QUADRATIC PROGRAMMING: QUANTITATIVE ANALYSIS AND POLYNOMIAL RUNNING TIME ALGORITHMS

By

JADRANKA SKORIN-KAPOV

B. Sc. The University of Zagreb, Yugoslavia, 1977

M. Sc. The University of Zagreb, Yugoslavia, 1983

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

(Department of Commerce and Business Administration)

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

August 1987

4 6

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Comerce and Business Administration

The University of British Columbia
1956 Main Mall
Vancouver, Canada
V6T 1Y3

Date October 9, 1987

Abstract

Many problems in economics, statistics and numerical analysis can be formulated as the optimization of a convex quadratic function over a polyhedral set. A polynomial algorithm for solving convex quadratic programming problems was first developed by Kozlov at al. (1979). Tardos (1986) was the first to present a polynomial algorithm for solving linear programming problems in which the number of arithmetic steps depends only on the size of the numbers in the constraint matrix and is independent of the size of the numbers in the right hand side and the cost coefficients. In the first part of the thesis we extended Tardos' results to strictly convex quadratic programming of the form $\max \{ c^T x - \frac{1}{2} x^T D x \; : \; Ax \leq b \,, x \geq 0\}$ with D being symmetric positive definite matrix. In our algorithm the number of arithmetic steps is independent of c and b but depends on the size of the entries of the matrices A and D.

Another part of the thesis is concerned with proximity and sensitivity of integer and mixed-integer quadratic programs. We have shown that for any optimal solution $\bar{z}$ for a given separable quadratic integer programming problem there exist an optimal solution $\bar{x}$ for its continuous relaxation such that $\|\bar{z} - \bar{x}\|_\infty \leq n\Delta(A)$ where n is the number of variables and $\Delta(A)$ is the largest absolute subdeterminant of the integer constraint matrix A. We have further shown that for any feasible solution z, which is not optimal for the separable quadratic integer programming problem, there exists a feasible solution $\bar{z}$ having greater objective function value and with $\|z - \bar{z}\|_\infty \leq n\Delta(A)$. Under some additional assumptions the distance between a pair of optimal solutions to the integer quadratic programming problem with right hand side vectors b and b', respectively, depends linearly on $\|b - b'\|_1$. The extension to the mixed-integer nonseparable quadratic

case is also given.

Some sensitivity analysis results for nonlinear integer programming problems are given. We assume that the nonlinear $0-1$ problem was solved by implicit enumeration and that some small changes have been made in the right hand side or objective function coefficients. We then established what additional information to keep in the implicit enumeration tree, when solving the original problem, in order to provide us with bounds on the optimal value of a perturbed problem. Also, suppose that after solving the original problem to optimality the problem was enlarged by introducing a new $0-1$ variable, say x_{n+1}. We determined a lower bound on the added objective function coefficients for which the new integer variable x_{n+1} remains at zero level in the optimal solution for the modified integer nonlinear program. We discuss the extensions to the mixed-integer case as well as to the case when integer variables are not resticted to be 0 or 1. The computational results for an example with quadratic objective function, linear constraints and $0-1$ variables are provided.

Finally, we have shown how to replace the objective function of a quadratic program with $0-1$ variables (by an integer objective function whose size is polynomially bounded by the number of variables) without changing the set of optimal solutions. This was done by making use of the algorithm given by Frank and Tardos (1985) which in turn uses the simultaneous approximation algorithm of Lenstra, Lenstra and Lovász (1982).

Table of Contents

List of Figures

List of Tables

Acknowledgement

I would like to express my sincere appreciation to the University of British Columbia and especially to the Faculty of Commerce and Business Administration for providing me with a financial support through : Dean Earle D. McPhee Memorial Fellowship in Commerce and Business Administration, University of British Columbia Summer Graduate Fellowship, University of British Columbia Graduate Fellowship and Graduate Travel Grant. My sincere thanks go to the Division of Management Science for giving me the opportunity to be one of their Ph.D. Students.

I am deeply indebted to Professor Frieda Granot, my supervisor, for her guidance, many valuable comments and suggestions and for her friendliness and financial support. Furthermore, I would like to thank other members of my dissertation committee Dr Richard Anstee, Dr Daniel Granot and Dr Maurice Queyranne for their helpful suggestions.

Finally, special thanks to my husband Darko for " sailing in the same boat " , and to our two little beauties Lea and Nina for sharing their *mama* with many hours of overtime work.

To the memory of my mother, Fumica Boljunčić,

and to my father, Ivan Boljunčić .

Chapter I

INTRODUCTION

Many problems in economics, statistics and numerical analysis can be formulated as the optimization of a convex quadratic function over a polyhedral set. Moreover, some algorithms for solving large scale mathematical programming problems minimize a quadratic function over a polyhedral set as a subroutine, e.g. Held at al. [27] , Kennington and Shalaby [29] . Several methods that are based on solving a quadratic programming subproblem to determine a direction of search were also suggested for optimization problems with nonlinear constraints, e.g. Biggs [5] , Garcia and Mangasarian [20] and Gill at al. [21]. The existence of efficient quadratic programming algorithms and the fact that nonlinear functions can be sometimes accurately approximated by quadratic functions led to the development of approximation methods that make use of quadratic subproblems, e.g. Fletcher [18]. The above mentioned are just some of the reasons why the quadratic programming arose as a very important part of the rich theory of Mathematical Programming.

We start by presenting in Section 1.1. some preliminary definitions from linear algebra and convexity theory to be used in this thesis. Many of the results in this thesis make use of duality. We will review the convex nonlinear programming problem and its dual as stated by Wolfe[49] and, as a special case in Section 1.2., a convex quadratic programming problem and its dual as stated by Dorn [13]. In Section 1.3.

we introduce integer quadratic programming problems. In Section 1.4. some transformations of quadratic programs which will be used in subsequent Chapters are summarized. We do not attempt to survey the algorithms suggested for optimizing a convex quadratic function, rather we restrict our attention to polynomially bounded algorithms and review them in Section 1.5. An introduction to lattices and the transformation of simultaneous Diophantine approximation problem to a short lattice vector problem will be given in Section 1.6.

1.1. PRELIMINARY DEFINITIONS FROM LINEAR ALGEBRA AND CONVEXITY THEORY

In this Section we first review some well known definitions and results the details on which can be found in any text-book of linear algebra.

We are considering in this thesis the vector space R^n . The elements of R^n are ordered tuples of real numbers denoted as $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ and refered to as vectors. The elements of R are called scalars and will be denoted by lowercase Greek letters α, β, γ, etc.

The function $N(x) : R^n \rightarrow R$ is called a norm if

1) $N(x) \geq 0$ for all $x \in R^n$, $N(x) = 0$ if and only if $x = 0$;

2) $N(\alpha x) = |\alpha| N(x)$ for all $x \in R^n$, $\alpha \in R$;

3) $N(x + y) \leq N(x) + N(y)$ for all x , $y \in R^n$.

We will use the standard notation $\| x \|$ to denote the norm of a vector x . For $1 \leq p < \infty$ the p-norm will be given by

$$\|x\|_p = (\sum_{i=1}^{n} |x_i|^p)^{\frac{1}{p}}, \quad x \epsilon R^n \quad . \tag{1.1}$$

In this thesis we will make use of

$\|x\|_1 = \sum_{i=1}^{n} |x_i|$ the l_1-norm ,

$\|x\|_2 = (\sum_{i=1}^{n} |x_i|^2)^{\frac{1}{2}}$ the l_2-norm ,

and

$\|x\|_\infty = \max\{|x_i| : i = 1, ...n\}$ the l_∞-norm .

These three norms are equivalent in the sense that for any $x \in R^n$,

$$\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty \tag{1.2}$$

$$\|x\|_\infty \leq \|x\|_1 \leq n\|x\|_\infty \tag{1.3}$$

$$\|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\|x\|_2. \tag{1.4}$$

For a scalar α, $\lceil \alpha \rceil$ (resp., $\lfloor \alpha \rfloor$) will denote the smallest (resp., greatest) integer not smaller (resp., not greater) than α .

The constraint sets of optimization problems to be treated in this thesis are sets of linear equalities and (or) inequalities which are usually stated in matrix form, therefore the second algebraic object of our interest are real matrices. Unless otherwise stated, in our thesis vectors are considered as column vectors, i.e. $x \in R^n$ is an $n \times 1$ matrix. Transposition applied to a vector x (resp., matrix A) will have the usual notation x^T (resp., A^T).

For our purposes , some special square matrices deserve to be mentioned here.

1) **Identity matrix** I : $I = (\delta_{ij})_{i,j=1}^{n}$, $\delta_{ij} = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{for } i \neq j \end{cases}$.

2) **Diagonal matrix** D : $D = (d_{ij})_{i,j=1}^{n}$, $d_{ij} = 0$ for $i \neq j$.

3) **Nonsingular matrix** A : A is nonsingular if and only if there exist a matrix B such that $AB = BA = I$. B is usually denoted as A^{-1} and is refered to as the *inverse* of A .

4) **Symmetric matrix** A : A is symmetric if and only if $A = A^T$.

5) **Positive Semidefinite matrix** A : A is positive semidefinite if and only if $x^T Ax \geq 0 \;\forall\, x \in R^n$. A is **positive definite** if in addition $x^T Ax = 0$ implies $x = 0$.

6) **Idempotent matrix** P : P is idempotent if and only if $P^2 = P$.

The vectors $x_1, ..., x_r$ are said to be **linearly independent** if and only if their linear combination vanishes in a trivial way only , i.e.

$$\alpha_1 x_1 + ... + \alpha_r x_r = 0$$

implies $\alpha_1 = ... = \alpha_r = 0$. Otherwise, the vectors are said to be **linearly dependent** and each vector x_i with $\alpha_i \neq 0$ can be expressed as a linear combination of the remaining vectors.

A **row** (resp., **column**) **rank** of an $n \times n$ matrix A is the number of its linearly independent rows (resp., columns). The row and column rank of a given matrix always coincide and will be denoted by $r(A)$ to be refered to as the **rank** of the matrix. A nonsingular $n \times n$ matrix A has a full rank , i.e. $r(A) = n$.

A **determinant** is a function that assigns to each $n \times n$ matrix A with columns A_1,...,A_n a scalar value denoted by $det\ A$ that has the following properties:

For each scalar α and each $i = 1, ..., n$

1) $det(A_1, ..., \alpha A_i, ..., A_n) = \alpha det(A_1, ..., A_i, ..., A_n)$,

2) $det(A_1, .., A_i, .., A_j, .., A_n) = det(A_1, .., A_i + \alpha A_j, .., A_j, .., A_n)$ for each $j \neq i$,

3) $det(I) = 1$.

As a consequence, it can be observed that $det A$ for a singular matrix A is equal to zero. For a nonsingular matrix with all entries integral , $det A$ is not less than one.

For $n > 3$, determinants are very inefficient as a computational tool, but they are useful to obtain some theoretical results. For example, if A is an $n \times n$ nonsingular matrix then the **cofactor** of any element a_{rs} of A is defined as

$$cof\, a_{rs} = (-1)^{r+s} det A_{rs}, \tag{1.5}$$

where A_{rs} is the $(n-1) \times (n-1)$ matrix obtained from A by deleting row r and column s. The inverse matrix A^{-1} can then be stated as

$$A^{-1} = \frac{1}{det A} com A \tag{1.6}$$

where $comA$ is the transpose of the matrix of cofactors of A . The unique solution $x^T = (x_1, ..., x_n)$ to a linear system $Ax = b$ where b is an $n \times 1$ vector is given by

$$x_i = \frac{det A_i}{det A} \quad , \quad i = 1, ..., n \tag{1.7}$$

where A_i is the $n \times n$ matrix obtained by replacing the i-th column of A by the vector b (*Cramer's* rule). An upper bound on the value of $det A$ is given by the following inequality

$$|det A| \leq \|a_1\|_2 \cdots \|a_n\|_2. \tag{1.8}$$

where $a_1, ..., a_n$ are columns of A (*Hadamard's* inequality).

Combining (1.2) with (1.8) implies that if $A = (a_{ij})$ with $|a_{ij}| \leq \alpha$ for all i , j , then

$$|det A| \leq \alpha^n n^{\frac{n}{2}}.$$

For an $m \times n$ matrix A we will denote by

$$\Delta(A) = \max \{ \; |det H| : \text{all square submatrices } H \text{ of } A\}.$$

The scalar $\Delta(A)$, where A is the constraint matrix of a quadratic optimization problem, will play an important role in many results of this thesis.

For an $m \times n$ matrix A , the matrix norms to be considered in this thesis are

$$\|A\|_1 = \max_{1 \le j \le n} \sum_{i=1}^{m} |a_{ij}| \quad , \tag{1.9}$$

$$\|A\|_\infty = \max_{1 \le i \le m} \sum_{j=1}^{n} |a_{ij}| \quad . \tag{1.10}$$

The matrix norms are consistent with the vector norms in the sense that

$$\|Ax\|_1 \le \|A\|_1 \|x\|_1 \quad ,$$

$$\|Ax\|_\infty \le \|A\|_\infty \|x\|_\infty \quad .$$

Next, we will review some definitions from convexity theory (see for example Stoer and Witzgall [44]).

A function $f : R^n \to R \cup \{+\infty, -\infty\}$ is **convex** if

i) $\{x : f(x) = -\infty\} = \emptyset$,

ii) $\{x : f(x) < +\infty\} \ne \emptyset$,

iii) $f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y)$ for $0 \le \lambda \le 1$ and $x,\ y \in R^n$.

EXAMPLE A quadratic function $f(x) = c^T x + \frac{1}{2} x^T D x$ with D symmetric positive semidefinite is a convex function. This since $(x-y)^T D (x-y) \ge 0$ for all x, y and therefore $x^T D x + y^T D y \ge 2 x^T D y$ which in turn implies

$$\begin{aligned} & f(\lambda x + (1-\lambda)y) \\ &= c^T(\lambda x + (1-\lambda)y) + \tfrac{1}{2}(\lambda x + (1-\lambda)y)^T D (\lambda x + (1-\lambda)y) \\ &= \lambda c^T x + (1-\lambda) c^T y + \tfrac{1}{2}\lambda^2 x^T D x + \tfrac{1}{2}(1-\lambda)^2 y^T D y + \lambda(1-\lambda) x^T D y \end{aligned}$$

$$\leq \lambda c^T x + (1-\lambda)c^T y + \tfrac{1}{2}\lambda^2 x^T Dx + \tfrac{1}{2}(1-\lambda)^2 y^T Dy$$
$$+\tfrac{1}{2}\lambda(1-\lambda)x^T Dx + \tfrac{1}{2}\lambda(1-\lambda)y^T Dy$$
$$= \lambda c^T x + (1-\lambda)c^T y + \tfrac{1}{2}\lambda x^T Dx + \tfrac{1}{2}(1-\lambda)y^T Dy$$
$$= \lambda f(x) + (1-\lambda)f(y) \ .$$

If D is positive definite, then $f(x) = c^T x + \frac{1}{2}x^T Dx$ is a **strictly convex function.** With regard to the optimization of a convex function, recall that every local minimum is a global one and that a strictly convex function has a unique minimum. If f is convex, then $-f$ is **concave.**

A set $S \subset R^n$ is said to be a **convex set** if it contains all *convex* combinations of its elements, i.e. for any $x^1, x^2, ..., x^s \ \epsilon \ S$, $\sum_{i=1}^{s} w_i x^i \ \epsilon \ S$ for all $w_i \geq 0$, $\forall \, i$ and $\sum_{i=1}^{s} w_i = 1$.

If S contains all *nonnegative* linear combinations of its elements , i.e. $\sum_{i=1}^{s} w_i x^i$, $w_i \geq 0$ for all i , then it is called a **convex cone** .

The solution set P of a finite system of linear inequalities $Ax \leq b$ is a convex set refered to as a **polyhedron** . A cone which is a polyhedron is called a **polyhedral cone** and it is the solution set of some homogeneous system of linear inequalities $Ax \leq 0$. In Chapter III we will use the following theorem due to Minkowski (which is a special case of Carathéodory's theorem see e.g. [44], page 55):

THEOREM 1.1.1 Every polyhedral cone has a finite set of generators.

PROOF See Stoer and Witzgall [44], Theorem 2.8.6., page 55.

□

1.2. DUALITY IN CONVEX NONLINEAR PROGRAMMING

Although our thesis is concerned mainly with quadratic programming, the results presented in Chapter IV were shown to be valid for a broader class of problems, namely for convex nonlinear programs in which some variables are restricted to be integral. In this Section we will review a duality for convex nonlinear problems as introduced by Wolfe [49] and will then state a duality theorem of Dorn [13] for a special case of quadratic convex programming problems.

Consider the following nonlinear programming problem

$$\min \{f(x) : g_i(x) \geq b_i, i = 1, ..., m \} \tag{1.11}$$

where x is an $n \times 1$ vector, f is (resp., g_i , $i = 1, ..., m$ are) real valued, differentiable convex (resp., concave) functions on R^n .

Denote by S a set of feasible solutions to (1.11). In the sequel we assume that on the boundary of the constraint set no singularities will occur, i.e. that some constraint qualification is satisfied. (For detailed discussion on constraint qualifications see e.g. [3] or [36].) For example, we can assume Abadie's constraint qualification (see [4]) of the following form to be valid.

Let $\bar{x} \in S$, then every z satisfying $\nabla g_i(\bar{x})^T z \geq 0$ for all i such that $g_i(\bar{x}) = b_i$ has to be an element of a *cone of tangents*

$$T = \{ z : z = \lim_{n \to \infty} \lambda_n (x^n - \bar{x}),\ x^n \in S,\ \lambda_n \geq 0 \quad \text{for} \quad \text{all} \quad n,\ \lim_{n \to \infty} x^n = \bar{x} \}.$$

If the constraints are all linear, this constraint qualification is automatically satisfied. (See Bazaraa [4], Lemma 5.1.4., page 164.) The Karush-Kuhn-Tucker necessary optimality conditions (which are also sufficient optimality conditions under suitable convexity assumptions) are as follows.

If $\bar{x}$ is an optimal solution to problem (1.11) under some constraint qualification condition, then there exists a vector $u \geq 0$ such that

$$\nabla f(\bar{x}) = u^T \nabla g(\bar{x})$$

and

$$u^T(g(x) - b) = 0 \ .$$

For problem (1.11) Wolfe's dual [49] is given by

$$\begin{aligned} \max \quad & b^T u + f(x) - u^T g(x) \\ \text{s.t.} \quad & u^T \nabla g(x) = \nabla f(x) \\ & u \geq 0 \end{aligned} \tag{1.12}$$

where $b^T = (b_1, ..., b_m)$ and $g(x)^T = (g_1(x), ..., g_m(x))$.

If all the constraints are linear, then the objective function of (1.12) can be equivalently written as $b^T u + f(x) - x^T \nabla f(x)$.

Consider now the convex quadratic programming problem

$$\min \{c^T x + \frac{1}{2} x^T D x : Ax \geq b, x \geq 0\} \tag{1.13}$$

where D is a symmetric positive semidefinite $n \times n$ matrix , c and x are $n \times 1$ vectors , b is an $m \times 1$ vector and A is an $m \times n$ matrix. Positive semidefiniteness of the matrix D implies convexity of the objective function. As stated by Dorn [13] a dual of problem (1.13) can be written as

$$\max \{b^T u - \frac{1}{2} x^T D x : A^T u - Dx \leq c, u \geq 0\} \ . \tag{1.14}$$

The Karush-Kuhn-Tucker optimality conditions for a pair of problems (1.13) and

(1.14) are the primal and dual feasibility conditions

$$
\begin{aligned}
Ax &\geq b \\
x &\geq 0 \\
A^T u - Dx &\leq c \\
u &\geq 0 \, ,
\end{aligned}
\tag{1.15}
$$

and the complementary slackness conditions

$$
\begin{aligned}
x^T(c - A^T u + Dx) &= 0 \\
u^T(Ax - b) &= 0 \, .
\end{aligned}
\tag{1.16}
$$

The existence theorem for quadratic programming states that the feasibility of both the primal and dual programs implies the existence of optimal solutions for each of them . The following theorem is taken from Dorn [13].

THEOREM 1.2.1 i) If $x = \bar{x}$ is a solution to (1.13) , then a solution $(u, x) = (u_0, x_0)$ exists to problem (1.14) such that $D\bar{x} = Dx_0$.

ii) Conversely , if a solution $(u, x) = (u_0, x_0)$ to problem (1.14) exists, then a solution $\bar{x}$ which satisfies $D\bar{x} = Dx_0$ exists to problem (1.13).

In either case the objective function values for (1.13) and (1.14) are equal.

PROOF See Dorn [13], page 156.

□

Also, if one of the problems (1.13) or (1.14) is feasible while the other is not, then on its constraint set the objective function of the feasible program is unbounded in the direction of extremization (see [13]).

