ROB 101 - Fall 2021

A Recap: Chapters 1-10

November 1 , 2021

UNIVERSITY OF MICHIGAN

A system of linear equations can have

- No solution;
- Unique solution (one and only one solution);
- Infinite number of solutions.

Linear Systems of Equations: Case I

$$x + y = 4$$
$$2x - y = -1$$

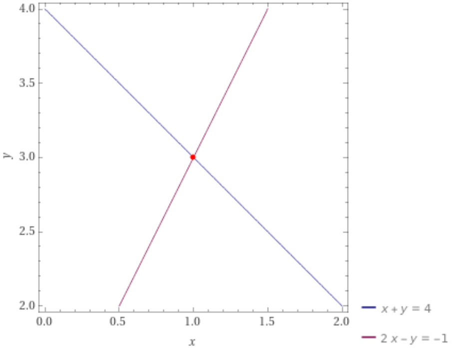


Figure: Graphical solution. Unique solution $\Leftrightarrow$ intersecting lines!

Linear Systems of Equations: Case II

$$x - y = 1$$
$$2x - 2y = -1$$

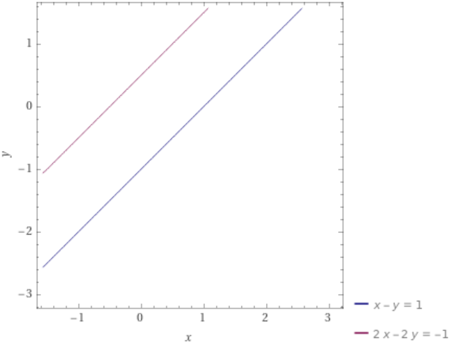


Figure: No solution $\Leftrightarrow$ parallel lines!

Linear Systems of Equations: Case III

$$\begin{aligned} x - y &= 1 \\ 2x - 2y &= 2 \end{aligned}$$

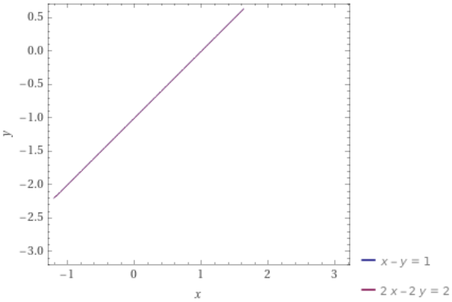


Figure: Infinite number of solutions ⇔ coincident lines!

We now write a general system of linear equations as

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad &= \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned}$$

We can write this system as:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix},$$

where:

$$A := \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, x := \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, b := \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

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

- All terms above the diagonal of the matrix A are zero.
- More precisely, the condition is $a_{ij} = 0$ for all $j > i$.
- Such matrices are called *lower-triangular*.

Lower Triangular Systems and Forward Substitution

We will solve this example using a method called *forward substitution*.

$$
\begin{aligned}
3x_1 &= 6 \\
2x_1 - x_2 &= -2 \\
x_1 - 2x_2 + 3x_3 &= 2
\end{aligned}
\iff
\underbrace{\begin{bmatrix} 3 & 0 & 0 \\ 2 & -1 & 0 \\ 1 & -2 & 3 \end{bmatrix}}_{A}
\underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}_{x}
=
\underbrace{\begin{bmatrix} 6 \\ -2 \\ 2 \end{bmatrix}}_{b}.
$$

Upper Triangular Matrices

$$A = \begin{bmatrix} 1 & 3 & 2 \\ \mathbf{0} & 2 & 1 \\ \mathbf{0} & \mathbf{0} & 3 \end{bmatrix}$$

- All terms below the diagonal of the matrix A are zero.
- More precisely, the condition is $a_{ij} = 0$ for $i > j$.
- Such matrices are called *upper-triangular*.

