5-2011

Solution Theory for Systems of Bilinear Equations

Dian Yang
College of William and Mary

Solution Theory for Systems of Bilinear Equations

A thesis submitted in partial fulfillment of the requirement
for the degree of Bachelor of Science in Mathematics from
The College of William and Mary

by

Dian Yang

Accepted for Honors

Charles R. Johnson, Advisor

Ryan Vinroot

John Delos

Williamsburg, VA

April 15, 2011

ABSTRACT

Solution Theory for Systems of Bilinear Equations

Dian Yang
Department of Mathematics
Bachelor of Science in Mathematics

For $A_1,\ldots,A_m \in M_{p,q}(\mathbb{F})$ and $g \in \mathbb{F}^m$, any system of equations of the form $y^T A_i x = g_i$, $i = 1,\ldots,m$, with y varying over $\mathbb{F}^p$ and x varying over $\mathbb{F}^q$ is called bilinear. A solution theory for complete systems ($m = pq$) is given in [1]. In this paper we give a general solution theory for bilinear systems of equations. For this, we notice a relationship between bilinear systems and linear systems. In particular we prove that the problem of solving a bilinear system is equivalent to finding rank one points of an affine matrix function. And we study how in general the rank one completion problem can be solved. We also study systems with certain left hand side matrices $\{A_i\}_{i=1}^m$ such that a solution exist no matter what right hand side g is. Criteria are given to distinguish such $\{A_i\}_{i=1}^m$.

Keywords: bilinear system of equations, bilinear forms, rank one completion problem

ACKNOWLEDGMENTS

First and foremost, I would like to acknowledge my advisor Professor Charles R. Johnson, who has been extremely helpful in teaching me how to do research, how to speak proper English and how to be humorous, as well as how to be a decent person. I would also like to thank Mr. Michael Shilling and Ms. Olivia Walch who gave me good inspiration during the research. And last, but not least, I thank Professors John Delos and Ryan Vinroot, who kindly agreed to be on my honors committee and have been very nice to me during my career at William and Mary.

Contents

Chapter 1

Introduction

We are well acquainted with linear systems which are most commonly seen in sciences and engineering. In most occasions they are written in vector form

$$Ax = b, \tag{1.1}$$

in which $A \in M_{m,n}(\mathbb{F})$, $x \in \mathbb{F}^n$, $b \in \mathbb{F}^m$. Such a system consists of m individual equations

$$a_i^T x = b_i, \quad i = 1, 2 \ldots, m, \tag{1.2}$$

with $a_i \in \mathbb{F}^n$, $x \in \mathbb{F}^n$, $b_i \in \mathbb{F}$. In each equation, there is a linear function on the left hand side and a constant on the right hand side.

One might extend this formalism and conceive a system in which the variables may be partitioned into two disjoint subsets, such that the left hand sides are linear in each set separately, and the right hand sides remain constants. Such left hand sides are called bilinear forms. It is known that all bilinear forms can be written in the form of $y^T Ax$, in which A is a matrix and y and x are two independent vectors of variables. Hence a system of m bilinear equations can be written as

$$y^T A_i x = g_i, \quad i = 1, 2 \ldots, m, \tag{1.3}$$

in which $A_i \in M_{p,q}(\mathbb{F})$, $y \in \mathbb{F}^p$, $x \in \mathbb{F}^q$, $g_i \in \mathbb{F}$. Accordingly, we call such a system a *system of bilinear equations (BLS)*. (We will fix and use this notation (p,q,m,x,y) throughout.)

We cannot help but notice the similarity between a bilinear system and linear system. For example, for both systems there are homogeneous cases, in which the right hand side constants are identically zero. Also, just as one can add a multiple of a particular equation to another without changing the solution of a linear system, one can do the same to a bilinear system. This, in fact means that, if we see the A_i's in the bilinear system (1.3) as vectors (just like a_i's in the linear system (1.2)), we can perform Gaussian elimination on the bilinear system as well.

With this tool in mind, we now restrict our attention to only bilinear systems whose matrices A_i's in the left hand side bilinear forms (which from now on we shall refer to as *LHS matrices*) are linearly independent as vectors.

Why should we not care about other kinds of bilinear systems? If we pick a BLS whose LHS matrices are linearly dependent, we can always perform Gaussian elimination on the set of equations. As a result we shall get a new set of equations whose LHS matrices consists of a smaller set of linearly independent matrices and a number of zero matrices. Depending on whether the right hand side constants are zeros or not, those equations with zero LHS matrices could either have form "$0 = 0$" or "$0 = *$", where $*$ is a nonzero constant. If all of these equations with zero LHS matrices have form "$0 = 0$", which are redundant, we can discard them and a new BLS whose LHS matrices are linearly independent remains. If one of these equations has the form "$0 = *$", we have a contradiction, and the set of solutions of this BLS is empty. In either case we need to do nothing more than solving a BLS whose LHS matrices are linearly independent. Therefore, it suffices to restrict our attention to only such systems. From now on we only care about linear or bilinear systems whose left hand side are linearly independent.

Further more, there is a notion of a complete and incomplete linear system of equation. This classification distinctly separate linear systems into two kinds: in the complete case we have as many equations as variables ($m = n$), and then have an unique solution. In the incomplete case we have fewer equations than variables ($m < n$), in which case we have an affine solution space of dimension $n - m$. Here we note that the case in which we have more equations than variables ($m > n$) does not occur under our pre-processing assumption.

Do we expect to see a similar classification for bilinear systems? As we shall see in the next chapter, the answer is yes. We shall also have the complete case where $m = pq$ and the incomplete case where $m < pq$. And we shall show that the case $m > pq$ cannot occur (under our assumption) and that the solution of a complete bilinear system is in some way also "unique".

Simple as the formalism might be, such systems have been rarely studied with exception of two papers. One is by Cohen and Tomasi [2], in whose paper an iterative algorithm was suggested, which under some cases, converges to a solution of a homogeneous bilinear system if a solution exists. The other one was by Johnson & Link [1], in whose paper complete BLS's are thoroughly studied and a complete solution was found. We shall clarify the concept of "completeness" in the following chapter, and give a more general method that help solve the incomplete case as well.

Bilinear systems may arise in many ways, but the authors of [1] has been motivated to study them because of their connection with the analysis of whether two patterns $\mathscr{P}$ and $\mathscr{Q}$ commute. This study is elaborated in this paper [3]. There are other applications as well. Together they will be discussed in the last chapter.

Chapter 2

Solution Theory for Bilinear Systems

A *solution* to a bilinear system is a pair of vectors $x \in \mathbb{F}^q$, $y \in \mathbb{F}^p$ simultaneously satisfying all m bilinear equations. Our purpose is to develop a theory that both determines whether there is any solution to a bilinear system (the *solvability* problem) and finds all the solutions of this system (*solving* a BLS). We shall note that this is generally very hard to do (but we have developed a theory that helps).

2.1 Observations

There is a few observations about the solution set of a BLS we can make before actually solving the system.

We shall first notice, if (x,y) is a solution to a BLS, so is $(tx, \frac{1}{t}y)$ for all $0 \neq t \in \mathbb{F}$. Hence the set of solution can be partitioned in to equivalent classes: "$[(x,y)]$", defined by the equivalent relation

$$(x,y) \sim (tx, \frac{1}{t}y).$$

This observation tells that the total number of variables is slightly less than it appears at first glance. (more like $p+q-1$ instead of $p+q$.) It also tells that there is no BLS that can have a unique solution as a linear system can, the best one can do is to have a "unique class", as will be pointed out soon.

Secondly, as pointed out by Cohen and Tomasi [2], we can view a bilinear system

$$y^T A_i x = g_i, \quad i = 1, 2 \ldots, m,$$

as a linear system, by fixing either x or y. In some sense, it means slicing the p-by-q-by-m 3-dimensional array formed by stacking LHS matrices $A_1, A_2, \ldots, A_m$ in the two other directions. For example, if the vector y or x is fixed, the bilinear system becomes a linear system $Yx = g$ or $X^T y = g$. Here, $g = (g_1, g_2, \ldots, g_m)^T$ and

$$Y = \begin{pmatrix} y^T A_1 \\ y^T A_2 \\ \vdots \\ y^T A_m \end{pmatrix} = y_1 R_1 + y_2 R_2 + \ldots + y_p R_p$$

in which R_i is an m-by-q matrix and is the slicing of $\mathbb{A}$ with the i-th rows of $A_1, \ldots, A_m$. Similarly,

$$X = \begin{pmatrix} A_1 x & A_2 x & \ldots & A_m x \end{pmatrix} = x_1 S_1 + \ldots x_q S_q$$

in which S_j is a p-by-m matrix and has the j-th columns of $A_1, \ldots A_m$ in order. If (x, y) is a solution to a bilinear system with right hand side g, then x will be a solution of the linear system $Yx = g$, and conversely y will be a solution to the linear system $X^T y = g$.

In this way, the solutions of a bilinear system can be expressed as the union of solution sets of linear systems with the other variable as the index. However, this method doesn't show us the true structure of the solution set of a bilinear system. We shall not use this approach in our paper.

Further more, if we look at any homogeneous system

$$y^T A_i x = 0, \quad i = 1, 2 \ldots, m, \tag{2.1}$$

where $A_i \in M_{p,q}(\mathbb{F})$, $y \in \mathbb{F}^p$, $x \in \mathbb{F}^q$. We shall notice it always has the *trivial solutions*: $x = 0$, y arbitrary, and x arbitrary, $y = 0$. On the other hand, if a BLS has a trivial solution, then it must be homogeneous ($g_i = 0$ results by plugging in $x = 0$ or $y = 0$), and hence have all the trivial solutions. In other words, inhomogeneous systems have only *nontrivial solutions* (solutions such that $x \neq 0$ and $y \neq 0$). This much is clear, it's natural to focus our attention to only nontrivial solutions to homogeneous systems and solutions to inhomogeneous systems. (From now on, when we speak of the solution of a BLS, we assume we are talking about nontrivial solutions.)

2.2 Symmetries of BLS

Some bilinear systems looks different but they are essentially the same. To minimize our work we need to identify such systems.

First, as mentioned in the introduction, we can alter a bilinear system in several ways without changing its solution set, they are analogous to elementary operations on linear systems. Notice linear operations on a bilinear equation correspond to linear operations on the LHS matrices A_i's and right hand side constants g_i's. Hence we may use pairs "(A_i, g_i)" to represent equations in a BLS and summarize these transformations below:

(i) The (A_i, g_i) pairs may be permuted.

(ii) An (A_i, g_i) pair may be multiplied by a nonzero scalar.

(iii) An (A_i, g_i) pair may be replaced by itself plus a linear combination of the other (A_j, g_j) pairs.

Notice that $A_1, A_2, \ldots, A_m$ may be taken to be any basis of the space that they span using the above transformations, with appropriate modification of g_i's on the right hand side. This greatly reduces the number of bilinear systems we need to work with.

As a side note, the *elementary linear operations* (i), (ii), (iii) could also be used to transform g^T to $(1,0,\ldots,0)$ in the cases of inhomogeneous systems. This transformation was used in [2] to transform any inhomogeneous bilinear system to an "almost homogeneous" one, so that they could utilize the algorithm they developed for homogeneous system on inhomogeneous systems as well.

Secondly, we point out an additional type of transformation that doesn't maintain the solution set, but preserves solvability of a bilinear system. This transformation is a change of coordinates,

$$(x,y) \to (Q^{-1}x, P^{-1}y),$$

or equivalently, simultaneous equivalence on the matrices A_i,

$$A_i \to PA_iQ,\ i = 1,2,\ldots,m,$$

in which $P \in M_p(\mathbb{F})$ and $Q \in M_q(\mathbb{F})$ are nonsingular. In this case, the right hand side g is not changed, and the new bilinear system is solvable if and only if the original one was. The relation above also gives a one-to-one correspondence between solutions of the original bilinear system and of the new bilinear system. We call such transformations *equivalence* transformations on bilinear systems, they could be used to make the solvability (or non-solvability) of bilinear systems more transparent.

2.3 Some Notations and Definitions

We shall introduce some notations and definitions that we are about to use.

First, we use $(\mathbb{A}, g)$ as a short hand of a bilinear system, where

$$\mathbb{A} := \{A_1, \ldots, A_m\}$$

is the set of LHS matrices of this bilinear system, and

$$g := (g_1, \ldots, g_m)^T$$

is the vector of right hand side constants. In situations where more than one sets of LHS matrices are present, we shall denote them as: $\mathbb{B} := \{B_1, \ldots, B_m\}$, $\mathbb{C} := \{C_1, \ldots, C_m\}$, etc.

“vec" operator transforms a p-by-q matrix into a pq column vector by stacking columns of the matrix from left to right. (This notation is used in standard Matrix Analysis books, for instance the one by Horn and Johnson [4]) For example:

$$\text{vec}\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$$

and

$$\text{vec}\begin{pmatrix} a & d & g \\ b & e & h \\ c & f & i \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \\ i \end{pmatrix}$$

Similarly, we can also define its inverse “unv" provided the dimensions of the matrix we want to transform into. For example:

$$\text{unv}\begin{pmatrix}1\\2\\3\\4\end{pmatrix} = \begin{pmatrix}1 & 3\\2 & 4\end{pmatrix}$$

Now using the definition above, for each set of LHS matrices $\mathbb{A}$, we define a unique *script matrix*:

$$\mathscr{A} := (\mathscr{A}_1, \mathscr{A}_2, \ldots, \mathscr{A}_m),$$

where

$$\mathscr{A}_i := \text{vec}A_i \in \mathbb{F}^{pq}.$$

Each script matrix $\mathscr{A}$ is a pq-by-m matrix that contain all the information of the LHS matrices of a BLS. Similarly, for sets of LHS matrices $\mathbb{B}$, $\mathbb{C}$, etc., the script matrices are denoted as $\mathscr{B}$, $\mathscr{C}$, etc., accordingly.

