AE 525a/ME 525 Lecture #1: Linear Equations and Gauss-Jordan Reduction

Consider a set of m linear equations in n unknowns

$$\begin{aligned} \mathsf{Ax} &= \mathsf{c} \\ \mathsf{A} &\in R^{\mathsf{m}\times\mathsf{n}}; \mathsf{x} \in R^{\mathsf{n}\times 1}; \mathsf{c} \in R^{\mathsf{m}\times 1} \end{aligned} \tag{1}$$

or using index notation

$$a_{ik}x_k = c_i; \quad i = 1:m; k = 1:n \tag{2}$$

where summation over repeated indices is understood. Often, this system of equation is defined in terms of its augmented matrix

$$[\mathsf{A}|\mathsf{c}] \tag{3}$$

or

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & c_1 \\ a_{21} & a_{22} & \dots & a_{2n} & c_2 \\ \vdots & & & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & c_m \end{bmatrix} \tag{4}$$

Now for $a_{11} \neq 0$ we divide the first row by a_{11}. Subsequently we multiply the first row by $(-a_{21})$ and subtract it from the second row, multiply the first row by $(-a_{31})$ and subtract it from the third row, etc., thus eliminating x_1 from rows $2:m$. The new system becomes

$$\begin{bmatrix} 1 & a'_{12} & \dots & a'_{1n} & c'_1 \\ 0 & a'_{22} & \dots & a'_{2n} & c'_2 \\ \vdots & & & & \vdots \\ 0 & a'_{m2} & \dots & a'_{mn} & c'_m \end{bmatrix}$$

where the primes denote the changed variables and the unknown x_1 has been eliminated from equations two to m. Now we divide row two by a'_{22} (we can always rearrange the rows $2:m$ so that $a'_{22} \neq 0$). Multiplying the new row two by $(-a'_{12})$ and subtracting it from row one and similarly multiplying the new row two by $(-a'_{32})$ and subtracting it from row three, etc., results in the following

$$\begin{bmatrix} 1 & 0 & a''_{13} & \dots & a''_{1n} & c''_1 \\ 0 & 1 & a''_{23} & \dots & a''_{2n} & c''_2 \\ \vdots & & & & & \vdots \\ 0 & a'_{m2} & a''_{m3} & \dots & a''_{mn} & c''_m \end{bmatrix}$$

where double primes denote that the corresponding variables have been twice updated (changed). We continue this process r times until it terminates, i.e., until $r = m$ or until the coefficients of all $x's$ are zero in all equations following the r-th equation. We call these $m-r$ equations the **residual equations.** When $m > r$, we have that the augmented system becomes

$$\begin{bmatrix} 1 & 0 & 0 & \dots & 0 & \alpha_{1,r+1} & \dots & \alpha_{1,n} & \gamma_1 \\ 0 & 1 & 0 & \dots & 0 & \alpha_{2,r+1} & \dots & \alpha_{2,n} & \gamma_2 \\ 0 & 0 & 1 & \dots & 0 & \alpha_{3,r+1} & \dots & \alpha_{3,n} & \gamma_3 \\ \vdots & & & & \vdots & & & \vdots & \\ 0 & 0 & 0 & \dots & 1 & \alpha_{r,r+1} & \dots & \alpha_{r,n} & \gamma_r \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & c^{(r)}_{r+1} \\ \vdots & & & & \vdots & & & \vdots & \\ 0 & 0 & 0 & & 0 & 0 & \dots & 0 & c^{(r)}_m \end{bmatrix} \tag{5}$$

Therefore, there are two possibilities:

1. With $m > r$, one or more residual equations have a non-zero RHS, i.e., $c^{(r)}_k \neq 0; k = r+1 : m$, then **no solution exists.**

2. When $c^{(r)}_k = 0; k = r+1 : m$, then the set of m equations is reduced to a set of r equations

$$\begin{aligned} x_1 &= \gamma_1 - \alpha_{1,r+1}x_{r+1} - \dots - \alpha_{1,n}x_n \\ x_2 &= \gamma_2 - \alpha_{2,r+1}x_{r+1} - \dots - \alpha_{2,n}x_n \\ &\vdots \\ x_r &= \gamma_r - \alpha_{r,r+1}x_{r+1} - \dots - \alpha_{r,n}x_n \end{aligned} \tag{6}$$

   called a **consistent system.** Thus the most general solution expresses each of the r unknowns $x_1, ..., x_r$ as a linear combination of the remaining $(n-r)$ variables (and a specific constant γ) each of which can be assigned arbitrarily. If $r = n$, we get a unique solution while for $r \neq n$, we get an $(n-r)-$parameter family of solutions. The number $n - r = d$ is called the defect of the system.

