A Primal-Dual Active-Set Method for Convex Quadratic Programming *

Ekaterina Kostina Olga Kostyukova

Abstract

The paper deals with a method for solving general convex quadratic programming problems with equality and inequality constraints. The interest in such problems comes from at least two facts. First, quadratic models are widely used in real-life applications. Second, in many algorithms for nonlinear programming, a search direction is determined at each iteration as a solution of a quadratic problem. The method uses information about dual and primal variables to effectively manage active sets. At each iteration the duality gap is decreasing, and the process can be stopped earlier on a suboptimal solution. Numerical experiments show that the method is effective on problems with many box constraints and range inequalities.

Keywords: quadratic programming, active set strategy, primal-dual method

1 Introduction

In this paper we consider the general quadratic programming problem with range and box constraints. The interest in such problems comes from at least two facts. First, quadratic models are widely used in real-life applications. Second, in Sequential Quadratic Programming (SQP) methods for nonlinear programming, a search direction at each iteration is determined as a solution of a quadratic problem.

There is a lot of publications devoted to quadratic programming, see e.g. [8]. The method suggested in the paper is a generalization of the adaptive method [4], [3] for linear programming and belongs to the class of primal-dual active set methods, that is active set strategies which are based on the primal and dual information. This feature makes the method especially effective for solving quadratic programming problems with many inequality constraints.

The paper is organized as follows. In the next section we give some basic definitions and a motivation for our approach. Section 3 presents the description of the algorithm. In section 4 we discuss the results of numerical experiments.

*The preprint is an extended version of the paper: *E.A. Kostina and O.I. Kostjukova*, An Algorithm for Solving Convex Quadratic Programming Problems with Linear Equality and Inequality Constraints, *Zhurnal vychislitelnoy matematiki i matematicheskoy fiziki*, **42** (2001), 7, 1012 - 1026, (in Russian, transl. to English in *Computational Mathematics and Mathematical Physics.*)

2 Basic Definitions and Optimality Conditions

We consider a general quadratic programming problem with range and box constraints

$$\begin{aligned}\min f(x) &= c^T x + \frac{1}{2} x^T D x \\ & L_i \le A_i x \le U_i,\ i = 1, ..., m, \\ & l_j \le x_j \le u_j,\ j = 1, ..., n\end{aligned} \tag{1}$$

where $L = (L_i,\, i = 1, ..., m)^T$ and $U = (U_i,\, i = 1,\, ..., m)^T$ are m-vectors, $c = (c_j,\, j = 1, ..., n)^T$, $l = (l_j,\, j = 1, ..., n)^T$, $u = (u_j,\, j = 1, ..., n)^T$ and $x = (x_j,\, j = 1, ..., n)^T$ are n-vectors; $A_i = (A_{ij},\, j = 1, ..., n)$ is a n-row, $A = \begin{pmatrix} A_i \\ i = 1, ..., m \end{pmatrix}$ is an $m \times n$ matrix, D is a symmetric $(D^T = D)$ and positive semidefinite $n \times n$ matrix. We denote row and column index sets by $I = \{1, 2, ..., m\}$ and $J = \{1, 2, ..., n\}$. Let us note that box constraints in problem (1) can be interpreted as an l_∞ trust-region constraint in SQP methods.

The following vector-matrix notations are used throughout the paper: let S and Q denote index sets $S = \{s_1, s_2, ..., s_k\}$ and $Q = \{q_1, q_2, ..., q_r\}$, then vector $v(Q)$ and matrix $M(S, Q)$ are defined as follows:

$$v(Q) = (v_j,\ j \in Q)^T, \quad M(S, Q) = \begin{bmatrix} m_{ij}, j \in Q \\ i \in S \end{bmatrix} = \begin{bmatrix} M_i \\ i \in S \end{bmatrix}$$

with $M_i = M(i, Q) = (m_{ij}, j \in Q)$ denoting the $i-$th row of the matrix $M(S, Q)$.

Definition 1 *Any n-vector x satisfying all the constraints of problem (1) is called a feasible solution of problem (1).*

A feasible solution $x^\star$ is optimal if $f(x^\star) \le f(x)$ for all feasible solutions x of problem (1).

Given any $\varepsilon > 0$ an ε-optimal (suboptimal) feasible solution x^ε satisfies $f(x^\varepsilon) \le f(x^\star) + \varepsilon$.

2.1 Classical Optimality Conditions and Dual Problem

Following theory of nonlinear programming we can formulate *optimality conditions* for problem (1). We introduce Lagrange multipliers $\lambda_1,\, \lambda_2 \in R^m$, and $\mu_1,\, \mu_2 \in R^n$ and the Lagrange function $L = L(x, \lambda_1, \lambda_2, \mu_1, \mu_2)$:

$$\begin{aligned} L &= L(x, \lambda_1, \lambda_2, \mu_1, \mu_2) \\ &= c^T x + \frac{1}{2} x^T D x - \lambda_1^T (U - Ax) - \lambda_2^T (-L + Ax) - \mu_1^T (u - x) - \mu_2^T (-l + x) \end{aligned}$$

Theorem 1 (First-Order Necessary Optimality Conditions) *Let $x^\star$ be an optimal solution. Then there exist non-negative vectors $\lambda_1^\star,\, \lambda_2^\star \in R^m$, and $\mu_1^\star,\, \mu_2^\star \in R^n$ such that the following conditions are true:*

$$\frac{\partial L}{\partial x}(x^\star, \lambda_1^\star, \lambda_2^\star, \mu_1^\star, \mu_2^\star) = c + Dx^\star + A^T(\lambda_1^\star - \lambda_2^\star) + (\mu_1^\star - \mu_2^\star) = 0 \tag{2}$$

$$\begin{array}{llll} \mu_1^{\star T}(u - x^\star) = 0, & \mu_2^{\star T}(x^\star - l) = 0, & \mu_1^\star \ge 0, & \mu_2^\star \ge 0, \\ \lambda_1^{\star T}(U - Ax^\star) = 0, & \lambda_2^{\star T}(Ax^\star - L) = 0, & \lambda_1^\star \ge 0, & \lambda_2^\star \ge 0. \end{array} \tag{3}$$

$\diamond$

With the notation

$$\lambda^\star = \lambda_1^\star - \lambda_2^\star,\ \mu^\star = \mu_1^\star - \mu_2^\star \tag{4}$$

stationarity condition (2) and complementarity conditions (3) can be rewritten as

$$c + Dx^\star + A^T\lambda^\star = -\mu^\star, \tag{5}$$

$$\mu_j^\star \begin{cases} \geq 0, & \text{if } x_j^\star = u_j \\ \leq 0, & \text{if } x_j^\star = l_j \\ = 0, & \text{if } l_j < \ x_j^\star < u_j \end{cases} \quad j \in J, \quad \lambda_i^\star \begin{cases} \geq 0, & \text{if } A_i x^\star = U_i \\ \leq 0, & \text{if } A_i x^\star = L_i \\ = 0, & \text{if } L_i < A_i x^\star < U_i \end{cases} \quad i \in I. \tag{6}$$

Thus we can formulate

Theorem 2 (Second Form of the First-Order Necessary Optimality Conditions) *Let* $x^\star$ *be an optimal solution. Then there exists a vector* $\lambda^\star \in R^m$ *such that conditions (6) are true for* $x^\star$*,* $\lambda^\star$ *and* $\mu^\star = -c - Dx^\star - A^T\lambda^\star$.

