Chapter 3 Quadratic Programming

3.1 Constrained quadratic programming problems

A special case of the NLP arises when the objective functional f is quadratic and the constraints h, g are linear in $x \in \mathbb{R}^n$. Such an NLP is called a Quadratic Programming (QP) problem. Its general form is

$$\text{minimize} \quad f(x) := \frac{1}{2}\, x^T B x \;-\; x^T b \tag{3.1a}$$
$$\text{over} \quad x \in \mathbb{R}^n$$
$$\text{subject to} \quad A_1 x = c \tag{3.1b}$$
$$A_2 x \le d \;, \tag{3.1c}$$

where $B \in \mathbb{R}^{n\times n}$ is symmetric, $A_1 \in \mathbb{R}^{m\times n}$, $A_2 \in \mathbb{R}^{p\times n}$, and $b \in \mathbb{R}^n$, $c \in \mathbb{R}^m$, $d \in \mathbb{R}^p$.
As we shall see in this chapter, the QP (3.1a)-(3.1c) can be solved iteratively by active set strategies or interior point methods where each iteration requires the solution of an equality constrained QP problem.

3.2 Equality constrained quadratic programming

If only equality constraints are imposed, the QP (3.1a)-(3.1c) reduces to

$$\text{minimize} \quad f(x) := \frac{1}{2}\, x^T B x \;-\; x^T b \tag{3.2a}$$
$$\text{over} \quad x \in \mathbb{R}^n$$
$$\text{subject to} \quad Ax = c \;, \tag{3.2b}$$

where $A \in \mathbb{R}^{m\times n}$, $m \le n$. For the time being we assume that A has full row rank m.
The KKT conditions for the solution $x^* \in \mathbb{R}^n$ of the QP (3.2a),(3.2b) give rise to the following linear system

$$\underbrace{\begin{pmatrix} B & A^T \\ A & 0 \end{pmatrix}}_{=:\,K} \begin{pmatrix} x^* \\ \lambda^* \end{pmatrix} = \begin{pmatrix} b \\ c \end{pmatrix} \;, \tag{3.3}$$

where $\lambda^* \in \mathbb{R}^m$ is the associated Lagrange multiplier.
We denote by $Z \in \mathbb{R}^{n\times(n-m)}$ the matrix whose columns span $\mathrm{Ker} A$, i.e., $AZ = 0$.

Definition 3.1 KKT matrix and reduced Hessian
The matrix K in (3.3) is called the KKT matrix and the matrix $Z^T BZ$ is referred to as the reduced Hessian.

Lemma 3.2 Existence and uniqueness
Assume that $A \in \mathbb{R}^{m\times n}$ has full row rank $m \le n$ and that the reduced Hessian $Z^T BZ$ is positive definite. Then, the KKT matrix K is nonsingular. Hence, the KKT system (3.3) has a unique solution (x^*, λ^*).

Proof: The proof is left as an exercise.

•

Under the conditions of the previous lemma, it follows that the second order sufficient optimality conditions are satisfied so that x^* is a strict local minimizer of the QP (3.2a),(3.2b). A direct argument shows that x^* is in fact a global minimizer.

Theorem 3.3 Global minimizer
Let the assumptions of Lemma 3.2 be satisfied and let (x^*, λ^*) be the unique solution of the KKT system (3.3). Then, x^* is the unique global solution of the QP (3.2a),(3.2b).

Proof: Let $x \in \mathcal{F}$ be a feasible point, i.e., $Ax = c$, and $p := x^* - x$. Then, $Ap = 0$. Substituting $x = x^* - p$ into the objective functional, we get

$$\begin{aligned} f(x) &= \frac{1}{2}\,(x^* - p)^T B(x^* - p) \;-\; (x^* - p)^T b \;= \\ &= \frac{1}{2}\,p^T Bp \;-\; p^T Bx^* \;+\; p^T b \;+\; f(x^*) \;. \end{aligned}$$

Now, (3.3) implies $Bx^* = b - A^T\lambda^*$. Observing $Ap = 0$, we have

$$p^T Bx^* \;=\; p^T(b \;-\; A^T\lambda^*) \;=\; p^T b \;-\; \underbrace{(Ap)^T\lambda^*}_{=\,0} \;,$$

whence

$$f(x) \;=\; \frac{1}{2}\,p^T Bp \;+\; f(x^*) \;.$$

In view of $p \in \text{Ker } A$, we can write $p = Zu\;,\; u \in \mathbb{R}^{n-m}$, and hence,

$$f(x) \;=\; \frac{1}{2}\,u^T Z^T BZu \;+\; f(x^*) \;.$$

Since $Z^T BZ$ is positive definite, we deduce $f(x) > f(x^*)$. Consequently, x^* is the unique global minimizer of the QP (3.2a),(3.2b).

•

3.3 Direct solution of the KKT system

As far as the direct solution of the KKT system (3.3) is concerned, we distinguish between symmetric factorization and the range-space and null-space approach.

3.3.1 Symmetric indefinite factorization

A possible way to solve the KKT system (3.3) is to provide a symmetric factorization of the KKT matrix according to

$$P^T K P \; = \; L D L^T \; , \tag{3.4}$$

where P is an appropriately chosen permutation matrix, L is lower triangular with $\mathrm{diag}(L) = I$, and D is block diagonal.
Based on (3.4), the KKT system (3.3) is solved as follows:

$$\text{solve} \quad Ly \; = \; P^T \begin{pmatrix} b \\ c \end{pmatrix} \; , \tag{3.5a}$$

$$\text{solve} \quad D\hat{y} \; = \; y \; , \tag{3.5b}$$

$$\text{solve} \quad L^T \tilde{y} \; = \; \hat{y} \; , \tag{3.5c}$$

$$\text{set} \quad \begin{pmatrix} x^* \\ \lambda^* \end{pmatrix} \; = \; P\tilde{y} \; . \tag{3.5d}$$

3.3.2 Range-space approach

The range-space approach applies, if $B \in \mathbb{R}^{n \times n}$ is symmetric positive definite. Block Gauss elimination of the primal variable x^* leads to the Schur complement system

$$AB^{-1}A^T \lambda^* \; = \; AB^{-1}b \; - \; c \tag{3.6}$$

with the Schur complement $S \in \mathbb{R}^{m \times m}$ given by $S := AB^{-1}A^T$. The range-space approach is particularly effective, if

- B is well conditioned and easily invertible (e.g., B is diagonal or block-diagonal),
- B^{-1} is known explicitly (e.g., by means of a quasi-Newton updating formula),
- the number m of equality constraints is small.

