Supplement to "Matrices and Matroids for Systems Analysis"
Detail of the Proof of Theorem 5.1.8 (page 275)

Kazuo Murota, April 5, 2016

5.1.3 Matrix Pencil and Kronecker Form

A polynomial matrix $A(s) = (A_{ij}(s))$ with $\deg_s A_{ij}(s) \le 1$ for all (i,j) is called a *pencil*. Obviously, a pencil $A(s)$ can be represented as $A(s) = sX + Y$ in terms of a pair of constant matrices X and Y. A pencil $A(s)$ is said to be *regular* if it is square and $\det A(s)$ is a nonvanishing polynomial. A pencil is called *singular* if it is not regular.

A pencil can be brought into a canonical block-diagonal matrix by means of *strict equivalence* $PA(s)Q$ using constant nonsingular matrices P and Q. The block-diagonal matrix is known as the Kronecker form (Gantmacher [87, Chap. XII]). For $m \ge 1$ and $\varepsilon \ge 0$, we define an $m \times m$ bidiagonal matrix $N_m(s)$ and an $\varepsilon \times (\varepsilon + 1)$ bidiagonal matrix $L_\varepsilon(s)$ by

$$N_m(s) = \begin{pmatrix} 1 & s & & & \\ & 1 & s & & \\ & & \ddots & \ddots & \\ & & & \ddots & s \\ & & & & 1 \end{pmatrix}, \quad L_\varepsilon(s) = \begin{pmatrix} 1 & s & & & & \\ & 1 & s & & & \\ & & \ddots & \ddots & & \\ & & & \ddots & s & \\ & & & & 1 & s \end{pmatrix}.$$

For $\eta \ge 0$ we define $U_\eta(s)$ to be the transpose of $L_\eta(s)$.

Theorem 5.1.8 (Kronecker form). *For a pencil $A(s)$ over a field $\boldsymbol{F}$, there exist nonsingular matrices P and Q over $\boldsymbol{F}$ such that*

$$\begin{aligned} PA(s)Q = \text{block-diag}\,(&sI_{m_0} + B; N_{m_1}(s), \cdots, N_{m_b}(s); \\ &L_{\varepsilon_1}(s), \cdots, L_{\varepsilon_c}(s); U_{\eta_1}(s), \cdots, U_{\eta_d}(s)), \end{aligned} \tag{5.1}$$

where

$$m_1 \ge \cdots \ge m_b \ge 1, \quad \varepsilon_1 \ge \cdots \ge \varepsilon_c \ge 0, \quad \eta_1 \ge \cdots \ge \eta_d \ge 0,$$

and B is an $m_0 \times m_0$ matrix over $\boldsymbol{F}$. The indices, m_0; b, $m_1, \cdots, m_b$; c, $\varepsilon_1, \cdots, \varepsilon_c$; d, $\eta_1, \cdots, \eta_d$, are uniquely determined. Denoting $r = \operatorname{rank} A$ and using $\delta_k(A)$ $(k = 0, 1, 2, \cdots)$ in (5.3), we have

$$b = r - \max_{k \ge 0} \delta_k(A), \quad c = |\mathrm{Col}(A)| - r, \quad d = |\mathrm{Row}(A)| - r, \tag{5.2}$$

$$m_0 = \delta_r(A) - \sum_{i=1}^{c} \varepsilon_i - \sum_{j=1}^{d} \eta_j, \tag{5.3}$$

$$m_k = \delta_{r-k}(A) - \delta_{r-k+1}(A) + 1 \qquad (k = 1, \cdots, b). \tag{5.4}$$

For a regular pencil, in particular, the indices, m_0; b, $m_1, \cdots, m_b$, are determined by $\delta_k(A)$ $(k = 0, 1, 2, \cdots)$.

Proof.[1] The formulas (5.2)–(5.4) can be derived from the block diagonal structure (5.1). The uniqueness of the indices $\varepsilon_1, \cdots, \varepsilon_c$ and $\eta_1, \cdots, \eta_d$ is not difficult to establish.