Remark 1 *For a consistent system and $r < m$, the residual $(m-r)$ equations are actually ignorable, and hence must be implied by the remaining r equations.*

Example 2 *Consider the following system of equations*

$$\begin{bmatrix} 1 & 2 & -1 & -2 & -1 \\ 2 & 1 & 1 & -1 & 4 \\ 1 & -1 & 2 & 1 & 5 \\ 1 & 3 & -2 & -3 & -3 \end{bmatrix}$$

After Gauss-Jordan reduction the augmented matrix becomes

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 3 \\ 0 & 1 & -1 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

or

$$x_1 = 3 - x_3$$
$$x_2 = -2 + x_3 + x_4$$

which is a two-parameter solution. By choosing, say, $x_3 = C_1$ *and* $x_4 = C_2$ *we get*

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 3 \\ -2 \\ 0 \\ 0 \end{bmatrix} + C_1 \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix} + C_2 \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

0.1 Cramer's Rule

Definition 3 *Let* M_{ij} *denotes a subdeterminant of a square matrix* A *where the* $i-th$ *row and the* $j-th$ *column of* A *are deleted. Then* M_{ij} *is called a minor of* a_{ij}*.*

Definition 4 *The cofactor of* a_{ij} *is defined as*

$$A_{ij} = (-1)^{\underline{i+j}} M_{\underline{ij}} \tag{7}$$

where underlined indices denote no summation over repeated indices.

Remark 5 *The cofactor of* a_{ij} *equals to the coefficient of* a_{ij} *in the expansion of* $\det \mathsf{A}$.

0.1.1 Laplace Expansion Formulas for $\det \mathsf{A}$

Case 6 *$i-th$ Row Expansion*

$$\det \mathsf{A} = a_{\underline{i}k} A_{\underline{i}k} \tag{8}$$

where summation over repeated (not underlined) indices is understood.

Case 7 *$j-th$ Column Expansion*

$$\det \mathsf{A} = a_{k\underline{j}} A_{k\underline{j}} \tag{9}$$

Remark 8 *Replace* a_{ik} *in (8) by* $a_{rk}, (r \neq i)$ *is equivalent of replacing the elements of the* $i-th$ *row with the corresponding elements of the* $r-th$ *row. Since now two rows are the same the determinant of the resulting matrix must be zero, i.e.,*

$$a_{rk} A_{ik} = \delta_{ir} \det \mathsf{A} \tag{10}$$

Similarly, for the column expansion (9) we get

$$a_{ks} A_{kj} = \delta_{js} \det \mathsf{A} \tag{11}$$

If we consider the linear system

$$a_{ik} x_k = c_i; \quad i, k = 1 : n \tag{12}$$

with $\det \mathsf{A} \neq 0$, we get

$$(a_{ik} A_{ir}) x_k = c_i A_{ir}; \quad r = 1 : n$$

which becomes

$$\delta_{kr} \det \mathsf{A} = c_i A_{ir}$$

and thus

$$x_r = \frac{c_i A_{ir}}{\det \mathsf{A}} \tag{13}$$

which is the Cramer's rule.

Remark 9 *For the homogeneous case* $\mathsf{c} = \mathsf{0}$, *one solution is the trivial one* $\mathsf{x} = \mathsf{0}$. *This is the only possible solution if* $\det \mathsf{A} \neq 0$.

Remark 10 *A set of linear homogeneous equations in* n *unknowns cannot possess a nontrivial solution unless* $\det \mathsf{A} = 0$.

0.2 The Inverse Matrix

Recall from Eqns.(10) and (11) that

$$\begin{aligned} a_{ik} A_{jk} &= \delta_{ij} \det \mathsf{A} \\ a_{kj} A_{ki} &= \delta_{ij} \det \mathsf{A} \end{aligned}$$

Let

$$m_{ij} = \frac{A_{ji}}{\det \mathsf{A}}; \det \mathsf{A} \neq 0 \tag{14}$$

then

$$\begin{aligned} a_{ik} m_{kj} &= \delta_{ij} \\ m_{ik} a_{kj} &= \delta_{ij} \end{aligned}$$

or

$$\mathsf{A}\,\mathsf{M} = \mathsf{I} = \mathsf{M}\,\mathsf{I}$$

Therefore $\mathsf{M} = \mathsf{A}^{-1}$.