2.4 Ideas for Solving a BLS

Our main idea of solving a bilinear system is to transform it into a linear system with an additional condition. In short, it can be summarized as follows:

Theorem 1 *The set of solutions of a BLS*

$$y^T A_i x = g_i,\ i = 1, \ldots, m,$$

is equal to the set of solutions of the equation:

$$\mathscr{A}^T \text{vec}K = g,$$

in which $K = yx^T$*, and* $\mathscr{A} = (\text{vec}A_1, \ldots, \text{vec}A_m)$.

Proof. First, we define a new variable $K = yx^T$. (We shall keep this notation and call it the *K-matrix*.) If we write the left hand side of a bilinear equation $y^T Ax = g$ in coordinates we get:

$$\sum_{j=1}^{p}\sum_{k=1}^{q} y_j A_{jk} x_k = \sum_{j=1}^{p}\sum_{k=1}^{q} A_{jk} y_j x_k = \sum_{j=1}^{p}\sum_{k=1}^{q} A_{jk} K_{jk}. \tag{2.2}$$

Using the "vec" operator, we can write this bilinear equation $y^T Ax = g$ equivalently as

$$(\text{vec}A)^T \text{vec}K = g. \tag{2.3}$$

Hence a system of bilinear equations

$$y^T A_i x = g_i, \quad i = 1, 2 \ldots, m, \tag{2.4}$$

can be written alternatively as

$$(\text{vec}A_i)^T \text{vec}K = g_i, \quad i = 1, 2 \ldots, m. \tag{2.5}$$

We use notation $\mathscr{A} = (\text{vec}A_1, \ldots, \text{vec}A_m)$, and the system simplifies to

$$\mathscr{A}^T \text{vec}K = g, \tag{2.6}$$

with $g = (g_1, \ldots, g_m)^T$. □

Using Theorem 1 we reduces the BLS almost to a linear system

$$\mathscr{A}^T v = g, \tag{2.7}$$

with one additional condition on v. There are several ways to view the equation above. First, it could be seen as a set of linear equation in variables $y_j x_k$, $j = 1, \ldots, p$, $k = 1, \ldots, q$., or in tensor algebra words, v must be separable. (*i.e.* $v = y \otimes x$) A more inspiring way is to view v, after and *unv* operation, as a matrix that can be written as a product of two vectors: $K = yx^T$. Linear algebra tells us K must be a rank one matrix to yield a nontrivial solution for the BLS. (This result can be found in standard linear algebra books such as [5].) This allows a criterion for the solvability of bilinear systems:

Theorem 2 *A bilinear system*

$$y^T A_i x = g_i,\ i = 1, \ldots, m,$$

is solvable if and only if equation

$$\mathscr{A}^T \mathrm{vec} K = g, \tag{2.8}$$

has a rank one solution. (i.e. K which satisfy the equation 2.8 while being rank one.)

It's important to note the following fact:

Theorem 3 *There is a one-to-one relation between rank one solutions of equation 2.8 (K's) and classes of solutions of the original BLS. ($[(x,y)]$'s)*

Proof. On one hand, if $K = yx^T$ hold, then $K = \tilde{y}\tilde{x}^T$ hold for any $(\tilde{y}, \tilde{x}) \in [(x,y)]$, as

$$K = yx^T = \frac{1}{t} y \cdot tx^T = \tilde{y}\tilde{x}^T$$

On the other hand, if $K = \tilde{y}\tilde{x}^T$ and $K = yx^T$, then $(\tilde{y}, \tilde{x}) \in [(x,y)]$. This is because both x^T and $\tilde{x}^T$ (y and $\tilde{y}$) are proportional to a nonzero row (column) of K, and hence proportional to each other. Without loss of generality, let $\tilde{x} = tx$, $\tilde{y} = sy$. Relation $yx^T = K = \tilde{y}\tilde{x}^T$ implies $ts = 1$, which means

$$\tilde{x} = tx, \quad \tilde{y} = \frac{1}{t} y.$$

□

Theorem 1 to 3 together completely reduce the matter of solving a bilinear system to that of solving the equation 2.8 with the *rank one condition*. (all solutions need to have rank one.) This is nice. However, we do still need to solve equation 2.8. Essentially, equation 2.8 is the following linear system:

$$\mathscr{A}^T v = g. \tag{2.9}$$

We call it the *associated linear system* of the original BLS. It follows from the assumption that A_i's are linearly independent, the columns of $\mathscr{A}$ are also linearly independent. Like other linear systems, system 2.9 can be either complete or incomplete. Accordingly, we define a bilinear system to be *complete* if the associated linear system is complete ($pq = m$), and to be *incomplete* if the associated linear system is incomplete ($pq > m$). Note if $pq < m$, system 2.9 is not solvable due to our linear independence assumption, nor does the original BLS. We shall expand on these two cases respectively on section 2.5 and section 2.6.

Before we move on, we shall note that one can use the same method to cope with *multilinear systems*. Consider a system of equations whose left hand sides are linear to s distinct sets of variables:

$$\sum_{j_1=1}^{p_1}\sum_{j_2=1}^{p_2}\cdots\sum_{j_s=1}^{p_s} A_{ij_1j_2\cdots j_s}x^1_{j_1}x^2_{j_2}\cdots x^s_{j_s} = g_i,$$

where $A_1,\ldots,A_m$ are 0-s type tensors, $g \in \mathbb{F}^m$, $x_i \in \mathbb{F}^{p_i}$ for i from 1 to s.

Extend the definition of “vec" to an operator that transforms a tensor into a column vector by arranging the entries of the tensor lexicographically. Inherit the following notations:

$$\mathbb{A} := \{A_1,\ldots,A\}$$

$$\mathscr{A}_i := \mathrm{vec}A_i$$

$$\mathscr{A} := (\mathrm{vec}A_1,\mathrm{vec}A_2,\ldots,\mathrm{vec}A_m)$$

Define $K_{j_1j_2\cdots j_s} := x^1_{j_1}x^2_{j_2}\cdots x^s_{j_s}$. It follows:

$$\begin{aligned}
\mathscr{A}_i^T\mathrm{vec}K &= (\mathrm{vec}A_i)^T\mathrm{vec}K \\
&= \sum_{j_1=1}^{p_1}\sum_{j_2=1}^{p_2}\cdots\sum_{j_s=1}^{p_s} A_{ij_1j_2\cdots j_s}K_{j_1j_2\cdots j_s} \\
&= \sum_{j_1=1}^{p_1}\sum_{j_2=1}^{p_2}\cdots\sum_{j_s=1}^{p_s} A_{ij_1j_2\cdots j_s}x^1_{j_1}x^2_{j_2}\cdots x^s_{j_s} \\
&= g_i
\end{aligned}$$

Therefore, the mutilinear system can also be reduced to the following equation:

$$\mathscr{A}^T \mathrm{vec} K = g \tag{2.10}$$

This enables us to write down a parallel of theorem 1:

Theorem 4 *The set of solutions of a multilinear system*

$$\sum_{j_1=1}^{p_1} \sum_{j_2=1}^{p_2} \cdots \sum_{j_s=1}^{p_s} A_{i j_1 j_2 \cdots j_s} x^1_{j_1} x^2_{j_2} \cdots x^s_{j_s} = g_i \, i = 1, \dots, m,$$

is equal to the set of solutions of the equation:

$$\mathscr{A}^T v = g,$$

in which $v = x^1 \otimes x^2 \otimes \cdots \otimes x^s$*, and* $\mathscr{A} = (\mathrm{vec} A_1, \dots, \mathrm{vec} A_m)$.

2.5 Complete Bilinear Systems

Consider a complete system $(\mathbb{A}, g)$. By theorem 1 in section 2.4, it can be rewritten as equation:

$$\mathscr{A}^T \mathrm{vec} K = g, \tag{2.11}$$

where

$$K = yx^T. \tag{2.12}$$

Since BLS $(\mathbb{A}, g)$ is complete, $\mathscr{A}^T$ is a square matrix. Hence for any K we have

$$\mathrm{vec} K = (\mathscr{A}^T)^{-1} g, \tag{2.13}$$

and

$$K = \mathrm{unv}((\mathscr{A}^T)^{-1} g), \tag{2.14}$$

is unique, if exists.

By Theorem 3, BLS $(\mathbb{A}, g)$ has a nontrivial solution *if and only if* K has rank one. If so, pick any (x,y) such that $K = yx^T$, the set of nontrivial solutions is

$$[(x,y)] := \{(\tilde{x},\tilde{y}) : \tilde{x} = tx, \tilde{y} = \frac{1}{t}y, 0 \neq t \in \mathbb{F}\}, \tag{2.15}$$

which is a unique class. If K is not rank one, then BLS $(\mathbb{A}, g)$ has no nontrivial solutions.

It's interesting to see how this result is different from that of a complete linear system, which always yield a unique solution. For a complete bilinear system however, a solution is not always guaranteed. But if there is any solution, there must be a unique class of solutions.

As a side note, an alternative way of solving a complete system is suggested in the paper of Johnson and Link [1]. What happened there essentially is using the elementary linear operations mentioned in section 2.2 to perform a Gaussian elimination on the LHS matrices A_i's. This changes the LHS matrices to matrices with one 1 and 0's everywhere else. The left hand sides now become $y_i x_j$'s. Writing the equations in vector form directly gives an expression for yx^T. This process is equivalent to the step of taking inverse of $\mathscr{A}^T$ in our method above.

2.6 Incomplete Bilinear Systems

The matter of solving an incomplete bilinear system is little more delicate than that of solving a complete one. There are two ways we are going to do it. In what follows, we will introduce both and then prove their equivalences.

2.6.1 Solution Through Vector Parametic Forms

Consider an incomplete system $(\mathbb{A}, g)$. By theorem 1 in section 2.4, it can be rewritten as equation

$$\mathscr{A}^T \mathrm{vec} K = g, \tag{2.16}$$

where we have new variable

$$K = yx^T. \tag{2.17}$$

As we known from the theory of linear systems, the solutions of linear system $\mathscr{A}^T v = g$ has form

$$v = v_0 + z_1 v_1 + \cdots + z_r v_r, \tag{2.18}$$

where $r = pq - m$, v_0 is a particular solution to the problem, and $v_1, \ldots, v_r$ span the solution space of the associated homogeneous system $\mathscr{A}^T v = 0$. (There is a standard algorithm to compute $v_1, \ldots, v_r$, which is given in most Linear Algebra books [5]. But this render these vectors fixed. In principle they could be any set of linearly independent values in the space they span.) This form consists of a vector function with unknown z_i for $i = 1, \ldots, r$ as parameters. By applying "unv" operator to form 2.18 we obtain the solution set of $\mathscr{A}^T \mathrm{vec} K = g$:

$$K = K_0 + z_1 K_1 + \cdots + z_r K_r, \tag{2.19}$$

where $K_i := \mathrm{unv}(v_i)$ for $i = 0, \ldots, r$. We define $r = pq - m$ as the *codimension* of a BLS, as opposed to the number of equations "m" which can be under stood as the *dimension* of the system. Then we define the following matrix function:

$$K(z) := K_0 + z_1 K_1 + z_2 K_2 + \ldots z_r K_r. \tag{2.20}$$

By Theorem 3, a nonzero solution to the linear system will give a class of solutions to the bilinear system if and only if the matrix $K(z)$ has rank one for some choice of z. In this way we have a one-to-one relationship between rank one values of matrix function $K(z)$ and solution to the bilinear system. In particular, for each rank one value of $K(z)$, we pick a nonzero row x^T and non zero column y^T. Then we use relation $K(z) = yx^T$ to obtain a proper scale of x^T and y, and label this class of solutions $[(x,y)]_z$. The solutions of the original BLS can be then expressed as a union of classes of solutions associated with all z's such that $K(z)$ has rank one. *i.e.*

$$\bigcup_{z,\,\text{rank}K(x)=1} [(x,y)]_z$$

or

$$\bigcup_{z,\,\text{rank}K(x)=1} \{(x,y) : K(z) = yx^T\}. \tag{2.21}$$

We shall see examples of solving incomplete systems in this fashion later in this section.

Here we note that the z_i's belong not necessarily to field $\mathbb{F}$. One can also assume them to be in an extension field of $\mathbb{F}$. This implies that, in cases where there is no solution to a BLS in the field of parameters, it may well be that there are solutions in an extension field. This is very different compared to linear systems. We shall later show an example of a BLS with real parameters but only complex solutions.

This method reduces a system of bilinear equations to a system of linear equations. Since the solvabilities of systems of linear equations are well understood, this method seems to simplify the problem. However, after solving a linear system we need to decide if an affine space (2.19) contains a rank one matrix, which is in general a difficult problem.

2.6.2 Solution Through Completion

Another way of solving an incomplete bilinear system is to *complete* it into a complete bilinear system by adding additional equations. For example, if we have an incomplete

BLS $(\mathbb{A}, g)$, where $\mathbb{A} = \{A_1, \ldots, A_m\}$ and $g = (g_1, \ldots, g_m)$, we can add additional equations

$$y^T C_i x = z_i, \quad i = 1, \ldots, r, \tag{2.22}$$

where $r = pq - m$, and get a system of equations $(\mathbb{B}, \tilde{g}(z))$, where $\mathbb{B} = \{A_1, \ldots, A_m, C_1, \ldots, C_r\}$ and $\tilde{g}(z) = (g_1, \ldots, g_m, z_1, \ldots, z_r)$. Here, the matrices C_i's are added in a way that does not disrupt our linear independence assumption.