$\diamond$

There is a unique correspondence between the vectors $\lambda^\star$, $\mu^\star$ occuring in Theorem 2 and the vectors $\lambda_1^\star$, $\lambda_2^\star$, $\mu_1^\star$, and $\mu_2^\star$ occuring in Theorem 1 which is given by (4) and the relations

$$\lambda_{1i}^\star = \lambda_i^\star,\ \lambda_{2i}^\star = 0 \text{ if } \lambda_i^\star \geq 0, \quad \lambda_{1i}^\star = 0,\ \lambda_{2i}^\star = -\lambda_i^\star, \text{ if } \lambda_i^\star < 0,\ i \in I,$$

$$\mu_{1j}^\star = \mu_j^\star,\ \mu_{2j}^\star = 0, \text{ if } \mu_j^\star \geq 0, \quad \mu_{1j}^\star = 0,\ \mu_{2j}^\star = -\mu_j^\star, \text{ if } \mu_j^\star < 0,\ j \in J.$$

Obviously, optimal vectors $x^\star$, $\lambda_1^\star$, $\lambda_2^\star$, $\mu_1^\star$, $\mu_2^\star$ are feasible for the *dual* problem of problem (1) which can be written as follows

$$\begin{aligned} \max \varphi(\chi, s, t, w, v) \ &= \ -\frac{1}{2}\chi^T D\chi - U^T s + L^T t + l^T v - u^T w \\ & \quad D\chi + A^T(s - t) + c - v + w = 0 \\ & \quad s, t, w, v \geq 0 \end{aligned} \tag{7}$$

and we have for the primal and dual cost functions

$$c^T x^\star + \frac{1}{2}x^{\star T} D x^\star = -\frac{1}{2}x^{\star T} D x^\star - U^T\lambda_1^\star + L^T\lambda_2^\star + l^T\mu_2^\star - u^T\mu_1^\star.$$

It follows from duality theory for the pair of problems (1) and (7) that for any primal feasible solution x and for any dual feasible solution $\Lambda = (\chi, s, t, w, v)$ the inequality

$$f(x) \geq \varphi(\Lambda)$$

is true. Moreover, the equality

$$f(x) = \varphi(\Lambda)$$

holds if and only if x is optimal in (1) and Λ is optimal in (7). Having a primal feasible solution x and a dual feasible solution Λ we can compute the duality gap

$$\beta(x, \Lambda) = f(x) - \varphi(\Lambda)$$

and get the estimation $f(x) - f(x^\star) \leq \beta(x, \Lambda)$.

2.2 Karush-Kuhn-Tucker System

Let us introduce the sets $I_A(x^\star)$, $J_A(x^\star)$ of active indices and the sets $I_N(x^\star)$, $J_N(x^\star)$ of non-active indices at the solution $x^\star$, namely

$$\begin{aligned} I_A(x^\star) &= \{i \in I : \; A_i x^\star = L_i \text{ or } U_i\}, \; I_N(x^\star) = I \backslash I_A(x^\star), \\ J_A(x^\star) &= \{j \in J : \; x_j^\star = l_j \text{ or } u_j\}, \; J_N(x^\star) = J \backslash J_A(x^\star), \end{aligned}$$

and denote $B_i = A_i x^\star = L_i \vee U_i$, $i \in I_A(x^\star)$, and $b_j = x_j^\star = l_j \vee u_j$, $j \in J_A(x^\star)$. It follows from Theorem 2 that optimal $x^\star$ and $\lambda^\star$ satisfy the so-called Karush-Kuhn-Tucker (KKT) system of linear equations

$$\begin{bmatrix} D(J_N(x^\star), J_N(x^\star)) & A(I_A(x^\star), J_N(x^\star))^T \\ A(I_A(x^\star), J_N(x^\star)) & 0 \end{bmatrix} \begin{bmatrix} x^\star(J_N(x^\star)) \\ \lambda^\star(I_A(x^\star)) \end{bmatrix} = \begin{bmatrix} r_1(J_N(x^\star)) \\ r_2(I_A(x^\star)) \end{bmatrix}$$

where

$$\begin{aligned} & r_1(J_N(x^\star)) = -c(J_N(x^\star)) - D(J_N(x^\star), J_A(x^\star)) b(J_A(x^\star)), \\ & r_2(I_A(x^\star)) = B(I_A(x^\star)) - A(I_A(x^\star), J_A(x^\star)) b(J_A(x^\star)), \\ & \lambda_i^\star = 0, \; i \in I_N(x^\star), \; x_j^\star = b_j, \; j \in J_A(x^\star). \end{aligned}$$

Obviously, the KKT system is solvable, or equivalently, the KKT matrix

$$\begin{bmatrix} D(J_N(x^\star), J_N(x^\star)) & A(I_A(x^\star), J_N(x^\star))^T \\ A(I_A(x^\star), J_N(x^\star)) & 0 \end{bmatrix}$$

is nonsingular if $D(J_N(x^\star), J_N(x^\star))$ is positive definite and the matrix $A(I_A(x^\star), J_N(x^\star))$ has full row rank.

2.3 Working Basis and Working Basis Optimality Conditions

We have seen that we need to know optimal Lagrange vectors $\lambda^\star$ and $\mu^\star$ in order to define an optimal solution. The index sets $\{i \in I | \lambda_i^\star = 0, i = 1, ..., m\}$ and $\{j \in J | \mu_j^\star = 0, i = j, ..., n\}$ will define the non-active constraints and non-active variables at the optimal solution. Since we cannot know this information in advance we need to work with "guesses" for the optimal sets and optimal Lagrange vectors. This is done with the help of a *working basis*.

Definition 2 *A pair $S_{WB} = \{I_{WB}, J_{WB}\}$ with $J_{WB} \subset J$, $I_{WB} \subset I$, $|J_{WB}| = |I_{WB}|$, is called a working basis if the matrix $A_{WB} = A(I_{WB}, J_{WB})$ is nonsingular.*

During the iterations of the method a set $I_{WB} \subset I$ will be accumulated only from those indices $i \in I$ for which $A_i x$ is equal to one of the ranges. Contrary, we will try to keep nonbasic variables $x_j, \; j \in J_N = J \setminus J_{WB}$, at their bounds. As an initial working basis one can take the empty working basis

$$J_{WB} = \emptyset, I_{WB} = \emptyset.$$

Let a feasible solution x and a working basis S_{WB} be given. This information is sufficient for checking whether the feasible solution x is optimal. Similar to linear programming, at the feasible solution x we calculate a vector of multipliers $\pi = \pi(I) = \pi(x|I)$ from the linear system

$$A_{WB}^T \pi(I_{WB}) = -D(J_{WB}, J)x - c(J_{WB}), \quad \pi_i = 0, i \in I_N = I \backslash I_{WB},$$

and reduced costs $\Delta = \Delta(J) = \Delta(x|J)$ as

$$\Delta = -Dx - A^T(I_{WB}, J)\pi(I_{WB}) - c.$$

Definition 3 *A feasible solution x is called nondegenerate with respect to a working basis S_{WB}, if basic components of x and nonbasic constraints are non-active, that is*

$$\begin{array}{l} l_j < x_j < u_j \ \textit{ for all } \ j \in J_{WB}, \\ L_i < A_i x < U_i \ \textit{ for all } \ i \in I_N. \end{array}$$

The following theorem can be proved analoguosly to [1], [5].

