Les Cahiers du GERAD ISSN: 0711–2440

An *SQP* Adapted Simple Decomposition

Mehdi Lachiheb
Hichem Smaoui

G–2003–73

October 2003

An SQP Adapted Simple Decomposition

Mehdi Lachiheb
Laboratoire de Mécanique Appliquée
École Polytechnique de Tunisie

Hichem Smaoui
École Nationale d'Ingénieurs de Tunis

October, 2003

Les Cahiers du GERAD
G–2003–73

Abstract

In the present work Sacher's simple decomposition, originally developed for quadratic programming problems, is incorporated into a sequential quadratic programming algorithm in order to handle large scale nonlinear programming problems. The resulting algorithm is tested on several example problems. Results indicate good convergence of the sequence of quadratic problems and excellent precision in the solution by the decomposition method. Furthermore, analysis of the evolution of the optimum set of extreme points of the sequence of quadratic programming problems gave way to the development of a procedure for initiating the decomposition with a whole set of extreme points. This set is determined at the start of each new iteration, based on the results of the preceding one, bypassing the solution of many master problems. Considerable computational saving is shown to be achieved by the modified algorithm.

Keywords: Sequential quadratic programming, Nonlinear programming, Simple decomposition, Large scale, Extreme point.

Résumé

Dans ce travail la méthode de décomposition simple de Sacher, développée pour la résolution de problèmes quadratiques, est incorporé dans un algorithme de programmation quadratique séquentielle pour résoudre des problèmes de programmation non linéaire généralisée de grande dimension. Cette méthode a été essayée sur plusieurs exemples de problèmes tests. Les résultats montrent une convergence satisfaisante et une excellente précision. Par ailleurs, l'analyse de l'évolution du groupe de points extrêmes retenus dans chaque problème quadratique nous conduit à développer une procédure permettant de générer un groupe de points extrêmes initial pour chaque nouvelle itération à partir des points extrêmes de l'itération précédente sans passer par le programme maître.

Mots Clés : décomposition, programmation non linéaire, programmation quadratique séquentielle, problèmes de grande dimension.

Acknowledgments: A part of this work was done at GERAD. The first author thanks J.L. Goffin for his invitation and for many interesting discussions. I am also grateful to P. Hansen for his help.

1 Introduction

The sequential quadratic programming method (SQP) was developed by Biggs, Han and Powell [1–6] for solving nonlinear optimization problems. The solution of a general nonlinear programming problem

$$(P) \qquad \begin{cases} \min \ f(x) \\ g_i(x) \leq 0 \quad i \in I \\ h_i(x) = 0 \quad i \in L \\ x \in IR^n \end{cases}$$

where $I = \{1, 2, \ldots, m\}$ and $L = \{1, 2, \ldots, l\}$, is carried out by iteratively solving a sequence of quadratic programming (QP) problems of the form

$$(QP_k) \begin{cases} \min \ Q(d) = \frac{1}{2} d^t B_k d + d^t \nabla f(x^k) \\ \nabla g_i(x^k)^t d + g_i(x^k) \leq 0 \quad i \in I \\ \nabla h_i(x^k)^t d + h_i(x^k) = 0 \quad i \in L \\ d \in IR^n \end{cases}$$

Quadratic programming problems can be solved by a variety of algorithms. Being interested in the treatment of large scale problems their solution is sought in the present work via a decomposition technique [7], precisely Sacher's simple decomposition [8–15]. It consists in transforming the original quadratic programming problem QP_k, whose variables form the space vector, into a problem whose variables are the coefficients of the convex combinations expressing the space vector in terms of the extreme points of the feasible set. Solving a quadratic programming problem is then achieved via the iterative solution of two problems: a master problem and a subproblem. Among the attractive features of this decomposition method are the following advantages:

– the feasible set of the master problem is always the convex hull of a set of affinely independent extreme points, therefore, the dimension of the master problem never exceeds $n + 1$, and actually does not exceed $n + 1 - m'$ where m' is the rank of the Jacobian of the active constraints, including equalities,

– the constraints of the original quadratic problem appear only in the subproblem which is a linear programming problem,

– there is no need for Lagrange multipliers to be used in coupling master problem and subproblem.

