Supplements 7 to Chapter III

E. Gekeler
28/10/06

Dual Method of Linear-Quadratic Programming

(a) We apply the notations of § 3.3 and observe first that the minimum problem

$$\begin{array}{lclcl} f(x) & = & \frac{1}{2}x^T A x - a^T x & = & \text{Min!},\ x \in \mathbb{R}^n\,, \\ g(x) & = & Bx + b & \geq & 0 \in \mathbb{R}^m\,, \end{array} \tag{1}$$

is equivalent to the LAGRANGE *problem*

$$\begin{array}{l} \text{Find a } (x^*, y^*) \in \mathbb{R}^n \times \mathbb{R}^m \text{ such that} \\ L(x^*, y^*) = \max_{y\geq 0} \operatorname{Inf}_x \left\{ \frac{1}{2} x^T A x - a^T x - y^T(Bx + b) \right\} \end{array} . \tag{2}$$

Of course the problem is convex now as before therefore the multiplier rule is sufficient for a solution. Starting from the absolute minimum $x = A^{-1}a$, the objective function $f(x)$ is enlarged until the multiplier rule is fulfilled or, in other words until x is as near as possible to the feasible domain.

Notations and Conventions:

(1°) $g^i(x) > 0 \implies y^i = 0$.
(2°) $\mathcal{K} := \{1, \ldots, m\}$ und $\mathcal{J} \subset \mathcal{K}$ index sets; at the beginning normally $\mathcal{J} = \emptyset$.
(3°) The problem

$$\{f(x)\,;\ g^j(x) := b^j x + \beta^j \geq 0,\ j \in \mathcal{J}\} = \text{Min!}$$

is briefly called LQ($\mathcal{J}$).
(4°) $(x, \mathcal{B})$ is called S(olution) pair of LQ($\mathcal{J}$), if x is solution of LQ($\mathcal{J}$), if $g^j(x) = 0$ for all $j \in \mathcal{B} \subset \mathcal{J}$ and if $B^{\mathcal{B}} = [\nabla g^j(x)]_{j\in\mathcal{B}}$ has maximum row rank.

There is always $\mathcal{B} \subset \mathcal{A}(x)$ but not necessarily $\mathcal{A}(x) \subset \mathcal{J}$. Conversely, if (4°) for some $(x, \mathcal{B})$ then there is a index set $\mathcal{J} \supset \mathcal{B}$ such that $(x, \mathcal{B})$ is S-pair of LQ($\mathcal{J}$); the index set $\mathcal{J}$ however plays a minor role in the sequel. By (4°) and Theorem 3.8 we obtain immediately the following intermediate result:
Let $(d, y^{\mathcal{B}})$ be solution of

$$\begin{bmatrix} A & [B^{\mathcal{B}}]^T \\ B^{\mathcal{B}} & O \end{bmatrix} \begin{bmatrix} d \\ y^{\mathcal{B}} \end{bmatrix} = \begin{bmatrix} \nabla f(x)^T \\ 0\ (\equiv g^{\mathcal{B}}(x)) \end{bmatrix},$$

and let $y^{\mathcal{B}} \geq 0$ as well as $g^{\mathcal{B}}(x) = 0$ then $(x+d, \mathcal{B})$ is S-pair of LQ($\mathcal{B}$) and solution of LQ($\mathcal{J}$) for all $\mathcal{J}$ such that $\mathcal{B} \subset \mathcal{J}$ and $g^j(x) \geq 0,\ j \in \mathcal{J}$. Therefore $S(\mathcal{J}) := \{x \in \mathbb{R}^n,\ g^j(x) \geq 0,\ j \in \mathcal{J}\}$ is the set of feasible points for LQ($\mathcal{B}$).

But, in activating a restriction with index p, some or even all conditions of $\mathcal{B}$ may become infeasible again! The latter case corresponds practically to a restart of the method with a different start position.

