Lecture II
Linear Equation Theory

Back to the motivating problem:

Solving m equations for n unknowns:

$$\left.\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned}\right\} \qquad Ax = b.$$

Both *n* and *m* are large

When does a solution exist? When unique?

Three Cases

- **Case 1:** The same number of unknowns as the number of equations ($n = m$)

Extensions:

- **Case 2:** More unknowns ($n > m$)

- **Case 3:** Less unknowns ($n < m$)

A Basic Result for Case 1

If the square matrix A is nonsingular, i.e. $\det A \neq 0,$ then the linear equation $Ax = b$ has the unique solution

$$x = A^{-1}b.$$

Recall that the inverse A^{-1} of a nonsingular matrix A is defined as

$$A^{-1}A = AA^{-1} = I.$$

About the Matrix Inverse

- If a (square) matrix A is nonsingular, then its inverse is unique. That is,

$$\boxed{\begin{aligned} AB = I &\Leftrightarrow B = A^{-1} \\ BA = I &\Leftrightarrow B = A^{-1}. \end{aligned}}$$

Computation of the Matrix Inverse

- The inverse of a nonsingular matrix *A* is defined as

$$A^{-1} = (\det A)^{-1} (\operatorname{cof} A)^T ,$$

where cof A is the cofactor matrix of A:

$$\operatorname{cof} A \doteq \left[(-1)^{i+j} \det A_{ij} \right]_{n \times n} ,$$

A_{ij} = the matrix of order $n-1$, after deleting row i and column j from A.

Proof of $\boxed{A^{-1} = (\det A)^{-1} (\operatorname{cof} A)^T}$

Use the row and column expansions of $\det A$.

An Example

Solve the linear equation, *using the above basic result*:

$$\begin{pmatrix} 1 & 2 \\ 4 & 4 \end{pmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$

Another Computational Method: Cramer's Rule

For any $n \times n$ nonsingular matrix $A = (a_{ij})$,

the linear equation $Ax = b$ has the unique solution:

$$x_j = \frac{\Delta_j}{\det A}, \quad j = 1,\ 2,\ \ldots,\ n$$

where Δ_j is the determinant of the matrix formed by replacing the j-th column of A by b. For example,

$$\Delta_1 = \det\begin{pmatrix} b_1 & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ b_n & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \text{ etc}$$

Proof of Cramer's Rule

Consider the solution

$$x = A^{-1}b = [\text{cof}\, A]^T b \,/\, \det A.$$

$$So, \quad x_j = \left(\det A\right)^{-1} \sum_{i=1}^{n} (-1)^{i+j} (\det A_{ij}) b_i$$

$$= \left(\det A\right)^{-1} \Delta_j$$

because, by the j-column expansion of Δ_j,

$$\Delta_j = \sum_{i=1}^{n} (-1)^{i+j} (\det A_{ij}) b_i .$$

An Example

Solve the linear equation, *using Cramer's Rule*:

$$\begin{pmatrix} 1 & 2 \\ 4 & 4 \end{pmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$

Question

When is matrix *A* invertible?

A Necessary and Sufficient Condition

The linear equation $Ax = b$ is solvable for every b, if and only if $\det A \neq 0$.

Proof

The sufficiency is proved above.

For the necessity, take the basis vectors:

$$e^1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad e^2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots, \quad e^n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

Then, for each $1 \le i \le n$, $Ax = e^i$ has solution x^i.

$So,\ AX = I$, with $X \doteq [x^1, \ldots, x^n]$,

implying $\det A \neq 0$, because $\det(A)\det(X) = 1$.

Comments

- In the proof, the following important fact was used:

$$\det(AB) = \det A \bullet \det B$$

$$\text{for any } n \times n \text{ matrices } A \text{ and } B.$$

If $\det A = 0$, *then*

- for some vectors b, $Ax = b$ has no solution;
- for other vectors b, *the equation may have an infinite number of solutions*!

Homogenous Equations (b=0)

Question:

When does a general homogenous equation

$$Ax = 0$$

have a *nonzero* solution $x \neq 0$?

In other words, when are the column vectors of *A* linearly dependent?