Theorem 3 (Working Basis Optimality Criterion) *For optimality of a feasible solution x in problem (1) it is sufficient that the relations*

$$\Delta_j \begin{cases} \geq 0 & \textit{if} \quad x_j = u_j, \\ \leq 0 & \textit{if} \quad x_j = l_j, \\ = 0 & \textit{if} \quad l_j < x_j < u_j, \end{cases} \qquad \textit{for } j \in J_N, \textit{ and} \tag{8}$$

$$\pi_i \begin{cases} \geq 0, & A_i x = U_i, \\ \leq 0, & A_i x = L_i, \\ = 0, & L_i < A_i x < U_i, \end{cases} \qquad \textit{for} \quad i \in I_{WB}, \tag{9}$$

are true. The conditions (8) and (9) are necessary for optimality of feasible solution x, which is nondegenerate with respect to a working basis S_{WB}.

Given a feasible solution x and a working basis S_{WB} we can check ε-optimality of a feasible solution by computing a duality gap. Obviously, x, π and Δ define a special *dual feasible solution* $\Lambda^{acc} = \{\chi^{acc}, s^{acc}, t^{acc}, w^{acc}, v^{acc}\}$ which we will call in the future "dual solution accompanying working basis" or briefly "accompanying the dual solution":

$$\chi^{acc} = x,$$

$$s_i^{acc} = \begin{cases} \pi_i & \text{if} \quad \pi_i \geq 0, \ i \in I_{WB}, \\ 0 & \text{if} \quad \pi_i < 0, \ i \in I_{WB}, \\ 0 & \text{if} \quad i \in I_N, \end{cases} \qquad t_i^{acc} = \begin{cases} 0 & \text{if} \quad \pi_i \geq 0, \ i \in I_{WB}, \\ -\pi_i & \text{if} \quad \pi_i < 0, \ i \in I_{WB}, \\ 0 & \text{if} \quad i \in I_N, \end{cases}$$

$$w_j^{acc} = \begin{cases} \Delta_j & \text{if} \quad \Delta_j \geq 0, \ j \in J_N, \\ 0 & \text{if} \quad \Delta_j < 0, \ j \in J_N, \\ 0 & \text{if} \quad j \in J_{WB}, \end{cases} \qquad v_j^{acc} = \begin{cases} 0 & \text{if} \quad \Delta_j \geq 0, \ j \in J_N, \\ -\Delta_j & \text{if} \quad \Delta_j < 0 \ j \in J_N, \\ 0 & \text{if} \quad j \in J_{WB}. \end{cases}$$

Thus we can compute a duality gap for x and Λ^{acc} as

$$\beta(x, S_{WB}) \quad = \quad f(x) - \varphi(\Lambda^{acc}). \tag{10}$$

Taking into account the definition of the cost functions and the accompanying dual solution, we can show that

$$\begin{aligned}\beta = \beta(x, S_{WB}) &= \sum_{\Delta_j>0, j\in J_N} \Delta_j(u_j - x_j) + \sum_{\Delta_j<0, j\in J_N} \Delta_j(l_j - x_j) \\ &\quad + \sum_{\pi_i>0, i\in I_{WB}} \pi_i(U_i - A_i x) + \sum_{\pi_i<0, i\in I_{WB}} \pi_i(L_i - A_i x).\end{aligned} \tag{11}$$

Indeed,

$$\begin{aligned}\beta &= f(x) - \varphi(\Lambda^{acc}) \\ &= \frac{1}{2}x^T Dx + c^T x + \frac{1}{2}x^T Dx + U^T s^{acc} - L^T t^{acc} - l^T v^{acc} + u^T w^{acc} \\ &= (x^T D + \pi^T A + c^T)x - \pi^T Ax + U^T s^{acc} - L^T t^{acc} - l^T v^{acc} + u^T w^{acc} \\ &= -\Delta^T x + U^T s^{acc} - L^T t^{acc} - \pi^T Ax - l^T v^{acc} + u^T w^{acc} \\ &= \sum_{\Delta_j>0, j\in J_N} \Delta_j(u_j - x_j) + \sum_{\Delta_j<0, j\in J_N} \Delta_j(l_j - x_j) \\ &\quad + \sum_{\pi_i>0, i\in I_{WB}} \pi_i(U_i - A_i x) + \sum_{\pi_i<0, i\in I_{WB}} \pi_i(L_i - A_i x).\end{aligned}$$

Definition 4 *The number β defined by (11) is called a suboptimality estimate of a feasible solution x with respect to a working basis S_{WB}.*

It is evident that the estimate β (11) can be decomposed as

$$\beta = \beta_1 + \beta_2, \tag{12}$$

where

$$\begin{aligned}\beta_1 &= f(x) - f(x^\star) \geq 0, \\ \beta_2 &= \varphi(\Lambda^\star) - \varphi(\Lambda^{acc}) \geq 0,\end{aligned} \tag{13}$$

and $\Lambda^\star$ is an optimal dual feasible solution. It follows from (12) and (13) that the inequality

$$\beta_1 := f(x) - f(x^\star) \leq \beta \tag{14}$$

holds. Note that the components β_1 and β_2 are not independent because the number β_2 depends on x. It is evident that a feasible solution x is ε-optimal for a given $\varepsilon > 0$ if $\beta \leq \varepsilon$ for some working basis S_{WB}.

2.4 Classical and Working Basis Optimality Conditions

We discuss now the difference between two groups of optimality conditions, namely Theorem 2 and Theorem 3. Theorem 3 is formulated for a pair x and S_{WB} and contains the method of computing Lagrange vectors. Theorem 2 is formulated only for a vector x and does not contain a method of computing Lagrange vectors. Let us prove the next theorem as an analogy to the Karush-Kuhn-Tucker conditions:

Theorem 4 (Second Form of the Working Basis Optimality Conditions) *A feasible x is optimal in (1) if and only if there exists a working basis S_{WB} such the the conditions (8) and (9) hold for x and the working basis S_{WB}.*

Proof. Sufficiency. Let x be a feasible solution and a working basis S_{WB} exists such the the conditions (8) and (9) hold for x and the working basis S_{WB}. Then according to Theorem 3 x is optimal in (1).

Necessity. Let x be an optimal feasible solution of problem (1). It follows from Theorem 2 that there exist Lagrange vectors λ and $\mu = -Dx - A^T\lambda - c$ satisfying the Karush-Kuhn-Tucker conditions. Let

$$I_0 = \{i \in I : \lambda_i = 0\}, I_{WB} = I \backslash I_0;$$

$$J_0 = \{j \in J : \mu_j = 0\}, J_n = J \backslash J_0.$$

Denote

$$k := \text{rank} A(I_{WB}, J_0). \tag{15}$$

Two cases are possible

1. $k = |I_{WB}|$, and
2. $k < |I_{WB}|$.

Suppose that case 1) occurs. Obviously, a subset J_{WB} of the set J_0 exists such that the matrix $A(I_{WB}, J_{WB})$ is nonsingular. The feasible solution x and S_{WB} with $S_{WB} = \{I_{WB}, J_{WB}\}$, satisfies (8) and (9) with the vector $\pi = \lambda$.