We solve the upper triangular systems using a method called *back substitution*.

$$\begin{aligned} x_1 + 3x_2 + 2x_3 &= 6 \\ 2x_2 + x_3 &= -2 \\ 3x_3 &= 4, \end{aligned} \quad \Longleftrightarrow \quad \underbrace{\begin{bmatrix} 1 & 3 & 2 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} 6 \\ -2 \\ 4 \end{bmatrix}}_{b}.$$

General Case: Standard Matrix Multiplication

$$A \cdot B := \begin{bmatrix} \boxed{a_{11} \quad a_{12} \quad \cdots \quad a_{1k}} \\ \boxed{a_{21} \quad a_{22} \quad \cdots \quad a_{2k}} \\ \vdots \\ \boxed{a_{n1} \quad a_{n2} \quad \cdots \quad a_{nk}} \end{bmatrix} \cdot \begin{bmatrix} \boxed{\begin{matrix} b_{11} \\ b_{21} \\ \vdots \\ b_{k1} \end{matrix}} & \boxed{\begin{matrix} b_{12} \\ b_{22} \\ \vdots \\ b_{k2} \end{matrix}} & \cdots & \boxed{\begin{matrix} b_{1m} \\ b_{2m} \\ \vdots \\ b_{km} \end{matrix}} \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{i=1}^{k} a_{1i}b_{i1} & \sum_{i=1}^{k} a_{1i}b_{i2} & \cdots & \sum_{i=1}^{k} a_{1i}b_{im} \\ \sum_{i=1}^{k} a_{2i}b_{i1} & \sum_{i=1}^{k} a_{2i}b_{i2} & \cdots & \sum_{i=1}^{k} a_{2i}b_{im} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{k} a_{ni}b_{i1} & \sum_{i=1}^{k} a_{ni}b_{i2} & \cdots & \sum_{i=1}^{k} a_{ni}b_{im} \end{bmatrix}.$$

Matrix Multiplication in the Form of Columns Times Rows

Suppose that A is $n \times k$ and B is $k \times m$ so that the two matrices are compatible for matrix multiplication.

Then

$$A \cdot B = \sum_{i=1}^{k} a_i^{\mathrm{col}} \cdot b_i^{\mathrm{row}},$$

the sum of the columns of A multiplied by the rows of B.

Form the matrix product of $A = \begin{bmatrix} 1 & 0 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 5 & 2 \\ 0 & -1 \end{bmatrix}$.

$$A = \begin{bmatrix} 1 & 0 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} a_1^{\text{col}} & a_2^{\text{col}} \end{bmatrix} = \left[\boxed{\begin{matrix} 1 \\ 3 \end{matrix}} \quad \boxed{\begin{matrix} 0 \\ 4 \end{matrix}} \right]$$

and

$$B = \begin{bmatrix} 5 & 2 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} b_1^{\text{row}} \\ b_2^{\text{row}} \end{bmatrix} = \left[\begin{matrix} \boxed{5 \qquad 2} \\ \boxed{0 \quad -1} \end{matrix} \right].$$

Example

$$a_1^{\text{col}} \cdot b_1^{\text{row}} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 5 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 2 \\ 15 & 6 \end{bmatrix}$$

$$a_2^{\text{col}} \cdot b_2^{\text{row}} = \begin{bmatrix} 0 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & -4 \end{bmatrix}$$

$$\begin{aligned} A \cdot B &= \sum_{i=1}^{2} a_i^{\text{col}} \cdot b_i^{\text{row}} = a_1^{\text{col}} \cdot b_1^{\text{row}} + a_2^{\text{col}} \cdot b_2^{\text{row}} \\ &= \begin{bmatrix} 5 & 2 \\ 15 & 6 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & -4 \end{bmatrix} = \begin{bmatrix} 5 & 2 \\ 15 & 2 \end{bmatrix}. \end{aligned}$$

LU Factorization: Peeling the Onion

$$M = C_1 \cdot R_1 + C_2 \cdot R_2 + C_3 \cdot R_3 = \left[\begin{array}{ccc} C_1 & C_2 & C_3 \end{array}\right] \cdot \left[\begin{array}{c} R_1 \\ R_2 \\ R_3 \end{array}\right]$$

- L is *lower triangular*.

