A NEW, SOLVABLE, PRIMAL RELAXATION

FOR CONVEX NONLINEAR INTEGER PROGRAMMING PROBLEMS

by

Monique GUIGNARD[1,2]
Department of OPIM
The Wharton School, University of Pennsylvania

Keywords: nonlinear integer programming, relaxation, simplicial decomposition, convex hull.

Abstract

The paper describes a new primal relaxation (PR) for computing bounds on nonlinear integer programming (NLIP) problems. It is a natural extension to NLIP problems of the geometric interpretation of Lagrangean relaxation presented by Geoffrion (1974) for linear problems, and it is based on the same assumption that some constraints are complicating and are treated separately from the others. In the nonlinear case, however, this relaxation is not equivalent any more to Lagrangean relaxation, and it does not use Lagrangean multipliers. It consists in replacing the non-complicating constraint set by the convex hull of its integer points. It was introduced in Guignard [10], and described briefly in Guignard [11] for the case of linear constraints. Contrary to Outer Approximation ([18],[7]), it does not construct a superset of the continuous constraint set, but rather a subset of that set. After writing the complicating constraints as equality constraints, the relaxed problem can be shown to be asymptotically equivalent to a penalized nonlinear continuous problem as the penalty factor goes to infinity. Its constraint set is defined only implicitly, but is known to be a polytope. When the non-complicating constraints are linear, the penalized problem can be solved efficiently by using a linearization method. At each so-called major iteration until convergence has been achieved, the

[1] This research was partially supported by NSF Grants DDM-9014901, DMI-9900183 and DMI-0400155.

[2] guignard@wharton.upenn.edu, or guignard_monique@yahoo.fr

penalty coefficient is adjusted upward, and the corresponding penalized problem can be solved iteratively by simplicial decomposition, alternating between a priori much simpler linear integer programming problems and small continuous nonlinear problems over a simplex. Improved solution methods based on augmented Lagrangeans for the linear constraint case have been studied in [6], and successful implementations have been reported in [1], [2], [3] and [4][3].

The relaxation itself must be designed so as to yield linear integer programming problems that are relatively easy to solve. As in the linear MIP case, these subproblems yield Lagrangean-like integer solutions that are often only mildly infeasible in the complicating constraints, and can be used as starting points for Lagrangean heuristics.

We also describe a primal decomposition (PD), similar in spirit to Lagrangean decomposition in the linear case [12], for problems with several structured subsets of constraints.

To illustrate the concepts, and show that, like in Lagrangean relaxation for linear MIP problems, the PR bound can be anywhere between the continuous bound and the integer optimum, we solve several small examples explicitly, for both PR and PD, by a simplified version of Simplicial Decomposition.

Maybe the most interesting aspect of this primal relaxation is the following very promising special case. When one keeps all constraints as non-complicating, and uses the entire convex hull of all integer feasible solutions of the problem, one obtains the convex hull relaxation, or CHR, (see Albornoz 1998, and Ahlatcioglu and Guignard 2007, 2010). In this case there is no need for a penalization, and the solution of the relaxed problem requires only one major iteration. This special relaxation shows great promise first as a tool for computing very efficiently strong bounds in the pseudoconvex case, and also as a powerful heuristic, even for nonconvex problems. Computational evidence is presented in (Ahlatcioglu and Guignard, 2010) for several variants of the very difficult QAP and GQAP problems.

[3] several of these papers are available at http://opim.wharton.upenn.edu/~guignard/publications

.

Introduction

Lagrangean relaxation ([14], [15], [9]), has been used for decades as a powerful tool in solving difficult linear integer programming problems. Its main advantages over the continuous relaxation are that (1) it may yield a tighter bound than the continuous relaxation if the Lagrangean subproblem does not have the Integrality Property, and (2) it produces integer, rather than fractional, solutions that are often only mildly infeasible, and therefore can be used as good starting points for Lagrangean heuristics. While one usually solves the Lagrangean dual in the dual space by searching for a best set of Lagrangean multipliers, this is not the only method possible. Michelon and Maculan [17] showed that one can also solve the primal equivalent of the Lagrangean dual by placing the dualized (equality, but it would work as well for inequality) constraints into the objective function with a large penalty coefficient, and then using a linearization method such as Frank and Wolfe to solve the resulting nonlinear problem. A key realization here is an idea that had already been used in particular by Geoffrion [9]: when one maximizes a linear function over the integer points of a polytope, one optimal solution at least is an extreme point of the convex hull of these integer points. Once a nonlinear objective function is linearized, one can therefore equivalently optimize it over the integer points of a polytope, or over the convex hull of these integer points, whichever is easier. In the case of Michelon and Maculan’s approach, then, each iteration of Frank and Wolfe involves solving a linear integer Lagrangean-like subproblem and performing a nonlinear line search.

Consider now the case of an integer programming problem with a nonlinear objective function. It is usually very difficult to obtain strong bounds for such problems. Indeed it is not easy to use standard Lagrangean relaxation in this case, as the Lagrangean subproblem is still a nonlinear integer problem, and a priori not easier to solve than the original one. We introduced in [10], and briefly described in [11] a new relaxation, which is primal in nature, and can be used with nonlinear objective functions and linear constraints. We extend it here to problems with general constraint sets. It coincides with the standard Lagrangean relaxation in the linear case, but it is new for the nonlinear case.

We assume that part of the constraints have been identified as "complicating." The primal relaxation (PR) consists in replacing the non-complicating constraint set by the convex hull of its integer points. In the definition, no assumption is made concerning the convexity of the functions, nor the nature of the non-complicating constraints. Assumptions will have to be made, however, when considering the algorithms chosen to solve problem (PR), to guarantee that (1) they converge to global minima and that (2) the final value obtained is indeed a valid lower bound on the optimum of the MINLP problem.

