Linear Algebra Notes

Mario L. Gutierrez Abed

Elementary Matrix Operations and Systems of Linear Equations

ELEMENTARY MATRICES

• Theorem:
Let $A \in M_{m\times n}(\mathbb{F})$, and suppose matrix B is obtained from A by applying an elementary row/column operation. Then
a) $\exists$ an $m \times m$ elementary matrix E such that $B = E\,A$, if B is obtained by a row operation.
b) $\exists$ an $n \times n$ elementary matrix E' such that $B = A\,E'$, if B is obtained by a column operation.

• Theorem:
Elementary matrices are invertible. Moreover, the inverse of an elementary matrix is an elementary matrix of the same type.

RANK OF A MATRIX AND MATRIX INVERSES

Definition: Let $A \in M_{m\times n}(\mathbb{F})$. Then $\operatorname{rank}(A) = \operatorname{rank}(L_A)$, where $L_A : \mathbb{F}^n \longrightarrow \mathbb{F}^m$, $x \longmapsto A\,x$. Also $\operatorname{nullity}(A) = \operatorname{nullity}(L_A)$.

• Theorem:
An $n \times n$ matrix A is invertible iff $\operatorname{rank}(A) = n$.

Proof:
$(\Rightarrow)$
Suppose A is invertible. Then L_A is invertible $\Longrightarrow$ L_A is bijective $\Longrightarrow$

$\underbrace{\text{rank}(L_A) = \text{rank}(A) = \dim(\mathbb{F}^n)}_{\text{By the definition stated above}} = n$. ✓

($\Leftarrow$)

Suppose rank $(A) = n$. By definition $\text{rank}(L_A) = n$.

By the Rank-Nullity theorem,

$$\dim(\mathbb{F}^n) = \text{nullity}(L_A) + \text{rank}(L_A) \Longrightarrow n = \text{nullity}(L_A) + n \Longrightarrow \text{nullity}(L_A) = 0.$$

Thus L_A is injective. L_A is also surjective since the codomain is also $\mathbb{F}^n$.

So L_A is invertible $\Longrightarrow [L_A]_\beta^\gamma$ is invertible. ✓ ■

• Theorem:

Let V^n, W^m be finite dimensional VS's. Let β and γ be bases for V and W respectively and let $T \in \mathcal{L}(V, W)$. Then $\text{rank}(T) = \text{rank}[T]_\beta^\gamma$.

• Lemma:

Let $T : V_\beta \longrightarrow W_\gamma$ and $L_A : \mathbb{F}^n \longrightarrow \mathbb{F}^m$, where $A = [T]_\beta^\gamma$. Then,

$\text{nullity}(T) = \text{nullity}(L_A)$ and $\text{rank}(T) = \text{rank}(L_A)$. (Use the diagram for guidance)

$$\begin{tikzcd}
V_\beta^n \arrow[r, "T"] \arrow[d, "\mathcal{M}_\beta"'] \arrow[dr, phantom, "\iota"] & W_\gamma^m \arrow[d, "\mathcal{M}_\gamma"] \\
\mathbb{F}^n \arrow[r, "L_A"'] & \mathbb{F}^m
\end{tikzcd}$$

Proof:

Using problem #17 of section 2.4 (from the HW), let $V_0 = \mathcal{N}(T) \subseteq V$.

Then,

$$\mathcal{M}_\beta(V_0) \cong V_0 \Longrightarrow \dim\big(\mathcal{M}_\beta(V_0)\big) = \dim(V_0).$$

Now we only need to show that $\mathcal{M}_\beta(V_0) = \mathcal{N}(L_A)$:

($\subseteq$)

Let $x \in \mathcal{M}_\beta(V_0)$. That implies that $\exists\ \hat{x} \in V_0 = \mathcal{N}(T)$ such that $\mathcal{M}_\beta(\hat{x}) = x$.