However, by mere adding equations we lose many solutions. Hence the equations 2.22 we add, different from bilinear equations, actually contains free parameters z_i, $i = 1, 2, \ldots, r$. We define this way of adding equations be a *completion of a bilinear system*. Notice a completion $(\mathbb{B}, \tilde{g}(z))$ is not a bilinear system any more, but a "bilinear system" with extra variables on the right hand side. In this way we do not decrease the amount of solutions we have, but instead increase the number of variable. However, for each fixed z system $(\mathbb{B}, \tilde{g}(z))$ is a complete BLS and becomes easy for us to solve. This precisely is described by the following theorem.

Theorem 5 *Let $(\mathbb{A}, g)$ be an incomplete bilinear system and $(\mathbb{B}, \tilde{g}(z))$ be a completion of this system. Then the set of solutions of $(\mathbb{A}, g)$ is*

$$\{(x, y) : (x, y, z) \text{ is a solution of equation system } (\mathbb{B}, \tilde{g}(z))\}, \tag{2.23}$$

or equivalently:

$$\bigcup_z \{\text{solutions of complete BLS } (\mathbb{B}, \tilde{g}(z))\}. \tag{2.24}$$

Proof. On one hand, for each (x, y) that is a solution of $(\mathbb{A}, g)$, we may take $z_i = y^T C_i x$, for $i = 1, \ldots, r$, where C_i's are the matrices used to obtain completion $(\mathbb{B}, \tilde{g}(z))$. Notice this (x, y, z) satisfy all the equations of system $(\mathbb{B}, \tilde{g}(z))$. Hence there exist z such that (x, y, z) is a solutions of system $(\mathbb{B}, \tilde{g}(z))$. (Here entries of z belong to $\mathbb{F}$ if entries in x and y are restricted to the original field $\mathbb{F}$, and entries of z belong to an extension field if entries in x and y are allowed to take value from this extension field)

On the other hand, take any solution (x,y,z) of system $(\mathbb{B},\tilde{g}(z))$, (x,y) satisfy the first m equations of system $(\mathbb{B},\tilde{g}(z))$, which exactly means (x,y) is a solution of BLS $(\mathbb{A},g)$. □

Note that there are many ways to complete a bilinear system, as there are far more than one choices of C_i matrices that satisfy our linear independence assumption.

Also note what completion to a BLS means in terms of its associated linear system. It means augmenting the pq-by-m matrix $\mathscr{A}$ with linearly independent columns to a pq-by-pq invertible matrix $\mathscr{B} = (\mathscr{A},\mathscr{C})$ by adding extra columns vecC_i's to the right hand side of $\mathscr{A}$, and extending the right hand side g to $\begin{pmatrix} g \\ z \end{pmatrix}$. For each fixed z we have a complete linear system

$$\mathscr{B}^T v = \begin{pmatrix} g \\ z \end{pmatrix}, \tag{2.25}$$

or equation system

$$\mathscr{B}^T \mathrm{vec}K = \begin{pmatrix} g \\ z \end{pmatrix}, \tag{2.26}$$

in terms of variable K.

Now we are ready to solve the original BLS. For each fixed z, equation 2.26 has unique solution

$$K = \mathrm{unv}((\mathscr{B}^T)^{-1} \begin{pmatrix} g \\ z \end{pmatrix}). \tag{2.27}$$

Equations 2.27 is in fact a matrix function linear to z. Hence we can write it in the form of a linear combination with z_i's as coefficients:

$$K'(z) = K_0 + z_1 K_1 + \ldots + z_r K_r.$$

By Theorem 3, the set of nontrivial solution of BLS $(\mathbb{B}, \tilde{g}(z))$ is

$$\{(x,y) : K = yx^T\},$$

which is non empty if and only if K has rank one. Finally, by Theorem 5 we just proved, we obtain the entirety of the set of nontrivial solutions of the original BLS:

$$\bigcup_{z,\mathrm{rank}K'(z)=1} \{(x,y) : K'(z) = yx^T\}. \tag{2.28}$$

2.6.3 Equivalence of the Two Methods

Compare the two results of the solution of a BLS: set 2.21 and set 2.28. Formally they are entirely identical. But $K(z)$ and $K'(z)$ signifies different content in each expression. The first method is a more natural parallel to the theory of complete BLS, as it uses the same method to deal with the associated linear system but applied it to the solution of incomplete linear systems. It is very nice in a theoretical point of view. But the second method provide a more explicit form of matrix function $K'(z)$ in equation 2.27, which helps a lot in the real solving process. In terms of obtaining a solution, it doesn't matter whether $K(z)$ and $K'(z)$ are really identical for the same BLS. But as a matter of fact they are, and it can be demonstrated in the following theorem.

Theorem 6 *Let $(\mathbb{A}, g)$ be an incomplete BLS, $\mathscr{A}^T v = g$ be its associated linear system, and $\mathscr{A}^T v = 0$ be the corresponding homogeneous linear system.*

For each completion $(\mathbb{B}, \tilde{g}(z))$, there exists a basis $\{v_i\}_{i=1}^r$ that span the solution space of $\mathscr{A}^T v = 0$ and a specific solution v_0 to $\mathscr{A}^T v = g$ such that $K(z)$ (as in formula 2.19) and $K'(z)$ (as in formula 2.27) are identical matrix functions, the choice of $\{v_i\}_{i=0}^r$ is unique.

One the other hand, for each basis $\{v_i\}_{i=1}^r$ that span the solution space of $\mathscr{A}^T v = 0$ and a specific solution v_0 to $\mathscr{A}^T v = g$ there exists a completion $(\mathbb{B}, \tilde{g}(z))$ such that $K(z)$ (as in *formula 2.19) and $K'(z)$ (as in formula 2.27) are identical matrix functions. (This choice is not unique)*

Proof. On one hand, for each completion $(\mathbb{B}, \tilde{g}(z))$, there is an unique

$$K'(z) = \mathrm{unv}((\mathscr{B}^T)^{-1} \begin{pmatrix} g \\ z \end{pmatrix}) = K_0 + z_1 K_1 + \ldots + z_r K_r,$$

which is the solution of equation

$$\mathscr{B}^T \mathrm{vec} K = \begin{pmatrix} g \\ z \end{pmatrix}.$$

Hence if we define

$$v_i := \mathrm{unv} K_i, \quad i = 0, \ldots, r, \tag{2.29}$$

and

$$v(z) = v_0 + z_1 v_1 + \ldots + z_r v_r,$$

$v(z)$ must be the solution of equation

$$\mathscr{B}^T v = \begin{pmatrix} g \\ z \end{pmatrix}. \tag{2.30}$$

Take only the first m equations of system 2.30, we have:

$$\mathscr{A}^T v = g. \tag{2.31}$$

which is exactly the associated linear function of the original BLS. We know $v(z)$ is a solution for this system any z.

Take $z = 0$, we see v_0 is a specific solution to $\mathscr{A}^T v = g$, by linearity

$$v(z) - v_0 = z_1 v_1 + \ldots + z_r v_r$$

solves $\mathscr{A}^T v = 0$ for any z. Take $z = e_i$, for $i = 1, \ldots, r$, we see $\{v_i\}_{i=1}^r$ are all solutions of $\mathscr{A}^T v = 0$. Notice,

$$v(z) = (\mathscr{B}^T)^{-1} \begin{pmatrix} g \\ z \end{pmatrix}$$

$\{v_i\}_{i=1}^r$'s linear independence follows from that of columns of $(\mathscr{B}^T)^{-1}$. Since the solutions space of $\mathscr{A}^T v = 0$ is r dimensional, we know $\{v_i\}_{i=1}^r$ must span the solutions space of $\mathscr{A}^T v = 0$.

By definition this choice of $\{v_i\}_{i=0}^r$ is unique.

On the other hand, for any $\{v_i\}_{i=0}^r$ such that $\{v_i\}_{i=1}^r$ span the solution space of $\mathscr{A}^T v = 0$ and v_0 is a specific solution to $\mathscr{A}^T v = g$, we can find $\{u_i\}_{i=1}^m$ such that

$$\sum_{i=1}^m u_i g_i = v_0. \tag{2.32}$$

and

$$\mathscr{A}^T u_i = e_i, \quad i = 1, \ldots, m. \tag{2.33}$$

First we solve linear systems 2.33 and get an arbitrary set of solutions $\{u_i\}_{i=1}^m$ that doesn't necessarily satisfy condition 2.32. If $g = 0$ (original BLS is homogeneous), condition 2.32 is automatically met. If $g \neq 0$ (orignial BLS is inhomogeneous), notice

$$\mathscr{A}^T (\sum_{i=1}^m u_i g_i) = \sum_{i=1}^m g_i \mathscr{A}^T u_i = \sum_{i=1}^m e_i g_i = g. \tag{2.34}$$

Hence

$$v^* = \sum_{i=1}^m u_i g_i \tag{2.35}$$

is a solution to the associated linear system $\mathscr{A}^T v = g$. By linearity $v_0 - v_*$ is a solution to the associated homogeneous equation $\mathscr{A}^T v = 0$. Without loss of generality we assume $g_1 \neq 0$, then replace u_1 by

$$u_1^* = u_1 + \frac{1}{g_1}(v_0 - v^*) \tag{2.36}$$

Notice now set $\{u_i\}_{i=1}^m$ satisfy both condition 2.32 and condition 2.33. Also note that this choice is not unique.

Observe that $\{u_i\}_{i=1}^m \cup \{v_i\}_{i=1}^r$ form a linearly independent set. Hence we can define

$$\mathscr{B} = ((u_1, \ldots, u_m, v_1, \ldots, v_r)^{-1})^T. \tag{2.37}$$

Define a set of LHS matrices $\mathbb{B}$ accordingly as the set of columns of $\mathscr{B}$ after unv operation. Notice $(\mathbb{B}, \tilde{g}(z))$ is a completion of BLS $(\mathbb{A}, g)$. To verify this we just need to proof the first m columns of $\mathscr{B}$ coincides with those of $\mathscr{A}$, which is true since:

$$\begin{aligned} \mathscr{A}^T(\mathscr{B}^T)^{-1} &= \mathscr{A}^T(u_1, \ldots, u_m, v_1, \ldots, v_r) \\ &= (\mathscr{A}^T u_1, \ldots, \mathscr{A}^T u_m, \mathscr{A}^T v_1, \ldots, \mathscr{A}^T v_r) \\ &= (e_1, \ldots, e_m, 0, \ldots, 0) \\ &= (I, 0) \end{aligned} \tag{2.38}$$

There is a unique $K'(z)$ associated with completion $(\mathbb{B}, \tilde{g}(z))$. Now we need only to verify the fact that

$$K(z) = K'(z).$$

By formula 2.27,

$$\begin{aligned} K'(z) &= \mathrm{unv}((\mathscr{B}^T)^{-1} \begin{pmatrix} g \\ z \end{pmatrix}) \\ &= \mathrm{unv}((u_1, \ldots, u_m, v_1, \ldots, v_r) \begin{pmatrix} g \\ z \end{pmatrix}) \\ &= \mathrm{unv}(\sum_{i=1}^m u_i g_i + z_1 v_1 + \ldots + z_r v_r) \\ &= \mathrm{unv}(v_0 + z_1 v_1 + \ldots + z_r v_r) \\ &= K(z) \end{aligned} \tag{2.39}$$

Note that the choice of $(\mathbb{B}, \tilde{g}(z))$ is not unique since there are multiple ways to choose set $\{u_i\}_{i=1}^m$.

This concludes the proof. □

By theorem 6, we shall recognize $K(z)$ and $K'(z)$ as the same entity, and define it to be the *K-function* of an incomplete bilinear system. Both methods of solving incomplete BLS's have reduced the problem further to that of finding *rank one points* of a matrix function $K(z)$. We shall refer to this problem as the *rank one completion* problem, since we are completing a matrix that contains unknown variable to a rank one matrix.

Do notice, by theorem 6, the K-functions of an incomplete bilinear system is not unique, as we can choose different sets $\{v_i\}_{i=0}^r$ or use different completions.

Also notice that not all sets of matrices can be candidates of the the coefficients $K_0, \ldots, K_r$ in a K-function. Since $\mathrm{vec} K_i = v_i$ for $i = 1, \ldots, r$, which are linearly independent, so are $K_1, \ldots, K_r$. Furthermore, $K_0, \ldots, K_r$ must also be linearly independent if the original BLS is inhomogeneous. This is because v_0 is not in the kernel of $\mathscr{A}^T$, hence it cannot be a linear combination of $v_1, \ldots, v_r$.

