RC 20316(12/18/95)
Mathematics

IBM Research Report

Implementation of dual decomposition in OSL

Alan J. King

IBM Research Division
T.J. Watson Research Center
Yorktown Heights, New York

LIMITED DISTRIBUTION NOTICE

This report has been submitted for publication outside of IBM and will probably be copyrighted if accepted for publication. It has been issued as a Research Report for early dissemination of its contents. In view of the transfer of copyright to the outside publisher, its distribution outside of IBM prior to publication should be limited to peer communications and specific requests. After outside publication, requests should be filled only by reprints or legally obtained copies of the article (e.g., payment of royalties).

Research Division
Almaden · T.J. Watson · Tokyo · Zurich

IMPLEMENTATION OF DUAL DECOMPOSITION IN OSL

Alan J. King[†]

Abstract. This paper discusses the implementation in the Optimization Subroutine Library of dual decomposition, also known as *Benders decomposition* in mixed integer programming and the *L-shaped method* in stochastic programming. The formulation of optimality and feasibility cuts is explained and the correspondence to OSL variables is indicated.

1. Introduction.

Decomposition is a well-known technique for taking advantage of structural properties of mathematical programs. Consider a linear program where the variables can be divided into two sets as follows

$$(P) \qquad \begin{array}{ll} \min\limits_{x_1,x_2} & c_1x_1 + c_2x_2 \\ \text{subject to} & A_{11}x_1 \in [y_1^-, y_1^+] \\ & A_{21}x_1 + A_{22}x_2 \in [y_2^-, y_2^+] \\ & x_1 \in [x_1^-, x_1^+] \\ & x_2 \in [x_2^-, x_2^+] \end{array}$$

The matrix of this problem is lower block triangular. If it is relatively easy to optimize over the first or second set of variables considered separately, some advantage might be gained by partitioning the variables of the problem.

[†] IBM Research Division, Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, NY 10598, U.S.A.

Decomposition for linear programmming was first discussed by Dantzig and Wolfe in [2] as a scheme for partitioning the rows of a linear program. Dual decomposition partitions the variables of a linear program. It was studied by Benders [1] for application in mixed integer programming, and in the form of the L-shaped method, by van Slyke and Wets [5] for stochastic programming.

This brief paper discusses the particular implementation of the dual decomposition in the structures of the Optimization Subroutine Library [4]. A multicut version has been encoded in the stochastic programming software package [3].

2. Optimality and Feasibility Cuts

Dual decomposition forms the *master problem* from (P) by minimizing out the second set of variables x_2

$$(P_1) \qquad \begin{aligned} \min_{x_1,\theta_1} \quad & c_1 x_1 + \theta_1 \\ \text{subject to} \quad & A_{11} x_1 \in [y_1^-, y_1^+] \\ & x_1 \in [x_1^-, x_1^+] \\ & \theta_1 \geq f_2(x_1) \end{aligned}$$

The impact of the x_2 variables is represented abstractly in the master by the constraint

$$\theta_1 \geq f_2(x_1)$$

involving the function $f_2()$, which is the value of the *slave problem*

$$(P_2) \qquad \begin{aligned} f_2(\hat{x}_1) = \min_{x_1,x_2} \quad & c_2 x_2 \\ \text{subject to} \quad & A_{21} x_1 + A_{22} x_2 \in [y_2^-, y_2^+] \\ & x_1 = \hat{x}_1 \\ & x_2 \in [x_2^-, x_2^+] \end{aligned}$$

The abstract constraint has two effects: first, when a master proposal x_1 is feasible it communicates the optimal value of the slave to the master objective, and second, it represents the implied feasibility conditions imposed on master variables by the slave problem.

Dual decomposition iteratively approximates the abstract constraint by computing *optimality cuts* (involving θ_1 and x_1 master problem variables) and *feasibility cuts* (involving only x_1 variables) from the optimality conditions of the slave problem for successive master proposals.