Thus,

$$T(\hat{x}) = 0 \Longrightarrow \mathcal{M}_\gamma(T\,\hat{x}) = 0 \Longrightarrow (\mathcal{M}_\gamma\,T)\,\hat{x} = 0.$$

Thus by commutativity of the diagram
above,

$(\mathcal{M}_\gamma\,T)\,\hat{x} = \left(L_A\,\mathcal{M}_\beta\right)x = 0 \Longrightarrow L_A\left(\mathcal{M}_\beta\,\hat{x}\right) = 0 \Longrightarrow \mathcal{M}_\beta\,\hat{x} = \mathcal{M}_\beta\,x \in \mathcal{N}(L_A).$ ✓

(⊇)

Let $y \in \mathcal{N}(L_A) \subseteq \mathbb{F}^n$. So $L_A(y) = 0$. Also $\exists\ \hat{y} \in V$ such that $\mathcal{M}_\beta(\hat{y}) = y$.

So,

$$L_A(y) = L_A\left(\mathcal{M}_\beta\,\hat{y}\right) = 0 \Longrightarrow \left(L_A\,\mathcal{M}_\beta\right)\hat{y} = 0.$$

By commutativity of the diagram above,

$\left(L_A\,\mathcal{M}_\beta\right)\hat{y} =$.

$(\mathcal{M}_\gamma\,T)\,\hat{y} = 0 \Longrightarrow \mathcal{M}_\gamma\left(T\,\hat{y}\right) = 0 \Longrightarrow T\,\hat{y} = 0 \Longrightarrow \hat{y} \in \mathcal{N}(T) \Longrightarrow \mathcal{M}_\beta\,\hat{y} = y \in \mathcal{M}_\beta(\mathcal{N}(T))$

Thus,

$$\mathcal{M}_\beta(\mathcal{N}(T)) = \mathcal{N}(L_A) \Longrightarrow \mathcal{N}(T) \cong \mathcal{N}(L_A) \Longrightarrow \text{nullity}(T) = \text{nullity}(L_A).$$

Similarly, rank $(T) = \text{rank}(L_A)$. ✓ ■

• Theorem:

Let $A \in M_{m\times n}(\mathbb{F})$ and let $P \in M_{m\times m}(\mathbb{F})$ and $Q \in M_{n\times n}(\mathbb{F})$ be invertible. Then,

a) $\text{rank}(A\,Q) = \text{rank}(A)$

b) $\text{rank}(P\,A) = \text{rank}\,(A)$

c) $\text{rank}(P\,A\,Q) = \text{rank}\,(A)$

Proof:

a) We want to prove that $\text{rank}(A\,Q) = \text{rank}\,(A)$.

First observe that

$$R\left(L_{AQ}\right) = R\left(L_A\,L_Q\right) = L_A\,L_Q\,(\mathbb{F}^n) = L_A\left(L_Q\,(\mathbb{F}^n)\right) = L_A(\mathbb{F}^n) = R(L_A) \quad (\text{since } L_Q \text{ is onto}).$$

Therefore,

$$\text{rank}(A\,Q) = \dim\left(R\left(L_{AQ}\right)\right) = \dim(R(L_A)) = \text{rank}(A).$$ ✓

For b) and c) the proof is similar. .... ■

• Corollary:

Elementary row/column operations on a matrix are rank-preserving.

Definition:

1) The column space of an $m \times n$ matrix A, denoted Col (A), is the subspace of $\mathbb{F}^m$ that is generated by the columns of A.

2) The row space of A, Row (A), is the subspace of $\mathbb{F}^n$ that is generated by the rows of A.

Example:

Given $A = \begin{pmatrix} 3 & 1 & 2 & 5 \\ 1 & 2 & 0 & 1 \\ 2 & 1 & 3 & 2 \end{pmatrix}$, we have

$$\text{Col}(A) = \text{span}\left\{\begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix}, \begin{pmatrix} 5 \\ 1 \\ 2 \end{pmatrix}\right\} \quad \text{and}$$

$$\text{Row}(A) = \text{span}\{(3, 1, 2, 5), (1, 2, 0, 1), (2, 1, 3, 2)\}.$$

Note (we use this fact to prove the theorem below):

$A \in M_{m\times n}(\mathbb{F}) \Longrightarrow L_A : \mathbb{F}^n \longrightarrow \mathbb{F}^m$ with $x \longmapsto A\,x$. Let β be the standard ordered basis for $\mathbb{F}^n$, i.e. $\beta = \{e_1, \ldots, e_n\}$.