Problem (PR) is computationally feasible for instance when the non-complicating constraints are linear: it can indeed be solved by penalizing the complicating constraints, written as equations, in the objective function, and then using a linearization method, extending the idea of Michelon and Maculan [17] to the nonlinear case. As the penalty factor goes to infinity, the penalized problem is asymptotically equivalent to (PR). A major iteration consists of increasing the penalty coefficient if the algorithm has not converged yet, i.e., if a satisfactory feasibility is not achieved, and solving the new resulting penalized problem. This means solving alternatingly a linear integer programming problem over the non-complicating constraints, and performing either a simple line search if using Frank and Wolfe [8], or a search over a simplex, in the case of simplicial decomposition ([19],[13]). There are no more *nonlinear* integer subproblems to solve. We show that the bound obtained in this manner is at least as good as the continuous relaxation bound, and may be substantially stronger. It is also possible to define a primal decomposition that splits the constraint set into several subsets like in Lagrangean decomposition [12] (also called variable splitting [16]).

In the case of linear constraints, this primal relaxation is very attractive, as its implementation requires solving integer subproblems that are *linear* and for which one can select a good (or several for primal decomposition) structured subset(s) of constraints, exactly as in Lagrangean relaxation for *linear* integer programming problems. Finally there are better choices than a penalty method for solving the relaxation, and Contesse and Guignard [6] propose instead to use a (Proximal) Augmented Lagrangean (PAL) scheme, for its improved convergence for a finite value of the penalty factor, and better conditioning properties.

This paper concentrates on the concepts and properties of the primal relaxation, and leaves algorithmic implementation issues to other publications. In section 1, we review the approach of Michelon and Maculan to solve Lagrangean problems in the linear case. In section 2, we introduce the primal relaxation, and primal decomposition in section 3. In section 4, we apply a simplified algorithm to a small numerical example. Section 5 presents some conclusions, and some thoughts about future research.

Notation

For an optimization problem (P), FS(P) denotes the feasible set, V(P) the optimal value and OS(P) the optimal set of (P). If (P) is a (mixed-)integer programming problem, CR(P) (or (CR) if it is not ambiguous) denotes the continuous relaxation of (P). If K is a set in $\Re^n$, Co(K) denotes the convex hull of K. If x is a vector of $\Re^n$, $|x|$ denotes a norm of x, and x^+ means max $\{0,x\}$.

1. Primal Equivalent of Lagrangean Relaxation for Linear Integer Problems

We shall first recall Michelon and Maculan's approach [17] for solving Lagrangean duals in the linear integer problem case. Consider a *linear* integer programming problem

(LIP) $\quad \text{Min}_x \{ fx \mid Ax=b, Cx\leq d, x\in X\}$

where X specifies in particular the integrality requirements on x, and a Lagrangean relaxation of (LIP) :

LR(u) $\quad \text{Min}_x \{fx+u(Ax-b) \mid Cx\leq d, x\in X\}$

with the corresponding Lagrangean dual

(LR) $\quad \text{Max}_u \text{ Min}_x \{fx+u(Ax-b) \mid Cx\leq d, x\in X\}$

and its primal equivalent problem (Geoffrion [9])

(PLR) $\quad \text{Min}_x \{fx \mid Ax=b, x\in \text{Co}\{x|Cx\leq d, x\in X\}\}$.

As ρ approaches infinity, (PLR) becomes asymptotically equivalent to the penalized problem

(PP) $\quad \text{Min}_x \{\varphi (x) = fx + (½) \rho \, |Ax-b|^2 \mid x\in \text{Co}\{ x|Cx\leq d, x \in X\}\}$.

Notice that $\varphi (x)$ is a convex function. (PP) can be solved by a linearization method such as the method of Frank and Wolfe or, even better, simplicial decomposition. For simplicity, let us describe the approach using Frank and Wolfe. At iteration k, one has a current iterate $x(k)$ in whose vicinity one creates a linearization of the function $\varphi(x)$:

$$\psi_k . x = \varphi [x(k)] + \nabla\varphi [x(k)].[x-x(k)].$$

One solves the linearized problem

(LPP$_k$) $\quad \text{Min}_x \{\psi_k x \mid x\in \text{Co}\{ x|Cx\leq d, x\in X\}\}$

or equivalently, because the objective function is linear,

(LPP$_k$) $\quad \text{Min}_x \{\psi_k x \mid Cx\leq d, x\in X\}$.

Let $y(k)$ be its optimal solution. Then $x(k+1)$ is obtained by minimizing $\varphi(x)$ on the half-line $x=x(k)+\lambda [y(k)-x(k)]$, $\lambda\geq 0$. The process is repeated until either a convergence criterion is satisfied or a limit on the iteration number is reached.

The idea is attractive because

(1) while one cannot eliminate the convex hull in (PP), one can do so after the linearization, in other words, the convex hull computation is not necessary any more after (PP) has been transformed into a sequence of problems (LPP$_k$). Notice

too that (LPP_k) has the same constraint set as LR(u), i.e., it must be solvable if (LR(u)) is.

(2) even in case (LR(u)) decomposes into a family of smaller subproblems, this is usually not the case for (PP). (LPP_k), though, will also decompose, and the primal approach is fully as attractive as the original Lagrangean relaxation.

The slow convergence of Frank and Wolfe's algorithm, however, may make one prefer a faster linearization method, such as simplicial decomposition [19], or restricted simplicial decomposition [13].

2.Primal Relaxation for Nonlinear Integer Programming Problems

Consider now an integer programming problem

(IP) $\quad \text{Min}_x \ \{f(x) \mid g(x)=b, \ x \in Y\}$,

where the nonlinear functions f and g are differentiable, $g(x)=b$ are the complicating constraints, and Y is a bounded set of integer (or mixed-integer) points satisfying some additional restrictions.

We could try to solve (IP) directly by noticing that as the positive scalar ρ goes to infinity, (IP) becomes asymptotically equivalent to

(P1) $\quad \text{Min}_x \ \{f(x) + (½) \rho \mid g(x)-b \mid^2 \ \mid x \in Y\}$.

Unfortunately (P1) is almost always as difficult to solve as (IP). The constraint set of (P1) is not a polygon, and the objective function of (P1) is still nonlinear, so (P1) is still a nonlinear integer problem. We could consider the other problem

(P2) $\quad \text{Min}_x \ \{ f(x) + (½) \rho \mid g(x)-b \mid^2 \ \mid x \in \text{Co}(Y)\}$.

which is a relaxation of (P1), but in general is not equivalent to (P1).