Consider a master proposal $\hat{x}_1$ that has finite optimal value in the slave problem. The dual of the slave problem may be written

$$(D_2) \qquad \begin{array}{ll} \max\limits_{\delta_1,\delta_2,\pi_2} & \delta_1\hat{x}_1 - [\delta_2]^- x_2^+ + [\delta_2]^+ x_2^- - [\pi]^- y_2^+ + [\pi]^+ y_2^- \\ \text{subject to} & \delta_1 = -\pi_2 A_{21} \\ & \delta_2 = c_2 - \pi_2 A_{22} \end{array}$$

(The notation $[u]^-$ or $[u]^+$ equals the absolute value of the argument when u is respectively negative or positive and is equal to zero otherwise. Any infinite primal bound corresponds to a dual constraint. For example: if x_{2j}^+ is $+\infty$ then the dual objective term $-[\delta_{2j}]^- x_{2j}^+$ requires $[\delta_{2j}]^- = 0$, implying the constraint $\delta_{2j} \geq 0$. The contribution to the dual objective from such a term is zero.) This form of the dual highlights the *reduced costs* δ_1 and δ_2.

Note that the dual feasibility of the optimal solution $(\hat{\delta}_1, \hat{\delta}_2, \hat{\pi}_2)$ is independent of the master proposal, so for any master proposal x_1 one has

$$f_2(x_1) \geq \hat{\delta}_1 x_1 - [\hat{\delta}_2]^- x_2^+ + [\hat{\delta}_2]^+ x_2^- [\hat{\pi}]^- y_2^+ + [\hat{\pi}]^+ y_2^-$$

Therefore one can add the *optimality cut*

$$\theta_1 - \hat{\delta}_1 x_1 \geq f_2(\hat{x}_1) - \hat{\delta}_1 \hat{x}_1 \tag{2.1}$$

to the master problem without affecting the optimal value. This inequality expresses the fact that the reduced-cost vector $\hat{\delta}_1$ is a subgradient of the value function f_2 at the master proposal $\hat{x}_1$.

A proposal that leads to an infeasible slave problem generates a positive optimal value to the phase I dual problem

$$(D_2^I) \qquad \begin{array}{lll} f_2^I(\hat{x}_1) = \max\limits_{\delta_1,\delta_2,\pi_2} & \delta_1\hat{x}_1 - [\delta_2]^- x_2^+ + [\delta_2]^+ x_2^- - [\pi]^- y_2^+ + [\pi]^+ y_2^- \\ \text{subject to} & \delta_1 = -\pi_2 A_{21} \\ & \delta_2 = -\pi_2 A_{22} \\ & (\delta_1, \delta_2, \pi_2) \in [-e, e] \end{array}$$

The symbol e denotes the vector (of the appropriate dimension) whose entries are all ones. The primal phase I objective value $f_2^I(\hat{x}_1)$ turns out to be the sum of the primal infeasibilities, each infeasibilitiy being computed as the distance from the primal variable value or primal constraint value to its nearest bound. The dual phase I problem always has an optimal solution $(\hat{\delta}_1, \hat{\delta}_2, \hat{\pi}_2)$. Any master proposal that is feasible for the slave problem has phase I optimal value equal to zero and hence must satisfy

$$\hat{\delta}_1 x_1 - [\hat{\delta}_2]^- x_2^+ + [\hat{\delta}_2]^+ x_2^- - [\hat{\pi}]^- y_2^+ + [\hat{\pi}]^+ y_2^- \leq 0$$

But the infeasible $\hat{x}_1$ has

$$\hat{\delta}_1\hat{x}_1 - [\hat{\delta}_2]^- x_2^+ + [\hat{\delta}_2]^+ x_2^- - [\hat{\pi}]^- y_2^+ + [\hat{\pi}]^+ y_2^- > 0$$

The *feasibility cut*

$$-\hat{\delta}_1 x_1 \geq f_2^I(\hat{x}_1) - \hat{\delta}_1\hat{x}_1 \tag{2.2}$$

cuts off the infeasible $\hat{x}_1$ but leaves feasible all master proposals that are feasible for the slave problem.