$$L := \begin{bmatrix} C_1 & C_2 & C_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{bmatrix}.$$

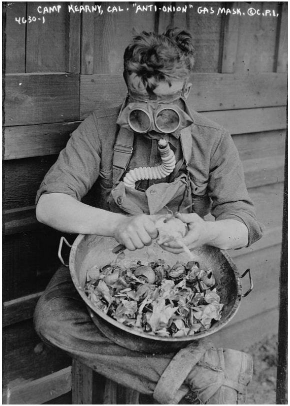
- U is *upper triangular*.

$$U := \begin{bmatrix} R_1 \\ R_2 \\ R_3 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 5 \\ 0 & 3 & 7 \\ 0 & 0 & 43 \end{bmatrix}.$$

- $M = L \cdot U$, the product of a lower triangular matrix and an upper triangular matrix.

A typical ROB 101 student working on LU factorization Homework:

This was the common practice before ROB 101 (ask graduate students):



LU Factorization for Solving Linear Equations

- We wish to solve the system of linear equations $Ax = b$.
- If we can factor $A = L \cdot U$, where U is upper triangular and L is lower triangular. Then

$$L \cdot Ux = b.$$

- Define $U \cdot x =: y$, then

$$\begin{aligned} Ly &= b \\ Ux &= y. \end{aligned}$$

- We first solve for y via forward substitution. Given y, we solve for x via back substitution.

Solving $Ax = b$ via LU with Row Permutation

Solving $Ax = b$ when A is square and $P \cdot A = L \cdot U$.

- $Ax = b \iff P \cdot Ax = P \cdot b \iff L \cdot Ux = P \cdot b$.
- Define $Ux =: y$, then

$$\begin{aligned} Ly &= P \cdot b \\ Ux &= y. \end{aligned}$$

- We first solve for y via forward substitution. Given y, we solve for x via back substitution.

Uber Pro-Tip: Number of Linearly Independent Vectors via an Enhanced LU Factorization

Fact

The matrix $A^{\mathsf{T}} \cdot A$ *always has an LDLT Factorization*

$$P \cdot A^{\mathsf{T}} \cdot A \cdot P^{\mathsf{T}} = L \cdot D \cdot L^{\mathsf{T}}.$$

Moreover,

- *the number of linearly independent columns of* A *is equal to the number of non-zero entries on the diagonal of* D*.*

Least Squares Solutions to $A_{n\times m} \cdot x_{m\times 1} = b_{n\times 1}$

- Assume $A^{\mathsf{T}}A$ is invertible, i.e., the columns of A are linearly independent.

- Then there is a unique vector $x^* \in \mathbb{R}^m$ achieving $\min_{x\in\mathbb{R}^m} ||Ax - b||^2$ and it satisfies the equation (called *the normal equations*)

$$\left(A^{\mathsf{T}}A\right) x^* = A^{\mathsf{T}}b.$$

-

$$x^* = (A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}b \iff x^* = \operatorname*{arg\,min}_{x\in\mathbb{R}^m} \|Ax - b\|^2 \iff \left(A^{\mathsf{T}}A\right) x^* = A^{\mathsf{T}}b.$$

Moreover, we identify $\mathbb{R}^n$ with the set of all n-column vectors with real entries

$$\{(x_1, x_2, \ldots, x_n) \mid x_i \in \mathbb{R}, 1 \le i \le n\}$$
$$\iff \mathbb{R}^n \iff \left\{ \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \;\middle|\; x_i \in \mathbb{R}, 1 \le i \le n \right\}$$

For all real numbers α and β, and all vectors x and y in $\mathbb{R}^n$

$$\alpha x + \beta y = \alpha \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \beta \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$= \begin{bmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{bmatrix} + \begin{bmatrix} \beta y_1 \\ \beta y_2 \\ \vdots \\ \beta y_n \end{bmatrix} = \begin{bmatrix} \alpha x_1 + \beta y_1 \\ \alpha x_2 + \beta y_2 \\ \vdots \\ \alpha x_n + \beta y_n \end{bmatrix}.$$

Suppose that $V \subset \mathbb{R}^n$ is a nonempty subset of $\mathbb{R}^n$.