The Fundamental Theorem of linear programming states that if a linear program has an optimal solution, then it has one which is a basic solution of a linear system of constraints. For a quadratic programming problem this is, however, not the case. An optimal solution for a quadratic programming problem may occur everywhere in the feasible region, in the interior as well as on the boundary. Consideration of nonbasic solutions makes quadratic programming more difficult than the linear one. However, if a quadratic program has an optimal pair (x^T, u^T) of primal and dual solutions satisfying (1.15) and (1.16)), then it has one which is a basic solution for the system of linear equalities and inequalities (1.15) or, equivalently, a solution that is a vertex of a polyhedron defined by (1.15). Combined with Cramer's rule (see (1.7)) this fact gives us a way to bound the values of the primal and dual variables. This will be discussed in more detail in Section 1.5.

1.3. INTEGER QUADRATIC PROGRAMMING

Many real world problems require a mathematical programming formulation in which all or some of the variables are restricted to be integral. Moreover, a quadratic objective function enables one to take into account the interactions between variables. The applications in, for example, finance [34], capital budgeting [31] or scheduling [40] have natural representations as $0-1$ quadratic programming (i.e. integer quadratic programming in which the variables are restricted to be zero or one).

In this thesis we will consider a general mixed-integer quadratic programming problem and will discuss some sensitivity aspects of it (see Chapter III and IV) as well as a transformation of the objective function coefficients in Chapter V.

1.4. SOME TRANSFORMATIONS OF QUADRATIC PROGRAMS

In some cases the form of the objective function $c^T x + x^T Dx$ of a quadratic programming problem is not suitable for our purposes. In this event some transformations are performed to obtain an equivalent quadratic programming problem of suitable form. In this section we will list the transformations of the objective function to be used in subsequent chapters.

Consider, for example, the quadratic cost matrix D . Without loss of generality one can assume that D is symmetric since, if not, $\tilde{D} = \frac{1}{2}(D + D^T)$ is symmetric and replacing D by $\tilde{D}$ in a quadratic programming problem of the form $\min \{c^T x + x^T Dx : Ax \geq b, x \geq 0\}$ will not change the objective function value.

Suppose, next, that we have a quadratic $0-1$ minimization problem of the form

$$\begin{aligned} \min\ & f(x) = c^T x + x^T Dx \\ \text{s.t}\ & Ax \geq b \\ & 0 \leq x \leq 1 \\ & x \text{ integer} \end{aligned} \tag{1.17}$$

If we want to solve this problem by implicit enumeration where at each node the continuous relaxation of a corresponding integer subproblem is solved, then we would like to ensure the convexity of the objective function (in order to avoid local minimum points). If D is not positive semidefinite, it is shown in [25] that problem (1.17) can be replaced by an equivalent problem in which the objective function is given by

$$f'(x) = (c^T - \lambda e^T)x + x^T(D + \lambda I)x \tag{1.18}$$

where $e^T = (1, ..., 1)$ and λ is a positive scalar such that $D + \lambda I$ is positive semidefinite. This is due to the fact that $x^2 = x$ for any vector x of zeros and ones.

It is often desirable to have a homogeneous quadratic form in the objective function of a quadratic programming problem. This can be achieved by adding a new variable and a new constraint. For example, for problem (1.13) an equivalent homogeneous problem is

$$\begin{aligned} \min \quad & \frac{1}{2}(x^T,\alpha)\begin{pmatrix} D & c \\ c^T & 0 \end{pmatrix}\begin{pmatrix} x \\ \alpha \end{pmatrix} \\ \text{s.t.} \quad & Ax \geq 0 \\ & \alpha = 1 \\ & x \geq 0 \,. \end{aligned} \tag{1.19}$$

Note that if D is positive definite adding a constant $\frac{1}{2}c^T D^{-1}c$ to the objective function results with a convex quadratic program since the matrix $\begin{pmatrix} D & c \\ c^T & c^T D^{-1} c \end{pmatrix}$ is positive definite (observe that $(x^T,\alpha)\begin{pmatrix} D & c \\ c^T & c^T D^{-1} c \end{pmatrix}\begin{pmatrix} x \\ \alpha \end{pmatrix}$ can be written as $(x + \alpha c^T D^{-1})^T D(x + \alpha c^T D^{-1})$).

An alternative way to homogenize the objective function of a strictly convex quadratic programming problems is to leave the matrix D unchanged, but to change the right hand side vector. This by replacing the vector x in (1.13) by $z - D^{-1}c$. The resulting problem is

$$\begin{aligned} -\frac{1}{2}c^T D^{-1}c \quad + \quad \min \quad & \frac{1}{2}z^T Dz \\ \text{s.t.} \quad & Az \geq b + AD^{-1}c \\ & Iz - Ix = D^{-1}c \\ & x \geq 0 \,. \end{aligned} \tag{1.20}$$

We will employ this transformation in Chapter II.

Due to the fact that a positive semidefinite matrix can be diagonalized, we shall see in Chapter III how a convex quadratic programming problem can be replaced by an equivalent separable convex quadratic programming problem.

Without loss of generality assume that a positive semidefinite matrix C is of the form $\left(\begin{smallmatrix} A & 0 \\ 0 & 0 \end{smallmatrix}\right)$ where $A = (a_{ij})$ is positive definite (this can be done by a change of coordinates). Using the diagonalization for positive definite matrices proposed in [48], we obtain

$$A = LDL^T \tag{1.21}$$

where D is a positive diagonal matrix and L is lower triangular with ones on the diagonal. The entries in D and L are determined using the following equations (which are modifications of the equations from Cholesky's decomposition)

$$\sum_{k=1}^{j} \ell_{ik} d_k \ell_{jk} = a_{ij} \quad \text{giving} \quad \ell_{ij} d_j = a_{ij} - \sum_{k=1}^{j-1} \ell_{ik} d_k \ell_{jk} \ , \ j = 1, \ldots, i-1 \ , \tag{1.22}$$

$$\sum_{k=1}^{i} \ell_{ik} d_k \ell_{ik} = a_{ii} \quad \text{giving} \quad d_i = a_{ii} - \sum_{k=1}^{i-1} \ell_{ik} d_k \ell_{ik}. \tag{1.23}$$

From (1.21) if follows that C can be written as $\left(\begin{smallmatrix} L \\ 0 \end{smallmatrix}\right) D(L^T, 0^T)$. Observe that no square roots appear in the process of diagonalization.

1.5. A REVIEW OF POLYNOMIAL ALGORITHMS FOR QUADRATIC PROGRAMS

Let us start this Section with some preliminary definitions (see for example [43] or [45]).

1) The **dimension of the input** is the number of data in the input of a given optimization problem. (Therefore, each number adds one to the dimension of the input).

2) The **size of a number** is the length of its binary description (i.e. the number of bits needed to record a given number in a binary format). For a rational number $\frac{p}{q}$ its size $s(\frac{p}{q})$ is given by

$$s(\frac{p}{q}) = \lceil log_2(p+1) \rceil + \lceil log_2(q+1) \rceil + 1 \; . \tag{1.24}$$

A size of a rational vector $r^T = (r_1, ..., r_n)$ (resp., matrix $A = (a_{ij})_{i=1,j=1}^{m \;\; n}$) is given by $s(r) = n + \sum_{i=1}^{n} s(r_i)$ (resp., $s(A) = mn + \sum_{i,j} s(a_{ij})$).

3) The **elementary arithmetic operations** are additions, comparisons, multiplications and divisions.

4) An algorithm's **running time** is the number of arithmetic operations performed in it.

5) An algorithm is **polynomial** if it has running time polynomial in the dimension and in the size of the input and, when applied to rational input, the size of the numbers occuring in it is polynomially bounded by the dimension of the problem and the size of the input numbers.

6) An algorithm is **strongly polynomial** if it has running time polynomial in the dimension of the input and, when applied to rational input, the size of the numbers occuring in it is polynomially bounded by the dimension of the problem and the size of the input numbers.

Note that the merit of a strongly polynomial algorithm lays in the fact that its running time is independent of the size of the input data.

In 1979 Hačijan [24] proved that a linear programming problem is polynomially solvable. In his result the size of the input was a very important constant. For the sake of completeness, let us state some well known results with respect to the size of the input.

THEOREM 1.5.1 Let A be a square rational matrix of size σ . Then the size of $det(A)$ is less then 2σ .

PROOF See Schrijver [43], Theorem 3.2., page 29.

$\square$

Consider a polyhedron $P = \{Ax \leq b\}$ and assume that A and b are integral. Denote by L the size of the input for P .

LEMMA 1.5.2 If the system $Ax \leq b$ is consistent, then there exists a solution x^0 in the Euclidean ball $S = \{x : \|x\| \leq 2^L\}$.

PROOF See Hačijan [24], Lemma 1.

$\square$

As a consequence of Hačijan's result, in the same year Kozlov, Tarasov and Hačijan [30] gave a polynomial algorithm for quadratic programs with n variables and m constraints of the form

$$\min\ f(x) = d^T x + \frac{1}{2} x^T C x \qquad (1.25)$$
$$\text{s.t. } Ax \leq b$$

where C is an integral symmetric positive semidefinite matrix and all other input data are integrals as well. The size of the input was defined as

$$\begin{aligned} L = & \sum_{i,j=1}^{n} log_2(|c_{ij}| + 1) + \sum_{j=1}^{n} log_2(|d_j| + 1) \\ & + \sum_{i,j=1}^{m,n} log_2(|a_{ij}| + 1) + \sum_{i=1}^{m} log_2(|b_i| + 1) + log_2 mn + 1 \, . \end{aligned} \tag{1.26}$$

Now, if problem (1.25) has a solution, it has a pair of optimal primal and dual variables $(\bar{x}^T, \bar{y}^T)$ which is a vertex of the polyhedron P' defined by

$$\begin{aligned} Ax &\leq b \\ A^T y - Dx &= d \\ y &\geq 0 \, . \end{aligned} \tag{1.27}$$

Observe that the length L' of the input data for P' is not greater than $2L$. Using Lemma 1.5.2, one can then give an upper bound to the components of $(\bar{x}^T, \bar{y}^T)$. Namely, any component of $(\bar{x}^T, \bar{y}^T)$ will have the form $\frac{t}{s}$ where t and s are integers such that $|t|$, $|s| \leq 2^{2L}$. Since the objective function is quadratic, the smallest rational number will be $\frac{1}{(2^{2L})^2} = \frac{1}{2^{4L}}$ and hence

$$|d^T x + \frac{1}{2} x^T D x| \leq 2^{5L} \, .$$

After checking the compatibility of a system of linear equalities and inequalities (1.27), Kozlov at al. found the exact value $f_0 = \frac{t}{s}$ of the objective function. This was done by checking the compatibility of $13L + 2$ systems P_k of the form

$$\begin{aligned} Ax &\leq b \\ f(x) &\leq \frac{t_k}{s_k} \end{aligned}$$

where t_k and s_k are integers such that $|t_k| \leq 2^{13L+2}$ and $|s_k| \leq 2^{8L+2}$. Having the objective function value of (1.25), an exact optimal point was obtained.

Helgason, Kennington and Lall [28] gave a polynomially bounded algorithm for a class of singly constrained quadratic programming problems of the form

$$\begin{aligned} \min\ & \frac{1}{2}x^T D x - a^T x \\ \text{s.t.}\ & \sum_{j=1}^{n} x_j = c \\ & 0 \leq x \leq b \end{aligned} \tag{1.28}$$

where D is a positive diagonal matrix and b is a nonnegative vector. This very specific problem arised in a decomposition procedure to solve multicommodity network flow problems using subgradient optimization. The structure of the objective function and constraints enables one to explicitly represent the primal and dual variables as a function of a single dual variable λ (the one associated with the constraint $\sum_{j=1}^{n} x_j = c$). The optimal value of λ can be found by using a binary search on the interval with $2n$ points. According to definition, since the number of elementary arithmetic operations is polynomially bounded by the number of variables, Helgason at al.'s algorithm for solving the restrictive class of problems (1.28) is strongly polynomial.

A more general class of problems was studied by Minoux [39]. He devised a polynomial algorithm for minimum quadratic cost flow problems of the form

$$\begin{aligned} \min\ & \sum_{u \epsilon U} \frac{1}{2}\omega_u(\varphi_u - \bar{\varphi}_u)^2 \\ \text{s.t.}\ & A\varphi = 0 \\ & b_u \leq \varphi_u \leq c_u \quad , \quad \forall u \epsilon U \end{aligned} \tag{1.29}$$

where U is the set of arcs and A is the node-arc incidence matrix of a given graph,

b_u (resp., c_u) is the integer lower (resp., upper) bound on the flow value φ_u on arc u , $\bar{\varphi}_u$ are integer constants and ω_u are positive numbers for each u .

Minoux's algorithm consists of solving a number of successive approximations of the initial problem, obtained by replacing the quadratic cost function with a piecewise linear convex cost function, using the out-of-kilter algorithm. The approximations are iteratively refined until a point, sufficiently close to the optimal solution of the initial problem, is obtained. Using the fact that for convex quadratic programs with nonempty feasible sets there exist an optimal point which is a basic solution of the polyhedron defined by the primal and dual constraints, the minimal distance between two basic solutions defines the sufficient approximation to locate the optimal solution. Since the number of approximations is polynomially bounded by the input size and the out-of-kilter is a polynomial time algorithm, the polynomial bound on the running time of Minoux's algorithm follows.

A drawbeck of Kozlov at al. and Minoux's algorithms is that their running time depends on the input size. It is still an open question whether there exist an algorithm for solving convex quadratic programming in strongly polynomial time. In Chapter II we will present an algorithm for strictly convex quadratic programming problems which runs in time independent of the size of the linear cost coefficients and the right hand side vector.

1.6. BASIS REDUCTION ALGORITHM AND APPLICATION TO THE SIMULTANEOUS DIOPHANTINE APPROXIMATION PROBLEM

We start this Section by introducing a lattice and a problem of finding a short vector in it. We then state the simultaneous Diophantine approximation problem and show how it can be transformed into a short lattice vector problem. More detailed discussion can be found in e.g. Lovász [35] or Schrijver [43].

Let $a_1, ..., a_n$ be linearly independent real vectors in R^n . The set of all points $x = z_1 a_1 + ... + z_n a_n$ with *integral* $z_1, ..., z_n$ is called a **lattice** with basis $A = (a_1, ..., a_n)$ and denoted by $\Lambda(A)$. A lattice is an additive group and its importance follows from the fact that it is the most general group of vectors in an $n-$dimensional space which contains n linearly independent vectors and which further satisfies the property that there is some sphere around the origin which contains no other vector of the group except the origin.

The basis is not uniquely determined by the lattice. For example, we can define

$$a_i' = \sum_{j=1}^{n} v_{ij} a_j \quad , \quad i = 1, ..., n \tag{1.30}$$

where v_{ij} are any integers with $det(v_{ij}) = \pm 1$. Then each a_i can be written as $\sum_{j=1}^{n} w_{ij} a_j'$ with integral w_{ij} . Substituting this expression into (1.30) and using the linear independence of a_i , it follows that

$$\sum w_{ij} v_{jl} = \begin{cases} 1 & \text{if } i = l \, , \\ 0 & \text{otherwise} \, . \end{cases}$$

Hence, $det(w_{ij}) det(v_{ij}) = 1$ implying $det(w_{ij}) = det(v_{ij}) = \pm 1$. It follows then that $det(a_1, ..., a_n) = \pm det(a_1', ..., a_n')$. (See for example Cassels [7]). We can, therefore, define the **determinant of a lattice**, $det\Lambda = |det(A)|$ where $A = (a_1, ..., a_n)$ is *any* basis of the lattice. Geometrically, $det\Lambda$ denotes the volume of a parallelopiped whose vertices are lattice points and which contains no other lattice point.

An upper bound of $det\Lambda$ is given by Hadamard's inequality

$$det\Lambda \leq \|a_1\|_2 \cdots \|a_n\|_2 \quad .$$

A lower bound of $det\Lambda$ is due to Hermite (1850) and is given as follows.

Every $n-$dimensional lattice Λ has a basis $(b_1, ..., b_n)$ such that

$$\|b_1\|_2 \cdots \|b_n\|_2 \leq c_n det\Lambda \tag{1.31}$$

where c_n is a constant depending only on n .

This result led to the question of finding such a basis. As stated by Minkowski, the solution to the above problem always exists provided $c_n \geq \left(\frac{2n}{\pi e}\right)^{\frac{n}{2}}$. To find a basis in the lattice for which the product of the euclidean norms of its vectors is minimal is $NP-$hard. However, for a weaker bound, taking $c_n = 2^{\frac{n(n-1)}{4}}$, Lovász gave a polynomial algorithm to find a basis satisfying (1.31). (See Lenstra at al. [33]).

A related problem is the following (Short Lattice Vector Problem) :

Given an $n-$dimensional lattice Λ and a number λ , find a vector $b \in \Lambda$, $b \neq 0$ such that $\|b\|_2 \leq \lambda$.

A classical result of Minkowski implies that for $\lambda \geq 2\sqrt{\frac{n}{2\pi e}}\sqrt[n]{det\Lambda}$ such a vector always exists, but no polynomial algorithm to find it is known to date. The shortest vector in the reduced basis obtained by Lovász's basis reduction algorithm has a length at most $2^{\frac{n-1}{4}}\sqrt[n]{det\Lambda}$. This is not the shortest vector in the lattice, however it is very useful in some applications.

Consider, for example, the Simultaneous Diophantine Approximation Problem (see e.g. Lovász [35]) :

Given $\alpha_1, ..., \alpha_n \in Q$, $0 < \varepsilon < 1$ and $\mathcal{Q} > 0$ find integers $p_1, .., p_n$ and q such that

$$0 < q \leq \mathcal{Q} \quad , \quad |\alpha_i - \frac{p_i}{q}| \leq \frac{\varepsilon}{\mathcal{Q}} \quad , \quad i = 1, ..., n . \tag{1.32}$$

A classical result due to Dirichlet (1842) is the proof of the existence of integers q and p_i, $i = 1, ... n$ such that (1.32) is satisfied, whenever $\mathcal{Q} \geq \varepsilon^{-n}$. No polynomial algorithm to find these integers is known (except for the case $n = 1$ when the method of continued fractions can be applied). However, a weaker approximation (for $\mathcal{Q} \geq 2^{\frac{n(n+1)}{4}} \varepsilon^{-n}$) can be found by transforming this problem into a short lattice vector problem and using Lovász's basis reduction algorithm as follows [35] :

Consider the lattice $\Lambda(A)$ generated by the columns of the $(n+1) \times (n+1)$ nonsingular matrix

$$A = \begin{pmatrix} 1 & \dots & 0 & \alpha_1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 1 & \alpha_n \\ 0 & \dots & 0 & \frac{\varepsilon}{\mathcal{Q}} \end{pmatrix} .$$

Any vector $b \,\epsilon \Lambda(A)$ has the form $b^T = (p_1 + p_{n+1}\alpha_1, ..., p_n + p_{n+1}\alpha_n, p_{n+1}\frac{\varepsilon}{\mathcal{Q}})$, where $p^T = (p_1, ..., p_{n+1}) \,\epsilon\, Z^{n+1}$. Suppose that $b \neq 0$ but $\|b\|_2 \leq \varepsilon$. This implies $p_{n+1} \neq 0$. Without loss of generality assume $p_{n+1} < 0$ and denote $q = -p_{n+1}$. It follows then that

$$|b_i| = |p_i - q\alpha_i| \leq \varepsilon \quad , \quad i = 1, ..., n \tag{1.33}$$

$$|b_{n+1}| = \frac{\varepsilon}{\mathcal{Q}} q \leq \varepsilon \quad \text{or} \quad q \leq \mathcal{Q} . \tag{1.34}$$

The shortest vector in the reduced basis of a lattice $\Lambda(A)$ obtained by Lovász's algorithm satisfies $\|b\|_2 \leq 2^{\frac{n}{4}} \sqrt[n+1]{det\Lambda(A)} = 2^{\frac{n}{4}} \left(\frac{\varepsilon}{\mathcal{Q}}\right)^{\frac{1}{n+1}}$. For $\mathcal{Q} = 2^{\frac{n(n+1)}{4}} \varepsilon^{-n}$, $\|b\|_2 \leq \varepsilon$ and hence $\|b\|_\infty \leq \|b\|_2 \leq \varepsilon$.

We shall make use of Lovász's basis reduction algorithm in Chapter 5 in order to find simultaneous approximations of the objective function coefficients of a given quadratic integer programming problem.

Chapter II

TOWARDS A STRONGLY POLYNOMIAL ALGORITHM FOR STRICTLY CONVEX QUADRATIC PROGRAMS

In [45] Tardos was the first to present a polynomial algorithm for solving linear programming problems in which the number of arithmetic steps depends only on the size of the numbers in the constraint matrix and is independent of the size of the numbers in the right hand side and the cost coefficients.