Since in any case neither (P1) nor (IP) is easy to solve, we will build a *new* primal relaxation of (IP) which will use (P2) as a subproblem for fixed ρ. We will then show in detail when and how the relaxed problem can actually be solved. We will more specifically show that if an integer programming problem of the form

(LIP) $\quad \text{Min}_x \ \{gx \mid x \in Y\}$

with $Y = \{x \mid Cx \le d,\ x \in X\}$ where X contains the integrality restrictions on at least some of the components of x, can be solved relatively easily, then we can design and solve a relaxation approach similar to the one described above for the integer *linear* case.

2.1.Definition of the Primal Relaxation.

We now formally define the new relaxation in the broadest possible way.

Definition 1.

We define the Primal Relaxation problem of problem

(IP) $\quad \text{Min}_x \ \{f(x) \,/\, g(x)=b,\ x \in Y\}$,

as the problem

(PR) $\quad \text{Min}_x \ \{f(x) \mid g(x)=b,\ x \in \text{Co}(Y)\}$.

(PR) is indeed a relaxation of (IP):

$\{x \mid g(x)=b,\ x \in \text{Co}(Y)\} \ \supseteq \ \{x \mid g(x)=b,\ x \in Y\}$.

If $Y = \{x \mid Cx \le d,\ x \in X\}$, then the so-called *continuous* relaxation of (IP),

(CR) $\text{Min}_x \ \{ f(x) \mid g(x)=b,\ Cx \le d,\ x \in \text{Co}(X)\}$,

is itself a relaxation of (PR), since in that case

$\{ x \in \text{Co}(X) \mid g(x)=b,\ Cx \le d \} \ \supseteq \ \{ x \in \text{Co}\{x \in X \mid Cx \le d \} \mid g(x)=b\}$.

(PR) cannot in general be solved directly, even in that case, since $\text{Co}\{x \in X \mid Cx \le d\}$ is usually not known explicitly, and even if it were, (PR) would probably be of the same level of difficulty as (IP) because it is an integer programming problem with a nonlinear objective function.

Roughly speaking, though, for ρ large enough, (PR) is asymptotically equivalent to the penalized problem

(PP) $\quad \text{Min}_x \{\varphi(x) = f(x)+\rho \,\| g(x)-b \|^2 \mid \ x \in \text{Co}\{x \in X \mid Cx \le d\}\}$,

where $\varphi(x)$ is a nonlinear function. (PP) can be solved by a linearization method such as

Frank and Wolfe. This method unfortunately is known to converge rather slowly. Another linearization method, called Simplicial Decomposition, should be used instead, and the overall convergence would be improved further if one used an augmented Lagrangean method instead of the penalization method described above. Such an approach was studied in Contesse and Guignard [6] and successful implementations were described in [1] and [2], and in more recent papers by Ahn, Contesse and Guignard [2],[3] and Ahlatcioglu and Guignard [4].

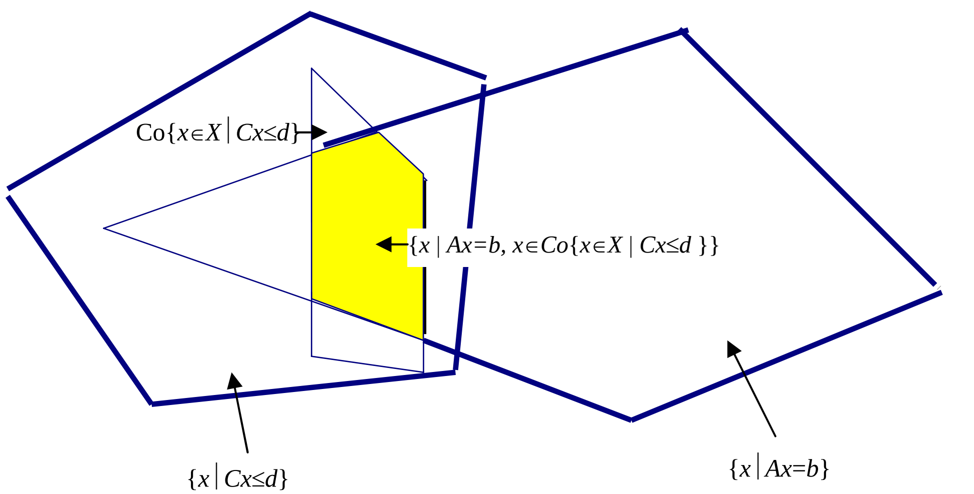

Figure 1

2.2 Properties of the Primal Relaxation.

We concentrate in this paper on the characteristics of the primal relaxation and not on algorithmic details or on obtaining an efficient implementation. This is why we choose to describe the approach based on a penalization method and on Frank and Wolfe's linearization method, to illustrate the relaxation and the general idea of its solution, rather than a more efficient implementation with augmented Lagrangeans and simplicial decomposition.

At iteration k, one has a current iterate $x(k)$ in whose vicinity one creates a linearization of the function $\varphi(x)$:

$$\psi_k \,.\, x = \varphi[x(k)] + \nabla\varphi\, [x(k)][x-x(k)].$$

One solves the linearized problem

(LP$_k$) $\quad \text{Min}_x \ \{\psi_k.x \mid x \in \text{Co}\{x \mid Cx \le d,\ x \in X\}\}$

or equivalently, because the objective function is linear,

(LP$_k$) $\quad \text{Min}_x \ \{\psi_k.x \mid Cx \le d,\ x \in X\}$.

Let $y(k)$ be its optimal solution. Then $x(k+1)$, the new linearization point, is obtained by minimizing $\varphi(x)$ on the half-line $x=x(k) + \lambda\, [y(k)-x(k)]$, $\lambda \ge 0$. The process is repeated until either a convergence criterion is satisfied or a limit on the iteration number is reached.

