HONAM MATHEMATICAL JOURNAL
Volume 4, Number 1, July. 1982

A Note on Linear Equations over a Commutative Ring

By Hideyuki Matsumura

1. Consider a system of linear equations

(1) $$\sum_{j=1}^{n} a_{ij} x_j = b_i \qquad (1 \leqslant i \leqslant r).$$

Let $A=(a_{ij})$ be the $r \times n$ matrix of the coefficients and $B=(A, \beta)$ be the $r \times (n+1)$ matrix obtained by adding the column $\beta = {}^t(b_1, \cdots, b_r)$ to A.

According to what we learned in the college linear algebra the system (1) is solvable iff

(2) $$\textit{rank } A = \textit{rank } B.$$

But ***this criterion is valid only when we work over a field.*** If we work over a commutative ring R instead of a field, it is another story. For instance $2x+4y=1$ is not solvable in Z, because for $x, y \in Z$ the left hand side is an even number.

Let R be a commutative ring (we assume that R has a unit element 1, and that $1 \neq 0$), and suppose a_{ij}, $b_i \in R$ in (1). Let

(3) $$\phi : R^n \to R^r$$

be the linear mapping from the free module R^n with basis $e_1, \cdots, e_n$ to the free module R^r with basis $e'_1, \cdots, e'_r$ defined by

(4) $$\phi(e_j) = \sum_{i=1}^{r} a_{ij} e'_i \qquad (1 \leqslant j \leqslant n).$$

Then (1) is solvable in R iff the element $\beta = \sum_{i=1}^{r} b_i e'_i \in R^r$ is in $\mathrm{Im}(\phi)$, or what is the same thing, the image $\bar{\beta}$ of β in $R^r / \textit{Im}(\phi)$ is zero.

For every prime ideal P of R let a_P denote the canonical image of $a \in R$ in the local ring R_P. If the system of equations $\sum_j a_{ij} x_j = b_i$ $(1 \leqslant i \leqslant r)$ is solvable in R_P we will simply say that (1) is solvable in R_P. On the other hand, the matrix (a_{ij}) defines the localization $\phi_P : R_P^n \to R_P^r$ of ϕ, and we have $R_P^r / \textit{Im}(\phi_P) = (R^r)_P / \textit{Im}(\phi)_P = (R^r / \textit{Im}(\phi))_P$. Thus (1) is solvable in R_P iff $\bar{\beta}_P = 0$, where $\bar{\beta}_P$ is the canonical image of $\bar{\beta}$ in $(R^r / \textit{Im}(\phi))_P$.

In general, if M is an R-module and $m \in M$, then we know ([6] p. 7 Lemma

1) that

$$m=0 \iff m_P=O \text{ in } M_P \text{ for all } P \in Max\,(R)$$

where Max (R) denotes the set of the maximal ideals of R. Therefore we have the following theorem.

THEOREM 1. The system of linear equations

$$\sum_{j=1}^{n} a_{ij}x_j = b_i \qquad (1 \leqslant i \leqslant r) \tag{1}$$

over a commutative ring R is solvable in R iff it is solvable in R_P for every $P \in Max\,(R)$.

2. Let $A=(a_{ij})$ be an $r \times n$ matrix over R ($i.e.\ a_{ij} \in R$) and let t be a positive integer. The ideal of R generated by the $t \times t$ minors of A is denoted by $I_t(A)$. Fo $t > min\,(r,n)$ we put $I_t(A)=(0)$. We have $I_1(A) \supseteq I_2(A) \supseteq \cdots$. These ideals are called the determinantal ideals and have been studied by many mathematicians. As in the case of a field we define the rank of the matrix A by

$$rank\ A = sup\ \{\, t \mid I_t(A) \neq 0 \,\}. \tag{5}$$

By an elementary operation on A we mean

(a) a permutation of the rows (or the columns),

(b) multiplication of a row (or column) by a unit of R,

(c) replacing the i- th row (or column) a_i by $a_i + ca_j$, where $c \in R$ and $j \neq i$

Then $I_t(A)$ does not change when we perform elementary operations on A. If a row (or colum) is a linear combination of the other rows (respectively, columns) with coefficients in R, then we can remove it from A without changing I_t (A). Returning to the system of linear equations (1), if it has a solution in R then the last column β of the matrix B is a linear combination of the columns of A. Therefore:

THEOREM 2. *In order that* (1) *is solvable in R, it is necessary that we have*

$$I_t(A)=I_t(B) \qquad t=1,\ 2,\ \cdots,\ r. \tag{6}$$

Remark. Unfortunately (6) is not sufficient in general, as the following example shows.