Definition

V is a subspace of $\mathbb{R}^n$ if any linear combination constructed from elements of V and scalars in $\mathbb{R}$ is once again an element of V. One says that V is *closed under linear combinations.*

In symbols, $V \subset \mathbb{R}^n$ is a subspace of $\mathbb{R}^n$ if for all real numbers α and β, and all vectors v_1 and v_2 in V

$$\boxed{\alpha v_1 + \beta v_2 \in V.}$$

Definition

Suppose that $S \subset \mathbb{R}^n$, then S is a set of vectors. The set of all possible linear combinations of elements of S is called the span of S,

$$\mathrm{span}\{S\} := \{\text{all possible linear combinations of elements of } S\}.$$

Suppose that that the set of vectors $\{u_1, u_2, \ldots, u_m\}$ is linearly independent and you generate a new set of vectors by

$$
\begin{aligned}
v_1 &= u_1, \\
v_2 &= u_2 - \left(\frac{u_2 \bullet v_1}{v_1 \bullet v_1}\right) v_1, \\
v_3 &= u_3 - \left(\frac{u_3 \bullet v_1}{v_1 \bullet v_1}\right) v_1 - \left(\frac{u_3 \bullet v_2}{v_2 \bullet v_2}\right) v_2, \\
&\vdots \\
v_k &= u_k - \sum_{i=1}^{k-1} \left(\frac{u_k \bullet v_i}{v_i \bullet v_i}\right) v_i, \quad \text{(General Step)}
\end{aligned}
$$

Then the set of vectors $\{v_1, v_2, \ldots, v_m\}$ is

- orthogonal, meaning, $i \neq j \implies v_i \bullet v_j = 0$
- span preserving, meaning that, for all $1 \leq k \leq m$,

$$\text{span}\{v_1, v_2, \ldots, v_k\} = \text{span}\{u_1, u_2, \ldots, u_k\},$$

- and linearly independent.

Remark

The unit vectors $\{e_1 = \frac{v_1}{\|v_1\|}, e_2 = \frac{v_2}{\|v_2\|}, \ldots, e_m = \frac{v_m}{\|v_m\|}\}$ *form an orthonormal set.*

Suppose that A is an $n \times m$ matrix with linearly independent columns.

Fact

Then there exists an $n \times m$ matrix Q with orthonormal columns and an upper triangular, $m \times m$, invertible matrix R such that $A = Q \cdot R$.

Least Squares via the QR Factorization

Whenever the columns of A are linearly independent, a least squared error solution to $Ax = b$ is computed as

- factor $A =: Q \cdot R$,
- compute $\bar{b} := Q^{\mathsf{T}} b$, and then
- solve $Rx = \bar{b}$ via back substitution.

Suppose that V is a subspace of $\mathbb{R}^n$. Then $\{v_1, v_2, \ldots, v_k\}$ is a *basis for* V if

- the set $\{v_1, v_2, \ldots, v_k\}$ is linearly independent, and
- $\operatorname{span}\{v_1, v_2, \ldots, v_k\} = V$.
- The maximum number of vectors in any linearly independent set contained in V is the *dimension* of V (here k).

Definition

Let $n \geq 1$ and, as before, define $e_i :=$ i-th column of the $n \times n$ identity matrix, I_n. Then

$$\{e_1, e_2, \ldots, e_n\}$$

is a basis for the vector space $\mathbb{R}^n$.
Its elements e_i are called both natural (standard) basis vectors and canonical basis vectors.

A Function View of a Matrix

- A function (or a map) view of a matrix defines two subspaces:
  1. its *null space* and
  2. its *range*.

- Let A be an $n \times m$ matrix.

- We can then define a function $f : \mathbb{R}^m \to \mathbb{R}^n$ by, for each $x \in \mathbb{R}^m$

$$f(x) := Ax \in \mathbb{R}^n.$$

The following subsets are naturally motivated by the function view of a matrix.

