Supplements 2 to Chapter III

E. Gekeler
11/10/06

Duality in Linear Optimization

Let $C \in \mathbb{R}^m{}_n$ denote a matrix with m rows and n columns, let $\mathbb{R}^n := \mathbb{R}^n{}_1$, $\mathbb{R}_n := \mathbb{R}^1{}_n$, and let $a \in \mathbb{R}_n$, $B \in \mathbb{R}^p{}_n$, $c \in \mathbb{R}^p$, $d \in \mathbb{R}^m$. We consider the pair of linear problems

$$\begin{array}{ll} (P) & \mathrm{Max}\{ax,\ Bx = c,\ Cx \le d\}, \\ (D) & \mathrm{Min}\{yc + zd,\ yB + zC = a,\ z \ge 0\} \end{array}$$

and compare the tableaus of both problems without using slack variables; cf. e.g. [BeRi].
An application of the Lemma of Farkas, cf. [Spe], [BeRi], yields the following characterization of a solution of (P) and (D):

Theorem 1 *(Satz 3.7) (a)* $x^* \in \mathbb{R}^n$ *is a solution of (P) iff there exists a tripel* $(x^*, y^*, z^*) \in \mathbb{R}^n \times \mathbb{R}_p \times \mathbb{R}_m)$ *satisfying*

$$\begin{array}{lll} (i) & Bx^* = c,\ Cx^* \le d & \textit{primal feasibility,} \\ (ii) & y^*B + z^*C = a,\ z^* \ge 0 & \textit{dual feasibility,} \\ (iii) & z^*(Cx^* - d) = 0 & \textit{complementary slackness.} \end{array}$$

(b) $(y^*, z^*) \in \mathbb{R}_p \times \mathbb{R}_m$ *is a solution of (D) iff there exists a quadrupel* $(y^*, z^*, u^*, v^*) \in \mathbb{R}_p \times \mathbb{R}_m \times \mathbb{R}^n \times \mathbb{R}^m$ *satisfying*

$$\begin{array}{lll} (i) & z^*C + y^*B = a,\ y^* \ge 0 & \textit{primal feasibility,} \\ (ii) & Bu^* + c = 0,\ Cu^* + d = v^* \ge 0 & \textit{dual feasibility,} \\ (iii) & z^*v^* = 0 & \textit{complementary slackness.} \end{array}$$

In the Problem (P), y^* and z^* are called the Lagrange multipliers of the solution x^*, and in the problem (D), u^* and v^* are called the Lagrange multipliers of the solution (y^*, z^*). Writing

$$x^* = -u^*,\ d - Bx^* = v^*, \tag{1}$$

we see that the conditions in (a) and in (b) do coincide and, moreover,

$$ax^* = y^*Bx^* + z^*Bx^* = y^*c + z^*d. \tag{2}$$

Therefore we have the following inference to Theorem 1:

Theorem 2 *The problem (P) has a solution* x^* *iff the problem (D) has a solution* (y^*, z^*) *and then (2) holds.*

In the sequel we suppose that the rank condition is fulfilled:

$$\mathrm{rank}(B) = p \text{ and } \mathrm{rank}\begin{bmatrix} B \\ C \end{bmatrix} = n.$$

Instead of (P) we now consider the problem

$$\mathrm{Max}\{ax,\ Cx \overset{[p]}{\le} d\} \tag{3}$$

where "$\overset{[p]}{\leq}$" indicates that in $Cx \leq d$ always the first p inequalities are active, i.e., hold as equations. Using the MATLAB-convention we write $C^{\mathcal{A}} := C(\mathcal{A},:)$ and $C_{\mathcal{A}} := C(:,\mathcal{A})$ with any suitable index set $\mathcal{A}$. Let then x be an extreme point of (P), let $\mathcal{A}(x)$ be the index set of a basis of x,

$$\mathcal{A}(x) = \{1,\ldots,p,\varrho_1,\ldots,\varrho_{n-p}\},\ \mathcal{N}(x) = \{1,\ldots,m\}\backslash\mathcal{A}(x),$$

and let

$$C^{\mathcal{A}} = \begin{bmatrix} c^{\varrho_1} \\ \vdots \\ c^{\varrho_{n-p}} \end{bmatrix}, \begin{bmatrix} B \\ C^{\mathcal{A}} \end{bmatrix}^{-1} =: A = [A_B,\ A_C] = [[a_1,\ldots,a_p],\ [a_{p+1},\ldots,a_n]], \tag{4}$$

where row vectors are denoted with upper indices and column vectors with lower indices. The columns of A are the edges of the feasible set in the extreme point x with direction pointing to x. Supposing that the problem (P) is solvable and omitting BLAND's rule we choose with $d = [\delta^1,\ldots,\delta^m]^T$

$$\begin{aligned} j &:= \operatorname{Min}\operatorname{Arg}\operatorname{Min}\{\varphi(k) := aa_k,\ k = p,\ldots,n\}, \\ i &:= \operatorname{Min}\operatorname{Arg}\operatorname{Min}\{\psi(k) := \frac{c^k x - \delta^k}{c^k a_j},\ c^k a_j < 0,\ k \in \mathcal{N}(x)\}. \end{aligned}$$