Example 1. Let $R=k[u,v]$, where K is a field and u, v are indeterminates. Consider the system

$$\begin{cases} u^2x_1 + vx_2 + ux_3 = v \\ vx_1 \qquad + ux_3 = u^2. \end{cases} \tag{7}$$

Then $$A=\begin{pmatrix} u^2 & v & u \\ v & o & u \end{pmatrix}, \qquad B=\begin{pmatrix} u^2 & v & u & v \\ v & o & u & u^2 \end{pmatrix}$$

and $I_1(A)=(u,\ v)R=I_1(B)$, $I_2(A)=(u^3, uv,\ v^2)R=I_2(B)$. But from the second equation $vx_1+ux_3=u^2$ it follows that x_1 must be divisible by u in R. So we put $x_1=ux_1'$. Then (7) is equivalent to

$$\text{(8)} \qquad \begin{cases} u^3x_1'+vx_2+ux_3=v \\ vx_1' \qquad\quad +x_3 =u. \end{cases}$$

This time we have

$$A=\begin{pmatrix} u^3 & v & u \\ v & o & 1 \end{pmatrix}, \qquad I_2(A)=(v,\ u^3)R,$$

$$B=\begin{pmatrix} u^3 & v & u & v \\ v & o & 1 & u \end{pmatrix}, \qquad u^2\in I_2(B).$$

Hence $I_2(B)\neq I_2(A)$ and (8) is not solvable. Thus (7) satisfies (6), but is not solvable.

THEOREM 3. *Let $A=(a_{ij})$ be an $r\times n$ matrix over a commutative ring R with $I_r(A)=R$. Then*

$$\text{(1)} \qquad \sum_j a_{ij}x_j=b_i \qquad (1\leqslant i\leqslant r)$$

is solvable in R for any constant terms b_i; in other words the linear mapping ϕ: $R^n\to R^r$ defined by the matrix A is surjective. Conversely, if ϕ is surjective then

$$I_r(A)=R.$$

Proof. If ϕ is surjective then the columns of the $r\times r$ unit matrix E are linear combinations of the columns of A, hence $I_r(A)=I_r(C)$ where C is the $r\times(n+r)$ matrix $(A,\ E)$. Clearly, $I_r(C)=R$.

Conversely, assume $I_r(A)=R$. We have to prove that (1) is solvable in R, and by Theorem 1 it suffices to show that (1) solvable in R_P for every $P\in Max(R)$. Thus we may assume that R is a local ring with maximal ideal m. (Note that the hypothesis $I_r(A)=R$ is preserved by localization.) Then $I_r(A)=R$ implies that some $r\times r$ minor of A is a unit of R. Suppose, say, that $\det(a_{ij})_{1\leqslant i\leqslant r,\ 1\leqslant j\leqslant r}$ is a unit of R. Then, if we put $x_{r+1}=\cdots=x_n=0$ and solve the equations

$$\text{(9)} \qquad \sum_{j=1}^{r} a_{ij}x_j=b_i \qquad (1\leqslant i\leqslant r)$$

by Cramer's rule, we get a solution of (1) in R. Q. E. D.

3. In this section we assume that R is an integral domain and $rank\,(A)=rank\,(B)=r'$. If $r'<r$ and if, say, the first r' rows of A are linearly independent, then a solution of the first r' equations in (1) automatically satisfies the remaining $r-r'$ equations. Therefore we can throw away the last $r-r'$ equations and assume $r=r'$.

Recall that an integrally closed noetherian domain of Krull dimension one is called a Dedekind domain. The ring of the algebraic integers in an algebraic number field K is a Dedekind domain. If P is a maximal ideal of a Dedekind domain R then the local ring R_P is a DVR (=discrete valuation ring).

THEOREM 4. *If R is a Dedekind domain and $rank\,(A)=rank\,(B)=r'$, then system (1) is solvable in R iff*

$$I_{r'}(A)=I_{r'}(B). \tag{10}$$

Proof. We have already seen that the condition is necessary (Th. 2). To prove the sufficiency, we may localize at a maximal ideal (Th. 1) because the condition (10) is preserved by localization. Therefore we assume that R is a DVR with prime element t. Then every nonzero ideal of R is of the form t^kR, $k\geqslant 0$. By the remark at the beginning of this section we may assume that $r=r'$.