Definition

1. $\mathrm{null}(A) := \{x \in \mathbb{R}^m \mid Ax = 0_{n\times 1}\}$ is the *null space* of A.
2. $\mathrm{range}(A) := \{y \in \mathbb{R}^n \mid y = Ax \text{ for some } x \in \mathbb{R}^m\}$ is the *range* of A.

Range of A Equals Column Span of A

Let A be an $n \times m$ matrix, its columns are vectors in $\mathbb{R}^n$,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} =: \begin{bmatrix} a_1^{\mathrm{col}} & a_2^{\mathrm{col}} & \ldots & a_m^{\mathrm{col}} \end{bmatrix}$$

Then

$$\boxed{\mathrm{range}(A) := \{Ax \mid x \in \mathbb{R}^m\} = \mathrm{span}\{a_1^{\mathrm{col}}, a_2^{\mathrm{col}}, \ldots, a_m^{\mathrm{col}}\} =: \mathrm{col\ span}\{A\}.}$$

Relation of Null Space and Range to Solutions of Linear Equations

Suppose that A is $n \times m$. Here are the key relations between solutions of $Ax = b$ and the null space and range of A.

1. $Ax = b$ has a solution if, and only if, $b \in \mathrm{range}(A)$.

Relation of Null Space and Range to Solutions of Linear Equations

Suppose that A is $n \times m$. Here are the key relations between solutions of $Ax = b$ and the null space and range of A.

2. If $Ax = b$ has a solution, then it is unique if, and only if, $\mathrm{null}(A) = \{0_{m\times 1}\}$.

Relation of Null Space and Range to Solutions of Linear Equations

Suppose that A is $n \times m$. Here are the key relations between solutions of $Ax = b$ and the null space and range of A.

3. Suppose that $\tilde{x}$ is a solution of $Ax = b$, so that $A\tilde{x} = b$. Then the set of all solutions is

$$\begin{aligned}\{x \in \mathbb{R}^m \mid Ax = b\} &= \tilde{x} + \mathrm{null}(A) \\ &:= \{\hat{x} \in \mathbb{R}^m \mid \hat{x} = \tilde{x} + \eta, \eta \in \mathrm{null}(A)\}.\end{aligned}$$

Relation of Null Space and Range to Solutions of Linear Equations

Suppose that A is $n \times m$. Here are the key relations between solutions of $Ax = b$ and the null space and range of A.

4. $Ax = b$ has a unique solution if, and only if $b \in \mathrm{range}(A)$ and $\mathrm{null}(A) = \{0_{m\times 1}\}$.

Relation of Null Space and Range to Solutions of Linear Equations

Suppose that A is $n \times m$. Here are the key relations between solutions of $Ax = b$ and the null space and range of A.

5. When $b = 0_{n\times 1}$, then it is always true that $b \in \mathrm{range}(A)$. Hence we deduce that $Ax = 0_{n\times 1}$ has a unique solution if, and only if, $\mathrm{null}(A) = \{0_{m\times 1}\}$.

Definition

For an $n \times m$ matrix A,

1. $\text{rank}(A) := \dim \text{range}(A)$.

2. $\text{nullity}(A) := \dim \text{null}(A)$.

Because $\text{range}(A) \subset \mathbb{R}^n$, we see that $\text{rank}(A) \le n$.

Theorem

For an $n \times m$ matrix A, we have the property

$$\mathrm{rank}(A) + \mathrm{nullity}(A) = m \quad \textit{number of columns of } A.$$

- *Since* $\mathrm{rank}(A)$ *is equal to the number of linearly independent columns of* A*, it follows that* $\mathrm{nullity}(A)$ *is counting the number of linearly dependent columns of* A.

- *If all of the columns of* A *are linearly independent, then none are dependent, and hence* $\mathrm{null}(A) = \{0_{m\times 1}\}$.

Proof.

See Chapter 10.6 of ROB 101 Book. □

Next Time

- Changing Gears!
- Chapters 11