(b) V-Triple and S-pair A triple $(v, \mathcal{B}, p)$ is called V(iolated)-triple, if
(1°) $g^{\mathcal{B}}(v) = 0$ (i.e. $\mathcal{B} \subset \mathcal{A}(v)$)
(2°) $\gamma := g^p(v) < 0$ (p-th constraint violated)
(3°) Die Matrix $\begin{bmatrix} B^{\mathcal{B}} \\ b^p \end{bmatrix}$ has maximum row rank.
(4°) In the solution $(s, y_{\mathcal{B}}, y_p)$ of the linear system

$$\begin{bmatrix} A & [B^{\mathcal{B}}]^T & [b^p]^T \\ B^{\mathcal{B}} & O & O \\ b^p & 0 & 0 \end{bmatrix} \begin{bmatrix} s \\ y^{\mathcal{B}} \\ y^p \end{bmatrix} = \begin{bmatrix} \nabla f(v)^T \\ 0 \\ 0 \end{bmatrix} \equiv \begin{bmatrix} Av - a \\ g^{\mathcal{B}}(v) = 0 \\ g^p(v) - \gamma = 0 \end{bmatrix} \tag{3}$$

the components satisfy $s = 0,\ (y_{\mathcal{B}}, y_p) \geq 0,\ y_{\mathcal{B}} > 0$.

(Rows and columns of the system with index $j \in \mathcal{B},\ y_j = 0$ can be omitted.)

Suppose that $\gamma = g^p(v) = 0$ then the V-triple would be a S-pair for some index set $\mathcal{J} \supset \{B \cup \{p\}\}$ by (2°), but as $g_p(x) < 0$ the V-triple is only a S-pair for a *modified* problem $\widetilde{LQ}(\{\mathcal{B}, p\})$ where

$$\widetilde{g}(x) := B^{\mathcal{B}\cup p}x + b^{\mathcal{B}\cup p} - \gamma\, e^p,\ e^p = [\delta^i{}_p] \in \mathbb{R}^{|\mathcal{B}|+1}.$$

This relation says that v lies on the boundary of the feasible domain $\mathcal{S}$ being shifted by the vector $\gamma\, e^p$. Or, in other words, if $(v, \mathcal{B}, p)$ is a V-triple then v is solution of the problem

$$f(x) = \min!,\ x \in \{x \in \mathbb{R}^n,\ g^j(x) = 0,\ j \in \mathcal{B},\ g^p - \gamma = 0\}.$$

The question is now how we come from a V-triple to a new S-pair.

Let for the time being A a symmetric and positive matrix and let

$$D^{-1} = \begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix}^{-1} = \begin{bmatrix} H & N^T \\ N & F \end{bmatrix},$$

then by the "bordering lemma"

$$F = -(BA^{-1}B^T)^{-1},\ N = -FBA^{-1},\ H = A^{-1}(I + B^T FBA^{-1}).$$

Therefore the solution of the linear system

$$D\begin{bmatrix} z \\ r^{\mathcal{B}} \end{bmatrix} = \begin{bmatrix} q \\ 0 \end{bmatrix}$$

(later $q = b^p$) satisfies

$$\boxed{\begin{array}{rcl} z & = & Hq, \quad \text{primal corrector direction,} \\ r^{\mathcal{B}} & = & Nq, \quad (r^{\mathcal{B}}, -1) \text{ dual corrector direction} \end{array}}\ .$$

Lemma 1 *Let $(v, \mathcal{B}, p)$ be V-triple and let*

$$[b^p]^T \equiv B^T e^p,\ z = H\,[b^p]^T,\ r^{\mathcal{B}} = N\,[b^p]^T,\ \alpha = g^p(v) - g^p(x),$$

then for all $t \in \mathbb{R}$

$$\begin{bmatrix} A & [B^{\mathcal{B}}]^T & [b^p]^T \\ B^{\mathcal{B}} & O & O \\ b^p & 0 & 0 \end{bmatrix} \left[\begin{bmatrix} 0 \\ y^{\mathcal{B}} \\ y^p \end{bmatrix} + \begin{bmatrix} -t\,z \\ -t\,r^{\mathcal{B}} \\ t \end{bmatrix}\right] = \left[\begin{bmatrix} Av - a \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \alpha \end{bmatrix}\right] \quad . \tag{4}$$