Consider case 2). In view of condition $k < |I_{WB}|$ there exists a nonzero vector $\Delta\lambda(I_{WB})$ such that

$$\Delta\lambda^T(I_{WB})A(I_{WB}, J_0) = 0.$$

Let $\Delta\lambda(I_0) = 0$ and

$$\delta^T(J) \;\; = \;\; -\Delta\lambda^T(I)A(I, J). \tag{16}$$

We calculate $\sigma_0 = \min\{\sigma_{i_0}, \sigma_{j_0}\}$, where

$$\begin{aligned} \sigma_{i_0} &= \min_{i \in I_{WB}} \sigma_i, \sigma_i = \begin{cases} -\lambda_i/\Delta\lambda_i, & \text{if } \lambda_i\Delta\lambda_i < 0; \\ \infty, & \text{otherwise,} \end{cases} \quad i \in I_{WB}; \\ \sigma_{j_0} &= \min_{j \in J_n} \sigma_j, \sigma_j = \begin{cases} -\mu_j/\delta_j, & \text{if } \mu_j\delta_j < 0; \\ \infty, & \text{otherwise ,} \end{cases} \quad j \in J_n, \end{aligned} \tag{17}$$

and δ is calculated according to (16). By construction, $\sigma_0 > 0$. If $\sigma_0 = \infty$ then we replace $\Delta\lambda$ with $-\Delta\lambda$ and calculate σ_0 by (16) and (17) with the new $\Delta\lambda$. As a result we get $0 < \sigma_0 < \infty$.

We compute $\bar{\lambda} = \lambda + \sigma_0\Delta\lambda$ and $\bar{\mu} = \mu + \sigma_0\delta$. Note that for the new Lagrange multipliers relations (5) and (6) hold.

Suppose $\sigma_0 = \sigma_{i_0}$.Then by construction $\bar{\lambda}_{i_0} = 0$. Put $\bar{I}_{WB} = I_{WB} \backslash i_0$, $\bar{J}_0 = J_0$ and calculate the corresponding number $\bar{k}$ according to (15), and consider cases 1), 2) as described above.

If $\sigma_0 = \sigma_{j_0}$ then by construction $\delta_{j_0} \neq 0$ and $\bar{\mu}_{j_0} = 0$. Put $\bar{I}_{WB} = I_{WB}$, $\bar{J}_0 = J_0 \cup j_0$, and calculate the corresponding number $\bar{k}$ according to (15). Since $\delta_{j_0} \neq 0$, we get $\bar{k} = k + 1$.

We show that the process of constructing the working basis S_{WB} is finite. Indeed, whenever case 2) occur we either decrease the cardinality of the set $|I_{WB}|$ or increase the number k. Consequently, after finite number of iterations case 1) will occur and the working basis S_{WB} will be constructed. Theorem 4 is proved.

$\diamond$

The second form of the working basis optimality conditions refines the classical optimality conditions. It allows to find in the whole set of Lagrange multipliers the vector that corresponds to a working basis.

2.5 Superbasis

Contrary to LP where there is always a solution at a vertex of the feasible polyhedron provided a problem has a solution, in QP problems a solution may lie on a facet. This happens when, for example, the number of active constraints is smaller than the number of nonactive variables. Curvature information or a possibility that a solution may be not at a vertex of the feasible polyhedron must be taken into account in the method. This principle is characteristic for methods of nonlinear programming. In our method this is implemented by a superbasis.

Let us supplement the notion of a working basis S_{WB} by a superbasis J_{SB}. Denote $J_0 = \{j \in J_N : \Delta_j = 0\}$ and compose the KKT matrix $R = R(J \cup I, J \cup I)$ of the blocks

$$R = R(J \cup I, J \cup I) = \begin{pmatrix} D(J,J) & A^T(I,J) \\ A(I,J) & \mathbf{0} \end{pmatrix}.$$

Let J_{SB} be a subset of J_0 and R_{SB} denote submatrix $R(J_{WB} \cup J_{SB} \cup I_{WB},\, J_{WB} \cup J_{SB} \cup I_{WB})$ of R.

Definition 5 *A set $J_{SB} \subset J_0$ is called a superbasis if the matrix R_{SB} is nonsingular.*

If $J_{SB} \subset J_0$ is a superbasis then the KKT system

$$\begin{bmatrix} D(J_{WB} \cup J_{SB}, J_{WB} \cup J_{SB}) & A^T(I_{WB}, J_{WB} \cup J_{SB}) \\ A(I_{WB}, J_{WB} \cup J_{SB}) & 0 \end{bmatrix} \begin{bmatrix} x(J_{WB} \cup J_{SB}) \\ \pi(I_{WB}) \end{bmatrix} = \begin{bmatrix} r_1(J_{WB}) \\ r_2(I_{WB}) \end{bmatrix}$$

is solvable with respect to the basic and superbasic variables and the multiplier vector.
Given a superbasis we can keep zero values of the reduced costs $\bar{\Delta}_j$, $j \in J_{SB}$, calculated at a vector $\bar{x} = x + \Delta x$ and a working basis S_{WB}, for any nonbasic values Δx_j, $j \in J_{NN} = J_N \backslash J_{SB}$, and basic values Δx_j, $j \in J_{WB}$ from the tangent space $A(I_{WB}, J)\Delta x = 0$.

Thus we take curvature information into account with the help of the superbasis.

3 Iteration

We now describe an iteration of the method. The method starts from a feasible solution x, a working basis S_{WB} and a superbasis J_{SB}. The feasible solution x is provided by some feasible point procedure. As the initial working sets we can take the empty sets

$$I_{WB} = \emptyset, J_{WB} = \emptyset, J_{SB} = \emptyset.$$

An iteration consists in the transformation of the triple $\{x, S_{WB}, J_{SB}\}$ to $\{\bar{x}, \bar{S}_{WB}, \bar{J}_{SB}\}$. The method is based on several principles.

- At each iteration the current primal point x is feasible.
- At each iteration the dual solution accompanying the working basis is feasible.
- The duality gap does not increase during the solution process, and the solution process is stopped when the duality gap is zero (or less then some suboptimality tolerance ε).
- The basis is changed by changing the accompanying dual solution.
- Dual information is used for an active set strategy.
- Curvature is taken into account by the superbasis.

Following the principle of decreasing the suboptimality estimate and using decomposition (12) we compose the iteration of two parts: decreasing a suboptimality estimate by changing a feasible solution and decreasing a suboptimality estimate by changing working basis sets.

A new feasible solution is given by

$$\bar{x} = x + \Theta d,$$

where d is a search direction and Θ is a suitable steplength.

The components d_j, $j \in J_{NN}$, of the direction d are taken as in the adaptive method of linear programming [4], [3]. The idea of the adaptive method is as follows:

- assume that current working basis provides correct guesses of optimal active sets at a primal solution $x^\star$;
- based on working basis construct primal solution κ (possibly unfeasible) and dual solution Π (possibly unfeasible) such that $f(\kappa) = \varphi(\kappa, \Pi)$;
- take step towards κ and Π in the feasibility region;
- if κ, Π are feasible then κ is optimal in the QP (1).

Note that contrary to the simplex method of LP, where we iterate with feasible primal solution and infeasible dual solution that satisfy complementarity condition until we get dual feasibility, in the adaptive method of LP we iterate with feasible primal and feasible dual solutions that do not satisfy complementarity condition until we get the complementarity condition fulfilled.