The objective of the present work is to take advantage of Sacher's decomposition by incorporating it into the SQP algorithm in order to enhance its large scale capabilities. First, the decomposition algorithm is applied to quadratic programming problems with unbounded feasible set by expressing the solutions as combinations of extreme points and extreme rays. The barrier function used in solving the master problem [15] is modified to accommodate unbounded feasible domain. The algorithm has been subjected to a significant number of tests on example problems. A number of these test problems have been constructed in a way to exhibit specific features such as ill-conditioning of the objective function [16]. Second, the decomposition method thus implemented is integrated into a sequential quadratic programming algorithm to form a general nonlinear programming code [16] that will be denoted $SQPD$. The latter has been validated through a number of numerical tests, each problem being subjected to many runs using different starting solutions. Examination of the evolution of the optimum set of extreme points (SEP) from a SQP iteration to another led to the development of a procedure that aims at reducing the computational effort devoted to the generation of intermediate extreme points. The underlying idea consists in initiating the decomposition process with a whole SEP instead of a single extreme point. The initial SEP is determined from the results of the preceding iteration of the SQP sequence without solving a series of master problems and subproblems. In the present paper, first the simple decomposition for solving QP problems is presented and extended to unbounded feasible sets. This extension is made by introducing extreme rays in addition to extreme points. Then, evolution of the sequence of the optimum SEP's is analyzed for the purpose of predicting a relevant SEP and reducing the global effort required for its generation. Finally, numerical results are presented for several nonlinear programming example problems that demonstrate the computational saving achieved by the proposed procedure.

2 Generalities

2.1 Sequential quadratic programming

The sequential quadratic programming method [1–6] combines the advantages of variable metric methods for unconstrained optimization with the rapid convergence of Newton's method for solving nonlinear systems of equations. It is based on the works of Biggs, Han and Powell [2,4–6]. The algorithm consists in solving a sequence of quadratic programming problems of the form

$$(PQ)_k \quad \left\{ \begin{array}{c} \min \;\; Q(d) = \frac{1}{2} d^t B_k d + d^t \nabla f(x^k) \\ \nabla\, g_i(x^k)^t d + \, g_i(x^k) \leq 0 \quad i \in I \\ \nabla h_i(x^k)^t d + h_i(x^k) = 0 \quad i \in L \\ d \in I\!R^n \end{array} \right.$$

where B_k is an approximation of the Hessian $L(x^*, \lambda^*, \mu^*)$, of the Lagrangian function:

$$l(x, \lambda^*, \mu^*) = f(x) + \sum_{i \in I} \lambda_i^* g_i(x) + \sum_{j \in L} \mu_j^* h_j(x)$$

over the set of feasible directions at the solution x^* of problem (P), λ^* and μ^* being the optimal Lagrange multipliers.

2.2 Sacher's simple decomposition:

Sacher's simple decomposition [13] is applied to quadratic programming problems of the form:

$$(\text{PQ}): \quad \begin{cases} \min \ \frac{1}{2}x^tBx + c^tx \\ A_1x \geq b_1 \\ A_2x = b_2 \\ x \geq 0 \end{cases}$$

where $x = (x_1, x_2, \ldots, x_n) \in I\!R^n$ is the vector of variables, B is a $n \times n$ positive semi-definite matrix, A_1 and A_2 are respectively $m_1 \times n$ and $m_2 \times n$ matrices, c, b_1 and b_2 are vectors of dimensions n, m_1 and m_2 respectively.

Let $S = \{x \in I\!R^n, A_1x \geq b_1, A_2x = b_2 \text{ and } x \geq 0\}$ be the feasible set for problem (PQ). S is a convex polytope [11], therefore there exist p extreme points $x^1, x^2, x^3, \ldots, x^p$ $(p \geq 1)$ and q extreme rays $d^1, d^2, d^3, \ldots, d^q$, $(q \geq 0)$ such that

$$\forall\, x \in S, \exists \ u_1, \ldots, u_p, \ v_1, \ldots, v_q \in I\!R^+ \text{ such that}$$
$$\sum_{i=1}^{p} u_i = 1 \quad \text{and} \quad x = \sum_{i=1}^{p} u_i.x^i + \sum_{i=1}^{q} v_j.d^j .$$

or, in matrix notation

$$x = Uu + Vv$$

where

$$U = \begin{pmatrix} x_1^1 & x_1^2 & \ldots & x_1^p \\ \\ x_2^1 & x_2^2 & \ldots & x_2^p \\ \ldots & \ldots & \ldots & \ldots \\ x_n^1 & x_n^2 & \ldots & x_n^p \end{pmatrix} \quad \text{and} \quad V = \begin{pmatrix} d_1^1 & x_1^2 & \ldots & d_1^q \\ \\ d_2^1 & d_2^2 & \ldots & d_2^q \\ \ldots & \ldots & \ldots & \ldots \\ d_n^1 & d_n^2 & \ldots & d_n^q \end{pmatrix}.$$

For simplicity of notation, we introduce the $n \times (p+q)$ matrix $W = (U, V)$ and the $(p+q)$-vector $w = \binom{u}{v}$ so that x can be written in the form $x = Ww$. Substituting Ww for x in problem (PQ) gives rise to an equivalent problem (MP) defined by

$$(MP): \quad \begin{cases} \min \ \frac{1}{2}w^tQw + s^tw \\ \sum_{i=1}^{p} u_i = 1 \\ w \geq 0 \end{cases}$$

where $Q = W^tBW$ is a $(p+q) \times (p+q)$- positive semi-definite matrix and $s = W^tc$ is a $(p+q)$-vector.