The existence of block diagonal form (5.1) is proven here. Put $A(s) = sX + Y$, where X and Y are matrices over $\boldsymbol{F}$.

Claim: There exist nonsingular matrices $\bar{P}$ and $\bar{Q}$ over $\boldsymbol{F}$ and partitions $(R_1, \cdots, R_\mu; R_\infty)$ and $(C_1, \cdots, C_\mu; C_\infty)$ of the row set R and the column set C of $\bar{A}(s) = s\bar{X} + \bar{Y} = \bar{P}(sX + Y)\bar{Q}$ such that

$$
\begin{aligned}
&\operatorname{rank} \bar{X}[R_i, C_j] = 0 \qquad (1 \le j \le \mu, j \le i \le \infty),\\
&\operatorname{rank} \bar{Y}[R_i, C_j] = 0 \qquad (1 \le j \le \mu, j+1 \le i \le \infty),\\
&\operatorname{rank} \bar{X}[R_{j-1}, C_j] = |C_j| \qquad (2 \le j \le \mu),\\
&\operatorname{rank} \bar{X}[R_\infty, C_\infty] = |C_\infty|,\\
&\operatorname{rank} \bar{Y}[R_i, C_i] = |R_i| \qquad (1 \le i \le \mu).
\end{aligned}
$$

Here R_μ, R_∞, and C_∞ can be empty, whereas other blocks are nonempty. Note that $\bar{A}(s)$ is an upper block-triangular matrix and that the rank conditions imply

$$|C_1| \ge |R_1| \ge |C_2| \ge |R_2| \ge \cdots \ge |C_\mu| \ge |R_\mu|, \quad |R_\infty| \ge |C_\infty|.$$

The upper block-triangular form in the claim can be constructed as follows. The column set C_1 is determined by a column-transformation for X, since $\bar{X}[R, C_1] = O$ and $\bar{X}[R, C \setminus C_1]$ is of full-column rank by the rank conditions. Then the row set R_1 is determined by a row-transformation for the submatrix $Y[R, C_1]$, since $\bar{Y}[R \setminus R_1, C_1] = O$ and $\bar{Y}[R_1, C_1]$ is of full-row rank. Next, C_2 is determined by a column-transformation for the submatrix $X[R \setminus R_1, C \setminus C_1]$ (with X denoting the modified X), since $\bar{X}[R \setminus R_1, C_2] = O$ and $\bar{X}[R \setminus R_1, (C \setminus C_1) \setminus C_2)]$ is of full-column rank. Then the row set R_2 is determined from the submatrix $Y[R \setminus R_1, C_2]$. Continuing this way, we eventually arrive at $C_{\mu+1} = \emptyset$ for some $\mu \ge 0$;

> **NOTE**: The condition "$C_{\mu+1} = \emptyset$" is easier to understand when it is rephrased as "$\bar{X}[R \setminus \bigcup_{i=1}^{\mu} R_i, C \setminus \bigcup_{j=1}^{\mu} C_j]$ is of full-column rank"

then we terminate by defining R_∞ and C_∞ to be the complements of $\bigcup_{i=1}^{\mu} R_i$ and $\bigcup_{j=1}^{\mu} C_j$, respectively.

In the above claim we may further assume (**see Section 5.1.4**) that

$$
\begin{aligned}
&\bar{X}[R_i, C_j] = O \qquad \text{unless } 2 \le j = i+1 \le \mu \text{ or } i = j = \infty,\\
&\bar{Y}[R_i, C_j] = O \qquad \text{unless } 1 \le i = j \le \infty,\\
&\bar{X}[R_{j-1}, C_j] = \begin{bmatrix} I_{|C_j|} \\ O \end{bmatrix} \qquad (2 \le j \le \mu),
\end{aligned}
$$

[1] The present proof, valid for an arbitrary $\boldsymbol{F}$, is communicated by S. Iwata. See Gantmacher [87, Chap. XII, §2] for an alternative proof in the case of $\boldsymbol{F} = \mathbf{C}$.

$$\bar{X}[R_\infty, C_\infty] = \begin{bmatrix} I_{|C_\infty|} \\ O \end{bmatrix},$$
$$\bar{Y}[R_i, C_i] = \begin{bmatrix} I_{|R_i|} & O \end{bmatrix} \qquad (1 \le i \le \mu),$$

where I_N denotes the identity matrix of order N. Then the matrix $\bar{A}(s)$ takes the form depicted in Fig. 5.1 for $\mu = 4$.