Let

$$adj(\mathsf{A}) = [A_{ji}] \tag{15}$$

then

$$\mathsf{A}^{-1} = \frac{adj\,\mathsf{A}}{\det \mathsf{A}} \tag{16}$$

Remark 11 *For a system* $\mathsf{A}\mathsf{x} = \mathsf{c}$, *we have that* $\mathsf{x} = \mathsf{A}^{-1}\mathsf{c}$ *or*

$$x_i = \frac{1}{\det \mathsf{A}}(A_{1i}c_1 + A_{2i}c_2 + ... + A_{ni}c_n)$$
$$= \frac{c_k A_{ki}}{\det \mathsf{A}}$$

which is the Cramer's rule.

Example 12

$$\mathsf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 1 & -1 & 3 \end{bmatrix}; \det \mathsf{A} = 6$$
$$adj\,\mathsf{A} = \begin{bmatrix} 5 & -9 & 1 \\ 2 & 0 & -2 \\ -1 & 3 & 1 \end{bmatrix}$$
$$\mathsf{A}^{-1} = \frac{1}{6}\begin{bmatrix} 5 & -9 & 1 \\ 2 & 0 & -2 \\ -1 & 3 & 1 \end{bmatrix}$$

0.3 Rank of a Matrix

Definition 13 *Rank of any matrix* A *is the order of the largest square submatrix of A (formed by deleting certain rows and/or columns of* A*) whose determinant does not vanish.*

Example 14

$$rank\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix} = 1$$

Claim 15 *Suppose that* A *is of rank* r. *we shall demonstrate that if set of* r *rows of* A, *containing a nonsingular* $r \times r$ *submatrix* R *is selected, then, any other row of* A *is a linear combination of those* r *rows.*

Proof. Suppose that R is the upper left corner of A. Then, consider the following submatrix of A

$$\mathsf{M} = \begin{bmatrix} a_{11} & a_{12} & ... & a_{1r} & a_{1s} \\ \vdots & & & & \\ a_{r1} & a_{r2} & ... & a_{rr} & a_{rs} \\ a_{q1} & a_{q2} & ... & a_{qr} & a_{qs} \end{bmatrix}; q > r; s > r \tag{17}$$

Since $\det \mathsf{R} \neq 0$, the rows of R are linearly independent. Furthermore, $rank(\mathsf{A}) = r$ implies that $\det \mathsf{M} = 0$ for all $q > r$ and $s > r$. But then the portion of the

$q-th$ row $a_{q1},...,a_{qr}$ can be expressed in terms of the rows of R. Let the rows of R be denoted by $\mathsf{a}_1,...,\mathsf{a}_r$. Thus

$$\begin{aligned} [a_{q1},...,a_{qr}]^{\mathsf{T}} &= \lambda_1 \mathsf{a}_1^{\mathsf{T}} + ... + \lambda_r \mathsf{a}_r^{\mathsf{T}} = \mathsf{R}^{\mathsf{T}} \boldsymbol{\lambda} \\ \boldsymbol{\lambda} &= \end{aligned} \tag{18}$$

But $\det \mathsf{R} \neq 0$ implies that there exists $\boldsymbol{\lambda} \neq 0$, or that the $(r+1)$st row is expressed as a linear combination of the r rows of R. Similar arguments apply for the columns. We should note that we still need to take care of the element a_{qs}. Suppose that the linear combination

$$\lambda_1 a_{1s} + \lambda_2 a_{2s} + ... + \lambda_r a_{rs} = a'_{qs}$$

Recall that in evaluating $\det \mathsf{M}$ we can subtract the linear combination of the first r rows (as stated by (18)) from the last row, without changing $\det \mathsf{M}$. Therefore, we get

$$\det \mathsf{M} = (a_{qs} - a'_{qs}) \det \mathsf{R} = 0$$

and thus

$$a'_{qs} = a_{qs}$$

which means that the entire last row of M is a linear combination of the first r rows. ■

Since the above is true for any $q, s > r$, we can state the following:

- If a matrix A is of rank r, and a set of r rows containing the nonsingular submatrix R of order r is selected, then any other row in the matrix is a linear combination of these r rows.
- The same statement is true if the word "row" is replaced by "column" throughout.
- Special Case: If a square matrix A is singular, then at least one of the rows of A must be a linear combination of the others. (The same applies to columns.)