Then a better — at least not worse — extreme point $\widetilde{x}$ is found from x by $\widetilde{x} = x - \psi(i)a_j$ which means the row vector c^{ϱ_j} is removed from and c^i is taken into the row basis of x yielding the basis of $\widetilde{x}$. The tableau of the primal problem (P) has thus the following form in the extreme point x:

$$\mathbf{P}(x) = [p^k{}_l] := \begin{bmatrix} A & x \\ C^{\mathcal{N}}A & r \\ w & \zeta \end{bmatrix} = \begin{bmatrix} A_B & A_C & x \\ C^{\mathcal{N}}A_B & C^{\mathcal{N}}A_C & r \\ y & z_{\mathcal{A}} & \zeta \end{bmatrix}, \begin{bmatrix} x = Ac \\ r = B^{\mathcal{N}}x - c^{\mathcal{N}} \\ y = aA_B \\ z_{\mathcal{A}} = aA_C \\ \zeta = aAc \end{bmatrix}. \tag{5}$$

The tableau $\mathbf{Q}(\widetilde{x}) = [q^k{}_l]$ of the extreme point $\widetilde{x}$ is obtained from $\mathbf{P}(x)$ by the well-known GAUSS-JORDAN step

$$\begin{array}{llll} q^i{}_j = 1/p^i{}_j & \text{(pivot element)}, & q^k{}_j = p^k{}_j/p^i{}_j,\ k \neq i & \text{(pivot column)}, \\ q^i{}_l = -p^i{}_l/p^i{}_j,\ l \neq j & \text{(pivot row)}, & q^k{}_l = p^k{}_l - p^k_j p^i{}_l/p^i{}_j & \text{(others)}. \end{array}$$

Let us now turn to the dual problem (D) which is is written in row form here for convenience. This problem has $m + p$ variables and $m + n$ side conditions which can be written in the form

$$[y,z]\begin{bmatrix} B & 0 \\ C & -I_m \end{bmatrix} \leq [a,0] \in \mathbb{R}_{m+n} \tag{6}$$

where the first n conditions are always active. The matrix of these side conditions has the full rank $m + p$. A row vector (y,z) is extreme point iff besides the n active restrictions $yB + zC = a \in \mathbb{R}_n$ at least further $m - n + p$ conditions $y^i \leq 0$ are active, i.e., the equality sign holds there. The gradients of these conditions are independent by (6). Let (y,z) be an extreme point of (D) and let — by historical reasons —

$$\begin{array}{lll} \mathcal{N}(z) & = \{i \in \{1,\ldots,m\},\ z^i = 0\}, & |\mathcal{N}(z)| = m - n + p, \\ \mathcal{A}(z) & = \{1,\ldots,m\}\backslash\mathcal{N}(z), & |\mathcal{A}(z)| = n - p, \\ z_A & := z(\mathcal{A}(y)),\ z_N := z(\mathcal{N}(z)). & \end{array}$$

In the present "row problem" the gradients of active side conditions are column vectors and the matrix $\widetilde{C}_{\mathcal{N}}$ of the gradients of the active conditions — corresponding to the matrix $C^{\mathcal{A}}$ in the problem (P) — has the following form after a suitable row permutation Q

$$Q\widetilde{C}_{\mathcal{N}} = \begin{bmatrix} B & 0 \\ C^{\mathcal{A}} & 0 \\ C^{\mathcal{N}} & -I_{m-n+p} \end{bmatrix}, \; Q\widetilde{B}_{\mathcal{A}} = \begin{bmatrix} 0 \\ -I_{n-p} \\ 0 \end{bmatrix}.$$

We write again briefly

$$\begin{bmatrix} B \\ C^{\mathcal{A}} \end{bmatrix}^{-1} =: A = [A_B, A_C], \; A_B \in \mathbb{R}^n{}_p, \; A_C \in \mathbb{R}^n{}_{n-p}.$$