2.6.4 Complete An Incomplete BLS In Different Ways

What changes for $K(z)$ if we use a different set of C_i's to obtain completion $(\mathbb{B}, \tilde{g}(z))$? The conclusion is the following:

Theorem 7 *When a BLS is completed in another way, the K-function $K(z)$ is reparametrized in terms of the z variable, which undergoes an affine transformation. (or a linear transformation for homogeneous systems)*

Proof. Assume we have two different augmentation $\mathscr{B}$ and $\mathscr{B}'$, where $\mathscr{B} = (\mathscr{A}, \mathscr{C})$ and $\mathscr{B}' = (\mathscr{A}, \mathscr{C}')$. Then by formula (2.27), we have:

$$K(z) = \text{unv}((\begin{pmatrix} \mathscr{A}^T \\ \mathscr{C}^T \end{pmatrix})^{-1} \begin{pmatrix} g \\ z \end{pmatrix}); \tag{2.40}$$

$$K'(z') = \text{unv}((\begin{pmatrix} \mathscr{A}^T \\ \mathscr{C}'^T \end{pmatrix})^{-1} \begin{pmatrix} g \\ z' \end{pmatrix}). \tag{2.41}$$

However, we know they represent the same set of matrices, which means there is a one-to-one correspondence between z and z' defined by relation $K(z) = K'(z')$. It follows:

$$(\begin{pmatrix} \mathscr{A}^T \\ \mathscr{C}^T \end{pmatrix})^{-1} \begin{pmatrix} g \\ z \end{pmatrix} = (\begin{pmatrix} \mathscr{A}^T \\ \mathscr{C}'^T \end{pmatrix})^{-1} \begin{pmatrix} g \\ z' \end{pmatrix}. \tag{2.42}$$

$$\begin{pmatrix} g \\ z' \end{pmatrix} = \begin{pmatrix} \mathscr{A}^T \\ \mathscr{C}'^T \end{pmatrix} \begin{pmatrix} \mathscr{A}^T \\ \mathscr{C}^T \end{pmatrix}^{-1} \begin{pmatrix} g \\ z \end{pmatrix} = \begin{pmatrix} I & 0 \\ S_1 & S_2 \end{pmatrix} \begin{pmatrix} g \\ z \end{pmatrix}, \tag{2.43}$$

where $\begin{pmatrix} S_1 & S_2 \end{pmatrix} = \mathscr{C}'^T \begin{pmatrix} \mathscr{A}^T \\ \mathscr{C}^T \end{pmatrix}^{-1}$. Hence we have the relations between z and z':

$$z' = S_1 g + S_2 z. \tag{2.44}$$

i.e. $K'(z')$ can be transformed to $K(z)$ if we perform the above affine transformation on the variable z, or a linear transformation if the original BLS is homogeneous. □

We have a lot of freedom in regard of what completion to use. The question is: is there any completion that is “better" than others? In a theoretical point of view, the answer should be no. (The answer may change if we talk about better in when practically solving a specific bilinear system.) The following theorem demonstrates why this might be the case.

Theorem 8 *For any specific solution* (x_0, y_0) *that corresponds to point* z_0 *of K-function:*

$$y_0 x_0^T = K(z_0),$$

we can perform transformation $K \longrightarrow K'$ *by using another completion, such that the new solution* (x_0, y_0) *corresponds instead to the origin of the new K-function*

$$y_0 x_0^T = K'(0).$$

Proof. Let

$$K(z) = K_0 + z_1 K_1 + \ldots + z_r K_r$$

be the K-function of a BLS for some completion. By Theorem 6, this corresponds to a choice of $v_i = \text{vec} K_i$, $i = 0, \ldots, r$, which are parameters of the general solution of the associated linear system, which has solution:

$$\{v : v = v_0 + z_1 v_1 + \ldots + z_r v_r\}.$$

Notice $\text{vec}K(z_0)$ belong to this set, by solution theory of linear systems, we can choose another set of vectors $\{v_i'\}_{i=0}^r$ such that $v_0' = \text{vec}K(z_0)$, and the general solution of the associated linear system can be expressed as:

$$\{v : v = v_0' + z_1 v_1' + \ldots + z_r v_r'\}.$$

Now define matrix function:

$$K'(z) = K_0' + z_1 K_1' + \ldots + z_r K_r',$$

where $K_i' = \text{unv} v_i'$, $i = 0, \ldots, r$. By Theorem 6, $K'(z)$ is the K-function of the original BLS for another completion. Hence

$$y_0 x_0^T = K(z_0) = K'(0).$$

□

This result looks a lot like the fact that one cannot choose a "right" representative solution v_0 in an incomplete linear system, as any solution is just a plane shift of another.

However, there are better completions to use if we are talking about simplifying the computation. In our examples we tend to use matrices with only one nonzero entry, the E_{ij}'s. From experiments, they significantly reduce the number of nonzero entries in K_i's.

2.7 Equivalence to Rank-One Completion Problems

In last section we have shown that for any bilinear system $(\mathbb{A}, g)$, we can find a K-function $K(z)$ (not unique), such that the problem of solving the BLS is equivalent to that of finding rank one points of the $K(z)$. We call the problem of finding rank one points of a matrix function a rank one completion problem. Hence any bilinear system can be reduced to a rank one completion problem.

We are curious if the BLS problem and the rank one completion problem are in general equivalent. That is asking, for an arbitrary matrix function $K(z)$, is there always a bilinear system whose K-function is exactly $K(z)$? This statement is not yet true, since a K-function must be an affine Matrix function and that $\{K_i\}_{i=0}^r$ (or $\{K_i\}_{i=1}^r$ for a homogeneous BLS) must be linearly independent. But will this condition for a matrix function suffice for it to be the K-function of some BLS? The answer is yes.

Theorem 9 *Let $K(z) = K_0 + z_1K_1 + \ldots + z_rK_r$ be an affine matrix function. If $\{K_i\}_{i=0}^r$ are linearly independent, (or $\{K_i\}_{i=1}^r$ are linearly independent and $K_0 = 0$) then $K(z)$ is the K-function of a certain inhomogeneous (or homogeneous) bilinear system .*

Proof. If $\{K_i\}_{i=0}^r$ are linearly independent, so are $\{\text{vec}K_i\}_{i=0}^r$. We can add $m-1$ extra vectors and complete $\{v\}_{i=2}^m$ to a basis of $\mathbb{F}^{pq}$, so that $\{\text{vec}K_i\}_{i=0}^r \cup \{v\}_{i=2}^m$ is still a linearly independent set. Let

$$\mathscr{B} = ((\mathrm{vec}K_0, v_2, \ldots, v_m, \mathrm{vec}K_1, \ldots, \mathrm{vec}K_r)^{-1})^T,$$

and

$$\mathscr{B} = (\mathscr{A}, \mathscr{C}),$$

where $\mathscr{A}$ represent the first m columns of $\mathscr{B}$ and $\mathscr{C}$ represent the other r columns of $\mathscr{B}$. Then let

$$g = (1, 0, \ldots, 0)^T.$$

Notice

$$K(z) = K_0 + \sum_{i=1}^{r} K_i z_i = \mathrm{unv}(\mathrm{vec}K_0 + \sum_{i=2}^{m} v_i g_i + \sum_{i=1}^{r} \mathrm{vec}K_i z_i) = \mathrm{unv}((\mathscr{B}^T)^{-1} \begin{pmatrix} g \\ z \end{pmatrix}).$$

Hence, by definition, $K(z)$ is the K-function of an inhomogeneous BLS $(\mathbb{A}, g)$ associated with completion $(\mathbb{B}, \tilde{g}(z))$.

If $\{K_i\}_{i=1}^r$ are linearly independent and $K_0 = 0$, we can add m extra vectors $\{v\}_{i=1}^m$ to this set $\{\mathrm{vec}K_i\}_{i=1}^r$ so that together they form a linearly independent set. Let

$$\mathscr{B} = ((v_1, \ldots, v_m, \mathrm{vec}K_1, \ldots, \mathrm{vec}K_r)^{-1})^T,$$

$$\mathscr{B} = (\mathscr{A}, \mathscr{C}),$$

and

$$g = (0, \ldots, 0)^T.$$

Notice

$$K(z) = \sum_{i=1}^{r} K_i z_i = \mathrm{unv}(\sum_{i=1}^{m} v_i g_i + \sum_{i=1}^{r} \mathrm{vec}K_i z_i) = \mathrm{unv}((\mathscr{B}^T)^{-1} \begin{pmatrix} g \\ z \end{pmatrix}).$$

Hence, by definition, $K(z)$ is the K-function of a homogeneous BLS $(\mathbb{A}, g)$ associated with completion $(\mathbb{B}, \tilde{g}(z))$. $\square$

The result of Theorem 9 indicates that the problem of solving incomplete bilinear systems is equivalent to the rank one completion problem of affine matrix functions. A complete solution to the latter will guarantee a complete solution to the former. Hence from now on we need only to focus on solving rank one completion problems.

How does one solves a rank one completion problem? From linear algebra we know, a nonzero matrix is rank one if and only if all its 2-by-2 minors are equal to zero [5]. The 2-by-2 minors of $K(z)$ are quadratic functions in the z_i's, hence solving a bilinear system comes down to solving a system of $\binom{p}{2}\binom{q}{2}$ quadratic equations in r variables.

Here we don't need to throw away point $z=0$ for which $K(z)$ is 0, since they correspond exactly to the trivial solutions of a homogeneous BLS. (Notice $z=0$ is the only point where $K(z)=0$, as $\{K_i\}_{i=1}^m$ are linearly independent. Also notice $K(z)$ can't be 0 for inhomogeneous system, as $\{K_i\}_{i=0}^m$ are linearly independent.)

For small q, p and r, our method simplifies the solvability problem and enables us to find complete solutions to bilinear systems. After all those steps, we are finally able to give some examples of solving bilinear systems, which we shall see in the next section.

However, solving a system of $\binom{p}{2}\binom{q}{2}$ quadratic equations in r variables could be awfully difficult when q, p, r are large. We will talk about how to get around this problem in the next chapter.

2.8 Examples

Here are some examples of solving incomplete bilinear systems. For simplicity we consider $p=q=2$, $m=3$, and $r=1$ case. In this case, the only 2-by-2 minor of $K(z)$ is its determinant.

Example 1 *Consider bilinear system $(\mathbb{A}, g)$, with coefficients in field $\mathbb{F}$ (for example $\mathbb{F}=\mathbb{R}$*

or $\mathbb{C}$), where

$$\mathbb{A} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \right\},$$

and

$$g = (-2, 1, 1)^T.$$

By equation (2.27):

$$K(z) = \begin{pmatrix} -2+z & 1-z \\ 1-z & z \end{pmatrix}.$$

Check the determinant:

$$\det K(z) = (-2+z)z - (1-z)^2 = -1 \neq 0.$$

Hence bilinear system $(\mathbb{A}, g)$ is not solvable.

In fact, this example works for any field $\mathbb{F}$, where "0","1" represent the naught and the unity, $2 := 1+1$, and $-1 := 0-1$. It's interesting that this example is not solvable in any field.

Example 2 *Consider another bilinear system $(\mathbb{A}', g)$ over field $\mathbb{F}$, where*

$$\mathbb{A}' = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & -2 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \right\},$$

and

$$g = (-2, 1, -2)^T.$$

By equation (2.27):

$$K(z) = \begin{pmatrix} -2+2z & -2-z \\ 1-z & z \end{pmatrix}.$$

Check the determinant:

$$\det K(z) = (-2+2z)z - (-2-z)(1-z) = z^2 - 3z + 2.$$

There are two zeros $z = 1$ or 2. Plug them back into $K(z)$, we get two rank one matrices:

$$K(z) = \begin{pmatrix} 0 & -3 \\ 0 & 1 \end{pmatrix} \text{ or } \begin{pmatrix} 2 & -4 \\ -1 & 2 \end{pmatrix}.$$

Find a particular solution of each:

$$y = (-3,1)^T, \qquad x = (0,1)^T$$
$$y = (2,-1)^T, \qquad x = (1,-2)^T.$$

Therefore the solution set of bilinear system $(\mathbb{A}, g)$ is

$$\{(x,y) : y = \lambda(-3,1)^T, x = \frac{1}{\lambda}(0,1)^T\} \bigcup \{(x,y) : y = \lambda(2,-1)^T, x = \frac{1}{\lambda}(1,-2)^T\}.$$

where $\lambda \in \mathbb{F}^*$.

Again $\mathbb{F}$ here can be any field. We shall see 1, 2, 3, ..., in the system as multiples of unity. This particular example even works for field $\mathbb{F}$ with characteristic 2, in which case $-3 = -1 = 1$, $\pm 2 = 0$, but there will still be 2 classes of equations.

Example 3 *Consider another bilinear system $(\mathbb{A}'', g)$, with coefficients in field $\mathbb{R}$, whose LHS matrices are*

$$\mathbb{A}'' = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & -2 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \right\},$$

and right hand side is

$$g = (0,1,-1)^T.$$

By equation (2.27), we have:

$$K(z) = \begin{pmatrix} 2z & -1-z \\ 1-z & z \end{pmatrix}.$$

Check the determinant:

$$\det K(z) = (2z)z - (-1-z)(1-z) = z^2+1.$$

It has no zeros in $\mathbb{R}$ *but two zeros in* $\mathbb{C}$*:* $z = i$ *or* $-i$*. Plug them back in* $K(z)$*, we get two rank one matrices:*

$$K(z) = \begin{pmatrix} 2i & -1-i \\ 1-i & i \end{pmatrix} \text{ or } \begin{pmatrix} -2i & -1+i \\ 1+i & -i \end{pmatrix}.$$

Find a particular solution of each:

$$y = (-1+i,1)^T, \qquad x = (1-i,i)^T$$
$$y = (-1-i,-1)^T, \qquad x = (1+i,-i)^T.$$

Bilinear system $(\mathbb{A},g)$ *is not solvable in* $\mathbb{R}$*. But in* $\mathbb{C}$ *the set of solutions of* $(\mathbb{A},g)$ *is*

$$\{(x,y) : y = \lambda(-1+i,1)^T, x = \frac{1}{\lambda}(1-i,i)^T\} \bigcup \{(x,y) : y = \lambda(-1-i,-1)^T, x = \frac{1}{\lambda}(1+i,-i)^T\}.$$

where $\lambda \in \mathbb{C}^*$*.*

This is an example of a bilinear system with real coefficients, but no real roots.

Chapter 3

The Rank One Completion Problem

In the previous chapter we have seen that the real difficulty in solving a bilinear system lies in finding all rank one points of an affine matrix function:

$$K(z) = K_0 + z_1 K_1 + \ldots + z_r K_r,$$

which we define as the K-function of the BLS. However, this problem could be very difficult for large bilinear systems. In this chapter we will explore different ways to solve the rank one completion problem and find cases where this problem can be completely solved.

