ANNALES SOCIETATIS MATHEMATICAE POLONAE
Series I: COMMENTATIONES MATHEMATICAE XXIII (1983)
ROCZNIKI POLSKIEGO TOWARZYSTWA MATEMATYCZNEGO
Seria I: PRACE MATEMATYCZNE XXIII (1983)

G. MAJCHER and T. STYRYLSKA (Kraków)

On a multistage optimization problem II

1. The purpose of this paper is the multistage solving of the following problem: we are looking for $r\times 1$ matrix V and $s\times 1$ matrix X which satisfy the system of the equations

$$AV+BX = C \tag{1}$$

and realize the minimum

$$\min V^T V. \tag{2}$$

The numerical matrices A, B, C, $m\times r$, $m\times s$, $m\times 1$, respectively, are given.

We assume

$$1\leqslant s < m < r, \quad \varrho_A = m, \quad \varrho_B = s, \tag{Z_1}$$

where ϱ_A, ϱ_B denote the rank of the matrix A, B, respectively.

Under the assumptions (Z_1) the problem (1) (2) has the unique solution (given, e.g., by the method of Lagrange's factors) of the form

$$V = A^T(AA^T)^{-1}\left[I-B\left(B^T(AA^T)^{-1}B\right)^{-1}B^T(AA^T)^{-1}\right]C, \tag{3}$$

$$X = (B^TB)^{-1}B^T(C-AV); \tag{4}$$

here I is the identity $m\times m$ matrix. The assumption (Z_1) guarantees the reversibility of the symmetric matrix $B^T(AA^T)^{-1}B$.

In the multistage problem instead of the system (1) the following system (5) is considered:

$$A_jV+B_jX = C_j, \quad j = 1, 2, \ldots, n, \tag{5}$$

where A_j, B_j, C_j are respectively $m_j\times r$, $m_j\times s$, $m_j\times 1$ matrices. We assume (Z_1) and

$$\sum_{j=1}^{n} m_j = m < r, \quad \varrho_{A_j} = m_j, \quad 1\leqslant n\leqslant m. \tag{Z_2}$$

Remarks. 1. In the case of $s = 0$, the problem (1) (2) was solved in [1]. Also there an information on the genesis of this problem was given.

In the compensating computation the problem (2) (5) is called "compensation of the direct conditioned observations with unknowns" and it often happens that the particular conditions (5) (in this case V denotes corrections) are obtained successively as investigations develop. The presented multistage procedure allows to solve the problem (2) (5) in the j-th stage by a suitable correction of the solution obtained in the preceding stage, which is practically more convenient than to compute all from the beginning.

2. Under some additional assumptions, the case $s \geqslant 1$ may be reduced to the case $s = 0$. This way will be used in our paper.

3. We shall distinguish two cases: (a) $s = \text{const}$ for $j = 1, 2, \ldots, n$ (Sections 2 and 3); (b) $s \neq \text{const}$; then system (5) will be of the form (19) (Sections 4 and 5).

4. If system (5) is not at once fully known, but the particular conditions are obtained successively, then it may happen that the number of the observations and thereby the number r of the corrections increases, $r \neq \text{const}$. However, it is not necessary to deal with this case (see remark, Section 3).

5. A comparison of (1) and (5) gives the following block-matrices

$$A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_n \end{bmatrix}, \quad C = \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_n \end{bmatrix}. \tag{6}$$

6. If $\varrho_C = 0$ (system (1) is homogeneous), the unique solution of the problem (1) (2) is $V = 0$, $X = 0$.

7. If $s = m$, the solution of the problem (1) (2) is $V = 0$, $X = B^{-1}C$.

2. The case of $s = \text{const}$. We assume

$$0 < s < m_1 < r, \quad \varrho_{B_1} = s. \tag{Z_3}$$

We divide $A_1V + B_1X = C_1$ into two subsystems (possibly after the preceding renumeration of the equations in system (5) for $j = 1$)

$$\begin{aligned} \mathring{A}_1V + \mathring{B}_1X &= \mathring{C}_1 \quad (s \text{ equations}), \\ \check{A}_1V + \check{B}_1X &= \check{C}_1 \quad (m_1 - s \text{ equations}), \end{aligned} \tag{7}$$

so that $\det \mathring{B}_1 \neq 0$. From the first subsystem (7) we get

$$X = \mathring{B}_1^{-1}(\mathring{C}_1 - \mathring{A}_1V), \tag{8}$$

and hence and from the second subsystem (7) we get the following system of equations:

$$K_1 V = Q_1, \tag{9}$$

where

$K_1 = \check{A}_1 - \check{B}_1 \mathring{B}_1^{-1} \mathring{A}_1$ is the $(m_1 - s) \times r$ matrix with the rank $\varrho_{K_1} = m_1 - s$,

$Q_1 = \check{C}_1 - \check{B}_1 \mathring{B}_1^{-1} \mathring{C}_1$ is the $(m_1 - s) \times 1$ matrix.