Then the matrix $\widetilde{A}$ of the edges in the extreme point (y, z) has here the form

$$\widetilde{A} \equiv \begin{bmatrix} \widetilde{a}^1 \\ \vdots \\ \widetilde{a}^{m+p} \end{bmatrix} := [\widetilde{C}_{\mathcal{N}}]^{-1} = \begin{bmatrix} A & 0 \\ C^{\mathcal{N}} A & -I_{m-n+p} \end{bmatrix} Q^T.$$

The first n rows of $\widetilde{A}$ cannot be chosen for descent directions because they leave the feasible set. In the present minimum problem, the optimality condition thus has the form

$$v^{\varrho_k} := -\widetilde{a}^{n+k} \begin{bmatrix} c \\ d \end{bmatrix} \geq 0, \; k = 1, \ldots, m-n+p, \tag{7}$$

i.e. $-v^{\varrho_k} \leq 0, \; k = 1, \ldots, m-n+p$, where

$$\widetilde{a}^{n+k} = [c^{\varrho_k} A, \; -[\delta^k{}_l]_{l=1}^{m-n+p}] Q^T \in \mathbb{R}_{m+p}. \tag{8}$$

($\delta^k{}_l$ KRONECKER symbol). Writing

$$Q\begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} c \\ d^{\mathcal{A}} \\ d^{\mathcal{N}} \end{bmatrix}, \begin{bmatrix} c \\ d \end{bmatrix} = Q^T \begin{bmatrix} c \\ d^{\mathcal{A}} \\ d^{\mathcal{N}} \end{bmatrix}$$

we obtain from (7)

$$v^{\mathcal{N}} := d^{\mathcal{N}} - C^{\mathcal{N}} A_C d^{\mathcal{A}} - C^{\mathcal{N}} A_B c \geq 0 \in \mathbb{R}^{m-n+p}. \tag{9}$$

If (y, z) is not optimal then (9) is violated and

a column e_r of $\begin{bmatrix} 0 \\ 0 \\ -I_{m-n+p} \end{bmatrix}$ and a column e_s of $\begin{bmatrix} 0 \\ -I_{n-p} \\ 0 \end{bmatrix}$ are to be exchanged.

In the matrix $Q\widetilde{C}_{\mathcal{N}}$ this corresponds to an exchange of **row** of $C^{\mathcal{N}}$ with a row of $C^{\mathcal{A}}$. We choose $e_r, \; r = \mathcal{N}(z)_j$ with

$$j = \operatorname{Max Arg}_k \operatorname{Max}\{\varphi(k) := \widetilde{a}^{n+k} \begin{bmatrix} c \\ d \end{bmatrix}, \; k \in \{1, \ldots, m-n+p\}\}, \tag{10}$$

i.e. $j = \operatorname{Max Arg}_k \operatorname{Max}\{-v^{\varrho_k}, \; k \in \{1, \ldots, m-n+p\}\}$. The search direction $\widetilde{a}^{n+j}$ yields for $(\widetilde{y}, \widetilde{z}) = (y, z) - \tau\widetilde{a}^{n+j}$ and $\tau > 0$

$$[(y, z) - \tau\widetilde{a}^{n+j}] \begin{bmatrix} c \\ d \end{bmatrix} < yc + zd.$$

For the computation of the optimum step length τ we have to substitute $(\widetilde{y}, \widetilde{z})$ into the inactive conditions being simple sign conditions in here,

$$z_{\sigma_k} \geq 0 \Longleftrightarrow -(y, z)\, e_{p+\sigma_k} \leq 0,\ k = 1, \ldots, n-p, \tag{11}$$

recalling that $\mathcal{A}(z) = \{\sigma_1, \ldots, \sigma_{n-p}\}$. Substitution of $(\widetilde{y}, \widetilde{z}) = (y, z) - \tau \widetilde{a}^{n+j}$ into (11) yields with (8) the condition for feasibility

$$-z_{\sigma_k} + \tau [c^{\varrho_j} A]_{p+k} \leq 0,\ k = 1, \ldots, n-p.$$

Supposing that the problem is solvable and omitting BLAND's rule we choose

$$i = \text{Min Arg Min}\{\psi(k) := \frac{z_{\sigma_k}}{[c^{\varrho_j} A]_{p+k}},\ [c^{\varrho_j} A]_{p+k} > 0,\ k \in \{1, \ldots, n-p\}\}. \tag{12}$$

Then $\tau^* := \psi(i) \geq 0$ is the optimum step length.