3.3.3 Null-space approach

The null-space approach does not require regularity of B and thus has a wider range of applicability than the range-space approach.
We assume that $A \in \mathbb{R}^{m\times n}$ has full row rank m and that $Z^T BZ$ is positive definite, where $Z \in \mathbb{R}^{n\times(n-m)}$ is the matrix whose columns span Ker A which can be computed by QR factorization (cf. Chapter 2.4).
We partition the vector x^* according to

$$x^* = Yw_Y + Zw_Z , \tag{3.7}$$

where $Y \in \mathbb{R}^{n\times m}$ is such that $[Y\ Z] \in \mathbb{R}^{n\times n}$ is nonsingular and $w_Y \in \mathbb{R}^m$, $w_Z \in \mathbb{R}^{n-m}$.
Substituting (3.7) into the second equation of (3.3), we obtain

$$Ax^* = AYw_Y + \underbrace{AZ}_{=\,0} w_Z = c , \tag{3.8}$$

i.e., Yw_Y is a particular solution of $Ax = c$.
Since $A \in \mathbb{R}^{m\times n}$ has rank m and $[Y\ Z] \in \mathbb{R}^{n\times n}$ is nonsingular, the product matrix $A[Y\ Z] = [AY\ 0] \in \mathbb{R}^{m\times m}$ is nonsingular. Hence, w_Y is well determined by (3.8).
On the other hand, substituting (3.7) into the first equation of (3.3), we get

$$BYw_Y + BZw_Z + A^T\lambda^* = b .$$

Multiplying by Z^T and observing $Z^T A^T = (AZ)^T = 0$ yields

$$Z^T BZw_Z = Z^T b - Z^T BYw_Y . \tag{3.9}$$

The reduced KKT system (3.9) can be solved by a Cholesky factorization of the reduced Hessian $Z^T BZ \in \mathbb{R}^{(n-m)\times(n-m)}$. Once w_Y and w_Z have been computed as the solutions of (3.8) and (3.9), x^* is obtained according to (3.7).
Finally, the Lagrange multiplier turns out to be the solution of the linear system arising from the multiplication of the first equation in (3.7) by Y^T:

$$(AY)^T\lambda^* = Y^T b - Y^T Bx^* . \tag{3.10}$$

3.4 Iterative solution of the KKT system

If the direct solution of the KKT system (3.3) is computationally too costly, the alternative is to use an iterative method. An iterative solver can be applied either to the entire KKT system or, as in the range-space and null-space approach, use the special structure of the KKT matrix.

3.4.1 Krylov methods

The KKT matrix $K \in \mathbb{R}^{(n+m)\times(n+m)}$ is indefinite. In fact, if A has full row rank m, K has n positive and m negative eigenvalues. Therefore, for the iterative solution of (3.3) Krylov subspace methods like GMRES (Generalized Minimum RESidual) and QMR (Quasi Minimum Residual) are appropriate candidates.

3.4.2 Transforming range-space iterations

We assume $B \in \mathbb{R}^{n\times n}$ to be symmetric positive definite and suppose that $\tilde{B}$ is some symmetric positive definite and easily invertible approximation of B such that $\tilde{B}^{-1}B \sim I$.
We choose $K_L \in \mathbb{R}^{(n+m)\times(n+m)}$ as the lower triangular block matrix

$$K_L \; = \; \begin{pmatrix} I & 0 \\ -A\tilde{B}^{-1} & I \end{pmatrix} , \tag{3.11}$$

which gives rise to the regular splitting

$$K_L K \; = \; \underbrace{\begin{pmatrix} \tilde{B} & A^T \\ 0 & \tilde{S} \end{pmatrix}}_{=:\; M_1} - \underbrace{\begin{pmatrix} \tilde{B}(I - \tilde{B}^{-1}B) & 0 \\ A(I - \tilde{B}^{-1}B) & 0 \end{pmatrix}}_{=:\; M_2 \sim 0} , \tag{3.12}$$

where $\tilde{S} \in \mathbb{R}^{m\times m}$ is given by

$$\tilde{S} \; := \; - \, A\tilde{B}^{-1}A^T \; . \tag{3.13}$$

We set

$$\psi \; := \; (x,\lambda)^T \quad , \quad \alpha \; := \; (b,c)^T \; .$$

Given an iterate $\psi^{(0)} \in \mathbb{R}^{n+m}$, we compute $\psi^{(k)}$, $k \in \mathbb{N}$, by means of the transforming range-space iterations

$$\begin{aligned} \psi^{(k+1)} \; &= \; \psi^{(k)} \; + \; M_1^{-1}K_L(\alpha \; - \; K\psi^{(k)}) \; = \\ &= \; (I \; - \; M_1^{-1}K_LK)\psi^{(k)} \; + \; M_1^{-1}K_L\alpha \quad , \quad k \geq 0 \; . \end{aligned} \tag{3.14}$$

The transforming range-space iteration (3.14) will be implemented as follows:

$$d^{(k)} \; = \; (d_1^{(k)}, d_2^{(k)})^T \; := \; \alpha \; - \; K\psi^{(k)} \; , \tag{3.15a}$$

$$K_L d^{(k)} \; = \; (d_1^{(k)}, \; - \, A\tilde{B}^{-1}d_1^{(k)} \; + \; d_2^{(k)})^T \; , \tag{3.15b}$$

$$M_1\varphi^{(k)} \; = \; K_L d^{(k)} \; , \tag{3.15c}$$

$$\psi^{(k+1)} \; = \; \psi^{(k)} \; + \; \varphi^{(k)} \; . \tag{3.15d}$$

3.4.3 Transforming null-space iterations

We assume that $x \in \mathbb{R}^n$ and $\lambda \in R^m$ admit the decomposition

$$x \;=\; (x_1, x_2)^T \quad , \quad x_1 \in \mathbb{R}^{m_1} \;,\; x_2 \in \mathbb{R}^{n-m_1} \;, \tag{3.16a}$$

$$\lambda \;=\; (\lambda_1, \lambda_2)^T \quad , \quad \lambda_1 \in \mathbb{R}^{m_1} \;,\; \lambda_2 \in \mathbb{R}^{m-m_1} \;, \tag{3.16b}$$

and that $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{n\times n}$ can be partitioned by means of

$$A \;=\; \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \quad , \quad B \;=\; \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \;, \tag{3.17}$$

where $A_{11}, B_{11} \in \mathbb{R}^{m_1\times m_1}$ with nonsingular A_{11}.
Partitioning the right-hand side in (3.3) accordingly, the KKT system takes the form

$$\begin{pmatrix} B_{11} & B_{12} & | & A_{11}^T & A_{21}^T \\ B_{21} & B_{22} & | & A_{12}^T & A_{22}^T \\ -- & -- & | & -- & -- \\ A_{11} & A_{12} & | & 0 & 0 \\ A_{21} & A_{22} & | & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1^* \\ x_2^* \\ - \\ \lambda_1^* \\ \lambda_2^* \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ - \\ c_1 \\ c_2 \end{pmatrix} . \tag{3.18}$$

We rearrange (3.18) by exchanging the second and third rows and columns

$$\begin{pmatrix} B_{11} & A_{11}^T & | & B_{12} & A_{21}^T \\ A_{11} & 0 & | & A_{12} & 0 \\ -- & -- & | & -- & -- \\ B_{21} & A_{12}^T & | & B_{22} & A_{22}^T \\ A_{21} & 0 & | & A_{22} & 0 \end{pmatrix} \begin{pmatrix} x_1^* \\ \lambda_1^* \\ - \\ x_2^* \\ \lambda_2^* \end{pmatrix} = \begin{pmatrix} b_1 \\ c_1 \\ - \\ b_2 \\ c_2 \end{pmatrix} . \tag{3.19}$$

