Lecture 6

Elementary matrices

Definition 11.1: Fix an integer $n > 0$. We define $n \times n$ matrices as follows.

(a) Suppose that $1 \leq p \leq n$ and that λ is a nonzero real number. We then let $D_p(\lambda)$ be the matrix that is the same as I_n except that $(D_p(\lambda))_{pp} = \lambda$.

(b) Supose that $1 \leq p, q \leq n$ with $p \neq q$, and that μ is an arbitrary real number. We then let $E_{pq}(\mu)$ be the matrix that is the same as I_n except that $(E_{pq}(\lambda))_{pq} = \mu$.

(c) Supose again that $1 \leq p, q \leq n$ with $p \neq q$. We let F_{pq} be the matrix that is the same as I_n except that $(F_{pq})_{pp} = (F_{pq})_{qq} = 0$ and $(F_{pq})_{pq} = (F_{pq})_{qp} = 1$.

An *elementary matrix* is a matrix of one of these types.

Example 11.2: In the case $n = 4$, we have

$$D_2(\lambda) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & & \\ 0 & & & \end{bmatrix} \qquad E_{24}(\mu) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & & \\ 0 & & & \end{bmatrix} \qquad F_{24} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Elementary matrices and row operations

Proposition 11.3: Let A be an $n \times n$ matrix, and let A' be obtained from A by a single row operation. Then $A' = UA$ for some elementary matrix U. In more detail:

(a) Let A' be obtained from A by multiplying the p'th row by λ. Then $A' = D_p(\lambda)A$.

(b) Let A' be obtained from A by adding μ times the q'th row to the p'th row. Then $A' = E_{pq}(\mu)A$.

(c) Let A' be obtained from A by exchanging the p'th row and the q'th row. Then $A' = F_{pq}A$.

Elementary matrices and row operations

Corollary 11.4: Let A and B be $n \times n$ matrices, and suppose that A can be converted to B by a sequence of row operations. Then $B = UA$ for some matrix U that can be expressed as a product of elementary matrices.

Proof.
The assumption is that there is a sequence of matrices $A_0, A_1, \dots, A_r$ starting with $A_0 = A$ and ending with $A_r = B$ such that A_i is obtained from A_{i-1} by a single row operation. By the Proposition, this means that there is an elementary matrix U_i such that $A_i = U_i A_{i-1}$. This gives

$$A_1 = U_1 A_0 = U_1 A$$
$$A_2 = U_2 A_1 = U_2 U_1 A$$
$$A_3 = \quad A_2 = \quad A$$

and so on. Eventually we get $B = A_r = U_r U_{r-1} \cdots U_1 A$. We can thus take $U = U_r U_{r-1} \cdots U_1$ and we have $B = UA$ as required. □

Invertibility

Theorem 11.5: Let A be an $n \times n$ matrix. Then the following statements are equivalent: if any one of them is true then they are all true, and if any one of them is false then they are all false.

(a) A can be row-reduced to I_n.
(b) The columns of A are linearly independent.
(c) The columns of A span $\mathbb{R}^n$.
(d) The columns of A form a basis for $\mathbb{R}^n$.
(e) A^T can be row-reduced .
(f) The columns of A^T are
(g) The columns of A^T
(h) The columns of A^T
(i) There is a matrix U such that $UA = I_n$.
(j) There is a matrix V such that $AV = I_n$.

Moreover, if these statements are all true then there is a unique matrix U that satisfies $UA = I_n$, and this is also the unique matrix that satisfies $AU = I_n$ (so the matrix V in (j) is necessarily the same as the matrix U in (i)).

Invertibility — what we already know

(a) A can be row-reduced to I_n.
(b) The columns of A are linearly independent.
(c) The columns of A span $\mathbb{R}^n$.
(d) The columns of A form a basis for $\mathbb{R}^n$.
(e),(f),(g),(h): same for A^T
(i) There is a matrix U such that $UA = I_n$.
(j) There is a matrix V such that $AV = I_n$.