Observe the following: $A\,e_1 = c_1$, where $c_1 = 1^{\text{st}}$ column of A. In general, we have that $A\,e_n = c_n$.

• Theorem:

Col $(A) = R(L_A)$.

This implies that dim(Col(A)) = rank(A), i.e. rank(A) is the maximum number of linearly independent columns of A (rank (A) = cardinality of a basis for Col(A)).

Proof:

($\subseteq$)

We want to show that $R(L_A) \subseteq \text{Col}(A)$.

Let $b \in R(L_A)$.

Then $\exists\, x \in \mathbb{F}^n$ such that $L_A(x) = A\,x = b$.

Let β be the standard ordered basis for $\mathbb{F}^n$. Then $x = x_1\,e_1 + \ldots + x_n\,e_n \;\; \forall\; x_i \in \mathbb{F}$.

Then,

$$\begin{aligned} b = A\,x = A(x_1\,e_1 + \ldots + x_n\,e_n) &= L_A(x_1\,e_1 + \ldots + x_n\,e_n) \\ &= x_1\,L_A\,e_1 + \ldots + x_n\,L_A\,e_n \\ &= x_1\,A\,e_1 + \ldots + x_n\,A\,e_n = x_1\,c_1 + \ldots + x_n\,c_n. \end{aligned}$$

Thus $b \in \text{span}(c_1, \ldots, c_n) = \text{Col}(A)$. ✓

(⊇)

We want to show that $R(L_A) \supseteq \text{Col}(A)$. (alternatively $\text{Col}(A) \subseteq R(L_A)$)

Let $\hat{b} \in \text{Col}(A)$.

Then $\exists\ d_i \in \mathbb{F}$ such that

$$\begin{aligned}\hat{b} = d_1\, c_1 + \ldots + d_n\, c_n = d_1\, c_1 + \ldots + d_n\, c_n \\ &= d_1\, A\, e_1 + \ldots + d_n\, A\, e_n \\ &= d_1\, L_A\, e_1 + \ldots + d_n\, L_A\, e_n \\ &= L_A(d_1\, e_1 + \ldots + d_n\, e_n) \\ &= A(d_1\, e_1 + \ldots + d_n\, A\, e_n)\ .\end{aligned}$$

So we have

$$\hat{b} = A\, d = L_A(d) \Longrightarrow \hat{b} \in\ R(L_A). \qquad \checkmark$$

■

• Corollary:

$\text{Col}(A) \cong R(T)$, where $A = [T]_\beta^\gamma$.

• Theorem:

Let $A \in M_{m\times n}(\mathbb{F})$ with $\text{rank}(A) = r$. Then $r \leq m$ and $r \leq n$. Also by applying a finite number of row/column operations to A, it can be transformed into the following matrix:

$$D = \begin{pmatrix} I_r & 0_1 \\ 0_2 & 0_3 \end{pmatrix},$$

where the 0_i are zero matrices and $D \in M_{m\times n}(\mathbb{F})$ $\left(D_{i\,j} = \begin{cases} 1 & \text{if } i = j \leq r \\ 0 & \text{otherwise} \end{cases} \right)$.

** For instance, for $r = 3$, D looks like $\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ **

• Corollary 1:

Let $A \in M_{m\times n}(\mathbb{F})$ with $\text{rank}(A) = r$. Then $\exists$ invertible matrices $B \in M_{m\times m}(\mathbb{F})$ and $C \in M_{n\times n}(\mathbb{F})$ such that $D = B\, A\, C$.

