization problem.

Example (CO):

$$\begin{array}{ll} \underset{s}{\text{minimize}} & f(s) = s \\ \text{subject to} & 0 \leq s \leq 1. \end{array}$$

Subsystem problems:

$$\begin{array}{cc} \underset{\sigma_1}{\min} \quad \frac{1}{2} \| \sigma_1 - s \|^2 & \underset{\sigma_2}{\min} \quad \frac{1}{2} \| \sigma_2 - s \|^2 \\ \text{s.t.} \qquad \sigma_1 \geq 0 & \text{s.t.} \qquad \sigma_2 \leq 1. \end{array}$$

Subsystem solutions:

$$\bar{\sigma}_1(s) = \begin{cases} 0 & \text{if } s \leq 0 \\ s & \text{if } s \geq 0 \end{cases} \qquad \bar{\sigma}_2(s) = \begin{cases} s & \text{if } s \leq 1 \\ 1 & \text{if } s \geq 1. \end{cases}$$

CO_2 system-level problem:

$$\begin{array}{ll} \underset{s}{\text{minimize}} & s \\ \text{subject to} & c_1(s) = \frac{1}{2} \| \bar{\sigma}_1(s) - s \|^2 = 0 \\ & c_2(s) = \frac{1}{2} \| \bar{\sigma}_2(s) - s \|^2 = 0, \end{array}$$

CO_1 system-level problem:

$$\begin{array}{ll} \underset{s}{\text{minimize}} & s \\ \text{subject to} & c_1(s) = \bar{\sigma}_1(s) - s = 0 \\ & c_2(s) = \bar{\sigma}_2(s) - s = 0. \end{array}$$

Example (CO):

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}(a_1^2(l_1, l_2) + 10\, a_2^2(l_1, l_2)) \\ \text{subject to} & s + l_1 \leq 1 \\ & -s + l_2 \leq -2, \end{array}$$

Subsystem problems:

$$\begin{array}{ll} \underset{\sigma_1, l_1}{\text{minimize}} & \frac{1}{2}\left[\| \sigma_1 - s \|^2 + \| a_1(\sigma_1, l_1, t_2) - t_1 \|^2\right] \\ \text{subject to} & \sigma_1 + l_1 \leq 1, \\ \text{where} & 2a_1 + t_2 = l_1. \end{array}$$

$$\begin{array}{ll} \underset{\sigma_2, l_2}{\text{minimize}} & \frac{1}{2}\left[\| \sigma_2 - s \|^2 + \| a_2(\sigma_2, l_2, t_1) - t_2 \|^2\right] \\ \text{subject to} & -\sigma_2 + l_2 \leq -2, \\ \text{where} & t_1 + 2a_2 = l_2. \end{array}$$

Subsystem solutions:

$$
\begin{aligned}
\bar{\sigma}_1(s, t_1, t_2) &= s + 1/5 \ \ \mathsf{min}((-s - 2t_1 - t_2 + 1), 0) \\
\bar{l}_1(s, t_1, t_2) &= 2t_1 + t_2 + 4/5 \ \ \mathsf{min}((-s - 2t_1 - t_2 + 1), 0) \\
\bar{\sigma}_2(s, t_1, t_2) &= s + 1/5 \ \ \mathsf{max}((-s + t_1 + 2t_2 + 2), 0) \\
\bar{l}_2(s, t_1, t_2) &= t_1 + 2t_2 - 4/5 \ \ \mathsf{max}((-s + t_1 + 2t_2 + 2), 0).
\end{aligned}
$$

Scalar discrepancy functions: singularity and breakdown of KKT conditions

For a scalar discrepancy function $c_i \geq 0$, we necessarily have $\nabla c_i(s, t_1, t_2) = 0$ at all feasible points.

$\Rightarrow$ Nonexistence of Lagrange multipliers is unavoidable.

$\Rightarrow$ Trouble for conventional optimization algorithms.

Illustration (CO_2):

$$\begin{array}{ll} \underset{s}{\text{minimize}} & s \\ \text{subject to} & c_1(s) = \frac{1}{2} \| \bar{\sigma}_1(s) - s \|^2 = 0 \\ & c_2(s) = \frac{1}{2} \| \bar{\sigma}_2(s) - s \|^2 = 0, \end{array}$$

$$\nabla c_1(s) = \begin{cases} s & \text{if } s \leq 0 \\ 0 & \text{if } s \geq 0 \end{cases}, \quad \nabla c_2(s) = \begin{cases} 0 & \text{if } s \leq 1 \\ s - 1 & \text{if } s \geq 1. \end{cases}$$

Violation of KKT condition (multiplier rule) at $s = 0$:

$$\nabla f(s) + \lambda_1 \nabla c_1(s) + \lambda_2 \nabla c_2(s) \quad = \quad 1 \quad \neq \quad 0.$$

Effect of singular constraints

NPSOL applied to CO_2 reformulation of one-variable LP:

Iteration	s	Penalty	Cumulative work
0	1.000e-03	0.0e+00	1
1	-9.990e-01	4.2e+00	2
2	-9.847e-01	5.7e+00	4
3	-8.282e-01	7.4e+00	6
4	-4.142e-01	2.7e+01	7
5	-3.430e-01	5.9e+01	9
6	-1.718e-01	4.0e+02	10
7	-1.436e-01	8.2e+02	12
8	-7.251e-02	5.4e+03	13
9	-6.076e-02	1.1e+04	15
10	-3.203e-02	6.5e+04	16
11	-2.717e-02	1.2e+05	18
12	-1.727e-02	5.1e+05	19
13	-1.442e-02	1.9e+06	20
14	-1.414e-02	4.7e+06	21

For QP example, convergence to non-solution can occur.