The aim of this Chapter is to extend Tardos' results to convex quadratic programming problems of the form $\max\{c^T x - \frac{1}{2}x^T Dx : Ax \leq b, x \geq 0\}$ with D being a positive definite matrix. We assume, without loss of generality, that A and D are integral. We develop a polynomially bounded algorithm for solving the strictly convex quadratic problem where the number of arithmetic steps is independent of c and b but depends on the size of the entries of the matrices A and D. If in particular the size of the entries in A and D is polynomially bounded in the dimension of the input, the algorithm is strongly polynomial, e.g., when the quadratic term corresponds to a least squares and A is a node arc incidence matrix of a directed graph.

Following Tardos [45] the algorithm presented here finds optimal primal and dual solutions to the quadratic programming problem (if they exist) by solving a sequence of simple quadratic programming problems using the polynomial algorithm for solving quadratic programming problems given in [30] and by checking feasibility of a linear system in time independent of the right hand side using Tardos' feasibility algorithm [45].

2.1. SETUP OF THE PROBLEM

For simplicity of exposition we will first consider the quadratic programming problem of the form

$$\max \{ c^T x - \frac{1}{2} x^T D x : Ax = b, \quad x \geq 0 \} \tag{2.1}$$

where

A is an integral $m \times n$ matrix with rank $(A) = m$;

D is an integral $n \times n$ symmetric positive definite matrix;

c and x are n-vectors and b is an m-vector.

The validity of the algorithm for quadratic programming problems with inequality constraints will be discussed in Section 2.3.

Using the fact that D is nonsingular we can substitute $z = x - D^{-1}c$ in (2.1) resulting with the following equivalent problem

$$\frac{1}{2} c^T D^{-1} c + \max \{ -\frac{1}{2} z^T D z : Az = b - AD^{-1}c \, , \; Iz - Ix = -D^{-1}c \, , \; x \geq 0 \}. \tag{2.2}$$

Recall that positive definiteness of D implies that the objective function of (2.1) is strictly concave which in turn implies the uniqueness of the optimal solution of (2.1) (if one exists). Moreover, if the set $\{x : Ax = b, x \geq 0\}$ is not empty, (2.1) will be bounded. The uniqueness of the optimal value of x implies also the uniqueness of the optimal value of z; however observe that z is not restricted to be non-negative any longer. Using now Dorn's duality [13], we can state a dual of the maximization problem given in (2.2) as

$$\min \{y^T(b - AD^{-1}c) - v^T D^{-1}c + \frac{1}{2}z^T Dz : A^T y + Iv + Dz = 0,\ v \leq 0\}. \qquad (2.3)$$

Substituting $v = -Dz - A^T y$ results with (see Section 1.4)

$$\min \{y^T(b - AD^{-1}c) + (y^T A + z^T D)D^{-1}c + \frac{1}{2}z^T Dz : A^T y + Dz \geq 0\} =$$
$$\min \{y^T b - y^T AD^{-1}c + y^T AD^{-1}c + z^T c + \frac{1}{2}z^T Dz : A^T y + Dz \geq 0\}.$$

Finally, after adding slack variables s , we get a dual of (2.2) of the form

$$\frac{1}{2}c^T D^{-1}c + \min \{c^T z + b^T y + \frac{1}{2}z^T Dz : A^T y + Dz - Is = 0,\ s \geq 0\}. \qquad (2.4)$$

It is easy to see that replacing z by $x - D^{-1}c$ will give us the following dual of (2.1)

$$\min \{b^T y + \frac{1}{2}x^T Dx : A^T y + Dx - Is = c, \quad s \geq 0\}. \qquad (2.5)$$

It is important to note that the same slack variables appear both in (2.4) and (2.5). Since the algorithm to be described in the following is significantly simpler when applied to problems with zero right hand side, we will always use the above transformation to replace a pair of primal and dual problems of the form (2.1) and (2.5) by an equivalent pair of the form (2.2) and (2.4).

Recall that the Karush-Kuhn-Tucker optimality conditions for a pair of problems (2.1) and (2.5) will have the form

$$\begin{pmatrix} A & 0 & 0 \\ D & A^T & -I \end{pmatrix} \begin{pmatrix} x \\ y \\ s \end{pmatrix} = \begin{pmatrix} b \\ c \end{pmatrix} \tag{2.6}$$

$$x, s \geq 0 \text{ and } x^T s = 0 .$$

At each iteration of the algorithm to follow we will detect at least one new slack variable which is equal to zero in all optimal pairs of primal and dual solutions for the pair of problems (2.1) and (2.5). Or equivalently, at each iteration we will detect at least one new dual constraint which is tight at optimality. After each iteration we will add constraints of the form $s_i = 0$, $i \in I$ (where I is the set of slack variables detected to be zero at the current iteration) to the linear system given in (2.6) and perform Tardos' feasibility algorithm which will give us a feasible solution (x^T, y^T, s^T). We will check if $x^T s = 0$. If so, the algorithm will terminate since an optimal pair of primal and dual solutions was determined. If on the other hand $x^T s \neq 0$, we will perform another iteration of our algorithm. In at most n iterations we will find a pair of optimal primal and dual solutions.

As stated above the algorithm will be applied to a minimization problem of the form

$$\min \{c^T z + b^T y + \frac{1}{2} z^T D z : Dz + A^T y - Is = 0,\ s \geq 0\}. \tag{2.7}$$

Before stating the algorithm we will give two preliminary lemmas. The first one is a direct generalization of Lemma 0.1 in [45], while the second is a special case of Lemma 2, p. 707 in [15].

LEMMA 2.1.1 Replacing $(c^T, b^T, 0^T)$ in (2.7) by $(c'^T, b'^T, a'^T) = (c^T, b^T, 0^T) - p^T(D, A^T, -I)$ for some n-vector p will not change the set of optimal solutions.

PROOF If $(\bar{z}^T, \bar{y}^T, \bar{s}^T)$ solves (2.7), then it also solves the problem obtained from (2.7) by replacing the linear cost coefficient $(c^T, b^T, 0^T)$ by (c'^T, b'^T, a'^T) and vice versa. This since

$$(c'^T, b'^T, a'^T)\begin{pmatrix} z \\ y \\ s \end{pmatrix} + \frac{1}{2}z^T Dz - (c^T, b^T, 0^T)\begin{pmatrix} z \\ y \\ s \end{pmatrix}$$

$$-\frac{1}{2}z^T Dz = -p^T(D, A^T, -I)\begin{pmatrix} z \\ y \\ s \end{pmatrix} = -p^T 0 = 0$$

which is a constant.

$\square$

LEMMA 2.1.2 If $(\bar{z}^T, \bar{y}^T, \bar{s}^T)$ solve (2.7), then $(\alpha\bar{z}^T, \alpha\bar{y}^T, \alpha\bar{s}^T)$ solves the quadratic problem in which $(c^T, b^T, 0^T)$ in (2.7) is replaced by $\alpha(c^T, b^T, 0^T)$ for any scalar $\alpha > 0$.

PROOF Let $(\bar{z}^T, \bar{y}^T, \bar{s}^T)$ solve (2.7). Then for any scalar $\alpha > 0$, $\alpha c^T(\alpha\bar{z}) + \alpha b^T(\alpha\bar{y}) + \frac{1}{2}(\alpha\bar{z})^T D(\alpha\bar{z}) = \alpha^2(c^T\bar{z} + b^T\bar{y} + \frac{1}{2}\bar{z}^T D\bar{z})$ and, therefore, $\alpha(\bar{z}^T, \bar{y}^T, \bar{s}^T)$ solves (2.7) in which $(c^T, b^T, 0^T)$ was replaced by $\alpha(c^T, b^T, 0^T)$.

$\square$

We are now ready to describe the polynomial quadratic programming algorithm whose number of arithmetic steps is independent of the size of the numbers in the vectors c and b.

2.2. THE QUADRATIC PROGRAMMING ALGORITHM

The QUADRATIC PROGRAMMING ALGORITHM (QPA) described below is a direct generalization to a class of strictly convex quadratic programming problems of Tardos' linear programming algorithm. It uses as input a strictly convex quadratic program (2.1), Tardos' feasibility algorithm, which is a polynomial algorithm to check the feasibility of a system of linear inequalities in time independent of the right hand side and if feasible it generates a basic solution, and a polynomial algorithm for solving convex quadratic programming problems, e.g., Kozlov et al.'s algorithm. The output from QPA is a pair of optimal primal and dual solutions for (2.1).

THE QUADRATIC PROGRAMMING ALGORITHM

STEP 1. Use Tardos' feasibility algorithm to check whether $\{Ax = b, \quad x \geq 0\}$ is feasible. If not, terminate since (2.1) is infeasible. If feasible, then the positive definiteness of D guarantees boundedness which in turn implies that the dual constraint set is feasible. Set $K = \emptyset$.

STEP 2. Let $D^\circ x + A^{\circ T} y - E^\circ s = c^\circ$ denote the equality system $Dx + A^T y - Is = c$ together with $s_i = 0$ for $i \in K$ and let $P^\circ = \{(x^T, y^T, s^T) : Ax = b,$ $Dx + A^T y - Is = c,\ s_i = 0$ for $i \in K$, $x \geq 0,\ s \geq 0\}$. Use Tardos' feasibility algorithm to find a point say $(\bar{x}^T, \bar{y}^T, \bar{s}^T)$ in P°. If $\bar{x}^T \bar{s} = 0$ terminate with $(\bar{x}^T, \bar{y}^T, \bar{s}^T)$ as an optimal solution to (2.1) and (2.5).

STEP 3. Find the projection (c'^T, b'^T, a'^T) of $(c^T, b^T, 0^T)$ onto the null space of $(D^\circ, A^{\circ T}, -E^\circ)$. Since the rows are linearly independent, this can be done using Gaussian elimination. Recall that for a matrix G with full row rank, the projection onto its null space $\{x : Gx = 0\}$ is determined by the idempotent matrix $P \equiv I - G^T(GG^T)^{-1}G$. I.e., for some vector x its projection onto the null space of G is Px $(G(Px) = G(x - G^T(GG^T)^{-1}Gx) = Gx - Gx = 0)$ where $Px = x - G^T(GG^T)^{-1}Gx = x - G^T p$ and $p \equiv (GG^T)^{-1}Gx$. Applying this to our case we obtain

$$(c'^T, b'^T, a'^T) = (c^T, b^T, 0^T) - p^T(D^\circ, A^{\circ T}, -E^\circ).$$

If $(c'^T, b'^T, a'^T) = 0$ then solve the problem $\min\{\frac{1}{2}z^T Dz : D^\circ z + A^{\circ T} y - E^\circ s = 0$, $s \geq 0\}$ by Kozlov et al.'s algorithm to obtain an optimal solution $(x^T = z^T + c^T D^{-1}, y^T, s^T)$.

STEP 4. Let

$$\alpha = \frac{(3n+m)(2n+m)\Delta}{\| (c'^T, b'^T, a'^T) \|_\infty}$$

where Δ is a maximum absolute subdeterminant of $(D, A^T, -I)$,

$$\bar{c}_i = \lceil \alpha c'_i \rceil \qquad i = 1, ..., n \; ,$$

$$\bar{b}_i = \lceil \alpha b'_i \rceil \qquad i = 1, ..., m \; ,$$

$$\bar{a}_i = \lceil \alpha a'_i \rceil \qquad i = 1, ..., n \; .$$

STEP 5. Use Kozlov et al.'s algorithm to find an optimal solution $(\bar{v}^T, \bar{z}^T)$ to

$$\max \{-\frac{1}{2}z^T Dz : \; D^{\circ T} v - Dz = \bar{c}, \; A^\circ v = \bar{b}, \; -E^{\circ T} v \leq \bar{a}\}$$

which is a dual of

$$\min \{\bar{c}^T z + \bar{b}^T y + \bar{a}^T s + \frac{1}{2}z^T Dz : D^\circ z + A^{\circ T} y - E^\circ s = 0, \quad s \geq 0\}. \tag{2.8}$$

Let

$$I = \{i : (-E^{\circ T}\bar{v})_i \leq \alpha a'_i - (2n+m)\Delta\}.$$

Add the set I to K and go to Step 2.

The following lemmas which are extensions of corresponding lemmas in [45] will be used in order to verify the validity of QPA.

LEMMA 2.2.1 Let $(\bar{z}^T, \bar{y}^T, \bar{s}^T)$ be an optimal solution to the problem (2.8) where $(\bar{c}^T, \bar{b}^T, \bar{a}^T) = \lceil (c^T, b^T, a^T) \rceil$ of some vector $(c^T, b^T, a^T) \in R^{n+m+n}$, D is an integral $n \times n$ positive definite matrix and D°, A° and E° are integral matrices of appropriate dimensions. Let $(\bar{v}^T, \bar{z}^T)$ be an optimal dual solution for (2.8). Then for any optimal solution $(\hat{z}^T, \hat{y}^T, \hat{s}^T)$ to

$$\min \{c^T z + b^T y + a^T s + \frac{1}{2} z^T D z : D^\circ z + A^{\circ T} y - E^\circ s = 0, \quad s \geq 0\} \tag{2.9}$$

(i.e., "unrounded" problem), we have

$$(-E^{\circ T} \bar{v})_i \leq a_i - (2n+m)\Delta \text{ implies } \hat{s}_i = 0. \tag{2.10}$$

PROOF The dual of (2.8) has the following set of constraints

$$D^{\circ T} v - Dz = \bar{c} \leq c + e$$
$$A^\circ v = \bar{b} \leq b + e$$
$$-E^{\circ T} v \leq \bar{a} \leq a + e$$

where e is a vector of ones of appropriate dimension. Moreover, since $(\bar{v}^T, \bar{z}^T)$ is an optimal dual solution for (2.8) and $(\bar{z}^T, \bar{y}^T, \bar{s}^T)$ an optimal primal solution, we have from complementary slackness that $\bar{s}^T(\bar{a} + E^{\circ T}\bar{v}) = 0$. Coupling with the fact that $a \leq \bar{a}$ we obtain

$$(-E^{\circ T} \bar{v})_i < a_i \leq \bar{a}_i \text{ implies } \bar{s}_i = 0. \tag{2.11}$$

Now, suppose that $(\hat{z}^T, \hat{y}^T, \hat{s}^T)$ is an optimal solution to the "unrounded" problem (2.9). Since $(\bar{z}^T, \bar{y}^T, \bar{s}^T)$ is a feasible solution for (2.9) we have

$$c^T(\hat{z} - \bar{z}) + b^T(\hat{y} - \bar{y}) + a^T(\hat{s} - \bar{s}) + \frac{1}{2}\hat{z}^T D \hat{z} - \frac{1}{2}\bar{z}^T D \bar{z} \leq 0. \tag{2.12}$$

Substituting $u = \hat{z} - \bar{z}$, $u^1 = \hat{y} - \bar{y}$, $u^2 = \hat{s} - \bar{s}$ in (2.12) and rearranging terms results with

$$c^T u + b^T u^1 + a^T u^2 + \frac{1}{2} u^T D u + \bar{z}^T D u \leq 0. \tag{2.13}$$

Since $\frac{1}{2} u^T D u \geq 0$ for each u, we obtain from (2.13) that

$$c^T u + b^T u^1 + a^T u^2 + \bar{z}^T D u \leq 0. \tag{2.14}$$

We will prove the validity of (2.10) by contradiction. Suppose that (2.10) is not valid, i.e., there exists an index i for which $(-E^{\circ T} \bar{v})_i \leq a_i - (2n+m)\Delta$ and $\hat{s}_i > 0$. Now since by the optimality condition $\bar{s}_i = 0$ we have that $u_i^2 > 0$. Moreover, $u_j^2 \geq 0$ for all j with $\bar{s}_j = 0$. If for some j, $u_j^2 < 0$, then $\bar{s}_j > 0$ which on the other hand implies (see (2.11)) that $(-E^{\circ T}\bar{v})_j \geq a_j$ or $-a_j - (E^{\circ T}\bar{v})_j \geq 0$. Now observe that the vector (u^T, u^{1T}, u^{2T}) satisfies the system $D^\circ u + A^{\circ T} u^1 - E^\circ u^2 = 0$. Thus there also exists an integral basic solution to $D^\circ u + A^{\circ T} u^1 - E^\circ u^2 = 0$, $u_i^2 > 0$ and $u_j^2 \geq 0$ for j with $\bar{s}_j = 0$ that satisfies (2.14). Denote this solution $(\bar{u}^T, \bar{u}^{1T}, \bar{u}^{2T})$ which by Cramer's rule satisfies $\| (\bar{u}^T, \bar{u}^{1T}, \bar{u}^{2T}) \|_\infty \leq \Delta$. Note that for the dual constraints $D^{\circ T} v - Dz = \bar{c}$ and $A^\circ v = \bar{b}$ corresponding to unconstrained primal variables z and y we have $c \leq D^{\circ T} v - Dz \leq c + e$ and $b \leq A^\circ v \leq b + e$ which in turn implies $\| -c + D^{\circ T} v - Dz \|_\infty \leq 1$, $\| -b + A^\circ v \|_\infty \leq 1$, and for j with $u_j^2 < 0$ also $| (-a^T - \bar{v}^T E^\circ)_j | \leq 1$. Combining the above facts and the conditions on $(\bar{u}^T, \bar{u}^{1T}, \bar{u}^{2T})$ we obtain from (2.14) that

$$0 \leq -c^T u - b^T u^1 - a^T u^2 - \bar{z}^T D u = (-c^T - \bar{z}^T D) u - b^T u^1 - a^T u^2)$$

$$= (-c^T - \bar{z}^T D) u - b^T u^1 - a^T u^2 + \bar{v}^T (D^\circ u + A^{\circ T} u^1 - E^\circ u^2)$$

$$= (-c^T - \bar{z}^T D + \bar{v}^T D^\circ) u + (-b^T + \bar{v}^T A^{\circ T}) u^1 + (-a^T - \bar{v}^T E^\circ) u^2$$

$$\leq \sum_{k=1}^n | (-c^T - z^T D + \bar{v}^T D^\circ)_k | \, | u_k |$$

$$+\sum_{k=1}^{m} \mid (-b^T + \bar{v}^T A^{\circ T})_k \mid \mid u_k^1 \mid + \sum_{k=1,k\neq i}^{n} \mid (-a^T - \bar{v}^T E^{\circ})_k \mid \mid u_k^2 \mid$$

$$+(-a^T - \bar{v}^T E^{\circ})_i u_i^2 \leq \sum_{k=1}^{n} \mid u_k \mid + \sum_{k=1}^{n} \mid u_k^1 \mid - (2n+m)\Delta$$

$$+\sum_{k=1,k\neq i}^{n} \mid u_k^2 \mid \leq n\Delta + m\Delta - (2n+m)\Delta + (n-1)\Delta = -\Delta$$

which is a contradiction.

□

Before stating the following lemma (which will ensure the finiteness of the algorithm), we will briefly explain the ideas used in the proof.

Consider a full row rank matrix H of the form $\begin{pmatrix} H_1 & H_2 \\ I & 0 \end{pmatrix}$ and a vector x of the same dimension as the columns of H. Let us partition $x^T = (x_1^T, x_2^T)$ corresponding to the partition of H to H_1 and H_2. Then the projection $(x_1'^T, x_2'^T)$ of x^T onto the null space of H is given by $(0^T, x_2'^T)$. This since $H_1 x_1' + H_2 x_2' = 0$, $x_1' = 0$ has to be valid.

In our case this tells us that wherever $s_k = 0$ is a row of $D^{\circ} z + A^{\circ T} y - E^{\circ} s = 0$ for $k \in \{1, ..., n\}$, the component a_k' of the projection a' determined in Step 3 of the algorithm is equal to zero.

Next, if we have a vector $x = G^T u$ for some full row rank matrix G its projection to the null space of G is zero since $Px = (I - G^T(GG^T)^{-1}G)x = G^T u - G^T(GG^T)^{-1}GG^T u = G^T u - G^T u = 0$.

If we have a vector y which is a projection of some vector w (i.e., $y = Pw$) onto the null space of G, then the projection of y onto the null space of G is the same vector y. This since the projection matrix $P = I - G^T(GG^T)^{-1}G$ is idempotent (i.e., $P \cdot P = P$) and therefore $Py = P(Pw) = Pw = y$.

Now, combining the above facts it is easy to see that for a vector $d = c' + A^T v$, where c' is a projection of some vector c onto the null space of a full row rank matrix A, $Pd = c'$.

LEMMA 2.2.2 The set I found in Step 5 of QPA contains at least one index i such that $s_i = 0$ is not a row of $D^\circ z + A^{\circ T} y - E^\circ s = 0$. (Recall that the latter system is equivalent to $D^\circ x + A^{\circ T} y - E^\circ s = c^\circ$ under the transformation $z = x - D^{-1}c$.)

PROOF The proof will follow by constructing a vector d and looking at its max-norm. We will use the notation from Step 4 and Step 5 of QPA and distinguish between the two cases.