Since the converse was already established earlier we have proved the following theorem:

Theorem 16 *A square matrix is singular if and only if one of its rows (columns) is a linear combination of the others.*

It is obvious that by interchanging rows and columns does not affect the rank of a matrix. Therefore,

$$rank \mathsf{A} = rank \mathsf{A}^{\mathsf{T}}$$

Next we examine certain operations on matrices which leave the rank invariant.

0.4 Elementary Operations

Consider a system of equations $\mathsf{Ax} = \mathsf{c}$, where $\mathsf{A} \in R^{\mathsf{m}\times\mathsf{n}}$; $\mathsf{x} \in R^{\mathsf{n}\times 1}$; $\mathsf{c} \in R^{\mathsf{m}\times 1}$, with an augmented matrix $[\mathsf{A}|\mathsf{c}] \in R^{\mathsf{m}\times(\mathsf{n}+1)}$. if we use Gauss-Jordan reduction, we get the system

$$
\begin{aligned}
x_1 - \alpha_{1,r+1}x_{r+1} - ... - \alpha_{1,n} &= \gamma_1 \\
x_2 - \alpha_{2,r+1}x_{r+1} - ... - \alpha_{2,n} &= \gamma_2 \\
&\vdots \\
x_1 r - \alpha_{r,r+1}x_{r+1} - ... - \alpha_{r,n} &= \gamma_r \\
0 &= \gamma_{r+1} \\
&\vdots \\
0 &= \gamma_m
\end{aligned} \tag{19}
$$

Note, we may renumber the equations (partial pivoting). Thus the augmented matrix after Gauss-Jordan reduction becomes

$$
\begin{bmatrix}
1 & 0 & 0 & ... & 0 & \alpha_{1,r+1} & ... & \alpha_{1,n} & \gamma_1 \\
0 & 1 & 0 & ... & 0 & \alpha_{2,r+1} & ... & \alpha_{2,n} & \gamma_2 \\
0 & 0 & 1 & ... & 0 & \alpha_{3,r+1} & ... & \alpha_{3,n} & \gamma_3 \\
\vdots & & & & \vdots & & & \vdots & \\
0 & 0 & 0 & ... & 1 & \alpha_{r,r+1} & ... & \alpha_{r,n} & \gamma_r \\
0 & 0 & 0 & ... & 0 & 0 & ... & 0 & \gamma_{r+1} \\
\vdots & & & & \vdots & & & \vdots & \\
0 & 0 & 0 & & 0 & 0 & ... & 0 & \gamma_m
\end{bmatrix} \tag{20}
$$

The steps in Gauss-Jordan reduction involve the following so called elementary operations:

1. Interchange of two rows (columns),
2. multiplication of the elements of a row (column) by a nonzero number, and
3. addition of the elements of one row (column) k times the corresponding elements of another row (column).

If $\gamma_{r+1} = ... = \gamma_m = 0$, then $rank[\mathsf{A}|\mathsf{c}] = r$. However, if one of γ_{r+1} to γ_m are different from zero we have that $rank[\mathsf{A}|\mathsf{c}] = r + 1$.

Claim 17 *The rank of the augmented matrix* $[\mathsf{A}|\mathsf{c}]$ *associated with the system* $\mathsf{Ax} = \mathsf{c}$, *is the same as the rank of the corresponding augmented matrix* $[\mathsf{A}'|\mathsf{c}']$ *obtained after Gauss-Jordan reduction since the rank of a matrix is not changed by elementary operations.*

Proof. From the properties of determinants it is obvious that elementary operations (1) and (2) do not change the rank of a matrix. To show that the elementary operation (3) does not change the rank of a matrix we proceed as follows. Suppose that the nonsingular matrix $\mathsf{R} \in R^{r \times r}$ is in the upper left corner of A and that k times the qth row is to be added to the ith row, where $i \leq r < q$. If we consider a submatrix S

$$\mathsf{S} = \begin{bmatrix} & \mathsf{R} & \\ a_{q1} & \ldots & a_{qr} \end{bmatrix}$$

We know that $rank\mathsf{S} = rank\mathsf{R} = r$, so that the last row of S is a linear combination of the rows of R. Thus by adding a linear combination of the rows of R to the ith row of R does not change the $\det \mathsf{R}$ and thus the rank remains unchanged. ■

0.5 Solvability of Sets of Linear Equations

We have shown that since Gauss-Jordan reduction involves only the elementary (row) operations we have that both augmented matrices of the system $\mathsf{Ax} = \mathsf{c}$ and the system obtained after the Gauss-Jordan reduction are of equal rank. The same is true of the coefficient matrices.