The solution of the first stage of the problem (2) (5) ($j = 1$) is (see [1] and (8), (9) above) of the form

$$V_1 = V^{(1)} = K_1^T (K_1 K_1^T)^{-1} Q_1, \quad X_1 = \mathring{B}_1^{-1} (\mathring{C}_1 - \mathring{A}_1 V_1). \tag{10}$$

Likewise, using formula (8), we eliminate X in all other stages. So we get in the second stage (see (5) for $j = 1, 2$; see also (9)) the system

$$K_1 V = Q_1, \quad K_2 V = Q_2, \tag{11}$$

where

$K_2 = A_2 - B_2 \mathring{B}_1^{-1} \mathring{A}_1$ is the $m_2 \times r$ matrix of the rank $\varrho_{K_2} = m_2$,

$Q_2 = C_2 - B_2 \mathring{B}_1^{-1} \mathring{C}_1$ is the $m_2 \times 1$ matrix.

But we want to get $V_2 = V^{(1)} + V^{(2)}$, so the corrections $V^{(2)}$ must satisfy the following system of equations (see (9) (11)):

$$K_1 V = 0, \quad K_2 V = Q_2 - K_2 V_1. \tag{12}$$

According to the results presented in [1], we find the matrices $F_{21} = K_2 K_1^T (K_1 K_1^T)^{-1}$, $\tilde{K}_2 = K_2 - F_{21} K_1$, $\tilde{Q}_2 = Q_2 - F_{21} Q_1$ and then

$$V^{(2)} = \tilde{K}_2^T (K_2 \tilde{K}_2^T)^{-1} \tilde{Q}_2, \quad V_2 = V^{(1)} + V^{(2)}, \quad X_2 = \mathring{B}_1^{-1} (\mathring{C}_1 - \mathring{A}_1 V_2),$$

etc. In an analogous way we get in the kth stage (see (5) for $j = 1, 2, \ldots, k$), $k \leqslant n$ the system (see (12)):

$$\begin{cases} K_j V = 0 \quad \text{for} \quad j = 1, 2, \ldots, k-1, \\ K_k V = Q_k - K_k V_{k-1}, \end{cases} \tag{13}$$

where K_1, Q_1, V_1 are as in (9), (10) and for $j = 2, 3, \ldots, k$

$$K_j = A_j - B_j \mathring{B}_1^{-1} \mathring{A}_1, \quad Q_j = C_j - B_j \mathring{B}_1^{-1} \mathring{C}_1, \quad V_j = \sum_{i=1}^{j} V^{(i)}. \tag{14}$$

So we have reduced the kth stage of the problem (2) (5) to the kth stage of the problem solved in [1]. Write, as it was in [1],

I the $r\times r$ identity matrix,

$$\Delta_j = \begin{cases} I & \text{for} \quad j = 0, \\ \Delta_{j-1}[I - K_j^T(K_j\Delta_{j-1}K_j^T)^{-1}K_j\Delta_{j-1}] & \text{for} \quad j = 1, 2, \ldots, n, \end{cases} \tag{15}$$

$$\tilde{K}_j = K_j\Delta_{j-1}, \quad \tilde{Q}_j = Q_j - K_jV_{j-1}, \tag{16}$$

and let us denote by P_k the problem (2) (5) (with $j = 1, 2, \ldots, k$). Then we have

THEOREM 1. *If the assumptions* (Z_1)–(Z_3) *are fulfilled, then the problem* P_k *has the* (*unique*) *solution which is of the form*

$$V_k = \sum_{j=1}^{k} V^{(j)}, \quad X_k = \mathring{B}_1^{-1}(\mathring{C}_1 - \mathring{A}_1V_k), \tag{17}$$

where by notations (14)–(16) *and taking into account that* $\tilde{K}_1 = K_1$, $\tilde{Q}_1 = Q_1$,

$$V^{(j)} = \tilde{K}_j^T(K_j\tilde{K}_j^T)^{-1}\tilde{Q}_j \quad \textit{for} \quad j = 1, 2, \ldots, k. \tag{18}$$

3. Algorithmic process for the case of $s =$ const. Formulae (14)–(18) are here applied (see also [1], § 8).

The 1st stage. (1) We find the matrix $\mathring{B}_1$. We calculate the matrices: (2) $\mathring{B}_1^{-1}$; (3) $\tilde{K}_1 = K_1$; (4) $\tilde{Q}_1 = Q_1$; (5) $K_1\tilde{K}_1^T$ (here $K_1K_1^T$); (6) $(K_1\tilde{K}_1^T)^{-1}$; (7) $L_1 = \tilde{K}_1^T(K\tilde{K}_1^T)^{-1}$; (8) $V_1 = V^{(1)} = L_1\tilde{Q}_1$; (9) X_1.