It is easy to transform the matrix A to the form

$$\begin{pmatrix} t^{e_1} & & & 0 \\ & t^{e_2} & & \\ & & \ddots & \\ 0 & & & t^{e_r} \end{pmatrix} \qquad 0\leqslant e_1\leqslant e_2\leqslant\cdots\leqslant e_r, \tag{11}$$

by elementary operations. In fact, if e_1 is the largest integer such that t^{e_1} divides all a_{ij}, we may assum, by permutations of rows and columns, that $a_{11}R=t^{e_1}R$. Multiplying the first column by a unit we may assume that $a_{11}=t^{e_1}$. Then, since all a_{ij} are divisible by a_{11}, we can reduce $a_{12}, \cdots, a_{1n}$ to zero by suitable elementary column operations. Then by elementary row operations we reduce $a_{21}, \cdots, a_{n1}$ to zero, to obtain

$$\begin{pmatrix} t^{e_1} & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{pmatrix}$$

where the entries a'_{ij} of the $(r-1)\times(n-1)$ matrix A' are divisible by t^{e_1}. We

repeat the process with A' instead of A, and so on. In the end A is transformed into the form (11). It should be remembered, however, that when we perform an elementary row operation on A, the same operation must be performed on B, because an elementary row operation on A corresponds to the transition from the original system of equations to an equivalent system. On the other hand, an elementary column operation corresponds to an invertible change of variables.

If A is of the form (11), then we have $I_r(A)=t^{e_1+\cdots+e_r}R$ and

(12) $$B=\begin{pmatrix} t^{e_1} & & & b_1 \\ & \ddots & & \vdots \\ & & 0 & \vdots \\ 0 & & t^{e_r} & b_r \end{pmatrix}$$

and the $r\times r$ minors of B other than those of A are :

$$b_i\, t^{e_1+\cdots+e_{i-1}+e_{i+1}+\cdots+e_r} \qquad i=1,\cdots,r.$$

These elements are in $I_r(A)$ iff

$$t^{e_i} \mid b_i \qquad i=1,\cdots,r,$$

i. e. iff the system

(13) $$t^{e_i}x_i=b_i, \qquad i=1,\cdots,r$$

is solvable in R. Q. E. D.

Remark 2. It is well known that, if R is a principal ideal domain, then for any matrix A over R there exist invertible matrices P and Q such that PAQ is of the form

$$\begin{pmatrix} a_1 & & & & \\ & a_2 & & 0 & \\ & & \ddots & & 0 \\ & & & a_r & \\ 0 & & & 0 & \end{pmatrix} \qquad a_1 \mid a_2 \mid \cdots \mid a_r,$$

with $a_i\neq 0$ and $r=\mathrm{rank}\,(A)$. (*Cf.* *e. g.* Bourbaki, Algébre Ch. 7.)

The proof is particularly simple in the case of DVR. as we have seen above.

Remark 3. According to Dickson ([2] p. 82−83), Th. 3 in the case $R=\mathbf{Z}$ was first proved by J. Heger [5] in 1856. I don't know if the case of a Dedekind domain has appeared in the literature.

4. When R is a noetherian ring, the solutions of a given system of homogeneous linear equations constitute a finitely generated R-module. Therefore, assuming that we can

I) decide whether a single linear equation

$$\sum_{i=1}^{n} a_i x_i = b$$

has a solution in R or not; and

II) find a system of generators (i.e. so-called fundamental system)

$$\xi^\alpha = (\xi_1^\alpha, \xi_2^\alpha, \cdots, \xi_n^\alpha), \quad 1 \leqslant \alpha \leqslant n_1$$

of the R-module of solutions of the homogeneous equation

$$\sum a_i x_i = 0,$$

we can theoretically decide the solvability of (1) as follows:

First we take up the first equation

(14) $$\sum a_{1i} x_i = b_1.$$

If it is not solvable, we are done. Otherwise we find the complete set of solutions of (14) in the form

(15) $$(x_1, \cdots, x_n) = \eta + \sum_{\alpha=1}^{n_1} t_\alpha \xi^\alpha,$$

where η is a solution of (14), $t_1, \cdots, t_{n_1}$ are parameters and $\{ \xi^\alpha \mid 1 \leqslant \alpha \leqslant n_1 \}$ is a system of generators of the solution module of $\sum a_{1i} x_i = 0$.