Implementation of this idea leads to the following computations:

$$d = \kappa - x,$$

where "optimal choice" of bounds for nonbasic κ_j, $j \in J_{NN} = J_N \backslash J_{SB}$, and for basic constraints B_i, $i \in I_{WB}$, is as follows

$$\begin{aligned} \kappa_j = l_j, &\quad \text{if} \quad \Delta_j < 0, \\ \kappa_j = u_j, &\quad \text{if} \quad \Delta_j \geq 0,\ j \in J_{NN}; \\ B_i = L_i, &\quad \text{if} \quad \pi_i < 0, \\ B_i = U_i, &\quad \text{if} \quad \pi_i \geq 0,\ i \in I_{WB}; \end{aligned}$$

basic components $\kappa(J_{SB})$ and $\kappa(J_{WB})$ and new multiplier vector $\Pi(I_{WB})$ solve the linear (KKT) system:

$$\begin{aligned} D(J_{WB},J)\kappa + A^T(I_{WB},J_{WB})\Pi(I_{WB}) &= -c(J_{WB}), \\ D(J_{SB},J)\kappa + A^T(I_{WB},J_{SB})\Pi(I_{WB}) &= -c(J_{SB}), \\ A(I_{WB},J)\kappa &= B(I_{WB}) \end{aligned} \tag{18}$$

One possible form of solving system (18) yields

$$\begin{aligned} H_{SB}\kappa(J_{SB}) &= \tilde{r}_1, \\ A_{WB}\kappa(J_{WB}) &= \tilde{r}_2, \\ A^T_{WB}\Pi(I_{WB}) &= \tilde{r}_3, \end{aligned}$$

for some right hand side $\tilde{r}$ where $H = H(J,J)$ is the reduced Hessian

$$\begin{aligned} H(J_N,J_N) &= D(J_N,J_N) - A^T(I_{WB},J_N)(A^{-1}_{WB})^T D(J_{WB},J_N) \\ &\quad -D(J_N,J_{WB})A^{-1}_{WB}A(I_{WB},J_N) \\ &\quad +A^T(J_{WB},J_N)(A^{-1}_{WB})^T D(J_{WB},J_{WB})A^{-1}_{WB}A(I_{WB},J_N), \\ H_{SB} &= H(J_{SB},J_{SB}),\ J_{SB} \subset J_N. \end{aligned}$$

The matrix H_{SB} is nonsingular as a result of the decomposition of the nonsingular matrix R_{SB} with respect to the nonsingular blocks A_{WB} and A^T_{WB}.

The steplength Θ along d is determined as

$$\Theta = \min\{1, \Theta_{j_0}, \Theta_{i_0}, \sigma_{j_0}, \sigma_{i_0}\}.$$

The five possible steplengths can be derived as follows.

i) The number $\Theta = 1$ gives the maximal step allowed by the box constraints at the components of the feasible solution with the indices from J_{NN}.

ii) $\Theta_{j_0} = \min\{\Theta_j,\ j \in J_{SB} \cup J_{WB}\}$ is the maximal step allowed by the box constraints at the components of the feasible solution with the indices from $J_{SB} \cup J_{WB}$

$$\Theta_j = \begin{cases} (u_j - x_j)/d_j, & \text{if } d_j > 0; \\ (l_j - x_j)/d_j, & \text{if } d_j < 0; \\ \infty, & \text{if } d_j = 0, \end{cases} \quad \text{for } j \in J_{SB} \cup J_{WB}.$$

iii) $\Theta_{i_0} = \min\{\Theta_i,\ i \in I_N\}$ is the maximal step allowed by the main constraints with the indices from I_N

$$\Theta_i = \begin{cases} (U_i - A_i x)/A_i d, & \text{if } A_i d > 0; \\ (L_i - A_i x)/A_i d, & \text{if } A_i d < 0; \\ \infty, & \text{if } A_i d = 0 \end{cases} \quad \text{for } i \in I_N = I \backslash I_{WB}.$$

iv) $\sigma_{j_0} = \min\ \sigma_j,\ j \in J_{NN}$, is the maximal step for which reduced costs with the indices $j \in J_{NN}$ keep their signs and the accompanying dual solution remains feasible

$$\sigma_j = \begin{cases} -\Delta_j/\delta_j, & \Delta_j\delta_j < 0; \\ 0, & \Delta_j = 0,\ \delta_j > 0,\ x_j \neq u_j; \\ 0, & \Delta_j = 0,\ \delta_j < 0,\ x_j \neq l_j; \\ \infty, & \text{in other cases}, \end{cases} \quad j \in J_{NN}. \tag{19}$$

v) $\sigma_{i_0} = \min\{\sigma_i,\ i \in I_{WB}\}$ is the maximal step for which multipliers $\pi_i,\ i \in I_{WB}$, keep their signs and the accompanying dual solution remains feasible

$$\sigma_i = \begin{cases} -\pi_i/\nu_i, & \pi_i\nu_i < 0, \\ 0, & \pi_i = 0,\ \nu > 0,\ A_i x \neq U_i, \\ 0, & \pi_i = 0,\ \nu < 0,\ A_i x \neq L_i, \\ \infty, & \text{in other cases}, \end{cases} \quad i \in I_{WB}. \tag{20}$$

In relations (19), (20)

$$\begin{aligned} \delta = (\delta_j, j \in J) &= -Dd - A^T(I_{WB}, J)\nu(I_{WB}), \\ \nu(I_{WB}) &= \Pi(I_{WB}) - \pi(I_{WB}) \end{aligned}$$

define the search directions for the vectors of reduced costs and multipliers corresponding to the direction d.

We set new feasible primal and dual solutions $\bar{x} = x + \Theta d$, $\bar{\Delta} = \Delta + \Theta\delta$ and $\bar{\pi} = \pi + \Theta v$. The suboptimality estimate of the feasible solution $\bar{x}$ with respect to the working basis S_{WB} (or duality gap) is equal to

$$\bar{\beta} = (1 - \Theta)\beta - \Theta(1 - \Theta)d^T Dd \leq \beta.$$

The inequalities $\bar{\beta}_1 \leq \beta_1$ and $\bar{\beta}_2 \leq \beta_2$ hold for the components $\bar{\beta}_1 = f(\bar{x}) - f(x^\star)$, $\bar{\beta}_2 = \varphi(\Lambda^\star) - \varphi(\bar{\Lambda}^{acc})$, where $\bar{\Lambda}^{acc}$ is the dual feasible solution accompanying $\bar{x}$ and S_{WB}. In fact $\bar{\beta}_1 \leq \beta_1$ according to the rules of constructing the variation Θd. The inequality $\bar{\beta}_2 \leq \beta_2$ follows from

$$\Delta\beta_2 = \bar{\beta}_2 - \beta_2 = \varphi(\Lambda^{acc}) - \varphi(\bar{\Lambda}^{acc}) = -\Theta(1 - \frac{1}{2}\Theta)d^T Dd \leq 0.$$

If $\Theta = 1$ the solution process stops at the optimal feasible solution $\bar{x}$. If $\bar{\beta} \leq \varepsilon$, then $\bar{x}$ is ε-optimal in (1). Otherwise we continue the iteration by changing the basis sets.

The following five cases are possible after the first part of the iteration

1. $\Theta = \Theta_{j_0},\ j_0 \in J_{SB}$,
2. $\Theta = \Theta_{j_0},\ j_0 \in J_{WB}$,
3. $\Theta = \Theta_{i_0},\ i_0 \in I_N$,
4. $\Theta = \sigma_{j_0},\ j_0 \in J_{NN}$,
5. $\Theta = \sigma_{i_0},\ i_0 \in I_{WB}$.

Let us formulate the rules for changing the basis sets for each case.

1. We drop index $j_0 \in J_{SB}$ from the superbasis:

$$\bar{J}_{SB} = J_{SB} \backslash j_0,$$

and do not change the working basis

$$\bar{S}_{WB} = S_{WB}.$$

The matrix $\bar{H}_{SB} = H(\bar{J}_{SB}, \bar{J}_{SB})$ is nonsingular as it is obtained from the positive definite matrix H_{SB} by removing the j_0-th row and the j_0-th column. We pass to a new iteration with the new working sets $\bar{S}_{WB}, \bar{J}_{SB}$.