Observing $B_{12} = B_{21}^T$, in block form (3.19) can be written as follows

$$\underbrace{\begin{pmatrix} \mathcal{A} & \mathcal{B}^T \\ \mathcal{B} & \mathcal{D} \end{pmatrix}}_{=:\; K} \underbrace{\begin{pmatrix} \psi_1^* \\ \psi_2^* \end{pmatrix}}_{=:\; \psi^*} \;=\; \underbrace{\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}}_{=:\; \alpha} , \tag{3.20}$$

where $\psi_i^* := (x_i^*, \lambda_i^*)^T \;,\; \alpha_i := (b_i, c_i)^T \;,\; 1 \le i \le 2$.
We note that block Gauss elimination in (3.20) leads to the Schur complement system

$$\begin{pmatrix} \mathcal{A} & \mathcal{B}^T \\ 0 & \mathcal{S} \end{pmatrix} \begin{pmatrix} \psi_1^* \\ \psi_2^* \end{pmatrix} \;=\; \begin{pmatrix} \alpha_1 \\ \alpha_2 - \mathcal{B}\dashv^{-1}\alpha_1 \end{pmatrix} . \tag{3.21}$$

The Schur complement $\mathcal{S}$ is given by

$$\mathcal{S} \;=\; \mathcal{D} \;-\; \mathcal{B}\mathcal{A}^{-1}\mathcal{B}^T \;, \tag{3.22}$$

where

$$\mathcal{A}^{-1} \;=\; \begin{pmatrix} 0 & A_{11}^{-1} \\ A_{11}^{-T} & -A_{11}^{-T} B_{11} A_{11}^{-1} \end{pmatrix} \;. \tag{3.23}$$

We assume that $\tilde{A}_{11} \in \mathbb{R}^{m_1 \times m_1}$ is an easily invertible approximation of A_{11} and define

$$\tilde{\mathcal{A}} \;=\; \begin{pmatrix} B_{11} & \tilde{A}_{11}^T \\ \tilde{A}_{11} & 0 \end{pmatrix} \;. \tag{3.24}$$

We remark that the inverse of $\tilde{\mathcal{A}}$ is given as in (3.23) with A_{11}^{-1}, A_{11}^{-T} replaced by $\tilde{A}_{11}^{-1}$ and $\tilde{A}_{11}^{-T}$, respectively.
We introduce the right transform

$$K_R \;=\; \begin{pmatrix} I & -\tilde{\mathcal{A}}^{-1}\mathcal{B}^T \\ 0 & I \end{pmatrix} \;, \tag{3.25}$$

which gives rise to the regular splitting

$$KK_R \;=\; \underbrace{\begin{pmatrix} \tilde{\mathcal{A}} & 0 \\ \mathcal{B} & \tilde{\mathcal{S}} \end{pmatrix}}_{=:\; M_1} \;-\; \underbrace{\begin{pmatrix} (I - \mathcal{A}\tilde{\mathcal{A}}^{-1})\tilde{\mathcal{A}} & (-I + \mathcal{A}\tilde{\mathcal{A}}^{-1})\mathcal{B}^T \\ 0 & 0 \end{pmatrix}}_{=:\; M_2 \sim 0} \;, \tag{3.26}$$

where

$$\tilde{\mathcal{S}} \;:=\; \mathcal{D} \;-\; \mathcal{B}\tilde{\mathcal{A}}^{-1}\mathcal{B}^T \;. \tag{3.27}$$

Given a start iterate $\psi^{(0)} = (\psi_1^{(0)}, \psi_2^{(0)})^T$, we solve the KKT system (3.20) by the transforming null-space iterations

$$\begin{aligned} \psi^{(k+1)} \;&=\; \psi^{(k)} \;+\; K_R M_1^{-1}(\alpha \;-\; K\psi^{(k)}) \;= \\ &=\; (I \;-\; K_R M_1^{-1} K)\psi^{(k)} \;+\; K_R M_1^{-1}\alpha \quad , \quad k \geq 0 \;. \end{aligned} \tag{3.28}$$

3.5 Active set strategies for convex QP problems

3.5.1 Primal active set strategies

We consider the constrained QP problem

$$\text{minimize} \quad f(x) := \frac{1}{2}\, x^T B x \;-\; x^T b \tag{3.29a}$$

$$\text{over} \quad x \in \mathbb{R}^n$$

$$\text{subject to} \quad Cx = c \tag{3.29b}$$

$$Ax \leq d \;, \tag{3.29c}$$

where $B \in \mathbb{R}^{n\times n}$ is symmetric positive definite, $C \in \mathbb{R}^{m\times n}$, $A \in \mathbb{R}^{p\times n}$, and $b \in \mathbb{R}^n$, $c \in \mathbb{R}^m$, $d \in \mathbb{R}^p$.
We write the matrices A and C in the form

$$A \;=\; \begin{bmatrix} a_1 \\ \cdot \\ a_p \end{bmatrix} \;,\; a_i \in \mathbb{R}^n \quad , \quad C \;=\; \begin{bmatrix} c_1 \\ \cdot \\ c_m \end{bmatrix} \;,\; c_i \in \mathbb{R}^n \;. \tag{3.30}$$

The inequality constraints (3.29c) can be equivalently stated as

$$a_i^T x \;\le\; d_i \quad , \quad 1 \le i \le m \;. \tag{3.31}$$

The primal active set strategy is an iterative procedure:
Given a feasible iterate $x^{(\nu)}$, $\nu \ge 0$, we determine an active set

$$\mathcal{I}_{ac}(x^{(\nu)}) \subset \{1, ..., p\} \tag{3.32}$$

and consider the corresponding constraints as equality constraints, whereas the remaining inequality constraints are disregarded. Setting

$$p \;=\; x^{(\nu)} \;-\; x \quad , \quad b^{(\nu)} \;=\; Bx^{(\nu)} \;-\; b \;, \tag{3.33}$$

we find

$$f(x) \;=\; f(x^{(\nu)} - p) \;=\; \frac{1}{2} p^T B p \;-\; (b^{(\nu)})^T p \;+\; g \;,$$

where $g := \frac{1}{2}(x^{(\nu)})^T B x^{(\nu)} - b^T x^{(\nu)}$.
The equality constrained QP problem to be solved at the $(\nu+1)$-st iteration step is then:

$$\text{minimize} \quad \frac{1}{2}\, p^T B p \;-\; (b^{(\nu)})^T p \tag{3.34a}$$

$$\text{over} \quad p \in \mathbb{R}^n$$

$$\text{subject to} \quad Cp = 0 \tag{3.34b}$$

$$a_i^T p = 0 \;,\; i \in \mathcal{I}_{ac}(x^{(\nu)}) \;, \tag{3.34c}$$

We denote the solution of (3.34a)-(3.34c) by $p^{(\nu)}$. The new iterate $x^{(\nu+1)}$ is then obtained according to

$$x^{(\nu+1)} \;=\; x^{(\nu)} \;-\; \alpha_\nu p^{(\nu)} \quad , \quad \alpha_\nu \in [0, 1] \;, \tag{3.35}$$

where α_ν is chosen such that $x^{(\nu+1)}$ stays feasible.
In particular, for $i \in \mathcal{I}_{ac}(x^{(\nu)})$ we have

$$a_i^T x^{(\nu+1)} \;=\; a_i^T x^{(\nu)} \;-\alpha_\nu\; a_i^T p^{(\nu)} \;=\; a_i^T x^{(\nu)} \;\le\; d_i \;.$$