This process has roughly the same advantages as in the linear case:

(1) while one cannot eliminate the convex hull in (PP), one can do it for (LP$_k$). We made the assumption earlier that a problem with a structure such as (LP$_k$) is solvable.

(2) in case the constraints of (IP) decompose into a family of smaller subproblems if the constraints $g(x)=b$ are removed, this property allows (LP$_k$) to decompose as well, even though this is not the case for (PP). The linearization of the objective function thus allows one to solve the problem via a sequence of decomposable linear integer programs and line searches. This is very attractive if it reduces substantially the size of the integer problems one has to solve. It is usually much easier to solve ten problems with thirty 0-1 variables each than a single problem with three hundred 0-1 variables.

One can also handle the case of inequality constraints of the form $h_j(x) \le d_j$ with some minor modification, the best known being by adding the square of a new continuous variable $x'j$ to each $h_j(x)$ before constructing the penalty function (see for instance [5], p. 318). One could also compute the penalty function slightly differently as $\varphi(x) = f(x) + \rho \{ \| g(x)-b \|^2 + \sum_j [h_i(x)-d_j]^{+}]^2 \}$.

2.3. Convergence of the algorithm.

The method of Frank and Wolfe, also called conditional gradient method, when properly implemented (see [5], p. 222) guarantees that every limit point is stationary. A sufficient condition for this stationary point to be a minimum for the penalized problem for given ρ, is that the function $\varphi(x)$ be convex. If using simplicial decomposition, pseudoconvexity is sufficient.

These conditions are clearly satisfied when g is a linear function of x.

2.4. A special case.

As in standard Lagrangean relaxation, the "extreme" case of subproblems with the Integrality Property (see [9] for the linear case of Lagrangean relaxation) will not yield any improvement over the continuous nonlinear programming relaxation.

Definition 2

Problem (PR) $\text{Min}_x \{f(x) \,/\, Ax=b,\ x \in Co\{x \,/\, Cx \le d,\ x \in X\}\}$ *is said to have the* ***Integrality Property*** *if the polyhedron* $P = Co\,\{x \,/\, Cx \le d,\ x \in X\}$ *coincides with the set* $\{x \,/\, Cx \le d,\ x \in Co(X)\}$.

In a somewhat simplified way, one can say that (PR) has the integrality property if the extreme points of the polytope $Cx \le d$ are in X.

Proposition 1.

If the Primal Relaxation problem

(PR) $\quad \text{Min}_x \{f(x) \,/\, Ax=b,\ x \in Co\{x \,/\, Cx \le d,\ x \in X\}\}$

has the integrality property, then V(PR) = V(CR).

In that case, (PR) is no improvement over the continuous relaxation. Yet, as in the linear case, one might still want to use it if solving (CR) requires using an exponential number of constraints. One such example is the TSP, for which Held and Karp [14], [15], showed that Lagrangean relaxation was nevertheless an attractive option.

2.5. An Example.

The following example illustrates that the bound V(PR) can be anywhere between V(IP) and V(CR), depending on the problem parameters, as happens for standard Lagrangean relaxation bounds.

Consider the following very simple 2-dimensional problem. One wants to minimize the distance to the point A(1,1) subject to the constraints $x_1 = 2x_2$ and $ax_1 + bx_2 \le c$, where x_1 and x_2 are (0-1) variables. We will write z(M) to denote the value of the objective function at the point M(x_1,x_2). The problems under consideration are:

$\text{(IP)}\ \text{Min}\ (1-x_1)^2 + (1-x_2)^2$	$\text{(PR)}\ \text{Min}\ (1-x_1)^2 + (1-x_2)^2$	$\text{(CR)}\ \text{Min}\ (1-x_1)^2 + (1-x_2)^2$
s.t. $x_1 - 2x_2 = 0$	s.t. $x_1 - 2x_2 = 0$	s.t. $x_1 - 2x_2 = 0$
$ax_1 + bx_2 \le c$	$x \in \text{Co}\{x \mid ax_1 + bx_2 \le c,$	$ax_1 + bx_2 \le c$
$x_1, x_2 \in \{0,1\}$	$x_1, x_2 \in \{0,1\}\}$	$x_1, x_2 \in [0,1]$

We will place $x_1 = 2x_2$ in the objective function as a penalty term. We will consider several cases. All problems are represented on Figure 2.

Case 1. *a*=10, *b*=1, *c*=9.

Then Co $\{x \mid 10x_1 + x_2 \le 9, x_1, x_2 \in\{0, 1\}\}$ is the line segment OD, and $\{x \mid Ax=b, x\in\text{Co}\{x \mid Cx\le d, x\in X\}\}$ is the origin O. Thus V(PR) = V(IP) = z(O) = $(1-0)^2 + (1-0)^2$ = 2, while V(CR) is reached at point P(18/21, 9/21) and is equal to z(P) = 0.35.

Case 2. *a*=2, *b*=1, *c*=2.

Then Co $\{x \mid 2x_1 + x_2 \le 2, x_1, x_2 \in\{0,1\}\}$ is the triangle ODF, and $\{x \mid Ax=b, x\in\text{Co}\{x \mid Cx\le d, x\in X\}\}$ is the line segment OS. Thus V(IP) = $(1-0)^2 + (1-0)^2$ = 2, while V(PR) = z(S) = $(1-2/3)^2+(1-1/3)^2$ and V(CR) is reached at the point Q (2/5, 4/5) and is equal to z(Q) = $(1-4/5)^2 + (1-2/5)^2$ =0.4.

Case 3. ***a*=1, *b*=1, *c*=1.**

Then Co $\{x \mid x_1 + x_2 \leq 1, x_1, x_2 \in \{0, 1\}\}$ is the triangle ODF, and $\{x \mid Ax=b, x \in \text{Co}\{x \mid Cx \leq d, x \in X\}\}$ is the line segment OS. Thus V(IP) = $(1\text{-}0)^2 + (1\text{-}0)^2 = 2$, while V(PR) = V(CR) = z(S) = $(1\text{-}2/3)^2 + (1\text{-}1/3)^2 = 0.55$.