2.2.1 Simple decomposition algorithm: Sacher's simple decomposition algorithm can be summarized in the following steps [13].

Step 1: Let U and V be two matrices made up columnwise of extreme points and extreme rays respectively. U has at least one column whereas V may be empty.

Step 2: Solve the master problem (PE). If it is unbounded the problem (PQ) is also unbounded. Otherwise, let $\binom{c}{u}$ denote the solution of the master problem and let

$$\tilde{x} = Uu + Vv.$$

Step 3: Solve the subproblem

$$(SP) \begin{cases} \min h^t x \\ A_1 x \geq b_1 \\ A_2 x = b_2 \\ x \geq 0 \end{cases}$$

where $h = BUu + BVv + c = B\tilde{x} + c$. If the solution of (SP) is bounded, then it must coincide with an extreme point which will be denoted by x^k. Otherwise let d^k be a feasible descent direction ($h^t d^k < 0$).

Step 4: If (SP) is bounded and has a solution x^k such that $h^t\tilde{x} = h^t x^k$, then $\tilde{x}$ is the solution of problem (PQ). Otherwise go to Step 5.

Step 5: If there exists $i \in IN/u_i = 0$ (resp. $v_i = 0$) then eliminate extreme point x^i (resp. extreme ray d^i). If subproblem (SP) is bounded, then replace $Uby(U, x^k)$. Otherwise replace V by (V, d^k). Go to Step 1.

2.2.2 Solution of the master problem: The structure of the master problem makes it suitable for solution by a penalty method. When the feasible set for the original quadratic problem is bounded the vector w in problem (PE) is made up solely of the components u_i verifying $\sum_{i=1}^{p} u_i = 1$. The barrier function used is [15]:

$$K(x, r) = -r \sum_{i=1}^{n_k} \log x_i.$$

where n_k is the current number of extreme points and extreme rays. The above function is not usable in general if the feasible set is unbounded. However, it is applicable under the assumption of positive definiteness of matrix B. Indeed, let

$$S = \left\{ x \in IR^p \times IR^q / \quad x \geq 0 \text{ and } \sum_{i=1}^{p} x_i = 1 \right\}.$$

Define the penalty function $\Psi : (x,r) \mapsto \Psi(x,r) = \frac{1}{2}x^tQx + s^tx - r\sum_{i=1}^{n_k}\log x_i$. Assuming B is positive definite, one has

$$\lim_{\|x^2\|\longrightarrow+\infty} \Psi(x,r) = +\infty$$

where $x = (x^1, x^2) \in I\!R^{n_k}$, $x^1 \in I\!R^p$ and $x^2 \in I\!R^q$,
hence $\forall\ A \geq 0\,, \exists\, D \geq 0\ \ /\ \ \forall x/\ \ \left\|x^2\right\| \geq D, \forall\ r \leq 1$, one has $\Psi(x,r) \geq A$.
Therefore there exists a compact set $I\!K$ in $I\!R^{n_k}$ such that

$$\min_{x\in S} \Psi(x,r) = \min_{x\in S\cap I\!K} \Psi(x,r).$$

$\tilde{x}$ being the optimum of $f(x)$ over S. Then $f(\tilde{x}) \leq f(x)\ \forall\ x \in S$. Let $\epsilon > 0$. Since the sum $\sum_{i=1}^{n_k}\log x_i$ is bounded from above over $I\!K$, it follows that

$$\overline{\lim_{r\longrightarrow 0}}\ r\sum_{i=1}^{n_k}\log x_i = 0, \forall x \in S\cap I\!K$$

therefore $\exists\ r_1 > 0\ /\ \forall\ r \leq r_1,\ \forall\, x \in S\cap I\!K$ one has $r\sum_{i=1}^{n_k}\log x_i \leq \epsilon$
which implies $\forall\ r \leq r_1\,,\ \forall x \in S\cap I\!K\,,\ f(\tilde{x}) - \epsilon \leq f(x) - r\sum_{i=1}^{n_k}\log x_i \Longrightarrow \forall\ r \leq r_1\,,\quad f(\tilde{x}) - \epsilon \leq f(\tilde{x}(r)) - \epsilon \leq \Psi(\tilde{x}(r), r)$, where $\tilde{x}(r)$ is the optimum of $\Psi(x,r)$ over S.
Since f is continuous there exists x such that: $f(x) - \epsilon \leq\ f(\tilde{x})$. Thus,

$$\begin{aligned} &\exists\, r_2 \geq 0\ /\ \forall\ r < r_2\,,\quad f(x) - r\sum_{i=1}^{n_k}\log x_i \leq\ f(\tilde{x}) + 2\epsilon \\ &\Longrightarrow \forall\ r < r_2\,,\quad \Psi(\tilde{x}(r), r) \leq\ f(\tilde{x}) + 2\epsilon\,. \end{aligned}$$