Review of Terminologies

- **Linear combination of vectors:**

$$\sum_{i=1}^{n} \alpha_i x_i \text{ is a linear combination of vectors } x_1, \ \ldots, \ x_n.$$

- **Linear independency:**

$$\sum_{i=1}^{n} \alpha_i x_i = 0 \Rightarrow \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0.$$

Review of Terminologies

- **Linear dependency:**

$$\boxed{\sum_{i=1}^{n} \alpha_i x_i = 0 \Rightarrow \exists \alpha_j \neq 0 \text{ for at least one } j.}$$

Review of A Basic Result

The vectors $x_1,\ x_2,\ \ldots,\ x_n$ are dependent **if and only if** one of the vectors is some linear combination of the other vectors. That is, $\exists j$ and constants α_i so that

$$x_j = \sum_{i \neq j} \alpha_i x_i .$$

Examples Revisited

Are the following vectors linearly dependent or independent?

1) Consider the vectors

$$x = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad y = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

2) Consider the vectors

$$x = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad y = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \quad z = \begin{bmatrix} 5 \\ 8 \end{bmatrix}$$

Homogenous Equations

A general homogenous equation

$$Ax = 0 \ , \quad A \in \mathbb{R}^{n \times n}$$

has a nonzero solution $x \neq 0$.

$\Leftrightarrow$

$$\det A = 0.$$

Sketch of Proof

Let $a^1,\ a^2,\ \ldots,\ a^n$ be the columns of A.

So, we can rewrite Ax as

$$Ax = x_1 a^1 + x_2 a^2 + \cdots + x_n a^n$$

Thus, $Ax = 0$ has a nonzero solution iff the columns of A are dependent.

Using Fact 4 of determinants, it follows that $\det A = 0$.

Case 2: More Unknowns

In this case, consider

$$Ax = 0$$

for *nonsquare* $A \in \mathbb{R}^{m\times n},$ with $m < n.$

A Fundamental Result

The linear homogeneous equation with more unknowns,

$$Ax = 0,\ A \in \mathbb{R}^{m\times n},\ m < n$$

always has a solution $x \neq 0 \in \mathbb{R}^n$.

Sketch of Proof ($m < n$)

Lemma: If p linearly independent vectors $\{x_i\}_{i=1}^{p}$ are linear combination of q vectors $\{y_j\}_{j=1}^{q}$, i.e.,

$$x_i = \sum_{j=1}^{q} \alpha_{ij} y_j, \quad 1 \le i \le p$$

then, $q \ge p$.

Sketch of Proof (m<n)

By means of this lemma, the columns $\left\{a^i\right\}_{i=1}^{n}$ of A in $\mathbb{R}^m$ $(m<n)$ must be linearly dependent.

$$Thus,\ Ax=0 \Rightarrow x_1a^1+\cdots+x_na^n=0$$

$$\Rightarrow \exists\, x=\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \neq \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

End of Proof

Numerical Example

Find <u>all</u> nonzero solutions for

$$x_1 + 2x_2 + 3x_3 = 0,$$

$$x_1 + 9x_2 + 28x_3 = 0.$$

Comment

The set of solutions to Ax=0 is called null space of $A \in \mathbb{R}^{m \times n}$, and often denoted as null(A):

$$null(A) = \left\{ x \in \mathbb{R}^n : Ax = 0 \right\}$$

It is easy to show that null(A) is a linear vector space with dimension less than or equal to n.

Question: What is the dimension of this null space?

Case 3: Fewer Unknowns

In this case, consider

$$Ax = 0$$

for nonsquare $A \in \mathbb{R}^{m\times n}$, with $m > n$.

A Fundamental Result

The homogeneous equation with fewer unknowns,

$Ax = 0,\ A \in \mathbb{R}^{m\times n},\ m > n$

has a solution $x \neq 0 \in \mathbb{R}^n$

$\Leftrightarrow$

every $n \times n$ determinant formed from n rows of A be zero. In other words, $\text{rank}(A) < n$.

Sketch of Proof (m>n)

- Necessity: If one $n \times n$ submatrix A_1 of A is nonsingular, we can rearrange A so that

$$A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}, \quad \text{with } A_1 \in \mathbb{R}^{n\times n},\ A_2 \in \mathbb{R}^{(m-n)\times n}$$

Then, $Ax = 0$ implies $A_1 x = 0$ and thus $x = 0$.