It can be seen on the above examples that the value of V(PR) can be arbitrarily close to either the integer optimum or the continuous optimum. This is rather similar to what happens for Lagrangean relaxation bounds in linear integer programming.

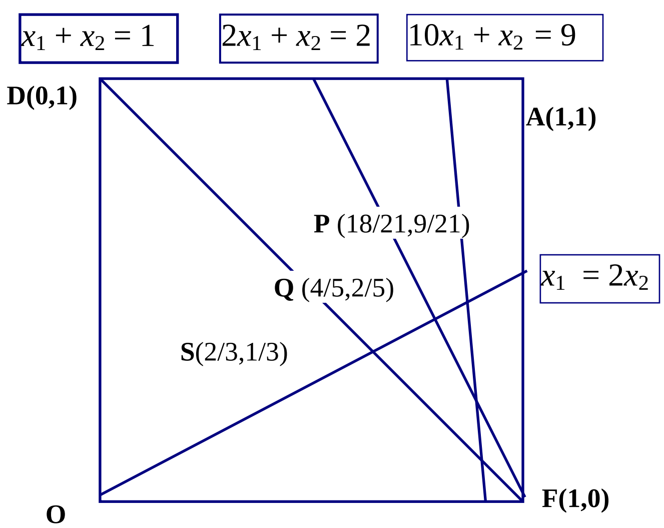

Figure 2

3. Primal Decomposition for Nonlinear Integer Programming Problems

We will now show that one can similarly define a primal decomposition, similar in spirit to that described for instance in Guignard and Kim [12].

Consider an integer programming problem with a nonlinear objective function and linear constraints in which one has replaced x by y in some of the constraints, after adding the copy constraint $x=y$:

(IP) $\quad \text{Min}_x \ \{f(x) \mid Ay \le b, y \in X, Cx \le d, x \in X, x=y\}.$

We will show that if *linear* integer programming problems of the form

(LIP_x) $\quad \text{Min}_x \ \{gx \mid Cx \le d, x \in X\}$

and

(LIP_y) $\quad \text{Min}_y \ \{hy \mid Ay \le b, y \in X\}$

can be solved relatively easily, then we can design a primal decomposition approach similar to the primal relaxation approach described above. The linearization procedure allows us to replace *linear programs* with *implicitly* defined polyhedral constraint sets by *linear integer programs* with well structured discrete constraint sets. If we applied the decomposition idea directly to (IP), we would obtain a nonlinear integer program for which the Frank and Wolfe algorithm would be meaningless. This is why we consider a convex hull relaxation of the constraint set *first* before introducing a penalty function.

3.1. Definition of Primal Decomposition.

Definition 3

*We define the **primal decomposition** of problem (IP) to be problem*

(PD) $\quad \text{Min}_x \ \{f(x) \ / \ x \in Co\{x \ / \ Ax \le b, x \in X\} \cap Co\{x \ / \ Cx \le d, x \in X\}\}.$

Problem (PD) is indeed a relaxation of (IP), since

$\text{Co}\{x \mid Ax \le b, x \in X\} \cap \text{Co}\{x \mid Cx \le d, x \in X\} \supseteq \{x \mid Ax \le b, Cx \le d, x \in X\}.$

At the same time, problem

(PR) $\quad \text{Min}_x \ \{f(x) \mid Ax \le b, x \in Co\{x \mid Cx \le d, x \in X\}\}.$

is a relaxation of (PD), since

$\{x \mid Ax \le b, x \in Co\{x \mid Cx \le d, x \in X\} \supseteq \text{Co}\{x \mid Ax \le b, x \in X\} \cap \text{Co}\{x \mid Cx \le d, x \in X\}$

and finally problem

(CR) $\quad\quad \text{Min}_x \ \{ f(x) \mid Ax \leq b,\ Cx \leq d,\ x \in \text{Co}(X)\}$,

the so-called *continuous* relaxation of (IP), is itself a relaxation of (PD), since

$\{x \mid Ax \leq b,\ Cx \leq d,\ x \in \text{Co}(X)\} \supseteq \text{Co}\{x \mid Ax \leq b,\ x \in X\} \cap \text{Co}\{x \mid Cx \leq d,\ x \in X\}$

(PD) cannot in general be solved directly, since on the one hand $\text{Co}\{x \mid Cx \leq d,\ x \in X\}$ and $\text{Co}\{x \mid Ax \leq b,\ x \in X\}$ are usually not known explicitly, and on the other hand, even if they were, (PD) would probably be of the same level of difficulty as (IP).

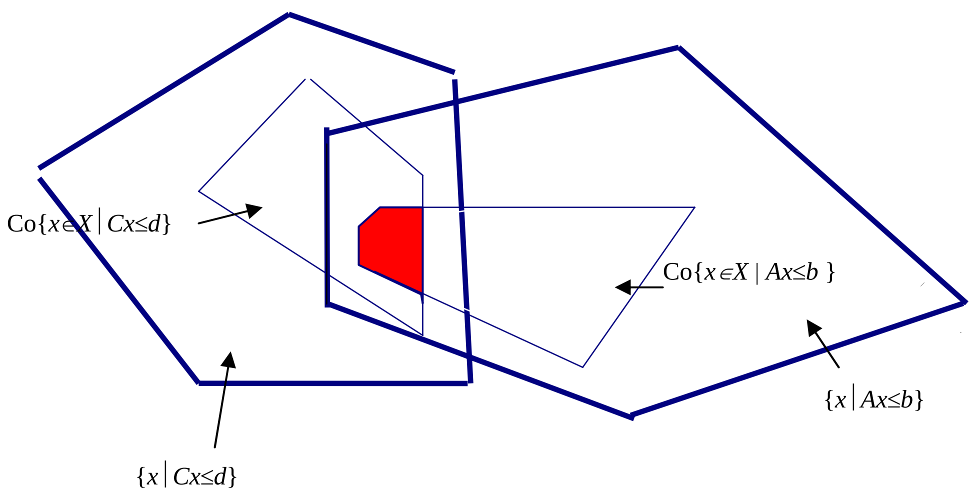

Figure 3

Again, roughly speaking, for ρ large enough, (PD) is asymptotically equivalent to the penalized problem

(PP) $\quad \text{Min}_x \ \{\varphi(x,y) = f(x) + \rho \,|x-y|^2 \ \mid y \in \text{Co}\{y \mid Ay \leq b,\ y \in X\},\ x \in \text{Co}\{x \mid Cx \leq d,\ x \in X\}\}$.