Therefore,
$\forall\varepsilon > 0,\ \exists\ \ r_0 > 0\ \ /\ \ \forall\ \ r \leq r_0\,,\quad f(\tilde{x}) - \epsilon \leq \Psi(\tilde{x}(r), r) \leq\ f(\tilde{x}) + \epsilon$ hence

$$\lim_{r\rightarrow 0} \Psi(\tilde{x}(r), r) = f(\tilde{x}).$$

Consider now a nonnegative, decreasing sequence $(r_k)_{k\in I\!N}$ such that

$$\lim_{k\rightarrow\infty} r_k = 0$$

and let $(x_k)_{k\in I\!N}$ be the sequence of corresponding solutions $x_k = \tilde{x}(r_k)$. Then

$$\lim_{k\rightarrow\infty} \Psi(x_k, r_k) = f(\tilde{x}).$$

Since, $x_k \in IK$, one can extract a subsequence $(x_{\Phi(k)})_{k\in IN}$ converging to x^*. From continuity of f it follows that $f(x^*) = f(\tilde{x})$; and since $X \cap IK$ is closed, $x_{\Phi(k)} \in X \cap IK \Longrightarrow x^* \in X$. Consequently, every accumulation point of the sequence $(x_k)_{k\in IN}$ is an optimum for (P).

Remarks:

– in case the function f is not strictly convex one can choose another penalty function $K(x,r)$ defined by

$$K(x,r) = -r \sum_{i=1}^{n_k} H(x_i)$$

where

$$H(x_i) = \begin{cases} \log x_i & \text{if } x_i \leq 1 \\ 1 - \frac{1}{x_i} & \text{if } x_i \geq 1 \end{cases}$$

which is continuous and differentiable over S.

– an advantage of the adopted choice for the penalty function is in that the barrier function is strictly convex even when the original function is nonconvex. This ensures uniqueness of the optimum for any value of r. In the following, the objective function of the problem (PQ) is assumed to be strictly convex. The penalized problem is written as

$$(\mathrm{MPP}_r)$$

$$\begin{cases} \min \frac{1}{2} \begin{pmatrix} u \\ v \end{pmatrix}^t Q \begin{pmatrix} u \\ v \end{pmatrix} + s^t \begin{pmatrix} u \\ v \end{pmatrix} - r \sum_{i=1}^{n_k} \log w_i \\ \sum_{i=1}^{p} u_i = 1 \end{cases}$$

For every solution $w = \binom{u}{v}$ let:

– D denote the diagonal matrix of dimension n_k having w_i as components.

– e denote the n_k vector whose first p components are ones and the remaining are zeroes.

– $f_r(w) = \frac{1}{2} w^t Q w + s^t w - r \sum_{i=1}^{n_k} \log w_i$.

– $g_r(w) = \nabla f_r(w) = Qw + s - r\, D^{-1} e$.

– $H_r(w) = \nabla^2 f_r(w) = Q + r\, D^{-2}$.

Lemma: [15]

For each penalty coefficient $r^j > 0$, let λ_j be the Lagrange multiplier associated with the unique constraint of problem (MP_{r^j}). Then

$$\lambda_j = \frac{e^t {H^{-1}}_{r^j} g_{r^j}}{e^t {H^{-1}}_{r^j} e}$$

and the Newton direction for problem MP_r at w is given by:

$$d_j = -{H^{-1}}_{r^j}(g_{r^j} - \lambda_j e).$$

3 Perturbation of extreme points

3.1 Extreme point characterization

In case the feasible set

$$S = \{x \in IR^n / \ \ A_1 x \geq b_1,\ A_2 x = b_2 \text{ and } \ x \geq 0\}.$$

of problem (P) is unbounded one may change it into a bounded set without altering the optimum solution, simply by imposing supplementary constraints $x_i \leq a, i = 1, \ldots, n$ where a is a sufficiently large real number. In the following, the assumption of bounded feasible set will be made. Feasible solutions are, therefore, written as convex combinations of extreme points only. The feasible set of a generic quadratic programming problem in the SQP sequence is defined by $S = \{x \in IR^n / \, \exists\, A_1 x \geq b_1, \quad A_2\, x = b_2 \text{ and } x \geq 0\}$. In order to characterize the extreme points of S we introduce slack variables and rewrite it as $S = \{x \in IR^n / \, \exists\, h \in IR^m \ / \ (x,h) \in H\}$ where

$$H = \left\{(x,h) \in IR^n \times IR^m / \ A \begin{pmatrix} x \\ h \end{pmatrix} = b\,, \quad x \geq 0 \text{ and } \ h \geq 0\right\},$$

$A = \begin{pmatrix} A_1 & -I \\ A_2 & 0 \end{pmatrix}$, $b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$, I denoting the $m \times m$ identity matrix. Thus, each extreme point is defined by an $m \times m$ nonsingular submatrix of A, or simply by a set of m columns of A.