Statements (a) to (d) are equivalent to each other by the "numerical criteria" (Proposition 10.12).

Similarly statements (e) to (h) are equivalent to each other.

Moreover, (a) to (d) are equivalent to (e) to (h) by "duality for bases" (Proposition 10.11).

The real issue is to prove that (a) to (h) are equivalent to (i) and (j).

Invertibility

(a) A can be row-reduced to I_n.
(b) The columns of A are linearly independent.
(c) The columns of A span $\mathbb{R}^n$.
(d) The columns of A form a basis for $\mathbb{R}^n$.
(e),(f),(g),(h): same for A^T
(i) There is a matrix U such that $UA = I_n$.
(j) There is a matrix V such that $AV = I_n$.

- If (a) holds then each row operation corresponds to an elementary matrix, and the product of those is a matrix U with $UA = I_n$; so (i) holds.
- Similarly, if (e) holds then there exists W with $WA^T = I_n$, so $AW^T = I_n$, so can take $V = W^T$ to see that (j) holds.
- Conversely, suppose that (i) holds. Let $v_1, \dots, v_r$ be the columns of A. A linear relation $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$ gives a a vector λ with $A\lambda = 0$. As $UA = I_n$ this gives $\lambda = UA\lambda = U0 = 0$, so our linear relation is the trivial one. Thus the columns v_i are linearly independent, so (b) holds.
- Similarly, (j) implies (f).
- Now (a)$\Leftrightarrow \dots \Leftrightarrow$(h) and (a)$\Rightarrow$(i)$\Rightarrow$(b) and (e)$\Rightarrow$(j)$\Rightarrow$(f); so (a) to (j) are all equivalent. □

Invertibility

(a) A can be row-reduced to I_n.
(b) The columns of A are linearly independent.
(c) The columns of A span $\mathbb{R}^n$.
(d) The columns of A form a basis for $\mathbb{R}^n$.
(e),(f),(g),(h): same for A^T
(i) There is a matrix U such that $UA = I_n$.
(j) There is a matrix V such that $AV = I_n$.

Definition 11.6:
We say that A is *invertible* if (any one of) the conditions (a) to (j) hold.
If so, we write A^{-1} for the unique matrix satisfying $A^{-1}A = I_n = AA^{-1}$ (which exists by the Theorem).

Remark 11.7: It is clear that A is invertible if and only if A^T is invertible.

Elementary matrices are invertible

All elementary matrices are invertible. More precisely:

(a) $D_p(\lambda^{-1})D_p(\lambda) = I_n$, so $D_p(\lambda)$ is invertible with inverse   .
For example, when $n = 4$ and $p = 2$ we have

$$D_2(\lambda)D_2(\lambda^{-1}) = \begin{bmatrix}1&0&0&0\\0&\lambda&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}\begin{bmatrix}1&0&0&0\\0&\lambda^{-1}&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix} = \begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix} = I_4.$$

(b) $E_{pq}(\mu)E_{pq}(-\mu) = I_n$, so $E_{pq}(\mu)$ is invertible with inverse   .
For example, when $n = 4$ and $p = 2$ and $q = 4$ we have

$$E_{24}(\mu)E_{24}(-\mu) = \begin{bmatrix}1&0&0&0\\0&1&0&\mu\\0&0&1&0\\0&0&0&1\end{bmatrix}\begin{bmatrix}1&0&0&0\\0&1&0&-\mu\\0&0&1&0\\0&0&0&1\end{bmatrix} = \begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix} = I_4.$$

(c) $F_{pq}^2 = I_n$, so F_{pq} is invertible and is its own inverse. For example, when $n = 4$ and $p = 2$ and $q = 4$ we have

$$F_{24}^2 = \begin{bmatrix}1&0&0&0\\0&0&0&1\\0&0&1&0\\0&1&0&0\end{bmatrix}\begin{bmatrix}1&0&0&0\\0&0&0&1\\0&0&1&0\\0&1&0&0\end{bmatrix} = \begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix} = I_4.$$

Row reduction and invertible matrices

Corollary 11.10: Let A and B be $n \times n$ matrices, and suppose that A can be converted to B by a sequence of row operations. Then $B = UA$ for some invertible matrix U.