(PP) can be solved by a linearization method. We describe here the approach based on Frank and Wolfe. At iteration k, one has a current iterate $(x(k),y(k))$ in whose vicinity one creates a linearization of the function $\varphi(x,y)$:

$$\psi_k\,.\,(x,y) = \varphi[x(k),y(k)] + \nabla\varphi[x(k),y(k)].[x-x(k),y-y(k)].$$

One solves the linearized problem

(LP$_k$) $\quad \text{Min}_{x,y}\,\{\psi_k(x,y) \mid x \in \text{Co}\{x \mid Cx \le d,\ x \in X\},\ y \in \text{Co}\{y \mid Ay \le b,\ y \in X\}\}$

which separates as follows, because the objective function is linear:

(LP$_k$) $\quad \text{Min}_x\ \{\psi^{x}{}_{k}\,.\,x \mid Cx \le d,\ x \in X\} + \text{Min}_y\ \{\psi^{y}{}_{k}\,.\,y \mid Ay \le b,\ y \in X\}.$

and again the relaxed problem separates into two linear subproblems of a type which we assumed we can solve. Decomposition in this case is achieved at each iteration of Frank and Wolfe where LP's with implicit constraints are replaced by IP's with a good structure.

3.2. An Example

Consider again the very simple example considered earlier. One wants to minimize the distance to the point A (1,1) subject to the constraints $x_1 = 2x_2$ and $ax_1 + bx_2 \le c$, where x_1 and x_2 are (0-1) variables. The problems under consideration are:

(IP)	(PD)	(CR)
$\text{Min } (1-x_1)^2 + (1-x_2)^2$	$\text{Min } (1-x_1)^2 + (1-x_2)^2$	$\text{Min } (1-x_1)^2 + (1-x_2)^2$
s.t. $x_1 - 2x_2 = 0$	s.t. $x \in \text{Co}\{x \mid x_1 - 2x_2 = 0$	s.t. $x_1 - 2x_2 = 0$
$ax_1 + bx_2 \le c$	$x_1,\ x_2 \in \{0,1\}\}$	$ax_1 + bx_2 \le c$
$x_1,\ x_2 \in \{0,1\}$	$x \in \text{Co}\{x \mid ax_1 + bx_2 \le c,$	$x_1,\ x_2 \in [0,1]$
	$x_1,\ x_2 \in \{0,1\}\}$	

We will call z(M) the value of the objective function at M(x_1,x_2).

We will reformulate (PD), creating a copy y of the variable x and adding the constraint $x = y$.

We will place $x = y$ in the objective function as a penalty term. We will consider several cases :

1. a=10, b=1, c=9. Then Co $\{x \mid 10x_1 + x_2 \le 9,\ x_1, x_2 \in \{0, 1\}\}$ is OD, and

Co{x | $Ax \le b$, $x \in X$} is O. Thus V(PR) = V(IP) = z(O) = $(1\text{-}0)^2 + (1\text{-}0)^2$ = 2, while V(CR) is reached at P(18/21, 9/21) and is equal to z(P) = 0.35.

0.35	2
V(CR)	V(PR)=V(IP)=V(PD)

2. a=2, b=1, c=2. Then Co {x | $2x_1 + x_2 \le 2$, $x_1, x_2 \in \{0,1\}$} is ODF, and Co{x | $Ax \le b$, $x \in X$} is O. Thus V(IP) = $(1\text{-}0)^2 + (1\text{-}0)^2$ = 2 = z(O) = V(PD), while V(PR) = z(S) = $(1\text{-}2/3)^2 + (1\text{-}1/3)^2$ and V(CR) is reached at Q (2/5, 4/5) and is equal to z(Q) = $(1\text{-}4/5)^2 + (1\text{-}2/5)^2$ =0.4.

0.4	0.55	2
V(CR)	V(PR)	V(IP)=V(PD)

3. a=1, b=1, c=1. Then Co {x | $x_1 + x_2 \le 1$, $x_1, x_2 \in \{0, 1\}$} is ODF, and

Co{x | $Ax \le b$, $x \in X$} is O. Thus V(IP) = $(1\text{-}0)^2 + (1\text{-}0)^2$ = 2=V(PD), while V(PR) = V(CR) = z(S) = $(1\text{-}2/3)^2 + (1\text{-}1/3)^2$ = 0.55.

. 55	2
V(CR)=V(PR)	V(IP)=V(PD)

It can be seen on the above examples that the value of V(PD) can be equal to the integer optimum, even when V(PR) is equal to the continuous optimum, V(CR) is always weaker than V(PR) which is itself weaker than V(PD), given than FS(CR) contains FS(PR) which in turn contains FS(PR) which itself contains FS(PD).

4. Bound computation: an example

We will now consider a three dimensional example on which we will demonstrate what bound computation involves. We shall use a slight modification of the algorithm of Frank and Wolfe, in which instead of a one-dimensional line search one performs a 2-dimensional triangular search in the triangle formed by the current linearization point and the last two solutions of linearized subproblems. It is actually almost a form of restricted simplicial decomposition [13].

The problem, represented in figure 4, is as follows:

Min $\{(2-x_2)^2 \mid x_1 - 2x_2 + x_3 = 0,\ 10x_1 + x_2 - x_3 \le 9,\ x_1, x_2, x_3 \in \{0,1\}\}$.

We will use the notation ABC…H to denote the convex hull of the points A, B, C, …, and H in R^3. For instance AB is the line segment AB, ABC the triangle ABC, etc.

In (PR) and (PD), we let $x_1 - 2x_2 + x_3 = 0$ stand for $Ax \le b$, and $10x_1 + x_2 - x_3 \le 9$ for $Cx \le d$.