Proof:

By the previous theorem, A can be transformed into D as defined above by elementary row/column operations to A. Thus,

$$D = E_p \ldots E_3\, E_2\, E_1\, A\, G_1\, G_2\, G_3 \ldots G_q.$$

Since the product of invertible matrices is also invertible, we can can define $B_{\text{inv}} = E_p \ldots E_3\, E_2\, E_1$

and $C_{\text{inv}} = G_1 \, G_2 \, G_3 \ldots G_q$, and thus we have $D = B\,A\,C$. ■

• Corollary 2:

Let $A \in M_{m\times n}(\mathbb{F})$. Then

a) $\text{rank}\left(A^{\text{T}}\right) = \text{rank}(A)$.

b) $\text{rank}(A)$ = max number of linearly independent rows = $\dim(\text{Row}(A))$.

c) $\text{Row}(A) \cong \text{Col}(A)$.

Proof:

a) By Corollary 1, $\exists$ invertible matrices B and C such that $D = B\,A\,C$.

Then $D^{\text{T}} = (B\,A\,C)^{\text{T}} = C^{\text{T}} A^{\text{T}} B^{\text{T}}$.

Since B, C are invertible, B^{T}, C^{T} are invertible.

Note that $\text{rank}\left(D^{\text{T}}\right) = \text{rank}(D) = \text{rank}(A)$.

But $\text{rank}\left(D^{\text{T}}\right) = \text{rank}\left(A^{\text{T}}\right)$, since A^{T} can be transformed into D^{T} as shown on the theorem.

Hence $\text{rank}\left(A^{\text{T}}\right) = \text{rank}(A)$. ✓

b) By part a),

$\text{rank}(A) = \dim(\text{Col}(A))$ and

$\text{rank}\left(A^{\text{T}}\right) = \dim(\text{Row}(A))$. ✓

c) By part b),

$\text{Row}\,(A) \cong \text{Col}\,(A)$. ✓ ■

• Corollary 3:

Every invertible matrix is a product of elementary matrices.

Proof:

Let $A \in M_{n\times n}(\mathbb{F})$ be invertible. Then $\text{rank}(A) = n$.

By a previous theorem, A can be transformed into $D = I_n$.

Then by Corollary 1, this matrix $D = I_n$ can be written as $I_n = B\,A\,C$, with B and C invertible.

Since B, C are invertible, B^{-1}, C^{-1} exist.

Then we have

$$B^{-1}(I_n)\,C^{-1} = B^{-1}(B\,A\,C)\,C^{-1}$$

$$\Longrightarrow B^{-1}\,C^{-1} = A \Longrightarrow (E_p \ldots E_1)^{-1}\,(G_1 \ldots G_q)^{-1} = A\,. \quad \text{(from proof of corollary 1)}$$

So we have that

$$A = \left(E_1^{-1} \ldots E_p^{-1}\right)\left(G_q^{-1} \ldots G_1^{-1}\right).$$

Since the inverse of an elementary matrix is also an elementary matrix, we have that the above is a product of elementary matrices. ■

• Theorem:

Let $T \in \mathcal{L}(V, W)$ and $U \in \mathcal{L}(W, Z)$, and let A, B be matrices such that AB is defined.
Then,
a) $\text{rank}(UT) \leq \text{rank}(U)$
b) $\text{rank}(UT) \leq \text{rank}(T)$
c) $\text{rank}(AB) \leq \text{rank}(A)$
d) $\text{rank}(AB) \leq \text{rank}(B)$

SYSTEMS OF LINEAR EQUATIONS

• Lemma:

If M is appropriately defined, we have that $M(A \mid B) = (M\,A \mid M\,B)$.
However, $(A \mid B)\,M \neq (A\,M \mid B\,M)$.
** This is the reason why we only use row operations when we're looking for the inverse of a matrix **

Note: Let A be an invertible $n \times n$ matrix and consider the augmented matrix $C = (A \mid I_n)$. Then
$A^{-1}\,C = A^{-1}(A \mid I_n) = (A^{-1}\,A \mid A^{-1}\,I_n) = (I_n \mid A^{-1})$

• Theorem:

Let A be an $m \times n$ matrix and let $A\,x = 0$ be a homogenous system. Let K be the solution set to the system $A\,x = 0$. Then $K = \mathcal{N}(L_A)$.
Also $\dim(K) = n - \text{rank}(A)$.