2. Construct the vector

$$h^T(J_{SB}) = A_{WB}^{-1}(j_0, I_{WB})A(I_{WB}, J_{SB}),$$

if $J_{SB} \neq \emptyset$. Two situations are possible:

(a) $J_{SB} \neq \emptyset$ and there exists an index $k \in J_{SB}$ such that $h_k \neq 0$;

(b) $J_{SB} = \emptyset$ or $h(J_{SB}) \equiv 0$.

In case 2a) the index j_0 is dropped from the working basis in two stages. In the first stage we exchange the indices j_0 and k between the working basis and superbasis:

$$\tilde{J}_{WB} = (J_{WB} \backslash j_0) \cup k, \tilde{J}_{SB} = (J_{SB} \backslash k) \cup j_0.$$

These operations mean that in the matrix R_{SB} rows and columns with indices k and j_0 are exchanged. The matrix $\tilde{A}_{WB} = A(I_{WB}, \tilde{J}_{WB})$ is nonsingular because $h_k \neq 0$. The matrix $\tilde{H}_{SB} = \tilde{H}(\tilde{J}_{SB}, \tilde{J}_{SB})$ is nonsingular as a result of decomposing the nonsingular matrix R_{SB} with respect to the nonsingular blocks $\tilde{A}_{WB}$ and $\tilde{A}_{WB}^T$.

After the index exchange we obtain the situation of case 1, i.e. $j_0 \in \tilde{J}_{SB}$. The second stage of working basis sets change consists then in the operations described for case 1.

The change of the working basis in case 2b) is based on the information about the dual problem (7). As it was mentioned in the previous section, we can construct the dual feasible solution $\bar{\Lambda} = \bar{\Lambda}(\bar{x}, S_{WB}) = (\bar{x}, \bar{s}, \bar{t}, \bar{w}, \bar{s})$ accompanying the pair $\bar{x}, S_{WB}$. It follows from (10) that by improving the accompanying dual feasible solution (and hence increasing the value of the dual cost function) we decrease the suboptimality estimate. Fixing the component $\bar{x}$ of the accompanying dual feasible solution we get a linear problem and use the rule of the "long dual step" [4], [3], [10] to improve the dual feasible solution $\bar{\Lambda}$. The search directions p, q for the vectors of reduced costs and multipliers are determined as follows

$$\begin{array}{rcl}
p_{j_0} & = & -\text{sign } d_{j_0}, \\
p(J_{WB} \cup J_{SB} \backslash j_0) & \equiv & 0, \\
p^T(J_{NN}) & = & -q^T(I_{WB})A(I_{WB}, J_{NN}), \\
q^T(I_{WB}) & = & -p_{j_0} A_{WB}^{-1}(j_0, I_{WB}); \\
q(I_N) & \equiv & 0.
\end{array}$$

Calculate the steps $\tilde{\sigma}_j,\ j \in J_{NN}$; $\tilde{\sigma}_i,\ i \in I_{WB}$:

$$\tilde{\sigma}_j = \begin{cases} -\bar{\Delta}_j/p_j, & \bar{\Delta}_j p_j < 0, \\ 0, & \bar{\Delta}_j = 0,\ p_j < 0,\ \bar{x}_j \neq l_j, \\ & \text{or } \bar{\Delta}_j = 0,\ p_j > 0,\ \bar{x}_j \neq u_j, \\ \infty, & \text{in other cases,} \end{cases} \tag{21}$$

for all $j \in J_{NN}$;

$$\tilde{\sigma}_i = \begin{cases} -\bar{\pi}_i/q_i, & \pi_i q_i < 0, \\ 0, & \bar{\pi}_i = 0,\ q_i < 0,\ A_i\bar{x} \neq L_i, \\ & \text{or } \bar{\pi}_i = 0,\ q_i > 0,\ A_i\bar{x} \neq U_i, \\ \infty, & \text{in other cases,} \end{cases} \tag{22}$$

for all $i \in I_{WB}$,

and sort finite steplengths in nondecreasing order, i.e.

$$\tilde{\sigma}_{k_1} \leq \tilde{\sigma}_{k_2} \leq \ldots \leq \tilde{\sigma}_{k_s} < \infty,\ s \leq |J_{NN}| + |I_{WB}|.$$

We set $\alpha_0 = |d_{j_0}| > 0$ and compute α_i, $i = 1, \ldots, s$, by

$$\alpha_{i+1} = \alpha_i - \begin{cases} |p_{k_{i+1}}|(u_{k_{i+1}} - l_{k_{i+1}}), & \text{if } k_{i+1} \in J_{NN}, \\ |q_{k_{i+1}}|(U_{k_{i+1}} - L_{k_{i+1}}), & \text{if } k_{i+1} \in I_{WB}, \end{cases} \quad i = 0, 1, \ldots, s-1.$$

Since the problem (1) is feasible we have $\alpha_s \leq 0$. Consequently, there exists an index r, $0 < r \leq s$, such that $\alpha_{r-1} > 0$ and $\alpha_r \leq 0$.

We compute new multipliers and reduced costs

$$\begin{aligned} \bar{\bar{\Delta}} &= \bar{\Delta} + \tilde{\sigma}_{k_r} p, \\ \bar{\bar{\pi}} &= \bar{\pi} + \tilde{\sigma}_{k_r} q. \end{aligned} \tag{23}$$

If $k_r \in I_{WB}$, we correct the working basis sets:

$$\begin{aligned} \bar{J}_{WB} &= J_{WB} \backslash j_0, \\ \bar{I}_{WB} &= I_{WB} \backslash k_r, \\ \bar{J}_{SB} &= J_{SB}. \end{aligned}$$

If $k_r \in J_{NN}$ we change the working basis sets by the rules

$$\begin{aligned} \bar{J}_{WB} &= (J_{WB} \backslash j_0) \cup k_r, \\ \bar{I}_{WB} &= I_{WB}, \\ \bar{J}_{SB} &= J_{SB}. \end{aligned}$$

We can show that the matrix $\bar{R}_{SB}$ is nonsingular, where

$$\bar{R}_{SB} = R(\bar{J}_{WB} \cup \bar{J}_{SB} \cup I_{WB}, \bar{J}_{WB} \cup \bar{J}_{SB} \cup I_{WB}). \tag{24}$$

The suboptimality estimate of the feasible solution $\bar{x}$ with respect to $\bar{S}_{WB}$ is equal to

$$\begin{aligned} \bar{\bar{\beta}} &= \bar{\beta}_1 + \bar{\bar{\beta}}_2, \\ \bar{\bar{\beta}}_2 &= \bar{\beta}_2 - \sum_{i=1}^{r-1} \alpha_i \tilde{\sigma}_{k_i}, \end{aligned} \tag{25}$$

and the following inequalities are true:

$$\begin{aligned} \bar{\bar{\beta}} &\leq \bar{\beta}, \\ \bar{\bar{\beta}}_2 &\leq \bar{\beta}_2. \end{aligned}$$

This finishes the iteration.

3. Construct the vector

$$h^T(J_{SB}) = A(i_0, J_{SB}) - A(i_0, J_{WB}) A_{WB}^{-1} A(I_{WB}, J_{SB}),$$

if $J_{SB} \neq \emptyset$. Two situations are possible:

(a) $J_{SB} \neq \emptyset$ and there exists an index $k \in J_{SB}$, such that $h_k \neq 0$;

(b) $J_{SB} \neq \emptyset$ or $h(J_{SB}) \equiv 0$.