For the tableau of the present row problem $\text{Max}\{\widetilde{x}\widetilde{a},\ \widetilde{x}\widetilde{C} \leq \widetilde{d}\}$ we have in complete analogy to (5)

$$\widetilde{\mathbf{P}} := \begin{bmatrix} \widetilde{A} & \widetilde{A}\widetilde{C}_{\mathcal{A}} & \widetilde{w} \\ \widetilde{x} & \widetilde{r} & \widetilde{\zeta} \end{bmatrix} \tag{13}$$

where $\widetilde{x}$ is now the actual extreme point and $\widetilde{w}$ contains the relevant parts of the multipliers:

$$\widetilde{x} = [y,\ z_{\mathcal{A}},\ z_{\mathcal{N}}] \in \mathbb{R}_{m+p},\ \widetilde{w} = -\begin{bmatrix} u \\ v^{\mathcal{N}} \end{bmatrix} \in \mathbb{R}^{m+p},\ \widetilde{\zeta} = zd + yc \in \mathbb{R}.$$

Moreover, we have

$$\widetilde{r} = \widetilde{x}\widetilde{C}_{\mathcal{A}} - \widetilde{d}_{\mathcal{A}} = \widetilde{x} \begin{bmatrix} 0 \\ -I_{n-p} \\ 0 \end{bmatrix} = -z_{\mathcal{A}}$$

and

$$\widetilde{A}\widetilde{C}_{\mathcal{A}} = \begin{bmatrix} A_B & A_C & 0 \\ C^{\mathcal{N}} A_B & C^{\mathcal{N}} A_C & -I_{m-n+p} \end{bmatrix} \begin{bmatrix} 0 \\ -I_{n-p} \\ 0 \end{bmatrix} = \begin{bmatrix} -A_C \\ -C^{\mathcal{N}} A_C \end{bmatrix}.$$

Therefore we obtain for the tableau of the dual problem

$$\widetilde{\mathbf{P}}(y, z) = \begin{bmatrix} A_B & A_C & 0 & -A_C & -u \\ C^{\mathcal{N}} A_B & C^{\mathcal{N}} A_C & -I_{m-n+p} & -C^{\mathcal{N}} A_C & -v^{\mathcal{N}} \\ y & z_{\mathcal{A}} & z_{\mathcal{N}} & -z_{\mathcal{A}} & \widetilde{\zeta} \end{bmatrix}.$$

In this tableau the second block column appears in the fourth block column again with negative sign therefore the fourth block column can be cancelled. The third block row can be cancelled, too, because the index set $\mathcal{A}(z)$ has always to be updated and $z_{\mathcal{N}}$ equals zero by definition of the index set $\mathcal{N}(z)$. Hence recalling (1) we obtain the desired result namely

$$\mathbf{P}(x^*) = \widetilde{\mathbf{P}}_{red}(y^*, z^*)$$

if x^* is a unique and nondegenerate solution of the problem (P). (Then (y^*, z^*) is a unique and nondegenerate solution of the problem (D).) In particular, the matrix A has the same dimension in both problems. Usually the last column of $\widetilde{\mathbf{P}}(y, z)$ is multiplied by -1 in most presentations but a sign change of a not–pivot column does not affect the global GAUSS-JORDAN step.

If x is a non–optimum extreme point of (P) then (y, z) in the tableau $\mathbf{P}(x)$ is not feasible for (D) and, vice versa, if (y, z) is a non–optimum extreme point of (D) then the point x associated to this problem by (1) is not feasible for (P). Therefore, if e.g. a primal problem is solved by the dual method, i.e. the method for solving (D), then the solution is approximated from the unfeasible domain and hence an approximation is an unfeasible point.

Bibliography

[BeRi] Best, M.J., und Ritter, K.:*Linear Programming.* Englewood Cliffs, N.Y.: Prentice Hall 1985.

[Bla] Bland, R.G.: *New finit pivot rules for simplex method.* Math. Operation Research 2 (1977), 103-107.

[Las] Lasdon, L.S.: *Optimization Theory for Large Systems.* London: Macmillan 1970.

[Naz] Nazareth, J.L.: *Computer Solutions of Linear Programs.* New York–Oxford: Oxford University Press 1987.

[Pad] Padberg, M.: *Linear Optimization and Extensions.* Berlin–Heidelberg–New York: Springer–Verlag 1995.

[Schr] Schrijver, A.: *Theory of Linear and Integer Programming.* New York: Wiley 1986.

[Spe] Spellucci, P.: *Numerische Verfahren der Nichtlinearen Optimierung.* Basel–Boston–Berlin: Birkhäuser 1992.