Then we substitute $x_i = \eta_i + \sum t_\alpha \xi_i^\alpha \ (1 \leqslant i \leqslant n)$ into the remaining $r-1$ equations of (1), to get a system of $r-1$ linear equations in n_1 unknowns $t_\alpha \ (1 \leqslant \alpha \leqslant n_1)$. Reducing the number of equations in this way, we will be able, not only to decide the solvability, but also to obtain (*if it is solvable*) a complete set of solutions of (1). The disavantage of this procedure is that we cannot decide the solvability until we really find the solution.

In II) we need some criterion to decide whether the solutions $\xi^1, \cdots, \xi^{n_1}$ generate the complete set of solutions of $\sum a_i x_i = 0$. In other words, given n_1 solutions $\xi^i = (\xi_1^i, \cdots, \xi_n^i)$ of $\sum a_i x_i = 0$, we must decide whether the sequence of Rmodules

(16) $$R^{n_1} \xrightarrow{\psi_2} R^n \xrightarrow{\psi_1} R$$

is exact, where ψ_1 and ψ_2 are defined by the matrices $(a_1, \cdots, a_n)$ and (ξ_j^i) respectively. We can ask, more generally, the problem of deciding the exactness of a sequence of the form

(17) $$R^p \xrightarrow{\psi_2} R^q \xrightarrow{\psi_1} R^r$$

where $\psi_2 \cdot \psi_1 = 0$ and ψ_i is defined by a matrix $A_i \ (i = 1, 2)$.

THEOREM 5. *If R is a Dedekind domain, then the sequence* (17) *is exact iff the*

following conditions hold :

(i) $rank\ A_1+rank\ A_2=q$,

and (ii) $I_{p'}(A_2)=R$, *where* $p'=rank\ A_2$.

Proof. A sequence of R-modules is exact iff it is exact after localization at P at every $P\in Max\ (R)$. The rank of a matrix does not change by localization, since R is an integral domain. An ideal I of R is equal to R iff $IR_P=R_P$ for every $P\in Max\ (R)$. Therefore we may localize and assume that R is a DVR. Then we may assume that A_1 is of the form $A_1=\begin{pmatrix}D&0\\0&0\end{pmatrix}$, where D is a diagonal $q'\times q'$ matrix similar to (11) with $q'=rank\ A_1$. Then the necessity of the conditions (i), (ii) is easy to check. To see the sufficiency we note that, since R is a local ring, (ii) holds iff some $p'\times p'$ minor of A_2 is a unit of R. Since $A_1A_2=0$ the matrix A_2 must be of the form $\begin{pmatrix}0\\C\end{pmatrix}$, where C is a $(q-q')\times p$ matrix. By (i) we have $p'=q-q'$. Thus we may suppose that C is of the form $(E,\ 0)$ where E is the $p'\times p'$ unit matrix. Then (17) is exact. Q. E. D.

Remark 4. According to Dickson ([2] p. 84), this theorem was proved by Frobenius [3] in the case $R=\mathbf{Z}$.

Remark 5. If R is a principal ideal domain, every submodule of a free module is free. Therefore the solution module of an equation $a_1x_1+\cdots+a_nx_n=0$ can be generated by $n-1$ vectors.

Example 2. Consider the equation

$$3x+4y+5z=0$$

over $\mathbf{Z}$. Denote the solution module by M. put $\xi^1=(0,\ 5,\ -4)$, $\xi^2=(5,\ 0,\ -3)$, $\xi^3=(4,\ -3,\ 0)$, $\eta=\xi^2-\xi^3=(1,\ 3,\ -3)$. Applying the theorem we see that no two of ξ^1, ξ^2, ξ^3 can generate M, and that ξ^1 and η generate M.

Unfortunately it is difficult to generalize the theorems 4 and 5 to rings of dimension >1. In the problem of deciding the exactness of (17) the most powerful tool is perhaps the criterion of Buchsbaum-Eisenbud [1], which can be used, e. g., in proving the following partial result.

If I is an ideal of a noetherian ring R and if $I\neq R$, we let depth (I, R) be the length of a maximal R-sequence contained in I. (When R is Cohen-Macaulay we have depth $(I, R)=ht\ I$.) If $I=R$ we set depth $(I, R)=\infty$.

A noetherian domain is called a regular domain if R_P is a regular local ring for

every maximal ideal P.

THEOREM 6. *Let R be a regular domain. Then* (17) *is exact only if the following condition hold* : (i) *rank* $(A_1)+$*rank* $(A_2)=q$, *and* (ii) *depth* $(I_{p^*}(A_2), R)>2$, where $p^*=$*rank* (A_2). Conversely, if these conditions are satisfied, if $\psi_1 \neq 0$ and if ψ_2 is injective, then (17) is exact.