3.1 Standard Checking

To check if a constant nonzero matrix has rank one, it suffices to check if all of its 2-by-2 minors are equal to zero. Finding all rank one points of a matrix function is almost the same.

First note when $K(z)$, as a K-function of some BLS, could be a zero matrix. Since $\{K_i\}_{i=0}^m$ are linearly independent for an inhomogeneous BLS, $K(z)$ cannot be zero in this case. For a homogeneous BLS, $K_0 = 0$. Since $\{K_i\}_{i=1}^m$ are linearly independent, $K(z) = 0$ if and only if $z = 0$. Recall $K(z) = 0$ corresponds exactly to the trivial solutions, which only happen in homogeneous systems.

For each choice of z, $K(z)$ is a constant matrix. The 2-by-2 minors of $K(z)$ are at-most-quadratic functions in the z_i's. (By *at-most-quadratic* we mean quadratic, linear, or constant) By the criterion for constant matrices, a nonzero $K(z)$ is rank one if and only if all its minors go to zero for this choice of z, in other words, z is a zero to these at-most-quadratic functions. Therefore, the rank one points of $K(z)$ coincide with the set of zeros of these at-most-quadratic function, excluding $z = 0$ if $K_0 = 0$. Since there is a one-to-one relationship between classes of nontrivial solutions to the BLS and rank one points of $K(z)$ (Theorem 3), the above discussion gives a one-to-one relationship between classes of nontrivial solutions to the BLS and solutions to a set of at-most-quadratic equations, where we exclude answer $z = 0$ if $K_0 = 0$.

We refer to the method of solving a bilinear system through solving a system of $\binom{p}{2}\binom{q}{2}$ at-most-quadratic equations as *standard checking*. By *checking* a set of minors of a matrix function we mean solving a system of at-most-quadratic equations associated with this set of minors, just as if checking whether the minors go to zero in a constant matrix. We refer to the method above as *standard checking* since we need to "check" all $\binom{p}{2}\binom{q}{2}$ minors of the K-function.

3.2 Necessity of Checking All Minors

It seems an awful amount of work to check all $\binom{p}{2}\binom{q}{2}$ minors. A natural question to ask is: is this really necessary? In the most general case, the answer is yes. In other words, checking all $\binom{p}{2}\binom{q}{2}$ minors is necessary to determine precisely all the rank one points of a K-function $K(z)$ without further restrictions on the structure of $K(z)$. This fact is demonstrated by the following theorem:

Theorem 10 *For any integers $p \geq 2$, $q \geq 2$, and any position for a 2-by-2 minor in a p-by-q matrix, there exist a K-function $K(z)$ of size p-by-q, such that checking all $\binom{p}{2}\binom{q}{2}$ minors but the one in this position will yield at least one extra solution z that render $K(z)$ having rank more than one.*

In other words, it is necessary to check all 2-by-2 minors to characterize all rank one points of an arbitrary K-function of any size.

Proof. Without loss of generality we assume $p \leq q$, as any scheme for checking minors in a p-by-q K-function can be transformed to one for a q-by-p K-function by transposition.

For $p = q = 2$, there is only one 2-by-2 minor. Without checking any minor, the set of rank one points cannot be determined. Any $K(z)$ of size 2-by-2 that can achieve a rank of two can serve as an example. The statement holds trivially.

For $p = 2$, $q = 3$, consider the following $K(z)$:

$$K(z) = \begin{pmatrix} 1 & z & -z \\ z & z & 1 \end{pmatrix}.$$

It suffices to consider only the not contiguous minor to be unchecked, as we can permute the columns of $K(z)$ to obtain the examples for the other two cases.

The two checked minors are $z - z^2 = 0$ and $z + z^2 = 0$, whose common solution is $z = 0$. Notice the result

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

is not rank one or less. This means, by not checking this specific minor, we yield an extra solution $z = 0$ that render $K(z)$ having rank more than one. It's easy to confirm that the matrix coefficients $\{K_i\}_{i=0}^{1}$ in $K(z)$ are linearly independent, hence by Theorem 9, $K(z)$ is a K-function of some BLS.

For $p \geq 2$, $q \geq 3$, we can extend the 2-by-3 example above by adding zeros in all additional entries. This creates the example for minors of 3 specific positions. Examples for minors in other positions can be obtained by permuting the columns and rows of the extended $K(z)$.

Hence, it is necessary to check all 2-by-2 minors in order to guarantee that an arbitrary $K(z)$ has rank one or zero.

(Notice the counter examples we give above are only for $r = 1$ cases. In fact, it's possible to create examples for any reasonable r. One way to do so is to extend $K(z)$ by adding not just zeros, but also $\{z_i\}_{i=2}^{r}$, each occupying an unique position.) □

3.3 When $r = 1$

In last chapter we have shown examples of solving incomplete systems where $r = 1$. In fact, we have a complete solution theory for this case, as rank one completion problems with $r = 1$ can be completely solved using standard checking.

In this case, $z = z_1$ is a scaler. The minors of $K(z)$ are at-most-quadratic functions over one variable, whose set of zeros contains at most two elements. The set of rank one points of $K(z)$ is the intersection of these sets, excluding 0 if $K_0 = 0$. Hence, in light of the derivation in the previous chapter, a complete solution of the the original BLS can be obtained.

(We may call $r = 1$ the "almost complete case", as in the complete case we have $r = 0$. However, it would be superfluous for us to use this wording.)

Now we switch our attention to cases where $r > 1$. How hard are these cases? First we point out that, even when r is small, each z_i may appear in many positions. When r is large, the at-most-quadratic functions may each have multiple variables. Using standard checking, we have to solve a system of $\binom{p}{2}\binom{q}{2}$ quadratic equations in r variables, which may be even more difficult than the problem of solving bilinear system. Therefore, the rank one completion problem is in general a very hard problem itself. We seem to have, sparing the $r = 0$ and 1 case, just transformed one very hard problem to another.

However, this is not the end of the story. We are now restricting ourselves to using standard checking to solve the rank one completion problem. We shall introduce another way of checking and a class of bilinear system that we can completely solve using this checking.

3.4 Fast Checking

Using standard checking one need to solve a system of $\binom{p}{2}\binom{q}{2}$ at-most-quadratic equations in r variables. $\binom{p}{2}\binom{q}{2}$ is a large number when one is talking about the number of equations. However, we have also shown that checking this large a number of minors is in a sense necessary without any extra restrictions on $K(z)$. Since our ambition of tackling the entire rank one completion problem with just standard checking failed, we venture to lose a little bit generality by putting a small restriction on $K(z)$, and see if we can reduce the number of minors we need to check. There are many ways to do this. In the following definition we shall introduce a rather efficient one.

Definition 11 (Center Checking) *Pick a particular entry in a matrix function $K(z)$ as the "center". In* center checking*, instead of checking all the minors of a matrix function, one merely check those that contain this center.*

In other words, instead of solving a system of $\binom{p}{2}\binom{q}{2}$ at-most-quadratic equations, one now need only to solve a system of $(p-1)(q-1)$ at-most-quadratic equations.

Notice this is an improvement in orders of magnitude. However, center checking would be of no use if it doesn't determine precisely all the rank one points of K-functions. We know from Theorem 10, that it doesn't for all K-functions. However, it works for a large group (in some sense, even majority) of K-functions. This is demonstrated in the following theorem.

Theorem 12 *If a K-function has a constant entry that is nonzero, then center checking with this entry as center determines precisely all the rank one points of this K-function.*

Proof. We need to prove that for each z, $K(z)$ has rank one if and only if all $(p-1)(q-1)$ minors in the center checking are zero. (Note since the "center" is nonzero, the possibility that $K(z)$ has rank zero is forbidden.) We already know that $K(z)$ has rank one (or zero) only if all $\binom{p}{2}\binom{q}{2}$ minors of $K(z)$ are zero. Hence the "only if" direction is clear, we need only to prove the "if" direction.

Let K_{ij}, $i=1,\ldots,p$, $j=1,\ldots,q$ be the entries of matrix K. Without loss of generality, we assume the center is K_{11}. Since the first column is nonzero, to prove that K has rank one, it suffices to prove that all columns in K are proportional to the first column. We shall prove that the j^{th} column is proportional to the first column. Again since the first column is nonzero, it suffices to prove that the p-by-2 matrix formed by the these two columns has rank one.

We may now apply the same argument on the rows. To prove this p-by-2 matrix has rank one, since the first row (K_{11}, K_{1j}) is nonzero, it suffices to prove that all rows in this p-by-2 matrix are proportional to the first row. Again since the first row is nonzero, for each row (K_{i1}, K_{ij}), it suffices to prove that the 2-by-2 matrix formed by the this row and the first row

$$\begin{pmatrix} K_{11} & K_{1j} \\ K_{i1} & K_{ij} \end{pmatrix}$$

has rank one. Notice this 2-by-2 matrix has rank one if and only if its only minor $K_{11}K_{ij} - K_{1j}K_{i1}$ is zero. However, minors $K_{11}K_{ij} - K_{1j}K_{i1}$, $i = 2,\dots,p$, $j = 2,\dots,q$ are exactly the minors of K that are checked in the center checking, and therefore, are zero by our premises.

Hence the j^{th} column is proportional to the first column, for $j = 2,\dots,q$. Therefore, matrix K has rank one. □

This additional condition is very minimal, while it necessitates checking of a much smaller set of minors. In fact, $(p-1)(q-1)$ is, in a way, the least number of minors one could check to still be able to determine if a matrix is rank one. For an arbitrary matrix, we can set the first column and the first row of a p-by-q matrix to be independent variables and the other entries dependent variables. To ensure this matrix has rank one, all other columns have to be proportion to the first one, and the proportions are predetermined by the first row. Therefore, we have at most one choice for all $(p-1)(q-1)$ entries not in the first column or the first row, which takes away $(p-1)(q-1)$ degrees of freedom from the matrix. However, we need to impose at least $(p-1)(q-1)$ equations to take away $(p-1)(q-1)$ degrees of freedom, which means the checking of at least $(p-1)(q-1)$ minors are required to determine the rank one points of a matrix function. (Note that knowing an entry is a nonzero constant doesn't take away any degree of freedom, as we are not specifying what this entry should be.)

Center checking significantly simplifies the rank one completion problem, even if it's still hard. However, we'd still like to know exactly how often do we see K-functions satisfy the condition of having a constant entry and, therefore, making center check applicable. In order to see it, we need to see what this condition for the K-function means for the original BLS. This relation is revealed in the following theorem.

Theorem 13 *The (i,j) entry of a K-function is constant if and only if E_{ij} belongs to the LHS matrices of the original BLS after Gaussian elimination.*

(Note linear operations on a bilinear system don't change its K-function, as long as the same set of LHS matrices $\{C_i\}_{i=1}^r$ are added to perform the completion. One way to understand it is by noticing that K-functions care only about the solution set, not how the BLS looks like. We will include its proof in the proof for the theorem above. Also note that the form after Gaussian elimination is a very good normal form for bilinear systems.)

Proof. First we shall prove that linear operations on a bilinear system don't change its K-function, so that the K-function we mentioned in the theorem above stay well defined.

Take bilinear system $(\mathbb{A}, g)$ and a completion $(\mathbb{B}, \tilde{g}(z))$ through adding LHS matrices $\mathbb{C}$. By formula 2.27, the K-function associated to this completion is

$$K(z) = \text{unv}((\mathscr{B}^T)^{-1} \begin{pmatrix} g \\ z \end{pmatrix}), \tag{3.1}$$

where $\mathscr{B} = (\mathscr{A}, \mathscr{C})$. Now perform a certain linear operation on bilinear system $(\mathbb{A}, g)$, and we obtain a new system $(\mathbb{A}', g')$. This operation can be represented by a left multiplication by a nonsingular m-by-m matrix M on both $\mathscr{A}^T = (\text{vec}A_1, \ldots, \text{vec}A_m)^T$ and $g = (g_1, \ldots, g_m)^T$. In specific:

$$\mathscr{A}'^T = M\mathscr{A}^T, \quad g' = Mg. \tag{3.2}$$

Now we use the same set of additional LHS matrices $\mathbb{C}$ to perform a completion and obtain new completion $(\mathbb{B}', \tilde{g}'(z))$. By formula 2.27, the K-function associated to this completion is

$$K'(z) = \text{unv}((\mathscr{B}'^T)^{-1} \begin{pmatrix} g' \\ z \end{pmatrix}), \tag{3.3}$$

where $\mathscr{B}' = (\mathscr{A}', \mathscr{C})$.

We claim $K(z) = K'(z)$.

First by formula 3.1 and 3.3

$$\mathscr{B}^T \text{vec} K(z) = \begin{pmatrix} g \\ z \end{pmatrix}, \quad \mathscr{B}'^T \text{vec} K'(z) = \begin{pmatrix} g' \\ z \end{pmatrix} \tag{3.4}$$

i.e.

$$\begin{pmatrix} \mathscr{A}^T \\ \mathscr{C}^T \end{pmatrix} \text{vec} K(z) = \begin{pmatrix} g \\ z \end{pmatrix} \tag{3.5}$$

and

$$\begin{pmatrix} \mathscr{A}'^T \\ \mathscr{C}^T \end{pmatrix} \text{vec} K'(z) = \begin{pmatrix} g' \\ z \end{pmatrix}. \tag{3.6}$$

Let's left multiply matrix

$$\begin{pmatrix} M & 0 \\ 0 & I \end{pmatrix}$$

to both side of equation 3.5, we have

$$\begin{pmatrix} M & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} \mathscr{A}^T \\ \mathscr{C}^T \end{pmatrix} \text{vec} K(z) = \begin{pmatrix} M & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} g \\ z \end{pmatrix}, \tag{3.7}$$

$$\begin{pmatrix} M\mathscr{A}^T \\ \mathscr{C}^T \end{pmatrix} \text{vec} K(z) = \begin{pmatrix} Mg \\ z \end{pmatrix}, \tag{3.8}$$

$$\begin{pmatrix} \mathscr{A}'^T \\ \mathscr{C}^T \end{pmatrix} \text{vec} K(z) = \begin{pmatrix} g' \\ z \end{pmatrix}. \tag{3.9}$$

Combine equation 3.6 and 3.9, we have:

$$\text{vec} K(z) = \begin{pmatrix} \mathscr{A}'^T \\ \mathscr{C}^T \end{pmatrix}^{-1} \begin{pmatrix} g' \\ z \end{pmatrix} = \text{vec} K'(z). \tag{3.10}$$

Hence $K(z) = K'(z)$.