CASE 1. If $\bar{z}$ is not a zero vector, then define the $(3n+m)$ vector d as follows

$$\begin{aligned}
d_i &= -\alpha c_i' + \bar{v}^T D_i^\circ - \bar{z}^T D_i, && i = 1, ..., n \\
d_{n+j} &= -\alpha b_j' + \bar{v}^T A_j^{\circ T}, && j = 1, ..., m \\
d_{n+m+k} &= \begin{cases} 0 \\ -\alpha a_k' - \bar{v}^T E_k^\circ \end{cases} && \begin{array}{l} \text{if } s_k = 0 \text{ is a row of } D^\circ z + A^{\circ T} y - E^\circ s = 0, \\ \text{otherwise for } k = 1, ..., n. \end{array} \\
d_{2n+m+\ell} &= \frac{(D\bar{z})_\ell}{\|D\bar{z}\|_\infty}, && \ell = 1, ..., n .
\end{aligned}$$

Since $\bar{v}, \bar{z}$ are dual variables for the "rounded" problem (2.8) we know that the following is valid:

$$\begin{aligned}
\alpha c' \leq & D^{\circ T}\bar{v} - D\bar{z} = \bar{c} = \lceil \alpha c' \rceil \leq \alpha c' + 1 \\
\alpha b' \leq & A^\circ \bar{v} = \bar{b} = \lceil \alpha b' \rceil \leq \alpha b' + 1 \\
& -E^{\circ T}\bar{v} \leq \bar{a} = \lceil \alpha a' \rceil \leq \alpha a' + 1.
\end{aligned}$$

Therefore $0 \leq -\alpha c' + D^{\circ T}\bar{v} - D\bar{z} \leq 1$, $0 \leq -\alpha b' + A^\circ \bar{v} \leq 1$ and $-\alpha a' - E^{\circ T}\bar{v} \leq 1$.

For the last n components of d we know that $-1 \leq \frac{(D\bar{z})_\ell}{\|D\bar{z}\|_\infty} \leq 1$ for $\ell = 1, .., n$. Now since all the components of d are bounded by 1 from above and, with the exception of components $d_{n+m+1}, ..., d_{2n+m}$ all others are bounded from below by -1, the validity of the Lemma will follow if we can show that

$$\| d \|_\infty \geq (2n+m)\Delta .$$

To that end note that d can be written as

$$d = \alpha \begin{pmatrix} -c' \\ -b' \\ -a' \\ 0 \end{pmatrix} + \begin{pmatrix} D^0 & A^{0T} & -E^0 & 0 \\ -D & 0 & 0 & \frac{1}{\|D\bar{z}\|_\infty} D \end{pmatrix}^T \begin{pmatrix} \bar{v} \\ \bar{z} \end{pmatrix} .$$

It is easy to see that the projection of d onto the null space of

$$\tilde{H} = \begin{pmatrix} D^0 & A^{0T} & -E^0 & 0 \\ -D & 0 & 0 & \frac{1}{\|D\bar{z}\|_\infty} D \end{pmatrix}$$

is given by the vector $\tilde{d} = \alpha(-c'^T, -b'^T, -a'^T, - \| D\bar{z} \|_\infty c'^T)$ (i.e. $\tilde{H} \cdot \tilde{d} = 0$).

Note also that $\| (c'^T, b'^T, a'^T) \|_\infty \leq \| (c'^T, b'^T, a'^T, \| D\bar{z} \|_\infty c'^T) \|_\infty$ and that for any n-vector x, $\| x \|_\infty \geq \frac{1}{n} \| x \|_2$ and $\| x \|_2 \geq \| x \|_\infty$ are valid.

Next, since $-\alpha(c'^T, b'^T, a'^T, \| D\bar{z} \|_\infty c'^T)$ is a projection of d,

$$\| d \|_2 \geq \alpha \| (c'^T, b'^T, a'^T, \| D\bar{z} \|_\infty c'^T) \|_2 .$$

Therefore,

$$\| d \|_\infty \geq \frac{1}{(3n+m)} \| d \|_2 \geq \frac{(3n+m)(2n+m)\Delta}{(3n+m)) \| (c'^T, b'^T, a'^T) \|_\infty} \| (c'^T, b'^T, a'^T, \| D\bar{z} \|_\infty c'^T) \|_2$$

$$\geq (2n+m)\Delta \frac{\| (c'^T, b'^T, a'^T, \| D\bar{z} \|_\infty c'^T) \|_2}{\| (c'^T, b'^T, a'^T, \| D\bar{z} \|_\infty c'^T) \|_\infty} \geq (2n+m)\Delta.$$

CASE 2. It $\bar{z}$ is a zero vector, then define the $(2n+m)$ vector d as follows

$$
\begin{aligned}
d_i &= -\alpha c'_i + \bar{v}^T D^\circ_i, && i = 1, \dots, n \\
d_{n+j} &= -\alpha b'_j + \bar{v}^T A^{\circ T}_j, && j = 1, \dots, m \\
d_{n+m+k} &= \begin{cases} 0 & \text{if } s_k = 0 \text{ is a row of } D^\circ z + A^{\circ T} y - E^\circ s = 0, \\ -\alpha a'_k - \bar{v}^T E^\circ_k & \text{otherwise for } k = 1, \dots, n \,. \end{cases}
\end{aligned}
$$

or, written in matrix form

$$d = \alpha \begin{pmatrix} -c' \\ -b' \\ -a' \end{pmatrix} + \begin{pmatrix} D^\circ & A^{\circ T} & -E^\circ \end{pmatrix}^T \bar{v}.$$

The proof now follows along the same lines as for CASE 1.

□

LEMMA 2.2.3 Every optimal pair of primal and dual variables $(\bar{x}^T, \bar{y}^T, \bar{s}^T)$ satisfies $\bar{s}_i = 0$ for $i \in I$.

PROOF By Lemma 2.1.1 replacing $(c^T, b^T, 0^T)$ by (c'^T, b'^T, a'^T) will not change the set of optimal solutions $(\bar{z}^T, \bar{y}^T, \bar{s}^T)$ for $\min\{c^T z + b^T y + \frac{1}{2} z^T D z : D^\circ z + A^{\circ T} y - E^\circ s = 0^\circ,\ s \geq 0\}$. By Lemma 2.1.2 multiplying the linear part of the objective function by a positive scalar α, one obtains that the set of optimal solutions $(\alpha \bar{z}^T, \alpha \bar{y}^T, \alpha \bar{s}^T)$ and the set of variables that are equal to zero in all optimal solutions is unchanged. Finally, Lemma 2.2.1 holds with c replaced by $\alpha c'$, b replaced by $\alpha b'$, and a replaced by $\alpha a'$.

□

LEMMA 2.2.4 After at most n iterations of QPA one gets a pair of optimal primal and dual solutions for problem (2.1).

PROOF By Lemma 2.2.3 adding constraints $s_i = 0, i \in I$ where I was determined in Step 5 of QPA does not affect the set of optimal solutions $(\bar{x}^T, \bar{y}^T, \bar{s}^T)$. Recall that the set of optimal solutions form a face of a polyhedron $\{Ax = b,$ $Dx + A^T y - Is = c,\ x \geq 0,\ s \geq 0\}$ for which $x^T s = 0$. By Lemma 2.2.2 no more than n iterations of QPA are possible. (In the worst case, where all $\bar{x}_i > 0$ and Tardos' feasibility algorithm in every iteration "missed" the desired face, i.e., face with $x^T s = 0$, and where I is a singleton in every iteration, one will have exactly n iterations.)

$\square$

Let us now calculate the complexity of QPA. Denote by $T(A)$ the complexity of Tardos' feasiblity algorithm when applied to the system $\{Ax = b,\ x \geq 0\}$, and by $T(A, D)$ its complexity when applied to the system $\{Ax = b,\ Dx + A^T y - Is = c,$ $x \geq 0,\ s \geq 0,\ s_i = 0$ for $i \in K\}$. Denote by $K(A, D)$ the complexity of Kozlov et al.'s algorithm. Note that we will apply Kozlov et al.'s algorithm only to quadratic programs for which the linear part of the objective function is integral and polynomially bounded by the matrices A and D and with right hand side vector which is zero.

THEOREM 2.2.5 The QUADRATIC PROGRAMMING ALGORITHM has running time polynomial in the size of the matrices A and D and independent of the sizes of the vectors c and b. It runs in

$$O(n(2n+m)^3 + n(2n+m)log(2n+m)(3n+m)\Delta + T(A) + nT(A,D) + nK(A,D)).$$

PROOF Step 1 takes $T(A)$ time (i.e., time of Tardos' feasibility algorithm when applied to a linear system with constraint matrix A). Consequently Step 2 takes $T(A, D)$ plus at most $2n$ comparisons to verify $x^T s = 0$. The Gaussian elimination in Step 3 takes $O((2n + m)^3)$ time (see Edmonds [16]). Step 4 takes $O((2n + m) log(2n + m)(3n + m)\Delta)$ comparisons to find $(\bar{c}^T, \bar{b}^T, \bar{a}^T)$ since $\| (\bar{c}^T, \bar{b}^T, \bar{a}^T) \|_\infty = (2n + m)(3n + m)\Delta$ and one can use binary search to obtain $(\bar{c}^T, \bar{b}^T, \bar{a}^T)$. Step 5 takes $K(A, D)$ time. Finally, we need at most n iterations and therefore the claimed complexity follows.

□

Recall (see Section 1.5) that an algorithm is termed strongly polynomial if all its operations consist of additions, comparisons, multiplications and divisions and if the number of such steps is polynomially bounded in the dimension of the input, where the dimension of the input is the number of data in the input. Further, when the algorithm is applied to rational input, then the size of the numbers occurring during the algorithm is polynomially bounded in the dimension of the input and the size of the input numbers.

Thus the polynomial algorithm described in this paper becomes strongly polynomial if the size of the entries in A and D are polynomially bounded in the dimension of the data. This clearly provides a strongly polynomial algorithm for, e.g., problems where one minimizes the norm over flow (transportation) type constraints [1,12].

2.3. EXTENSION TO THE INEQUALITY CONSTRAINED CASE

The Algorithm can be generalized in a straightforward way to work on strictly convex quadratic programs with inequality constraints. To show this consider the quadratic programming problem of the form.

$$\max \{c^T x - \frac{1}{2} x^T D x : Ax \leq b,\ x \geq 0\} \tag{2.15}$$

with the same assumptions on the dimension of the problem and the input data as for problem (2.1). Observe that the only difference between problems (2.15) and (2.1) is the existence of inequality instead of equality constraints.

The dual of (2.15) has the form

$$\min \{b^T y + \frac{1}{2} x^T D x :\ A^T y - Dx \geq c,\ y \geq 0\}\ . \tag{2.16}$$

Using the fact that D is positive definite and applying the transformation $z = x - D^{-1}c$ we obtain the equivalent pair of primal and dual problems

$$\begin{aligned} \frac{1}{2} c^T D^{-1} c + \max\{-\frac{1}{2} z^T D z : Az + Iw &= b - AD^{-1}c, \\ Iz - Ix &= -D^{-1}c, x \geq 0, w \geq 0\} \end{aligned} \tag{2.17}$$

and

$$\frac{1}{2} c^T D^{-1} c + \min\{c^T z + b^T y + \frac{1}{2} z^T D z : A^T y + Dz - Is = 0, y \geq 0, s \geq 0\}\ . \tag{2.18}$$

Note that an optimal solution $(\bar{z}^T, \bar{y}^T, \bar{s}^T)$ for (2.18) provides an optimal solution $\bar{x} = \bar{z} + D^{-1}c$ for (2.15) and $(\bar{x}^T, \bar{y}^T)$ an optimal solution for (2.16).

The Karush-Kuhn-Tucker optimality conditions for the pair of primal-dual problems (2.15), (2.16) will have the form

$$\begin{pmatrix} A & 0 & I & 0 \\ D & A^T & 0 & -I \end{pmatrix} \begin{pmatrix} x \\ y \\ w \\ s \end{pmatrix} = \begin{pmatrix} b \\ c \end{pmatrix} \tag{2.19}$$

$$x, y, w, s \geq 0 \qquad \text{and} \qquad x^T s = 0, \quad y^T w = 0.$$

In this case, at each iteration the algorithm will detect at least one new variable y or s which has to equal zero in all optimal pairs of primal and dual solutions for (2.15) and (2.16). After each iteration one will perform Tardos' feasibility algorithm to detect a basic point from the linear system (2.19). The conditions $x^T s = 0$ and $y^T w = 0$ will then be checked. If satisfied the algorithm will terminate since an optimal pair of primal and dual solutions were found. Otherwise, another integration of QPA will be performed.

Instead of problem (2.7) (in the equality case), we will work with the problem of the form

$$\min \{c^T z + b^T y + \frac{1}{2} z^T D z : Dz + A^T y - Is = 0, y \geq 0, s \geq 0\}. \tag{2.20}$$

The algorithm will have the same form, except that now STEP 1 will read "Set $K_1 = \phi$ and $K_2 = \phi$", in STEP 2 the system $\{D^o x + A^{oT} y - E^o s = c^o,\ y \geq 0,\ s \geq 0\}$ will be equivalent to the system $\{Dx + A^T y - Is = c,\ y \geq 0,\ s \geq 0,\ s_i = 0$ for $i \in K_1$ and $y_j = 0$ for $j \in K_2\}$, and the polyhedron P^o will be

$$P^o = \{(x^T, y^T, w^T, s^T) : \begin{pmatrix} A & 0 & I & 0 \\ D & A^T & 0 & -I \end{pmatrix} \begin{pmatrix} x \\ y \\ w \\ s \end{pmatrix} = \begin{pmatrix} b \\ c \end{pmatrix}, x \geq 0, y \geq 0, w \geq 0,$$

$$s \geq 0,\ s_i = 0 \quad \text{for} \quad i \in K_1, \quad y_j = 0 \quad \text{for} \quad j \in K_2\}.$$

STEP 5 will now read: "Use Kozlov et al.'s algorithm to find an optimal solution

$(\bar{v}^T, \bar{z}^T)$ to

$$\max\,\{-\frac{1}{2}z^T Dz : D^{\circ T}v - Dz = \bar{c}, A^{\circ}v \leq \bar{b}, -E^{\circ T}v \leq \bar{a}\}$$

which is the dual of

$$\min\,\{\bar{c}^T z + \bar{b}^T y + \bar{a}^T s + \frac{1}{2}z^T Dz : D^0 z + A^{0T} y - E^0 s = 0,\ y \geq 0,\ s \geq 0\}\,.$$

Let

$$I = \{i : (-E^{0T}\bar{v})_i \leq \alpha a'_i - (2n+m)\Delta\},$$

and

$$J = \{j : (A^0\bar{v})_i \ \leq \alpha b'_j - (2n+m)\Delta\}.$$

Add the set I to K_1, the set J to K_2 and go to Step 2."

The complexity of the algorithm will be affected in the sense that at most $(n+m)$ iterations are now possible.

Chapter III

PROXIMITY AND SENSITIVITY RESULTS FOR QUADRATIC INTEGER PROGRAMMING

In a recent paper, Cook et al. [8], obtained many proximity results for integer linear programming problems with a fixed constraint matrix and varying objective function and right-hand side vectors. In this Chapter we will extend their main proximity results to quadratic integer programming problems of the form

$$\begin{aligned} \max \;\; & c^T x + x^T D x \\ \text{s.t.} \;\; & Ax \leq b \\ & x \text{ integer} \end{aligned} \tag{3.1}$$

where c and x are n-vectors, b is an m-vector, A is an integral $m \times n$ matrix and D is a negative semidefinite $n \times n$ matrix.

In the sequel we will assume that the set $\{x : \; Ax \leq b, x \text{ integer}\}$ is non-empty, that $\max\{c^T x + x^T Dx : \; Ax \leq b\}$ exists, and that D is rational.

As stated before, for any matrix A let $\Delta(A)$ denote the maximum of the absolute values of the determinants of the square submatrices of A .

For simplicity of exposition we first consider problem (3.1) with diagonal matrix D , i.e. the separable case. We start by showing in Theorem 3.1.1 that, in this case, for any optimal solution $\bar{z}$ for (3.1) there exists an optimal solution x^* for its continuous relaxation such that $\| \bar{z} - x^* \|_\infty \leq n\Delta(A)$. Also, if z is a feasible solution to (3.1) which is not optimal then we show in Theorem 3.1.3 that there exists another feasible solution $\bar{z}$ to (3.1) having greater objective function value and with $\| z - \bar{z} \|_\infty \leq n\Delta(A)$. With some additional assumptions we show in Theorem 3.2.2 that if $\bar{z}$ and $\bar{z}'$ are optimal solutions for (3.1) with right hand side vectors b and b' respectively then $\| \bar{z} - \bar{z}' \|_\infty \leq \alpha \| b - b' \|_1 + \beta$ where α and β are parameters which depend only on A, D and n . Finally, we show how to extend the above results to mixed-integer quadratic programs.

3.1. PROXIMITY RESULTS FOR SEPARABLE QUADRATIC INTEGER PROGRAMMING

Theorem 3.1.1, to follow, provides an upper bound on the distance between a pair of optimal solutions for problem (3.1) and its continuous relaxation respectively. The bound obtained depends only on the number of variables n and the largest absolute subdeterminant of the matrix A , and is independent of the vectors b, c and the matrix D .

THEOREM 3.1.1 Let A be an integral $m \times n$ matrix, D a diagonal negative semidefinite $n \times n$ matrix, b an m-vector and c an n-vector such that the set $\{x : Ax \leq b \text{ , } x \text{ integer}\}$ is non-empty and $\max\{c^T x + x^T Dx : \; Ax \leq b\}$ exists. Then

(i) For each optimal solution $\bar{x}$ for $\max\{c^T x + x^T Dx : \; Ax \leq b\}$ there exists an optimal solution z^* for $\max\{c^T x + x^T Dx : \; Ax \leq b, x \text{ integer}\}$ with $\| z^* - \bar{x} \|_\infty \leq n\Delta(A)$, and

(ii) For each optimal solution $\bar{z}$ for $\max\{c^T x + x^T Dx : \; Ax \leq b, x \text{ integer}\}$ there exists an optimal solution x^* for $\max\{c^T x + x^T Dx : Ax \leq b\}$ with $\| x^* - \bar{z} \|_\infty \leq n\Delta(A)$.

PROOF Let $\bar{z}$ and $\bar{x}$ be optimal solutions for problems (3.1) and its continuous relaxation respectively. Further assume without loss of generality that the first ℓ components of $\bar{x} - \bar{z}$ are nonpositive and the last $n - \ell$ components are positive. Partition the rows of the matrix A into submatrices A_1 and A_2 such that $A_1\bar{x} \geq A_1\bar{z}$ and $A_2\bar{x} < A_2\bar{z}$. Consider the cone $C = \{x = (x_1, \ldots, x_\ell, x_{\ell+1}, \ldots, x_n) : A_1 x \geq 0,\ A_2 x \leq 0, x^1 \equiv (x_1, \ldots, x_\ell) \leq 0, x^2 \equiv (x_{\ell+1}, \ldots, x_n) \geq 0\}$. Clearly, C is nonempty since $\bar{x} - \bar{z} \in C$. Let G be a finite set of integral vectors which generate C . By Cramer's rule for each $g \in G$, $\| g \|_\infty \leq \Delta(A') = \Delta(A)$, where $A' = \begin{pmatrix} A \\ -I_1 \ \ 0 \\ 0 \ \ I_2 \end{pmatrix}$ and I_1 (resp., I_2) is an $\ell \times \ell$ (resp., $(n-\ell) \times (n-\ell)$) identity matrix. Now, since $\bar{x} - \bar{z} \in C$ there exists a finite set $\{g_1, \ldots, g_t\} \subseteq G$ and scalars $\alpha_i \geq 0, i = 1, \ldots, t$ such that $\bar{x} - \bar{z} = \sum_{i=1}^t \alpha_i g_i$.