Proof. Since $(v, \mathcal{B}, p)$ is a V-triple we have

$$\begin{bmatrix} A & [B^{\mathcal{B}}]^T & [b^p]^T \\ B^{\mathcal{B}} & O & O \\ b^p & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ y^{\mathcal{B}} \\ y^p \end{bmatrix} = \begin{bmatrix} \nabla f(v)^T \\ 0 \\ 0 \end{bmatrix} \equiv \begin{bmatrix} Av - a \\ 0 \\ 0 \end{bmatrix}. \quad (5)$$

Therefore we have to show that

$$\begin{bmatrix} A & [B^{\mathcal{B}}]^T & [b^p]^T \\ B^{\mathcal{B}} & O & O \\ b^p & 0 & 0 \end{bmatrix} \begin{bmatrix} -tz \\ -tr^{\mathcal{B}} \\ t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \alpha \end{bmatrix}. \quad (6)$$

Now indices are omitted!
Equation 1: Show that $Az + B^T r = q,\ q = b^T,\ z = Hq,\ r = Nq.$

$$A(A^{-1}(I - B^T N))q + B^T N q = q$$

Equation 2:

$$\begin{aligned} Bz &= BHq = BA^{-1}(I - B^T N)q \\ &= [BA^{-1} - (BA^{-1}B^T)(BA^{-1}B^T)^{-1}BA^{-1}]q = 0. \end{aligned}$$

Equation 3: Clear! See above! □

The linear system (6) differs from the system (3) only by the value

$$0 \neq \alpha = -g^p(x) + g^p(v) = -g^p(v) + g^p(v) - b^p t\, z = -t\, b^p z.$$

We now consider the primal and the dual correcture

$$x = v + t\, z\,, \quad y^T = \begin{bmatrix} y^{\mathcal{B}} - t\, r^{\mathcal{B}} \\ y^p + t \end{bmatrix},$$

then we have to choose t such that g^p becomes active if possible (g^p linear):

$$\begin{aligned} & g^p(x) = g^p(x + t\, z) \equiv g^p(v) + t\, b^p z = 0 \\ & \Longleftrightarrow t = t_2 := -g^p(v)/b^p z. \end{aligned}$$

Here $t_2 > 0$ since on the one hand $g^p(v) < 0$ by the assumption on the V-triple, and on the other hand, $b^p z = b^p H [b^p]^T > 0$ because H positive definit for rank-maximal $B^{\mathcal{B}}$.
Conclusion:
If $t = t_2$ and

$$y^{\mathcal{B}} - t_2 r^{\mathcal{B}} \geq 0,\ y^p + t_2 \geq 0, \quad (7)$$

then $(x + t_2 z, \mathcal{B} \cup \{p\})$ is a new S-pair for some $\mathcal{J} \supset \{\mathcal{B} \cup \{p\}\}$.
Suppose that (7) is violated then we must choose $t < t_2$ and a suitable index of the index set $\mathcal{B}$ must be cancelled in order to obtain a new V-triple, but with the additional condition that
(1°) the infeasibility $g^p(v) < 0$ decreases, i.e. $g^p(x) > g^p(v)$,
(2°) $f(x) \geq f(v)$ i.e. the cost functional does *not* decrease.
This procedure can be performed at most $|\mathcal{B}|$-times then $\mathcal{B} = \emptyset$, $y^{\mathcal{B}}$ is disappeared and $y^p \geq 0$ by assumption (cf. (2°). Thus we obtain after at most $|\mathcal{B}|$ steps of inactivation and one step of activation with $t = t_2$ a new S-pair where $\mathcal{B} = \{p\}$ and $f(x) > f(v)$. The cost functional is enlarged in every new L-pair and only a finite number of index subsets do exist therefore the algorithm breakso off some time.