On the other hand, if $a_i^T p^{(\nu)} \geq 0$ for some $i \notin \mathcal{I}_{ac}(x^{(\nu)})$, it follows that

$$a_i^T x^{(\nu+1)} \;=\; a_i^T x^{(\nu)} \;-\alpha_\nu \; a_i^T p^{(\nu)} \;\leq\; a_i^T x_{(\nu)} \;\leq\; d_i \; .$$

Finally, if $a_i^T p^{(\nu)} < 0$ for $i \notin \mathcal{I}_{ac}(x^{(\nu)})$, we have

$$a_i^T x^{(\nu+1)} \;=\; a_i^T x^{(\nu)} \;-\alpha_\nu \; a_i^T p^{(\nu)} \;\leq\; d_i \iff \alpha_\nu \;\leq\; \frac{a_i^T x^{(\nu)} - d_i}{a_i^T p^{(\nu)}} \; .$$

Consequently, in order to guarantee feasibility, we choose

$$\alpha_\nu \;:=\; \min\,(1, \min_{\substack{i \notin \mathcal{I}_{ax}(x^{(\nu)}) \\ a_i^T p^{(\nu)} < 0}} \frac{a_i^T x^{(\nu)} - d_i}{a_i^T p^{(\nu)}}) \; . \tag{3.36}$$

We define

$$\mathcal{I}_{b\ell}(p^{(\nu)}) := \{i \notin \mathcal{I}_{ac}(x^{(\nu)}) \mid a_i^T p^{(\nu)} < 0 \; , \; \min_{i \notin \mathcal{I}_{ax}(x^{(\nu)})} \frac{a_i^T x^{(\nu)-d_i}}{a_i^T p^{(\nu)}} < 1\} \; . \tag{3.37}$$

Clearly,

$$a_i^T x^{(\nu+1)} \;=\; a_i^T \big(x^{(\nu)} - \alpha_\nu p^{(\nu)}\big) \;=\; d_i \; , \; i \in \mathcal{I}_{b\ell}(x^{(\nu)}) \; .$$

Therefore, $\mathcal{I}_{b\ell}(p^{(\nu)})$ is referred to as the set of blocking constraints.
We specify $\mathcal{I}_{ac}(x^{(\nu+1)})$ by adding to $\mathcal{I}_{ac}(x^{(\nu)})$ the most restrictive blocking constraint:

$$\begin{aligned} &\mathcal{I}_{ac}(x^{(\nu+1)}) \;:= \\ &\mathcal{I}_{ac}(x^{(\nu)}) \cup \{j \in \mathcal{I}_{b\ell}(x^{(\nu)} \mid \frac{a_j^T x^{(\nu)} - d_j}{a_j^T p^{(\nu)}} = \min_{\substack{i \notin \mathcal{I}_{ac}(x^{(\nu)}) \\ a_i^T p^{(\nu)} < 0}} \frac{a_i^T x^{(\nu)} - d_i}{a_i^T p^{(\nu)}}\}. \end{aligned} \tag{3.38}$$

Further information with respect to a proper specification of the set of active constraints is provided by systematically checking the KKT conditions:
Assume that $p^{(\nu)} = 0$ is the solution of the QP problem (3.34a)-(3.34c). Since $p^{(\nu)}$ satisfies the KKT conditions associated with that QP problem, there exist Lagrange multipliers $\lambda^{(\nu)} \in \mathbb{R}^m$ and $\mu_i^{(\nu)}, i \in \mathcal{I}_{ac}(x^{(\nu)})$, such that

$$-b^{(\nu)} \;=\; -(Bx^{(\nu)} - b) \;=\; -\sum_{i=1}^m \lambda_i^{(\nu)} c_i \;-\; \sum_{i \in \mathcal{I}_{ac}(x^{(\nu)})} \mu_i^{(\nu)} a_i \; . \tag{3.39}$$

If we set

$$\mu_i^{(\nu)} \;:=\; 0 \quad , \quad i \in \{1, ..., p\} \setminus \mathcal{I}_{ac}(x^{(\nu)}) \; ,$$

it is clear that $x^{(\nu)}, \lambda^{(\nu)}$, and $\mu^{(\nu)}$ satisfy the first KKT condition with respect to the original QP problem (3.29a)-(3.29c).
Since $x^{(\nu)}$ is feasible, the second and third KKT conditions also hold true.
We check the fourth KKT condition in terms of the sign of the multiplier $\mu^{(\nu)}$:
If

$$\mu_i^{(\nu)} \ \geq \ 0 \quad , \quad i \in \mathcal{I}_{ac}(x^{(\nu)}) \ ,$$

the fourth KKT condition holds true. Consequently, $x^{(\nu)}$ is a strict local minimum, since B is symmetric positive definite.
On the other hand, if there exists $j \in \mathcal{I}_{ac}(x^{(\nu)})$ such that

$$\mu_j^{(\nu)} \ < \ 0 \ ,$$

we remove that constraint from the active set. We show that this strategy produces a direction p in the subsequent iteration step that is feasible with respect to the dropped constraint:

Theorem 3.4 Feasible and descent direction for active set strategy

We assume that $\hat{x} \in \mathbb{R}^n$ satisfies the KKT conditions for the QP problem (3.34a)-(3.34c), i.e., in particular (3.39) holds true. Further, assume that the constraint gradients $c_i, 1 \leq i \leq m, a_i, i \in \mathcal{I}_{ac}(\hat{x})$ are linearly independent. Finally, suppose there is $j \in \mathcal{I}_{ac}(\hat{x})$ such that $\hat{\mu}_j < 0$.
Let p is the solution of the QP problem

$$\text{minimize} \quad \frac{1}{2} \, p^T B p \ - \ \hat{b}^T p \tag{3.40a}$$
$$\text{over} \quad p \in \mathbb{R}^n$$
$$\text{subject to} \quad Cp = 0 \tag{3.40b}$$
$$a_i^T p = 0 \ , \ i \in \mathcal{I}_{ac}(\hat{x}) \setminus \{j\} \ . \tag{3.40c}$$

where $\hat{b} := B\hat{x} - b$.
Then, p is a feasible direction for the constraint j, i.e.,

$$a_j^T p \ \leq \ 0 \ .$$

Moreover, if p satisfies the second order sufficient optimality conditions for (3.40a)-(3.40c), then $a_j^T p < 0$, i.e., p is a descent direction.