• Sufficiency: Assume now all $n \times n$ submatrices of A are singular. Let $r < n$ be the largest number of rows of A that are linearly independent, i.e., rank(A). Let's decompose A into

$$A = \begin{bmatrix} B \\ C \end{bmatrix}, \text{ with } B \in \mathbb{R}^{r \times n},\ C \in \mathbb{R}^{(m-r)\times n},$$

and the r rows of B are linearly independent.

Clearly, $Bx = 0$ has a nonzero solution $x \neq 0$, *which* is also solution to $Cx = 0$, because each row of C is linear combination of the rows of B.

Case 2

Corollary

The dimension of the null space of $A \in \mathbb{R}^{m\times n}$ is $n - rank(A) := n - r$. That is,

$$\dim\left\{x \in \mathbb{R}^n :\ Ax = 0\right\} = n - rank(A).$$

Remark:

$$\boxed{\mathbb{R}^n = N(A) \oplus R(A^T)}$$

To prove $\dim\left\{x \in \mathbb{R}^n : \ Ax = 0\right\} = n - r$,

let us decompose $A = \begin{bmatrix} B \\ C \end{bmatrix}$, with the r rows of $B \in \mathbb{R}^{r \times n}$

linearly independent. Thus, $Ax = 0 \Leftrightarrow \ Bx = 0$

Rearrange B so that $\begin{bmatrix} B_1 & B_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$, with $\det B_1 \neq 0$.

with $B_1 \in \mathbb{R}^{r \times r}$, $B_2 \in \mathbb{R}^{r \times (n-r)}$, $x_1 \in \mathbb{R}^r$, $x_2 \in \mathbb{R}^{n-r}$.

Then, $B_1 x_1 + B_2 x_2 = 0$, or equivalently,

$x_1 = -B_1^{-1} B_2 x_2$, $\ x_2$ free parameters

$\Rightarrow \quad x = \begin{bmatrix} -B_1^{-1} B_2 \\ I_{(n-r)\times(n-r)} \end{bmatrix} x_2$, which completes the proof.

Inhomogeneous Equations

Given $A=\left(a_{ij}\right)_{m\times n}$ and $b=\left(b_i\right)_{m\times 1}$, solve x for

$$\left.\begin{aligned} a_{11}x_1+\cdots+a_{1n}x_n &= b_1\\ a_{21}x_1+\cdots+a_{2n}x_n &= b_2\\ &\vdots\\ a_{m1}x_1+\cdots+a_{mn}x_n &= b_m \end{aligned}\right\}\quad Ax=b.$$

A Fundamental Result

Consider $Ax = b$.

- It has a solution $x \in \mathbb{R}^n$ if and only if $\text{rank}A = \text{rank}B,\ \text{for } B \doteq \begin{pmatrix} A & b \end{pmatrix} \in \mathbb{R}^{m\times(n+1)}$.
- When $\text{rank}A = \text{rank}B$, all the solutions x take the form:

$$x \doteq x_p + x_h$$

where x_p = any particular solution of $Ax = b$;

x_h = solutions to the homogeneous eq. $Ax = 0$.

Proof of the Main Theorem

1) As seen previously, $Ax = b$ can be rewritten as:

$$x_1 a^1 + x_2 a^2 + \cdots + x_n a^n = b.$$

When $rank(A) = rank(B)$, b is linear combination of the columns $\left\{a^i\right\}_{i=1}^{n}$ of A, so the above eq. has a solution.

The converse is also true.

Proof of the Main Theorem

2) For any general solution x of $Ax = b$ and for any special solution x_p of $Ax = b$, it is easily seen that

$$A\left(x - x_p\right) = 0.$$

So, $x - x_p \in N(A)$, i.e., $x - x_p = x_h$, or equivalently

$$x \doteq x_p + x_h.$$

Comments

- Unlike the homogeneous case, an inhomogeneous equation may have no solution (trivial or nontrivial), because of the rank condition.

- When it has one solution x_p, then it may have an infinite number of solutions.

Example 1

The following inhomogeneous equation

$$x_1 + 2x_2 = 1$$

$$2x_1 + 4x_2 = 0$$

$$3x_1 + 6x_2 = 0$$

has no solution $x \in \mathbb{R}^2$.