Lemma 2 *Let* $(v, \mathcal{B}, p)$ *be a V-triple, let* $b := b^p$ *and*

$$x = v + tz,\ z = Hb^T,\ r = Nb^T,$$
$$t_1 = \mathrm{Min}\{y^i/r^i,\ i \in \mathcal{B},\ r^i > 0\},\ t_2 = -g^p(v)/b\,z,\ t = \mathrm{Min}\{t_1, t_2\}.$$

Then
(1°) $g^p(x) \geq g^p(v)$ *for* $0 \leq t \leq t_2$ *(*$>$ *if* $t > 0$*),*
(2°) $f(x) - f(v) = t(b\,z)\left(\frac{1}{2}t + y^p\right)$ *(*$>$ *if* $t > 0$*),*
(2°) *(Inactivation step) Suppose that*

$$t = t_1 = y^j/r^j < t_2 \textit{ for a } j \in \mathcal{B},\ j \neq p,$$

then $(x, \mathcal{B}\backslash\{j\}, p)$ *is a V-triple,*
(4°) *(Activation step) Suppose* $t = t_2$ *then* $(x, \mathcal{B} \cup \{p\})$ *is a S-pair.*

Proof. The matrix H is regular as long as the matrix $[^[B^{\mathcal{B}}]^T, b^T]$ has maximum rank. We have $b\,z = b\,Hb^T > 0$ hence

$$g^p(x)g^p(v) + tb\,z \geq g^p(v)\ (> 0\ f\text{"}ur\, t > 0)$$

According to Taylor

$$f(x) - f(v) = t\nabla f(x)z + \frac{1}{2}t^2 z^T Az$$

Because the assumption to be a V-triple, by the first row of (3)

$$\nabla f(x) = [[B^{\mathcal{B}}]^T, b^T]\begin{bmatrix} y^{\mathcal{B}} \\ y^p \end{bmatrix} \quad z = A^{-1}(I - B^T N)b^T = Hb^T$$

therefore, since $B^{\mathcal{B}}z = 0$,

$$\begin{aligned} & f(x) - f(v) = ty^p b\,z + ty^{\mathcal{B}}B^{\mathcal{B}}z \\ & +\frac{1}{2}t^2(b(I - N^T B)A^{-1}AA^{-1}(I - B^T N)b^T \\ & = ty^p b\,z + \frac{1}{2}t^2(b\,z)\ \ (correct),\ \ b\,z > 0 \\ & = t(b\,z)(\frac{1}{2}t + y^p) \end{aligned}$$

where $y^p > 0$ by assumption.
By definition of t_1 and t we have $y^{\mathcal{B}} - tr^{\mathcal{B}} \geq 0$ hence $g^i(x) = 0,\ i \in \mathcal{B},\ x = v + tz$ and

$$y_+ = \begin{bmatrix} y^{\mathcal{B}} \\ y^p \end{bmatrix} + t\begin{bmatrix} -r^{\mathcal{B}} \\ 1 \end{bmatrix} \geq 0$$

By construction we have

$$\begin{bmatrix} A & [B^{\mathcal{B}}]^T & [b^p]^T \\ B^{\mathcal{B}} & O & O \\ b^p & 0 & 0 \end{bmatrix}\begin{bmatrix} 0 \\ \\ y_+ \end{bmatrix} = \begin{bmatrix} Ax - a \\ 0 \\ g^p(x) \end{bmatrix}$$

where $g^p(x) = 0$ for $t = t_2$. Therefore, by definition
$t = t_2$, then $(x, \mathcal{B} \cup \{p\})$ S-pair

$0 < t < t_2$, then $(x, \mathcal{B}, p)$V-triple
$t = t_1$, then $(y_+)_j = 0$, hence also $(x, \mathcal{B}\backslash\{j\})$ V-triple (b^j to be cancelled!). Let $y^{\mathcal{B}} > 0$ without loss of generalization (other rows/columns cancelled), then always $t = \min\{t_1, t_2\} > 0$.□
Therefore we obtain a *new* S-pair after $q = |\mathcal{B}|$ inactivation steps at the latest and one activation step with $t = t_2$ since f always increases; *in the case where*

$$g_p(x) < 0\,, \text{ and} \begin{bmatrix} B^{\mathcal{B}} \\ b^p \end{bmatrix} \text{ row regular}\,.$$