The 2nd stage. We calculate (1) K_2; (2) Q_2; (3) $\Delta_1 = \Delta_0 - L_1\tilde{K}_1$; (4) $\tilde{K}_2 = K_2\Delta_1$; (5) $\tilde{Q}_2 = Q_2 - K_2V_1$; (6) $K_2\tilde{K}_2^T$; (7) $(K_2\tilde{K}_2^T)^{-1}$; (8) $L_2 = \tilde{K}_2^T(K_2\tilde{K}_2^T)^{-1}$; (9) $V^{(2)} = L_2\tilde{Q}_2$; (10) $V_2 = V^{(1)} + V^{(2)}$; (11) X_2.

The kth stage. We have already from the $(k-1)$th stage: Δ_{k-2}, $\tilde{K}_{k-1}$, L_{k-1}, V_{k-1}. We calculate: (1) K_k; (2) Q_k; (3) $\Delta_{k-1} = \Delta_{k-2} - L_{k-1}\tilde{K}_{k-1}$; (4) $\tilde{K}_k = K_k\Delta_{k-1}$; (5) $\tilde{Q}_k = Q_k - K_kV_{k-1}$; (6) $K_k\tilde{K}_k^T$; (7) $(K_k\tilde{K}_k^T)^{-1}$; (8) $L_k = \tilde{K}_k^T(K_k\tilde{K}_k^T)^{-1}$; (9) $V^{(k)} = L_k\tilde{Q}_k$; (10) $V_k = V_{k-1} + V^{(k)}$; (11) X_k.

Remark. If at a certain stage the number r changes and there appear e.g. $\tilde{r}$ corrections more (considering the assumptions on r), then it is necessary to increase the matrices $\mathring{A}_1$, $\check{A}_1$, A_j, K_j, $\tilde{K}_j$ of the previous stages by adding $\tilde{r}$ zero columns. One should also increase Δ_j to the form

$$\left[\begin{array}{c|c} \Delta_j & 0 \\ \hline 0 & I_{\tilde{r}\times\tilde{r}} \end{array}\right].$$

The matrices Q_j, $\tilde{Q}_j$, $K_j\tilde{K}_j^T$, $(K_j\tilde{K}_j^T)^{-1}$ do not change their dimensions.

4. The case of $s \neq$ const. In the most general case, each stage can have more unknowns. We shall write them separately, so, instead of (5), we shall have the system

$$(19)\qquad A_jV+B_{j1}X_1+B_{j2}X_2+\ldots+B_{jj}X_j = C_j, \quad j=1,2,\ldots,n,$$

where A_j, V, C_j are the same matrices as in (5),

B_{ji} are for $j=1,2,\ldots,n,\ i=1,2,\ldots,j,\ m_j\times s_i$ given matrices,
X_i $s_i\times 1$ matrix of unknowns,
$X^T=[X_1^T, X_2^T, \ldots, X_n^T]$.

We assume (Z_2) and

$$(Z_4)\qquad 0<s_j<m_j<r,\quad \varrho_{B_{jj}}=s_j \quad \text{for} \quad j=1,2,\ldots,n.$$

The 1st stage is the same as in Section 2.

The 2nd stage. $X_1=\mathring{B}_{11}^{-1}(\mathring{C}_1-\mathring{A}_1V)$ is put into the equation $A_2V+B_{21}X_1+B_{22}X_2=C_2$ (see (19)) and we get

$$(20)\qquad \underline{A}_2V+B_{22}X_2=\underline{C}_2,$$

where

$$(21)\quad \underline{A}_1=A_1,\quad \underline{C}_1=C_1,\quad \underline{A}_2=A_2-B_{21}\mathring{B}_{11}^{-1}\mathring{\underline{A}}_1,\quad \underline{C}_2=C_2-B_{21}\mathring{B}_{11}^{-1}\mathring{\underline{C}}_1.$$

We divide the system (20) like (7), namely

$$(22)\qquad \begin{aligned} &\mathring{\underline{A}}_2V+\mathring{B}_{22}X_2=\mathring{\underline{C}}_2 \quad (s_2 \text{ equations, } \det\mathring{B}_{22}\neq 0),\\ &\check{\underline{A}}_2V+\check{B}_{22}X_2=\check{\underline{C}}_2 \quad (m_2-s_2 \text{ equations}).\end{aligned}$$

Hence $X_2=\mathring{B}_{22}^{-1}(\mathring{\underline{C}}_2-\mathring{\underline{A}}_2V)$ and after substituting it into the second subsystem of (22) we get

$$(23)\qquad K_2V=Q_2,$$

where

$$(24)\qquad K_2=\check{\underline{A}}_2-\check{B}_{22}\mathring{B}_{22}^{-1}\mathring{\underline{A}}_2,\quad Q_2=\check{\underline{C}}_2-\check{B}_{22}\mathring{B}_{22}^{-1}\mathring{C}_2.$$