3.2 Influence of conditioning

Let B be a nonsingular submatrix of A and x the solution of the equation $Bx = b$. A small perturbation in the matrix A and the right hand side b results in a perturbation in the set H, and possibly in a change in the topology and number of its extreme points. The following cases may occur for a given extreme point characterized by a matrix B:

i/ the perturbed matrix $B+\delta B$ is singular, therefore no extreme point can be associated to it. In other words, at least one extreme point leaves the SEP as a result of the perturbation. This may happen when the matrix B is ill-conditioned

ii/ The matrix $(B + \delta B)$ is nonsingular and the equation

$$(B + \delta B)x = b + \delta b$$

has no nonnegative solution, which implies that the extreme point associated with matrix B transforms into a point which is not a vertex of the perturbed domain $H^{'} = \{x \in IR^n / \ \ (A + \delta A)\, x \geq b + \delta b \text{ and } \ x \geq 0\}$. In this case at least one extreme point enters the SEP. This may occur either at a nondegenerate extreme point with an ill-conditioned associated matrix B, or at a degenerate point independently of the conditioning of its associated matrix.

iii/ The matrix $(B + \delta B)$ is nonsingular and the equation

$$(B + \delta B)x = b + \delta b$$

has a nonnegative solution, which defines an extreme point $(x^i + \delta x^i)$. If the matrix B is well conditioned the perturbed extreme point should be close to x^i according to the following proposition.

Proposition 1. [18]. Let $\| \,.\, \|$ denote a subordinate matrix norm. If $\|\delta B\| < \frac{1}{\|B\|}$ then

$$\frac{\|\delta x\|}{\|x\|} \leq \frac{1}{1-\|B^{-1}\|\|\delta B\|}(Cond(B)(\frac{\|\delta B\|}{\|B\|} + \frac{\|\delta b\|}{\|b\|}))$$

where

$$Cond(B) = \left\|B^{-1}\right\| \|B\|$$

3.3 Approximation of the optimum *SEP*

The solution of the quadratic programming problem by the standard simple decomposition algorithm has been subjected to testing on many example problems. Examination of the variations of the extreme points through the SQP iterations has shown that, in some problems, particularly those exhibiting ill-conditioning, the number of extreme points getting in and out of the SEP is very large. Considering that the generation of each extreme point requires the solution of a large LP problem in addition to that of the master problem, the overall computational effort could be improved significantly if the number of extreme point generations were reduced. On the other hand, it has been noted that, in most cases and especially at the tail of the sequence, to each point in the optimum SEP of problem QP_k is associated a point in the optimum SEP of problem QP_{k+1} defined by the same columns in the coefficient matrix. In such cases the $k+1^{st}$ optimum SEP can be viewed as the result of the k^{th} optimum SEP by a smooth mapping T. This leads to the idea of obtaining the entire optimum SEP for a new QP directly from the previous one, at least in an approximate way, in general. In an attempt to construct an approximation of the $k+1^{st}$ optimum SEP the following approach is considered. Let $\{x^i, i = 1, \ldots, n_k\}$ be the optimum SEP for the k^{th} iteration. For each point x^j we seek a corresponding extreme point, for the feasible set S_{k+1}, that we characterize as the closest one to x^j. The new extreme point, denoted by y^j, is sought as the solution of the problem:

$$(SPL_{jk}) \quad \begin{cases} \min \sum\limits_{i \in L_j} x_i \\ x \in S^{k+1} \end{cases}$$

$L_j = \{i \in IN \ / \ \left\|x_i^j\right\| \leq \epsilon\}$, where ϵ a small nonnegative real number.

A drawback of the above formulation is that, due to ill-conditioning or degeneracy, the new points $y^i, i = 1, \ldots, n_k$ are not necessarily affinely independent, which may cause the number of points in the SEP to exceed the limit $n + 1 - m'$ in subsequent steps of the decomposition procedure. The largest affinely independent subset can be determined by applying the simplex algorithm to the following problem:

$$(R_k) \quad \begin{cases} \min \ x \ = \ \sum\limits_{i=1}^{n_k} u_i. \\ \sum\limits_{i=1}^{n_k} u_i y^i = x^* \\ \sum\limits_{i=1}^{n_k} u_i = 1 \\ u \geq 0 \end{cases}$$

where $x^* = \sum\limits_{i=1}^{n_k} u_i^* y^i$, u_i^* are the components of the optimum solution of the master problem corresponding to the SEP $\left\{y^i,\ i = 1, \ldots n_k\right\}$. The set of independent extreme points is obtained by retaining solely extreme points whose corresponding optimal coefficients are positive. The resulting set forms the initial group of extreme points for problem $(QP)_{k+1}$.