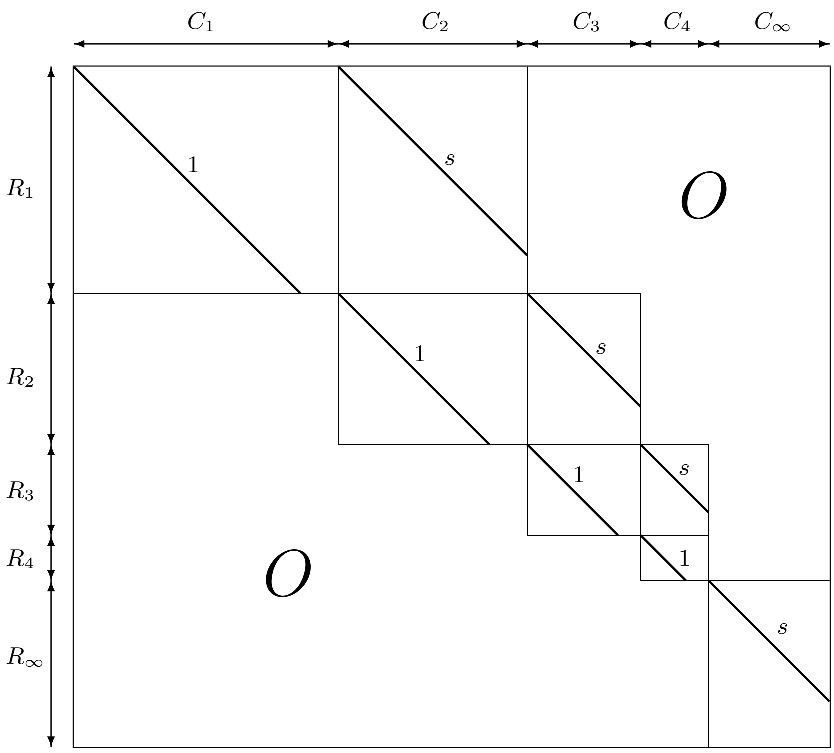

Fig. 5.1. Matrix $\bar{A}(s)$ in the proof for the Kronecker form ($\mu = 4$)

Consider the submatrix $\bar{A}[\bigcup_{i=1}^{\mu} R_i, \bigcup_{j=1}^{\mu} C_j]$. With suitable permutations of rows and columns, it can be put into a block-diagonal form with each diagonal block being equal to $N_m(s)$ with $1 \le m \le \mu$ or $L_\varepsilon(s)$ with $0 \le \varepsilon \le \mu-1$, where $N_m(s)$ appears with multiplicity $|R_m| - |C_{m+1}|$ (note: $|C_{\mu+1}| = 0$) and $L_\varepsilon(s)$ with multiplicity $|C_{\varepsilon+1}| - |R_{\varepsilon+1}|$. Namely,

$$\bar{A}[\bigcup_{i=1}^{\mu} R_i, \bigcup_{j=1}^{\mu} C_j] = \text{block-diag}\,(N_{m_1}(s), \cdots, N_{m_b}(s); L_{\varepsilon_1}(s), \cdots, L_{\varepsilon_c}(s))$$

with

$$b = \sum_{i=1}^{\mu} |R_i| - \sum_{j=2}^{\mu} |C_j|, \qquad c = \sum_{j=1}^{\mu} |C_j| - \sum_{i=1}^{\mu} |R_i|,$$
$$|\{k \mid m_k = m\}| = |R_m| - |C_{m+1}| \qquad (1 \leq m \leq \mu),$$
$$|\{k \mid \varepsilon_k = \varepsilon\}| = |C_{\varepsilon+1}| - |R_{\varepsilon+1}| \qquad (0 \leq \varepsilon \leq \mu - 1).$$