In case 3a) the working basis sets are changed by the rules:

$$\begin{aligned} \bar{J}_{WB} &= J_{WB} \cup k, \\ \bar{I}_{WB} &= I_{WB} \cup i_0, \\ \bar{J}_{SB} &= J_{SB} \backslash k. \end{aligned}$$

To change the working basis sets in case 3b) we use the information about the dual problem and improve the accompanying dual feasible solution $\bar{\Lambda}$. Construct the directions p, q and the number α_0:

$$\begin{aligned} q_{i_0} &= -\text{sign} A_{i_0} d, \\ q(I_N \backslash i_0) &\equiv 0; \\ q^T(I_{WB}) &= -q_{i_0} A(i_0, J_{WB}) A_{WB}^{-1}; \\ p^T(J_{NN}) &= -q^T(I) A(I, J_{NN}); \\ p(J_{WB} \cup J_{SB}) &\equiv 0; \\ \alpha_0 &= |A_{i_0} d|, \end{aligned}$$

and calculate the numbers $\tilde{\sigma}_j$, $j \in J_{NN}$, by (21) and $\tilde{\sigma}_i$, $i \in I_{WB}$, by (22). Following the operations of case 2b) we compute the numbers α_i, $i = 1, \ldots, s$, and find the index k_r. Then we compute new multipliers and reduced costs by (23). If $k_r \in I_{WB}$ we set

$$\begin{aligned} \bar{J}_{WB} &= J_{WB}, \\ \bar{I}_{WB} &= (I_{WB} \backslash k_r) \cup i_0, \\ \bar{J}_{SB} &= J_{SB}. \end{aligned}$$

If $k_r \in J_{NN}$ we set

$$\begin{aligned}\bar{J}_{WB} &= J_{WB} \cup k_r,\\ \bar{I}_{WB} &= I_{WB} \cup i_0,\\ \bar{J}_{SB} &= J_{SB}.\end{aligned}$$

It can be shown that new matrix R_{SB} (24) is nonsingular in both cases 3a) and 3b). The suboptimality estimate of the feasible solution $\bar{x}$ with respect to $\bar{S}_{WB}$ is given by (25). We pass to a new iteration.

4. Following the principle of taking into account curvature information we include the index j_0 into the superbasis, i.e.

$$\bar{J}_{SB} = J_{SB} \cup j_0.$$

The working basis is not changed, i.e.

$$\bar{J}_{WB} = J_{WB},\ \bar{I}_{WB} = I_{WB}.$$

The new matrix $\bar{R}_{SB}$ (24) is nonsingular.

5. We find an index $j_0 \in J_{WB}$ such that $A^{-1}_{WB}(j_0, i_0) \neq 0$. The new basis is given by

$$\begin{aligned}\bar{J}_{WB} &= J_{WB} \backslash j_0,\\ \bar{I}_{WB} &= I_{WB} \backslash i_0,\\ \bar{J}_{SB} &= J_{SB} \cup j_0.\end{aligned}$$

The new matrix $\bar{R}_{SB} = R(\bar{J}_{WB} \cup \bar{J}_{SB} \cup I_{WB}, \bar{J}_{WB} \cup \bar{J}_{SB} \cup I_{WB})$ is nonsingular.

This completes the description of an iteration of our method.

Remark 1 *In case of linear programming ($D = 0$) the described algorithm coincides with the "adaptive method" for linear programming [3, 4].*

Remark 2 *The basis up-dates correspond to the exchange of rows and/or columns in the KKT matrix. Thus updating techniques for matrix factorizations may be efficiently used.*

4 Computational experiments

The method has been coded in FORTRAN77. To test ideas and principles we have conducted several numerical experiments.

In the experiments, test problems

$$\begin{aligned}\min f(x) &= c^T x + \frac{1}{2} x^T D x \\ & A_i x \leq U_i,\ i = 1, ..., m,\\ & l_j \leq x_j \leq u_j,\ j = 1, ..., n\end{aligned} \tag{26}$$

with known optimal solution $x^\star$ and known optimal solution characteristics such as the number of active variables ($x^\star_j = l_j$ or $x^\star_j = u_j$) and the number of active inequality constraints ($A_i x^\star = U_i$), have been generated.

The following generator was used. The input data for each problem is:

n — the number of variables,

m — the number of constraints ("less than" type),

m_* — the number of active inequality constraints at the point $x^\star$, $0 \le m_* \le m$,

n_* — the number of active variables at the optimal solution $x^\star$.

For each problem the generator starts with generating values of an optimal solution, e.g.

$$x_j^\star = 1, j = 1, \ldots, n.$$

Then we generate reduced costs $\Delta_j(x^\star), j = 1, \ldots, n,$ and multipliers $\pi_i(x^\star), i = 1, \ldots, m,$ corresponding to the optimal solution $x^\star$, namely

$$\begin{aligned} \Delta_j(x^\star) &\neq 0, \ j = 1, \ldots, n_*, \\ \Delta_j(x^\star) &= 0, \ j = n_* + 1, \ldots, n; \\ \pi_i(x^\star) &> 0, \ i = 1, \ldots, m_*, \\ \pi_i(x^\star) &= 0, \ i = m_* + 1, \ldots, m. \end{aligned}$$

Using this information we generate vectors of box constraints l and u and some initial feasible solution x^I as follows

$$\begin{aligned} &l_j = x_j^\star, \ u_j > x_j^\star, \quad \text{if} \quad \Delta_j(x^\star) < 0; \\ &l_j < x_j^\star, \ u_j = x_j^\star, \quad \text{if} \quad \Delta_j(x^\star) > 0; \\ &l_j < x_j^\star, \ u_j > x_j^\star, \quad \text{if} \quad \Delta_j(x^\star) = 0; \ j \in J; \\ &x^I = \alpha u + (1 - \alpha) l \end{aligned}$$

where $\alpha \in [0, 1]$ is some random number. The positive definite matrix D is given by

$$D = G^T G$$

where G is a randomly generated matrix. Now we construct the matrix A and the vector U so that the relations

$$\begin{aligned} &A_i x^\star = U_i, \ A_i x^I = U_i, \ i = 1, \ldots, m_*; \\ &A_i x^\star < U_i, \ A_i x^I < U_i, \ i = m_* + 1, \ldots, m, \end{aligned}$$

hold for the optimal and for the initial feasible solutions. We obtain this in the following way: all the elements of the matrix A except for the first column are randomly generated, the elements of the first column and the vector U are given by:

$$\begin{aligned} &A(i,1) = \frac{1}{x_1^\star - x_1^I}(A(i, J\backslash 1) x^I(J\backslash 1) - A(i, J\backslash 1) x^\star(J\backslash 1)), \ i = 1, \ldots, m; \\ &U_i = A_i x^\star, \ i = 1, \ldots, m_*, \\ &U_i > A_i x^\star, \ i = m_* + 1, \ldots, m. \end{aligned}$$

The vector c is set by

$$c = -D x^\star - A^T \pi(x^\star) - \Delta(x^\star).$$

Concrete numbers have been generated from the interval $[-10, 10]$ using the random number generator URAND [2].

The aim of the experiments is to compare our method with Powell's implementation (ZQPCVX) [11] of the dual method [9], that has been developed for solving medium size problems with dense matrices, and with the implementation QPSOL [6] of a primal active set method. As our method is a primal one, a feasible point procedure has to be executed before starting the solution process.

In the experiments, we consider two types of available initial information about the problem:
a) The initial feasible solution x^I is given.
b) There is no information about any feasible solution.