Now we shall prove the theorem itself.

Take $(\mathbb{A}, g)$ and perform a Gaussian elimination. We get a new system $(\mathbb{A}', g')$, for which $\mathscr{A}'^T$ is in reduced row echelon form. As proven above, if we use the same additional LHS matrices $\mathbb{C}$ to complete both systems, their K-functions will be the same. Hence

$$\mathrm{vec}K(z) = \begin{pmatrix} \mathscr{A}'^T \\ \mathscr{C}^T \end{pmatrix}^{-1} \begin{pmatrix} g' \\ z \end{pmatrix}. \tag{3.11}$$

We denote $\mathbb{A}' = \{A'_1, \ldots, A'_m\}$. Note that following statement is equivalent to the statement of the theorem:

The i^{th} entry of $\mathrm{vec}K(z)$ is a constant if and only if e_i belongs to the set $\{\mathrm{vec}A'_1, \ldots, \mathrm{vec}A'_m\}$.

By formula 3.11, the i^{th} entry of $\mathrm{vec}K(z)$ is constant if and only if the i^{th} entry of vector

$$\begin{pmatrix} \mathscr{A}'^T \\ \mathscr{C}^T \end{pmatrix}^{-1} \begin{pmatrix} g' \\ z \end{pmatrix}$$

is constant, which happens if and only if the rear r entries in the i^{th} row of matrix

$$\begin{pmatrix} \mathscr{A}'^T \\ \mathscr{C}^T \end{pmatrix}^{-1}$$

are zeros. Now what are the the rear r entries in the i^{th} row of the above matrix in terms of $\mathscr{A}'^T$ and $\mathscr{C}^T$? To see their relation, we recall a common way to compute inverses. *i.e.* if M is a nonsingular square matrix, perform a row reduction on matrix (M, I) and transform it to $(I, \tilde{M})$, then $\tilde{M} = M^{-1}$. For convenience, in an intermediate state (M_1, M_2) during the row reduction, we refer to M_1 as the *left matrix* and M_2 as the *right matrix*. We have relation

$$\left(\begin{pmatrix} \mathscr{A}'^T \\ \mathscr{C}^T \end{pmatrix} \quad \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \right) \sim \left(\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \quad \begin{pmatrix} \mathscr{A}'^T \\ \mathscr{C}^T \end{pmatrix}^{-1} \right),$$

where "$\sim$" represent row reduction. The transformation from left to right can be done by performing Gaussian elimination on the left matrix. Now we are ready to prove the equivalent statement of the theorem.

On one hand, if e_i belongs to the set $\{\text{vec}A'_1, \ldots, \text{vec}A'_m\}$. then e_i^T is among the first m rows of the left matrix, which correspond to a row of the same position in right matrix (let's call it r_i^T, which equals to e_j^T for some j in $\{1, \ldots, m\}$.), which has all zeros for its rear r entries. Notice after Gaussian elimination, row (e_i^T, r_i^T) is unaltered besides being moved to the i^{th} row, as the i^{th} row of I is exactly e_i^T. Hence the i^{th} row of matrix

$$\begin{pmatrix} \mathscr{A}'^T \\ \mathscr{C}^T \end{pmatrix}^{-1}$$

is r_i^T, which has zeros for all of its rear r entries. By formula 3.11, i^{th} entry of $\text{vec}K(z)$ is a constant.

One the other hand, if the i^{th} entry of $\text{vec}K(z)$ is a constant, the i^{th} row of matrix

$$\begin{pmatrix} \mathscr{A}'^T \\ \mathscr{C}^T \end{pmatrix}^{-1},$$

which define as r_i^T, has zeros for all of its rear r entries. Therefore, r_i^T is a linear combination of e_i^T for $i = 1, \ldots, m$, which are the first m rows of the beginning right matrix I. Notice both left and right matrix have gone through the same linear operation, therefore e_i^T, the i^{th} row of resulting left matrix I, must be a linear combination of $\text{vec}{A'_1}^T, \ldots, \text{vec}{A'_m}^T$, first m rows of the beginning left matrix I. However $\text{vec}{A'_1}^T, \ldots, \text{vec}{A'_m}^T$ is already in their Gaussian elimination form. The only way for e_i^T to be a linear combination of $\text{vec}{A'_1}^T, \ldots, \text{vec}{A'_m}^T$ is for e_i^T to be exactly one of them, which means in other words, e_i belongs to the set $\{\text{vec}A'_1, \ldots, \text{vec}A'_m\}$. $\square$

In a probabilistic sense, seeing a bilinear system with some E_{ij} in its LHS matrices after Gaussian elimination is not often. In fact, the probability is 0. However, in real life situations, we are more likely to encounter LHS matrices with rational or even integral entires, making the appearance of such situations much more probable.

3.5 When a column or a row of $K(z)$ is constant

In light of Theorem 12, we immediately get complete solutions of bilinear systems whose K-function have a constant and nonzero column or row. Note only one entry of this column or row needs to be nonzero. Pick any nonzero entry in this column or row as center. Notice now all the minors needed for center checking are now *at-most-linear* equations instead of at-most-quadratic equations. By *at-most-linear equations* we mean linear equations or equations of type "$* = 0$", where $*$ is a constant. It's easier to see in an example:

Let $K(z)$ be a 3-by-3 K-function whose first column is constant. Since $K(z)$ is an affine matrix function, it can be written as

$$K(z) = \begin{pmatrix} c_1 & a_1(z) & a_4(z) \\ c_2 & a_2(z) & a_5(z) \\ c_3 & a_3(z) & a_6(z) \end{pmatrix}, \tag{3.12}$$

where c_i, $i = 1, 2, 3$ are constants, and $a_i(z)$, $i = 1, \ldots, 6$ are affine functions of z. Without loss of generality we assume c_1 is nonzero. Taking c_1 as the center, the minors checked in center checking are:

$$\begin{aligned} c_1 a_2(z) - c_2 a_1(z) &= 0 \\ c_1 a_3(z) - c_3 a_1(z) &= 0 \\ c_1 a_5(z) - c_2 a_4(z) &= 0 \\ c_1 a_6(z) - c_3 a_4(z) &= 0, \end{aligned} \tag{3.13}$$

which forms a linear system of equations with some "$* = 0$" equations mixed inside, as $a_i(z)$'s are affine functions. Equations of type "$* = 0$" can be thrown away if all $*$'s are zeros, and render the set of rank one point of $K(z)$ empty if one of the $*$ is nonzero. In either case, we need to solve no more than a linear system, which we know how to. Hence a complete solution of bilinear systems whose K-function has a constant column or row is obtained.

If there is a constant row or column in one K-function of a BLS, we immediately know how to obtain a complete solution to this BLS. However, there are more than one K-functions associated with each bilinear system. If one of them doesn't have a constant row or column, how do we know if none of the others have a constant row or column? In fact we know this fact from the following theorem:

Theorem 14 *Let $(\mathbb{A}, g)$ be a bilinear system and $K(z)$ be a K-function of $(\mathbb{A}, g)$. If $K(z)$ doesn't have a constant column or row, neither does any other K-functions of $(\mathbb{A}, g)$.*

Proof. This result follows immediately from Theorem 13, according to which whether any entry of a K-function of a BLS is constant depends entirely on the set of LHS matrices $\mathbb{A}$, and hence independent of which K-function of the BLS we are talking about. In specific, if the i^{th} row (or j^{th} column) of a K-function $K(z)$ is constant, then set $\{E_{ij}\}_{j=1}^{q}$ (or $\{E_{ij}\}_{i=1}^{p}$) is a subset of the LHS matrices of the BLS after Gaussian elimination. Therefore, the i^{th} row (or j^{th} column) of any other K-function $K'(z)$ is also constant. $\square$

Now, what kind of bilinear systems have K-functions that have a constant column or row? The answer is obvious in light of Theorem 13. For example, by Theorem 13, the i^{th} row of $K(z)$ is constant if and only if E_{ij} for $j = 1, \ldots, q$ belong to the LHS matrices of the original BLS after Gaussian elimination.

As a side note, recall from section 2.2 that we can perform equivalence operation on a bilinear system, that is, changing the basis of our unknown (x, y). This means our method above solves more bilinear systems than it appears. In specific, if we use new variables

$$\tilde{x} = Qx \quad \tilde{y} = Py,$$

we shall have a new K-function

$$\tilde{K}(z) = \tilde{y}\tilde{x}^T = Pyx^T Q^T = PK(z)Q^T.$$

That is saying, our method obtains a complete solution to the original bilinear system if matrix function $PK(z)Q^T$ has a constant column or row for some nonsingular Q and P.

Chapter 4

Solvability of Bilinear Systems for all Right Hand Sides

Here we introduce a new type of problem that is related to solving bilinear system: for want kind of LHS matrices $\mathbb{A} = \{A_1, \ldots, A_m\}$ is bilinear system $(\mathbb{A}, g)$ solvable for all right hand sides g's? We refer to such sets of LHS matrices *always-solvable* ones. As we shall see, this can only happen for certain value of m's.

4.1 When $m \leq 2$

It is clear that when $m = 1$, every bilinear system is solvable because $A_1 \neq 0$ due to our linear independence hypothesis. Interestingly this remains true for $m = 2$. (and we know already that it is not so for $m = 3$, as we saw in Example 1)

Theorem 15 *Under the linear independence hypothesis, every bilinear system with $m \leq 2$ is solvable. (In other words, any set of LHS matrices with $m \leq 2$ is always-solvable.)*

Proof. Let A_1 and A_2 be linearly dependent matrices. We want to show that the bilinear system

$$y^T A_1 x = g_1, y^T A_2 x = g_2 \tag{4.1}$$

is solvable for all right hand sides (g_1, g_2). Since homogeneous systems always have trivial solutions, we assume g_1 and g_2 are not both equal to zero. As pointed out in by Cohen and Tomasi [2], we may apply a linear operation on the system that takes $g = (g_1, g_2)^T$ to $(1,0)^T$. In this process the linear independence of matrices A_1 and A_2 is preserved. Hence, it suffices to consider the system

$$y^T A_1 x = 1, y^T A_2 x = 0. \tag{4.2}$$

If there exist an x such that vector $A_1 x$ does not lies in the span of vector $A_2 x$, we may find a y normal to $A_2 x$ but not to $A_1 x$, and obtain a solution by normalizing y according to $y^T A_1 x = 1$. Hence, system (4.2) is solvable unless $A_1 x$ lies in the span of $A_2 x$ for all x. Let's assume so and derive a contradiction.

If $A_1 x$ lies in the span of $A_2 x$ for all x, then $A_2 x = 0$ implies $A_1 x = 0$, thus the kernel of A_2 is contained in the kernel of A_1.

Let $b_1, \ldots, b_s$ be a basis of the kernel of A_2, and let us complete this basis with $b_{s+1}, \ldots, b_q$ to a basis for $\mathbb{R}^q$. By our assumption,

$$A_1 b_k = \alpha_k A_2 b_k, \quad k = s+1, \ldots, q.$$

If the α_k's are all equal, then $A_1 = \alpha_{s+1} A_2$, which contradict our linear independence assumption. If not, without loss of generality we assume $\alpha_{s+1} \neq \alpha_{s+2}$. There exist an α such that,

$$A_1(b_{s+1} + b_{s+2}) = \alpha A_2(b_{s+1} + b_{s+2}).$$

On the other hand,

$$A_1(b_{s+1} + b_{s+2}) = \alpha_{s+1} A_2 b_{s+1} + \alpha_{s+2} A_2 b_{s+2}.$$

Therefore,

$$(\alpha_{s+1}-\alpha)A_2b_{s+1}+(\alpha_{s+2}-\alpha)A_2b_{s+2}=0,$$

$$A_2[(\alpha_{s+1}-\alpha)b_{s+1}+(\alpha_{s+2}-\alpha)b_{s+2}]=0.$$

Hence $(\alpha_{s+1}-\alpha)b_{s+1}+(\alpha_{s+2}-\alpha)b_{s+2}$ is in the kernel of A_2. The linear independence of $b_1,\ldots,b_{s+2}$ implies that $(\alpha_{s+1}-\alpha)=(\alpha_{s+2}-\alpha)=0$, which contradict our assumption $\alpha_{s+1}\neq\alpha_{s+2}$. □

4.2 When $m \geq p+q$

As suggested by Theorem 15 and Example 1, not all sets of LHS matrices $\mathbb{A}$ are always-solvable when m is greater than 2. However, will there be an always-solvable set of LHS matrices $\mathbb{A}$ for all $m>2$? For the fields we worry about, the answer is no. For $m \geq p+q$, an always-solvable set of LHS matrices doesn't exist.

Notice if $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$, the left hand side bilinear forms define a smooth map from $\mathbb{F}^{p+q}$ to $\mathbb{F}^m$. When m is much larger than $p+q$, Sard's Lemma forbid this map to be surjective. When $\mathbb{F}$ is a finite field, similar dimensionality issues are involved. These points are made rigorous in the following theorem, which gives an upper bound to the largest m that allow an always-solvable situation.