In order to prove part (i) we will define the vector $z^* = \bar{z} + \sum_{i=1}^t \lfloor \alpha_i \rfloor g_i$. We will show first that z^* is feasible and optimal for (3.1). Coupling with the fact that $z^* = \bar{x} - \sum_{i=1}^t (\alpha_i - \lfloor \alpha_i \rfloor) g_i$ we will obtain that $\| z^* - \bar{x} \|_\infty \leq t\Delta(A) \leq n\Delta(A)$. The proof of the feasibility of z^* is identical to the proof given in Cook et al., but we will state it for the sake of completeness. Observe that z^* is integral and satisfies $A_1 z^* = A_1(\bar{x} - \sum_{i=1}^t (\alpha_i - \lfloor \alpha_i \rfloor g_i)) \leq A_1\bar{x}$, $A_2 z^* = A_2(\bar{z} + \sum_{i=1}^t \lfloor \alpha_i \rfloor g_i) \leq A_2\bar{z}$ and therefore $Az^* \leq b$. To prove optimality, recall (see e.g. Dorn [13]) that the dual of the continuous relaxation of (3.1) is given by

$$\min\{u^T b - x^T Dx : u^T A - 2x^T D = c^T, u \geq 0\} \tag{3.2}$$

with complementary slackness conditions $u^T(b - Ax) = 0$. Now, if we partition b^T to (b_1^T, b_2^T) corresponding to the partition of A to (A_1, A_2) we have $A_2\bar{x} < A_2\bar{z} \leq$

b_2 , and thus it follows from the above condition that for a vector $\bar{u}^T = (\bar{u}_1^T, \bar{u}_2^T)$ of optimal dual variables for (3.2) we have $\bar{u}_2 = 0$ and $\bar{u}_1^T A_1 = c^T + 2\bar{x}^T D$. Also, since for every $y \in C,\ A_1 y \geq 0$ we have

$$c^T y + 2\bar{x}^T D y \geq 0. \tag{3.3}$$

Finally, using (3.3) and the fact that $\sum_{i=1}^t \lfloor \alpha_i \rfloor g_i \in C$ we obtain

$$\begin{aligned}
c^T z^* + z^{*T} D z^* &= c^T(\bar{z} + \sum_{i=1}^t \lfloor \alpha_i \rfloor g_i) + \bar{z}^T D \bar{z} + 2\bar{z}^T D(\sum_{i=1}^t \lfloor \alpha_i \rfloor g_i) \\
&+ (\sum_{i=1}^t \lfloor \alpha_i \rfloor g_i)^T D(\sum_{i=1}^t \lfloor \alpha_i \rfloor g_i) \\
&= c^T \bar{z} + \bar{z}^T D \bar{z} + c^T(\sum_{i=1}^t \lfloor \alpha_i \rfloor g_i) + 2\bar{x}^T D(\sum_{i=1}^t \lfloor \alpha_i \rfloor g_i) \\
&- 2(\sum_{i=1}^t \alpha_i g_i)^T D(\sum_{i=1}^t \lfloor \alpha_i \rfloor g_i) + (\sum_{i=1}^t \lfloor \alpha_i \rfloor g_i)^T D(\sum_{i=1}^t \lfloor \alpha_i \rfloor g_i) \\
&\geq c^T \bar{z} + \bar{z}^T D \bar{z} + (\sum_{i=1}^t (\lfloor \alpha_i \rfloor - 2\alpha_i) g_i)^T D(\sum_{i=1}^t \lfloor \alpha_i \rfloor g_i).
\end{aligned}$$

It remains to show that $(\sum_{i=1}^t (\lfloor \alpha_i \rfloor - 2\alpha_i) g_i)^T D(\sum_{i=1}^t \lfloor \alpha_i \rfloor g_i) \geq 0$. To that end recall that $g_i \in C$ and hence can be partitioned into (g_i^1, g_i^2) where g_i^1 contains the first ℓ components and g_i^2 the last $n - \ell$ components of g_i . Clearly, $g_i^1 \leq 0,\ g_i^2 \geq 0$. Diagonality and negative semidefiniteness of D implies

$$g_i^T D g_j \leq 0 \quad \text{for all} \quad i, j = 1, \ldots, t\ .$$

Coupled with the fact that $\lfloor \alpha_i \rfloor - \alpha_i \leq 0$ and $\lfloor \alpha_i \rfloor \geq 0$ for all i we obtain

$$(\textstyle\sum_{i=1}^t (\lfloor \alpha_i \rfloor - 2\alpha_i) g_i)^T D(\sum_{i=1}^t \lfloor \alpha_i \rfloor g_i) = \sum_{i=1}^t (\lfloor \alpha_i \rfloor - 2\alpha_i) \lfloor \alpha_i \rfloor g_i^T D g_i$$

$$+ \textstyle\sum_{i=1}^t \sum_{j>i}^t ((\lfloor \alpha_i \rfloor - 2\alpha_i) \lfloor \alpha_j \rfloor + (\lfloor \alpha_j \rfloor - 2\alpha_j) \lfloor \alpha_i \rfloor) g_i^T D g_j \geq 0\ .$$

Thus, $c^T z^* + z^{*T} D z^* \geq c^T \bar{z} + \bar{z}^T D \bar{z}$ and part (i) is valid. Also, the optimality of $\bar{z}$ and z^* implies

$$c^T(\sum_{i=1}^{t} \lfloor \alpha_i \rfloor g_i) + (2\bar{z} + \sum_{i=1}^{t} \lfloor \alpha_i \rfloor g_i)^T D(\sum_{i=1}^{t} \lfloor \alpha_i \rfloor g_i) = 0. \tag{3.4}$$

In order to prove (ii) define first the vector

$$x^* = \bar{x} - \sum_{i=1}^{t} \lfloor \alpha_i \rfloor g_i = \bar{z} + \sum_{i=1}^{t} (\alpha_i - \lfloor \alpha_i \rfloor) g_i \tag{3.5}$$

which will be shown to be feasible and optimal for the continuous relaxation of (3.1). The feasibility of x^* can be shown in a similar manner to the way it was done for z^*. To show optimality observe that

$$\begin{aligned}
c^T x^* + x^{*T} D x^* &= c^T \bar{x} - c^T(\sum_{i=1}^{t} \lfloor \alpha_i \rfloor g_i) + \bar{x}^T D \bar{x} - 2\bar{x}^T D(\sum_{i=1}^{t} \lfloor \alpha_i \rfloor g_i) \\
&+ (\sum_{i=1}^{t} \lfloor \alpha_i \rfloor g_i)^T D(\sum_{i=1}^{t} \lfloor \alpha_i \rfloor g_i) \\
&= c^T \bar{x} + \bar{x}^T D \bar{x} - (c^T(\sum_{i=1}^{t} \lfloor \alpha_i \rfloor g_i) + (2\bar{z} + \sum_{i=1}^{t} \lfloor \alpha_i \rfloor g_i)^T D(\sum_{i=1}^{t} \lfloor \alpha_i \rfloor g_i)) \\
&+ 2(\sum_{i=1}^{t} (\lfloor \alpha_i \rfloor - \alpha_i) g_i)^T D(\sum_{i=1}^{t} \lfloor \alpha_i \rfloor g_i) = c^T \bar{x} + \bar{x}^T D \bar{x} \\
&+ 2(\sum_{i=1}^{t} (\lfloor \alpha_i \rfloor - \alpha_i) g_i)^T D(\sum_{i=1}^{t} \lfloor \alpha_i \rfloor g_i) \geq c^T \bar{x} + \bar{x}^T D \bar{x},
\end{aligned}$$

where the equalities follow from (3.5) and (3.4) and the last inequality follows from the fact that

$$(\sum_{i=1}^{t} (\lfloor \alpha_i \rfloor - \alpha_i) g_i)^T D(\sum_{i=1}^{t} \lfloor \alpha_i \rfloor g_i) \geq 0.$$

Finally, from (3.5) $\| x^* - \bar{z} \|_\infty \leq t\Delta(A) \leq n\Delta(A)$.

□

Using the example provided by L. Lovász for the linear case (see Schrijver [43], page 241) it can be verified that the bound is tight. Observe that if lower and upper bounds on the variables x_i are known one can naturally improve the bound on the difference between an integral and continuous optimal solution. Indeed, if $P = \{x : Ax \le b,\ \ell_i \le x_i \le u_i\}$ then denote by $B = \min\{\ n\Delta(A),\ \max\{|\ u_i - \ell_i\ | : i = 1, ..., n\}\ \}$. Clearly B is a valid bound. Trivially, in the case of a $0-1$ quadratic programming problem we have $B = 1$.

As a consequence of Theorem 3.1.1 we obtain Corollary 3.1.2 which provides a bound on the difference between the optimal objective function value of (3.1) and its continuous relaxation.

COROLLARY 3.1.2 Let A be an integral $m \times n$ matrix and b an m-vector. Let $P = \{x :\ Ax \le b\}$ be a nonempty bounded polyhedron having an integral point. Then,

$$\max\{c^T x + x^T Dx :\ Ax \le b\} - \max\{c^T x + x^T Dx :\ Ax \le b, x \text{ integer}\}$$

$$\le n\Delta(A)(\|\ c\ \|_1 + 2n\Delta(A \mid b)\ \|\ D\ \|_\infty).$$

PROOF Denote by $\bar{z}$ (resp., QI) an optimal solution (resp., the optimal objective function value) for $\max\{c^T x + x^T Dx :\ Ax \le b, x \text{ integer}\}$ and by $\bar{x}$ (resp., QC) an optimal solution (resp., the optimal objective function value) for its continuous relaxation with $\|\ \bar{x} - \bar{z}\ \|_\infty \le n\Delta(A)$ which clearly exists by Theorem 3.1.1. Then,

$$QC - QI = c^T\bar{x} + \bar{x}^T D\bar{x} - c^T\bar{z} - \bar{z}^T D\bar{z}$$

$$= c^T(\bar{x} - \bar{z}) + (\bar{x} + \bar{z})^T D(\bar{x} - \bar{z})$$

$$\le \|\ c\ \|_1 \|\ \bar{x} - \bar{z}\ \|_\infty + \|\ \bar{x} + \bar{z}\ \|_1 \|\ D(\bar{x} - \bar{z})\ \|_\infty$$

$$\le \|\ c\ \|_1 \|\ \bar{x} - \bar{z}\ \|_\infty + \|\ \bar{x} + \bar{z}\ \|_1 \|\ D\ \|_\infty \|\ \bar{x} - \bar{z}\ \|_\infty$$

where $\| D \|_\infty = \max\{|d_{ii}| : i = 1, \ldots, n\}$. Now since for a vector x , $\| x \|_1 \leq n \| x \|_\infty$ and for a pair of vectors x, y $\| x + y \|_\infty \leq \| x \|_\infty + \| y \|_\infty$, we have $\| \bar{x} + \bar{z} \|_1 \leq n(\| \bar{x} \|_\infty + \| \bar{z} \|_\infty) \leq n(\Delta(A \mid b) + \Delta(A \mid b))$ by the boundedness of the polyhedrons. Therefore,

$$
\begin{aligned}
QC - QI &\leq \| \bar{x} - \bar{z} \|_\infty (\| c \|_1 + 2n\Delta(A \mid b) \| D \|_\infty) \\
&\leq n\Delta(A)(\| c \|_1 + 2n\Delta(A \mid b) \| D \|_\infty) .
\end{aligned}
$$

□

Theorem 3.1.3, which follows, will show that for any integral solution z of $Ax \leq b$ which is not optimal for (3.1), there exists an integral solution $\bar{z}$ of $Ax \leq b$ which has greater objective function value and with $\| \bar{z} - z \|_\infty \leq n\Delta(A)$.

THEOREM 3.1.3 For each integral solution z of $Ax \leq b$, either z is an optimal solution for (3.1) with a diagonal matrix D or there exists an integral solution $\bar{z}$ of $Ax \leq b$ with $\| z - \bar{z} \|_\infty \leq n\Delta(A)$ and $c^T\bar{z} + \bar{z}^T D\bar{z} > c^T z + z^T Dz$.

PROOF Let z be an integral solution of $Ax \leq b$ which is not optimal for (3.1) with a diagonal matrix D . Then there exists an integral solution z^* of $Ax \leq b$ with $c^T z^* + z^{*T} Dz^* > c^T z + z^T Dz$. As in the proof of Theorem 3.1.1, without loss of generality, assume that the first ℓ components of $z^* - z$ are nonpositive and that the last $n - \ell$ components are positive. Partition the rows of A into submatrices A_1 and A_2 such that $A_1 z^* \geq A_1 z$, $A_2 z^* < A_2 z$ and consider the cone $C = \{x = (x_1, \ldots, x_\ell, x_{\ell+1}, \ldots, x_n) : A_1 x \geq 0 , A_2 x \leq 0 , x^1 \equiv (x_1, \ldots, x_\ell) \leq 0 , x_2 \equiv (x_{\ell+1}, \ldots, x_n) \geq 0\}$. Since $z^* - z \in C$, we can write $z^* - z = \sum_{i=1}^t \alpha_i g_i$ where $\alpha_i \geq 0$ for $i = 1, \ldots, t$ and $\{g_1, \ldots, g_t\}$ is a subset of integral generators of C . By Cramer's rule $\| g_i \|_\infty \leq \Delta(A)$ for every i . If $\alpha_i \geq 1$ for some i , then $z + g_i = z^* - \sum_{j=1_{j \neq i}}^t \alpha_j g_j - (\alpha_i - 1)g_i$ is an integer feasible solution of $Ax \leq b$.

Now, if in addition $c^T g_i + (2z+g_i)^T D g_i > 0$, then $c^T(z+g_i) + (z+g_i)^T D(z+g_i) > c^T z + z^T D z$ and by choosing $\bar{z} = z + g_i$ we are done. Thus, let us assume that for all i for which $\alpha_i \geq 1$ we have

$$c^T g_i + (2z+g_i)^T D g_i \leq 0. \tag{3.6}$$

Next, define the integral vector

$$\bar{z} = z^* - \sum_{i=1}^{t} \lfloor \alpha_i \rfloor g_i = z + \sum_{i=1}^{t} (\alpha_i - \lfloor \alpha_i \rfloor) g_i$$

which is a feasible solution of $Ax \leq b$ and satisfies $\| \bar{z} - z \|_\infty \leq t\Delta(A) \leq n\Delta(A)$. To complete the proof we will show next that the objective function value at $\bar{z}$ is larger than at z^*. Indeed,

$$c^T\bar{z} + \bar{z}^T D\bar{z} = c^T z^* + z^{*T} D z^* - c^T(\textstyle\sum_{i=1}^t \lfloor \alpha_i \rfloor g_i) - 2z^{*T} D(\sum_{i=1}^t \lfloor \alpha_i \rfloor g_i)$$

$$+(\textstyle\sum_{i=1}^t \lfloor \alpha_i \rfloor g_i)^T D(\sum_{i=1}^t \lfloor \alpha_i \rfloor g_i) = c^T z^* + z^{*T} D z^*$$

$$-(c^T(\textstyle\sum_{i=1}^t \lfloor \alpha_i \rfloor g_i) + (2z + 2\sum_{i=1}^t \alpha_i g_i - \sum_{i=1}^t \lfloor \alpha_i \rfloor g_i)^T D(\sum_{i=1}^t \lfloor \alpha_i \rfloor g_i))$$

$$= c^T z^* + z^{*T} D z^* - (c^T(\textstyle\sum_{i=1}^t \lfloor \alpha_i \rfloor g_i) + 2z^T D(\sum_{i=1}^t \lfloor \alpha_i \rfloor g_i))$$

$$-(\textstyle\sum_{i=1}^t (2\alpha_i - \lfloor \alpha_i \rfloor) g_i)^T D(\sum_{i=1}^t \lfloor \alpha \rfloor g_i) = c^T z^* + z^{*T} D z^*$$

$$-(c^T(\textstyle\sum_{i=1}^t \lfloor \alpha_i \rfloor g_i) + 2z^T D(\sum_{i=1}^t \lfloor \alpha_i \rfloor g_i)) - \sum_{i=1}^t (2\alpha_i - \lfloor \alpha_i \rfloor)\lfloor \alpha_i \rfloor g_i^T D g_i$$

$$-\textstyle\sum_{i=1}^t \sum_{j>i}^t ((2\alpha_i - \lfloor \alpha_i \rfloor)\lfloor \alpha_j \rfloor + (2\alpha_j - \lfloor \alpha_j \rfloor)\lfloor \alpha_i \rfloor) g_i^T D g_j \geq c^T z^* + z^{*T} D z^*$$

$$-(c^T(\textstyle\sum_{i=1}^t \lfloor \alpha_i \rfloor g_i) + 2z^T D(\sum_{i=1}^t \lfloor \alpha_i \rfloor g_i)) - \sum_{i=1}^t (2\alpha_i - \lfloor \alpha_i \rfloor)\lfloor \alpha_i \rfloor g_i^T D g_i$$

where the inequality follows from the fact that $g_i \in C$, $\lfloor \alpha_i \rfloor \geq 0$ and $2\alpha_i - \lfloor \alpha_i \rfloor \geq 0$ for all i which in turn implies $((2\alpha_i - \lfloor \alpha_i \rfloor)\lfloor \alpha_j \rfloor + (2\alpha_j - \lfloor \alpha_j \rfloor)\lfloor \alpha_i \rfloor) g_i^T D g_j \leq 0$.

Therefore

$$c^T\bar{z} + \bar{z}^T D\bar{z} \geq c^T z^* + z^{*T} Dz^* - c^T(\sum_{i=1}^{t} \lfloor \alpha_i \rfloor g_i)$$
$$- \sum_{i=1}^{t} \lfloor \alpha_i \rfloor (2z + (2\alpha_i - \lfloor \alpha_i \rfloor) g_i)^T D g_i \ .$$

Now, since for $\alpha_i < 1$, $\lfloor \alpha_i \rfloor = 0$, the only nonvanishing terms in the above summation correspond to $\alpha_i \geq 1$. Further, since D is negative semidefinite and $2\alpha_i - \lfloor \alpha_i \rfloor = \alpha_i + (\alpha_i - \lfloor \alpha_i \rfloor) \geq \alpha_i \geq 1$, $(2\alpha_i - \lfloor \alpha_i \rfloor) g_i^T D g_i \leq g_i^T D g_i$. This implies, using (3.6), that

$$c^T\bar{z} + \bar{z}^T D\bar{z} \geq c^T z^* + z^{*T} Dz^* - \sum_{\substack{i=1 \\ \alpha_i \geq 1}}^{t} \lfloor \alpha_i \rfloor (c^T g_i + (2z + g_i)^T D g_i)$$
$$\geq c^T z^* + z^{*T} D z^* > c^T z + z^T D z \ .$$

□

Observe that the above Theorem assures finite "test set" for problem (3.1). In other words, one has to check only a finite number (which depends only on the constraint matrix and the number of variables) of integral vectors in order to obtain a better solution or to verify the optimality of the current solution. Although the set is finite, it might require the exponential (in n) number of comparisons in order to get better solution.

3.2. SENSITIVITY FOR QUADRATIC INTEGER PROGRAMMING : THE RIGHT HAND SIDE CASE

In this Section we show that if $\bar{z}$ and $\bar{z}'$ are optimal solutions for (3.1) with negative definite matrix D , constraint matrix A of full row rank and right hand side vectors b and b' respectively, then $\| \bar{z} - \bar{z}' \|_\infty \leq \alpha \| b - b' \|_1 + \beta$ where

α and β are parameters which depend only on A, D and n. Observe that restricting D to be negative definite implies that the continuous relaxation of the quadratic integer program has a unique optimal solution whenever the polyhedron $P = \{x : Ax \leq b\}$ is nonempty.

For simplicity of exposition Theorem 3.2.2, to follow, will be stated and proved for the separable case. The general case will be discussed in Section 3.3. Theorem 3.2.2 is using a special case of Theorem 2.1 in [11] in which changes in the linear cost coefficients are considered. For completeness we will state the part of Theorem 2.1 in [11] we use.

LEMMA 3.2.1 Let A be an $m \times n$ matrix, b an m-vector, c and c' n-vectors and $D = diag(d_i)$ an $n \times n$ diagonal negative definite matrix. Let $\bar{x}$ (resp., $\bar{x}'$) be the optimal solution for $\max\{c^T x + x^T D x \ : Ax \leq b\}$ (resp., $\max\{c'^T x + x^T D x \ : Ax \leq b\}$) . Then

$$\| \bar{x} - \bar{x}' \|_\infty \leq \frac{1}{2K} \| c - c' \|_1$$

where $K = -\max\{d_i \ : i = 1, \ldots, n\} > 0$.

PROOF See Daniel [11], Theorem 2.1 .

□

THEOREM 3.2.2 Let $\{Ax \leq b\} \equiv \{A_E x = b_E, A_I x \leq b_I\}$ and $\{Ax \leq b'\} \equiv \{A_E x = b'_E, A_I x \leq b'_I\}$ have integral solutions where A is an integral $m \times n$ matrix of full row rank and b, b' are m-vectors. Let c be an n-vector and $D = diag(d_i)$ a diagonal, negative definite $n \times n$ matrix. Then

(i) For every optimal solution $\bar{z}$ for $\max\{c^T x + x^T D x : \ Ax \leq b, x \text{ integral}\}$ and every optimal solution $\bar{z}'$, for $\max\{c^Tx + x^TDx : Ax \leq b', x \text{ integral}\}$ we have $\| \bar{z} - \bar{z}' \|_\infty \leq n\Delta(A)(H(D) \| b - b' \|_1 + 2)$ where $H(D) = \max\{|d_i| : i = 1, \ldots, n\} / \min\{|d_i| : i = 1, \ldots, n\}$.

(ii) Assume, in addition, that D is integral. If $f(b)$ (resp., $f(b')$) is the optimal value for $\max\{c^Tx + x^TDx : Ax \leq b, x \text{ integral}\}$ (resp., $\max\{c^Tx + x^TDx : Ax \leq b', x \text{ integral}\}$), then $|f(b) - f(b')| \leq n\Delta(A)(\| c \|_1 + n \| D \|_\infty (2n\Delta(A) + \Delta(\tilde{A}|_c^b) + \Delta(\tilde{A}|_c^{b'})))(H(D) \| b - b' \|_1 + 2)$ where $\tilde{A} = \begin{pmatrix} A & 0 \\ -2D & A^T \end{pmatrix}$.