In case b), we solve an augmented quadratic programming problem instead of the initial one:

$$\begin{aligned} \min \qquad & c^T x + M \sum_{i=1}^{m} e_i x_{n+i} + \frac{1}{2} x^T D x, \\ & A_i x + x_{n+i} \le U_i, i = 1, ..., m, \\ & l \le x \le u, \\ & l_{n+i} \le x_{n+i} \le u_{n+i}, i = 1, ..., m, \end{aligned} \tag{27}$$

where

$$\begin{array}{lllll} e_i = -1, & l_{n+i} = x^{I,2}_{n+i}, & u_{n+i} = 0, & \text{if} & x^{I,2}_{n+i} < 0, \\ e_i = 0, & l_{n+i} = 0, & u_{n+i} = 0, & \text{if} & x^{I,2}_{n+i} \ge 0, \quad \text{for } i = 1, ..., m, \end{array}$$

$x^{I,2}$ is some point satisfying $l \le x^{I,2} \le u,\ x^{I,2}_{n+i} = U_i - A_i x^{I,2},\ i = 1, ..., m,$ and $M > 0$ is some large number chosen so that at the solution $(x^\star, x^\star_{n+i},\ i \in I)$ of (27) the residuals of the constraints become equal to zero:

$$x^\star_{n+i} = 0, i \in I.$$

In this case, $x^\star$ is optimal for (26). Thus instead of the original problem with m constraints and n variables, we solve the augmented $m \times (n + m)$ problem (27) with a special structure.

The results of the experiments are given in Table 1. The columns "QP-0" and "QP-1" of the table present numbers of iterations of QPSOL necessary to solve problems (26) starting from the initial feasible point x^I and infeasible point $x^{I,2} = (u + l)/2$ respectively. The column "A-0" of the table contains the number of iterations necessary to solve problems of type (26) by the suggested method starting from the initial feasible point x^I and the empty working basis. Numbers of iteration necessary to solve problems of type (27) by the described method starting from the point $x^{I,2}$ and the "empty" working basis set are given in the column "A-1". For the comparison we include the column "QP-2" with numbers of iterations of QPSOL necessary to solve (1) starting from not feasible point x^I and the "empty" active sets.

n	m	m_*	n_*	ZQPCVX	QP-0	QP-1	QP-2	A-0	A-1
50	150	5	45	150	218	201	727	138	179
50	100	5	45	138	159	405	541	179	170
50	50	5	45	105	89	93	310	105	129
50	150	10	40	155	988	911	393	163	180
50	100	10	40	134	581	532	486	136	130
50	50	10	40	106	82	63	394	113	125
50	150	20	30	200	722	237	744	158	192
50	100	20	30	165	128	222	580	158	172
50	50	20	30	162	82	57	432	136	142
50	150	30	20	188	593	493	1336	238	310
50	100	30	20	216	115	42	755	187	214
50	50	30	20	158	60	37	504	147	172
50	150	5	40	161	239	794	397	146	212
50	100	5	40	115	735	104	691	161	175
50	50	5	40	121	82	67	392	133	123
50	150	10	35	168	760	141	1397	149	179
50	100	10	35	129	263	134	1174	111	133
50	50	10	35	106	70	73	402	137	139
50	150	20	25	192	247	611	824	155	236
50	100	20	25	155	432	64	646	150	199
50	50	20	25	138	68	71	471	134	149
50	150	30	15	185	569	180	945	217	313
50	100	30	15	167	267	47	775	194	189
50	50	30	15	146	52	34	526	136	163
50	150	5	50	157	953	932	487	116	141
50	100	5	50	148	491	715	478	90	143
50	50	5	50	103	98	111	276	86	77
50	150	10	45	177	229	735	1222	155	187
50	100	10	45	143	153	97	329	114	119
50	50	10	45	123	83	78	313	131	130
50	150	20	35	259	186	70	767	167	194
50	100	20	35	228	150	79	557	171	185
50	50	20	35	192	100	94	435	152	142
50	150	30	25	222	683	300	1169	189	298
50	100	30	25	194	556	55	801	165	240
50	50	30	25	157	69	43	476	126	170
				5763	11352	8922	23152	5343	6351

Table 1: Numerical results for problems with inequality and box constraints

The suggested method can also be applied to problems with only box constraints ($m = 0$). The results of experiments with such problems are given by Table 2.

n	n_*	ZQPCVX	QP-0	QP-1	A-0	A-1
50	50	55	50	50	24	24
50	45	48	50	50	33	31
50	40	42	43	43	32	34
100	100	105	102	101	44	46
100	95	101	95	95	47	51
100	90	94	96	93	60	62
200	200	209	196	197	110	106
200	195	202	193	193	109	105
200	190	194	191	191	112	114
		1050	1016	1013	571	573

Table 2: Numerical results for problems with box constraints

The results of the experiments can be summarized as follows. Comparing columns A-0 and A-1 we conclude that the absence of an initial feasible solution does not necessarily lead to the increase of the number of iterations. Comparing columns “QP-1”, “QP-2” and “A-1” we should admit how important are the better implementation of the first phase as well as construction of initial working sets rather than starting from empty ones. Also it is necessary to note that the method for solving (27) could be considerably improved because it can be implemented in a special way taking the structure of the augmented problem into account. The method shows better results for problems with many inequalities and for problems in which there are many active components in the optimal solution and relatively few constraints are active at the solution. We should admit that for problems with many equality constraints the primal active set methods are better than primal-dual or dual active set methods.

5 Acknowledgements

We thank the Deutsche Forschungsgemeinschaft (DFG) for financial support through SFB 539 and the Fund of Fundamental Research of Republic Belarus through “Mathematical Structures 21”.

References

[1] A. S. CHERNUSHEVICH, *Algorithms for solving linear-quadratic extremal problems*, Ph.D. thesis, Institute of Mathematics, BSSR Academy of Sciences, Minsk, 1987 (in Russian).

[2] G. E. FORSYTHE, M. A. MALCOLM AND C. B. MOLER, *Computer Methods for Mathematical Computations,* Prentice-hall, London, 1977.

[3] R. GABASOV, *Adaptive method of linear programming*, Preprints of the University of Karlsruhe, Institute of Statistics and Mathematics, Karlsruhe, Germany (1993).

[4] R. GABASOV, F. M. KIRILLOVA AND O. I. KOSTYUKOVA, *A method of solving general linear programming problems*, Doklady AN BSSR, **23**, 3 (1979), pp. 197–200 (in Russian).

[5] R. Gabasov, F. M. Kirillova and O. I. Kostyukova, *Solution of linear - quadratic extremal problems*, Doklady AN SSSR, **280**, 3 (1985), pp. 529-533 (in Rissian).

[6] P. E. Gill, W. Murray, M. A. Saunders and M. H. Wright, *User's Guide for SOL/QPSOL*, Report SOL 83-7, Department of Operations Research, Stanford University, California (1983).

[7] P. E. Gill, W. Murray and M. H. Wright, *Practical Optimization*, Academic Press Inc., London (1981).

[8] N. I. M. Gould and Ph. L. Toint, *A Quadratic Programming Bibliography*, `http://www.optimization-online.org/DB_HTML/2001/02/285.html`, (2001).

[9] D. Goldfarb and A. Idnani, *A numerically stable dual method for solving strictly convex quadratic programs*, Math. Progr., 27 (1983), pp. 1–33.

[10] E. Kostina, *The Long Step Rule in the Bounded-Variable Dual Simplex Method: Numerical Experiments*, Mathematical Methods of Operations Research, **55** (2002), I. 3.

[11] M. J. D. Powell, *ZQPCVX - A Fortran subroutine for convex quadratic programming*, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Report DAMTP/1983/NA17 (1983).