Since we want to have the solution V_2 of these two stages together of the form $V_2=V^{(1)}+V^{(2)}$, so $V^{(2)}$ must satisfy the system of equations

$$(25)\qquad K_1V=0,\quad K_2V=Q_2-K_2V_1$$

(see (12)). $V^{(2)}$ and V_2 are computed by the same calculation method as in Section 2. However, $X^T=[X_1^T, X_2^T]$, where for $i=1,2$, $X_i=\mathring{B}_{ii}^{-1}(\mathring{\underline{C}}_i-\mathring{\underline{A}}_iV_2)$.

In an analogous way we get in the kth stage, $k\leqslant n$, the following system of equations (see (13) (14)):

$$(26)\qquad \begin{aligned} &K_jV=0 \quad \text{for} \quad j=1,2,\ldots,k-1,\\ &K_kV=Q_k-K_kV_{k-1},\end{aligned}$$

where for $j = 1, 2, \ldots, k$ (see (21) (24)):

$$K_j = \check{\underline{A}}_j - \check{B}_{jj}\mathring{B}_{jj}^{-1}\mathring{\underline{A}}_j, \qquad Q_j = \check{\underline{C}}_j - \check{B}_{jj}\mathring{B}_{jj}^{-1}\mathring{\underline{C}}_j,$$

(27)

$$\underline{A}_j = A_j - \sum_{\nu=1}^{j-1} B_{j\nu}\mathring{B}_{\nu\nu}\mathring{\underline{A}}_\nu, \quad \underline{C}_j = C_j - \sum_{\nu=1}^{j-1} B_{j\nu}\mathring{B}_{\nu\nu}^{-1}\mathring{\underline{C}}_\nu.$$

The further procedure is like that in Section 2 (using formulae (15) (16)).

Denote by P_k the problem (2) (19) with $j = 1, 2, \ldots, k$.

THEOREM 2. *If the assumptions* (Z_2), (Z_4) *are fulfilled and if* $\varrho_A = m$ (see (6)), *then the problem* P_k *has the* (*unique*) *solution, which is of the form*

$$V_k = \sum_{j=1}^{k} V^{(j)}, \qquad X^T = [X_1^T, X_2^T, \ldots, X_k^T], \tag{28}$$

where (see (14)–(16))

$$X_i = \mathring{B}_{ii}^{-1}(\mathring{\underline{C}}_i - \mathring{\underline{A}}_i V_k) \quad for \quad i = 1, 2, \ldots, k,$$

$$\tilde{K}_1 = K_1, \quad \tilde{Q}_1 = Q_1, \quad V^{(j)} = \tilde{K}_j^T (K_j \tilde{K}_j^T)^{-1} \tilde{Q}_j \quad for \quad j = 1, 2, \ldots, k.$$

5. Algorithmic process for the case of $s \neq$ const.

The 1st stage. We calculate: (1) $\mathring{B}_{11}$; (2) $\mathring{B}_{11}^{-1}$ and afterwards as in Section 3, the 1st stage, (3)–(9).

The 2nd stage. We calculate: (1) $\underline{A}_2$; (2) $\underline{C}_2$; (3) $\mathring{B}_{22}$; (4) $\mathring{B}_{22}^{-1}$ and then as in Section 3, the 2nd stage, (1)–(10) (of course, using formulae (24)). Finally we calculate X_1, X_2.

The kth stage. From the $(k-1)$th stage we have already $\mathring{\underline{A}}_\nu$, $\mathring{\underline{C}}_\nu$, $\mathring{B}_{\nu\nu}^{-1}$ for $\nu = 1, 2, \ldots, k-1$, and also Δ_{k-2}, $\tilde{K}_{k-1}$, L_{k-1}, V_{k-1}. We calculate: (1) $\underline{A}_k$; (2) $\underline{C}_k$; (3) $\mathring{B}_{kk}$; (4) $\mathring{B}_{kk}^{-1}$ and then as in Section 3, the kth stage, (1)–(10) (of course, using formulae (27)). Finally we calculate $X_1, X_2, \ldots, X_k$ (see (28)).

References

[1] G. Majcher and T. Styrylska, *On a multistage optimization problem I*, Comment. Math. 21 (1979), 163–171.

[2] T. Styrylska, *Calculation of the enthalpy of creating by the method of equalizing the enthalpy of chemical reactions*, Zesz. Nauk. Politechniki Krakowskiej 9 (1973), 1–47 (in Polish).

[3] G. Majcher and T. Styrylska, *Multistage compensation of the direct conditioned observations with unknowns*, Matematyka Stosowana 17(1981), 5–10 (in Polish).