For the submatrix $\bar{A}[R_\infty, C_\infty]$, we apply the above argument to its transpose. Let $(\hat{R}_1, \cdots, \hat{R}_{\hat{\mu}}; \hat{R}_\infty)$ and $(\hat{C}_1, \cdots, \hat{C}_{\hat{\mu}}; \hat{C}_\infty)$ be the resulting partitions of R_∞ and C_∞, respectively. Since $\operatorname{rank} \bar{X}[R_\infty, C_\infty] = |C_\infty|$, we have $|\hat{C}_i| = |\hat{R}_{i+1}|$ for $i = 1, \cdots, \hat{\mu} - 1$ and $|\hat{C}_{\hat{\mu}}| = 0$. This means that there exist nonsingular matrices $\hat{P}$ and $\hat{Q}$ such that

$$\hat{P}\bar{A}[R_\infty, C_\infty]\hat{Q} = \text{block-diag}\,(U_{\eta_1}(s), \cdots, U_{\eta_d}(s); s\hat{X}_\infty + \hat{Y}_\infty),$$

where $\hat{X}_\infty$ is nonsingular. Finally, we can transform $\hat{X}_\infty$ to the identity matrix to obtain the desired block-diagonal form (5.1). ■

5.1.4 Detail of the Second Half of the Proof

Suppose that we are given a pencil $s\bar{X}+\bar{Y}$ that satisfies the conditions:

$$\begin{aligned}
&\operatorname{rank}\bar{X}[R_i,C_j]=0 \qquad (1\le j\le\mu, j\le i\le\infty),\\
&\operatorname{rank}\bar{Y}[R_i,C_j]=0 \qquad (1\le j\le\mu, j+1\le i\le\infty),\\
&\operatorname{rank}\bar{X}[R_{j-1},C_j]=|C_j| \qquad (2\le j\le\mu),\\
&\operatorname{rank}\bar{X}[R_\infty,C_\infty]=|C_\infty|,\\
&\operatorname{rank}\bar{Y}[R_i,C_i]=|R_i| \qquad (1\le i\le\mu)
\end{aligned}$$

in Claim (page 276). We will transform it to another pencil $s\bar{X}+\bar{Y} := P(s\bar{X}+\bar{Y})Q$ that satisfies the following conditions:

$$\begin{aligned}
&\bar{X}[R_i,C_j]=O \qquad \text{unless } 2\le j=i+1\le\mu \text{ or } i=j=\infty,\\
&\bar{Y}[R_i,C_j]=O \qquad \text{unless } 1\le i=j\le\infty,\\
&\bar{X}[R_{j-1},C_j]=\begin{bmatrix} I_{|C_j|}\\ O\end{bmatrix} \qquad (2\le j\le\mu),\\
&\bar{X}[R_\infty,C_\infty]=\begin{bmatrix} I_{|C_\infty|}\\ O\end{bmatrix},\\
&\bar{Y}[R_i,C_i]=\begin{bmatrix} I_{|R_i|} & O\end{bmatrix} \qquad (1\le i\le\mu),
\end{aligned}$$

shown at the bottom of page 276. In the following we assume $\mu=4$.

1. Suppose that $(\bar{X},\bar{Y})$ satisfies the conditions in Claim. This is depicted in Fig. 5.2, where $\equiv$ means a submatrix of full-row rank and $|||$ a submatrix of full-column rank.