We have also shown that if one or more numbers $\gamma_{r+1}, ..., \gamma_m$ is nonzero, then the set of equations $\mathsf{Ax} = \mathsf{c}$ has no solution. In terms of the rank we can rephrase that statement as follows.

Claim 18 *If and only if the rank of augmented matrix is greater than the rank of the coefficient matrix, i.e.,* $rank[\mathsf{A}|\mathsf{c}] > rank\mathsf{A}$, *the system* $\mathsf{Ax} = \mathsf{c}$ *has no solution.*

Therefore, we have proved the following theorem.

Theorem 19 *A set of linear equations* $\mathsf{Ax} = \mathsf{c}$ *possesses a solution if and only if* $rank[\mathsf{A}|\mathsf{c}] = rank\mathsf{A}$.

- If the two ranks are both equal to r, and if we select the set of r equations whose coefficient matrix contains a nonsingular submatrix of order r, then we may ignore all the other equations, since they are implied by the r basic equations. The $n - r$ unknowns whose coefficients are not involved in the nonsingular $r \times r$ submatrix can be assigned **arbitrary values**, after which the remaining r unknowns can be determined in terms of them.
  - If $n = r$, the unknowns are uniquely determined.
  - If $n > r$, the most general solution involves $n - r$ independent arbitrary constants.

Example 20 *Consider the system of equations*

$$\begin{aligned} x_1 + 2x_2 &= b_1 \\ 2x_1 + 4x_2 &= b_2 \\ 3x_1 + 6x_2 &= b_3 \end{aligned} \tag{21}$$

Then

$$\begin{aligned} [\mathsf{A}|\mathsf{c}] &= \begin{bmatrix} 1 & 2 & b_1 \\ 2 & 4 & b_2 \\ 3 & 6 & b_3 \end{bmatrix} \\ rank\mathsf{A} &= 1 \end{aligned}$$

If $\mathsf{b}^\mathsf{T} = [1, 0, 0]$ *we have that* $rank[\mathsf{A}|\mathsf{c}] = 2 > rank\mathsf{A}$, *and there is no solution. If* $\mathsf{b}^\mathsf{T} = [5, 10, 15]$ *we have that* $rank[\mathsf{A}|\mathsf{c}] = 1 = rank\mathsf{A}$, *and there exists a solution. In fact any vector of the form*

$$\mathsf{x} = \begin{bmatrix} 5 \\ 0 \end{bmatrix} + \alpha \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

with α *arbitrary, is a solution.*

If $\mathsf{c} = \mathsf{0}$ we have a homogeneous system of equations $\mathsf{A}\mathsf{x} = \mathsf{0}$. In that case we have always that $rank[\mathsf{A}|\mathsf{0}] = rank\mathsf{A}$, and thus solution exists (i.e., $\mathsf{x} = \mathsf{0}$). However, if $rank\mathsf{A} = n$, then $\mathsf{x} = \mathsf{0}$ is the only solution (recall if $\det \mathsf{A} \neq 0$, $\mathsf{x} = \mathsf{0}$ is the only solution).

If $rank\mathsf{A} = r < n$, then infinitely many solutions exist with $n - r$ number of independent variables.

Case 21 $\mathsf{A}\mathsf{x} = \mathsf{0}$; $\mathsf{A} \in R^{\mathsf{n}\times\mathsf{n}}$ *and* $rank\mathsf{A} = n - 1$. *In that case* $\det \mathsf{A} = 0$ *and we have that*

$$a_{\mathsf{ik}} x_{\mathsf{k}} = 0; k, i = 1 : n \tag{22}$$

Recall now that

$$a_{\mathsf{ik}} A_{\mathsf{jk}} = \delta_{\mathsf{ij}} \det \mathsf{A}; \mathsf{i}, \mathsf{j}, \mathsf{k} = 1 : \mathsf{n}$$

But since $\det \mathsf{A} = 0$ *we have*

$$a_{\mathsf{ik}} A_{\mathsf{jk}} = 0 \tag{23}$$

By comparing Eqns.(22) and (23) we get

$$x_{\mathsf{i}} = C A_{\mathsf{si}}; s, i = 1 : n \tag{24}$$

where C *is arbitrary. Since* $n - r = 1$ *and since (24) contains only one parameter, then (24) must represent the most general solution of (22) (unless for a particular* s *all coefficients of* $A_{\mathsf{si}} = 0$, *which cannot happen for all values of* s *since* $rank\mathsf{A} = n-1$). *Therefore, the unknowns are proportional to the cofactors of their coefficients in any row of* A.