Proof. Exactness and rank do not change by localization. Moreover, If I is an ideal of R and P is a prime ideal containing I, then in general we have depth $(I, R) \leqslant$ depth $(IR_P, R_P) \leqslant$ depth $(PR_P, R_P)=$depth (R_P), and the equalities hold for at least one P (*cf.* [6] first edition p. 101, 2 nd edition p. 105). Therefore we may localize and assume that R is a regular local ring. Let (17) be exact. Then (i) holds since it holds over the quotient field of R. As for (ii) we use the famous theorem of *J.-P.* Serre which says that every module over a regular local ring has a finite projective dimension. Thus, taking a free resolution of ker (ψ_2) of finite length and adding it to the left of (17), we get a complex

$$0 \to R^h \xrightarrow{\psi_d} R^k \to \cdots \xrightarrow{\psi_3} R^p \xrightarrow{\psi_2} R^q \xrightarrow{\psi_1} R^r, \tag{18}$$

which is exact. According to an important theorem of Buchsbaum-Eisenbud [1] (*cf*, also [7]), a complex

$$0 \to F_n \xrightarrow{\phi_n} F_{n-1} \to \cdots \xrightarrow{\phi_2} F_1 \xrightarrow{\phi_1} F_0, \tag{19}$$

where F_i are free modules of finite rank, is exact iff, for $1 \leqslant k \leqslant n$, we have (i)* *rank* $\phi_{k+1}+$ *rank* $\phi_k=$*rank* F_k and (ii)* depth $(I_{r_k}(\phi_k)), R) \geqslant k$ where $r_k=$ *rank* (ϕ_k). (By $I_t(\phi)$ we mean $I_t(A)$, where A is a matrix representing the linear map ϕ with respect to some bases. This definition is obviously independent of the choice of the bases. Similarly for rank (ϕ).) Applying this theorem to (18) we see that depth $(I_{p^*}(A_2), R) \geq 2$, as wanted.

If we do not assume that (17) is exact, but assume that $\psi_1 \neq 0$ and ψ_2 is injective, then the sequence

$$0 \to R^p \xrightarrow{\psi_2} R^q \xrightarrow{\psi_1} R^r \tag{20}$$

is a complex of the type (19), and if the conditions (i), (ii) of the theorem are satisfied then the conditions (i)*, (ii)* of the Buchsbaum-Eisenbud theorem hold. Thus (20) (hence also (17)) is exact. Q. E. D.

Example 3. Let $R=k[u, v]$ be the polynomial ring in u, v over a field k, and consider the equation

$$u^2 x + uvy + v^2 z = 0. \tag{21}$$

The vectors $\xi=(-v,\ u,\ o)$, $\eta=(o,\ -v,\ u)$ generate the solution module of (21), as one can easily check by Th. 6. On the other hand the matrix

$$A=\begin{pmatrix} -v^2 & -uv & o \\ uv & u^2 & -v \\ o & o & u \end{pmatrix}$$

has rank 2 and depth $(I_2(A), R)=2$, but its columns $v\xi$, $u\xi$, η do not generate the solution module.

Question. *Are there any convenient criteria for the solvability of* (1) *over, say,* $k[u, v]$?

References

[1] D. Buchsbaum-D. Eisenbud *What makes a complex exact* ? J. of Algebra **25** (1973), pp. 259−268.

[2] L. E. Dickson, *History of The Theory of Numbers,* Vol. II. ca. 1920. Reprint from Chelsea Publ. Co. 1952.

[3] G. Frobenius, *Theorie der lineare Formen mit ganzen Coefficienten.* J. reine u. angew. Math. **86** (1878), pp. 146−208.

[4] G. Frobenius, *Theorie der lineare Formen mit ganzen Coefficienten (Fortsetzung).* ibid. **88** (1879), pp. 96−116.

[5] J. Heger, *Über die Auflösung eines Systemes von mehreren unbestimmten Gleichungen des erstes Grades in ganzen Zahlen, welche eine grössere Anzahl von Unbekannten in sich schliesen, als sie zu bestimmen vermögen.* Sitzungsberichte Akad. Wissen. Wien **21** (1856), pp. 550−560.

[6] H. Matsumura, *Commutative Algebra.* Benjamin, first ed. 1970, second ed. 1980.

[7] D. G. Northcott, *Finite Free Resolutions,* Cambridge Univ. Press 1976.

Department of Mathematics
Faculty of Sciences
Nagoya University
Nagoya, 464, Japan.