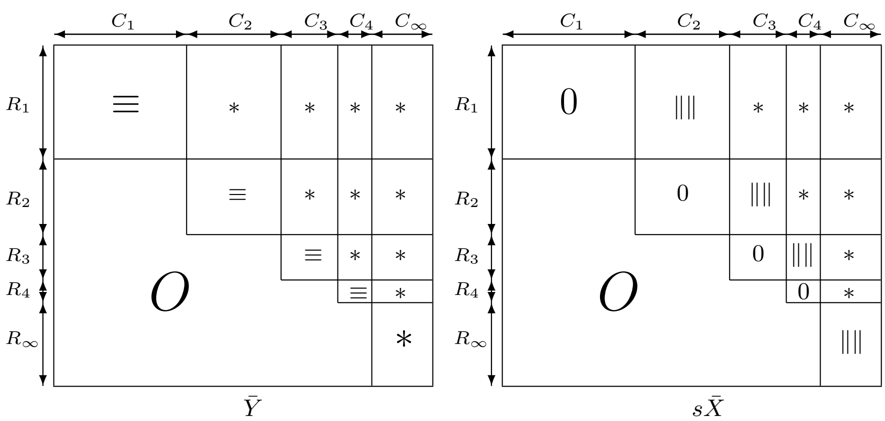

Fig. 5.2. Proof for the Kronecker form (1); initial form

2. The submatrix $\bar{Y}[R_1, C_1]$ is of full-row rank, and therefore it can be transformed to $\bar{Y}[R_1, C_1] = \begin{bmatrix} I_{|R_1|} & O \end{bmatrix}$ (rank normal form). The entire matrix looks like Fig. 5.3. The matrix $\bar{X}$ retains its structure (zero/nonzero pattern, full-rankness of submatrices).

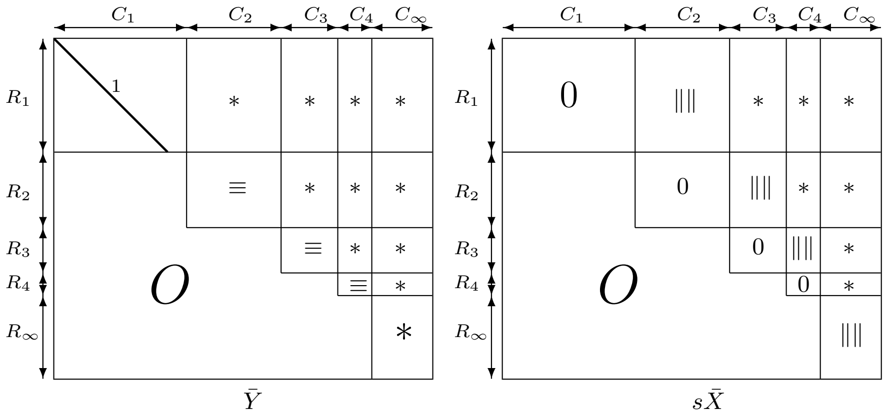

Fig. 5.3. Proof for the Kronecker form (2)

3. The submatrix $\bar{X}[R_1, C_2]$ is of full-column rank, and therefore it can be transformed to $\bar{X}[R_1, C_2] = \begin{bmatrix} I_{|C_2|} \\ O \end{bmatrix}$ (rank normal form).

   This process involves a row-transformation for R_1; let P denote the matrix representing this row-transformation. The row-transformation for R_1 destroys the rank normal form of $\bar{Y}[R_1, C_1]$ constructed above, but we can restore the normal form by applying the inverse column transformation to C_1, i.e, by multiplying $\begin{bmatrix} P^{-1} & O \\ O & I \end{bmatrix}$ from the right.

   Then the entire matrix looks like Fig. 5.4.