That is,

(IP) Min $(2-x_2)^2$

s.t. $x_1 - 2x_2 + x_3 = 0$

$10x_1 + x_2 - x_3 \le 9$

$x_1, x_2, x_3 \in \{0,1\}$

(PD) Min $(2-x_2)^2$

s.t. $x \in Co\{x \mid x_1 - 2x_2 + x_3 = 0,$ and $x_1, x_2, x_3 \in \{0,1\}\}$

$x \in Co\{x \mid 10x_1 + x_2 - x_3 \le 9,$ and $x_1, x_2, x_3 \in \{0,1\}\}$

(CR) Min $(2-x_2)^2$

s.t. $x_1 - 2x_2 + x_3 = 0$

$10x_1 + x_2 - x_3 \le 9$

$x_1, x_2, x_3 \in [0,1]$

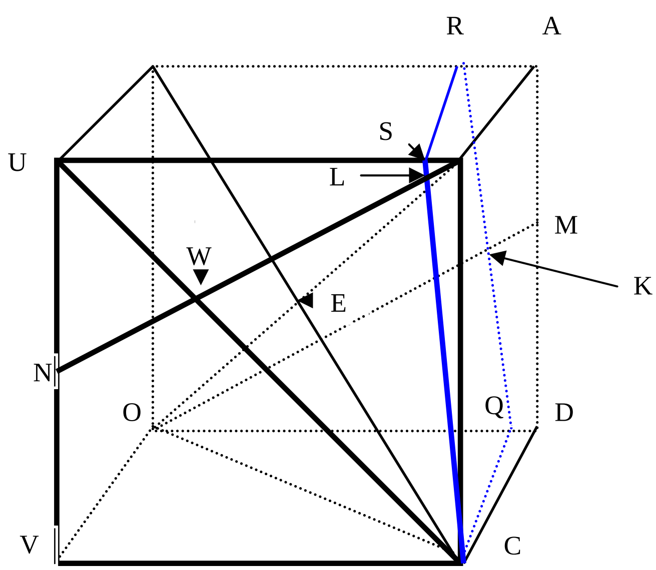

Figure 4

and (PR) Min $(2-x_2)^2$

s.t. $x_1 - 2x_2 + x_3 = 0$

$x \in Co\{x \mid 10x_1 + x_2 - x_3 \le 9,\ x_1, x_2, x_3 \in \{0,1\}\}$.

Then FS(CR) = OKLN, FS(PR) = OEWN, FS(PD) = OE, and FS(IP)=O, and

V(CR) = z (L) = 1.1, V(PR) = z (W) =1.7, V(PD) = z (E) =2.25, V(IP) = z (O) =4.

V(CR)	V(PR)	V(PD)	V(IP)
1.1	1.7	2.25	4

We will show the computation for (PR):

(PR) Min $(2-x_2)^2$

s.t. $x_1 - 2x_2 + x_3 = 0$

$x \in \mathrm{Co}\{x \mid 10x_1 + x_2 - x_3 \le 9,\ x_1, x_2, x_3 \in \{0,1\}\}$.

(PR) is asymptotically equivalent, as ρ goes to infinity, to

Min $\varphi(x) = (2-x_2)^2 + \rho\,(x_1 - 2x_2 + x_3)^2$

s.t. $x \in \mathrm{Co}\,\{x \mid 10x_1 + x_2 - x_3 \le 9,\ x_1, x_2, x_3 \in \{0,1\}\}$.

The linearization of the objective function at $x^{(0)}$ yields the function

$[2\rho\,(x_1 - 2x_2 + x_3),\ -2\,(2-x_2) - 4\rho\,(x_1 - 2x_2 + x_3),\ 2\rho\,(x_1 - 2x_2 + x_3)]_{x=x0}\ [x_1, x_2, x_3]$

The initial point, $x^{(1)} = (0.5, 1, 0)$, is chosen arbitrarily. The slack in the equality constraint at $x(1)$, i.e., the amount of violation in the penalized constraint, is $s(1) = -1.5$. The first linearized problem is

Min $-3\rho\, x_1 + (-2 + 6\rho)\, x_2 - 3\rho\, x_3$

s.t. $10x_1 + x_2 - x_3 \le 9$,

$x_1, x_2, x_3 \in \{0,1\}$.

We choose to take $\rho = 5000$.

Iteration 1

The gradient at $x(1)$ is (-15000, 29998, -15000). The solution of the linearized problem is $y(1)$= (1, 0, 1). Since this is the first iteration, one only does a line search, in the direction $da(1) = y(1) - x(1) = (0.5, -1, 1)$. The line search yields a stepsize of 0.429. The corresponding solution is $x(2)$ =(0.714, 0.571, 0.429). The slack in the equality constraint at $x(2)$ is $s(2)$ = -8.16313E-5. The nonlinear objective function value is 2.041, and the penalty term is 6.66367E-9.

Iteration 2

The current linearization point is $x(2)$ = (0.714, 0.571, 0.429). The gradient at $x(2)$ is (-0.816,1.224, -0.81). The solution of the linearized problem is $y(2)$ = (0, 1, 1). The directions of triangular search are $da(2)$ = $y(2)$ - $x(2)$ =(-0.714, 0.429, 0.571) and $db(2)$ = $y(2)$ - $x(2)$ = (0.286, -0.571, 0.571). The search is over the triangle formed by $x(2)$, $y(1)$ and $y(2)$, with sides $da(2)$ and $db(2)$. The stepsizes are $step_a$ = 0.667 in the direction $da(2)$ and $step_b$ = 0.333 in the direction $db(2)$. The sum of the stepsizes must be less than or equal to 1 if one wants to stay within the triangle. The solution of the search is $x(3)$ = (0.333, 0.667, 1), and the slack in the equality constraint at $x(3)$ is $s(3)$ = 8.88869E-5.
The nonlinear objective function value is 1.778 and the penalty term value is 7.90088E-9.