Proof. Since p is a solution of the QP problem (3.40a)-(3.40c), there exist Lagrange multipliers $\tilde{\lambda}_i, 1 \leq i \leq m$, and $\tilde{\mu}_i, i \in \mathcal{I}_{ac}(\hat{x}), i \neq j$, such that

$$Bp \ - \ \hat{b} \ = \ - \sum_{i=1}^{m} \tilde{\lambda}_i c_i \ - \sum_{i \in \mathcal{I}_{ac}(\hat{x}), i \neq j} \tilde{\mu}_i a_i \ . \tag{3.41}$$

Let Z be the null space basis of the matrix

$$\left[\, [c_i]^T_{1 \le i \le m} \; [a_i]^T_{i \in \mathcal{I}_{ac}(\hat{x}), i \neq j} \, \right] \; .$$

The second order necessary optimality conditions imply that

$$Z^T B Z$$

is positive semidefinite. Since p has the form $p = Zp_Z$ for some vector p_Z, we deduce

$$p^T B p \;\ge\; 0 \; .$$

Since we have assumed that (3.39) holds true (with $b^{(\nu)}, x^{(\nu)}, \lambda_i^{(\nu)}, \mu_i^{(\nu)}$ replaced by $\hat{b}, \hat{x}, \hat{\lambda}_i, \hat{\mu}_i$), subtraction of (3.39) from (3.41) yields

$$Bp \;=\; - \sum_{i=1}^{m} (\tilde{\lambda}_i - \hat{\lambda}_i) c_i \;-\; \sum_{i \in \mathcal{I}_{ac}(\hat{x}), i \neq j} (\tilde{\mu}_i - \hat{\mu}_i) a_i \;+\; \hat{\mu}_j a_j \; . \tag{3.42}$$

Forming the inner product with p and observing $c_i^T p = 0, 1 \le i \le m$, and $a_i^T p = 0, i \in \mathcal{I}_{ac}(\hat{x}), i \neq j$, we find

$$p^T B p \;=\; \hat{\mu}_j a_j^T p \; . \tag{3.43}$$

Since $\hat{\mu}_j < 0$, we must have $a_j^T p \le 0$.
If the second order sufficient optimality conditions are satisfied , it follows from (3.43) that

$$a_j^T p = 0 \iff p^T B p = p_Z^T Z^T B Z p_Z = 0 \iff p_Z = 0 \; ,$$

which implies $p = 0$. Due to the linear independence of the constraint gradients, then (3.42) gives us $\hat{\mu}_j = 0$, which is a contradiction. Hence, we must have $p^T B p > 0$, whence $a_j^T p < 0$.

•

Corollary 3.5 Strictly decreasing objective functional

Suppose that $p^{(\nu)} \neq 0$ is a solution of the quadratic subprogram (3.34a)-(3.34c) and satisfies the second order sufficient optimality conditions. Then, the objective functional of the original QP problem is strictly decreasing along the direction $p^{(\nu)}$.

Proof. Let us denote by Z the null space basis matrix associated with (3.34a)-(3.34c). Then $Z^T B Z$ is positive definite, and we find that $p^{(\nu)}$ is the unique global minimizer of (3.34a)-(3.34c).
On the other hand, $p = 0$ is also a feasible point for (3.34a)-(3.34c). Consequently, the value of the objective functional at $p = 0$ must be larger, i.e.,

$$\frac{1}{2} \, (p^{(\nu)})^T B p^{(\nu)} \;-\; (b^{(\nu)})^T p^{(\nu)} \;<\; 0 \; .$$

Since $(p^{(\nu)})^T B p^{(\nu)} \geq 0$ and $\alpha_\nu \in [0,1]$, we obtain

$$\frac{1}{2}\ \alpha_\nu (p^{(\nu)})^T B p^{(\nu)} \ - \ (b^{(\nu)})^T p^{(\nu)} \ < \ 0 \ .$$

It follows that

$$\begin{aligned} & \frac{1}{2}\ (x^{(\nu)} - \alpha_\nu p^{(\nu)})^T) B (x^{\nu)} - \alpha_\nu p^{(\nu)}) \ - \ b^T (x^{(\nu)} + \alpha_\nu p^{(\nu)}) \ = \\ = \ & \frac{1}{2}(x^{(\nu)})^T B x^{(\nu)} - b^T x^{(\nu)} + \alpha_\nu \big[\frac{1}{2}\alpha_\nu (p^{(\nu)})^T B p^{(\nu)} - (b^{(\nu)})^T p^{(\nu)}\big] \ < \\ < \ & \frac{1}{2}(x^{(\nu)})^T B x^{(\nu)} - b^T x^{(\nu)} \ . \end{aligned}$$

•

As far as the specification of the active set is concerned, after having computed the Lagrange multipliers $\lambda^{(\nu)}, \mu^{(\nu)}$, if $p^{(\nu)} = 0$, usually the most negative multiplier $\mu_i^{(\nu)}$ is removed from the active set. This leads us to the following algorithm:

Primal active set strategy

Step 1: Compute a feasible starting point $x^{(0)}$ and determine the set $\mathcal{I}_{ac}(x^{(0)})$ of active inequality constraints.

Step 2: For $\nu \geq 0$, proceed as follows:
Compute $p^{(\nu)}$ as the solution of (3.34a)-(3.34c).

Case 1: If $p^{(\nu)} = 0$, compute the multipliers $\lambda_i^{(\nu)}, 1 \leq i \leq m, \mu_i^{(\nu)}, i \in \mathcal{I}_{ac}(x^{(\nu)})$ that satisfy (3.39).
If $\mu_i^{(\nu)} \geq 0, i \in \mathcal{I}_{ac}(x^{(\nu)})$, then stop the algorithm.
The solution is $x^* = x^{(\nu)}$.
Otherwise, determine $j \in \mathcal{I}_{ac}(x^{(\nu)})$ such that
$\mu_j^{(\nu)} = \min_{i \in \mathcal{I}_{ac}(x^{(\nu)})} \mu_i^{(\nu)}$.
Set $x^{(\nu+1)} = x^{(\nu)}$ and $\mathcal{I}_{ac}(x^{(\nu+1)}) := \mathcal{I}_{ac}(x^{(\nu)}) \setminus \{j\}$.

Case 2: If $p^{(\nu)} \neq 0$, compute α_ν and set

$$x^{(\nu+1)} := x^{(\nu)} - \alpha_\nu p^{(\nu)}.$$

In case of blocking constraints, compute $\mathcal{I}_{ac}(x^{(\nu+1)})$ according to (3.38).

There are several techniques to compute an initial feasible point $x^{(0)} \in \mathcal{F}$. A common one requires the knowledge of some approximation $\tilde{x}$ of a feasible point which should not be "too infeasible". It amounts to the solution of the following linear programming problem:

$$\begin{aligned} &\text{minimize} \quad e^T z \ , && (3.44a)\\ &\text{over} \quad (x,z) \in \mathbb{R}^n \times \mathbb{R}^{m+p} \\ &\text{subject to} \quad c_i^T x + \gamma_i z_i = c_i \ , \ 1 \le i \le m \ , && (3.44b)\\ &\qquad a_i^T x - \gamma_{m+i} z_{m+i} \le d_i \ , \ 1 \le i \le p \ , && (3.44c)\\ &\qquad z \ge 0 \ , && (3.44d) \end{aligned}$$

where

$$e = (1,...,1)^T \quad , \quad \gamma_i = \begin{cases} -\text{sign}(c_i^T \tilde{x} - c_i) & , \quad 1 \le i \le m \\ 1 & , \quad m+1 \le i \le m+p \end{cases} . \quad (3.45)$$

A feasible starting point for this linear problem is given by

$$x^{(0)} = \tilde{x} \quad , \quad z_i^{(0)} = \begin{cases} |c_i^T \tilde{x} - c_i| & , \quad 1 \le i \le m \\ \max(a_{i-m}^T \tilde{x} - d_{i-m}, 0) & , \quad m+1 \le i \le m+p \end{cases} . \quad (3.46)$$

Obviously, the optimal value of the linear programming subproblem is zero, and any solution provides a feasible point for the original one.