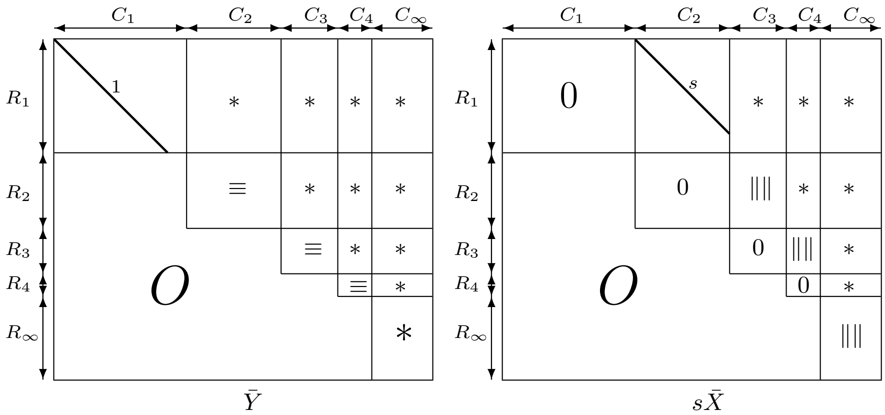

Fig. 5.4. Proof for the Kronecker form (3)

4. The submatrix $\bar{Y}[R_2, C_2]$ is of full-row rank, and therefore it can be transformed to $\bar{Y}[R_2, C_2] = \begin{bmatrix} I_{|R_2|} & O \end{bmatrix}$ (rank normal form).

   This process involves a column-transformation for C_2; let Q denote the matrix representing this column-transformation. The column-transformation for C_2 destroys the rank normal form of $\bar{X}[R_1, C_2]$ constructed above, but we can restore the normal form by applying the inverse row transformation to R_1, i.e, by multiplying $\begin{bmatrix} Q^{-1} & O \\ O & I \end{bmatrix}$ from the left.

   This row-transformation for R_1, however, destroys the normal form of $\bar{Y}[R_1, C_1]$, and we restore its normal form by applying the inverse column transformation to C_1, i.e, by multiplying $\begin{bmatrix} Q & O \\ O & I \end{bmatrix}$ from the right.

   Then the entire matrix looks like Fig. 5.5.

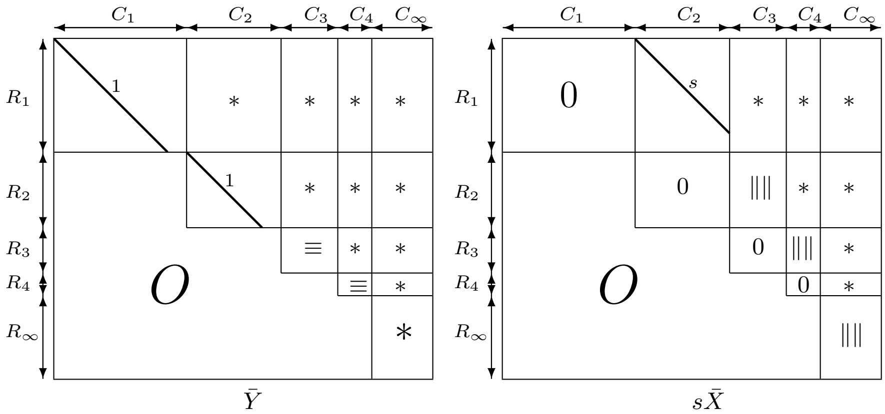

Fig. 5.5. Proof for the Kronecker form (4)

5. Continuing in this way we arrive at a matrix of the form of Fig. 5.6. We have successfully realized the submatrices I in $\bar{Y}$ and sI in $s\bar{X}$. We are going to eliminate the upper off-diagonal entries.

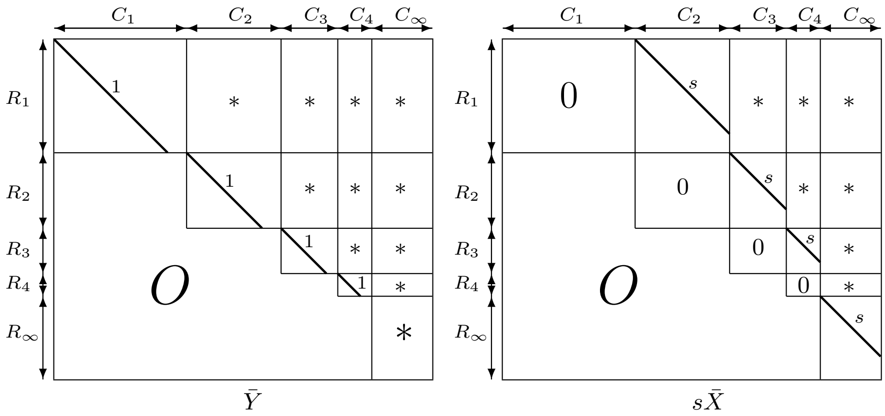

Fig. 5.6. Proof for the Kronecker form (5)