If the master problem is unbounded, then the master proposal may come in the form of a ray: $\hat{x}_1 + t\hat{u}_1, t \geq 0$. One must determine whether the ray direction $\hat{u}_1$ can be combined with a ray u_2 for the slave problem that together form a direction of unboundedness for the original problem. To do this, transform the primal bounds $[x_2^-, x_2^+]$ and $[y_2^-, y_2^+]$ into their *recession* bounds $[\xi_2^-, \xi_2^+]$ and $[\eta_2^-, \eta_2^+]$ by setting all finite bounds to zero, for example

$$x_{2j}^+ \mapsto \xi_{2j}^+ := \begin{cases} 0 & \text{if} \quad x_{2j}^+ < +\infty \\ +\infty & \text{if} \quad x_{2j}^+ = +\infty \end{cases}$$

for all j and all bounds $(x_2^-, x_2^+, y_2^-, y_2^+)$, and solve the slave *recession problem*

$$(P_2^R) \qquad \begin{aligned} f_2^R(\hat{u}_1) = \min_{u_1, u_2} \quad & c_2 u_2 \\ \text{subject to} \quad & A_{21}u_1 + A_{22}u_2 \in [\eta_2^-, \eta_2^+] \\ & u_1 = \hat{u}_1 \\ & u_2 \in [\xi_2^-, \xi_2^+] \end{aligned}$$

If this problem is feasible and $c_1\hat{u}_1 + c_2\hat{u}_2 < 0$ then the original problem is unbounded along the feasible ray $(\hat{x}_1, \hat{x}_2) + t(\hat{u}_1, \hat{u}_2), t \geq 0$. Otherwise, the dual information from the solution of the recession problem may be used to construct cuts for the master.

The dual recession problem is

$$(D_2^R) \qquad \begin{aligned} \max_{\delta_1, \delta_2, \pi_2} \quad & \delta_1\hat{u}_1 - [\delta_2]^-\xi_2^+ + [\delta_2]^+\xi_2^- - [\pi]^-\eta_2^+ + [\pi]^+\eta_2^- \\ \text{subject to} \quad & \delta_1 = -\pi_2 A_{21} \\ & \delta_2 = c_2 - \pi_2 A_{22} \end{aligned}$$

Observe that the feasibility region is the same as the slave dual (D_2); the only differences are that the finite-valued objective coefficients in (D_2) have been zeroed in (D_2^R). An

optimal solution $(\hat{\delta}_1, \hat{\delta}_2, \hat{\pi}_2)$ for the recession problem is also feasible for (D_2) so by the same argument justifying (2.1) it forms an *optimality cut*

$$\theta_1 - \hat{\delta}_1 x_1 \geq -[\hat{\delta}_2]^- x_2^+ + [\hat{\delta}_2]^+ x_2^- - [\hat{\pi}]^- y_2^+ + [\hat{\pi}]^+ y_2^- \tag{2.3}$$

When the recession problem is infeasible, the optimal phase I solution may be used to form a *feasibility cut*

$$-\hat{\delta}_1 x_1 \geq -[\hat{\delta}_2]^- x_2^+ + [\hat{\delta}_2]^+ x_2^- - [\hat{\pi}]^- y_2^+ + [\hat{\pi}]^+ y_2^- \tag{2.4}$$

To verify this, note that $\hat{u}_1$ is infeasible so $\hat{\delta}_1 \hat{u}_1 > 0$ and hence (2.4) is violated by $\hat{x}_1 + t\hat{u}_1$ for all sufficiently large t. On the other hand any master proposal x_1 that is feasible for the slave must satisfy (2.4) since the optimal phase I dual solution is feasible for (D_2^I). The original primal bounds, not the recession bounds, are used as coefficients in computing both cuts.

3. The Algorithm

Dual decomposition proceeds iteratively by generating master proposals for the slave problem, computing the relevant cuts, adding the cuts to the master, and solving the resulting master problem to obtain a new master proposal.

The master problem after the k-th step of the method has the form

$$(P_1^k) \qquad \begin{array}{ll} \min\limits_{x_1, \theta_1} & c_1 x_1 + \theta_1 \\ \text{subject to} & A_{11} x_1 \in [y_1^-, y_1^+] \\ & x_1 \in [x_1^-, x_1^+] \\ & \theta_1 e + F^k x_1 \geq f^k \\ & G^k x_1 \geq g^k \end{array}$$

The rows of the matrices F^k and G^k and the right-hand side vectors f^k and g^k are formed by the collection of optimality and infeasibility cuts, respectively.