Proof:

Since K is the solution set to $A\,x = 0$, by definition $K = \{s \in \mathbb{F}^n \mid A\,s = 0\}$. We have that $L_A(s) = A\,s$.
So $K = \{s \in \mathbb{F}^n \mid L_A(s) = 0\}$.
Hence $K = \mathcal{N}(L_A)$ by definition. ✓
Now we have that

$$\begin{aligned} \dim(K) &= \text{nullity}(L_A) \\ &= n - \text{rank}(L_A) \\ &= n - \text{rank}(A) \quad ✓ \end{aligned}$$ ■

• Corollary:

If $m < n$, $A\,x = 0$ has a nonzero solution.

Proof:
If $m < n$, L_A cannot be injective. So we have that nulity$(L_A) > 0$.
This implies that $\exists$ a nonzero n-tuple in $\mathcal{N}(L_A) = K$. ■

Note: Since K is a subspace, we can find a basis of solutions to the system $A\,x = 0$.

We are going to use homogenous systems to get solutions for nonhomogenous systems. Let $A\,x = b$ be a nonhomogenous system. Then, $A\,x = 0$ is the associated homogenous system to $A\,x = b$.

- Theorem:

Let K be the solution set to $A\,x = b$ and K_H be the solution set to the associated homogenous system $A\,x = 0$.
Then for any $s \in K$,

$$K = \{s\} + K_H = \{s + k : k \in K_H\}.$$

Proof:
Let $s \in K$.
$(k \subseteq \{s\} + K_H)$
Let $w \in K$. Then $A\,w = b$.
Then consider

$$A(w - s) = A\,w - A\,s = b - b = 0.$$

This implies that $w - s \in K_H$.
This in turn implies that $\exists$ a $k = w - s \in K_H$ such that

$$w - s = k \Longrightarrow w = s + k \Longrightarrow w \in \{s\} + K_H \qquad \checkmark$$

$(\{s\} + K_H \subseteq k)$
Let $v \in \{s\} + K_H$. Then $\exists\, \hat{k} \in K_H$ such that $v = s + \hat{k}$.
Now consider

$$A\,v = A\left(s + \hat{k}\right) = A\,s + A\,\hat{k} = b + 0 = b.$$

$$\Longrightarrow v \in K. \quad \checkmark$$ ■

Example:
Solve the following system of linear equations:

$$x_1 + x_2 - x_3 = 1$$
$$4\,x_1 + x_2 - 2\,x_3 = 3$$

Solution:

$$\begin{pmatrix} 1 & 1 & -1 \\ 4 & 1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

First find a solution to the system:

$$S = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \qquad \checkmark$$

Now we look at the associated homogenous system:

$x_1 + x_2 - x_3 = 0$

$4\,x_1 + x_2 - 2\,x_3 = 0$

$$\begin{pmatrix} 1 & 1 & -1 \\ 4 & 1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$K_H = \mathcal{N}(L_A) \longrightarrow \dim(K) = 3 - \text{rank}(A) = 3 - 2 = 1$

Since $\dim(K) = 1$, we have $\text{nullity}(L_A) = 1 \Longrightarrow \mathcal{N}(L_A)$ is one dimensional.

Hence every solution in K_H is a scalar multiple of one particular solution.

In this case we have

$$k = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

Thus $\left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \right\}$ is a basis for K_H. ✓

Thus we have that

$$K = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + t \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} : t \in \mathbb{R} \right\} \qquad \checkmark$$

• Theorem:

Let $A\,x = b$ be a system of n equations and n unknowns. Then A is invertible iff $A\,x = b$ has a unique solution.

Proof:

$(\Rightarrow)$

Suppose A is invertible. Then A^{-1} exists and

$$A^{-1}(A\,x) = A^{-1}\,b \implies x = A^{-1}\,b.$$

Now suppose that s is another solution to $A\,x = b$, such that $s \neq x$.