Another technique introduces a measure of infeasibility in the objective functional in terms of a penalty parameter $\beta > 0$:

$$\begin{aligned} &\text{minimize} \quad \frac{1}{2} x^T B x - x^T b + \beta t \ , && (3.47a)\\ &\text{over} \quad (x,t) \in \mathbb{R}^n \times \mathbb{R} \\ &\text{subject to} \quad c_i^T x - c_i \le t \ , \ 1 \le i \le m \ , && (3.47b)\\ &\qquad -(c_i^T x - c_i) \le t \ , \ 1 \le i \le m \ , && (3.47c)\\ &\qquad a_i^T x - d_i \le t \ , \ 1 \le i \le p \ , && (3.47d)\\ &\qquad t \ge 0 \ . && (3.47e) \end{aligned}$$

For sufficiently large penalty parameter $\beta > 0$, the solution of (3.47a)-(3.47e) is $(x, 0)$ with x solving the original quadratic programming problem.

3.5.2 Primal-dual active set strategies

We consider a primal-dual active set strategy which does not require feasibility of the iterates. It is based on a Moreau-Yosida type approximation of the indicator function of the convex set of inequality constraints

$$K \ := \ \{ v \in \mathbb{R}^p \mid v_i \le 0 \ , \ 1 \le i \le p \} \ . \quad (3.48)$$

The indicator function $I_K : K \to \overline{\mathbb{R}}$ of K is given by

$$I_K(v) \ := \ \begin{cases} 0 & , \quad v \in K \\ +\infty & , \quad v \notin K \end{cases} . \quad (3.49)$$

The complementarity conditions

$$\begin{aligned} a_i^T x^* \ - \ d_i \ &\leq \ 0 \ , \ \mu_i^* \ \geq \ 0 \ , \\ \mu_i^* \ (a_i^T x^* \ - \ d_i) \ &= \ 0 \ , \ 1 \leq i \leq p \end{aligned}$$

can be equivalently stated as

$$\mu^* \in \partial I_K(v^*) \ , \ v_i^* := a_i^T x^* \ - \ d_i \ , \ 1 \leq i \leq p \ , \tag{3.50}$$

where ∂I_K denotes the subdifferential of the indicator function I_K as given by

$$\mu \in \partial I_K(v) \iff I_K(v) \ + \ \mu^T(w - v) \ \leq \ I_K(w) \ , \ w \in \mathbb{R}^p \ . \tag{3.51}$$

Using the generalized Moreau-Yosida approximation of the indicator function I_K, (3.50) can be replaced by the computationally more feasible condition

$$\mu^* \ = \ \sigma \ \left[v^* \ + \ \sigma^{-1}\mu^* \ - \ P_K(v^* \ + \ \sigma^{-1}\mu^*)\right] \ , \tag{3.52}$$

where σ is an appropriately chosen positive constant and P_K denotes the projection onto K as given by

$$P_K(w) := \begin{cases} w_i & , \quad w_i < 0 \\ 0 & , \quad w_i \geq 0 \end{cases} \ , \ 1 \leq i \leq p \ .$$

Note that (3.52) can be equivalently written as

$$\mu_i^* \ = \ \sigma \ \max\left(0, a_i^T x^* - d_i + \frac{\mu_i^*}{\sigma}\right) \ . \tag{3.53}$$

Now, given startiterates $x^{(0)}, \lambda^{(0)}, \mu^{(0)}$, the primal-dual active set strategy proceeds as follows: For $\nu \geq 1$ we determine the set $\mathcal{I}_{ac}(x^{(\nu)})$ of active constraints according to

$$\mathcal{I}_{ac}(x^{(\nu)}) \ := \ \{1 \leq i \leq p \mid a_i^T x^{(\nu)} - d_i + \frac{\mu_i^{(\nu)}}{\sigma} > 0\} \ . \tag{3.54}$$

We define

$$\mathcal{I}_{in}(x^{(\nu)}) \ := \ \{1, ..., p\} \setminus \mathcal{I}_{ac}(x^{(\nu)}) \tag{3.55}$$

and set

$$p^{(\nu)} \ := \ \text{card } \mathcal{I}_{ac}(x^{(\nu)}) \ , \tag{3.56}$$

$$\mu_i^{(\nu+1)} \ := \ 0 \ , \ i \in \mathcal{I}_{in}(x^{(\nu)}) \ . \tag{3.57}$$

We compute $(x^{(\nu+1)}, \lambda^{(\nu+1)}, \tilde{\mu}^{(\nu+1)}) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{p^{(\nu)}}$ as the solution of the KKT system associated with the equality constrained quadratic programming problem

$$\text{minimize} \quad Q(x) := \frac{1}{2}\, x^T B x \;-\; x^T b \tag{3.58a}$$
$$\text{over} \quad x \in \mathbb{R}^n$$
$$\text{subject to} \quad c_i^T x = c_i \ , \ 1 \le i \le m \ , \tag{3.58b}$$
$$a_i^T x = d_i \ , \ i \in \mathcal{I}_{ac}(x^{(\nu)}) \ . \tag{3.58c}$$

Since feasibility is not required, any startiterate $(x^{(0)}, \lambda^{(0)}, \mu^{(0)})$ can be chosen.

The following results show that we can expect convergence of the primal-dual active set strategy, provided the matrix B is symmetric positive definite.

Theorem 3.3 Reduction of the objective functional
Assume $B \in \mathbb{R}^{n \times n}$ to be symmetric, positive definite and refer to $\| \cdot \|_E := ((\cdot)^T B(\cdot))^{1/2}$ as the associated energy norm. Let $x^{(\nu)}, \nu \ge 0$, be the iterates generated by the primal-dual active set strategy. Then, for $Q(x) := \frac{1}{2} x^T Bx - x^T b$ and $\nu \ge 1$ there holds

$$\begin{aligned} & Q(x^{(\nu)}) \;-\; Q(x^{(\nu-1)}) \;= \\ = \;& -\frac{1}{2}\, \|x^{(\nu)} - x^{(\nu-1)}\|_E^2 - \sum_{\substack{i \in \mathcal{I}_{ac}(x^{(\nu)}) \\ i \notin \mathcal{I}_{ac}(x^{(\nu-1)})}} \mu_i^{(\nu)} (d_i - a_i^T x^{(\nu-1)}) \;\le\; 0 \ . \end{aligned} \tag{3.59}$$

Proof. Observing the KKT conditions for (3.58a)-(3.58c), we obtain

$$\begin{aligned} & Q(x^{(\nu)}) \;-\; Q(x^{(\nu-1)}) \;= \\ = \;& -\,\frac{1}{2}\, \|x^{(\nu)} - x^{(\nu-1)}\|_E^2 \;+\; (x^{(\nu)} - x^{(\nu-1)})^T (Bx^{(\nu)} - b) \;= \\ = \;& -\,\frac{1}{2}\, \|x^{(\nu)} - x^{(\nu-1)}\|_E^2 \;-\; \sum_{i \in \mathcal{I}_{ac}(x^{(\nu)})} \mu_i^{(\nu)} a_i^T (x^{(\nu)} - x^{(\nu-1)}) \ . \end{aligned}$$