Theorem 16 *Let $\mathbb{F}$ be either $\mathbb{R}$, $\mathbb{C}$ or a finite field, and $\mathbb{A}=\{A_1,\ldots,A_m\}\subset M_{p,q}(\mathbb{F})$ be a set of LHS matrices. If $\mathbb{A}$ is always-solvable, then $m \leq p+q-1$.*

Proof. Assume $m \geq p+q$. Define a degree-two polynomial map on $\mathbb{F}^q\times\mathbb{F}^p$ to $\mathbb{F}^m$ by:

$$\begin{aligned} F:\ \mathbb{F}^q\times\mathbb{F}^p &\longrightarrow \mathbb{F}^m \\ (x,y) &\longmapsto (y^TA_1x,\ldots,y^TA_mx) \end{aligned}$$

It suffices to prove that the image of F is strictly contained in $\mathbb{F}^m$. We do so respectively for the cases $\mathbb{F}=\mathbb{R}$, $\mathbb{F}=\mathbb{C}$ and $|\mathbb{F}|<\infty$.

Case 1 **($\mathbb{F}=\mathbb{R}$)** *Function F is $C^\infty(\mathbb{R}^{p+q})$ smooth. By Sard's Lemma, the subset of $\mathbb{R}^{p+q}$ where* $\operatorname{rank}(dF)<m$ *has an image of measure 0 in $\mathbb{R}^m$. Therefore, it suffices to show* $\operatorname{rank}(dF)<m$ *for all pairs (x,y).*

Denote $y^TA_ix = y_jA_{ijk}x_k$ (here we start using Einstein's notation to avoid writing $\sum_{j=1}^{p}\sum_{k=1}^{q}$ repetitively). It follows that

$$dF=\begin{pmatrix} y_jA_{1j1} & \dots & y_jA_{1jq} & A_{11k}x_k & \dots & A_{1pk}x_k\\ y_jA_{2j1} & \dots & y_jA_{2jq} & A_{21k}x_k & \dots & A_{2pk}x_k\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ y_jA_{mj1} & \dots & y_jA_{mjq} & A_{m1k}x_k & \dots & A_{mpk}x_k \end{pmatrix}=\begin{pmatrix} y^TA_1 & x^TA_1^T\\ \vdots & \vdots\\ y^TA_m & x^TA_m^T\end{pmatrix}.$$

Since

$$dF\begin{pmatrix}x\\-y\end{pmatrix}=\begin{pmatrix} y^TA_1 & x^TA_1^T\\ \vdots & \vdots\\ y^TA_m & x^TA_m^T\end{pmatrix}\begin{pmatrix}x\\-y\end{pmatrix}=\begin{pmatrix} y^TA_1x-x^TA_1^Ty\\ \vdots\\ y^TA_mx-x^TA_m^Ty\end{pmatrix}=\begin{pmatrix}0\\ \vdots\\ 0\end{pmatrix},$$

the columns of matrix dF are linearly dependent. Therefore, $\operatorname{rank}(dF)<p+q\leq m$ *holds for any $(x,y)\neq 0$. For the $(x,y)=(0,0)$ case, $dF=(0)$ and* $\operatorname{rank}(dF)=0$. *Therefore, $F(\mathbb{F}^q\times\mathbb{F}^p)\subsetneq\mathbb{F}^m$.*

Case 2 **($\mathbb{F}=\mathbb{C}$)** *Function F can be viewed as a function from $\mathbb{R}^{2p+2q}$ to $\mathbb{R}^{2m}$ (denote as $\widetilde{F}$):*

$$\begin{aligned}\widetilde{F}:\quad & \mathbb{R}^{2p+2q} && \longrightarrow \mathbb{R}^{2m}\\ & (\operatorname{Re}(x),\operatorname{Im}(x),\operatorname{Re}(y),\operatorname{Im}(y)) && \longmapsto (\operatorname{Re}(F(x,y)),\operatorname{Im}(F(y,x)))\end{aligned}$$

Since $\widetilde{F}$ is still a polynomial, it is $C^\infty(\mathbb{R}^{2p+2q})$ smooth. By Sard's Lemma, the subset of $\mathbb{R}^{2p+2q}$ where $\mathrm{rank}(dF) < 2m$ has an image of measure 0 in $\mathbb{R}^{2m}$. Therefore, it suffices to show that $\mathrm{rank}(dF) < 2m$ for all pairs (x,y). The proof is as follows:

Denote $a := \mathrm{Re}(x), b := \mathrm{Im}(x), c := \mathrm{Re}(y), d := \mathrm{Im}(y), B_i := \mathrm{Re}(A_i)$, and $C_i := \mathrm{Im}(A_i)$. By definition

$$\begin{aligned}
\widetilde{F}(a,b,c,d) &= (\mathrm{Re}(F(x,y)), \mathrm{Im}(F(y,x))) \\
&= (\mathrm{Re}((c+Id)(B_i+IC_i)(a+Ib), \ldots, \mathrm{Im}((c+Id)(B_i+IC_i)(a+Ib), \ldots) \\
&= (c^T B_i a - d^T C_i a - c^T C_i b - d^T B_i b, \ldots, -d^T C_i b + d^T B_i a + c^T C_i a + c^T B_i b, \ldots).
\end{aligned}$$

Therefore,

$$\begin{aligned}
d\widetilde{F} &= \begin{pmatrix} \frac{\partial \mathrm{Re}(F)}{\partial a} & \frac{\partial \mathrm{Re}(F)}{\partial b} & \frac{\partial \mathrm{Re}(F)}{\partial c} & \frac{\partial \mathrm{Re}(F)}{\partial d} \\ \frac{\partial \mathrm{Im}(F)}{\partial a} & \frac{\partial \mathrm{Im}(F)}{\partial b} & \frac{\partial \mathrm{Im}(F)}{\partial c} & \frac{\partial \mathrm{Im}(F)}{\partial d} \end{pmatrix} \\
&= \begin{pmatrix} c^T B_1 - d^T C_1 & a^T B_1^T - b^T C_1^T & -c^T C_1 - d^T B_1 & -a^T C_1^T - b^T B_1^T \\ \vdots & \vdots & \vdots & \vdots \\ c^T B_1 - d^T C_m & a^T B_m^T - b^T C_m^T & -c^T C_m - d^T B_m & -a^T C_m^T - b^T B_m^T \\ d^T B_1 + c^T C_1 & a^T C_1^T + b^T B_1^T & -d^T C_1 + c^T B_1 & -b^T C_1^T + a^T B_1^T \\ \vdots & \vdots & \vdots & \vdots \\ d^T B_m + c^T C_m & a^T C_m^T + b^T B_m^T & -d^T C_m + c^T B_m & -b^T C_m^T + a^T B_m^T \end{pmatrix}.
\end{aligned}$$

Since

$$d\widetilde{F} \begin{pmatrix} a \\ -c \\ b \\ -d \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

the columns of matrix dF are linearly dependent. Therefore, for any $(a,b,c,d) \neq 0$, $\mathrm{rank}(d\widetilde{F}) < 2p+2q \leq 2m$. For $(a,b,c,d) = (0,0,0,0)$ case, $d\widetilde{F} = (0)$, $\mathrm{rank}(d\widetilde{F}) = 0$. Hence $F(\mathbb{F}^q \times \mathbb{F}^p) \subsetneq \mathbb{F}^m$.

Case 3 ($|\mathbb{F}| = N < \infty$) *It suffices to count the number of elements in $F(\mathbb{F}^q \times \mathbb{F}^p)$ and $\mathbb{F}^m$. Define relation $(cx, y) \sim (x, cy)$ on $\mathbb{F}^q \backslash \{0\} \times \mathbb{F}^p \backslash \{0\}$. Notice $F(cx, y) = F(x, cy)$ $(c \in \mathbb{F}^*)$, the quotient map:*

$$\begin{aligned} \hat{F}: \quad \mathbb{F}^q \backslash \{0\} \times \mathbb{F}^p \backslash \{0\} / \sim \quad &\longrightarrow \mathbb{F}^m \\ [(x,y)] \qquad\qquad &\longmapsto F(x,y) \end{aligned}$$

is well defined. Since

$$F(\{0\} \times \mathbb{F}^p \cup \mathbb{F}^q \times \{0\}) = \{0\},$$

we have

$$\begin{aligned} |F(\mathbb{F}^q \times \mathbb{F}^p)| &= |F(\mathbb{F}^q \backslash \{0\} \times \mathbb{F}^p \backslash \{0\})| + 1 \\ &\le \frac{(N^q - 1)(N^p - 1)}{N - 1} + 1 \\ &\le (N^q - 1)(N^p - 1) + 1 = N^{p+q} - N^p - N^q + 2 \\ &< N^{p+q} \le N^m = |\mathbb{F}^m| \end{aligned}$$

Hence $F(\mathbb{F}^q \times \mathbb{F}^p) \subsetneq \mathbb{F}^m$. (We assumed $|\mathbb{F}| > 1$, which is reasonable.)

□

4.3 When $3 \le m \le p+q-1$

The only range that we have left behind is $3 \le m \le p+q-1$. For any m in this range, examples of both always-solvable and not always-solvable sets of LHS matrices exist. This implies the distinction lies within the structure of the LHS matrices. We will analyze the structural origins of always solvability and present several sufficient conditions.

4.3.1 Examples of always-solvable Cases and Not always-solvable Cases

We can easily construct examples of not always-solvable cases for $m > 2$ by generalizing Example 1. In fact, this unsolvable example can be extended to one of any size (p,q,m), such that $p+q > m > 2$ and $p,q \geq 2$.

Example 4 *First, we extend the three 2-by-2 LHS matrices to p-by-q ones by adding zeros. Since $p+q > m$, it's always possible to add $m-3$ additional matrices so that the set of LHS matrices stay linearly independent. Notice this extended system is not solvable if the first three entries of the new right hand side g coincide with the right hand side constants in Example 1. For such g's, the first three equations in the new BLS coincide with the three equations in Example 1, which have no solutions. Hence the whole system doesn't have a solution either.*

Now we will construct examples of always-solvable bilinear systems.

Here is a way to construct such examples trivially. Take $p = 1$ (we can do the same for $q = 1$), The A_i's in bilinear system

$$yA_ix = g_i, \quad ,i = 1,\ldots,m$$

becomes row vectors. Let

$$A = \begin{pmatrix} A_1 \\ \vdots \\ A_m \end{pmatrix},$$

and we can write the BLS in form

$$yAx = g,$$

where $A \in M_{m,q}(\mathbb{F})$ and $m \leq q$ ($m \leq p+q-1$). BLS of this type is always-solvable, as we can take $y = 1$ and x a solution to linear system

$$Ax = g.$$

Examples of smaller m's can be obtained by taking away a proper number of equations out of the system.

We care more about nontrivial examples where $p,q > 1$. In such cases, it is not difficult to construct always-solvable sets of LHS matrices whenever $m \le p+q-2$. The following is an example of case $m = p+q-2$.

Example 5 *Consider a bilinear system defined by LHS matrices $A_i = E_{1i}$, $i = 1,\dots,q-1$ and $A_{q+i-2} = E_{in}$, $i = 2,\dots,p$ and an arbitray g. If we use the other E_{ij} matrices to perform a completion, we obtain K-function:*

$$K(z) = \begin{pmatrix} g_1 & g_2 & \cdots & g_{q-1} & z_1 \\ z_2 & z_3 & \cdots & z_q & g_q \\ \vdots & & \ddots & \vdots & \\ z_{r-q+2} & & \cdots & z_r & g_{p+q-2} \end{pmatrix},$$

which may be completed to a rank one matrix, whatever the g.

Again, examples of $m < p+q-2$ cases can be obtained by taking away some equations from the example above.

When $p,q > 1$ and $m = p+q-1$, we may construct matrices A_i for which the bilinear system is solvable for "almost" all g's with few exceptions.

Example 6 *Consider a bilinear system defined by LHS matrices $E_{11}, E_{12}, \dots, E_{1q}$ and matrices $E_{2q}, E_{3q}, \dots, E_{pq}$ and an arbitrary g. If we use the other E_{ij} matrices to perform a completion, we obtain K-function:*

$$K(z) = \begin{pmatrix} g_1 & g_2 & \cdots & g_{q-1} & g_q \\ z_1 & z_2 & \cdots & z_{q-1} & g_{q+1} \\ \vdots & & \ddots & \vdots & \\ z_{r-q+2} & & \cdots & z_r & g_{p+q-1} \end{pmatrix}.$$

This bilinear system is solvable whenever $g_q \neq 0$, but has no solution when $g_q = 0$ while $g_i \neq 0$ for some $i < q$ and some $i > q$.

In the latter case, without loss of generality, we assume $g_1 \neq 0$ and $g_{q+1} \neq 0$. Notice there is no way to make minor

$$\begin{pmatrix} g_1 & g_q \\ z_1 & g_{q+1} \end{pmatrix}$$

zero.

We do not know of an always-solvable set of LHS matrices when $p,q > 1$ and $m = p+q-1$. We conjecture that they do not exist.

In general, then, whether a bilinear system is always-solvable for $m \leq p+q-1$ depends upon the data $A_1, \ldots, A_m$. Even when p and q are large and $m = 3$, a bilinear system may not be always-solvable for "local reasons", as demonstrated by Example 4.

4.3.2 First Sufficient Condition for Always Solvability

How, then, can we determine always solvability when m is in range $3 \leq m \leq p+q-1$? Here we provide two sufficient conditions.

First note that a bilinear system becomes a linear system when enough of the variables are taken to have particular values. If the resulting linear system has m linearly independent equations, then it is solvable for all right hand sides. The solution of the linear system for each right hand side also provides a solution to the corresponding bilinear system with that right hand side, thus rendering the original LHS matrices always-solvable.