Then

$$A^{-1}(A\,s) = A^{-1}\,b \implies s = A^{-1}\,b\,.$$

Thus $s = x$. ($\Rightarrow\Leftarrow$) ✓

($\Leftarrow$)

Suppose $A\,x = b$ has a unique solution s.

Then by the preceding theorem,

$$K = \{s\} + K_H \implies \{s\} = \{s\} + K_H \implies K_H = \{\vec{0}\}$$
$$\implies \dim(K_H) = 0 = n - \mathrm{rank}(A)$$
$$\implies \mathrm{rank}(A) = n$$
$$\implies A \text{ is invertible.} \quad ✓ \quad \blacksquare$$

• Theorem:

Let $A\,x = b$ be a system. Then, the system is consistent iff $\mathrm{rank}(A) = \mathrm{rank}(A \mid b)$.

Proof:

($\Rightarrow$)

Note that if $A\,x = b$, then $b \in R(L_A) = \mathrm{Col}(A)$.

This implies that

$$b \in \mathrm{span}(c_1, \ \ldots, \ c_n) = \mathrm{span}(c_1, \ \ldots, \ c_n, \ b)\,.$$
$$\implies \dim(\mathrm{span}(c_1, \ \ldots, \ c_n)) = \dim(\mathrm{span}(c_1, \ \ldots, \ c_n, \ b))$$
$$\implies \mathrm{rank}(A) = \mathrm{rank}(A \mid b). \quad ✓$$

($\Leftarrow$)

To prove in this direction we simply prove the above backwards, (i.e. Assume $\mathrm{rank}(A) = \mathrm{rank}(A \mid b)$, then $\dim(\mathrm{span}(c_1, \ \ldots, \ c_n)) = \dim(\mathrm{span}(c_1, \ \ldots, \ c_n, \ b))$, then blah blah ....) ✓ $\blacksquare$

Definition: Two systems are said to be equivalent if they have the same solution set.

• Theorem:

Let $A\,x = b$ be a system of m linear equations in m unknowns, and let C be an invertible $m \times m$ matrix.

Then the system $(CA)\,x = C\,b$ is equivalent to $A\,x = b$.

Proof:

Let K be a solution set of $A\,x = b$ and let K' be a solution set of $(CA)\,x = C\,b$. We wish to show that $K = K'$.

($K \subseteq K'$)

Let $s \in K$.

Then

$$A\,s = b$$

$\Longrightarrow C(A\,s) = C\,b$
$\Longrightarrow (CA)\,s = C\,b$
$\Longrightarrow s \in K'$ ✓

$(K' \subseteq K)$
Let $\tilde{s} \in K'$.
Then
$(CA)\,\tilde{s} = C\,b$
$\Longrightarrow C^{-1}(CA)\,\tilde{s} = C^{-1}(C\,b)$
$\Longrightarrow A\,\tilde{s} = b$
$\Longrightarrow \tilde{s} \in K$. ✓ ■

• Corollary:
Let $A\,x = b$ be a nonhomogenous system of m linear equations in n unknowns.
If $(A' \mid b')$ is obtained from $(A \mid b)$ by a finite number of elementary row operations, then the system $A'\,x = b'$ is equivalent to $A\,x = b$.

Proof:
If $(A' \mid b')$ is obtained from $(A \mid b)$ by finitely many elementary row operations, then

$$E_p \ldots. E_2\,E_1(A \mid b) = (A' \mid b'),$$

where E_i are elementary matrices of the appropriate type.
Then by a previous theorem, $E_p \ldots. E_2\,E_1 = C$, where C is invertible.
By the lemma that states that $C(A \mid B) = (C\,A \mid C\,B)$,we have

$$E_p \ldots. E_2\,E_1(A \mid b) = C(A \mid b) = (C\,A \mid C\,b) = (A' \mid b').$$

Thus $A' = CA$ and $b' = C\,b$. Then, by the theorem to which this is a corollary, $(C\,A)\,x = C\,b$ is equivalent to $A\,x = b$, since $(C\,A)\,x = C\,b \Leftrightarrow A'\,x = b'$. ■