PROOF We will start by determining an upper bound on the difference between the optimal solutions to the corresponding continuous quadratic programs. To that end let us first write the dual of the continuous relaxation of (3.1) in primal form

$$
\begin{aligned}
&- \max\{(0^T, -b^T)\begin{pmatrix} x \\ u \end{pmatrix} + (x^T, u^T)\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ 0 \end{pmatrix}\} \\
&\text{s.t.} \begin{pmatrix} -2D & A^T \\ 2D & -A^T \\ 0 & -I \end{pmatrix}\begin{pmatrix} x \\ u \end{pmatrix} \leq \begin{pmatrix} c \\ -c \\ 0 \end{pmatrix}.
\end{aligned}
\tag{3.7}
$$

Now, we will treat changes in the right hand side of (3.1) by considering changes in part of the linear cost coefficient of (3.7). To that end, recall that if $\bar{x}$ is an optimal solution for $\max\{c^Tx + x^TDx : Ax \leq b\}$ then $\bar{x}$ is also an optimal solution for the linear programming problem $\max\{(c^T + 2\bar{x}^TD)x : Ax \leq b\}$ see e.g. Dorn [13]. Thus if $(\bar{x}, \bar{u})$ and $(\bar{x}', \bar{u}')$ are, respectively, optimal of (3.7) and the quadratic programming problem obtained from (3.7) by replacing b by b' then, for every feasible solution (x, u) of (3.7) we have

$$
\begin{aligned}
&- u^Tb + 2\bar{x}^TDx \leq -\bar{u}^Tb + 2\bar{x}^TD\bar{x} \\
&- u^Tb' + 2\bar{x}'^TDx \leq -\bar{u}'^Tb' + 2\bar{x}'^TD\bar{x}'.
\end{aligned}
\tag{3.8}
$$

Substituting $(\bar{x}', \bar{u}')$ (resp., $(\bar{x}, \bar{u})$) in the first (resp., second) inequality of (3.8) adding and rearranging terms results with

$$2(\bar{x}' - \bar{x})(-D)(\bar{x}' - \bar{x}) \leq (\bar{u}' - \bar{u})^T(b - b') \leq \| \bar{u}' - \bar{u} \|_\infty \| b - b' \|_1. \tag{3.9}$$

Now, since both $(\bar{x}, \bar{u})$ and $(\bar{x}', \bar{u}')$ are feasible for (3.7), $(\bar{x}' - \bar{x}, \bar{u}' - \bar{u}) \equiv (v, w)$ satisfies $2Dv - A^T w = 0$ and hence

$$A^T(\bar{u}' - \bar{u}) = 2D(\bar{x}' - \bar{x}).$$

Using the fact that A^T is integral and has full column rank, Cramer's rule, and a property of determinants we obtain

$$\| \bar{u}' - \bar{u} \|_\infty \leq \Delta(A^T \mid 2D(\bar{x}' - \bar{x})) \leq 2n\Delta(A) \| D \|_\infty \| \bar{x}' - \bar{x} \|_\infty . \tag{3.10}$$

Now let $K \equiv \min\{-d_i : i = 1, ..., n\} = -\max\{d_i : i = 1, ..., n\}$. Substituting (3.10) into (3.9) and observing that $K > 0$ results with

$$2K \| \bar{x}' - \bar{x} \|_\infty^2 \leq 2K \| \bar{x}' - \bar{x} \|_2^2 \leq 2n\Delta(A) \| D \|_\infty \| \bar{x}' - \bar{x} \|_\infty \| b' - b \|_1 \tag{3.11}$$

or

$$\| \bar{x}' - \bar{x} \|_\infty \leq \frac{n\Delta(A) \| D \|_\infty}{K} \| b' - b \|_1 = n\Delta(A)H(D) \| b' - b \|_1 .$$

Using the fact that the continuous relaxation of our quadratic integer program has a unique solution and applying part (ii) of Theorem 3.1 it follows that for every pair of optimal solutions $\bar{z}$ and $\bar{z}'$ for the quadratic integer programming problem with right hand side vectors b and b' respectively

$$\| \bar{z} - \bar{z}' \|_\infty \leq \| \bar{z} - \bar{x} \|_\infty + \| \bar{x} - \bar{x}' \|_\infty + \| \bar{x}' - \bar{z}' \|_\infty \leq$$
$$\leq n\Delta(A)((H(D) \| b - b' \|_1 + 2).$$

To complete the proof we will show that part (ii) of the theorem follows from the above bound, the existence of the optimal solutions and Cramer's rule.

Denote by $\tilde{A} = \begin{pmatrix} A & 0 \\ -2D & A^T \end{pmatrix}$. Observe that an optimal pair of primal and dual solutions $(\bar{x}, \bar{u})$ for the continuous relaxation of (3.1) is a basic solution of a polyhedron defined by $Ax \leq b$, $A^T u - 2Dx = c$. Now, by Cramer's rule and the integrality of A and D we get

$$\| \bar{x} \|_\infty \leq \| \begin{pmatrix} \bar{x} \\ \bar{u} \end{pmatrix} \|_\infty \leq \Delta(\tilde{A} \mid_c^b) \, .$$

Thus, for an optimal integral solution we have (by Theorem 3.1)

$$\| \bar{z} \|_\infty \leq \| \bar{z} - \bar{x} \|_\infty + \| \bar{x} \|_\infty \leq n\Delta(A) + \Delta(\tilde{A} \mid_c^b) \, .$$

Therefore

$$\| \bar{z} \|_\infty + \| \bar{z}' \|_\infty \leq 2n\Delta(A) + \Delta(A \mid_c^b) + \Delta(A \mid_c^{b'}) \, .$$

Now,

$$| f(b) - f(b') | = | c^T(\bar{z} - \bar{z}') + (\bar{z} - \bar{z}')^T D(\bar{z} + \bar{z}') | \leq$$

$$\leq \| c \|_1 \| \bar{z} - \bar{z}' \|_\infty + n \| D \|_\infty \| \bar{z} + \bar{z}' \|_\infty \| \bar{z} - \bar{z}' \|_\infty \leq$$

$$\leq n\Delta(A)(\| c \|_1 + n \| D \|_\infty (2n\Delta(A) + \Delta(\tilde{A} \mid_c^b |) + \Delta(\tilde{A} \mid_c^{b'} |))(H(D) \| b - b' \|_1 + 2).$$

$\square$

3.3. EXTENSION TO NONSEPARABLE QUADRATIC MIXED-INTEGER PROGRAMMING

The results given so far can be easily extended to separable quadratic mixed-integer problems of the form

$$\begin{aligned} \max \quad & c_1^T x + c_2^T y + x^T D_1 x + y^T D_2 y \\ \text{s.t.} \quad & Ax + By \le b \qquad\qquad (3.12) \\ & x \ \text{ integer} \end{aligned}$$

where c_1 and x are k–vectors, c_2 and y are $(n-k)$–vectors, A (resp., B) is an $m \times k$ (resp., $m \times (n-k)$) integer matrix, D_1 (resp., D_2) is a $k \times k$ (resp., $(n-k) \times (n-k)$) diagonal, negative semidefinite matrix and b is an m-vector. For example, Theorem 3.1.1 for the mixed–integer case should read "...

(i) For each optimal solution $(\bar{x}^T, \bar{y}^T)$ for the continuous relaxation of (3.12) there exists an optimal solution $(\hat{z}^T, \hat{w}^T)$ for (3.12) with $\| (\hat{z}^T, \hat{w}^T) - (\bar{x}^T, \bar{y}^T) \|_\infty \le n\Delta(A \mid B)$, and

(ii) For each optimal solution $(\bar{z}^T, \bar{w}^T)$ for problem (3.12) there exists an optimal solution $(\hat{x}^T, \hat{y}^T)$ for its continuous relaxation with $\| (\hat{x}^T, \hat{y}^T) - (\bar{x}^T, \bar{y}^T) \|_\infty \le n\Delta(A \mid B)$."

Consequently, Corollary 3.1.2 and Theorem 3.1.3 have analogous statements for the mixed-integer case and all the proofs follow in a straightforward way from the proofs given for the pure integer case. Unfortunately, Theorem 3.1.3 loses its significance in the mixed-integer case because it does not imply the existence of a finite "test set" any longer.

The validity of the theory for the mixed-integer case allows us to consider a broader class of quadratic integer programming problems namely when the matrix D in (3.1) is not necessarily diagonal. This since D can be diagonalized (see Section 1.4) using an $\ell \times n$ matrix B (where ℓ is the rank of D) such that

$D = B^T CB$, with C being a diagonal negative definite matrix. Recall here that, without loss of generality, we can assume that $D = \begin{pmatrix} D' & 0 \\ 0 & 0 \end{pmatrix}$, where D' is an $\ell \times \ell$ negative definite matrix. Using the diagonalization for definite matrices proposed in [48], we obtain that $D' = B'^T CB'$ or $D = B^T CB = (B', 0)^T C(B', 0)$ without dealing with square roots, i.e. no irrationalities will occur. Thus, problem (3.1) can be equivalently written as

$$\begin{aligned} \max \quad & c^T x + x_1^T C x_1 \\ \text{s.t.} \quad & Ax \quad \leq b \\ & Bx - Ix_1 = 0 \\ & x \ \text{ integer} \end{aligned} \tag{3.13}$$

or

$$\begin{aligned} \max \quad & (c^T, 0^T) \begin{pmatrix} x \\ x_1 \end{pmatrix} + (x^T, x_1^T) \begin{pmatrix} 0 & 0 \\ 0 & C \end{pmatrix} \begin{pmatrix} x \\ x_1 \end{pmatrix} \\ \text{s.t.} \quad & \bar{A} \begin{pmatrix} x \\ x_1 \end{pmatrix} \equiv \begin{pmatrix} A & 0 \\ B & -I \\ -B & I \end{pmatrix} \begin{pmatrix} x \\ x_1 \end{pmatrix} \leq \begin{pmatrix} b \\ 0 \\ 0 \end{pmatrix} \\ & x \text{ integer .} \end{aligned} \tag{3.14}$$

Problem (3.14) is a mixed-integer, separable, quadratic programming problem for which the theory developed above is valid. Observe, however, that the bounds should be expressed in terms of $\bar{n}$ and $\Delta(\bar{A})$, where $\bar{n} = n + \ell$ and $\bar{A}$ as defined in (3.14). Also, if B is not integral, we will have to multiply $Bx - Ix_1 = 0$ by a large enough constant which will not restrict generality but will enlarge the bound. Note, however, that the size of a constant might not be bounded by a polynomial in the size of A .

A similar transformation can be carried out for nonseparable mixed-integer quadratic programs. The stability of mixed-integer quadratic minimization programs in the absence of boundedness on the feasible region and convexity on the objective function was studied by Bank and Hansel [2].

In order to generalize Theorem 3.2.2 to the nonseparable case observe that the proof of Theorem 3.2.2 up to equation (3.11) is valid without the diagonality assumption on the (negative definite) matrix D. However, in this case the constant K will be the smallest absolute eigenvalue of D. In order to obtain the bound on the difference between the two optimal integral solutions $\bar{z}$ and $\bar{z}'$ we used Theorem 3.1.1 in the Proof of Theorem 3.2.2 and, therefore, required separability of the objective function. At this point we can transform problem (3.1) to an equivalent separable mixed–integer problem (3.14). Using the fact that the optimal solutions $\bar{x}$ and $\bar{x}'$ to the continuous relaxation of (3.1) with right hand side vectors b and b', respectively, are unique (since they are optimal solutions for strictly convex quadratic problems) and that the matrix B obtained in the diagonalization of D is nonsingular (and, as stated before, can be assumed to be integral), we can obtain the bound on the difference between the optimal solutions $\bar{z}$ and $\bar{z}'$ for (3.1) with right hand side vectors b and b' respectively, as follows :

$$\| \bar{z} - \bar{z}' \|_\infty \leq \| \binom{\bar{z}}{B\bar{z}} - \binom{\bar{z}'}{B\bar{z}'} \|_\infty$$

$$\leq \| \binom{\bar{z}}{B\bar{z}} - \binom{\bar{x}}{B\bar{x}} \|_\infty + \| \binom{\bar{x}}{B\bar{x}} - \binom{\bar{x}'}{B\bar{x}'} \|_\infty + \| \binom{\bar{x}'}{B\bar{x}'} - \binom{\bar{z}'}{B\bar{z}'} \|_\infty$$

$$\leq 2(2n)\Delta(\bar{A}) + (1+ \| B \|_\infty) \| \bar{x} - \bar{x}' \|_\infty$$

$$\leq (1+ \| B \|_\infty)\Delta(A)H(C) \| b - b' \|_1 + 4n\Delta(\bar{A})$$

where $\bar{A}$, B and C are given in (3.14) and $\binom{\bar{z}}{B\bar{z}}$ (resp., $\binom{\bar{z}'}{B\bar{z}'}$) are optimal for (3.14) with right hand side vectors b (resp., b').

Chapter IV

NONLINEAR INTEGER PROGRAMS: SENSITIVITY ANALYSIS FOR BRANCH AND BOUND

Recently, Schrage and Wolsey [42] studied the effect of small changes in the right hand side or objective function coefficients of a linear integer program. In this Chapter we will naturally extend their results to nonlinear integer programming problems. Although much attention has been given to sensitivity analysis for linear integer programming, unfortunately this is not the case for nonlinear integer programming. Some results for the latter problem can be found for e.g. in Radke [41] who developed a continuity analysis for nonlinear bounded integer programming, in [38] where McBride and Yormark solved a class of parametric quadratic integer programming problems which were obtained by changing the right hand side of a single constraint, or in [9] were Cooper solved a parametric family, with respect to the right hand side, of a pure integer nonlinear program with separable objective function and constraints. We will restrict our attention to a nonlinear integer program whose continuous relaxation is a convex programming problem satisfying the Kuhn-Tucker constraint qualification. This since we will make use of the duality theory for convex nonlinear programming problems as introduced by Wolfe [49]. In Section 4.1. we will consider pure $0-1$ nonlinear programs. In Section 4.2. we will discuss the extension to the mixed-integer case as well as to the case when integer variables are not restricted to be 0 or 1. Finally, the computational results for an example with quadratic objective function, linear constraints and $0-1$ variables will be given in Section 4.3.

4.1. THE PURE NONLINEAR 0-1 PROGRAM

Consider the nonlinear integer programming problem

$$\begin{aligned} N(b) = \min \ & f(x) \\ \text{s.t. } & g(x) \geq b \\ & 0 \leq x \leq 1 \\ & x \text{ integer} \end{aligned} \tag{4.1}$$

where x is an $n \times 1$ vector, b is an $m \times 1$ vector, $g(x)^T = (g_1(x), ..., g_m(x))$, f and g_i, $i = 1, ..., m$ are real-valued, differentiable functions and, furthermore, f is convex and g_i, $i = 1, ...m$ are concave on R^n. We will also assume that the continuous relaxation of (4.1) satisfies the Kuhn-Tucker constraint qualification (see e.g. [49]), which is automatically satisfied if the constraints are all linear.

Explicitly assume the nonlinear integer programming problem (4.1) was solved by implicit enumeration and some small changes have been made in the right hand side or objective function coefficients of (4.1). The question we would like to answer is what information from the implicit enumeration tree, if at hand, will provide us with bounds on the optimal value of the perturbed problem.

Before attempting to answer the above question observe that the continuous relaxation of (4.1) is a convex nonlinear optimization problem for which we can state the dual (see Wolfe [49])

$$\max \ b^T u + e^T v^2 + f(x) - u^T g(x) - (v^1 + v^2)^T x$$

$$\text{s.t. } u^T \nabla g(x) + v^{1T} I + v^{2T} I = \nabla f(x) \tag{4.2}$$

$$u\ ,\ v^1 \geq 0\ ,\ v^2 \leq 0$$

where $e^T = (1, ..., 1)$ and u (resp., v^1 , v^2) denotes the vector of dual variables associated with the constraints $g(x) \geq b$ (resp., $x \geq 0$, $-x \geq -1$). Solving the primal (i.e. the continuous relaxation of (4.1)) with some readily available computer code, one can obtain the value of these dual variables as a byproduct (e.g. using MINOS -Modular In-core Nonlinear Optimization System).

REMARK 4.1.1 If all the constraints of (4.1) are linear, then the objective function of (4.2) can be written as (see e.g. Dorn [14]) $b^T u + e^T v^2 + f(x) - x^T \nabla f(x)$.

Let us start by solving (4.1) using implicit enumeration. In doing so we construct a tree with node 1 corresponding to the continuous relaxation of the original problem. At each node t of the tree one solves

$$\begin{aligned} R^t(b) = \min\ & f(x) \\ \text{s.t. }\ & g(x) \geq b \\ & L_j^t \leq x_j \leq U_j^t\ , \quad j = 1, ..., n \end{aligned} \tag{4.3}$$

where

$$\begin{aligned} L_j^t = U_j^t = 0 \quad & , j \epsilon F_0^t \\ L_j^t = U_j^t = 1 \quad & , j \epsilon F_1^t \\ L_j^t = 0, U_j^t = 1 \quad & , j \epsilon F \backslash (F_0^t \cup F_1^t) \end{aligned}$$

with $F = \{1, ..., n\}$ and F_0^t (resp., F_1^t) is the set of indices of variables fixed to zero (resp., one). Notice that (4.3) is the continuous relaxation of the integer nonlinear subproblem $I^t(b)$ associated with node t . For simplicity of exposition we will use

$N(b)$, $I^t(b)$ and $R^t(b)$ to denote the respective problems as well as their optimal value.

Now, since for each j at most one of the constraints $L_j^t \le x_j$ or $x_j \le U_j^t$ can be binding at any time, we will associate the same dual variable v_j with the two constraints. If $L_j^t \le x_j$ is binding then the associated dual variable satisfies $v_j \ge 0$ otherwise $v_j \le 0$. Let (u_t^T, x_t^T, v_t^T) be the associated dual solution obtained by solving (4.3). Using (4.2) the optimal objective function value of (4.3) equals

$$\begin{aligned} R^t(b) &= b^T u_t + \sum_{j \epsilon F_1^t} (v_t)_j + \sum_{j \epsilon F \setminus (F_1^t \cup F_0^t)} \min\{0, (v_t)_j\} + f(x_t) \\ &\quad - u_t^T g(x_t) - \sum_{j \epsilon F_1^t} (v_t)_j - \sum_{j \epsilon F \setminus (F_1^t \cup F_0^t)} \min\{0, (v_t)_j\} \\ &= b^T u_t + f(x_t) - u_t^T g(x_t) \,. \end{aligned}$$

REMARK 4.1.2 If the constraints in (4.1) are all linear the optimal objective function value of (4.3) can be alternatively written as

$$R^t(b) = b^T u_t + \sum_{j \epsilon F_1^t} (v_t)_j + \sum_{j \epsilon F \setminus (F_1^t \cup F_0^t)} \min\{0, (v_t)_j\} + f(x_t) - x_t^T \nabla f(x_t) \,.$$

Assume now that (4.1) is perturbed by replacing b by a new vector d . We would like to use the information from the implicit enumeration tree to derive a lower bound on the value of the perturbed problem. To this end notice that the dual variables u_t, x_t, v_t derived at node t of the tree remain feasible, but not necessarily optimal, for the perturbed problem. Therefore, a lower bound on the objective function of the perturbed problem is given by

$$\underline{\mathrm{R}}^t(d) = d^T u_t + f(x_t) - u_t^T g(x_t) \,. \tag{4.4}$$

Let $\underline{\mathrm{I}}^t(d)$ denote a lower bounding function on the objective function value

$I^t(d)$. For terminal nodes with feasible solutions for $R^t(b)$ we can set $\underline{I}^t(d) = \underline{R}^t(d)$. For terminal nodes with no feasible solutions set

$$\underline{I}^t(d) = \begin{cases} -\infty & \text{if } \underline{R}^t(d) \leq 0 \\ +\infty & \text{if } \underline{R}^t(d) > 0 \end{cases}$$

where u_t and x_t used in this case in the evaluation of $\underline{R}^t(d)$ are part of the dual variables $(u_t^T,\ x_t^T,\ v_t^T)$ associated with the minimization of the sum of the infeasibilities. (Having a problem with only linear constraints and using Remark 4.1.2, both u_t and v_t and $x_t \equiv 0$ can be used in the evaluation of $\underline{R}^t(d)$.)

For each nonterminal node t define

$$\underline{I}^t(d) = \max \{\underline{R}^t(d)\ , \min \{\underline{I}^{L(t)}(d),\ \underline{I}^{R(t)}(d)\}\} \tag{4.5}$$

where $L(t)$ and $R(t)$ are the two offsprings of t .

Theorem 4.1.3 below provides us with a lower bound on the objective function value of the perturbed problem obtained from (4.1).

THEOREM 4.1.3 $\underline{I}^1 d)$ is a lower bound for $N(d)$.