4 Numerical Examples

4.1 Powell's Problem

Powell's problem [4] is an example with a small number of variables and exhibiting pronounced nonlinearity. Table I presents the sequence of optimum SEP corresponding to a run of the SQP algorithm started at the solution $x_0 = (0, -2, 2, 0, -1)$ using the unmodified version of the simple decomposition. It can be seen that the maximum number of extreme points used at a given step is 4, that is less than $n + 1 = 6$. The basic columns stabilize from the third iteration for extreme points x^3, from the fourth iteration for x^1 and from the sixth for x^4. It can be noted that the latter leaves the optimum SEP at iteration 4 and reenters it at the sixth iteration.

The optimum solution obtained is $x^* = (-0.699034, -0.869963, 2.789922, 0.6968791, -0.69657065)$ and the objective value is 0.4388502. On the other hand it should be noted that convergence of the SQP sequence is achieved within 8 iterations with a tolerance of 10^{-5} on the norm of direction d, i.e., the same number of iterations as reported in [4].

Similarly, Table II presents the sequence of optimum SEP using the starting point $x_0 = (-2, 2, 2, -1, -1)$. In this example, the basic columns are seen to stabilize from the first iteration for all extreme points. Extreme point x^2 leaves the SEP at the fourth iteration.The optimum solution obtained is $x^* = (-1.71714, 1.59571, 1.82723, -0.76364, -0.76364)$ and the objective value is 0.0539495.

Table I: Sequence of optimum SEP and $\|d\|$ for Powell's problem $x_0 = (0, -2, 2, 0, -1)$.

Itr.	1	2	3	4	5	6	7	8
	.00000	0.0000	.00000	.00000	.00000	.00000	.00000	0.0000
	49.416	49.570	59.557	68.526	75.862	82.122	82.283	82.282
$x1$	24.166	27.685	35.386	42.095	49.420	53.821	53.782	53.770
	39.100	25.445	12.031	.00000	.00000	.00000	.00000	.00000
	0.0000	0.0000	0.0000	.74450	15.535	25.021	24.958	24.933
	$1.E+5$	$1.E+5$	309.49	184.783	146.59	.00000	.00000	.00000
	49.416	49.570	.00000	.00000	.00000	82.122	82.283	82.282
$x2$	50024.	44132.	23028.	13505.	5334.4	.00000	.00000	27.242
	20039.	49049.	64808.	76366.	90345.	16968.	90231.	$1.E+5$
	$1.E+5$	$1.E+5$	$1.E+5$	$1.E+5$	$1.E+5$	16892.	90155.	99975.
	.00000	0.0000	309.49	184.78	146.59	127.80	127.43	127.44
	49.416	49.570	.00000	.00000	.00000	.00000	.00000	.00000
$x3$	50024.	44132.	55.283	46.717	44.959	44.336	44.194	44.183
	20039.	49049.	47.509	45.336	41.332	38.406	38.554	38.573
	$1.E+5$	$1.E+5$	.00000	.00000	.00000	0.0000	.00000	.00000
	$1.E+5$	$1.E+5$				50.599	50.112	50.049
	49.416	49.570				49.609	49.927	49.967
$x4$	24.166	27.699				50.115	50.021	50.009
	39.100	25.449				.00000	.00000	.00000
	0.0000	0.0000				.00000	.00000	.00000
$\|d\|$	0.8074	0.3264	0.2187	0.1700	0.0679	0.0044	0.0002	.00004

4.2 Ten Bar Truss Design Problem

In this example the optimum design problem for a ten bar truss structure is considered. The detailed problem statement is given in [19]. The truss is to be designed for minimum self weight subject to stress constraints and minimum gage restraints on the cross sectional areas which constitute the design variables of the problem. The problem is solved by the SQP algorithm using the unmodified simple decomposition. The sequence converges within 6 iterations with a tolerance of 10^{-6} on $\|d\|$.

The optimal solution obtained is $x^* = (7.937867, 0.1, 8.0621, 3.9379, 0.1, 0.1, 5.7447, 5.5690, 5.5690, 0.1)$ and the optimum volume is 15931,8. Table III shows the sequence of optimum SEP. It is interesting to note that, except for the first iteration, the optimum SEP reduces to a singleton. Indeed, the number m' of active constraints, including lower bound constraints on the variables, is 10, so that $n + 1 - m' = 1$. As a consequence, there is no master problem to solve. Moreover, the unique extreme point corresponds to a constant set of basic columns with respect to both the original design variables and the slack variables.

