Submitted to the Bulletin of the Australian Mathematical Society
doi:10.1017/S ...

RANK MINIMIZATION FOR THE GENERALIZED SYLVESTER EQUATION

JUN XU

Abstract

The well known Roth's similarity theorem is a special case of a rank minimization theorem.

2010 *Mathematics subject classification:* primary 15A24.
Keywords and phrases: Roth's similarity theorem, rank minimization, Sylvester equation.

1. Introduction

Denote by $\boldsymbol{A} \in \mathbb{F}^{m\times m}$, $\boldsymbol{B} \in \mathbb{F}^{n\times n}$, and $\boldsymbol{C} \in \mathbb{F}^{m\times n}$ three given matrices over some field $\mathbb{F}$. Denote by $\mathrm{GL}(n, \mathbb{F}) = \{\boldsymbol{M} \in \mathbb{F}^{n\times n} | \det \boldsymbol{M} \neq 0\}$ the general linear group of degree n over the field $\mathbb{F}$. The following theorem is the well-known Roth's similarity theorem, which provides a sufficient and necessary condition for guaranting the consistency of the Sylvester equation (1.1) in terms of an equivalence between two associated matrices.

Theorem 1.1 ([1]). *The matrix equation*

$$\boldsymbol{A}\boldsymbol{X} - \boldsymbol{X}\boldsymbol{B} = \boldsymbol{C} \tag{1.1}$$

is solvable w.r.t $\boldsymbol{X} \in \mathbb{F}^{m\times n}$ *if and only if there exists a matrix* $\boldsymbol{P} \in GL(m+n, \mathbb{F})$ *such that*

$$\boldsymbol{P}\begin{pmatrix} \boldsymbol{A} & \boldsymbol{C} \\ \boldsymbol{0} & \boldsymbol{B} \end{pmatrix} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{B} \end{pmatrix}\boldsymbol{P}$$

We now generalize the Roth's similarity theorem to a general result based on rank minimization. The purpose of this note is to prove the following theorem.

Theorem 1.2. *Given* $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ *defined above, denote by we have*

$$\min\{\mathrm{rank}(\boldsymbol{A}\boldsymbol{X} - \boldsymbol{X}\boldsymbol{B} - \boldsymbol{C}) | \boldsymbol{X} \in \mathbb{F}^{m\times n}\}$$

$$= \min\left\{\mathrm{rank}\left[\boldsymbol{P}\begin{pmatrix} \boldsymbol{A} & \boldsymbol{C} \\ \boldsymbol{0} & \boldsymbol{B} \end{pmatrix} - \begin{pmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{B} \end{pmatrix}\boldsymbol{P}\right] | \boldsymbol{P} \in \mathrm{GL}(m+n, \mathbb{F})\right\}.$$

Include acknowledgement of support here

PROOF. Denote by

$$\alpha(\boldsymbol{X}) = \boldsymbol{AX} - \boldsymbol{XB} - \boldsymbol{C}, \tag{1.2}$$

and

$$\beta(\boldsymbol{P}) = \boldsymbol{P}\begin{pmatrix} \boldsymbol{A} & \boldsymbol{C} \\ \boldsymbol{0} & \boldsymbol{B} \end{pmatrix} - \begin{pmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{B} \end{pmatrix}\boldsymbol{P}. \tag{1.3}$$

Define

$$R_\alpha = \min \operatorname{rank}\{\alpha(\boldsymbol{X}) | \boldsymbol{X} \in \mathbb{F}^{m\times n}\} \tag{1.4}$$

and

$$R_\beta = \min \operatorname{rank}\{\beta(\boldsymbol{P}) | \boldsymbol{P} \in \mathrm{GL}(m+n, \mathbb{F}).\} \tag{1.5}$$

From [2, 3], we can obtain that

$$R_\alpha = \operatorname{rank}\begin{pmatrix} \boldsymbol{A} & \boldsymbol{C} \\ \boldsymbol{0} & \boldsymbol{B} \end{pmatrix} - \operatorname{rank}\begin{pmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{B} \end{pmatrix}. \tag{1.6}$$

For $\boldsymbol{P} \in \mathrm{GL}(m+n, \mathbb{F})$, we can obtain that

$$\begin{aligned} \operatorname{rank}\beta(\boldsymbol{P}) &\geq \operatorname{rank}\boldsymbol{P}\begin{pmatrix} \boldsymbol{A} & \boldsymbol{C} \\ \boldsymbol{0} & \boldsymbol{B} \end{pmatrix} - \operatorname{rank}\begin{pmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{B} \end{pmatrix}\boldsymbol{P} \\ &= \operatorname{rank}\begin{pmatrix} \boldsymbol{A} & \boldsymbol{C} \\ \boldsymbol{0} & \boldsymbol{B} \end{pmatrix} - \operatorname{rank}\begin{pmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{B} \end{pmatrix} \\ &= R_\alpha. \end{aligned} \tag{1.7}$$

Hence, we have that $R_\beta \geq R_\alpha$. On the other hand, we denote by

$$\boldsymbol{P_X} = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{X} \\ \boldsymbol{0} & \boldsymbol{I} \end{pmatrix}, \tag{1.8}$$

then we have

$$\beta(\boldsymbol{P_X}) = \begin{pmatrix} \boldsymbol{0} & -\alpha(\boldsymbol{X}) \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}. \tag{1.9}$$

Therefore,

$$R_\beta \leq \min\{\operatorname{rank}\beta(\boldsymbol{P_X}) | \boldsymbol{X} \in \mathbb{F}^{m\times n}\} = \min\{\operatorname{rank}\alpha(\boldsymbol{X})\} = R_\alpha. \tag{1.10}$$

This completes the proof. □

References

[1] R. E. Roth, 'The equations $AX - BY = C$ and $AX - XB = C$ in equations', *Proc. Amer. Math. Soc.* **3** (1952), 392-396.
[2] Y. Tian, 'The minimal rank of matrix expression $A - BX - YC$', *Missouri J. Math. Sci.* **14** (2002), 40-48.
[3] Y. Tian, 'Rank equalities related to generalized inverses of matrices and their applications', Master Thesis, Montreal, Quebec, Canada (2000).

Jun Xu, Department of Computing, The Hong Kong Polytechnic University, Hung Hom, Hong Kong, China
e-mail: csjunxu@comp.polyu.edu.hk