PROOF We show first that for any node t and any d , $\underline{I}^t(d) \leq I^t(d)$ is valid. If t is a terminal node then the inequality follows from the definition of $\underline{I}^t(d)$ and the convention that an infeasible minimization problem has objective function value $+\infty$. Now, if t is not a terminal node, $I^t(d) \geq R^t(d) \geq \underline{R}^t(d)$. Further, from the implicit enumeration

$$I^t(d) = \min \{I^{L(t)}(d)\ ,\ I^{R(t)}(d)\ \} \geq \min \{\underline{I}^{L(t)}(d)\ ,\ \underline{I}^{R(t)}(d)\ \} .$$

Therefore, $I^t(d) \geq \max \{\underline{R}^t(d)\ ,\ \min \{\underline{I}^{L(t)}(d)\ ,\ \underline{I}^{R(t)}(d)\}$ and by (4.5), $I^t(d) \geq$

$\underline{I}^t(d)$. By induction towards node 1 of the tree we get $N(d) = I^1(d) \geq \underline{I}^1(d)$.

$\square$

Theorem 4.1.4 to follow provides an answer to the following question. Assume we augment problem (4.1) by adding a new $0-1$ variable, say x_{n+1} , resulting in the addition of new linear terms in the constraints and some terms (not necessarily linear) in the objective function. Under what conditions will x_{n+1} remain at zero level in the optimal solution to the modified integer nonlinear program?

THEOREM 4.1.4. Suppose after solving (4.1) to optimality the problem was enlarged by introducing a new $0-1$ variable, say x_{n+1} , resulting in the addition of a new linear term, say $a_i x_{n+1}$, to each constraint $i = 1, ..., m$, and a number of new terms in the objective function given by $\tilde{f}(x)x_{n+1}$. Then there exist an optimal solution to the new problem with $x_{n+1} = 0$ if

$$\tilde{f}_{F_1} \geq N(b) - \underline{I}^1(b-a)$$

where $a^T \equiv (a_1, ..., a_m)$, $\tilde{f}_{F_1} \equiv \tilde{f}(\bar{x})$ and $\bar{x}$ is the optimal solution to the initial problem (4.1).

PROOF Suppose $x_{n+1} = 1$ is in an optimal solution at node 1 . The remaining optimal values can thus be found by solving problem (4.1) with right hand side $b-a$. By Theorem 4.1.3, $N(b-a) \geq \underline{I}^1(b-a)$ and therefore a solution to the enlarged problem with $x_{n+1} = 1$ has objective function not less than

$$N \equiv \tilde{f}_{F_1} + \underline{I}^1(b-a) .$$

Now, if $N \geq N(b)$ the solution to the initial problem remains optimal.

$\square$

4.2. THE MIXED-INTEGER CASE

Consider now problem (4.1) where only a subset of variables is restricted to be integer. Theorem 4.1.3 can be carried over without any changes. For Theorem 4.1.4 to be valid the assumptions remain unchanged, while the result should read : "Then there exist an optimal solution to the new problem with $x_{n+1} = 0$ if

$$\tilde{f}(\bar{x}) \geq N(b) - \underline{\mathrm{I}}^1(b-a)$$

where $\tilde{f}(\bar{x})$ denotes the sum of the new terms in the objective function evaluated at $\bar{x}$, the optimal mixed-integer solution to the original problem, and $x_{n+1} = 1$."

Next, consider the nonlinear mixed–integer program of the form (4.1) except that the integer variables are not necessarily restricted to be zero or one. The bound (4.4) can be derived in a straightforward manner as for the $0-1$ case and Theorem 4.1.3 will be valid without any changes. However, if x_{n+1} is an integer variable restricted to the interval $[0, U]$, then we will restrict ourselves in Theorem 4.1.4 to the case in which the added terms in the objective function are linear in x_{n+1} . In this event the bound can be improved as follows. If

$$\min\{\tilde{f}(\bar{x})x_{n+1} + \underline{\mathrm{I}}^1(b - ax_{n+1}) : x_{n+1} = 1, ..., U\} \geq N(b)$$

then the solution to the initial problem remains optimal.

4.3 THE QUADRATIC 0-1 PROBLEM: AN EXAMPLE

Consider the quadratic integer programming problem

$$Q(b) = \min\ c^T x + \frac{1}{2} x^T D x$$

$$\text{s.t. } Ax \geq b \tag{4.6}$$

$$0 \leq x \leq 1$$

$$x \text{ integer}$$

where A is an $m \times n$ and D an $n \times n$ matrix, c and x are $n-$vectors and b is an $m-$vector.

Observe that, although no restrictions are imposed on D , we can assume without loss of generality that D in (4.6) is symmetric and positive semidefinite. This will ensure the existence of a global minimal solution to the continuous relaxation of (4.6) whenever the polyhedron $P = \{x : Ax \geq b, 0 \leq x \leq 1\}$ is nonempty. Indeed, as stated in Section 1.4, if D is not of the desired form (4.6) can be replaced by an equivalent problem in which the objective function of (4.6) is replaced by

$$(c^T - \frac{1}{2}\lambda e^T)x + \frac{1}{2}x^T(\frac{1}{2}(D + D^T) + \lambda I)x$$

where λ is a positive scalar such that $\frac{1}{2}(D + D^T) + \lambda I$ is positive semidefinite, see for example [25].

Taking into account Remark 4.1.1, the dual of the continuous relaxation of (4.6) can be written as (see e.g. Dorn [13])

$$\begin{aligned} \max\ & b^T u + e^T v - \frac{1}{2}x^T Dx \\ \text{s.t. } & A^T u - Dx + Iv \leq c \\ & u \geq 0\,,\ -v \geq 0\,. \end{aligned} \tag{4.7}$$

This since the continuous relaxation of problem (4.6) can be written as

$$\min\{\frac{1}{2}x^T Dx + c^T x : \tilde{A}x \geq \tilde{b}, x \geq 0\}$$

where $\tilde{A} = \begin{pmatrix} A \\ -I \end{pmatrix}$ and $\tilde{b} = \begin{pmatrix} b \\ -e \end{pmatrix}$ with $e^T = (1, ..., 1)$. Following Dorn [13], Type I, page 60, its dual is

$$\max \{-\frac{1}{2}w^T Dw + \tilde{b}^T z : \tilde{A}^T z - Dw \leq c,\ z \geq 0\}$$

which is equivalent to

$$\begin{aligned} \max \quad & -\frac{1}{2}w^T Dw + b^T z_1 - e^T z_2 \\ \text{s.t. } & A^T z_1 - I z_2 - Dw \leq c \\ & z_1 \geq 0 \quad , \quad z_2 \geq 0 \, . \end{aligned}$$

Now, replacing w by x , z_1 by u and z_2 by $-v$ results in (4.7).

From Remark 4.1.2, the bound (4.4) in this case becomes

$$\underline{\mathrm{R}}^t(d) = d^T u_t + \sum_{j \epsilon F_1^t} (v_t)_j + \sum_{j \epsilon F \setminus (F_0^t \cup F_1^t)} \min\{0,\ (v_t)_j\} - \frac{1}{2} x_t^T D x_t \, . \tag{4.8}$$

It is easy to see that $\underline{\mathrm{R}}^t(d)$ can be further improved using information obtained from other nodes in the tree (see also [42] for the linear case). Indeed if the dual prices of all nodes, say N , of the implicit enumeration tree are used, $\underline{\mathrm{R}}^t(d)$ can be improved to

$$\underline{\mathrm{R}}^t(d) = \max_{s \epsilon N} \{d^T u_s + \sum_{j \epsilon F_1^t} (v_s)_j + \sum_{j \epsilon F \setminus (F_0^t \cup F_1^t)} \min\{0, (v_s)_j\} - \frac{1}{2} x_s^T D x_s \} \, .$$

For problem (4.6), Theorem 4.1.2 specializes to " Suppose after solving (4.6) to optimality the problem was enlarged by adding a new column, say a_{n+1} to A . Then, there exist an optimal solution to the new problem with $x_{n+1} = 0$ if

$$c_{n+1} + \frac{1}{2}(d_{n+1,n+1} + \sum_{i \epsilon F_1} d_{i,n+1}) \geq Q(b) - \underline{\mathrm{I}}^1(b - a_{n+1}) \, ,$$

where F_1 is the set of variables fixed to 1 in the optimal solution for (4.6) with $Q(b)$ its optimal objective function value."

EXAMPLE 4.1 Consider the quadratic integer programming problem of the form

$$Q = \min 65x_1 - 10x_2 + 7x_3 + 58x_4 - 8x_5 + 23x_6 - 8x_1x_4 + 16x_1x_6 + 4x_4x_6$$

$$\text{s.t. } 70x_1 - 20x_2 + 30x_3 - 20x_4 + 90x_5 + 100x_6 \geq 200$$

$$100x_1 + 30x_2 - 30x_3 + 80x_4 - 5x_5 + 70x_6 \geq 100$$

$$0 \leq x_i \leq 1 \,, x_i \text{ integer} \,, \ i = 1, ..., 6 \,.$$

The equivalent objective function of the form $c^T x + \frac{1}{2} x^T D x$ with positive semidefinite matrix D has

$$c^T = (15, -10, 7, 50, -8, 5) \quad \text{and} \quad D = \begin{pmatrix} 100 & 0 & 0 & -8 & 0 & 16 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -8 & 0 & 0 & 16 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 16 & 0 & 0 & 4 & 0 & 36 \end{pmatrix} .$$

The optimal solution to the continuous relaxation equals (0.20588,1,0.5196,0,1,1) with objective function value of 17.139 . The corresponding $0-1$ optimal solution equals (0,0,1,1,1,1) with objective function value of 84 . The continuous relaxation subproblems were solved using QPSOL, a FORTRAN package for Quadratic Programming developed at Systems Optimization Laboratory, Department of Operations Research, Stanford University, and implemented on AMDAHL 470 V-6 computer model II. QPSOL minimizes an arbitrary quadratic function subject to linear constraints where upper and lower bounds on the variables are handled separately. It requires an initial estimate of the solution and a subroutine to define the quadratic part of the objective function. Among output arguments, the Lagrangian multipliers for each constraint are given. In the case of an infeasible constraint set , the minimum of the sum of the infeasibilities was determined using the LINDO package.

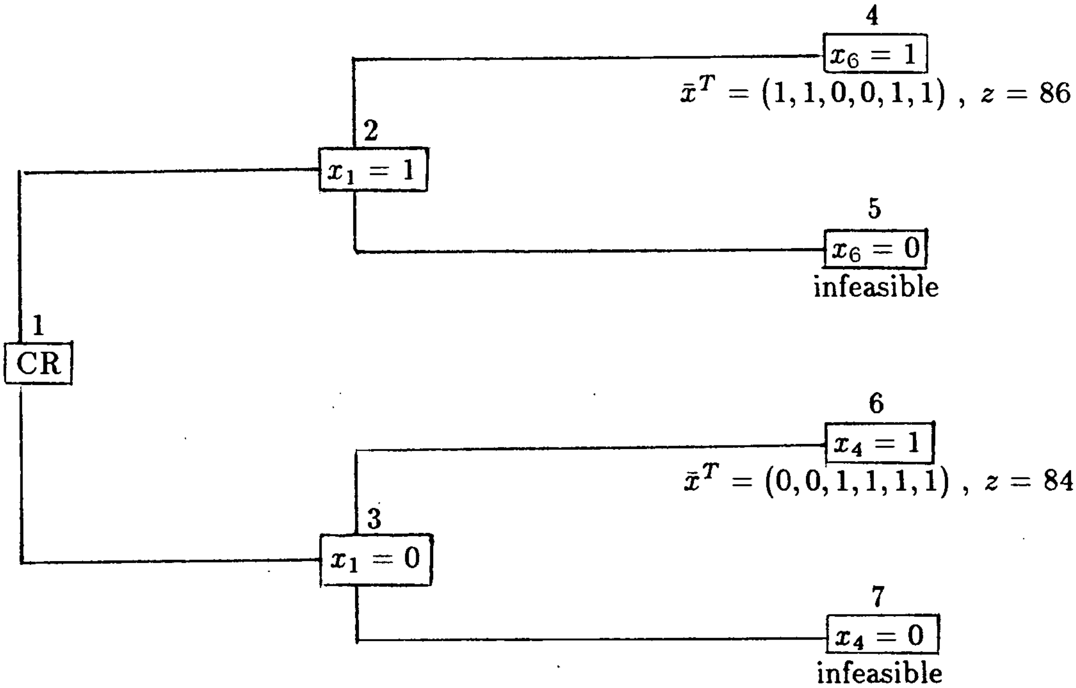

Figure 4.1: Branch and Bound Tree for Example 4.1

Figure 4.1 describes the enumeration tree associated with example 4.1. The number above each node corresponds to the node index while the entry in each node represents the branching choice. For terminal nodes with feasible solution an optimal solution and the optimal objective function value are denoted by $\bar{x}$ and z , respectively.

The lower bounding functions $\underline{R}^t(d)$ for the continuous subproblems solved at each node are given by

$$\underline{R}^7(d) = d_1 + d_2 - 265 ,$$

$$\underline{R}^6(d) = 0.5d_1 - 16 ,$$

$$\underline{R}^5(d) = d_1 + 0.25d_2 - 206.25 \ ,$$

$$\underline{R}^4(d) = 86 \ ,$$

$$\underline{R}^3(d) = 2.756d_1 + 1.504d_2 - 640.88 \ ,$$

$$\underline{R}^2(d) = 0.318d_1 - 1.68 \ ,$$

$$\underline{R}^1(d) = 0.441d_1 + 0.207d_2 - 91.75 \ .$$

The lower bounding functions for the corresponding integer problems are

$$\underline{I}^7(d) = \begin{cases} -\infty & \text{if } \underline{R}^7(d) \leq 0 \\ +\infty & \text{if } \underline{R}^7(d) > 0 \end{cases} ,$$

$$\underline{I}^6(d) = \underline{R}^6(d) \ ,$$

$$\underline{I}^5(d) = \begin{cases} -\infty & \text{if } \underline{R}^5(d) \leq 0 \\ +\infty & \text{if } \underline{R}^5(d) > 0 \end{cases} ,$$

$$\underline{I}^4(d) = \underline{R}^4(d) \ ,$$

$$\underline{I}^3(d) = \max \{\underline{R}^3(d) \ , \ \min \{\underline{I}^6(d), \underline{I}^7(d)\}\} \ ,$$

$$\underline{I}^2(d) = \max \{\underline{R}^2(d) \ , \ \min \{\underline{I}^4(d), \underline{I}^5(d)\}\} \ ,$$

$$\underline{I}^1(d) = \max \{\underline{R}^1(d) \ , \ \min \{\underline{I}^2(d), \underline{I}^3(d)\}\} \ .$$

A sample of sensitivity analysis for $d_1 \in (180, 240)$ and $d_2 \in (60, 140)$ is given in Table 4.1. The first number in each cell equals the optimal value $Q(d)$, the second number equals $\underline{I}^1(d)$, while the third number equals $\underline{R}^1(d)$.

d_1 / d_2	180	190	200	210	220	230	240
60	12 -0.106 -0.106	12 4.29 4.29	12 8.69 8.69	86 86 13.09	86 86 17.49	86 86 22.1	86 86 26.51
90	74 55.56 6.26	84 79 10.67	84 84 15.08	86 86 19.49	86 86 19.76	86 86 28.31	86 86 32.72
100	74 74 8.33	84 79 12.74	84 84 17.139	86 86 21.56	86 86 21.83	86 86 30.38	86 86 34.79
110	74 74 10.4	84 79 14.81	84 84 19.21	86 86 23.63	86 86 23.9	86 86 32.45	86 86 36.86
120	74 74 12.47	86 79 16.88	86 84 21.29	86 86 25.7	86 86 30.11	86 86 34.52	86 86 38.93
130	74 74 14.54	86 79 18.95	86 84 23.36	86 86 27.77	86 86 32.18	86 86 36.59	86 86 41
140	74 74 16.61	86 79 21.02	86 84 25.43	86 86 29.84	86 86 34.25	86 86 38.66	86 86 43.07

Table 4.1: Sensitivity Analysis Sample for Example 4.1

Chapter V

SIMULTANEOUS APPROXIMATION IN QUADRATIC 0-1 PROGRAMMING

Consider the following quadratic programming problem

$$\begin{aligned} \min \quad & \tilde{c}^T\tilde{x} + \frac{1}{2}\tilde{x}^T\tilde{D}\tilde{x} \\ \text{s.t.} \quad & \tilde{A}\tilde{x} \geq \tilde{b} \\ & 0 \leq \tilde{x} \leq 1 \\ & \tilde{x} \quad \text{integer} \end{aligned} \tag{5.1}$$

where $\tilde{A}$ is an $m \times n$ matrix, $\tilde{D}$ is an $n \times n$ symmetric matrix, $\tilde{c}$ and $\tilde{x}$ are n-vectors and $\tilde{b}$ is an m-vector. Problem (5.1) is a natural representation of many problems in, for example, finance [34] and capital budgeting [31]. Different approaches for solving the above problem can be found in the literature, e.g., linearization methods where the quadratic problem is transformed into a linear $0-1$ or a mixed-integer program can be found, respectively, in Watters [47] and Glover [22]. Algorithms based on a branch and bound method have been proposed by many authors, e.g., Mao and Wallingford [37], Laughhunn [31] and Hansen [26]. McBride and Yormark [38] gave an implicit enumeration algorithm in which at each node they solve a quadratic programming relaxation of a corresponding integer subproblem using Lemke-Howston's complementary pivoting algorithm. It is conceivable that a success of such an implicit enumeration algorithm depends greatly on the efficiency of the quadratic programming algorithm used. Although, at present, the polynomiality of an algorithm can not always be identified with real world computational efficiency or practicality, it is an important theoretical result which leads the research efforts in the direction of constructing efficient problem oriented polynomial algorithms. As stated in Section 1.5. Kozlov, Tarasov and Hačijan [30] provided the first polynomial time algorithm for convex quadratic programming problems. For a class of strictly convex quadratic programming problems, in Chapter II we proposed a polynomially bounded algorithm in which the number of arithmetic steps is independent on the size of the numbers in the linear cost coefficients and in the right hand side vector.

We show in this Chapter how to replace the objective function of a quadratic $0-1$ programming problem with n variables by an objective function with integral coefficients whose size is polynomially bounded by n, without changing the set of optimal solutions. We will use Frank and Tardos' [19] algorithm which in turn uses the simultaneous approximation algorithm from Lenstra at al. [33]. The above result assures that the running time of any algorithm for solving quadratic $0-1$ programming problems can be made independent of the size of the objective function coefficients. This since the equivalent problem can then be solved by e.g. an implicit enumeration algorithm in which at each node the continuous relaxation of the corresponding integer subproblem is solved in polynomial time independent of the size of the objective function coefficients.

Observe that since (5.1) is a $0-1$ programming problem then a constraint i with $b_i > \sum_{j=1}^{n} |a_{ij}|$ is clearly infeasible. This since $\sum_{j=1}^{n} a_{ij}x_j \leq \sum_{j=1}^{n} |a_{ij}| < b_i$. Therefore, assuming that (5.1) is feasible automatically assures that the entries of b can be bounded by the entries of the constraint matrix A. Note, however, that this does not imply that the *input size* of b is polynomially bounded by the input size of A since b can be a rational vector $\frac{p}{q}$ ($p^T = (p_1, ..., p_n)$, $q^T = (q_1, ..., q_n)$) where the size of p and (or) q is not polynomially bounded by the size of the entries of A . In any event, if the entries in the constraint matrix are polynomially bounded by the number of variables and/or constraints, then the continuous relaxations of the integer subproblems can be solved in strongly polynomial time using the algorithm presented in Chapter II. Recall that in a strongly polynomial algorithm the number of elementary arithmetic operations (i.e., additions, comparisons, multiplications and divisions) is independent of the size of the input and is polynomially bounded in the dimension of the input (i.e., the number of data in the input).

In Section 5.1. we state the problem and give some preliminary definitions. The extension of Frank and Tardos' preprocesing algorithm to quadratic $0-1$ problems is given in Section 5.2.

5.1. SETUP OF THE PROBLEM

For simplicity of exposition we will consider a problem with an objective function in homogeneous quadratic form

$$\begin{array}{ll} \min & \frac{1}{2}x^T Dx \\ s.t. & Ax \geq b \\ & 0 \leq x \leq 1 \\ & x \quad \text{integer} . \end{array} \tag{5.2}$$

This can be done without loss of generality since the transformation of the objective function of (5.1) to the homogeneous quadratic form given in (5.2) can be easily achieved as stated in Section 1.4. For the sake of completeness we will give here some details.