6. **(Rows in R_4)** Since $\bar{X}[R_\infty, C_\infty]$ contains an identity matrix, we can change $\bar{X}[R_4, C_\infty]$ to O by a row transformation. Next, since $\bar{Y}[R_4, C_4]$ contains an identity matrix, we can change $\bar{Y}[R_4, C_\infty]$ to O by a column transformation. Note that $\bar{X}[R_4, C_\infty]$ is not affected, since $\bar{X}[R_4, C_4] = O$. Then the entire matrix looks like Fig. 5.7.

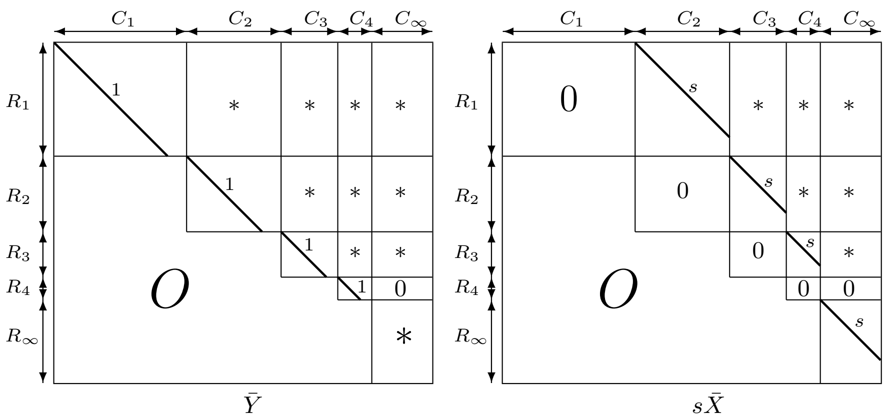

Fig. 5.7. Proof for the Kronecker form (6)

7. **(Rows in R_3)** Since $\bar{X}[R_\infty, C_\infty]$ contains an identity matrix, we can change $\bar{X}[R_3, C_\infty]$ to O by a row transformation. Next, since $\bar{Y}[R_3, C_3]$ contains an identity matrix, we can change $\bar{Y}[R_3, C_4 \cup C_\infty]$ to O by a column transformation. Note that $\bar{X}[R_3, C_4 \cup C_\infty]$ is not affected, since $\bar{X}[R_3, C_3] = O$. Then the entire matrix looks like Fig. 5.8.

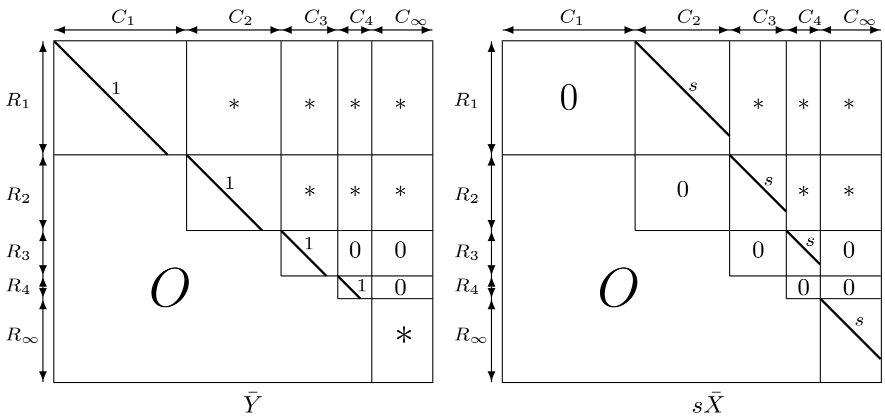

Fig. 5.8. Proof for the Kronecker form (7)