Example 2

The following inhomogeneous equation

$$x_1 + 2x_2 = 5$$
$$2x_1 + 4x_2 = 10$$
$$3x_1 + 6x_2 = 15$$

has an infinite number of solutions

$$x = x_p + x_h$$
$$= \begin{bmatrix} 5 \\ 0 \end{bmatrix} + \lambda \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \qquad \lambda \in \mathbb{R}.$$

Application to an Optimization Problem

Given m (noisy) observations $b_1, \ldots, b_m$, and (experimental) variables $a_i = (a_{i1}, \ldots, a_{in})$, find the best possible values $x_0,\ x_1,\ \ldots,\ x_n$ to match

$$b_i = x_0 + x_1 a_{i1} + \cdots + x_n a_{in} \ , \quad 1 \le i \le m.$$

Or, equivalently, to minimize

$$P = \sum_{i=1}^{m} \left(b_i - x_0 - x_1 a_{i1} - \cdots - x_n a_{in} \right)^2 .$$

Necessary Condition

A solution $x = \left(x_0 \ x_1 \ \ldots \ x_n \right)$ to the (nonlinear) optimization problem is often called "least-squares solution".

It must satisfy the 1st-order necessary conditions:

$$\frac{\partial P}{\partial x_j} = 0 \ , \quad j = 0, 1, \ldots, n$$

$$\Leftrightarrow \quad \sum_{i=1}^{m} a_{ij} \left(b_i - x_0 - x_1 a_{i1} - \cdots - x_n a_{in} \right) = 0,$$

$$with \quad a_{i0} = 1.$$

Normal Equation

The necessary conditions can be written in compact matrix form:

$$A^T A x = A^T b \qquad \text{normal equation}$$

where

$$A = \begin{pmatrix} 1 & a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_{m1} & \cdots & a_{mn} \end{pmatrix} \in \mathbb{R}^{m\times(n+1)}, \quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}.$$

Comment

It is interesting to note that finding an (optimal) least-squares solution *x boils down to solving the inhomogeneous normal equation!*

Sufficiency

A solution x to the normal equation $A^T Ax = A^T b$ *does* minimize the sum of squares, P.

Indeed, for any other vector $y := x + z$,

$$\|Ay - b\|^2 = \|(Ax - b) + Az\|^2$$

$$= \|Ax - b\|^2 + 2(Az)^T (Ax - b) + \|Az\|^2$$

$$= \|Ax - b\|^2 + \|Az\|^2$$

$$\geq \|Ax - b\|^2 .$$

Further Comments

1) If $\det\left(A^T A\right) \neq 0,$ *i.e.*, $A^T A \in \mathbb{R}^{(n+1)\times(n+1)}$ is nonsingular, then the least-squares solution x to the best linear fit problem is unique.

2) If $\det\left(A^T A\right) = 0,$ many possible best fits; because

$A^T Az = 0$ has infinitely many nontrivial solutions $z \neq 0$,

$thus, z^T A^T Az = \left\|Az\right\|^2 = 0,$ for many $z \neq 0$.

An Example

Find the best linear fit $b = x_0 col(1) + a^1 x_1 + a^2 x_2$

for the data

$$b = \begin{bmatrix} 1 \\ 2 \\ 0 \\ -1 \end{bmatrix}, \quad a^1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \quad a^2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ -1 \end{bmatrix}.$$

Solution

First, note that there is no (exact) solution to the linear equation *Ax*=*b*.

However, there is a unique (least-squares) best linear fit:

$$b \cong \frac{17}{6}\operatorname{col}(1,\ldots,1) - \frac{13}{6}a^1 - \frac{2}{3}a^2.$$

Remark

If you want to know more about optimization, it is a good idea to take the sequence class ECE-GY 6233 "Systems Optimization Methods".

Homework #2

1. Consider the matrix

$$A = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}.$$

What is the null space of A? What is the rank of A? What is the dimension of the null space?

Homework #2

2. For any pair of $n \times n$ matrices $A,\ B$, show that $\det(AB) = \det(BA) = \det A \det B.$

3. Give some simple examples to show that $AB \neq BA.$

Homework #2

4. Consider linear equations of the form

$$x_1 + 2x_2 + 3x_3 + 4x_4 = 0,$$
$$2x_1 + 4x_2 + \lambda_1 x_3 + \lambda_2 x_4 = 0.$$

What is the range of parameters (λ_1, λ_2) for which the equations have nonzero solutions?

Also, find all nonzero solutions.