For example, by adding a new variable $y = 1$ problem (5.1) can be restated as (5.2) with

$$x = \begin{pmatrix} \tilde{x} \\ y \end{pmatrix}, \quad D = \begin{pmatrix} \tilde{D} & \tilde{c} \\ \tilde{c}^T & 0 \end{pmatrix}, \quad A = \begin{pmatrix} \tilde{A} & 0 \\ 0 & 1 \\ 0 & -1 \end{pmatrix} \text{ and } \quad b = \begin{pmatrix} \tilde{b} \\ 1 \\ -1 \end{pmatrix}.$$

Alternatively, since $x_i^2 = x_i$ for every $i = 1, ..., n$, we have

$$\tilde{c}^T x + \frac{1}{2}\tilde{x}^T \tilde{D}\tilde{x} = \frac{1}{2}\tilde{x}^T(\tilde{D} + 2C)\tilde{x}$$

where C is a diagonal matrix with $c_{ii} = \tilde{c}_i$, $i = 1, ..., n$.

In the sequel we will use some vector and matrix norms as defined in Section 1.1.

Let $S = \{x \in B_2^n : Ax \geq b\}$, where $B_2 = \{0, 1\}$. A vector $v \in S$ is said to be a **feasible solution** of (5.2) while a vector $z \in S$ for which $z^T Dz \leq v^T Dv$ for every $v \in S$ is said to be an **optimal solution** for (5.2).

The following lemma will justify the algorithm to follow.

LEMMA 5.1.1 If for every $u, v \in S$ we have

$$\text{sign}(u - v)^T D(u + v) = \text{sign}(u - v)^T \bar{D}(u + v)$$

for some symmetric matrices D and $\bar{D}$, then problems (5.2) with matrices D and $\bar{D}$, respectively, have the same set of optimal solutions.

PROOF We will show that every optimal solution of (5.2) with matrix D (resp., $\bar{D}$) in the objective function is optimal for problem (5.2) with $\bar{D}$ (resp., D) in the objective function. To that end suppose that for some $u \in S$ $u^T Du \leq$

v^TDv is valid for every $v \in S$. Then it follows from the symmetricity of D that $u^TDu - v^TDv = (u-v)^TD(u+v) \leq 0$ for all $v \in S$. Now

$$\text{sign } (u-v)^TD(u+v) = \begin{cases} -1 & \text{if } u^TDu < v^TDv \\ 0 & \text{if } u^TDu = v^TDv \ . \end{cases}$$

By the assumption, $\text{sign}(u-v)^T\bar{D}(u+v) = \text{sign}(u-v)^TD(u+v)$. This means that $u^T\bar{D}u \leq v^T\bar{D}v$ for every $v \in S$, which in turn implies optimality of u for problem (5.2) with the matrix $\bar{D}$ in the objective function.

□

5.2. SIMULTANEOUS APPROXIMATION OF OBJECTIVE FUNCTION COEFFICIENTS

Frank and Tardos [19] presented an algorithm which replaces a rational cost coefficient vector w of a linear programming problem with an integral vector $\bar{w}$, without changing the set of optimal solutions. Their algorithm uses a revised version of Lenstra, Lenstra and Lovász's simultaneous approximation algorithm (LLL algorithm) which is strongly polynomial. For the sake of completeness we will state Frank and Tardos' algorithm (F-T algorithm) [19] :

INPUT $w = (w(1), ..., w(n))$ rational vector and an integer N with $1 \leq N \leq 2n!$
OUTPUT $\bar{w} = (\bar{w}(1), ..., \bar{w}(n))$ integral such that $\| \bar{w} \|_\infty \leq 2^{n^3+2n^2+2n}N^{n(n+2)}$ and sign $(w, b) =$ sign $(\bar{w}, b)$ whenever b is an integral vector with $\| b \|_1 \leq N$.

0. Let $M = 2^{n^2+n+1}N^{n+1}$, $w_1 = w$, $\bar{w} = 0$ and $i = 1$.

1. Let $w'_i = \frac{1}{\|w_i\|_\infty} w_i$.

2. Apply the revised LLL algorithm to N and $w'_i(1), ..., w'_i(n)$. Let $p_i =$

$(p_i(1), \ldots, p_i(n))$ and q_i denote the output. Then $\| q_i w_i' - p_i \|_\infty < 1/N$ and $1 \leq q_i \leq 2^{n^2+n} N^n$.

3. Let $w_{i+1} = q_i w_i - p_i$ and $\bar{w} = M\bar{w} + p_i$. If $w_{i+1} \equiv 0$, HALT. Otherwise let $i = i+1$ and GOTO 1.

END.

The algorithm presented above can be generalized into a preprocessing algorithm that will transform the objective function coefficients of a $0-1$ quadratic programming problem into integer coefficients whose size will be bounded by a polynomial function of n and for which the set of optimal solutions remains unchanged. As stated above, without loss of generality, we will assume the homogeneous form in the objective function.

PREPROCESSING ALGORITHM FOR QUADRATIC 0-1 PROBLEMS

INPUT $D = (d_{ij})$ an $n \times n$ symmetric rational matrix and an integer $N = 4n^2$.

OUTPUT $\bar{D} = (\bar{d_{ij}})$ an $n \times n$ symmetric integer matrix with $\bar{\delta} \leq 2^{\tilde{n}^3 + 2\tilde{n}^2 + 2\tilde{n}} N^{\tilde{n}(\tilde{n}+2)}$, where $\tilde{n} = \frac{n(n+1)}{2}$ and $\bar{\delta} = \max\{\bar{d_{ij}}\ i,j = 1, \ldots, n\}$ and for which $\text{sign}(u-v)^T D(u+v) = \text{sign}(u-v)^T \bar{D}(u+v)$ for every integral u, v with $\| u \|_\infty \leq 1$ and $\| v \|_\infty \leq 1$.

STEP 1. Construct a rational vector $d = (d_{11}, \ldots, d_{1n}, d_{22}, \ldots, d_{2n}, \ldots, d_{nn})$ where d_{ij} is the entry in the i^{th} row and j^{th} column of the matrix D. Recall that since D is a symmetric matrix, we only need to approximate $\frac{n(n+1)}{2}$ elements of D.

STEP 2. Apply F-T algorithm to the vector d and integer N obtaining the integral vector $\bar{d}$.

STEP 3. Construct the integer, symmetric $n \times n$ matrix $\bar{D} = (\bar{d_{ij}})$ using the entries of the output vector $\bar{d}$.

END .

Theorem 5.2.1 to follow is a generalization of Theorem 3.1 in [19].

THEOREM 5.2.1 The matrix $\bar{D}$ satisfies the output criteria.

PROOF Using the entries of the vector d_i (resp., d'_i, p_i) from F-T algorithm, construct a symmetric matrix D_i (resp., D'_i, P_i). Denote by δ_i (resp., π_i) the largest absolute value of the entries in D_i (resp., P_i) and by r the number of iterations in F-T algorithm. The first part of the output criteria (i.e., the bound on the entries of $\bar{D}$) is satisfied by construction of the matrix $\bar{D}$ and since the F-T algorithm is valid. The validity of the second assertion can be shown in a similar way it was done in [19] as follows. Recall that for a matrix D its max-norm is given by $\| D \|_\infty = \max_{1 \le i \le n} \sum_{j=1}^{n} |d_{ij}|$. Now, we first show that $(u-v)^T D_i(u+v) \ge 0$ implies $(u-v)^T P_i(u+v) \ge 0$. Suppose, on the contrary, that $(u-v)^T P_i(u+v) < 0$. Then, from the integrality of u, v and P_i , $(u-v)^T P_i(u+v) \le -1$. Therefore

$$(u-v)^T D_i(u+v) = \delta_i(u-v)^T D'_i(u+v)$$

$$= \delta_i(u-v)^T \{\tfrac{1}{q_i} P_i + \tfrac{1}{q_i}(q_i D'_i - P_i)\}(u+v)$$

$$\le \delta_i\{\tfrac{-1}{q_i} + \tfrac{1}{q_i}(u-v)^T(q_i D'_i - P_i)(u+v)\}$$

$$\le \delta_i\{\tfrac{-1}{q_i} + \tfrac{1}{q_i} \| u-v \|_1 \| q_i D'_i - P_i \|_\infty \| u+v \|_\infty\}$$

$$\le \delta_i\{\tfrac{-1}{q_i} + \tfrac{1}{q_i} n \| u-v \|_\infty \, n \, \| q_i d'_i - p_i \|_\infty \| u+v \|_\infty\}$$

$$< \delta_i\{\tfrac{-1}{q_i} + \tfrac{1}{q_i}\tfrac{n^2}{N}(\| u \|_\infty + \| v \|_\infty)^2\}$$

$$\leq \delta_i\{\tfrac{-1}{q_i} + \tfrac{1}{q_i}\tfrac{4n^2}{N}\} = 0$$

which is a contradiction. Interchanging the roles of u and v will result in a reversed inequality which in turn proves that $(u-v)^T D_i(u+v) = 0$ implies $(u-v)^T P_i(u+v) = 0$. From F-T algorithm and the construction of the vector d and the matrices $\bar{D}, P_1, ..., P_r$ it follows that D as well as $\bar{D}$ are linearly dependent on $P_1, ..., P_r$. Therefore, if $(u-v)^T P_i(u+v) = 0$ for each i, then $(u-v)^T D(u+v) = (u-v)^T \bar{D}(u+v) = 0$ and in this case the theorem is proved. Now, suppose that this is not the case and that j is the smallest index such that $(u-v)^T P_j(u+v) \neq 0$. F-T algorithm implies that $\text{sign}(u-v)^T D(u+v) = \text{sign}(u-v)^T D_j(u+v)$. Since $\text{sign}(u-v)^T D_j(u+v)$ is equal to $\text{sign}(u-v)^T P_j(u+v)$, it remains to show that $\text{sign}(u-v)^T P_j(u+v) = \text{sign}(u-v)^T \bar{D}(u+v)$. To that end recall that $\bar{D} = \sum_{i=1}^r M^{r-i} P_i$, where M is given in F-T algorithm. By induction on k we will prove that for $j \leq k \leq r$, $\text{sign}(u-v)^T P_j(u+v) = \text{sign} \sum_{i=1}^k M^{k-i}(u-v)^T P_i(u+v)$. For $k = j$ this follows because $(u-v)^T P_i(u+v) = 0$ for $i < j$. Assume that the induction hypothesis is true for $k-1$. Without loss of generality assume $\text{sign}(u-v)^T P_j(u+v) = +1$. Then $\text{sign} \sum_{i=1}^{k-1} M^{k-1-i}(u-v)^T P_i(u+v) = +1$, which together with the integrality of u, v and P_i implies

$$\sum_{i=1}^{k-1} M^{k-1-i}(u-v)^T P_i(u+v) \geq 1.$$

Now,

$$\sum_{i=1}^{k} M^{k-i}(u-v)^T P_i(u+v) = M \sum_{i=1}^{k-1} M^{k-1-i}(u-v)^T P_i(u+v)$$
$$+ (u-v)^T P_k(u+v) \geq M - 4n^2 \| p_k \|_\infty$$
$$\geq M - 2^{\tilde{n}^2+\tilde{n}} N^{\tilde{n}} N > 0.$$

The last inequality follows from F-T algorithm since $\| d_i' \|_\infty = 1$ implies $\| p_i \|_\infty \leq q_i \leq 2^{\tilde{n}^2+\tilde{n}} N^{\tilde{n}}$. This completes the proof.

□

The preprocessing algorithm described above can precede, for example, an implicit enumeration algorithm for solving quadratic $0-1$ programming problems. At each node, due to the above transformation, the continuous relaxation of the corresponding integer subproblem can then be solved in time independent of the objective function coefficients. Observe that we can always assume, without loss of generality, that the continuous subproblems are convex quadratic programming problems, i.e. the matrix associated with the quadratic terms is positive semidefinite (see e.g. [25]).

Chapter VI

AREAS FOR FURTHER RESEARCH

In this thesis we extended a number of recent results for linear programming problems to quadratic programming problems. Moreover, the results from Chapter IV were shown to be valid for a broader class of problems, namely for nonlinear integer programming problems whose convex continuous relaxations satisfy a given constraint qualification.

One possible avenue of further research is to try and extend the results obtained in Chapter III to a broader class of problems, for example to separable convex problems in which some or all of the variables are restricted to be integral.

As far as Chapter V is concerned, one might try to extend the given result to quadratic integer programming problems in which the integral variables are not necessarily restricted to be 0 or 1 .

In Chapter II a polynomial algorithm (whose running time is independent of the size of the linear cost coefficients and the right hand side vectors) was proposed for a class of strictly convex quadratic programming problems. It is an open question whether there exists a *strongly* polynomial algorithm for the above class of problems, as well as whether there exists such an algorithm for a class of linear programs.

Finally, although the results in this thesis have mainly theoretical significance, one might investigate the practical benefits in some cases. For example, the calculation of a lower bound on the objective function value of a problem with perturbed right hand side vector (see Chapter IV) might help a decision maker in deciding on a suitable changes of the initial right hand side vector.

Bibliography

[1] A.Bachem and B.Korte, An Algorithm for Quadratic Optimization Over Transportation Polytopes , *Z. Angew. Math. Mech.* 58 (1978) 456–461 .

[2] B. Bank and R. Hansel, Stability of Mixed–Integer Quadratic Programming Problems , *Mathematical Programming Study* 21 (1984) 1–17 .

[3] M.S. Bazaraa, J.J.Goode and C.M. Shetty, Constraint Qualifications Revisited, *Management Science,* 18 (1972) 567–573 .

[4] M.S. Bazaraa and C.M. Shetty, *Nonlinear Programming, Theory and Algorithms,* John Willey, New York (1979) .

[5] M.C. Biggs, Constrained Minimization Using Recursive Equality Quadratic Programming , *Numerical Methods for Nonlinear Optimization,* ed. F.A. Lootsma, Academic Press (1972) 411–428 .

[6] J.C.G. Boot, *Quadratic Programming : Algorithms, Anomalies, Applications,* Studies in Mathematical and Managerial Economics, Vol 2, ed. H. Theil, North-Holland Publishing Company (1964) .

[7] J.W.S. Cassels, *An Introduction to the Geometry of Numbers*, Die Grundlehren Der Mathematischen Wissenschaften In Einzeldarstellungen, Band 99, Springer-Verlag (1959).

[8] W. Cook, A.M.H. Gerards, A. Schrijver and E. Tardos, Sensitivity Theorems in Integer Linear Programming , *Mathematical Programming* 34 (1986) 251–264 .

[9] M.W. Cooper, Postoptimality Analysis in Nonlinear Integer Programming: The Right–Hand Side Case , *Naval Res. Logist. Quart.* 28 (1981) 301–307 .

[10] R.W. Cottle, Symmetric Dual Quadratic Programs , *Quart. Appl. Math.* 21 (1963) 237–243 .

[11] J.W. Daniel, Stability of the Solution of Definite Quadratic Programs , *Mathematical Programming* 5 (1973) 41–53.

[12] J.L. Debiesse and G. Matignon, Comparison of Different Methods for Calculation of Traffic Matrices , *Ann. Télécommunic.* 35 (1980) 91–102 .

[13] W.S. Dorn, Duality in Quadratic Programming , *Quart. Appl. Math.* 18 (1960) 155–162 .

[14] ______, A Duality Theorem for Convex Programs, *IBM J. of Res. and Devel.* 4 (1960) 407–413.

[15] B.C. Eaves, On Quadratic Programming , *Management Science* Vol. 17 No. 11 (1971) 698–711 .

[16] J. Edmonds, System of Distinct Representatives and Linear Algebra , *J. of Res. Nat. Bur. Standards* 71-B (1967) 241–245 .

[17] D.T. Finkbeiner II, *Introduction to Matrices and Linear Transformations,* A Series of Books in the Mathematical Sciences, Third Edition, ed. V. Klee, W.H. Freedman and Company (1978) .

[18] R. Fletcher, An Algorithm for Solving Linearly Constrained Optimization Problems , *Mathematical Programming* 2 (1972) 133–165 .

[19] A. Frank and E. Tardos, An Application of Simultaneous Approximation in Combinatorial Optimization , *26-th Annual Symposium on Foundations of Computer Science, IEEE,* New York (1985) 459–463 .

[20] U.M. Garcia-Palomares and O.L Mangasarian , Superlinearly Convergent Quasi-Newton Algorithms for Nonlinearly Constrained Optimization Problems , *Mathematical Programming* 11 (1976) 1–13 .

[21] P.E. Gill, W. Murray, M.A. Saunders and M.H. Wright, QP–Based Methods for Large–Scale Nonlinearly Constrained Optimization , *Nonlinear Programming* 4, ed. Mangasarian, Meyer and Robinson , Academic Press (1981) 57–98 .

[22] F. Glover, Improved Linear Integer Programming Formulations of Nonlinear Integer Programs, *Management Science* 22 (1975) 455–460.

[23] M. Grötschel, L. Lovász and A. Schrijver, *The Ellipsoid Method and Combinatorial Optimization,* Springer, Heidelberg, to appear.

[24] L.G. Hačijan, A Polynomial Algorithm in Linear Programming , *Soviet Math. Dokl.* 20 (1979) 191–194 .

[25] P.L.Hammer and A.A. Rubin, Some Remarks on Quadratic Programming with 0–1 variables , *Revue Francaise d'Informatique et de Recherche Operationalle* 4 (1970) 67–79 .

[26] P. Hansen, Quadratic Zero–One Programming by Implicit Enumeration, *Numerical Methods in Nonlinear Optimization,* ed. F.A. Lootsma, Academic Press, New York (1972) 265–278.

[27] M. Held, P. Wolfe and H. Crowder, Validation of Subgradient Optimization , *Mathematical Programming* 6 (1974) 62–88 .

[28] R. Helgason, J. Kennington and H. Lall, A Polynomially Bounded Algorithm for a Singly Constrained Quadratic Program, *Mathematical Programming* 18 (1980) 338–343 .

[29] J. Kennington and M. Shalaby, An Effective Subgradient Procedure for Minimal Cost Multy–Commodity Flow Problems , *Management Science* 23 (1977) 994–1004 .

[30] M.K. Kozlov, S.P. Tarasov and L.G. Hačijan, Polynomial Solvability of Convex Quadratic Programming , *Soviet Math. Dokl.* 20 (1979) 1108–1111 .

[31] D.J. Laughhunn, Quadratic Binary Programming with Applications to Capital Budgeting Problems , *Operations Research* 10 (1970) 454–467 .

[32] R. Lazimy, Improved Algorithm for Mixed–Integer Quadratic Programs and a Computational Study , *Mathematical Programming* 32 (1985) 100–113 .

[33] A.K. Lenstra, H.W. Lenstra,Jr and L.Lovász, Factoring Polynomials with Rational Coefficients , *Math. Ann.* 261 (1982) 515–534 .

[34] J. Lintner, The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets , *Rev. Econ. Statist.* 47 (1965) 13–37 .

[35] L. Lovász, An Algoritmic Theory of Numbers, Graphs and Convexity , *Report No.85368-OR, Institut für Ökonometrie und Operations Research, Universität Bonn,* (1985) .

[36] O.L. Mangasarian, *Nonlinear Programming,* McGraw Hill Series in Systems Science (1969) .

[37] J.C. Mao and B.A. Wallingford, An Extension of Lawler and Bell's Method of Discrete Optimization with Examples from Capital Budgeting, *Management Science* 15 (1969) 851–860.

[38] R.D. McBride and J.S. Yormark, Finding all Solutions for a Class of Parametric Quadratic Integer Programming Problems, *Management Science* 26 (1980) 784–795.

[39] M. Minoux, A Polynomial Algorithm for Minimum Quadratic Cost Flow Problems , *European Journal of Operational Research* 18 (1984) 377–387 .

[40] J.J. Moder and C.R. Phillips, *Project Management with CPM and PERT* , Reinhold Ind. Eng. and Man. Sci. Texbook Series, New York (1964) .

[41] M.A. Radke, *Sensitivity Analysis in Discrete Optimization* , Ph.D. Dissertation, University of California , Los Angeles (1974) .

[42] L. Schrage and L. Wolsey, Sensitivity Analysis for Branch and Bound Integer Programming , *Operations Research* 33 (1985) 1008–1023 .

[43] A. Schrijver, *Theory of Linear and Integer Programming* , Wiley-Interscience

Series in Discrete Mathematics and Optimization (1986) .

[44] J. Stoer and C. Witzgall, *Convexity and Optimization in Finite Dimensions I,* Springer-Verlag (1970) .

[45] E. Tardos, A Strongly Polynomial Algorithm to Solve Combinatorial Linear Programs , *Operations Research* 34 (1986) 250–256 .

[46] C. Van de Panne, *Methods for Linear and Quadratic Programming,* Studies in Mathematical and Managerial Economics (ed. H. Theil), North-Holland Publishing Company (1975) .

[47] L.G. Watters, Reduction of Integer Polynomial Problems to Zero–One Linear Programming Problems , *Operations Research* 15 (1967) 1171–1174.

[48] J. Wilkinson and C. Reinsch, *Linear Algebra, Handbook for Automatic Computation, Vol.2* , Springer-Verlag, New York (1971) .

[49] P. Wolfe, A Duality Theorem for Nonlinear Programming, *Quart. Appl. Math.* 19 (1961) 239–244.