Table II: Sequence of optimum SEP and $\|d\|$ for Powell's problem $x_0 = (-2, 2, 2, -1, -1)$.

Itr.	1	2	3	4	5
	22.3636	24.2513	24.3366	23.1894	$23,0914$
	77.7196	78.7143	79.3742	81.0056	$81,0056$
$x1$	.000000	.000000	.000000	.000000	.000000
	108.712	108.738	108.171	106.168	105.985
	.000000	.000000	.000000	.000000	.000000
	100.083	94.8945	93.6834		
	.000000	000000	.000000		
$x2$	142.486	137.057	135.109		
	.000000	.000000	.000000		
	82.8055	84.7330	86.0907		
	22.3636	24.2513	24.3366	23.1894	23.0914
	77.7196	78.7143	79.3742	81.0056	81.1590
$x3$	.000000	000000	.000000	.000000	.000000
	.000000	000000	.000000	.000000	.000000
	108.712	108.738	108.171	106.168	105.985
	100.083	94.8945	93.6834	93.2352	93.1792
	.000000	000000	.000000	.000000	.000000
$x4$	142.486	137.057	135.109	130.634	130.233
	82.8054	84.7330	86.0907	90.0489	90.3946
	.000000	.000000	.000000	.000000	.000000
$\|d\|$	0.30900	0.02544	0.02190	0.00625	0.00006

Table III: Sequence of optimum SEP and $\|d\|$ for ten bar truss problem.

Itr.	1		2	3	4	5	6
	$x1$	$x2$	$x1$	$x1$	$x1$	$x1$	$x1$
	4.76881	4.35332	6.39218	7.57217	7.82922	7.83786	7.83787
	.000000	.000000	.000000	.000000	.000000	.000000	.000000
	8.09764	8.89494	7.88101	7.96076	7.96213	7.96213	7.96213
	3.52469	3.39200	3.79158	3.83744	3.83787	3.83787	3.83787
	.000000	.000000	.000000	.000000	.000000	.000000	.000000
	.000000	.000000	.000000	.000000	.000000	.000000	.000000
	4.39436	5.29781	5.60752	5.64397	5.64472	5.64472	5.74472
	4.18233	3.25080	4.66185	5.36208	5.46714	5.46899	5.46899
	5.19629	4.99725	5.43361	5.46885	5.46899	5.46899	5.46899
	.000000	.000000	.000000	.000000	.000000	.000000	.000000
$\|d\|$	2.090		2.008	1.210	0.0257	0.0086	0.00001

4.3 Large size analytical examples

A family of example analytical problems are now constructed in the following form:

$$\begin{cases} \min \sum_{i=1}^{n-p} exp((x_i^2-4)(x_i-4)) + \sum_{i=n+1-p}^{n} (x_i^2-1)(x_i-1). \\ g_i(x) \ = x_i^2 + x_{i+1}^2 - 5 \leq 0, \quad i = 1, \ldots, n-1 \\ g_n(x) = x_n^2 + x_{n-1}^2 - 5 \leq 0 \\ x \geq 0.1 \end{cases}$$

where p is a positive integer which controls the number of active constraints at the optimum.

The analytical solution of these problems is trivial. Many example problems have been solved using both the unmodified and the modified decomposition procedures (SQPD) in order to assess the incidence of the initial SEP approach on the computational effort as the problem size increases. The computational load, justifiably measured by the total number of pivots involved in the generation of extreme points, is plotted in Figure 1 as a function of the number of variables for $p = 20$. The saving achieved by the modified SQPD method is clearly demonstrated. It is noted that the computational advantage improves to greater proportions as the problem size increases. This is essentially explained by two factors. The first is the redundant generation of intermediate extreme points carried out in the SQPD algorithm, which is avoided by the modified method. The second is the difference in the nature of the linear programming subproblems of the SQPD algorithm and those that generate the initial SEP in the modified algorithm. The second factor is clearly illustrated by the example problem with $n = 450$ where the total number of generated extreme points is nearly the same for both algorithms whereas the modified algorithm requires only half the number of pivots.