In van Slyke and Wets [5] it is shown that the following algorithm terminates in a finite number of steps.

Dual Decomposition Algorithm.

0. Solve the master problem (P_1^k).
   - 0.1 If infeasible, problem (P) is infeasible. STOP.
   - 0.2 If unbounded, go to step 2 with ray $\hat{x}_1^k + t\hat{u}_1^k, t \geq 0$.

0.3 If optimal, go to step 1 with solution $(\hat{x}_1^k, \hat{\theta}_1^k)$.

1. Solve the slave problem (P_2) with master proposal $\hat{x}_1^k$.
   1.1 If unbounded, problem (P) is unbounded. STOP.
   1.2 If infeasible, generate infeasibility cut (2.2) and add coefficients to matrix G^k and right-hand side g^k. Go to step 0 with $k \mapsto k+1$.
   1.3 (Optimality test) If optimal and the optimality gap $f_2(\hat{x}_1^k) - \hat{\theta}_1^k$ is within tolerance, optimal solution is found. STOP.
   1.4 Otherwise, generate optimality cut (2.1) and add coefficients to matrix F^k and right-hand side f^k. Go to step 0 with $k \mapsto k+1$.
2. (Unboundedness procedure) Solve the slave recession problem (P_2^R) with master ray $\hat{u}_1^k$.
   2.1 If infeasible, generate infeasibility cut (2.4) and add coefficients to matrix G^k and right-hand side g^k. Go to step 0 with $k \mapsto k+1$.
   2.2 If optimal and $c_1\hat{u}_1^k + c_2\hat{u}_2^k \geq 0$, generate optimality cut (2.3) and add coefficients to matrix F^k and right-hand side f^k. Go to step 3 with $k \mapsto k+1$.
3. Solve the slave problem (P_2) with master proposal $\hat{x}_1^k$.
   3.1 If infeasible, generate infeasibility cut (2.2) and add coefficients to matrix G^k and right-hand side g^k. Go to step 0 with $k \mapsto k+1$.
   3.2 If feasible and $c_1\hat{u}_1^k + c_2\hat{u}_2^k < 0$, problem (P) is unbounded. STOP.

4. OSL Correspondences.

The problem data referenced in the paper correspond to the following indices in OSL (see the documentation for the OSL subroutine `ekknget`). A word of warning: OSL returns the *negative* of the dual solutions.

objective coefficients	c	`NOBJECTIVE`
column lower bounds	x^-	`NCOLLOWER`
column upper bounds	x^+	`NCOLUPPER`
row lower bounds	y^-	`NROWLOWER`
row upper bounds	y^+	`NROWUPPER`
primal solution	x	`NCOLSOL`
reduced costs	δ	`NCOLRCOSTS`
negative dual solution	$-\pi$	`NROWDUALS`

When a problem is infeasible OSL returns the phase I reduced costs δ and negative duals $-\pi$ at the pointers `NCOLAUX` and `NROWAUX` respectively.

The optimal value of a linear program (for example, as required in (2.1)) is returned in the OSL variable `ROBJVALUE`. The sum of primal infeasibilities for an infeasible problem (as required in (2.2)) is returned in `RSUMPINF`. See the documentation for the OSL subroutine `ekkrget`.

References.

[1] J.F. Benders, "Partitioning methods for solving mixed variables programming problems", *Numerische Mathematik* **4** (1962) 238-252.

[2] G.B. Dantzig and P. Wolfe, "The decomposition algorithm for linear programs", *Econometrica* **29** (1961).

[3] A.J. King, "SP/OSL version 1.0, stochastic programming interface user's guide", (Research Report RC 19757, IBM T.J. Watson Research Center, P.O. Box 218, Yorktown Heights, NY 10598, 1994).

[4] IBM Corporation, *Optimization Subroutine Library Guide and Reference* (Publication SC23-0519-0, 1990).

[5] R.M. van Slyke and R. J-B Wets, "L-shaped linear programs with application to optimal control and stochastic programming", *SIAM J. Appl. Math.* **17** (1969) 638–663.