Vector-valued discrepancy functions

For a vector-valued discrepancy function c_i of the type that arises in CO_1, the system-level constraint derivatives are necessarily discontinuous on multiple surfaces of design variables, including the solution.

$\Rightarrow$ Nonsmoothness at solutions is unavoidable.

Vector-valued discrepancy functions also suffer from singular Jacobians.

$\Rightarrow$ Trouble for conventional (smooth) optimization algorithms.

Illustration (CO_1):

$$\nabla c_1(s) = \begin{cases} -1 & \text{if } s \leq 0 \\ 0 & \text{if } s \geq 0 \end{cases}, \quad \nabla c_2(s) = \begin{cases} 0 & \text{if } s \leq 1 \\ 1 & \text{if } s \geq 1. \end{cases}$$

The constraints are discontinuous at $s = 0$.

Generic properties of approaches based on discrepancy functions: recapitulation

- Scalar discrepancy functions:
  - Singularity of system-level constraints at feasible points.
  - Nonexistence of Lagrange multipliers.
- Vector discrepancy functions:
  - Nonsmoothness of system-level constraints at solutions.
  - Singularity of system-level constraints at feasible points.

System-level constraints are more nonlinear than those in a single-level approach.

$\Rightarrow$ Trouble for conventional (smooth) optimization algorithms.

In general, there is an increase in the computational effort (compared to a single-level approach).

Possible remedies

1. Single-level formulations.
2. ⋆ Relaxation or approximation of system-level constraints.
3. ⋆ Approximation of disciplinary problems.
4. ⋆ Trilevel approaches.
5. Nonsmooth optimization algorithms
6. Ignore the problem(s).

Relaxation of system-level constraints

Treat the system-level scalar discrepancy constraints as inequalities: $c_i \leq \varepsilon$, for ε suitably close to 0.

Challenge: Choose ε sufficiently small that the solution of the relaxed problem is close to the solution of the real problem, but not so close that the computational difficulties re-emerge (see paper).

The tolerance ε involves the MDA, not just direct design variable mismatch.

Related penalty function approach:

$$\text{minimize} \quad f + w(c_1 + c_2).$$

Alternatively: response surface (e.g., polynomial, spline) approximations of system-level constraints.

System-level problem with KS approximation

System-level problem:

$$\begin{array}{ll} \underset{(s,l_1,l_2)}{\text{minimize}} & f(s,t_1,t_2) \\ \text{subject to} & KS_1(s,t_1,t_2;\ \rho) \leq \varepsilon \\ & KS_2(s,t_1,t_2;\ \rho) \leq \varepsilon. \end{array}$$

KS_i: optimal value of unconstrained disciplinary problem with Kreisselmeier–Steinhauser cumulative objective.

Constraints are (in general) smooth and non-singular.

Moreover, approximation is conservative—disciplinary design are satisfied by disciplinary design variables.

Sobieski (1993), Balling and Sobieski (1994)

Trilevel approaches

Trilevel OLD/CO approach: Solve a sequence of "nice" bilevel problems that are increasingly like the exact non-smooth OLD/CO problem.

Two simple schemes:

1. Treat the system-level scalar discrepancy constraints as inequalities: $c_i \leq \varepsilon$. Solve a sequence of bilevel problems, letting $\varepsilon \to 0$.

   Treat the scalar discrepancy constraints as penalty terms, and let the penalty weight $\to \infty$.

2. Solve a sequence of bilevel problems using the KS approximation to the disciplinary optimization problem, and let $\rho \to \infty$.

Alternative scheme: DeMiguel and Murray (later talk).

In all cases, the outermost sequence of problems becomes increasingly badly behaved.

Trilevel primal-dual algorithms

Penalty term in augmented Lagrangian couples disciplinary calculations.

Proposed trilevel solutions:

- Stephanopoulos and Westerberg (1975): Separable approximation of penalty term.
- Watanabe, Nishimura, and Matsubara (1978): Non-separable penalty terms computed as solution of (yet another) optimization problem.
- Bertsekas (1979): Minimization of Yosida–Moreau regularization of the single-level problem.
- De Luca and Di Pillo (1987): Exact penalty function with trilevel computation of primal and dual variables.

All appear to suffer a loss of computational efficiency.

Lessons learned

Problem formulation has practical consequences for computation.

Analytical and computational difficulties may be introduced by the choice of MDO reformulation.

Approaches based on discrepancy functions illustrate this observation.

Some conjectures

Conjecture: There is no approach to MDO problem formulation that is

> Bilevel AND efficient AND robust AND highly autonomous in execution (parallel across disciplines) AND exact.

Corollary: We seem to be forced to single-level or trilevel schemes.

Conjecture: Parallel autonomy of execution may be generally at odds with overall efficiency.

Conjecture: At least when MDO = nonlinear programming, autonomy of implementation is at least as important as autonomy of execution.