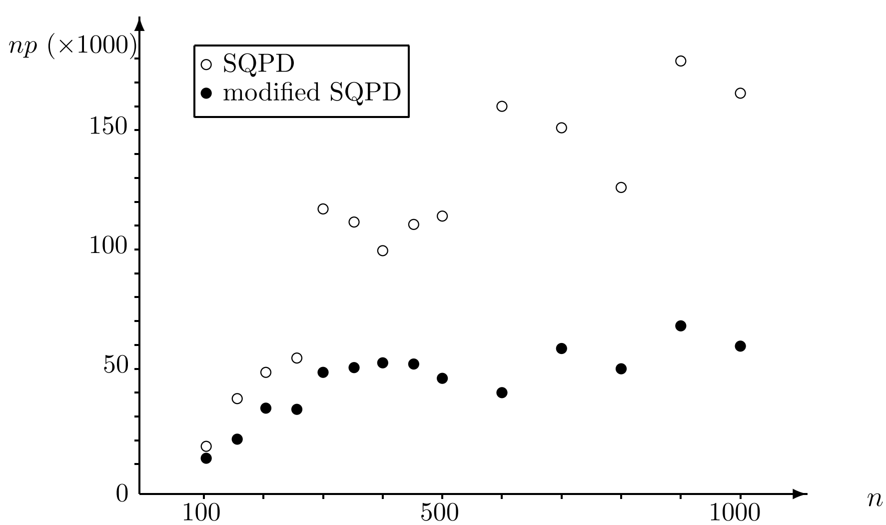

Figure 1: Evolution of total number of pivots versus problem size.

5 Conclusion

In the present work Sacher's simple decomposition is applied in solving the quadratic programming problems of the sequence of the SQP algorithm for nonlinear programming. The resulting algorithm naturally preserves the superlinear convergence of the sequential quadratic programming method, and has the advantage of providing improved accuracy and a capability for handling large scale problems. Furthermore, a procedure is developed that aims at reducing the computational effort devoted to the generation of intermediate extreme points. It consists in initiating the decomposition process with a whole set of extreme points, determined from the results of the preceding iteration, without solving a series of master problems and subproblems. Numerical results are presented for several nonlinear programming example problems that demonstrate computational saving up to 60% achieved by the proposed procedure. Possible improvements are under study. Future work will be devoted to reduction of storage requirement.

References

[1] A.V. Fiacco and G.P. McCormick (1968). *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, John Wiley and Sons, New York, SEC.2.4.

[2] J.V. Burke and S.P. Han (1989). A robust sequential quadratic programming method, *Mathematical Programming* 43, pp. 277–303.

[3] David F. Shanno and Kang Hoh Phua (1989). Numerical experience with sequential quadratic programming algorithms for equality constrained nonlinear programming, *ACM Transactions on Mathematical Software*, Vol. 15, No. 1, pp. 49–63.

[4] M.J.D. Powell (1978). Algorithms for nonlinear constraints that use lagrangian functions, *Math. Prog.*, Vol. 14, no.2.

[5] M.J.D. Powell (1977). The Convergence of Variable Metric Methods for Nonlinear Constrained Optimization Calculations, Proceedings of the Special Interest Group on Mathematical Programming Symposium-Univ. of Wisconsin-Madison.

[6] M.J.D. Powell (1977). A Fast Algorithm for Nonlinear Constrained Optimization Calculations, no.DAMPTP77/NA 382, University of Cambridge, England.

[7] G.B Dantzig (1963). *Linear Programming and Extensions*, Princeton University Press.

[8] B. Von Hohenbalken (1977). Simplicial decomposition in nonlinear programming algorithms, *Mathematical Programming* 13, pp. 49–68.

[9] R.S. Sacher (1980). A decomposition algorithm for quadratic programming, *Mathematical Programming* 18, pp. 16–30.

[10] C.M. Shetty and M. Ben Daya (1988). A decomposition procedure for convex quadratic programs, *Naval Research Logistics Quarterly* 35, pp. 111–118.

[11] J.M. Mulvey, S.A. Zenios and D.P. Ahlfeld (1990). Simplicial decomposition for convex generalized networks, *Journal of Information and Optimization Sciences* 11, pp. 359–387.

[12] J.A. Ventura and D.W. Hearn (1993). Restricted simplicial decomposition for convex constrained problems, *Mathematical Programming* 59, pp. 71–85.

[13] A. Mar'in (1995). Restricted simplicial decomposition with side constraints, *Networks* 26, pp. 199–215.

[14] D.W. Hearn, S. Lawphongpanich and J.A. Ventura (1987). Restricted simplicial decomposition: Computation and extensions, *Mathematical Programming Study*, 31, pp. 99–118.

[15] M. Ben Daya (1994). A Hybrid Decomposition Approach for Convex Quadratic Programming, King Fahd University of Petroleum and Minerals, under code SE/LINPROG/138.

[16] M. Lachiheb (1997). Extension de la décomposition hybride pour la programmation non linéaire, Mémoire de D.E.A. Mathématiques appliquées, ENIT, Tunis.

[17] R.Tyrrell Rockafellar (1970). *Convex Analysis*, Princeton, New Jersey, Princeton University Press.

[18] P.G.Ciarlet (1982). *Introduction à l'analyse numérique matricielle et à l'optimisation*, Masson, Paris.

[19] Uri Kirsch (1981). *Optimum Structural Design*, McGraw-Hill Book Company.