8. **(Rows in R_2)** Since $\bar{X}[R_3 \cup R_\infty, C_4 \cup C_\infty]$ contains an identity matrix, we can change $\bar{X}[R_2, C_4 \cup C_\infty]$ to O by a row transformation. Next, since $\bar{Y}[R_2, C_2]$ contains an identity matrix, we can change $\bar{Y}[R_2, C_3 \cup C_4 \cup C_\infty]$ to O by a column transformation. Note that $\bar{X}[R_2, C_3 \cup C_4 \cup C_\infty]$ is not affected, since $\bar{X}[R_2, C_2] = O$. Then the entire matrix looks like Fig. 5.9.

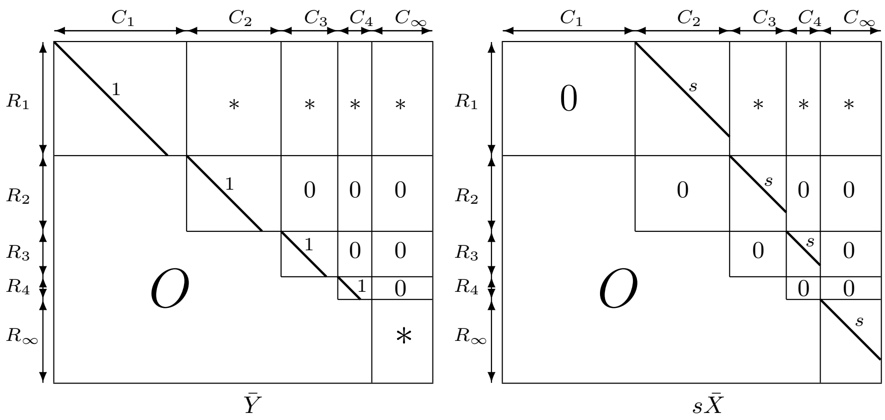

Fig. 5.9. Proof for the Kronecker form (8)

9. **(Rows in R_1)** Since $\bar{X}[R_2 \cup R_3 \cup R_\infty, C_3 \cup C_4 \cup C_\infty]$ contains an identity matrix, we can change $\bar{X}[R_1, C_3 \cup C_4 \cup C_\infty]$ to O by a row transformation. Next, since $\bar{Y}[R_1, C_1]$ contains an identity matrix, we can change $\bar{Y}[R_1, C_2 \cup C_3 \cup C_4 \cup C_\infty]$ to O by a column transformation. Note that $\bar{X}[R_1, C_2 \cup C_3 \cup C_4 \cup C_\infty]$ is not affected, since $\bar{X}[R_1, C_1] = O$. Then the entire matrix looks like Fig. 5.10, which is in the desired form. (END)

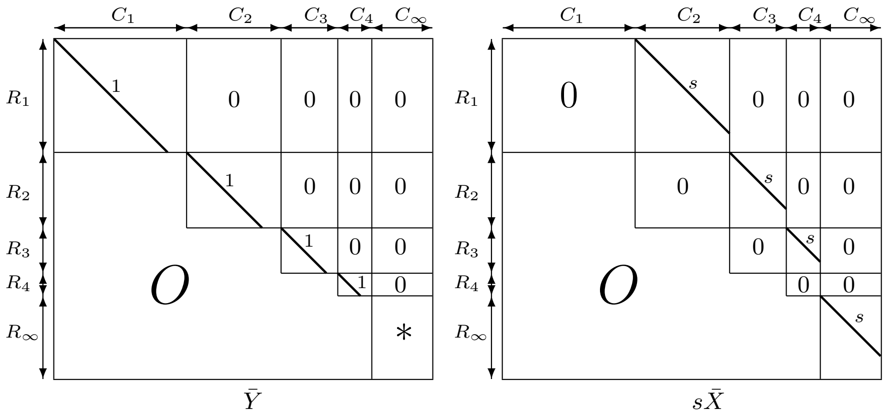

Fig. 5.10. Proof for the Kronecker form (9); the desired form