Finally we have to weaken the maximum rank condition of the matrix $\begin{bmatrix} B^{\mathcal{B}} \\ b^p \end{bmatrix}$ in the assumption on the V-triple. This is managed by the following result which constitutes the keystone of the method `dlqp`:

Lemma 3 *Let $(x, \mathcal{B})$ be a S-pair and $p \in \mathcal{K}\backslash\mathcal{B}$ such that*

$$g^p(x) < 0, \quad r^T B^{\mathcal{B}} = b^p.$$

Then
(1°) *If $r \leq 0$, the problem LQ($\mathcal{B} \cup \{p\}$) and thus the original problem (1) has no solution.*
(2°) *If there exists a $r^i > 0$, then ist $(x, \mathcal{B}\backslash\{j\}, p)$ is a V-triple where*

$$j = \arg\min \left\{ \frac{y^i}{r^i},\ r^i > 0, i \in \mathcal{B} \right\}. \tag{8}$$

Proof. (1°) Let u be arbitrary feasible for $QP(\mathcal{B} \cup \{p\})$. Since $(x, \mathcal{B})$ shall be a S-pair we then have $g^i(x) = 0$, $i \in \mathcal{B}$ and $g^i(u) \geq 0$, $i \in \mathcal{B}$, therefore $b^i(u - x) \geq 0$, $i \in \mathcal{B}$, i.e. $B^{\mathcal{B}}(u - x) \geq 0$. Furthermore, by assumption,

$$g^p(x) < 0\,,\ g^p(u) \geq 0\,, \text{ hence } b^p(u - x) \geq 0\,.$$

Let now $r \leq 0$ then

$$b^p(u - x) = r^T B^{\mathcal{B}}(u - x) \leq 0$$

which is a contradictin. Therefore the feasible set is empty in this case.
(2°) Accordingly, let $r_j > 0$ for a $j \in \mathcal{B} = \{1, \ldots, q\}$ and let without loss of generality $j = q$. Since $b^p = r^T B^{\mathcal{B}}$ then

$$b^q = \frac{1}{r_q} \left(-\sum_{i=1}^{q-1} r_i t b^i + b^p \right).$$

Since $(x, \mathcal{B})$ shall be a S-pair and since

$$\begin{bmatrix} A & [B^{\mathcal{B}}]^T \\ B^{\mathcal{B}} & 0 \end{bmatrix} \begin{bmatrix} 0 \\ y^{\mathcal{B}} \end{bmatrix} = \begin{bmatrix} \nabla f^T \\ 0 \end{bmatrix}$$

we then have

$$\begin{aligned} \nabla f(x) &= \sum_{i=1}^{q} y_i b^i = \sum_{i=1}^{q-1} y_i b^i + b^p \\ &= \sum_{i=1}^{q-1} y_i b^i + \frac{y_q}{r_q} \left(-\sum_{i=1}^{q-1} r_i b^i + b^q \right) \\ &= \frac{y_q}{r_q} b^p + \sum_{i=1}^{q-1} b^i \left(y_i - \frac{y_q}{r_q} r_i \right) \end{aligned}$$

Assume $r_i = 0$ then $y_i \geq 0$ since $(x, \mathcal{B})$ shall be a S-pair.
Assume $r_i \geq 0$ then $(....) \geq 0$ because condition (8).
Therefore, altogether

$$[b^1, \ldots, b^{q-1}, b_p]w = \nabla f(x)^T \,, \; w \geq 0 \,.$$

Let $\widetilde{\mathcal{B}} = \{\mathcal{B} \backslash \{q\}\} \cup \{p\}$ then, as an assumption for a V-triple, it is to be shown that $B^{\widetilde{\mathcal{B}}}$ is row-regular.
But now $B^{\mathcal{B}}$ is row-regular therefore also $B^{\mathcal{B}\backslash\{q\}}$. Suppose that $B^{\tilde{\mathcal{B}}}$ is not row-regular then (assumption)

$$b^p = \widetilde{r}^T B^{\mathcal{B}\backslash\{q\}} = [r^{\mathcal{B}}]^T B^{\mathcal{B}}$$

hence $r_q = 0$ which is a contradiction.