For $i \in \mathcal{I}_{ac}(x^{(\nu)})$ we have $a_i^T x^{(\nu)} = d_i$, whereas $a_i^T x^{(\nu-1)} = d_i$ for $i \in \mathcal{I}_{ac}(x^{(\nu-1)})$. we thus get

$$\begin{aligned} & Q(x^{(\nu)}) \;-\; Q(x^{(\nu-1)}) \;= \\ = \;& -\,\frac{1}{2}\, \|x^{(\nu)} - x^{(\nu-1)}\|_E^2 \;-\; \sum_{\substack{i \in \mathcal{I}_{ac}(x^{(\nu)}) \\ i \notin \mathcal{I}_{ac}(x^{(\nu-1)})}} \mu_i^{(\nu)} (d_i - a_i^T x^{(\nu-1)}) \ . \end{aligned}$$

But $\mu_i^{(\nu)} \ge 0, i \in \mathcal{I}_{ac}(x^{(\nu)})$ and $a_i^T x^{(\nu-1)} \le d_i, i \notin \mathcal{I}_{ac}(x^{(\nu-1)})$ which gives the assertion.

•

Corollary 3.4 Convergence of a subsequence to a local minimum

Let $x^{(\nu)}, \nu \geq 0$, be the iterates generated by the primal-dual active set strategy. Then, there exist a subsequence $\mathbb{N}' \subset \mathbb{N}$ and $x^* \in \mathbb{R}^n$ such that $x^{(\nu)} \to x^*, \nu \to \infty, \nu \in \mathbb{N}'$. Moreover, x^* is a local minimizer of (3.29a)-(3.29c).

Proof. The sequence $(x^{(\nu)})_{\mathbb{N}}$ is bounded, since otherwise we would have $Q(x^{(\nu)}) \to +\infty$ as $\nu \to \infty$ in contrast to the result of the previous theorem. Consequently, there exist a subsequence $\mathbb{N}' \subset \mathbb{N}$ and $x^* \in \mathbb{R}^n$ such that $x^{(\nu)} \to x^*, \nu \to \infty, \nu \in \mathbb{N}'$. Passing to the limit in the KKT system for (3.58a)-(3.58c) shows that x^* is a local minimizer.

•

The following result gives an a priori estimate in the energy norm.

Theorem 3.5 A priori error estimate in the energy norm

Let $x^{(\nu)}, \nu \geq 0$, be the iterates generated by the primal-dual active set strategy and $\mathbb{N}' \subset \mathbb{N}$ such that $x^{(\nu)} \to x^*, \nu \to \infty, \nu \in \mathbb{N}'$. Then, there holds

$$\|x^{(\nu)} - x^*\|_E^2 \;\leq\; 2 \sum_{\substack{i \in \mathcal{I}_{ac}(x^*) \\ i \notin \mathcal{I}_{ac}(x^{(\nu)})}} \mu_i^* (d_i - a_i^T x^{(\nu)}) \ . \tag{3.60}$$

Proof. The proof is left as an exercise.

3.6 Interior-point methods

Interior-point methods are iterative schemes where the iterates approximate a local minimum from inside the feasible set. For ease of exposition, we restrict ourselves to inequality constrained quadratic programming problems of the form

$$\text{minimize} \quad Q(x) := \frac{1}{2}\, x^T B x \;-\; x^T b \tag{3.61a}$$

$$\text{over} \quad x \in \mathbb{R}^n$$

$$\text{subject to} \quad Ax \leq d, \tag{3.61b}$$

where $B \in \mathbb{R}^{n \times n}$ is symmetric, positive semidefinite, $A = [a_i]_{1 \leq i \leq p} \in \mathbb{R}^{p \times n}$, $b \in \mathbb{R}^n$, and $d \in \mathbb{R}^p$.

We already know from Chapter 2 that the KKT conditions for (3.61a)-(3.61b) can be stated as follows:

If $x^* \in \mathbb{R}^n$ is a solution of (3.61a)-(3.61b), there exists a multiplier $\mu^* \in \mathbb{R}^p$ such that

$$Bx^* \;+\; A^T \mu^* \;-\; b \;=\; 0\ , \tag{3.62a}$$

$$Ax^* \;-\; d \;\leq\; 0\ , \tag{3.62b}$$

$$\mu_i^*\ (Ax^* \;-\; d)_i \;=\; 0\ ,\ 1 \leq i \leq p\ , \tag{3.62c}$$

$$\mu_i^* \;\geq\; 0\ ,\ 1 \leq i \leq p\ . \tag{3.62d}$$

By introducing the slack variable $z := d - Ax$, the above conditions can be equivalently formulated as follows

$$\begin{aligned} Bx^* \; + \; A^T\mu^* \; - \; b \; &= \; 0 \; , && (3.63a) \\ Ax^* \; + \; z^* \; - \; d \; &= \; 0 \; , && (3.63b) \\ \mu_i^* \; z_i \; &= \; 0 \; , \; 1 \leq i \leq p \; , && (3.63c) \\ z_i^*, \mu_i^* \; &\geq \; 0 \; , \; 1 \leq i \leq p \; . && (3.63d) \end{aligned}$$

We can rewrite (3.63a)-(3.63d) as a constrained system of nonlinear equations. We define the nonlinear map

$$F(x,\mu,z) \; = \; \begin{bmatrix} Bx \; + \; A^T\mu - b \\ Ax \; + \; z \; - \; d \\ ZD_\mu e \end{bmatrix} \quad , \quad z,\mu \geq 0 \; , \qquad (3.64)$$

where

$$Z \; := \; \mathrm{diag}(z_1,...,z_p) \quad , \quad D_\mu \; := \; \mathrm{diag}(\mu_1,...,\mu_p) \quad , \quad e \; := \; (1,...,1)^T \; .$$

Given a feasible iterate (x, μ, z), we introduce a duality measure according to

$$\kappa \; := \; \frac{1}{p} \sum_{i=1}^{p} z_i\mu_i \; = \; \frac{z^T\mu}{p} \; . \qquad (3.65)$$

Definition 3.6 Central path

The set of points $(x_\tau, \mu_\tau, z_\tau)$, $\tau > 0$, satisfying

$$F(x_\tau, \mu_\tau, z_\tau) \; = \; \begin{bmatrix} 0 \\ 0 \\ \tau e \end{bmatrix} \quad , \quad z_\tau, \mu_\tau > 0 \qquad (3.66)$$

is called the central path.