Proof.

- Corollary 11.4 tells us that $B = UA$ for some matrix U that is a product of elementary matrices.
- Example 11.8 tells us that elementary matrices are invertible.
- Proposition 11.9 tells us that products of invertible matrices are invertible.
- Thus, U is invertible.

□

Products of invertible matrices are invertible

Proposition 11.9:
If A and B are invertible $n \times n$ matrices, then AB is also invertible, and $(AB)^{-1} = B^{-1}A^{-1}$.

Proof.
Put $C = AB$ and $D = B^{-1}A^{-1}$.

$$DC = B^{-1}A^{-1}AB = B^{-1}I_nB = B^{-1}B = I_n$$

$$CD = ABB^{-1}A = AI_nA^{-1} = AA^{-1} = I_n$$

This shows that D is an inverse for C, so C is invertible with $C^{-1} = D$ as claimed. □

More generally, if $A_1, A_2, \dots, A_r$ are invertible $n \times n$ matrices, then the product $A_1A_2 \cdots A_r$ is also invertible, with

$$(A_1A_2 \cdots A_r)^{-1} = A_r^{-1} \cdots A_2^{-1}A_1^{-1}.$$

The proof is similar.

Finding inverses by row-reduction

To check whether A is invertible, row-reduce it and see whether you get the identity. We can find the inverse by a closely related procedure.

Method 11.11: Let A be an $n \times n$ matrix.

(a) Form the augmented matrix $[A|I_n]$ and row-reduce it.
(b) If the result has the form $[I_n|B]$, then A is invertible with $A^{-1} = B$.
(c) If the result has any other form then A is not invertible.

Proof of correctness.
Let $[T|B]$ be the row-reduction of $[A|I_n]$.
Then T is the row-reduction of A, so A is invertible if and only if $T = I_n$.
Suppose that this holds, so $[A|I_n]$ reduces to $[I_n|B]$. As in Corollary 11.4 we see that there is a matrix U such that $[I_n|B] = U[A|I_n] = [UA|U]$. This gives $B = U$ and $UA = I_n$ so $BA = I_n$, so $B = A^{-1}$. □

Example of finding an inverse

Consider the matrix $A = \begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix}$. We have the following row-reduction:

$$[A|I_3] = \left[\begin{array}{ccc|ccc} 1 & a & b & 1 & 0 & 0 \\ 0 & 1 & c & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & & 1 & -a & 0 \\ 0 & 1 & c & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right]$$

$$\to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -a & \\ 0 & 1 & 0 & 0 & 1 & \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right]$$

We conclude that $A^{-1} = \begin{bmatrix} 1 & & \\ 0 & 1 & \\ 0 & 0 & 1 \end{bmatrix}$.

It is a straightforward exercise to check this directly:

$$\begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -a & ac-b \\ 0 & 1 & -c \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -a & ac-b \\ 0 & 1 & -c \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix}.$$

Example of finding an inverse

Consider the matrix $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{bmatrix}$. We have the following row-reduction:

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 4 & 0 & 1 & 0 \\ 1 & 3 & 9 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & & & \\ 0 & 1 & 3 & & & \\ 0 & 2 & 8 & & & \end{array}\right] \to$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & -2 & 2 & -1 & 0 \\ 0 & 1 & 3 & -1 & 1 & 0 \\ 0 & 0 & 2 & 1 & -2 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & -3 & 1 \\ 0 & 1 & 0 & -5/2 & 4 & -3/2 \\ 0 & 0 & 1 & 1/2 & -1 & 1/2 \end{array}\right]$$

We conclude that

$$A^{-1} = \begin{bmatrix} 3 & -3 & 1 \\ -5/2 & 4 & -3/2 \\ 1/2 & -1 & 1/2 \end{bmatrix}.$$