Iteration 3

The current linearization point $x(3)$ is (0.333, 0.667, 1). The gradient at $x(3)$ is (-0.889, -0.889, -0.889), and the solution $y(3)$ of the linearized problem is (1, 0, 1). The directions of triangular search $da(3)$ and $db(3)$ are respectively (0.667, 0.667, 0) and (-0.333, 0.333, 0). The stepsizes are respectively 0.282 in the direction $da(3)$ and 0.564 in the direction $db(3)$. The solution is $x(4)$ = (0.333, 0667, 1), and the slack in the equality constraint at $x(4)$ is $s(4)$ = -8.88869E-5. The nonlinear objective function value is 1.778 and the penalty value is 7.900883E-9. Since $x(3)$ and $x(4)$ are identical, the algorithm stops. Since the penalty value does not affect the objective function value any more, we can consider that problem (PR) is solved, with V(PR) = 1.778.

Conclusion

Even though in case of a nonlinear objective function the primal relaxation proposed above may not always be equivalent to a Lagrangean relaxation, it will work in a manner quite similar to Lagrangean relaxation. The subproblems solved in the linearization steps have the same constraints one would have chosen in the Lagrangean relaxation. If the constraints are separable, so will be the subproblems. The relaxation proposed here is always at least as good as the continuous relaxation. and possibly much stronger as demonstrated by some of the examples presented.

Lagrangean relaxation has been a favorite tool of many IP researchers for LIP, even though there is no guarantee that the bounds obtained strongly dominate continuous bounds for a specific instance. This depends on both problem structure and data instance. Except for the Integrality Property, there is no "theoretical result" in Geoffrion's paper related to the strength of the LR bound, there could not have been any. In the same spirit, there can be no "theoretical result" related to the strength of the PR bound. The purpose of the small examples was to show that like for LR, the bound can be as bad or as good as possible (equal to either the continuous bound or the integer optimum).

While PR is equivalent to Lagrangean relaxation (LR) for <u>linear</u> integer programs (LIPs), in the nonlinear case, PR is a <u>new relaxation</u>, different from LR in its very definition. The main advantage over LR is algorithmic: if, with a <u>linear</u> objective function, some subproblem of the original MINLP problem is much easier to solve than the MINLP, then the corresponding PR bound can be computed (relatively) easily, while in LR, the Lagrangean subproblems, being in general nonlinear, are still a priori difficult. This is most likely while LR is used so little for MINLPs.

The same PR idea can be applied to yield relaxations akin, but not necessarily equivalent, to Lagrangean decompositions or substitutions.

References

[1] Ahn S., "On solving some optimization problems in stochastic and integer programming with applications in finance and banking," Ph.D. dissertation, University

of Pennsylvania, OPIM Department, June 1997.

[2] Ahn S., Contesse L. and Guignard M., "A proximal augmented Lagrangean relaxation for linear and nonlinear integer programming: application to nonlinear capacitated facility location," University of Pennsylvania, Department of OPIM Research Report, 1996.

[3] Ahn S., Contesse L. and Guignard M., "A primal relaxation for nonlinear integer programming solved by the method of multipliers, *Part I*: Theory and algorithm," University of Pennsylvania, Department of OPIM Research Report, 2007. "A primal relaxation for nonlinear integer programming solved by the method of multipliers, *Part II:* Application to nonlinear capacitated facility location," University of Pennsylvania, Department of OPIM Research Report, 2006.

[4] Ahlatcioglu A. and Guignard M., "Application of Primal Relaxation For Nonlinear Integer Programs Solved By The Method Of Multipliers To Generalized Quadratic Assignment Problems," OPIM Dept. Report, University of Pennsylvania, Sept. 2007

[5] Bertsekas, D. "Nonlinear Programming," Athena Press, 2d edition, 2d printing, 2003.

[6] Contesse L. and Guignard M., "A proximal augmented Lagrangean relaxation for linear and nonlinear integer programming," University of Pennsylvania, Department of OPIM Research Report 95-03-06, March 1995.

[7] Fletcher, R. and Leyffer, S., "Solving mixed integer nonlinear programs by outer approximation," *Mathematical Programming* 66, 327-349, 1994.

[8] Frank M. and Wolfe P., "An algorithm for quadra1ic programming," *Naval Research Quarterly*, 3(1,2), 95-109, 1956.

[9] Geoffrion A., "Lagrangean relaxation and its uses in integer programming," *Mathematical Programming Study* 2 , 82-114, 1974.

[10] Guignard M., "Primal relaxations for integer programming," University of Pennsylvania, Department of Operations and Information Management Report 94-02-01, also presented as an invited tutorial at CLAIO, Santiago, 1994.

[11] Guignard M., "Lagrangean Relaxation," *TOP*, 11(2), 151-228, Dec. 2003.

[12] Guignard M. and Kim S., "Lagrangean decomposition: A model yielding stronger Lagrangean bounds," *Mathematical Programming*, 39, 215-228, 1987.

[13] Hearn, D.W., Lawphongpanich S. and Ventura J.A., "Restricted simplicial decomposition: computation and extensions, " *Mathematical Programming Study* 31, 99-118, 1987.

[14] Held M. and Karp R. M., "The Traveling-Salesman Problem and Minimum Spanning Trees," *Operations Research* 18, No. 6, 1138-1162, 1970.

[15] Held M. and Karp R. M., The Traveling-Salesman Problem and Minimum Spanning Trees: Part II," *Mathematical Programming* 1, 6–25, 1971.

[16] Jornsten, K. And Nasberg, M., "A new Lagrangean relaxation approach to the generalized assignment problem," *European J of Operational Research* 27, 313-323, 1986.

[17] Michelon P. and Maculan N., "Solving the Lagrangean dual problem in ınteger programming," Departement d'Informatique et de Recherche Operationnelle, Universite de Montreal, Publication 822, May 1992.

[18] Viswanathan, J. and Grossmann, I.E., " A combined penalty function and outer approximation method for MINLP optimization," *Comp Chem Eng* 14(7) 769-782, 1990.

[19] Von Hohenbalken B., "Simplicial decomposition in nonlinear programming algorithms," *Mathematical Programming*, 13, 49-68, 1977.