The idea is to apply Newton's method to (3.64) to compute $(x_{\sigma\kappa}, \mu_{\sigma\kappa}, z_{\sigma\kappa})$ on the central path, where $\sigma \in [0, 1]$ is a parameter chosen by the algorithm. The Newton increments $(\Delta x, \Delta z, \Delta\mu)$ are the solution of the linear system

$$\begin{pmatrix} B & A^T & 0 \\ A & 0 & I \\ 0 & Z & D_\mu \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta\mu \\ \Delta z \end{pmatrix} \; = \; \begin{pmatrix} -r_b \\ -r_d \\ -ZD_\mu e \; + \; \sigma\kappa e \end{pmatrix} \; , \qquad (3.67)$$

where

$$r_b \; := \; Bx \; + \; A^T\mu \; - \; b \quad , \quad r_d \; := \; Ax \; + \; z \; - \; d \; .$$

The new iterate $(\overline{x}, \overline{\mu}, \overline{z})$ is then determined by means of

$$(\overline{x}, \overline{\mu}, \overline{z}) \; = \; (x, \mu, z) \; + \; \alpha \; (\Delta x, \Delta\mu, \Delta z) \qquad (3.68)$$

with α chosen such that $(\overline{x}, \overline{\mu}, \overline{z})$ stays feasible.
Since D_μ is a nonsingular diagonal matrix, the increment Δz in the slack variable can be easily eliminated resulting in

$$\begin{pmatrix} B & A^T \\ A & -D_\mu^{-1} Z \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta \mu \end{pmatrix} = \begin{pmatrix} -r_b \\ -r_d - D_\mu^{-1} g \end{pmatrix} , \tag{3.69}$$

where $g := -ZD_\mu e + \sigma \kappa e$.

3.7 Logarithmic barrier functions

We consider the inequality constrained quadratic programming problem (3.62a)-(3.62b). Algorithms based on barrier functions are iterative methods where the iterates are forced to stay within the interior

$$\mathcal{F}^{int} \ := \ \{ \ x \in \mathbb{R}^n \mid a_i^T x \ - \ c_i \ < \ 0 \ , \ 1 \le i \le p \ \} \ . \tag{3.70}$$

Barrier functions have the following properties:

- They are infinite outside $\mathcal{F}^{int}$.
- They are smooth within $\mathcal{F}^{int}$.
- They approach ∞ as x approaches the boundary of $\mathcal{F}^{int}$.

Definition 3.7 Logarithmic barrier function

For the quadratic programming problem (3.62a)-(3.62b) the objective functionals

$$B^{(\beta)}(x) \ := \ Q(x) \ - \ \beta \sum_{i=1}^{p} \log(d_i - a_i^T x) \ , \ \beta > 0 \tag{3.71}$$

are called logarithmic barrier functions. The parameter β is referred to as the barrier parameter.

Theorem 3.8 Properties of the logarithmic barrier function

Assume that the set $\mathcal{S}$ of solutions of (3.62a)-(3.62b) is nonempty and bounded and that the interior $\mathcal{F}^{int}$ of the feasible set is nonempty. Let $\{\beta_k\}_{\mathbb{N}}$ be a decreasing sequence of barrier parameters with $\beta_k \to 0$ as $k \to \infty$. Then there holds:

(i) For any $\beta > 0$ the logarithmic barrier function $B^{(\beta)}(x)$ is convex in $\mathcal{F}^{Int}$ and attains a minimizer $x(\beta)$ on $\mathcal{F}^{int}$. Any local minimizer $x(\beta)$ is also a global minimizer of $B^{(\beta)}(x)$.

(ii) If $\{x(\beta_k)\}_{\mathbb{N}}$ is a sequence of minimizers, then there exists $\mathbb{N}' \subset \mathbb{N}$ such that $x(\beta_k) \to x^* \in \mathcal{S}, k \in \mathbb{N}'$.

(iii) If Q^* is the optimal value of the objective functional Q in (3.62a)-(3.62b), then for any sequence $\{x(\beta_k)\}_{\mathbb{N}}$ of minimizers there holds

$$Q(x(\beta_k)) \to Q^* \quad , \quad B^{(\beta_k)}(x(\beta_k)) \to Q^* \quad \text{as } k \to \infty \ .$$

Proof. We refer to *M.H. Wright; Interior methods for constrained optimization. Acta Numerica, 341-407, 1992.*

•

We will have a closer look at the relationship between a minimizer of $B^{(\beta)}(x)$ and a point (x, μ) satisfying the KKT conditions for (3.62a)-(3.62b).
If $x(\beta)$ is a minimizer of $B^{(\beta)}(x)$, we obviously have

$$\nabla_x B^{(\beta)}(x(\beta)) \;=\; \nabla_x Q(x(\beta)) \;+\; \sum_{i=1}^{p} \frac{\beta}{d_i - a_i^T x(\beta)}\; a_i \;=\; 0\;. \tag{3.72}$$

Definition 3.9 Perturbed complementarity
The vector $z^{(\beta)} \in \mathbb{R}^p$ with components

$$z_i^{(\beta)} \;:=\; \frac{\beta}{d_i - a_i^T x(\beta)} \quad , \quad 1 \le i \le p \tag{3.73}$$

is called perturbed or approximate complementarity.

The reason for the above definition will become obvious shortly.
Indeed, in terms of the perturbed complementarity, (3.72) can be equivalently stated as

$$\nabla_x Q(x(\beta)) \;+\; \sum_{i=1}^{p} z_i^{(\beta)}\; a_i \;=\; 0\;. \tag{3.74}$$

We have to compare (3.74) with the first of the KKT conditions for (3.62a)-(3.62b) which is given by

$$\nabla_x L(x,\mu) \;=\; \nabla_x Q(x) \;+\; \sum_{i=1}^{p} \mu_i a_i \;=\; 0\;. \tag{3.75}$$

Obviously, (3.75) looks very much the same as (3.74).
The other KKT conditions are as follows:

$$a_i^T x \;-\; d_i \;\le\; 0\;,\; 1 \le i \le p\;, \tag{3.76a}$$

$$\mu_i \;\ge\; 0\;,\; 1 \le i \le p\;, \tag{3.76b}$$

$$\mu_i\;(a_i^T x \;-\; d_i) \;=\; 0\;,\; 1 \le i \le p\;. \tag{3.76c}$$

Apparently, (3.76a) and (3.76b) are satisfied by $x = x(\beta)$ and $\mu = z^{(\beta)}$.
However, (3.76c) does not hold true, since it follows readily from (3.73) that

$$z_i^{(\beta)}(d_i \;-\; a_i^T x(\beta)) \;=\; \beta \;>\; 0\;,\; 1 \le i \le p\;. \tag{3.77}$$

On the other hand, as $\beta \to 0$ a minimizer $x(\beta)$ and the associated $z^{(\beta)}$ come closer and closer to satisfying (3.76c). This is reason why $z^{(\beta)}$ is called perturbed (approximate) complementarity.

Theorem 3.10 Further properties of the logarithmic barrier function

Assume that $\mathcal{F}^{int} \neq \emptyset$ and that x^* is a local solution of (3.62a)-(3.62b) with multiplier μ^* such that the KKT conditions are satisfied.
Suppose further that the LICQ, strict complementarity, and the second order sufficient optimality conditions fold true at (x^*, μ^*). Then there holds:

(i) For sufficiently small $\beta > 0$ there exists a strict local minimizer $x(\beta)$ of $B^{(\beta)}(x)$ such that the function $x(\beta)$ is continuously differentiable in some neighborhood of x^* and $x(\beta) \to x^*$ as $\beta \to 0$.

(ii) For the perturbed complementarity $z^{(\beta)}$ there holds:

$$z^{(\beta)} \to \mu^* \quad \text{as } \beta \to 0 \ . \tag{3.78}$$

(iii) For sufficiently small $\beta > 0$, the Hessian $\nabla_{xx} B^{(\beta)}(x)$ is positive definite.

Proof. We refer to *M.H. Wright; Interior methods for constrained optimization. Acta Numerica, 341-407, 1992.*

•