Example 7 *Take this bilinear system with an arbitrary right hand side:*

$$x_1y_3 + x_2y_2 = g_1$$
$$x_2y_1 + x_2y_3 = g_2$$
$$x_2y_2 = g_3$$

If we take $x_2 = 1$ *and* $y_3 = 1$, *we get a linear system*

$$x_1 = g_1 - 1$$
$$y_1 = g_2 - 1$$
$$y_2 = g_3,$$

whose equations are linearly independent, and hence has a solution for all right hand sides. This gives a solution to the bilinear system:

$$x = \begin{pmatrix} g_1 - 1 \\ 1 \end{pmatrix} \text{ and } y = \begin{pmatrix} g_2 - 1 \\ g_3 \\ 1 \end{pmatrix},$$

which means this set of LHS matrices is always-solvable.

Note that if the m equations of the obtained linear system are linearly dependent, there always exists a right hand side g that render the linear system not solvable, and thus preclude us from making a conclusion to the associated bilinear system.

As a side note, if precisely all the variables in either vector x or y are specified, to ensure that the resulting linear system have linearly independent equation, it suffices to find a x or y such that X or Y has full rank. Of course, for this to happen we must have $m \leq p$ (q). Here X and Y are as defined in section 2.1.

4.3.3 Second Sufficient Condition for Always Solvability

Now we present a way of to determine always solvability for sparse LHS matrices.

This method is based on the arrangements of the collective support of the matrices A_i. By support of a matrix we mean the pattern of the nonzero entries in this matrix. By collective support of matrices we mean the union of their supports. For example, matrix

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}$$

has support

$$\begin{pmatrix} * & * & 0 \\ * & * & 0 \end{pmatrix},$$

and matrix

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

has support

$$\begin{pmatrix} * & 0 & * \\ 0 & * & 0 \end{pmatrix}.$$

Their collective support is

$$\begin{pmatrix} * & * & * \\ * & * & 0 \end{pmatrix}.$$

Let $\mathscr{P}$ be the collective support LHS matrices A_i's. We say that the bilinear system satisfies the 3–*corner property (TCP)* if $\mathscr{P}$ has no 2-by-2 subpattern with 3 or more entries nonzero, like

$$\begin{pmatrix} * & * \\ * & 0 \end{pmatrix}$$

or

$$\begin{pmatrix} * & * \\ 0 & * \end{pmatrix}.$$

i.e. $\mathscr{P}$ is a pattern in which the nonzero entries in a row do not intersect nonzero entries in a column unless the intersecting entry is unique in a least one of the row or the column. Note that this happens if and only if, $\mathscr{P}$ has form (all blank entries are zero)

$$\begin{pmatrix}
\cdots & 0 & & & & & & \\
 & * & & & & & & \\
 & \vdots & & & & & & \\
 & * & & & & & & \\
 & 0 & *\cdots* & & & & & \\
 & & & *\cdots* & 0 & & & \\
 & & & & * & & & \\
 & & & & \vdots & & & \\
 & & & & * & & & \\
 & & & & 0 & *\cdots* & 0 & \\
 & & & & & & 0 & \\
 & & & & & & & \ddots \\
 & & & & & & & & 0
\end{pmatrix},$$

up to permutation equivalence.

This statement can be proven through the following algorithm: permute any nonzero entry to the top left. Now it must be the only nonzero entry in the first row (or column) due to TCP. Now permute the columns so that the nonzero entries in the first row are contiguous, hence we have obtained the following pattern:

$$\begin{pmatrix} *\cdots* & 0 \\ & ? \end{pmatrix},$$

where "?" represent the block that contain the rest of the nonzero entries. Now repeat the algorithm for block "?", and we shall obtain the pattern above.

We call such forms *a deleted echelon forms (DEF)*, which are by definition, forms of type $diag\{L_1, \ldots, L_s, 0, \ldots, 0\}$ where L_i's are blocks of type

$$\begin{pmatrix} * & \cdots & * \end{pmatrix}$$

or

$$\begin{pmatrix} * \\ \vdots \\ * \end{pmatrix}.$$

We have following sufficient condition for always solvability:

Theorem 17 *A set of LHS matrices is always-solvable if its collective support satisfies TCP (or satisfies TCP after an equivalent transformation $A_i \to PA_iQ$, $i = 1, \ldots, m$), and that the number of nonzero entries in the support are equal to m.*

Proof. It suffices to proof the theorem for LHS matrices with collective support of form DEF.

Take an arbitray right hand side g. We have a bilinear system $(\mathbb{A}, g)$. Since the LHS matrices $\mathbb{A}$ has supports in a m dimensional subspace of $M_{p,q}(\mathbb{F})$, by their linear independence, $\mathbb{A}$ becomes E_{ij} type matrices after Gaussian elimination on the equations. Notice this set of m E_{ij}'s type matrices still have the same support. Each of them must correspond to a nonzero position in this support.

Now we may perform a completion using the rest of the E_{ij} type matrices. Since all the LHS matrices in the completion are already in forms E_{ij}, the associated K-function preserve the form where constants lies on the collective support of $\mathbb{A}$ and z_i's lies on the where the zero entries are on the pattern of the support. Hence we have obtained a K-function of "deleted echelon form" up to permutation of columns and rows, which looks like, for example:

$$\begin{pmatrix}
\cdots & z_{i_1} & & & & & & & \\
 & * & & & & & & & \\
 & \vdots & & & & & & & \\
 & * & & & & & & & \\
 & z_{i_2} & *\cdots* & & & & & & \\
 & & & *\cdots* & z_{i_3} & & & & \\
 & & & & * & & & & \\
 & & & & \vdots & & & & \\
 & & & & * & & & & \\
 & & & & z_{i_4} & *\cdots* & z_{i_5} & & \\
 & & & & & & z_{i_6} & & \\
 & & & & & & & \ddots & \\
 & & & & & & & & z_{i_7}
\end{pmatrix},$$

where all the $*$ stands for constants, and all the blank entries are other z_i's.

We claim that this type of K-function always have a rank one completion. We shall prove it by induction:

It suffices to prove for fixed rank one matrix A and row $\mu := (\mu_1, \ldots, \mu_t)$, matrix $\begin{pmatrix} A & B \\ \gamma^T & \mu \end{pmatrix}$ can be completed to a rank one matrix by a proper choice of γ and B. The case where μ is a column follows from transposition.

By basic matrix theory, rank one matrix A can be written as: $(\lambda_1 v, \ldots, \lambda_s v)$. Therefore, the matrix $\begin{pmatrix} A & B \\ \gamma^T & \mu_1, \ldots, \mu_t \end{pmatrix}$ can be complete to $\begin{pmatrix} \lambda_1 v, \ldots, \lambda_s v & \mu_1 v, \ldots, \mu_t v \\ \lambda_1, \ldots, \lambda_s & \mu_1, \ldots, \mu_t \end{pmatrix}$, which has rank at most one. $\square$

Note if $p, q > 1$,the number of nonzero entries in a DEF (or a pattern that satisfies TCP) is at most $p+q-2$. It could be $p+q-1$, however, if $p = 1$ or $q = 1$. But that would the only case that this value is attained.

4.3.4 Exceptions

However, conditions we mentioned above are not complete. Note that a set of LHS matrices may be always-solvable without satisfying any of the above sufficient conditions. Here is an example:

Example 8 *Take a the bilinear system of equations defined by LHS matrices*

$$A_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, A_2 = \begin{pmatrix} -1 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix} \text{ and } A_3 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}$$

and an arbitrary right hand side $g = (g_1, g_2, g_3)^T$. We solve this BLS and obtain the following K-function:

$$K = \begin{pmatrix} z_2 & z_3 & -g_2 - z_2 \\ -g_3 + z_3 & z_1 & g_1 - z_3 \end{pmatrix}.$$

Standard checking gives a system of three quadratic equations:

$$z_1 z_2 - z_3^2 + z_3 g_3 = 0$$

$$z_1 z_2 - z_3^2 + z_1 g_2 + z_3 g_1 = 0$$

$$z_2(g_1 - g_3) + z_3 g_2 - g_2 g_3 = 0.$$

If $g_1 \neq g_3$ we have solution: $z_1 = 0, z_3 = 0, z_2 = \frac{g_2 g_3}{g_1 - g_3}$. If $g_1 = g_3$, we have solution $z_3 = g_1$, $z_1 = 0$ and arbitrary z_2. This shows that bilinear system of equations defined by matrices A_1, A_2, A_3 is solvable for any right hand side g.

Note that this example doesn't satisfy either of the sufficient conditions.

In this case the TCP does not hold, as the support is

$$\begin{pmatrix} * & * & * \\ * & 0 & * \end{pmatrix}.$$

The rank of X is always less than or equal to 2 since $q = 2$, and

$$Y = \begin{pmatrix} 0 & y_1 & y_2 \\ -y_1 & 0 & -y_1 \\ -y_2 & y_1 & 0 \end{pmatrix}$$

is singular for any choice of y. Write the bilinear forms as a linear combination of the dyads $x_i y_j$'s:

$$x_1 y_2 + x_2 y_3$$

$$x_1 y_1 + x_1 y_3$$

$$x_1 y_2 - x_2 y_1,$$

by enumeration, we shall see that no specification of just one variable gives a linear system and no specialization of two variables gives an invertible one. However, this set of LHS matrices is always-solvable regardless.

Chapter 5

Applications

In this chapter we shall discuss some examples of how the solution theory for bilinear systems could be applied.

5.1 Commutativity of Patterns

As we mentioned in the introduction, the study of commutativity of patterns has motivated the study of bilinear systems.

By a pattern $\mathscr{P}$ we mean an array of $*$'s and 0's in which a $*$ indicates a nonzero entry. We say a real matrix $A = (a_{i,j})$ belongs to pattern $\mathscr{P}$ if its dimensions agree with those of $\mathscr{P}$, and $a_{i,j} \neq 0$ if and only if the (i, j) entry of $\mathscr{P}$ is a $*$. We say that two n-by-n patterns $\mathscr{P}$ and $\mathscr{Q}$ commute if there exist matrices $A \in \mathscr{P}$, $B \in \mathscr{Q}$ that commute, *i.e.* , $AB = BA$.

Notice the problem of commutativity of patterns can be reduced to the problem of solving a certain kind of homogeneous bilinear systems. One can replace the $*$'s in $\mathscr{P}$ by a variable vector x and the $*$'s in $\mathscr{Q}$ by a variable vector y. The equation $\mathscr{P}\mathscr{Q} - \mathscr{Q}\mathscr{P} = 0$ gives a special homogeneous BLS in as many variables as the total number of nonzero entries of $\mathscr{P}$ and $\mathscr{Q}$. Here, a totally nonzero solution is required.

For example, let

$$\mathscr{P} = \begin{pmatrix} * & * \\ 0 & * \end{pmatrix}, \quad \mathscr{Q} = \begin{pmatrix} * & 0 \\ * & * \end{pmatrix}. \tag{5.1}$$

We want to know if the two patterns commute. We replace the $*$'s by variables, and we have:

$$\mathscr{P} = \begin{pmatrix} x_1 & x_2 \\ 0 & x_3 \end{pmatrix}, \quad \mathscr{Q} = \begin{pmatrix} y_1 & 0 \\ y_2 & y_3 \end{pmatrix}. \tag{5.2}$$

The the two pattern commute if and only if there exist totally nonzero (x, y), such that

$$\begin{pmatrix} x_1 & x_2 \\ 0 & x_3 \end{pmatrix} \begin{pmatrix} y_1 & 0 \\ y_2 & y_3 \end{pmatrix} - \begin{pmatrix} y_1 & 0 \\ y_2 & y_3 \end{pmatrix} \begin{pmatrix} x_1 & x_2 \\ 0 & x_3 \end{pmatrix} = 0 \tag{5.3}$$

i.e.

$$\begin{pmatrix} x_2y_2 & x_2y_3 - x_2y_1 \\ x_3y_2 - x_1y_2 & -x_2y_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \tag{5.4}$$

which is, as a system of equations:

$$\begin{aligned} x_2y_2 &= 0 \\ x_2y_3 - x_2y_1 &= 0 \\ x_3y_2 - x_1y_2 &= 0. \end{aligned} \tag{5.5}$$

The left hand side is linear to both x and y, hence the system above is a BLS. In this specific example, $\mathscr{P}$ and $\mathscr{Q}$ do not commute, as equation $x_2y_2 = 0$ forbid x_2 and y_2 to be both nonzero, and hence there's no $A \in \mathscr{P}$, $B \in \mathscr{Q}$ such that $AB = BA$.

Details of this research can be found in [3].

5.2 Quaternions

In a recent paper [6], it was proposed to use pairs of three vectors to represent quaternions. It is proven in the paper that, formula

$$T(v,w) = [v \cdot w, v \times w]$$

gives a ring isomorphism between equivalent classes of pairs of three vectors and quaternions. Here, the process of finding the class of pairs of three vectors associated with each quaterion is exactly the process of solving a bilinear system.

Bibliography

[1] Charles R. Johnson and Joshua A. Link. Solution theory for complete bilinear systems of equations. *Numer. Linear Algebra Appl.*, 16(11-12):929–934, 2009.

[2] Scott Cohen and Carlo Tomasi. Systems of bilinear equations. Technical report, Stanford, CA, USA, 1997.

[3] Charles R. Johnson and Maria da Graça Marques. Patterns of commutativity: the commutant of the full pattern. *Electron. J. Linear Algebra*, 14:43–50 (electronic), 2005.

[4] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1990.

[5] David C. Lay. *Linear Algebra and Its Applications*. Addison Wesley, 3 edition, September 2005.

[6] Bob Palais. Quaternions in three dimensions. *arXiv:1011.6279v1*, 2010.